

PROTEIN SYNTHESIS

Rqc2p and 60S ribosomal subunits mediate mRNA-independent elongation of nascent chains

Peter S. Shen,¹ Joseph Park,² Yidan Qin,^{8,9} Xueming Li,^{7*} Krishna Parsawar,¹⁰ Matthew H. Larson,^{3,4,5,6} James Cox,^{1,10} Yifan Cheng,⁷ Alan M. Lambowitz,^{8,9} Jonathan S. Weissman,^{3,4,5,6†} Onn Brandman,^{2†} Adam Frost^{1,7†}

In Eukarya, stalled translation induces 40S dissociation and recruitment of the ribosome quality control complex (RQC) to the 60S subunit, which mediates nascent chain degradation. Here we report cryo-electron microscopy structures revealing that the RQC components Rqc2p (YPL009C/Tae2) and Ltn1p (YMR247C/Rkr1) bind to the 60S subunit at sites exposed after 40S dissociation, placing the Ltn1p RING (Really Interesting New Gene) domain near the exit channel and Rqc2p over the P-site transfer RNA (tRNA). We further demonstrate that Rqc2p recruits alanine- and threonine-charged tRNA to the A site and directs the elongation of nascent chains independently of mRNA or 40S subunits. Our work uncovers an unexpected mechanism of protein synthesis, in which a protein—not an mRNA—determines tRNA recruitment and the tagging of nascent chains with carboxy-terminal Ala and Thr extensions (“CAT tails”).

Despite the processivity of protein synthesis, faulty messages or defective ribosomes can result in translational stalling and incomplete nascent chains. In Eukarya, this leads to recruitment of the ribosome quality control complex (RQC), which mediates the ubiquitylation and degradation of incompletely synthesized nascent chains (1–4). The molecular components of the RQC include the AAA adenosine triphosphatase Cdc48p and its ubiquitin-binding cofactors, the RING-domain E3 ligase Ltn1p, and two proteins of unknown function, Rqc1p and Rqc2p. We set out to determine the mechanism(s) by which relatively rare (5) proteins such as Ltn1p, Rqc1p, and Rqc2p recognize and rescue stalled 60S ribosome nascent chain complexes, which are vastly outnumbered by ribosomes translating normally or in stages of assembly.

To reduce structural heterogeneity and enrich for complexes still occupied by stalled nascent chains, we immunoprecipitated Rqc1p-bound RQC assemblies from *Saccharomyces cerevisiae* strains

lacking the C-terminal RING domain of Ltn1p, which prevents substrate ubiquitylation and Cdc48 recruitment (7). Three-dimensional (3D) classification of Ltn1 Δ RING particles revealed

60S ribosomes with nascent chains in the exit tunnel and extraribosomal densities (Fig. 1). These extraribosomal features were resolved between 5 and 14 Å and proved to be either Tif6p or RQC components as characterized below (Fig. 1 and figs. S1 to S7). Tif6p was not observed bound to the same 60S particles bound by RQC factors (figs. S1 to S3). We repeated the purification, imaging, and 3D classification from *rqc2 Δ* cells and computed difference maps. This analysis did not reveal density attributable to Rqc1p but did identify Rqc2p as a transfer RNA (tRNA)-binding protein that occupies the 40S binding surface and Ltn1p as the elongated molecule that meets Rqc2p at the sarcin-ricin loop (SRL) (Figs. 1 and 2 and figs. S1 to S5). Comparison of the 60S-bound Ltn1p with reconstructions of isolated Ltn1p suggests that the N terminus of Ltn1p engages the SRL with Rqc2p and that the middle region—which contains long HEAT/Armadillo repeats that adopt an elongated superhelical structure—reaches around the 60S (6). This conformation probably positions the C-terminal RING domain near the exit tunnel to ubiquitylate stalled nascent chains [figs. S5 and S6 and (7)].

A refined reconstruction of the Rqc2p-occupied class demonstrated that Rqc2p makes extensive contacts with an approximately P-site positioned (~P-site) tRNA (Figs. 1 and 2 and fig. S7). Rqc2p has a long coiled coil that makes direct contact

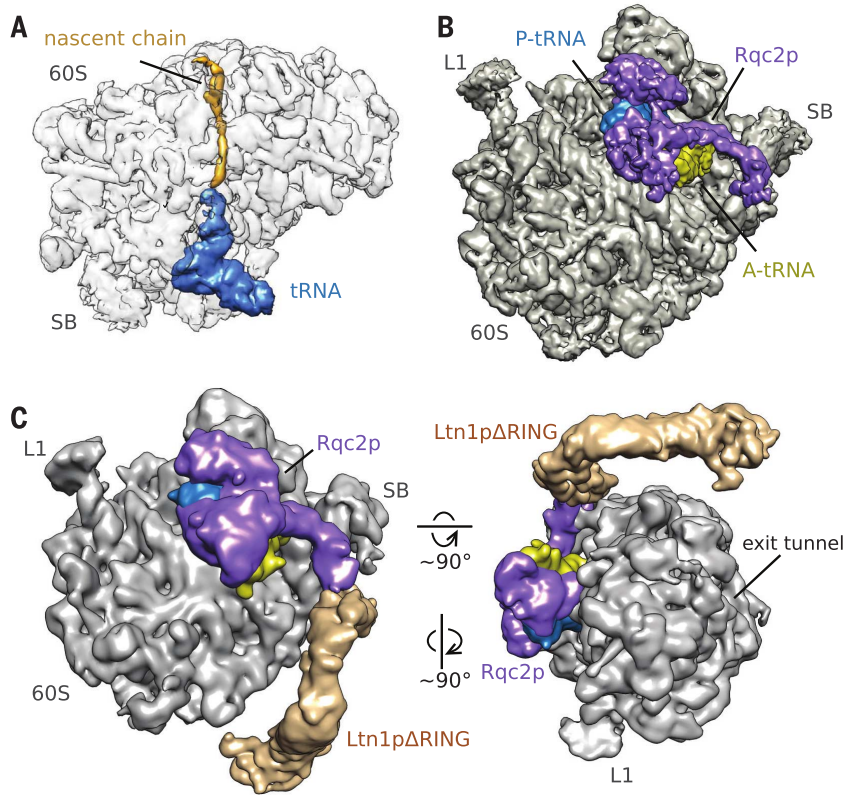


Fig. 1. Cryo-EM reconstructions of peptidyl-tRNA-60S ribosomes bound by the RQC components Rqc2p and Ltn1p. (A) A peptidyl-tRNA-60S complex isolated by immunoprecipitation of Rqc1p. The ribosome density is transparent to visualize the nascent chain. (B) Rqc2p (purple) and an A-site tRNA (yellow) bound to peptidyl-tRNA-60S complexes. Landmarks are indicated (L1, L1 stalk; SB, P-stalk base). (C) Ltn1p (tan) bound to Rqc2p-peptidyl-tRNA-60S complexes (B).

¹Department of Biochemistry, University of Utah, UT 84112, USA. ²Department of Biochemistry, Stanford University, Palo Alto, CA 94305, USA. ³Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, CA 94158, USA. ⁴Howard Hughes Medical Institute, University of California, San Francisco, San Francisco, CA 94158, USA. ⁵California Institute for Quantitative Biomedical Research, University of California, San Francisco, San Francisco, CA 94158, USA. ⁶Center for RNA Systems Biology, University of California, San Francisco, San Francisco, CA 94158, USA. ⁷Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94158, USA. ⁸Institute for Cellular and Molecular Biology, University of Texas at Austin, Austin, TX 78712, USA. ⁹Department of Molecular Biosciences, University of Texas at Austin, Austin, TX 78712, USA. ¹⁰Mass Spectrometry and Proteomics Core Facility, University of Utah, UT 84112, USA. *Present address: School of Life Sciences, Tsinghua University, Beijing 100084, China. †Corresponding author. E-mail: jonathan.weissman@ucsf.edu (J.S.W.), onn@stanford.edu (O.B.), adam.frost@ucsf.edu (A.F.)

with the SRL and the 60S P-stalk base (Fig. 2A). This structure also revealed Rqc2p binding to an \sim A-site tRNA, whose 3'-CCA tail is within the peptidyl transferase center of the 60S (Fig. 2B and fig. S7). This observation was unexpected, because A-site tRNA interactions with the large ribosomal subunit are typically unstable and require mRNA templates and elongation factors (8). Rqc2p's interactions with the \sim A-site tRNA appeared to involve recognition between the anticodon loop and a globular N-terminal domain, as well as D-loop and T-loop interactions along Rqc2p's coiled coil (Figs. 2 and 3).

To determine whether Rqc2p binds specific tRNA molecules, we extracted total RNA after RQC purification from strains with intact *RQC2* versus *rqc2Δ* strains. Deep sequencing by a new method using a thermostable group II intron reverse transcriptase (9) revealed that the presence of Rqc2p leads to an \sim 10-fold enrichment of tRNA^{Ala(AGC)} and tRNA^{Thr(AGT)} in the RQC (Fig. 3A). In complexes isolated from strains with intact *RQC2*, Ala(AGC) and Thr(AGT) are the most abundant tRNA molecules, even though they are less abundant than a number of other tRNAs in yeast (10).

Our structure suggested that Rqc2p's specificity for these tRNAs is due in part to direct interactions between Rqc2p and positions 32 to 36 of the anticodon loop, some of which are edited in the mature tRNA (Fig. 3). Adenosine 34 in the anticodon of both tRNA^{Ala(AGC)} and tRNA^{Thr(AGT)} is deaminated to inosine (11–13), leading to a diagnostic guanosine upon reverse transcription (13, 14) (Fig. 3, B and C). Further analysis of the sequencing data revealed that cytosine 32 in tRNA^{Thr(AGT)} is also deaminated to uracil in \sim 70% of the Rqc2p-enriched reads [Fig. 3 and (15)]. Together with the structure, this suggests that Rqc2p binds to the D-, T-, and anticodon loop of the \sim A-site tRNA, and that recognition of the 32-UUIGY-36 edited motif accounts for Rqc2p's specificity for these two tRNAs (Fig. 3, C and D). The pyrimidine at position 36 could explain the discrimination between the otherwise similar anticodon loops that harbor purines at base 36.

While assessing why Rqc2p evolved to bind these specific tRNA molecules, we considered these observations: First, our structural and biochemical data indicate that Rqc2p binds the 60S subunit after a stalled ribosome dissociates [fig. S6 (1, 2)]. Second, stalled nascent chains accumulate as higher-molecular-weight species in the presence of Rqc2p than in its absence [Fig. 4A, also seen in Fig. 3E of (1)]. Finally, amino acid addition to a nascent chain can be mediated by the large ribosomal subunit in vitro even when decoupled from an mRNA template and the small subunit (16). Together, these facts led us to hypothesize that Rqc2p may promote the extension of stalled nascent chains with alanine and threonine residues in an elongation reaction that is mRNA- and 40S-free. This hypothesis makes specific predictions. First, the Rqc2p-dependent increase in the molecular weight of the nascent chain should occur from the C terminus exclu-

sively. Second, the C-terminal extension should consist entirely of alanine and threonine residues that start immediately at the stalling sequence. Finally, the alanine and threonine extension should not have a defined sequence.

To test these predictions, we expressed a series of reporters containing a stalling sequence [tracts of up to 12 consecutive arginine codons, including pairs of the difficult-to-decode CGA codon (17)], inserted between the coding regions of green fluorescent protein (GFP) and red fluorescent

protein (Fig. 4A). Null mutations in RQC components or inhibition of the proteasome led to the accumulation of nascent chain fragments that are normally degraded in wild-type cells (Fig. 4A) (1–4, 18). Furthermore, *ltm1Δ* and *rqc2Δ* cells have different phenotypes: Expression of the stalling reporter in *ltm1Δ* led to the formation and accumulation of higher-molecular-weight species that resolve as a smear \sim 1.5 to 5 kD above the expected position of GFP (Fig. 4A). GFP mass-shifted products are observable in *rqc1Δltm1Δ*

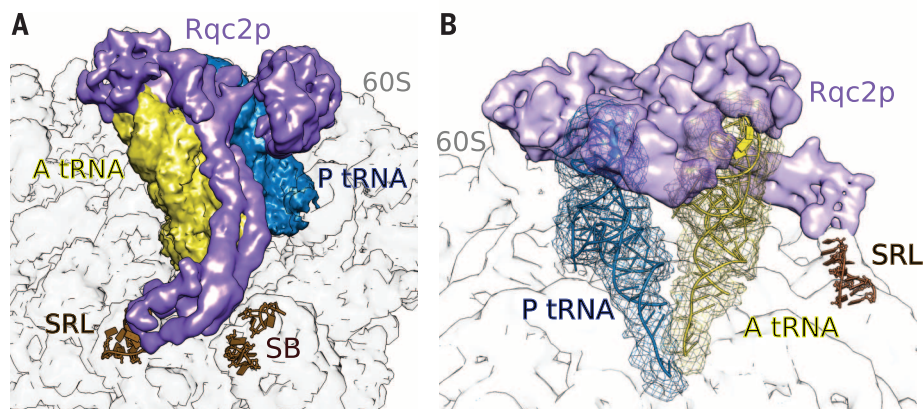


Fig. 2. Rqc2p binding to the 60S ribosome and \sim P-site, and \sim A-site tRNAs. (A) Rqc2p contacts \sim P- and \sim A-site tRNAs, the SRL, and P-stalk base ribosomal RNA (SB). (B) Rigid body fitting of tRNAs structures (ribbons) into EM densities (mesh).

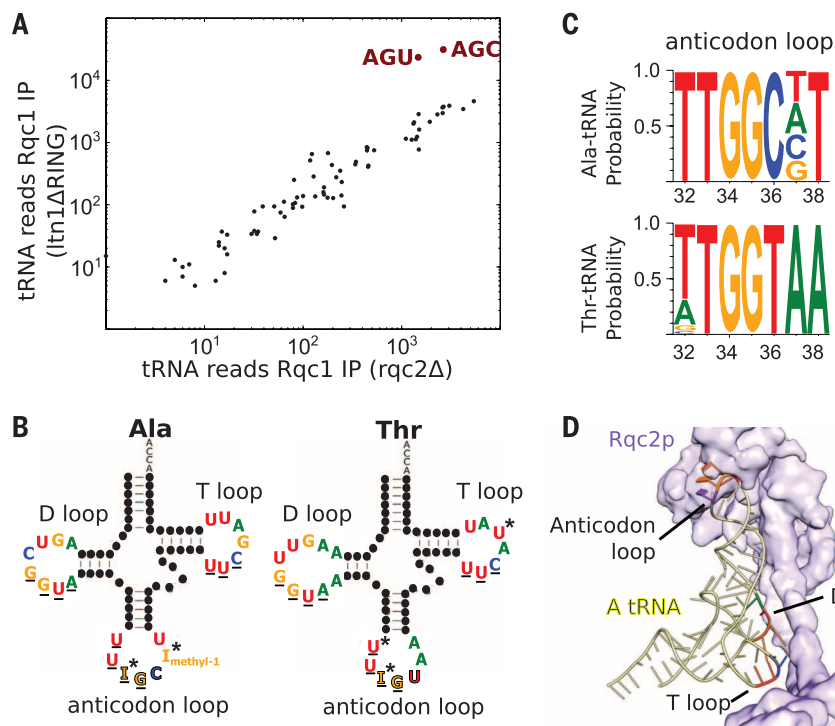


Fig. 3. Rqc2p-dependent enrichment of tRNA^{Ala(AGC)} and tRNA^{Thr(IGU)}. (A) tRNA cDNA reads extracted from purified RQC particles and summed per unique anticodon, with versus without Rqc2p. (B) Secondary structures of tRNA^{Ala(AGC)} and tRNA^{Thr(IGU)}. Identical nucleotides are underlined. Edited nucleotides are indicated with asterisks (24, 25). (C) Weblogo representation of cDNA sequencing reads related to shared sequences found in anticodon loops (positions 32 to 38) of mature tRNA^{Ala(AGC)} and tRNA^{Thr(IGU)} (26). (D) \sim A-tRNA contacts with Rqc2p at the D-, T-, and anticodon loops. Identical nucleotides between tRNA^{Ala(AGC)} and tRNA^{Thr(IGU)} are colored as in (B) (A, green; U, red; C, blue; G, orange) and pyrimidine, purple. Anticodon nucleotides are indicated as slabs.

double mutants, less prominent but still observable in *rqc1Δ* single mutants, but absent in all *rqc2Δ* single and double mutants (Fig. 4A). Thus, Rqc2p is necessary for the production of these higher-molecular-weight GFP species.

We probed the location of the extra mass along the GFP by inserting a tobacco etch virus (TEV) protease cleavage site upstream of the stalling tract (Fig. 4B). GFP resolved as a single band of the expected size with TEV treatment, indicating that the extra mass is located at or after the stall sequence. To pinpoint the location of the extra mass along the GFP, we moved the TEV cleavage site after the R12 stalling sequence. This created a mass-shifted GFP that was insensitive to TEV treatment, suggesting that the post-R12 TEV cleavage site was not synthesized. One possible model is that a translational frameshift occurs near the R12 sequence, which causes the mRNA to be mistranslated until the next out-of-frame stop codon. We falsified this model in two ways. First, we detected an Rqc2p-dependent GFP mass shift using a shorter R4 reporter in which multiple STOP codons were engineered in the +1 and +2 frames following the polyarginine tract (fig. S8). Second, we detected the Rqc2p-dependent GFP mass shift in a construct encoding a hammerhead ribozyme. The ribozyme cleaves the coding sequence of the GFP mRNA, leaving a truncated non-stop mRNA that causes a stall during translation of its final codon [fig. S8 (19)]. Thus, the GFP mass shift is located at or after the stall sequence but cannot be explained by mRNA translation past the stalling tract in any frame.

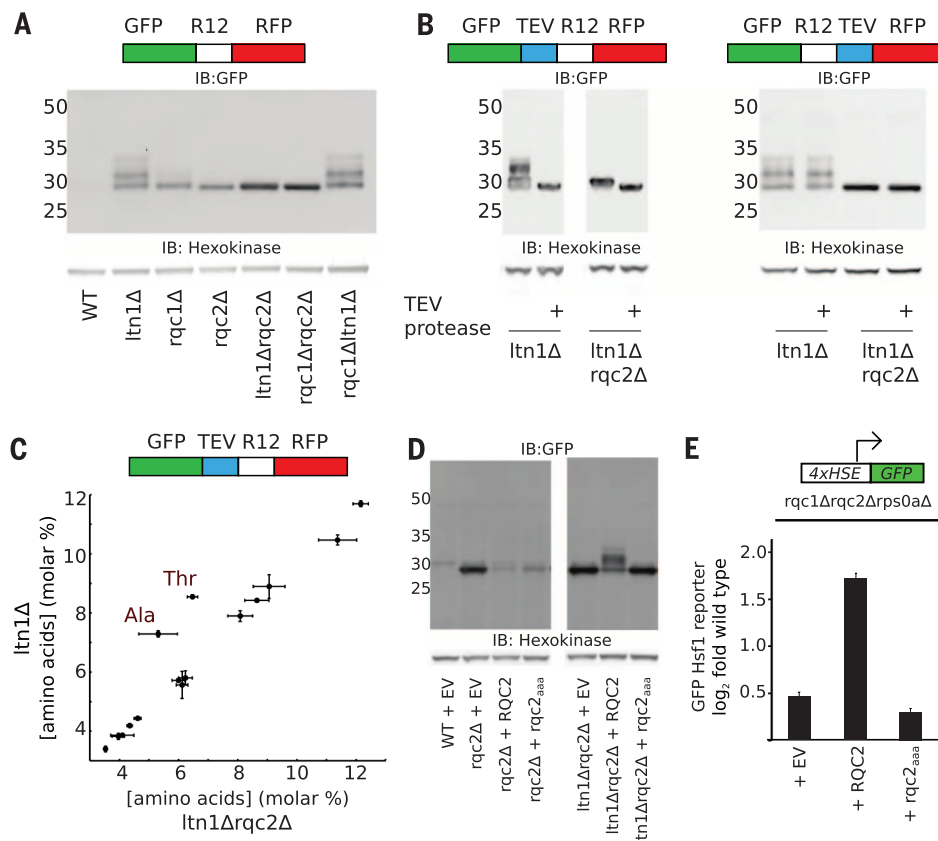
In order to determine the composition of the GFP mass-shifted products, we performed total amino acid analysis of immunopurified GFP from strains expressing the stalling reporter. Purified GFP from *ltn1Δ* (Fig. 4C) or *rqc1Δ* strains (fig. S9) is enriched in alanine and threonine as compared to purified GFP from double mutants with *rqc2Δ* which do not produce extended GFP. We then used Edman degradation to sequence TEV release fragments after purification of the stalled GFP reporter from the *ltn1Δ* strain. The first three codons in the R12 sequence are CGG-CGA-CGA, and Edman degradation suggested that the ribosome stalls at the first pair of the challenging-to-decode CGA codons (fig. S10). Following the encoded arginine residues, rising levels of alanine and threonine were detected at the C terminus (fig. S10). We further characterized these fragments by mass spectrometry and detected diverse poly-Ala and poly-Thr species ranging from 5 to 19 residues, with no defined sequence (table S1). Together, these observations demonstrate that Rqc2p directs the elongation of stalled nascent chains with nontemplated carboxy-terminal Ala and Thr extensions, or “CAT tails.”

Earlier work (1) revealed that the accumulation of stalled nascent chains (e.g., by deletion of *LTN1*) led to a robust heat shock response that is fully dependent on Rqc2p, although the mechanism by which Rqc2p enabled this stress response was unclear. We hypothesized that CAT tails may be required for activation of heat shock factor 1 (Hsf1p). To isolate the effect of CAT tails in this context, we sought an *rqc2*

allele that could not support CAT tail synthesis but could still bind to the 60S and facilitate Ltn1p-dependent ubiquitylation of the nascent chains. Rqc2p belongs to the conserved NFACT family of nucleic acid-binding proteins (20), and the N-terminal NFACT-N domain of Rqc2p is 22% identical to the NFACT-N domain of the *Staphylococcus aureus* protein Fbp (PDB:3DOA). Based on sequence and predicted secondary structure conservation, we fit this structure into a portion of the cryo-electron microscopy (cryo-EM) density ascribed to Rqc2p (figs. S11 and S12). This modeling exercise predicts that Rqc2p's NFACT-N domain recognizes features of both the P- and A-site tRNA molecules and that conserved residues D9, D98, and R99, which have been hypothesized to play roles in nucleic acid-binding or -modifying reactions (20), may contact the ~A-site tRNA (20) (fig. S12). An Rqc2p variant in which these residues were mutated to alanine (*rqc2_{aaa}*) rescued 60S recognition and the clearance of the stalling reporter almost as effectively as wild-type Rqc2p but did not support CAT tail synthesis (Fig. 4D and fig. S12). This CAT tail-deficient *rqc2_{aaa}* allele also failed to rescue Hsf1p transcriptional activation (Fig. 4E), indicating that CAT tails may promote Hsf1p activation.

Integrating our observations, we propose the model schematized in fig. S13. Ribosome stalling leads to dissociation of the 60S and 40S subunits, followed by recognition of the peptidyl-tRNA-60S species by Rqc2p and Ltn1p. Ltn1p ubiquitylates the stalled nascent chain, and this

Fig. 4. Rqc2p-dependent formation of CAT tails. (A, B, and D) Immunoblots of stalling reporters in RQC deletion strains. (C) Total amino acid analysis of immunoprecipitated GFP expressed in *ltn1Δ* and *ltn1Δrqc2Δ* strains, $n = 3$ independent immunoprecipitations. (E) Triplicate GFP levels measured with a flow cytometer and normalized to a wild-type control. EV, empty vector. All error bars are standard deviations.



leads to Cdc48 recruitment for extraction and degradation of the incomplete translation product. Rqc2p, through specific binding to Ala(IGC) and Thr(IGU) tRNAs, directs the template-free and 40S-free elongation of the incomplete translation product with CAT tails. CAT tails induce a heat shock response through a mechanism that is yet to be determined.

Hypomorphic mutations in the mammalian homolog of *LTNI* cause neurodegeneration in mice (21). Similarly, mice with mutations in a central nervous system-specific isoform of tRNA^{Arg} and GTPBP2, a homolog of yeast Hbs1 which works with PELOTA/Dom34 to dissociate stalled 80S ribosomes, suffer from neurodegeneration (22). These observations reveal the consequences that ribosome stalls impose on the cellular economy. Eubacteria rescue stalled ribosomes with the transfer-messenger RNA (tmRNA)-SmpB system, which appends nascent chains with a unique C-terminal tag that targets the incomplete protein product for proteolysis (23). The mechanisms used by eukaryotes, which lack tmRNA, to recognize and rescue stalled ribosomes and their incomplete translation products have been unclear. The RQC—and Rqc2p's CAT tail tagging mechanism in particular—bear both similarities and contrasts to the tmRNA *trans*-translation system. The evolutionary convergence upon distinct mechanisms for extending incomplete nascent chains at the C terminus argues for their importance in maintaining proteostasis. One advantage of tagging stalled chains is that it may distinguish them from normal translation products and facilitate their removal from the protein pool. An alternate, not mutually exclusive, possibility is that the extension serves to test the functional integrity of large ribosomal subunits, so that the cell can detect and dispose of defective large subunits that induce stalling.

REFERENCES AND NOTES

1. O. Brandman *et al.*, *Cell* **151**, 1042–1054 (2012).
2. Q. Defenuillere *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **10**, 1073/pnas.1221724110 (2013).
3. R. Verma, R. S. Oania, N. J. Kolawa, R. J. Deshaies, *eLife* **2**, e00308 (2013).
4. S. Shao, K. von der Malsburg, R. S. Hegde, *Mol. Cell* **50**, 637–648 (2013).
5. G.-W. Li, D. Burkhardt, C. Gross, J. S. Weissman, *Cell* **157**, 624–635 (2014).
6. D. Lyumkis *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 1702–1707 (2013).
7. S. Shao, R. S. Hegde, *Mol. Cell* **55**, 880–890 (2014).
8. R. Lill, J. M. Robertson, W. Wintermeyer, *Biochemistry* **25**, 3245–3255 (1986).
9. G. E. Katibah *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 12025–12030 (2014).
10. D. Chu, D. J. Barnes, T. von der Haar, *Nucleic Acids Res.* **39**, 6705–6714 (2011).
11. P. F. Agris, F. A. P. Vendeix, W. D. Graham, *J. Mol. Biol.* **366**, 1–13 (2007).
12. F. H. C. Crick, *J. Mol. Biol.* **19**, 548–555 (1966).
13. A. P. Gerber, W. Keller, *Science* **286**, 1146–1149 (1999).
14. E. Delannoy *et al.*, *Plant Cell* **21**, 2058–2071 (2009).
15. M. A. T. Rubio, F. L. Ragona, K. W. Gaston, M. Ibba, J. D. Alfonzo, *J. Biol. Chem.* **281**, 115–120 (2006).
16. R. E. Monro, *Nature* **223**, 903–905 (1969).
17. D. P. Letzring, K. M. Dean, E. J. Grayhack, *RNA* **16**, 2516–2528 (2010).
18. S. Ito-Harashima, K. Kuroha, T. Tatematsu, T. Inada, *Genes Dev.* **21**, 519–524 (2007).

19. K. Kobayashi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 17575–17579 (2010).
20. A. M. Burroughs, L. Aravind, *RNA Biol.* **11**, 360–372 (2014).
21. J. Chu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 2097–2103 (2009).
22. R. Ishimura *et al.*, *Science* **345**, 455–459 (2014).
23. S. D. Moore, R. T. Sauer, *Annu. Rev. Biochem.* **76**, 101–124 (2007).
24. A. Gerber, H. Grosjean, T. Melcher, W. Keller, *EMBO J.* **17**, 4780–4789 (1998).
25. K. W. Gaston *et al.*, *Nucleic Acids Res.* **35**, 6740–6749 (2007).
26. G. E. Crooks, G. Hon, J. M. Chandonia, S. E. Brenner, *Genome Res.* **14**, 1188–1190 (2004).

ACKNOWLEDGMENTS

Electron microscopy was performed at the University of Utah and the University of California. We thank D. Belnap (University of Utah) and M. Braumfeld (University of California, San Francisco) for supervision of the electron microscopes; A. Orendt and the Utah Center for High Performance Computing and the NSF Extreme Science and Engineering Discovery Environment consortium for computational support; D. Sidote (University of Texas at Austin)

for help processing RNA-seq data; and D. Herschlag and P. Harbury for helpful comments. Amino acid analysis was performed by J. Shulze at the University of California, Davis Proteomics Core. Edman sequencing was performed at Stanford University's Protein and Nucleic Acid Facility by D. Winant. This work was supported by the Searle Scholars Program (A.F.); Stanford University (O.B.); NIH grants 1DP2GM10772-01 (A.F.), GM37949, and GM37951 (A.M.L.); the Center for RNA Systems Biology grants P50 GM102706 (J.S.W.) and U01 GM098254 (J.S.W.); and the Howard Hughes Medical Institute (J.S.W.). The authors declare no competing financial interests. The cryo-EM structures have been deposited at the Electron Microscopy Data Bank (accession codes 2811, 2812, 6169, 6170, 6171, 6172, 6176, and 6201).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6217/75/suppl/DC1
Materials and Methods
Figs. S1 to S13
Table S1
References (27–41)

7 August 2014; accepted 14 November 2014
10.1126/science.1259724

CANCER ETIOLOGY

Variation in cancer risk among tissues can be explained by the number of stem cell divisions

Cristian Tomasetti^{1*} and Bert Vogelstein^{2*}

Some tissue types give rise to human cancers millions of times more often than other tissue types. Although this has been recognized for more than a century, it has never been explained. Here, we show that the lifetime risk of cancers of many different types is strongly correlated (0.81) with the total number of divisions of the normal self-renewing cells maintaining that tissue's homeostasis. These results suggest that only a third of the variation in cancer risk among tissues is attributable to environmental factors or inherited predispositions. The majority is due to “bad luck,” that is, random mutations arising during DNA replication in normal, noncancerous stem cells. This is important not only for understanding the disease but also for designing strategies to limit the mortality it causes.

Extreme variation in cancer incidence across different tissues is well known; for example, the lifetime risk of being diagnosed with cancer is 6.9% for lung, 1.08% for thyroid, 0.6% for brain and the rest of the nervous system, 0.003% for pelvic bone and 0.00072% for laryngeal cartilage (1–3). Some of these differences are associated with well-known risk factors such as smoking, alcohol use, ultraviolet light, or human papilloma virus (HPV) (4, 5), but this applies only to specific populations

exposed to potent mutagens or viruses. And such exposures cannot explain why cancer risk in tissues within the alimentary tract can differ by as much as a factor of 24 [esophagus (0.51%), large intestine (4.82%), small intestine (0.20%), and stomach (0.86%)] (3). Moreover, cancers of the small intestinal epithelium are three times less common than brain tumors (3), even though small intestinal epithelial cells are exposed to much higher levels of environmental mutagens than are cells within the brain, which are protected by the blood-brain barrier.

Another well-studied contributor to cancer is inherited genetic variation. However, only 5 to 10% of cancers have a heritable component (6–8), and even when hereditary factors in predisposed individuals can be identified, the way in which these factors contribute to differences in cancer incidences among different organs is obscure. For example, the same, inherited mutant *APC* gene is responsible for both the predisposition to colorectal and small intestinal cancers

¹Division of Biostatistics and Bioinformatics, Department of Oncology, Sidney Kimmel Cancer Center, Johns Hopkins University School of Medicine and Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, 550 North Broadway, Baltimore, MD 21205, USA.

²Ludwig Center for Cancer Genetics and Therapeutics and Howard Hughes Medical Institute, Johns Hopkins Kimmel Cancer Center, 1650 Orleans Street, Baltimore, MD 21205, USA.

*Corresponding author. E-mail: ctomasetti@jhmi.edu (C.T.); vogelbe@jhmi.edu (B.V.)

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of February 13, 2015):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/347/6217/75.full.html>

Supporting Online Material can be found at:

<http://www.sciencemag.org/content/suppl/2014/12/31/347.6217.75.DC1.html>

This article **cites 40 articles**, 20 of which can be accessed free:

<http://www.sciencemag.org/content/347/6217/75.full.html#ref-list-1>

This article appears in the following **subject collections**:

Biochemistry

<http://www.sciencemag.org/cgi/collection/biochem>