# UCSF
## UC San Francisco Electronic Theses and Dissertations

**Title**

The role of the Ribosome Quality Control Complex (RQC) in Maintaining Protein Homeostasis in Yeast and Higher Eukaryotes

**Permalink**

https://escholarship.org/uc/item/6b38941p

**Author**

Kostova, Kamena Kamenova

**Publication Date**

2018

Peer reviewed|Thesis/dissertation

The role of the Ribosome Quality Control Complex (RQC) in Maintaining Protein Homeostasis

in Yeast and Higher Eukaryotes

by

Kamena Kamenova Kostova

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

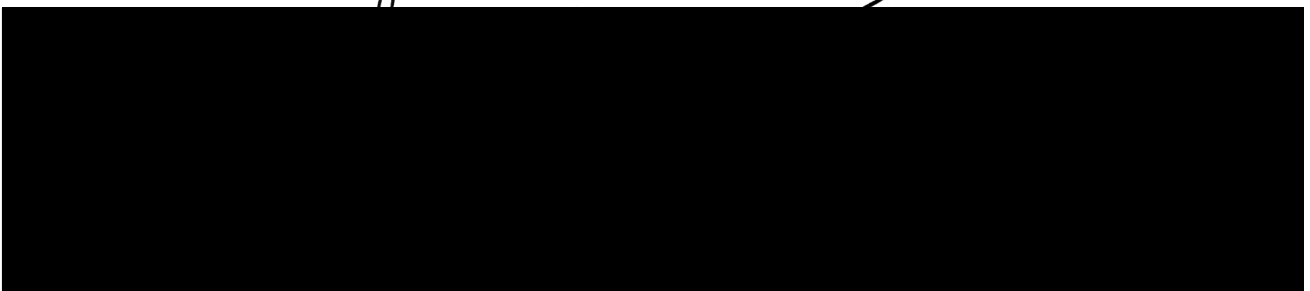DOCTOR OF PHILOSOPHY

in

Biomedical Sciences

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

Committee in Charge

ACKNOWLEDGEMENTS

This work would not have been possible without the mentorship, generosity, and encouragement of many people.

First, I want to thank my mentor, Jonathan Weissman. He helped me grow as a scientist. He allowed me the freedom to pursue my own ideas, and gave me guidance when I felt stuck. He always taught by example and showed me how to be a creative scientist, a fair collaborator, and a good labmate. He was always there to celebrate my successes with a glass of scotch and help me through my failures (with a glass of Bulgarian moonshine).

I want to thank all members of the Weissman Lab for creating a supportive and stimulating atmosphere. Every single person in the lab has contributed to my PhD work with experiments, reagents, or ideas. I feel honored to be part of such an incredible group of scientists. I want to thank Marco Jost for editing countless drafts of proposals, papers, cover letters, etc. Calvin Jan, Elizabeth Costa, and Christopher Williams taught me everything I know about yeast genetics. Jeff Hussmann has been a great mentor about anything computational. Kelsey Hickey has been a good collaborator and an RQC enthusiast.

My thesis committee members, Peter Walter, Davide Ruggero and Carol Gross, have guided me and supported me through my PhD. They always had time to talk to me and give me honest advice.

I would not be the person I am today without the unconditional love and support of my mother, Radka Kostova. She has sacrificed so much for me and I am forever grateful. She taught me the value of education and believed in me even when I did not believe in myself.

Finally, I want to thank my husband, TJ Hu. Even in my wildest dreams I never imagined that I will meet my soulmate in the mouse house of UCSF. He lightens my day with humor,

gives me a shoulder to cry on, and cheers me on. He always pushes me to give my best and keep

going even when I am ready to give up. I cannot wait to grow old with you.

The role of the Ribosome Quality Control Complex (RQC) in Maintaining Protein Homeostasis

in Yeast and Higher Eukaryotes

Kamena Kamenova Kostova

ABSTRACT

Protein biosynthesis is the most energy-consuming process during cellular proliferation

and any event that interferes with protein production jeopardizes cell viability. It has been

recently appreciated that cells have a number of surveillance mechanisms to counter the threat

posed by translation failure. One such mechanism is provided by the ribosome quality control

complex (RQC), which engages stalled ribosomes and facilitates the degradation of the partially

synthesized nascent polypeptide. Mechanistic characterization of the RQC in yeast had

demonstrated that the RQC tags partially synthesized polypeptides with C-terminal alanine and

threonine (CAT) tails in a remarkable untemplated elongation reaction. However, the biological

significance of the addition of CAT-tails had remained poorly understood. For my thesis work, I

elucidated the role of CAT-tails in yeast and explored the mechanisms of translation quality

control in higher eukaryotes.

I discovered a surprising property of the core RQC ubiquitin ligase, Ltn1p, that marks the

stalled nascent polypeptide for degradation. In contrast to most ubiquitin ligases, Ltn1p is

extremely rigid and can only access a small number of amino acids right outside the ribosome

exit tunnel. If this narrow window is depleted of lysine residues, the sites of ubiquitin addition,

CAT-tailing becomes critical for polypeptide degradation. In particular, the addition of CAT-

tails to the C-terminus of the stalled polypeptide can push out lysines sequestered in the

ribosome exit tunnel, making those residues accessible to Ltn1p. My work has established that

CAT-tailing provides a fail-safe mechanism for nascent polypeptide degradation that greatly expands the range of RQC-degradable substrates.

Although the RQC has been well characterized in yeast, the composition and function of the mammalian complex remains poorly understood. I engineered an RQC reporter and used it to perform a genome-wide CRISPR screen. This unbiased approach allowed me to gain a comprehensive view of the mammalian RQC pathway, and to identify novel RQC components. I identified a mammalian-specific complex that translationally silences faulty mRNAs by blocking ribosome initiation. This novel quality control mechanism could prevent the formation of ribosome "traffic jams" on defective messages, alleviating the stalling burden in mammalian cells.

CONTRIBUTIONS

Excluding the contributions specified here, this dissertation constitutes the work of Kamena Kostova. Portions of Chapters I, II, and IV are reproduced from a peer-reviewed publication (Kostova et al., 2017). The work on the mammalian RQC complex was carried out in a collaboration with Kelsey Hickey and Kimberley Dickson. Jonathan Weissman supervised this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

that stabilize the stalling reporter.

**CHAPTER ONE**

Introduction and review

Introduction

Protein synthesis is a complex and vital mechanism mediated by the ribosome. In eukaryotes, ribosomes start translating at mRNA start codon, AUG. During the elongation phase, the mRNA nucleotide sequence is decoded into its corresponding sequence of amino acids. This process implies the transit of appropriate tRNAs through A, P, and E ribosomal sites (A for aminoacyl, P for peptidyl, and E for exit). Translation terminates when the ribosome encounters a stop codon, at which point termination factors rather than tRNAs bind to the ribosome. In response, the polypeptide is released and the ribosome is recycled to start a new round of translation (Schmeing and Ramakrishnan, 2009).

However, if a ribosome stops translating before it reaches a stop codon, it cannot passively dissociate from the mRNA and becomes trapped. This process, known as ribosome stalling, can be detrimental for the cell for the following reasons. First, the stalled ribosome blocks further translation of the bound mRNA, decreasing protein output. Second, the incomplete polypeptide stuck on the ribosomal P site might be insoluble or exhibit disregulated activity if released. Finally, because messages are typically translated by multiple ribosomes, all ribosomes following a stalled one will also be arrested, leading to global translation inhibition.

Since protein biosynthesis is the most energy-consuming process during cellular proliferation, events that interferes with protein production, such as ribosome stalling, jeopardize cell viability. To counter these challenges, cells have evolved quality control pathways that degrade the mRNA, release the trapped ribosome, and target the nascent polypeptide for proteasomal degradation. This latter process is mediated by the Ribosome Quality control Complex (RQC) (Brandman et al., 2012; Defenouillère et al., 2013; Shao and Hegde, 2014; Shao et al., 2013; Verma et al., 2013).

## Yeast RQC complex

The RQC has been initially identified in *S. cerevisiae*. The complex is comprised of four core components – an ubiquitin ligase, Ltn1p, two poorly characterized proteins at the time, YDR333C (RQC1) and TAE2 (RQC2), as well as the AAA+ ATPase Cdc48 (Fig. 1-1). Subsequent work has shed light on the sequence of events of the RQC engagement, as well as the function of the various components. The RQC engages a 60S subunit carrying a nascent polypeptide-bound P-site transfer RNA (tRNA), after the dissociation of the 40S and the mRNA. It is believed that Rqc2p binds first the 60S subunit near the tRNA binding site, and helps recruit Ltn1p. The ubiquitin ligase loops around the 60S ribosome and positions its active site close to the ribosome exit tunnel. The binding site of Rqc1p has not been conclusively identified. It is believed that the protein binds near the ribosome exit tunnel, preventing aggregation of the partially synthesized polypeptide. Finally, ubiquitination of the nascent polypeptide leads to the recruitment of Cdc48p and its adapter. This molecular motor facilitates the extraction of the nascent polypeptide from the ribosome exit tunnel and its delivery to the proteasome, which eventually degrades the stalling polypeptide product (Brandman and Hegde, 2016a).

## Carboxy-terminal alanine and threonine (CAT) tails

Remarkably, Rqc2p is also involved in an unprecedented elongation reaction that occurs independently of both the 40S ribosomal subunit and an RNA template. Cryoelectron microscopy studied revealed that the RQC engaged 60S subunit contains not only P-site tRNA, but also A-site tRNA (Shen et al., 2015). Since the 60S is the product of a stalled ribosome, the A-site of the subunit was expected to be empty. Sequencing from the purified 60S–RQC

complexes revealed that the bound tRNAs are markedly enriched in alanyl- and threonyl-tRNAs. In addition, the molecular weight of model stalling substrates stabilized by deletion of the ubiquitin ligase Ltn1p was shown to be larger than expected from the primary sequence. Total amino acid analysis of such stalled nascent polypeptide showed overrepresentation of alanine and threonine, suggesting that these two amino acids are added to the nascent polypeptide without them being encoded in the mRNA. The observed extra molecular weight was shown to be due to amino acid extensions added to the C-terminus of the protein in an Rqc2-dependent manner. Cumulatively, these observations led to a model where Rqc2 facilitates the recruitment of charged alanine and threonine tRNAs to stalled 60S ribosomes. The amino acids that these tRNAs carry are subsequently added to the nascent polypeptide. This fascinated process of tagging the nascent chains with Carboxy-terminal Ala and Thr extensions has been named CAT-tailing.

The structure of the Rqc2p bound to the A-site tRNA, allowed for the identification of key residues that facilitate the interaction. Mutating these residues in Rqc2p (Rqc2p[mut]) abolishes the ability of the protein to bind tRNAs and facilitate CAT-tailing. However, the mutant is still able to engage the complex, recruit Ltn1p, and more importantly, facilitate the efficient degradation of the limited number of stalling substrates studied so far (Choe et al., 2016a; Defenouillère et al., 2016; Shen et al., 2015; Yonashiro et al., 2016). In addition, recent work has shown that when RQC-mediated degradation is compromised (e.g., in yeast strains with Ltn1p deleted), the presence of CAT tails on poly-lysine containing substrates causes protein aggregation, which leads to chaperone sequestration and proteotoxic stress (Choe et al., 2016a; Defenouillère et al., 2016; Yonashiro et al., 2016). The biological role, if any, of CAT tails in the context of the functional RQC remains poorly defined.

My thesis work has focused on defining the biological role of the CAT tails. In Chapter two I show that the Ltn1p ubiquitin ligase is surprisingly rigid and can only access residues right outside the ribosome exit tunnel. Therefore, for substrates that have no Ltn1p-accessible lysines, CAT tail addition can enable degradation by exposing lysines sequestered in the ribosome exit tunnel. In Chapter three I explore the role of CAT-tailing for substrates stalled during co-translational translocation at the ER lumen. Overall, my studies show that CAT tails provide a fail-safe mechanism that greatly expands the range of RQC-degradable substrates.

Mammalian RQC

Despite the rapidly expanding molecular knowledge of the yeast RQC, the systematic characterization of the mammalian RQC pathway has been lagging behind. Components of the pathway have been linked to nascent polypeptide degradation even before the identification of the complex in yeast. For example, in a forward genetic screen for factors causing neurodegeneration, Chu et al. identified a hypomorphic mutation in the murine homolog of Ltn1, listerin. Mice carrying this mutation displayed motor defects later in life due to motor neuron death (Chu et al., 2009). The later studies in yeast led to the hypothesis that the neurodegenerative phenotype was caused by accumulation of toxic translation product in mice harboring defective Ltn1 ubiquitin ligase.

Studying the mammalian pathway has been hindered by the lack of model stalling substrates. The substrate used for the majority of the yeast studies, a stretch of non-optimal arginine codon pairs sandwiched between two fluorescent proteins, does not cause stalling in higher eukaryotes. Alternative substrates, such as reporters lacking a stop codon or truncated messages without a polyA tail, become rapidly degraded and the steady state RNA and protein

levels are too low to be used for biochemical or genetic experiments. Messages caring strong secondary structures, such as stem loops, are surprisingly readily unwinded by the ribosome, suggesting that the mammalian ribosome is more processive than the yeast one.

Recently, a stalling substrate containing a polyA stretch has been used to study the early steps of the RQC pathway, i.e. the detection of the stall ribosome and ubiquitinaton of ribosomal proteins by the ubiquitin ligase ZNF598 (Juszkiewicz and Hegde, 2017a; Sundaramoorthy et al., 2017). Although both papers show that the RQC pathway is also capable of coping with stalled ribosomes in mammalian cells, they do not explore the presence of mammalian specific components. In Chapter five I develop a new mammalian RQC stalling reporter and use it for a high throughput genetic screen to identify novel mammalian-specific RQC factors (Chapter six). This screen yielded a comprehensive view of the mammalian RQC pathway and enabled me to identify a novel complex that translationally silences messages harboring stalled ribosomes.
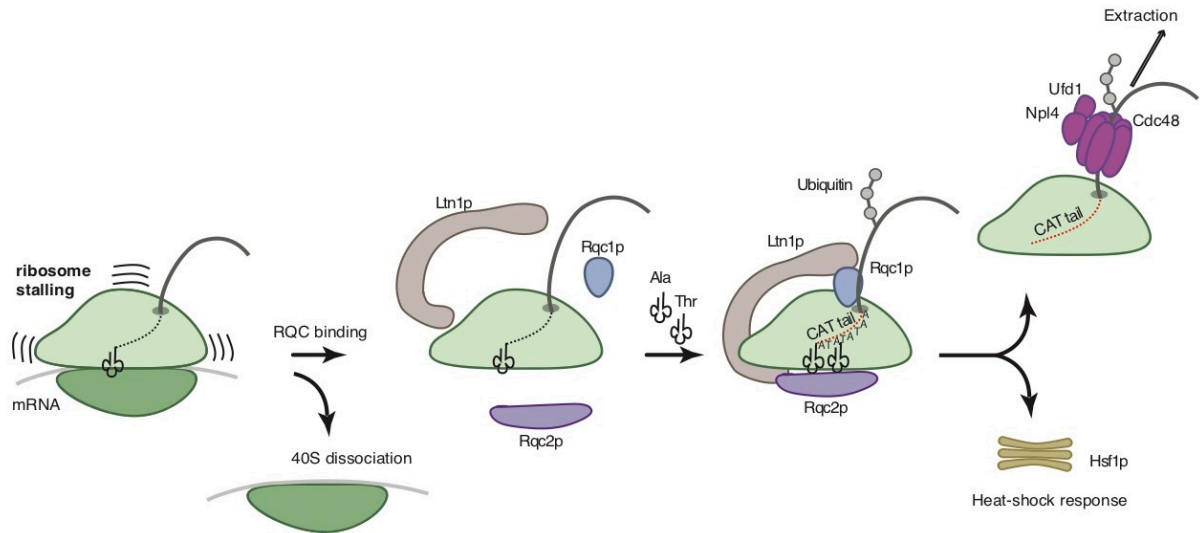
**Figure 1-1 Model of RQC engagement and function.** Stalled ribosomes are detected by early detection factors (not shown) and removed from the mRNA. The 40S subunit is released and recycled, and the 60S subunit that still carries P-site tRNA is the substrate of the RQC. The complex has three core components in yeast. It contains a ubiquitin ligase, Ltn1p, a protein that enables the recruitment of charged tRNAs and the synthesis of CAT tails, Rqc2p, and a poorly characterized protein, Rqc1p. The RQC facilitates the ubiquitination of the stalled nascent polypeptide, its extraction from the 60S subunit, and subsequent proteasomal degradation. In addition, the pathway can also activate the heat-shock response via a poorly characterized mechanism.

**CHAPTER TWO**

CAT-tailing as a fail-safe mechanism for efficient degradation of stalled nascent polypeptides in

the cytoplasm

Introduction

A surprising result from the original paper describing the RQC2-mediated addition of CAT-tails to the stalled nascent polypeptide (Shen et al., 2015) was that an RQC2 mutant that was specifically defective in its ability to recruit charged tRNAs, but could still recruit Ltn1p to the 60S ribosomes and therefore facilitate assembly of the complex, had little effect on the degradation of the nascent polypeptide. The only phenotype that could be assigned to the CAT tails at the time was their ability to induce the heat-shock response via an unknown mechanism. Later studies showed that when the RQC is not active, the presence of a CAT-tail, in addition to a homopolymeric stretch, such as a poly lysine track, leads the aggregation of the stalled nascent polypeptide (Choe et al., 2016b; Defenouillère et al., 2016; Shen et al., 2015; Yonashiro et al., 2016). The observed aggregates can sequester chaperones, such as Sis1p, which disrupts the protein homeostasis in the cell and is believed to induce the heat-shock response. However, the role of the CAT-tails in the context of the functional RQC remained poorly understood.

Results

To investigate the role of CAT-tailing in nascent polypeptide degradation, I designed two stalling constructs that differed in whether or not they encoded lysines, the canonical ubiquitin acceptor. I used green fluorescent protein (GFP) with all 20 lysines intact (20 Lys GFP) or mutated to arginine (Lys-free GFP). Both constructs contained a naturally lysine-free 3xHA epitope tag, a tobacco etch virus (TEV) protease site, and a well-characterized polyarginine-encoding stalling sequence (12R) (Dimitrova et al., 2009; Letzring et al., 2010). I also introduced C-terminal red fluorescent protein (RFP), which allowed us to monitor the efficiency of ribosome stalling (Fig. 2-1A, B). The constructs were expressed in *S. cerevisiae* cells with an

intact RQC, cells with single or double deletions of *LTN1* or *RQC2*, or cells harboring the CAT tail defective Rqc2p$^{mut}$. The 12R sequence efficiently induced ribosome stalling as no full-length GFP-12R-RFP fusion protein was detected (Fig.2-2). The 20 Lys GFP was efficiently degraded in wild-type (*wt*) cells, but accumulated in RQC-deletion strains (Fig. 2-1A). I also detected a range of Rqc2p-dependent higher molecular weight species upon Ltn1p deletion, indicative of CAT-tailing. In these and subsequent measurements of substrate degradation, I treated samples with TEV protease to remove the C-terminus of the stalled product, which collapses CAT tail-elongated species into a discrete band, facilitating quantitation (Fig. 2-1A, B, middle panel). Compared to *rqc2Δ* and *ltn1Δ* strains, in *rqc2$^{mut}$/ltn1Δ* and *rqc2Δ/ltn1Δ* strains I observed additional accumulation of the stalling reporter, which has been previously observed (Choe et al., 2016b; Defenouillère et al., 2016; Shen et al., 2015; Yonashiro et al., 2016), and is likely a secondary effect, caused by the inability of those strains to induce a heat-shock response (Brandman et al., 2012; Shen et al., 2015). Importantly, no additional accumulation of 20 Lys GFP stalling substrate from *ltn1Δ* cells was observed upon proteasomal inhibition (Fig. 2-3), which argues against the possibility of Ltn1p-independent process that leads to ubiquitin- and proteasome-mediated degradation of RQC substrates. Finally, consistent with previous work (Shen et al., 2015), the 20 Lys GFP substrate was efficiently degraded in the Rqc2p$^{mut}$ strain, confirming that CAT tails are not necessary for degradation of this particular substrate (Fig. 2-1A, C).

In contrast to 20 Lys GFP, Lys-free GFP was heavily CAT-tailed even in *wt* cells (Fig. 2-1B), consistent with previous reports (Yonashiro et al., 2016). I next compared protein levels of Lys-free GFP in *wt* cells and cells with deletions of RQC components (Fig. 2-1B, D). I detected comparable Lys-free GFP accumulation in all strains, indicating that, although Lys-free GFP

engages the RQC complex and is CAT-tailed, it is not degraded. This apparent lack of degradation was not confounded by differences in mRNA levels between the different strains (Fig. 2-1D). These data establish that the presence of lysines as ubiquitin acceptors is a key requirement for nascent polypeptide degradation. In addition, the absence of further Lys-free GFP accumulation upon proteasomal inhibition indicates that the CAT tail itself does not serve as a degron to promote proteasomal degradation (Fig. 2-1E).

Cryo-electron microscopy structures of Ltn1p bound to the 60S ribosomal subunit have revealed that the RING domain of Ltn1p, which is required for ubiquitin transfer, is held in close proximity to the ribosome exit tunnel (Lyumkis et al., 2014; Malsburg et al., 2015). This observation suggests that the position of lysines along the nascent polypeptide may determine the ability of Ltn1p to ubiquitinate them and that lysines may be optimally positioned as they emerge from the exit tunnel. To explore this hypothesis, I designed three stalling constructs based on the 20 Lys GFP construct (Fig. 2-4A). I either decreased the distance between the last lysine on GFP and the stalling sequence by removing linkers or increased it by mutating the final lysine to arginine. I then measured protein degradation in *wt* versus *rqc2Δ* cells. Although these stalling constructs were similar in their number of lysines (19 or 20) and presumably nascent polypeptide fold, the proximity of the lysines to the Ltn1p active site was different, which resulted in a dramatic difference in their RQC-mediated degradation (Fig. 2-4A). The stalling construct with the shortest linker of 37 amino acids, placing the last lysine just outside the ribosome exit tunnel, was robustly degraded in *wt* cells. However, extending the distance between the last lysine on GFP and the stall site to 51 and 75 amino acids rendered the stalling constructs increasingly resistant to RQC-mediated degradation.

These results suggest that the ability of the RQC to efficiently degrade nascent polypeptides depends not only on the presence of lysines, but also on their position relative to the Ltn1p RING domain. Alternatively, the lack of degradation may result from the Lys residues being embedded within the folded structure of GFP. To distinguish between these possibilities, I designed a stalling construct that contained four lysines in an unstructured region (as part of a 3xFLAG tag) that were systematically displaced by 0, 10, 20, 40, or 80 amino acids from the 12R stalling site (XTEN0, 10, 20, 40, and 80, respectively) (Fig. 2-4B). The extension is derived from the XTEN protein (Schellenberger et al., 2009), which has been previously used as a neutral, lysine-free, unstructured linker that minimally impacts protein stability and solubility (Komor et al., 2016). Depending on the length of the XTEN linker, the four lysines of the 3xFLAG tag will be at a different distance from the Ltn1p active site when RQC engages the stalled ribosome. Therefore, the relative degradation efficiency provides a proxy for the ability of Ltn1p to ubiquitinate those lysines. Because CAT tails are C-terminal extensions of various length (Shen et al., 2015), CAT-tailing can alter the number of amino acids between the C-terminus of the stalling substrate and the 3xFLAG tag and thus complicates our analysis. To control for this effect, I expressed the XTEN constructs in *wt* cells, CAT-tailing-deficient (expressing Rqc2p[mut]), and *rqc2Δ* cells (Fig. 2-4B, Fig.2-5), as well as *Ltn1ΔRING* cells (Fig. 2-6A).

In contrast to previous RQC substrates, the XTEN0 construct was efficiently degraded only in cells capable of CAT-tailing (Fig. 2-4B, Fig. Fig. 2-7), in an Ltn1p-dependent manner (Fig. 2-6A). The degradation of the XTEN0 substrate was not improved by overexpressing Rqc2p[mut], suggesting that the defect was caused by the lack of CAT tails and not by lower Rqc2p[mut] levels or inefficient complex engagement (Fig. 2-6B). Because the 3xFLAG tag is at

the extreme C-terminus of the stalled XTEN0 substrate, all four lysines of the stalled nascent chain are fully sequestered in the ribosome exit tunnel and thus inaccessible to Ltn1p (Fig. 2-7). However, the addition of CAT tails exposes these "hidden" lysines, pushing them out of the exit tunnel and making them available for Ltn1p-mediated ubiquitination, as previously hypothesized (Brandman and Hegde, 2016b; Simms et al., 2017). A similar effect was seen for a construct analogous to XTEN0 based on a lysine-free mutant of yeast His3p (Fig. 2-8A, B), as well as constructs with two consecutive lysines before the stalling site (Fig. 2-8C, D). Thus, the degradation of stalled substrates in which the only available lysine residues are sequestered within the ribosome may generally require CAT tail elongation.

Compared to the XTEN0 substrate, degradation of the XTEN10 and XTEN20 substrates, in which the lysine stretch is predicted to be at least partially present immediately outside of the exit tunnel, became less dependent on CAT-tailing, as evidenced by partial degradation of the nascent polypeptide in $rqc2^{mut}$ cells (Fig. 2-4B, Fig. 2-7B). The degradation of these substrates was somewhat more robust in $wt$ cells, likely because not all lysines in the 3xFLAG tag are outside the exit tunnel, and thus CAT-tailing enhanced their accessibility to Ltn1p. Critically, when the 3xFLAG was positioned past a certain distance (XTEN linker of 40 amino acids or longer) beyond the exit tunnel, the substrate was completely resistant to RQC-mediated degradation. Taken together, these data provide strong support for a model in which Ltn1p can only access a limited window of amino acids on the nascent polypeptide. If this stretch lacks ubiquitin acceptors, CAT tail elongation can expose lysines sequestered in the ribosome exit tunnel to facilitate polypeptide degradation.

In order to confirm that the lack of RQC-mediated degradation of substrates with lysines located past a critical distance from the stalling site was the result of the differential ability of

Ltn1p to ubiquitinate those substrates, I examined the ubiquitination of stalling substrates with lysines sequestered in the ribosome exit tunnel (XTEN0), within (XTEN10) or past (XTEN80) the Ltn1p-accessible region (Fig. 2-4C, Fig. 2-9). Consistent with the degradation results described above, I observed Ltn1p-dependent ubiquitination when lysines were positioned in the exit tunnel or within the Ltn1p-accessible region, and no ubiquitination when the lysines were past the Ltn1p-accessible window (Fig. 2-4C). In addition, I observed similar levels of ubiquitination of a substrate with lysines residues positioned in the Ltn1p-acccesible region in cells expressing wild type or mutant Rqc2p (Fig. 2-4D). These data suggest that the process of CAT-tailing does not strongly impact the ability of Ltn1p to directly ubiquitinate stalling substrates but rather enables Ltn1p to gain access to lysines sequestered in the exit tunnel.

I next explored how the sequences surrounding the Lys residues impact the ability of Ltn1p to ubiquitinate those residues. I utilized a Lys-free GFP construct in which the Lys-containing 3xFLAG tag was replaced with a Lys-free 3xHA tag upstream of a Lys residue positioned at different distances from the 12R stalling site (Fig. 2-10A, Fig. 2-11A). This construct was degraded in a CAT-tail-dependent manner when the lysine was 29 amino acids away from the 12R site (i.e. sequestered in the exit tunnel), but was no longer degraded when it was 49 amino acids away. In addition, I confirmed that the inability of RQC to degrade a substrate with a Lys residue 49 amino acids away from the stalling site was not caused by steric hindrance, since positioning the Lys residue between two unstructured regions did not rescue degradation (Fig. 2-10B). These experiments establish that the ability of Ltn1p to effectively access residues in a limited window proximal to the exit tunnel is determined by the distance of the Lys residues from the C-terminus, rather than by the sequences or local structure surrounding these residues.

A further advantage of the 3xHA constructs is that the Lys residues are at a discrete position rather than being distributed over 24 amino acids (as part of a 3xFLAG tag), thus allowing me to more precisely define the Ltn1p accessible window. I therefore systematically extended the distance between the Lys residues and the stalling site in five amino acid increments from 29 (sequestered in the exit tunnel) to 44 amino acids (past Ltn1p's accessible region) (Fig. 2-10C). In marked contrast to the substrate with Lys residues in the ribosome exit tunnel, which was fully dependent on CAT-tailing for degradation, the stalling construct containing lysines positioned between 34-41 amino acids away from the stall site was efficiently degraded in a CAT tail-independent manner (Fig. 2-10C, D). As observed with the FLAG-based constructs, these differences in degradation were consistent with differences in Ltn1p-mediated ubiquitination of the model substrates (Fig. 2-12). Importantly, the Ltn1p accessible window defined by using four lysines (Fig. 2-4B) is fully consistent with the refined window (Fig. 2-11B). These data show that Ltn1p can efficiently access ~12 amino acids outside the ribosome exit tunnel, and that CAT-tailing expands the range of suitable Ltn1p substrates by exposing lysines sequestered in the ribosome exit tunnel.

Non-stop mRNAs missing an in-frame stop codon (e.g., as the result of premature polyadenylation) present an important source of natural RQC substrates with lysines sequestered in the exit tunnel (Brandman and Hegde, 2016b). For such messages, the ribosome continues translating into the poly(A) tail, which appends a series of AAA-encoded lysines to the C-terminus. Since the poly(A) tails of *S. cerevisiae* mRNAs have a median length of 27 nucleotides (Subtelny et al., 2014), which would encode ~9 lysines and possibly fewer due to ribosome sliding (Koutmou et al., 2015), the resulting lysines will be sequestered entirely within the exit tunnel. Indeed, I find that Lys-free GFP encoded without a stop codon and with a defined

poly(A) tail of 30 nucleotides was degraded in a CAT tail-dependent manner (Fig. 2-10E).

Importantly, a stalling construct harboring a single lysine residue hidden in the exit tunnel can be

efficiently ubiquitinated by Ltn1p (Fig. 2-13) and degraded by the proteasome (Fig. 2-10A) only

in the presence of CAT tails.  Therefore, CAT-tailing exposes Lys residues that result from

translation through the poly(A) tail, allowing for efficient ubiquitination and degradation of non-

stop decay substrates.

Conclusions

Collectively, my studies reveal a remarkable and unanticipated feature of Ltn1p, the key

ubiquitin ligase responsible for RQC-mediated degradation of incomplete nascent chains: Ltn1p

is only able to access those lysines that lie within a very narrow window of the exit tunnel, likely

due to the fact that the RING domain of Ltn1p is held in close proximity to the exit tunnel as

well as interactions of the nascent polypeptide with chaperones, processing and targeting factors

(Kramer et al., 2009). This spatial specificity could protect the cell from collateral damage (e.g.,

degradation of ribosomal proteins or the translocation machinery, as well as unregulated quality

control signaling (Higgins et al., 2015; Juszkiewicz and Hegde, 2017a)) caused by incidental

ubiquitination by Ltn1p. However, it also suggests that without a mechanism to relieve the

lysine-positioning restriction, many endogenous RQC substrates would be resistant to Ltn1p-

mediated degradation. Therefore, Ltn1p's limited reach presents a challenge to the RQC

machinery that must deal with a diverse range of substrates. My studies reveal that, in addition to

its previously described role in promoting aggregation and inducing a heat shock response when

RQC function is compromised (Choe et al., 2016b; Defenouillère et al., 2016; Yonashiro et al.,

2016), CAT-tailing acts as a fail-safe mechanism that enables a far broader range of substrates to

be degraded. Specifically, CAT-tailing can expand the number of RQC-degradable substrates by

exposing lysines sequestered in the ribosome exit tunnel, including those that result from

translation through a poly(A) tail.

Figures
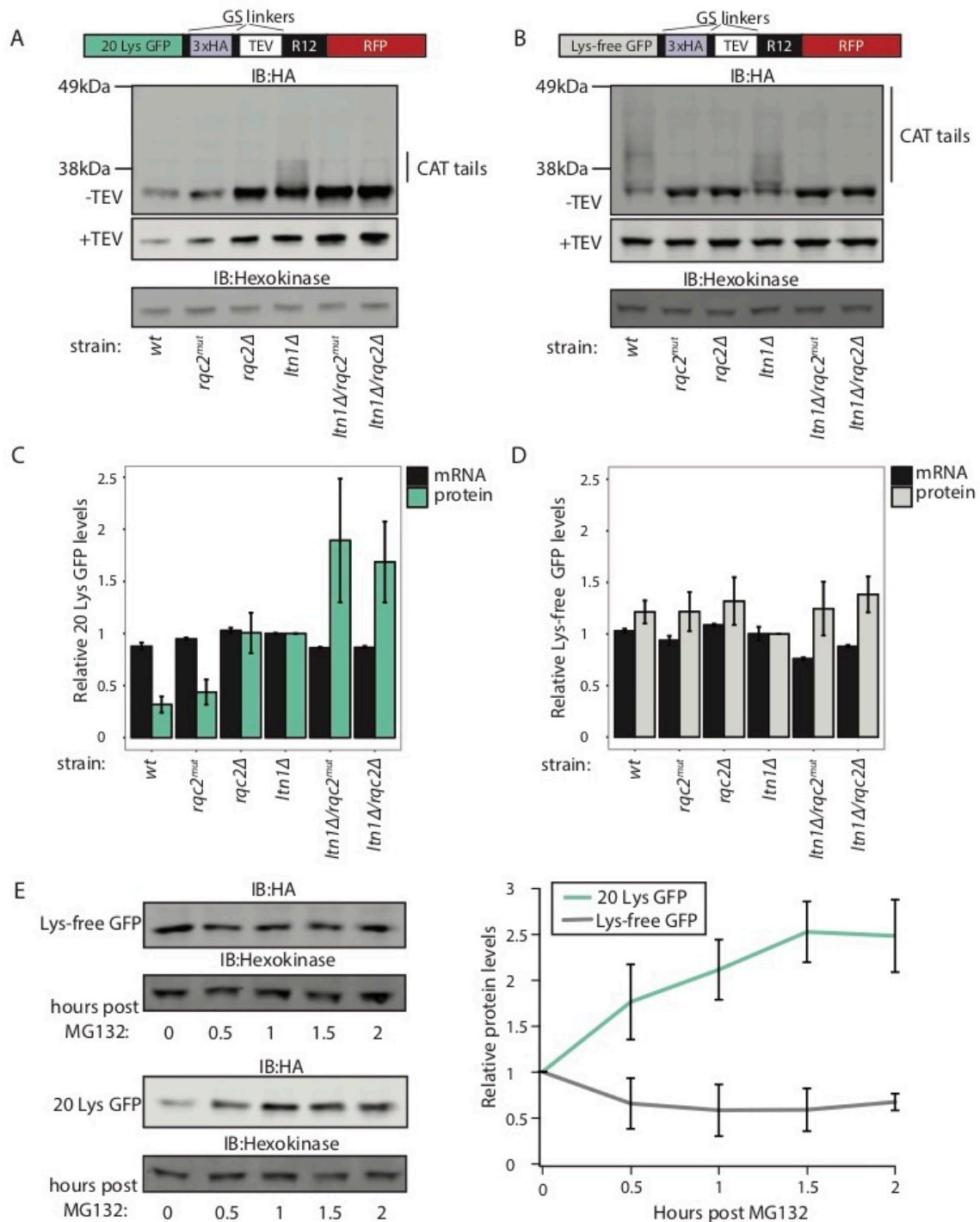


**Fig. 2-1. Degradation of RQC substrates requires lysines.** (A, B) Stalling reporters with (20 Lys GFP) or without (Lys-free GFP) Lys residues were expressed in wild-type (wt) cells or cells harboring mutations or deletions of the indicated RQC components. Cell lysates were

18

incubated with TEV protease or mock digested, and analyzed by SDS-PAGE and immunoblot (IB). Hexokinase was used as a loading control. (C, D) Protein and mRNA levels of 20 Lys GFP and Lys-free GFP stalling substrates were quantified by densitometry (green and grey bars) and quantitative PCR (black bars). Error bars indicate s.d. from three independent IBs or three qPCR samples. (E) 20 Lys GFP and Lys-free GFP stalling constructs were expressed in pdr5Δ cells. At time point 0 the proteasome inhibitor MG132 was added to the media (60 μM) and samples were harvested every 30 minutes. Total cell lysate was incubated with TEV protease and analyzed by SDS-PAGE and IB. The relative amount of stalling substrate accumulating over time was quantified by densitometry (graph on the right). Error bars indicate s.d. from three independent experiments.
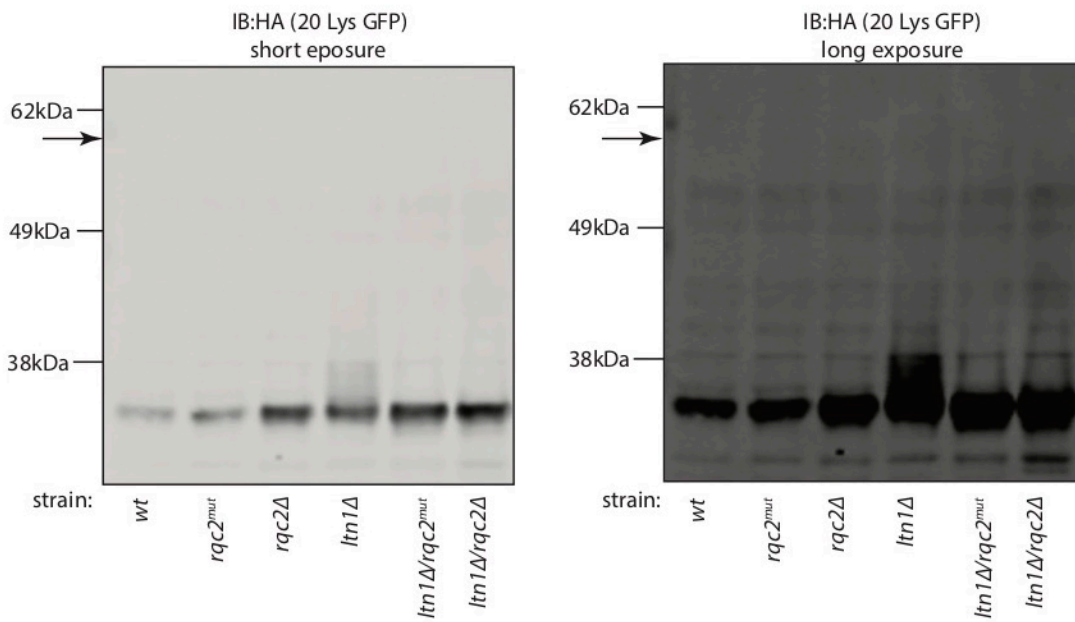
**Fig. 2-2. A polybasic region (12R) causes efficient ribosome stalling.** Immunoblot (IB) shown in Fig. 2-1A at normal exposure (left) and overexposed (right). The arrow indicates the approximate location of the readthrough product (GFP-3xHA-TEV-12R-RFP).
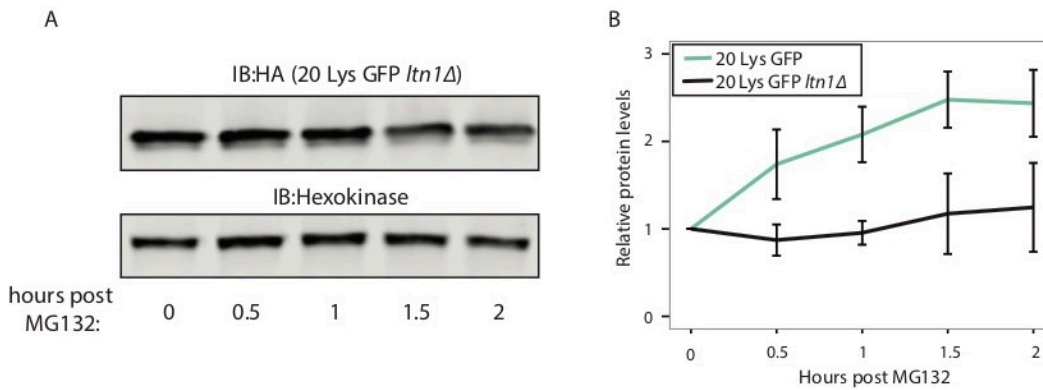
**Fig. 2-3. RQC substrates are strictly Ltn1p-dependent for ubiquitin- and proteasome-mediated degradation.** (**A**) 20 Lys GFP stalling construct was expressed in *pdr5Δ* or *pdr5Δ/ltn1Δ* cells. At time point 0 the proteasome inhibitor MG132 was added to the media (60μM) and samples were harvested every 30 minutes. Total cell lysate was incubated with TEV protease and analyzed by SDS-PAGE and immunoblot (IB). (**B**) The relative amount of stalling substrate accumulating over time was quantified by densitometry. The lines show 20 Lys GFP accumulation upon proteasomal inhibition from *pdr5Δ* cells (green) (representative IB shown in Fig. 2-1E) and *pdr5Δ/ltn1Δ* cells (black). Error bars indicate s.d. from three independent experiments.
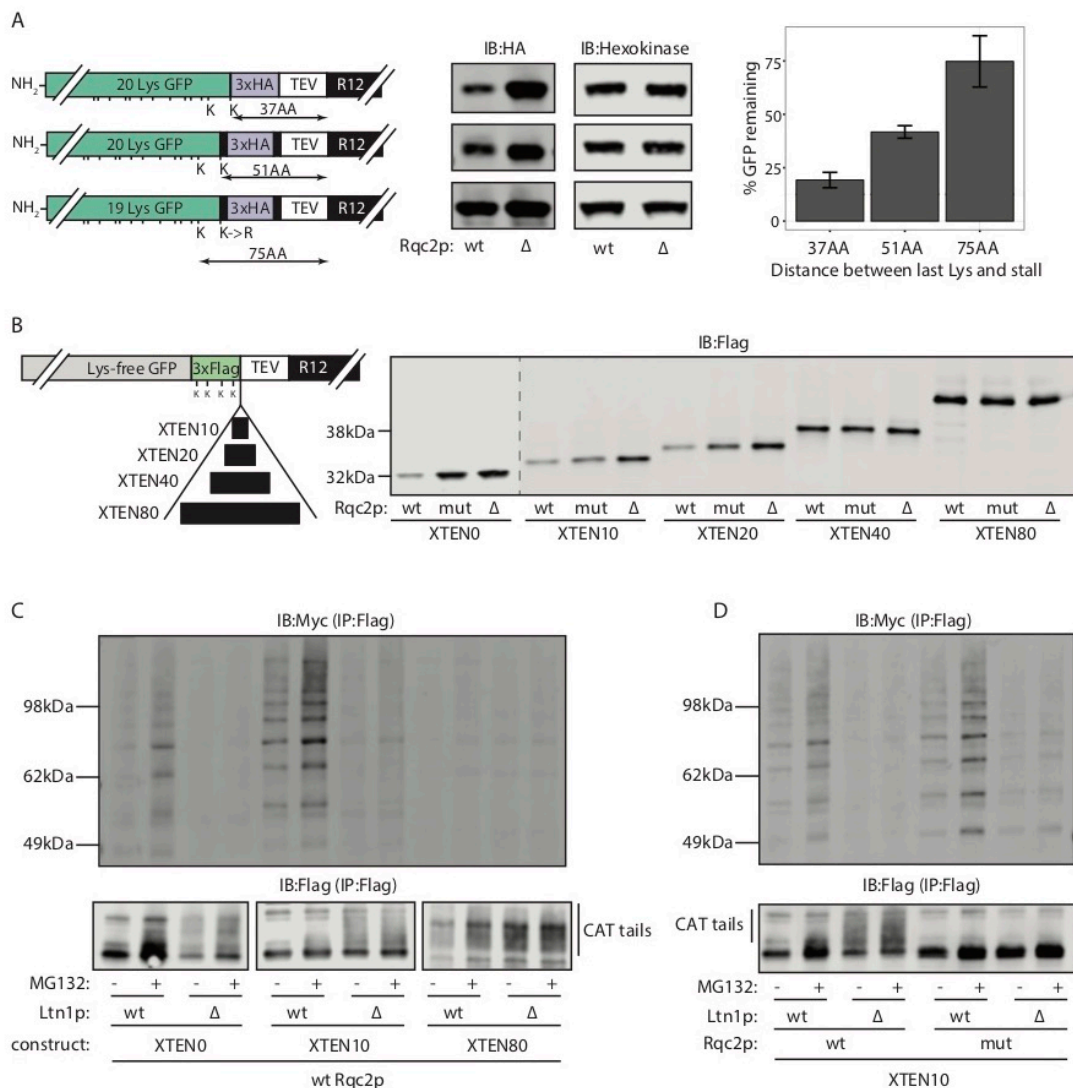
**Fig. 2-4. Lysine positioning is critical for Ltn1p-mediated ubiquitination and proteasomal degradation.** (A) Variants of 20 Lys GFP with the indicated number of amino acids between the most C-terminal lysine of GFP and the first Arg of the R12 stall site were expressed in wt or rqc2Δ cells and analyzed by SDS-PAGE and immunoblot (IB) (middle). The relative amount of GFP substrate remaining in wt cells compared to rqc2Δ cells was quantified by densitometry (right). Error bars indicate s.d. from three independent experiments. (B) XTEN linkers of the indicated length were inserted between the 3xFLAG tag and the TEV site of a Lys-free GFP stalling construct (left). Constructs were expressed in cells harboring wild-type (wt), mutant (mut) or no (Δ) Rqc2p. Total cell lysates were incubated with TEV protease and analyzed by SDS-PAGE and immunoblot (IB) (right). (C) Cells expressing Myc-tagged ubiquitin and XTEN0, 10 or 80 stalling construct were treated (+) or not (−) with the proteasome inhibitor MG132. SDS-boiled lysates of wt or ltn1Δ strains were used for FLAG immunoprecipitation (IP). (D) FLAG IP for XTEN10 reporter from wt or ltn1Δ cells harboring wt or mut Rqc2p.
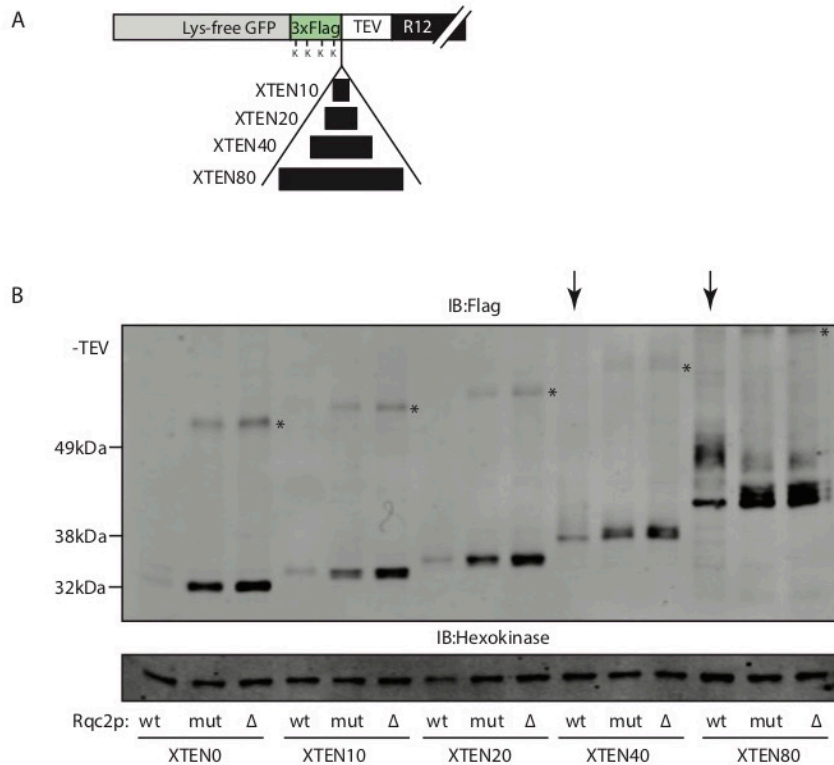
**Fig. 2-5. Lysine position is critical for RQC-mediated degradation.** (**A**) XTEN linkers of the indicated length were inserted between the 3xFLAG tag and the TEV site of a Lys-free GFP stalling construct. (**B**) Constructs were expressed in *wt*, *rqc2^mut^* or *rqc2Δ* cells. Total cell lysates were analyzed by SDS-PAGE and immunoblot (IB). Hexokinase was used as a loading control. Arrows point at substrates with visible CAT tails in *wt* cells. Asterisks (*) indicate a higher molecular weight-band of unknown origin that is present in all tested strains, including *wt* cells. Since it is TEV- and RNase-resistant, this band does not represent a read- through product or tRNA-linked nascent polypeptide; it is likely a stable oligomer of the substrate.
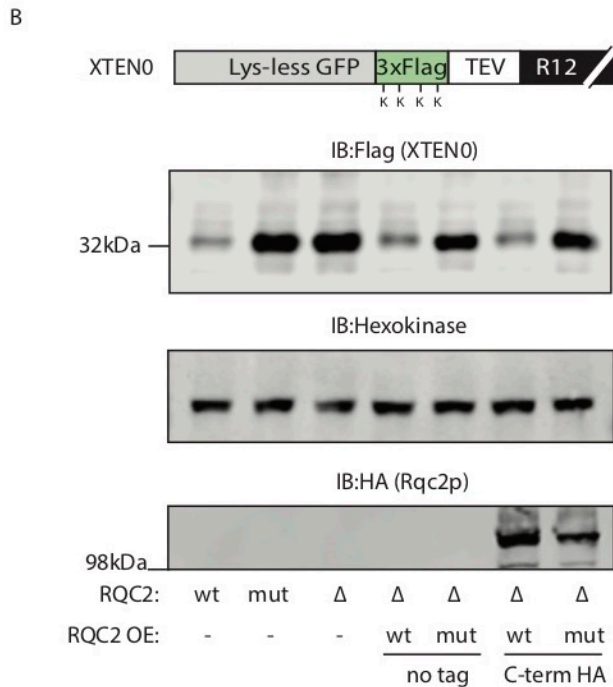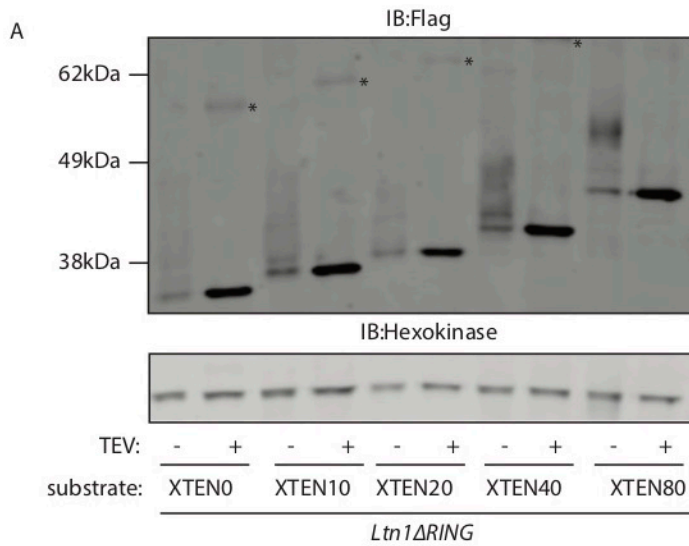
**Fig. 2-6. CAT tail elongation enables nascent polypeptide degradation.** (**A**) XTEN0-80 constructs were expressed in *Ltn1ΔRING* cells. Total cell lysate was incubated with TEV protease (+) or mock digested (-) and then analyzed by SDS-PAGE and immunoblot (IB). The asterisks (**\***) indicate the same species described in Fig. S2. (**B**) Wild-type (wt) or mutant (mut) Rqc2p with or without a carboxy terminal HA tag was overexpressed (OE) in *rqc2Δ* cells. Protein levels of the XTEN0 stalling reporter were compared between the overexpression cells and *wt*, *rqc2^mut^* or *rqc2Δ* cells.
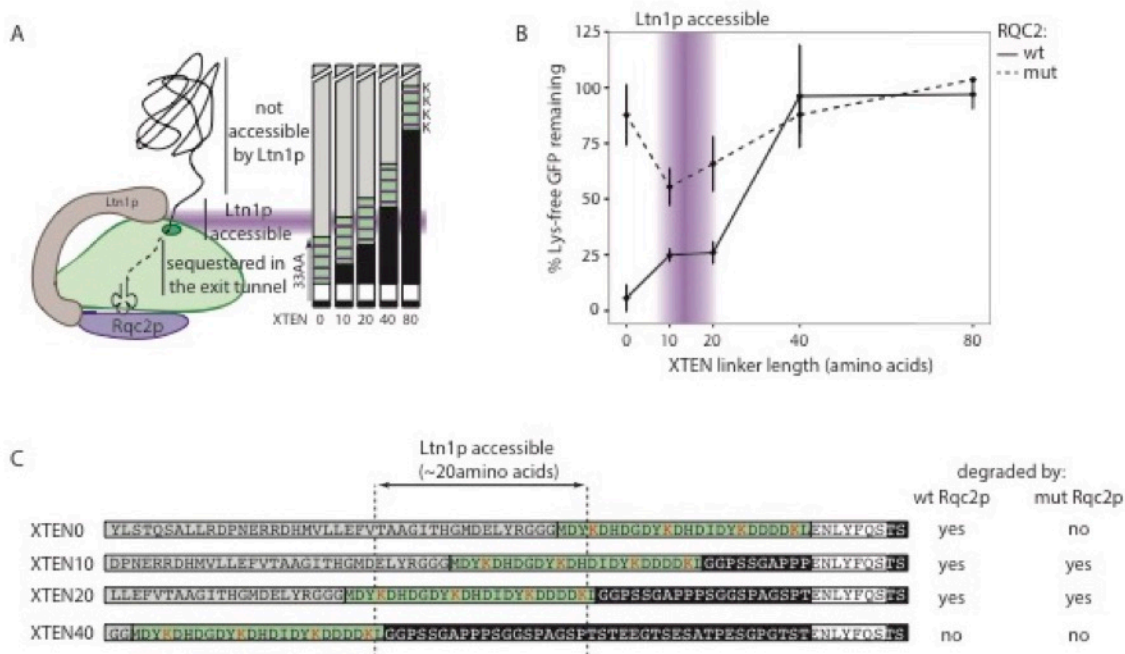
**Fig. 2-7. Ltn1p can access a limited number of amino acids outside the ribosome exit tunnel.** (**A**) Schematic of Ltn1p's accessible region (shaded area). (**B**) Quantification of protein levels from the stalling constructs in Fig. 2B. Shown is the percent stalling construct remaining in *wt* or *rqc2^mut^* or cells relative to *rqc2Δ* cells as a function of the XTEN linker length. The error bars represent s.d. from three independent experiments. (**C**) Alignment of the last 75 amino acids before the 12R stalling site of constructs XTEN0-40. Highlighted are the C-terminus of the Lys-less GFP (grey), 3xFLAG tag (green), linkers (black), and TEV site (white). The Lys residues are highlighted in red. The dashed line points to the approximate Ltn1p's accessible region.
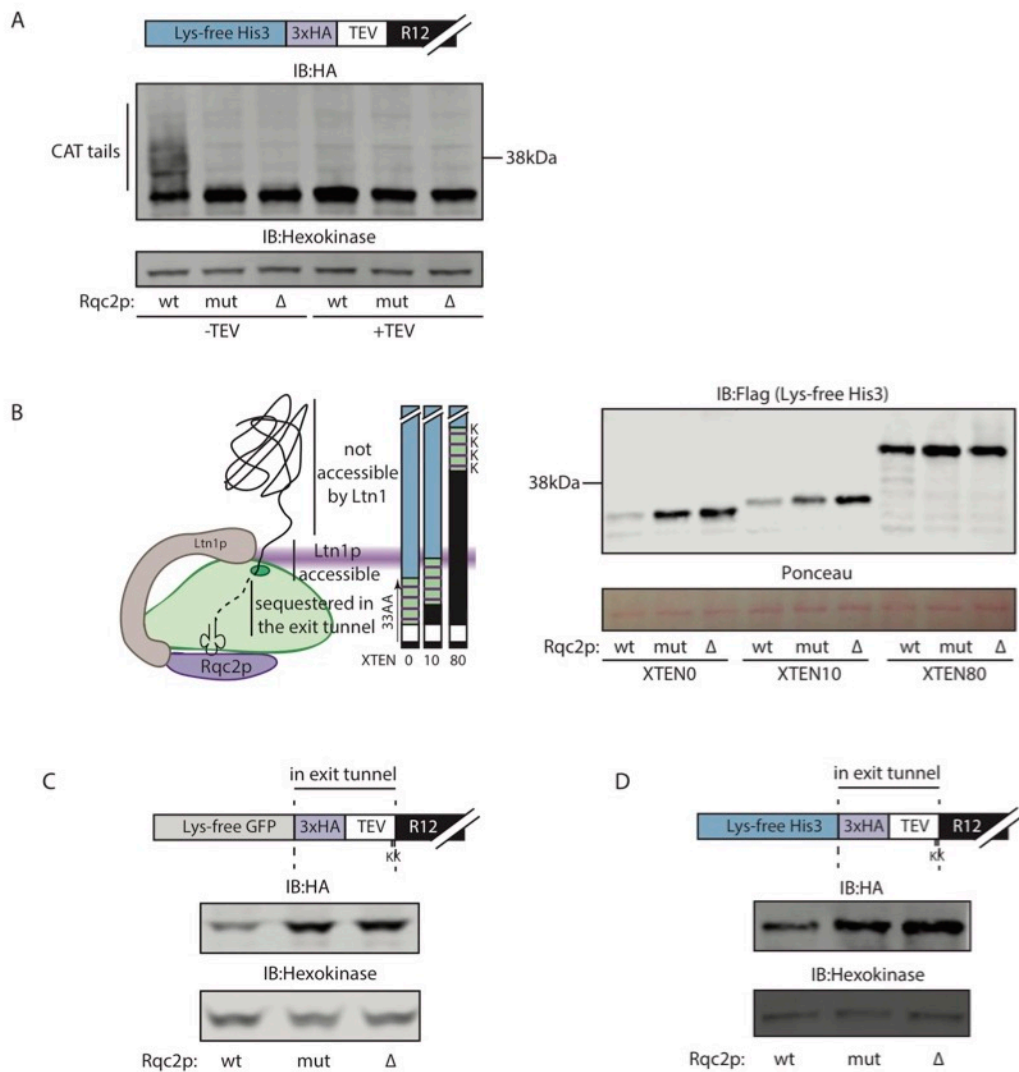
**Fig. 2-8. Lysine availability dictates the degradation of a His3p-based stalling reporter.** (**A**) Lysine free His3 (Lys-free His3) was expressed in *wt*, *rqc2^mut*, or *rqc2Δ* cells. Total cell lysate was incubated with TEV protease (+TEV) or mock digested (-TEV) and analyzed by SDS-PAGE and immunoblot (IB). (**B**) Lys-free His3 stalling constructs harboring Lys-rich 3xFLAG tag located at different distances from the 12R stalling site (left) were expressed in *wt*, *rqc2^mut*, or *rqc2Δ* cells. Total cell lysate was incubated with TEV protease and analyzed by SDS-PAGE and IB (right). (**C, D**) Lys-free stalling constructs (Lys-free GFP or Lys-free His3) with two lysines introduced before the 12R stalling site were expressed in *wt*, *rqc2^mut*, or *rqc2Δ* cells. Total cell lysate was incubated with TEV protease and analyzed by SDS- PAGE and IB.
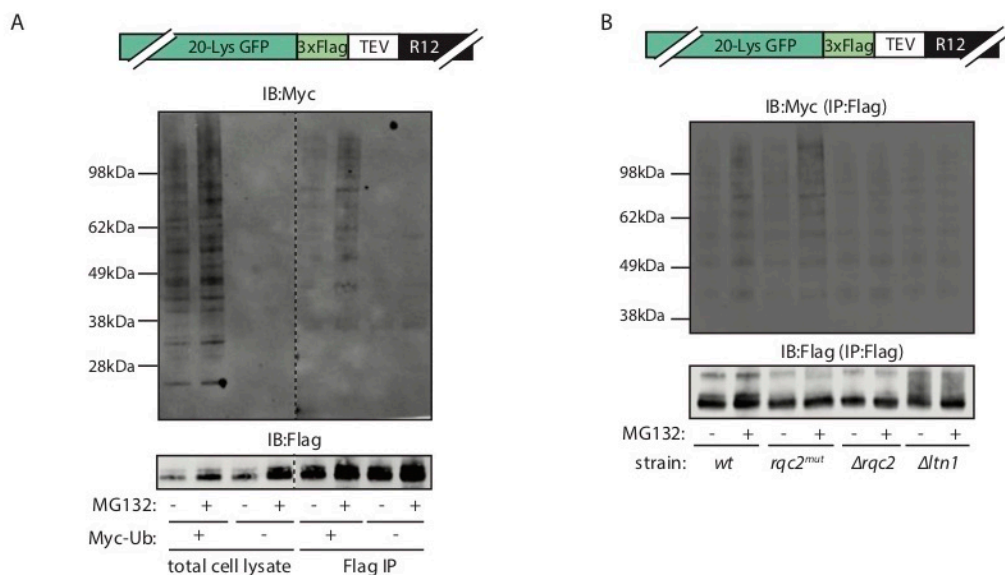
**Fig. 2-9. Ubiquitin can be detected on the 20 Lys GFP stalling substrate.** (**A**) *prd5Δ* cells expressing 20-Lys GFP stalling reporter and copper-inducible Myc-tagged ubiquitin were treated with (+) or without (-) $Cu_2SO_4$. Samples from total cell lysate or denaturing immunoprecipitation (Flag IP) were analyzed on SDS-PAGE and immunoblot (IB). (**B**) Cells expressing Myc-tagged ubiquitin and 20-Lys GFP stalling reporter were treated with (+) or without (-) the proteasome inhibitor MG132. SDS-boiled lysates from wt, *rqc2^{mut}*, *rqc2Δ,* and *ltn1Δ* cells were used for Flag IP. The eluate was analyzed by SDS-PAGE and IB. All cells had the multidrug transporter gene *PDR5* deleted.
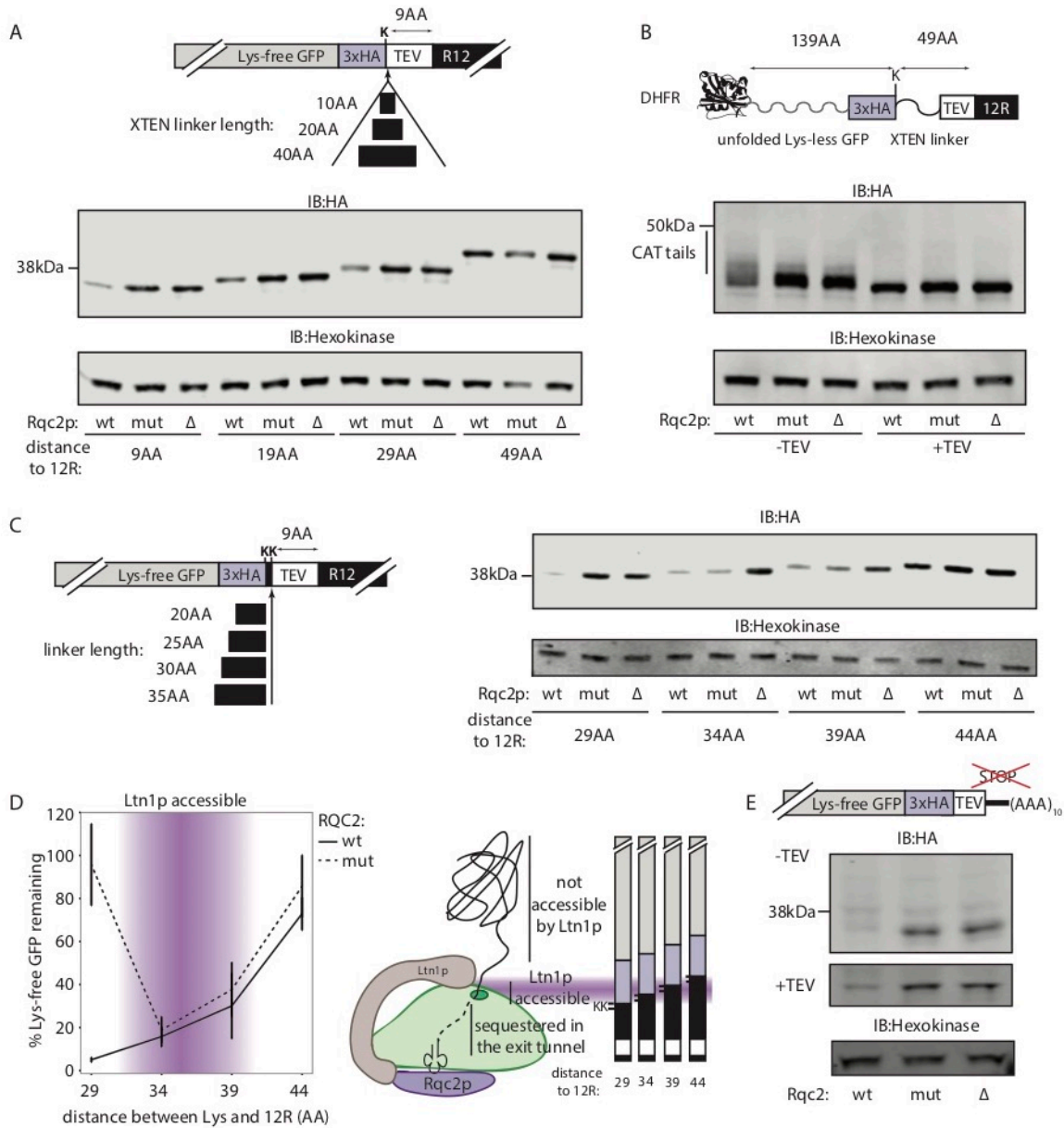
**Fig. 2-10. The ability of Ltn1p to access lysines outside the exit tunnel is independent of sequence or structural context.** (A) XTEN linkers of the indicated length were inserted after the single Lys residue of a Lys-free GFP stalling construct. Constructs were expressed in cells harboring wild-type (wt), mutant (mut) or no (Δ) Rqc2p. Total cell lysates were incubated with TEV protease and analyzed by SDS-PAGE and immunoblot (IB). (B) Stalling construct with a Lys residue positioned between two unstructured sequences (139 amino acids of Lys-free GFP and 40 amino acids of XTEN linker) was expressed in wt, rqc2mut, and rqc2Δ cells. Cell lysates were treated with TEV protease or mock digested and analyzed by SDS-PAGE and IB. (C) XTEN linkers of the indicated length were inserted after the Lys residues of a Lys-free GFP stalling construct (left). Constructs were expressed in cells harboring wt, mut or Δ Rqc2p. Total cell lysates were incubated with TEV protease and

analyzed by SDS-PAGE and IB (right). (D) Quantification of protein levels from the stalling constructs in (C) (left). Shown is the percent stalling construct remaining in wt or rqc2mut cells relative to rqc2Δ cells as a function of the distance between the Lys residues and the 12R stalling site. The error bars represent s.d. from three independent experiments. Schematic of Ltn1p's accessible region (shaded area) (right). (E) Lys-free GFP with no in-frame stop codon and a (AAA)10 tail was expressed in wt, rqc2mut, and rqc2Δ cells. Cell lysates were analyzed by SDS-PAGE and IB.
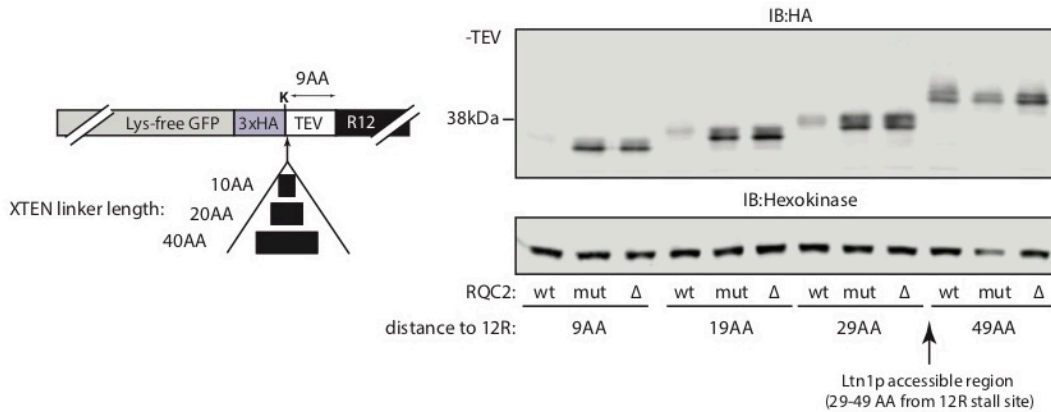
**Fig. 2-11. Ltn1p's ability to access Lys residues outside the exit tunnel is context independent.** (**A**) Lys-free GFP containing one lysine before the TEV protease site and variable XTEN linker lengths was expressed in *wt*, *rqc2^{mut}*, or *rqc2Δ* cells. Total cell lysates were analyzed by SDS-PAGE and immunoblot (IB). Hexokinase was used as a loading control. Arrow points at the transition from a substrate that has CAT-tail-dependent degradation to a non-degradable substrate. (**B**) Alignment of the last 53 amino acids before the 12R stalling site of Lys-free GFP 3xFLAG (4 Lys) and Lys-free GFP 3xHA (2 Lys). Highlighted are the C-terminus of the Lys-less GFP (grey), 3xFLAG tag (green), 3xHA (purple), linkers (black), and TEV site (white). The Ltn1p-accessible Lys residues are highlighted in red. The dashed line points to the approximate Ltn1p-accessible region.
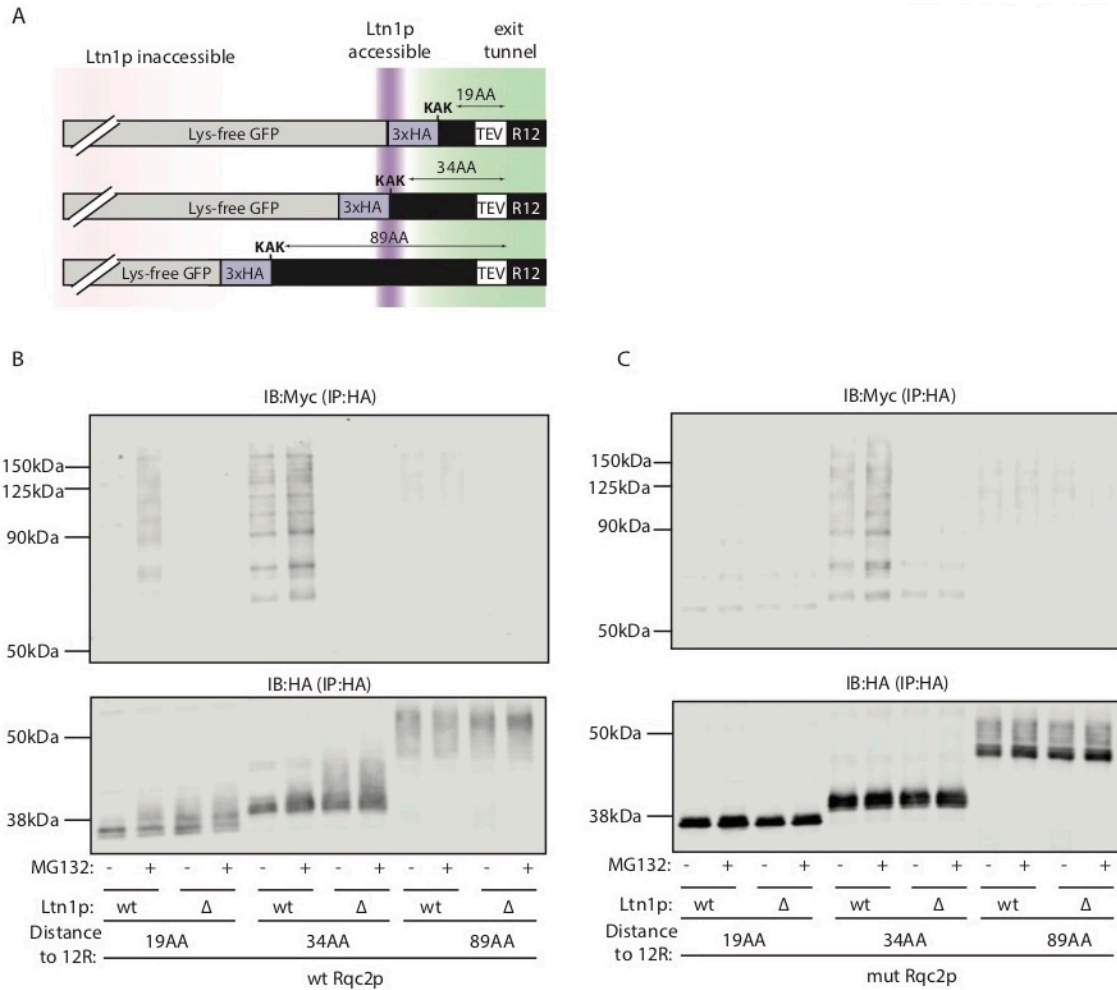
**Fig. 2-12. Ubiquitination of stalling constructs is influenced by lysine positioning. (A)** Schematic of the stalling constructs used for ubiquitin detection. Ribosome exit tunnel (~32 amino acids) and Ltn1p-accessible window (~12 amino acids) are highlighted in green and purple respectively. (**B**, **C**) Lys-free GFP with two lysines positioned at 19, 34 or 89 amino acids away from the 12R stalling site were expressed in *wt* or *ltn1Δ* cells harboring wild type (wt) or mutant (mut) Rqc2p. All cells expressed Myc-tagged ubiquitin and had the multidrug transporter gene *PDR5* deleted. SDS-boiled lysates were used for HA immunoprecipitation (IP). The eluate was analyzed by SDS-PAGE and immunoblot (IB).
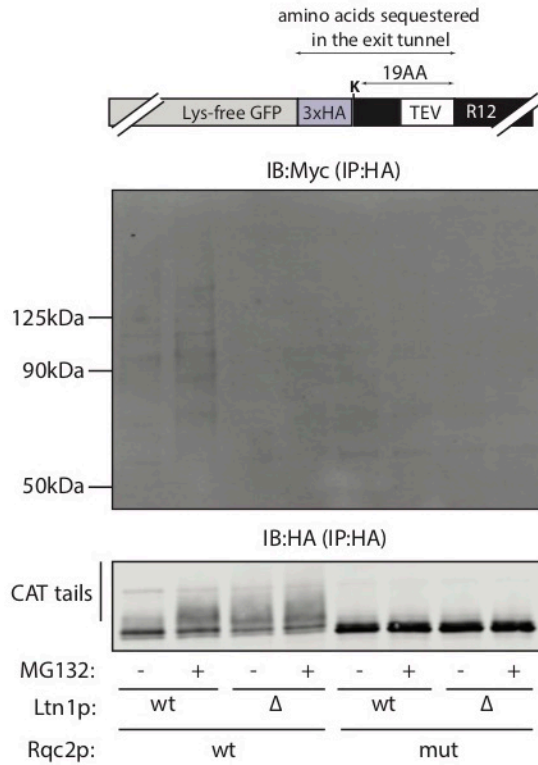
**Fig. 2-13. CAT-tailing facilitates Ltn1p-mediated ubiquitination of a single Lys residue sequestered in the ribosome exit tunnel.** Lys-free GFP with a single lysine sequestered in the ribosome exit tunnel (diagram above) was expressed in *wt*, *ltn1Δ*, *rqc2$^{mut}$* or *rqc2$^{mut}$/ltn1Δ* cells. All cells expressed Myc-tagged ubiquitin and had the multidrug transporter gene *PDR5* deleted. SDS-boiled lysates were used for HA immunoprecipitation (IP). The eluate was analyzed on SDS-PAGE and immunoblot (IB).

Materials and Methods

**Yeast strains**

All strains used in this study are listed in supplemental Table 2-1. Strain BY4741 was used as the wild-type parental strain. Genomic knockouts were generated by one-step gene replacement as previously described (Rothstein, 1991). The knockouts were confirmed by genomic PCR. Rqc2p$^{mut}$ strain was generated by replacing the endogenous *RQC2* open reading frame with the mutant version via scarless pop-in/pop-out method (Longtine et al., 1998).

Stalling constructs (GFP, Lys-less GFP, XTEN0-80, Lys-His3) were expressed on yeast centromere plasmids. Rqc2p constructs were overexpressed from yeast episomal plasmids. All plasmids are detailed in supplemental Table 2-2.

**Cell Lysis**

Cells were grown to final OD$_{600}$ between 0.6 and 1, and harvested by centrifugation. The pellets were washed once in water and resuspended in yeast lysis buffer (10 mM Tris pH7.5, 150 mM NaCl, 0.5 mM EDTA, 0.5% NP-40, 1x Halt$^{TM}$Protease Inhibitor Cocktail (Thermo Fisher Scientific)). An equal volume of glass beads (BioSpec, 0.5 mm, acid washed) was added to the cells and the cells were lysed by vortexing for 1min and chilling on ice for 1min, five times total. The lysate was clarified by centrifugation at 20 000 g for 2 min. Protein concentration was measured with BCA Assay (TheroFisher Scientific).

**Western Blot**

Cell lysates were denatured at 70°C for 10 minutes in Laemmli buffer. Proteins were separated on Bolt® 4-12% Bis-tris gels (Thermo Fisher Scientific), transferred to nitrocellulose membrane

using the Trans-Blot® Turbo™ Transfer System (Bio-Rad) according to the manufacturer's instructions, blocked with 5% Milk in TBS, and subsequently probed. The HA epitope tag was detected using the high-affinity rat anti-HA antibody at a 1:1,000 dilution (Roche 3F10). The FLAG epitope tag was detected using the high-affinity mouse anti-FLAG antibody at a 1:3,000 dilution (Sigma F1804). Hexokinase was detected by rabbit anti-hexokinase antibody at 1:10,000 dilution (United States Biological H2035-02). Licor IRDye700 anti-mouse (Odyssey), IRDye800/IRDye700 anti-rabbit (Odyssey), or IR800 anti-rat (Rockland) secondary antibodies were then used at 1:10,000 dilution. All blots were visualized using the Licor (Odyssey) system. ImageJ was used for protein quantification.

**QPCR**

RNA was isolated from yeast cells in mid log phase as previously described (Ares, 2012). Residual DNA was removed with DNA-free™ DNA Removal Kit (Thermo Fisher Scientific). RNA was converted to cDNA using AMV Reverse Transcriptase (Promega) under standard conditions with oligo dT. Quantitative PCR reactions were prepared with 1X master mix containing 1X Colorless GoTaq® Reaction Buffer (Promega, M792A), dNTPs (0.2 mM each), primers (0.75 μM each), and 1000X SYBR Green with GoTaq® DNA polymerase (Promega, M830B) in 22 μL reactions. Reactions were run on a LightCycler thermal cycler (Roche). Three technical replicas were used for the mRNA quantification in Fig. 1. Primer sequences are listed in supplemental table S3.

**Proteasome inhibition**

MG132 (Sellechem) was dissolved in DMSO and added to yeast cells at 60 μM final concentration. Cells were harvested every 30 min and lysed via bead beating. Clarified lysates were incubated with TEV protease overnight at 4°C to remove any C-terminal extensions (CAT

34

tails). Stalling substrate stabilization was evaluated with quantitative western blot using the Licor (Odyssey) system. The yeast cells used in this experiment had the multidrug transporter gene *PDR5* deleted.

**Myc-Ubiquitin induction**

Myc-tagged ubiquitin was expressed ectopically under copper-inducible promoter (Table S2). Yeast cells were grown for 2h in the presence of 200 μM $Cu_2SO_4$ to induce Myc-Ubiquitin expression. MG132 proteasom inhibitor was added at 60 μM final concentration. Cells were harvested via centrifugation in 1 hour.

**Denaturing Immunoprecipitation**

Stalling substrates were immunoprecipitated as previously described (Bengtson and Joazeiro, 2010). Briefly, yeast cells were treated with 0.1 M NaOH for 5 min at room temperature, pelleted, resuspended and boiled in the presence of 1% SDS, 50 mM Tris-HCl pH 7.5, 5 mM EDTA, 5 mM NEM and 1x Halt[TM] Protease Inhibitor Cocktail (ThermoFisher Scientific). 100 μl of boiled lysate was further diluted with 900 μl of IP buffer (50 mM Tris-HCl pH 7.5, 250 mM NaCl, 5 mM EDTA, 0.5% NP40, 1 mM NEM), and incubated with 25 ul bead-conjugated anti-Flag antibody (Anti-FLAG® M2 Magnetic Beads, Sigma-Aldrich) or 20 ul bead-conjugated anti-HA antibody (Pierce[TM] Anti-HA Magnetic Beads, ThermoFisher Scientific) for 2h at room temperature. After three washes in IP buffer, beads were boiled in 2x Laemmli Buffer with no BME, and the eluates were used for immunoblot.

Tables

**Table 2-1.** Yeast strains used in this study.

| name | strain | Genotype | source |
|---|---|---|---|
| BY4741 | BY4741 | *MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0* | this study |
| *pdr5Δ* | yKK205 | BY4741 *pdr5Δ::KanMX* | this study |
| *ltn1Δ* | yKK178 | BY4741 *ltn1Δ::KanMX* | this study |
| *rqc2Δ* | yKK179 | BY4741 *rqc2Δ::KanMX* | this study |
| *RQC2mut/ltn1Δ* | yKK272 | BY4741 *rqc2Δ::RQC2mut ltn1Δ::KanMX* | this study |
| *rqc2Δ/ltn1Δ* | yKK287 | BY4741 *rqc2Δ::KanMX ltn1Δ::URA3* | this study |
| *RQC2mut* | yKK108 | BY4741 *rqc2Δ::RQC2mut* | this study |
| *ltn1ΔRING* | yKK102 | BY4741 *Ltn1ΔRING::NatMX* | Ref. 11 |
| *pdr5Δ/ltn1Δ* | yKH16 | BY4741 *pdr5Δ::Kan ltn1Δ::Nat* | this study |
| *pdr5Δ/rqc2Δ* | yKK226 | BY4741 *pdr5Δ::Kan rqc2Δ:: Nat* | this study |
| *pdr5Δ/RQC2mut* | yKH21 | BY4741 *pdr5Δ::Kan rqc2Δ::RQC2mut* | this study |
| *pdr5Δ/ltn1Δ/RQC2mut* | yKH22 | BY4741 *pdr5Δ::Kan ltn1Δ::Nat rqc2Δ::RQC2mut* | this study |
| *ski2Δ* | yKK327 | BY4741 *ski2Δ::Ura* | this study |
| *ski2Δ/RQC2mut* | yKK328 | BY4741 *ski2Δ::Ura rqc2Δ::RQC2mut* | this study |
| *ski2Δ/rqc2Δ* | yKK329 | BY4741 *ski2Δ::Ura rqc2Δ::Kan* | this study |
| *ski2Δ/ltn1Δ* | yKK330 | BY4741 *ski2Δ::Ura ltn1Δ::Kan* | this study |

**Table 2-2.** Plasmids used in this study.

| Name | ID | Backbone | Expression | Source |
|---|---|---|---|---|
| $GFP_{20Lys}$ | pKK148 | pRS313 | pTDH3-$GFP_{20Lys}$-3xHA-TEV-R12-RFP | this study |
| $GFP_{Lys-free}$ | pKK135 | pRS313 | pTDH3-$GFP_{Lys-free}$-3xHA-TEV-R12-RFP | this study |
| $GFP_{20Lys}$ (37 linker) | pKK168 | pRS313 | pTDH3-$GFP_{20Lys}$-3xHA-TEV(no linkers) - R12-RFP | this study |
| $GFP_{19Lys}$ (51 linker) | pKK140 | pRS313 | pTDH3-$GFP_{19Lys}$(K238R)-3xHA-TEV-R12-RFP | this study |
| 3xFLAG XTEN0 | pKK151 | pRS313 | pTDH3-$GFP_{Lys-free}$-3xFlag-TEV-R12-RFP | this study |
| 3xFLAG XTEN10 | pKK190 | pRS313 | pTDH3-$GFP_{Lys-free}$-3xFlag-XTEN10-TEV-R12-RFP | this study |
| 3xFLAG XTEN20 | pKK191 | pRS313 | pTDH3-$GFP_{Lys-free}$-3xFlag-XTEN20-TEV-R12-RFP | this study |
| 3xFLAG XTEN40 | pKK192 | pRS313 | pTDH3-$GFP_{Lys-free}$-3xFlag-XTEN40-TEV-R12-RFP | this study |
| 3xFLAG XTEN80 | pKK198 | pRS313 | pTDH3-$GFP_{Lys-free}$-3xFlag-XTEN80-TEV-R12-RFP | this study |
| $His3_{Lys-free}$ | pKK172 | pRS316 | pTDH3-$His3_{Lys-free}$-3xHA-TEV-R12-RFP | this study |
| $His3_{Lys-free}$ 3xFLAG | pKK175 | pRS316 | pTDH3-$His3_{Lys-free}$-3xFlag-TEV-R12-RFP | this study |
| $His3_{Lys-free}$ 2xLys | pKK174 | pRS316 | pTDH3-$His3_{Lys-free}$-3xHA-2Lys-TEV-R12-RFP | this study |
| $GFP_{Lys-free}$ 2xLys | pKK136 | pRS313 | pTDH3-$GFP_{Lys-free}$-3xHA-TEV-2Lys-R12-RFP | this study |
| 1K XTEN0 | pKK218 | pRS313 | pTDH3-$GFP_{Lys-free}$-3xHA-1K-TEV-R12-RFP | this study |
| 1K XTEN10 | pKK219 | pRS313 | pTDH3-$GFP_{Lys-free}$-3xHA-1K-XTEN10-TEV-R12-RFP | this study |
| 1K XTEN20 | pKK220 | pRS313 | pTDH3-$GFP_{Lys-free}$-3xHA-1K-XTEN20-TEV-R12-RFP | this study |
| 1K XTEN40 | pKK221 | pRS313 | pTDH3-$GFP_{Lys-free}$-3xHA-1K-XTEN40-TEV-R12-RFP | this study |
| 2K XTEN20 | pKK242 | pRS313 | pTDH3-$GFP_{Lys-free}$-3xHA-2K-XTEN20-TEV-R12-RFP | this study |
| 2K XTEN25 | pKK243 | pRS313 | pTDH3-$GFP_{Lys-free}$-3xHA-2K-XTEN25-TEV-R12-RFP | this study |
| 2K XTEN30 | pKK244 | pRS313 | pTDH3-$GFP_{Lys-free}$-3xHA-2K-XTEN30-TEV-R12-RFP | this study |
| 2K XTEN35 | pKK245 | pRS313 | pTDH3-$GFP_{Lys-free}$-3xHA-2K-XTEN35-TEV-R12-RFP | this study |
| NSD $GFP_{Lys-free}$ | pKK189 | pRS313 | pTDH3-$GFP_{Lys-free}$-3xHA_TEV_(AAA)x10_Rz | this study |
| RQC2 | pKK187 | pRS426 | pPGK-RQC2 | this study |
| RQC2mut | pKK182 | pRS426 | pPGK-RQC2mut | this study |
| RQC2-HA | pKK183 | pRS426 | pPGK-RQC2-HA | this study |
| RQC2mut-HA | pKK184 | pRS426 | pPGK-RQC2mut-HA | this study |
| Myc-Ub | pUb221 | | pCUP1-3xHis-Myc-Ub | Finley Lab |
| $GFP_{20Lys}$-3xFLAG | pKK150 | pRS313 | pTDH3-$GFP_{20Lys}$-3xFLAG-TEV-R12-RFP | this study |
| unfolded GFP | pKK260 | pRS313 | pTDH3-DHFR-$GFP_{Lys-free}$(100AA)-3xHA-1K-XTEN40-TEV-R12-RFP | this study |

**CHAPTER THREE**

The role of CAT-tailing in degradation of nascent polypeptide stalled at the endoplasmic

reticulum

## Introduction

Although recent studies have advanced our understanding of the role of the RQC in degrading nascent polypeptides stalled in the cytoplasm, far less is known of how the complex engages ribosomes stalled on the endoplasmic reticulum (ER). Approximately one-fourth of the synthesized proteins are co-translationally targeted to organelles (Costa et al., 2018). Therefore, it is likely that a substantial fraction of the translational stall will occur when a ribosome is engaged at the ER translocon. Ribosome stalling at the ER poses its own challenges. First, the highly crowded region around the ribosome exit tunnel (Pfeffer et al., 2014) may sterically preclude Ltn1p access to the nascent polypeptide engaged at the translocon. Second, the ubiquitination and extraction of these nascent peptides is further hindered by the observation that secretory proteins are transferred from the ribosomal exit tunnel into the Sec61 channel with little or no cytosolic exposure (Park and Rapoport, 2012).

Recent studies have shown that the RQC can also target translocon-engaged 60S subunits on the (ER), and that the stalled nascent polypeptides can be ubiquitinated by Ltn1p (Malsburg et al., 2015). However, whether and how CAT-tailing facilitates nascent peptide degradation at the ER has not been explored. In Chapter two, I have established lysine availability as a key determinant for nascent polypeptide degradation. The access to lysine residues by Ltn1p may be particularly challenging for substrates stalled at the ER, as the majority of the translated polypeptide will be sequestered within the ER lumen.

## Results

To better understand RQC-mediated degradation at the ER, I designed a stalling substrate that is co-translationally translocated into the ER lumen (Jan et al., 2014). I used the coding

sequence of the ER chaperone Pdi1p, followed by a 3xFLAG tag, a TEV protease site, a stalling sequence (6R), an HA tag and an HDEL ER-retention sequence (Fig. 3-1A). As a control, I constructed a Pdi1p substrate without the stalling sequence. We expressed the two constructs in wt cells and compared the steady-state protein levels in total cell lysate and cell media (Fig. 3-1B). When ribosomes stall at the stalling sequence, the C-terminal ER retention signal is not translated. Thus, stalling products not degraded by the RQC complex will escape the ER and be secreted in the media. The lower levels of the stalling substrate compared to the control Pdi1p in both cell lysate and media suggest that the ribosome stalls at the 6R sequence and that the nascent polypeptide is efficiently degraded.

I then compared the levels of Pdi1p 6R secreted in the media by wt cells or cells with deletions of RQC components (Fig. 3-1C). I detected accumulation of stalling product in the media of *ltn1Δ* and *rqc2Δ* cells. These nascent polypeptides were sensitive to Endoglycosidase H, indicating that they were glycosylated and trafficked through the secretory pathway. In addition, deletion of *LTN1* led to the accumulation of heavily CAT-tailed species of Pdi1p, confirming that CAT-tailing is not limited to cytoplasmic substrates. The length of the observed CAT tails, up to ~350 amino acids (estimated by observed molecular weight shift of ~40 kDa), was substantially longer than any reported cytoplasmic CAT tails (maximum of ~50 amino acids). However, these long CAT tails did not affect protein stability in the ER, as CAT-tailed Pdi1p accumulated to the same extent in the media as the non-CAT-tailed Pdi1p from *rqc2Δ* cells (Fig. 3-1C). Furthermore, the stalling construct was not stabilized by deletions of various components of the ER-associated degradation pathway (Fig. 3-2A), suggesting that the CAT tails do not serve as a degron in the ER.

When a ribosome stalls at the ER, only about 10-15 amino acids of the translocating nascent polypeptide are exposed on the cytoplasmic side, accessible to Ltn1p (Malsburg et al., 2015). To explore the role of CAT-tailing in nascent polypeptide degradation at the ER, I modified the Pdi1p 6R stalling construct to remove any lysines from the suggested Ltn1p-accessible region by replacing the 3xFLAG tag with a 3xHA tag and mutating the C-terminal lysine of Pdi1p to arginine (Fig. 3-1D). I then compared the accumulation of the original (Lys-rich linker) and modified (Lys-free linker) Pdi1p 6R in media of wt, Rqc2p$^{mut}$, and *rqc2Δ* cells (Fig. 3-1E). The construct with a Lys-rich linker showed largely CAT tail-independent degradation. However, the construct that had Lys-free linker was stabilized in Rqc2p$^{mut}$ cells, but was still efficiently degraded in wt cells, in a manner dependent on the functional RING domain of Ltn1p (Fig. 3-1E, 3-2B).

Based on these results, I hypothesized that CAT tails can serve as a cytosolically accessible flexible loop that facilitates retrotranslocation (Wiertz et al., 1996) and subsequent degradation by exposing lysines "hidden" in the ER. To further explore the role of the CAT tail as a flexible loop, I introduced a genetically encoded CAT tail (containing 60 alanine and threonine residues) at the C-terminus of Pdi1p and expressed the construct in wt, Rqc2p$^{mut}$, and *rqc2Δ* cells (Fig. 3-1E). I observed a significant increase in Pdi1p degradation in Rqc2p$^{mut}$ cells (Fig. 3-1E, bar graph). The lack of complete restoration of degradation could be due to the limited length of the artificial CAT tail relative to the long endogenous ones observed for ER substrates. Degradation in the Rqc2p$^{mut}$ strain was blocked when the genetically encoded CAT tail was sequestered in the ER by introducing a strong glycosylation site (Fig. 3-1E, 3-2C, E). This result suggests that the CAT tail-mediated degradation relies on the CAT tail being available on the cytoplasmic side, where it could engage cytoplasmic molecular machinery

capable of extracting polypeptides out of the ER. Retrotranslocation of the stalled polypeptide would present lysines previously sequestered in the ER to Ltn1p.

Conclusions

My work shows that in the context of ribosomes stalled on the translocon, CAT tails can serve as flexible linkers that may alleviate steric hindrance and facilitate retrotranslocation, exposing lysines sequestered in the ER lumen. It is tempting to speculate that the low complexity CAT tail is recognized by molecular motors, such as Cdc48p, that can facilitate extraction from the ER. Indeed, Cdc48p has already been implicated in extracting faulty polypeptides from the ER lumen as part of the ERAD pathway (Wolf and Stolz, 2012). However, how the CAT tails are recognized and what is the minimum length of the linker that can result in efficient extraction is yet to be determined.

Surprisingly, in my study we were able to detect CAT tails that are nearly as long as the protein they are attached to. However, these long unstructured regions were still able to pass through the secretory pathway without triggering degradation by the ERAD. This observation might tell us something about the interaction of the CAT tails with other proteins and with each other. Although the presence of CAT tails has been linked to protein aggregation, these long extensions seem not to affect protein solubility in the context of the ER. Indeed, if a stalled protein fails to be degraded when engaged at the Sec61p translocon, the addition of neutral CAT tails could in fact help solubility, prevent sequestration of chaperones and allow a faulty polypeptide without an ER retention sequence to be secreted into the media.
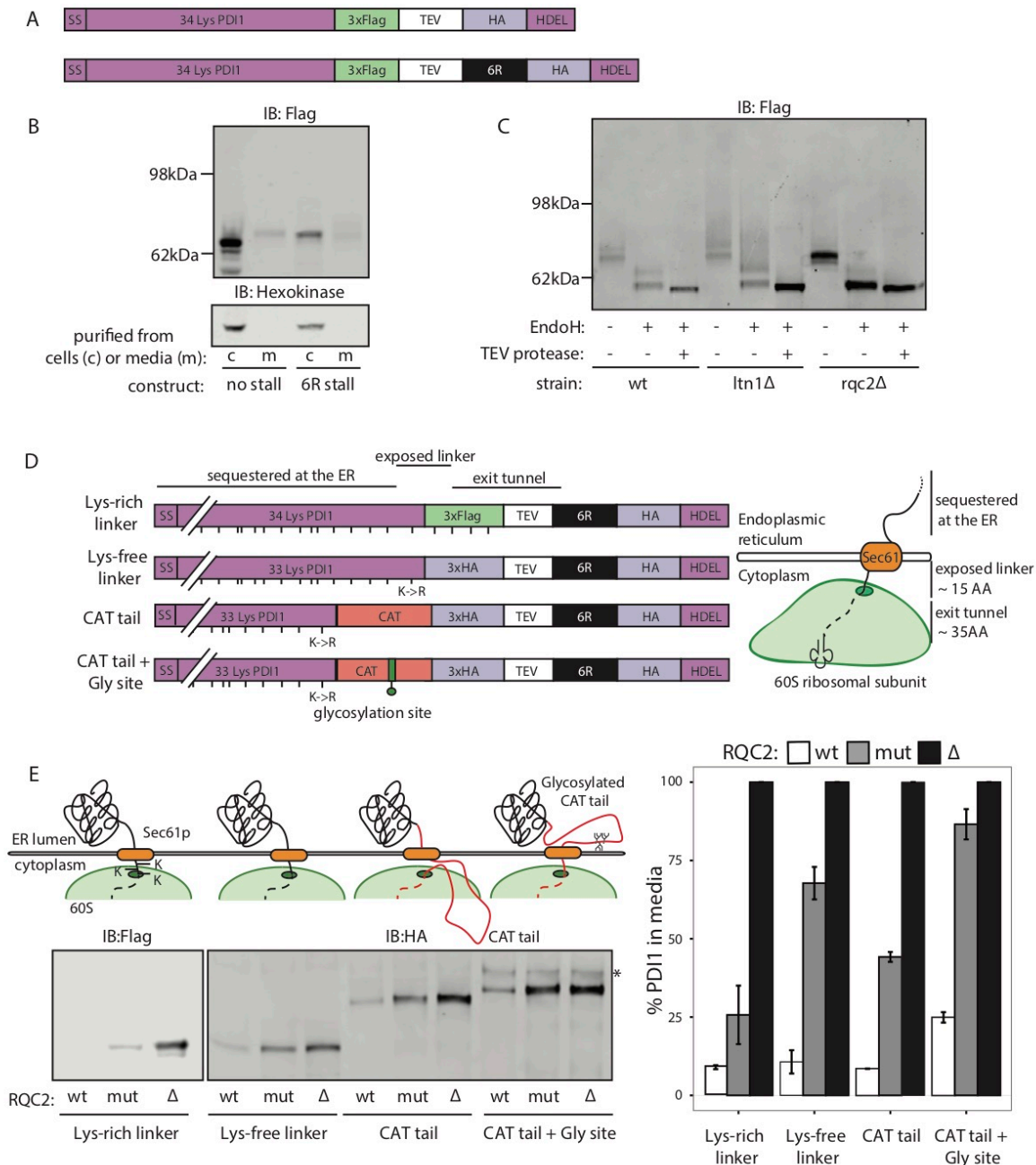
**Fig. 3-1. CAT tail dependence for secretory RQC substrates.** (**A**) Pdi1p as a stalling construct at the ER (**B**) IB of tagged Pdi1p with and without 6R stalling sequence from total cell lysate (c) and media (m). (**C**) IB of Pdi1p 6R purified from media from wt, *ltn1Δ*, and *rqc2Δ* cells. The purified protein was incubated with TEV protease (+) or mock digested (-). Samples were also deglycosylated (+EndoH). (**D**) Pdi1p stalling constructs. Vertical lines mark the location of lysines (K). (**E**) IB of Pdi1p stalling constructs shown in (D). Protein was purified from the media of cells with wt, mut or Δ Rqc2p and incubated with TEV protease and Endoglycosidase

H. Pdi1p was analyzed by SDS-PAGE and IB (left) and quantified by densitometry. The bar graph shows the amount of Pdi1p accumulating in media relative to *rqc2Δ* for each construct. Error bars represent s.d. from two independent experiments. Asterisk (*) indicates non-specific bands.
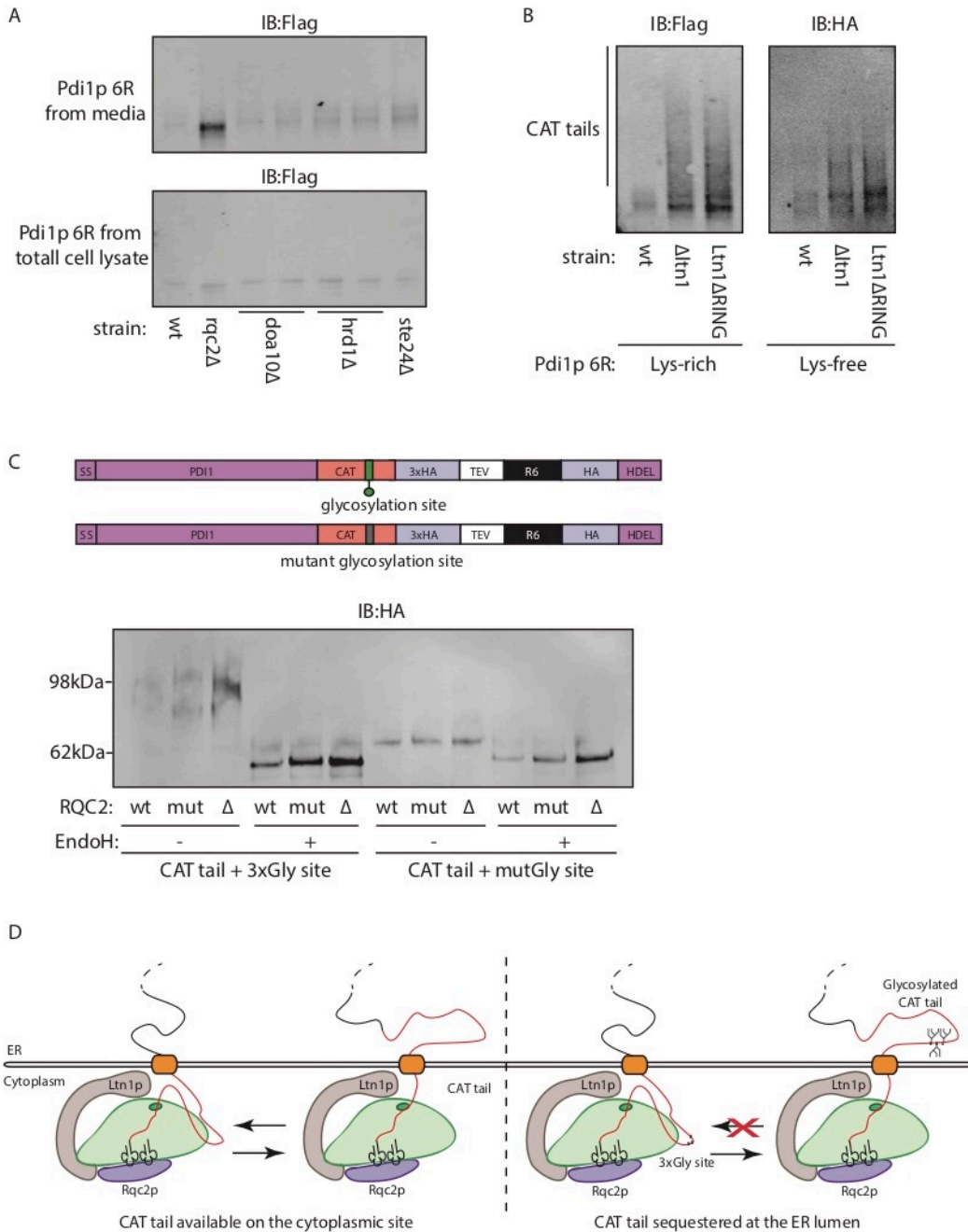
**Fig. 3-2. CAT-tailing at the endoplasmic reticulum. (A)** Pdi1p 6R stalling construct was expressed in wild type (wt) cells and cells with deletions of the RQC component (*rqc2Δ*) or ERAD components (*doa10Δ*, *hrd1Δ*, *ste24Δ*). Protein from total cell lysate or cell-free media was analyzed on SDS-PAGE and IB. Note: the IB shows two biological replicas for *Δdoa10* and *Δhrd1* strains. **(B)** Pdi1p 6R stalling constructs with lysine-free or lysine-rich linker were expressed in wt, *ltn1Δ*, and *Ltn1ΔRING* cells. Protein was precipitated from cell-free media and analyzed on SDS-PAGE and IB. **(C)** Pdi1p 6R stalling construct with a CAT tail containing a strong glycosylation site (+ 3xGly) or a mutated site (+ mutGly) was expressed in wt, mut or Δ

Rqc2p cells. Pdi1p 6R was purified from cell-free media and incubated with Endoglycosydase H (+) or mock digested (-). All samples were incubated with TEV protease. Samples were analyzed on SDS-PAGE and IB. (**D**) Model of CAT tail movement through the Sec61p translocon. An unmodified CAT tail (such as CAT + mutGly) can freely toggle in and out of the ER lumen (left). A CAT tail harboring a glycosylation site (CAT + 3xGly) is modified once it translocates into the ER (right). The bulky sugar modification does not allow retrotranslocation of the CAT tail to the cytoplasmic site.

Materials and Methods

**Yeast strains**

Pdi1p constructs were integrated genomically in the LEU2 locus and were driven by a Sec63

promoter.

**Protein purification from media**

Yeast cells were grown to mid log and 600ul of cell culture at $OD_{600} = 1$ was harvested. The

cells were pelleted and the cell-free media was subjected to Chloroform/Methanol precipitation.

**Endo H treatment**

Samples were treated with Endoglycosidase H (NEB) according to manufacturer's instructions.

**CHAPTER FOUR**

Computationally and experimentally evaluating the role of CAT-tailing in degradation of

endogenous substrates

Introduction

The majority of RQC studies in both yeast and mammals used model stalling construct, such as mRNAs containing polybasic stretches (Juszkiewicz and Hegde, 2017a; Sundaramoorthy et al., 2017), non-optimal codon pairs (Brandman et al., 2012), stem loops (Doma and Parker, 2006), as well as messages lacking a stop codon (Choe et al., 2016a). However, the frequency of these events *in vivo* has not been established, and ribosome profiling experiments in yeast strains lacking ribosome recycling factors are only now emerging. While the sites of endogenous stalling remain poorly defined, a number of processes, such as mRNA fragmentation, oxidative damage (Simms et al., 2014), premature polyadenylation, or stress from translation inhibitors can cause stalling at any position along a message. Having established lysine positioning as a critical determinant of CAT-tail-dependent degradation, I computationally evaluated the frequency with which endogenous substrates would be expected to rely on CAT-tailing for efficient degradation.

Results

For simplicity, I considered the nascent chains produced if ribosomes stall with uniform probability at each codon along every annotated coding sequence in the yeast genome (Engel et al., 2014). I calculated the fraction of potential stalling sites for which there are no lysines accessible to Ltn1p but at least one lysine "hidden" in the ribosome exit tunnel for ranges of possible combinations of ribosome exit tunnel length and number of amino acids outside the exit tunnel accessible by Ltn1p (Fig. 4-1). Based on my experimental estimates of Ltn1p's reach (~12 amino acids) and previous measurements of the length of the ribosome exit tunnel (~35 amino acids) (Ban et al., 2000; Harms et al., 2001; Nissen et al., 2000), I estimate that up to 40% of potential stalling sites will generate nascent polypeptides that will depend on CAT tail elongation

for efficient degradation (Fig. 4-1). Therefore, CAT-tailing substantially increases the fraction of RQC-degradable substrates from ~60% to ~ 95% (Fig. 4-2).

A different class of RQC substrates results from mRNAs devoid of in-frame stop codons due to premature polyadenylation. As shown in Chapter two, CAT-tailing can facilitate the degradation of these substrates by exposing the lysines produced by translation of the poly(A) tail. Assuming, for simplicity, that premature polyadenylation is independent of message identity, I calculated the fraction of premature poly(A) sites that will lead to substrates with no lysines within Ltn1p's reach as a function of the number of amino acids outside the exit tunnel and accessible by Ltn1p (Fig. 4-3, black line). My analysis suggests that ~35% of such substrates do not have any lysines available for Ltn1p-mediated ubiquitylation, and therefore their degradation would be CAT-tail dependent.

Finally, I calculated the fraction of stall positions on messages encoding secreted or transmembrane proteins (Ast et al., 2013) that lead to a nascent polypeptide without any lysine accessible to Ltn1p (Fig. 4-3, purple line). Due to the higher hydrophobicity of these proteins relative to the soluble proteome, a larger fraction of the stalling sites, nearly 50%, have no lysines within Ltn1p's reach. Therefore, these stalling products will depend on the addition of CAT tails for extraction from the ER and subsequent degradation.

If correct, this estimate suggests that for cells in which CAT tail formation is compromised by the Rqc2p$^{mut}$, the accumulation of non-degradable endogenous substrates will lead to growth defects. A direct comparison of the growth rates between cells with functional RQC and CAT-tail deficient RQC is hindered by the lack of growth phenotype of yeast lacking RQC components under standard growth conditions (Choe et al., 2016b; Defenouillère et al., 2013). However, strains with deletions of RQC components have been previously shown to have

increased sensitivity to the elongation inhibitor cycloheximide (CHX) (Alamgir et al., 2010) and loss of exosome components (Defenouillère et al., 2013). Therefore, I hypothesized that inducing widespread stochastic ribosome stalling by cycloheximide (CHX) treatment, combined with mRNA stabilization in a *ski2Δ* background would cause a CAT-tail dependent growth phenotype. Indeed, under these conditions, *rqc2^{mut}* cells exhibited an intermediate growth defect between *wt* and RQC-deletion cells (*ltn1Δ* or *rqc2Δ*), consistent with the hypothesis that a significant fraction of the stalled nascent polypeptides in these cells are not efficiently degraded without CAT-tailing (Fig. 4-4).

Conclusions

The described above analysis for the first time estimates the fraction of stalling substrates that relies on CAT-tailing for efficient degradation. It is believed that, at least in yeast, ribosome stalling is stochastic and can occur on any message, anywhere along the message. Therefore, any mRNA could become a substrate of the RQC pathway. As a result, the RQC pathway needs to cope with a diverse set of substrates that are not always optimal for degradation, either due to the lack of lysines in the Ltn1p accessible region, or due to steric hindrance. However, the ability of the complex to add CAT tails to the stalled nascent polypeptide alleviates some of these restrictions and allows the pathway to degrade almost any stalled polypeptide it encounters.
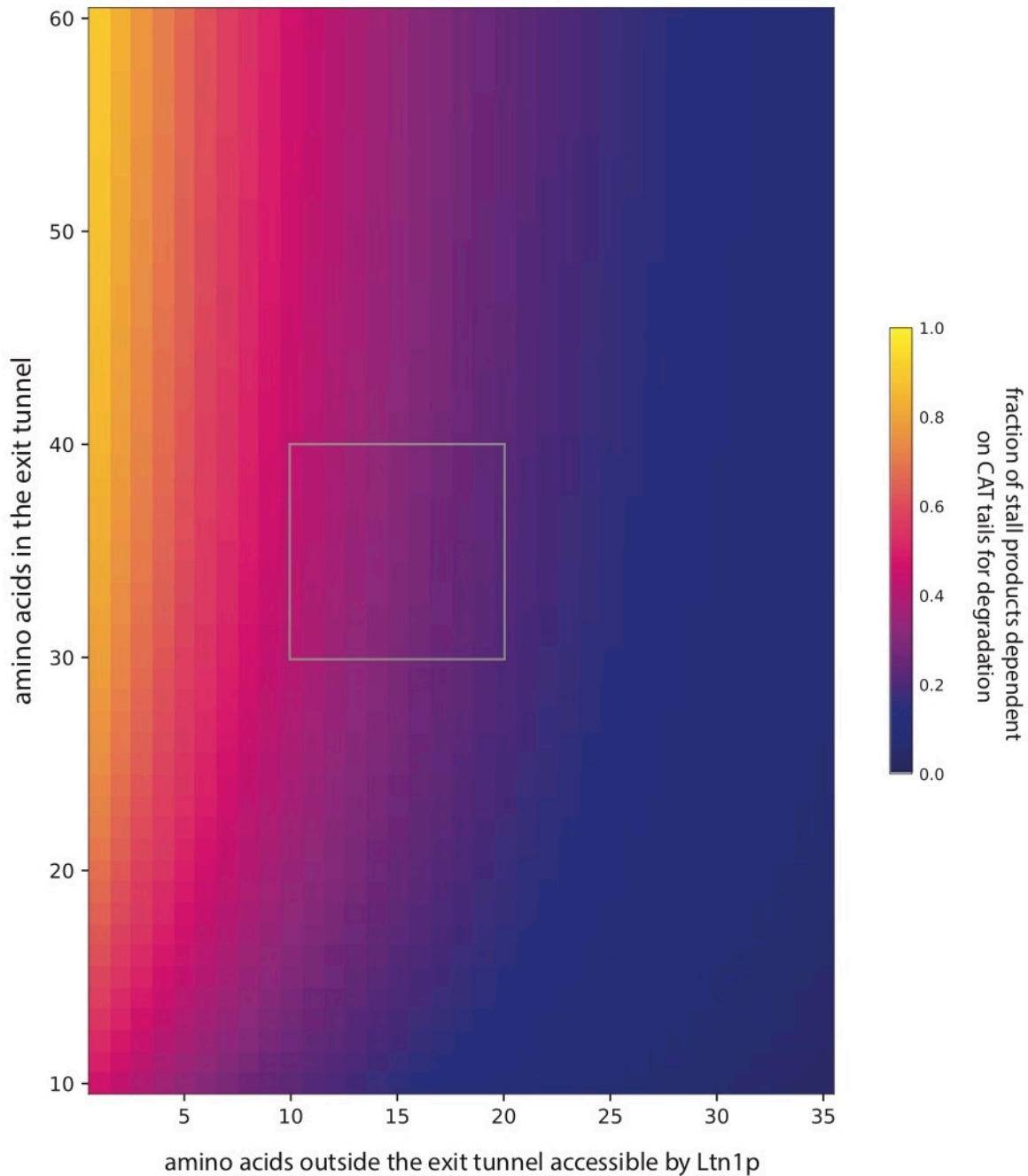
**Fig. 4-1. CAT tail-dependent degradation of endogenous RQC substrates.** Heat map showing the fraction of potential stalling sites for which there is no lysine accessible by Ltn1p but at least one lysine "hidden" in the ribosome exit tunnel, as a function of the number of amino acids outside the exit tunnel within Ltn1p's reach and the number of amino acids sequestered in the exit tunnel. The grey box highlights a subset of the parameter space most consistent with the experimental measurements of Ltn1p reach and ribosome exit tunnel length.
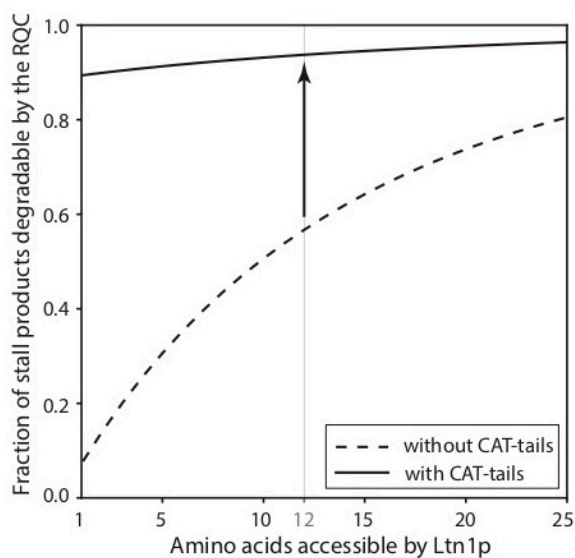
**Fig. 4-2**. **CAT-tailing increases the number of RQC degradable substrates.** The fraction of stalling positions leading to a RQC-degradable nascent polypeptide in the presence (solid line) or absence (dashed line) of CAT tails as a function of the number of amino acids accessible to Ltn1p and a fixed exit tunnel length of 35 amino acids. The arrow shows the increase of RQC-degradable substrates in the presence of CAT tails at the estimated Ltn1p reach of 12 amino acids.
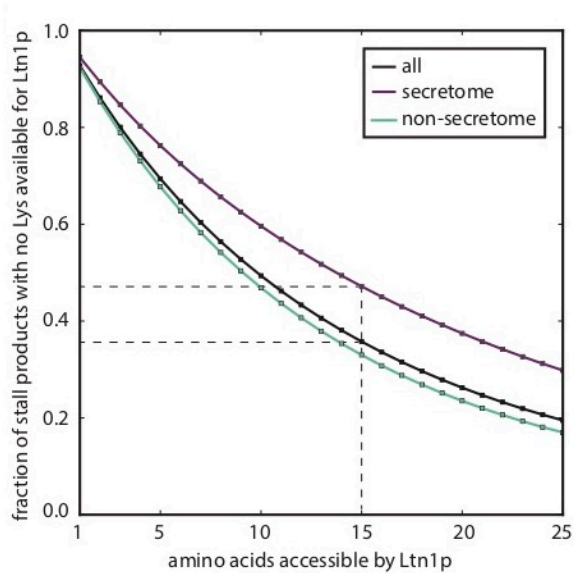
**Fig. 4-3. CAT tail-dependent degradation of cytoplasmic and transmembrane RQC substrates.** The fraction of stalling products with no lysine in Ltn1p's reach as a function of the number of amino acids outside the exit tunnel accessible to Ltn1p for all yeast coding sequences (black), the yeast secretome (purple), and non-secretome (green). Dotted lines indicate the average number of amino acids accessible to Ltn1p and the corresponding fractions of stall products requiring CAT-tailing for degradation.
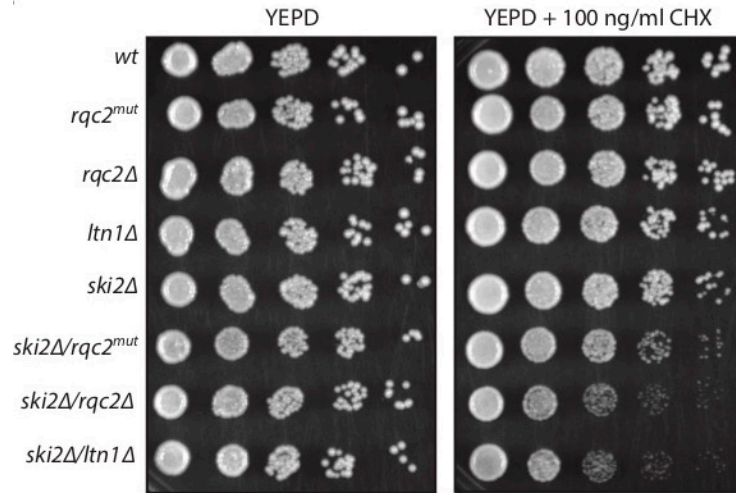
**Fig. 4-4. Growth defect of yeast cells unable to CAT-tail.** Wild-type (*wt*) and RQC mutant cells were grown to exponential phase in liquid YEPD medium, serially diluted fivefold and spotted on YEPD plates with or without 100 ng/ml cycloheximide (CHX).
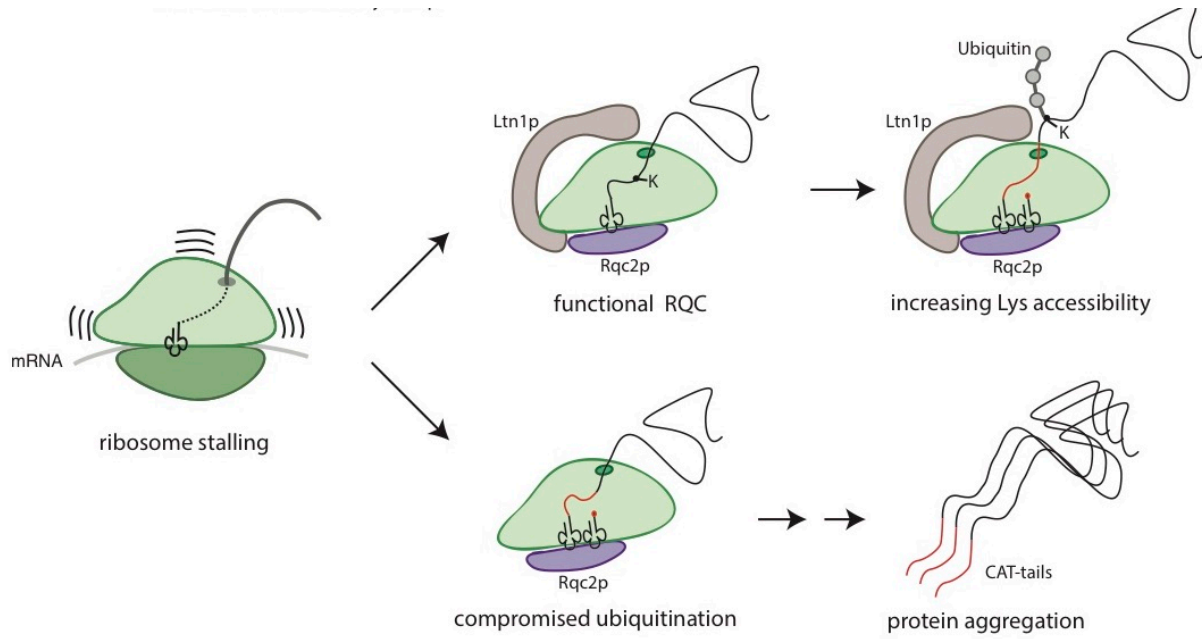
**Fig. 4-5. Model for the function of CAT tails *in vivo*.** If the RQC is not functional, the addition of CAT tails could lead to protein aggregation and proteotoxic stress (bottom). However, in the context of the intact complex, the addition of CAT tails can expose Lys residues sequestered in the ribosome exit tunnel, enabling their ubiquitination and the degradation of the stalled nascent polypeptide (top).

<u>Materials and Methods</u>

**Frogging**

Cells from *wt* and RQC mutant strains were grown to exponential phase in liquid YEPD medium, serially diluted fivefold and spotted onto YEPD plates with or without cycloheximide (CHX) (100 ng/ml). Plates without CHX were incubated for 2 days and with CHX for 3 days at 30°C.

**Lys availability prediction**

Code and data used to calculate lysine availability is included as a supplementary file S1 in the published manuscript (Kostova et al., 2017).

**CHAPTER FIVE**

Mammalian stalling reporter

<u>Introduction</u>

The study of the mammalian RQC pathway has been hindered by the lack of a model stalling substrate. Traditionally, studies in yeast have utilized a substrate containing a stretch of non-optimal codon pairs (CGA-CGG) sandwiched between two fluorescent proteins (GFP and RFP) (Dimitrova et al., 2009). These pairs cause stalling because in yeast they are decoded via a purine-purine I·A wobble base pairing between the tRNA anticodon and the mRNA codon (Gamble et al., 2016). However, these codon pairs do not cause ribosome stalling in higher eukaryotes, likely because they are now decoded by a designated tRNA that forms Watson-Crick pairing with the mRNA codon.

Messages that do not contain an in-frame stop codon, e.i. non-stop decay substrates, are an alternative source of stalling substrates. These messages could result from premature polyadenylation, mutations in the stop codon, or fragmentation of the mRNA. These substrates have been extremely valuable in studying the non-stop decay pathway in yeast and are usually artificially stabilized in cell by using strains defective in RNA degradation, such as strains harboring deletions in various exosomal components. However, using these substrates in mammalian cells has proven challenging. First, these messages are rapidly detected and degraded by the mammalian exosome, and as a result the substrate levels are extremely low. Second, stabilizing the substrate by inhibiting the exosome is not possible since this complex is essential in higher eukaryotes.

We aimed at developing a mammalian RQC reporter that induces robust ribosome stalling and is expressed at high enough levels to allow for high throughput genetic screening.

<u>Results</u>

First, we developed a flow cytometry-based assay to quantitatively ribosome stalling at

single-cell resolution in mammalian cells. We designed a series of reporters containing three different fluorescent proteins – BFP, GFP, and RFP (Fig. 5-1A). We introduced a viral t2a sequence between the BFP and the GFP. This sequence prevents the formation of a peptide bond without interrupting translation elongation, resulting in two independent fluorescent proteins. Therefore, the BFP fluorescence can be used as a proxy for the mRNA levels and the translation efficiency of the construct, without interfering with RQC engagement. We introduced various stalling sequences between the GFP and the RFP, which will be explained in more detail below. Complete translation of the reporter with no ribosome stalling, will generate three proteins (BFP, GFP, and RFP) in equal amounts. However, if the ribosome terminally stalls at the introduced stalling sequence, no RFP would be made, since the ribosome cannot proceed with translation. In addition, the GFP will also be targeted for degradation since it is the product of a stalled ribosome. Therefore, in wild type cells we will observe high BFP levels and low GFP and RFP levels.

First, we used a well-characterized stem loop that has been shown to block the elongating ribosome in yeast (Doma and Parker, 2006) (Fig. 5-1). However, for this construct we obtained stoichiometric BFP:GFP:RFP rations, indicative of ribosome read through. In fact, the levels of the three fluorescent proteins were identical to the ones obtained from a control reporter with no stalling sequence. Therefore, although the stem loop can serve as a strong stalling sequence in yeast, the mammalian ribosome is more processive and can easily melt the strong RNA structure.

We also designed a construct that contained a mammalian pausing sequence PPE (Ingolia et al., 2011). However, this sequence did not induce any ribosome stalling. This observation suggests that ribosome pausing and ribosome stalling are two different events during translation that can trigger distinct cellular responses.

We also introduced various RNA pseudoknots with viral origin that are believed to serve as roadblocks for the ribosome (Staple et al., 2005). The human ribosome successfully translated through all engineered pseudoknots leading to no change in the GFP:RFP ratio (Fig. 5-2C)

In order to introduce terminal stalls in the constructs, we cloned the U1 and U2 motives and the tRNA-like structure from the MALAT1 long non-coding RNA (Wilusz et al., 2012). The 5' leader sequence of the MALAT1 tRNA structures is recognized by ribonuclease P (RNase P), which leads to the excitation of the tRNA from the message, leaving a reporter lacking a stop codon, UTR, or polyA tail. After excision of the tRNA-like structure, the upstream structure, containing the two motives, folds into a unique triple helix (Wilusz et al., 2012) that protects the 3' end of the reporter from exosomal degradation. Therefore, this construct is stably expressed in cells despite the lack of a polyA tail. The MALAT1 t-RNA-like reporter led to significantly lower GFP levels compared to the control reporter (Fig. 5-1 B, C, E), indicative of stalling. In addition, the GFP fluorescence was stabilized approximately two-fold in cells lacking NEMF, indicating that the RQC facilitates the degradation of the construct (Fig. 6-1).

Conclusions

We have engineered a robust fluorescent stalling reporter that can be used in mammalian cells lines to evaluate the function of the mammalian RQC, as well as to identify mammalian-specific RQC components. While this work was being conducted, two papers were published describing an alternative mammalian stalling reporters (Juszkiewicz and Hegde, 2017b; Sundaramoorthy et al., 2017). That reporter contained a polyA stretch (21 AAA codons), separated from GFP and RFP by t2a ribosome skipping sequences. Although this construct has proven useful in studying the early stages of stalled ribosome detection, it has several limitations. First, it contains an extremely long homopolymeric stretch that is not found in mammalian

messages. Second, the cause of ribosome stalling is not clear, i.e. it is not known whether the tail stretch binds polyA-binding proteins (PABs) that block the ribosome or the translation of the AAA codons is inherently slow. Finally, translation through the polyA sequence leads to ribosome slippage, which changes the reading frame of the ribosome, and therefore low RFP fluorescence that could be misinterpreted as stalling. Our construct contains a fundamentally different type of stall, i.e. a non-stop decay mRNA on which the ribosomes translates to the end of the message without encountering a stop codon. Therefore, it will be valuable to compare the dependencies of these two reporters on mammalian RQC components. It is important to note that neither the polyA reporter, nor the MALAT1 one contain stalling sites found in mammalian messages. Therefore, a key future direction is to explore the endogenous clients of the RQC.
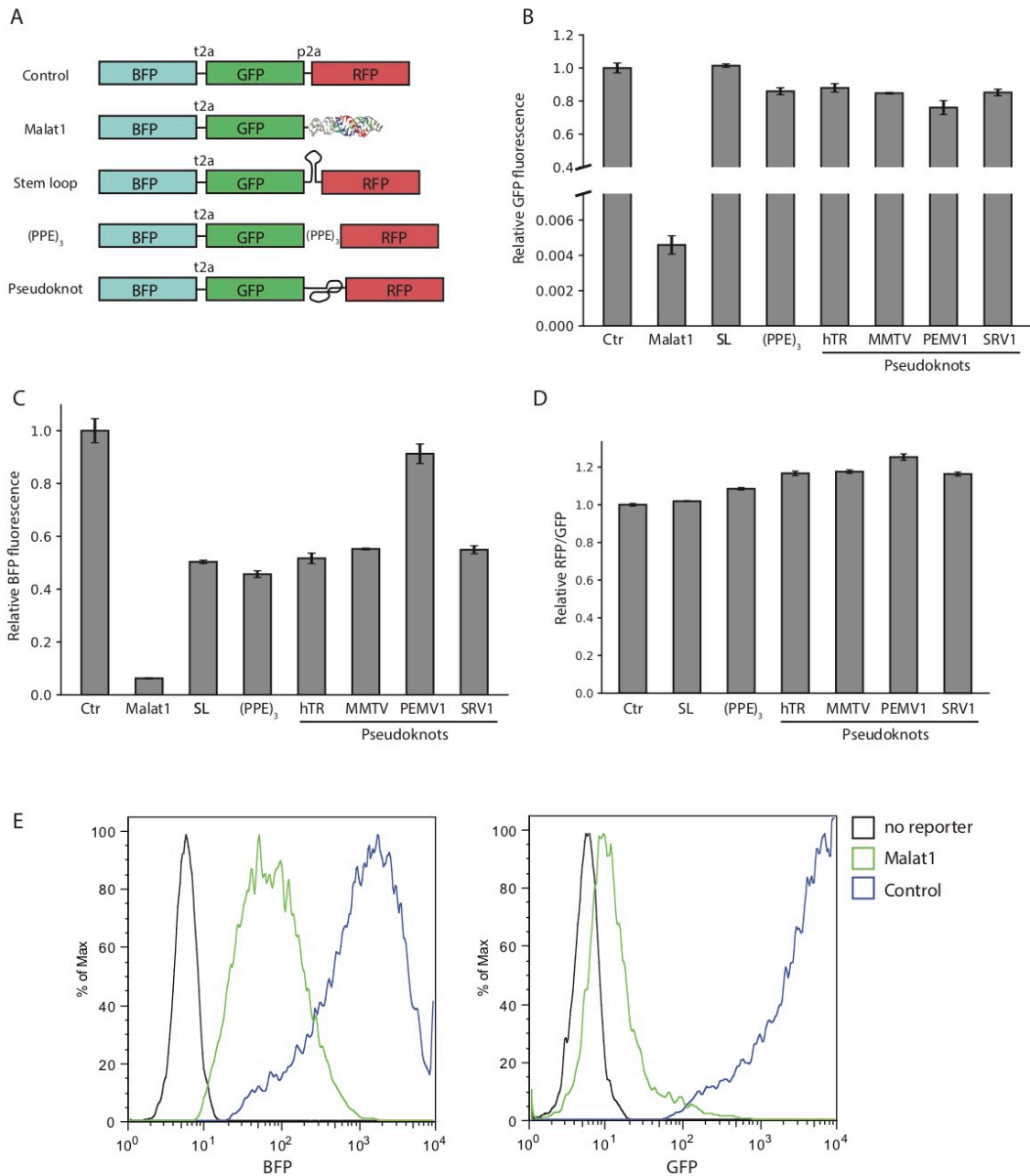
**Fig. 5-1. Mammalian stalling constructs. (A)** Diagrams of mammalian stalling reporters. **(B)** Relative GFP fluorescence of HEK293T cells transiently transfected with the reporter construct containing the indicated test sequences. **(C)** Relative BFP fluorescence of HEK293T cells transiently transfected with the reporter construct containing the indicated test sequences. **(D)** Median RFP:GFP ratio of HEK293T cells transiently transfected with a subset of the reporters. **(E)** BFP and GFP fluorescence of cells expressing no reporter, a control non-stalling reporter, or the MALAT1 stalling reporter.

Materials and Methods

**Constructs**

Reporter constructs for transient expression were generated starting with the pcDNA3 vector. All used plasmids are listed in Table 5-1.

**Cell Culture**

HEK293T cells were cultured in DMEM medium supplemented with 10% fetal bovine serum (FBS), 100 U/ml penicillin, 100 mg/ml streptomycin and 0.292 mg/ml L-glutamine (all from Invitrogen, Carlsbad, California) at 378C in 5% $CO_2$ Transient transfections were performed using TransIt-293 (Mirus) according to manufacturer's protocol. The NEMF knock our cell line was generated via CRISPR-Cas9 genome editing as previously described (Lin et al., 2014), and confirmed via Sanger Sequencing.

**Flow Cytometry Analysis**

Trypsinized cells were sedimented (1000 rpm for 5 min at room temperature), resuspended in PBS, and analyzed using the Becton Dickinson LSR II and FlowJo software.

Tables

**Table 5-1. Mammalain constructs**

| name | construct |
| --- | --- |
| pKK89 | pCMV-BFP-T2A-GFP-P2A-mCherry |
| pKK100 | pCMV-BFP-t2a-GFP-MALAT1 |
| pKK268 | pCMV-BFP-T2A-GFP-(PPE)3-RFP-STOP_polyA |
| pKK269 | pCMV-BFP-T2A-GFP-hTR-RFP-STOP_polyA |
| pKK270 | pCMV-BFP-T2A-GFP-MMTV-RFP-STOP_polyA |
| pKK271 | pCMV-BFP-T2A-GFP-PEMV1-RFP-STOP_polyA |
| pKK272 | pCMV-BFP-T2A-GFP-SRV1-RFP-STOP_polyA |
| pKK261 | pCMV_BFP_t2a_GFP_SL_RFP |

**CHAPTER SIX**

High throughput CRISPRi screen for mammalian specific RQC components

All known components of the yeast RQC pathway are conserved up to humans. Some early studies of the mammalian RQC have shown that the complex is intact and functional in higher eukaryotes (Juszkiewicz and Hegde, 2017b; Sundaramoorthy et al., 2017). However, it is not clear how the composition of the complex has evolved. As many multi component complexes, such as the ribosome, it is likely that the RQC has also acquired extra components and/or functionalities. The study presented here is the first attempt to obtain a comprehensive view of the mammalian RQC pathway.

Results

We used a PiggyBac Transposon method (Zhao et al., 2016) to stably integrate the MALAT1 stalling reporter in K562 cells harboring dCas9-KRAB fusion protein (Gilbert et al., 2014). We used Fluorescence Activated Cell Sorting (FACS) to isolate a population of cells that expressed high BFP and low GFP. The BFP marker was used as a proxy for the integration and expression of the reporter. The low GFP:BFP ratio is indicative of ribosome stalling and degradation of the nascent polypeptide. To confirm that the GFP was degraded in RQC-dependent manner, we used CRISPRi to knock down core RQC components, such as LTN1 and NEMF (Fig. 6-1). Downregulation of either RQC component led to stabilization of the stalling reporter, showing that a functional RQC was necessary for the degradation of the stalled protein. Surprisingly, knocking down the two core RQC components has differential effect on the stabilization of the reporter, i.e. RQC2 knock down led to a four-fold stabilization, whereas LTN1 knock down led to two fold increase in GFP fluorescence. This observation could stem from differences in the knock down levels of the two proteins. Indeed, the sgRNAs used against RQC2 led to lower mRNA levels compared to LTN1. In addition, it is possible that mammalian

cells have alternative and/or complementary E3 ubiquitin ligases that can partially rescue degradation if LTN1 is absent. Finally, based on work in yeast that argues that stalling substrates often aggregate when CAT-tails are added and Ltn1p is not functioning (Choe et al., 2016a; Yonashiro et al., 2016), it is possible that the lack of LTN1 leads to aggregates in mammalian cells that are degraded via an RQC-independent pathway. Many of these possibilities are yet to be explored.

Next, we performed a FACS-based whole genome CRISPRi screen (Fig. 6-2). Using our reporter cell line, we screened a genome-scale CRISPRi library (CRISPRi-v2) that targets 18,905 genes (20,526 TSSs) with 5 sgRNAs per TSS (Horlbeck et al., 2016). Briefly, reporter cells transduced with the library were grown for 8 days and then separated into bins according to their GFP signal by FACS. Cells in the top and bottom thirds of this reporter distribution were collected and processed to measure the frequencies of sgRNAs contained within each, from which we calculated sgRNA and gene-level reporter signal phenotypes. Our CRISPRi-v2 screen identified more that 200 hit genes that stabilize GFP upon knock down.

Based on the screen hits, we were able to build a comprehensive model of the mammalian RQC pathway. For example, the mammalian homologs of all three core components of the yeast RQT pathway scored as hits in the screen (Fig. 6-3A). ASCC3 (yeast Slh1p), ZNF598 (yeast Hel2p), and GNB2L1 (yeast Rack1p) are believe to detect stalled ribosomes and trigger the downstream pathway (Juszkiewicz and Hegde, 2017b; Matsuo et al., 2017; Sundaramoorthy et al., 2017). In addition, we identified PELO and HBS1L, the two factors known to play a role in splitting stalled ribosomes (Fig. 6-3B) (Tsuboi et al., 2012). The strongest hit from the screen was NEMF (yeast Rqc2p) (Fig. 6-3C). The remaining two core components of the complex, LTN1and TCF25 (yeast Rqc1p) also scored in the top 10 hits. We

also identified components of the mammalian exosome, which degraded mRNA, and the proteasome, which ultimately degrades stalled nascent polypeptides. Consistent with results from yeast studies, we were able to confirm the role of the RQC pathway in coping with stalled ribosomes in higher eukaryotes.

In addition to the identified conserved RQC components, among the hits from the screen were various proteins with either no yeast homologs or homologs that have not been previously implicated in the RQC. One of these hits was GIGYF2 (Fig. 6-3D). Previous studies have suggested that GIGYF2, together with 4EHP and ZNF598, inhibits translation initiation (Morita et al., 2012). Therefore, it is tempting to speculate that GIGYF2 blocks initiation of more ribosome on a defective message once stalling is detected, therefore decreasing the number of stalled ribosomes.

Conclusions

Although the RQC has been well characterized in yeast, the composition and function of the mammalian complex remains poorly understood. Our most recent work has focused on elucidating the role of the mammalian RQC pathway. We utilized the RQC reporter described in chapter five to performed a genome-wide CRISPRi screen that identified a number of mammalian-specific RQC components. We were able to reconstruct the mammalian RQC pathway, which has conserved components from yeast. In addition, we were able to identify more than 200 additional components. We are currently in the process of dissecting the role of these novel components in mammalian translation quality control. We have designed a series of follow-up experiment that would allow us to place these proteins in different stages of the pathway, such as stalled ribosome detection, mRNA degradation, initiation inhibition, nascent polypeptide degradation, etc. The outlined here screen allowed us to gain a comprehensive view

of the mammalian RQC pathway and is the first step towards unraveling how the pathway has

evolved in higher eukaryotes.
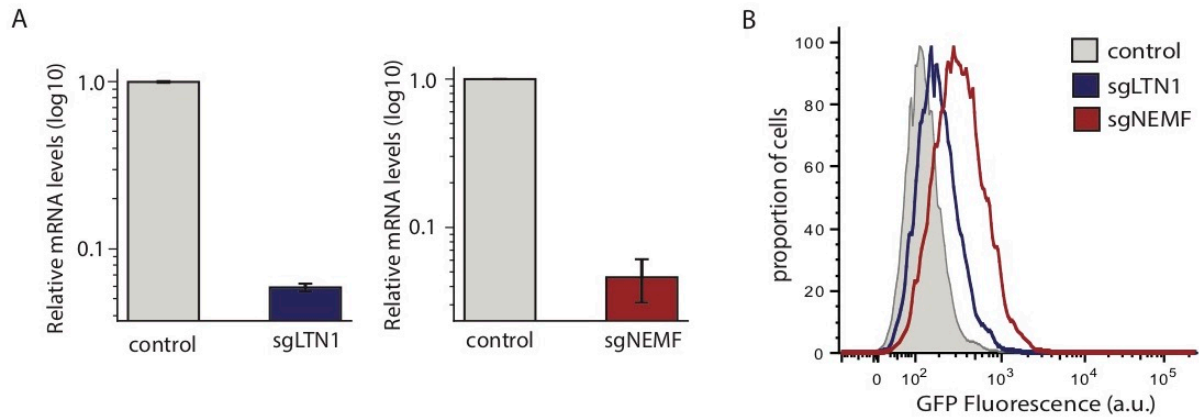
**Fig. 6-1. Characterization of the screening cell line. A.** Relative mRNA levels of RQC components in cells expressing sgRNAs against *NEMF* or *LTN1*. Data represent means relative to *ACTB* mRNA and normalized to cells expressing a non-targeting sgRNA (*GAL4*) sgRNA ± standard error of biological replicates (n = 3). **B.** A histogram of GFP fluorescence for the stalling reporter cell line expressing a control sgRNA or sgRNAs targeting core RQC components.
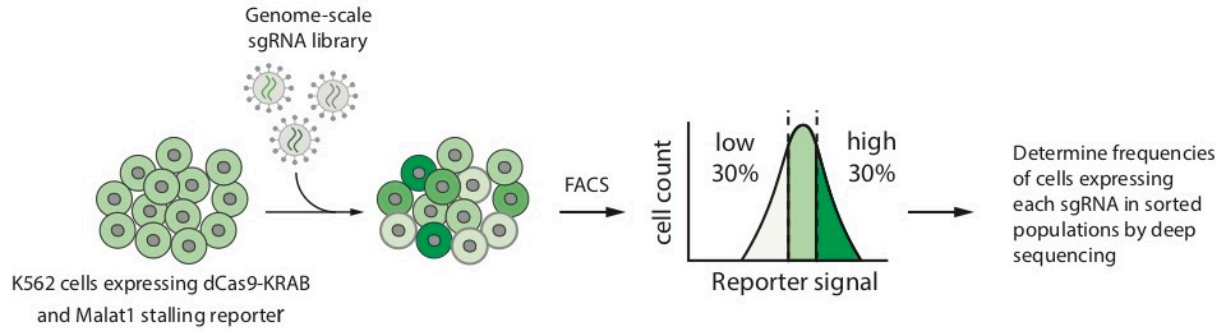
**Fig. 6-2 Schematic of FACS-based CRISPRi screen.** K562 expressing dCas9-KRAB and the MALAT1 stalling reporter are infected with a whole genome CRISPRi library. Fluorescent Activated Cell Sorting (FACS) is used to purify the population of cells with high and low GFP fluorescence. Deep sequencing is used to identify the sgRNAs expressed in the purified cell populations.
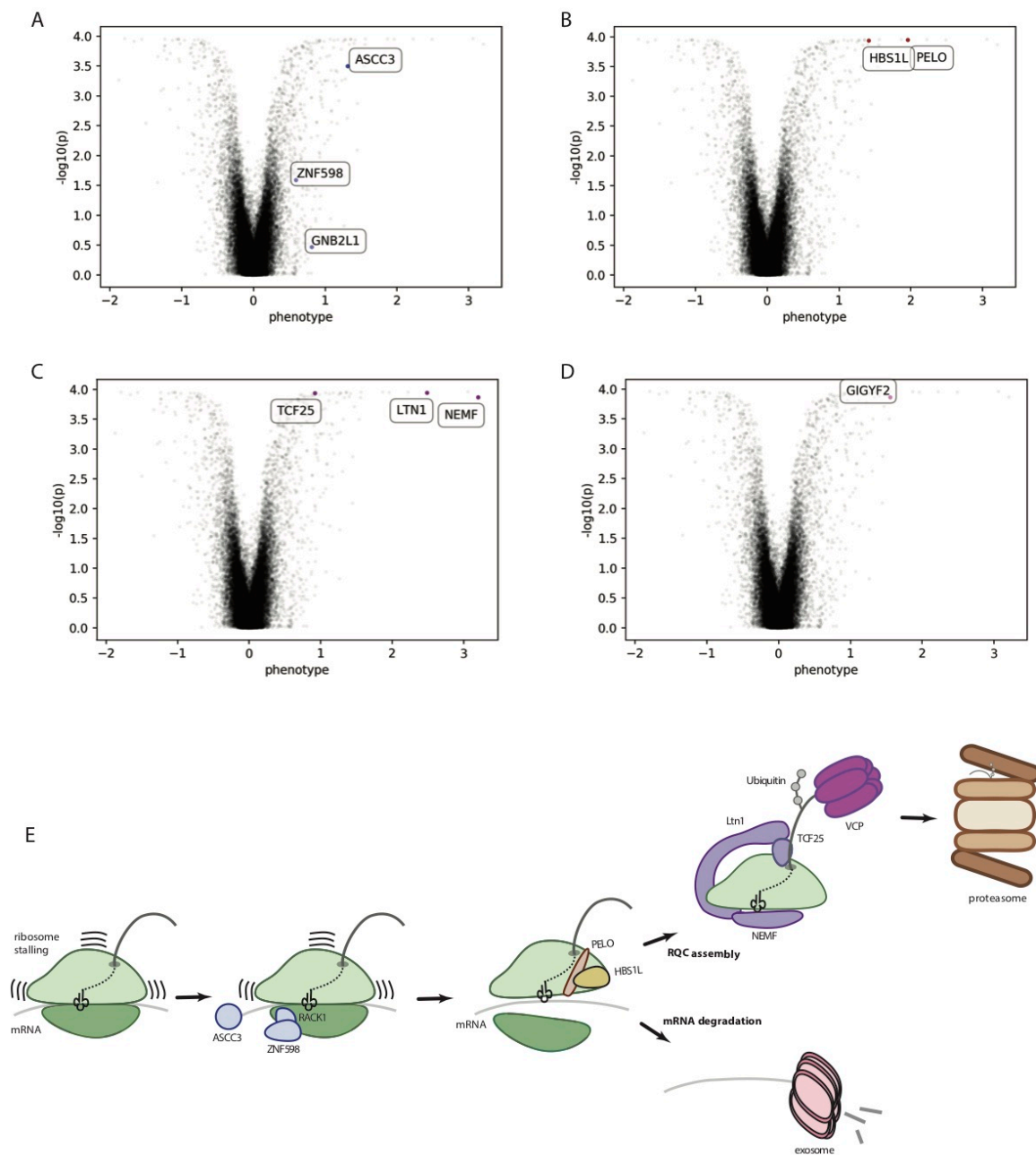
**Fig. 6-3. Genome-scale CRISPRi screening to identify gene-depletion events that stabilize the stalling reporter. (A-D)** Volcano plot of reporter phenotypes and p values from CRISPRi-v2 screen. The phenotype is defined as the log2 enrichment of sgRNA sequences identified within the high GFP cells over the low GFP cells. **E.** Model of the mammalian RQC pathway.

Materials and Methods

**Cell Culture**

K562 cells were grown in RPMI-1640 with 25mM HEPES, 2.0 g/L NaHCO3, 0.3 g/L L-Glutamine supplemented with 10% FBS, 2 mM glutamine, 100 units/mL penicillin and 100 μg/mL streptomycin. HEK293T cells were grown in Dulbecco's modified eagle medium (DMEM) in 10% FBS, 100 units/mL penicillin and 100 μg/mL streptomycin. Lentivirus was produced by transfecting HEK293T with standard packaging vectors using TransIT-LTI Transfection Reagent (Mirus, MIR 2306). Viral supernatant was harvested at least ~2 days after transfection and filtered through a PVDF syringe filter and/or frozen prior to infection.

To construct the stalling reporter cell line, K562 cells stably expressing dCas9-KRAB (Gilbert et al., 2014) were electroporated with the MALAT1 stalling reporter and a PiggyBac transposase (System Biosciences, PB210PA-1) according to the manufacturer's instructions. This polyclonal line was used for the CRISPRi screen.

**Whole-genome CRISPRi Screen**

Reporter screens were conducted using protocols similar to those previously described (Gilbert et al., 2014; Horlbeck et al., 2016). The CRISPRi-v2 (Addgene, Cat#83969) sgRNA libraries were transduced into the reporter cell line at an MOI < 1. Cells were grown in spinner flasks for 2 days without selection, followed by 2 days of selection with 1 μg/mL puromycin. Screen replicates were split prior infection and carried separately throughout the remainder of the experiment. Cells were separated into those with the highest (~28%–33%) and lowest (~30%–35%) GFP signal 8 days post transduction by fluorescence-activated cell sorting (FACS). Cell pellets were frozen after collection. Approximately 23-30 million cells were collected per bin. Genomic DNA

was isolated from frozen cells and the sgRNA-encoded regions were enriched, amplified, and prepared for sequencing. CRISPRi-v2 samples were sequenced with greater coverage. Sequenced protospacer sequences were aligned and data were processed as described (Horlbeck et al., 2016) with custom Python scripts (ScreenProcessing, available at https://github.com/mhorlbeck/ScreenProcessing). Reporter phenotypes for library sgRNAs were calculated as the log2 enrichment of sgRNA sequences identified within the high GFP cells over the low GFP cells. Phenotypes for each transcription start site were then calculated as the average reporter phenotype of the 3 sgRNAs with the strongest phenotype by absolute value (most active sgRNAs). Mann-Whitney test p values were calculated by comparing all sgRNAs targeting a given TSS to the full set of negative control sgRNAs.

**CHAPTER SEVEN**

Discussion and future perspectives

Protein synthesis is fundamental cellular process that is regulated at many levels and subject to stringent quality control. Although that pathways that monitor cellular proteins have been extensively studied, quality control of nascent polypeptides during translation has only recently emerged (Brandman and Hegde, 2016b). Indeed, protein synthesis is an error prone process and various factors can induce translation failure in the form of stalled ribosome. Such stalled ribosomes trigger various quality control pathways that ultimately lead to the degradation of the associated mRNA (Doma and Parker, 2006), recycling of the ribosome (Pisareva et al., 2011) and degradation of the partially synthesized nascent polypeptide (Ito-Harashima et al., 2007). At the heart of the last step is the Ribosome Quality control Complex (RQC).

Yeast RQC

1. Rqc1p

Over the past years we have gained detailed knowledge of the structure and function of the RQC complex in yeast. However, there are still components of the complex that remain poorly characterized, such as Rqc1p. Although this protein is a core component of the complex, is function us unknown. Some studies suggest that Rqc1p prevents nascent polypeptide aggregation (Defenouillère et al., 2016). However, it is still unclear whether and how this protein can chaperone the stalled nascent polypeptide and facilitate its targeting to the proteasome.

2. CAT tails

The discovery of the CAT tails is likely the most surprising result that came from studying the RQC. Indeed, Rqc2p's ability to recruit charged tRNAs to the RQC-bound 60S ribosomal subunit and facilitate the addition of alanine and threonine amino acids to the C-terminus of the stalled nascent polypeptide with no RNA template is simply fascinating. Our research clearly shows that the addition of CAT-tails can facilitate the degradation of nascent

polypeptides lacking Lys residues in the Ltn1p's accessible region. Therefore, the CAT tails are not a degron themselves, but serve as a fail-safe mechanism that vastly expands the range of RQC degradable substrates (Kostova et al., 2017). In addition, when the complex is not functional, CAT-tailing can induce aggregation, sequester away chaperones, and cause proteotoxic stress (Choe et al., 2016a; Defenouillère et al., 2016; Yonashiro et al., 2016).

Although we have a better understanding of the function of the CAT tails, their structure remains a mystery. The CAT tails are stretches of random alanine and threonine residues. It is not known why the system evolved to utilize precisely these two amino acids. Our studies in nascent polypeptide degradation suggest that a more advantageous amino acid to add to a stalled nascent polypeptide is Lysine, since the addition of that amino acid ensures the ability of the RQC to degrade the failed polypeptide. However, stretches of lysine residue could potentially interact with the ribosome exit tunnel and hinder the release of the stalled nascent polypeptide. In this sense, maybe the mixture of alanine and threonine serves as "nature's Teflon", allowing smooth extraction from the ribosome exit tunnel. Additionally, it is possible that these two amino acids form weak bonds with the tRNAs that deliver them to the ribosome (Peacock et al., 2014), which can also be advantageous for release of the nascent polypeptide. Finally, the ability of the CAT tails to aggregate might be advantageous, especially if nascent polypeptide cannot be properly degraded either due to the lack of Lys residues or defects in the RQC itself.

3. Endogenous substrates

The RQC field was built on artificial reporters. None of the used constructs and the stalling sites they utilized are found in nature. Although these reporters have been tremendously useful for dissecting the pathway, the spectrum of endogenous substrates that the RQC serves have not been explored. Deletions of RQC components in yeast have no effect on the growth of

the yeast, suggesting that ribosome stalling is rare in rapidly dividing yeast under standard growth conditions. However, yeast strains occurring in nature live under various growth conditions that might not be as favorable. Various studies have shown that environmental factors such as UV radiation and oxidative stress can damage mRNA and components of the ribosome, which can result in translation failure. In addition, naturally occurring translation inhibitors, such as emetine and cyclohieximide, can also induce ribosome stalling. Therefore, it is important to explore the role of the RQC in the fitness of cells under various growth conditions.

<u>Mammalian</u>

1. Mammalian Specific components

All known components of the RQC are conserved in higher eukaryotes. However, our knowledge of the structure and function of the mammalian complex is lagging behind. Initially, studying the mammalian complex has been hindered by the lack stalling substrates. However, the designed by us MALAT1-based reporter, as well as the reporters engineered in the Hegde Lab have primed the field for mechanistic studies. Indeed, the described above CRISPRi screen is the first comprehensive approach to understand the complexity of the mammalian pathway. In addition to factors that have been characterized in yeast, we discovered a number of mammalian-specific components that could play a role in coping with stalled ribosomes. However, more detailed mechanistic studies are yet to characterize these putative RQC components.

2. Endogenous substrates

In contrast to yeast, human cells have a growth defects when components of the RQC are deleted, mutated, or downregulated. This observation has two potential explanations. First, ribosome stalling is more frequent in higher eukaryotes. It is likely that transcription and splicing are more error-prone in human cell lines, which would lead to more faulty mRNA that would

result in ribosome stalling. Indeed, most genes in *S. cerevisiae* do not contain introns, so mistakes during splicing could be rare. Alternatively, it is possible that there are proteins being degraded via the RQC in a regulatory manner in mammalian cells. Therefore, inactivation of the RQC could lead to accumulation of these proteins and potential perturbation of the cellular homeostasis.

3. CAT tails

As mentioned above, one of the most fascinating aspects of the RQC is the ability of the complex to tag stalled nascent polypeptides with non-templated amino acids, i.e. CAT-tailing. However, CAT tails have only been observed in yeast and to date there is no report of mammalian CAT tails. The mammalian homolog of Rqc2p, NEMF, has similar structure and also binds the bottom 60S subunit close to the anticodons of the A- and P-site tRNAs. Based on the conservation of the complex from yeast to mammalian cells, it is surprising that mammalian CAT tails have not been observed. It is possible that the CAT-tails are a yeast-specific phenomenon. Our studies suggest that the CTA tails are a fail-safe mechanism that compensates for the rigidity of the Ltn1p ubiquitin ligase. It is possible that higher eukaryotes no longer need this compensatory mechanism, for example because the mammalian LTN1 is more flexible and can reach and uiquitinate a large fraction of the stalled nascent peptide. Another possible explanation is that mammalian CAT tails exist, but they are difficult to observe. For example, the mammalian CAT tails might serve as a degron and any protein species that acquires a CAT tail is rapidly degraded by another pathway. Similar to CAT tails in yeast, the mammalian ones might be prone to aggregation and big aggregates are difficult to detect via Western blot. The aggregation itself could also induce the degradation of the CAT-tailed species. We are yet to discover which of these hypotheses is true.

4. Human disease

Finally, mutations in RQC components have been associated with neurodegenerative disorders (Chu et al., 2009). However, how the inability to degrade stalled nascent polypetides translates into neuronal death has not been explored. The core RQC ubiquitin ligase, LTN1, was initially identified in a mouse forward genetic screen for neurological phenotypes. Although LTN1 is widely expressed in all tissues, motor and sensory neurons and neuronal processes in the brainstem and spinal cord are primarily affected in the mutant mouse, which suggests that the RQC might have differential functions in various tissues and organs or during the lifetime of the organism. Indeed, many neurodegenerative disordered are age related. It is believed that as neurons age their ability to maintain protein homeostasis diminished. Therefore, it is possible that mutations in the RQC could add an insult to an already strained system and push it towards disease. Since the RQC was only recently discovered, we are yet to explore its role in human diseases.

References:

Alamgir, M., Erukova, V., Jessulat, M., Azizi, A., and Golshani, A. (2010). Chemical-genetic profile analysis of five inhibitory compounds in yeast. BMC Chem. Biol. *10:6*.

Ares, M. (2012). Isolation of total RNA from yeast cell cultures. Cold Spring Harb. Protoc. *2012*, 1082–1086.

Ast, T., Cohen, G., and Schuldiner, M. (2013). A Network of Cytosolic Factors Targets SRP-Independent Proteins to the Endoplasmic Reticulum. Cell *152*, 1134–1145.

Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. (2000). The Complete Atomic Structure of the Large Ribosomal Subunit at 2.4 Å Resolution. Science (80-. ). *289*.

Bengtson, M.H., and Joazeiro, C.A.P. (2010). Role of a ribosome-associated E3 ubiquitin ligase in protein quality control. Nature *467*, 470–473.

Brandman, O., and Hegde, R.S. (2016a). Ribosome-associated protein quality control. Nat. Struct. Mol. Biol. *23*, 7–15.

Brandman, O., and Hegde, R.S. (2016b). Ribosome-associated protein quality control. Nat. Struct. Mol. Biol. *23*, 7–15.

Brandman, O., Stewart-Ornstein, J., Wong, D., Larson, A., Williams, C.C., Li, G.-W., Zhou, S., King, D., Shen, P.S., Weibezahn, J., et al. (2012). A ribosome-bound quality control complex triggers degradation of nascent peptides and signals translation stress. Cell *151*, 1042–1054.

Choe, Y.-J., Park, S.-H., Hassemer, T., Körner, R., Vincenz-Donnelly, L., Hayer-Hartl, M., and Hartl, F.U. (2016a). Failure of RQC machinery causes protein aggregation and proteotoxic stress. Nature *531*, 191–195.

Choe, Y.-J., Park, S.-H., Hassemer, T., Körner, R., Vincenz-Donnelly, L., Hayer-Hartl, M., and Hartl, F.U. (2016b). Failure of RQC machinery causes protein aggregation and proteotoxic

stress. Nature *531*, 191–195.

Chu, J., Hong, N.A., Masuda, C.A., Jenkins, B. V, Nelms, K.A., Goodnow, C.C., Glynne, R.J., Wu, H., Masliah, E., Joazeiro, C.A.P., et al. (2009). A mouse forward genetics screen identifies LISTERIN as an E3 ubiquitin ligase involved in neurodegeneration. Proc. Natl. Acad. Sci. U. S. A. *106*, 2097–2103.

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.

Defenouillère, Q., Yao, Y., Mouaikel, J., Namane, A., Galopier, A., Decourty, L., Doyen, A., Malabat, C., Saveanu, C., Jacquier, A., et al. (2013). Cdc48-associated complex bound to 60S particles is required for the clearance of aberrant translation products. Proc. Natl. Acad. Sci. U. S. A. *110*, 5046–5051.

Defenouillère, Q., Zhang, E., Namane, A., Mouaikel, J., Jacquier, A., and Fromont-Racine, M. (2016). Rqc1 and Ltn1 Prevent C-terminal Alanine-Threonine Tail (CAT-tail)-induced Protein Aggregation by Efficient Recruitment of Cdc48 on Stalled 60S Subunits. J. Biol. Chem. *291*, 12245–12253.

Dimitrova, L.N., Kuroha, K., Tatematsu, T., and Inada, T. (2009). Nascent peptide-dependent translation arrest leads to Not4p-mediated protein degradation by the proteasome. J. Biol. Chem. *284*, 10343–10352.

Doma, M., and Parker, R. (2006). Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. Nature.

Engel, S.R., Dietrich, F.S., Fisk, D.G., Binkley, G., Balakrishnan, R., Costanzo, M.C., Dwight, S.S., Hitz, B.C., Karra, K., Nash, R.S., et al. (2014). The reference genome sequence of Saccharomyces cerevisiae: then and now. G3 *4*, 389–398.

Gamble, C.E., Brule, C.E., Dean, K.M., Fields, S., and Grayhack, E.J. (2016). Adjacent Codons Act in Concert to Modulate Translation Efficiency in Yeast. Cell *166*, 679–690.

Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C., et al. (2014). Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. Cell *159*, 647–661.

Harms, J., Schluenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I., Bartels, H., Franceschi, F., and Yonath, A. (2001). High Resolution Structure of the Large Ribosomal Subunit from a Mesophilic Eubacterium. Cell *107*, 679–688.

Higgins, R., Gendron, J.M., Rising, L., Mak, R., Webb, K., Kaiser, S.E., Zuzow, N., Riviere, P., Yang, B., Fenech, E., et al. (2015). The Unfolded Protein Response Triggers Site-Specific Regulatory Ubiquitylation of 40S Ribosomal Proteins. Mol. Cell *59*, 35–49.

Horlbeck, M.A., Gilbert, L.A., Villalta, J.E., Adamson, B., Pak, R.A., Chen, Y., Fields, A.P., Park, C.Y., Corn, J.E., Kampmann, M., et al. (2016). Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. Elife *5*, e19760.

Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell *147*, 789–802.

Ito-Harashima, S., Kuroha, K., Tatematsu, T., and Inada, T. (2007). Translation of the poly(A) tail plays crucial roles in nonstop mRNA surveillance via translation repression and protein destabilization by proteasome in yeast. Genes Dev. *21*, 519–524.

Jan, C.H., Williams, C.C., and Weissman, J.S. (2014). Principles of ER cotranslational translocation revealed by proximity-specific ribosome profiling. Science (80-. ). *346*.

Juszkiewicz, S., and Hegde, R.S. (2017a). Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination. Mol. Cell 743–750.

Juszkiewicz, S., and Hegde, R.S. (2017b). Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination. Mol. Cell *65*, 743–750.e4.

Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature *533*, 420–424.

Kostova, K.K., Hickey, K.L., Osuna, B.A., Hussmann, J.A., Frost, A., Weinberg, D.E., and Weissman, J.S. (2017). CAT-tailing as a fail-safe mechanism for efficient degradation of stalled nascent polypeptides. Science *357*, 414–417.

Koutmou, K.S., Schuller, A.P., Brunelle, J.L., Radhakrishnan, A., Djuranovic, S., Green, R., Arenz, S., Meydan, S., Starosta, A., Berninghausen, O., et al. (2015). Ribosomes slide on lysine-encoding homopolymeric A stretches. Elife *4*, 446–452.

Kramer, G., Boehringer, D., Ban, N., and Bukau, B. (2009). The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. Nat. Struct. Mol. Biol. *16*, 589–597.

Letzring, D.P., Dean, K.M., and Grayhack, E.J. (2010). Control of translation efficiency in yeast by codon-anticodon interactions. RNA *16*, 2516–2528.

Lin, S., Staahl, B.T., Alla, R.K., and Doudna, J.A. (2014). Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. Elife *3*, e04766.

Longtine, M.S., Mckenzie III, A., Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P., and Pringle, J.R. (1998). Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. Yeast *14*, 953–961.

Lyumkis, D., Oliveira dos Passos, D., Tahara, E.B., Webb, K., Bennett, E.J., Vinterbo, S., Potter, C.S., Carragher, B., and Joazeiro, C.A.P. (2014). Structural basis for translational surveillance by the large ribosomal subunit-associated protein quality control complex. Proc. Natl. Acad. Sci. U.

S. A. *111*, 15981–15986.

Malsburg, K. v. d., Shao, S., and Hegde, R.S. (2015). The ribosome quality control pathway can access nascent polypeptides stalled at the Sec61 translocon. Mol. Biol. Cell *26*, 2168–2180.

Matsuo, Y., Ikeuchi, K., Saeki, Y., Iwasaki, S., Schmidt, C., Udagawa, T., Sato, F., Tsuchiya, H., Becker, T., Tanaka, K., et al. (2017). Ubiquitination of stalled ribosome triggers ribosome-associated quality control. Nat. Commun. *8*, 159.

Mefferd, A.L., Kornepati, A.V.R., Bogerd, H.P., Kennedy, E.M., and Cullen, B.R. (2015). Expression of CRISPR/Cas single guide RNAs using small tRNA promoters. RNA *21*, 1683–1689.

Morita, M., Ler, L.W., Fabian, M.R., Siddiqui, N., Mullin, M., Henderson, V.C., Alain, T., Fonseca, B.D., Karashchuk, G., Bennett, C.F., et al. (2012). A novel 4EHP-GIGYF2 translational repressor complex is essential for mammalian development. Mol. Cell. Biol. *32*, 3585–3593.

Nissen, P., Hansen, J., Ban, N., Moore, P.B., and Steitz, T.A. (2000). The Structural Basis of Ribosome Activity in Peptide Bond Synthesis. Science (80-. ). *289*.

Park, E., and Rapoport, T.A. (2012). Mechanisms of Sec61/SecY-Mediated Protein Translocation Across Membranes. Annu. Rev. Biophys. *41*, 21–40.

Peacock, J.R., Walvoord, R.R., Chang, A.Y., Kozlowski, M.C., Gamper, H., and Hou, Y.-M. (2014). Amino acid-dependent stability of the acyl linkage in aminoacyl-tRNA. RNA *20*, 758–764.

Pfeffer, S., Dudek, J., Gogala, M., Schorr, S., Linxweiler, J., Lang, S., Becker, T., Beckmann, R., Zimmermann, R., and Förster, F. (2014). Structure of the mammalian oligosaccharyl-transferase complex in the native ER protein translocon. Nat. Commun. *5*, 3072.

Pisareva, V.P., Skabkin, M.A., Hellen, C.U.T., Pestova, T. V, and Pisarev, A. V (2011). Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. EMBO J. *30*, 1804–1817.

Rothstein, R. (1991). Targeting, disruption, replacement, and allele rescue: integrative DNA transformation in yeast. Methods Enzymol. *194*, 281–301.

Schellenberger, V., Wang, C., Geething, N.C., Spink, B.J., Campbell, A., To, W., Scholle, M.D., Yin, Y., Yao, Y., Bogin, O., et al. (2009). A recombinant polypeptide extends the in vivo half-life of peptides and proteins in a tunable manner. Nat. Biotechnol. *27*.

Schmeing, T.M., and Ramakrishnan, V. (2009). What recent ribosome structures have revealed about the mechanism of translation. Nature *461*, 1234–1242.

Shao, S., and Hegde, R.S. (2014). Reconstitution of a Minimal Ribosome-Associated Ubiquitination Pathway with Purified Factors. Mol. Cell *55*, 880–890.

Shao, S., von der Malsburg, K., and Hegde, R.S. (2013). Listerin-Dependent Nascent Protein Ubiquitination Relies on Ribosome Subunit Dissociation. Mol. Cell *50*, 637–648.

Shen, P.S., Park, J., Qin, Y., Li, X., Parsawar, K., Larson, M.H., Cox, J., Cheng, Y., Lambowitz, A.M., Weissman, J.S., et al. (2015). Rqc2p and 60S ribosomal subunits mediate mRNA-independent elongation of nascent chains. Science (80-. ). *347*.

Simms, C.L., Hudson, B.H., Mosior, J.W., Rangwala, A.S., and Zaher, H.S. (2014). An Active Role for the Ribosome in Determining the Fate of Oxidized mRNA. Cell Rep. *9*, 1256–1264.

Simms, C.L., Thomas, E.N., and Zaher, H.S. (2017). Ribosome-based quality control of mRNA and nascent peptides. Wiley Interdiscip. Rev. RNA *8*.

Staple, D.W., Butcher, S.E., Jacobs, J., Mukhopadhyay, B., and Hennig, M. (2005). Pseudoknots: RNA Structures with Diverse Functions. PLoS Biol. *3*, e213.

Subtelny, A.O., Eichhorn, S.W., Chen, G.R., Sive, H., and Bartel, D.P. (2014). Poly(A)-tail profiling reveals an embryonic switch in translational control. Nature *508*, 66–71.

Sundaramoorthy, E., Leonard, M., Mak, R., Liao, J., Fulzele, A., and Bennett, E.J. (2017). ZNF598 and RACK1 Regulate Mammalian Ribosome-Associated Quality Control Function by Mediating Regulatory 40S Ribosomal Ubiquitylation. Mol. Cell *65*, 751–760.e4.

Tsuboi, T., Kuroha, K., Kudo, K., Makino, S., Inoue, E., Kashima, I., and Inada, T. (2012). Dom34:Hbs1 Plays a General Role in Quality-Control Systems by Dissociation of a Stalled Ribosome at the 3′ End of Aberrant mRNA. Mol. Cell *46*, 518–529.

Verma, R., Oania, R.S., Kolawa, N.J., and Deshaies, R.J. (2013). Cdc48/p97 promotes degradation of aberrant nascent polypeptides bound to the ribosome. Elife *2*, e00308.

Wiertz, E., Tortorella, D., Bogyo, M., Yu, J., and Mothes, W. (1996). Sec61-mediated transfer of a membrane protein from the endoplasmic reticulum to the proteasome for destruction. Nature.

Wilusz, J.E., JnBaptiste, C.K., Lu, L.Y., Kuhn, C.-D., Joshua-Tor, L., and Sharp, P.A. (2012). A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. Genes Dev. *26*, 2392–2407.

Wolf, D.H., and Stolz, A. (2012). The Cdc48 machine in endoplasmic reticulum associated protein degradation. Biochim. Biophys. Acta - Mol. Cell Res. *1823*, 117–124.

Yonashiro, R., Tahara, E.B., Bengtson, M.H., Khokhrina, M., Lorenz, H., Chen, K.-C., Kigoshi-Tansho, Y., Savas, J.N., Yates, J.R., Kay, S.A., et al. (2016). The Rqc2/Tae2 subunit of the ribosome-associated quality control (RQC) complex marks ribosome-stalled nascent polypeptide chains for aggregation. Elife *5*, e11794.

Zhao, S., Jiang, E., Chen, S., Gu, Y., Shangguan, A.J., Lv, T., Luo, L., and Yu, Z. (2016). PiggyBac transposon vectors: the tools of the human gene encoding. Transl. Lung Cancer Res. *5*,

120–125.

## Publishing Agreement

*It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.*

### Please sign the following statement:

*I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.*

_____

Author Signature

10/23/18

Date