

# Quality Control of Orphaned Proteins

Szymon Juskiewicz<sup>1</sup> and Ramanujan S. Hegde<sup>1,\*</sup>

<sup>1</sup>MRC Laboratory of Molecular Biology, Cambridge CB2 0QH, UK

\*Correspondence: [rhegde@mrc-lmb.cam.ac.uk](mailto:rhegde@mrc-lmb.cam.ac.uk)

<https://doi.org/10.1016/j.molcel.2018.07.001>

The billions of proteins inside a eukaryotic cell are organized among dozens of sub-cellular compartments, within which they are further organized into protein complexes. The maintenance of both levels of organization is crucial for normal cellular function. Newly made proteins that fail to be segregated to the correct compartment or assembled into the appropriate complex are defined as orphans. In this review, we discuss the challenges faced by a cell of minimizing orphaned proteins, the quality control systems that recognize orphans, and the consequences of excess orphans for protein homeostasis and disease.

## Introduction

The advent of methods to observe the interior of cells by electron microscopy in the 1950s revealed far more compartmentalization than was apparent previously using light microscopy with vital dyes and stains (Porter, 1955–1956). This led to the definition of many membrane-bound organelles in the cell and our first views of many non-membrane-bound structures such as the ribosome (Palade, 1955). In parallel with direct visualization, biochemical fractionation was able to isolate these various morphologically identifiable structures (Claude, 1943, 1946). The structures were soon shown to have distinct functions (De Duve, 1965) and eventually associated with specific subsets of cellular proteins that are often part of larger complexes with defined composition and stoichiometry. These studies initiated modern cell biology, a major goal of which is to understand how intracellular organization is generated and maintained to facilitate cellular function.

The most extensive efforts have been aimed at determining how newly made proteins are segregated to the appropriate organelle (Blobel, 1980; Wickner and Schekman, 2005). Although the accuracy of intracellular protein targeting was usually assumed to be high, one could anticipate that it cannot be perfect. As detailed below, it is increasingly clear that the recognition and disposal of mistargeted copies of a protein is an important facet of achieving effective net segregation. Similarly, the assembly of proteins into complexes is likely to be imperfect (Harper and Bennett, 2016), necessitating the degradation of unassembled components. Thus, the prompt degradation of proteins that fail to be correctly localized or assembled is critical to the maintenance of intracellular organization. We refer to proteins that are terminally separated from their correct location or partners as orphans.

In this review, we discuss our understanding of how cells recognize orphans and selectively route them for degradation. We begin by providing an accounting of the eukaryotic proteome and the extent of its sub-cellular organization into compartments and complexes. After defining the challenges to achieving a well-organized proteome, we consider successively the mechanisms cells use to identify mislocalized and misassembled proteins. We end with a consideration of how pathways for the quality control of orphaned proteins impact cellular physiology and disease.

## Organization of the Eukaryotic Proteome

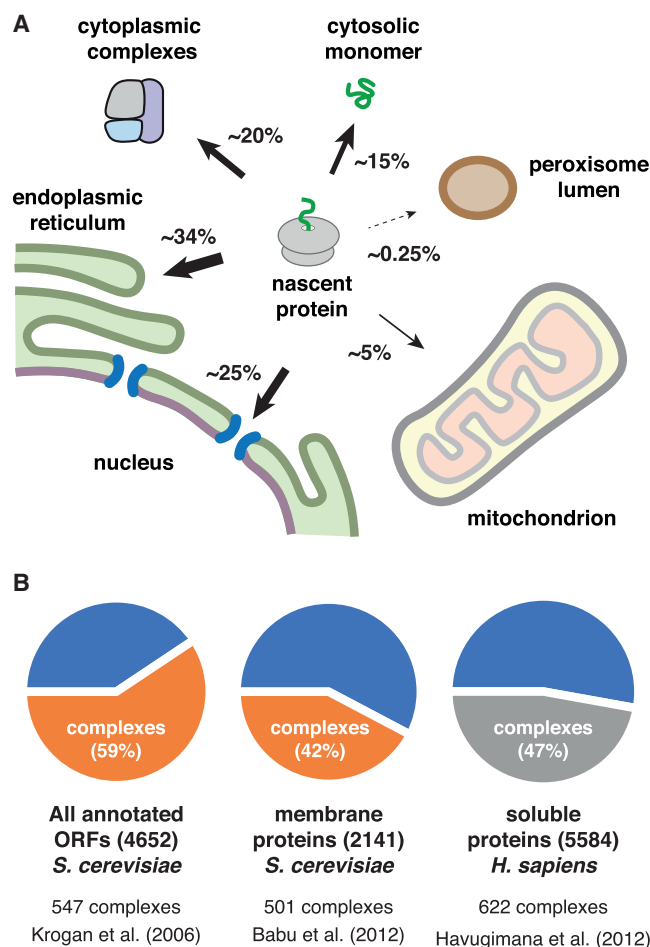
Bioinformatic tools to predict protein location based on known trafficking signals (Emanuelsson et al., 2007), together with increasingly thorough proteome-wide analytic tools (Itzhak et al., 2016; Mulvey et al., 2017), provide estimates of how many proteins are delivered to non-cytosolic destinations. In parallel, systematic analyses of protein complexes (Babu et al., 2012; Gavin et al., 2006; Krogan et al., 2006) and their relative abundances (Kulak et al., 2014) inform on what proportion of nascent proteins need additional assembly. These studies show that the vast majority of newly made proteins are destined for a different location, assembled with other cellular factors, or both (Figure 1).

Of the 20,341 reference proteins of the human genome (UniProt Consortium, 2018), ~7,000 are targeted to the endoplasmic reticulum for eventual residence in the endomembrane system, nuclear envelope, plasma membrane, outside the cell, and peroxisomal membrane. Approximately 1,000 proteins are destined for mitochondria, ~50 for the peroxisomal lumen, and a relative handful for more specialized structures such as lipid droplets. Around 5,000 proteins operate primarily in the nucleus, although the entry and exit of many of these are often dynamic and regulated. Thus, ~65% of genes encode for proteins that must be recognized for selective trafficking to a membrane-enclosed compartment (Figure 1A).

Analysis of protein interactions by mass spectrometry across the yeast proteome (Gavin et al., 2006; Krogan et al., 2006) indicates that over half of all proteins may be in stable multi-protein complexes (Figure 1B). While the proportion is somewhat less for membrane proteins (Babu et al., 2012), this may be due to the additional challenge of retaining interactions during detergent solubilization. Although data on the human proteome is less complete, analysis of the relatively abundant cytosolic proteins in HeLa and HEK293 cells indicate that a little less than half of them are in stable complexes (Havugimana et al., 2012). This means that of the ~35% of proteins that do not traffic to another destination (see above), fewer than half are unaccompanied by partners. Thus, only ~15% of genes encode proteins that simply fold and function without first requiring sub-cellular localization and assembly.

When abundance, cell type, and growth rate are taken into account, there are many circumstances where the vast majority





**Figure 1. Accounting of Protein Localization and Assembly**  
(A) The approximate percentages of human genes whose proteins are destined for various intracellular compartments are shown.  
(B) Data from three proteomic studies indicating that roughly half of all proteins in yeast and mammals are in protein complexes.

of synthesized proteins must be localized or assembled. For example, in highly secretory cell types such as pancreatic exocrine cells or antibody-secreting plasma cells, nearly all ribosomes are synthesizing proteins destined for the ER (Brewer and Hendershot, 2005; Pfeffer et al., 2016). In reticulocytes, almost all protein synthesis is dedicated to  $\alpha$ - and  $\beta$ -globin, which assemble into a heterotetramer (Benz and Forget, 1974). Even in a less specialized cell type making a diverse proteome, the cell's most abundant proteins are localized (e.g., histones to the nucleus) or part of complexes (e.g., tubulin, ribosomes, and proteasomes).

These considerations indicate that even if the rate of failure for localization or assembly is less than 0.1%, the absolute number of polypeptides that must be recognized by quality control is high. As argued below, the actual failure rates are likely to be a few percent, imposing a substantial constitutive burden on quality control pathways. Based on these estimates, we propose that orphans are the major source of substrates for most cellular quality control pathways in normal unstressed cells. Thus, a

mechanistic framework for their selective recognition and degradation is of substantial importance to understanding how protein homeostasis is maintained in cells to avoid various disease states.

### Challenges to Accurate Protein Localization

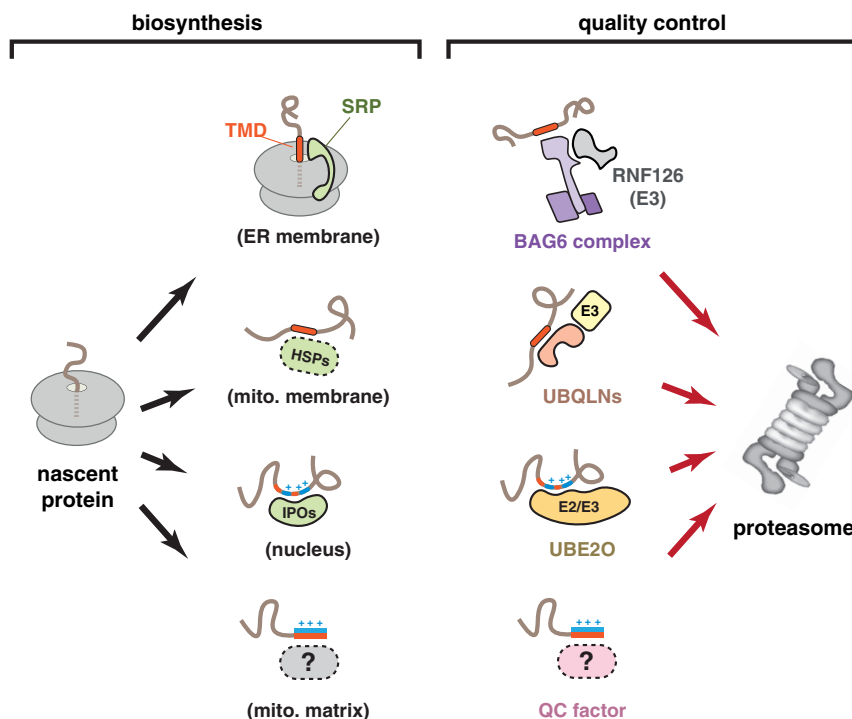
The fundamental principle underlying protein segregation to organelles is the recognition of signal sequences within a nascent protein by targeting factors that specify the appropriate destination (Blobel, 1980). With few exceptions, the targeting signal for a particular organelle is not a specific sequence. Instead, it is typically specified by more general properties such as hydrophobicity, length, and charge (von Heijne, 1995). Thus, the set of proteins that are recognized by a targeting factor have signals that can differ substantially in sequence as long as they share the relevant underlying biochemical feature(s).

While molecular recognition of a distinct sequence can be extremely specific, recognition of a diverse set of sequences with loosely shared features will necessarily be limited in specificity. The actual rate of failure has been studied only cursorily and is best understood for secretory protein segregation to the ER. Very early studies anecdotally noted that the signal sequences of some proteins are more efficient than others in cell-free translocation assays. It was later appreciated that not only do signals differ in their relative efficiencies (Kim et al., 2002; Levine et al., 2005), but also in their requirements for translocation machinery (Fons et al., 2003; Ng et al., 1996; Voigt et al., 1996).

Mammalian cell culture experiments comparing different ER signal sequences showed that even the best signals fail ~5% of the time (Levine et al., 2005; Rane et al., 2004). In mouse brain, it has been estimated that the efficient and well-characterized signal from prolactin fails ~1%–2% of the time, while the more average signal from prion protein fails ~5% of the time (Rane et al., 2010). These *in vitro*, cell culture, and *in vivo* experiments all indicate that failure rates of protein segregation to the ER are on the order of 1%–10%. Although similar measurements of mitochondrial import efficiency remain to be performed, the comparable diversity of signal sequences (von Heijne, 1995) and analogous mechanisms of recognition suggest that failure rates may be comparable.

In addition to intrinsic failure of localization under normal conditions, acute organelle stress has been shown to impair protein import into the ER and mitochondria (Kang et al., 2006; Wright et al., 2001). The step that fails appears to involve the actual translocation reaction across the membrane, albeit by different mechanisms. In the case of ER translocation, it is thought that chaperones in the lumen bind partially translocated polypeptides to prevent their back-sliding (Brodsky et al., 1995; Matlack et al., 1999). Engagement of these chaperones with misfolded proteins during ER stress reduces their availability for driving translocation. This is thought to be a protective mechanism to minimize the load of misfolded proteins in the ER during stress (Kang et al., 2006) but results in an increased number of mislocalized proteins orphaned from the ER.

In the case of mitochondria, import of many proteins relies on the electrochemical gradient across the inner membrane (Wiedemann and Pfanner, 2017). Stress conditions that impair this



**Figure 2. Recognition Factors for Protein Localization and Mislocalization**

Sequence features in a nascent protein can be recognized by various factors that mediate localization to the destinations indicated in parentheses (left side). Not all factors or destinations are shown. Orange indicates hydrophobicity, and blue indicates basic residues. SRP, signal recognition particle; IPO, importin family member; HSP, heat shock protein. The right side shows that the same sequence features used by biosynthesis factors can also be recognized by quality control factors that ultimately lead to degradation at the proteasome.

Various studies had informally noted that cell-free translation of a protein intended for the ER results in the protein's ubiquitination. This observation was exploited to identify factors needed for recognition and ubiquitination of the mislocalized protein (Hessa et al., 2011). These experiments led to Bag6 (also called Bat3 or Scythe), a large, widely conserved cytosolic protein capable of recognizing mislocalized proteins. A role for Bag6 in quality control was consistent with earlier studies that reported a role in

the degradation of newly synthesized polypeptides prematurely released from the ribosome with puromycin (Minami et al., 2010). A ubiquitin-like (UBL) domain in Bag6 associates with RNF126, a ubiquitin ligase that works with the E2 enzyme UbcH5 to polyubiquitinate substrates associated with Bag6 (Rodrigo-Brenni et al., 2014). Knockdown of either Bag6 or RNF126 partially impairs degradation of proteins that fail to be successfully imported into the ER.

gradient will result in reduced translocation. At least one protein whose import is impaired initiates a stress response (Nargund et al., 2012), while the general increase in mislocalized proteins seems to initiate a compensatory increase in cytosolic degradation capacity (Wang and Chen, 2015; Wrobel et al., 2015). Other cellular states, such as cell division, may also result in temporary attenuation of mitochondrial protein import (Harbauer et al., 2014; Schmidt et al., 2011; Weidberg and Amon, 2018). Thus, cells active in protein synthesis are constantly challenged with proteins orphaned in the cytosol due to a combination of intrinsic and stress-dependent failures in segregating proteins to the ER and mitochondria. The extent of import failure for proteins that function exclusively in the nucleus (e.g., histones) remains poorly studied and may further contribute to orphans.

### Recognition of Mislocalized Protein Orphans

A number of early observations indicated that mislocalized proteins in the cytosol are degraded. First, rare human mutations that disrupt the function of a signal peptide result in lower overall expression levels (Cassanelli et al., 1998; Karaplis et al., 1995; Seppen et al., 1996). Second, proteasome inhibition led to the appearance of a non-glycosylated form of prion protein (Ma and Lindquist, 2001), which later proved to be generated by stabilization of the non-targeted population (Chakrabarti et al., 2011; Drisaldi et al., 2003; Rane et al., 2004). Third, inhibition of ER translocation by small molecules (Besemer et al., 2005; Garrison et al., 2005) or acute ER stress (Kang et al., 2006) resulted in rapid proteasome-dependent degradation of the non-translocated protein. While these findings showed that mislocalized proteins are degraded in the cytosol, the factors involved were not known.

The key feature of mislocalized proteins that are recognized by Bag6 proved to be precisely the elements that are also recognized by the targeting and insertion machinery: transmembrane domains (TMDs) and signal sequences (Figure 2). Thus, deleting these hydrophobic domains precludes Bag6 recognition and ubiquitination despite the fact that the resulting protein is still mislocalized and presumably unable to fold (Hessa et al., 2011). It is instead apparently recognized by cytosolic quality control pathways that recognize protein misfolding. Indeed, enforced mislocalization by deletion of a signal sequence in yeast results in recognition by the general chaperone Hsp70 (Park et al., 2007).

Analogous studies of a mitochondrial membrane protein showed that prior to its insertion into mitochondria, it associates with members of the Ubiquilin (UBQLN) family (Itakura et al., 2016). Structure-function analysis of UBQLN1 identified a middle domain that specifically associates with the TMDs of membrane proteins in the cytosol. Unlike substrate interaction with Bag6 (Shao et al., 2017), the UBQLN1 association is dynamic (Itakura et al., 2016). However, if the interaction is prolonged, as would be the case if import into mitochondria fails, the membrane protein becomes ubiquitinated in a UBQLN-dependent manner. This ubiquitination is mediated by a yet-unidentified ligase that is

recruited to the ubiquitin-associating (UBA) domain in UBQLN. Although other members of the UBQLN family (mammals have four) have been less well studied, they have similar middle domains that may also interact with TMDs of mislocalized proteins (Itakura et al., 2016; Suzuki and Kawahara, 2016). It is attractive to speculate that their specificities are somewhat different from each other to collectively cover the wide range of sequences that define TMDs. Indeed, Bag6 appears to prefer more hydrophobic TMDs than UBQLN1 (Itakura et al., 2016). Additional studies are needed to investigate the substrate ranges and molecular basis of specificity of these factors.

Unlike membrane proteins whose exposed TMDs provide a distinctive cue for recognition of mislocalization, the soluble proteins of membrane-bound compartments more closely resemble cytosolic proteins. In the case of proteins destined for the ER, signal sequences are sufficiently similar to TMDs that they seem to be recognized by Bag6 (Hessa et al., 2011). Soluble proteins destined for mitochondria or the nucleus have amphipathic or basic signals, respectively. Whether dedicated factors recognize persistent residence of these signals in the cytosol for degradation remains largely unexplored. One candidate factor for this role is UBE2O, a hybrid E2-E3 enzyme that was recently shown to recognize and ubiquitinate ribosomal proteins in the cytosol (Nguyen et al., 2017; Yanagitani et al., 2017). Under normal circumstances, newly made ribosomal proteins are recognized by dedicated nuclear import factors (Jäkel and Görlich, 1998), delivered to the nucleolus, and assembled with rRNA into pre-ribosomal subunits (Peña et al., 2017). It appears that failure of nuclear import leads to recognition by UBE2O, multi-mono-ubiquitination, and proteasomal degradation.

Although the precise sequence features recognized by UBE2O are unclear, it appears to be the juxtaposition of hydrophobic and basic residues (Yanagitani et al., 2017). This is noteworthy because it is similar in some ways to both nuclear import (Jäkel and Görlich, 1998) and mitochondrial import signals (von Heijne, 1995), as well as a common feature of nucleic acid-binding proteins (Hentze et al., 2018; Nelson, 1995). Ribosomal protein interactions with UBE2O and nuclear import factors appear to be mutually exclusive (Yanagitani et al., 2017), suggesting that UBE2O may interact with the nuclear import signal. It will therefore be interesting to determine whether this is a general property of UBE2O and whether other nucleic acid binding proteins, such as histones, subunits of the signal recognition particle (SRP), and mitoribosomal proteins, are recognized by UBE2O when they are orphaned in the cytosol.

Thus, the cytosol appears to be patrolled by a set of factors that can interact with signal sequences and TMDs on the one hand and ubiquitin ligases on the other (Figure 2). These factors collectively recognize proteins that expose organelle targeting domains, an indicator that targeting may have failed, and mediate their tagging for degradation. It is noteworthy that the substrate specificity of these factors, while likely to be rather broad, appears to be especially well suited to the same elements recognized by protein targeting factors such as SRP and TRC40 (also called Get3), which mediate membrane protein targeting to the ER (Shao and Hegde, 2011), and Importins 5 and 7, which mediate nuclear import of ribosomal proteins (Jäkel and Görlich, 1998). In order to not interfere with targeting, Bag6, UBQLNs,

and UBE2O must act after substrates have attempted targeting, the mechanisms of which are discussed later. Even though these factors collectively recognize a broad range of orphans mislocalized to the cytosol, additional quality control machinery that recognize nuclear and mitochondrial targeting signals may remain to be discovered.

### Proteins Orphaned to the Wrong Organelle

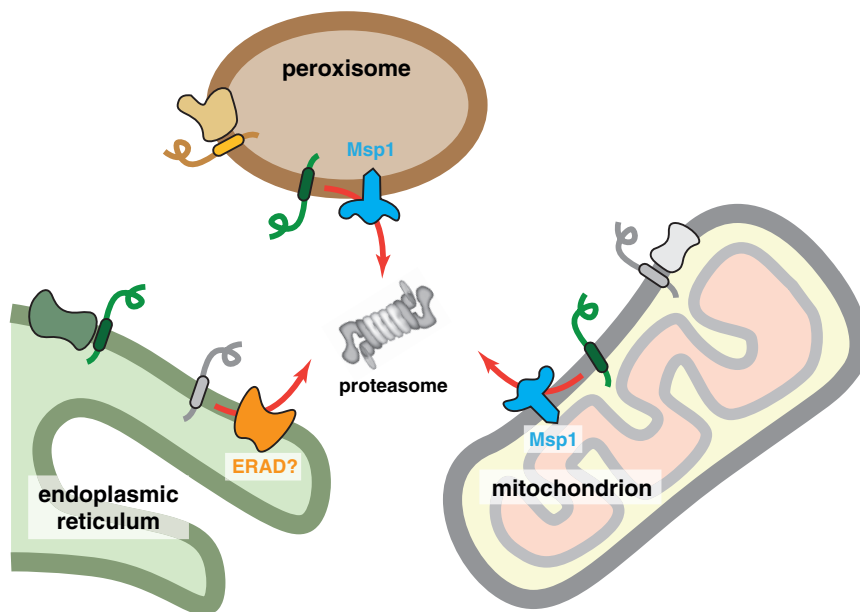
The TMDs of membrane proteins targeted to the ER and mitochondria are very similar in their biophysical properties (Guna and Hegde, 2018), consistent with their eventual residence in a lipid bilayer. For many of these proteins, the TMD(s) serve as the main or sole targeting signal. Due to the similarities among these TMDs, membrane proteins can be routed to the wrong organelle. How cells deal with this problem is only partially understood.

The simplest and best-studied case involves tail-anchored (TA) membrane proteins (Hegde and Keenan, 2011). These proteins contain a single TMD close to the C terminus, which serves as the sole targeting signal. TA proteins destined for the mitochondria tend to have slightly less hydrophobic TMDs (Wattenberg et al., 2007) and often have downstream basic residues in the C-terminal tail (Horie et al., 2002). Nevertheless, there is substantial overlap in their properties and a clear risk of mistargeting.

The best-characterized TA targeting pathway to the ER is known as the TRC (TMD recognition complex) pathway in mammals and the homologous GET (guided entry of TA proteins) pathway in yeast (Hegde and Keenan, 2011). Deletion of GET pathway components in yeast results in mislocalization to the cytosol (often as protein aggregates) and mitochondria (Jonikas et al., 2009; Schuldiner et al., 2008). A membrane-embedded AAA-ATPase Msp1 in the mitochondrial outer membrane is thought to be involved in removing TA proteins mislocalized to mitochondria (Chen et al., 2014; Okreglak and Walter, 2014). A sub-population of Msp1 is also found in the peroxisomal membrane, where it is thought to serve a similar role (Weir et al., 2017). Hence, the levels of mislocalized TA proteins in mitochondria or peroxisomes increases when Msp1 (or its human homolog, ATAD1) is deleted. As this stabilization is seen even in wild-type cells containing an intact TRC/GET pathway (Chen et al., 2014), mislocalization must be a constitutive problem in cells and tissues, consistent with the challenges of achieving precise targeting specificity.

The mechanistic basis of Msp1 function in clearance of mislocalized protein substrates is not well understood. Purified Msp1 reconstituted into synthetic liposomes was shown to drive ATP-dependent extraction of a co-reconstituted TA protein to the cytosol (Wohlever et al., 2017). The ATPase domain of Msp1, which faces the cytosol, was shown to be a hexamer containing a central pore. Because mutations within the pore impair the TA protein extraction activity, it has been posited that the substrate is pulled through the pore into the cytosol (Wohlever et al., 2017). Such a model would be consistent with the mechanism of action of other AAA protein unfoldases for which translocation through the central pore is strongly supported (Sauer and Baker, 2011). How the substrate is kept in a soluble form after extraction, how it is ubiquitinated, and how it is delivered to the proteasome all remain unknown. One possibility is that cytosolic factors such





**Figure 3. Recognition of Protein Mislocalization at Organelles**

Membrane protein insertion into the wrong organelle can be recognized by its lack of association with an interaction partner. In mitochondria and peroxisomes, the hexameric ATPase Msp1 is involved in extracting such mislocalized proteins to the cytosol for degradation by the proteasome. In the ER, the specific machinery for mislocalized proteins has not been studied, but probably involves known ER-associated degradation (ERAD) pathways. The molecular basis of recognition is not known, but may involve the same features that facilitate assembly into protein complexes, explaining why successful assembly would prevent recognition by quality control factors.

as Bag6 or UBQLNs may mediate these downstream steps similarly to how they recognize and degrade cytosolic orphans generated by failed targeting. Consistent with this concept, Bag6 has been shown to maintain the solubility of membrane proteins extracted from the ER until their degradation at the proteasome (Wang et al., 2011).

The basis of substrate selection by Msp1 is also only partially understood. Studies of the peroxisomal TA protein Pex15 have provided initial insights into this key issue (Weir et al., 2017). Pex15 is normally found in complex with Pex3. It was observed that excess Pex15 in peroxisomes is degraded in an Msp1-dependent manner, while Pex15 in complex with Pex3 evades Msp1 (Weir et al., 2017). As discussed in detail later, orphans of multi-protein complexes are often targets for quality control via recognition of the regions that are normally shielded by interaction partner(s). It appears that Msp1 may use such a mechanism to recognize orphaned Pex15 and perhaps other substrates (Figure 3).

This finding suggests that when Pex15 is mistargeted to mitochondria, its recognition by Msp1 for degradation is due to the absence of Pex3. Whether all Msp1 substrates are normally part of larger complexes whose absence cues their recognition remains to be determined. Furthermore, it is unclear whether Msp1 directly recognizes substrates or uses adaptors. In this context, it is noteworthy that Msp1 associates with Cis1, a protein that is upregulated by excessive mitochondrial import failure and plays a yet-unknown role in clearance of failed import products (Weidberg and Amon, 2018). Whether Cis1 is involved in identifying substrates for delivery to Msp1 remains unclear. Thus, both the specific features of mistargeted proteins that are recognized and the factors that mediate recognition remain to be elucidated in molecular terms.

The factor(s) that play the analogous role to Msp1 in the ER are unknown. It has recently been shown that mitochondrial TA proteins can be mistargeted to the ER, particularly under

overexpression conditions (Vitali et al., 2018). Furthermore, the ribosome-associated chaperone-like protein NAC appears to antagonize SRP and limit its promiscuous recognition of nascent mitochondrial proteins (Gamerding et al., 2015; Wiedmann et al., 1994).

Thus, when NAC levels are reduced, mitochondrial proteins can be detected at the ER. Although these examples of ER mistargeting are seen under perturbed conditions, it is likely that a low level of mistargeting occurs in normal cells. How these mistargeted proteins are recognized and degraded is unclear. One possibility is that known ER-associated degradation pathways are involved (Vembar and Brodsky, 2008). This problem warrants investigation. Tools to induce excessive ER mistargeting should facilitate this line of inquiry.

### Challenges to Assembly of Protein Complexes

After segregating proteins to the correct compartment, most proteins must additionally assemble with one or more partners (typically other proteins, but also RNAs). The assembly of multi-protein complexes imposes three major challenges to the cell. First, the components of the complex must be synthesized in the appropriate stoichiometry. Second, the unassembled subunits of a complex must avoid inappropriate interactions until their assembly. Third, partners must find each other within a crowded cell. As with protein localization, limitations to the overall efficiency of achieving these essential steps in complex assembly results in orphans that must be recognized and degraded.

Direct analysis of translation rates by sequencing of ribosome-protected footprints (ribosome profiling) indicates that cells typically express subunits of a multi-protein complex at close to the stoichiometry found in the final complex (Li et al., 2014). A combination of several mechanisms contributes to this concordance. First, the genes for proteins that function together are frequently organized in a single operon in prokaryotes (Ames and Martin, 1964). This allows the transcript levels to necessarily be increased and decreased in unison. Differences in the order of genes in an operon and their respective Shine-Dalgarno sequences can presumably further tune their relative levels of translation (Bonde et al., 2016; Lim et al., 2011).

In eukaryotes, gene order (Dávila López et al., 2010) and the similarities of promoters engaged by the same transcription factors (Lee et al., 2009) can also coordinate transcript levels of protein subunits. Beyond this, further fine-tuning of mRNA half-lives and translation rates via untranslated regions and codon optimality probably contribute to expression at the desired stoichiometry. For example, the 5' end of mRNAs encoding ribosomal proteins and many translation factors contain terminal oligo-pyrimidine sequences (Meyuhas and Kahan, 2015). These sequences allow the translation of these mRNAs to be repressed and de-repressed in concert, thereby maintaining their approximate stoichiometry. Nevertheless, these mechanisms are all subject to stochasticity (Munsky et al., 2012), meaning that the inevitable imbalances in synthesis levels generate orphans. It is reasonable to assume that achieving expression stoichiometry to greater than 90% precision is very challenging given the inherent noisiness of gene expression. For complexes with many subunits, such as ribosomes and proteasomes, matching expression levels may be particularly difficult.

The mechanisms by which proteins are assembled into complexes has been studied in a number of contexts, including ribosome biogenesis in the nucleolus (Peña et al., 2017), hemoglobin assembly in pre-erythrocytes (Feng et al., 2004; Kihm et al., 2002), and histone assembly (Hammond et al., 2017). Each of these processes is distinctive in its use of specific chaperones and assembly factors. The main underlying principle from these studies is that unassembled subunits of a complex are held temporarily by an assembly factor until it is displaced by the appropriate interacting subunit. In many cases, the precise order of protein complex assembly is important and under selective pressure to be maintained, presumably to minimize misassembled products (Marsh et al., 2013). However, the efficiency of these processes is unknown in most cases, especially *in vivo*, so their relative contributions to the generation of orphans remains to be determined.

Recent proteomic pulse-chase experiments indicate a substantial constitutive burden of unassembled orphans in cultured cells (McShane et al., 2016). In this study, kinetic analysis of protein degradation found ~10% of proteins are degraded non-exponentially, with a proportion (ranging from 10% to 70%) being degraded immediately after synthesis. Approximately 70% of these non-exponentially degraded proteins are subunits of known multi-protein complexes, suggesting that the rapidly degraded populations are those polypeptides that failed assembly. As this analysis identified only the most abundant proteins, it appears that cells must constantly identify and degrade unassembled orphans.

### Recognition of Unassembled Cytosolic Protein Orphans

It has been apparent for many decades that unassembled orphans of multi-protein complexes are selectively degraded. The earliest appreciation of this phenomenon comes from the study of hemoglobin, a hetero-tetramer normally formed of two  $\alpha$  and two  $\beta$  subunits (Perutz et al., 1960). Mutations in hemoglobin that markedly reduce its unusually long half-life were correlated with an increased tendency of tetramer dissociation into subunits and decreased binding to heme (Jacob et al., 1968; Rieder, 1974). Many years later, studies of the multi-subunit

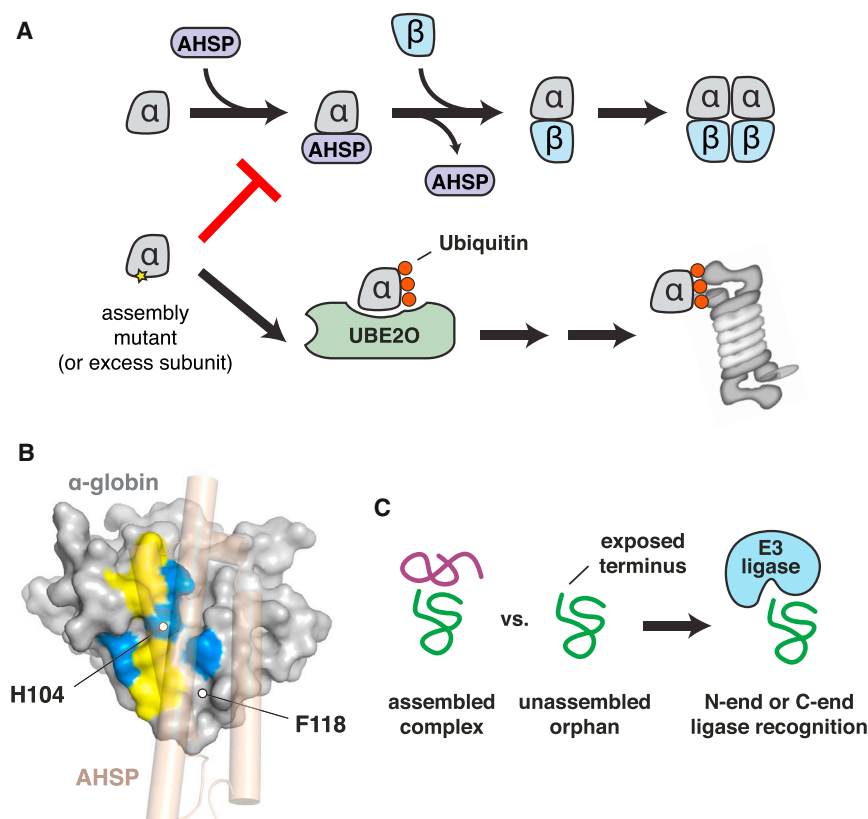
T cell receptor (TCR) showed that the  $\alpha$  subunit is rapidly degraded when expressed without its interaction partners (Lippincott-Schwartz et al., 1988). Conversely, it is well appreciated that knockdown or knockout of one subunit of a multi-protein complex often destabilizes the other subunits. Thus, in numerous contexts and cellular compartments, cells have mechanisms to selectively identify orphaned subunits and target them for degradation.

Two general mechanisms can explain how orphans are recognized. First, some proteins may not be able to achieve a stable folded state in the absence of its interaction partner. This instability would result in its unfolding and recognition by general quality control pathways that monitor protein folding. Second, a subunit interface that is normally shielded in the intact complex is recognized by a factor that ultimately leads to degradation. Factors for both mechanisms have been identified in different biological contexts.

The most challenging multi-subunit complex to assemble in cells is probably the ribosome. Its biogenesis from 80 proteins and pre-rRNA involves dozens of dedicated assembly factors, numerous processing reactions of both proteins and rRNA, a variety of post-translational modifications, and multiple transport steps across the nuclear envelope (Peña et al., 2017). Individual ribosomal proteins have long been known to be refractory to overexpression, suggesting efficient mechanism(s) for degrading excess copies (Abovich et al., 1985; Warner et al., 1985). Despite the fact that ribosomes have a very long half-life, ribosomal proteins are a major set of proteins that accumulate when proteasome activity is diminished (Mayor et al., 2005, 2007). This suggests that ribosomal proteins may be a major source of orphans in actively growing cells. The need to destroy orphaned ribosomal proteins may be especially critical as these factors are aggregation-prone in the absence of an RNA scaffold or a ribosome assembly chaperone (Jäkel et al., 2002).

As discussed above, a ribosomal protein orphaned at the step of import into the nucleus is recognized in the cytosol by UBE2O (Yanagitani et al., 2017). After nuclear import, failure at the step of assembly with rRNA is recognized by a different mechanism. In yeast, the large nuclear-localized E3 ligase Tom1 was shown to interact with unassembled ribosomal proteins and mediate their ubiquitination (Sung et al., 2016). The sites of ubiquitination suggest that Tom1 probably interacts with basic regions that would ordinarily interact with rRNA. These regions are inaccessible in intact ribosomes, which are not recognized by Tom1. HUWE1 is the closest homolog of Tom1 in mammals. Its knockdown partially increases the levels of overexpressed uL24/RPL26, but the basis for this effect was not explored in sufficient depth to conclude a direct effect (Sung et al., 2016).

In separate studies, however, it was observed that HUWE1 is required for efficient degradation of farnesyltransferase alpha (FTNA) and UBL4A when their respective binding partners are absent (Xu et al., 2016). Although interaction studies were not performed, HUWE1 presumably recognizes the exposed interfaces of these and other proteins to mediate their ubiquitination. Proteomic analysis identified 72 candidate HUWE1 substrates (although, curiously, not ribosomal proteins), almost all of which are known subunits of multi-protein complexes. Most of these candidate substrates operate in the nucleus, yet HUWE1



**Figure 4. Assembly Interfaces Serve as Degradation Signals**

(A) Schematic of hemoglobin assembly in which the  $\alpha$  subunit is temporarily shielded by  $\alpha$ -hemoglobin-stabilizing protein (AHSP) until it assembles with the  $\beta$  subunit. Free  $\alpha$  subunit, generated due to excess or a mutation that prevents assembly, is instead recognized by UBE20 for ubiquitination and targeting to the proteasome for degradation. (B) The assembly interface of  $\alpha$ -globin showing hydrophobic (yellow) and basic (blue) surfaces. AHSP is shown in transparent tan to illustrate how this interface is shielded. The positions of two assembly mutants that cause human anemia are indicated. UBE20 is known to bind composite basic/hydrophobic peptides, suggesting that the interface is recognized via these features when assembly fails.

(C) An exposed N or C terminus with a destabilizing residue can be shielded in a correctly assembled complex (left) but exposed as an unassembled orphan. N-end and C-end E3 ubiquitin ligases that recognize such residues would then selectively target the orphan for degradation.

appears to function in the cytosol. Thus, HUWE1 may identify orphans that fail localization, assembly, or both. The molecular features shared by HUWE1 clients and the basis of their recognition remain to be addressed.

In addition to its role in recognizing mislocalized ribosomal proteins discussed above, UBE20 can also recognize folded but unassembled proteins. This conclusion comes from the observation that *in vitro* and *in vivo*, UBE20 is needed for efficient ubiquitination and degradation of unassembled  $\alpha$ -globin (Nguyen et al., 2017; Yanagitani et al., 2017). The interface of  $\alpha$ -globin that interacts with  $\beta$ -globin is usually shielded by  $\alpha$ -hemoglobin-stabilizing protein (AHSP) until assembly with  $\beta$ -globin (Feng et al., 2004, 2005) (Figure 4A). Mutations in  $\alpha$ -globin that impair interaction with AHSP (and  $\beta$ -globin) are more effectively recognized and ubiquitinated by UBE20 (Yanagitani et al., 2017). This indirectly suggests that UBE20 recognizes the same region as AHSP. This interface is basic and hydrophobic (Figure 4B), which is similar to other established UBE20 targets such as unassembled ribosomal proteins (Nguyen et al., 2017; Yanagitani et al., 2017). Thus, orphan  $\alpha$ -globin appears to be identified for degradation by its exposed assembly interface.

Notably, UBE20 is highly upregulated during erythrocyte development concordantly with increased globin production (Wefes et al., 1995). Whether this upregulation is a programmed event in preparation for the inevitable increase in substrate load or a response to increased orphans remains to be determined.

In support of the latter possibility, acute induction of protein misfolding has been shown to cause UBE20 upregulation in cultured cells (Miyazaki et al., 2015). Structure-function analysis of UBE20 indicates that it has two major substrate interaction domains with different binding properties (Nguyen et al., 2017; Yanagitani et al., 2017). Identifying the client

ranges of these domains and the structural basis of their interactions remains an important goal.

The general concept of orphan recognition via exposure of a normally buried interface has been exploited by other ubiquitination factors. The N-end rule pathways of substrate ubiquitination recognize certain residues at an exposed N terminus and selectively target these proteins for degradation (Tasaki et al., 2012). One mechanism involves recognition of N-terminal acetylation by the ubiquitin ligase CNOT4 (Shemorry et al., 2013). Several multi-protein complexes appear to contain N-acetylated subunits whose N terminus is shielded by other subunits in the intact complex. Failure to assemble into a complex (or complex disassembly) exposes the N-terminal degron, resulting in ubiquitination by an N-end rule ubiquitin ligase (Figure 4C). Examples include the Hcn1 subunit of the APC/C complex and the Cog1 subunit of the yeast COG complex.

It is possible that other N-end rule ligases exploit a similar mechanism to recognize different unassembled proteins. For example, the N-end rule ubiquitin ligase Ubr1 was shown to be needed for efficient clearance of unassembled Fas2, a subunit of the fatty acid synthase complex (Scazzari et al., 2015). Whether Ubr1 uses its N-end rule recognition domain or uses an adaptor such as Hsp70 to identify unassembled Fas2 is unclear at present. Hsp70 is needed for Ubr1-dependent Fas2 ubiquitination, supporting the latter hypothesis. Presumably, an interface of Fas2 shielded by its interaction partner, Fas1, is exploited for its recognition by chaperones, which over time

recruits a ubiquitin ligase to initiate degradation. The recent discovery of a “C-end rule” (Koren et al., 2018; Lin et al., 2018) raises the possibility that exposure of a normally shielded C terminus might also be exploited by cells to recognize orphans (Figure 4C). This idea remains to be investigated.

### Recognition of Unassembled Protein Orphans at the ER

Each of the examples of unassembled protein degradation discussed thus far occur in the cytosol. Analogous processes operate in the ER membrane and lumen and probably also different mitochondrial compartments. In the ER, the most extensively studied multi-protein assemblies have been the immunoglobulins and TCR. TCR $\alpha$  and CD3 $\delta$ , two subunits of the TCR, have been widely used as models for ERAD (Bonifacino and Lippincott-Schwartz, 1991; Bonifacino et al., 1989; Huppa and Ploegh, 1997; Lippincott-Schwartz et al., 1988). Both subunits are single-spanning transmembrane proteins that are recognized for degradation via their TMDs (Bonifacino et al., 1991; Cosson et al., 1991; Manolios et al., 1990). The prevailing model is that the orphaned TMDs of TCR $\alpha$  and CD3 $\delta$  are recognized by the membrane-embedded regions of a ER-resident ubiquitin ligase complex centered around Hrd1 (Kikkert et al., 2004). Although the structural basis of their recognition remains unknown, it depends critically on charged residues within their TMDs that are shielded by other subunits in the assembled TCR. A similar mode of recognition by Hrd1 was also shown for other integral membrane substrates such as Hmg2 (Sato et al., 2009).

Based on these examples, one might speculate that orphans of multi-protein complexes assembled via their TMDs are recognized by the exposure of polar residues within the lipid bilayer. This would provide a relatively universal cue for unassembled proteins without relying on sequence-specific recognition. In this light, it is intriguing that the recently determined cryo-EM structure of Hrd1 shows a hydrophilic pocket that might serve as the site of recognition of such orphaned TMDs (Schoebel et al., 2017). Even though the functional role of this putative binding region of Hrd1 remains to be investigated experimentally, the model is consistent with previous observations that mutations within this hydrophilic region of Hrd1 affect recognition of some integral membrane ERAD substrates (Sato et al., 2009).

Although recognition of orphaned TMDs within the membrane has long been presumed, experimental evidence for this idea is generally lacking. An alternative idea is that orphaned TMDs, due to their exposed polar or charged residues, are unstable in the lipid bilayer and translocate into the ER lumen (Feige and Hendershot, 2013). This results in their recognition by luminal chaperones, particularly BiP, which are proposed to deliver them for degradation. Successful assembly would prevent this translocation, providing an explanation for why orphans are selectively degraded. Although both models ultimately involve Hrd1 (or another ubiquitin ligase complex) mediating retrotranslocation into the cytosol for degradation, one posits recognition in the ER lumen while the other involves recognition in the membrane.

Although the two views have yet to be reconciled, one attraction of the luminal recognition model is its similarity to how mis-assembled luminal protein complexes might be monitored. One example of protein complex assembly in the ER lumen is immunoglobulins typically built from two heavy and two light chains

(Huber et al., 1976). After translocation into the ER lumen, the first heavy chain immunoglobulin domain (CH1) binds to the chaperone BiP in an unfolded and reduced state (Bole et al., 1986; Haas and Wabl, 1983; Hendershot et al., 1987; Vanhove et al., 2001). CH1 can only fold upon interaction with light chain (Feige et al., 2009). Hence, in the absence of light chains, heavy chains are retained in the ER and are eventually dislocated to the cytosol to be degraded by the proteasome (Mancini et al., 2000). Even though the assembly process varies between different immunoglobulin isotypes (Baumal et al., 1971), BiP-mediated retention of heavy chains in the ER appears to be universal (Hendershot et al., 1987). Prolonged association with BiP, whether of orphan heavy chains or TCR subunits, may result in delivery to ERAD machinery via adaptors such as specific J-proteins (Shen and Hendershot, 2005).

### Distinguishing Intermediates from Orphans

Proteins should be recognized as orphans to be degraded only after reasonable attempt(s) at localization and assembly have failed. Thus, the biosynthesis machinery should have higher priority for access to nascent chains than the quality control machinery, but this priority should be time limited. The mechanistic basis of how priority and timing are determined is important to understand because this critical decision balances promiscuous degradation of normal biosynthetic intermediates versus excessive persistence of potentially toxic and aggregation-prone orphans.

One mechanism by which priority can be conferred is by spatial segregation of the biosynthetic and quality control factors. The best illustration of this principle is SRP interaction with the ribosome. The majority of eukaryotic proteins destined for the ER are recognized co-translationally by SRP (Chartron et al., 2016; Costa et al., 2018; Keenan et al., 2001). This interaction is facilitated by the ability of SRP to interact with the ribosome such that the signal sequence binding domain is positioned precisely at the ribosomal exit tunnel (Halic et al., 2004; Voorhees and Hegde, 2015). Not only is SRP thought to sample ribosomes for the presence of an emerging signal, but it may be stabilized there while the signal is still inside the ribosomal tunnel (Berndt et al., 2009; Voorhees and Hegde, 2015).

In this manner, nascent secretory and membrane proteins would not have an opportunity to engage quality control factors such as Bag6 or UBQLNs unless SRP recognition has failed. Thus, despite its ~5- to 10-fold lower abundance (Kulak et al., 2014), SRP nevertheless has priority due to its precise localization at the ribosome. Once a protein has been released from the ribosome, SRP's relatively low abundance and poor capacity to bind proteins in solution precludes its interference with other processes.

At a later step in biogenesis, secretory and membrane proteins are translocated into the ER through the ribosome-associated Sec61 translocation channel (Rapoport et al., 2017). Any proteins that selectively associate, even dynamically, with Sec61 will get priority for interaction with the nascent chain over those that do not. Indeed, processing enzymes such as oligosaccharyl transferase and signal peptidase enjoy such priority (Braunger et al., 2018), as might the chaperone BiP (via its recruitment by the



translocon component Sec63 [Brodsky et al., 1995]) and Calnexin [Lakkaraju et al., 2012]. In this manner, biosynthetic factors could be given priority before degradation factors such as Hrd1.

A second mechanism of conferring priority is to use a combination of abundance and fast binding. This appears to be how newly made TA proteins are prioritized for ER targeting ahead of proteasomal degradation. Recent studies show that among the three TMD binding factors involved in TA protein targeting, SGTA is the fastest and first interactor [Shao et al., 2017]. Its competitive advantage may be further increased by its interaction with the Bag6 complex or Hsp70, both of which can associate with the ribosome. TA proteins can transfer from SGTA to TRC40 [Mock et al., 2015; Shao et al., 2017] by a mechanism that does not require release of TA protein into the bulk cytosol [Shao et al., 2017]. Loading on TRC40 is a commitment to targeting because TA protein dissociation is very slow relative to the rate of targeting. TA proteins have a limited time to complete these events as dictated by their off-rate from SGTA. Dissociation from SGTA permits an opportunity to be captured by Bag6, which recruits an E3 ligase for substrate ubiquitination [Rodrigo-Brenni et al., 2014]. Very slow dissociation from Bag6 means TA binding to this factor is effectively a commitment for degradation. Thus, triage by multiple factors of seemingly identical specificity can be accomplished simply by their differential on- and off-rates combined with committed downstream events such as membrane insertion or ubiquitination [Shao et al., 2017].

The final mechanism is based on an intrinsically slow reaction followed by a specific commitment step. This is best exemplified by glycoprotein quality control in which a specific irreversible trimming event of an N-linked glycan generates a product that is recognized by a lectin coupled to quality control pathways [Caramelo and Parodi, 2015; Tannous et al., 2015]. The trimming event is carried out by intrinsically slow mannosidases, effectively placing a time limit on protein maturation attempts. As predicted by this model, overexpression of the mannosidase accelerates the rate of degradation [Hosokawa et al., 2001].

An analogous mechanism of a slow enzymatic reaction linked to a commitment step seems to be used for some orphans in the cytosol. Membrane proteins bound to UBQLNs are not initially committed for degradation [Itakura et al., 2016]. Instead, their dynamic release and re-binding provides opportunities at successful insertion. However, a relatively slow ubiquitination step seems to be the key commitment step. When ubiquitin is added to the substrate, a UBA domain in UBQLNs binds to ubiquitin, preventing substrate release and thereby ending insertion attempts [Itakura et al., 2016]. Presumably, the time allowed for insertion is dependent on the affinity of the ligase for UBQLNs, its abundance, and speed of ubiquitination. Identifying the ligase should allow these aspects of the model to be tested.

### Can Failures Be Detected Co-translationally?

The systems for detection and degradation of orphans discussed so far all operate post-translationally. This is consistent with the idea that biosynthesis must necessarily have an opportunity to succeed before a polypeptide is deemed an orphan. However, many biosynthetic reactions are initiated or occur co-translationally: recognition by SRP, the initial stages of polypeptide folding, various modifications, and even multi-subunit

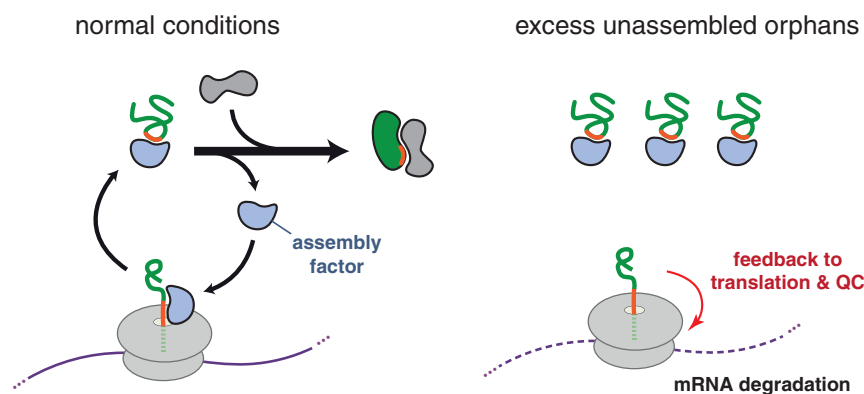
assembly. It is therefore possible that cells have evolved mechanisms to detect orphans co-translationally via failure of a decisive co-translational biosynthetic step. Early detection is advantageous because minimizing the time an aberrant polypeptide resides in a cell reduces the risk of inappropriate interactions, aggregation, and other adverse consequences. Although co-translational orphan detection has not been established unambiguously, two examples are noteworthy.

The first example concerns an autoregulatory mechanism to control  $\beta$ -tubulin expression. As noted already,  $\alpha$ - and  $\beta$ -tubulin form a constitutive complex [Feit et al., 1971]. In the presence of excess unpolymerized tubulin subunits, synthesis of both  $\alpha$ -tubulin and  $\beta$ -tubulin polypeptides is reduced [Cleveland et al., 1981]. One regulatory mechanism involves selective degradation of  $\beta$ -tubulin mRNA [Pittenger and Cleveland, 1985]. This degradation is strictly dependent on translation and more specifically on the first four amino acids in the nascent polypeptide [Bachurski et al., 1994; Gay et al., 1989]. Remarkably, a monoclonal antibody that binds these four residues abolished translation-dependent mRNA degradation [Theodorakis and Cleveland, 1992]. These findings led to the idea that co-translational association of nascent  $\beta$ -tubulin with some yet-unidentified factor (which is not  $\alpha$ -tubulin) is needed to escape mRNA degradation.

In the intervening time since these findings, various pathways of mRNA decay have been characterized, including ones dependent on ribosome stalling [Shoemaker and Green, 2012]. Thus, one attractive model is that a limiting assembly factor binds nascent  $\beta$ -tubulin, the absence of which leads to translation arrest and mRNA degradation. Assuming that this factor is displaced upon assembly with  $\alpha$ -tubulin (or tubulin incorporation into microtubules),  $\beta$ -tubulin orphans would sequester the assembly factor and lead to mRNA decay. In this way, the production of orphan  $\beta$ -tubulin would be tightly restricted via a co-translational detection mechanism.

The second example concerns the consequences of failed co-translational recognition by SRP. It has been observed that knockdown of SRP or mutating a signal peptide such that it cannot be recognized by SRP causes a reduction in the corresponding mRNA and protein [Karamyshev et al., 2014]. This effect was dependent on Argonaute2, which was also observed to interact with the SRP-deficient nascent chain at the ribosome [Karamyshev et al., 2014]. These findings suggest a model in which co-translational failure of a hydrophobic sequence to be recognized by SRP triggers Argonaute2-dependent mRNA decay.

These two examples illustrate a potentially important principle (Figure 5). If an interaction critical for biogenesis occurs co-translationally, its failure would necessarily result in an orphan. There may be mechanisms to detect such failures co-translationally, which would provide the cellular quality control systems access to both the nascent polypeptide and associated mRNA. Although still poorly studied, there is evidence that many protein complexes may initiate assembly during translation [Duncan and Mata, 2011; Williams and Dichtl, 2018]. If such co-translational interactions were coupled to translation elongation, the mRNA decay pathways downstream of ribosome stalling could be exploited to fine-tune the balanced expression of subunits and minimize the generation



**Figure 5. Hypothetical Model for Co-translational Monitoring of Protein Assembly**

A nascent polypeptide emerging from the ribosome is recognized by an assembly factor, which is recycled when assembly occurs correctly. When assembly fails, the assembly factor is titrated by excess unassembled orphans (right side). The unavailability of this nascent chain binding protein at the ribosome is proposed to trigger ribosome-associated quality control pathway via effects on translation (red arrow), perhaps via factors (not shown) that bind to the sequence motif normally recognized by the assembly factor.

of orphans. Such an early-detection system may represent one of the physiologic roles of ribosome-associated quality control pathways. Intriguingly, up to ~12%–15% of nascent proteins might be ubiquitinated co-translationally at the ribosome for reasons that remain to be investigated (Duttler et al., 2013; Wang et al., 2013).

### Orphan-Related Pathologies

As with other types of quality control, such as protein misfolding or processing, excessive production of orphans can be dominantly detrimental to cellular homeostasis. This is observed in the disease phenotypes of various inherited mutations. Rare mutations have been described in the signal sequences of a wide range of proteins (Cassanelli et al., 1998; Karaplis et al., 1995; Seppen et al., 1996). In most cases, the mutation disrupts the hydrophobic core of the signal to reduce or eliminate its targeting function. This would lead to an increase of orphans in the cytosol of cells expressing the mutant protein. Accordingly, the phenotypes are typically dominant and tissue-specific to the most highly expressing cell type. This suggests that the increase in orphans, not simply a loss of function, is the basis of cellular damage. In mouse studies of the non-essential prion protein (PrP), introduction of a version containing a signal sequence that is only ~50% efficient led to a dominant neurodegenerative phenotype (Rane et al., 2008). Thus, over long times in a complex organism, even a modest increase in the load of mislocalized orphans in the cytosol can be detrimental.

Perhaps the most common set of diseases involving orphans are the Thalassemias, a group of hereditary blood diseases characterized by anemia, hemolysis, and various downstream consequences (Olivieri, 1999; Piel and Weatherall, 2014). These diseases result when one or more of the alleles encoding  $\alpha$ - or  $\beta$ -globin is deficient, resulting in imbalanced synthesis of hemoglobin subunits. Mutations in the  $\alpha$ - $\beta$  interface that impair assembly of hemoglobin also cause rare variants of anemia related to the Thalassemias (Clarke and Higgins, 2000; Kohne, 2011). While a reduced level of mature hemoglobin is certainly a major contributor to the pathogenesis of these diseases, substantial evidence indicates that the unpaired globin proteins are dominantly toxic in many ways. For example, unpaired  $\alpha$ - or  $\beta$ -globin has been suggested to generate increased reactive oxygen species, form intracellular aggregates, or interfere with other cellular

processes such as assembly of a membrane cytoskeleton (Weatherall and Clegg, 2001). As noted above, unassembled  $\alpha$ -globin is selectively recognized, ubiquitinated, and degraded by UBE2O. However, this does not seem to be the only degradation pathway, as some  $\alpha$ -globin ubiquitination was observed even in UBE2O knockout cells (Nguyen et al., 2017). Furthermore, the pathway of unpaired  $\beta$ -globin degradation remains unknown. These may prove to be important facets of the pathophysiology of the Thalassemias.

Given the detrimental consequences of orphans, it is noteworthy that aneuploidy is a prominent feature in a large proportion of cancers (Rajagopalan and Lengauer, 2004). As the genes for different subunits of multi-protein complexes are often on different chromosomes, aneuploidy would necessarily unbalance expression. Experiments in yeast have shown that duplication of a single chromosome in an otherwise haploid cell is detrimental, while an additional chromosome in a diploid cell is less so (Dephoure et al., 2014; Dodgson et al., 2016; Oromendia et al., 2012). This suggests that proteome imbalance may underlie the fitness cost, an idea consistent with the observed protein homeostasis defects seen in aneuploid yeast. Experiments in human cells induced to be aneuploid support the findings in yeast (Ben-David et al., 2014; Stingle et al., 2012; Williams et al., 2008), yet aneuploid cancer cells paradoxically grow unrestrained. This might imply that they manage to grow robustly despite aneuploidy, not because of it. Consistent with this idea, many cancer cells have higher levels of the major cytosolic chaperone Hsp70 (Murphy, 2013), elevated proteasome activity (Chen and Madura, 2005), and increased autophagy (Singh et al., 2018). It is also noteworthy that UBE2O is amplified in many cancers, while reduction of UBE2O in different mouse cancer models provides a benefit by attenuating tumor growth (Liang et al., 2017; Vila et al., 2017). Thus, understanding the mechanistic basis of orphan recognition and degradation may provide new therapeutic opportunities in a number of common diseases ranging from Thalassemias to cancers.

### ACKNOWLEDGMENTS

We thank Hegde lab members for useful discussions. Work in the authors' lab is supported by the UK Medical Research Council (MC\_UP\_A022\_1007 to R.S.H.).

## REFERENCES

- Abovich, N., Gritz, L., Tung, L., and Rosbash, M. (1985). Effect of RP51 gene dosage alterations on ribosome synthesis in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 5, 3429–3435.
- Ames, B.N., and Martin, R.G. (1964). Biochemical Aspects of Genetics: The Operon. *Annu. Rev. Biochem.* 33, 235–258.
- Babu, M., Vlasblom, J., Pu, S., Guo, X., Graham, C., Bean, B.D.M., Burston, H.E., Vizeacumar, F.J., Snider, J., Phanse, S., et al. (2012). Interaction landscape of membrane-protein complexes in *Saccharomyces cerevisiae*. *Nature* 489, 585–589.
- Bachurski, C.J., Theodorakis, N.G., Coulson, R.M., and Cleveland, D.W. (1994). An amino-terminal tetrapeptide specifies cotranslational degradation of beta-tubulin but not alpha-tubulin mRNAs. *Mol. Cell. Biol.* 14, 4076–4086.
- Baumal, R., Potter, M., and Scharff, M.D. (1971). Synthesis, assembly, and secretion of gamma globulin by mouse myeloma cells. 3. Assembly of the three subclasses of IgG. *J. Exp. Med.* 134, 1316–1334.
- Ben-David, U., Arad, G., Weissbein, U., Mandefro, B., Maimon, A., Golan-Lev, T., Narwani, K., Clark, A.T., Andrews, P.W., Benvenisty, N., and Carlos Biancotti, J. (2014). Aneuploidy induces profound changes in gene expression, proliferation and tumorigenicity of human pluripotent stem cells. *Nat. Commun.* 5, 4825.
- Benz, E.J., Jr., and Forget, B.G. (1974). The biosynthesis of hemoglobin. *Semin. Hematol.* 11, 463–523.
- Berndt, U., Oellerer, S., Zhang, Y., Johnson, A.E., and Rospert, S. (2009). A signal-anchor sequence stimulates signal recognition particle binding to ribosomes from inside the exit tunnel. *Proc. Natl. Acad. Sci. USA* 106, 1398–1403.
- Besemer, J., Harant, H., Wang, S., Oberhauser, B., Marquardt, K., Foster, C.A., Schreiner, E.P., de Vries, J.E., Dascher-Nadel, C., and Lindley, I.J.D. (2005). Selective inhibition of cotranslational translocation of vascular cell adhesion molecule 1. *Nature* 436, 290–293.
- Blöbel, G. (1980). Intracellular protein topogenesis. *Proc. Natl. Acad. Sci. USA* 77, 1496–1500.
- Bole, D.G., Hendershot, L.M., and Kearney, J.F. (1986). Posttranslational association of immunoglobulin heavy chain binding protein with nascent heavy chains in nonsecreting and secreting hybridomas. *J. Cell Biol.* 102, 1558–1566.
- Bonde, M.T., Pedersen, M., Klausen, M.S., Jensen, S.I., Wulff, T., Harrison, S., Nielsen, A.T., Herrgård, M.J., and Sommer, M.O.A. (2016). Predictable tuning of protein expression in bacteria. *Nat. Methods* 13, 233–236.
- Bonifacino, J.S., and Lippincott-Schwartz, J. (1991). Degradation of proteins within the endoplasmic reticulum. *Curr. Opin. Cell Biol.* 3, 592–600.
- Bonifacino, J.S., Suzuki, C.K., Lippincott-Schwartz, J., Weissman, A.M., and Klausner, R.D. (1989). Pre-Golgi degradation of newly synthesized T-cell antigen receptor chains: intrinsic sensitivity and the role of subunit assembly. *J. Cell Biol.* 109, 73–83.
- Bonifacino, J.S., Cosson, P., Shah, N., and Klausner, R.D. (1991). Role of potentially charged transmembrane residues in targeting proteins for retention and degradation within the endoplasmic reticulum. *EMBO J.* 10, 2783–2793.
- Braunger, K., Pfeffer, S., Shrima, S., Gilmore, R., Berninghausen, O., Mandon, E.C., Becker, T., Förster, F., and Beckmann, R. (2018). Structural basis for coupling protein transport and N-glycosylation at the mammalian endoplasmic reticulum. *Science* 360, 215–219.
- Brewer, J.W., and Hendershot, L.M. (2005). Building an antibody factory: a job for the unfolded protein response. *Nat. Immunol.* 6, 23–29.
- Brodsky, J.L., Goeckeler, J., and Schekman, R. (1995). BiP and Sec63p are required for both co- and posttranslational protein translocation into the yeast endoplasmic reticulum. *Proc. Natl. Acad. Sci. USA* 92, 9643–9646.
- Caramelo, J.J., and Parodi, A.J. (2015). A sweet code for glycoprotein folding. *FEBS Lett.* 589, 3379–3387.
- Cassanelli, S., Bertolini, S., Roller, M., De Stefano, F., Casarino, L., Elicio, N., Naselli, A., and Calandra, S. (1998). A 'de novo' point mutation of the low-density lipoprotein receptor gene in an Italian subject with primary hypercholesterolemia. *Clin. Genet.* 53, 391–395.
- Chakrabarti, O., Rane, N.S., and Hegde, R.S. (2011). Cytosolic aggregates perturb the degradation of nontranslocated secretory and membrane proteins. *Mol. Biol. Cell* 22, 1625–1637.
- Chartron, J.W., Hunt, K.C.L., and Frydman, J. (2016). Cotranslational signal-independent SRP preloading during membrane targeting. *Nature* 536, 224–228.
- Chen, L., and Madura, K. (2005). Increased proteasome activity, ubiquitin-conjugating enzymes, and eEF1A translation factor detected in breast cancer tissue. *Cancer Res.* 65, 5599–5606.
- Chen, Y.-C., Umanah, G.K.E., Dephoure, N., Andrabi, S.A., Gygi, S.P., Dawson, T.M., Dawson, V.L., and Rutter, J. (2014). Msp1/ATAD1 maintains mitochondrial function by facilitating the degradation of mislocalized tail-anchored proteins. *EMBO J.* 33, 1548–1564.
- Clarke, G.M., and Higgins, T.N. (2000). Laboratory investigation of hemoglobinopathies and thalassemias: review and update. *Clin. Chem.* 46, 1284–1290.
- Claude, A. (1943). The constitution of protoplasm. *Science* 97, 451–456.
- Claude, A. (1946). Fractionation of Mammalian Liver Cells By Differential Centrifugation: II. Experimental Procedures and Results. *J. Exp. Med.* 84, 61–89.
- Cleveland, D.W., Lopata, M.A., Sherline, P., and Kirschner, M.W. (1981). Unpolymerized tubulin modulates the level of tubulin mRNAs. *Cell* 25, 537–546.
- Cosson, P., Lankford, S.P., Bonifacino, J.S., and Klausner, R.D. (1991). Membrane protein association by potential intramembrane charge pairs. *Nature* 351, 414–416.
- Costa, E.A., Subramanian, K., Nunnari, J., and Weissman, J.S. (2018). Defining the physiological role of SRP in protein-targeting efficiency and specificity. *Science* 359, 689–692.
- Dávila López, M., Martínez Guerra, J.J., and Samuelsson, T. (2010). Analysis of gene order conservation in eukaryotes identifies transcriptionally and functionally linked genes. *PLoS ONE* 5, e10654.
- De Duve, C. (1965). The separation and characterization of subcellular particles. *Harvey Lect.* 59, 49–87.
- Dephoure, N., Hwang, S., O'Sullivan, C., Dodgson, S.E., Gygi, S.P., Amon, A., and Torres, E.M. (2014). Quantitative proteomic analysis reveals posttranslational responses to aneuploidy in yeast. *eLife* 3, e03023.
- Dodgson, S.E., Kim, S., Costanzo, M., Baryshnikova, A., Morse, D.L., Kaiser, C.A., Boone, C., and Amon, A. (2016). Chromosome-specific and global effects of aneuploidy in *Saccharomyces cerevisiae*. *Genetics* 202, 1395–1409.
- Driscaldi, B., Stewart, R.S., Adles, C., Stewart, L.R., Quaglio, E., Biasini, E., Fioriti, L., Chiesa, R., and Harris, D.A. (2003). Mutant PrP is delayed in its exit from the endoplasmic reticulum, but neither wild-type nor mutant PrP undergoes retrotranslocation prior to proteasomal degradation. *J. Biol. Chem.* 278, 21732–21743.
- Duncan, C.D.S., and Mata, J. (2011). Widespread cotranslational formation of protein complexes. *PLoS Genet.* 7, e1002398.
- Duttler, S., Pechmann, S., and Frydman, J. (2013). Principles of cotranslational ubiquitination and quality control at the ribosome. *Mol. Cell* 50, 379–393.
- Emanuelsson, O., Brunak, S., von Heijne, G., and Nielsen, H. (2007). Locating proteins in the cell using TargetP, SignalP and related tools. *Nat. Protoc.* 2, 953–971.
- Feige, M.J., and Hendershot, L.M. (2013). Quality control of integral membrane proteins by assembly-dependent membrane integration. *Mol. Cell* 51, 297–309.
- Feige, M.J., Groscurth, S., Marcinowski, M., Shimizu, Y., Kessler, H., Hendershot, L.M., and Buchner, J. (2009). An unfolded CH1 domain controls the assembly and secretion of IgG antibodies. *Mol. Cell* 34, 569–579.
- Feit, H., Slusarek, L., and Shelanski, M.L. (1971). Heterogeneity of tubulin subunits. *Proc. Natl. Acad. Sci. USA* 68, 2028–2031.

- Feng, L., Gell, D.A., Zhou, S., Gu, L., Kong, Y., Li, J., Hu, M., Yan, N., Lee, C., Rich, A.M., et al. (2004). Molecular mechanism of AHSP-mediated stabilization of  $\alpha$ -hemoglobin. *Cell* 119, 629–640.
- Feng, L., Zhou, S., Gu, L., Gell, D.A., Mackay, J.P., Weiss, M.J., Gow, A.J., and Shi, Y. (2005). Structure of oxidized  $\alpha$ -haemoglobin bound to AHSP reveals a protective mechanism for haem. *Nature* 435, 697–701.
- Fons, R.D., Bogert, B.A., and Hegde, R.S. (2003). Substrate-specific function of the translocon-associated protein complex during translocation across the ER membrane. *J. Cell Biol.* 160, 529–539.
- Gamerding, M., Hanebuth, M.A., Frickey, T., and Deuerling, E. (2015). The principle of antagonism ensures protein targeting specificity at the endoplasmic reticulum. *Science* 348, 201–207.
- Garrison, J.L., Kunkel, E.J., Hegde, R.S., and Taunton, J. (2005). A substrate-specific inhibitor of protein translocation into the endoplasmic reticulum. *Nature* 436, 285–289.
- Gavin, A.C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dümpelfeld, B., et al. (2006). Proteome survey reveals modularity of the yeast cell machinery. *Nature* 440, 631–636.
- Gay, D.A., Sisodia, S.S., and Cleveland, D.W. (1989). Autoregulatory control of beta-tubulin mRNA stability is linked to translation elongation. *Proc. Natl. Acad. Sci. USA* 86, 5763–5767.
- Guna, A., and Hegde, R.S. (2018). Transmembrane domain recognition during membrane protein biogenesis and quality control. *Curr. Biol.* 28, R498–R511.
- Haas, I.G., and Wabl, M. (1983). Immunoglobulin heavy chain binding protein. *Nature* 306, 387–389.
- Halic, M., Becker, T., Pool, M.R., Spahn, C.M.T., Grassucci, R.A., Frank, J., and Beckmann, R. (2004). Structure of the signal recognition particle interacting with the elongation-arrested ribosome. *Nature* 427, 808–814.
- Hammond, C.M., Strømme, C.B., Huang, H., Patel, D.J., and Groth, A. (2017). Histone chaperone networks shaping chromatin function. *Nat. Rev. Mol. Cell Biol.* 18, 141–158.
- Harbauer, A.B., Opalińska, M., Gerbeth, C., Herman, J.S., Rao, S., Schönfisch, B., Guiard, B., Schmidt, O., Pfanner, N., and Meisinger, C. (2014). Mitochondria. Cell cycle-dependent regulation of mitochondrial preprotein translocase. *Science* 346, 1109–1113.
- Harper, J.W., and Bennett, E.J. (2016). Proteome complexity and the forces that drive proteome imbalance. *Nature* 537, 328–338.
- Havugimana, P.C., Hart, G.T., Nepusz, T., Yang, H., Turinsky, A.L., Li, Z., Wang, P.I., Boutz, D.R., Fong, V., Phanse, S., et al. (2012). A census of human soluble protein complexes. *Cell* 150, 1068–1081.
- Hegde, R.S., and Keenan, R.J. (2011). Tail-anchored membrane protein insertion into the endoplasmic reticulum. *Nat. Rev. Mol. Cell Biol.* 12, 787–798.
- Hendershot, L., Bole, D., Köhler, G., and Kearney, J.F. (1987). Assembly and secretion of heavy chains that do not associate posttranslationally with immunoglobulin heavy chain-binding protein. *J. Cell Biol.* 104, 761–767.
- Hentze, M.W., Castello, A., Schwarzl, T., and Preiss, T. (2018). A brave new world of RNA-binding proteins. *Nat. Rev. Mol. Cell Biol.* 19, 327–341.
- Hessa, T., Sharma, A., Mariappan, M., Eshleman, H.D., Gutierrez, E., and Hegde, R.S. (2011). Protein targeting and degradation are coupled for elimination of mislocalized proteins. *Nature* 475, 394–397.
- Horie, C., Suzuki, H., Sakaguchi, M., and Mihara, K. (2002). Characterization of signal that directs C-tail-anchored proteins to mammalian mitochondrial outer membrane. *Mol. Biol. Cell* 13, 1615–1625.
- Hosokawa, N., Wada, I., Hasegawa, K., Yorihozi, T., Tremblay, L.O., Herscovics, A., and Nagata, K. (2001). A novel ER alpha-mannosidase-like protein accelerates ER-associated degradation. *EMBO Rep.* 2, 415–422.
- Huber, R., Deisenhofer, J., Colman, P.M., Matsushima, M., and Palm, W. (1976). Crystallographic structure studies of an IgG molecule and an Fc fragment. *Nature* 264, 415–420.
- Huppa, J.B., and Ploegh, H.L. (1997). The alpha chain of the T cell antigen receptor is degraded in the cytosol. *Immunity* 7, 113–122.
- Itakura, E., Zavodszky, E., Shao, S., Wohlever, M.L., Keenan, R.J., and Hegde, R.S. (2016). Ubiquitins chaperone and triage mitochondrial membrane proteins for degradation. *Mol. Cell* 63, 21–33.
- Itzhak, D.N., Tyanova, S., Cox, J., and Borner, G.H. (2016). Global, quantitative and dynamic mapping of protein subcellular localization. *eLife* 5, <https://doi.org/10.7554/eLife.16950>.
- Jacob, H.S., Brain, M.C., Dacie, J.V., Carrell, R.W., and Lehmann, H. (1968). Abnormal haem binding and globin SH group blockade in unstable haemoglobins. *Nature* 218, 1214–1217.
- Jäkel, S., and Görlich, D. (1998). Importin beta, transportin, RanBP5 and RanBP7 mediate nuclear import of ribosomal proteins in mammalian cells. *EMBO J.* 17, 4491–4502.
- Jäkel, S., Mingot, J.-M., Schwarzmaier, P., Hartmann, E., and Görlich, D. (2002). Importins fulfil a dual function as nuclear import receptors and cytoplasmic chaperones for exposed basic domains. *EMBO J.* 21, 377–386.
- Jonikas, M.C., Collins, S.R., Denic, V., Oh, E., Quan, E.M., Schmid, V., Weibezahn, J., Schwappach, B., Walter, P., Weissman, J.S., and Schuldiner, M. (2009). Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum. *Science* 323, 1693–1697.
- Kang, S.-W., Rane, N.S., Kim, S.J., Garrison, J.L., Taunton, J., and Hegde, R.S. (2006). Substrate-specific translocational attenuation during ER stress defines a pre-emptive quality control pathway. *Cell* 127, 999–1013.
- Karamyshev, A.L., Patrick, A.E., Karamysheva, Z.N., Griesemer, D.S., Hudson, H., Tjon-Kon-Sang, S., Nilsson, I., Otto, H., Liu, Q., Rospert, S., et al. (2014). Inefficient SRP interaction with a nascent chain triggers a mRNA quality control pathway. *Cell* 156, 146–157.
- Karaplis, A.C., Lim, S.K., Baba, H., Arnold, A., and Kronenberg, H.M. (1995). Inefficient membrane targeting, translocation, and proteolytic processing by signal peptidase of a mutant preproparathyroid hormone protein. *J. Biol. Chem.* 270, 1629–1635.
- Keenan, R.J., Freymann, D.M., Stroud, R.M., and Walter, P. (2001). The signal recognition particle. *Annu. Rev. Biochem.* 70, 755–775.
- Kihm, A.J., Kong, Y., Hong, W., Russell, J.E., Rouda, S., Adachi, K., Simon, M.C., Blobel, G.A., and Weiss, M.J. (2002). An abundant erythroid protein that stabilizes free alpha-haemoglobin. *Nature* 417, 758–763.
- Kikkert, M., Doolman, R., Dai, M., Avner, R., Hassink, G., van Voorden, S., Thandavar, S., Roitelman, J., Chau, V., and Wiertz, E. (2004). Human HRD1 is an E3 ubiquitin ligase involved in degradation of proteins from the endoplasmic reticulum. *J. Biol. Chem.* 279, 3525–3534.
- Kim, S.J., Mitra, D., Salerno, J.R., and Hegde, R.S. (2002). Signal sequences control gating of the protein translocation channel in a substrate-specific manner. *Dev. Cell* 2, 207–217.
- Kohne, E. (2011). Hemoglobinopathies: clinical manifestations, diagnosis, and treatment. *Dtsch. Arztebl. Int.* 108, 532–540.
- Koren, I., Timms, R.T., Kula, T., Xu, Q., Li, M.Z., and Elledge, S.J. (2018). The Eukaryotic Proteome Is Shaped by E3 Ubiquitin Ligases Targeting C-Terminal Degrons. *Cell* 173, 1622–1635.e14.
- Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P., et al. (2006). Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* 440, 637–643.
- Kulak, N.A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014). Minimal, encapsulated proteomic sample processing applied to copy-number estimation in eukaryotic cells. *Nat. Methods* 11, 319–324.
- Lakkaraju, A.K., Abrami, L., Lemmin, T., Blaskovic, S., Kunz, B., Kihara, A., Dal Peraro, M., and van der Goot, F.G. (2012). Palmitoylated calnexin is a key component of the ribosome-translocon complex. *EMBO J.* 31, 1823–1835.
- Lee, J.W., Zemojtel, T., and Shakhnovich, E. (2009). Systems-level evidence of transcriptional co-regulation of yeast protein complexes. *J. Comput. Biol.* 16, 331–339.
- Levine, C.G., Mitra, D., Sharma, A., Smith, C.L., and Hegde, R.S. (2005). The efficiency of protein compartmentalization into the secretory pathway. *Mol. Biol. Cell* 16, 279–291.



- Li, G.-W., Burkhardt, D., Gross, C., and Weissman, J.S. (2014). Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell* 157, 624–635.
- Liang, K., Volk, A.G., Haug, J.S., Marshall, S.A., Woodfin, A.R., Bartom, E.T., Gilmore, J.M., Florens, L., Washburn, M.P., Sullivan, K.D., et al. (2017). Therapeutic Targeting of MLL Degradation Pathways in MLL-Rearranged Leukemia. *Cell* 168, 59–72.e13.
- Lim, H.N., Lee, Y., and Hussein, R. (2011). Fundamental relationship between operon organization and gene expression. *Proc. Natl. Acad. Sci. USA* 108, 10626–10631.
- Lin, H.-C., Yeh, C.-W., Chen, Y.-F., Lee, T.-T., Hsieh, P.-Y., Rusnac, D.V., Lin, S.-Y., Elledge, S.J., Zheng, N., and Yen, H.S. (2018). C-Terminal End-Directed Protein Elimination by CRL2 Ubiquitin Ligases. *Mol. Cell* 70, 602–613.e3.
- Lippincott-Schwartz, J., Bonifacino, J.S., Yuan, L.C., and Klausner, R.D. (1988). Degradation from the endoplasmic reticulum: disposing of newly synthesized proteins. *Cell* 54, 209–220.
- Ma, J., and Lindquist, S. (2001). Wild-type PrP and a mutant associated with prion disease are subject to retrograde transport and proteasome degradation. *Proc. Natl. Acad. Sci. USA* 98, 14955–14960.
- Mancini, R., Fagioli, C., Fra, A.M., Maggioni, C., and Sitia, R. (2000). Degradation of unassembled soluble Ig subunits by cytosolic proteasomes: evidence that retrotranslocation and degradation are coupled events. *FASEB J.* 14, 769–778.
- Manolios, N., Bonifacino, J.S., and Klausner, R.D. (1990). Transmembrane helical interactions and the assembly of the T cell receptor complex. *Science* 249, 274–277.
- Marsh, J.A., Hernández, H., Hall, Z., Ahnert, S.E., Perica, T., Robinson, C.V., and Teichmann, S.A. (2013). Protein complexes are under evolutionary selection to assemble via ordered pathways. *Cell* 153, 461–470.
- Matlack, K.E., Misselwitz, B., Plath, K., and Rapoport, T.A. (1999). BiP acts as a molecular ratchet during posttranslational transport of prepro- $\alpha$  factor across the ER membrane. *Cell* 97, 553–564.
- Mayor, T., Lipford, J.R., Graumann, J., Smith, G.T., and Deshaies, R.J. (2005). Analysis of polyubiquitin conjugates reveals that the Rpn10 substrate receptor contributes to the turnover of multiple proteasome targets. *Mol. Cell. Proteomics* 4, 741–751.
- Mayor, T., Graumann, J., Bryan, J., MacCoss, M.J., and Deshaies, R.J. (2007). Quantitative profiling of ubiquitylated proteins reveals proteasome substrates and the substrate repertoire influenced by the Rpn10 receptor pathway. *Mol. Cell. Proteomics* 6, 1885–1895.
- McShane, E., Sin, C., Zuber, H., Wells, J.N., Donnelly, N., Wang, X., Hou, J., Chen, W., Storchova, Z., Marsh, J.A., et al. (2016). Kinetic analysis of protein stability reveals age-dependent degradation. *Cell* 167, 803–815.e21.
- Meyuhas, O., and Kahan, T. (2015). The race to decipher the top secrets of TOP mRNAs. *Biochim. Biophys. Acta* 1849, 801–811.
- Minami, R., Hayakawa, A., Kagawa, H., Yanagi, Y., Yokosawa, H., and Kawahara, H. (2010). BAG-6 is essential for selective elimination of defective proteasomal substrates. *J. Cell Biol.* 190, 637–650.
- Miyazaki, Y., Chen, L.C., Chu, B.W., Swigut, T., and Wandless, T.J. (2015). Distinct transcriptional responses elicited by unfolded nuclear or cytoplasmic protein in mammalian cells. *eLife* 4, <https://doi.org/10.7554/eLife.07687>.
- Mock, J.-Y., Chartron, J.W., Zaslaver, M., Xu, Y., Ye, Y., and Clemons, W.M., Jr. (2015). Bag6 complex contains a minimal tail-anchor-targeting module and a mock BAG domain. *Proc. Natl. Acad. Sci. USA* 112, 106–111.
- Mulvey, C.M., Breckels, L.M., Geladaki, A., Britovšek, N.K., Nightingale, D.J.H., Christoforou, A., Elzek, M., Deery, M.J., Gatto, L., and Lilley, K.S. (2017). Using hyperLOPIT to perform high-resolution mapping of the spatial proteome. *Nat. Protoc.* 12, 1110–1135.
- Munsky, B., Neuert, G., and van Oudenaarden, A. (2012). Using gene expression noise to understand gene regulation. *Science* 336, 183–187.
- Murphy, M.E. (2013). The HSP70 family and cancer. *Carcinogenesis* 34, 1181–1188.
- Nargund, A.M., Pellegrino, M.W., Fiorese, C.J., Baker, B.M., and Haynes, C.M. (2012). Mitochondrial import efficiency of ATFS-1 regulates mitochondrial UPR activation. *Science* 337, 587–590.
- Nelson, H.C. (1995). Structure and function of DNA-binding proteins. *Curr. Opin. Genet. Dev.* 5, 180–189.
- Ng, D.T., Brown, J.D., and Walter, P. (1996). Signal sequences specify the targeting route to the endoplasmic reticulum membrane. *J. Cell Biol.* 134, 269–278.
- Nguyen, A.T., Prado, M.A., Schmidt, P.J., Sendamarai, A.K., Wilson-Grady, J.T., Min, M., Campagna, D.R., Tian, G., Shi, Y., Dederer, V., et al. (2017). UBE2O remodels the proteome during terminal erythroid differentiation. *Science* 357, eaan0218.
- Okreglak, V., and Walter, P. (2014). The conserved AAA-ATPase Msp1 confers organelle specificity to tail-anchored proteins. *Proc. Natl. Acad. Sci. USA* 111, 8019–8024.
- Olivieri, N.F. (1999). The  $\beta$ -thalassemias. *N. Engl. J. Med.* 341, 99–109.
- Oromendia, A.B., Dodgson, S.E., and Amon, A. (2012). Aneuploidy causes proteotoxic stress in yeast. *Genes Dev.* 26, 2696–2708.
- Palade, G.E. (1955). A small particulate component of the cytoplasm. *J. Biophys. Biochem. Cytol.* 1, 59–68.
- Park, S.-H., Bolender, N., Eisele, F., Kostova, Z., Takeuchi, J., Coffino, P., and Wolf, D.H. (2007). The cytoplasmic Hsp70 chaperone machinery subjects misfolded and endoplasmic reticulum import-incompetent proteins to degradation via the ubiquitin-proteasome system. *Mol. Biol. Cell* 18, 153–165.
- Peña, C., Hurt, E., and Panse, V.G. (2017). Eukaryotic ribosome assembly, transport and quality control. *Nat. Struct. Mol. Biol.* 24, 689–699.
- Perutz, M.F., Rossmann, M.G., Cullis, A.F., Muirhead, H., Will, G., and North, A.C.T. (1960). Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5-Å resolution, obtained by X-ray analysis. *Nature* 185, 416–422.
- Pfeffer, S., Dudek, J., Zimmermann, R., and Förster, F. (2016). Organization of the native ribosome-translocon complex at the mammalian endoplasmic reticulum membrane. *Biochim. Biophys. Acta* 1860, 2122–2129.
- Piel, F.B., and Weatherall, D.J. (2014). The  $\alpha$ -thalassemias. *N. Engl. J. Med.* 371, 1908–1916.
- Pittenger, M.F., and Cleveland, D.W. (1985). Retention of autoregulatory control of tubulin synthesis in cytoplasts: demonstration of a cytoplasmic mechanism that regulates the level of tubulin expression. *J. Cell Biol.* 101, 1941–1952.
- Porter, K.R. (1955–1956). The submicroscopic morphology of protoplasm. *Harvey Lect.* 51, 175–228.
- Rajagopalan, H., and Lengauer, C. (2004). Aneuploidy and cancer. *Nature* 432, 338–341.
- Rane, N.S., Yonkovich, J.L., and Hegde, R.S. (2004). Protection from cytosolic prion protein toxicity by modulation of protein translocation. *EMBO J.* 23, 4550–4559.
- Rane, N.S., Kang, S.-W., Chakrabarti, O., Feigenbaum, L., and Hegde, R.S. (2008). Reduced translocation of nascent prion protein during ER stress contributes to neurodegeneration. *Dev. Cell* 15, 359–370.
- Rane, N.S., Chakrabarti, O., Feigenbaum, L., and Hegde, R.S. (2010). Signal sequence insufficiency contributes to neurodegeneration caused by transmembrane prion protein. *J. Cell Biol.* 188, 515–526.
- Rapoport, T.A., Li, L., and Park, E. (2017). Structural and mechanistic insights into protein translocation. *Annu. Rev. Cell Dev. Biol.* 33, 369–390.
- Rieder, R.F. (1974). Human hemoglobin stability and instability: molecular mechanisms and some clinical correlations. *Semin. Hematol.* 11, 423–440.
- Rodrigo-Brenni, M.C., Gutierrez, E., and Hegde, R.S. (2014). Cytosolic quality control of mislocalized proteins requires RNF126 recruitment to Bag6. *Mol. Cell* 55, 227–237.

- Sato, B.K., Schulz, D., Do, P.H., and Hampton, R.Y. (2009). Misfolded membrane proteins are specifically recognized by the transmembrane domain of the Hrd1p ubiquitin ligase. *Mol. Cell* 34, 212–222.
- Sauer, R.T., and Baker, T.A. (2011). AAA+ proteases: ATP-fueled machines of protein destruction. *Annu. Rev. Biochem.* 80, 587–612.
- Scazzari, M., Amm, I., and Wolf, D.H. (2015). Quality control of a cytoplasmic protein complex: chaperone motors and the ubiquitin-proteasome system govern the fate of orphan fatty acid synthase subunit Fas2 of yeast. *J. Biol. Chem.* 290, 4677–4687.
- Schmidt, O., Harbauer, A.B., Rao, S., Eyrich, B., Zahedi, R.P., Stojanovski, D., Schönfisch, B., Guiard, B., Sickmann, A., Pfanner, N., and Meisinger, C. (2011). Regulation of mitochondrial protein import by cytosolic kinases. *Cell* 144, 227–239.
- Schoebel, S., Mi, W., Stein, A., Ovchinnikov, S., Pavlovicz, R., DiMaio, F., Baker, D., Chambers, M.G., Su, H., Li, D., et al. (2017). Cryo-EM structure of the protein-conducting ERAD channel Hrd1 in complex with Hrd3. *Nature* 548, 352–355.
- Schuldiner, M., Metz, J., Schmid, V., Denic, V., Rakwalska, M., Schmitt, H.D., Schwappach, B., and Weissman, J.S. (2008). The GET complex mediates insertion of tail-anchored proteins into the ER membrane. *Cell* 134, 634–645.
- Seppen, J., Steenken, E., Lindhout, D., Bosma, P.J., and Elferink, R.P. (1996). A mutation which disrupts the hydrophobic core of the signal peptide of bilirubin UDP-glucuronosyltransferase, an endoplasmic reticulum membrane protein, causes Crigler-Najjar type II. *FEBS Lett.* 390, 294–298.
- Shao, S., and Hegde, R.S. (2011). Membrane protein insertion at the endoplasmic reticulum. *Annu. Rev. Cell Dev. Biol.* 27, 25–56.
- Shao, S., Rodrigo-Brenni, M.C., Kivlen, M.H., and Hegde, R.S. (2017). Mechanistic basis for a molecular triage reaction. *Science* 355, 298–302.
- Shemorry, A., Hwang, C.S., and Varshavsky, A. (2013). Control of protein quality and stoichiometries by N-terminal acetylation and the N-end rule pathway. *Mol. Cell* 50, 540–551.
- Shen, Y., and Hendershot, L.M. (2005). ERdj3, a stress-inducible endoplasmic reticulum DnaJ homologue, serves as a cofactor for BiP's interactions with unfolded substrates. *Mol. Biol. Cell* 16, 40–50.
- Shoemaker, C.J., and Green, R. (2012). Translation drives mRNA quality control. *Nat. Struct. Mol. Biol.* 19, 594–601.
- Singh, S.S., Vats, S., Chia, A.Y.-Q., Tan, T.Z., Deng, S., Ong, M.S., Arfuso, F., Yap, C.T., Goh, B.C., Sethi, G., et al. (2018). Dual role of autophagy in hallmarks of cancer. *Oncogene* 37, 1142–1158.
- Stingle, S., Stoehr, G., Peplowska, K., Cox, J., Mann, M., and Storchova, Z. (2012). Global analysis of genome, transcriptome and proteome reveals the response to aneuploidy in human cells. *Mol. Syst. Biol.* 8, 608.
- Sung, M.-K., Porras-Yakushi, T.R., Reitsma, J.M., Huber, F.M., Sweredoski, M.J., Hoelz, A., Hess, S., and Deshaies, R.J. (2016). A conserved quality-control pathway that mediates degradation of unassembled ribosomal proteins. *eLife* 5, <https://doi.org/10.7554/eLife.19105>.
- Suzuki, R., and Kawahara, H. (2016). UBQLN4 recognizes mislocalized transmembrane domain proteins and targets these to proteasomal degradation. *EMBO Rep.* 17, 842–857.
- Tannous, A., Pisoni, G.B., Hebert, D.N., and Molinari, M. (2015). N-linked sugar-regulated protein folding and quality control in the ER. *Semin. Cell Dev. Biol.* 41, 79–89.
- Tasaki, T., Sriram, S.M., Park, K.S., and Kwon, Y.T. (2012). The N-end rule pathway. *Annu. Rev. Biochem.* 81, 261–289.
- Theodorakis, N.G., and Cleveland, D.W. (1992). Physical evidence for cotranslational regulation of beta-tubulin mRNA degradation. *Mol. Cell. Biol.* 12, 791–799.
- UniProt Consortium (2018). UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* 46, 2699–2699.
- Vanhove, M., Usherwood, Y.K., and Hendershot, L.M. (2001). Unassembled Ig heavy chains do not cycle from BiP in vivo but require light chains to trigger their release. *Immunity* 15, 105–114.
- Vembar, S.S., and Brodsky, J.L. (2008). One step at a time: endoplasmic reticulum-associated degradation. *Nat. Rev. Mol. Cell Biol.* 9, 944–957.
- Vila, I.K., Yao, Y., Kim, G., Xia, W., Kim, H., Kim, S.J., Park, M.K., Hwang, J.P., González-Billalabeitia, E., Hung, M.C., et al. (2017). A UBE2O-AMPK $\alpha$ 2 axis that promotes tumor initiation and progression offers opportunities for therapy. *Cancer Cell* 31, 208–224.
- Vitali, D.G., Sinzel, M., Bulthuis, E.P., Kolb, A., Zabel, S., Mehlhorn, D.G., Figueiredo Costa, B., Farkas, A., Clancy, A., Schuldiner, M., et al. (2018). The GET pathway can increase the risk of mitochondrial outer membrane proteins to be mistargeted to the ER. *J. Cell Sci.* 131, jcs211110.
- Voigt, S., Jungnickel, B., Hartmann, E., and Rapoport, T.A. (1996). Signal sequence-dependent function of the TRAM protein during early phases of protein transport across the endoplasmic reticulum membrane. *J. Cell Biol.* 134, 25–35.
- von Heijne, G. (1995). Protein sorting signals: simple peptides with complex functions. *EXS* 73, 67–76.
- Voorhees, R.M., and Hegde, R.S. (2015). Structures of the scanning and engaged states of the mammalian SRP-ribosome complex. *eLife* 4, <https://doi.org/10.7554/eLife.07975>.
- Wang, X., and Chen, X.J. (2015). A cytosolic network suppressing mitochondria-mediated proteostatic stress and cell death. *Nature* 524, 481–484.
- Wang, Q., Liu, Y., Soetandyo, N., Baek, K., Hegde, R., and Ye, Y. (2011). A ubiquitin ligase-associated chaperone holdase maintains polypeptides in soluble states for proteasome degradation. *Mol. Cell* 42, 758–770.
- Wang, F., Durfee, L.A., and Hübregtse, J.M. (2013). A cotranslational ubiquitination pathway for quality control of misfolded proteins. *Mol. Cell* 50, 368–378.
- Warner, J.R., Mitra, G., Schwindinger, W.F., Studeny, M., and Fried, H.M. (1985). *Saccharomyces cerevisiae* coordinates accumulation of yeast ribosomal proteins by modulating mRNA splicing, translational initiation, and protein turnover. *Mol. Cell. Biol.* 5, 1512–1521.
- Wattenberg, B.W., Clark, D., and Brock, S. (2007). An artificial mitochondrial tail signal/anchor sequence confirms a requirement for moderate hydrophobicity for targeting. *Biosci. Rep.* 27, 385–401.
- Weatherall, D.J., and Clegg, J.B. (2001). The thalassaemia syndromes, 4th edition (Blackwell Publ. Ltd.), pp. 132–174.
- Wefes, I., Mastrandrea, L.D., Haldeman, M., Koury, S.T., Tamburlin, J., Pickart, C.M., and Finley, D. (1995). Induction of ubiquitin-conjugating enzymes during terminal erythroid differentiation. *Proc. Natl. Acad. Sci. USA* 92, 4982–4986.
- Weidberg, H., and Amon, A. (2018). MitoCPR-A surveillance pathway that protects mitochondria in response to protein import stress. *Science* 360, eaan4146.
- Weir, N.R., Kamber, R.A., Martenson, J.S., and Denic, V. (2017). The AAA protein Msp1 mediates clearance of excess tail-anchored proteins from the peroxisomal membrane. *eLife* 6, <https://doi.org/10.7554/eLife.28507>.
- Wickner, W., and Schekman, R. (2005). Protein translocation across biological membranes. *Science* 310, 1452–1456.
- Wiedemann, N., and Pfanner, N. (2017). Mitochondrial machineries for protein import and assembly. *Annu. Rev. Biochem.* 86, 685–714.
- Wiedmann, B., Sakai, H., Davis, T.A., and Wiedmann, M. (1994). A protein complex required for signal-sequence-specific sorting and translocation. *Nature* 370, 434–440.
- Williams, N.K., and Dichtl, B. (2018). Co-translational control of protein complex formation: a fundamental pathway of cellular organization? *Biochem. Soc. Trans.* 46, 197–206.
- Williams, B.R., Prabhu, V.R., Hunter, K.E., Glazier, C.M., Whittaker, C.A., Housman, D.E., and Amon, A. (2008). Aneuploidy affects proliferation and spontaneous immortalization in mammalian cells. *Science* 322, 703–709.

Wohlever, M.L., Mateja, A., McGilvray, P.T., Day, K.J., and Keenan, R.J. (2017). Msp1 is a membrane protein dislocase for tail-anchored proteins. *Mol. Cell* 67, 194–202.e6.

Wright, G., Terada, K., Yano, M., Sergeev, I., and Mori, M. (2001). Oxidative stress inhibits the mitochondrial import of preproteins and leads to their degradation. *Exp. Cell Res.* 263, 107–117.

Wrobel, L., Topf, U., Bragoszewski, P., Wiese, S., Sztolsztener, M.E., Oeljeklaus, S., Varabyova, A., Lirski, M., Chroscicki, P., Mroczek, S., et al. (2015).

Mistargeted mitochondrial proteins activate a proteostatic response in the cytosol. *Nature* 524, 485–488.

Xu, Y., Anderson, D.E., and Ye, Y. (2016). The HECT domain ubiquitin ligase HUWE1 targets unassembled soluble proteins for degradation. *Cell Discov.* 2, 16040.

Yanagitani, K., Juszkievicz, S., and Hegde, R.S. (2017). UBE2O is a quality control factor for orphans of multiprotein complexes. *Science* 357, 472–475.