Published in final edited form as:

Wiley Interdiscip Rev RNA. 2018 September ; 9(5): e1486. doi:10.1002/wrna.1486.

PolyA tracks, polybasic peptides, poly-translational hurdles

Laura L. Arthur¹, Sergej Djuranovic¹

¹Department of Cell Biology and Physiology, Washington University School of Medicine, 600 South Euclid Avenue, Campus Box 8228, St. Louis, MO 63110, USA

Abstract

The abundance of mRNA is one of the major determinants of protein synthesis. As such, factors that influence mRNA stability often contribute to gene regulation. Polyadenylation of the 3' end of mRNA transcripts, the poly(A) tail, has long been recognized as one of these regulatory elements given its influence on translation efficiency and mRNA stability. Unwanted translation of the poly(A) tail signals to the cell an aberrant polyadenylation event or the lack of stop codons which makes this sequence an important element in translation fidelity and mRNA surveillance response. Consequently, investigations into the effects of the poly(A) tail lead to the discoveries that poly-lysine as well as other poly-basic peptide sequences and, to a much greater extent, polyA mRNA sequences within the open reading frame influence mRNA stability and translational efficiency. Conservation and evolutionary selection of codon usage in polyA track sequences across multiple organisms suggests a biological significance for coding polyA tracks in the regulation of gene expression. Here we discuss the cellular responses and consequences of coding polyA track translation and synthesis of polybasic peptides.

Graphical Abstract

Introduction

All organisms must be able to regulate gene expression and abundance to develop, reproduce, and respond to their environments. In the flow of information from DNA to RNA to protein, regulatory mechanisms have evolved to increase or decrease the rate of essentially every biochemical reaction. As the intermediate molecule in this process, the abundance and translational efficiency of mRNA plays a central role in the final abundance of protein. Changes to translation result in more rapid changes in protein expression

than changes to transcription. Consequently, the life cycle of mRNA and its translational efficiency is under strict regulation.

Much of the research into these post-transcriptional regulation mechanisms focuses on modulation of translation initiation and mRNA stability by *cis*-acting elements in the 5' and 3' untranslated regions (UTRs) and the *trans*-acting elements that interact with them. For example, the 5' terminal oligopyrimidine (TOP) motif consists of 4–15 pyrimidine nucleotides near the transcription start site that is bound by La-related protein 1 (LARP1) to stimulate translation in response to mammalian target of rapamycin (mTOR) to promote cell growth (Amaldi & Pierandrei-Amaldi, 1997). TOP motifs are found in most ribosomal proteins and may play a role in translational upregulation of these genes during the M phase of mitotic cell division (Park et al., 2016; Yamashita et al., 2008). Similarly, the Pumilio and FBF (PUF) family of proteins bind specific sequences in the 3' UTRs of many transcripts to decrease expression of those genes (Wickens, Bernstein, Kimble, & Parker, 2002). See references (Hinnebusch, Ivanov, & Sonenberg, 2016) and (Szostak & Gebauer, 2012) for a more in depth review of 5' and 3' UTR mediated regulation, respectively. While control of translation initiation and modulation of mRNA stability through binding of *trans*-acting elements play an important role in specifying the amount of translated protein products, the contribution of translation elongation in overall gene regulation has been largely overlooked. Over the last several years, however, regulation of the rate and fidelity of elongation has emerged as one of the key determinants of mRNA stability, translation efficiency, and thus protein abundance (Presnyak et al., 2015).

Here, we discuss a novel translation elongation regulatory mechanism mediated by short stretches of consecutive adenosine residues in the open reading frame (ORF) of an mRNA, termed polyA tracks. Coding polyA tracks are *cis*-acting elements that cause ribosomal stalling and programmed ribosomal frameshifting, decreasing mRNA stability and reducing protein synthesis. PolyA tracks are found endogenously in 1–2% of genes in most organisms, including humans, meaning a significant portion of the genome is potentially regulated by this mechanism (Arthur et al., 2015). The similarities of coding polyA tracks to both the poly(A) tail and poly-basic amino acid tracks warrant a brief discussion and comparison of these mechanisms as well.

Poly(A) tail mediated regulation

The poly(A) tail is a key regulator of gene expression

The most well-known polyadenylate sequence, the poly(A) tail, plays an essential role in the regulation of mRNA stability and translation. Nearly all eukaryotic mRNAs undergo 3' end cleavage and polyadenylation to generate the poly(A) tail. The poly(A) tail globally stimulates translation by binding cytosolic poly(A)-binding protein 1 (PABPC1), which interacts with eukaryotic initiation factor 4G (eIF4G) to form a pseudo-ring structure and promote efficient translation initiation (Tarun & Sachs, 1996; Tarun, Wells, Deardorff, & Sachs, 1997).

Poly(A) tail length can be dynamically regulated by cytoplasmic deadenylases and polyadenylases in a developmental, tissue specific, or cell cycle dependent manner to

influence translation efficiency. Many groups have shown that during early embryogenesis in *Xenopus* and *Drosophila* oocytes, translationally silent maternal mRNA with short poly(A) tails are activated by extension of their poly(A) tails (Wickens, 1990). More recently, work in HeLa cells showed that the poly(A) tail on a subset of transcripts is shortened or extended during the M phase of the mitotic cell cycle. Transcripts with poly(A) tails shortened to fewer than 20 nt were less efficiently translated, possibly because of the reduced ability to bind PABPC1 (Park et al., 2016). Analysis of transcripts with poly(A) tail lengths shorter than 30 nt consistently finds that they are translationally silenced and destabilized (Chorghade et al., 2017; Lima et al., 2017; Park et al., 2016). However, studies on the translation of transcripts with poly(A) tails longer than 30 nt have come to conflicting conclusions. Initial studies have found a positive correlation of tail length to translational efficiency generally (Weill, Belloc, Bava, & Méndez, 2012) or only in embryonic cells (Chang, Lim, Ha, & Kim, 2014; Park et al., 2016; Subtelny, Eichhorn, Chen, Sive, & Bartel, 2014). More recently, contrasting work shows a negative correlation of tail length to translation efficiency was found in *C. elegans* (Lima et al., 2017). These discrepancies may be due to different experimental design or reflect a complexity of regulation that is cell type and developmentally dependent.

In yeast, transcripts that lack a poly(A) tail, such as those cleaved by an endonuclease, are rapidly degraded by the cytoplasmic exosome (Meaux & Hoof, 2006). Additionally, transcripts lacking a poly(A) tail, with or without a stop codon, are translated with lower efficiency than those that are polyadenylated (Meaux & Hoof, 2006). These observations are consistent with the role of the poly(A) tail in stimulating translation and indicates that the lack of poly(A) tail signals to the cell that the transcript is aberrant.

Alternative polyadenylation generates transcripts with alternative 3' coding sequences or 3' untranslated regions (UTRs) (Tian & Manley, 2017). This regulated process can influence the expression of the transcript by including or excluding protein binding sites in the alternative 3' UTR. Yet another way that alternative polyadenylation can influence expression is by placing the poly(A) tail before the canonical stop codon, effectively creating a transcript with no in-frame stop codon, also known as a nonstop transcript.

Translation of the poly(A) tail induces mRNA and nascent protein decay

Typically, the ribosome does not encounter the poly(A) tail since it is found downstream of the termination codon (Figure 1a). Aberrant or premature polyadenylation (APA) or mutations that eliminate the stop codon, however, allow the ribosome to continue translation into the poly(A) tail (Figure 1b). In fact, premature polyadenylation appears to occur in approximately 1% of yeast and human transcripts (Frischmeyer et al., 2002; Ozsolak et al., 2010). These nonstop transcripts are less stable and produce less protein than properly polyadenylated transcripts (Frischmeyer et al., 2002; van Hoof, Frischmeyer, Dietz, & Parker, 2002), but the molecular mechanisms underlying this instability and the role, if any, the poly(A) tail plays in it was unclear until recently. More recent studies have identified a role for ribosome associated quality control in eliminating those sequences, as discussed below. Furthermore, these studies revealed new mechanisms of ribosome mediated gene regulation, including regulation by short, coding polyA tracks.

Eukaryotic mRNAs lacking a stop codon are rapidly decayed by the translation dependent mRNA surveillance pathway nonstop decay (NSD). This pathway was initially characterized in yeast and requires the exosome complex of 3' to 5' exoribonucleases involved in general cytoplasmic mRNA decay (Frischmeyer et al., 2002; van Hoof et al., 2002). It is distinct from general mRNA decay because it does not require deadenylation prior to degradation and involves endonuclease activity of the exosome, rather than just the exonuclease activity (Frischmeyer et al., 2002; van Hoof et al., 2002). A cofactor of the yeast exosome, Ski7, recognizes ribosomes stalled on nonstop transcripts. This function requires the GTPase C-terminal domain of Ski7 in a manner homologous to the GTPase domain of eRF3 recognizing the stop codon of a normal transcript (van Hoof et al., 2002). Since Ski7 tightly binds to the exosome, it is recruited to transcripts recognized as nonstop (van Hoof et al., 2002). The C-terminal domain of Ski7 is the only known factor that is specific to the NSD pathway. Ski7, however, is found only in some species of *Saccharomyces*. A homologous protein, HBS1L, was found to function in the NSD pathway in other eukaryotes (Hashimoto, Takahashi, Sakota, & Nakamura, 2017; S. Saito, Hosoda, & Hoshino, 2013). Yeast Hbs1 protein (human Hbs1L), together with Dom34 (human Pelota), was also found to be one of the mediators of another mRNA surveillance pathway, termed No-go decay (NGD) (Doma & Parker, 2006).

NGD is an mRNA surveillance pathway that detects ribosomes stalled by mRNA or nascent protein features other than the poly(A) tail (Figure 1c). Stable secondary structure in the mRNA, rare codons, and chemical damage to the mRNA have all been shown to induce the NGD pathway (Doma & Parker, 2006; Gandhi, Manzoor, & Hudak, 2008). Stalled ribosomes are recognized by the Hbs1:Dom34 complex which promotes ribosome subunit dissociation from the mRNA to resolve the stall (Doma & Parker, 2006; Christopher J Shoemaker, Eyler, & Green, 2010). Complete ribosome recycling is enabled by action of Rli (human ABCE1) a highly conserved ABC-type ATPase (Pisareva, Skabkin, Hellen, Pestova, & Pisarev, 2011; C. J. Shoemaker & Green, 2011). Endonucleolytic cleavage of the stall-inducing mRNA by an unknown endonuclease and subsequent ribosome recycling are followed by further degradation of cleaved transcript using the canonical mRNA decay pathways with the exonuclease activity of exosome in the 3' to 5' direction and Xrn1 in the 5' to 3' direction (Doma & Parker, 2006).

Aberrant mRNA transcripts induce both NSD and NGD during translation while nascent peptides are synthesized. Like the destabilization of aberrant mRNA transcripts by NSD and NGD, the nascent protein must also be degraded to protect the cell from the potential harmful effects of aberrant or truncated proteins. The ribosome-associated quality control (RQC) complex, originally discovered and characterized in yeast, fulfills this role by targeting the nascent peptide to the proteasome for degradation (Brandman et al., 2012; Defenouillère et al., 2013; Yonashiro et al., 2016). The RQC consists of four known components - Ltn1p, Rqc1p, Rqc2p, and Cdc48p that act together to identify and target the stalled nascent peptide for degradation. Briefly, the Hbs1:Dom34 complex dissociates the large and small subunits of the ribosome without hydrolyzing the peptidyl-tRNA. The nascent peptide remains associated with the 60S subunit of the ribosome (Christopher J Shoemaker et al., 2010), creating a distinct substrate compared to normal translation termination in which eRF1 hydrolyses the peptidyl-tRNA and triggers its dissociation

from the 60S subunit. Rqc1 recognizes the 60S subunit associated with peptidyl-tRNA, and this complex promotes the association of the E3 ligase, Ltn1p (yeast homologue of mammalian Listerin), with the ribosome near the exit channel and stalled peptide. Ltn1p then ubiquitylates the stalled nascent peptide, targeting it to Cdc48p and subsequently to the proteasome (Bengtson & Joazeiro, 2010). Activation of the RQC pathway removes the peptidyl-tRNA from the 60S subunit to complete ribosome recycling and ensures the degradation of nascent peptides translated from aberrant, nonstop transcripts.

Evidence for ribosomal stalling by polybasic sequences

Translation of the poly(A) tail results in the synthesis of a poly-lysine sequence at the C-terminus of the nascent peptide. This poly-lysine tail was predicted to be the cause of ribosome stalling events observed on poly(A) tails. In fact, research into the poly(A) tail mediated activation of NSD and NGD, as well as studies on the RQC pathway, led to the discovery that clusters of basic amino acids, either lysine or arginine, can act as a stalling sequence even when encoded before the stop codon. It was experiments in yeast, prompted by the above question of non-stop transcripts, which revealed that polybasic sequences of greater than 10 residues in length in the coding sequence are sufficient to cause reduction in protein abundance (Ito-Harashima, Kuroha, Tatematsu, & Inada, 2007). In vitro studies, using rabbit reticulocyte lysates indicated that runs of positively charged lysine or arginine residues impede translation rates through electrostatic interactions of positively charged peptide with the negatively charged ribosome exit channel (Lu & Deutsch, 2008). As such, translation of the polybasic sequence are a cause for ribosomal stalling even in the context outside of poly(A) tail translation, targeting sequences found in endogenous genes in yeast (Dimitrova, Kuroha, Tatematsu, & Inada, 2009) Thus, stretches of poly-basic residues within protein coding genes are yet another mechanism that controls protein abundance.

Cellular factors involved in resolution of polybasic stalling

Research in several labs verified the regulatory importance of polybasic stretches. Brandman *et al.* identified putative RQC substrates with six or more basic amino acids in a stretch of 10. Notably, one of the putative RQC substrates identified in that study was Rqc1, which contains a well conserved polybasic region. Mutation analysis revealed that this polybasic region serves as an autoregulatory element. When Rqc1 expression is high, the RQC pathway is highly active and constrains expression of Rqc1 transcripts through its polybasic track (Brandman *et al.*, 2012). These polybasic stretches appear to trigger both transcript and nascent peptide degradation consistent with stalling on the poly(A) tail: via activation of mRNA surveillance and the RQC.

Several groups recently identified and further characterized mammalian homologs of yeast RQC factors (Garzia *et al.*, 2017; Juskiewicz & Hegde, 2017; Sundaramoorthy *et al.*, 2017). ZNF598 is a human RING domain protein that shares homology with Hel2, a yeast RING domain-containing protein. Deletion of Hel2 in yeast promotes read-through of polybasic sequences (Brandman *et al.*, 2012; Juskiewicz & Hegde, 2017). Knockdown of ZNF598 in a terminal stalling assay increased read-through of a sequence of at least 20 AAA codons to produce full length reporter proteins (Juskiewicz & Hegde, 2017; Sundaramoorthy *et*

al., 2017). The increase in full length protein rather than an increase in truncated protein fragments suggests that the ribosomes did not terminally stall in the absence of ZNF598. A similar experiment in ZNF598 knockout HEK cells finds consistent ZNF598-dependent repression of polyA reporters (Garzia et al., 2017). The N-terminal RING domain of ZNF598 acts as an E3 ligase that was shown to ubiquitylate RPS20, RPS10, and RPS3 (Garzia et al., 2017; Sundaramoorthy et al., 2017). These ubiquitylation events are necessary to promote terminal stalling on consecutive AAA codons.

A genetic screen in *S. cerevisiae* revealed that mutations in ASC1, the yeast homolog of human RACK1, increased expression of transcripts containing 12 consecutive arginine or lysine codons in a translation dependent manner (Kuroha et al., 2010). RACK1 is a highly conserved component of the eukaryotic 40S subunit that acts as a scaffold protein in a wide variety of signal transduction pathways (Gerbasi, Weaver, Hill, Friedman, & Link, 2004). Knockdown of RACK1 in human cells also increased expression of transcripts with consecutive AAA lysine codons (Sundaramoorthy et al., 2017). RACK1 facilitates ubiquitylation of certain ribosomal proteins in response to translation of consecutive AAA codons, including ZNF598 targets RPS3, RPS10 and RPS20. Since double knockdown of RACK1 and ZNF598 did not have an additive effect, it seems that these proteins act in the same pathway to promote stalling on consecutive AAA codons. The same conclusion was reached for arginine CGA repeats, albeit without a focus on ubiquitination of ribosomal proteins, in yeast studies following deletion of Asc1 and Hel2 (Letzring, Wolf, Brule, & Grayhack, 2013; K. Saito, Horikawa, & Ito, 2015). The discovery that knockdown of ZNF598 and RACK1 can promote read-through of consecutive lysine codons indicates that the ribosome does not inherently stall, but must be induced to stall by ZNF598-mediated and RACK1-assisted ubiquitylation of ribosomal proteins (Garzia et al., 2017; Juszkievicz & Hegde, 2017; Sundaramoorthy et al., 2017). After the ribosome is ubiquitylated by ZNF598 and RACK1, it is terminally stalled and recognized by the Hbs1:Dom34 complex which splits the ribosomal subunits (Tsuboi et al., 2012). Nuclear export mediator factor (NEMF, mammalian homologue of Rqc2p), stabilizes Listerin, the mammalian homologue of yeast Ltn1p, which ubiquitylates the nascent peptide to mark it for degradation by the proteasome, completing the RQC pathway (Shao, Brown, Santhanam, & Hegde, 2015) (list of all mentioned factors involved in NSD, NGD and polybasic stalling is attached to Figure 1).

Note that characterization of RQC factors in mammals cells relied on poly-lysine reporters encoded exclusively with AAA codons, since AAG lysine and poly arginine sequences did not cause efficient stalling in human cell culture. This distinction raises the possibility that stalling due to polybasic sequences in the nascent peptide does not affect mammalian ribosomes as strongly as yeast ribosomes and that consecutive AAA codons may be stalling the ribosome in a different manner. It is also possible that the nature of the reporters used in these studies accounts for the difference in stalling potentials. Both Hegde and Bennett's group used a dual fluorescence stalling system in which GFP and RFP are separated by two viral 2A peptide sequences flanking the stalling sequence (Juszkievicz & Hegde, 2017; Sundaramoorthy et al., 2017). The 2A peptides are great tool for expression of two or more proteins from one ORF (Ryan & Drew, 1994), however, 2A peptides also induce ribosome stalling (Doronina et al., 2008). Considering conflicting reports on the

importance of protein elongation (eEF2) and release factors (eRF1 and eRF3) in “stop-carry on” translation by 2A peptides, one should be cautious on how 2A peptides impact results, complicate analysis, and change overall context of tested polybasic sequences (Doronina et al., 2008; Machida et al., 2014; Roulston et al., 2016). A study that uses direct insertion of polybasic sequences between two reporter proteins has indicated ZNF598 contributing to overall protein quality control during stalling at consecutive AAA codons (Garzia et al., 2017). A recent study in yeast argues for a ribosome collision model as a trigger for ubiquitination of RPS3 and NGD response during ribosomal stalling on a different set of stalling sequences (Simms, Yan, & Zaher, 2017). This study implies different outcomes for mRNA stability coming from stalling sequences at different locations along reporter coding sequence. Given these various reports, it is clear that organismal and reporter differences will play a big role in future studies of polybasic- and polyA-ribosomal stalling.

Coding polyA sequences repress translation more than polybasic sequences

Codon usage in poly-lysine tracks show an underrepresentation of the AAA codon

Sequences of six or more consecutive polybasic amino acids, lysine and arginine, are underrepresented in the transcriptome of multiple organisms compared to runs of other amino acids, suggesting a selective pressure against polybasic amino acid sequences (Karlin, Brocchieri, Bergman, Mrazek, & Gentles, 2002). Further analysis of codon usage within poly-lysine peptides, however, reveals that there is lower than expected frequency of AAA codons compared to AAG codons in runs of four consecutive lysine residues in the coding regions of genes (Arthur et al., 2015). Similar discrepancies in lysine codon usage patterns are present in over 150 prokaryotic and eukaryotic genomes that have been analyzed (Habich, Djuranovic, & Szczesny, 2016). Interestingly, when a poly-lysine track does contain several consecutive AAA codons, orthologous proteins across species have comparable lysine codon usage for these tracks (Arthur et al., 2015; Koutmou et al., 2015). This apparent conservation and evolutionary selection of codon usage in such sequences suggests a functional significance for coding polyA tracks.

PolyA is the most efficient stalling sequence

Careful biochemical analysis of poly-lysine stalling sequences in both prokaryotic and eukaryotic systems revealed that consecutive AAA lysine codons exhibit a greater delay in ribosome movement than an equivalent number of AAG lysine codons (Arthur et al., 2015, 2017; Koutmou et al., 2015). Reporters with consecutive AAA codons produce significantly less protein than those with AAG, despite coding for the same polybasic amino acid sequence. This discordance in translation likely does not result from differences in tRNA abundance or binding affinity, as there are multiple cognate and isodecoder tRNAs for lysine codons in higher organisms. *E. coli* uses a single tRNA for both lysine codons (Chan & Lowe, 2009), a UUU* modified anti-codon for both AAA and AAG, but still shows reduced translational efficiency for iterated AAA codons (Koutmou et al., 2015). The presence of this phenomenon in *E. coli* similarly discounts the possibility that interactions with poly(A)-binding protein (PABP) decrease translation efficiency of consecutive AAA

codons, as PABP is not found in prokaryotes (Mangus, Evans, & Jacobson, 2003). Instead, it argues for the possibility that the interactions between the ribosome and the mRNA sequence itself cause the difference in translational efficiency. The role of the lysine residues was examined by kinetic analysis of lysine incorporation in an *in vitro* translation system which showed that the rate of second lysine addition in an iterated sequence of either AAA or AAG codons is slower than typical elongation rates of heteropolymeric sequences. Consecutive AAA codons, however, slowed incorporation to a greater extent than AAG codons (Koutmou et al., 2015). Together, these studies suggest that the polyA sequence itself contributes to decreased protein expression, but whether the lysine residues translated from the mRNA sequence facilitates or enables this decrease is unknown.

There are conflicting reports about the effect of polybasic stalling in human tissue culture, but consistent evidence suggests that polyA tracks are a potent negative regulator of expression. In two recent reports, HEK293 cell line ribosomes are resistant to stalling on many of the stalling sequences characterized in yeast, including poly arginine and selected stem loop structures (Juszkiewicz & Hegde, 2017; Sundaramoorthy et al., 2017). Even poly-lysine reporters encoded by as many as 20 AAG codons were reported to have little to no decrease in protein expression compared to no-insert controls (Juszkiewicz & Hegde, 2017; Sundaramoorthy et al., 2017). Early investigations of the difference between AAG vs AAA induced stalling in human dermal fibroblasts found that 12 iterated AAG codons were sufficient to induce some mRNA instability and decreased protein expression in tissue culture (Arthur et al., 2015), consistent with several studies of polybasic sequences in yeast (Bengtson & Joazeiro, 2010; Brandman et al., 2012; Dimitrova et al., 2009; Ito-Harashima et al., 2007). In all studies, however, the coding polyA sequence mediates more potent translational regulation across many species with as few as four consecutive AAA codons reducing protein expression by 50% of control constructs (Arthur et al., 2015, 2017; Juszkiewicz & Hegde, 2017; Sundaramoorthy et al., 2017). These studies indicate that the stalling effect is specific for lysine residues encoded by polyA sequences rather than for poly-lysine stretches in general (Figure 2). Four AAG codons had no effect on protein expression in any reports. Aggregate ribosome profiling data show that in a sequence of four consecutive lysine codons, with three or more AAA codons, one can see increased ribosomal occupancy, indicative of increased ribosomal stalling. The same effect is not seen with four lysine tracks with fewer than three AAA codons (Arthur et al., 2015). The polyA induced stalling likely triggers activation of mRNA surveillance and the RQC pathway to decrease mRNA stability and protein expression in a manner similar to polybasic induced stalling (Arthur et al., 2015).

PolyA sequences cause frameshifting in addition to stalling

Investigations of the stalling potential of consecutive AAA codons revealed that the ribosome can also frameshift while elongating over the polyA track (Arthur et al., 2015; Koutmou et al., 2015) (Figure 2). In an *in vitro* translation system, as few as six consecutive adenosine residues were sufficient to cause a change in translation frame and the addition of an extra lysine residue in the nascent peptide when programmed reactions lacked either downstream charged-tRNAs or termination factors associated with stop codons (Koutmou et al., 2015). When all factors are present, truncated products that result from frameshifting

were detected *in vitro* with a minimum of three consecutive AAA codons (Koutmou et al., 2015). Ribosomal frameshifting was also seen with reporters expressing human genes with endogenous polyA tracks (Arthur et al., 2015). This behavior is surprising because frameshifting typically requires the presence of a downstream stimulatory element such as a pseudoknot or stem loop structures that slow ribosome movement on mRNAs (Ketteler, 2012). It is possible that slowed incorporation of lysine residues in consecutive tracks may stall the ribosome sufficiently to allow frameshifting to occur, substituting for physical stall introduced by pseudoknot structure typically seen in programmed frameshifting.

When the ribosome frameshifts on a polyA track, it will continue the elongation cycle out of frame and likely encounter a premature termination codon downstream (Figure 3a). Like other programmed ribosomal frameshifting events, this can lead to the activation of nonsense mediated decay (NMD) which will result in degradation of the transcript and nascent peptide. Knockout of the major NMD factor, Upf1, in *S. cerevisiae* partially restores expression of reporters containing 12 AAA codons, demonstrating that polyA track transcripts are degraded by the NMD pathway in yeast (Koutmou et al., 2015). It is worth mentioning that mRNA levels of the reporter with polyA track in the second exon of beta-globin gene are notably similar to insertion of a stop codon at the same position (Arthur et al., 2015). The frameshifted or stalled product in this case would be efficiently degraded and was not observed. Moreover, no frameshifting has been reported for other polybasic stall sequences. In fact, it was shown that ribosomes do not normally frameshift on rare CGA codon repeats in wild type yeast even though they cause significant ribosomal stalling (Letzring et al., 2013). This characteristic is perhaps the reason that polyA tracks have a much stronger effect on protein expression and mRNA stability. PolyA tracks probably stall and elicit NGD to the same extent as other polybasic tracks, but then undergo additional destabilization by NMD because of frameshifting events.

RACK1, the ribosome associated protein discussed above for its role in stalling on AAA encoded poly-lysine tracks, is also known to promote frame maintenance during translation elongation. Knockdown of the yeast homolog of RACK1, Asc1, resulted in frameshifting on CGA codon repeats which stall in wild type yeast strains but do not frameshift (Wolf & Grayhack, 2015). If human RACK1 also functions to maintain the coding frame, then reduced levels would lead to more frameshifting, and thus, more degradation by NMD. However, as discussed above, it also allows more read through of the stall sequence and less degradation by NGD or NSD. The role of RACK1 in translation efficiency and possibly frame maintenance was further elucidated by a study of the effects of viral infection and translation initiation in human cells (Jha et al., 2017). The VacV DNA virus relies on host translation machinery to replicate and is able to selectively inhibit synthesis of host proteins (Dai et al., 2017; Moss, 1968). During VacV infection, RACK1 is phosphorylated on an extended loop that contacts the 18S ribosomal RNA at the mRNA channel. Pulldown of a phosphomimic RACK1 mutant showed a reduced recovery of host mRNA that lack a polyA-leader and enhanced recovery of viral mRNA with the polyA-leader (Jha et al., 2017). Authors hypothesize that modification of RACK1 may regulate ribosome sliding across difficult polyA sequences in 5'UTRs and give advantage to viral mRNAs with polyA-leaders. While this study does not illuminate contribution of RACK1 to overall sliding and stalling at coding polyA tracks it clearly shows that RACK1 protein may have certain

preference for adenosine rich sequences. Taken together these studies indicate that it is likely that RACK1 (yeast Asc1) has complex effects on the expression and translational fidelity of transcripts with polyA tracks.

Endogenous coding polyA sequences regulate gene expression

Bioinformatic analysis of the human genome revealed that more than 450 genes contain a polyA track – defined as a sequence of 12 nucleotides in which at least 11 are adenosine (Arthur et al., 2015). A similar proportion of polyA track genes were found in other sequenced and analyzed genomes (Arthur et al., 2015; Habich et al., 2016). Since this definition of polyA tracks depends on the nucleotide sequence and not the amino acid sequence, it does not necessarily encode four consecutive lysine residues. Additionally, the interrupting nucleotide in the sequence can fall anywhere among the adenosines. It is not known how the exact composition of the polyA track in endogenous genes affects stalling and frameshifting efficiencies given the possibilities for mRNA modifications and alternative splicing of these genes. Insertion of gene-derived polyA tracks into a reporter significantly decrease protein expression and mRNA stability compared to both no insertion and consecutive AAG controls (Arthur et al., 2015). These experiments demonstrate that polyA tracks play a role in controlling gene expression by attenuating translation efficiency and likely promoting degradation of the transcripts by NGD and NMD (Figure 2 and 3a).

In addition to degradation by one of the mRNA surveillance pathways, there is another possible outcome for transcripts harboring a polyA track. If a polyA track induces frameshifting in the last exon of a transcript, it is likely that the NMD pathway will be inefficient, since it is known that premature termination codons less than 50 nt upstream of the last exon-exon junction do not initiate strong NMD response (Nagy & Maquat, 1998). Instead, the translation cycle could terminate normally on the out of frame stop codon (Figure 3b). The protein synthesized in this scenario would have an altered, frameshifted amino acid sequence after the polyA track. Depending on position of the out of frame stop codon, the novel protein would have either a C-terminal extension or truncation compared to translation in the annotated 0 frame. No such proteins have been identified *in vivo* yet, however, bioinformatics analyses suggests that some polyA track genes may escape NMD and result in a novel C-termini of considerable length (Arthur et al., 2015). Altered C-termini could add regulatory or functional domains to known proteins, adding to the complexity of the proteome. It is expected that these hypothetical proteins would be produced in low amounts and hard to distinguish from their annotated 0 frame counterparts due to the minimal sequence specificity that comes only from the unique C-termini.

Synonymous mutations in polyA tracks significantly impact gene expression

The discrepancy between translational efficiency of consecutive AAA and AAG codons raises the possibility of changes in gene expression due to synonymous mutations (Figure 4). Synonymous mutations alter a gene sequence without changing the sequence of the encoded protein, therefore, typically leaving its function unchanged. For this reason, synonymous mutations are often considered inconsequential when analyzing sequencing data for cancer or other genetic diseases to identify causal variants. *Cis*-acting regulatory gene sequences, however, are sensitive to synonymous mutations. Mutations in exonic motifs that define

splice sites are a well-recognized example of synonymous variants causing changes to protein expression and function (Supek, Miñana, Valcárcel, Gabaldón, & Lehner, 2014). Similarly, synonymous mutations that interrupt a polyA track can significantly affect gene expression without affecting the protein sequence (Arthur et al., 2015). This effect has been demonstrated with both reporter genes with artificial polyA tracks inserted and with genes that endogenously contain a polyA track. When the lysine codons are mutated from AAA to AAG codons to decrease the length of the polyA track, protein expression increases significantly. The converse is true for AAG to AAA mutations, which increase the length of the polyA track (Arthur et al., 2015). In addition to changing gene dosage, a synonymous mutation that increased the length of the polyA track in the *ZCRB1* gene resulted in the production of a potential frameshifted protein (Arthur et al., 2015), a result which supports the prediction of proteins with novel c-terminals produced from polyA track genes.

Perspectives

The discovery of coding polyA track mediated gene regulation adds to the growing list of post-transcriptional gene regulatory mechanisms. It has become clear that ribosome-mediated translation control is a pathway commonly exploited by the cell to selectively modulate expression of genes. This is done through engagement of mRNAs, tRNAs, ribosomes, nascent polypeptide chains, and pathways that control both protein and mRNA quality. PolyA tracks are another mechanism that regulates translation of physiologically correct mRNA at the step of elongation by eliciting responses from both mRNA surveillance pathways and RQC. Nonetheless, open questions remain about the molecular mechanism and biological significance of polyA-induced stalling and frameshifting.

ZNF598 and RACK1 are the most upstream factors shown to act in the ribosome stalling and rescue pathway, meaning their recruitment and ubiquitylation activity may promote terminal stalling on sequences that would otherwise transiently pause the ribosome. If so, regulation of the abundance of these factors would change the stalling efficiency of polyA tracks, allowing for dynamic regulation of these genes. RACK1 is already known to function in cell signaling pathways and has increased translational efficiency during the M phase of the mitotic cell cycle when translation is globally suppressed (Park et al., 2016).

The possibility of alternative proteins rising from frameshifts on polyA tracks, as well as contribution of synonymous mutations in cellular health, are open questions with implication for human health. The frameshifted products as well as synonymous mutations in polyA tracks could be cell type specific as many of the endogenous polyA track genes are subject to alternative splicing, often excluding polyA track containing exons. This would further add to the complexity of gene regulation by polyA tracks resulting in differential gene expression across cell types.

Finally, a majority of the sequenced genomes contain 1–2% of polyA containing transcripts (Habich et al., 2016). This distribution represents a small, but significant, number of genes potentially regulated by this mechanism. However, organisms with very AT rich genomes, such as *Plasmodium* species with more than 60% of total transcripts containing polyA tracks, may have evolved different mechanisms that enable productive and correct synthesis

of lysine-rich peptides from long polyA tracks. If so, elucidating the differences between these ribosomes may provide insight into ribosome evolution and potential targets for therapeutics.

Acknowledgments

We thank S. Pavlovic-Djuranovic, H. Zaher, and J. Skeath for helpful comments. The work in Djuranovic lab is supported by the NIH grant T32 GM007067 to LLA and NIH Grant R01 GM112824 to SD.

References

- Amaldi F, Pierandrei-Amaldi P. 1997; TO P Genes : A Translationally Controlled Class of Genes Including Those Coding for Ribosomal Proteins. *Progress in Molecular and Subcellular Biology*. 18
- Arthur LL, Chung JJ, Jankirama P, Keefer KM, Kolotilin I, Pavlovic-djuranovic S, Djuranovic S. 2017; Rapid generation of hypomorphic mutations. *Nature Communications*. 8 :1–15. DOI: 10.1038/ncomms14112
- Arthur LL, Pavlovic-Djuranovic S, Koutmou KS, Green R, Szczesny P, Djuranovic S. 2015; Translational control by lysine-encoding A-rich sequences. *Science Advances*. 1 (6) :e1500154. doi: 10.1126/sciadv.1500154 [PubMed: 26322332]
- Bengtson MH, Joazeiro CAP. 2010; Role of a ribosome-associated E3 ubiquitin ligase in protein quality control. *Nature*. 467 (7314) :470–3. DOI: 10.1038/nature09371 [PubMed: 20835226]
- Brandman O, Stewart-Ornstein J, Wong D, Larson A, Williams CC, Li G-W, Weissman JS. 2012; A ribosome-bound quality control complex triggers degradation of nascent peptides and signals translation stress. *Cell*. 151 (5) :1042–54. DOI: 10.1016/j.cell.2012.10.044 [PubMed: 23178123]
- Chan PP, Lowe TM. 2009; GtRNAdb : A database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Research*. 37 (93–97) :93–97. DOI: 10.1093/nar/gkn787
- Chang H, Lim J, Ha M, Kim VN. 2014; TAIL-seq: Genome-wide determination of poly(A) tail length and 3' end modifications. *Molecular Cell*. 53 (6) :1044–1052. DOI: 10.1016/j.molcel.2014.02.007 [PubMed: 24582499]
- Chorghade S, Seimetz J, Emmons R, Yang J, Bresson SM, Lisio M, De, Conrad NK. 2017; Poly(A) tail length regulates PABPC1 expression to tune translation in the heart. *eLife*. e24139 :1–19. DOI: 10.7554/eLife.24139
- Dai A, Cao S, Dhungel P, Luan Y, Liu Y, Xie Z, Yang Z. 2017; Ribosome Profiling Reveals Translational Upregulation of Cellular Oxidative Phosphorylation mRNAs during Vaccinia Virus-Induced Host Shutoff. *Journal of Virology*. 91 (5) :e01858–16. DOI: 10.1128/JVI.01858-16 [PubMed: 28003488]
- Defenouillère, Q; Yao, Y; Mouaikel, J; Namane, A; Galopier, A; Decourty, L; Fromont-racine, M. Cdc48-associated complex bound to 60S particles is required for the clearance of aberrant translation products. 2013.
- Dimitrova LN, Kuroha K, Tatematsu T, Inada T. 2009; Nascent peptide-dependent translation arrest leads to Not4p-mediated protein degradation by the proteasome. *The Journal of Biological Chemistry*. 284 (16) :10343–52. DOI: 10.1074/jbc.M808840200 [PubMed: 19204001]
- Doma MK, Parker R. 2006; Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature*. 440 (7083) :561–4. DOI: 10.1038/nature04530 [PubMed: 16554824]
- Doronina VA, Wu C, de Felipe P, Sachs MS, Ryan MD, Brown JD. 2008; Site-Specific Release of Nascent Chains from Ribosomes at a Sense Codon. *Molecular and Cellular Biology*. 28 (13) :4227–4239. DOI: 10.1128/MCB.00421-08 [PubMed: 18458056]
- Frishmeyer, Pa; van Hoof, A; O'Donnell, K; Guerrero, AL; Parker, R; Dietz, HC. 2002; An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science*. 295 (5563) :2258–61. DOI: 10.1126/science.1067338 [PubMed: 11910109]
- Gandhi R, Manzoor M, Hudak Ka. 2008; Depurination of Brome mosaic virus RNA3 in vivo results in translation-dependent accelerated degradation of the viral RNA. *The Journal of Biological Chemistry*. 283 (47) :32218–28. DOI: 10.1074/jbc.M803785200 [PubMed: 18815133]

- Garzia A, Jafarnejad SM, Meyer C, Chapat C, Gogakos T, Morozov P, Sonenberg N. 2017 The E3 ubiquitin ligase and RNA-binding protein ZNF598 orchestrates ribosome quality control of premature polyadenylated mRNAs. *Nature Communications*. 8 May. doi: 10.1038/ncomms16056
- Gerbası VR, Weaver CM, Hill S, Friedman DB, Link AJ. 2004; Yeast Asc1p and mammalian RACK1 are functionally orthologous core 40S ribosomal proteins that repress gene expression. *Molecular and Cellular Biology*. 24 (18) :8276–8287. DOI: 10.1128/MCB.24.18.8276
- Habich M, Djuranovic S, Szczesny P. 2016; PATACSDb—The database of polyA translational attenuators in coding sequences. *PeerJ Computer Science*. 2 :e45. doi: 10.7717/peerj-cs.45
- Hashimoto Y, Takahashi M, Sakota E, Nakamura Y. 2017; Nonstop-mRNA decay machinery is involved in the clearance of mRNA 5′-fragments produced by RNAi and NMD in *Drosophila melanogaster* cells. *Biochemical and Biophysical Research Communications*. 484 (1) :1–7. DOI: 10.1016/j.bbrc.2017.01.092 [PubMed: 28115162]
- Hinnebusch AG, Ivanov IP, Sonenberg N. 2016; Translational control by 5′-untranslated regions of eukaryotic mRNAs. *Science*. 352 (6292) :1413–1416. [PubMed: 27313038]
- Ito-Harashima S, Kuroha K, Tatematsu T, Inada T. 2007; Translation of the poly(A) tail plays crucial roles in nonstop mRNA surveillance via translation repression and protein destabilization by proteasome in yeast. *Genes & Development*. 21 (5) :519–24. DOI: 10.1101/gad.1490207 [PubMed: 17344413]
- Jha S, Rollins MG, Fuchs G, Procter DJ, Hall EA, Cozzolino K, Walsh D. 2017; Trans-kingdom mimicry underlies ribosome customization by a poxvirus kinase. *Nature*. 546 (7660) :651–655. DOI: 10.1038/nature22814 [PubMed: 28636603]
- Juszkiewicz S, Hegde RS. 2017; Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination. *Molecular Cell*. 65 :1–8. [PubMed: 28061329]
- Karlin S, Brocchieri L, Bergman A, Mrazek J, Gentles AJ. 2002; Amino acid runs in eukaryotic proteomes and disease associations. *PNAS*. 99 (1) :333–338. DOI: 10.1073/pnas.012608599 [PubMed: 11782551]
- Ketteler R. 2012 On programmed ribosomal frameshifting: The alternative proteomes. *Frontiers in Genetics*. 3 Nov. :1–10. DOI: 10.3389/fgene.2012.00242 [PubMed: 22303408]
- Koutmou KS, Schuller AP, Brunelle JL, Radhakrishnan A, Djuranovic S, Green R. 2015; Ribosomes slide on lysine-encoding homopolymeric A stretches. *eLife*. 4 :1–18. DOI: 10.7554/eLife.05534
- Kuroha K, Akamatsu M, Dimitrova L, Ito T, Kato Y, Shirahige K, Inada T. 2010; Receptor for activated C kinase 1 stimulates nascent polypeptide-dependent translation arrest. *EMBO Reports*. 11 (12) :956–61. DOI: 10.1038/embor.2010.169 [PubMed: 21072063]
- Letzring DP, Wolf AS, Brule CE, Grayhack EJ. 2013; Translation of CGA codon repeats in yeast involves quality control components and ribosomal protein L1. *RNA*. 19 :1208–1217. DOI: 10.1261/rna.039446.113.3 [PubMed: 23825054]
- Lima, SA; Chipman, LB; Nicholson, AL; Chen, Y-H; Yee, BA; Yeo, GW; Pasquinelli, AE. Short poly(A) tails are a conserved feature of highly expressed genes. *Nature Structural & Molecular Biology*. 2017 Nov.
- Lu J, Deutsch C. 2008; Electrostatics in the ribosomal tunnel modulate chain elongation rates. *Journal of Molecular Biology*. 384 (1) :73–86. DOI: 10.1016/j.jmb.2008.08.089 [PubMed: 18822297]
- Machida K, Mikami S, Masutani M, Mishima K, Kobayashi T, Imataka H. 2014; A translation system reconstituted with human factors proves that processing of encephalomyocarditis virus proteins 2A and 2B occurs in the elongation phase of translation without eukaryotic release factors. *Journal of Biological Chemistry*. 289 (46) :31960–31971. DOI: 10.1074/jbc.M114.593343
- Mangus DA, Evans MC, Jacobson A. 2003; Poly(A)-binding proteins : multifunctional scaffolds for the post- transcriptional control of gene expression. *Genome Biology*. 4 (7)
- Meaux S, Van Hoof A. 2006; Yeast transcripts cleaved by an internal ribozyme provide new insight into the role of the cap and poly (A) tail in translation and mRNA decay. *RNA*. 12 :1323–1337. DOI: 10.1261/rna.46306.First [PubMed: 16714281]
- Moss B. 1968; Inhibition of HeLa cell protein synthesis by the vaccinia virion. *Journal of Virology*. 2 (10) :1028–1037. [PubMed: 5723706]

- Nagy E, Maquat LE. 1998; A rule for termination-codon position within intron-containing genes: When nonsense affects RNA abundance. *Trends in Biochemical Sciences*. 23 (6) :198–199. DOI: 10.1016/S0968-0004(98)01208-0 [PubMed: 9644970]
- Ozsolak F, Kapranov P, Foissac S, Kim SW, Fishilevich E, Monaghan AP, Milos PM. 2010; Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation. *Cell*. 143 (6) :1018–1029. DOI: 10.1016/j.cell.2010.11.020 [PubMed: 21145465]
- Park J, Yi H, Kim Y, Chang H, Kim VN, Park J, Kim VN. 2016; Regulation of Poly(A) Tail and Translation during the Somatic Cell Cycle. *Molecular Cell*. 62 (3) :462–471. DOI: 10.1016/j.molcel.2016.04.007 [PubMed: 27153541]
- Pisareva VP, Skabkin MA, Hellen CUT, Pestova TV, Pisarev AV. 2011; Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. *EMBO Journal*. 30 (9) :1804–1817. DOI: 10.1038/emboj.2011.93
- Presnyak V, Alhusaini N, Chen YH, Martin S, Morris N, Kline N, Collier J. 2015; Codon optimality is a major determinant of mRNA stability. *Cell*. 160 (6) :1111–1124. DOI: 10.1016/j.cell.2015.02.029 [PubMed: 25768907]
- Roulston C, Luke GA, de Felipe P, Ruan L, Cope J, Nicholson J, Ryan MD. 2016; “2A-Like” Signal Sequences Mediating Translational Recoding: A Novel Form of Dual Protein Targeting. *Traffic*. 17 (8) :923–939. DOI: 10.1111/tra.12411 [PubMed: 27161495]
- Ryan MD, Drew J. 1994; Foot-and-mouth disease virus 2A oligopeptide mediated cleavage of an artificial polyprotein. *The EMBO Journal*. 13 (4) :928–33. [PubMed: 8112307]
- Saito K, Horikawa W, Ito K. 2015; Inhibiting K63 Polyubiquitination Abolishes No-Go Type Stalled Translation Surveillance in *Saccharomyces cerevisiae*. *PLoS Genetics*. 11 (4) :1–18. DOI: 10.1371/journal.pgen.1005197
- Saito S, Hosoda N, Hoshino SI. 2013; The Hbs1-Dom34 protein complex functions in non-stop mRNA decay in mammalian cells. *Journal of Biological Chemistry*. 288 (24) :17832–17843. DOI: 10.1074/jbc.M112.448977
- Shao S, Brown A, Santhanam B, Hegde RS. 2015; Structure and assembly pathway of the ribosome quality control complex. *Molecular Cell*. 57 (3) :433–445. [PubMed: 25578875]
- Shoemaker CJ, Eyler DE, Green R. 2010; Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop off to initiate no-go decay. *Science*. 330 (6002) :369–372. DOI: 10.1126/science.1192430.Dom34 [PubMed: 20947765]
- Shoemaker CJ, Green R. 2011; Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. *Proceedings of the National Academy of Sciences*. 108 (51) :E1392–E1398. DOI: 10.1073/pnas.1113956108
- Simms CL, Yan LL, Zaher HS. 2017; Ribosome Collision Is Critical for Quality Control during No-Go Decay. *Molecular Cell*. 68 (2) :361–373. e5. doi: 10.1016/j.molcel.2017.08.019 [PubMed: 28943311]
- Subtelny AO, Eichhorn SW, Chen GR, Sive H, Bartel DP. 2014; Poly(A)-tail profiling reveals an embryonic switch in translational control. *Nature*. 508 (1) :66–71. DOI: 10.1038/nature13007 [PubMed: 24476825]
- Sundaramoorthy E, Leonard M, Mak R, Liao J, Fulzele A, Bennett EJ. 2017; ZNF598 and RACK1 Regulate Mammalian Ribosome-Associated Quality Control Function by Mediating Regulatory 40S Ribosomal Ubiquitylation. *Molecular Cell*. 65 :1–10. DOI: 10.1016/j.molcel.2016.12.026 [PubMed: 28061329]
- Supek F, Miñana B, Valcárcel J, Gabaldón T, Lehner B. 2014; Synonymous mutations frequently act as driver mutations in human cancers. *Cell*. 156 (6) :1324–35. DOI: 10.1016/j.cell.2014.01.051 [PubMed: 24630730]
- Szostak E, Gebauer F. 2012; Translational control by 3'-UTR-binding proteins. *Briefings in Functional Genomics*. 12 (1) :58–65. DOI: 10.1093/bfpg/els056 [PubMed: 23196851]
- Tarun SZ, Sachs AB. 1996; Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *The EMBO Journal*. 15 (24) :7168–77. [PubMed: 9003792]
- Tarun SZ, Wells SE, Deardorff JA, Sachs AB. 1997; Translation initiation factor eIF4G mediates in vitro poly(A) tail-dependent translation. *Proceedings of the National Academy of Sciences of the United States of America*. 94 (17) :9046–51. DOI: 10.1073/pnas.94.17.9046 [PubMed: 9256432]

- Tian B, Manley JL. 2017; Alternative polyadenylation of mRNA precursors. *Nat Rev Mol Cell Biol.* 18 (1) :165–187. DOI: 10.1007/128
- Tsuboi T, Kuroha K, Kudo K, Makino S, Inoue E, Kashima I, Inada T. 2012; Dom34: Hbs1 Plays a General Role in Quality-Control Systems by Dissociation of a Stalled Ribosome at the 3' End of Aberrant mRNA. *Molecular Cell.* 46 (4) :518–529. DOI: 10.1016/j.molcel.2012.03.013 [PubMed: 22503425]
- van Hoof A, Frischmeyer Pa, Dietz HC, Parker R. 2002; Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science.* 295 (5563) :2262–4. DOI: 10.1126/science.1067272 [PubMed: 11910110]
- Weill L, Belloc E, Bava FA, Méndez R. 2012; Translational control by changes in poly(A) tail length: Recycling mRNAs. *Nature Structural and Molecular Biology.* 19 (6) :577–585. DOI: 10.1038/nsmb.2311
- Wickens M. 1990; In the beginning is the end: regulation of poly(A) addition and removal during early development. *Trends in Biochemical Sciences.* 15 (8) :320–324. DOI: 10.1016/0968-0004(90)90022-4 [PubMed: 2204159]
- Wickens M, Bernstein DS, Kimble J, Parker R. 2002; A PUF family portrait: 3'UTR regulation as a way of life. *Trends in Genetics.* 18 (3) :150–157. [PubMed: 11858839]
- Wolf AS, Grayhack EJ. 2015; Asc1, homolog of human RACK1, prevents frameshifting in yeast by ribosomes stalled at CGA codon repeats. *RNA.* 21 :935–945. DOI: 10.1261/rna.049080.114.frameshift [PubMed: 25792604]
- Yamashita R, Suzuki Y, Takeuchi N, Wakaguri H, Ueda T, Sugano S, Nakai K. 2008; Comprehensive detection of human terminal oligo-pyrimidine (TOP) genes and analysis of their characteristics. *Nucleic Acids Research.* 36 (11) :3707–3715. DOI: 10.1093/nar/gkn248 [PubMed: 18480124]
- Yonashiro R, Tahara EB, Bengtson MH, Khokhrina M, Lorenz H, Chen K-C, Joazeiro CA. 2016; The Rqc2/Tae2 subunit of the Ribosome-Associated Quality Control (RQC) complex marks ribosome-stalled nascent polypeptide chains for aggregation. *eLife.* 5 doi: 10.7554/eLife.11794

Figure 1. Normal and aberrant translation

(a) Scheme of translation of normal mRNA with 5' cap (m^7GpppG), 3' poly(A) tail, 3' and 5' untranslated regions (UTR), and open reading frame (ORF) indicated. Translation proceeds in three phases; initiation, elongation, and termination. (b) In the absence of an in-frame stop codon, the ribosome continues elongation into the poly(A) tail. This activates nonstop decay and the recruitment of the GTPase Ski7 in *Saccharomyces cerevisiae*. In other eukaryotes, Dom34 and Hbs1 are recruited. (c) Peptide or mRNA induced stalls (indicated by gray box on mRNA scheme) during the elongation cycle activate the no-go mRNA decay pathway. The rescue factors Dom34 and Hbs1 are recruited to recycle the ribosome in this

pathway as well. Both nonstop decay and no-go decay result in endonucleolytic cleavage and degradation of the mRNA and ubiquitylation and degradation of the nascent peptide. (d) Table of factors involved in ribosome stall induction, ribosome release, and nascent protein degradation discussed in text (if different than human gene nomenclature, yeast homologs are indicated).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 2. Coding polyA tracks induce translational frameshifting and ribosome stalling

Translation of polyA tracks in coding sequences of mRNAs can result in two possible outcomes – frameshifting and ribosome stalling. Frameshifting on polyA tracks (depicted by change in the ORF and protein color) does not require mRNA structure elements seen in programmed frameshifting in viruses. Ribosome stalling (indicated by gray box on mRNA scheme) results in activation of NGD mechanisms with reduction in protein levels (illustrated by released ribosomes) but synthesis of the full length protein in correct frame.

(a) Degradation by Nonsense-mediated RNA decay

(b) Translation of alternative C-terminal

Figure 3. Potential outcomes of programmed ribosome frameshifting on coding polyA tracks

(a) Ribosomes that frameshift on a polyA track or other recoding element will continue elongation until reaching a stop codon in the new frame (protein sequence from non-zero frame indicated by red line). Encountering a premature termination codon will initiate the mRNA surveillance pathway, nonsense mediated decay to degrade both the mRNA and nascent peptide. If the premature termination codon occurs before the last exon junction complex (EJC), degradation by NMD is stimulated. Ubiquitination of the frameshifted nascent protein is indicated (*). (b) If the frameshift happens in the last exon, the transcript may evade NMD and instead synthesize a nascent peptide with an alternative C-terminus coded by the non-zero frame (indicated by UAA*) after the polyA track and either truncated (top) or extended (bottom) depending on the position of the out of frame termination codon.

mRNA with coding polyA track

mRNA with synonymous mutation (Lys AAG to AAA)

mRNA with synonymous (Lys AAA to AAG)

Figure 4. Model of effects of synonymous mutations within coding polyA tracks

(a) Scheme of translation of mRNA with polyA track indicated by the grey segment in the open reading frame. The translation efficiency of mRNAs with a polyA track are sensitive to synonymous mutations. (b) Mutations that increase the number of adenosine residues, Lys AAG to AAA, will increase stalling and frameshifting on the polyA track, leading to reduced WT protein expression from the mRNA and greater frequency of production of alternative C-terminus proteins. (c) The opposite mutation, Lys AAA to AAG, will increase

WT protein expression and decrease frequency of alternative C-terminus protein production by decreasing the frequency of stalling and frameshifting on the polyA tracks.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript