Rqc1 and other yeast proteins containing highly positively charged domains are not targets of the RQC complex

Géssica C. Barros[1#], Rodrigo D. Requião[1#], Rodolfo L. Carneiro[1#], Claudio A. Masuda[2], Mariana H. Moreira[1], Silvana Rossetto[3], Tatiana Domitrovic[4*] and Fernando L. Palhano[1*]

[1] Programa de Biologia Estrutural, Instituto de Bioquímica Médica Leopoldo de Meis, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, 21941-902, Brazil.
[2] Programa de Biologia Molecular e Biotecnologia, Instituto de Bioquímica Médica Leopoldo de Meis, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, 21941-902, Brazil.
[3] Departamento de Ciência da Computação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, 21941-902, Brazil.
[4] Departamento de Virologia, Instituto de Microbiologia Paulo de Góes, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 21941-902, Brazil.

* To whom correspondence should be addressed: Tel.: +55 21 3938-6761; Email: palhano@bioqmed.ufrj.br or domitrovic@micro.ufrj.br

# These authors contributed equally.

ABSTRACT

Highly positively charged protein segments are known to result in poor translation efficiency by the action of the ribosome quality control (RQC) complex. This effect may be explained by ribosome stalling caused by electrostatic interactions between the nascent peptide and the negatively charged ribosome exit tunnel, leading to translation termination followed by protein degradation mediated by the RQC complex. These poly-basic sequences are mainly studied with reporter systems, where artificial sequences are introduced into heterologous genes. The unique example of a yeast protein containing a polybasic domain that was described as a target of the RQC complex is Rqc1, a protein essential for RQC function. Rqc1 poly-basic sequences are thought to be necessary to activate the RQC, leading to protein down-regulation. We aimed to determine whether the RQC complex acts as a regulatory mechanism for other endogenous proteins containing similar poly-basic sequences. Here, we show by bioinformatics, ribosome profiling data, and western blot protein quantification that endogenous proteins containing poly-basic sequences similar to or even more positively charged than those of Rqc1 are not targeted by the RQC complex, suggesting that endogenous poly-basic sequences are not sufficient to induce this type of regulation. We further demonstrate that Rqc1, previously described as an RQC target, is not regulated by the RQC complex, but instead by Ltn1, the E3 component of the complex, alone. Our data suggest that Ltn1 is able to regulate the expression of Rqc1 post-translationally and independent of the RQC complex.

Keywords: ribosome quality control, ribosome profiling, poly (A) tracts, poly-basic domains, stalled polypeptides, protein quality control

INTRODUCTION

The first study to propose that the translation of positive amino acids could lead to translational stalling was published in 2008 by Lu and Deutsch [1]. It was already known by that time that the ribosomal exit tunnel, a narrow passage through which the nascent peptide passes as soon as it is synthesized, is negatively charged [2] and that the translation of the poly (A) tail from constructs lacking a stop codon, which leads to the incorporation of lysine residues (the codon AAA translates into lysine), causes translational repression [3-5]. This led to the hypothesis that a high concentration of positively charged residues inside the exit tunnel could lead to slower translation rates due to the presence of electrostatic interactions. Using a cell-free approach, Lu and Deutsch, 2008, demonstrated that the presence of lysine or arginine repeats is sufficient to decrease translation rates. They observed that the presence of as few as four nonconsecutive arginines in the construct was already enough to diminish the translational speed, while an equal concentration of negatively charged or uncharged residues did not affect the translation. This phenomenon was also observed with lysines and was independent of codon choice, suggesting that the retardation effect was explicitly caused by the charge itself, not the amino acid structure or tRNA availability [1].

These early experiments prompted the use of poly-basic sequences as stalling regions in many subsequent papers. Most of these publications focused on the characterization of the ribosome quality control (RQC) complex, a cellular pathway responsible for the co-translational degradation of nascent peptides [6, 7]. Poly-basic sequences were added to or removed from a reporter protein, and the amount of translated protein was detected by western blot [6, 8, 9]. Using this approach, several authors have confirmed that the presence of poly-arginine (R), poly-lysine (K), or mixtures of both amino acids can reduce protein production [10-12] and that this effect was linked to the action of the quality control machinery [6,13-16].

The RQC complex removes stalled nascent polypeptides from the ribosome and activates stress signals in *Saccharomyces cerevisiae* [6, 16]. A defective RQC complex causes proteotoxic stress in yeast cells [9] and neurodegeneration in mice [17]. The yeast RQC complex components are Rqc1, Rqc2, and Ltn1 proteins, all of which have a relatively low copy number when compared with the number of ribosomes in a cell, thus indicating that the system could easily reach saturation [6, 13, 18-20]. The RQC complex is modulated by Asc1 and Hel2. These proteins detect ribosome stalling and promote early termination, leading to ribosomal subunit dissociation and transcript degradation. [6]. After ribosome dissociation, the RQC complex can bind to the 60S subunit and degrade the nascent polypeptide through Ltn1 activity, an E3 ubiquitin ligase involved in co-translational degradation [6, 8]. Rqc1 and Ltn1 lead to the recruitment of Cdc48 and its co-factors, Npl4 and Ufd1, which extract the stalled ubiquitylated polypeptide, through ATP hydrolysis, leaving the 60S subunit free to engage in another round of translation [7]. Rqc2 is involved in stress signaling through the formation of the C-terminal alanine-threonine (CAT) tail. After attaching to the 60S subunit, Rqc2 can incorporate alanine and threonine residues to the C-terminal of the stalled peptide, creating the CAT tail [9, 21-24]. This incorporation is random and independent of mRNA, and it is important to induce stress and/or to expose lysine residues that may be trapped inside the ribosome exit tunnel to the

ubiquitin ligase activity of Ltn1 [23].

In addition to the action as a salvage pathway for ribosomes stuck in unproductive translation events, it was suggested that the RQC pathway could regulate the expression of transcripts containing stalling features. This idea was put forward by an observation that Rqc1 levels increase with the deletion of *ASC1*, *HEL2*, *LTN1*, or *RQC1*, suggesting that Rqc1 protein expression could be controlled by RQC activity, providing negative feedback for the system [6]. Rqc1 contains a well-conserved K/R poly-basic sequence, or poly-basic domain, in its N-terminal that, when substituted by alanine residues, led to increased protein expression. Based on these data, it was suggested that the RQC complex regulates Rqc1 protein levels through its poly-basic domain during Rqc1 translation [6].

It has been proposed that the trigger for RQC activity is ribosome collision, with the formation of disomes, trisomes, and even tetrasomes [14, 25-28,]. Different studies have shown that the existence of stalling features in endogenous sequences seems to be accompanied by unique structural characteristics, such as low initiation efficiency [29, 30] and/or the positioning of the stall sequence at the beginning of the transcript [31]. These adaptations are thought to result in fewer ribosomes per transcript, which would avoid ribosome collisions at the challenging segments, thus preventing RQC activation.

To test whether the RQC complex can regulate the expression of endogenous proteins through their poly-basic domains, we used a bioinformatics screen to identify the proteins with higher concentrations of K/R of the yeast proteome. Using previously published ribosome profiling data, we observed that, poly-basic domains led to slower ribosome movement during translation. However, we could not detect signatures of severe ribosome stalling, such as short mRNA fragments that characterize empty-A site ribosomes. These data suggested that poly-basic domains of endogenous proteins do not lead to RQC complex activation. To test this hypothesis, we selected genes containing extremely positively charged segments and, without altering the original promoter or the 5' UTR, we tested protein expression in yeast strains deleted for RQC components. In this experimental setting, we could not observe an increase in the final protein output, which is usually observed with reporters containing stalling sequences. This result suggested that RQC activity does not restrict protein production of endogenous genes with poly-basic domains. Our data indicate that poly-basic domains of endogenous proteins are not sufficient for the recruitment of the RQC complex. Most likely, the RQC machinery is dedicated to recycling ribosomes involved in aberrant mRNA translation and plays a small role in controlling endogenous protein levels under normal growth conditions. Rqc1 was the only target that showed up-regulation in an *ltn1Δ* background, but its levels were unperturbed when the complex was disabled at upstream steps (*asc1Δ*, *hel2Δ*) or at the CATylation step (*rqc2Δ*). These results are inconsistent with the current model of Rqc1 being regulated by the RQC complex. Additionally, the deletion of Ltn1 increased the expression of the full-length Rqc1, suggesting that this regulation could not be coupled to a putative stalling event at a polybasic sequence located at the N-terminal. Based on these data, we propose that Rqc1 regulation by Ltn1 is a post-translational event, independent of the RQC complex activity.

MATERIALS AND METHODS

Data	Source	Identifier
Translation efficiency	Weinberg et al., 2016	SRR1049521
S. cerevisiae drug-free Riboseq data	Pop et al., 2014	SRR1688545
S. cerevisiae + CHX + ANS Riboseq data	Wu et al., 2019	SRR7241922
S. cerevisiae + CHX + TIG Riboseq data	Wu et al., 2019	SRR7241919
S. cerevisiae + CHX Riboseq data	Matsuo et al., 2020	SRR9054693
Kozak score 50 pb	Cuperus et al., 2017	
Kozak score 5 pb	Cuperus et al., 2017	
Codon stabilization coefficient gene (CSCg)	Carneiro et al, 2019	
Time for translation initiation	Siwiak and Zielenkiewicz, 2010	
40S ribosomal small subunit (SSU) footprint profiling	Archer et al., 2016	
Software and algorithms		
Net charge calculation	Requião et al, 2017	
Poly (A) finder	This paper	https://github.com/ mhoyerm/adenosine_sequenc e
Consecutive K and/or R finder	This paper	https://github.com/ mhoyerm/arg_lys_sequence
Stretches of K and/or R finder	This paper	https://github.com/ mhoyerm/arg_lys_stretches
A-site estimation	This paper	https://github.com/RodolfoCar neiro/RPF- build/tree/master/Build

Data sources

Coding sequences and the annotation of *S. cerevisiae* were obtained from the *Saccharomyces* Genome Database (SGD; https://www.yeastgenome.org) and Ensembl Genomes (http://ensemblgenomes.org/). We excluded dubious ORFs as defined in the *Saccharomyces* Genome Database from our analysis.

Statistical analyses and raw data

The raw data used to create Figures 2 and 3, and the statistical analysis t-test calculations are presented in Supplementary Tables S1-2. All statistical analyses were performed with GraphPad Prism 7.

Net charge calculation

We developed a program that screens the primary sequence of a given protein and calculates the net charge in consecutive frames of a 30-amino-acid window, the approximate number of amino acids that fill the ribosome exit tunnel [32, 33]. For the net charge determination, K and R were considered +1; D and E were considered -1; and every other residue was considered 0. The N- and C-termini charges were disregarded. In a previous publication, we have shown that these simplified parameters are equivalent to a calculation using partial charges of individual amino acids at pH 7.4, according to

their p*K*a values and the Henderson-Hasselbach equation [33]. The algorithm initially establishes a stretch containing a predetermined number of amino acids (#1 to #30, for example). The stretch charge is calculated, and the charge value and the position of the first amino acid are temporarily saved to the memory. Then, our algorithm advances to the next stretch, from amino acid #2 to amino acid #31, and performs the same analysis until the end of the protein.

Codon stabilization coefficient (CSC) calculation

Based on Pearson's correlation between the frequency of occurrence of each codon in each mRNA and the mRNA half-lives, Coller and colleagues created the codon occurrence to mRNA stability coefficient (CSC) [34]. We designed and implemented an algorithm that calculated the mean value of the CSC for each gene (CSCg) [35].

Poly (A) finder

An algorithm was developed to locate the position of stretches containing only adenine (A) in an ORF. The algorithm finds adenine stretches of a chosen size for each ORF and generates as output a file containing the gene name, the stretch size, and its position in the ORF. For our analysis, we considered stretches of size equal to or larger than 10 adenines.

Stretches of K and/or R finder

We developed an algorithm that searches for stretches containing a certain number of arginines and/or lysines in a stretch of 10 amino acids in a protein. The chosen number of K and/or R in the stretches was equal to or greater than 8. The input file of the algorithm was a proteome fasta file, and the output file returned the protein name and the position of the stretch found for each protein.

Consecutive K and/or R finder

An algorithm was developed to find the location of stretches containing consecutive arginines and/or lysines in a protein. The stretches analyzed had a size equal to or larger than 6. The input file of the algorithm was a proteome fasta file, and the output file returned the protein name, the stretch size, and the position of the stretch in the protein.

Yeast strains and growth conditions

Since no commercial antibodies are available for the proteins used herein, we used a TAP-TAG collection of yeast [36]. In this collection, each ORF was tagged with a C-terminal TAG, and the same antibody could be used to detect the full collection. For each potential target of the RQC complex described before, a knockout of the *LTN1* or *ASC1* strain was created by homologous recombination. For the Rqc1 TAP-TAG strain, a knockout of *RQC2* and *HEL2* was also created. After the creation of the knockout strains, the level of the TAG proteins was compared with the wild-type strain by western blot analyses. The TAP-TAG collection strains were cultivated in YPD medium (1% yeast extract, 2% peptone, 2% glucose). The *S. cerevisiae* BY4741 strain transformed with GFP-R12-RFP vector was

cultivated in minimum synthetic medium, uracil dropout (0.67% Difco Yeast Nitrogen Base without amino acids, 2% glucose, and 0.003% methionine, leucine, and histidine).

Yeast strain constructions

For knockout creation, *S. cerevisiae* BY4741 yeast (TAP-TAG collection) was transformed using the PEG/Lithium Acetate method [37]. The product of the PCR used in the transformation was obtained by the amplification of the KanMX6 gene contained in the pFA6a-KanMx6 vector. All genes were knocked out by the insertion of the geneticin resistance gene into their original ORFs. The knockouts were confirmed using four sets of PCR. Two PCR sets combined primers for the target gene flank region (5'UTR or 3'UTR) with resistance gene primers. The other two PCR sets combined primers for the target gene flank region and ORF gene target primers. All the primer pairs used in the construction and confirmation of the mutants are depicted Support Information.

Protein extraction and SDS-PAGE

The protein extraction of the strains was conducted in the exponential phase of growth, treating equal quantities of cells using the NaOH/TCA method [38]. For the SDS-PAGE, equal sample volumes were applied on the acrylamide gels and submitted to an electrophoretic run. For the YHR131C TAP-TAG strain, protein extraction was conducted using a glass bead protein extraction protocol [39]. Total protein quantification was conducted using the BCA Protein Assay kit (Thermo Fisher), and equal protein quantities were applied on SDS-PAGE. The acrylamide gel concentration changed between 10 and 12.5% according to the size of the analyzed protein.

Western blotting

The proteins separated by SDS-PAGE were transferred to PVDF membranes using the TRANS-BLOT semidry system from BIO-RAD in transfer buffer (25 mM Tris, 192 mM glycine, 0.1% SDS, 20% methanol, pH 8.3). The membranes were blocked using a Li-Cor blocking buffer with an overnight incubation at room temperature. Incubation with primary antibodies was performed at 4ºC for at least 1 hour. The membranes were incubated in Li-Cor Odyssey secondary antibodies at room temperature for at least 1 hour and then revealed in an Odyssey scanner. The primary antibodies used were mouse anti-PGK1 (Invitrogen; dilution 1:10,000), rabbit anti-TAP-Tag (Thermo Scientific; dilution 1:10,000), mouse anti-GFP (Sigma; dilution 1:2,000), and rabbit anti-Asc1 (dilution 1:5,000) [40). The secondary antibodies that were used were anti-mouse 680 LT Li-Cor (dilution 1:10,000) and anti-rabbit 800 CW Li-Cor (dilution 1:1,000).

qPCR

For qPCR analysis, the Rqc1 TAP-TAG knockout strains constructed in this work were cultivated until the exponential phase (O.D ~ 0,6-0,8) and submitted to RNA extraction following the hot phenol method. The same protocol was used for the corresponding knockouts from a commercial collection as a control. After extraction, the RNAs were submitted to DNase I (Thermo Scientific) treatment for DNA contamination elimination. The reverse transcription was performed using random primers from

High-Capacity cDNA Reverse Transcription Kit (Thermo Scientific). For the cDNA amplification, we used SYBR Green PCR Master Mix on the StepOne Plus System (Applied Biosystems). The following protocol was used on the thermocycler: 95°C for 10 min and 40 cycles of 95°C for 15s and 60° C for 1 min, melting curve analysis: 95°C for 15s, 60°C for 1 min and 95°C for 15s. The primers used were: 5'CGGAATGCACCCGCAACATT and 5'CGGCCTTTCCCATAATTGCCA for *RQC1*, and 5' TTCCCAGGTATTGCCGAAA and 5' TTGTGGTGAACGATAGATGGA for *ACT1.* The analysis of differential expression was made by relative quantification using $2^{-\Delta\Delta Ct}$ method, and the statistical analysis was conducted on Prism.

Cycloheximide (CHX) treatment

For the experiments using CHX, concentrations ranging from 10^{-6} to 10^{-1} mg/ml were used, with concentration steps of 10 times the dilution factor. Strains grew until the exponential phase and were treated with CHX for 10 hours under shaking at 28ºC.

Ribosome profiling data

S. cerevisiae ribosome profiling data were treated as described previously [33, 41]. The data were analyzed as described by Ingolia and collaborators [42], except that the program used here was Geneious R11 (Biomatter Ltd., New Zealand) instead of the CASAVA 1.8 pipeline. The data were downloaded from GEO [43, 44], and the adaptor sequence was trimmed. The trimmed FASTA sequences were aligned to *S. cerevisiae* ribosomal and noncoding RNA sequences to remove rRNA reads. The unaligned reads were aligned to the *S. cerevisiae* S288C genome deposited in the *Saccharomyces* Genome Database. First, we removed any reads that mapped to multiple locations. Then, the reads were aligned to the *S. cerevisiae* coding sequence database, allowing two mismatches per read. For 27-29 nt footprint analyses, the ribosome profiling data were obtained without any drug, and just fragments with 27-29 nt were used [43]. For 20-22 nt footprint analyses, the ribosome profiling data were obtained with cells lysed with buffer containing 0.1 mg/ml of cycloheximide (CHX) and tigecycline (TIG) or 0.1 mg/ml of cycloheximide (CHX) and anisomycin (ANS) [44]. We normalized the coverage within the same transcript by dividing each nucleotide read by the sum of the number of reads for each gene in a given window. A window of 150 nucleotides before and after the beginning of a stretch of interest was used to calculate the number of reads of each gene. For Figures 3C-G, 5C, and S2, the ribosome profiling analysis was done following as described by Ingolia and collaborators [42]. The datasets used in our analyses were from Pop *et al.*, 2014 [43] (GEO accession GSE115162) and from Matsuo *et al.*, 2020 [25] (GEO accession GSE131214). The data were mapped to the R64.2.1 S288C Coding Sequence (CDS) from the Saccharomyces Genome Database Project. The alignment was done using Bowtie accepting two mismatches at most. The analysis was restricted to 28–32nt reads for Figure S2. The A-site was estimated based on the read length: it was defined as the codon containing the 16th nt for 28 nt reads; the 17th nt for 39, 30 and 31 nt reads; and the 18th nt for 32 nt reads, considering the first nt as 1. The A-site estimation and other Ribosome Profiling Analysis were performed by custom software written in Python 3.8.2. For Figure S2, we compared the ribosome profiles normalized for genes of the

five distinct groups with 2,000 aleatory genes. Statistical analyses were performed with multiple t-tests using the Holm-Sidak method, and the adjusted p-value was calculated for each point (GraphPad Prism 7).

Clustering analysis

The clustering analyze of Figure 2B was performed by the Euclidean distance using Orange 3 software [45].

Translation initiation score

For Figure 2B, the genome was divided into eight equal groups based on their absolute values, from highest to lowest. The groups with the highest and the lowest values were further divided, making a total of ten groups. Each group was normalized individually. Their normalized values were then adjusted to fit the whole genome, meaning that the percentages from the group with the lowest values went from 0-100% to 0-6.25%, and the percentages from the group with the highest values, from 0-100% to 93.75-100%, for example. This methodology was used to avoid the influence of a few genes with extremely high and low values that can make the visualization of the difference between the values of most proteins difficult.

RESULTS

Identification of poly-basic domains in *Saccharomyces cerevisiae*.

Our first step was to find endogenous yeast proteins containing sequences with different arrangements of poly-basic domains. The four distinct, but possibly overlapping, selected architectures were: a) sequences of 8 or more lysines/arginines residues (K/R) in a window of 10 amino acids, b) 6 or more consecutive K/R, c) adenine repeats longer than 10 nucleotide residues, and d) net charge ≥ +12 in a window of 30 residues. The architectures a) and b) were selected because they are often present in RQC activation reporters and lead to the accumulation of reads in ribosome profiling data [6, 9, 21-24, 26, 28, 41, 46-53]. The architecture c) was selected because poly (A) mRNA composition can lead to ribosome sliding, worsening the translation efficiency of these sequences [54, 55]. Architecture d) was selected because proteins containing a stretch of 30 residues with net charge ≥ +12 in (approximately the polypeptide length enclosed by the ribosome exit tunnel) are extremely rare in most proteomes [32, 33]. We also compared these architectures with a group of 1800 genes containing inhibitory codon pairs (ICP) that are known to impair translation with substantial down-regulation in protein output (e.g., CGA-CGA and CGA-CCG) [56].

We found that 354 proteins (approximately 6% of the yeast proteome) contained at least one of the characteristics described above (Figure 1A). The features most commonly found were high net charges and poly (A) sequences, with 170 and 121 sequences, respectively. Moreover, the Venn diagram (Figure 1A) shows that just 91 proteins have more than one feature at the same time, and

only 1 protein, the uncharacterized YNL143C, has all the characteristics. Localization analysis revealed that no bias was found for the different categories analyzed (Figure 1B). We selected YNL143C and other proteins with prominent poly-basic domains (Figures 1C-E), for further experimental analysis in the next sections.

Proteins with poly-basic domains show divergent patterns of translation efficiency parameters.

If the presence of poly-basic domains represents an important obstacle to translation, we expect to find signatures of poor translatability in proteins that contain them. Therefore, we asked whether the poly-basic and the ICP data sets would be significantly different from the rest of the yeast genes in experimental measurements that reflect translation activity. We chose to analyze their translation efficiency (TE), derived from ribosome profiling data and calculated from the ratio of ribosome-protected footprint, read counts to the total mRNA read counts for each gene [57]. We also analyzed and their codon stabilization coefficient (CSC), a metric derived from mRNA half-life that positively correlates with TE and protein abundance [35]. Apart from the TE and the CSC, another characteristic associated with the translational activity is the rate of translation initiation. Endogenous genes containing stall sequences have a slower or less efficient translation initiation; this is thought to be an adaptation to avoid ribosome collision during elongation [30]. Therefore, we also analyzed the time for translation initiation [58], the Kozak efficiency score of 5 nt- and 50 nt-long sequences [59], and the 5'UTR footprint density of the 40S ribosomal small subunits [60]. The cumulative distributions for each of the independent measurements described above are shown in Supplementary Figure S1. In Figure 2A, we show the p values obtained with the t-test statistical analyses comparing the individual data sets with the rest of the yeast genes. As expected, the ICP group had not only worse TE and CSC values but also slower initiation times associated with lower Kozak's efficiency scores, indicating that slower translation rates were associated with a less efficient initiation. When considered as a unique group (Fig 2A, poly-basic group), poly-basic domains had lower TE and CSC values than the other yeast genes, but no significant effect was observed when considering initiation parameters. The 5'UTR footprint density of the 40S ribosomal small subunits yielded no significant difference against ICP and poly-basic group, probably because this new technique still provides low coverage of the genome.

To evaluate the characteristics of each gene from the poly-basic group, we created a unified rank for all the measurements above. Using the values obtained with the entire yeast data set, each score system was organized in 10 bins; next, the maximum and minimum bin values were used to normalize the scale. We created a heat map where 1 refers to the maximum translation activity (TE, CSC) or optimal initiation (translation initiation time, Kozak's efficiency, and the 5'UTR footprint density of the 40S ribosomal small subunits). A cluster analysis of the 354 genes with poly-basic sequences revealed that the group is rather heterogeneous and displays all spectra of characteristics, with some sequences having excellent translatability indicators with poor initiation scores and vice versa (Figure 2B). Notably, the sequences with the highest positively charged domains or longest poly (A) tracts also showed diverse features (Figure 2B, lower panel).

Ribosome profiling of endogenous proteins with poly-basic domains.

To test if the selected poly-basic domains architectures have any effect on translation, we took advantage of previously published ribosome footprint profiling data. This methodology is based on the deep sequencing of ribosome-protected mRNA fragments [42]. During translation, each ribosome encloses a portion of mRNA, protecting it against RNAase digestion. This enclosure allows protected fragments to be sequenced, generating a map, at nucleotide resolution, of translational activity for transcribed genes. This map can report the occurrence of ribosome pausing or stalling, revealed by the accumulation of deep sequencing reads at specific positions of the transcript. Recently, it was shown that ribosomes with open A sites produce shorter mRNA fragments (20-22 nucleotides) than ribosomes with occupied A sites, which produce longer mRNA fragments (27-29 nucleotides) [44]. Structural and biochemical evidence demonstrated that stalled ribosomes are in the open A site conformation [61], suggesting that the 20-22 fragments could more accurately reflect translation efficiency.

Analyses of the ribosome profiling (RP) data of the poly-basic domains indicated an enrichment of 27-29 nucleotide long footprints, which indicates a reduction in translational speed (Figure 3A). We also plotted the number of reads at an approximate position of the ribosomal A site and observed an accumulation of reads downstream the first codon of the poly-basic domain, which is consistent with the result shown in Figure 3A (Supplementary Figure S2). However, when the 20-22 nucleotide long footprints were analyzed, the only group that showed an accumulation of reads was the ICP group (Figure 3B). The same result was observed with another RP dataset obtained with a different combination of translation inhibitors (cycloheximide and anisomycin) (Supplementary Figure S3). These results suggested that poly-basic domains can reduce elongation rates but is not enough to cause severe ribosome stalling leading to the accumulation of open-A site ribosomes.

Another way to evaluate if endogenous poly-basic domains can cause ribosome stalling is by the analysis of ribosome collisions. Ribosome collisions are the result of the contact between a stalled ribosome and a second translating ribosome and act as signals to the RQC activation pathway [26, 28]. If the endogenous poly-basic domains are stalling sequences, we should observe not only an increase in short nucleotide long footprints corresponding to an empty A-site but also an increase in long nucleotide long footprints upstream of the stalling site with a periodicity of roughly 10 codons, which corresponds to the collided ribosomes [25]. Among the 354 genes with one or more poly-basic domains identified in our bioinformatics screening, we selected the four top genes (Figure 1C-E) for subsequent analyses: YBR054W/Yro2 (net charge = +17), YGL078C/Dbp3 (net charge = +16), YHR131C (14 consecutive arginines) and YNL143C (32 consecutive adenines coding for 10 lysines) (Figures 2 and 3). To expand our analyses, we also included genes with the two longest poly-basic domains, stretches with 6 or more K/R in 10 amino acids of the yeast proteome: YLR197W/Nop56 and YOR310C/Nop58 (42 and 44 residues long poly-basic domains, respectively). As a positive control, we included the gene Sdd1, which has been recently described as an endogenous target of the RQC complex and shows the ribosome collision footprint pattern described above [25]. Our analyses revealed that, apart from Sdd1, none of the selected proteins had the periodic footprint

pattern that would indicate ribosome collisions (indicated by dashed lines), further confirming that their poly-basic domains are not sufficient to cause ribosome stalling (Figure 3C-G). YHR131C and YNL143C were not included in Figure 3 because of their low RP coverage. The ribosome collision footprint pattern was also carefully investigated for the remaining 348 proteins with poly-basic domains, with similar negative results (Supplementary Figure S4).

Proteins with poly-basic domains are not targets of the RQC complex.

To confirm that poly-basic domains do not cause ribosome stalling and translation termination, we examined whether the deletion of RQC components had any effect on the final protein output of our selected proteins, which would imply a regulatory role for the RQC complex on their expression. TAP-tagged versions of the 6 genes selected from the previous section were obtained from a yeast library containing a chromosomally integrated TAP-tag at the 3' region of each gene, thus preserving the natural promoter region and the 5' UTR of the transcripts [36]. We chose to delete Ltn1, the E3 ubiquitin ligase of the RQC, and Asc1, involved in the modulation of RQC activity because they represent different steps in the RQC pathway and have been consistently used in the investigation of its mechanism of action [6, 20, 25, 53, 62-64]. Each TAP-tag strain was then used to generate the *ltn1*Δ or *asc1*Δ variants, confirmed by PCR and western blot (Supplementary Figure S5). In Figure 4, we show the protein expression of each gene as probed by anti-TAP western blot in the 3 different backgrounds and the relative quantification of at least 3 independent experiments. For each gene, we also show the complete ribosome profile and the protein TE value in comparison to the yeast proteome. We observed that *ltn1*Δ caused no regulation of protein levels in any of the transcripts, what is expected since the TAP-tag is downstream the stalling sequences. The results with *asc1*Δ revealed no regulation in YGL078C/Dbp3, YNL197W/Nop56, and YOR310C/Nop58, and down-regulation in YBR054W/Yor2, YHR131C, and YNL143C (Figure 4). This pattern could not be explained by any feature, such as ribosome density on the poly-basic site, TE, or initiation time. All of these proteins differed from what was observed with reporters containing stalling sequences, which had their expression levels up-regulated in *asc1*Δ strains [6, 20, 53, 62-64]. Therefore, we conclude that, even though Asc1 seems to control the expression of some of these proteins, the complete RQC pathway is not actively degrading any of the endogenous proteins tested.

Ltn1 regulates rqc1, but the mechanism is independent of other RQC.

To investigate the mechanism of Rqc1 regulation by the RQC, we analyzed Rqc1 expression in the backgrounds *ltn1Δ*, *asc1Δ*, *rqc2Δ*, and *hel2Δ*. For this, we used a TAP-tagged version of Rqc1, meaning that we can only observe an increase in the levels of the full-length peptide. As previously described, the deletion of Ltn1 increased the expression levels of Rqc1, but, different from Brandman *et al.*, 2012, the deletion of *ASC1*, *HEL2* or *RQC2* had no impact on Rqc1 expression levels (Figure 5A). Rqc1 regulation was different from that of GFP-R12-RFP, an artificial reporter with a poly-basic sequence known to be a target of the RQC [6, 20, 21, 23], which presented increased levels of GFP (truncated peptide) when Ltn1 was deleted, and increased levels of both GFP and RFP (full-length peptide) when *ASC1* or *HEL2* was deleted (Supplementary Figure S6). It is noteworthy that the Rqc1

poly-basic sequence (blue arrow) is located at the protein N-terminal, meaning that only the up-stream poly-peptide fragment should be stabilized in an *ltn1*Δ strain. Therefore, the overexpression of the full-length Rqc1 (Figure 5A and Brandman et al., 2012) is inconsistent with the current model for RQC action. Moreover, the RQC1 mRNA levels were the same in all tested backgrounds, which is also unexpected for a transcript containing a stalling sequence (Figure 5B). To test if the presence of the C-terminal TAP-TAG influenced *RQC1* mRNA levels, we repeated the qPCR analysis with the non-tagged WT *RQC1*, and the same results were observed (Supplementary Figure S7). Moreover, ribosome profiling of Rqc1 did not show the characteristic footprint pattern of ribosome collisions and accumulation of short fragments (Figure 5C – blue dashed lines), further indicating that the Rqc1 basic polypeptide does not act as a strong stalling sequence. Another way to verify whether a transcript is a target of the RQC complex is through the saturation of the rescue system; if the protein of interest is a target, the saturation of the RQC will increase its expression levels. The saturation can be accomplished by treatment with low concentrations of an elongation inhibitory drug, such as cycloheximide (CHX), which leads to multiple events of ribosome stalling and RQC activation [6, 26]. Accordingly, Figure 5D shows that the GFP signal from the GFP-R12-RFP reporter increased at low CHX concentrations and that this effect was Ltn1 dependent (Figure 5D). However, when we tested the Rqc1-TAP-TAG strain, we could not observe the same result, further indicating that the RQC complex does not control the final Rqc1 protein output during translation (Figure 5E).

DISCUSSION

The hypothesis that the translation of positive amino acids leads to ribosome stalling and translation repression has been disputed since shortly after its proposal. The Grayhack group was the first to propose that codon bias was the major factor behind ribosome stalling observed with positive sequences [63, 65-67]. They examined the influence of all sequences of ten repeated codons on luciferase activity, and they demonstrated that the codon choice had a stronger effect than the final charge. Of particular interest is the codon CGA, which leads to a 99% reduction in luciferase activity in comparison with the codon AGA, both coding for arginine [67]. In recent works, they proposed that the CGA codon does not need to be highly concentrated to cause translational repression: the presence of as few as two consecutive CGA codons [64, 66] or specific combinations of CGA with a second codon, the so-called inhibitory pairs [56, 68], are sufficient to inhibit translation. In the same token, the Green and Djuranovic groups proposed that poly (A) sequences are the major cause of translation repression observed with poly-lysine stretches [11, 54, 61, 69]. They showed that the presence of six AAA codons in tandem leads to an approximately 30% to 70% reduction in protein expression in comparison to AAG in different organisms and that this reduction is proportional to the quantity of AAA [11, 54]. Letzring *et al.*, 2010, had already observed that the translation of the codon AAA is worse than that of AAG [67], but what both the Green and the Djuranovic groups demonstrated is that the translation disruption is caused by the presence of a poly (A) sequence in the mRNA, independent of codon choice. Sequences of AAA codons followed or preceded by codons rich in adenine (preceded by CAA and followed by AAC, for example) can cause greater translation repression than the same sequences of AAA codons alone [11]. The presence of an extended poly (A)

sequence can cause ribosomal sliding, a slow ribosome frameshifting movement that leads to the incorporation of extra lysine residues in the nascent peptide. More importantly, it leads to the translation of premature termination codons, which are targeted by nonsense-mediated decay and lead to translation termination [54]. This may be caused by the helical conformation that poly (A) sequences can adopt, which allows them to enter the ribosomal A site and inhibit translation [61].

On the other hand, there is still evidence that the charge itself can play a role in translation dynamics. Positively charged sequences, independent of codon choice, are translated at slower rates than negatively charged or non-charged sequences [33, 41], which is still considered a hallmark of ribosome stalling. Additionally, proteins with a high density of negatively charged amino acids in a 30-residue stretch are far more common than proteins with a high concentration of arginines and/or lysines in 30 residues, suggesting a negative selection pressure against long poly-basic sequences [33, 70]. Moreover, even in the papers that highlight the effect of codon choice on the inhibitory effect observed with poly-basic sequences, we can see evidence that charge alone is still a relevant factor for translation efficiency. Both Arthur *et al.*, 2015, and Koutmou *et al.*, 2015, show that the translation of twelve AGA codons leads to a reduction in protein amount of approximately 50% [11, 54]. Taken together, the literature suggests that the translation of rare codons (CGA) and poly (A) sequences are detrimental for translation but does not rule out the effect of positively charged segments as RQC complex activation factors.

Our search for endogenous targets of the RQC complex was motivated by the possibility of translation repression caused by positively charged segments, as described above. Rqc1 was the first example of an endogenous protein with co-translational regulation promoted by the RQC complex [6]. Our results show that poly-basic domains similar to Rqc1 are relatively common (Figure 1), and even though they are translated at slower rates (Figure 3A and S2) [41], it seems that this is not sufficient to cause a severe ribosome stalling with empty A site, as characterized by the stabilization of 20 to 22 nucleotide footprints [44, 61] (Figure 3B, S3 and S4). Consistent with our results that poly-basic domains do not cause premature termination, not even the most positively charged domains could be identified as endogenous targets of the RQC complex (Figure 4). Recently, the yeast gene *SDD1* was identified as a natural target of RQC. The staling sequence contains a rare combination of inhibitory features: a stretch of bulky amino acids, followed by CGA-CGA ICP, and a short poly (A) tract. The result is a combination of amino acid−specific inactivation of the Peptidyl- transferase center at the 60S subunit and mRNA-specific obstruction of the decoding center at the 40S subunit [25]. Based on these data, we believe that the RQC complex targets are mainly aberrant translation products and, therefore, play a minor role in the regulation of protein with poly-basic domains.

Similar to what was observed with other proteins with poly-basic domains, we showed that the RQC complex does not inhibit Rqc1 translation. In our hands, *rqc2*Δ, *hel2*Δ, and *asc1*Δ did not significantly affected relative Rqc1 levels, which speaks against an RQC complex-dependent regulation (Fig 5A). However, Rqc1 over expression in a *ltn1*Δ strain agree with the result from Brandman *et al.*, 2012, which, at the time, was interpreted as evidence that the Rqc1-Flag nascent protein was being targeted by the RQC complex. However, the deletion of Ltn1 alone is usually

associated with the stabilization of translation intermediates (up to the stalling point), not full-length constructs, as we observed in the GFP-R12-RFP reporter (Supplementary Figure S6). Taking our data into account, we speculate that the K-rich basic domain can function as the ubiquitylation site involved in Rqc1 regulation. Since it was previously reported that Ltn1 could ubiquitylate substrates independently of the RQC complex [71, 72], we speculate that Ltn1 might regulate Rqc1 in a process that is uncoupled from Rqc1 translation.

DATA AVAILABILITY

The algorithms used in this work are available in the GitHub repository (https://github.com/ mhoyerm).

ACKNOWLEDGMENT

We thank Dr. Wendy Gilbert for providing us with the Asc1 antibody and Dr. Onn Brandman for providing us with a GFP-R12-RFP stalling reporter. We also thank Santiago Alonso and Ramon Cid for their technical support.

FUNDING

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo a Pesquisa do Estado do Rio de Janeiro (FAPERJ) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest with the contents of this article.

Supplemental data for this article can be accessed on the publisher's website.

REFERENCES

1. Lu,J. and Deutsch,C. (2008) Electrostatics in the ribosomal tunnel modulate chain elongation rates. *J. Mol. Biol.*, **384**, 73–86.
2. Lu,J., Kobertz,W.R. and Deutsch,C. (2007) Mapping the electrostatic potential within the ribosomal exit tunnel. *J. Mol. Biol.*, **371**, 1378–1391

3. Frischmeyer,P.A., van Hoof,A., O'Donnell,K., Guerrerio,A.L., Parker,R. and Dietz,H.C. (2002) An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science*, **295**, 2258–2261
4. van Hoof,A., Frischmeyer,P.A., Dietz,H.C. and Parker,R. (2002) Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science*, **295**, 2262–2264.
5. Inada,T. and Aiba,H. (2005) Translation of aberrant mRNAs lacking a termination codon or with a shortened 3'-UTR is repressed after initiation in yeast. *EMBO J.*, **24**, 1584–1595.
6. 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.
7. 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.
8. 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.
9. Choe,Y.-J., Park,S.-H., Hassemer,T., Körner,R., Vincenz-Donnelly,L., Hayer-Hartl,M. and Hartl,F.U. (2016) Failure of RQC machinery causes protein aggregation and proteotoxic stress. *Nature*, **531**, 191–195.
10. 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.
11. Arthur,L.L., Pavlovic-Djuranovic,S., Koutmou,K.S., Green,R., Szczesny,P. and Djuranovic,S. (2015) Translational control by lysine-encoding A-rich sequences. *Sci. Adv.*, **1**, e1500154.
12. 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.
13. 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.
14. Juszkiewicz,S., Chandrasekaran,V., Lin,Z., Kraatz,S., Ramakrishnan,V. and Hegde,R.S. (2018) ZNF598 Is a Quality Control Sensor of Collided Ribosomes. *Mol. Cell*, **72**, 469–481.
15. Joazeiro,C.A.P. (2019) Mechanisms and functions of ribosome-associated protein quality control. *Nat. Rev. Mol. Cell Biol.*, **20**, 368–383.
16. Joazeiro,C.A.P. (2017) Ribosomal Stalling During Translation: Providing Substrates for Ribosome-Associated Protein Quality Control. *Annu. Rev. Cell Dev. Biol.*, **33**, 343–368.

17. 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.
18. Gamerdinger,M. (2016) Protein quality control at the ribosome: focus on RAC, NAC and RQC. *Essays Biochem.*, **60**, 203–212.
19. Juszkiewicz,S. and Hegde,R.S. (2017) Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination. *Mol. Cell*, **65**, 743–750.
20. Sitron,C.S., Park,J.H. and Brandman,O. (2017) Asc1, Hel2, and Slh1 couple translation arrest to nascent chain degradation. *RNA*, **23**, 798–810.
21. 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) Protein synthesis. Rqc2p and 60S ribosomal subunits mediate mRNA-independent elongation of nascent chains. *Science*, **347**, 75–78.
22. 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.
23. 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.
24. Sitron,C.S. and Brandman,O. (2019) CAT tails drive degradation of stalled polypeptides on and off the ribosome. *Nat. Struct. Mol. Biol.*, **26**, 450–459.
25. Matsuo Y, Tesina P, Nakajima S, Mizuno M, Endo A, Buschauer R, Cheng J, Shounai O, Ikeuchi K, Saeki Y, Becker T, Beckmann R, Inada T. (2020) RQT complex dissociates ribosomes collided on endogenous RQC substrate SDD1. *Nat Struct Mol Biol.*, **27**, 323-332.
26. Simms,C.L., Yan,L.L. and Zaher,H.S. (2017) Ribosome Collision Is Critical for Quality Control during No-Go Decay. *Mol. Cell*, **68**, 361–373.
27. Simms,C.L., Yan,L.L., Qiu,J.K. and Zaher,H.S. (2019) Ribosome Collisions Result in +1 Frameshifting in the Absence of No-Go Decay. *Cell Rep.*, **28**, 1679–1689.
28. Ikeuchi,K., Tesina,P., Matsuo,Y., Sugiyama,T., Cheng,J., Saeki,Y., Tanaka,K., Becker,T., Beckmann,R. and Inada,T. (2019) Collided ribosomes form a unique structural interface to induce Hel2-driven quality control pathways. *EMBO J.*, **38**, e100276.
29. Heyer,E.E. and Moore,M.J. (2016) Redefining the Translational Status of 80S Monosomes. *Cell*, **164**, 757–769.
30. Park,H. and Subramaniam,A.R. (2019) Inverted translational control of eukaryotic gene expression by ribosome collisions. *PLoS Biol.*, **17**, e3000396.
31. Tuller T, Veksler-Lublinsky I, Gazit N, Kupiec M, Ruppin E, Ziv-Ukelson M. (2011) Composite effects of gene determinants on the translation speed and density of ribosomes. *Genome Biol.*, **12**, R110.

32. Fedyukina,D.V. and Cavagnero,S. (2011) Protein folding at the exit tunnel. *Annu. Rev. Biophys.*, **40**, 337–359.
33. Requião,R.D., Fernandes,L., de Souza,H.J.A., Rossetto,S., Domitrovic,T. and Palhano,F.L. (2017) Protein charge distribution in proteomes and its impact on translation. *PLoS Comput. Biol.*, **13**, e1005549.
34. Presnyak,V., Alhusaini,N., Chen,Y.-H., Martin,S., Morris,N., Kline,N., Olson,S., Weinberg,D., Baker,K.E., Graveley,B.R., *et al.* (2015) Codon Optimality Is a Major Determinant of mRNA Stability. *Cell*, **160**, 1111–1124.
35. Carneiro,R.L., Requião,R.D., Rossetto,S., Domitrovic,T. and Palhano,F.L. (2019) Codon stabilization coefficient as a metric to gain insights into mRNA stability and codon bias and their relationships with translation. *Nucleic Acids Res.*, **47**, 2216–2228.
36. Puig,O., Caspary,F., Rigaut,G., Rutz,B., Bouveret,E., Bragado-Nilsson,E., Wilm,M. and Séraphin,B. (2001) The Tandem Affinity Purification (TAP) Method: A General Procedure of Protein Complex Purification. *Methods*, **24**, 218–229.
37. Gietz,R.D. and Schiestl,R.H. (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.*, **2**, 31–34.
38. Volland,C., Urban-Grimal,D., Géraud,G. and Haguenauer-Tsapis,R. (1994) Endocytosis and degradation of the yeast uracil permease under adverse conditions. *J. Biol. Chem.*, **269**, 9833–9841.
39. Boucherie,H. and Monribot-Espagne,C. (2006) Two-Dimensional Gel Electrophoresis of Total Yeast Proteins. In: Xiao W. (eds) *Yeast Protocol*, Methods in Molecular Biology, Humana Press, Totowa, NJ, pp. 47–64.
40. Coyle,S.M., Gilbert,W.V. and Doudna,J.A. (2009) Direct link between RACK1 function and localization at the ribosome in vivo. *Mol. Cell. Biol.*, **29**, 1626–1634.
41. Requião,R.D., de Souza,H.J.A., Rossetto,S., Domitrovic,T. and Palhano,F.L. (2016) Increased ribosome density associated to positively charged residues is evident in ribosome profiling experiments performed in the absence of translation inhibitors. *RNA Biol.*, **13**, 561–568.
42. Ingolia,N.T., Brar,G.A., Rouskin,S., McGeachy,A.M. and Weissman,J.S. (2012) The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat. Protoc.*, **7**, 1534–1550.
43. Pop,C., Rouskin,S., Ingolia,N.T., Han,L., Phizicky,E.M., Weissman,J.S. and Koller,D. (2014) Causal signals between codon bias, mRNA structure, and the efficiency of translation and elongation. *Mol. Syst. Biol.*, **10**, 770–770.
44. Wu,C.C.-C., Zinshteyn,B., Wehner,K.A. and Green,R. (2019) High-Resolution Ribosome Profiling Defines Discrete Ribosome Elongation States and Translational Regulation during Cellular Stress. *Mol. Cell*, **73**, 959–970.
45. Demšar,J., Curk,T., Erjavec,A., Gorup Č., Hočevar, T. Milutinovič, M., Možina, M., Polajnar, M., Toplak, M., Starič, A., *et al*. (2013) Orange: data mining toolbox in Python. *J. Mach. Learn. Res.,***14**, 2349-2353.

46. Chiabudini M, Conz C, Reckmann F, Rospert S. (2012) Ribosome-associated complex and Ssb are required for translational repression induced by polylysine segments within nascent chains. *Mol Cell Biol.*, **32**, 4769-79.
47. Shao S, von der Malsburg K, Hegde RS. (2013) Listerin-dependent nascent protein ubiquitination relies on ribosome subunit dissociation. *Mol Cell.*, **50**, 637-48.
48. Chiabudini M, Tais A, Zhang Y, Hayashi S, Wölfle T, Fitzke E, Rospert S. (2014) Release factor eRF3 mediates premature translation termination on polylysine-stalled ribosomes in Saccharomyces cerevisiae. *Mol Cell Biol.*, **34**, 4062-76.
49. Matsuda R, Ikeuchi K, Nomura S, Inada T. (2014) Protein quality control systems associated with no-go and nonstop mRNA surveillance in yeast. *Genes Cells.*, **19**, 1-12.
50. Crowder JJ, Geigges M, Gibson RT, Fults ES, Buchanan BW, Sachs N, Schink A, Kreft SG, Rubenstein EM. (2015) Rkr1/Ltn1 Ubiquitin Ligase-mediated Degradation of Translationally Stalled Endoplasmic Reticulum Proteins. *J Biol Chem.*, **290**, 18454-66.
51. Saito K, Horikawa W, Ito K. (2015) Inhibiting K63 polyubiquitination abolishes no-go type stalled translation surveillance in Saccharomyces cerevisiae. (2015) *PLoS Genet.*, **11**, e1005197.
52. Doamekpor SK, Lee JW, Hepowit NL, Wu C, Charenton C, Leonard M, Bengtson MH, Rajashankar KR, Sachs MS, Lima CD, Joazeiro CA. (2016) Structure and function of the yeast listerin (Ltn1) conserved N-terminal domain in binding to stalled 60S ribosomal subunits. *Proc Natl Acad Sci U S A.*, **113**, E4151-60.
53. Matsuo Y, Ikeuchi K, Saeki Y, Iwasaki S, Schmidt C, Udagawa T, Sato F, Tsuchiya H, Becker T, Tanaka K, Ingolia NT, Beckmann R, Inada T. (2017) Ubiquitination of stalled ribosome triggers ribosome-associated quality control. *Nat Commun.*, **8**, 159.
54. Koutmou,K.S., Schuller,A.P., Brunelle,J.L., Radhakrishnan,A., Djuranovic,S. and Green,R. (2015) Ribosomes slide on lysine-encoding homopolymeric A stretches. *Elife*, **4**, e05534.
55. Habich,M., Djuranovic,S. and Szczesny,P. (2016) PATACSDB—the database of polyA translational attenuators in coding sequences. *Peer J. Comput. Sci.*, **2**, e45.
56. 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.
57. Weinberg,D.E., Shah,P., Eichhorn,S.W., Hussmann,J.A., Plotkin,J.B. and Bartel,D.P. (2016) Improved Ribosome-Footprint and mRNA Measurements Provide Insights into Dynamics and Regulation of Yeast Translation. *Cell Rep.*, **14**, 1787–1799.
58. Siwiak,M. and Zielenkiewicz,P. (2010) A Comprehensive, Quantitative, and Genome-Wide Model of Translation. *PLoS Comput Biol*, **6**, e1000865–15.
59. Cuperus,J.T., Groves,B., Kuchina,A., Rosenberg,A.B., Jojic,N., Fields,S. and Seelig,G. (2017) Deep learning of the regulatory grammar of yeast 5′ untranslated regions from 500,000 random sequences. *Genome Res.*, **27**, 2015–2024.
60. Archer,S.K., Shirokikh,N.E., Beilharz,T.H. and Preiss,T. (2016) Dynamics of ribosome scanning and recycling revealed by translation complex profiling. *Nature*, **535**, 570–574.

61. Tesina P, Lessen LN, Buschauer R, Cheng J, Wu CC, Berninghausen O, Buskirk AR, Becker T, Beckmann R, Green R. (2020) Molecular mechanism of translational stalling by inhibitory codon combinations and poly(A) tracts. *EMBO J.*,**39**, e103365.
62. Kuroha K, Akamatsu M, Dimitrova L, Ito T, Kato Y, Shirahige K, Inada T. (2010) Receptor for activated C kinase 1 stimulates nascent polypeptide-dependent translation arrest. *EMBO Rep.*, **11**, 956-61.
63. Letzring,D.P., Wolf,A.S., Brule,C.E. and Grayhack,E.J. (2013) Translation of CGA codon repeats in yeast involves quality control components and ribosomal protein L1. *RNA*, **19**, 1208–1217.
64. Wang,J., Zhou,J., Yang,Q. and Grayhack,E.J. (2018) Multi-protein bridging factor 1(Mbf1), Rps3 and Asc1 prevent stalled ribosomes from frameshifting. *Elife*, **7**, 934.
65. Dean,K.M. and Grayhack,E.J. (2012) RNA-ID, a highly sensitive and robust method to identify cis-regulatory sequences using superfolder GFP and a fluorescence-based assay. *RNA*, **18**, 2335–2344.
66. Wolf,A.S. and Grayhack,E.J. (2015) Asc1, homolog of human RACK1, prevents frameshifting in yeast by ribosomes stalled at CGA codon repeats. *RNA*, **21**, 935–945.
67. 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.
68. Ghoneim,D.H., Zhang,X., Brule,C.E., Mathews,D.H. and Grayhack,E.J. (2018) Conservation of location of several specific inhibitory codon pairs in the Saccharomyces sensu stricto yeastsreveals translational selection. *Nucleic Acids Res.*, **59**, 149–14.
69. Arthur,L.L. and Djuranovic,S. (2018) PolyA tracks, polybasic peptides, poly-translational hurdles. *WIREs RNA*, **9**, e1486–12.
70. Karlin,S., Brocchieri,L., Bergman,A., Mrazek,J. and Gentles,A.J. (2002) Amino acid runs in eukaryotic proteomes and disease associations. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 333–338.
71. Maurer,M.J., Spear,E.D., Yu,A.T., Lee,E.J., Shahzad,S. and Michaelis,S. (2016) Degradation Signals for Ubiquitin-Proteasome Dependent Cytosolic Protein Quality Control (CytoQC) in Yeast. *G3 (Bethesda)*, **6**, 1853–1866.
72. Zuzow,N., Ghosh,A., Leonard,M., Liao,J., Yang,B. and Bennett,E.J. (2018) Mapping the mammalian ribosome quality control complex interactome using proximity labeling approaches. *Mol. Biol. Cell.*, **29**, 1258–1269.

FIGURES

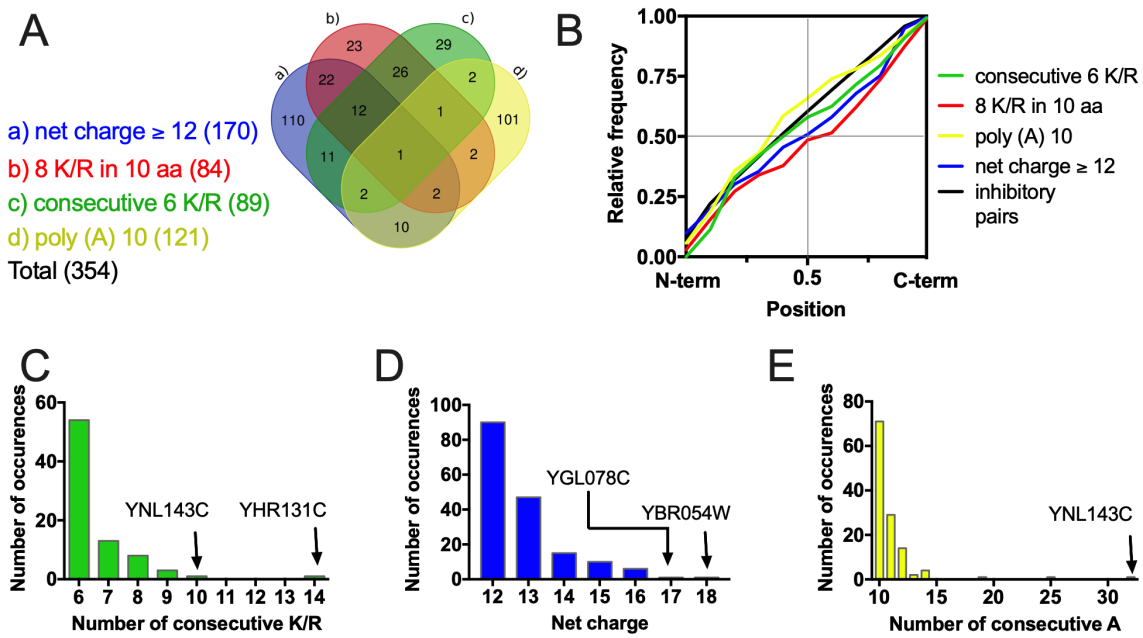

Figure 1. Bioinformatics screening to identify the polybasic domains in the yeast proteome. (A) Venn diagram showing the genes identified as potential targets of the RQC complex. The list of identified genes is presented in Supplemental Table S1. (B) Cumulative frequency distribution of the location of the poly-basic domains of each group. For each gene, the desired attribute was located, and its position was determined in relation to the length of the gene. The frequency of distribution of the genes is depicted in the following categories: (C) consecutive 6 K/R, (D) net charge equal to or higher than + 12, and (E) 10 or more consecutive adenines. The genes with the highest scores of each categories were highlighted.

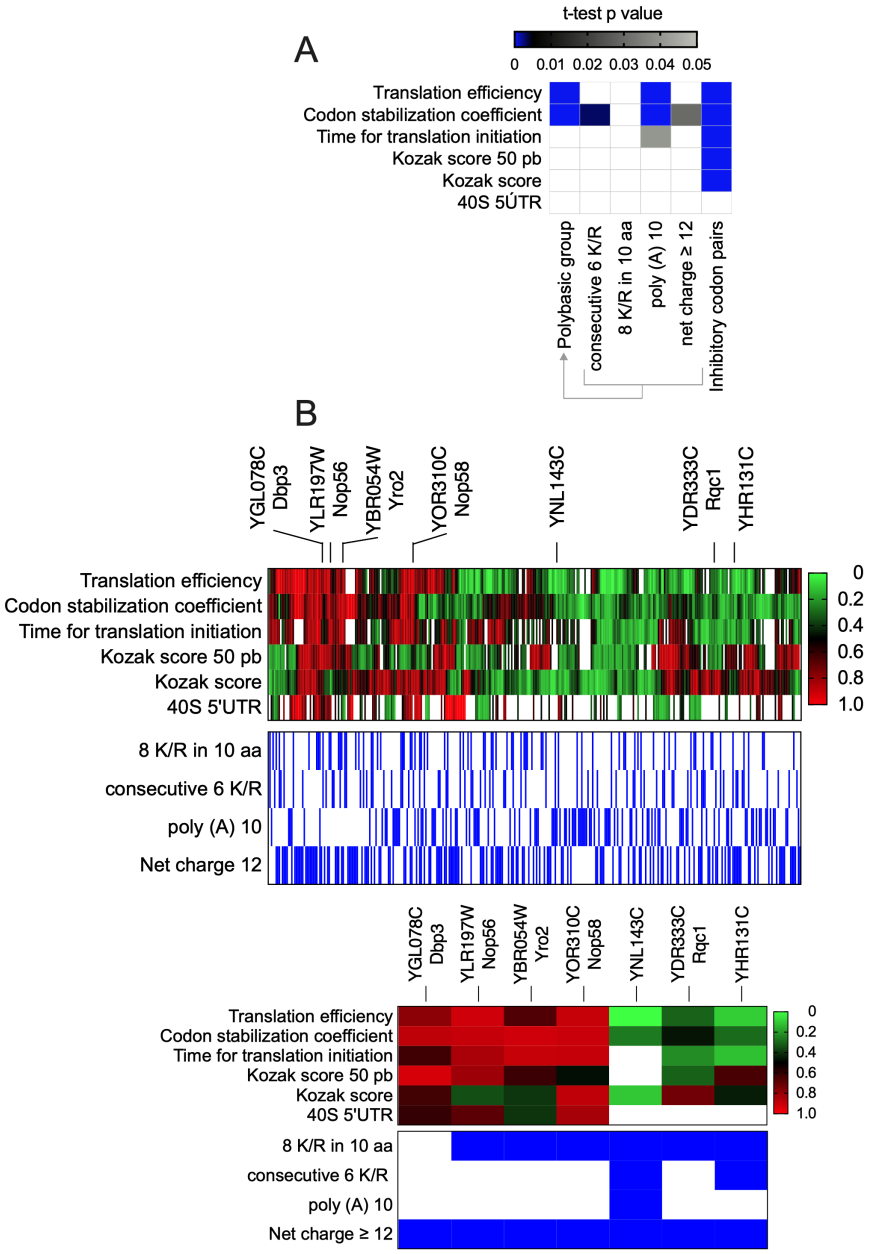

Figure 2. Translation initiation rate of genes with polybasic domains. (A) The different groups of genes with poly-basic domains were compared regarding parameters that, to some extent, reflect translation initiation rates. The Kolmogorov-Smirnov test p values are plotted for each comparison, and a white rectangle means that a nonsignificant difference was found (p values >0.05). As a control, genes with one of the 17 inhibitory codon pairs (ICP) were used. (B) The values of the translation initiation parameters were normalized (where 0 and 1 represent the lowest and highest initiation rates of the full genome, respectively), and then the 354 identified genes were clustered by the Euclidean distance. (C) Translation initiation parameters cluster of the genes with the greatest poly-basic domains (Figure 1C-E) and YDR333C/Rqc1 (included for comparison). The raw data, sample size, and p values are presented in Supplemental Table S2.

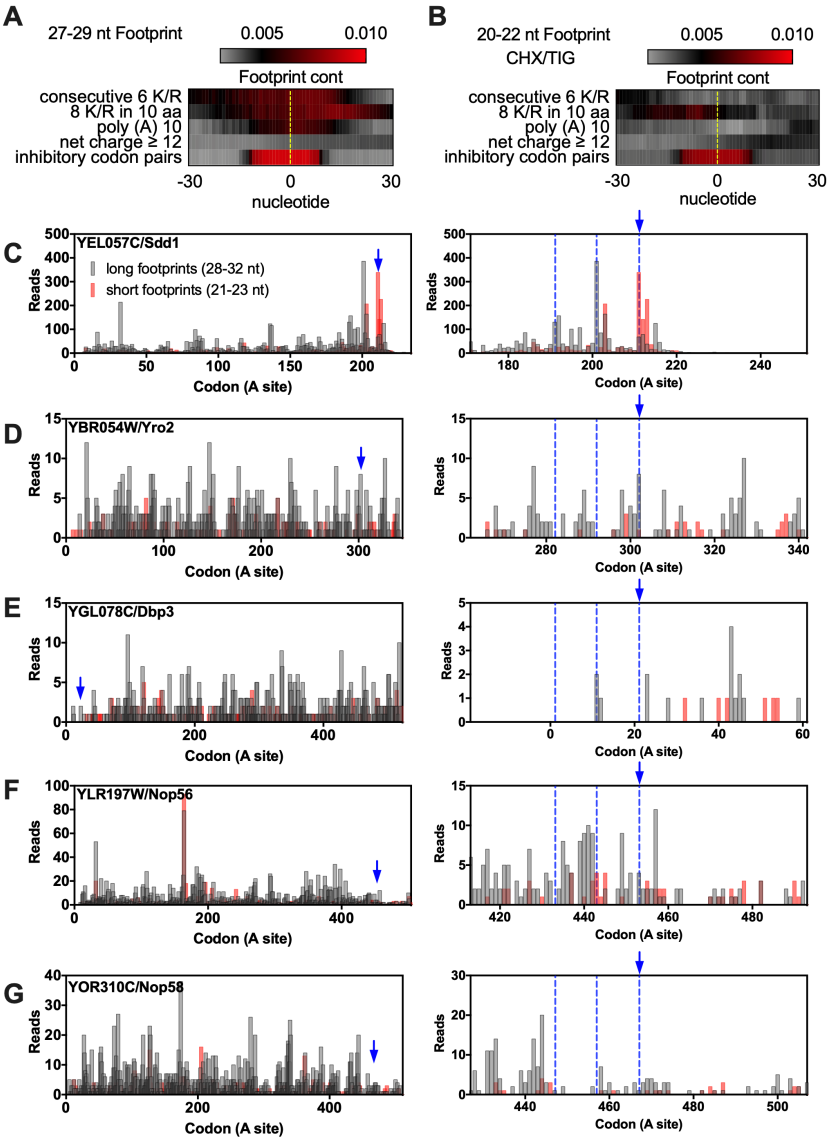

Figure 3. Ribosome profiling data of poly-basic domains. (A) 27-29 and (B) 20-22 nucleotide footprint analyses of genes containing poly-basic domains. As control, genes with one of the 17 inhibitory codon pairs (ICP) were used. The dotted yellow line represents the start of the poly-basic domain or of the ICP. (C) Sdd1 was used as standard endogenous target of RQC. 28-32 and 21-23 nucleotide footprint analysis of Sdd1 shows an accumulation of short reads at its stalling site, indicating ribosome stalling, and accumulation of long reads upstream of the stalling site with a periodicity of roughly 10 codons, indicating ribosome collisions, which are highlighted on the right panel. Blue arrow indicates the start of the staling domain, and the blue dotted line indicates the hypothetical position of the collided ribosomes. (D-G) 28-32 and 21-23 nucleotide footprint analysis of the proteins with the highest poly-basic domains (YHR131C and YNL143C were omitted because of their low RP coverage). For panels A-B, the full read was used while for panels C-G, just the ribosomal A site read was used. The reads were plotted at an approximate position of the ribosomal A site.

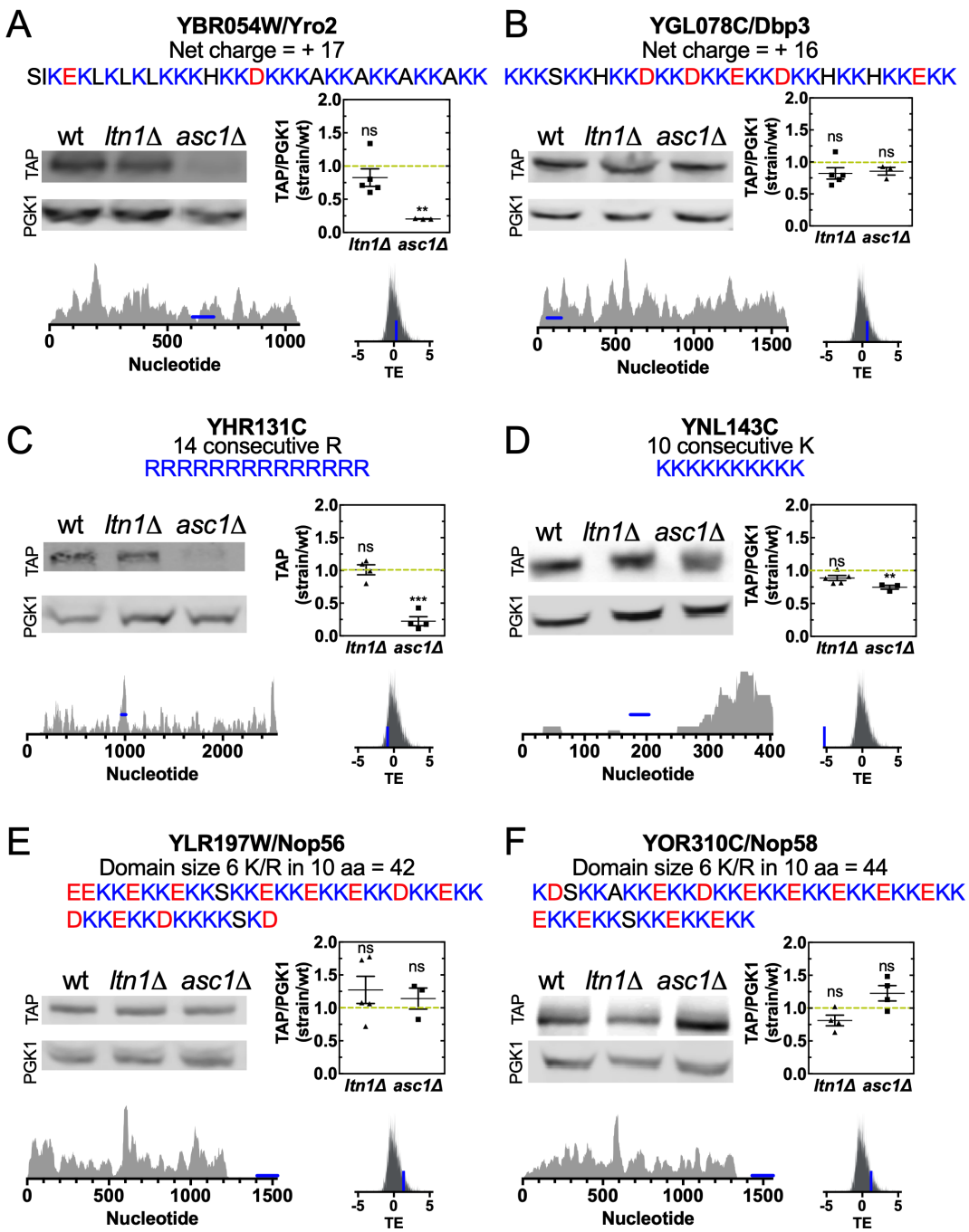

Figure 4. Proteins with the greatest poly-basic domains are not targets of the RQC complex in yeast. (A-F) Proteins with the greatest poly-basic domains of yeast were measured by western blot in wild type (wt), *lnt1*Δ, and *asc1*Δ strains. For each protein, the primary sequence of their poly-basic domain, their ribosome profiling data, and translation efficiency (TE) are presented. The blue bar in the ribosome profiling data indicates the position of the poly-basic sequence, while the blue bar in the TE data shows the TE of the specific gene inside the TE distribution of all genes. The band intensity was quantified and normalized in relation to the housekeeping gene phosphoglycerate kinase (*PGK1*). One-way ANOVA with Bonferroni's multiple comparison test was used for statistical analyses.

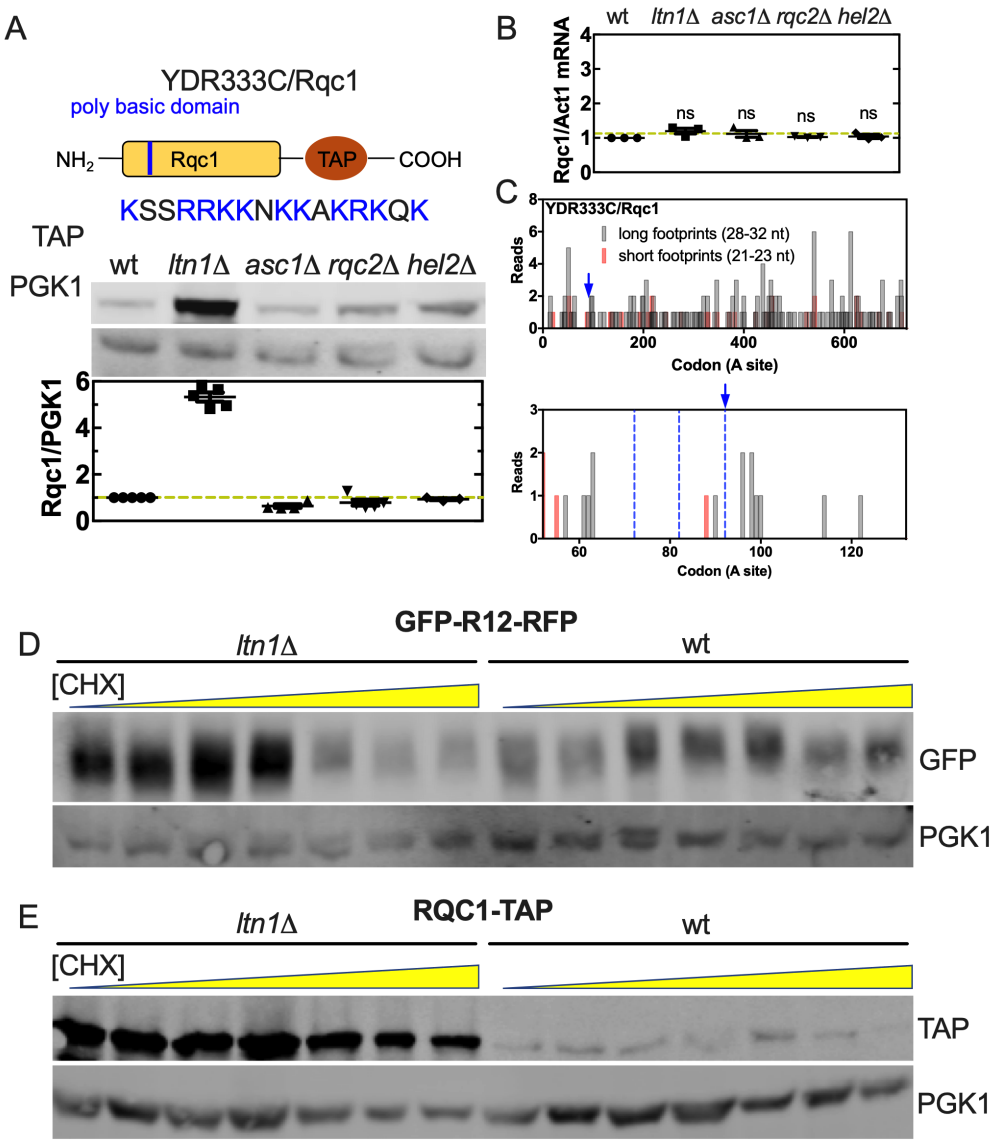

Figure 5. Rcq1 is not co-translationally degraded by the RQC complex. (A) A TAP-tagged version of Rqc1 was used to measure the protein levels in the wt, *lnt1*Δ, *asc1*Δ, *rqc2*Δ, and *hel2*Δ strains. The band intensity was quantified and normalized in relation to the housekeeping gene phosphoglycerate kinase (*PGK1*). (B) qPCR analysis of Rqc1-TAP mRNA levels in the wt, *lnt1*Δ, *asc1*Δ, *rqc2*Δ, and *hel2*Δ strains. The Rqc1 levels were normalized in relation to the housekeeping gene actin (Act1). The analysis of differential expression was made by relative quantification using $2^{-\Delta\Delta Ct}$ method. One-way ANOVA with Bonferroni's multiple comparison test was used for statistical analyses. (C) 28-32 and 21-23 nucleotide footprint analysis of Rqc1 shows neither accumulation of short reads at its poly-basic site nor accumulation of long reads upstream of the stalling site (left and right panels, respectively). Blue arrow indicates the start of the poly-basic domain, and the blue dotted line indicates the hypothetical position of the collided ribosomes. The reads were plotted at an approximate position of the ribosomal A site. (D) Levels of the poly-basic reporter GFP-R12-RFP subjected to 10 hours cycloheximide (CHX) treatment at increasing concentrations. (E) The same experiment as panel D was performed with an Rqc1-TAP strain.