Title
Global analysis of translation termination in E. coli

Author
Baggett, Natalie Elise

Publication Date
2017

Peer reviewed|Thesis/dissertation

Global analysis of translation termination in *E. coli*

by

Natalie Elise Baggett

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry and Molecular Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Acknowledgements

I would like to thank my advisor, Carol Gross, who consistently overestimated my abilities and thereby allowed me to accomplish more than I thought possible. Her scientific guidance and support were crucial to the development and follow-through of this project. As well as my thesis committee members, Jonathan Weissman and Jeffery Cox, who provided sound guidance during this scientific endeavor.

To my parents, Patricia and Steven Baggett, who were relentless in their support of me throughout my entire life, especially during my graduate career. Thank you for all the sacrifices you have made for my education and me.

To Sam Rodely, my partner in adventure and life, who constantly believed in me through this entire journey even when I did not. Thank you being a perpetual reminder about what matters.

My graduate career would not have been as bright if not for the members of the Carol Gross Lab, who provided both scientific insight and solace; Monica Guo, Virgil Rhodius, Andrew Gray, Byoung Mo Koo, Bentley Lim, David Burkhardt, Anthony Shiver, Jason Peters, Brian Sharon, John Hawkins, Melanie Silvis (Yay, Melanie!), Candy Lu, Hendrik Osadnik, Yan Zhang, Marco Jost, Horia Todor, Jeff Hussmann, Danh Le and Jennifer Thompson.

Parts of the text in this thesis are a reprint of the publication as it appears in Public Library of Science (PLOS)- Genetics. The co-author listed in this publication, Carol Gross, directed and supervised the research that forms the basis for the thesis.

Statement from Carol Gross

This thesis includes reprints of material published with coauthors other than Natalie Baggett. In Chapter 1, Natalie conceived of and performed all experiments, Natalie and Yan Zhang both performed computation analysis of ribosome profiling data. The manuscript was co written by myself, Natalie and Yan Zhang.

Global analysis of translation termination in *E. coli*

Natalie Baggett

Abstract

Terminating protein translation accurately and efficiently is critical for both protein fidelity and ribosome recycling for continued translation. The three bacterial release factors (RFs) play key roles: RF1 and 2 recognize stop codons and terminate translation; and RF3 promotes disassociation of bound release factors. Probing release factors mutations with reporter constructs containing programmed frameshifting sequences or premature stop codons had revealed a propensity for readthrough or frameshifting at these specific sites, but their effects on translation genome-wide have not been examined.

We provide the first global assessment of translation termination at native loci within the *E. coli* genome using a set of well-characterized release factor mutations. This approach allowed us to study loci within their native termination context and to simultaneously monitor changes in expression of key release factors in response to perturbations of termination. Strains with increasingly severe release factor defects exhibit increasingly severe accumulation of ribosomes over stop codons, indicative of an increased duration of the termination/release phase of translation. Release factor mutant strains also exhibit increased occupancy in the region following the stop codon at a significant number of genes. Our global analysis revealed that, as expected, translation termination is generally efficient and accurate, but that at a significant number of genes (≥ 50) the ribosome signature after the stop codon is suggestive of translation past the

stop codon, indicating some genes are natively susceptible to low termination efficiency. *E. coli* K-12 has been previously identified to be acutely sensitive to perturbations in translation termination. Here we show even native *E. coli* K-12 exhibits the ribosome signature suggestive of protein extension, especially at UGA codons, which rely exclusively on the reduced function RF2 allele of the K-12 strain for termination. Deletion of RF3 increases the severity of these defects. We unambiguously demonstrate readthrough and frameshifting protein extensions and their further accumulation in mutant strains for a few select cases. In addition to enhancing recoding, ribosome accumulation over stop codons disrupts attenuation control of biosynthetic operons, and may alter expression of some overlapping genes. Together, these functional alterations may either augment the protein repertoire or produce deleterious proteins.

Table of Contents

List of Tables

Chapter 1. Global analysis of translation termination in *E. coli*

List of Figures

Chapter 1. Global analysis of translation termination in *E. coli*

Introduction

Overview

Translation of mRNAs to synthesize proteins is an energetically costly process within the cell. At each stage of translation, initiation, elongation and termination, there are layers of mechanisms to ensure each step retains a high level of fidelity. Each amino acid addition to the polypeptide chain requires breakage of 4 phosphate bonds; meaning the cell expends a tremendous amount of energy to create its proteome. The desired result is the production of accurate, full-length proteins, which can be achieved due to the multi-step control mechanisms. Stresses to the cell through the manipulation of factors involved in translation test the boundaries of cellular tolerance to errors in translation fidelity. Translation is an essential process within the cell, therefore only some perturbations can be studied while the cell is still viable. Interestingly, a point mutation of an essential release factor in *E. coli* K-12 leaves it especially sensitive to further perturbations.

In my thesis work I focused on translation termination efficiency both in native *E. coli* K-12 and other release factor mutations. Using ribosome profiling I wanted to identify the molecular signatures of termination efficiency and how they change in strains with altered release factors. By using this global technology, I could build upon and expand existing knowledge of translation termination efficiency.

Translation initiation

In the initiation phase of translation, ribosomal subunits are assembled on mRNA transcripts at the initiation codon. The initiation codon is recognized by formyl methionine-tRNA (fMet-tRNA) in the peptidyl-site or P-site of the ribosome. The vast majority of transcripts utilize AUG as the initiation codon that pairs with fMet-tRNAs to initiate transcription [1]. While other initiation codons can be utilized, the efficiency of initiation drops at non-AUG codons [2]. In *E. coli*, 14% of genes use GUG as an alternative initiation codon [3]. All other codons are recognized by aminoacyl-tRNAs in the aminoacyl-site or A-site during translation.

Three initiation factors (IFs) together ensure assembly of the ribosomal units, increasing fidelity of initiation [4,5]. IF1, IF2 and IF3 all promote the binding of fMet-tRNA to the initiation codon and 30S subunit [4] to help to ensure fidelity of translation [6]. In addition to specific roles during translation initiation, IF1 and IF3 both play important roles after 70S ribosome dissociation following translation termination [7]. IF3 binds to newly released 30S subunits to enhance dissociation of bound tRNA and mRNA, it also prevents premature formation of the 70S ribosomal complex [7,8].

In addition to *trans* acting factors that increase the efficiency of initiation there are *cis* elements along the transcript itself. The primary *cis* element is the Shine-Dalgarno (SD) sequence, which is located 10-16 bp upstream of the initiation codon and interacts with the 3' end of 16S rRNA [9]. These interactions between the SD sequence on the mRNA and the 30S ribosomal subunit are the basis for SD-dependent translation initiation [5]. The addition of IF1, IF2, IF3 and fMet-tRNA forms a 30S pre-initiation complex [4–6]. Subsequent recruitment of the 50S subunit forms the 70S initiation complex, which is primed for the next step of translation, elongation [6,10,11]. The composition, length and spacing of the SD sequence from the initiation codon all contribute to the efficiency of translation [12].

Although the SD sequence was one of the earliest recognized determinants of initiation, its impact on translation is not completely clarified. Recent work shows that for endogenous mRNAs, the level of structure of that mRNA is the most informative variable for the rate of protein production/mRNA. Structure is inversely related to protein production, with the most structured mRNAs show the lowest rate of protein production [13]. Indeed, the fraction of SD usage of SD-dependent initiation varies in genomes between 11.6% and 90.8% [14], suggesting that initiation can occur in the absence of an SD sequence. These mechanisms are collectively known as SD-independent initiation. One mechanism utilizes an interaction between ribosomal protein S1 and single stranded regions of mRNA upstream of the initiation codon to enhance

SD-independent initiation [15]. The interactions between the S1 protein and 5'UTR sequences can also improve rates of SD-dependent initiation on transcripts with poor SD sequences [16].

Finally, some transcripts without a 5'UTR are translated by a leaderless mechanism [17]. In contrast to messages with 5'UTRs, where ribosomes assemble in a stepwise fashion of 30S then 50S subunit association, leaderless messages begin with the complete 70S complex assembled over the initiation codon [17,18]. The cellular levels of two initiation factors, IF2 and IF3, which can change in stressed conditions, heavily influence the translation efficiency of leaderless messages [19]. Notably leaderless mRNA transcripts are higher in prevalence in Gram-positive bacteria than Gram-negative bacteria suggesting that although a rare occurrence in *E. coli,* other organisms utilize this mechanism to control translation efficiency [17].

Translation elongation

Once the 70S ribosomal complex has been formed atop the initiation codon in the P-site, the codon in the A-site is recognized by an aminoacyl-tRNA and translation elongation begins. The process of elongation contains three main steps; decoding of the A-site codon by an aminoacyl-tRNA, peptide bond formation which adds the aminoacyl group to the growing polypeptide chain, then translocation of the ribosome forward along the mRNA.

Recognition of the A-site codon by the proper aminoacyl-tRNA occurs with the help of elongation factor, EF-Tu, and GTP. Once the proper aminoacyl-tRNA has bound to the codon in the A-site, GTP is hydrolyzed and EF-Tu is released. Fidelity at this decoding step is ensured by both the codon-anti codon interaction and by structural elements of the ribosome, resulting in a misincorporation rate of 1 in 10^4 [20]. Another level of fidelity at this step is achieved by assuring that tRNAs are charged with the proper amino acid. The protein family that attaches cognate amino acid and tRNA pairs is called aminoacyl-tRNA synthetases (ARSs). Through induced fit mechanisms, proofreading functionality and multiple checkpoints, ARSs only attach miscognate amino acids 1 in 10^5 events [21].

After decoding of the A-site codon by an aminoacyl-tRNA through the accommodation step, the aminoacyl group of the tRNA is oriented into the peptidyl transfer center (PTC). Miscognate tRNAs accommodation occurs much slower than cognate tRNAs, allowing for an additional level of fidelity [22]. The placement of the aminoacyl group into the PTC allows for peptidyl-transferase reaction, in which aminoacyl group of the A-site tRNA forms a new peptide bond with the existing polypeptide attached to the P-site peptidyl-tRNA, a reaction catalyzed by the ribosome itself [23]. This reaction adds one amino acid to the growing polypeptide chain and leaves the newly formed peptidyl-tRNA in the A-site.

Once the peptidyl-tRNA is formed in the A-site, translocation occurs with the help of elongation factor, EF-G and GTP. The presence of EF-G and the hydrolysis of GTP together enable a ratchet movement between the 30S and 50S subunits that moves the ribosomal complex one codon forward [24]. During this movement the tRNAs transition through hybrid states, known as A/P and P/E states, during which tRNA position on the 50S subunit has moved forward, but relative to the 30S subunit and mRNA the tRNAs are still in the starting position. These hybrid states are resolved by hydrolysis of GTP by EF-G [25]. Completed translocation results in the release of uncharged tRNA from the exit-site, or E-site, places the peptidyl-tRNA in the P-site and aligns the next codon in the now empty A-site. Returning the ribosome to a state where decoding of the A-site codon can begin again. Each round of elongation, decoding, peptide hydrolysis and translocation results in one addition amino acid being added to the growing polypeptide chain.

Although the process of translation elongation occurs with high accuracy, errors are made at a rate of 1 in 10^4 incorporations. All the mechanisms to ensure fidelity during translation explained up to this point prevent incorporation of miscognate amino acids into the polypeptide chain. A proofreading mechanism of translation elongation was proposed by Zaher et al. during which premature termination is triggered on transcripts that have misincorporated an amino acid [26]. This mechanism was coined post-peptidyl-transfer quality control, or post PT-QC, and

involves release factors, RF2 and RF3, to trigger early translation termination of peptides made by decoding errors [26]. Once misincorporation is identified through post PT-QC the transcript is prematurely terminated rather than continuing to synthesize an error filled protein [26].

Translation termination

Elongation continues until the ribosomal complex reaches a stop codon. *E. coli* has three stop codons: UAA, UGA and UAG. The recognition of the stop codon occurs by a class I release factor, either RF1 or RF2. Both class I release factors are tRNA mimics. They decode the stop codon in a similar manner to interactions achieved by standard codon and aminoacyl-tRNA binding, but instead of adding an amino acid to the growing polypeptide chain they catalyze peptide hydrolysis and release through their conserved GGQ motif located at the peptidyl transfer center (PTC) [27]. The two class 1 RFs recognize stop codons with overlapping specificity: RF1 recognizes UAA and UAG whereas RF2 recognizes UAA and UGA. Once a stop codon enters the A-site it is decoded by RF1 or RF2, a process that occurs with the same level of fidelity as tRNA selection during elongation, preventing premature termination [28].

A third release factor, RF3, a class II RF is a GTPase that binds in a similar position as EF-G, outside of the A-site and catalyzes the release of RF1 or RF2 by a conformational chain catalyzed by GTP hydrolysis [29]. Interestingly RF3 is non-essential in *E. coli* K-12, unlike eukaryotic eRF3 in eukaryotes that seems to carry out a similar activity but through a different mechanism [30,31]. The rapid removal of class I RFs by RF3 results in a post-termination complex, which allows ribosome recycling to begin and frees class I RFs maintaining their proper cellular concentration [32,33]. Along with ribosome release factor, RRF, and elongation factor, EF-G the post-termination complex is fully disassembled [34]. The deacylated tRNA in the P-site is released while the 30S and 50S ribosomal subunits are freed from the mRNA message. Once the complex is disassembled, initiation factor, IF3, binds to the freed 30S subunit to prevent immediate re-association with the larger subunit.

Fidelity at translation termination is ensured by rapid recognition of stop codons by class I RFs, subsequent peptide cleavage catalyzed by the GGQ motif, and prompt recycling of the ribosomal complex. Decreasing the efficiency of any of these steps can reduce the fidelity of termination. Many mutations of class I RFs, RF1 and RF2, result increase stop codon readthrough as measured in *lacZ* reporter constructs [35–39]. The GGQ motif catalyzing peptide cleavage is found in class I RFs is universally conserved in all kingdoms of life, emphasizing its importance in rapid and efficient protein termination [40]. Recently it was discovered that this motif is methylated by the methyltransferase, PrmC, and that this methylation markedly improves termination *in vivo* for both RF1 and RF2 [41–44]. There are several key proteins utilized in recycling of the post-termination complex that all serve to expedite ribosome turn over: RF3, RRF, EF-G and IF3. Reduction of ribosome recycling efficiency by mutation or depletion of these key proteins reduces the fidelity of translation and results in a build up of ribosomes at stop codons [45,46]. While a subset of ribosome complexes have all the proper accessory proteins (RF1/2, RF3, RRF, etc.) to accurate terminate they are not rapidly recycled which sequesters the available factors. The remaining ribosomes are positioned over the stop codon but lack a class I RF to begin termination, these complexes are at a greater risk for recoding events because the effective cellular concentration of RF1/2 is so low. The increased dwell time over stop codons also decreases the accuracy of termination.

Recoding

Reduction in termination fidelity manifests in an increase of recoding, such as stop codon readthrough or frameshifting. There is a delicate balance for accurate translation termination; too strong and the cell is at risk for premature termination, too weak and the cell is prone to recoding events which can create elongated proteins. The majority of extended proteins synthesized with reduced termination fidelity are likely deleterious to the cell, due to either the large energy expense to the cell to produce these proteins or direct negative effects of the

extended proteins. However there are some specific instances where reduced termination efficiency is programmed in the genome and utilized to expand the proteome. These cases are known as programmed ribosomal frameshift events and occur across many forms of life with varying frequency between species, including *E. coli*, Saccharomyces *cerevisae* and *Drosophila melanogaster* [47–49].

In addition to all of the *trans*-acting factors utilized to ensure effective translation termination, there are many *cis*-acting sequences in the mRNA that alter the efficacy of any given stop codon. These sequence-based *cis* elements are largely responsible for instances of programmed ribosome frameshifts or other programmed recoding events. There are two main classes of these elements surrounding stop codons: slippery sequences and pseudoknots. Slippery sequences refer to a region of mRNA where ribosomes are prone to a shift in reading frame [50]. In *E. coli,* slippery sequences include internal Shine-Dalgarno sequences that utilize interactions with the ribosome to shift mRNA into a new reading frame [51,52]. A pseudoknot is a structural element of RNA that upon formation, and often in conjunction with a slippery sequence, can cause translating ribosomes to shift into a new reading frame [53,54]. These elements acting alone or together can greatly increase the rate at which frameshifting occurs on a specific mRNA.

Thesis plan

Of particular interest in my work is how the translation termination machinery reacts to perturbations of release factors. We used ribosome profiling to identify the position of ribosomes along mRNA transcripts at a given time point for each strain. This snapshot of ribosomes within a cell allowed us to monitor the efficiency of termination. These studies also allowed us to examine consequences of reduced termination efficiency. We first looked at ribosome occupancy over stop codons, as this metric reveals whether ribosomes spend more time at a stop codon. Enhanced occupancy indicates that termination is taking longer than in the wild-

type cell. Also of particular interest was determining the pattern of ribosome occupancy in the region following the annotated stop codon for each gene. Ribosomes downstream of the annotated stop could indicate that ribosomes translate past the stop codon, especially when a drop-off in their frequency coincides with a stop codon in the post-ORF region. The current limitations of ribosome profiling in *E. coli* required follow-up biochemical experiments, which allowed us to conclude that these were translating ribosomes. Because we took a global approach to monitor termination efficiency we determined whether perturbations to release factors had differing impacts throughout the cell that were dependent on stop codon identity. We also looked specifically at genes that have a unique regulation (i.e. leader peptide attenuation) and structural organization (i.e. overlapping ORFs) to determine if disrupting termination efficiency could have more wide reaching impacts that change gene regulation. All of this work and discussion of our findings can be found in Chapter 1.

We attempted to modify the ribosome profiling protocol to preserve reading frame information of the ribosome footprint, which is currently possible in other organisms like *Saccharomyces cerevisae* and *Drosophila melanogaster* but not *E. coli* [48,49,55]. This would further enhance our studies into termination efficiency by easily allowing identification of translating ribosomes past the stop codon. A summary of this unsuccessful attempt is found in Appendix A.

As mentioned above there are some recoding events found in *E. coli* that are programmed within the genome to expand the functional repertoire. A strong candidate for a novel programmed frameshift event was identified in *cspG* during studies of cold temperature stress. The first observed phenotype suggested possible temperature dependent recoding, that resulted in an extended product a low temperature. Follow up experiments to identify functionality of an extended CspG protein and our findings are found in Appendix B.

References

1. Kozak M. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. Gene. 2005;361: 13–37. doi:10.1016/j.gene.2005.06.037
2. Sussman JK, Simons EL, Simons RW. Escherichia coli translation initiation factor 3 discriminates the initiation codon in vivo. Mol Microbiol. 1996;21: 347–360. doi:10.1046/j.1365-2958.1996.6371354.x
3. Blattner FR, Iii GP, Bloch CA, Perna NT, Burland V, Riley M, et al. The Complete Genome Sequence of Escherichia coli K-12. Science (80-). 1997;277: 1453–1462. doi:10.1126/science.277.5331.1453
4. Gualerzi CO, Pon CL. Initiation of mRNA translation in prokaryotes. Biochemistry. 1990;29: 5881–5889.
5. Malys N, McCarthy JEG. Translation initiation: variations in the mechanism can be anticipated. Cell Mol Life Sci. 2011;68: 991–1003. doi:10.1007/s00018-010-0588-z
6. Benne R, Ebes F, Voorma HO. Sequence of events in initiation of protein synthesis. Eur J Biochem. 1973;38: 265–273.
7. Antoun A, Pavlov MY, Lovmar M, Ehrenberg M. How initiation factors tune the rate of initiation of protein synthesis in bacteria. EMBO J. 2006;25: 2539–50. doi:10.1038/sj.emboj.7601140
8. Peske F, Rodnina M V., Wintermeyer W. Sequence of steps in ribosome recycling as defined by kinetic analysis. Mol Cell. 2005;18: 403–412. doi:10.1016/j.molcel.2005.04.009
9. Shine J, Dalgarno L. The 3’-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc Natl Acad Sci U S A. 1974;71: 1342–6. doi:10.1073/pnas.71.4.1342
10. Luchin S, Putzer H, Hershey JWB, Cenatiempo Y, Grunberg-manago M, Laalami S. In

Vitro Study of Two Dominant Inhibitory GTPase Mutants of Escherichia coli Translation Initiation Factor IF2. J Biol Chem. 1999;274: 6074–6079.

11. Antoun A, Pavlov MY, Andersson K, Tenson T, Ehrenberg M. The roles of initiation factor 2 and guanosine triphosphate in initiation of protein synthesis. EMBO J. 2003;22: 5593–5601. doi:10.1093/emboj/cdg525

12. Kozak M. Initiation of translation in prokaryotes and eukaryotes. Gene. 1999;234: 187–208. Available: http://www.ncbi.nlm.nih.gov/pubmed/10395892

13. Burkhardt DH, Rouskin S, Zhang Y, Li G-W, Weissman JS, Gross CA. Operon mRNAs are organized into ORF-centric structures that predict translation efficiency. Elife. 2017;6: e22037. doi:10.7554/eLife.22037

14. Chang B, Halgamuge S, Tang SL. Analysis of SD sequences in completed microbial genomes: Non-SD-led genes are as common as SD-led genes. Gene. 2006;373: 90–99. doi:10.1016/j.gene.2006.01.033

15. Tzareva N V., Makhno VI, Boni I V. Ribosome-messenger recognition in the absence of the Shine-Dalgarno interactions. FEBS Lett. 1994;337: 189–194. doi:10.1016/0014-5793(94)80271-8

16. Komarova A V, Tchufistova LS, Supina E V, Boni I V. Protein S1 counteracts the inhibitory effect of the extended Shine-Dalgarno sequence on translation. RNA. 2002;8: 1137–1147. doi:10.1017/S1355838202029990

17. Moll I, Grill S, Gualerzi CO, Bläsi U. Leaderless mRNAs in bacteria: Surprises in ribosomal recruitment and translational control. Mol Microbiol. 2002;43: 239–246. doi:10.1046/j.1365-2958.2002.02739.x

18. Balakin a G, Skripkin E a, Shatsky IN, Bogdanov a a. Unusual ribosome binding properties of mRNA encoding bacteriophage lambda repressor. Nucleic Acids Res. 1992;20: 563–71. doi:10.1093/nar/20.3.563

19. Grill S, Moll I, Hasenöhrl D, Gualerzi CO, Bläsi U. Modulation of ribosomal recruitment to

5′-terminal start codons by translation initiation factors IF2 and IF3. FEBS Lett. 2001;495: 167–171. doi:10.1016/S0014-5793(01)02378-X

20. Zaher HS, Green R. Fidelity at the Molecular Level: Lessons from Protein Synthesis. Cell. Elsevier Inc.; 2009;136: 746–762. doi:10.1016/j.cell.2009.01.036

21. Francklyn CS. DNA Polymerases and Aminoacyl-tRNA Synthetases Shared Mechanisms for Ensuring the Fidelity of Gene Expression. Biochemistry. 2008;47.

22. Pape T, Wintermeyer W, Rodnina M. Induced fit in initial selection and proofreading of aminoacyl-tRNA on the ribosome. EMBO J. 1999;18: 3800–3807. doi:10.1093/emboj/18.13.3800

23. Sievers A, Beringer M, Rodnina M V, Wolfenden R. The ribosome as an entropy trap. PNAS. 2004;101: 12397–12398.

24. Valle M, Zavialov A, Sengupta J, Rawat U, Ehrenberg M, Frank J. Locking and unlocking of ribosomal motions. Cell. 2003;114: 123–134. doi:10.1016/S0092-8674(03)00476-8

25. Ehrenberg M. Protein synthesis: Translocation in slow motion. Nature. 2010;466: 325–326. doi:10.1038/466325a

26. Zaher HS, Green R. A primary role for release factor 3 in quality control during translation elongation in Escherichia coli. Cell. Elsevier Inc.; 2011;147: 396–408. doi:10.1016/j.cell.2011.08.045

27. Trobro S, Aqvist J. Mechanism of the translation termination reaction on the ribosome. Biochemistry. 2009;48: 11296–11303.

28. Jørgensen F, Adamski F, Tate W, Kurland C. Release factor-dependent false stops are infrequent in Escherichia coli. J Mol Biol. 1993;230: 41–50.

29. Gao H, Zhou Z, Rawat U, Huang C, Bouakaz L, Wang C, et al. RF3 Induces Ribosomal Conformational Changes Responsible for Dissociation of Class I Release Factors. Cell. 2007;129: 929–941. doi:10.1016/j.cell.2007.03.050

30. Nakamura Y, Ito K, Isaksson L. Emerging Understanding of Translation Termination. Cell.

1996;87: 147–150.

31. Zhouravleva G, Frolova L, Le Goff X, Le Guellec R, Inge-Vechtomov S, Kisselev L, et al. Termination of translation in eukaryotes is governed by two interacting polypeptide chain release factors, eRF1 and eRF3. EMBO J. European Molecular Biology Organization; 1995;14: 4065–72.

32. Heurgue-Hamard V, Karimi R, Mora L, MacDougall J, Leboeuf C, Grentzmann G, et al. Ribosome release factor RF4 and termination factor RF3 are involved in dissociation of peptidyl-tRNA from the ribosome. EMBO J. 1998;17: 808–816. doi:10.1093/emboj/17.3.808

33. Zavialov A V., Hauryliuk V V., Ehrenberg M. Splitting of the posttermination ribosome into subunits by the concerted action of RRF and EF-G. Mol Cell. 2005;18: 675–686. doi:10.1016/j.molcel.2005.05.016

34. Hirokawa G, Demeshkina N, Iwakura N, Kaji H, Kaji A. The ribosome-recycling step: consensus or controversy? Trends Biochem Sci. 2006;31: 143–9. doi:10.1016/j.tibs.2006.01.007

35. Rydén SM, Isaksson LA. A temperature-sensitive mutant of Escherichia coli that shows enhanced misreading of UAG/A and increased efficiency for tRNA nonsense suppressors. MGG Mol Gen Genet. 1984;193: 38–45. doi:10.1007/BF00327411

36. Kawakami K, Inada T, Nakamura Y. Conditionally lethal and recessive UGA-suppressor mutations in the prfB gene encoding peptide chain release factor 2 of Escherichia coli. J Bacteriol. 1988;170: 5378–5381.

37. Kawakami K, Nakamura Y. Autogenous suppression of an opal mutation in the gene encoding peptide chain release factor 2. Proc Natl Acad Sci U S A. 1990;87: 8432–6.

38. Elliott T, Wang X. Salmonella typhimurium prfA Mutants Defective in Release Factor 1. J Bacteriol. 1991;173.

39. Mikuni O, Kawakami K, Nakamura Y. Sequence and functional analysis of mutations in

the gene encoding peptide-chain-release factor 2 of Escherichia coli. Biochimie. 1991;73: 1509–1516. doi:10.1016/0300-9084(91)90185-4

40. Frolova LY, Tsivkovskii RY, Sivolobova GF, Oparina NY, Serpinsky OI, Blinov VM, et al. Mutations in the highly conserved GGQ motif of class 1 polypeptide release factors abolish ability of human eRF1 to trigger peptidyl-tRNA hydrolysis.

41. Dinçbas-Renqvist V, Engström Å. A post-translational modification in the GGQ motif of RF2 from Escherichia coli stimulates termination of translation. EMBO …. 2000;19. doi:10.1093/emboj/19.24.6900

42. Heurgue-Hamard V, Champ S, Engstrom A, Ehrenberg M, Buckingham RH. The hemK gene in Escherichia coli encodes the N5-glutamine methyltransferase that modifies peptide release factors. EMBO J. 2002;21.

43. Nakahigashi K, Kubo N, Narita S, Shimaoka T, Goto S, Oshima T, et al. HemK, a class of protein methyl transferase with similarity to DNA methyl transferases, methylates polypeptide chain release factors, and hemK knockout induces defects in translational termination. Proc Natl Acad Sci U S A. 2002;99: 1473–8. doi:10.1073/pnas.032488499

44. Mora L, Heurgué-Hamard V, De Zamaroczy M, Kervestin S, Buckingham RH. Methylation of bacterial release factors RF1 and RF2 is required for normal translation termination in vivo. J Biol Chem. 2007;282: 35638–35645. doi:10.1074/jbc.M706076200

45. Freistroffer D V, Pavlov MY, MacDougall J, Buckingham RH, Ehrenberg M. Release factor RF3 in E.coli accelerates the dissociation of release factors RF1 and RF2 from the ribosome in a GTP-dependent manner. EMBO J. 1997;16: 4126–33. doi:10.1093/emboj/16.13.4126

46. Janssen BD, Hayes CS. Kinetics of paused ribosome recycling in Escherichia coli. J Mol Biol. Elsevier Ltd; 2009;394: 251–67. doi:10.1016/j.jmb.2009.09.020

47. Antonov I, Coakley A, Atkins JF, Baranov P V, Borodovsky M. Identification of the nature

of reading frame transitions observed in prokaryotic genomes. Nucleic Acids Res. 2013;41: 6514–30. doi:10.1093/nar/gkt274

48. Ingolia NT, Ghaemmaghami S, Newman JRS, Weissman JS. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science. 2009;324: 218–23. doi:10.1126/science.1168978

49. Dunn JG, Foo CK, Belletier NG, Gavis ER, Weissman JS. Ribosome profiling reveals pervasive and regulated stop codon readthrough in Drosophila melanogaster. Elife. 2013;2: e01179. doi:10.7554/eLife.01179

50. Gurvich O, Baranov P. Sequences that direct significant levels of frameshifting are frequent in coding regions of Escherichia coli. EMBO J. 2003;22: 5941–5950. doi:10.1093/emboj/cdg561

51. Larsen B, Peden J, Matsufuji S, Matsufuji T, Brady K, Maldonado R, et al. Upstream stimulators for recoding. Biochem Cell Biol. 1995;73: 1123–1129. doi:10.1139/o95-121

52. Márquez V, Wilson D, Tate W, Triana-Alonso F, Nierhaus KH. Maintaining the ribosomal reading frame: the influence of the E site during translational regulation of release factor 2. Cell. 2004;118: 45–55.

53. Tzeng TH, Tu CL, Bruenn JA. Ribosomal frameshifting requires a pseudoknot in the Saccharomyces cerevisiae double-stranded RNA virus. J Virol. 1992;66: 999–1006.

54. Staple DW, Butcher SE. Pseudoknots: RNA structures with diverse functions. PLoS Biol. 2005;3: e213. doi:10.1371/journal.pbio.0030213

55. Oh E, Becker AH, Sandikci A, Huber D, Chaba R, Gloge F, et al. Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. Cell. 2011;147: 1295–308. doi:10.1016/j.cell.2011.10.044

Chapter 1

Global analysis of translation termination in *E. coli*

Abstract

Terminating protein translation accurately and efficiently is critical for both protein fidelity and ribosome recycling for continued translation. The three bacterial release factors (RFs) play key roles: RF1 and 2 recognize stop codons and terminate translation; and RF3 promotes disassociation of bound release factors. Probing release factors mutations with reporter constructs containing programmed frameshifting sequences or premature stop codons had revealed a propensity for readthrough or frameshifting at these specific sites, but their effects on translation genome-wide have not been examined. We performed ribosome profiling on a set of isogenic strains with well-characterized release factor mutations to determine how they alter translation globally. Consistent with their known defects, strains with increasingly severe release factor defects exhibit increasingly severe accumulation of ribosomes over stop codons, indicative of an increased duration of the termination/release phase of translation. Release factor mutant strains also exhibit increased occupancy in the region following the stop codon at a significant number of genes. Our global analysis revealed that, as expected, translation termination is generally efficient and accurate, but that at a significant number of genes (≥ 50) the ribosome signature after the stop codon is suggestive of translation past the stop codon. Even native *E. coli* K-12 exhibits the ribosome signature suggestive of protein extension, especially at UGA codons, which rely exclusively on the reduced function RF2 allele of the K-12 strain for termination. Deletion of RF3 increases the severity of the defect. We unambiguously demonstrate readthrough and frameshifting protein extensions and their further accumulation in mutant strains for a few select cases. In addition to enhancing recoding, ribosome accumulation over stop codons disrupts attenuation control of biosynthetic operons, and may alter expression of some overlapping genes. Together, these functional alterations may either augment the protein repertoire or produce deleterious proteins.

Introduction

Ribosomes translate the genetic information in the mRNA to a linear sequence of amino acids in the polypeptide chain through a process consisting of initiation, elongation, termination, and recycling. During initiation, the 30S subunit of the bacterial ribosome and various initiation factors assemble at the initiation codon on the mRNA. Elongation commences after the 50S subunit of the ribosome joins the complex. Cognate aminoacyl tRNAs, together with elongation factors decode the mRNA sequentially, binding first at the acceptor site (A site), followed by movement to the P site after amino acid transfer to the polypeptide chain at the peptidyl-transferase center. Several layers of error correction minimize the misincorporation of non-cognate amino acids [1,2]. Termination is signaled when a stop codon (UAA, UAG, UGA) enters the A site of the ribosome, where it is recognized either by release factor (RF) 1 or 2 [3]. RF1 or RF2 hydrolyze the polypeptide chain to terminate translation, and are then dissociated from the ribosome by RF3 during recycling [4]. Peptide release is a high fidelity process (error frequency of approximately 10^{-5}), ensuring that stop codon recognition precedes peptide release [5–7]. Finally, the ribosome is dissociated to its 30S and 50S subunits by elongation factor EF-G, and ribosome recycling factor (RRF) [8–10].

There is increasing structural and biochemical understanding of the three bacterial release factors. RF1 and RF2 are structural mimics of an aminoacyl tRNA and both are essential in native *E. coli* K-12 [11,12]. They bind in the A site using conserved protein motifs in Domain 2 to recognize the 2nd and 3rd positions of the stop codons (RF1: U<u>AA</u> and U<u>AG</u>; RF2: U<u>AA</u> and U<u>GA</u>) [13]. Their universally conserved GGQ amino acid motif then reaches into the peptidyl transferase center to release the peptide chain by catalyzing its hydrolysis from the tRNA [12,13]. Methylation of RF1 and RF2 by PrmC at their GGQ motif enhances release factor activity [14].

Interestingly, *E. coli* K-12 strains have an RF2 variant with Thr at position 246 rather than the canonical Ala246 or Ser246 [15]. All other bacteria, including other *E. coli* lineages,

have Ala or Ser at position 246 [16]. In this work, we call the *E. coli* K-12 RF2 variant, $RF2^{K-12}$, and the *E. coli B* variant with Ala246, $RF2^{B}$. $RF2^{K-12}$ is discrepant from $RF2^{B}$ in its properties. First, $RF2^{K-12}$ has reduced ability to catalyze hydrolysis of the peptide bond to terminate translation relative to $RF2^{B}$ [15]. Second, $RF2^{K-12}$ but not $RF2^{B}$ is almost completely dependent on methylation for activity [14]. Third, UAA decoding is done primarily by RF1 in strains with $RF2^{K-12}$, but by both RF1 and RF2 in strains with $RF2^{B}$ [14,15]. Indeed, RF1 is non-essential in an *E. coli* K-12 strain with $RF2^{B}$ [17,18]. Finally, because the level of RF2 is tuned to need via an internal UGA frameshifting event necessary to produce the full-length protein, $RF2^{K-12}$ is present at a higher level than $RF2^{B}$ as expected from its lower activity [14,19,20].

RF3, a non-essential release factor of *E. coli*, promotes dissociation of RF1 and RF2 from the ribosome [21,22]. This reaction occurs slowly in the absence of RF3 [4,23]. RF3 is a homologue of EF-G, a GTPase translocation factor that catalyzes movement of tRNA on the ribosome and ribosome dissociation [24]. The current idea is that RF3-GTP binds to the ribosome in the same location as EF-G, and similarly induces inter-subunit rotation, which creates a steric clash with the bound RFs, promoting dissociation of RF1 and RF2 [25,26].

The role of the release factors in translation termination fidelity has been explored *in vivo* by measuring by release of small artificial peptides or *in vitro* by measuring frameshifting or stop codon readthrough in a synthetic constructs containing a known frameshifting site or premature stop codon [16,27–32]. These assays showed that $RF2^{B}$ and methylated RF1 and RF2 terminated translation better than $RF2^{K-12}$, or un-methylated RF1 and RF2 [9,14,16,33–35]. Additionally, temperature sensitive mutations within RF1 and RF2 increase stop codon readthrough on *in vitro* constructs [29,31,32,36,37]. Suppressors of these temperature sensitive mutations mapped to the *prfC* (RF3) locus [38]. Further studies indicate that strains lacking RF3 have increased stop codon readthrough of *lacZ* reporter constructs, and that suppression and increased frameshifting over the known *prfB* (RF2) frameshifting sequence [9,33–35,39]. Overexpression of RF3 was shown to decrease frameshifting over the *prfB* frameshifting

sequence [34,39]. These studies and RF3 overexpression studies indicated that cooperative interactions between RF1/2 and RF3 improved termination efficiency [38]. *In vitro* studies also implicate RF2 and RF3 in post-peptidyl transfer quality control (post PT-QC), a mechanism for selectively terminating translation of polypeptides that have misincorporated amino acids, and phenotypes suggestive of post PT-QC were found *in vitro* and *in vivo* [34,40], but were not reproduced in K-12 strains with $RF2^B$ [35]. Taken together, the work thus far indicates how release factor mutations alter translation termination at specific reporter constructs or known frameshifting sites. However, the effects of these mutations on translation termination have not been studied on a global scale or at physiologically relevant native gene loci.

To elucidate the global effects of these mutations and observe how they perturb the translatome, we used ribosome profiling to examine the behavior of ribosomes at stop codons, and compare the extent of recoding and readthrough events genome-wide in native *E. coli* K-12, with that in K-12 $RF2^B$ cells and in both strains lacking RF3. We also examined changes in protein expression globally among the strains. We find significant differences in the duration of termination/release dependent upon strain background, identify new recoding events and reveal the impact of altered termination on genes whose expression is regulated by transcription-translation coupling.

Results

Growth rates at various temperatures

We examined the effect of altered release factors on the growth of *E. coli* MG1655, the prototypical K-12 strain used in our studies. Previous results indicated inconsistent phenotypes for MG1655 and BW25113, another K-12 strain [34,35,41]. Here, we determine the growth rates of MG1655 (K-12 $RF2^{K-12}$) and isogenic single and double mutant release factor derivatives (Table 1). At 37˚C in MOPS-complete glucose medium, supplemented with all amino acids, K-12 $RF2^{K-12}$ΔRF3 has a significantly slower growth rate than K-12 $RF2^{K-12}$ (doubling times of 36 min and 28 min, respectively), but the slow growth phenotype of ΔRF3 is rescued by $RF2^{B}$ (Fig. 1A). This suggests that the enhanced activity of $RF2^{B}$ compensates for the ΔRF3 defects that result in reduced growth rate. K-12 $RF2^{B}$ and K-12 $RF2^{K-12}$ have indistinguishable growth rates.

Our previous global phenotyping screen of the *E. coli* BW25113 single gene deletion library indicated that ΔRF3 was quite cold-sensitive at 16˚C and 20˚C [42]. ΔRF3 strains in other backgrounds, both K-12 and non-K-12 derived, also grow slowly at 25°C, a phenotype that is reversed by $RF2^{B}$ΔRF3 [35]. We therefore characterized the growth rate of our isogenic *E. coli* K-12 MG1655 strains across an expanded temperature range. Consistent with previous results [35], K-12 $RF2^{K-12}$ ΔRF3 exhibits severe cold sensitivity at 20˚C, which was reversed in the K-12 $RF2^{B}$ΔRF3 strain (Fig 1B). Surprisingly, however, at 15˚C, the $RF2^{B}$ allele did not rescue ΔRF3 (Fig 1B). These results motivated us to examine strain viability at each temperature (Fig. 1C). We find that although it is slow growing, K-12 $RF2^{K-12}$ ΔRF3 maintains viability at all temperatures tested. In contrast, K-12 $RF2^{B}$ ΔRF3 exhibited a near 5-fold decrease in viability at 15˚C. Thus, at very low temperature the $RF2^{B}$ variant is more deleterious than $RF2^{K-12}$ variant when paired with RF3 deletion (see discussion).

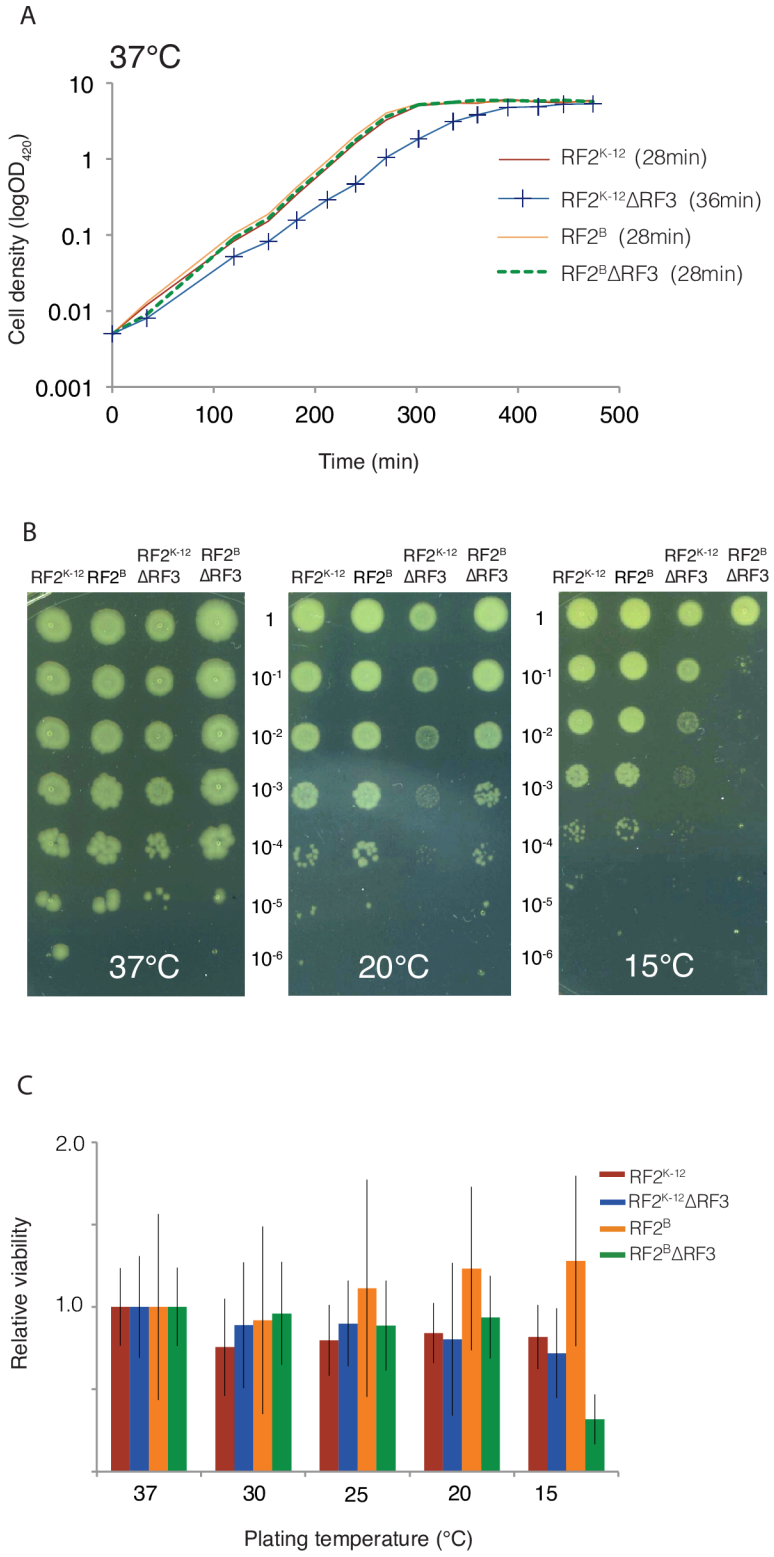
Table 1. *E. coli* strains and genotypes in this study.

Strain name	Genotype
K-12 $RF2^{K-12}$ (MG1655)	*F- λ- ilvG- rfb-50 rph-1*
K-12 $RF2^{K-12}$ ΔRF3	MG1655 Δ*prfC::frt*
K-12 $RF2^{B}$	MG1655 *prfB* [*E. coli* B]
K-12 $RF2^{B}$ ΔRF3	MG1655 *prfB* [*E. coli* B] Δ*prfC::frt*

Figure 1. Growth defect and cold sensitivity rescue of RF mutants.

(A) Growth curves of MG1655 K-12, K-12 RF2K-12ΔRF3, K-12 RF2B, K-12 RF2BΔRF3 growing in MOPS complete-glucose media at 37°C. Doubling times are indicated. (B) Spot dilutions of cultures growing exponentially at 37°C in LB were spotted onto LB plates and incubated at 37°C, 20°C, and 15°C. These indicate that RF2B rescues the slow growth of the K-12 RF2K-12ΔRF3 strain at 20°C. (C) After growth to mid-exponential phase in liquid LB cultures serial dilutions of each culture were plated on LB-agar in triplicate. Colony forming units (CFUs) of each strain were calculated for each temperature and plotted relative to the CFU of that strain at 37°C. Error bars are the standard error calculated between replicates (See materials and methods).

Figure 1


A
37°C
Cell density (logOD420)
10
1
0.1
0.01
0.001
0
100
200
300
400
500
Time (min)
RF2K-12 (28min)
RF2K-12ΔRF3 (36min)
RF2B (28min)
RF2BΔRF3 (28min)
B
RF2K-12
RF2B
RF2K-12 ΔRF3
RF2B ΔRF3
1
10-1
10-2
10-3
10-4
10-5
10-6
37°C
20°C
15°C
C
Relative viability
2.0
1.0
37
30
25
20
15
Plating temperature (°C)
RF2K-12
RF2K-12ΔRF3
RF2B
RF2BΔRF3

Global analysis of translation termination/release in strains with altered termination efficiency.

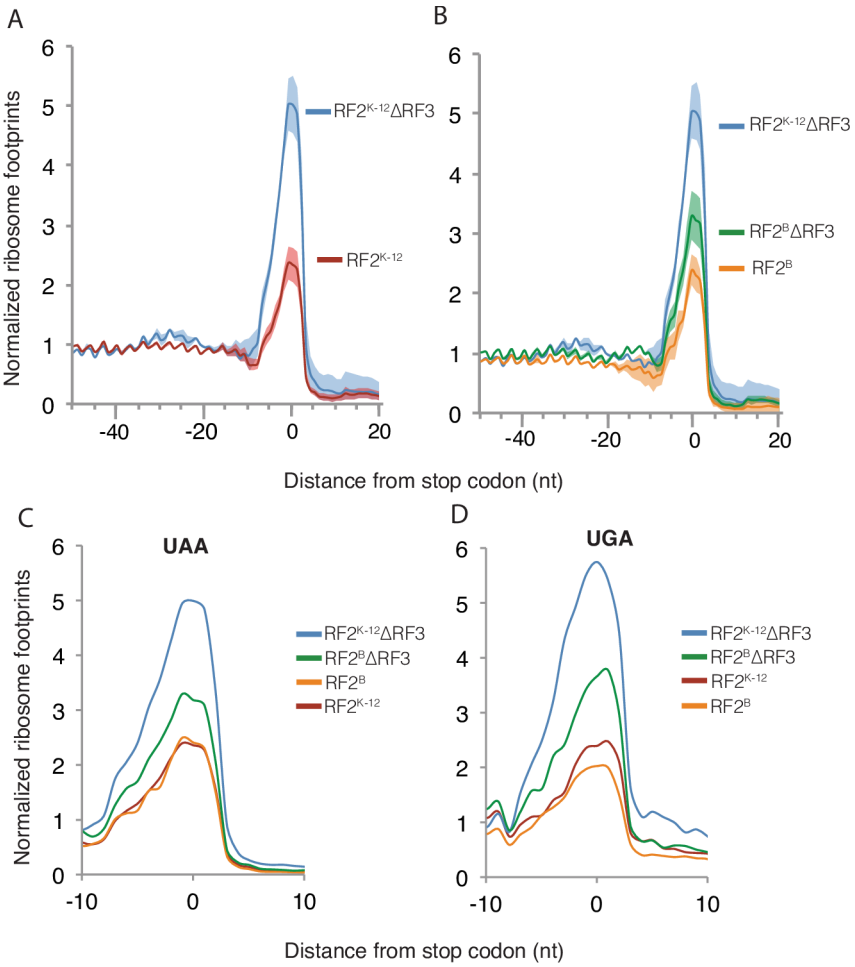
We performed ribosome profiling for K-12 $RF2^{K-12}$ and K-12 $RF2^{B}$ strains, with and without RF3. All ribosome profiling experiments were performed at 37˚C in MOPS- complete glucose medium because it contains a balanced complement of amino acids (see Methods). As a proxy for duration of ribosome termination and release, we quantified the extent of ribosome occupancy at stop codons. The density of footprints at any codon is related to the dwell time of the ribosome at that position [43]. Thus, higher stop codon occupancy is indicative of increased duration of termination/release. We note that this metric is a composite measurement that minimally consists of the rates of: release factor binding, polypeptide chain termination, release factor release, and ribosome recycling. Our initial analysis used a dataset of approximately 1200 well-expressed genes aligned at the stop codons of their open reading frames (ORFs), and is comprised of 947 UAA stop codons, recognized by both RF1 and RF2, and 231 UGA codons recognized only by RF2 (Fig 2, see Methods). For additional analyses, we expanded the dataset to 3390 genes so that we could also study the lower abundant UAG and UGA stop codons, as well as all of the 4-base stop codons (S1 Fig, see Methods).

We used our 1200 gene dataset to compare global stop-codon occupancy by native K-12 $RF2^{K-12}$ and K-12 $RF2^{K-12}$ ΔRF3 cells, as well as K-12 $RF2^{B}$ cells and K-12 $RF2^{B}$ ΔRF3 cells. We found that $RF2^{K-12}$ ΔRF3 cells have a global 2.5-fold increase in ribosome occupancy at stop codons, relative to native K-12 $RF2^{K-12}$ indicating a significant defect in termination/release (Fig 2A). K-12 $RF2^{B}$ ΔRF3 cells also have a global increase in ribosome occupancy at stop codons relative to K-12 $RF2^{B}$ (Fig 2B), but the magnitude of this increase is less (1.4-fold). It is likely that the more active $RF2^{B}$ variant limits the effect of ΔRF3. Changes in occupancy are also evident in ribosome footprint density plots of single genes (S2 Fig).

Figure 2. Ribosome occupancy over stop codons increases in the absence of RF3 in *E.coli* K-12.

Metagene analysis of ribosome footprint density in the region surrounding stop codons. The approximately 1200 well-expressed genes were aligned at their stop codons and the median normalized ribosome density at each position was calculated from ribosome profiling data of strains grown in MOPS complete-glucose media at 37°C. (A and B) Median normalized density (solid line) with standard error (shaded region) across repeat experiments. (A) Ribosome stop codon density for K-12 $RF2^{K-12}$ΔRF3 (2 replicates) versus K-12 $RF2^{K-12}$ (4 replicates). (B) Ribosome stop codon density for K-12 $RF2^{B}$ΔRF3 (5 replicates) versus K-12 $RF2^{B}$ (2 replicates). (C and D) Average normalized density across replicates at UAA stops (947 genes) and UGA stops (231 genes).

Figure 2


A
B
C
D
Normalized ribosome footprints
Distance from stop codon (nt)
RF2K-12ΔRF3
RF2K-12
RF2BΔRF3
RF2B
UAA
UGA

Because global stop codon occupancy could be driven by the behavior at the highly abundant UAA stop codons, we next compared the behavior at each of the three stop codons. We first compared K-12 $RF2^{K-12}$ ΔRF3 and K-12 $RF2^{B}$ ΔRF3 at all three stop codons using our larger 3390 gene dataset. Changes in occupancy of the ΔRF3 strains at all three stop codons, including the low abundance UAG stop codon recognized only by RF1, mirrors their change in global occupancy (Fig 2C, 2D and S1B Fig). Occupancies over UAG are very similar to those over UAA, although the low number of UAG stops increases the noise in the data (S1B Fig). The defects of translation termination/release seen at all stop codons for strains lacking RF3 is consistent with previous studies showing that RF3 is important for release of both RF1 and RF2 [22]. Notably, relative to K-12 $RF2^{K-12}$, the K-12 $RF2^{K-12}$ΔRF3 strain has a very slight increase in ribosome occupancy ~20-30 bp upstream of the stop, at the position expected for a 2^{nd} ribosome (Fig 2A). The height of this shadow peak increases in highly translated genes (S3 Fig), suggestive of ribosome pileup. However, further experiments are necessary to definitively establish this point. In summary, deleting RF3 results in a general defect in termination/release at all three stop codons, but the magnitude of this effect is much smaller in K-12 $RF2^{B}$ than in the native K-12 $RF2^{K-12}$ strain.

We next compared the stop codon occupancy of K-12 $RF2^{K-12}$ with that of K-12 $RF2^{B}$ both globally (Fig 2A, 2B) and at specific stop codons (Fig2C, 2D, S1B Fig). Both strains have similar stop codon occupancy globally, as well as at the UAA and UAG stop codons (Fig2A-C, S1B Fig). However, for UGA codons, recognized solely by RF2, there is a general trend towards decreased occupancy in K-12 $RF2^{B}$ versus K-12 $RF2^{K-12}$ strains, consistent with observations that suggesting that the $RF2^{B}$ protein is more efficient at mediating termination at UGA than $RF2^{K-12}$ variant (Fig 2C, 2D, S1B Fig) [15,16].

The base following the stop codon, called the fourth base, is known to impact the efficiency of stop codon termination and is recognized by both RF1 and RF2 [39,44,45]. We created a metagene plot for each 4-base stop codon to examine the occupancy preferences of

each of our strains. While these plots generally have more noise than the 3 base stop plots due a smaller dataset for each 4 base stop codon, particularly the lower abundance UAG and UGA stops, we can observe general trends of occupancy within each specific stop codon group (S1 Fig). Despite noise, behavior of strains at the 4-base codons generally mirrored their behavior at the 3-base codons, with some additional patterns. The K-12 $RF2^{B}$ strain has the most severe decrease in stop codon occupancy relative to K-12 $RF2^{K\text{-}12}$ at nearly all 4-base UGA stop codons (S1 Fig). Strikingly, at the UGAA stop, K-12 $RF2^{B}\Delta RF3$ has ribosome occupancy equivalent to that of native K-12 $RF2^{K\text{-}12}$, rather than the slightly elevated level characteristic at other stops. Together, these results suggest that $RF2^{B}$ variant is significantly more effective than the $RF2^{K\text{-}12}$ protein at UGAA stops (see discussion).

Release factor expression

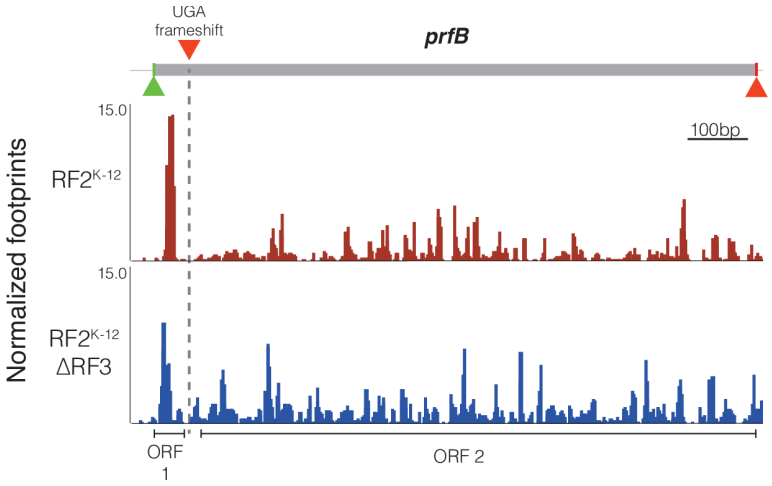
We also used the ribosome profiling data to examine release factor expression in each strain. Previous experiments using reporter constructs indicated that the rate of frameshifting at the internal *prfB* UGA stop codon, used to change expression of RF2, site increased in a ΔRF3 strain [34,39]. We show that this result also is true at the endogenous *prfB* locus by calculating the ratio of ribosome occupancy of the larger 2nd frameshifted ORF (encoding full length RF2) relative to that of the small 1st ORF. Native K-12 $RF2^{K\text{-}12}$ cells exhibit ~28% frameshifting at this locus and this value increased to ~47% in the K-12 $RF2^{K\text{-}12}$ ΔRF3 strain (Fig 3). K-12 $RF2^{B}$ cells appear to have slightly less frameshifting (21%) than K-12 $RF2^{K\text{-}12}$ cells. K-12 $RF2^{B}\Delta RF3$ cells do not exhibit the increased frameshifting as seen in $RF2^{K\text{-}12}\Delta RF3$ cells, but instead exhibit a similar amount of frameshifting as K-12 $RF2^{K\text{-}12}$.

Concomitantly, we measured the protein levels of RF2, finding general agreement with our calculated rate of frameshifting. K-12 $RF2^{K\text{-}12}$ cells have a higher level of RF2 than K-12 $RF2^{B}$ (S1 Table). Our data are consistent with a previously reported ~40% reduction of RF2 in K-12 $RF2^{B}$ calculated by quantitative western blot [14].

Figure 3: ΔRF3 increase rate of known programmed frameshift in *prfB*.

The translation of prfB is regulated through a programmed +1 frameshift at codon 26, shown as a UGA frameshift (dotted line). Successful frameshifting results in the complete RF2 protein. Ribosome profiling data was used to estimate frameshifting over the UGA at codon 26 by comparing the ribosome occupancy (RPKM) across the first small ORF prior to the UGA stop and second ORF encoding full-length RF2. This ratio gives a rough estimate of percent frameshift at the *prfB* locus for each RF mutant.

Figure 3



% frameshift based on ribosome footprints (ORF2/ORF1)

$RF2^{K-12}$	27.7%
$RF2^{K-12}\Delta RF3$	46.7%
$RF2^{B}$	21.8%
$RF2^{B}\Delta RF3$	28.2%

Additionally, as expected from increased frameshifting, K-12 $RF2^{K-12}$ ΔRF3 cells have increased expression of RF2 relative to K-12 $RF2^{K-12}$ cells (S1 Table). K-12 $RF2^{K-12}$ ΔRF3 cells also have increased expression of RF1 relative to K-12 $RF2^{K-12}$ cells The fact that K-12 $RF2^{K-12}$ ΔRF3 cells still exhibit severe termination/release defects suggests that increased release factor expression only partially mitigates this phenotype. Neither RF2 nor RF1 expression increases in K-12 $RF2^{B}$ΔRF3 cells relative to K-12 $RF2^{B}$, presumably because the higher activity of the $RF2^{B}$ relative to $RF2^{K12}$ prevents the signals triggering enhanced release factor expression.

ΔRF3 strains exhibit higher ribosome density in the post-ORF region

Extended stop-codon occupancy could have the downstream consequence of facilitating either stop codon readthrough or of frameshifting, collectively called here recoding events. Our finding of an increase in the programmed frameshifting at the *prfB* locus in certain strain backgrounds is consistent with that idea. Recoding events would result in an increase in ribosome density after the annotated stop codon. We therefore quantified the average ribosome density downstream of the open reading frame (ORF) relative to the average density within the ORF itself. We call this metric relative post-ORF ribosome occupancy (RPOR. We analyzed RPOR for the almost 1600 genes are well separated from their immediate downstream gene, i.e. having ≥ 65 nucleotides between the stop codon of the upstream gene and the start codon of the downstream gene (Fig 4A). The distance constraint is necessary to enable us to unambiguously examine ribosomes past the stop codon of the upstream gene without interference from translation of the downstream gene.

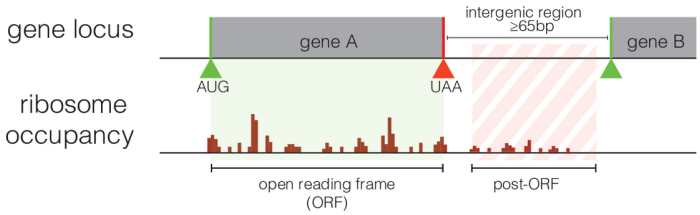
A plot of the cumulative distribution of RPOR values indicates that most genes have very low RPOR values across all strains. Indeed, nearly 60% of genes had an RPOR value under 0.1, indicative of very few ribosomes in the post-ORF region (Fig 4B). Thus, as expected, translation termination has high fidelity and is generally efficient [1,7].

Figure 4: Genome wide increase in post-ORF occupancy in K-12 ΔRF3.

(A) Ribosome occupancy in the region past the annotated stop codon was calculated for all genes that had an intergenic region of 65 bases or greater, measured from the stop codon of the upstream gene of interest to the start codon of the downstream gene, and met expression thresholds for both mRNA abundance and ribosome footprints (1656 genes for the deepest sequenced library). The metric relative post-ORF ribosome occupancy (RPOR) was calculated from the average occupancy over the annotated ORF and the average occupancy in a post-ORF window 20-60bp after the stop codon. By using the average occupancy for both the post-ORF and ORF, we reduce the impact of length in these calculations. (B) The distribution of RPOR values from 0 to 2.0 for K-12 $RF2^{K-12}$, K-12 $RF2^{K-12}$ΔRF3, and K-12 $RF2^{B}$ΔRF3 is shown as a cumulative distribution function for one experiment with 1139 genes after all zero RPOR values were removed. A shift towards higher RPOR values in the K-12 $RF2^{K-12}$ΔRF3 strain relative to the K-12 $RF2^{K-12}$ strain is statistically significant (p-value 0.04; Kolmogorov-Smirnov (K-S) test), but the small shift between the K-12 $RF2^{B}$ΔRF3 and K-12 $RF2^{K-12}$ strains is not (p-value 0.84, K-S test). (C) A histogram of the distribution of RPOR values ≥0.2 for K-12 $RF2^{K-12}$ and K-12 $RF2^{K-12}$ΔRF3 compares the number of genes in each category for each strain.

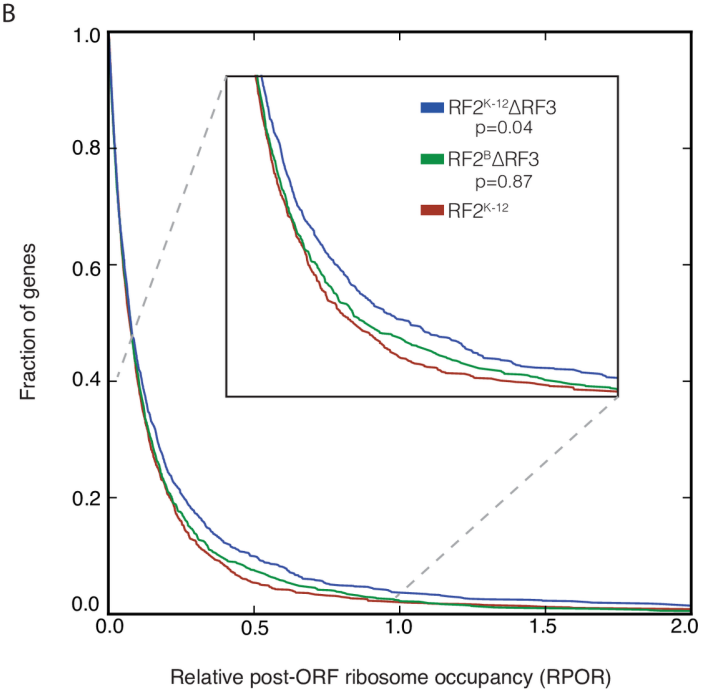
Figure 4

A

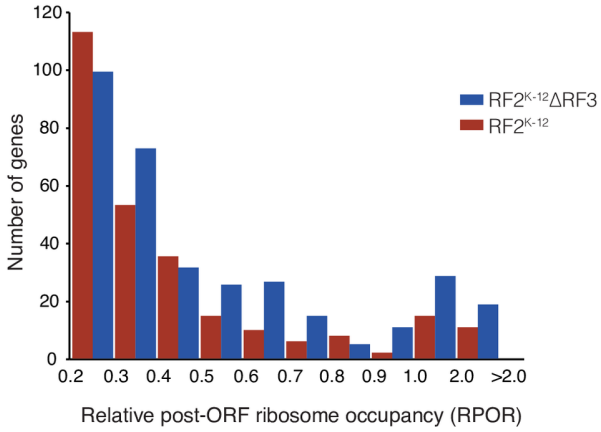


relative post-ORF ribosome occupancy (RPOR) = Average occupancy of post-ORF / Average occupancy of ORF

B



C



However, the K-12 RF2$^{K\text{-}12}$ΔRF3 strain trends towards higher RPORs than K-12 RF2$^{K\text{-}12}$ across the entire range of RPOR values (Fig 4B and S4A Fig), indicating a potential for globally reduced termination efficiency across the translatome. This shift is quite pronounced for those genes with the highest RPOR values: K-12 RF2$^{K\text{-}12}$ΔRF3 strains has nearly twice as many genes with RPOR > 1.0 (Fig 4C). Although the RPOR values of K-12 RF2^{B}ΔRF3 are not nearly as elevated relative to those of K-12$^{K\text{-}12}$ΔRF3 (Fig 4B), replicate experiments suggest that they are elevated relative to the native K-12 RF2$^{K\text{-}12}$ strain (S4B Fig). Finally, the RPOR values of K-12 RF2^{B} are very similar to those of K-12 RF2$^{K\text{-}12}$. Our most complete K-12 RF2^{B} dataset showed a very slight decrease in RPOR values relative to K-12 RF2$^{K\text{-}12}$ cells, but a smaller dataset did not exhibit this trend (S5 Fig).

Ribosome profiling in other organisms, including *Saccharomyces cerevisae* and *Drosophila melanogaster*, is capable of producing protected ribosome fragments with reading frame information more precisely allowing identification of recoding events [46,47]. However, bacterial ribosome footprints are generated with MNase, and reading frame information is lost because of the sequence specificities of this nuclease [48]. These specificities do not allow for perfect cutting of mRNA around the ribosome, resulting in variable ribosome footprint sizes, depending on sequence context, without reading frame information [49]. Therefore, we used a secondary criterion to identify those ORFs whose translation is likely extended into the post-ORF region as a result of either readthrough or frameshifting. If the ribosomes found in the post-ORF region are actually translating, then translation should terminate after ribosomes encounter a post-ORF stop codon in the translating reading frame. Therefore, we searched the data for ORFs exhibiting a reduction in ribosome density after a stop codon present in any frame of the post-ORF region relative to its density prior to that stop codon (Fig 5).

We hand-annotated the 100 genes with the highest relative post-ORF occupancy genes in both native K-12 RF2$^{K\text{-}12}$ and in K-12 RF2$^{K\text{-}12}$RF3 strains (121 total; S2-5 Tables) to identify cases where there was a decrease in ribosome density coincident with a stop codon located in

any of the three possible post-ORF reading frames. Using this criterion we classified 43 ORFs as likely recoding events (S2 Table). An additional 41 ORFs were classified as possible recoding events based on appropriate post-ORF termination with the addition of confounding factors such as low reads or other sequence elements that could contribute to high RPOR values (S3 Table). The vast majority of genes annotated as likely recoding events had a larger RPOR value in K-12 RF2$^{K\text{-}12}$ΔRF3 strain than in native K-12 RF2$^{K\text{-}12}$ cells (S2 Table). Interestingly, UGA codons were significantly over-represented (p-value = 1x10^{-4}) in both likely and possible recoding events, occurring at nearly double their expected frequency, while UAA codons were de-enriched (S6 Table). We classified the 37 ORFs without reduced ribosome density after post-ORF stop codons as non-recoding events (S4 Table). We additionally found 4 cases of exceptionally high RPOR, which appear to stem from mis-annotations (*ydcM, yeaP, wbbK, yebW*) (S5 Table) and 12 ORFs that had too few reads to classify. Taken together, our data strongly suggest that native K-12 RF2$^{K\text{-}12}$ has some recoding events and that deletion of RF3 in the K-12 background enhances recoding, likely due to poor termination/release efficiency.

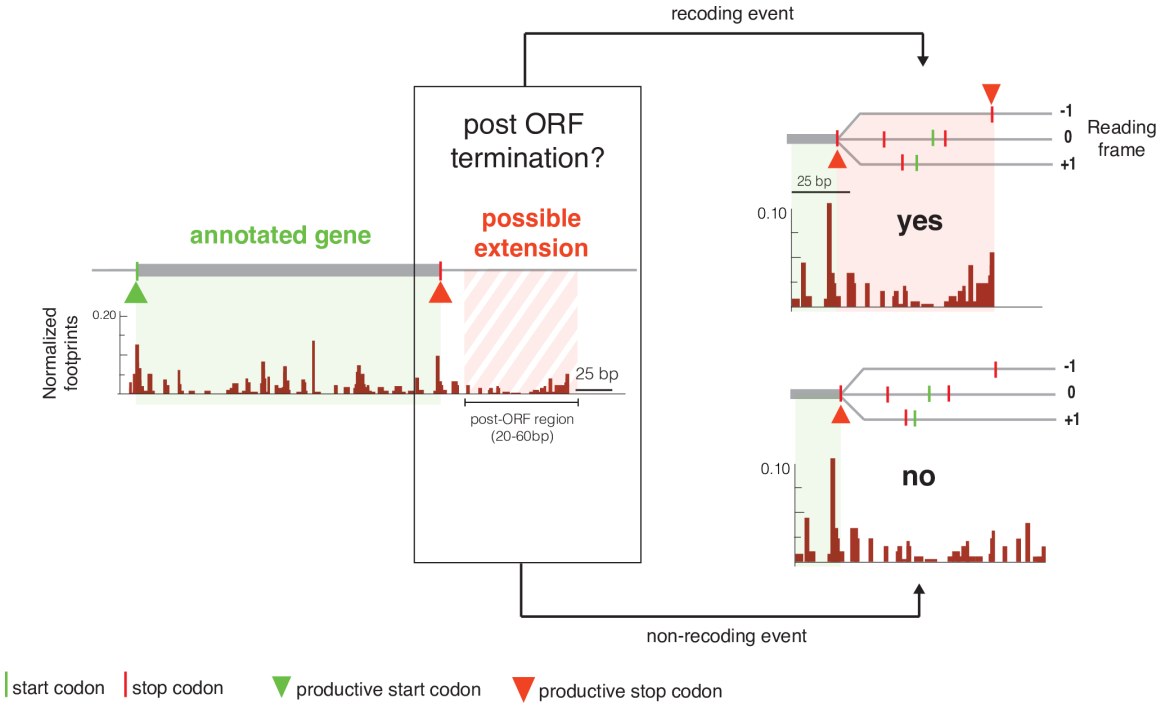
Genes with high post-ORF ribosome occupancy show extended protein products

We tested whether our screening criterion of reduced ribosome density after a stop codon in the post-ORF region identifies genes with a recoding event for three individual genes: *nudL* and *panZ* identified in this work, and *pheL*, a known frameshifter (S2 Table) [50]. We individually expressed the three genes in a multi-copy plasmid in all strain backgrounds; they all contained an N-terminus FLAG-tag, and the two newly identified putative recoding events were additionally tagged with streptavidin on the C-terminus of the suspected extension frames. We tested for protein extension products with quantitative Western blotting. This revealed that all three genes exhibited extended proteins in native K-12 RF2$^{K\text{-}12}$, and that two have increased extended product in K-12 RF2$^{K\text{-}12}$ΔRF3 (Fig 6B, S6 Fig).

Figure 5: Schematic explaining ORF classification as recoding or non-recoding.

Genes with high relative post-ORF ribosome occupancy (RPOR) were individually analyzed to determine the origin of the ribosomes in the post-ORF region. Left hand panel: Schematic of ribosome footprints for an annotated gene (shaded light green) and its post-ORF region (shaded red stripes). The ORF start (green triangle) and stop (red triangle) codons are indicated. Right hand panel: a zoomed-in post-ORF region with all possible start (green line) and stop (red line) codons in any reading frame indicated. Upper right hand panel: A post-ORF region where ribosome density abruptly decreases after the stop codon in the -1 frame, classified as a putative recoding event. The hypothesized extended region is shaded in red and the putative stop codon is marked with a red triangle. Lower right hand panel: A post-ORF region where ribosome density does not decrease after any possible the stop codon, classified as a non-recoding event.

Figure 5


recoding event
post ORF termination?
annotated gene
possible extension
Normalized footprints
0.20
25 bp
post-ORF region (20-60bp)
-1
0
+1
Reading frame
25 bp
0.10
yes
0.10
no
non-recoding event
start codon
stop codon
productive start codon
productive stop codon

The *nudL* post-ORF region ribosome density decreased significantly after two closely spaced stop codons, suggesting either readthrough (0 frame) or a -1 frameshift (Fig 6A). By constructing C-terminal streptavidin tags in both frames, we were able to show that the protein extension product resulted from stop codon readthrough in the 0 frame (S7 Fig). The extension product is approximately 6-fold more prevalent in the K-12 $RF2^{K-12}$ΔRF3 strain than in K-12 $RF2^{K-12}$ (Fig 6B) and is undetectable in strains with $RF2^{B}$ even when they lack RF3 (S8A Fig).

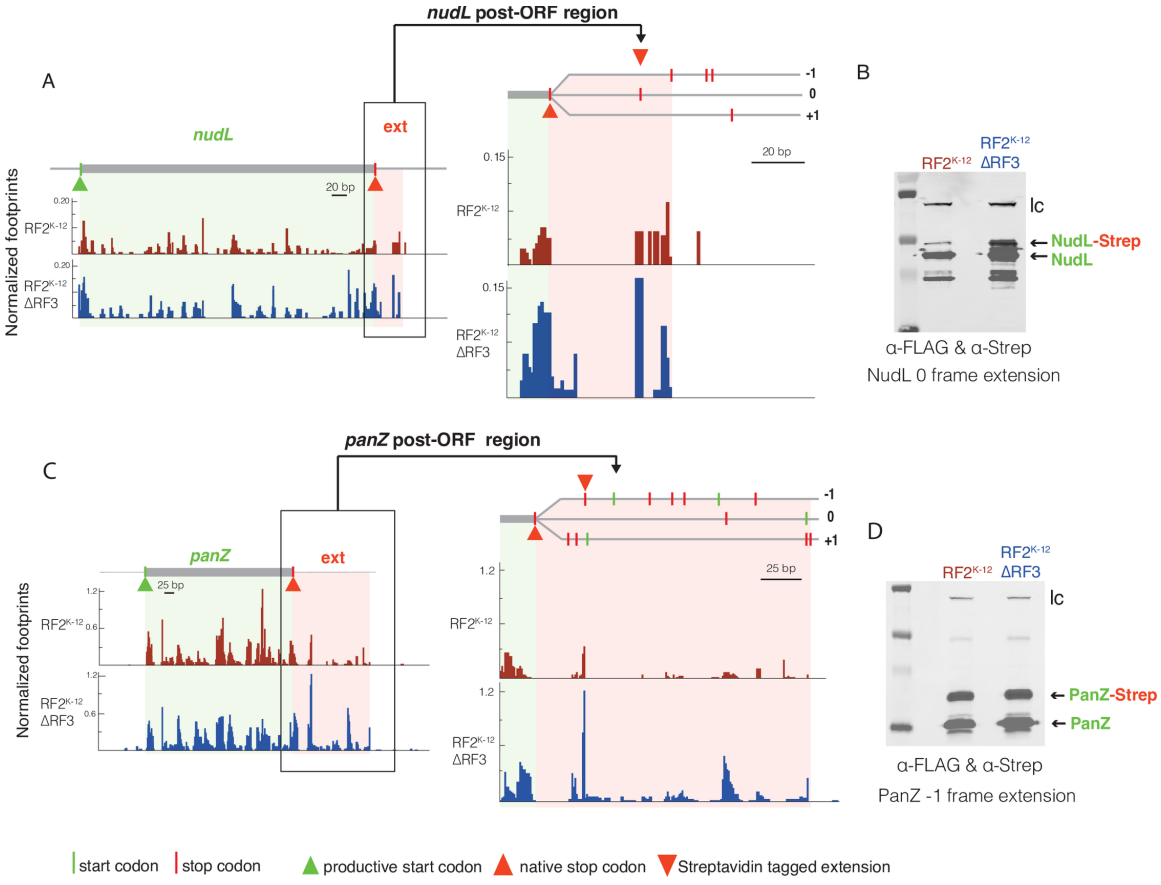
Ribosome density in the post-ORF region of *panZ* is complex with many possible points of termination (Fig 6C). We tested whether there was extension in the -1 frame, as predicted by the major decrease in ribosome density after the 1st stop codon in the post-ORF region. We observe a large amount of the -1 frameshift product in all strains, accounting for ~35% of the total PanZ production (Fig 6D, S8B Fig). The high level of PanZ extension product may result from unknown *cis* element(s) that leads to frameshifting.

The three well-established frameshifting events in native K-12 *E. coli* are RF2 (encoded by *prfB*), PheL, and DnaX [51]. We discussed the frameshift used to control the amount of full-length RF2 above (Fig 3). Here we describe our studies of PheL, the other established locus where a frameshift event produces an extended product. Our ribosome profiling data suggested decreased ribosome density after a stop in the 0 frame (S6A Fig), in addition to the previously identified +1 frameshift, which ends in the downstream *pheA* gene [52]. Using N-terminal FLAG-tagged PheL, we see both products in K-12 $RF2^{K-12}$ cells, with enhanced production in K-12 $RF2^{K-12}$ ΔRF3 cells (S6B Fig). However, we find no extended FLAG-PheL products in K-12 $RF2^{B}$ or K-12 $RF2^{B}$ΔRF3 cells (S6B Fig). Interestingly, native K-12 $RF2^{K-12}$ cells have an unusually large ribosome density over the *pheL* UGAA four-base stop codon (S6A Fig). This enhanced density is almost completely eliminated in both K-12 $RF2^{B}$ and K-12 $RF2^{B}$ΔRF3 cells (S6A Fig), likely because UGAA signals rapid termination in cells with $RF2^{B}$ protein variant.

Figure 6: ΔRF3 increases recoding in genes identified with high ribosome occupancy post-ORF.

(A and C) Left panels: Normalized ribosome footprints are shown across nudL (A) or panZ (C). The ORFs (shaded green) and post-ORF region, with hypothesized extensions (shaded red) are indicated. Right panels: zoomed-in post-ORF region with all possible start (green line) and stop (red line) codons in any reading frame indicated. The stop codon of the C-terminally tagged extension is indicated with a red triangle within the shaded red extension. (B and D) Western blots of strains containing N-terminally FLAG-tagged and C-terminally streptavadin tagged NudL (B) or PanZ (D) and use SurA as a loading control (lc). The NudL C-terminal streptavidin tag is in the 0 frame, and that of PanZ is in the -1 frame. (A) The post-ORF region of *nudL* has reduction in ribosome occupancy correlated with stop codons in both the 0 and -1 frame. We determined the 0 frame produces an extended NudL product (S7 Fig). (B) Blotting of α-FLAG indicated the NudL protein, 22.78kDa, and the readthrough extension product at 24.06kDa which was also tagged with streptavidin. (C) The post-ORF region of panZ reveals several possibly productive stop codons. (D) Blotting of α-FLAG indicated the PanZ protein, 15.84kDa, and the -1 frame extension product at 20kDa, which was tagged with both FLAG and streptavidin.

Figure 6


nudL post-ORF region
A
nudL
ext
20 bp
Normalized footprints
RF2K-12
RF2K-12 ΔRF3
0.20
0.15
-1
0
+1
B
RF2K-12
RF2K-12 ΔRF3
Ic
NudL-Strep
NudL
α-FLAG & α-Strep
NudL 0 frame extension
panZ post-ORF region
C
panZ
ext
25 bp
1.2
0.6
D
PanZ-Strep
PanZ
α-FLAG & α-Strep
PanZ -1 frame extension
start codon
stop codon
productive start codon
native stop codon
Streptavidin tagged extension

Coupled genes in polycistronic operons

Genes in *E. coli* are densely packed in the chromosome; with approximately 15% of adjacent ORF pairs having overlapping stop and start codons [53]. In some cases, this overlap has been shown to promote translational coupling, possibly by enabling upstream ribosomes to influence downstream ORF translation by unwinding mRNA structure or affecting ribosome dissociation-reinitiation cycle [54–58]. We asked whether the relative translation level of downstream genes (normalized to that of upstream genes) increased in K-12 $RF2^{K-12}$ΔRF3 strains as compared to native K-12 $RF2^{K-12}$ for adjacent ORF pairs with overlapping stop and start codons. We find that when the upstream gene has higher translation than the downstream gene in the native K-12 $RF2^{K-12}$ cells deleting RF3 can increase downstream gene translation (S9A Fig). The aberrantly high translation was rescued in K-12 $RF2^{B}$ ΔRF3 strains, and K-12 $RF2^{B}$ strain is similar to native K-12 $RF2^{K-12}$ strain (S9B and S9C Fig). Accumulation of ribosomes at the upstream stop codon may promote unwinding of the mRNA structure at the translation initiation region of the downstream gene to increase its translation level, or alternatively, may be an example of readthrough.

ΔRF3 increases attenuation of biosynthetic genes under control of leader peptide

Expression of many *E. coli* biosynthetic genes is controlled by regulated transcription termination, or attenuation. Attenuation is mediated by two competing RNA stem-loop structures, one signaling transcription termination and the other allowing transcription to continue [59]. Ribosome occupancy in the leader peptide mRNA determines the ratio of the two stem-loop structures. Enhanced ribosome occupancy in the leader peptide, indicative of a deficiency of the amino acid produced by the biosynthetic operon, leads to enhanced transcriptional readthrough, thus ensuring adequate production of the limiting amino acid. We asked whether altered ribosome occupancy at leader peptide stop codons in the K-12 $RF2^{K-12}$ΔRF3 strain altered attenuation. A comparison of ribosome footprint density of the native K-12

RF2^{K-12} with that of K-12 RF2^{K-12}ΔRF3 indicated that biosynthetic genes under the control of a leader peptide appeared to be down-regulated in the K-12 RF2^{K-12}ΔRF3 strain relative to native K-12 RF2^{K-12} in rich media (Fig 7A) and less so, if at all, in minimal media (Fig 7B).

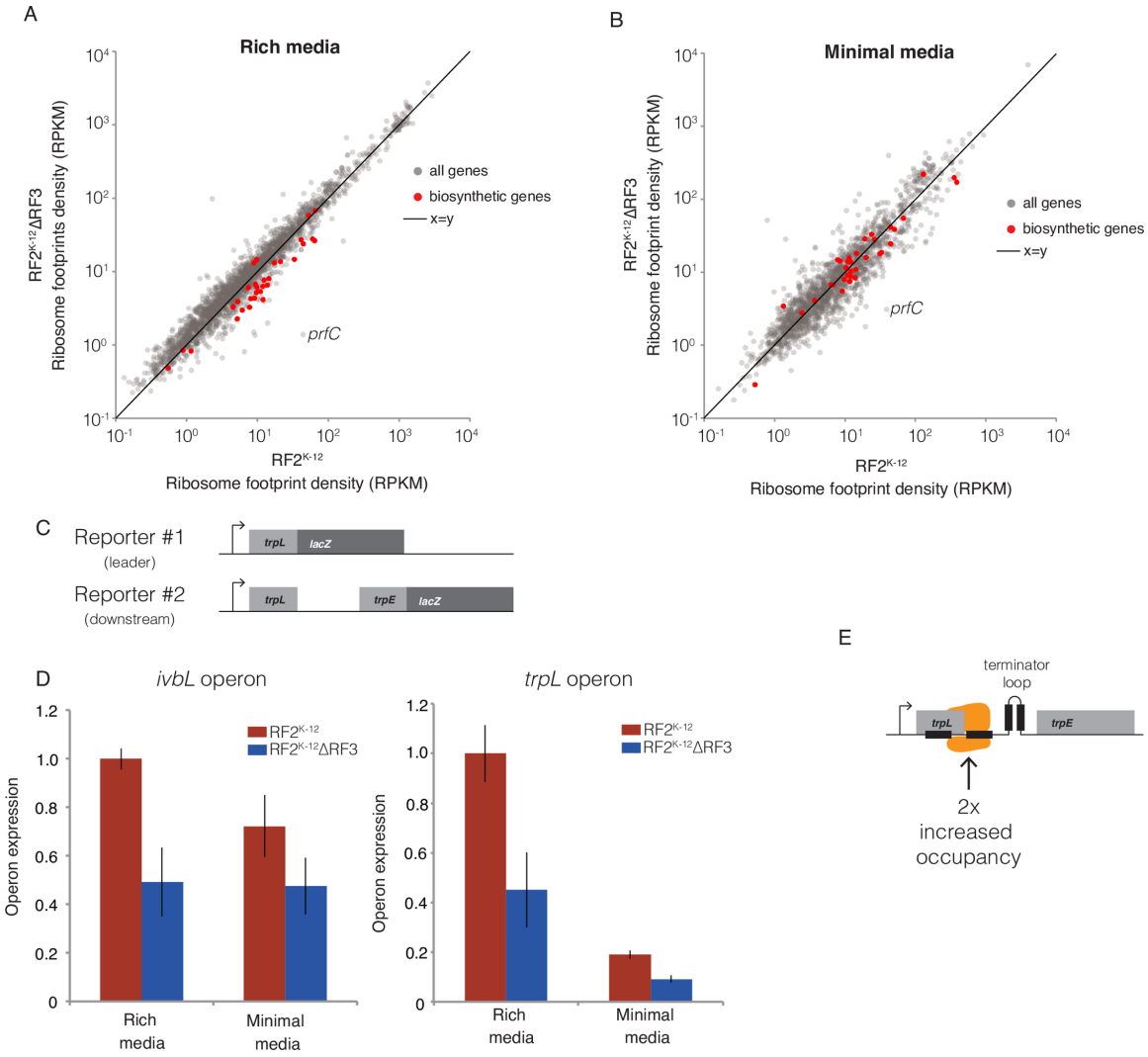
We quantified the attenuation of two leader peptide operons, *ivbL* and *trpL*, in both MOPS-complete glucose, a rich medium, and MOPS-glucose, a minimal medium using two plasmid-based *lacZ* reporters. Reporter #1 measured expression of the leader peptide to quantify the rate of transcription of the operon; and reporter #2 measured expression of downstream genes (Fig 7C). Relative to native K-12 RF2^{K-12} cells, K-12 RF2^{K-12} ΔRF3 cells have reduced expression of *trp* genes in both rich and minimal medium, while *ivbL* has reduced expression in rich medium (Fig 7D). As reporter #1 allows us to normalize for rate of transcription, we conclude that this decrease in expression is the result of increased attenuation due to increased ribosome occupancy exactly over the stop codon in K-12 RF2^{K-12} ΔRF3 cells (Fig 7E). We also performed these experiments for the *hisGDCBHAFI* operon but were unable to measure transcription, as reporter #1 was toxic. Results with reporter #2 indicated that expression of *his* was decreased in K-12 RF2^{K-12} ΔRF3 cells in rich but not minimal medium (S10 Fig). We believe that this also represents attenuation because expression of the *ivlBN* and *trpEDCBA* leader peptides were comparable in K-12 RF2^{K-12} and RF2^{K-12}ΔRF3 strains, suggesting that the lack of RF3 may not alter transcription from the leader peptide promoter.

The results of expression data from our ribosome profiling experiments, coupled with confirmatory *lacZ* fusion experiments indicate that altered ribosome density over the stop codon alters the outcome of the RNA structure competition, such that cells with enhanced stop codon occupancy have increased transcription termination. Most likely, increased occupancy over the stop codon shifts the equilibrium between the readthrough and termination stem-loop structures to one favoring the termination stem-loop. These findings are consistent with and extend an early study showing that some temperature sensitive RF2 strains increased transcription termination in the *trp* operon [60].

Figure 7. ΔRF3 reduces expression of biosynthetic genes controlled by leader peptide attenuation.

(A and B) Scatter plots of ribosome footprint density in RPKM for K-12 $RF2^{K-12}$ and K-12 $RF2^{K-12}\Delta RF3$ strains for all genes above minimum read threshold, with biosynthetic genes under the control of leader peptide attenuation in red are all others in grey. (A) Rich media samples of K-12 $RF2^{K-12}$ and K-12 $RF2^{K-12}\Delta RF3$ grown in MOPS complete-glucose media are averaged data across multiple replicates. (B) Minimal media ribosome footprint densities between K-12 $RF2^{K-12}$ and K-12 $RF2^{K-12}\Delta RF3$ in MOPS minimal-glucose media are plotted for a single replicate. (C) Schematic of the chromosomally integrated reporter constructs measuring expression of the leader peptide (reporter #1) or downstream gene expression (reporter #2), both under control of their native promoter. (D) Operon expression in K-12 $RF2^{K-12}$ and K-12 $RF2^{K-12}\Delta RF3$ strains was calculated by dividing the β-galactosidase activity of reporter #2 by activity of reporter #1 and normalizing to the K-12 $RF2^{K-12}$ strain in in MOPS complete-glucose rich media. (E) Increased attenuation in K-12 $RF2^{K-12}\Delta RF3$ strain may result from increased occupancy over the stop codon and post-ORF region. This mechanism would stabilize the formation of the downstream terminator loop over the anti-terminator loop.

Figure 7


A
Rich media
Ribosome footprints density (RPKM)
RF2K-12ΔRF3
RF2K-12
Ribosome footprint density (RPKM)
all genes
biosynthetic genes
x=y
prfC
B
Minimal media
Ribosome footprint density (RPKM)
RF2K-12ΔRF3
RF2K-12
Ribosome footprint density (RPKM)
all genes
biosynthetic genes
x=y
prfC
C
Reporter #1
(leader)
trpL
lacZ
Reporter #2
(downstream)
trpL
trpE
lacZ
D
ivbL operon
Operon expression
RF2K-12
RF2K-12ΔRF3
Rich media
Minimal media
trpL operon
Operon expression
RF2K-12
RF2K-12ΔRF3
Rich media
Minimal media
E
terminator loop
trpL
trpE
2x increased occupancy

In that study, enhanced termination was observed only at UGA stops, but in our study, possibly because of a stronger perturbation, we observe increased termination also at the UAG stop in the *ivbL* locus. These results indicate how the defects of release factor mutations can alter downstream gene expression depending on transcription-translation coupling.

Discussion

It is critically important for organisms to terminate the translation of proteins accurately and in an appropriate time frame. Release factors are central to this process. In this work, we compared the accuracy of translation termination and its downstream consequences at a global scale in native *E. coli* K-12 cells, which have a reduced function RF2 ($RF2^{K-12}$) with those in K-12 cells harboring a fully functional RF2 ($RF2^{B}$), and examined the consequences of the deletion of RF3 in both strain backgrounds. Our results provide the first picture of the genome wide consequences of stressing translation on ribosome behavior and release factor synthesis in a panel of isogenic strains with increasingly deficient release factor activity.

We used ribosome occupancy over stop codons as our *in vivo* composite metric for translation termination and ribosome release. In native K-12 $RF2^{K-12}$ cells, the average occupancy of ribosomes at stop codons is about 2-fold higher than that in the coding regions, reflecting the time required from stop codon recognition to ribosome release (Fig 2A). The faster rate of peptide hydrolysis by the $RF2^{B}$ variant relative to $RF2^{K-12}$ [16] is manifest as a reduction in ribosome stop codon occupancy in K-12 $RF2^{B}$ at UGA stop codons, which are fully dependent on RF2 for polypeptide release (Fig 2D). This effect is particularly strong at the UGAA codon which has relatively weak binding to RF2 (S1F Fig) [39,44,61]. Weak association of RF2 with a stop codon may allow immediate dissociation of RF2 following rapid hydrolysis, which is especially rapid in strains containing the more active $RF2^{B}$ variant, thus further decreasing the time to ribosome dissociation. On the other hand, both K-12 $RF2^{B}$ and K12 $RF2^{K-12}$ strains have similar stop codon occupancies at UAA codons. UAA is decoded by both RF1 and RF2, with previous studies indicating that UAA is predominantly recognized by RF1 in strains with $RF2^{K-12}$, but that RF2 plays a major role in recognizing UAA codons in strains with $RF2^{B}$ [14,15,30]. Thus, the higher activity of $RF2^{B}$ as compared to $RF2^{K-12}$ is manifest only at the UGA codon, recognized solely by RF2.

Deletion of RF3 globally perturbs stop codon occupancy by ribosomes, both for K-12 $RF2^{B}$ and K-12 $RF2^{K\text{-}12}$ strains. K-12 $RF2^{B}$ ΔRF3 cells exhibit about a 1.4-fold increase in occupancy at all stop codons relative to K-12 $RF2^{B}$ (Fig 2B-D), indicating that the absence of RF3-mediated release factor dissociation visibly increases the dwell time of ribosomes at stop codons. Because of the composite nature of our measurement, we cannot say whether the effect we see is commensurate with the expectation from *in vitro* studies that loss of RF3 decreases the rate of release factor dissociation by as much as 500-fold or whether our results indicate that EF-G and RRF may partially compensate for RF3, as has been suggested [4,62,63]. For K-12 $RF2^{K\text{-}12}$ ΔRF3 cells, the global perturbation of stop codon occupancy by ribosomes is more severe, exhibiting an almost 2.5-fold increase in ribosome occupancy at all stop codons relative to native K-12 cells (Fig 2A), a 1.5-fold increase in expression of both RF1 and RF2 (S1 Table), and ≥20% increase in doubling time (Fig 1A). As the level of RF2 is adjusted by an internal frameshift, the rate of frameshifting changes in concert with the change in protein levels (Fig 3).

RF1 and RF2 are sub-stoichiometric with respect to ribosomes: RF1 is 100-fold less abundant, and RF2 is 12-fold less abundant than the number of ribosomes in K-12 $RF2^{K\text{-}12}$cells [37,43,64]. This level of release factors is normally sufficient for efficient termination because release factors act catalytically on the small subpopulation of terminating ribosomes. However, in the absence of RF3, the much slower dissociation of RFs increases the dwell time of RFs at stop codons, which in turn leads to a decrease in the cellular concentration of free RFs. This effect is exacerbated when it is coupled with the defective peptide release $RF2^{K\text{-}12}$ protein, leading to even longer dwell time of release factors and further depletion of pool of free RFs. The cell responds by increasing expression of RF1 and RF2 in the K-12 $RF2^{K\text{-}12}$ΔRF3 strain, but this is insufficient to counteract sequestration of RFs as termination is globally slowed at all stop codons, including UAG, which is recognized solely by RF1 (Fig 2, S1 Fig).

Our studies stressed translation in K-12 $RF2^{K-12}$ by removing RF3. Previous studies indicated that perturbing translation in K-12 $RF2^{K-12}$ by inactivating the pseudouridine synthase, *rluD,* or the RF methyltransferase, *prmC,* were ameliorated by the $RF2^{B}$ allele [14,65]. In K-12 *E. coli* deletion of these modification enzymes leads to a very severe growth defect, which is lethal in the case of K-12 Δ*prmC*. Therefore the relatively gentler perturbation by ΔRF3 in K-12 allows us to identify these changes in ribosomal occupancy that would be difficult or impossible to identify in K-12 *rluD* or *prmC* deletions. Deletions of *prmC* or *rluD* are viable in backgrounds containing $RF2^{B}$ and they would likely exhibit the similar molecular signature as K-12 $RF2^{K-12}$ΔRF3, enhanced expression of RF1 and 2 coupled with a significant enhancement in ribosome occupancy at stop codons.

Two lines of evidence suggest that the enhanced ribosome occupancy seen at stop codons in cells with reduced RF2 function (e.g. $RF2^{K-12}$) and lacking RF3 leads to translation readthrough and/or frameshifting at a significant number of genes. First, for 121 ORFs with the highest relative post-ORF ribosome occupancy, ≥30% of the genes have a decrease in ribosomal density after a stop codon in the post-ORF region (S2 Table). This signature is indicative of bona fide translation termination at those stop codons. Second, we rigorously confirmed recoding events in three cases. For the 2 new cases, we visualized the extension products by performing Western blotting on proteins that had distinct tags at both their N-terminus and at the C-terminus of the putative protein extension (Fig 6 and S6 Fig). These experiments reveal that 2 of the 3, *nudL* and *pheL* show readthrough of the UGA stop codon, notable because in bacteria, readthrough is rare relative to frameshifting [51]. We favor the idea that near-cognate tRNAs for cysteine and tryptophan known to decode UGA stop codons [66,67] account for readthrough, as the sequence element required for using selenocysteine to decode UGA [68–70] is not present at this locus. Interestingly, PheL also exhibits the previously documented +1 frameshift [52], indicating that during their extended occupancy over the stop codon, both readthrough and frameshifting events are possible.

Programmed recoding can expand the repertoire of gene products and is utilized for gene regulation in bacteria, viruses, yeast and higher eukaryotes [51,71]. While most of the recoding previously found in *E. coli* produces functional proteins, as is the case for *prfB* and *dnaX*, a functional role for +1 PheL frameshift seen in K-12 *E. coli* has not been found [52]. We believe that the bulk of the suspected recoding we found in *E. coli* simply reflects the limitations of the reduced function $RF2^{K-12}$ protein, which terminates poorly at UGA stops relative to $RF2^{B}$, a phenotype that is exacerbated by the absence of RF3. This idea is consistent with over-representation of UGA codons amongst ORFs likely to exhibit the greatest amount of recoding and with the strain specific behavior of the readthrough products of NudL and PheL and +1 frameshift product of PheL. Both NudL and PheL extended products increase about 5-fold in a K-12 $RF2^{K-12}\Delta RF3$ strain but were not present in K-12 $RF2^{B}$. Remarkably, they are also absent in K-12 $RF2^{K-12}\Delta RF3$ suggesting that in these cases $\Delta RF3$ alone is not sufficient to drive recoding and it might require a second perturbation of the weaker $RF2^{K-12}$ variant. It is likely that many of the other UGA stop codons with very high post-ORF ribosome occupancy will show the same patterns as NudL and PheL. While we examined recoding only for ORFs with very high post-ORF ribosome occupancy, the shift towards higher occupancy in K-12 $RF2^{K-12}\Delta RF3$ strains occurs in majority of the 1600 genes examined (Fig 4B), potentially producing defective protein products, albeit at a very low level.

Why might the defects in translation termination of $\Delta RF3$ in a K-12 background lead to increased recoding? As a consequence of release factor sequestration, the effective concentration of free RF1/2 is reduced, reducing termination efficiency and increasing the likelihood of recoding. During the increased time spend at stop codons, conformational flexibility in the complex may allow transient slippage and/or partial binding of an amino-acyl tRNA, which would be followed by binding of EF-G and subsequent elongation. The chances for EF-G binding are increased when cells lack RF3 because both proteins bind to the same position on the ribosome [13]. Recent studies of the kinetics of intragenic frameshifting reveal that it is

accompanied by slower EF-G catalyzed ribosome translocation and a change in the fluctuating conformational states [72–74]. Because of the altered ribosome conformation at the stop codon in the absence of RF3, we speculate that such conditions may also occur during slippage at the stop codon, with its frequency determined by sequences surrounding the stop codon.

Given the sensitivity of *E. coli* K-12 to perturbations in translation termination, we wonder why this lineage does not have the more conserved, fully functional $RF2^{B}$ protein, given that extreme translational stress of Δ*rluD* in K-12 quickly acquires suppressors at the RF2 locus [65]. There are several possibilities First, the alterations in gene expression for K-12 $RF2^{K\text{-}12}$ relative to K-12 $RF2^{B}$, in the presence or absence of RF3, may in aggregate be beneficial to the host, resulting in maintenance of the allele. Second, the lab strains derived from *E. coli* K-12 may not been under strong enough selective pressure to select for enhanced function release factor mutants. In support of this idea, when *E. coli* K-12 $RF2^{K\text{-}12}$ is exposed to extreme translational stress by deleting Δ*rluD* (rRNA pseudouridine synthase), it acquires suppressors at the RF2 locus [65]. Finally, K-12 $RF2^{K\text{-}12}$ may have acquired a compensatory mutation(s) enabling the reduced function allele to perform relatively well. Similar compensatory mutations have been documented in studies of antibiotic resistance [75]. If so, one might imagine that the compensatory mutations might be incompatible with the fully functional allele in some circumstances. The loss of viability of K-12 $RF2^{B}$ ΔRF3 but not K-12 $RF2^{K\text{-}12}$ ΔRF3 at very low temperatures is consistent with possibility that contemporary K-12 strains have acquired compensatory mutations, which increase their fitness.

Materials and Methods

Bacterial strains and their construction

All experiments were performed in *E. coli* K-12 strain MG1655 or its derivatives. Our strain name and genotype of each strain in K-12 *E. coli* are as follows; K-12: MG1655 *prfB* [*E. coli* K-12], $RF2^{K\text{-}12}\Delta$RF3: MG1655 *prfB* [*E. coli* K-12] Δ*prfC::frt*, K-12 $RF2^{B}$: MG1655 *prfB* [*E. coli* B], $RF2^{B}\Delta$RF3: MG1655 *prfB* [*E. coli* B] Δ*prfC::frt*. The K-12 $RF2^{K\text{-}12}\Delta$RF3 strain was constructed by transducing the *prfC::kanR* locus from the KEIO collection into MG1655 and its *kanR* cassette was flipped out using pCP20 [76]. The K-12 $RF2^{B}$ strain was constructed by transducing the *prfB* ($RF2^{B}$) locus with *kanR* downstream into MG1655, The original $RF2^{B}$ Kan construct was a gift of Lei Wang at The Salk Institute for Biological Studies. The linked *kanR* marker was removed by λ Red assisted homologous recombination to obtain $RF2^{B}$ without markers [77]. K-12 $RF2^{B}\Delta$RF3 was constructed using standard P1 transduction of the KEIO locus into the K-12 $RF2^{B}$ background.

Growth experiments

Cultures for growth curve experiments were grown in rich defined liquid media, MOPS complete-glucose overnight at 37°C then diluted to OD_{420} 0.005 in 35mL of fresh MOPS complete-glucose media. They were grown in a shaking water-bath incubator at 37°C with cell density measured by optical density at 420nm wavelength every 30 minutes until stationary phase. Cell densities over OD_{420} 0.5 were diluted to ensure we obtained values within the linear range of the spectrophotometer.

Cold sensitivity experiments

The spotting assay of cold sensitivity was performed by culturing cells to mid-exponential growth in LB media at 37°C and making serial 1 in 10 dilutions up to 10^{-6}. We spot plated 3uL of each

dilution onto LB plates in duplicate and incubated at 37°C, 30°C, 25°C, 20°C, and 15°C. Plates were removed from the incubators and imaged when individual colonies were visible. Cells for the cold viability experiments were prepared in the same manner and dilution series were plated on LB plates by glass beads in triplicate. Plates were incubated in the same temperature series and viable cells were counted once colonies were visible.

Ribosome profiling

Ribosome profiling was performed as previously described by Oh et al [49]. All strains; K-12, K-12 RF2$^{K\text{-}12}$ΔRF3, K-12 RF2^{B} and K-12 RF2^{B}ΔRF3 were grown at 37°C in MOPS complete-glucose, a rich defined media, as a liquid culture [78]. Previous ribosome profiling experiments have shown that ribosome pausing increases over serine codons during exponential growth in Luria broth due to serine being the first catabolized amino after sugar is utilized [78]. Use of MOPS complete-glucose eliminates these pauses and provides a more stable growth media for *E. coli* during our experiments. Profiling was also performed in K-12 and K-12 RF2$^{K\text{-}12}$ΔRF3 strains in MOPS minimal-glucose, a minimal defined media, for the results seen in Figure 7B. Cells from each strain were rapidly harvested by filtration and lysate was produced by pulverization of liquid nitrogen cooled samples. From this lysate ribosome footprints were created using MNase and total RNA was extracted for a simultaneous RNA-seq library production. Ribosome footprints between 24-32nt were isolated from the initial polyacrylamide sizing gel. Total RNA purified samples for RNA-seq were fragmented by alkaline fragmentation and 20-35nt fragments were isolated from the sizing gel. After ligation to Linker-1 the fragments were converted to a cDNA library and subsequent library preparations were performed as previously described [46,49]. Libraries were sequenced on an Illuminia HiSeq 2000 or HiSeq4000. Raw and processed data are available on the NCBI Gene Expression Omnibus (GEO) under the accession number GSE88725.

Sequencing analysis

Generated sequencing reads were analyzed as previously described [46,49]. Reads were initially aligned to a genome file containing *E. coli* rRNA and tRNA sequences to computationally subtract these reads. Then the remaining unaligned reads were aligned to the *E. coli* genome, reads with more than two mismatches were excluded as were reads aligning to multiple portions of the genome. Aligned reads between 20 and 40bp in length were trimmed by 10bp on each side and center mapped on the genome as previously described [49]. Wiggle files were generated using the same read lengths and center mapping as described above and counts were normalized per million reads and visualized in IGV. To analyze post-ORF occupancies as outlined in Fig 4, we also aligned ribosome profiling reads to the 3' end in accordance with a recent publication to increase reading frame information [79]. We found both center alignment and 3' alignments of fragments led to the same conclusion. Gene expression for each gene was calculated in Plastid, a Python library for deep sequencing genomics, by masking occupancy at the first and last 5 codons within the annotated ORF, and normalized for readdepth and gene length by using RPKM (reads per kilobase of transcript per million mapped reads) [80]. Expression between two strains was only compared when genes each had at least 100 counts. Metagene analysis was performed by normalizing 70-100nt upstream of the stop codon within each gene then plotting the median for each position. Genes with less than 1 counts in this window and genes less than 50nt away from the downstream coding frame were excluded from the metagene analysis in Fig 2 (note: S1 Fig differs slightly with a lack of read count filter to include more genes). Signals from biological replicates of different samples were averaged and standard deviations were calculated respectively.

Post-ORF region calculations

Genome wide calculations of ribosomes in the post-ORF occupancy were performed for all genes with an intergenic region of 65bp or greater, overlapping genes were also excluded, that had average mRNA and ribosome footprint densities of 0.1 or greater.

Post-ORF occupancy was calculated in a window 20-60bp downstream of the annotated stop codon. The start point of 20bp was chosen in order to eliminate the impact of ribosomes over the stop codon contributing to post-ORF counts. This window was utilized for all genes regardless if possible extensions were shorter or longer than that window. The average post-ORF occupancy of a gene was divided by the average occupancy of gene, which was calculated as described above, to obtain the relative post-ORF ribosome occupancy metric (RPOR). The RPOR roughly estimates the fraction of recoding occurring at each gene. For the cumulative distribution analysis all RPOR values of 0.0 were removed because we are unable to resolve between zero recoding events and insufficient read depth, the presence of zeros also falsely inflated the K-S statistic.

Extended-ORF tagging and western blot

The pFLAG-MAC plasmid background was used to construct inducible N-terminal-FLAG, C-terminal streptavidin tagged plasmids for *nudL* and *panZ* while only the N-terminal-FLAG was used for *pheL*. We cloned in the gene locus from directly after the start codon extending far past the annotated stop codon using restriction digest cloning with HindIII and EcoRI. Utilizing our ribosome profiling data we were able to identify the possible extension and cloned in a sufficient segment of each 3'-UTR, using predicted stop codons to place the C-terminal Strep tag. Each plasmid sequence was confirmed by Sanger sequencing and TSS transformed into K-12, K-12 $RF2^{K\text{-}12}\Delta RF3$, K-12 $RF2^{B}$ and K-12 $RF2^{B}$ $\Delta RF3$ strains. These strains containing pFLAG were cultured to mid-exponential growth and induced with 1mM IPTG for 2 hours and total cell protein was harvested using TCA precipitation. NudL and PanZ proteins were resolved on 12% Bis-Tris

gels with MOPS running buffer. PheL protein was resolved on 16% Tricine-SDS gels with Tricine-SDS running buffer. Proteins were transferred onto nitrocellulose membranes via wet transfer then blocked using Li-Cor Odyssey PBS blocking buffer and incubated with primary antibodies α-FLAG (Sigma-F3165) and α-SurA at 1:5000 dilutions and α-strep (abcam, ab76950) at a 1:2500 dilution overnight at 4°C. Membranes were washed with PBS before incubation with Li-Cor α-mouse and α-rabbit secondary antibodies at 1:10000 dilutions and again washed before imaging. Membranes were imaged on a Li-Cor infrared imager. Quantifications were performed using Image-J software and SurA as a loading control.

Attenuation reporters and assay

Two reporter constructs were assembled for each biosynthetic locus; *hisL*, *trpL* and *ivbL* in an *attλ* integratable plasmid. The leader reporter was constructed by cloning in the entire 5'UTR and leader peptide without its stop codon upstream of *lacZ* creating a translational fusion. The downstream reporter was constructed by cloning the entire 5'UTR, leader peptide, intergenic region and a small segment of the next gene downstream into the same plasmid backbone upstream of *lacZ*. Once all six plasmids were generated and validated by Sanger sequencing, we integrated them into both K-12 and K-12 $RF2^{K\text{-}12}$ ΔRF3 cells. Integration was performed using CRIM helper plasmid pINT-ts and single integrants were confirmed by diagnostic PCR [81]. After multiple attempts we found the *hisL* leader peptide reporter to be toxic to the cell, likely due to the high production of the *lacZ*-fusion.

Cultures containing reporter plasmids were grown in either MOPS complete-glucose or MOPS minimal-glucose overnight at 37°C and dilutions were performed to OD_{420} of 0.005 prior to the experiment in their respective medias in triplicate. Once cultures had reached exponential growth samples of 500uL and 1mL were taken simultaneously to measure β-galactosidase (β-gal) activity and OD_{420} at three timepoints. β-gal assay samples were immediately added to a tube containing 500uL Z-buffer, 1.25uL beta-mercaptoethanol, 30uL 0.1% SDS and 40uL

chloroform then vortexed before being placed on ice. Once collection was completed all β-gal sample tubes were incubated at 28°C prior to induction with 200uL 4mg/mL ONPG. Development was stopped with 500uL of 1M sodium bicarbonate and samples were centrifuged for 5 minutes at max speed to remove cellular debris. Samples were then transferred into a 96-well plate to measure OD_{420} in a Varioskan plate reader. Cell density and OD_{420} of developed β-gal samples were used to calculate Miller units of activity.

References

1. Zaher HS, Green R. Fidelity at the Molecular Level: Lessons from Protein Synthesis. Cell. Elsevier Inc.; 2009;136: 746–762. doi:10.1016/j.cell.2009.01.036
2. Francklyn CS. DNA Polymerases and Aminoacyl-tRNA Synthetases Shared Mechanisms for Ensuring the Fidelity of Gene Expression. Biochemistry. 2008;47.
3. Dunkle J a, Cate JHD. Ribosome structure and dynamics during translocation and termination. Annu Rev Biophys. 2010;39: 227–44. doi:10.1146/annurev.biophys.37.032807.125954
4. Koutmou KS, McDonald ME, Brunelle JL, Green R. RF3:GTP promotes rapid dissociation of the class 1 termination factor. RNA. 2014;20: 609–20. doi:10.1261/rna.042523.113
5. Korostelev A. Structural aspects of translation termination on the ribosome. RNA. 2011;17: 1409–21. doi:10.1261/rna.2733411
6. Freistroffer D V, Kwiatkowski M, Buckingham RH, Ehrenberg M. The accuracy of codon recognition by polypeptide release factors. Proc Natl Acad Sci U S A. 2000;97: 2046–51. doi:10.1073/pnas.030541097
7. Jørgensen F, Adamski F, Tate W, Kurland C. Release factor-dependent false stops are infrequent in Escherichia coli. J Mol Biol. 1993;230: 41–50.
8. Hirokawa G, Demeshkina N, Iwakura N, Kaji H, Kaji A. The ribosome-recycling step: consensus or controversy? Trends Biochem Sci. 2006;31: 143–9. doi:10.1016/j.tibs.2006.01.007
9. Grentzmann G, Brechemier-Baey D, Heurgué-Hamard V, Buckingham RH. Function of Polypeptide Chain Release Factor RF-3 in Escherichia coli. J Biol Chem. 1995; doi:10.1074/jbc.270.18.10595
10. Hirashima A, Kaji A. Role of elongation factor G and a protein factor on the release of ribosomes from messenger ribonucleic acid. J Biol Chem. 1973;248: 7580–7.
11. Moffat JG, Tate WP. A single proteolytic cleavage in release factor 2 stabilizes ribosome

binding and abolishes peptidyl-tRNA hydrolysis activity. J Biol Chem. 1994;269: 18899–18903.

12. Youngman EM, McDonald ME, Green R. Peptide release on the ribosome: mechanism and implications for translational control. Annu Rev Microbiol. 2008;62: 353–73. doi:10.1146/annurev.micro.61.080706.093323

13. Zhou J, Korostelev A, Lancaster L, Noller HF. Crystal structures of 70S ribosomes bound to release factors RF1, RF2 and RF3. Curr Opin Struct Biol. Elsevier Ltd; 2012;22: 733–42. doi:10.1016/j.sbi.2012.08.004

14. Mora L, Heurgué-Hamard V, De Zamaroczy M, Kervestin S, Buckingham RH. Methylation of bacterial release factors RF1 and RF2 is required for normal translation termination in vivo. J Biol Chem. 2007;282: 35638–35645. doi:10.1074/jbc.M706076200

15. Uno M, Ito K, Nakamura Y. Functional specificity of amino acid at position 246 in the tRNA mimicry domain of bacterial release factor 2. Biochimie. 1996;78: 935–43.

16. Dinçbas-Renqvist V, Engström Å. A post-translational modification in the GGQ motif of RF2 from Escherichia coli stimulates termination of translation. EMBO J. 2000;19. doi:10.1093/emboj/19.24.6900

17. Johnson DBF, Wang C, Xu J, Schultz MD, Schmitz RJ, Ecker JR, et al. Release factor one is nonessential in Escherichia coli. ACS Chem Biol. 2012;7: 1337–44. doi:10.1021/cb300229q

18. Johnson D, Xu J, Shen Z. RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. Nat Chem Biol. 2011;7: 779–786. doi:10.1038/nchembio.657.RF1

19. Adamski F, Donly BC, Tate WP. Competition between frameshifting, termination and suppression at the frameshift site in the Escherichia coli release factor-2 mRNA. Nucleic Acids Res. 1993;21: 5074–8.

20. Craigen WJ, Caskey CT. Expression of peptide chain release factor 2 requires high-

efficiency frameshift. Nature. 1986;322: 273–275.

21. Goldstein JL, Caskey CT. Peptide chain termination: effect of protein S on ribosomal binding of release factors. Proc Natl Acad Sci U S A. 1970;67: 537–543. doi:10.1073/pnas.67.2.537

22. Freistroffer D V, Pavlov MY, MacDougall J, Buckingham RH, Ehrenberg M. Release factor RF3 in E.coli accelerates the dissociation of release factors RF1 and RF2 from the ribosome in a GTP-dependent manner. EMBO J. 1997;16: 4126–33. doi:10.1093/emboj/16.13.4126

23. Klaholz BP, Myasnikov AG, Van Heel M. Visualization of release factor 3 on the ribosome during termination of protein synthesis. Nature. 2004;427: 862–865. doi:10.1038/nature02332

24. Leipe DD, Wolf YI, Koonin E V, Aravind L. Classification and evolution of P-loop GTPases and related ATPases. J Mol Biol. 2002;317: 41–72. doi:10.1006/jmbi.2001.5378

25. Gao H, Zhou Z, Rawat U, Huang C, Bouakaz L, Wang C, et al. RF3 Induces Ribosomal Conformational Changes Responsible for Dissociation of Class I Release Factors. Cell. 2007;129: 929–941. doi:10.1016/j.cell.2007.03.050

26. Helgstrand M, Mandava CS, Mulder FAA, Liljas A, Sanyal S, Akke M. The Ribosomal Stalk Binds to Translation Factors IF2, EF-Tu, EF-G and RF3 via a Conserved Region of the L12 C-terminal Domain. J Mol Biol. 2007;365: 468–479. doi:10.1016/j.jmb.2006.10.025

27. Mora L, Zavialov A, Ehrenberg M, Buckingham RH. Stop codon recognition and interactions with peptide release factor RF3 of truncated and chimeric RF1 and RF2 from Escherichia coli. Mol Microbiol. 2003;50: 1467–1476. doi:10.1046/j.1365-2958.2003.03799.x

28. Zavialov A V, Mora L, Buckingham RH, Ehrenberg M. Release of peptide promoted by

the GGQ motif of class 1 release factors regulates the GTPase activity of RF3. Mol Cell. 2002;10: 789–98. Available: http://www.ncbi.nlm.nih.gov/pubmed/12419223

29. Mikuni O, Kawakami K, Nakamura Y. Sequence and functional analysis of mutations in the gene encoding peptide-chain-release factor 2 of Escherichia coli. Biochimie. 1991;73: 1509–1516. doi:10.1016/0300-9084(91)90185-4

30. Elliott T, Wang X. Salmonella typhimurium prfA Mutants Defective in Release Factor 1. J Bacteriol. 1991;173.

31. Kawakami K, Nakamura Y. Autogenous suppression of an opal mutation in the gene encoding peptide chain release factor 2. Proc Natl Acad Sci U S A. 1990;87: 8432–6.

32. Rydén SM, Isaksson LA. A temperature-sensitive mutant of Escherichia coli that shows enhanced misreading of UAG/A and increased efficiency for tRNA nonsense suppressors. MGG Mol Gen Genet. 1984;193: 38–45. doi:10.1007/BF00327411

33. Mikuni O, Ito K, Moffat J, Matsumura K, McCaughan K, Nobukuni T, et al. Identification of the prfC gene, which encodes peptide-chain-release factor 3 of Escherichia coli. Proc Natl Acad Sci U S A. 1994;91: 5798–802.

34. Zaher HS, Green R. A primary role for release factor 3 in quality control during translation elongation in Escherichia coli. Cell. Elsevier Inc.; 2011;147: 396–408. doi:10.1016/j.cell.2011.08.045

35. O'Connor M. Interactions of release factor RF3 with the translation machinery. Mol Genet Genomics. 2015; doi:10.1007/s00438-015-0994-x

36. Kawakami K, Inada T, Nakamura Y. Conditionally lethal and recessive UGA-suppressor mutations in the prfB gene encoding peptide chain release factor 2 of Escherichia coli. J Bacteriol. 1988;170: 5378–5381.

37. Elliott T, Wang X. Salmonella typhimurium prfA Mutants Defective in Release Factor 1. J Bacteriol. 1991;173.

38. Matsumura K, Ito K, Kawazu Y, Mikuni O, Nakamura Y. Suppression of temperature-

sensitive defects of polypeptide release factors RF-1 and RF-2 by mutations or by an excess of RF-3 in Escherichia coli. J Mol Biol. 1996;258: 588–99. doi:10.1006/jmbi.1996.0271

39. Crawford DJ, Ito K, Nakamura Y, Tate WP. Indirect regulation of translational termination efficiency at highly expressed genes and recoding sites by the factor recycling function of Escherichia coli release factor RF3. EMBO J. 1999;18: 727–32. doi:10.1093/emboj/18.3.727

40. Petropoulos AD, McDonald ME, Green R, Zaher HS. Distinct roles for release factor 1 and release factor 2 in translational quality control. J Biol Chem. 2014;289: 17589–96. doi:10.1074/jbc.M114.564989

41. Grentzmann G, Brechemier-baey D, Heurgue V, Mora L, Buckingham RH. Localization and characterization of the gene encoding release factor RF3 in Escherichia coli. Microbiology. 1994;91: 5848–5852.

42. Nichols RJ, Sen S, Choo YJ, Beltrao P, Zietek M, Chaba R, et al. Phenotypic landscape of a bacterial cell. Cell. Elsevier Inc.; 2011;144: 143–56. doi:10.1016/j.cell.2010.11.052

43. Li GW, Burkhardt D, Gross C, Weissman JS. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. Cell. Elsevier; 2014;157: 624–635. doi:10.1016/j.cell.2014.02.033

44. Poole E, Brown C, Tate W. The identity of the base following the stop codon determines the efficiency of in vivo translational termination in Escherichia coli. EMBO J. 1995;14: 151–158.

45. Ozawa Y, Saito R, Washio T, Tomita M. Comparative study of translation termination sites and release factors (RF1 and RF2) in procaryotes. J Mol Evol. 2003;56: 665–72. doi:10.1007/s00239-002-2435-9

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:

218–23. doi:10.1126/science.1168978

47. Dunn JG, Foo CK, Belletier NG, Gavis ER, Weissman JS. Ribosome profiling reveals pervasive and regulated stop codon readthrough in Drosophila melanogaster. Elife. 2013;2: e01179. doi:10.7554/eLife.01179

48. Dingwall C, Lomonossoff GP, Laskey RA. High sequence specificity of micrococcal nuclease. Nucleic Acids Res. 1981;9: 5287–5296.

49. Oh E, Becker AH, Sandikci A, Huber D, Chaba R, Gloge F, et al. Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. Cell. 2011;147: 1295–308. doi:10.1016/j.cell.2011.10.044

50. Gurvich O, Baranov P. Sequences that direct significant levels of frameshifting are frequent in coding regions of Escherichia coli. EMBO J. 2003;22: 5941–5950. doi:10.1093/emboj/cdg561

51. Namy O, Rousset J, Napthine S, Brierley I. Reprogrammed Genetic Decoding in Cellular Gene Expression. Mol Cell. 2004;13: 157–168.

52. Gurvich OL, Näsvall SJ, Baranov P V, Björk GR, Atkins JF. Two groups of phenylalanine biosynthetic operon leader peptides genes: a high level of apparently incidental frameshifting in decoding Escherichia coli pheL. Nucleic Acids Res. 2011;39: 3079–92. doi:10.1093/nar/gkq1272

53. Huvet M, Stumpf MP. Overlapping genes: a window on gene evolvability. BMC Genomics. 2014;15: 721. doi:10.1186/1471-2164-15-721

54. Aksoy S, Squires CL, Squires C. Translational coupling of the trpB and trpA genes in the Escherichia coli Tryptophan operon. J Bacteriol. 1984;157: 363–367.

55. Oppenheim DS, Yanofsky C. Translational coupling during expression of the tryptophan operon of Escherichia coli. Genetics. 1980;95: 785–795.

56. Schumperli D, McKenney K, Sobieski DA, Rosenberg M. Translational coupling at an intercistronic boundary of the Escherichia coli galactose operon. Cell. 1982;30: 865–871.

doi:10.1016/0092-8674(82)90291-4

57. Yates JL, Nomura M. Feedback regulation of ribosomal protein synthesis in E. coli: Localization of the mRNA target sites for repressor action of ribosomal protein L1. Cell. 1981;24: 243–249. doi:10.1016/0092-8674(81)90520-1

58. Burkhardt DH, Rouskin S, Zhang Y, Li G-W, Weissman JS, Gross CA. Operon mRNAs are organized into ORF-centric structures that predict translation efficiency. Elife. 2017;6: e22037. doi:10.7554/eLife.22037

59. Yanofsky C. Transcription Attenuation : Once Viewed as a Novel Regulatory Strategy. J Bacteriol. 2000;182. doi:10.1128/JB.182.1.1-8.2000.

60. Roesser JR, Nakamurall Y, Yanofsky C. Regulation of Basal Level Expression of the Tryptophan Operon of. J Biol Chem. 1989;264: 12284–12288.

61. Tate WP, Poole ES, Dalphin ME, Major LL, Crawford DJG, Mannering SA. The translational stop signal: Codon with a context, or extended factor recognition element? Biochimie. 1996;78: 945–952. doi:10.1016/S0300-9084(97)86716-8

62. Zavialov A V., Hauryliuk V V., Ehrenberg M. Splitting of the posttermination ribosome into subunits by the concerted action of RRF and EF-G. Mol Cell. 2005;18: 675–686. doi:10.1016/j.molcel.2005.05.016

63. Zavialov a V, Buckingham RH, Ehrenberg M. A posttermination ribosomal complex is the guanine nucleotide exchange factor for peptide release factor RF3. Cell. 2001;107: 115–24. doi:10.1016/S0092-8674(01)00508-6

64. Weiss RB, Murphy JP, Gallant J a. Genetic screen for cloned release factor genes. J Bacteriol. 1984;158: 362–4.

65. O'Connor M, Gregory ST. Inactivation of the RluD pseudouridine synthase has minimal effects on growth and ribosome function in wild-type Escherichia coli and Salmonella enterica. J Bacteriol. 2011;193: 154–162. doi:10.1128/JB.00970-10

66. Caskey CT, Beaudet A, Nirenberg M. RNA codons and protein synthesis. 15. Dissimilar

responses of mammalian and bacterial transfer RNA fractions to messenger RNA codons. J Mol Biol. 1968;37: 99–118.

67. MacBeath G, Kast P. UGA read-through artifacts-when popular gene expression systems needs a pATCH. Biotechniques. 1998;794: 789–794.

68. Engelberg-Kulka H, Liu Z, Li C, Reches M. An extended Escherichia coli “selenocysteine insertion sequence” (SECIS) as a multifunctional RNA structure. Biofactors. 2001;14: 61–68.

69. Zhang Y, Gladyshev VN. An algorithm for identification of bacterial selenocysteine insertion sequence elements and selenoprotein genes. Bioinformatics. 2005;21: 2580–2589. doi:10.1093/bioinformatics/bti400

70. Namy O, Lecointe F. Translational recoding and RNA modifications. Topics in Current Genetics. 2005;12. doi:10.1007/b106847

71. Atkins J, Gesteland R. Recoding: Expansion of Decoding Rules Enriches Gene Expression [Internet]. 2010. Available: http://link.springer.com/content/pdf/10.1007/978-3-642-83384-7.pdf

72. Caliskan N, Katunin VI, Belardinelli R, Peske F, Rodnina M V. Programmed -1 frameshifting by kinetic partitioning during impeded translocation. Cell. Elsevier Inc.; 2014;157: 1619–31. doi:10.1016/j.cell.2014.04.041

73. Kim H-K, Liu F, Fei J, Bustamante C, Gonzalez RL, Tinoco I. A frameshifting stimulatory stem loop destabilizes the hybrid state and impedes ribosomal translocation. Proc Nat Acad Sci USA. 2014;111: 5538–43. doi:10.1073/pnas.1403457111

74. Chen J, Petrov A, Johansson M, Tsai A, O’Leary SE, Puglisi JD. Dynamic pathways of -1 translational frameshifting. Nature. 2014;512: 328–32. doi:10.1038/nature13428

75. Levin BR, Perrot V, Walker N. Compensatory mutations, antibiotic resistance and the population genetics of adaptive evolution in bacteria. Genetics. 2000;154: 985–997.

76. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, et al. Construction of

Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol. 2006;2: 2006.0008. doi:10.1038/msb4100050

77. Thomason L, Court DL, Bubunenko M, Costantino N, Wilson H, Datta S, et al. SPECIALIZED TECHNIQUES UNIT 1.16 Recombineering: Genetic Engineering in Bacteria Using Homologous Recombination. Curr Protoc Mol Biol. 2007;1: 1.

78. Li G-W, Oh E, Weissman JS. The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. Nature. Nature Publishing Group; 2012;484: 538–41. doi:10.1038/nature10965

79. Woolstenhulme CJ, Guydosh NR, Green R, Buskirk AR. High-Precision analysis of translational pausing by ribosome profiling in bacteria lacking EFP. Cell Rep. The Authors; 2015;11: 13–21. doi:10.1016/j.celrep.2015.03.014

80. Dunn JG. plastid: a positional library for sequencing analysis [Internet]. 2016. Available: http://plastid.readthedocs.io

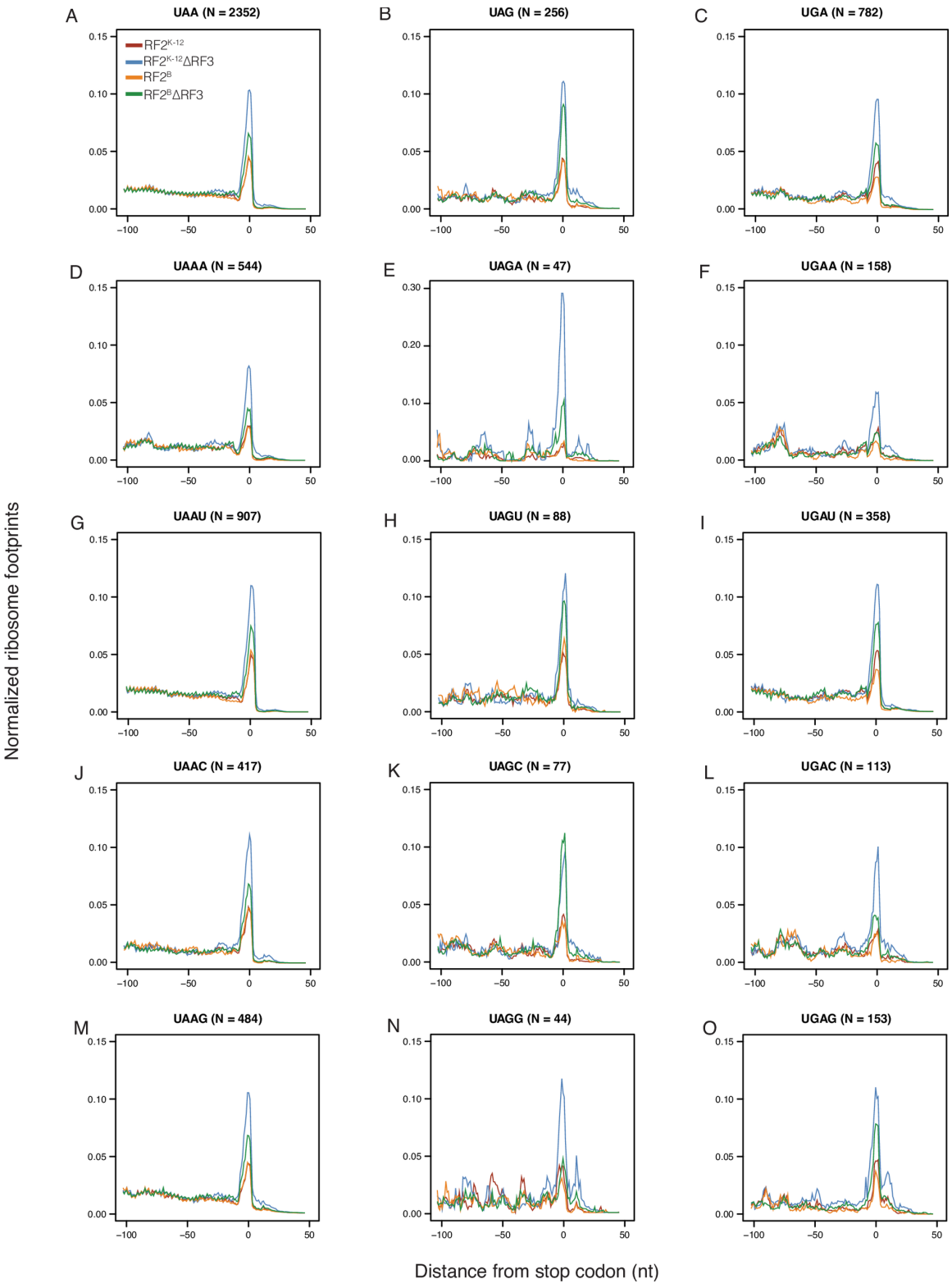
81. Haldimann A, Wanner BL. Conditional-Replication , Integration , Excision , and Retrieval Plasmid-Host Systems for Gene Structure-Function Studies of Bacteria. Society. 2001;183: 6384–6393. doi:10.1128/JB.183.21.6384

Supplemental Information

Supplemental figure S1. Ribosome occupancy over stop codons.

Metagene plots of ribosome density for each 3-base stop codon (UAA, UAG and UGA) and each four-base stop codon (each stop codon with a varying fourth position). Shown is the median ribosome footprint density in the region surrounding stop codons for K-12 $RF2^{K-12}$ (4), K-12 $RF2^{K-12}\Delta RF3$ (2), K-12 $RF2^{B}$ (2) and K-12 $RF2^{B}\Delta RF3$ (5) strains grown in MOPS complete-glucose media at 37°. The numbers in parentheses following each strain indicates the number of repeat determinations. In order to improve resolution for the rare UGA codon and across all possible four-base stop codons, we did not select genes with high read density, this dataset included 3390 genes.

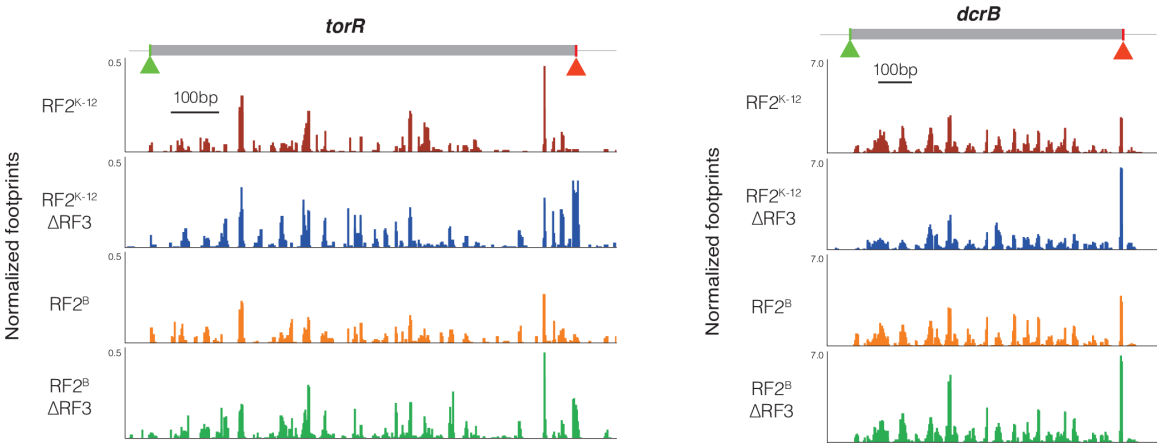
Supplemental figure S1



Supplemental figure S2. Increased ribosome occupancy over stop codons of single genes.

The normalized ribosome occupancy is shown over two randomly chosen genes, *torR* (A) and *dcrB* (B) for all strains; K-12 $RF2^{K-12}$, K-12 $RF2^{B}$, K-12 $RF2^{K-12}$ΔRF3 and K-12 $RF2^{B}$ΔRF3. The start codon of each gene is annotated with a green triangle and stop codon with a red triangle. Peak intensity between strains remains relatively consistent with the largest variability over the stop codon.

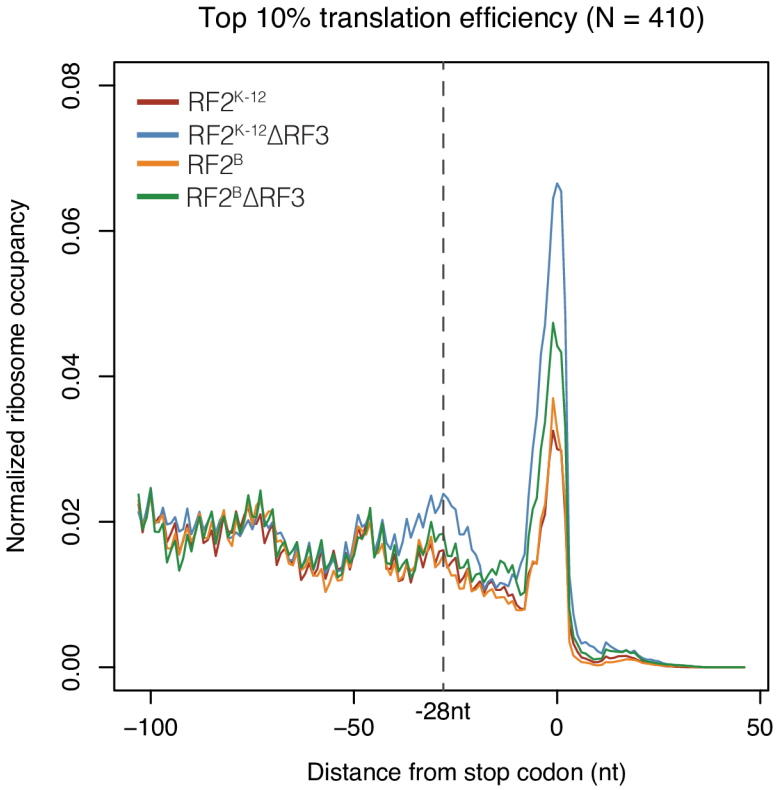
Supplemental figure S2



Supplemental figure S3. Ribosome occupancy over stop codons of genes with high translation efficiency.

Metagene analysis of ribosome footprint density in the region surrounding stop codons. The top 10% of genes with the highest translation efficiency were aligned at their stop codon and the median normalized ribosome density at each position was calculated from ribosome profiling data of strains grown in MOPS complete-glucose media at 37°C. Translation efficiency was defined as the rate of protein production per mRNA molecule (RPKM of ribosome profiling reads normalized by RPKM of mRNA-seq read). Average normalized density was calculated across repeat experiments for K-12 $RF2^{K-12}$, K-12 $RF2^{K-12}\Delta RF3$, K-12 $RF2^{B}$ and K-12 $RF2^{B}\Delta RF3$ containing 4, 2, 2 and 5 datasets respectively for each stop codon. In K-12 $RF2^{K-12}\Delta RF3$ a slight increase in ribosome occupancy is seen approximately 28nt upstream of the stop codon.

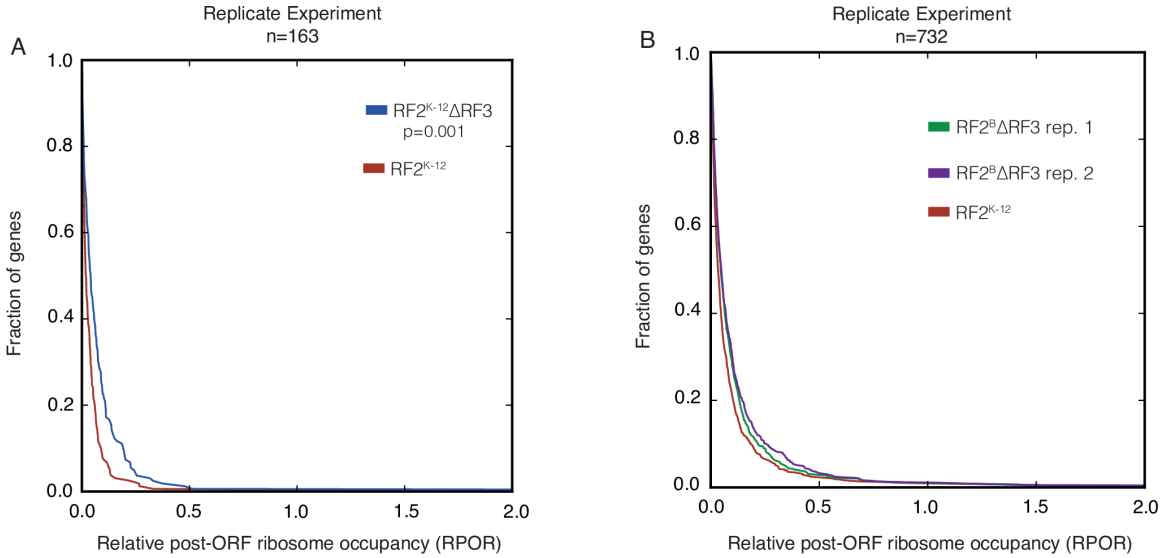
Supplemental figure S3



Supplemental figure S4. Post-ORF ribosome occupancy of ΔRF3 mutants for replicate experiments.

(A) The distribution of RPOR values from 0 to 2.0 for K-12 $RF2^{K-12}$ and K-12 $RF2^{K-12}$ΔRF3 strains are shown as a cumulative distribution function for a replicate experiment to Figure 3B. After removal of all zero RPOR values, 163 genes were analyzed. The shift of K-12 $RF2^{K-12}$ΔRF3 to higher RPOR values is statistically significant (p-value = 0.001; K-S test). (B) The distribution of RPOR values from 0 to 2.0 for K-12 $RF2^{K-12}$ and two K-12 $RF2^{B}$ΔRF3 replicates in the same experiment using a cumulative distribution function. This analysis included 732 genes after all zero value PORs were removed. The two K-12 $RF2^{B}$ΔRF3 replicates are both statistically significant when compared to $RF2^{K-12}$ (p-value < 0.005).

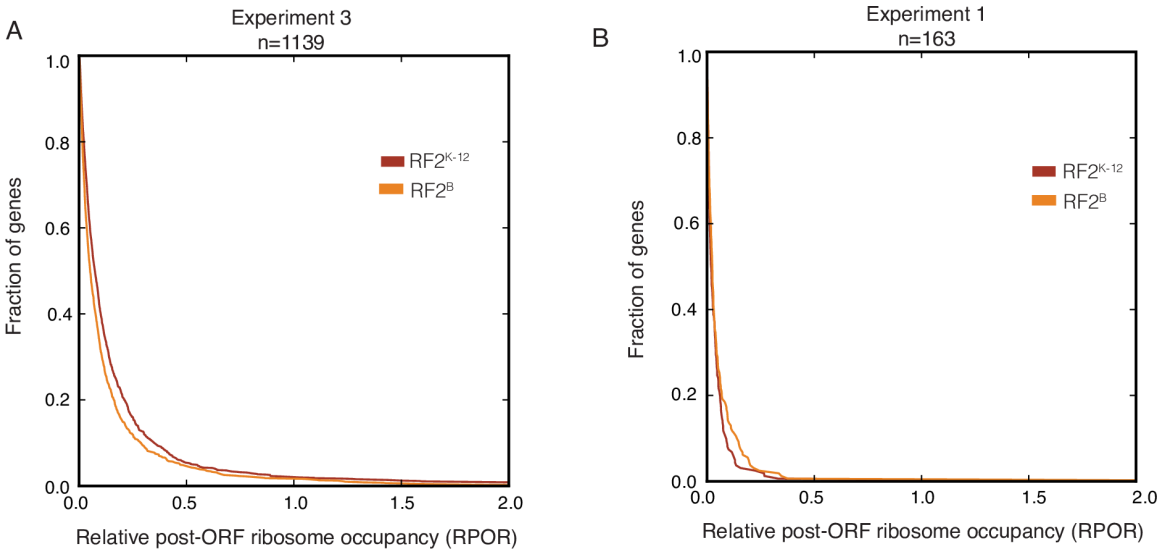
Supplemental figure S4



Supplemental figure S5. Post-ORF ribosome occupancy between $RF2^{B}$ mutants.

Relative post-ORF ribosome occupancy (RPOR) values between 0 and 2.0 of K-12 $RF2^{K-12}$ and K-12 $RF2^{B}$ replicate experiments are shown as cumulative distribution function. (A) Our largest dataset of 1139 genes analyzed shows a slight shift towards lower RPOR values in K-12 $RF2^{B}$ versus K-12 $RF2^{K-12}$, which is statistically significant (p-value = 5.3×10^{-6}; K-S test). (B) A small dataset comprised of 163 genes analyzed, shows a slight statistically insignificant shift towards higher RPOR values for K-12 $RF2^{B}$ versus K-12 $RF2^{K-12}$.

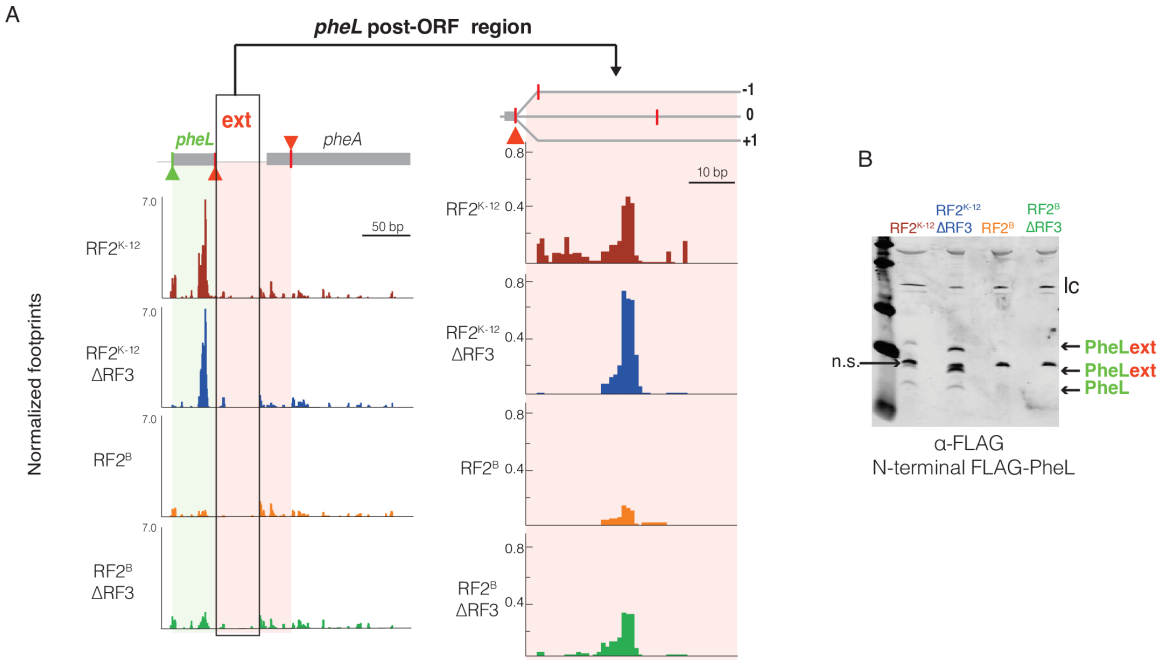
Supplemental figure S5



Supplemental figure S6. Frameshifting at the *pheL* locus in all RF mutants.

(A) Normalized ribosome footprints are shown across the locus of *pheL* (shaded green) and post-ORF region with hypothesized extensions shaded in red, which extends into *pheA*. Post-ORF stop codons are annotated for each reading frame with red bars; the known *pheL* +1 extension stop codon is marked with a red triangle within *pheA*. (A) A section of the *pheL* post-ORF region is enlarged and we see a reduction in ribosome density correlated with a 0 frame stop codon potentially signally a readthrough event. (B) Western blots of α-FLAG and α-SurA (loading control, lc) for all strains. In addition to a non-specific binding product (n.s.) we observe full length PheL at 3.26kDa, and two possible extended PheL products, a 4.85kDa 0-frame product and a 9.0kDa consistent with a +1 frameshift. We estimate the +1 frameshift efficiency in K-12 $RF2^{K\text{-}12}$ to be 52% and increases to 84% in K-12 $RF2^{K\text{-}12}$ΔRF3. We do not observe any PheL products in either K-12 $RF2^{B}$ strain; this is apparently consistent with the drastic shift in ribosome occupancy over *pheL* in K-12 $RF2^{B}$ strains (A).

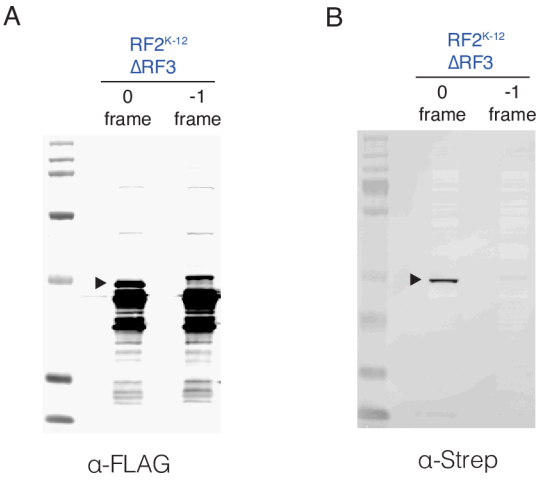
Supplemental figure S6



Supplemental figure S7. Western blots of C-terminally tagged NudL constructs.

Western blot of K-12 RF2$^{K\text{-}12}$ΔRF3 strain containing a plasmid encoding N-terminal-FLAG NudL, with a C-terminal streptavidin tag on either the 0 frame or -1 frame extension. Membranes were blotted for both FLAG (A) and Streptavidin (B). (A and B) Flag tagged NudL products are seen in both 0 and -1 frame constructs, however only the 0 frame product blots for streptavidin, as indicated by the black arrow. The 0 frame construct was utilized for further studies.

Supplemental figure S7

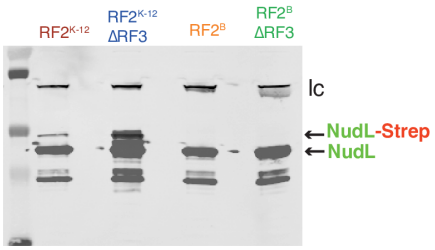


Supplemental figure S8. Western blots of NudL and PanZ constructs in all RF mutants.

Western blot of α-FLAG and α-streptavidin for induced constructs of N-terminal-FLAG-NudL (A) or PanZ (B) in K-12 $RF2^{K-12}$, K-12 $RF2^{B}$, K-12 $RF2^{K-12}$ΔRF3 and K-12 $RF2^{B}$ΔRF3. SurA was used as a loading control (lc). (A) The extended NudL protein visible in K-12 $RF2^{K-12}$ and K-12 $RF2^{K-12}$ ΔRF3 is no longer visible in either K-12 $RF2^{B}$ or K-12 $RF2^{B}$ΔRF3. (B) The presence of the streptavidin tagged -1 frameshift extended PanZ protein is seen in similar abundance in all strains, including K-12 $RF2^{B}$ and K-12 $RF2^{B}$ ΔRF3, at approximately 35% of total PanZ.

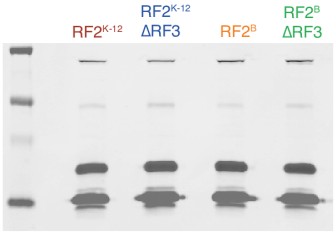
Supplemental figure S8

A



α-FLAG & α-Strep
NudL 0 frame extension

B

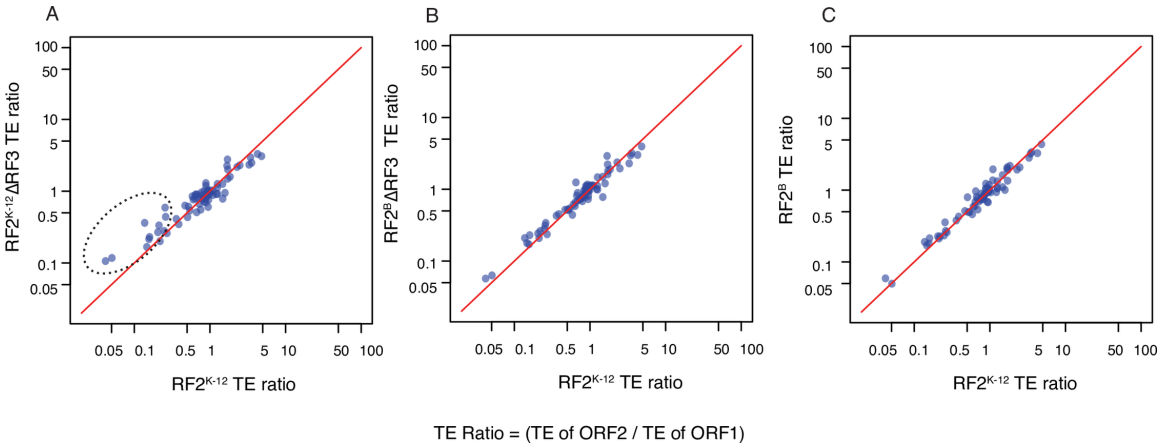


α-FLAG & α-Strep
PanZ -1 frame extension

Supplemental figure S9. Translation of overlapping ORF pairs.

The ratio of translation levels of overlapping ORFs [downstream (ORF2)/upstream (ORF1)] were compared between K-12 $RF2^{K-12}$, K-12 $RF2^{K-12}$ΔRF3, K-12 $RF2^{B}$, and K-12 $RF2^{B}$ΔRF3 strains. A total of 72 ORF pairs were analyzed. The translation level of each ORF was quantified by translation efficiency (TE), defined as the rate of protein production per mRNA molecule (RPKM of ribosome profiling reads normalized by RPKM of mRNA-seq read). The red diagonal lines represent equivalent expression of the two strains being compared. (A) Gene pairs with a TE ratio <0.5 have increased expression of downstream genes in the K-12 $RF2^{K-12}$ΔRF3 strain, indicated by the dotted ellipse. (B and C) Compared to the K-12 $RF2^{K-12}$ strain, K-12 $RF2^{B}$ ΔRF3 strain (B) and K-12 $RF2^{B}$ strain (C) show equivalent expression across all gene pairs.

Supplemental figure S9

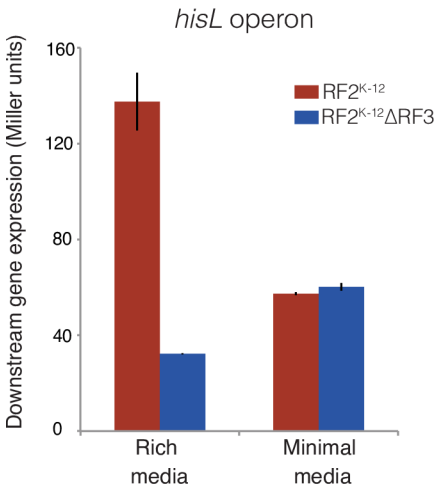


TE Ratio = (TE of ORF2 / TE of ORF1)

Supplemental figure S10. Attenuation of *hisL* operon in ΔRF3.

A reporter plasmid for *hisL* attenuation was constructed by fusing lacZ to the first gene downstream of the leader peptide, *hisG*. The promoter plasmid fused to the leader peptide used to normalize transcription from the operon promoter (see Fig 7C, bottom), was toxic under the control of the *hisL* promoter, and could not be used to normalize expression of the downstream gene reporter construct. We therefore report only the β-galactosidase activity of the downstream reporter in Miller Units for K-12 $RF2^{K\text{-}12}$ and K-12 $RF2^{K\text{-}12}$ΔRF3 backgrounds in MOPS complete-glucose and MOPS minimal-glucose media.

Supplemental figure S10



Supplemental table S1. Expression of release factors in K-12 and all RF mutants.

Gene expression as measured by ribosome footprint density (reads per kilobase of transcript per million mapped reads, RPKM) for the three release factors *prfA*, *prfB*, and *prfC.* The average ribosome density and standard error of the mean were calculated from replicate data sets: K-12 $RF2^{K\text{-}12}$ (n=4); K-12 $RF2^{K\text{-}12}\Delta$RF3 (n=2); K-12 $RF2^{B}$ (n=2); K-12 $RF2^{B}\Delta$RF3 (n=5). Ribosome profiling was performed in MOPS-complete medium at 37°C. As described briefly in the method sections, and in detail by Li et al. [40], ribosome density is proportional to the rate of synthesis of each protein.

Supplemental table S1

Locus	**RF2^{K-12}**	**RF2^{K-12}ΔRF3**	**RF2^{B}**	**RF2^{B}ΔRF3**
prfA (RF1)	82.8 ± 10.7	118.7 ± 15.0	87.1 ± 17.1	74.3 ± 14.8
prfB (RF2)	736.7 ± 113.7	1069.5 ± 77.7	526.2 ± 57.8	508.0 ± 99.3
prfC (RF3)	438.9 ± 80.7	--	461.7 ± 120.5	--

Gene expression (Ribosome density RPKM)

Supplemental table S2. Annotated likely recoding events.

Of the top 100 post-ORF ribosome occupancy (RPOR) values in K-12 $RF2^{K-12}$ and K-12 $RF2^{K-12}\Delta$RF3 (121 total), 43 were classified as likely recoding events because they experienced reduced ribosome density following a stop codon in any one of the three possible reading frames, a proxy for active post-ORF translation (See Fig 4). The RPOR value for strains; K-12 $RF2^{K-12}$, K-12 $RF2^{K-12}\Delta$RF3, K-12 $RF2^{B}$ and K-12 $RF2^{B}\Delta$RF3 are shown for each of the genes. The genes with asterisk (*) indicate those used in follow-up studies with N-terminal-FLAG, C-terminal streptavidin constructs to visualize extended products.

Supplemental table S2

gene	$RF2^{K-12}$ RPOR	$RF2^{K-12}$ ΔRF3 RPOR	$RF2^{B}$ RPOR	$RF2^{B}$ ΔRF3 RPOR
ydbL	5.75	4.65	2.33	3.42
ybbB	2.31	4.14	1.21	1.86
ydcA	1.91	3.11	1.43	1.78
nadR	1.73	2.63	1.51	1.85
barA	1.58	6.80	0.63	0.23
yegH	1.09	2.06	1.24	1.18
yiaU	1.00	2.69	0.80	0.72
*nudL**	0.89	0.93	1.20	0.81
yhiN	0.83	3.66	1.75	1.54
yneE	0.80	0.89	0.44	1.29
setA	0.75	0.84	0.09	0.46
ycbB	0.71	2.70	0.59	0.86
betA	0.64	1.64	0.60	1.14
ybiU	0.64	1.61	0.66	0.76
yidQ	0.63	0.91	0.09	0.20
hcaT	0.57	0.74	0.44	0.72
ybhK	0.56	0.31	0.64	0.36
*pheL**	0.51	1.10	0.41	0.72
ydiH	0.49	0.80	0.64	1.22
yjaB	0.48	0.67	0.05	0.02
yneG	0.47	0.52	0.29	0.27
ydgC	0.46	0.16	0.37	0.60
udk	0.46	1.21	0.48	0.67
flk	0.45	0.52	0.38	0.50
argO	0.44	0.69	0.25	0.27
fhuF	0.43	0.64	0.30	0.12
*panZ**	0.42	0.96	0.54	0.47
dbpA	0.42	0.64	0.19	0.48
pdxH	0.41	0.35	0.39	0.64
yadS	0.41	1.24	0.14	0.33
pspG	0.39	1.93	0.42	0.60
dusC	0.38	0.89	0.59	0.82
yiaJ	0.34	0.78	0.11	0.56
aat	0.32	1.98	0.86	1.25
zwf	0.27	0.65	0.15	0.22
yafJ	0.26	0.97	0.33	0.19
obgE	0.26	0.62	0.21	0.35

gene	$RF2^{K-12}$ RPOR	$RF2^{K-12}$ ΔRF3 RPOR	$RF2^{B}$ RPOR	$RF2^{B}$ ΔRF3 RPOR
yidE	0.23	0.66	0.23	0.08
mpaA	0.21	0.70	0.00	0.23
ygiH	0.12	0.73	0.13	0.14
dapB	0.12	0.75	0.16	0.25
dsdC	0.08	0.66	0.17	0.10
ydiA	0.02	0.66	0.16	0.46

Supplemental table S3. Annotated possible recoding events.

Of the top 100 post-ORF ribosome occupancy (RPOR) values in K-12 $RF2^{K\text{-}12}$ and K-12 $RF2^{K\text{-}12}\Delta$RF3 strains (121 total), 41 were classified as possible recoding events because in addition to the reduction in ribosome occupancy after a stop codon described above, possible confounding effects indicated in the column labeled Class were present. a: unannotated downstream ORF; b: REP-element in the post-ORF region; c: Shine-Dalgarno (SD) or downstream gene is located within the post-ORF region; d: ribosomes from another unknown source; e: Low reads in the post-ORF region.

Supplemental table S3

gene	$RF2^{K-12}$ RPOR	$RF2^{K-12}$ ΔRF3 RPOR	$RF2^{B}$ RPOR	$RF2^{B}$ ΔRF3 RPOR	class
ykgG	3.21	13.53	6.29	3.12	a
xapR	2.06	5.24	1.07	1.41	b & e
yjdP	1.51	1.22	1.09	1.10	e
cusS	1.47	2.02	0.67	5.62	e
yabP	1.29	2.53	0.47	0.55	c
pspD	1.21	4.41	0.56	0.88	c
chbG	1.03	1.90	0.67	1.08	b & e
polB	0.97	2.05	0.56	1.09	e
cvrA	0.89	0.51	0.51	0.37	e
ftsK	0.89	1.00	0.52	1.13	b
yaiA	0.87	1.14	0.50	0.77	c
ypfI	0.83	1.22	0.82	2.15	e
recC	0.81	0.71	0.66	0.94	e
intZ	0.78	0.61	0.32	0.26	c
yidZ	0.69	3.53	1.05	3.01	e
yiaG	0.63	0.19	0.46	0.84	e
hipA	0.56	0.33	0.15	0.34	d
yabP	0.56	0.95	0.47	0.55	c
yeeA	0.55	0.47	0.21	0.54	c
rmuC	0.55	0.25	0.52	0.16	e
fre	0.52	0.53	0.07	0.10	b
yjdJ	0.52	0.19	0.17	0.13	e
mltA	0.49	0.64	0.65	1.00	c
hpf	0.49	0.26	0.29	0.45	c
pitA	0.46	0.31	0.22	0.31	c
yhbQ	0.45	1.05	0.78	0.78	e
rcsC	0.44	0.19	0.49	0.38	e
rrmJ	0.44	0.07	0.15	0.22	e
yfiE	0.42	0.63	0.08	0.23	e
cobC	0.366	1.138	1.42	0.91	d
sbcC	0.317	2.273	0.27	0.58	d
mglC	0.301	0.709	0.25	0.34	e
xylR	0.275	0.841	1.55	0.45	b & e
pspF	0.257	0.758	0.12	0.47	c
gadX	0.241	0.907	0.07	0.83	e
eptB	0.200	0.999	0.24	0.41	d
cueR	0.179	1.166	0.13	0.39	d

gene	$RF2^{K-12}$ RPOR	$RF2^{K-12}$ ΔRF3 RPOR	$RF2^{B}$ RPOR	$RF2^{B}$ ΔRF3 RPOR	class
uspB	0.144	0.845	0.00	0.46	e
livF	0.110	0.971	0.05	0.34	c
appY	0.098	1.660	0.17	1.24	e
yhdP	0.080	0.962	0.37	0.64	c
yddE	0.075	1.157	0.00	0.17	b & e
yhgE	0.000	0.781	0.74	0.94	e

Classes:

a: unannotated downstream element

b: REP-element

c: Shine-Dalgarno (SD)

d: ribosomes from another unknown source

e: low reads within the post-ORF region

Supplemental table S4. Annotated non-recoding events.

Of the top 100 post-ORF ribosome occupancy (RPOR) values in K-12 $RF2^{K-12}$ and K-12 $RF2^{K-12}\Delta$RF3 strains (121 total), 37 were classified as non-recoding events. The post-ORF region of these genes did not exhibit a reduction in ribosome density after stop codons in any frame suggestive that ribosome occupancy was not due to active translation (See Fig 4). Sequence elements of these post-ORF region and possible sources of these post-ORF ribosomes were annotated in Column labeled Class as: a. small RNA; b. REP-element c: Shine-Dalgarno (SD) or downstream gene; d: ribosomes from an unknown source; e: ribosome binding region.

Supplemental table S4

gene	$RF2^{K-12}$ RPOR	$RF2^{K-12}$ ΔRF3 RPOR	$RF2^{B}$ RPOR	$RF2^{B}$ ΔRF3 RPOR	class
sgrR	2.16	1.49	1.12	1.07	a
yifN	2.05	0.74	0.95	0.71	a
yiiQ	2.04	1.09	1.23	0.98	c
yjdC	1.34	0.42	0.45	0.47	d
clcA	1.27	1.49	0.84	0.98	c
ybjN	1.10	1.18	0.32	0.36	d
iclR	0.89	0.61	0.22	0.25	c
htrL	0.79	0.47	0.55	0.82	d
ybjL	0.70	0.44	0.17	0.25	d
rpmJ	0.69	0.13	0.41	0.62	e
btuD	0.64	0.38	0.30	0.78	c
mtfA	0.62	1.86	1.29	2.08	d
ycjG	0.62	0.37	0.35	0.18	d
yqaA	0.57	0.90	0.87	0.40	c
ybbN	0.55	0.67	0.62	0.66	b
rnd	0.55	0.15	0.22	0.24	d
cmk	0.53	0.60	0.46	0.27	e
yjeT	0.52	0.62	0.47	0.56	d
pyrG	0.48	0.20	0.23	0.39	d
cpxP	0.48	1.29	0.41	0.44	a
infC	0.45	0.22	0.32	0.38	d
ycgX	0.44	1.73	1.11	1.27	d
ycgJ	0.43	0.62	0.41	0.49	d
yfbR	0.43	0.10	0.22	0.33	b
tolA	0.42	0.27	0.21	0.32	d
ynfA	0.42	0.21	0.58	0.34	d
mobB	0.41	0.13	0.05	0.52	b
iap	0.41	0.06	0.08	0.59	d
ycaK	0.33	0.65	0.25	0.86	d
ygjR	0.33	0.76	0.04	0.22	d
uup	0.32	0.65	0.20	0.31	c
csrD	0.24	0.64	0.34	0.53	d
cho	0.23	0.74	0.44	1.00	c
araC	0.22	0.74	0.25	0.18	d
yfcO	0.20	0.73	0.12	0.54	c
yhjJ	0.11	0.62	0.24	0.34	b
bssS	0.09	1.96	0.17	0.54	c

Classes:

a: small RNA

b: REP-element

c: Shine-Dalgarno (SD)

d: ribosomes from another unknown source

e: ribosome binding region

Supplemental table S5. Misannotations resulting in high RPOR values.

Of the top 100 post-ORF ribosome occupancy (RPOR) values in K-12 $RF2^{K\text{-}12}$ and K-12 $RF2^{K\text{-}12}\Delta$RF3 strains, four instances were the result of misannotations. The post-ORF region of the genes listed was located over an unidentified or recently identified ORF.

Supplemental table S5

gene	$RF2^{K-12}$ RPOR	$RF2^{K-12}$ ΔRF3 RPOR	$RF2^{B}$ RPOR	$RF2^{B}$ ΔRF3 RPOR
ydcM	4.07	6.70	4.80	7.68
yeaP	2.69	1.59	2.25	1.89
wbbK	2.54	1.15	1.92	1.68
yebW	1.60	4.11	1.32	0.92

Supplemental table S6. Stop codon distribution of annotated recoding events.

The distribution of stop codons of the 84 genes classified as likely or possible recoding events in S2 and 3 Tables is shown in the observed column. The expected values were calculated using the frequencies in which each codon, UAA, UAG and UGA occur across the *E.coli* K-12 genome. The Chi-squared value was calculated to be 26.277 with 2 degrees of freedom, which is statistically significant by two-tailed t-test with a p-value of less than 0.0001.

Supplemental table S6

Stop Codon	Observed	Expected
UAA	40.5% (n=34)	64.3% (n=54)
UAG	11.9% (n=10)	7.1% (n=6)
UGA	47.6% (n=40)	28.6% (n=24)
Total	100% (n=84)	100% (n=84)

Chi-squared = 26.277 df=2 p-value <0.0001

Appendix A

Protocol optimization for use of RNaseA/T1 nuclease for ribosome footprint digestion for bacterial ribosome profiling

Introduction

In Chapter 1, I established that K-12 strains without RF3 have a higher number of ribosomes located in the region down stream of the open reading frame (post-ORF region). I followed up these results using biochemical analysis of 2X-FLAG fusions but wanted to explore a more global approach to determine reading frame information in the post-ORF region.

The current bacterial ribosome profiling protocol uses MNase to generate ribosome footprints instead of the RNase I utilized in some eukaryotes, which does not digest effectively in *E.coli* [1,2]. Due to the sequence specificity of MNase [3], frame information of the ribosome is not generated. RNase I does not have strong sequence specificities therefore the enzyme is capable of generating a consistent footprint around the ribosome and frame information can be inferred. This frame information of ribosome footprints could yield more information about the ribosomes we find in the post-ORF region. We proposed testing the alternative nuclease cocktail of RNaseA and RNaseT1. Together they cleave the 3' end of C, G and U residues [4], and the mix is commercially available. We considered purifying a mutant version of RNaseA, T45G, which has increased activity for A residues. However, the activity of the mutant version is slower by over two orders of magnitude, and is worst over U residues [5], suggesting it would not improve the cocktail. We proceeded to test the commercially available RNaseA/T1 mix.

I first tested RNaseA/T1 for its ability to collapse polysomes, using sucrose gradients to visualize the distribution of ribosomes. Surprisingly, all initial concentrations we tested led to collapse of polysomes into monosomes, although the highest concentration seemed to result in a diminished monosome fraction (Figure A.1). I then tested a lower range of concentrations, which again showed complete collapse of polysomes. Unfortunately, due to the large amount of RNA used in this experiment we were unable to integrate the sucrose gradient curves (Figure A.2). Using less RNA and a more comprehensive range of RNaseA/T1 concentrations along with and MNase control I generated sucrose gradient curves and collected monosome fractions for further analysis (Figure A.3). To determine if RNaseA/T1 was destroying ribosome

complexes by breaking them into subunits I used integrations of curves from Figure A.1 and Figure A.3. I calculated the conservation of complete ribosome complexes from a large concentration range and we observe loss of ribosome complexes in RNaseA/T1 concentrations above 15U/7.5U (Figure A.4). The second test I performed using peak integration was to measure the total enrichment of the monosome peak after digestion since it is the only fraction utilized in ribosome profiling. Integrations of the monosome peaks in the sucrose gradient indicated that total amount of recoverable monosomes began to fall drop in RNaseA/T1 concentrations over 10U/5U (Figure A.5). All of these metrics derived from the sucrose gradient tests left me with a likely functional concentration range between 0.25U/0.125U and 10U/5U.

To further inform the best concentration of RNaseA/T1 to use in our profiling experiments I next performed a nuclease protection assay to estimate the size of the fragments generated after nuclease digestion (Figure A.6). While this assay did not provide perfect nucleotide resolution it did give me relative size distribution between the concentrations. Based on all of these results I chose to prepare a ribosome-profiling library using two RNaseA/T1 concentrations, 2U/1U and 5U/2.5U. I generated the libraries using the same protocol as standard MNase ribosome profiling protocol. After alignment to the genome I plotted the average read length of both libraries in comparison to an MNase generated library (Figure A.7). This analysis revealed that the average read length of both RNaseA/T1 libraries was approximately 20 nucleotides, much smaller than expected. The average MNase generated footprint is 26 nucleotides and while we expected a slightly smaller footprint due to the reduced sequence specificity, a 20 nucleotide footprint seemed far too small. In addition I did not see reading frame information, or phasing, above 50% for any fragment length in our dataset. Although initial experiments seemed very promising the final result suggests that RNaseA/T1 was not an effective method to generate frame information in ribosome profiling.

Materials and Methods

Strains, growth conditions and harvest. All RNaseA/T1 test samples were *E.coli* K-12 MG1655. Overnight cultures were diluted 200-fold and grown to mid-exponential phase (OD_{420}= 0.5) at 30°C in 1L baffled flasks with 200mL of MOPS-complete plus glucose medium. Cells were collected consistent with the harvest protocol for ribosome profiling [2].

Footprint generation using RNaseA/T1. Cell samples were lysed using a mixer mill, and lysate was clarified as according to previous profiling protocols [2]. These lysates were used for subsequent RNaseA/T1 and MNase digestion. The first test used 5 A260 units of RNA for each reaction and variable amounts of RNaseA/T1. The concentration of RNaseA/T1 per unit A260 was 10U/5U, 30U/15U, 40/20U and 100U/50U and a MNase control. All samples were digested for 1 hour at 25°C and RNaseA/T1 samples were quenched with 6.7uL of SuperaseIN while MNase samples were quenched with 2.4uL of 0.5M EGTA. The second test used 25 A260 units of RNA for each reaction and concentrations of RNaseA/T1 per unit A260 of 0.5U/0.25U, 1U/0.5U, 2U/1U, 3U/1.5U, 4U/2U and a MNase control. The third test used 7 A260 units of RNA for each reaction and concentrations of RNaseA/T1 per unit A260 of 0.25U/0.125U, 5U/2.5U, 10U/5U, 15U/7.5U and a MNase control. Samples were digested and quenched under the same conditions. In accordance to standard ribosome profiling protocols all samples were placed on a 7%-47% short sucrose gradient and centrifuged at 35000rpm for 2.5hours at 4°C. All samples were then fractionated and UV absorbance peaks were recorded. The monosome fractions for samples between 0.25/0.125U and 10U/5U of RNaseA/T1 per unit A260 were collected for RNA extraction prior to the nuclease protection assay.

RNA extraction from monosome fraction. RNA was extracted from the monosome fraction of the sucrose gradients using a standard hot phenol chloroform extraction used in previous publications [2].

Nuclease protection assay. Using the same gapA construct that was used for MNase protocol optimization [2] we generated a full length gapA PCR product. Using the MaxiScript T7 Transcription Kit and $^{\alpha 32}$P-UTP we created a radioactive probe of gapA for our nuclease protection assay. Excess radioactive nucleotides were removed using an Illustra ProbeQuant G-50 Micro-Column. We used this radioactive gapA probe along with the *mir*Vana™ miRNA Detection Kit to perform the nuclease protection assay. Using the RNA extracted from the monosome fraction of a range of RNaseA/T1 concentrations. We used 5ug of RNA from each sample to hybridize to the probe for 3minutes at 95°C then overnight at 42°C. Digestion of hybridized gapA probe and sample RNA was performed using 150uL of RNase digestion buffer and 1.5uL of RNaseA/T1 and incubated at 37°C for 30 minutes. RNases were inactivated using 225uL RNase inactivation buffer and 225uL 100% ethanol into each sample then incubating at -20°C for 30 minutes. The remaining sample was pelleted by centrifuge at max speed at 4°C for 15 minutes then the supernatant was aspirated off. Samples were prepared for gel using 5uL gel loading buffer II. A 10% TBE-Urea gel was pre-run for 1 hour at 200 volts, immediately afterwards samples were loaded and run at 200 volts for 48 minutes. The gel was imaged using a phosphor screen and developed overnight.

Ribosome profiling, library preparation and analysis of RNaseA/T1 samples. After digestion of cellular lysate with RNaseA/T1 we utilized the same protocol for library preparation as outlined in Chapter 1 and as previous described [2,6]**.** Libraries were sequenced on an Illuminia HiSeq 2000 or HiSeq4000. All sequencing reads were trimmed of the ligated linker, and aligned to a genome file containing *E.coli* rRNA and tRNA. All unaligned reads were then aligned to the *E.coli* genome as previously described [2,6]. The aligned reads were then analyzed for read length.

Figure A.1. Polysome collapse using wide range of RNaseA/T1 concentrations.

The first test of RNaseA/T1 as an alternative nuclease for ribosome profiling was to determine if it could collapse polysomes as measured by sucrose gradients. Using a wide range of RNaseA/T1 concentrations to digest standardly prepared lysate we saw efficient collapse of polysomes into the monosome peak of the sucrose gradient for all concentrations tested. The concentration of enzyme used is shown as units of RNaseA/T1 per unit A260.

Figure A.1

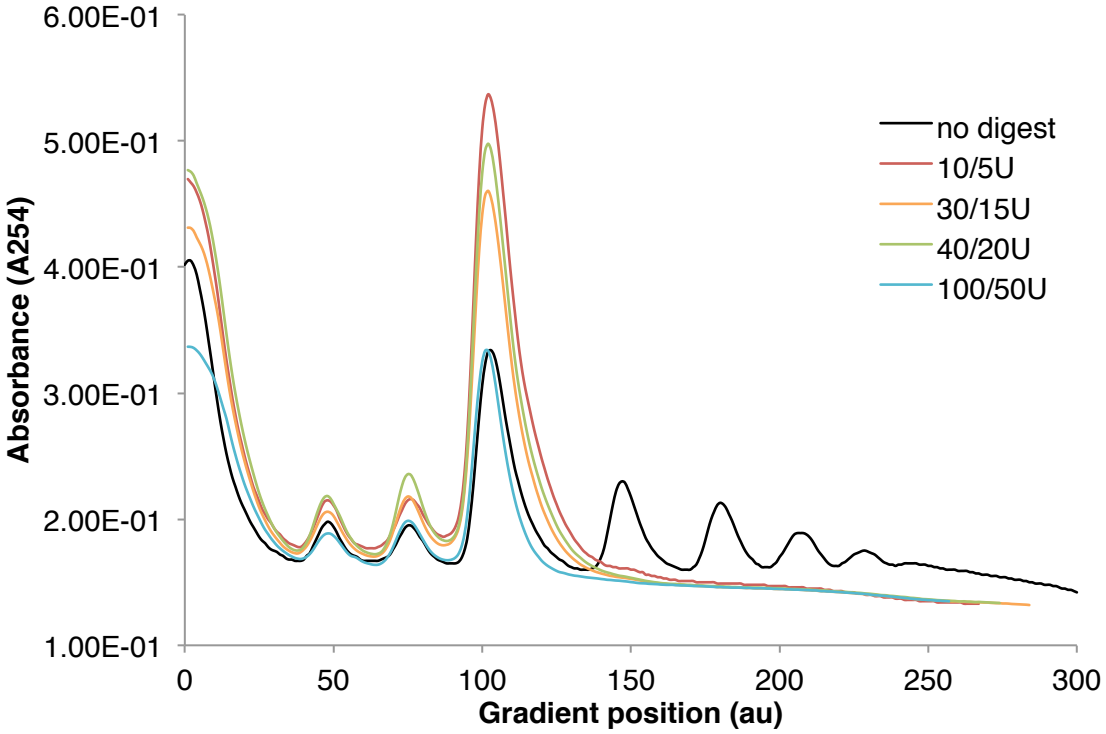

6.00E-01
5.00E-01
4.00E-01
3.00E-01
2.00E-01
1.00E-01
Absorbance (A254)
0
50
100
150
200
250
300
Gradient position (au)
no digest
10/5U
30/15U
40/20U
100/50U

Figure A.2. Polysome collapse using a lower range of RNaseA/T1 concentrations.

After we established that polysomes could be collapsed using RNaseA/T1, we tested a lower range of concentrations to determine the effective range of the collapse. Using a lower range of RNaseA/T1 concentrations to digest standardly prepared lysate we again saw efficient collapse of polysomes into the monosome peak of the sucrose gradient for all concentrations tested. The concentration of enzyme used is shown as units of RNaseA/T1 per unit A260. Due to the amount of RNA loaded onto the sucrose gradients the UV monitor maxed out its quantification and while we were able to gather monosome fractions we could not perform ribosome conversation analysis on these curves. The monosome peaks of these samples were collected for RNA extraction and further analysis.

Figure A.2

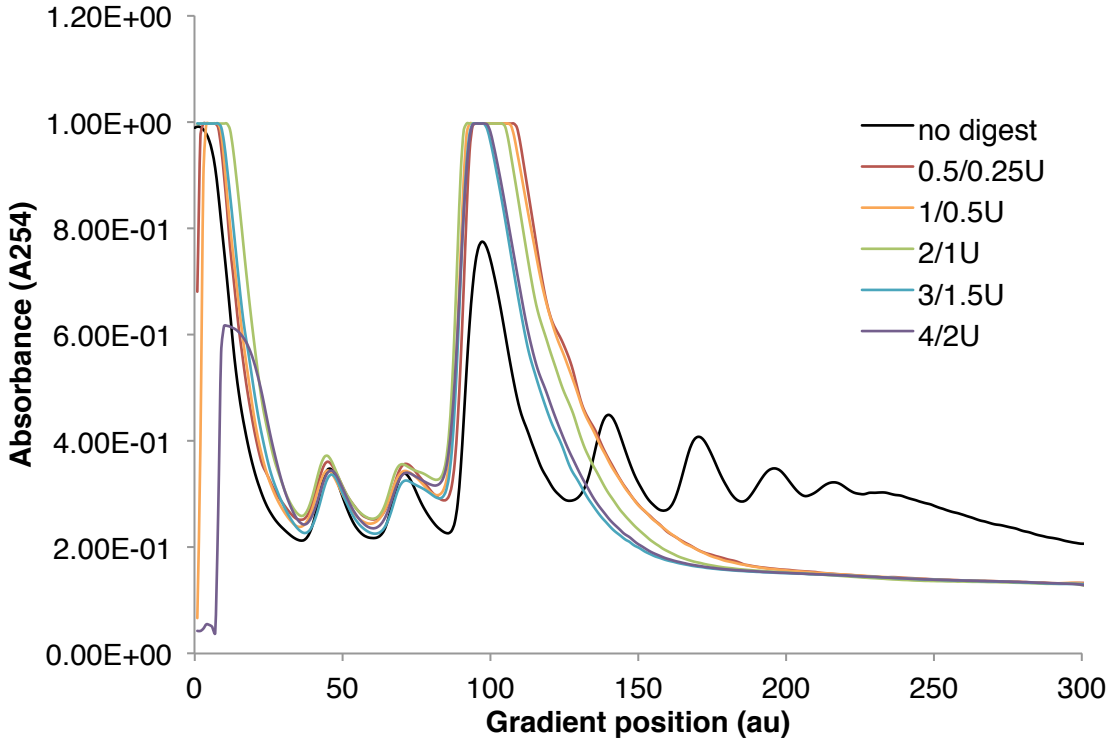

1.20E+00
1.00E+00
8.00E-01
6.00E-01
4.00E-01
2.00E-01
0.00E+00
Absorbance (A254)
0
50
100
150
200
250
300
Gradient position (au)
no digest
0.5/0.25U
1/0.5U
2/1U
3/1.5U
4/2U

Figure A.3. Polysome collapse using RNaseA/T1.

We repeated some previous conditions with the addition of an MNase control, the current ribosome profiling nuclease, to compare efficiency of ribosome collapse. Using a broad range of effective RNaseA/T1 concentrations to digest standardly prepared lysate we again saw efficient collapse of polysomes into the monosome peak of the sucrose gradient for all concentrations tested. Interestingly the lowest concentration looked very similar to the MNase sample which in this experiment in addition to the large monosome peak exhibited a “shoulder” peak of disomes. The concentration of enzyme used is shown as units of RNaseA/T1 per unit A260. The monosome peaks of these samples were collected for RNA extraction and further analysis.

Figure A.3

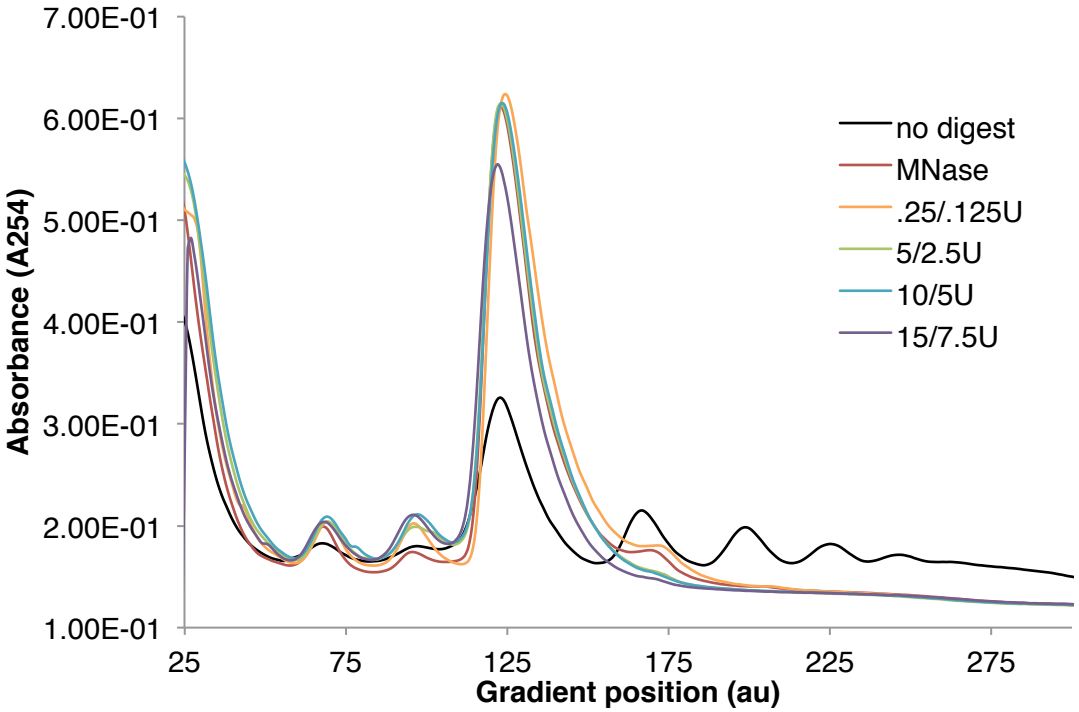

7.00E-01
6.00E-01
5.00E-01
4.00E-01
3.00E-01
2.00E-01
1.00E-01
Absorbance (A254)
25
75
125
175
225
275
Gradient position (au)
no digest
MNase
.25/.125U
5/2.5U
10/5U
15/7.5U

Figure A.4. Ribosome complex conservation across wide range of RNaseA/T1 concentrations.

Using integrations of the sucrose gradient data in Fig A.1 and A.3 the level of ribosome complex conservation was calculated. The total number of ribosome complexes was found by calculating the area under the curve from monosomes to the highest order polysomes in the non-digested control sample. This value was compared to the same integration in the digested samples. If polysomes are only collapsed to monosomes the conservation should be near 100 percent. If the enzyme breaks a ribosome complex into its subunits the conservation number will drop below 100 percent, as seen in the highest concentrations tested. The optimal concentration for ribosome profiling would only collapse polysomes into monosomes with no breaking apart of the ribosomal complex into subunits.

Figure A.4

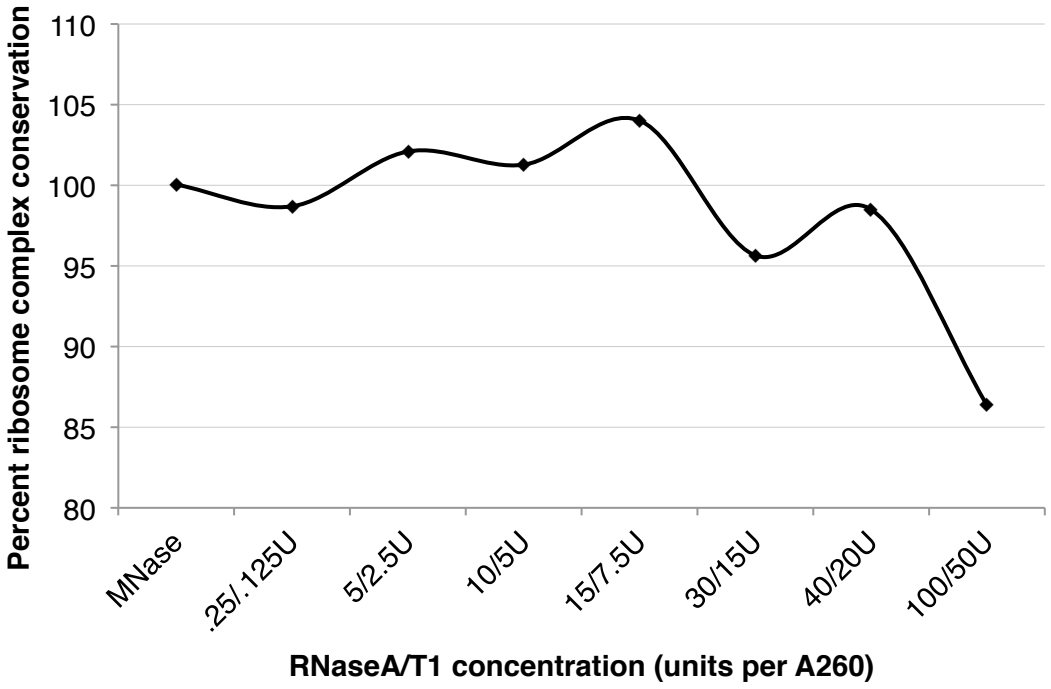

Percent ribosome complex conservation
110
105
100
95
90
85
80
MNase
.25/.125U
5/2.5U
10/5U
15/7.5U
30/15U
40/20U
100/50U
RNaseA/T1 concentration (units per A260)

Figure A.5. Monosome enrichment across wide range of RNaseA/T1 concentrations.

In addition to choosing a nuclease concentration that did not break apart a ribosome complex we wanted to optimize the monosome fraction we generated. The ideal concentration would maximize the monosome fraction since this is the fraction that is harvested and used for subsequent ribosome profiling. We integrated the area under the monosome peak and plotted it relative to the undigested control. We observed that while a larger range of concentrations conserved total ribosome complexes and smaller range up to 10U/5U RNaseA/T1 had peak monosome enrichment.

Figure A.5

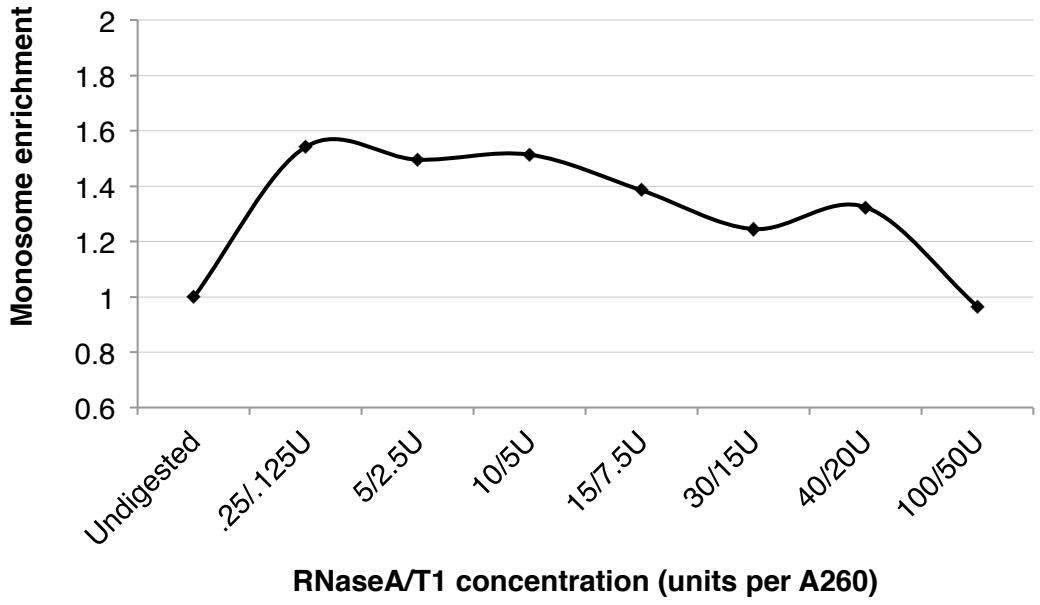

Monosome enrichment
2
1.8
1.6
1.4
1.2
1
0.8
0.6
Undigested
.25/.125U
5/2.5U
10/5U
15/7.5U
30/15U
40/20U
100/50U
RNaseA/T1 concentration (units per A260)

Figure A.6. Size distribution of ribosome footprints created by RNaseA/T1 digestion determined by nuclease protection assay.

Using a nuclease protection assay of a radioactive *gapA* probe we were able to determine the relative size distribution of ribosome footprints generated with varying RNaseA/T1 concentrations. While determining absolute size is difficult we were able to determine that the footprints generated were relatively close to the expected range and by comparing the relative size of the MNase sample we decided to continue ribosome profiling with two RNaseA/T1 concentrations, 2U/1U and 5U/2.5U.

Figure A.6

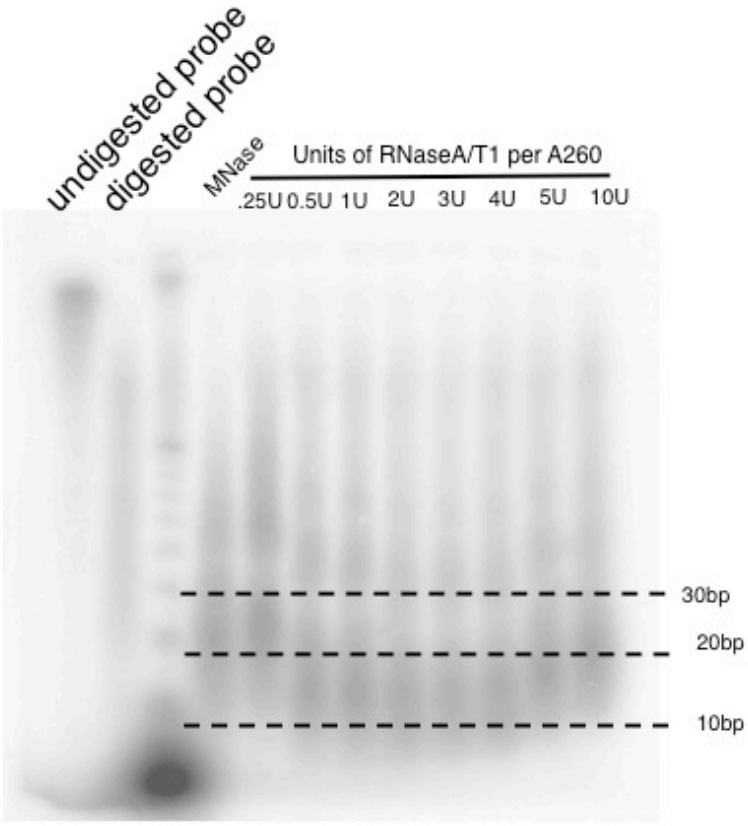
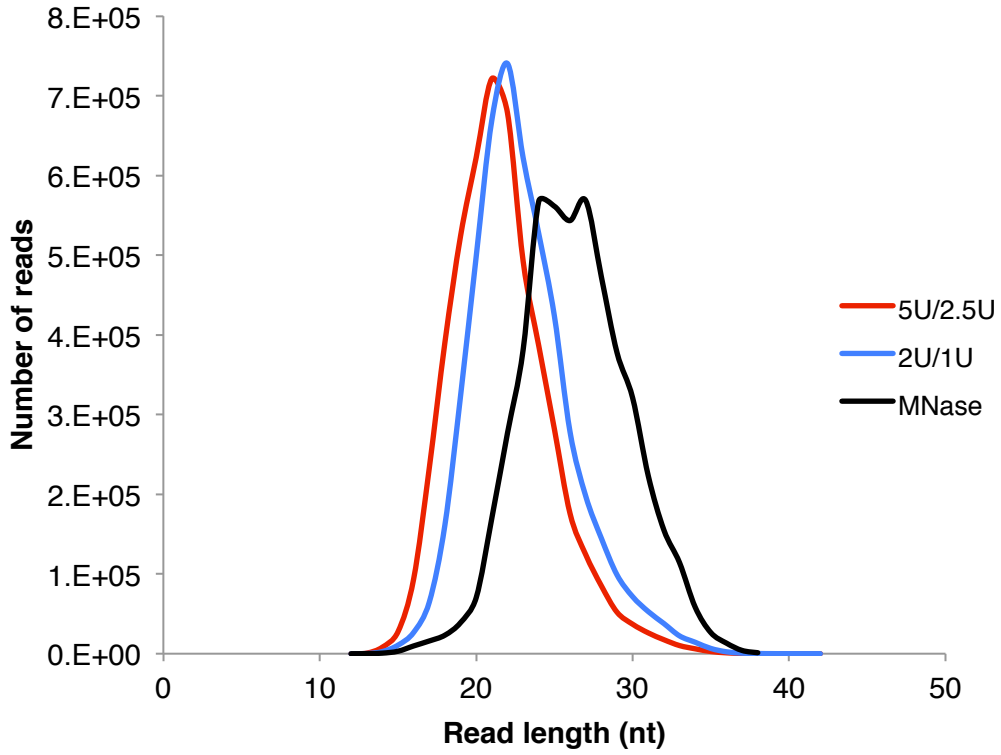

undigested probe
digested probe
MNase
Units of RNaseA/T1 per A260
.25U 0.5U 1U 2U 3U 4U 5U 10U
30bp
20bp
10bp

Figure A.7. Read length of ribosome profiling library.

After sequencing the ribosome profiling libraries generated using two different concentrations of RNaseA/T1 we discovered they had a much smaller ribosome footprint size than the MNase control library with an average read length of 21 nucleotides.

Figure A.7


8.E+05
7.E+05
6.E+05
5.E+05
4.E+05
3.E+05
2.E+05
1.E+05
0.E+00
Number of reads
0
10
20
30
40
50
Read length (nt)
5U/2.5U
2U/1U
MNase

References

1. Ingolia NT, Ghaemmaghami S, Newman JRS, Weissman JS. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science. 2009;324: 218–23. doi:10.1126/science.1168978

2. Oh E, Becker AH, Sandikci A, Huber D, Chaba R, Gloge F, et al. Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. Cell. 2011;147: 1295–308. doi:10.1016/j.cell.2011.10.044

3. Dingwall C, Lomonossoff GP, Laskey RA. High sequence specificity of micrococcal nuclease. Nucleic Acids Res. 1981;9: 5287–5296.

4. delCardayré SB, Ribó M, Yokel EM, Quirk DJ, Rutter WJ, Raines RT. Engineering ribonuclease A: production, purification and characterization of wild-type enzyme and mutants at Gln11. Protein Eng. 1995;8: 261–73.

5. Kelemen BR, Schultz LW, Sweeney RY, Raines RT. Excavating an active site: the nucleobase specificity of ribonuclease A. Biochemistry. 2000;39: 14487–94.

6. Ingolia NT, Ghaemmaghami S, Newman JRS, Weissman JS. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science. 2009;324: 218–23. doi:10.1126/science.1168978

Appendix B

Investigation of temperature dependent recoding at the *cspG* locus

Introduction

In previous experiments in the Carol Gross Lab, ribosome profiling was performed in cells that had been grown at 30°C instead of 37°C. One of the observed phenotypes was increased ribosome occupancy after the annotated stop codon of the open reading frame, or post-ORF region. This was similar to the occupancy seen at 37°C in Δ*prfC* cells as outlined in Chapter 1. One of the interesting cases of this extended occupancy occurred at the *cspG* locus, one member of the family of cold-shock proteins that is utilized during cold shock [1–3]. We were especially interested in this phenotype at lower temperatures given the involvement of cold shock proteins. If these ribosomes were in fact the result of a recoding event, a +1 frameshift to be exact, the resulting protein would be homologous to a csp in *Salmonella typhimurium*. This led us to hypothesize about increased functionality of CspG via programmed frameshift at cold temperatures.

My initial experiment was designed to detect a recoding event occurring over the stop codon in any frame. I created a chromosomally integrated reporter construct, which placed *lacZ* downstream *cspG* inside its hypothesized extension region in every possible frame as well as a full translational fusion of *cspG-lacZ* (Figure B.1). I then determined the frame resulting in production of the *cspG-lacZ,* which would signal the frame of the extension. I harvested proteins at 37°C and 10°C, size separated the proteins on denaturing SDS gels, and visualized the fusion with a western blot against LacZ. Unfortunately, at 37°C, all proteins were approximately the same size as LacZ*,* whereas a mixture of LacZ and CspG-LacZ products were found at 10°C (Figure B.2). This signified that either all LacZ products were a result of CspG-LacZ products and the CspG was quickly cleaved due to its small comparable size, or that LacZ could be produced from the construct independently from CspG. I believe the latter case to be unlikely due to the lack of LacZ initiation codon and the fact that we see variable levels of LacZ between constructs which is especially low in the 0 frame. Due to the likely degradation of CspG from the fusion protein, this construct was not useful for obtaining frame

information by visualizing the extended product. However, we decided it was still productive to measure β-galactosidase activity for all reporters in different temperature conditions.

I assayed all the *cspG-lacZ* reporters at several timepoints after cold shock. I found that the -1 frame had overall the highest amount of readthrough relative to the other reporters and that the time point 2 hours after cold shock had the highest amount of readthrough for all reporters (Figure B.3). The +1 frame, which results in the longest possible extension that is also homologous to the *Salmonella* csp, did have nearly 10% readthrough at 2 hours, but was less than the -1 frame rate of 35% (Figure B.3). There is a chance that a programmed +1 frameshift does occur at the *cspG* locus but cold shock response triggers widespread -1 frameshifting. We decided to test the same reporters in a wider range of temperatures to see if -1 frameshifting was increased as temperature decreased. I assayed the same reporters at 37°C, 30°C, 22°C, and 10°C at mid-exponential growth and determined that the -1 frame exhibited the highest amount of readthrough at all temperatures with increasing severity as temperature dropped (Figure B.4). Since these conclusions are solely the result of a single locus, more experiments need to be performed to determine if a -1 frameshift is the most common shift genome wide regardless of temperature.

In further attempts to visualize the extended *cspG* product and identify any additional functionality of the extended version I created new constructs using a 5' His-tag (Figure B.5). One construct contained only the annotated open reading frame and removed the entire 3'UTR; I refer to this as *cspG*-short. The other contains the entirety of the 3'UTR; I refer to this construct as *cspG*-extendable. I made both constructs with or without the 5'-His-tag to check for loss of functionality due to tagging (Figure B.5). The initial test of functionality was to place all the plasmid constructs in the quadruple deletion strain, Δ*cspABGE*, which looses viability at cold temperatures [2]. Complementation of the *cspG* should reduce the cold sensitivity. Due to extremely slow growth at 10°C instead of measuring cell growth, I decided to assay translation initiation using ^{35}S-methionine incorporation after shift to 10°C. The first experiment sampled

time points immediately after temperature shift to 10°C until 6 hours after cold shock. Using the triple deletion strain, Δ*cspABE*, as a benchmark I established that all plasmids except the 5'-His-tag of *cspG*-short restored translation initiation (Figure B.6.) There did not appear to be any additional functionality between short or extended versions, and in combination with later experiment we established the His tag did not impact functionality. For the second experiment we extended the window in which we assayed translation initiation, because 6 hours is generally how long it takes cells to acclimate to growth at 10°C. This extended study again did not reveal any additional functionality between short or extended versions (Figure B.7).

In a final attempt to reveal any difference in functionality between the short and extended versions of *cspG* I plated serial dilutions of Δ*cspABGE* strains containing each plasmid construct and incubated the plates at a wide range of temperatures. Cell viability was equivalent between all strains at 37°C but Δ*cspABGE* viability began to slightly reduce at 30°C (Figure B.8A and B). All *cspG* plasmids were able to equally complement this phenotype at 30°C, 25°C, 20°C, and 15°C (Figure B.8).

Using the 5'His-tagged *cspG* I hoped to visualize an extended CspG product. I harvested protein from cells after 20 hours at 10°C and purified CspG using its His-tag. A western blot revealed CspG at the expected size but also a non-specific product at the same size as our expected extended CspG protein. After failing to identify any functionality difference between versions of *cspG* and not being able to visualize an extended CspG protein I decided to cease experimentation. Due to the redundancy of *csp*s [2] I wonder if functionality might come with more *csp* extensions, or ratios of short versus extended products. Further investigation into combinations of *csp* extensions would be the logical next step given the known *csp* redundances.

Materials and Methods

Strains and growth conditions. All experiments were performed in the *E.coli* K-12 MG1655 background. We also utilized quadruple deletion, Δ*cspABGE*, which was constructed using the KEIO collection and Kan^R was flipped out before transduction of a new deletion locus, the final Kan^R remains at the *cspE* locus as we were unable to transform cells after deletion of *cspE* [4].

Construction of chromosomal cspG-lacZ fusion reporters. We started with the strain NC397 which has a kan^R-Cat^R-*sacB* cassette between *lacI* and *lacZ*. Using PCR amplification we generated variable forms of the *cspG* locus under control of a tetracycline promoter, including a small portion of the hypothesized extension region. We utilized spacer nucleotides to align the downstream *lacZ* in frame of a +1 frameshift using A, -1 frameshift using AA and no spacers for the 0 frame. We also constructed a full protein fusion *cspG-lacZ* in the same manner, which placed *lacZ* in frame of *cspG* and removed the original *cspG* stop codon. After amplification using primers with 40nt long homology arms we recombined the modified *cspG-lacZ* reporters into the genome at the native *lacZ* locus using λ-Red functions [5]. We transduced our these reporters and fusion construct into MG1655 along with the tet repressor from DH5αZ1.

Visualization of cspG-lacz fusions. Cultures containing the integrated reporter plasmids were grown in either MOPS complete-glucose overnight at 37°C and dilutions were performed to OD_{420} of 0.01 prior to the experiment then incubated at 37°C. Once cultures had reached OD_{420} = 0.1, expression of *cspG* was induced using 0.1uM anhydrous tetracycline (aTc). For the cold shocked sample, cells were moved from 37°C to 10°C at OD_{420} of 0.1 and induced with aTc. Proteins were harvested 50minutes after induction at 37°C and 26 hours after induction and shift to 10°C, 1.5mL of culture was combined with 200uL of 100% TCA to precipitate protein. The culture with TCA was inverted and incubated on ice for 30minutes and stored overnight at

4°C to complete precipitation. Samples were then centrifuged at 4°C at max speed for 15minutes, supernatant was aspirated, pellet was washed with 1mL cold acetone, again centrifuged for 15minutes. The aspirated cell pellet was dried at 30°C and resuspended in 50uL of TCA resuspension buffer (10mM Tris base pH 11, 3% SDS) with 7% β-mercaptoethanol and 16uL of 4X LDS sample loading dye. Samples were loaded onto a 7.5% Tris-HCl Biorad ReadyGel and run for 2 hours at 100 volts. After a wet-transfer was completed onto nitrocellulose, the membrane was blocked for 1 hour in Licor Odyssey Blocking Buffer, washed three times with 1X TBS, then incubated with primary antibody, α-LacZ rabbit (ab616), at 1:250 overnight at 4°C. After complete washing with 1X TBS, the membrane was then incubated with secondary antibodies, Licor α-mouse and α-rabbit, at 1:10,000 for 20 minutes. The membrane was washed again with 1X TBS before imaging on the Licor.

β-galactosidase assay of readthrough at the *cspG locus*. Each reporter plasmid in the WT background was cultured in either MOPS complete-glucose overnight at 37°C and dilutions were performed to OD_{420} of 0.01 prior to the experiment. Cultures were grown at either 37°C, 30°C or 22°C. Once cultures had reached OD_{420} = 0.1, expression of *cspG* was induced using 0.1uM anhydrous tetracycline (aTc). For the cold shocked sample, cells were moved from 37°C to 10°C at OD_{420} of 0.1 and induced with aTc. For each time point exponential growth samples of 500uL and 1mL were taken simultaneously to measure β-galactosidase (β-gal) activity and cell density at OD_{420}. β-gal assay samples were immediately added to a tube containing 500uL Z-buffer, 1.25uL beta-mercaptoethanol, 30uL 0.1% SDS and 40uL chloroform then vortexed before being placed on ice. Once collection was completed all β-gal sample tubes were incubated at 28°C prior to induction with 200uL 4mg/mL ONPG. Development was stopped with 500uL of 1M sodium bicarbonate and samples were centrifuged for 5 minutes at max speed to remove cellular debris. Samples were then transferred into a 96-well plate to measure OD_{420} in a Varioskan plate reader and βgal/mL was calculated for each sample. To approximate the

amount of readthrough occurring at the *cspG* locus we divided the activity of each frame reporter by the fusion *cspG-lacZ* reporter.

Construction of plasmid based *cspG* constructs. Using the backbone plasmid of pNG171, a low copy plasmid, we inserted *cspG* and its native promoter sequence using restriction digestion cloning and ligation at the Cla1 and Sac1 sites. A plasmid was made for the annotated open reading frame version, or short version, and with the mRNA of the hypothetical extended region, annotated here as the extendable version. The same plasmids were made with a 5'-HIS-tag so we could visualize extended proteins. All plasmids were placed in wild-type MG1655 and quadruple deletion Δ*cspABGE*.

^{35}S- methionine incorporation assay. Cultures were grown at 37°C in MOPS –methionine plus glucose medium with 25ug/mL chloramphenicol to maintain the pNG171 *cspG* plasmids. Overnight cultures were diluted to 25mL cultures of OD_{420} of 0.005 and grown at 37°C until OD_{420} of 0.8. To initiate the cold shock response these mid-exponential cultures were split into 12.5mL of 10°C MOPS–methionine plus glucose medium for each 12.5mL of culture and growth was continued in a shaking water bath at 10°C. Samples were taken at various time points after shift to 10°C. For each sample time point, 900mL of cultures were placed into a falcon tube containing 1uL of EasyTag™ L-^{35}S-methionine, 5uL of 300uM methionine and 24uL of MOPS–methionine medium. Each falcon tube was then immediately incubated at 10°C in the shaking water bath for 4 minutes before stopping the reaction with 100uL of 50% TCA, briefly vortexed then placed on ice. After all samples have been taken we assembled a vacuum filtration apparatus using 25mm glass fibre filters with a 1.2um pore size (Millipore APFC02500) and prepared a scintillation vial with 2mL of scintillation fluid. The filter was first wet with 750uL of 10% TCA, after a brief vortex of the sample 100uL was pipetted into the center of the filter paper. The filter was washed three times with 750uL of 10% TCA, allowing the liquid to be cleared by

the vacuum between each application, then three times with 750uL of 95% ethanol. The washed sample filter paper was tweezed into the scintillation vial and completely submerged in the fluid. After 15 minutes from the last sample being placed in the scintillation fluid and counts were measured using the ^{35}S channel on a Beckman Coulter LS-6500 scintillation counter.

Cold sensitivity of quadruple *csp* deletion strain and complementation with *cspG*. The variable length pNG171 *cspG* plasmids were transformed into the cold sensitive ΔcspABGE strain. Cultures were grown at 37°C in LB medium with 25ug/mL chloramphenicol to maintain the pNG171 *cspG* plasmids. Overnight cultures were diluted 1:50 and grown at 37°C for 1.5 hours, then diluted to 10^{-6} using serial 1:10 dilutions. The dilutions were plated, 2uL of each, onto LB agar chloramphenicol plates, and incubated at 15°C, 20°C, 25°C, 30°C, 37°C until colonies were visible (141hrs, 44hrs, 28hrs, 17hrs and 17hrs respectively).

Purification of HIS-CspG and visualization. Cultures were grown at 37°C in LB medium with 25ug/mL chloramphenicol to maintain the pNG171 *cspG* plasmids. Overnight cultures were diluted to 25mL cultures of OD_{600} of 0.05 and grown at 37°C until OD_{600} of 0.8. To initiate the cold shock response these mid-exponential cultures were split into 250mL of 10°C LB medium for each 250mL of culture and growth was continued in a shaking water bath at 10°C. After 20 hours at 10°C cells were pelleted at 4,500rpm for 10minutes at 4°C. The cell pellet was resuspended in 40mL of guanidine buffer (6M guanidine, 50mM NaPi pH 7.5, 380mM NaCl, 10mM imidazole) using a douncer. Cells were then lysed using a microfluidizer and centrifuged at max speed for 20 minutes at 4°C to remove cellular debris. In a fresh falcon tube 500mL of Ni-NTA agarose was added to the lysed supernatant and allowed to nutate at 4°C for 30 minutes. The product was then loaded into 1mL Biorad column and allowed to flow until agarose was settled. The column was then washed with 20mL of guanidine buffer, followed by 40mL of wash buffer (50mM NaPi pH 7.5, 380mM NaCl, 10mM imidazole). To elute 500uL of guanidine

buffer with 500mM imidazole was added to the column and used to resuspend the agarose, the column was parafilmed and allowed to nutate for 30 minutes at 4°C. After the first round of elution, another 500uL elution was performed and collected in a separate tube. A fraction of the sample, 20uL, was combined with 4X LDS sample buffer and 10% β-mercaptoethanol incubated at 95°C for 5 minutes then centrifuged at max speed for 5 minutes. The sample was loaded onto a 16% tris-tricine gel with a 4% sample gel on top and run at 200V until the sample dye reached the bottom of the gel. After a wet-transfer was completed onto nitrocellulose, the membrane was blocked for 1 hour in Licor Odyssey Blocking Buffer, washed three times with 1X TBS, then incubated with primary antibody, α-HIS rabbit (sc-803), at 1:250 overnight at 4°C. After complete washing with 1X TBS, the membrane was then incubated with secondary antibodies, Licor α-mouse and α-rabbit, at 1:10,000 for 20 minutes. The membrane was washed again with 1X TBS before imaging on the Licor.

Figure B.1. Chromosomally integrated *cspG* reporter and fusion constructs.

(A) The native locus of *cspG* is shown with its hypothesized extension region, with the first stop codon for each frame in the extension region annotated. (B) Readthrough reporters for each possible recoding event which could result in a final -1, +1 or 0 frame extension were constructed with the help of A nucleotide spacers. (C) To control for induction across experiments and growth conditions we constructed a fusion reporter locus which removes the native *cspG* stop codon and fuses in-frame *lacZ* immediately downstream. Estimates of readthrough in each frame were calculated by dividing β-galactosidase activity of each frame reporter by the activity of the fusion reporter construct with the same time and growth conditions.

Figure B.1

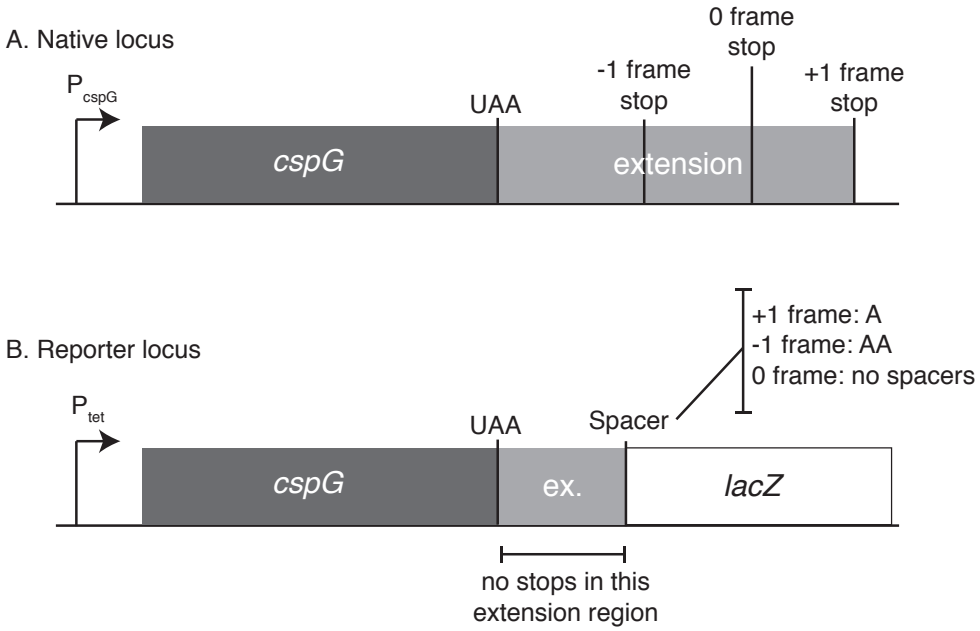

A. Native locus
P cspG
UAA
-1 frame stop
0 frame stop
+1 frame stop
cspG
extension
B. Reporter locus
+1 frame: A
-1 frame: AA
0 frame: no spacers
P tet
UAA
Spacer
cspG
ex.
lacZ
no stops in this extension region
C. Fusion Reporter locus
P tet
cspG
lacZ

Figure B.2. Visualization of CspG-LacZ fusions during normal and cold shock growth conditions.

(A) Proteins harvested at 37°C for each frame reporter, CspG-LacZ fusion and LacZ, induced from WT using IPTG. Variable expression was seen across all reporters but interestingly there was no size differential from native LacZ. (B) Proteins from all reporter strains were harvested 26hours after shift from 37°C to 10°C. In this experiment we were able to visualize products larger than native LacZ in addition to the native LacZ size.

Figure B.2

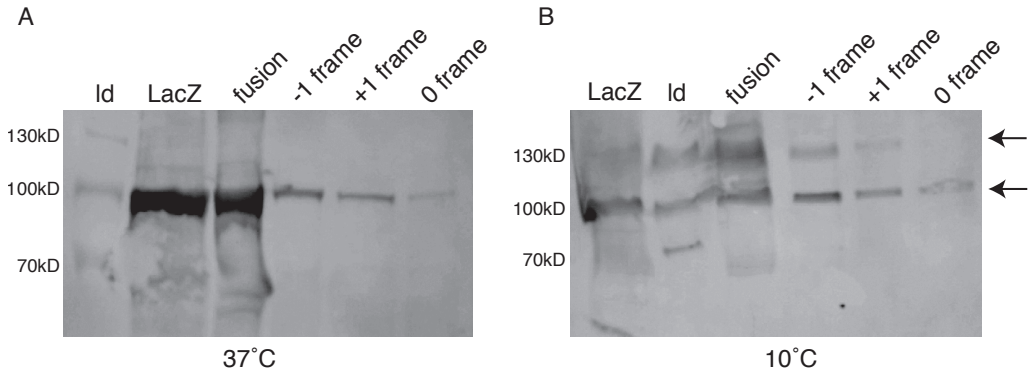

A
Id
LacZ
fusion
-1 frame
+1 frame
0 frame
130kD
100kD
70kD
37˚C
B
LacZ
Id
fusion
-1 frame
+1 frame
0 frame
130kD
100kD
70kD
10˚C

Figure B.3. Readthrough and recovery at the *cspG* locus after cold shock.

All *cspG* reporter strains constructed in Figure B.1 were assayed for β-galactosidase activity after shift from growth at 37°C to 10°C. Samples were taken immediately after the temperature shift, 0hrs, then at 2, 4, and 26hrs after the shift. The readthrough was calculated by dividing β-galactosidase activity of each frame reporter by the activity of the fusion reporter construct at the same time point.

Figure B.3

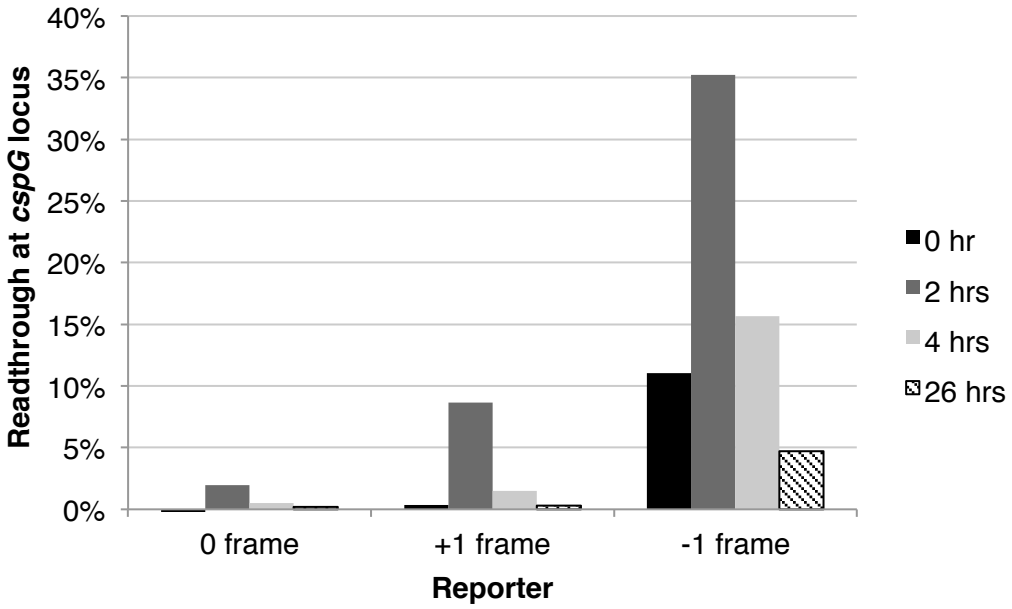

40%
35%
30%
25%
20%
15%
10%
5%
0%
Readthrough at cspG locus
0 frame
+1 frame
-1 frame
Reporter
0 hr
2 hrs
4 hrs
26 hrs

Figure B.4. Readthrough and recovery at the *cspG* locus after cold shock.

All *cspG* reporter strains constructed in Figure B.1 were assayed for β-galactosidase activity after growth at 37°C, 30°C, 22°C and shift from 37°C to 10°C. Reporter expression was induced at early exponential phase and samples were assayed at mid-exponential phase; 50minutes, 120minutes, 210minutes, and 26 hours respectively. The readthrough was calculated by dividing β-galactosidase activity of each frame reporter by the activity of the fusion reporter construct at the same time point and growth temperature.

Figure B.4

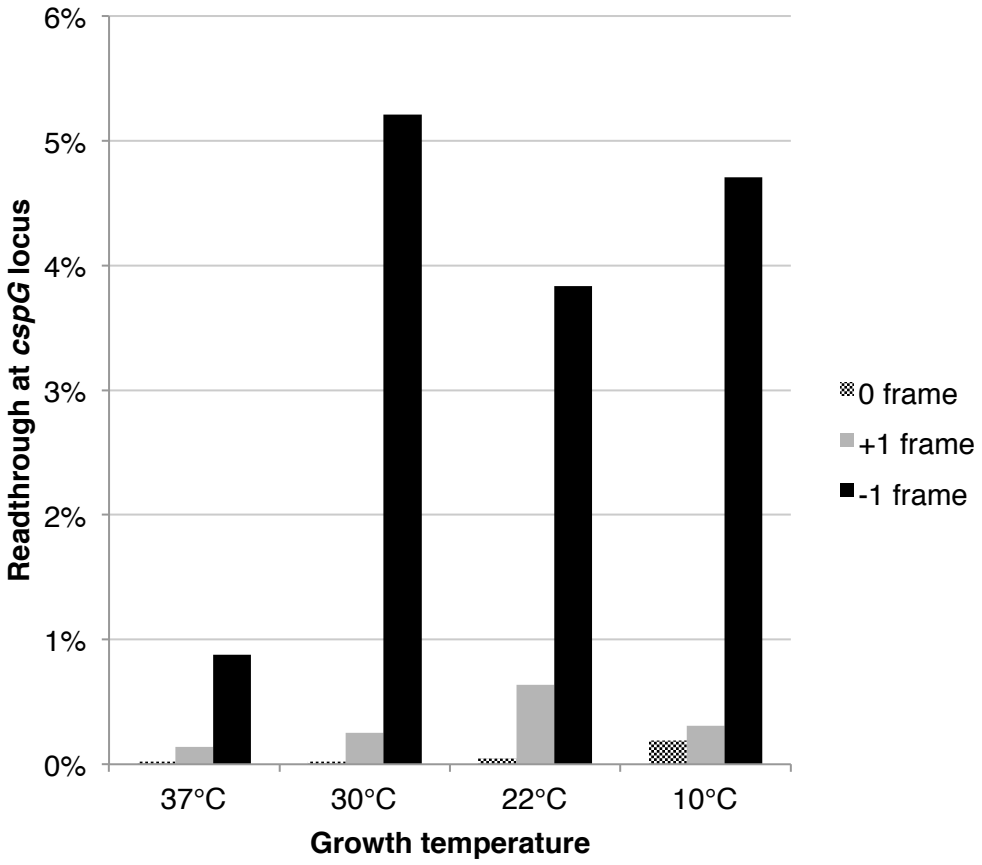

6%
5%
4%
3%
2%
1%
0%
Readthrough at cspG locus
37°C
30°C
22°C
10°C
Growth temperature
0 frame
+1 frame
-1 frame

Figure B.5. Plasmid based *cspG* short and extendable constructs.

Due to the apparent cleavage of CspG from the CspG-LacZ fusion reporters we constructed low copy plasmid based reporters for a short, or non-extended, version of *cspG* from which we removed the entire 3'UTR and replaced it with multiple stop codons (A) and an extendable version which contained the entire 3'UTR (B). We made both of these constructs with and without a 5'-HIS-tag.

Figure B.5

A. Plasmid based short *cspG*

B. Plasmid based extendable *cspG*

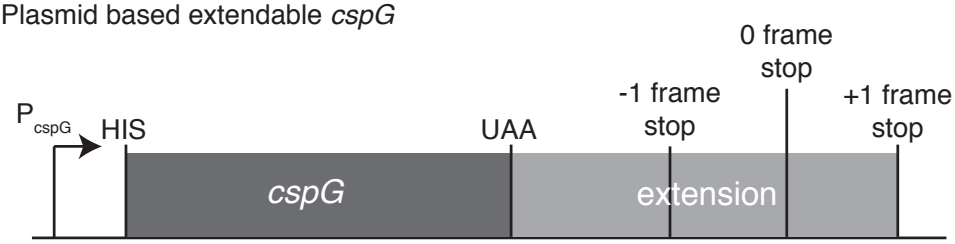


Figure B.6. Complementation of Δ*cspABGE* cold sensitivity with *cspG* plasmids.

As a metric of cold viability we measured translation through ^{35}S-methionine incorporation after shift from 37°C to 10°C. Over a timespan to 400 minutes after temperature shift, all *cspG* constructs except His-tagged *cspG*-short appeared to rescue the Δ*cspABGE* phenotype to Δ*cspABE* levels.

Figure B.6

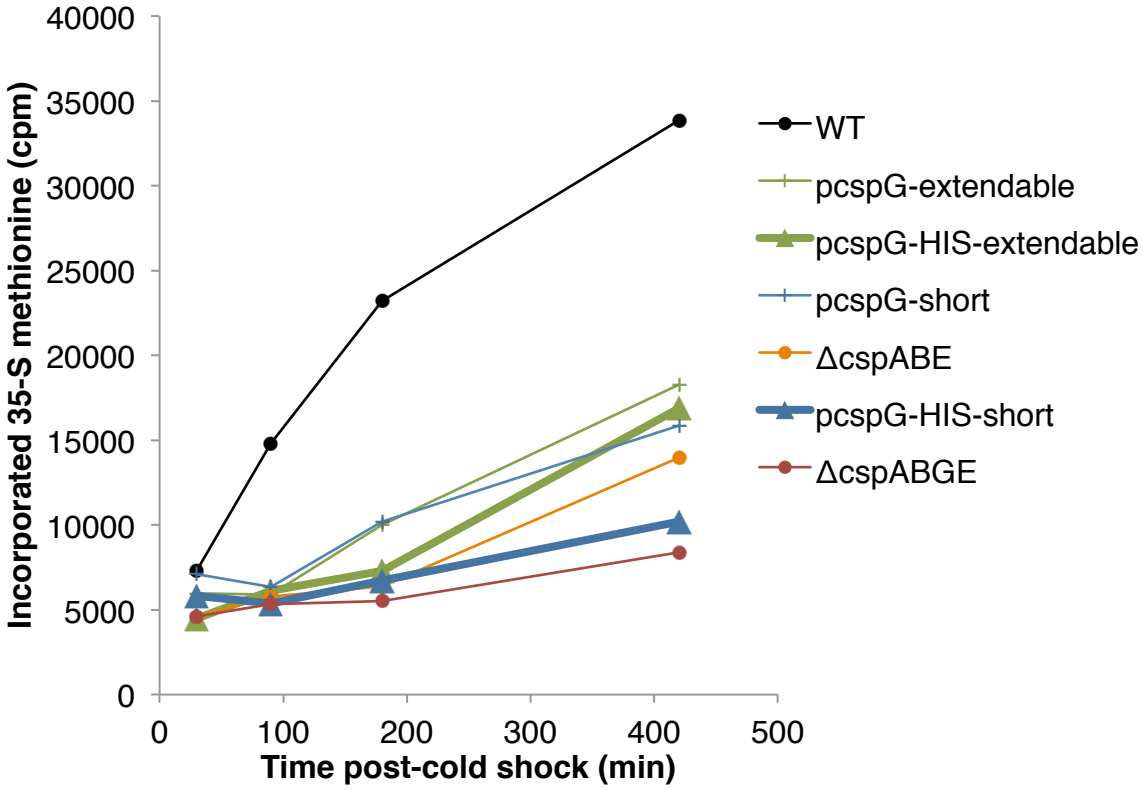

40000
35000
30000
25000
20000
15000
10000
5000
0
Incorporated 35-S methionine (cpm)
0
100
200
300
400
500
Time post-cold shock (min)
WT
pcspG-extendable
pcspG-HIS-extendable
pcspG-short
ΔcspABE
pcspG-HIS-short
ΔcspABGE

Figure B.7. Complementation of Δ*cspABGE* cold sensitivity with *cspG* plasmids.

As a follow up to our prior ^{35}S-methionine incorporation after shift from 37°C to 10°C, we performed the same experiment over an extended time period, to 26 hours after shift. All *cspG* plasmids rescued the quadruple deletion phenotype including His-tagged *cspG*-short, which did not exhibit recovering in the shorter experiment.

Figure B.7

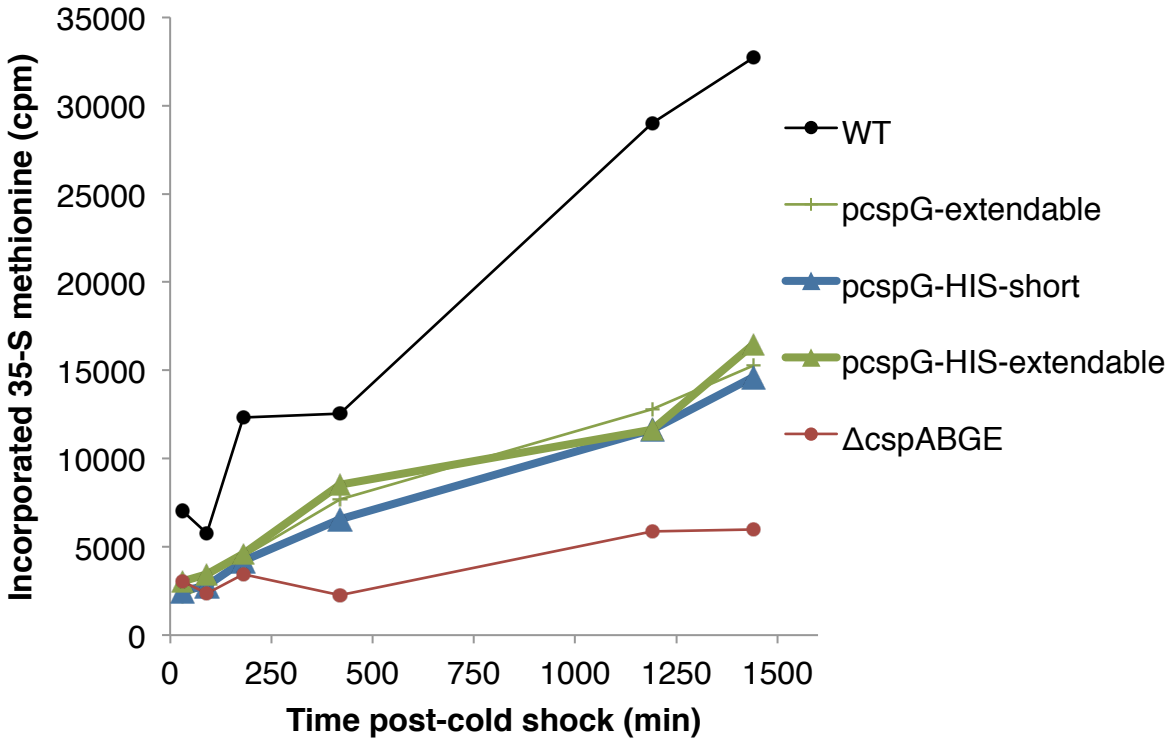

35000
30000
25000
20000
15000
10000
5000
0
Incorporated 35-S methionine (cpm)
0
250
500
750
1000
1250
1500
Time post-cold shock (min)
WT
pcspG-extendable
pcspG-HIS-short
pcspG-HIS-extendable
ΔcspABGE

Figure B.8. Cold sensitivity rescue of ∆*cspABGE* cold sensitivity by *cspG* complementation.

We complemented the cold sensitivity phenotype of ∆*cspABGE* using a low copy plasmid containing *cspG* with variable features (see Figure B.1). The unextended, short version of *cspG*, is labeled here as “s” and the extendable version is labeled here as “ex. The His-tag versions of each are labeled as “H-s” or “H-ex”. (A) The strains all grew normally at 37°C. (E) All forms of *cspG* were able to rescue the cold sensitivity of the quadruple deletion strain. Most notably rescue the non-viable phenotype seen at 15°C. The labeled H-s or H-ex strains appear to have full functionality as seen in panels C, D and E.

Figure B.8

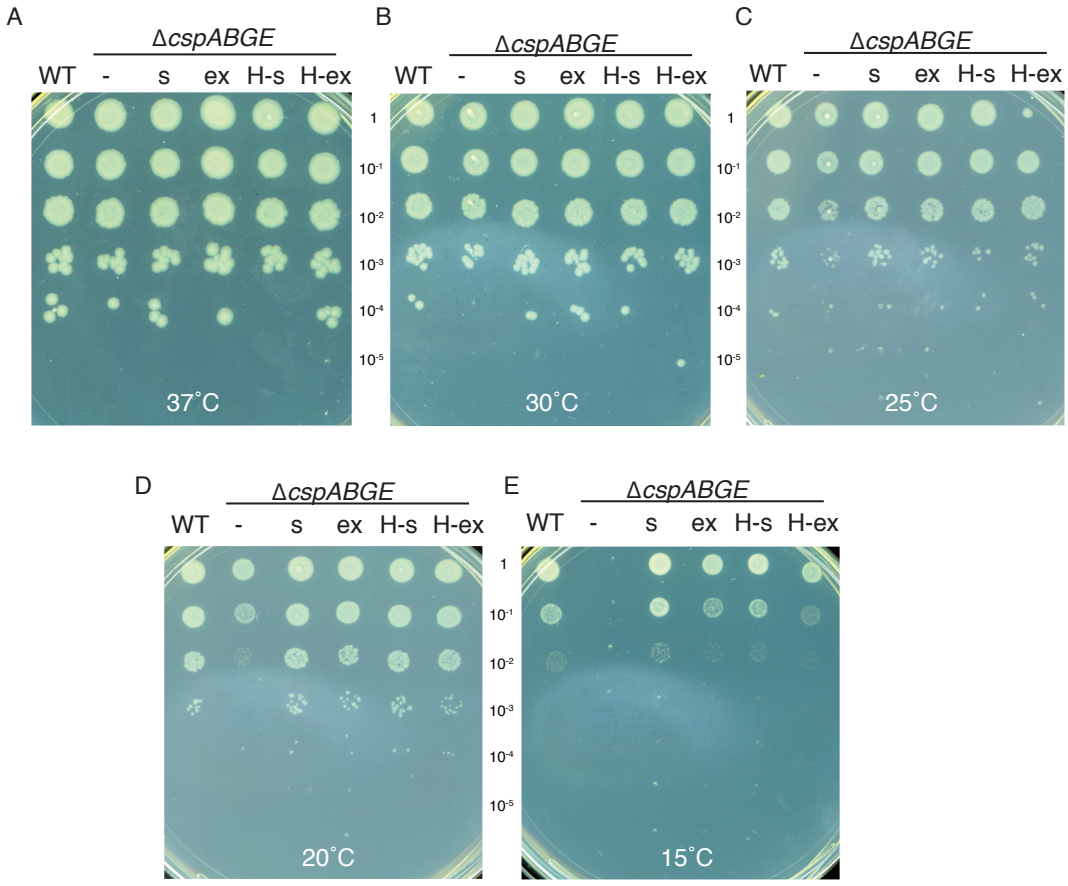
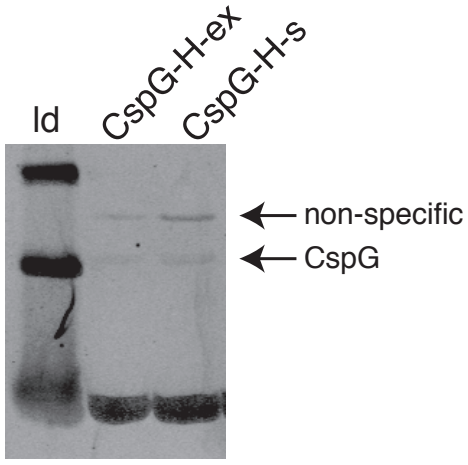

A
ΔcspABGE
WT - s ex H-s H-ex
37˚C
B
ΔcspABGE
WT - s ex H-s H-ex
30˚C
C
ΔcspABGE
WT - s ex H-s H-ex
25˚C
1
10-1
10-2
10-3
10-4
10-5
D
ΔcspABGE
WT - s ex H-s H-ex
20˚C
E
ΔcspABGE
WT - s ex H-s H-ex
15˚C

Figure B.9. Visualization of purified 5'His tagged CspG after cold shock.

Proteins were harvested 20 hours after cultures were shifted from 37°C to 10°C for strains containing either the plasmid for 5'His-*cspG*-short or 5'His-*cspG*-extendable. After purification using the His-tag, we visualized CspG and a non-specific protein, approximately the size of the expected extended CspG product.

Figure B.9



References

1. Graumann PL, Marahiel M a. A superfamily of proteins that contain the cold-shock domain. Trends Biochem Sci. 1998;23: 286–90.

2. Phadtare S, Inouye M. Genome-Wide Transcriptional Analysis of the Cold Shock Response in Wild-Type and Cold Sensitive, Quadruple-csp-deletion Strains of Escherichia coli. J Bacteriol. 2004;186: 7007–7014. doi:10.1128/JB.186.20.7007

3. Phadtare S, Alsina J, Inouye M. Cold-shock response and cold-shock proteins. Curr Opin Microbiol. 1999;2: 175–180.

4. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol. 2006;2: 2006.0008. doi:10.1038/msb4100050

5. Thomason L, Court DL, Bubunenko M, Costantino N, Wilson H, Datta S, et al. SPECIALIZED TECHNIQUES UNIT 1.16 Recombineering: Genetic Engineering in Bacteria Using Homologous Recombination. Curr Protoc Mol Biol. 2007;1: 1.

Appendix C

Future directions

Viability defect of the K-12 RF2^{B} ΔRF3 at 15°C

In our experiments we showed how growth defects of ΔRF3 in an *E. coli* K-12 background are rescued in the presence of the higher activity RF2^{B} variant. The cold sensitivity of K-12 RF2$^{K\text{-}12}$ ΔRF3 at 20°C was also rescued in the presence of RF2^{B}, consistent with previous observations [1]. In contrast to other temperatures, we found at 15°C the K-12 RF2^{B} ΔRF3 strain is more cold sensitive to the K-12 RF2$^{K\text{-}12}$ ΔRF3 strain. We further identified this sensitivity was a result of a 5-fold drop in viability rather than decreased doubled times found in the K-12 RF2$^{K\text{-}12}$ ΔRF3 strain. It is currently unclear to us why K-12 RF2^{B} ΔRF3 would suddenly loose viability at 15°C.

As proposed in our manuscript, *E. coli* K-12 may have a compensatory mutation in its genome that becomes deleterious at low temperature. I propose two experiments that could help understand this curious phenotype. The first experiment is to delete RF3 in *E. coli* B, which natively has the RF2^{B} variant, and test for viability at cold temperatures. If *E. coli* B ΔRF3 has normal viability at 15°C we can assume the original phenotype was derived from something specific to the K-12 strain. My top candidate for a K-12 specific variation is a point mutation in *rpsG*(S7) found only in *E. coli* K-12 [2]. This point mutation abolishes the original UAA stop codon of *rpsG* resulting in an extended protein, which terminates in the next in frame stop codon. This stop codon happens to be a UGA stop codon that is terminated with lower efficiency due to the lower activity of RF2$^{K\text{-}12}$ variant. Interestingly, *rspG* is located in an operon with *rpsL*(S1) upstream and *fusA*(EF-G) and *tufA*(EF-Tu). It has been identified that *rpsL* and *rpsG* use translational coupling to control expression of the operon [3]. In effect this results in two possible explanations for a K-12 change in translation; the first is the S7 protein that is extended by 23 amino acids in K-12 may alter functionality; the second is the misregulation of translation on the *rpsL-rpsG-fusA-tufA* operon altering expression of key translational proteins. To examine these possibilities the original UAA stop codon in *rpsG* can be restored in the K-12 RF2^{B} ΔRF3 strain and tested for viability at 15°C. Alternatively, or in addition to these experiments, the UGA stop

codon that *rpsG* relies on in *E. coli* K-12 could be changed to a more efficient UAA stop codon. Increased termination efficiency might eliminate misregulation of the operon. In addition to tests of viability, ribosome profiling could help identify specific misregulation of the *rpsL-rpsG-fusA-tufA* operon.

Other translation termination defects

Another potential future direction of my work would be manipulation of other translation termination genes followed by ribosome profiling. These studies could show if the phenotypes we identified by deletion of RF3 hold true for other defects. Given the inherent sensitivity of K-12 *E. coli* to translation termination stress, these experiments would need to be carried out in strains containing the $RF2^{B}$ variant to avoid accumulation of suppressors. Two particular genes, *prmC* and *rluD* are prime candidates for this study [1,4,5]. While I anticipate the phenotypes to be similar to those we observed in ΔRF3 strains the magnitude of these phenotypes and changes in expression of release factors or other termination machinery could change. This could potentially reveal regulatory mechanisms utilized by *E. coli* to maintain efficient termination.

References

1. O'Connor M. Interactions of release factor RF3 with the translation machinery. Mol Genet Genomics. 2015; doi:10.1007/s00438-015-0994-x
2. Dreyfus M, Heurgué-Hamard V. Termination troubles in Escherichia coli K12. Mol Microbiol. 2011;79: 288–91. doi:10.1111/j.1365-2958.2010.07468.x
3. Saito K, Nomura M. Post-transcriptional Regulation of the str Operon in Escherichia coli Structural and Mutational Analysis of the Target Site for Translational. 1994; 125–139.
4. Mora L, Heurgué-Hamard V, De Zamaroczy M, Kervestin S, Buckingham RH. Methylation of bacterial release factors RF1 and RF2 is required for normal translation termination in vivo. J Biol Chem. 2007;282: 35638–35645. doi:10.1074/jbc.M706076200
5. O'Connor M, Gregory ST. Inactivation of the RluD pseudouridine synthase has minimal effects on growth and ribosome function in wild-type Escherichia coli and Salmonella enterica. J Bacteriol. 2011;193: 154–162. doi:10.1128/JB.00970-10

Publishing Agreement

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

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

Author Signature Natalie Baggett Date 2/10/2017