RESEARCH ARTICLE

Inverted translational control of eukaryotic gene expression by ribosome collisions

Heungwon Park, Arvind R. Subramaniam [ID] *

Basic Sciences Division and Computational Biology Section of Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, United States of America

* rasi@fredhutch.org

Abstract

The canonical model of eukaryotic translation posits that efficient translation initiation increases protein expression and mRNA stability. Contrary to this model, we find that increasing initiation rate can decrease both protein expression and stability of certain mRNAs in the budding yeast *Saccharomyces cerevisiae*. These mRNAs encode a stretch of polybasic residues that cause ribosome stalling. Our computational modeling predicts that the observed decrease in gene expression at high initiation rates occurs when ribosome collisions at stalls stimulate abortive termination of the leading ribosome or cause endonucleolytic mRNA cleavage. Consistent with this prediction, the collision-associated quality-control factors Asc1 and Hel2 (orthologs of human RACK1 and ZNF598, respectively) decrease gene expression from stall-containing mRNAs only at high initiation rates. Remarkably, hundreds of *S. cerevisiae* mRNAs that contain ribosome stall sequences also exhibit lower translation efficiency. We propose that inefficient translation initiation allows these stall-containing endogenous mRNAs to escape collision-stimulated reduction in gene expression.

Introduction

Translation of mRNAs by ribosomes is a major point for regulation of eukaryotic gene expression. Initiation is usually the slowest kinetic step in the translation of eukaryotic mRNAs [1]. As a result, increasing the initiation rate of mRNAs typically results in higher expression of the encoded protein. Higher initiation rates can also protect eukaryotic mRNAs from decay. This protection can occur through preferential binding of translation initiation factors over mRNA decay factors to the 5′ cap of mRNAs [2–4]. Such mRNA stabilization amplifies the positive effect of high initiation rate on protein expression [5, 6]. Thus, efficient initiation is widely associated with increased mRNA stability and higher protein expression in eukaryotes [7–9].

Efficient initiation is also required for quality control at ribosomes that slow down or stall during elongation [10, 11]. Slowing down of ribosomes at nonoptimal codons accelerates mRNA decay [12–14]. The no-go mRNA decay (NGD) and the ribosome-associated quality-control (RQC) pathways target ribosomes stalled at polybasic stretches, rare codon repeats, or mRNA stem loops [15, 16]. These pathways together mediate cleavage of mRNAs, degradation

Abbreviations: a.u., arbitrary unit; CAT, collide and abortive termination; CSAT, collision-stimulated abortive termination; CSEC, collision-stimulated endonucleolytic cleavage; NGD, no-go mRNA decay; RNA-seq, RNA sequencing; RQC, ribosome-associated quality control; SAT, simple abortive termination; SEC, simple endonucleolytic cleavage; TE, translation efficiency; TJ, traffic jam; WT, wild-type; YFP, yellow fluorescent protein.

of nascent peptides, and recycling of ribosomes and tRNAs [17–25]. Mutations that destabilize mRNAs or block translation initiation also suppress the above quality-control events [13, 15, 26]. So how do eukaryotic cells balance normal translation and quality control to set the overall stability and protein expression of stall-containing mRNAs?

Computational modeling of translation can be useful for understanding the effect of ribosome stalling on protein expression [27, 28]. The widely used traffic jam (TJ) model predicts higher protein expression as initiation rate is increased until elongation at the stall becomes rate-limiting [29, 30]. In the elongation-limited regime of the TJ model, queues of collided ribosomes prevent further increase in protein expression [31]. Constant protein expression as a function of initiation rate has been observed for stall-encoding mRNAs in bacteria [32]. However, ribosome density measurements do not reveal long queues of collided ribosomes associated with the TJ model [33, 34]. Instead, our previous measurements of protein expression in bacteria support a kinetic model in which ribosome collisions stimulate abortive termination at elongation stall sites [35]. Recent work found that ribosome collisions in the budding yeast *S. cerevisiae* are also associated with increased endonucleolytic mRNA cleavage at elongation stall sites [26]. Indeed, biochemical and structural studies have revealed that collided ribosomes in eukaryotes have a unique structural conformation that might facilitate the binding and action of quality-control factors [36, 37]. These studies led us to investigate whether computational modeling of quality-control mechanisms at collided ribosomes can reveal new modes of regulating eukaryotic gene expression.

To constrain our kinetic models of eukaryotic quality control, we first experimentally characterized the effect of initiation rate on protein expression and stability of both normal and stall-containing mRNAs in the budding yeast *S. cerevisiae*. By designing 5′ UTR mutations that decouple translation initiation rate from normal mRNA stability, we find that high initiation rates can reduce gene expression of stall-containing mRNAs. To our knowledge, this inverse relation has not been previously observed in experiments. By computationally simulating different kinetic models of quality control, we predict that high initiation rates decrease protein expression and mRNA stability when ribosome collisions cause the leading ribosome to abort translation or cleave the mRNA. We tested our prediction by measuring the effect of various known quality-control factors on protein expression and mRNA stability. We find that Hel2 and Asc1, the only two identified factors associated with ribosome collisions [26, 36, 37], are also required for the decrease in gene expression at high initiation rates. Our analysis reveals that hundreds of endogenous *S. cerevisiae* mRNAs with ribosome stall sequences have unusually low translation efficiency (TE). We propose that inefficient translation initiation on these mRNAs allows them to escape from collision-stimulated reduction in gene expression. More generally, our work illustrates how computational modeling can be combined with quantitative experiments for understanding the in vivo kinetics of translational control mechanisms in eukaryotes.

Results

High initiation rates decrease protein output of stall-containing *S. cerevisiae* mRNAs

To measure the effects of initiation rate variation on protein expression, we used a fluorescent reporter system in the budding yeast *S. cerevisiae* (Fig 1A). Our reporters consist of the *PGK1* gene from *S. cerevisiae*, which has been extensively used in previous studies of mRNA translation and stability [5, 13]. We fused the *PGK1* coding sequence to a 3×FLAG-encoding sequence at the 5′ end and the yellow fluorescent protein (*YFP*) gene at the 3′ end. We integrated this 2-kb protein-coding region, under the control of the *GPD* promoter, the *GPD* 5′
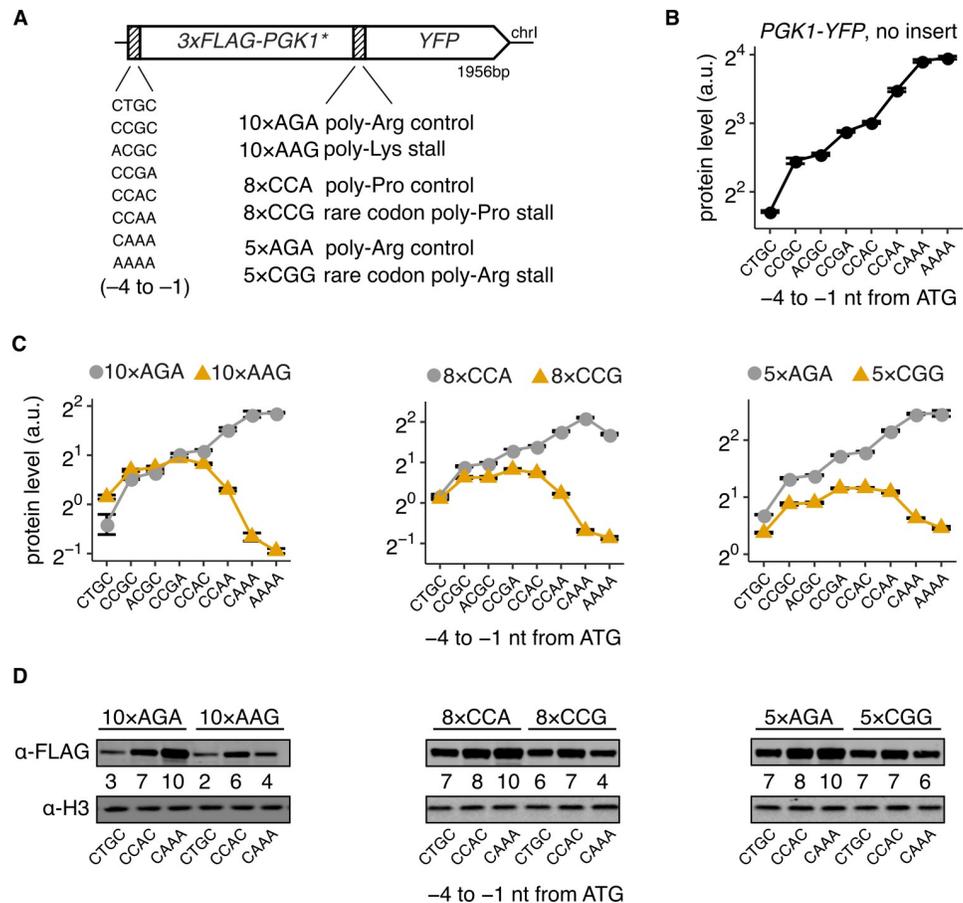
Fig 1. High initiation rates decrease protein output of stall-containing *S. cerevisiae* mRNAs. (A) Schematic of *3×FLAG-PGK1*-YFP* reporters used in (C) and (D). The −4 to −1 nt region preceding the ATG start codon had one of eight different Kozak sequences as indicated. One of three different stall sequences or their respective controls were inserted at the end of *PGK1**. *PGK1** was modified from wild-type *PGK1* sequence based on a previous study [11] (see Strain and plasmid construction). (B) Protein levels of *3×FLAG-PGK1-YFP* control reporters with no insert and wild-type *PGK1* sequence. (C) Protein levels of the constructs shown in (A). Protein levels are quantified as the mean YFP fluorescence of 10,000 cells for each strain as measured using flow cytometry. Protein levels are expressed as a.u. relative to the mean RFP levels from a constitutively expressed mKate2 control. Error bars in (B) and (C) show standard error of the mean over three or four independent transformants. Many error bars are smaller than data markers. Two of the total 192 strains were clear outliers and were removed after manual inspection. (D) Western blots of reporters with low (CTGC), medium (CCAC), or high (CAAA) initiation rates and with indicated stall sequences or controls using antibody against the FLAG epitope at the N terminus. Histone H3 level is shown as loading control. Numbers for each lane indicate the ratio of the FLAG signal to the H3 signal and are normalized to a maximum of 10 within each blot. The underlying data for panels B and C can be found at https://github.com/rasilab/ribosome_collisions_yeast. a.u., arbitrary unit; chrI, chromosome I; nt, nucleotide; YFP, yellow fluorescent protein.

https://doi.org/10.1371/journal.pbio.3000396.g001

UTR, and the *CYC1* 3′ UTR, as a single copy into the genome (see Strain and plasmid construction). We expressed a red fluorescent protein (mKate2) gene with identical regulatory sequences as our reporter from a different genomic locus for normalizing the measured YFP fluorescence.

To confirm our reporter system's utility for studying both normal and stall-containing mRNAs, we introduced five tandem arginine AGA codons (5×AGA) at five different locations along the *PGK1* gene (henceforth referred to as *PGK1**, S1 Fig panel A), following earlier work [11]. We individually mutated each of the five 5×AGA repeats to their synonymous CGG rare arginine codons and measured their effect on YFP expression using flow cytometry. The

5×CGG rare codons reduce YFP expression by 40%–70%, with the stalls away from the ATG start codon having a stronger effect (S1 Fig panel B).

To vary the initiation rate of our reporters, we designed eight different 5′ UTR variants in which we mutated the 4 nucleotides preceding the ATG start codon. These mutations vary 3×FLAG-PGK1-YFP expression over a 5-fold range (Fig 1B). This hierarchy of YFP expression among the 5′ UTR variants is concordant with previous measurements [38] (S1 Fig panel C). Because the 5′ UTR mutations are located over 30 nucleotides away from the 5′ end of the mRNA, we expect them to have minimal effects on the binding of translation initiation factors and mRNA decay factors to the 5′ cap.

We then combined the 5′ UTR mutations with various stall and control sequences to measure their combinatorial effect on protein expression (Fig 1A). We used the 5×CGG stall sequence from above, as well as 8×CCG rare proline codon repeats and 10×AAG lysine codon repeats, which are known to trigger ribosome stalling [39–41]. As controls, we used 5×AGA, 8×CCA, and 10×AGA repeats, respectively.

Strikingly, the three 5′ UTR variants (CCAA, CAAA, AAAA) that have up to 2-fold-higher protein level than the median for normal mRNAs (Fig 1B) instead cause up to a 4-fold reduction in protein level when the mRNA contains any of the three ribosome stalls (orange triangles, Fig 1C). This decrease in protein expression at high initiation rates is largely absent when the stalls are replaced by control sequences of synonymous common codons (5×AGA and 8×CCA, Fig 1C).

Similar to our above results with rare arginine and proline codon-mediated stalls, replacing the 10×AAG lysine repeat stall by a 10×AGA arginine control abrogates the decrease in protein expression at high initiation rates (Fig 1C, left panel). Importantly, the 10×AAG repeat differs from the 10×AGA control only by a single-nucleotide frameshift, and it is located over 1.3 kb away from the initiation codon. Hence, changes in mRNA secondary structure or other subtle physical interactions between the 5′ UTR mutations and the stall mutations are unlikely to account for the dramatically different behaviors of the 10×AAG and 10×AGA inserts at high initiation rates. Furthermore, because AAG is a common lysine codon with abundant cellular tRNA [42, 43] and our reporter is integrated as a single copy in the *S. cerevisiae* genome, the observed effects on protein expression at high initiation rates are unlikely to arise from titration of rare tRNAs by the stall-encoding repeats.

Western blotting against the N terminus 3×FLAG epitope qualitatively reproduced the differences in protein expression as measured by flow cytometry (Fig 1D). These results confirm that high initiation rates have distinct effects on expression of the full-length protein depending on the presence of ribosome stall sequences in the corresponding mRNA.

High initiation rates decrease lifetime of stall-containing *S. cerevisiae* mRNAs

We reasoned that the change in steady-state protein expression with initiation rate in our experiments could arise from change in stability of mRNAs, change in synthesis rate of the full-length protein from each mRNA, or a combination of both. To isolate the effect of initiation rate variation on mRNA stability, we developed a high-throughput sequencing–based assay to quantify mRNA levels of our *PGK1-YFP* reporters at high resolution (Fig 2A). In our assay, each combination of 5′ UTR mutation and codon repeat is tagged with four distinct 8-nucleotide barcodes in their 3′ UTR. We pooled these barcoded reporters and integrated them into the genome of *S. cerevisiae*. Reporter mRNA from the pooled strains was reverse transcribed to cDNA. The barcode region from both cDNA and genomic DNA was then amplified by PCR and counted by sequencing. The relative mRNA level of each reporter
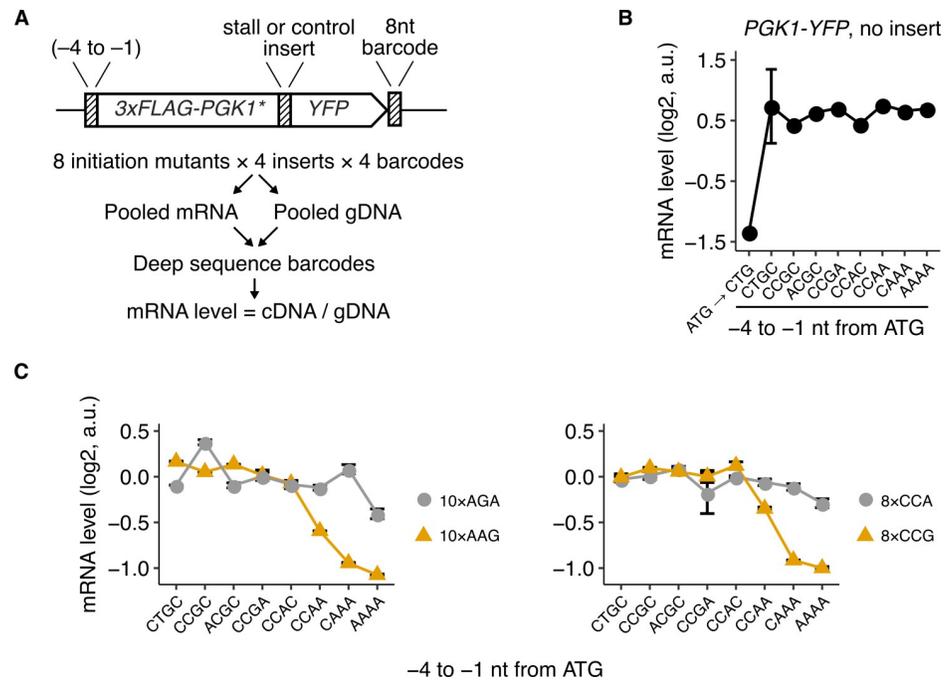
Fig 2. High initiation rates decrease stability of stall-containing *S. cerevisiae* mRNAs. (A) Schematic of deep-sequencing assay used for measuring mRNA levels. Reporters with different 5′ UTR mutations and stall or control inserts were tagged with four unique 8-nt barcodes in their 3′ UTR. The reporters were pooled and transformed into *S. cerevisiae*. cDNA and gDNA from the pooled reporters were amplified, and the barcodes were counted by high-throughput sequencing. Barcodes that had fewer than 100 counts for either cDNA or gDNA were discarded during analysis. mRNA level was calculated as cDNA counts normalized by gDNA counts for each barcode and was median normalized across all reporters in each pool. (B) mRNA levels of *3×FLAG-PGK1-YFP* control reporters with no insert and wild-type *PGK1* sequence. The leftmost point is for a reporter with the ATG start codon mutated to CTG. (C) mRNA levels of *3×FLAG-PGK1*-YFP* constructs with stall or control inserts. Error bars in (B) and (C) show standard error over 3 or 4 distinct barcodes for each reporter variant. One reporter variant (10×AGA, ACGC) had only two barcodes with read counts higher than the threshold for discarding. Most error bars are smaller than data markers. The underlying data for panels B and C can be found at https://github.com/rasilab/ribosome_collisions_yeast. a.u., arbitrary unit; gDNA, genomic DNA; nt, nucleotide; YFP, yellow fluorescent protein.

https://doi.org/10.1371/journal.pbio.3000396.g002

variant was obtained by normalizing the cDNA count of the corresponding barcode by its genomic DNA count.

We tested our experimental strategy by measuring steady-state mRNA levels of the *3×FLAG-PGK1-YFP* reporters with no ribosome stall sequences and eight different 5′ UTR mutations. mRNA levels differ less than 25% between the eight 5′ UTR variants (Fig 2B), even though protein levels differ more than 5-fold between these variants (Fig 1B). Because steady-state mRNA levels reflect the balance between mRNA synthesis and decay, we parsimoniously interpret these measurements to imply that our 5′ UTR mutations change only the translation initiation rate at the ATG start codon of our reporters without affecting their mRNA stability or transcription rate. As a positive control, mutating the ATG start codon of our reporter to CTG decreases mRNA levels by 4-fold (Fig 2B). We expect this decrease if absence of the main ATG start codon results in translation initiation at one of the four downstream out-of-frame start codons in the 3×FLAG-encoding region, which then destabilizes the mRNA by non-sense-mediated mRNA decay.

We then applied our experimental strategy to measure mRNA levels of the 5′ UTR variants of our *3×FLAG-PGK1*-YFP* reporters with various stall or control inserts (Fig 2C). The 8×CCG and 10×AAG stall sequences result in up to a 2-fold decrease in mRNA levels,

specifically in the high initiation rate regime (yellow triangles, Fig 2C). By contrast, 3×*FLAG-PGK1*-YFP* reporters with the 8×CCA and 10×AGA control inserts show little or no decrease in mRNA levels at high initiation rates (gray circles, Fig 2C), which is similar to the 3×*FLAG-PGK1-YFP* constructs without inserts (Fig 2B). We interpret the decrease in mRNA levels of the 8×CCG and 10×AAG stall-containing reporters at high initiation rates as changes in their mRNA lifetime. This interpretation is justified given that the 5′ UTR mutations, on their own, do not affect mRNA levels (Fig 2B).

Comparison between the protein and mRNA measurements from our *PGK1-YFP* reporters show that protein levels at high initiation rate decrease up to 4-fold from their peak value, whereas mRNA levels decrease only 2-fold (CCGA versus AAAA 5′ UTR mutations in Figs 1C and 2C). Thus, we conclude that the decrease in protein expression at high initiation rates arises from both steady-state change in mRNA levels and change in the synthesis rate of the full-length protein from each mRNA.

Collision-stimulated abortive termination model predicts reduced protein output at high initiation rates

We hypothesized that computational modeling can provide mechanistic insight into how high initiation can result in lower protein expression and reduced stability of stall-containing eukaryotic mRNAs. Because kinetic models that consider both translation and stability of eukaryotic mRNAs have not been formulated so far to our knowledge, we first defined a joint model for normal translation and canonical mRNA decay in eukaryotes (see Kinetic modeling). We modeled normal translation as a three-step process composed of initiation, elongation, and termination, similar to previous studies [44, 45]. Each ribosome occupies a footprint of 10 codons on mRNAs [46], and initiation and elongation occur only if the required footprint is not blocked by another leading ribosome. We modeled canonical mRNA decay in eukaryotes as a three-step process of deadenylation, decapping, and exonucleolysis [47]. Even though a more detailed kinetic model of canonical mRNA decay is available [47], because we aimed to predict overall mRNA lifetime, we considered only the three main steps of canonical mRNA decay. We chose the rate constants for these decay steps [47] (S3 Table) to result in a mean mRNA lifetime of 35 min, corresponding to that of the *PGK1* mRNA in *S. cerevisiae* [5]. We used a rule-based modeling approach and performed exact stochastic simulations using an agent-based framework to predict mRNA lifetime and protein expression (see Kinetic modeling).

We then systematically investigated the effect of quality control at elongation stalls on protein expression. We first studied kinetic models of how stalled ribosomes undergo abortive termination by setting the mRNA decay rate to zero (S3 Table). In our model, abortive termination includes all molecular processes that terminate polypeptide synthesis including splitting of the stalled ribosome, degradation of the nascent peptide, and hydrolysis of the peptidyl-tRNA. We did not distinguish between these different molecular events because all of them reduce synthesis rate of the full-length protein, which is the quantity that we measure in our experiments.

We considered three kinetic models of how abortive termination might occur at stalled ribosomes from intact mRNAs (Fig 3A, Kinetic modeling). In the simple abortive termination (SAT) model, stalled ribosomes are subject to kinetic competition between abortive termination and normal elongation (Fig 3A). This SAT model has been considered in recent modeling studies [35, 48, 49]. Because recent experiments found a role for ribosome collisions in quality control [26, 36, 37], we then considered two possible effects of ribosome collisions on abortive termination. In the collision-stimulated abortive termination (CSAT) model, only stalled
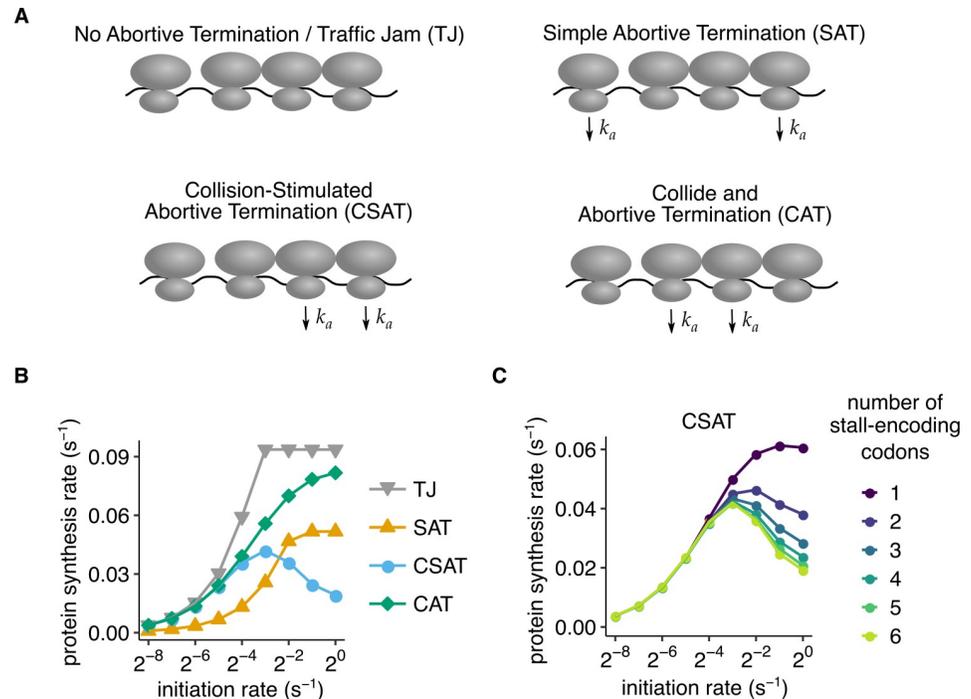
A

No Abortive Termination / Traffic Jam (TJ)

Simple Abortive Termination (SAT)

$\downarrow k_a$ $\downarrow k_a$

Collision-Stimulated
Abortive Termination (CSAT)

$\downarrow k_a$ $\downarrow k_a$

Collide and
Abortive Termination (CAT)

$\downarrow k_a$ $\downarrow k_a$

B

protein synthesis rate (s^{-1})

initiation rate (s^{-1})

- ▼ TJ
- ▲ SAT
- ● CSAT
- ◆ CAT

C

CSAT

protein synthesis rate (s^{-1})

initiation rate (s^{-1})

number of stall-encoding codons
- 1
- 2
- 3
- 4
- 5
- 6

Fig 3. Collision-stimulated abortive termination model predicts reduced protein output at high initiation rates.
(A) Schematic of different abortive termination models simulated in (B). k_a denotes the nonzero rate of abortive termination of ribosomes from indicated configurations. In the TJ model, ribosomes do not abort translation. In the SAT model, only ribosomes that have not undergone collision from the mRNA entry side abort. In the CAT model, only ribosomes that have undergone collision from the mRNA entry side ("trailing" ribosomes in a collision) abort. In the CSAT model, only ribosomes that have undergone collision from the mRNA exit side ("leading" ribosomes in a collision) abort. (B) Average protein synthesis rate as a function of initiation rate predicted using stochastic simulations of the four models in (A). The simulations were of a 650-codon mRNA corresponding to the 3×*FLAG-PGK1**-*YFP* reporters in our experiments. Ribosome stalls were simulated as a consecutive stretch of six slowly translated codons with a net elongation rate of 0.1 s^{-1}. All other codons had an elongation rate of 10 s^{-1}. The stalls were located after 400 codons, corresponding to their approximate location in our experiments. (C) Effect of varying the number of tandem (consecutive) slowly translated codons encoding the stall in the CSAT model. The elongation rate of each slowly translated codon was scaled to maintain a net elongation rate of 0.1 s^{-1} across the stall. mRNA decay rate was set to zero in these simulations to isolate the effect of abortive termination. All other model parameters are listed in S3 Table. Standard errors from repeating stochastic simulations with different initial random seeds are smaller than data markers in (B) and (C). The underlying data for panels B and C can be found at https://github.com/rasilab/ribosome_collisions_yeast. CAT, collide and abortive termination; CSAT, collision-stimulated abortive termination; SAT, simple abortive termination; TJ, traffic jam; YFP, yellow fluorescent protein.

ribosomes that get hit by trailing ribosomes undergo abortive termination (Fig 3A). Conversely, in the collide and abortive termination (CAT) model, only trailing ribosomes that run into leading stalled ribosomes undergo abortive termination (Fig 3A). Finally, as a control, we also considered the TJ model, in which ribosomes do not undergo abortive termination (Fig 3A).

We simulated the above-described kinetic models of abortive termination with varying initiation rate and analyzed the predicted protein synthesis rate (defined as the number of full proteins produced per second). We simulated elongation stalls by introducing a stretch of six tandem poorly translated codons after 400 codons in a 650-codon mRNA, which is similar in length to the *PGK1-YFP* reporters that we use in our experiments. In our simulations, the control TJ model without abortive termination exhibits a linear increase in protein synthesis rate with initiation rate until the elongation rate at the stall becomes rate limiting (gray triangles, Fig 3B; S2 Fig panel A). This observation is consistent with previous studies of this model [50,

51]. The SAT model exhibits a similar behavior to the TJ model, but with lower protein synthesis rate and a higher initiation rate at which protein synthesis rate reaches saturation (yellow triangles, Fig 3B, S2 Fig panel B, C).

Unexpectedly, the two kinetic models in which ribosome collisions cause abortive termination of either the trailing or the leading ribosome (Fig 3A) exhibit different behaviors as initiation rate is varied. In the CAT model in which the trailing ribosome abortively terminates upon collision, protein synthesis rate increases monotonically as the initiation rate is increased (green diamonds, Fig 3B). By contrast, in the CSAT model, in which the leading ribosome abortively terminates upon collision, protein synthesis rate increases initially, then reaches a maximum, and surprisingly decreases with further increase in initiation rate (blue circles, Fig 3B).

The above dichotomy between the CSAT and CAT models of abortive termination at ribosome collisions can be intuitively understood as follows: In the CAT model, even though the trailing ribosomes at collisions undergo abortive termination, every ribosome that reaches the stall eventually finishes protein synthesis at a rate determined by the elongation rate at the stall. In the CSAT model, the leading stalled ribosome is stimulated to undergo abortive termination upon collision; thus, at sufficiently high initiation rates and abortive termination rates, very few ribosomes will get past the stall. Stated differently, the difference between the CSAT and CAT models arises because all ribosomes that finish protein synthesis stall at the slow codons, whereas not all ribosomes collide with a leading ribosome before they finish protein synthesis.

The decrease in protein expression at high initiation rates in the CSAT model is counterintuitive, and to further probe its origin, we systematically varied the elongation rate at the stall and the number of tandem (consecutive) codons encoding the stall. The decrease in protein synthesis rate requires the initiation rate to exceed the total elongation rate past the stall (S2 Fig panel D) because ribosome collisions occur only in this regime. More surprisingly, the decrease in protein synthesis occurs only if ribosomes stall at multiple consecutive codons (Fig 3C). As a concrete example, the CSAT model predicts that two tandem stall codons with a 5 s ribosome dwell time at each codon decreases protein synthesis rate at high initiation rates, but a single stall codon with a 10 s ribosome dwell time does not. This nonintuitive prediction can be understood by observing that the number of kinetic partitions between normal elongation and abortive termination changes dynamically with initiation rate in the CSAT model with n consecutive stall codons. At low initiation rates ($k_{init} \ll k_{stall}$), there is no kinetic partitioning toward abortive termination, similar to the TJ model. At high initiation rates ($k_{init} \gg k_{stall}$), each ribosome that finishes protein synthesis undergoes n kinetic partitions toward normal elongation, similar to the SAT model. However, at intermediate initiation rates ($k_{init} \sim k_{stall}$), the average number of kinetic partitions increases from 0 to n as the initiation rate is increased. In this sense, the CSAT model "interpolates" between the TJ model and the SAT model.

Thus, our simulations of different kinetic models of abortive termination reveal distinct signatures of initiation rate variation on protein expression. The CSAT model, unlike the other models, predicts a decrease in protein expression at high initiation rate. Importantly, this decrease is not simply a consequence of ribosome collisions stimulating quality control, but it has two other essential ingredients in our simulations (Fig 3B and 3C): the leading ribosome in the collision undergoes abortive termination, and the stall itself is composed of multiple kinetic steps.

Collision-stimulated endonucleolytic cleavage model predicts decrease in mRNA lifetime at high initiation rates

Because endonucleolytic cleavage has been shown to reduce mRNA lifetime of stall-containing mRNAs [15], we next considered different kinetic models of how endonucleolytic mRNA
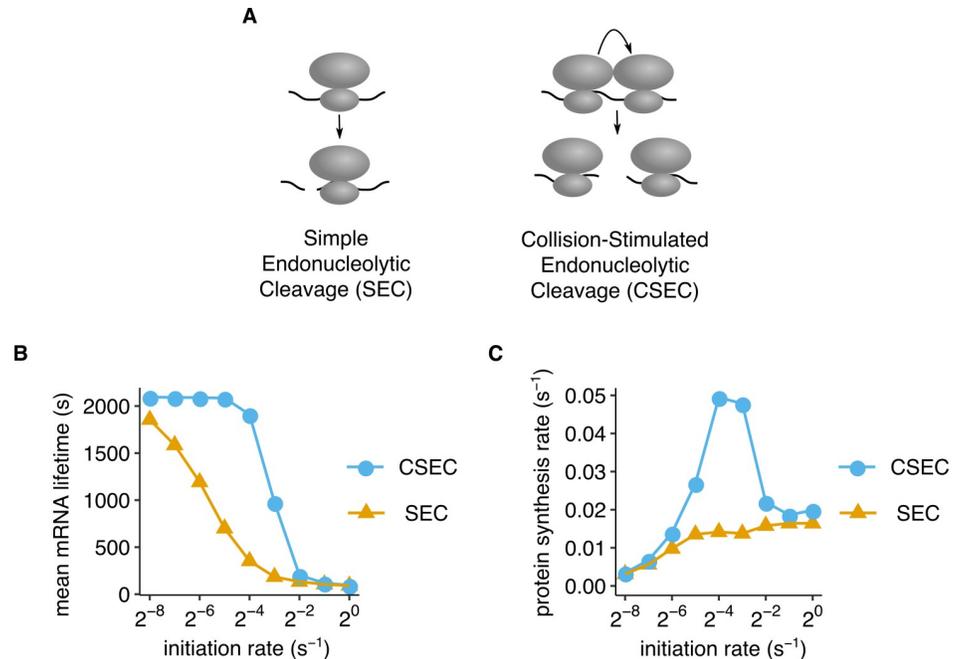
Fig 4. CSEC model predicts decrease in mRNA lifetime at high initiation rates. (A) Schematic of SEC and CSEC mRNA cleavage models simulated in (B) and (C). (B, C) Average mRNA lifetime and protein synthesis rate over 10^6 s as a function of initiation rate predicted using stochastic simulations of the models in (A). Reporters were simulated as in Fig 3, but with a nonzero endonucleolytic cleavage rate of 0.001 s^{-1} and a net elongation rate of 0.1 s^{-1} across the stall. Canonical mRNA decay was allowed, whereas abortive termination rate was set to zero. All other model parameters are listed in S3 Table. Standard errors from repeating stochastic simulations with different initial random seeds are smaller than data markers in (B) and (C). The underlying data for panels B and C can be found at https:// github.com/rasilab/ribosome_collisions_yeast. CSEC, collision-stimulated endonucleolytic cleavage; SEC, simple endonucleolytic cleavage.

https://doi.org/10.1371/journal.pbio.3000396.g004

cleavage at ribosome stalls decreases mRNA lifetime as a function of initiation rate (Fig 4A, Kinetic modeling). We modeled endonucleolytic mRNA cleavage as occurring on the 5′ side of the stalled ribosome [52, 53]. In our modeling, once mRNAs are endonucleolytically cleaved, translation initiation is immediately repressed through decapping, and ribosomes that have already initiated on the 5′ mRNA fragment are efficiently recycled once they reach the truncated 3′ end [19, 20]. We simulated ribosome stalling by introducing a stretch of six poorly translated codons after 400 codons in a 650-codon mRNA, which is similar to our *PGK1-YFP* reporters. To isolate the effect of endonucleolytic cleavage on our predictions, we set the abortive termination rate from intact mRNAs to be zero. We monitored the mean lifetime of mRNAs, which we define as the time interval between the end of transcription and the start of 5′–3′ exonucleolytic decay after decapping. As before, we also monitored protein synthesis rate as we varied the initiation rate of mRNAs in our simulations.

We considered two distinct kinetic models of how endonucleolytic mRNA cleavage might occur at ribosome stalls (Fig 4A, Kinetic modeling). In the simple endonucleolytic cleavage (SEC) model, cleavage occurs through simple kinetic competition with normal elongation [54]. In the collision-stimulated endonucleolytic cleavage (CSEC) model, cleavage occurs only between two collided ribosomes.

Both the simple and collision-stimulated models of endonucleolytic cleavage predict a decrease in mRNA lifetime at high initiation rates (Fig 4B). However, whereas the SEC model predicts a gradual decrease in mRNA lifetime (yellow triangles, Fig 4B), the CSEC model predicts a sharp decrease in mRNA lifetime as the initiation rate is increased (blue circles, Fig 4B).

The dependence of mRNA lifetime on initiation rate in the SEC model is surprising at first glance because there is no direct role for initiation rate in this model. However, even in the SEC model, the endonucleolytic cleavage probability of a given mRNA increases with the frequency of ribosome stalling, which in turn increases with the frequency of initiation. Thus, in both our kinetic models of endonucleolytic cleavage, mRNAs transition from being primarily degraded through the canonical decay pathway at low initiation rates to being endonucleolytically cleaved at high initiation rates.

Even though mRNA lifetimes decrease at high initiation rates in both the simple and collision-stimulated cleavage models, protein synthesis rates in these models display strikingly different behaviors as a function of initiation rate (Fig 4C). In the SEC model, protein synthesis rate increases monotonically with initiation rate (yellow triangles, Fig 4C): at low initiation rates, it increases linearly where mRNAs are degraded primarily through the canonical decay pathway, and it saturates at high initiation rates at a value determined by the endonucleolytic cleavage rate at stalls (S3 Fig panel A). In the high initiation rate regime of the SEC model, each mRNA is translated by a fixed number of ribosomes (on average) before it undergoes endonucleolytic cleavage. By contrast, in the CSEC model, protein synthesis rate exhibits a nonmonotonic behavior (blue circles, Fig 4C): it increases linearly until the initiation rate matches the elongation rate at the stall, at which point it decreases sharply and saturates at the same value as the SEC model. Unlike the CSAT model (blue circles, Fig 3C), the behavior of the CSEC model does not depend sensitively on the number of tandem slowly translated codons that the ribosome transits through at each stall (S3 Fig panel B).

Thus, our simulations of different kinetic models of endonucleolytic mRNA cleavage reveal distinct signatures of initiation rate variation on mRNA lifetime and protein expression. The CSEC model, unlike the SEC model, predicts the decrease in mRNA stability observed in our experiments only at high initiation rates (Fig 2C). The CSEC model also correctly predicts the observed decrease in protein expression at high initiation rates (Fig 1C), which is independent of the predicted effect of abortive termination on protein expression in the CSAT model.

Hel2 and Asc1 attenuate gene expression from stall-containing mRNAs only at high initiation rates

We then sought to test the prediction from our modeling that collision-stimulated quality control (CSEC and CSAT) directly contributes to the decrease in protein expression and mRNA stability at high initiation rates from stall-containing mRNAs. Toward this, we focused on several factors that have been previously implicated in quality control at ribosome stalls but whose effect on gene expression as a function of initiation rate has not been characterized. The ribosome-ubiquitin E3 ligase Hel2 (ZNF598 in humans) is activated by collided ribosomes [26, 36, 37]. The ribosome-associated protein Asc1 (RACK1 in humans) is located at the interface between collided ribosomes [36, 37] and, along with Hel2, couples translational arrest to nascent chain degradation [55–58]. We also consider the ribosome recycling factor Dom34 (PELO in humans), which plays a critical role in no-go decay [15, 54, 59]. Nascent chain degradation is mediated by the E3 ligase Ltn1 (LTN1 in humans), which is part of the RQC complex [16, 18, 21, 22, 60]. We measured protein expression and mRNA levels from our 3×*FLAG-PGK1**-*YFP* reporters in *S. cerevisiae* strains in which *HEL2*, *ASC1*, *DOM34*, and *LTN1* were individually deleted (Fig 5A–5C).

Deletion of *ASC1* completely rescues protein expression from the 8×CCG stall-containing reporters (Δ*ASC1* in Fig 5A). In fact, the 8×CCG reporter has slightly higher expression than the 8×CCA control reporter at high initiation rate (bottom versus top panel, Δ*ASC1* in Fig 5A). Similarly, deletion of *HEL2* increases protein expression of stall-containing reporters by
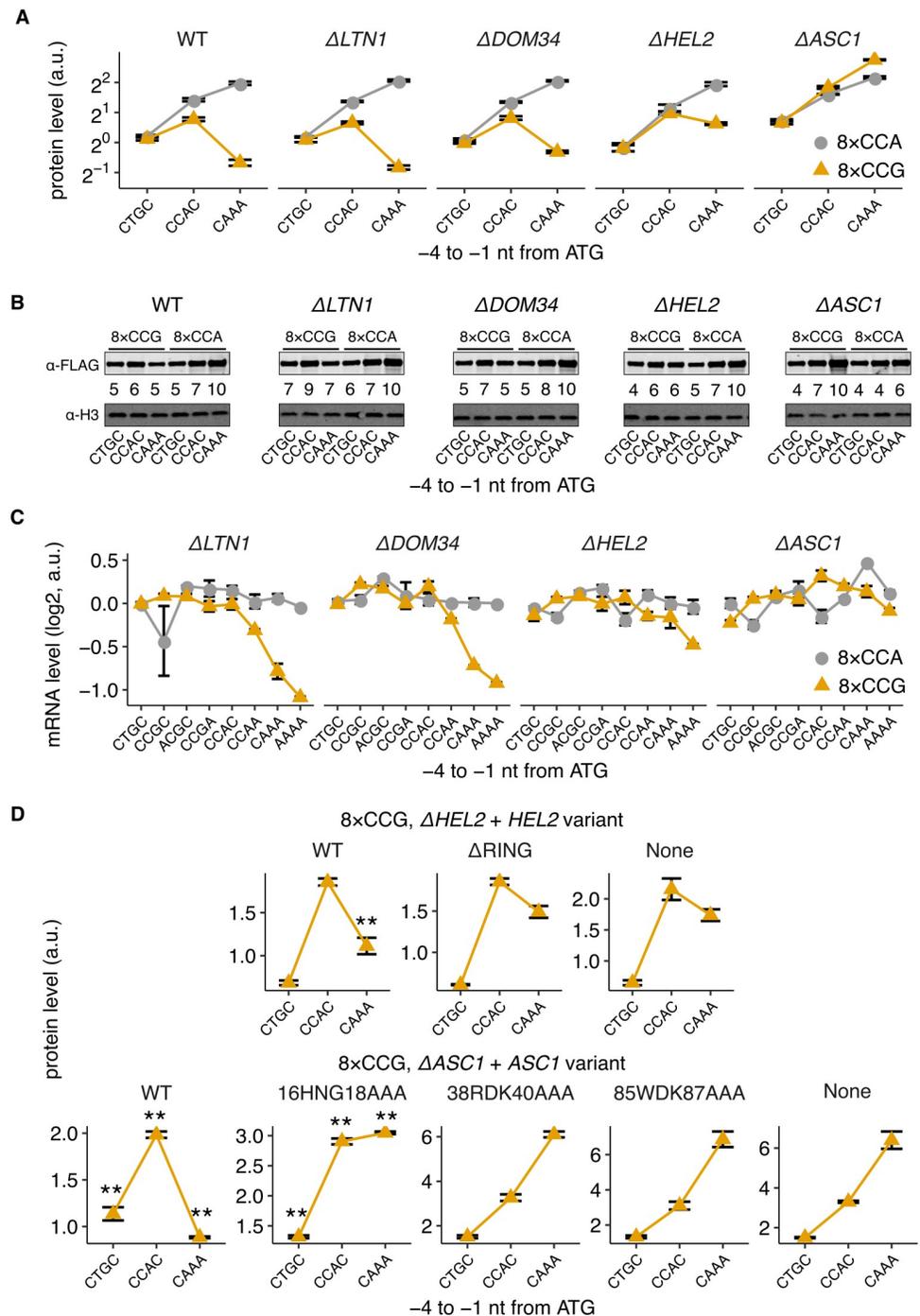
Fig 5. Hel2 and Asc1 attenuate gene expression from stall-containing mRNAs only at high initiation rates. (A) Protein levels of 3×*FLAG-PGK1**-*YFP* reporters (see Fig 1A) with low (CTGC), medium (CCAC), or high (CAAA) initiation rates and with stall (8×CCG) or control (8×CCA) repeats in different deletion backgrounds. (B) Western blots of reporters from (A) performed and quantified as in Fig 1. (C) mRNA levels of 3×*FLAG-PGK1**-*YFP* reporters. Mean mRNA levels and their standard errors quantified as in Fig 2. The Δ*LTN1*-8×CCA-AAAA variant has only two barcodes. (D) Protein levels of the 8×CCG stall reporter expressed in Δ*HEL2* (top) or Δ*ASC1* (bottom) strain and complemented with the indicated Hel2 or Asc1 variant. Error bars in (A) and (D) calculated as in Fig 1. ** in (D) denotes $P < 0.01$ (two-tailed t test) for comparison between indicated Hel2/Asc1 variant and the parent strain ("None") for the same 5′ UTR sequence. The underlying data for panels A, C, and D can be found at https://github.com/rasilab/ribosome_collisions_yeast. a.u., arbitrary unit; nt, nucleotide; WT, wild-type; YFP, yellow fluorescent protein.

https://doi.org/10.1371/journal.pbio.3000396.g005

3-fold at high initiation rate compared with the wild-type (WT) strains but has no effect at low initiation rate or in the control reporters (ΔHEL2 in Fig 5A). Deletion of either *DOM34* or *LTN1* has little to no effect on protein expression from either the stall-containing or control reporters (ΔDOM34 and ΔLTN1 in Fig 5A). However, deletion of *LTN1*, the E3 ligase necessary for nascent chain degradation, stabilizes the peptide resulting from abortive termination at the 8×CCG stalls only at high initiation rates (S4 Fig panel C). Western blotting against the N terminus 3×FLAG epitope qualitatively matches the above measurements of protein expression using flow cytometry (Fig 5B). The rescue of protein expression at high initiation rates in the *ASC1* and *HEL2* deletion strains is also observed with the weaker 5×CGG stalls (S4 Fig panel A).

We then used our pooled sequencing–based strategy to quantify the mRNA levels of our reporters (Fig 2A) in different deletion backgrounds. Deletion of either *ASC1* or *HEL2* selectively rescues the decreased mRNA levels of the 8×CCG stall-containing reporters at high initiation rates (ΔASC1 and ΔHEL2 in Fig 5C). By contrast, deletion of neither *DOM34* nor *LTN1* has any effect on the decreased mRNA levels at high initiation rates (ΔDOM34 and ΔLTN1 in Fig 5C).

We checked whether the observed rescue of protein expression in the *ASC1* and *HEL2* deletion strains can be reversed by constitutive expression of the respective proteins in *trans* from the *HIS3* locus (Fig 5D). Complementing by the WT Asc1 protein fully reverses the protein expression rescue at high initiation rates from 8×CCG stall-containing mRNAs in ΔASC1 strains (WT, Fig 5D, lower panel). This reversal is abrogated either partially or completely with Asc1 mutants that are defective in translational arrest [61] (16HNG18AAA, 38RDK40AAA, and 85WDK87AAA, Fig 5D, lower panel). Similarly, complementing by the WT Hel2 protein partially reverses the protein expression rescue at high initiation rates from stall-containing mRNAs in ΔHEL2 strains (WT, Fig 5D, upper panel). This reversal is not observed with a Hel2 mutant that cannot bind the interacting E2 enzyme Ubc4 [58] (ΔRING, Fig 5D, upper panel). In contrast to the 8×CCG stall reporters, complementing with various Asc1 and Hel2 mutants had little effect on protein expression from the 8×CCA-containing control reporters (S4 Fig panel B).

Based on the above measurements, we conclude that Asc1 and Hel2, which have been previously associated with ribosome collision–stimulated quality control [26, 36, 37], are necessary to reduce protein expression and mRNA stability at high initiation rates from stall-containing mRNAs. By contrast, neither Dom34 nor Ltn1 have a role in regulating gene expression at high initiation rates.

Endogenous mRNAs with stall sequences show signatures of inefficient translation initiation

Stall-inducing polybasic tracts decrease protein expression and mRNA stability at high initiation rates in our reporter-based experiments; do they also shape the translation of endogenous *S. cerevisiae* mRNAs? *S. cerevisiae* coding sequences that contain stretches of 6 or more lysine and arginine codons within a 10-codon window lead to ribosome stalling as measured using ribosome profiling [16]. Similarly, tandem repeats of 2 or more prolines also induce ribosome pausing [39, 40]. Over 1,250 *S. cerevisiae* protein-coding sequences contain either 10-codon stretches with at least 6 lysine and arginine codons or 10-codon stretches with at least 6 proline codons. Using a high-quality *S. cerevisiae* ribosome profiling dataset [62], we recapitulated previous observations [16] of increased ribosome density about 8 codons into the stall (Fig 6A). We did not observe a similar increase around control sequences that are enriched for glutamate or aspartate codons (S5 Fig panel A).
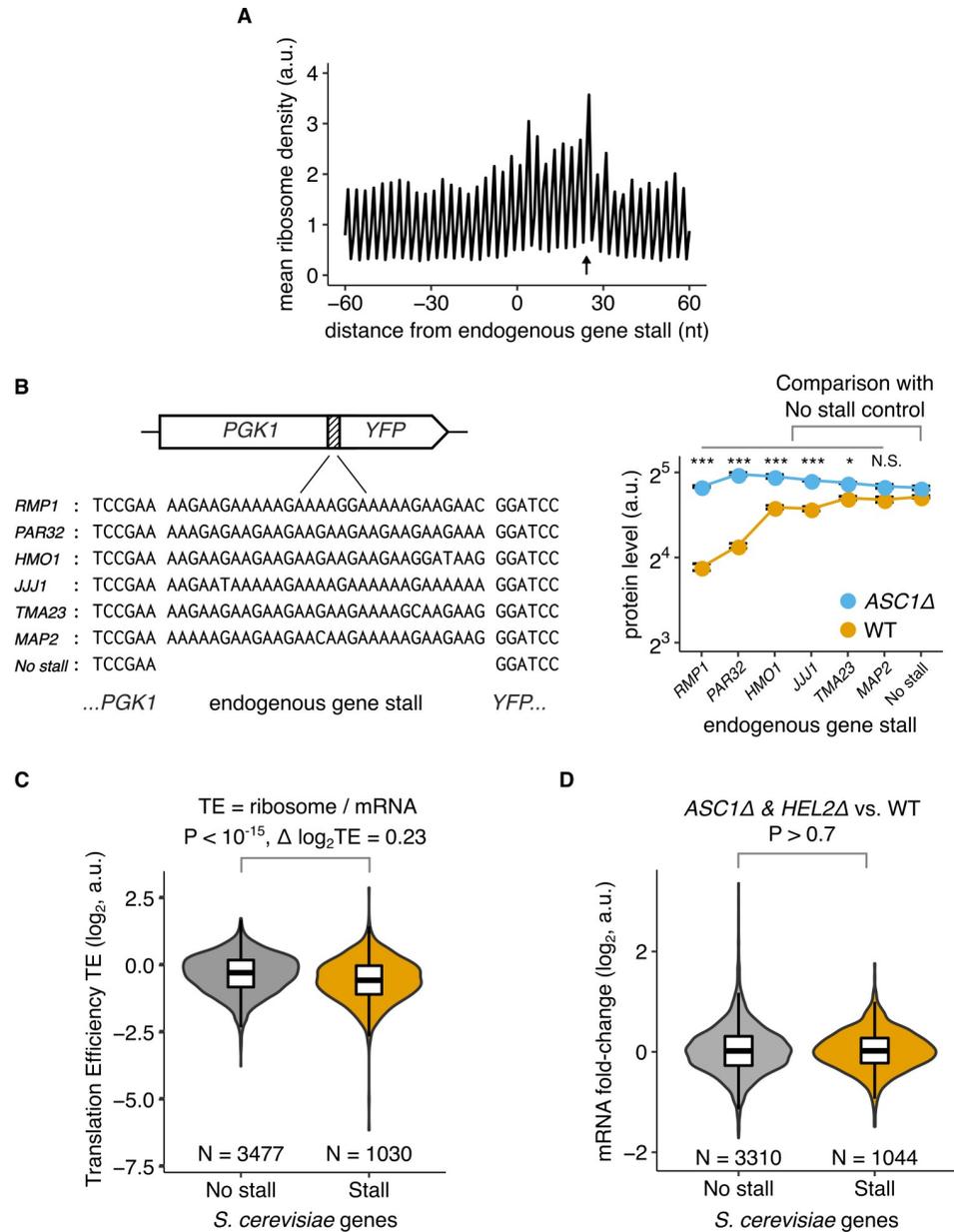
Fig 6. Endogenous mRNAs with stall sequences show signatures of inefficient translation initiation. (A) Mean ribosome density around stall-encoding sequences in *S. cerevisiae* mRNAs using data from Weinberg and colleagues [62]. Stall-encoding sequences are defined as 10-codon windows that have either a minimum of 6 lysine and arginine codons or a minimum of 6 proline codons. The first nucleotide of the 10-codon window is at distance 1 nt. In total, 1,251 *S. cerevisiae* mRNAs have at least one stall-encoding sequence. Arrow indicates peak in ribosome density at +24 nt that is consistent with Brandman and colleagues [16]. (B) Effect of endogenous stalls on protein expression from reporter mRNAs. Protein levels measured by flow cytometry and their mean and standard error quantified as in [Fig 1]. ***, *, N.S. indicate $P < 0.001$, $P < 0.05$, $P > 0.05$ (one-tailed *t* test) for higher $\Delta ASC1$/WT protein-level ratio of the stall-encoding reporters in comparison with the no-stall control. (C) TE of *S. cerevisiae* mRNAs that either contain or do not contain stall-encoding sequences. TE is defined as the normalized ratio of ribosome footprint counts to total mRNA counts as measured by Weinberg and colleagues [62]. (D) Fold change in mRNA levels [57] between both $\Delta HEL2$ and $\Delta ASC1$ strains and WT strain. (C, D) Box plots shows mean and standard deviation within each gene group; violin plot shows the gene density at each y-axis value; P values calculated using the two-sided Wilcoxon rank-sum test. The underlying data for panels A, B, C, and D can be found at https://github.com/rasilab/ribosome_collisions_yeast. a.u., arbitrary unit; nt, nucleotide; N.S., not significant; TE, translation efficiency; WT, wild-type.

https://doi.org/10.1371/journal.pbio.3000396.g006

To test whether ribosome stall sequences in endogenous genes can cause collision-stimulated decrease in gene expression, we cloned the 10-codon stall-encoding sequences from six endogenous mRNAs into our *PGK1-YFP* reporter system with an optimal initiation context (Fig 6B, left panel). Two of the six tested endogenous stall sequences have a significantly lower protein expression compared with a no-stall control (Fig 6B, right panel, yellow circles). However, this comparison does not account for the different amino acid sequences among the stall-containing reporters, which could alter the measured fluorescence. Because Asc1 is required for collision-stimulated decrease in gene expression (Fig 5A), we expect that protein expression from stall-containing reporters should increase upon deletion of *ASC1*. Consistent with this expectation, five of the six tested endogenous stall sequences have a significantly higher ratio of protein expression between the Δ*ASC1* and WT strains when compared to the no-stall control (Fig 6B, right panel, yellow versus blue circles).

The widespread nature of stall-encoding sequences in endogenous *S. cerevisiae* mRNAs suggests two distinct possibilities in light of the collision-stimulated decrease in gene expression uncovered in our work. First, an endogenous mRNA containing a stall could have a sufficiently high initiation rate such that it is constantly subject to collision-stimulated quality control. Indeed such a regulatory mechanism was proposed for *RQC1*, which encodes one of the components of the RQC complex [16]. Alternatively, a stall might be present because of functional requirements on the protein that are unrelated to collision-stimulated quality control. In this case, the presence of the stall might necessitate that the corresponding mRNA have sufficiently low initiation rate to prevent collision-stimulated decrease in mRNA stability and protein expression.

To distinguish between the above two possibilities, we tested whether stall-containing mRNAs have significant differences in initiation rate compared with other endogenous mRNAs in *S. cerevisiae*. We used TE as quantified by the ratio of ribosome footprints to mRNA abundance as a proxy for initiation rate [62]. We found that the TE of *S. cerevisiae* mRNAs containing stall-encoding sequences (as defined above, $N = 1,030$) is significantly lower than that of mRNAs ($N = 3,477$) that do not contain such sequences ($P < 10^{-15}$, Wilcoxon 2-sided test, $\Delta\log_2$ TE = 0.23, Fig 6C). One potential confounding factor in this analysis is that abortive termination at stalls can selectively decrease the apparent TE of stall-containing mRNAs. However, we observed a similar difference in TE between stall-containing and control mRNAs, even when TE was calculated using only mRNA regions 5′ to stalls (S5 Fig panel B). Another possible confounding factor is that the stall-encoding sequences cause an increase in local ribosome occupancy on the corresponding mRNA, thus possibly affecting the measured TE. However, this increase will result in a higher apparent TE, thereby lowering the measured difference between stall-containing and control mRNAs.

Based on our observation that stall-encoding *S. cerevisiae* mRNAs have lower TE, we hypothesize that inefficient translation initiation on these mRNAs allows them to avoid collision-stimulated reduction in protein expression and mRNA stability. This hypothesis predicts that deletion of *HEL2* or *ASC1* should have no effect on the levels of these endogenous mRNAs, unlike the increase in levels of our reporter mRNAs at high initiation rates (Fig 5C). To verify this prediction, we compared the levels of *S. cerevisiae* mRNAs between Δ*HEL2* or Δ*ASC1* strains and WT strains, as measured previously by RNA sequencing (RNA-seq) [57]. Because deletion of these factors can have indirect and pleiotropic effects unrelated to quality control [63], we looked for consistent differential regulation of mRNA levels in both Δ*HEL2* and Δ*ASC1* strains compared with the WT strain. We find that, on average, stall-containing *S. cerevisiae* mRNAs do not show a significantly higher up-regulation compared with control mRNAs in the Δ*HEL2* and Δ*ASC1* strains ($P > 0.7$, Wilcoxon 2-sided test, Fig 6D). This observation suggests that stall-containing *S. cerevisiae* mRNAs have been evolutionarily selected to

avoid collision-stimulated reduction in gene expression through a decrease in their translation initiation rates. An alternate possibility is that the stalls are not sufficiently strong to trigger quality control on many of the endogenous mRNAs. The latter possibility, however, does not explain why these group of stall-containing endogenous mRNAs should have significantly lower TE.

Discussion

In this work, we find that efficient translation initiation can decrease the protein expression and stability of certain eukaryotic mRNAs that undergo elongation stalls. To our knowledge, this is the first demonstration of an inverse relation between initiation rate and gene expression from mRNAs in *S. cerevisiae*. Although this observation is seemingly at odds with previous studies on normal eukaryotic mRNAs [2–4], our computational modeling suggests that it arises from the interplay between normal translation and collision-stimulated quality control at elongation stalls.

Recent studies have implicated ribosome collisions in triggering endonucleolytic mRNA cleavage and ribosomal ubiquitination in eukaryotes [26, 36, 37]. However, these studies did not delineate the relative effects of ribosome collisions and normal translation on overall mRNA stability and protein expression. Our quantitative experiments reveal that ribosome collisions decrease mRNA stability and protein expression at high initiation rates. As discussed below, our computational modeling illuminates a number of important kinetic constraints at ribosome collisions that are not readily apparent from previous genetic and structural studies.

Contrary to what one might intuitively expect, abortive termination by leading and trailing ribosomes at collisions have qualitatively distinct effects on protein expression (CSAT versus CAT models, Fig 3B). Specifically, the observed decrease in protein expression at high initiation rates (Fig 1C) can be explained only if the leading ribosome in a collision undergoes abortive termination, after accounting for changes in mRNA levels (Fig 2C). Furthermore, the decrease in protein expression at high initiation rates due to abortive termination requires ribosomes to stall repeatedly (Fig 3C). This observation suggests that ribosomes stall at multiple locations within each polybasic tract, which then leads to abortive termination. This is in contrast to elongation stalls due to amino acid starvation in bacteria, in which a single stall-inducing codon is sufficient to trigger ribosome collisions but does not cause a reduction in protein expression at high initiation rates [35]. It will be interesting to test whether a reduction in protein expression due to poly-A-mediated stalls at high initiation rates is also observed in mammalian cells [55, 56, 58].

Our modeling suggests that decrease in mRNA stability at high initiation rates is a general consequence of competition between endonucleolytic mRNA cleavage and canonical mRNA decay. Notably, we observe such a decrease in simulations of both collision-independent and collision-stimulated models of mRNA cleavage (SEC versus CSEC models, Fig 4B). However, the decrease in protein expression at high initiation rates occurs only when ribosome collisions stimulate endonucleolytic mRNA cleavage (SEC versus CSEC models, Fig 4C). Observation of endonucleolytic mRNA cleavage only 100 nucleotides or more 5′ to the stall site has been interepreted to imply that a queue of several stacked ribosomes might form an oligomer to trigger endonucleolytic mRNA cleavage [26]. Our measured decrease in protein and mRNA levels at high initiation rates can be explained by our simpler models of collision-stimulated quality control, which require sensing of only the disome interface and is consistent with recent structural work [36, 37]. Nevertheless, it remains possible that more complex kinetic models with higher-order interactions not considered in our work underlie our observed changes in gene expression.

We observed that deletion of Asc1 or Hel2 rescues the decreased protein expression and mRNA levels at high initiation rates from stall-containing mRNAs (Fig 5). Hel2/ZNF598 associates with collided ribosomes and stimulates RQC and NGD through ribosomal ubiquitination [36, 37]. Thus, ribosome collisions likely occur still in the absence of Hel2, but they do not stimulate ribosomal ubiquitination and downstream quality control. Asc1/RACK1 localizes to the interface between the collided ribosomes [36, 37] and is required for binding of Hel2 to 18S rRNA [64] and triggering RQC and NGD [16, 55, 57, 58]. Thus, one possibility is that ribosome collisions occur in the absence of Asc1 but are not detected by cells because of impaired binding of Hel2 to collided ribosomes. An alternate possibility is suggested by observations that ΔASC1 strains have perturbed translation initiation and exhibit a growth defect [63, 65]. Hence, the frequency of ribosome collisions could be lower on stall-containing mRNAs in the ΔASC1 strains, which would also prevent collision-stimulated quality control. However, ΔASC1 strains have similar polysome–monosome ratio to WT strains [63, 65], which argues against a substantial global reduction in translation initiation rate.

An important limitation of our model is that we did not explicitly consider the kinetics of ribosomal ubiquitination by Hel2 or its downstream effects on abortive termination and endonucleolytic mRNA cleavage [37]. Ribosomal ubiquitination can be readily incorporated in our rule-based modeling framework. However, we lack experimental estimates of the rates of ubiquitination and deubiquitination at collided ribosomes. Whether and how abortive termination and endonucleolytic mRNA cleavage compete with each other at ubiquitinated ribosomes also remains poorly understood. Genetic and biochemical studies of RQC and NGD mutants with separable effects on abortive termination and endonucleolytic mRNA cleavage [37, 66] will enable formulation of a joint kinetic model of these processes in the future.

Our measurements of normal mRNA levels (Fig 2B) do not support models of simple kinetic competition between translating ribosomes and canonical mRNA decay factors in *S. cerevisiae*. For example, our data are inconsistent with a model in which regions of the mRNA that are not covered by ribosomes are subject to random endonucleolytic cleavage [67]. Any competition between translation and canonical mRNA decay must occur prior to start codon recognition, such as during cap binding by translation initiation factors [2].

Our computational results allow us to draw four key conclusions regarding kinetic models of eukaryotic translation. First, the widely used TJ model [29, 30] (Fig 3B) does not accurately capture ribosome dynamics on stall-containing eukaryotic mRNAs. Specifically, our measured decrease in protein expression at high initiation rates (Fig 1C) is inconsistent with the TJ model. Second, the TJ model does not consider the interaction between translation and mRNA stability, which we observe experimentally (Fig 2C). Third, recent studies have applied the TJ model to infer the extent of ribosome collisions from ribosome profiling measurements [68–70]. These studies assume that collided ribosomes are underrepresented in ribosome profiling data because of protocol-related biases against isolating multiribosome footprints. Our work suggests that collided ribosomes can also be depleted from the pool of translated mRNAs because they stimulate abortive termination and mRNA decay. Fourth, modeling studies of abortive termination assume that ribosomes randomly drop off during elongation [48, 49] or that abortive termination due to ribosome collisions is negligible [71]. Our work instead reveals that abortive termination occurs in response to ribosome collisions caused by elongation stalls and that this process is highly regulated by cells using dedicated quality-control factors.

Our analysis of endogenous *S. cerevisiae* mRNAs that contain stall-encoding sequences reveals that the initiation rates of these mRNAs might have been evolutionarily tuned to values low enough to escape ribosome collision-driven reduction in gene expression (Fig 6). However, this aggregate analysis will not identify individual stall-containing mRNAs on which

ribosome collisions might have a regulatory role. Nor does our analysis consider mRNAs in which ribosome stalls are caused by sequences other than polybasic tracts such as mRNA stem loops or other peptide features [72]. A straightforward way by which ribosome collisions on individual mRNAs could play a functional role is through direct autoregulation, such as that previously described for the Rqc1 component of the RQC complex [16]. A more intriguing possibility is suggested by the fact that polybasic tracts occur in several mRNAs that encode components of either the translation apparatus or the ribosome biogenesis pathway. The stalls on these mRNAs could be part of a cellular feedback control that tunes the translational capacity of the cell by sensing ribosome collisions on mRNAs that encode the translation apparatus or enable its synthesis.

Materials and methods

Programming code and instructions for reproducing the simulations, data analyses, and figures in this manuscript are publicly available at https://github.com/rasilab/ribosome_collisions_yeast. Our customized versions of the simulation software PySB, BioNetGen, and NFSim can be accessed from our laboratory's GitHub page (https://github.com/rasilab).

Strain and plasmid construction

S1 and S2 Tables list the plasmid cloning vectors and the *S. cerevisiae* strains used in this work. DNA_ sequences.fasta (https://github.com/rasilab/ribosome_collisions_yeast/blob/master/data/dna_sequences.fasta) contains the sequences of the reference plasmid vectors and inserts used in this work. Plasmids and strains will be sent upon request. PCR oligo sequences are available upon request.

The BY4741 *S. cerevisiae* strain background (S288C, *MATa HIS3Δ1 LEU2Δ0 MET15Δ0 URA3Δ0*, Thermo Fisher) was used for all experiments in this study. The pHPSC16 vector containing *mKate2* and the *LEU2* marker was integrated as a single copy at the *HO* locus. The resulting strain (scHP15) was used as a parent for inserting all *PGK1-YFP* reporters and for deleting quality control–associated genes. The p41894 vector and its derivatives containing *PGK1-YFP* variants and the *URA3* marker were inserted as a single copy at a ChrI intergenic locus. The pSB2273 vector containing *HEL2* or *ASC1* variants and the *HIS3* marker were inserted as a single copy at the *HIS3* locus. Integration into the *S. cerevisiae* genome was performed by transforming 0.5–2 μg of the linearized (NotI or PmeI restriction) plasmid vector according to the LiAc/SS carrier DNA/PEG method [73]. Single yeast colonies were selected on synthetic complete (SC) agar plates lacking either leucine, uracil, or histidine.

LTN1, *DOM34*, *HEL2*, or *ASC1* was deleted from the *S. cerevisiae* genome using PCR-mediated gene disruption [74]. Deletion cassettes with flanking homology arms are provided in DNA_sequences.fasta. Primers for amplifying *KAN* or *NAT* resistance markers with 40–50-bp homology arms were designed using the sequences from the *S. cerevisiae* genome deletion project webpage (http://www-sequence.stanford.edu/group/yeast_deletion_project/downloads.html).

All plasmids were cloned by restricting the parent vectors and inserting 2–4 PCR-amplified fragments with 20–30-bp homology arms using isothermal assembly [75]. The inserted sequences were confirmed by Sanger sequencing. The 5′ UTR mutations for varying initiation rate and the ribosome stall/control repeats were introduced into the PCR primers used for amplifying inserts prior to isothermal assembly. These 5′ UTR and insert sequences are provided in DNA_sequences.fasta. Standard molecular biology procedures [76] were followed for all other steps of plasmid cloning. For cloning the endogenous stall inserts in Fig 6B, the last lysine codon in *PGK1* was removed so that it did not artificially amplify the observed effect of the endogenous stalls on protein expression.

The *3×FLAG-PGK1-YFP* reporter genes were inserted into the p41894 vector [77] between the *GPD* (*TDH3*) promoter + 5′ UTR and the *CYC1* 3′ UTR + transcriptional terminator using the XbaI and XhoI restriction sites. The sequence around the XbaI site of p41894 was modified to retain the native 5′ UTR sequence of *TDH3* while mutating the −12 to −7 nucleotide from ATG to XbaI (TAAACA to TCTAGA). *PGK1** protein-coding sequence was generated from WT *S. cerevisiae PGK1* by introducing 5×AGA repeats at the five locations used in a previous study [11]. The *S. cerevisiae*-optimized *YFP* (mCitrineV163A variant) sequence was taken from a previous work [78] and was fused with *PGK1* or *PGK1** with an intermediate BamHI site. The sequence of the resulting *p41894-3×FLAG-PGK1*-YFP* vector, pHPSC57, is provided in DNA_sequences.fasta. The 3×*FLAG-PGK1-YFP* insert sequence between XbaI and XhoI of pHPSC417 is also provided.

The *HO3-LEU2-pGPD-mKate2-Cyc1t* vector was constructed by inserting two 400-bp sequences from the *HO* locus of the *S. cerevisiae* genome into a pUC19 backbone. The *LEU2* gene from pRS315 [79] and a *GPDp-mKate2-CYC1t* cassette were inserted between the two 400-bp *HO* sequences. The *mKate2* coding sequence was taken from previous work [80] and was synonymously mutated to the most frequent codon for each amino acid in the *S. cerevisiae* genome. The *GPDp-Cyc1t* flanking regions are identical to the ones in our reporter vector, described above. The sequence of the resulting vector, pHPSC16, is provided in DNA_sequences.fasta.

The *HEL2* and *ASC1* WT or mutant coding sequences were inserted into pSB2273 between the *GPD* promoter and the *ADH1* terminator between the XhoI and BamHI restriction sites. The *HEL2* coding sequence was PCR-amplified from the *S. cerevisiae* genome, whereas the *ASC1* coding sequence was PCR-amplified from *S. cerevisiae* cDNA to exclude the intron. Sequence of the parent vector, pSB2273, is provided in DNA_sequences.fasta. All the *HEL2* and *ASC1* insert sequences cloned into pSB2273 are also provided.

For creating the barcoded reporter strains for mRNA measurements, each of the *PGK1-YFP* reporter coding sequences was PCR-amplified separately with four distinct barcodes in the 3′ UTR. The ATG → CTG control reporters were amplified with three distinct barcodes. The PCR products were separately inserted into pHPSC57 between the XbaI and XhoI restriction sites using isothermal assembly. The isothermal assembly products were pooled and transformed at high efficiency into electro-competent *Escherichia coli*. After overnight selection on LB-agar plates with carbenicillin, the bacterial lawn was scraped to extract the pooled plasmids. The pooled plasmids were integrated into either scHP15 WT strain or one of the four deletion strain backgrounds at the ChrI intergenic locus using the same transformation protocol as above. A minimum of 500 *S. cerevisiae* colonies were pooled for each strain and stored as glycerol stocks. The pooled reporter plasmids and their respective barcodes are provided as part of the data analysis code at https://github.com/rasilab/ribosome_collisions_yeast#high-throughput-sequencing.

Flow cytometry

In total, 3–8 single *S. cerevisiae* colonies from transformation were inoculated into separate wells of 96-well plates containing 150 μl of yeast extract peptone with 2% dextrose (YEPD) medium in each well and grown overnight for 16 h at 30°C with shaking at 800 rpm. The saturated cultures were diluted 100-fold into 150 μl of fresh YEPD medium. After growing for 6 h under the same conditions as overnight, the plates were placed on ice and analyzed using the 96-well attachment of a flow cytometer (BD FACS Aria or Symphony). Forward scatter (FSC), side scatter (SSC), YFP fluorescence (FITC), and mKate2 fluorescence (PE.Texas.Red) were measured for 10,000 cells in each well. The resulting data in individual .fcs files for each well were combined into a single tab-delimited text file. Mean YFP and mKate2 fluorescence in

each well was calculated, and the YFP background from the scHP15 strain and the mKate2 background from the BY4741 strain were subtracted. The ratio of the background-subtracted YFP to mKate2 signal was normalized by that of a standard internal control strain included in all experiments. The average and standard error of this ratio were calculated over all transformant replicate wells and is shown in Figs 1, 5 and 6. Analysis code for reproducing the figures from raw flow cytometry data are available at https://github.com/rasilab/ribosome_collisions_yeast. In total, 2–5 outliers among several hundred wells that were not fluorescent were discarded in some experiments, and they are indicated as such in the analysis code for that experiment. The mKate2 channel measurement was not steady during the flow cytometry measurement in Fig 5D, and hence the YFP fluorescence for this experiment is plotted without normalizing by mKate2 signal. In general, normalizing by mKate2 slightly reduced the fluctuations between biological replicates but did not have any qualitative effect in any experiment. Full analysis code for flow cytometry data is provided in https://github.com/rasilab/ribosome_collisions_yeast#flow-cytometry. The *P* values in Fig 5D are calculated using the t.test function in R. The analysis code for calculating these *P* values is provided in https://github.com/rasilab/ribosome_collisions_yeast/blob/master/analysis/flow/hel2_asc1_mutants.md.

Deep-sequencing quantification of mRNA levels

S. cerevisiae glycerol stocks containing pooled reporter strains ($> 10^7$ cells) were thawed and grown overnight in 3 ml of YEPD at 30˚C in roller drums. The saturated cultures were diluted 200-fold into 50 ml of fresh YEPD medium and grown for 4.5 h at 30˚C with shaking at 200 rpm. Each culture was split into two 25-ml aliquots (one for extracting RNA and the other for extracting genomic DNA) into prechilled 50-ml conical tubes and spun down at 3000*g*, 4˚C, 2 min. The supernatant was removed, and the cell pellets were flash frozen in a dry ice/ethanol bath and stored at −80˚C.

For RNA extraction, the cell pellets were resuspended in 400 μl of Trizol (15596-026, Thermo Fisher) in a 1.5-ml tube and vortexed with 500 μl of glass beads (G8772, Sigma) at 4˚C for 15 min. RNA was extracted from the resulting lysate using the Direct-zol RNA Miniprep Kit (R2070, Zymo) following manufacturer's instructions. For genomic DNA extraction, the cell pellets were processed using the YeaStar Genomic DNA Kit (D2002, Zymo) following manufacturer's instructions.

After RNA extraction, 1 μg of total RNA was reverse transcribed into cDNA using SuperScript IV (18090-010, Thermo Fisher) reverse transcriptase and a primer annealing to the 3′ UTR (oMF321: ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCGTGACATAACT AATTACAT) following manufacturer's instructions. A 252-nucleotide region surrounding the 8-nucleotide barcode was PCR-amplified from both the cDNA and genomic DNA in two rounds. Round 1 PCR was carried out for 10 cycles with oMF321 and oHP139 (GTGACTGG AGTTCAGACGTGTGCTCTTCCGATCTAAGGACCCAAACGAAAAG) primers using Phusion polymerase (F530, Thermo Fisher) following manufacturer's instructions. Round 2 PCR was carried out for 10 cycles for genomic DNA samples and 10 or 14 cycles for cDNA samples with a common reverse primer (oAS111: AATGATACGGCGACCACCGAGATCTA CACTCTTTCCCTACACGACGCTC) and indexed forward primers for pooled high-throughput sequencing of different samples (CAAGCAGAAGACGGCATACGAGATxxxxxxGTGAC TGGAGTTCAGACGTGTGCTC, xxxxxx refers to the 6-nucleotide sample barcode and is provided in a table as part of our analysis script). The pooled libraries were sequenced in a HiSeq 2000 (Illumina) in a 50-nucleotide single-end rapid-run mode.

The high-throughput sequencing data were obtained as demultiplexed .fastq files. The barcode corresponding to each reporter was identified in each 50-nucleotide read, and the

barcode counts for each reporter were tallied within each sample. The log2 barcode counts for the cDNA were normalized by those of the genomic DNA, after applying a minimum threshold of 100 counts per barcode within each sample. The average and standard error of this log2 ratio was calculated over all barcodes that crossed the 100-count threshold. The average log2 ratio was median-normalized within each set of 5′ UTR variants for a given coding sequence and sample and is shown in Figs 2 and 5. Analysis code for reproducing the figures from raw high-throughput sequencing data is available at https://github.com/rasilab/ribosome_collisions_yeast#high-throughput-sequencing.

Western blotting

Overnight cultures were grown from single *S. cerevisiae* colonies in 3 ml of YEPD at 30°C in roller drums. The saturated cultures were diluted 100-fold into 3 ml of fresh YEPD medium and grown for 5 h at 30°C in roller drums. The cultures were quickly chilled in an ice-water bath and spun down in 1.5-ml tubes. The cell pellet was washed in 500 μl of water and incubated in 200 μl of 0.1 M NaOH for 5 min at room temperature. The pellet was resuspended in 50 μl of 1× Laemmli buffer, and western blots were performed using standard molecular biology procedures [76]. The anti-FLAG antibody (F1804, Sigma) and the anti-H3 antibody (ab1791, Abcam) were visualized using IRDye antibodies (926-68072 and 926-32211, Licor) on an Odyssey imager. The raw signal in each lane was quantified using ImageJ using the rectangle-select tool followed by the Analyze-Measure tool. For each lane, the FLAG signal was divided by the H3 signal and normalized to a maximum of 10 within each blot. Uncropped images of blots and their quantification are provided in https://github.com/rasilab/ribosome_collisions_yeast/tree/master/data/western. The normalization of lanes is done in https://github.com/rasilab/ribosome_collisions_yeast/blob/master/analysis/western/western_analysis.md.

Kinetic modeling

All scripts and files mentioned below are available at https://github.com/rasilab/ribosome_collisions_yeast/modeling. We specify our kinetic models using the PySB interface [81] to the BioNetGen modeling language [82]. The resulting Python script tasep.py is parsed using the BioNetGen parser into the tasep.bngl file and converted into the tasep.xml file for input to the agent-based stochastic simulator, NFsim [83].

The molecules and the reactions in our model, along with the parameters and initial conditions used in our simulations are described below. A complete list of parameters in our model is provided in S3 Table. The tasep.bngl file provides a machine-readable summary of the description below using BioNetGen language.

Molecules. Our kinetic model of eukaryotic quality control has five molecule types: DNA, ribosome, mRNA, fully synthesized protein, and aborted protein. Below, we describe the components, their states, and binding partners of each molecule type.

The DNA molecule does not have any components. It is used as a template for generating new mRNA molecules through transcription in order to compensate for the mRNA molecules degraded through canonical mRNA decay or endonucleolytic mRNA cleavage and thus maintain a translatable mRNA pool in our simulation at steady state.

The ribosome molecule has three components: an A site (*a*), a back/mRNA exit site (*b*), and a front/mRNA entry site (*f*). These sites do not have distinct states but serve as bonding sites. The A site can bond to the codon site on the mRNA. The back/mRNA exit site can bond to a collided trailing ribosome on the mRNA. The front/mRNA entry site can bond to a collided

leading ribosome on the mRNA. The ribosome has an mRNA footprint of 10 codons in our simulation, which is the approximate size of a ribosome footprint in *S. cerevisiae* [46].

The mRNA molecule has the following components: cap, start region, codon sites (c_i), backbone sites (r_i), and poly-A tail sites (p_i).

The cap can be either present or absent. The cap has to be present for translation initiation and be absent for initiation of 5′–3′ exonucleolytic mRNA decay. mRNAs are transcribed with caps in our model because the kinetics of how mRNAs are synthesized and rendered translatable is not of interest in this work. mRNAs can be decapped either upon complete deadenylation of the poly-A tail [47] or after endonucleolytic mRNA cleavage [59].

The mRNA start region can be either clear or blocked. The start region has to be clear for translation initiation. Translation initiation renders the start region blocked, whereas elongation past a ribosome footprint distance from the start codon renders the start region clear. The inclusion of start region component in our model is not strictly necessary, but it allows compact specification of start codon occlusion by initiating ribosomes. The alternative is explicit specification of the ribosome occupancy of every codon within the ribosome footprint distance from the start codon.

The mRNA codon sites do not have distinct states. They serve as bonding sites for the ribosome A site. The number of mRNA codon sites Lm is specified by the length of the coding region in our simulation (650 corresponding to the *PGK1-YFP* reporter used in our experiments).
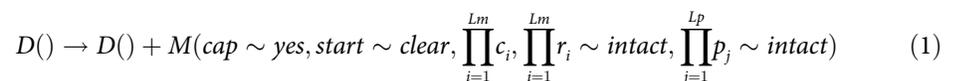
The mRNA backbone sites can be either intact, endonucleolytically cleaved, or exonucleolytically cleaved. These sites serve as substrates for cotranslational endonucleolytic cleavage or exonucleolytic degradation. The number of mRNA backbone sites Lm is specified by the length of the coding region in our simulation.

The mRNA poly-A tail sites can be either intact or exonucleolytically cleaved. The poly-A sites serve as substrates for deadenylation during canonical mRNA decay. The number of mRNA poly-A sites Lp is specified by the length of the poly-A region in our simulation (set to 60).

The fully synthesized protein P and aborted protein A molecules do not have any components. They are used for tracking the number of full proteins and aborted proteins produced during the course of the simulation. An equivalent strategy will be to exclude these molecules and just track the ribosomes that undergo normal and premature termination events.

Kinetic reactions.   Below, we describe each type of kinetic reaction in our model of quality control during eukaryotic translation. We use a syntax similar to that of BioNetGen [82] for illustrating kinetic reactions.

Transcription produces capped mRNAs without any ribosomes (Eq 1). All mRNA backbone sites and poly-A tail sites are intact. The transcription rate constant $k_{transcription}$ is set to 0.001 s^{-1}. The value of this parameter is arbitrary and is chosen to maintain a pool of 1–2 translatable mRNAs at any given time in the simulation and to produce around 1,000 mRNAs during a typical simulation run for 10^6 s.

$$D() \rightarrow D() + M(cap \sim yes, start \sim clear, \prod_{i=1}^{Lm} c_i, \prod_{i=1}^{Lm} r_i \sim intact, \prod_{j=1}^{Lp} p_j \sim intact) \qquad (1)$$

Translation initiation occurs on capped mRNAs when the start region is free, the start codon is not bonded, and the mRNA backbone at the start codon is intact (Eq 2). It results in the start region being blocked and the ribosome A site being bonded to the start codon ("!1" in Eq 2). The translation initiation rate k_{init} is systematically varied in our simulations. We use a default value of 0.1 s^{-1} to match previous estimates [44, 84]. The absolute value of this

parameter is not critical in our modeling. Rather, the dimensionless ratio of the translation initiation rate to elongation rate is the important parameter in our model because it determines the frequency of ribosome collisions at the stall.

$$R(a) + M(cap \sim yes, start \sim clear, c_1, r_1 \sim intact) \rightarrow$$
$$R(a!1) \circ M(cap \sim yes, start \sim blocked, c_1!1, r_1 \sim intact)$$

(2)

Elongation of a ribosome results in the ribosome A site moving from codon c_i to codon c_{i+1} (Eq 3). Elongation requires that there is no leading ribosome within a footprint distance at c_{i+10}. Additionally, when a ribosome elongates at the ninth codon from the initiation codon, the mRNA start region is cleared for a new round of initiation (Eq 4). Finally, elongation of ribosomes that have been hit from the back results in loss of the molecular interaction with the trailing ribosome ("!3" in Eq 5). The elongation rate in all these cases, k_{elong}, is set to 10 s^{-1} to match experimental estimates [85, 94].

$$R(a!1, f, b) \circ M(c_i!1, c_{i+1}, c_{i+10}, r_i \sim intact) \rightarrow$$
$$R(a!1, f, b) \circ M(c_i, c_{i+1}!1, c_{i+10}, r_i \sim intact)$$

(3)

$$R(a!1, f, b) \circ M(c_9!1, c_{10}, c_{19}, r_9 \sim intact, start \sim blocked) \rightarrow$$
$$R(a!1, f, b) \circ M(c_9, c_{10}!1, c_{19}, r_9 \sim intact, start \sim clear)$$

(4)

$$R(a!1, f, b!3) \circ M(c_{i-10}!2, c_i!1, c_{i+1}, c_{i+10}, r_i \sim intact) \circ R(a!2, f!3) \rightarrow$$
$$R(a!1, f, b) \circ M(c_{i-10}!2, c_i, c_{i+1}!1, c_{i+10}, r_i \sim intact) \circ R(a!2, f)$$

(5)

Termination results in dissociation of the ribosome–mRNA complex along with production of a protein molecule (Eq 6). Termination of ribosomes that have been hit from the back results in loss of the molecular interaction with the trailing ribosome (Eq 7). The termination rate, k_{term}, is set to 1 s^{-1} to be lower than the normal elongation rate of 10 s^{-1}. This choice reflects the observation that ribosome density at stop codons tends to be several-fold higher than at sense codons [40]. The absolute value of k_{term} is not important in our modeling as long as it is greater than the total rate of initiation and elongation across the mRNA.

$$R(a!1, f, b) \circ M(c_{Lm}!1) \rightarrow R(a, f, b) + M(c_{Lm})$$

(6)

$$R(a!1, f, b!3) \circ M(c_{Lm-10}!2, c_{Lm}!1, r_{Lm} \sim intact) \circ R(a!2, f!3) \rightarrow$$
$$R(a, f, b) + M(c_{Lm-10}!2, c_{Lm}, r_{Lm} \sim intact) \circ R(a!2, f)$$

(7)

Collision between two ribosomes requires them to be separated by exactly a footprint distance on an intact mRNA, and it results in a bond ("!3" in Eq 8) between the f site of the trailing ribosome and the b site of the leading ribosome. The collision rate constant is set equal to the elongation rate at the A site codon of the trailing ribosome. This choice reflects our assumption that collision occurs when the trailing ribosome tries to translocate to the next codon on the mRNA but is sterically blocked by the leading ribosome.

$$R(a!1, b) \circ M(c_{i-10}!2, c_i!1) \circ R(a!2, f) \rightarrow$$
$$R(a!1, b!3) \circ M(c_{i-10}!2, c_i!1) \circ R(a!2, f!3)$$

(8)

Deadenylation of the poly-A tail occurs in the 3′–5′ direction. Deadenylation at position j requires that the A at position j+1 has been removed (Eq 9). This constraint does not apply to the 3′ end of the poly-A tail (at position Lp), whose removal begins the process of

deadenylation (Eq 10). We set the deadenylation rate per nucleotide $k_{deadenylation}$ to be 0.03 s^{-1} and the poly-A tail to be 60 nucleotides to match previous estimates for the *PGK1* mRNA [47]. The deadenylation rate sets the overall rate of canonical mRNA decay in our modeling. Our model of deadenylation can be refined further based on recent biochemical and genome-wide studies of poly-A tail metabolism [14, 86]. We do not pursue this here because we do not monitor poly-A tail length, and our focus is on modeling cotranslational quality control.

$$M(p_j \sim intact, p_{j+1} \sim exocleaved) \rightarrow M(p_j \sim exocleaved, p_{j+1} \sim exocleaved) \tag{9}$$

$$M(p_{Lp} \sim intact) \rightarrow M(p_{Lp} \sim exocleaved) \tag{10}$$

Decapping during canonical mRNA decay occurs after the poly-A tail has been fully deadenylated (Eq 11). We set the decapping rate constant $k_{decapping}$ to be 0.01 s^{-1}, similar in magnitude to previous work [47]. Because we do not know the initiation rate of deadenylated but capped mRNAs, we assume that mRNAs are initiated at their normal efficiency right until the cap is removed (Eq 2).

$$M(cap \sim yes, p_1 \sim exocleaved) \rightarrow M(cap \sim no, p_1 \sim exocleaved) \tag{11}$$

5′–3′ exonucleolysis at position i requires that the mRNA backbone at position i−1 has been removed (Eq 12) and that the position is not blocked by a ribosome. 5′–3′ exonucleolysis can initiate during canonical mRNA decay only after the mRNA has been decapped (Eq 13). We set the 5′–3′ exonucleolysis rate constant k_{exo_53} to be 1 s^{-1} [47], such that it is faster than the total rate of deadenylation and decapping.

$$M(r_i \sim intact, r_{i-1} \sim exocleaved, c_i) \rightarrow$$
$$M(r_i \sim exocleaved, r_{i-1} \sim exocleaved, c_i) \tag{12}$$

$$M(r_1 \sim intact, c_1, cap \sim no) \rightarrow M(r_1 \sim exocleaved, c_1, cap \sim no) \tag{13}$$

3′–5′ exonucleolysis at position i requires that the mRNA backbone at position i+1 has been removed (Eq 14) and that the position i is not blocked by a ribosome. The 3′–5′ exonucleolysis can initiate during canonical mRNA decay only after the mRNA has been fully deadenylated (Eq 15). We set the 3′–5′ exonucleolysis rate constant k_{exo_35} to be 0. This choice is consistent with the much slower observed rate of 3′–5′ exonucleolysis in comparison with 5′–3′ exonucleolysis in *S. cerevisiae* [47]. Furthermore, any mRNA undergoing 3′–5′ exonucleolysis will not contribute to the pool of full-length proteins or mRNA molecules, which are the experimental observables of interest in this work.

$$M(r_i \sim intact, r_{i+1} \sim exocleaved, c_i) \rightarrow$$
$$M(r_i \sim exocleaved, r_{i+1} \sim exocleaved, c_i) \tag{14}$$

$$M(r_{Lm} \sim intact, c_{Lm}, p_1 \sim exocleaved) \rightarrow$$
$$M(r_{Lm} \sim exocleaved, c_{Lm}, p_1 \sim exocleaved) \tag{15}$$

Abortive (premature) termination results in dissociation of the aborting ribosome from the mRNA. Abortive termination of a ribosome also results in the loss of molecular interaction with any collided ribosome in the front or the back. In our modeling of collision-stimulated quality control from intact mRNAs, ribosomes have different rates of abortive termination depending on whether they have undergone collision from the front ($k_{abort\_front\_hit}$, Eq 18), back ($k_{abort\_back\_hit}$, Eq 17), or both ($k_{abort\_both\_hit}$, Eq 19) sides. Ribosomes that have not

undergone collisions also abortively terminate ($k_{abort\_no\_hit}$, Eq 16). The different models of abortive termination in our work (Fig 3A) correspond to different combinations of the four types of abortive termination. In the SAT model (Eq 8 in Fig 3A), $k_{abort\_no\_hit} = k_{abort\_back\_hit} > 0$ and $k_{abort\_front\_hit} = k_{abort\_both\_hit} = 0$. This choice implies that collisions do not have any stimulatory effect on abortive termination and that ribosomes that are stacked behind the leading stalled ribosome do not undergo abortive termination. In the CSAT model (Fig 3A), $k_{abort\_back\_hit} = k_{abort\_both\_hit} > 0$ and $k_{abort\_no\_hit} = k_{abort\_front\_hit} = 0$. This choice implies that collisions from the back are necessary for abortive termination, but ribosomes that are stalled either on their own or because of a leading stacked ribosome do not undergo abortive termination. In the CAT model (Fig 3A), $k_{abort\_front\_hit} = k_{abort\_both\_hit} > 0$ and $k_{abort\_no\_hit} = k_{abort\_back\_hit} = 0$. This choice implies that collisions from the front are necessary for abortive termination, but ribosomes that are stalled without a leading stacked ribosome do not undergo abortive termination. The in vivo abortive termination rates are not known. In each of the above three models, we chose the respective nonzero abortive termination rates to be such that they would decrease protein expression appreciably because of stalling.

$$R(a!1, f, b) \circ M(c_i!1, r_i \sim intact) \rightarrow R(a, f, b) + M(c_i, r_i \sim intact) + A() \tag{16}$$

$$\begin{aligned} R(a!1, f, b!2) \circ M(c_{i-10}!3, c_i!1, r_i \sim intact) \circ R(a!3, f!2) \rightarrow \\ R(a, f, b) + M(c_{i-10}!3, c_i, r_i \sim intact) \circ R(a!3, f) + A() \end{aligned} \tag{17}$$

$$\begin{aligned} R(a!1, f!2, b) \circ M(c_i!1, c_{i+10}!3, r_i \sim intact) \circ R(a!3, b!2) \rightarrow \\ R(a, f, b) + M(c_i, c_{i+10}!3, r_i \sim intact) \circ R(a!3, f) + A() \end{aligned} \tag{18}$$

$$\begin{aligned} R(a!3, f!4, b!2) \circ M(c_{i-10}!1, c_i!3, c_{i+10}!5, r_i \sim intact) \circ R(a!1, f!2) \circ R(a!5, b!4) \rightarrow \\ R(a, f, b) + M(c_{i-10}!1, c_i, c_{i+10}!5, r_i \sim intact) \circ R(a!1, f) \circ R(a!5, b) + A() \end{aligned} \tag{19}$$

We also modeled abortive termination of ribosomes with a truncated mRNA in its A site using the same rules as above but with $r_i \sim endocleaved$ instead of $r_i \sim intact$. This is necessary for releasing all the ribosomes that are 5′ to the endonucleolytic cleavage site. We assume that ribosomes abort with a uniformly high rate of 1 s^{-1} from mRNAs with cleaved A sites independent of whether they have undergone collision. This is consistent with previous in vitro experiments showing that Dom34/Hbs1 can efficiently recycle ribosomes that are stalled on a truncated mRNA [19, 20].

Endonucleolytic mRNA cleavage on translating ribosomes results in the mRNA being cleaved at a distance Lc 5′ to the A site. We further assume that the endonucleolytic cleavage triggers immediate decapping. This assumption is consistent with increased rates of degradation of mRNAs that have undergone endonucleolytic cleavage [59]. This assumption does not affect any of the conclusions in this work because further rounds of initiation of cleaved mRNAs will not result in the production of full-length proteins, the observable of interest to us. In our modeling of CSEC, we consider different rates of endonucleolytic mRNA cleavage depending on whether ribosomes have undergone collision from the front ($k_{cleave\_front\_hit}$, Eq 21), back ($k_{cleave\_back\_hit}$, Eq 22), or both ($k_{cleave\_both\_hit}$, Eq 23) sides. Ribosomes that have not undergone collisions also trigger endonucleolytic mRNA cleavage ($k_{cleave\_no\_hit}$, Eq 20). The different models of cotranslational endonucleolytic cleavage in our work (Fig 4A) correspond to different combinations of the four types of endonucleolytic cleavage. In the SEC model (Fig 4A), $k_{cleave\_no\_hit} = k_{cleave\_back\_hit} > 0$ and $k_{cleave\_front\_hit} = k_{cleave\_both\_hit} = 0$. This choice implies that ribosome collisions do not have any stimulatory effect on endonucleolytic mRNA cleavage

and that ribosomes that are stacked behind the leading stalled ribosome do not cause endonucleolytic mRNA cleavage. In the CSEC model (Fig 4A), $k_{cleave\_back\_hit} = k_{cleave\_both\_hit} > 0$ and $k_{cleave\_no\_hit} = k_{cleave\_front\_hit} = 0$. This choice implies that collisions from the back are necessary for stimulating endonucleolytic mRNA cleavage, but ribosomes that are stalled either on their own or because of a leading stacked ribosome do not result in endonucleolytic cleavage. Unlike the case of abortive termination, we do not separately model the effect of the trailing versus leading ribosome in a collision stimulating endonucleolytic cleavage, because both these models will result in a cleaved mRNA.

$$R(a!1, b, f) \circ M(c_i!1, r_{i-Lc} \sim intact, cap \sim yes) \rightarrow \rightarrow$$
$$R(a!1, b, f) \circ M(c_i!1, r_{i-Lc} \sim endocleaved, cap \sim no) \tag{20}$$

$$R(a!1, b, f!2) \circ M(c_i!1, c_{i+10}!3, r_{i-Lc} \sim intact, cap \sim yes) \circ R(a!3, b!2) \rightarrow$$
$$R(a!1, b, f) \circ M(c_i!1, c_{i+10}!3, r_{i-Lc} \sim endocleaved, cap \sim no) \circ R(a!3, b) \tag{21}$$

$$R(a!1, b!2, f) \circ M(c_{i-10}!3, c_i!1, r_{i-Lc} \sim intact, cap \sim yes) \circ R(a!3, f!2) \rightarrow$$
$$R(a!1, b, f) \circ M(c_{i-10}!3, c_i!1, r_{i-Lc} \sim endocleaved, cap \sim no) \circ R(a!3, f) \tag{22}$$

$$R(a!3, b!2, f!4) \circ M(c_{i-10}!1, c_i!3, c_{i+10}!5, r_{i-Lc} \sim intact, cap \sim yes) \circ$$
$$R(a!1, f!2) \circ R(a!5, b!4) \rightarrow$$
$$R(a!3, b, f!4) \circ M(c_{i-10}!1, c_i!3, c_{i+10}!5, r_{i-Lc} \sim endocleaved, cap \sim no) \circ$$
$$R(a!1, f) \circ R(a!5, b!4) \tag{23}$$

We set Lc = 10 codons in our simulations, even though experiments indicate that endonucleolytic mRNA cleavage occurs 14–15 nucleotides behind the stalled ribosome [52, 53]. This choice of Lc to be equal to the ribosome footprint size simplifies the specification of reaction rules but does not affect the predictions for either protein output or mRNA lifetime. This can be set to the more physiological value of 5 in case the model predictions for cleaved fragments are examined (such as when comparing against short ribosome footprints [52]). The in vivo endonucleolytic cleavage rates are not known. In the above two models, we chose the respective nonzero endonucleolytic mRNA cleavage rates to be such that they would decrease mRNA lifetimes appreciably because of stalling.

Simulation of quality-control model.   Each set of simulations for systematically varying a few parameters is performed inside one of the subfolders in https://github.com/rasilab/ribosome_collisions_yeast/tree/master/modeling/simulation_runs. See https://github.com/rasilab/ribosome_collisions_yeast#modeling for further details. The Python script choose_simulation_parameters.py specifies the combination of parameters that we simulate using the tasep.py model. The Python script run_simulation.py imports the PySB model specified in tasep.py and then invokes the NFsim simulation with a specific parameter combination determined by a command-line argument. The script run_simulation.py also invokes the R script get_mrna_lifetime_and_psr.R, which calculates the mean and standard deviation of mRNA lifetimes and protein synthesis rate in each simulation. Protein synthesis rate is defined as the ratio of the number of full proteins produced during the simulation to the biological time for which the simulation was run (typically 10^6 s). The lifetime of each mRNA molecule in the simulation is defined as the interval from the time of transcription to the initiation of 5′–3′ exonucleolysis of the mRNA. Because the transcription rate constant was set to 0.001 s^{-1} in our simulations, approximately 1,000 mRNA molecules are followed from birth to death in each

simulation lasting for 10^6 s. In the case of simulations in which mRNAs are not degraded (Fig 3C), the transcription rate is set to zero, and a single initial mRNA is tracked during the course of the full simulation. All the simulation steps are implemented as Snakemake workflows [87] in the Snakefile script.

Analysis of FACS-Seq data from Dvir and colleagues

S1 Table was downloaded from http://www.pnas.org/lookup/suppl/doi:10.1073/pnas. 1222534110/-/DCSupplemental/sd01.xlsx. The exp column in this table was averaged over all sequence_ variant rows that share the same last 4 nucleotides. The mean and standard error of this average are shown in S1 Fig panel C. This analysis is described at https://github.com/ rasilab/ribosome_collisions_yeast/#analysis-of-dvir-et-al-2013-data.

Ribosome profiling analysis

Raw .fastq files from a previous study [62] were downloaded for the SRA experiments, SRX391789 (RNA-seq) and SRX391790 (Ribo-seq). The adapters were removed using cutadapt [88] with the arguments –cut=8 –adapter=TCGTATGCCGTCTTCTGCTTG –minimum-length=15. The trimmed reads were first depleted of rRNA contaminant reads and then aligned to the *S. cerevisiae* transcriptome and genome (sacCer3) using tophat 2.0.14 [89] with the arguments –bowtie1 –num-threads 8 –no-novel-juncs –library-type fr-unstranded –keep-tmp –read-mismatches 2. The resulting .bam files were sorted and indexed using samtools [90]. The aligned reads were assigned to the 13th nucleotide from the 5′ end, and this was used to calculate coverage across the sacCer3 genome.

To identify stall sequences, the protein-coding sequences not marked as "Dubious" in the sacCer3 genome (annotation set: saccharomyces_cerevisiae_R64-1-1) were scanned for 10-codon windows that contained at least 6 lysine and arginine codons or 6 proline codons. The control sequences were similarly identified by scanning for windows that contained at least 6 glutamate and aspartate codons. This analysis is described at https://github.com/rasilab/ ribosome_collisions_yeast/#identification-of-putative-rqc-stalls.

The Ribo-seq coverage in the 150-nucleotide window for each stall or control sequence was normalized by the mean coverage within that window. This normalized coverage was then averaged at each position in the window across all stall or control sequences and shown in Fig 6A, S5 Fig panel A.

To calculate TE of *S. cerevisiae* genes, the Ribo-seq and RNA-seq rpkm values for each gene were downloaded from ftp://ftp.ncbi.nlm.nih.gov/geo/series/GSE75nnn/GSE75897/suppl/. TE for each gene was calculated as the ratio of the Ribo-seq rpkm to RNA-seq rpkm. Only genes with a minimum of 5 rpkm in each sample were used for the TE calculation.

To calculate TE of mRNA regions 5′ to stalls, we restricted the Ribo-seq and RNA-seq coverage to this region for stall-containing mRNAs and calculated TE for mRNAs that contained at least 100 reads in this region. For control mRNAs, we restricted the coverage to regions 5′ to the median-stall location on stall-containing mRNAs (215 codons). We used this restricted region of comparison to account for the 5′– 3′ skew in ribosome density along all *S. cerevisiae* mRNAs [62]. This analysis is described at https://github.com/rasilab/ribosome_collisions_ yeast/#analysis-of-weinberg-et-al-2016-data.

RNA-seq analysis

Raw .fastq files from a previous study [57] were used for differential gene expression analysis between Δ*HEL2*, Δ*ASC1*, and WT strains. Only RNA-seq data from strains containing control reporters were used for the analysis. Any ambiguous terminal nucleotides were trimmed using

cutadapt [88] with the argument –trim-n. The trimmed reads were first depleted of rRNA contaminant reads and then aligned to the *S. cerevisiae* genome (sacCer3) using bowtie 1.1.1 [91] with the arguments -v 2 –un /dev/null –threads 8 –sam. The resulting .sam files were converted to the binary .bam format, sorted, and indexed using samtools [90]. The alignments were used to calculate the number of reads aligning to each transcript in the TxDb.Scerevisiae. UCSC.sacCer3.sgdGene R package using the findOverlaps function from GenomicAlignments R package [92]. Transcripts that had a minimum of 100 alignments in each sample were used for differential gene expression analysis using DESeq2 with default parameters [93]. The Δ*HEL2* and Δ*ASC1* samples were treated as replicates for DESeq2 input and compared with the WT strain. The log2FoldChange output from the results function in DESeq2 was used for plotting Fig 6D. This analysis is described at https://github.com/rasilab/ribosome_collisions_yeast/#analysis-of-sitron-et-al-2017-data.

Supporting information

S1 Fig. Effect of stall location on protein expression. (A) Schematic of *3×FLAG-PGK1*-YFP* reporters used in (B). The hatched regions indicate the five locations where an $(AGA)_5$ is inserted into *PGK1*. One of these locations is synonymously mutated to $(CGG)_5$ in the constructs shown in (B) (along with the no-mutation control). (B) Protein levels of *3×FLAG-PGK1*-YFP* reporters with $(CGG)_5$ inserted at one of the five locations indicated in (A). The no-mutation control is shown as −. Protein levels are quantified as the mean fluorescence of 10,000 cells for each strain using flow cytometry. Error bars show standard error of the mean over 4 independent yeast transformants. Protein levels are expressed as a.u. relative to the mean RFP levels from a constitutively expressed mKate2 control. (C) Protein levels of *YFP* library with randomized 10-nucleotide region preceding the ATG start codon. Data are from Dvir and colleagues [38]. Measured protein levels of all constructs with the same −4 to −1 nucleotides preceding ATG are averaged, and the error bars represent standard error of this average. The underlying data for panels B and C can be found at https://github.com/rasilab/ribosome_collisions_yeast. a.u., arbitrary unit; RFP, red fluorescent protein; YFP, yellow fluorescent protein.
(PDF)

S2 Fig. Simulated effect of elongation and abortive termination rates in the TJ, SAT, and CSAT models. (A, B, D) Protein synthesis rate as a function of initiation rate for different rates of elongation at the ribosome stall in the TJ (A), SAT (B), and CSAT (D) models. The different elongation rates at the stall are indicated graphically as vertical dashed lines for comparison with initiation rate. (C) Protein synthesis rate as a function of initiation rate for different rates of abortive termination at the ribosome stall in the SAT model. The mRNA is 650 codons long, and the stall is encoded by six slowly translated codons located after 400 codons from the start. All other model parameters are listed in S3 Table. The underlying data for panels A, B, C, and D can be found at https://github.com/rasilab/ribosome_collisions_yeast. CSAT, collision-stimulated abortive termination; SAT, simple abortive termination; TJ, traffic jam.
(PDF)

S3 Fig. Simulated effect of cleavage rate and number of stalls in the SEC and CSEC models. (A) Protein synthesis rate as a function of initiation rate for different rates of cotranslational endonucleolytic cleavage in the SEC model. (B) Protein synthesis rate as a function of initiation rate for different number of codons encoding the ribosome stall in the CSEC model. The mRNA is 650 codons long, and the stall is encoded by six slowly translated codons located after 400 codons from the start in (A). All other model parameters are listed in S3 Table. The underlying data for panels A and B can be found at https://github.com/rasilab/ribosome_collisions_yeast.

CSEC, collision-stimulated endonucleolytic cleavage; SEC, simple endonucleolytic cleavage.
(PDF)

S4 Fig. Repressive effect of high initiation rate on gene expression requires Hel2/ZNF598 and Asc1/RACK1. (A) Protein levels of *3×FLAG-PGK1*-YFP* reporters (see Fig 1A) with varying initiation rates and with stall (5×CGG) or control (5×AGA) repeats. The reporters were integrated into the genome of isogenic strains with individual full deletions of *LTN1*, *HEL2*, or *ASC1*. (B) Protein levels of the 8×CCA control reporter expressed in either ΔHEL2 (top) or ΔASC1 (bottom) strain and complemented with the indicated HEL2 or ASC1 variant, respectively. (C) Western blots of 8×CCG and 8×CCA reporters with varying initiation rates and expressed in either ΔLTN1 or WT strain. Antibody against the FLAG epitope at the N terminus was used for detecting both full-length 3×FLAG-PGK1*-YFP and the truncated 3×FLAG-PGK1* due to abortive termination at the 8×CCG stall. * indicates the expected size (45 kD) of the truncated peptide. Histone H3 level is shown as loading control. Error bars in (A) and (B) show standard error over 3 or 4 independent transformants. The ΔASC1-CAAA-5×CGG variant alone has a single transformant. The underlying data for panels A and B can be found at https://github.com/rasilab/ribosome_collisions_yeast. WT, wild-type; YFP, yellow fluorescent protein. (PDF)

S5 Fig. Ribosome density around control sequences does not increase, and TE of mRNA regions 5′ to stalls. (A) Mean ribosome density around control sequences using data from Weinberg and colleagues [62]. Control sequences are defined as 10-codon windows that have a minimum of 6 glutamate or aspartate codons. A total of 1,552 *S. cerevisiae* mRNAs have at least one such control sequence. The ribosome density is normalized within the window around each control sequence before calculating the mean across all sequences. (B) TE of *S. cerevisiae* mRNA regions 5′ to stall sequences. TE is defined as the normalized ratio of ribosome footprint counts to total mRNA counts of regions 5′ to stalls for stall-containing mRNAs or the region on control mRNAs 5′ to the median stall location (215 codons) on stall-containing mRNAs. The underlying data for panels A and B can be found at https://github.com/rasilab/ribosome_collisions_yeast. TE, translation efficiency. (PDF)

S1 Table. List of plasmids used for this study.
(PDF)

S2 Table. List of *S. cerevisiae* strains used in this study.
(PDF)

S3 Table. Simulation parameters.
(PDF)

Acknowledgments

Author Contributions

Conceptualization: Heungwon Park, Arvind R. Subramaniam.

References

1. Sonenberg N, Hinnebusch AG. Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. Cell. 2009; 136(4):731–745. https://doi.org/10.1016/j.cell.2009.01.042 PMID: 19239892

2. Schwartz DC, Parker R. Mutations in Translation Initiation Factors Lead to Increased Rates of Deadenylation and Decapping of mRNAs in Saccharomyces Cerevisiae. Mol Cell Biol. 1999; 19(8):5247–5256. https://doi.org/10.1128/mcb.19.8.5247 PMID: 10409716

3. Schwartz DC, Parker R. mRNA Decapping in Yeast Requires Dissociation of the Cap Binding Protein, Eukaryotic Translation Initiation Factor 4E. Mol Cell Biol. 2000; 20(21):7933–7942. https://doi.org/10.1128/mcb.20.21.7933-7942.2000 PMID: 11027264

4. Chan LY, Mugler CF, Heinrich S, Vallotton P, Weis K. Non-Invasive Measurement of mRNA Decay Reveals Translation Initiation as the Major Determinant of mRNA Stability. eLife. 2018; 7:e32536. https://doi.org/10.7554/eLife.32536 PMID: 30192227

5. Muhlrad D, Decker CJ, Parker R. Turnover Mechanisms of the Stable Yeast PGK1 mRNA. Mol Cell Biol. 1995; 15(4):2145–2156. https://doi.org/10.1128/mcb.15.4.2145 PMID: 7891709

6. LaGrandeur T, Parker R. The Cis Acting Sequences Responsible for the Differential Decay of the Unstable MFA2 and Stable PGK1 Transcripts in Yeast Include the Context of the Translational Start Codon. RNA. 1999; 5(3):420–433. https://doi.org/10.1017/s1355838299981748 PMID: 10094310

7. Parker R. RNA Degradation in Saccharomyces Cerevisae. Genetics. 2012; 191(3):671–702. https://doi.org/10.1534/genetics.111.137265 PMID: 22785621

8. Roy B, Jacobson A. The Intimate Relationships of mRNA Decay and Translation. Trends in Genetics. 2013; 29(12):691–699. https://doi.org/10.1016/j.tig.2013.09.002 PMID: 24091060

9. Heck AM, Wilusz J. The Interplay between the RNA Decay and Translation Machinery in Eukaryotes. Cold Spring Harb Perspect Biol. 2018; 10(5):a032839. https://doi.org/10.1101/cshperspect.a032839 PMID: 29311343

10. Shoemaker CJ, Green R. Translation Drives mRNA Quality Control. Nat Struct Mol Biol. 2012; 19(6):594–601. https://doi.org/10.1038/nsmb.2301 PMID: 22664987

11. Radhakrishnan A, Green R. Connections Underlying Translation and mRNA Stability. Journal of Molecular Biology. 2016; 428(18):3558–3564. https://doi.org/10.1016/j.jmb.2016.05.025 PMID: 27261255

12. Presnyak V, Alhusaini N, Chen YH, Martin S, Morris N, Kline N, et al. Codon Optimality Is a Major Determinant of mRNA Stability. Cell. 2015; 160(6):1111–1124. https://doi.org/10.1016/j.cell.2015.02.029 PMID: 25768907

13. Radhakrishnan A, Chen YH, Martin S, Alhusaini N, Green R, Coller J. The DEAD-Box Protein Dhh1p Couples mRNA Decay and Translation by Monitoring Codon Optimality. Cell. 2016; 167(1):122–132. e9. https://doi.org/10.1016/j.cell.2016.08.053 PMID: 27641505

14. Webster MW, Chen YH, Stowell JAW, Alhusaini N, Sweet T, Graveley BR, et al. mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases. Molecular Cell. 2018; 70(6):1089–1100. https://doi.org/10.1016/j.molcel.2018.05.033 PMID: 29932902

15. Doma MK, Parker R. Endonucleolytic Cleavage of Eukaryotic mRNAs with Stalls in Translation Elongation. Nature. 2006; 440(7083):561–564. https://doi.org/10.1038/nature04530 PMID: 16554824

16. Brandman O, Stewart-Ornstein J, Wong D, Larson A, Williams CC, Li GW, et al. A Ribosome-Bound Quality Control Complex Triggers Degradation of Nascent Peptides and Signals Translation Stress. Cell. 2012; 151(5):1042–1054. https://doi.org/10.1016/j.cell.2012.10.044 PMID: 23178123

17. Kuroha K, Akamatsu M, Dimitrova L, Ito T, Kato Y, Shirahige K, et al. Receptor for Activated C Kinase 1 Stimulates Nascent Polypeptide-dependent Translation Arrest. EMBO Rep. 2010; 11(12):956–961. https://doi.org/10.1038/embor.2010.169 PMID: 21072063

18. Bengtson MH, Joazeiro CAP. Role of a Ribosome-Associated E3 Ubiquitin Ligase in Protein Quality Control. Nature. 2010; 467(7314):470–473. https://doi.org/10.1038/nature09371 PMID: 20835226

19. Shoemaker CJ, Eyler DE, Green R. Dom34:Hbs1 Promotes Subunit Dissociation and Peptidyl-tRNA Drop-Off to Initiate No-Go Decay. Science. 2010; 330(6002):369–372. https://doi.org/10.1126/science.1192430 PMID: 20947765

20. Pisareva VP, Skabkin MA, Hellen CUT, Pestova TV, Pisarev AV. Dissociation by Pelota, Hbs1 and ABCE1 of Mammalian Vacant 80S Ribosomes and Stalled Elongation Complexes. EMBO J. 2011; 30 (9):1804–1817. https://doi.org/10.1038/emboj.2011.93 PMID: 21448132

21. Defenouillère Q, Yao Y, Mouaikel J, Namane A, Galopier A, Decourty L, et al. Cdc48-Associated Complex Bound to 60S Particles Is Required for the Clearance of Aberrant Translation Products. PNAS. 2013; 110(13):5046–5051. https://doi.org/10.1073/pnas.1221724110 PMID: 23479637

22. Shao S, von der Malsburg K, Hegde RS. Listerin-Dependent Nascent Protein Ubiquitination Relies on Ribosome Subunit Dissociation. Molecular Cell. 2013; 50(5):637–648. https://doi.org/10.1016/j.molcel.2013.04.015 PMID: 23685075

23. Verma R, Reichermeier KM, Burroughs AM, Oania RS, Reitsma JM, Aravind L, et al. Vms1 and ANKZF1 Peptidyl-tRNA Hydrolases Release Nascent Chains from Stalled Ribosomes. Nature. 2018; 557(7705):446–451. https://doi.org/10.1038/s41586-018-0022-5 PMID: 29632312

24. Kuroha K, Zinoviev A, Hellen CUT, Pestova TV. Release of Ubiquitinated and Non-Ubiquitinated Nascent Chains from Stalled Mammalian Ribosomal Complexes by ANKZF1 and Ptrh1. Molecular Cell. 2018; 72(2):286–302.e8. https://doi.org/10.1016/j.molcel.2018.08.022 PMID: 30244831

25. Yip MCJ, Keszei AFA, Feng Q, Chu V, McKenna MJ, Shao S. Mechanism for Recycling tRNAs on Stalled Ribosomes. Nat Struct Mol Biol. 2019; 26(5):343–349. https://doi.org/10.1038/s41594-019-0211-4 PMID: 31011209

26. Simms CL, Yan LL, Zaher HS. Ribosome Collision Is Critical for Quality Control during No-Go Decay. Molecular Cell. 2017; 68(2):361–373. https://doi.org/10.1016/j.molcel.2017.08.019 PMID: 28943311

27. von der Haar T. Mathematical and Computational Modelling of Ribosomal Movement and Protein Synthesis: An Overview. Comput Struct Biotechnol J. 2012; 1:e201204002. https://doi.org/10.5936/csbj.201204002 PMID: 24688632

28. Zur H, Tuller T. Predictive Biophysical Modeling and Understanding of the Dynamics of mRNA Translation and Its Evolution. Nucleic Acids Res. 2016; 44(19):9031–9049. https://doi.org/10.1093/nar/gkw764 PMID: 27591251

29. MacDonald CT, Gibbs JH, Pipkin AC. Kinetics of Biopolymerization on Nucleic Acid Templates. Biopolymers. 1968; 6(1):1–25. https://doi.org/10.1002/bip.1968.360060102 PMID: 5641411

30. Zia R, Dong J, Schmittmann B. Modeling Translation in Protein Synthesis with TASEP: A Tutorial and Recent Developments. J Stat Phys. 2011; 144(2):405–428.

31. Chu D, Kazana E, Bellanger N, Singh T, Tuite MF, von der Haar T. Translation Elongation Can Control Translation Initiation on Eukaryotic mRNAs. EMBO J. 2014; 33(1):21–34. https://doi.org/10.1002/embj.201385651 PMID: 24357599

32. Hersch SJ, Elgamal S, Katz A, Ibba M, Navarre WW. Translation Initiation Rate Determines the Impact of Ribosome Stalling on Bacterial Protein Synthesis. J Biol Chem. 2014; 289(41):28160–28171. https://doi.org/10.1074/jbc.M114.593277 PMID: 25148683

33. Elgamal S, Katz A, Hersch SJ, Newsom D, White P, Navarre WW, et al. EF-P Dependent Pauses Integrate Proximal and Distal Signals during Translation. PLoS Genet. 2014; 10(8):e1004553. https://doi.org/10.1371/journal.pgen.1004553 PMID: 25144653

34. Woolstenhulme CJ, Guydosh NR, Green R, Buskirk AR. High-Precision Analysis of Translational Pausing by Ribosome Profiling in Bacteria Lacking EFP. Cell Rep. 2015; 11(1):13–21. https://doi.org/10.1016/j.celrep.2015.03.014 PMID: 25843707

35. Ferrin MA, Subramaniam AR. Kinetic Modeling Predicts a Stimulatory Role for Ribosome Collisions at Elongation Stall Sites in Bacteria. eLife. 2017; 6:e23629. https://doi.org/10.7554/eLife.23629 PMID: 28498106

36. Juszkiewicz S, Chandrasekaran V, Lin Z, Kraatz S, Ramakrishnan V, Hegde RS. ZNF598 Is a Quality Control Sensor of Collided Ribosomes. Mol Cell. 2018; 72(3):469–481. https://doi.org/10.1016/j.molcel.2018.08.037 PMID: 30293783

37. Ikeuchi K, Tesina P, Matsuo Y, Sugiyama T, Cheng J, Saeki Y, et al. Collided Ribosomes Form a Unique Structural Interface to Induce Hel2-driven Quality Control Pathways. EMBO J. 2019; 38: e100276. https://doi.org/10.15252/embj.2018100276 PMID: 30609991

38. Dvir S, Velten L, Sharon E, Zeevi D, Carey LB, Weinberger A, et al. Deciphering the Rules by Which 5′-UTR Sequences Affect Protein Expression in Yeast. PNAS. 2013; 110(30):E2792–E2801. https://doi.org/10.1073/pnas.1222534110 PMID: 23832786

39. Gutierrez E, Shin BS, Woolstenhulme CJ, Kim JR, Saini P, Buskirk AR, et al. eIF5A Promotes Translation of Polyproline Motifs. Molecular Cell. 2013; 51(1):35–45. https://doi.org/10.1016/j.molcel.2013.04.021 PMID: 23727016

40. Schuller AP, Wu CCC, Dever TE, Buskirk AR, Green R. eIF5A Functions Globally in Translation Elongation and Termination. Molecular Cell. 2017; 66(2):194–205. https://doi.org/10.1016/j.molcel.2017.03.003 PMID: 28392174

41. Dimitrova LN, Kuroha K, Tatematsu T, Inada T. Nascent Peptide-Dependent Translation Arrest Leads to Not4p-Mediated Protein Degradation by the Proteasome. J Biol Chem. 2009; 284(16):10343–10352. https://doi.org/10.1074/jbc.M808840200 PMID: 19204001

42. Ikemura T. Codon Usage and tRNA Content in Unicellular and Multicellular Organisms. Mol Biol Evol. 1985; 2(1):13–34. https://doi.org/10.1093/oxfordjournals.molbev.a040335 PMID: 3916708

43. Tuller T, Waldman YY, Kupiec M, Ruppin E. Translation Efficiency Is Determined by Both Codon Bias and Folding Energy. Proc Natl Acad Sci USA. 2010; 107(8):3645–3650. https://doi.org/10.1073/pnas.0909910107 PMID: 20133581

44. Shah P, Ding Y, Niemczyk M, Kudla G, Plotkin JB. Rate-Limiting Steps in Yeast Protein Translation. Cell. 2013; 153(7):1589–1601. https://doi.org/10.1016/j.cell.2013.05.049 PMID: 23791185

45. Subramaniam AR, Zid BM, O'Shea EK. An Integrated Approach Reveals Regulatory Controls on Bacterial Translation Elongation. Cell. 2014; 159(5):1200–1211. https://doi.org/10.1016/j.cell.2014.10.043 PMID: 25416955

46. Ingolia NT, Ghaemmaghami S, Newman JRS, Weissman JS. Genome-Wide Analysis in Vivo of Translation with Nucleotide Resolution Using Ribosome Profiling. Science. 2009; 324(5924):218–223. https://doi.org/10.1126/science.1168978 PMID: 19213877

47. Cao D, Parker R. Computational Modeling of Eukaryotic mRNA Turnover. RNA. 2001; 7(9):1192–1212. https://doi.org/10.1017/s1355838201010330 PMID: 11565744

48. Zarai Y, Margaliot M, Tuller T. A Deterministic Mathematical Model for Bidirectional Excluded Flow with Langmuir Kinetics. PLoS ONE. 2017; 12(8):e0182178. https://doi.org/10.1371/journal.pone.0182178 PMID: 28832591

49. Bonnin P, Kern N, Young NT, Stansfield I, Romano MC. Novel mRNA-Specific Effects of Ribosome Drop-off on Translation Rate and Polysome Profile. PLoS Comput Biol. 2017; 13(5):e1005555. https://doi.org/10.1371/journal.pcbi.1005555 PMID: 28558053

50. Shaw LB, Zia RKP, Lee KH. Totally Asymmetric Exclusion Process with Extended Objects: A Model for Protein Synthesis. Phys Rev E. 2003; 68(2):021910.

51. Chou T, Lakatos G. Clustered Bottlenecks in mRNA Translation and Protein Synthesis. Phys Rev Lett. 2004; 93(19):198101. https://doi.org/10.1103/PhysRevLett.93.198101 PMID: 15600884

52. Guydosh NR, Green R. Translation of Poly(A) Tails Leads to Precise mRNA Cleavage. RNA. 2017; 23:749–761. https://doi.org/10.1261/rna.060418.116 PMID: 28193672

53. Guydosh NR, Kimmig P, Walter P, Green R. Regulated Ire1-Dependent mRNA Decay Requires No-Go mRNA Degradation to Maintain Endoplasmic Reticulum Homeostasis in S. Pombe. eLife Sciences. 2017; 6:e29216.

54. Passos DO, Doma MK, Shoemaker CJ, Muhlrad D, Green R, Weissman J, et al. Analysis of Dom34 and Its Function in No-Go Decay. MBoC. 2009; 20(13):3025–3032. https://doi.org/10.1091/mbc.E09-01-0028 PMID: 19420139

55. Sundaramoorthy E, Leonard M, Mak R, Liao J, Fulzele A, Bennett EJ. ZNF598 and RACK1 Regulate Mammalian Ribosome-Associated Quality Control Function by Mediating Regulatory 40S Ribosomal Ubiquitylation. Molecular Cell. 2017; 65(4):751–760.e4. https://doi.org/10.1016/j.molcel.2016.12.026 PMID: 28132843

56. Juszkiewicz S, Hegde RS. Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination. Molecular Cell. 2017; 65(4):743–750.e4. https://doi.org/10.1016/j.molcel.2016.11.039 PMID: 28065601

57. Sitron CS, Park JH, Brandman O. Asc1, Hel2, and Slh1 Couple Translation Arrest to Nascent Chain Degradation. RNA. 2017; 23(5):798–810. https://doi.org/10.1261/rna.060897.117 PMID: 28223409

58. Matsuo Y, Ikeuchi K, Saeki Y, Iwasaki S, Schmidt C, Udagawa T, et al. Ubiquitination of Stalled Ribosome Triggers Ribosome-Associated Quality Control. Nat Commun. 2017; 8(1):159. https://doi.org/10.1038/s41467-017-00188-1 PMID: 28757607

59. Tsuboi T, Kuroha K, Kudo K, Makino S, Inoue E, Kashima I, et al. Dom34:Hbs1 Plays a General Role in Quality-Control Systems by Dissociation of a Stalled Ribosome at the 3′ End of Aberrant mRNA. Molecular Cell. 2012; 46(4):518–529. https://doi.org/10.1016/j.molcel.2012.03.013 PMID: 22503425

60. Verma R, Oania RS, Kolawa NJ, Deshaies RJ. Cdc48/P97 Promotes Degradation of Aberrant Nascent Polypeptides Bound to the Ribosome. eLife. 2013; 2:e00308. https://doi.org/10.7554/eLife.00308 PMID: 23358411

61. Ikeuchi K, Inada T. Ribosome-Associated Asc1/RACK1 Is Required for Endonucleolytic Cleavage Induced by Stalled Ribosome at the 3′ End of Nonstop mRNA. Sci Rep. 2016; 6:28234. https://doi.org/10.1038/srep28234 PMID: 27312062

62. Weinberg DE, Shah P, Eichhorn SW, Hussmann JA, Plotkin JB, Bartel DP. Improved Ribosome-Footprint and mRNA Measurements Provide Insights into Dynamics and Regulation of Yeast Translation. Cell Reports. 2016; 14(7):1787–1799. https://doi.org/10.1016/j.celrep.2016.01.043 PMID: 26876183

63. Thompson MK, Rojas-Duran MF, Gangaramani P, Gilbert WV. The Ribosomal Protein Asc1/RACK1 Is Required for Efficient Translation of Short mRNAs. eLife Sciences. 2016; 5:e11154.

64. Winz ML, Peil L, Turowski TW, Rappsilber J, Tollervey D. Molecular Interactions between Hel2 and RNA Supporting Ribosome-Associated Quality Control. Nat Commun. 2019; 10(1):563. https://doi.org/10.1038/s41467-019-08382-z PMID: 30718516

65. Kouba T, Rutkai E, Karásková M, Valášek LS. The eIF3c/NIP1 PCI Domain Interacts with RNA and RACK1/ASC1 and Promotes Assembly of Translation Preinitiation Complexes. Nucleic Acids Res. 2012; 40(6):2683–2699. https://doi.org/10.1093/nar/gkr1083 PMID: 22123745

66. Simms CL, Kim KQ, Yan LL, Qiu J, Zaher HS. Interactions between the mRNA and Rps3/uS3 at the Entry Tunnel of the Ribosomal Small Subunit Are Important for No-Go Decay. PLoS Genet. 2018; 14 (11):e1007818. https://doi.org/10.1371/journal.pgen.1007818 PMID: 30475795

67. Deneke C, Lipowsky R, Valleriani A. Effect of Ribosome Shielding on mRNA Stability. Phys Biol. 2013; 10(4):046008. https://doi.org/10.1088/1478-3975/10/4/046008 PMID: 23883670

68. Dao Duc K, Saleem ZH, Song YS. Theoretical Analysis of the Distribution of Isolated Particles in Totally Asymmetric Exclusion Processes: Application to mRNA Translation Rate Estimation. Phys Rev E. 2018; 97(1):012106.

69. Duc KD, Song YS. The Impact of Ribosomal Interference, Codon Usage, and Exit Tunnel Interactions on Translation Elongation Rate Variation. PLoS Genet. 2018; 14(1):e1007166. https://doi.org/10.1371/journal.pgen.1007166 PMID: 29337993

70. Diament A, Feldman A, Schochet E, Kupiec M, Arava Y, Tuller T. The Extent of Ribosome Queuing in Budding Yeast. PLoS Comput Biol. 2018; 14(1):e1005951. https://doi.org/10.1371/journal.pcbi.1005951 PMID: 29377894

71. Reuveni S, Meilijson I, Kupiec M, Ruppin E, Tuller T. Genome-Scale Analysis of Translation Elongation with a Ribosome Flow Model. PLoS Comput Biol. 2011; 7(9):e1002127. https://doi.org/10.1371/journal.pcbi.1002127 PMID: 21909250

72. Sabi R, Tuller T. Computational Analysis of Nascent Peptides That Induce Ribosome Stalling and Their Proteomic Distribution in Saccharomyces Cerevisiae. RNA. 2017; 23(7):983–994. https://doi.org/10.1261/rna.059188.116 PMID: 28363900

73. Gietz RD, Schiestl RH. High-Efficiency Yeast Transformation Using the LiAc/SS Carrier DNA/PEG Method. Nat Protocols. 2007; 2(1):31–34. https://doi.org/10.1038/nprot.2007.13 PMID: 17401334

74. Wach A, Brachat A, Pöhlmann R, Philippsen P. New Heterologous Modules for Classical or PCR-based Gene Disruptions in Saccharomyces Cerevisiae. Yeast. 1994; 10(13):1793–1808. PMID: 7747518

75. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, Smith HO. Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. Nat Meth. 2009; 6(5):343–345.

76. Sambrook J. Molecular Cloning: A Laboratory Manual, Third Edition. 3rd ed. Woodbury, NY: Cold Spring Harbor Laboratory Press; 2001.

77. Hughes AL, Gottschling DE. An Early Age Increase in Vacuolar pH Limits Mitochondrial Function and Lifespan in Yeast. Nature. 2012; 492(7428):261–265. https://doi.org/10.1038/nature11654 PMID: 23172144

78. Hansen AS, O'Shea EK. Promoter Decoding of Transcription Factor Dynamics Involves a Trade-off between Noise and Control of Gene Expression. Mol Syst Biol. 2013; 9(1):704.

79. Sikorski RS, Hieter P. A System of Shuttle Vectors and Yeast Host Strains Designed for Efficient Manipulation of DNA in Saccharomyces Cerevisiae. Genetics. 1989; 122(1):19–27. PMID: 2659436

80. Lee S, Lim WA, Thorn KS. Improved Blue, Green, and Red Fluorescent Protein Tagging Vectors for S. Cerevisiae. PLoS ONE. 2013; 8(7):e67902. https://doi.org/10.1371/journal.pone.0067902 PMID: 23844123

81. Lopez CF, Muhlich JL, Bachman JA, Sorger PK. Programming Biological Models in Python Using PySB. Mol Syst Biol. 2013; 9:646. https://doi.org/10.1038/msb.2013.1 PMID: 23423320

82. Harris LA, Hogg JS, Tapia JJ, Sekar JAP, Gupta S, Korsunsky I, et al. BioNetGen 2.2: Advances in Rule-Based Modeling. Bioinformatics. 2016; 32(21):3366–3368. https://doi.org/10.1093/bioinformatics/btw469 PMID: 27402907

83. Sneddon MW, Faeder JR, Emonet T. Efficient Modeling, Simulation and Coarse-Graining of Biological Complexity with NFsim. Nat Methods. 2011; 8(2):177–183. https://doi.org/10.1038/nmeth.1546 PMID: 21186362

84. Arava Y, Wang Y, Storey JD, Liu CL, Brown PO, Herschlag D. Genome-Wide Analysis of mRNA Translation Profiles in Saccharomyces Cerevisiae. PNAS. 2003; 100(7):3889–3894. https://doi.org/10.1073/pnas.0635171100 PMID: 12660367

85. Bonven B, Gulløv K. Peptide Chain Elongation Rate and Ribosomal Activity in Saccharomyces Cerevisiae as a Function of the Growth Rate. Mol Gen Genet. 1979; 170(2):225–230. https://doi.org/10.1007/bf00337800 PMID: 372763

86. Subtelny AO, Eichhorn SW, Chen GR, Sive H, Bartel DP. Poly(A)-Tail Profiling Reveals an Embryonic Switch in Translational Control. Nature. 2014; 508(7494):66–71. https://doi.org/10.1038/nature13007 PMID: 24476825

87. Köster J, Rahmann S. Snakemake—a Scalable Bioinformatics Workflow Engine. Bioinformatics. 2012; 28(19):2520–2522. https://doi.org/10.1093/bioinformatics/bts480 PMID: 22908215

88. Martin M. Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads. EMBnet-journal. 2011; 17(1):10–12.

89. Trapnell C, Pachter L, Salzberg SL. TopHat: Discovering Splice Junctions with RNA-Seq. Bioinformatics. 2009; 25(9):1105–1111. https://doi.org/10.1093/bioinformatics/btp120 PMID: 19289445

90. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map Format and SAMtools. Bioinformatics. 2009; 25(16):2078–2079. https://doi.org/10.1093/bioinformatics/btp352 PMID: 19505943

91. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and Memory-Efficient Alignment of Short DNA Sequences to the Human Genome. Genome Biol. 2009; 10(3):R25. https://doi.org/10.1186/gb-2009-10-3-r25 PMID: 19261174

92. Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, et al. Software for Computing and Annotating Genomic Ranges. PLoS Comput Biol. 2013; 9(8):e1003118. https://doi.org/10.1371/journal.pcbi.1003118 PMID: 23950696

93. Love MI, Huber W, Anders S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. Genome Biol. 2014; 15(12):550. https://doi.org/10.1186/s13059-014-0550-8 PMID: 25516281

94. Boehlke KW, Friesen JD. Cellular Content of Ribonucleic Acid and Protein in Saccharomyces Cerevisiae as a Function of Exponential Growth Rate: Calculation of the Apparent Peptide Chain Elongation Rate. J Bacteriol. 1975; 121(2):429–433. PMID: 1089627