CHARACTERIZATION OF SMALL RNAS AS RIBOSOMAL REGULATORS DURING GENE EXPRESSION IN BACTERIA

SYDNEE H. CALHOUN
Bachelor of Science, University of Lethbridge, 2019

A thesis submitted in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in

BIOCHEMISTRY

Department of Chemistry and Biochemistry
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

CHARACTERIZATION OF SMALL RNAS AS RIBOSOMAL REGULATORS DURING GENE EXPRESSION IN BACTERIA

SYDNEE H. CALHOUN

Date of Defense: June 30, 2022

Dr. H.-J. Wieden	Professor	Ph.D.
Dr. T. Patel	Associate Professor	Ph.D.
Thesis Co-Supervisors		
Dr. A. Russell	Associate Professor	Ph.D.
Thesis Examination Committee Member		
Dr. U. Kothe	Professor	Ph.D.
Thesis Examination Committee Member		
Dr. N. Thakor	Associate Professor	Ph.D.
Chair, Thesis Examination Committee		

ABSTRACT

Ribosome-associated non-coding RNAs (rancRNAs) are a recently discovered class of small RNA (sRNA) that regulate translation via direct interactions with the ribosome. rancRNA-dependent mechanisms have been confirmed in all domains of life but are poorly understood in bacteria. Since the ribosome is highly conserved among the three domains of life, common rancRNA mechanisms likely exist. To identify such potential regulatory rancRNAs, the sRNA pools of *Escherichia coli* ribosomes isolated under different stress conditions were analyzed using next-generation sequencing. Bioinformatic analysis detected 67 sRNAs, including eight novel sRNAs. In parallel, rancRNA candidates were identified using previously published RNAseq datasets. One candidate, rttR, was confirmed to preferentially bind to the 70S ribosome *in vitro*. However, rttR does not influence protein production *in vitro* or *in vivo.* Structural *in silico* analysis predicts that rttR interacts with 16S ribosomal RNA helix 17, suggesting a function for rttR during ribosome biogenesis, which requires further investigation.

PREFACE

The review article in Chapter 2, “Non-coding RNAs: What are we missing?”, was written by Cristina Carvalho Barbosa, Sydnee Calhoun, and HJ Wieden (1). This review discusses non-coding RNA biogenesis and their role during post-transcriptional regulation in eukaryotes and prokaryotes. Additionally, there is a focus on how the ribosome was understudied in relation to ncRNA regulation. The initial manuscript was written by Cristina Carvalho Barbosa. My contributions include the sections on ncRNA biogenesis, long non-coding RNAs, the history of ncRNAs interacting with the ribosome, and completed revisions during the publication process. HJ Wieden made contributions to editing and providing ideas for the manuscript.

The work completed in Chapter 3, “Investigation of annotated bacterial small RNAs functioning as ribosomal regulators,” was completed by Cristina Carvalho Barbosa, Sydnee Calhoun, and HJ Wieden. Cristina Carvalho Barbosa initially conducted the bioinformatic analysis and selected the rancRNA candidates. She also designed the constructs used for experiments and contributed to initial experimental design. I repeated the bioinformatic analysis to include additional studies and completed *in silico* analyses on the rancRNA candidates. I prepared the RNA for experiments and performed filter binding, PURE, RT PCR, and BONCAT experiments. HJ Wieden designed the project and provided help with the experiments.

The work completed in Chapter 4, “Identification of potential rancRNA candidates under stress conditions in *Escherichia coli*,” was completed by Sydnee Calhoun, Emily Wilton, Cristina Carvalho Barbosa, and HJ Wieden. Cristina Carvalho Barbosa planned the experimental design and selected the stress conditions to be tested. I made modifications to the protocols for cell growth and ribosome isolation and completed these experiments. cDNA library generation, sequencing, and DESeq2 analysis was completed by Azenta/Genewiz. I completed additional analyses on the DESeq2 data. Emily Wilton conducted the Rockhopper analysis on the sequencing data and further analysis of the predicted RNAs. HJ Wieden aided in project planning and provided feedback during experiments.

ACKNOWLEDGEMENTS

I would like to thank Dr. Hans-Joachim Wieden for supervising me during my Honours Thesis and Masters, Dr. Trushar Patel for agreeing to be my co-supervisor, and my committee members Dr. Tony Russell and Dr. Ute Kothe for providing insightful comments on my project. I would also like to thank Cristina Carvalho Barbosa de Souza for being my Honours Thesis supervisor and all the previous work she contributed to this project. Thank you to Dr. Borries Demeler and Amy Henrickson for teaching me how to conduct AUC experiments as well. Additionally, I would like to thank the students I supervised, Jingyi Li and Rebecca Avileli, for their contributions and letting me gain teaching experience. This project would not be possible without the help and feedback of many people that I would like to thank – Emily Wilton, Justin Vigar, Jalyce Heller, Luke Saville, Harland Brandon, Luc Roberts, and Dr. Govardhan Reddy Veerareddygari. In addition, thank you to all Wieden, Kothe, Patel, and Demeler lab members for additional laboratory assistance and feedback during laboratory meetings. Special thank you to Fan Mo for running the lab and many discussions and laughs about my project and non-related things. Lastly, thank you to my family and friends for their support during my studies, and all the friends I have made during my schooling.

TABLE OF CONTENTS

ABSTRACT iii

PREFACE iv

ACKNOWLEDGEMENTS v

LIST OF TABLES viii

LIST OF FIGURES ix

LIST OF ABBREVIATIONS xi

CHAPTER 1: Introduction 1

CHAPTER 2: Non-coding RNAs: What are we missing? 5

2.1: Abstract 5

2.2: Introduction 5

2.3: Translational regulation by ncRNAs 8

2.4: ncRNAs derived from functional RNAs 13

2.5: A new ncRNA target: the ribosome 14

2.6: What are we missing? 16

2.7: Acknowledgements 18

CHAPTER 3: Investigation of annotated bacterial small RNAs functioning as ribosomal regulators 19

3.1: Introduction 19

3.2: Methods 22

3.2.1: Identification of rancRNA candidates 22

3.2.2: *In silico* analysis of rancRNA candidates 23

3.2.3: RNA preparation and purification 24

3.2.4: Nitrocellulose filter binding 26

3.2.5: *In vitro* translation 27

3.2.6: *In vivo* translation 29

3.3: Results 31

3.3.1: rancRNA candidate identification 31

3.3.2: *In silico* Analysis of rttR 33

3.3.3: rancRNA candidates preferentially bind to ribosomes and ribosomal subunits 37
3.3.4: Assessing rancRNA candidate effects on translation 38
3.4: Discussion 43
3.5: Supplementary Information 47
CHAPTER 4: Identification of potential rancRNA candidates under stress conditions in *Escherichia coli*. 61
4.1: Introduction 61
4.2: Methods 62
4.2.1: Ribosome isolation 62
4.2.2: Preparation of RNA for next generation sequencing 64
4.2.3: Bioinformatic analysis 66
4.3 Results 67
4.3.1: sRNA isolation and sequencing 67
4.3.2: Ribosomal RNA is present in all sequencing reads 69
4.3.3: Sequencing reads primarily map to Coding RNA genes 72
4.3.4: Few ncRNAs are present in ribosomal particles under minimal medium stress 75
4.3.5: ncRNAs are abundant with 50S and 30S ribosomal subunits under hyperosmotic stress.. 78
4.3.6: ncRNAs are enriched in the Tetracycline 30S2 peak 80
4.3.7: Comparison between conditions reveals a common response to stress 84
4.4 Discussion 91
4.5 Supplementary Information 97
CHAPTER 5: Conclusions and Future Directions 126
REFERENCES 127

LIST OF TABLES

Table 2.1. Summary of previously identified rancRNAs in different organisms and their mechanism of action. 16
Table 3.1. PCR conditions for a 20 µL reaction using Phusion Polymerase. 24
Table 3.2. PCR cycle conditions for rancRNA candidates. 25
Table 3.3. *In vitro* transcription components for 100 µL reaction. 26
Table 3.4. Reverse transcription components for 20 µL reaction. 28
Table 3.5. List of rancRNAs identified. 32
Table 3.6. Apparent dissociation constants determined for rancRNA candidates. 38
Table S3.1. List of identified sRNAs found to associate with ribosomes in previously published datasets. 47
Table S3.2. *In silico* analysis predicting base pairing interactions of rancRNA candidates and 23S rRNA or 16S rRNA. 50
Table S3.3. rancRNA candidate sequences and primers. 54
Table S3.4. Dunnett's multiple comparison test for quantitative analysis of rttR's effect on *in vivo* translation within each induction condition. 57
Table S3.5. Dunnett's multiple comparison test for quantitative analysis of rttR's effect on *in vivo* translation within each induction condition. 58
Table S3.6. Dunnett's multiple comparison test for quantitative analysis of the scramble RNA's effect on *in vivo* translation between induction conditions. 60
Table S3.7. Dunnett's multiple comparison test for quantitative analysis of the scramble RNA's effect on *in vivo* translation within the induction condition. 60
Table 4.1. Significant ncRNAs identified in all conditions through DESeq2 analysis. 87
Table 4.2. Significant ncRNAs identified in all conditions through Rockhopper analysis. 88
Table S4.1. Sequencing statistics for 70S, 50S, and 30S ribosomes and subunits from LB condition. 99
Table S4.2. Sequencing statistics for 70S, 50S, and 30S ribosomes and subunits from hyperosmotic condition. 100
Table S4.3. Sequencing statistics for 70S, 50S, and 30S ribosomes and subunits from minimal medium condition. 102
Table S4.4. Sequencing statistics for 70S, 50S, and 30S ribosomes and subunits from tetracycline condition. 103
Table S4.5. List of ncRNAs identified in the minimal medium condition through DESeq2 analysis. 104
Table S4.6. List of ncRNAs identified in the hyperosmotic condition through DESeq2 analysis. 106
Table S4.7. List of ncRNAs identified in the tetracycline condition through DESeq2 analysis. 107
Table S4.8. List of ncRNAs identified in the minimal medium condition through Rockhopper analysis. 110
Table S4.9. List of ncRNAs identified in the hyperosmotic condition through Rockhopper analysis. 114
Table S4.10. List of ncRNAs identified in the tetracycline condition through Rockhopper analysis. 119

LIST OF FIGURES

Figure 1.1. Overview of known rancRNA mechanisms. 4
Figure 2.1. Biogenesis of small non-coding RNAs (sncRNAs) in prokaryotes and eukaryotes. 11
Figure 2.2. Mechanisms of base-pairing ncRNAs involved in modulation of protein synthesis. 12
Figure 3.1. Overview of characterization methodology for rancRNA candidates *in silico* and experimentally. 21
Figure 3.2. Overview of Biorthogonal non-canonical amino acid tagging (BONCAT) assay. 29
Figure 3.3. Identification of rancRNA candidates. 32
Figure 3.4. rttR is conserved in Enterobacteriaceae. 34
Figure 3.5. rttR is predicted to bind to 16S rRNA helix 17. 36
Figure 3.6. Binding of rancRNA candidates to ribosomes and ribosomal subunits. 38
Figure 3.7. rancRNA candidates, except for omrB, have minimal effects on translation *in vitro.* 40
Figure 3.8. *In vivo* monitoring of global *de novo* protein synthesis. 43
Figure S3.1. Predicted secondary structure of rancRNA candidates using RNAfold and generated with VARNA. 49
Figure S3.2. rttR genomic location in *E. coli* MG1655. 53
Figure S3.3. *In vitro* transcribed RNA of all rancRNA candidates and a scramble RNA control analyzed on a 10% urea PAGE and stained with ethidium bromide. 55
Figure S3.4. RT PCR of RNA added into PURExpress system analyzed on a 10% native PAGE and stained with ethidium bromide. 56
Figure S3.5. Biological replicates of extracted proteins from BONCAT assay with rttR induction. 57
Figure S3.6. Control experiment of *in vivo* monitoring of global *de novo* protein synthesis with scramble RNA induction. 58
Figure S3.7. Biological replicates of extracted proteins from BONCAT assay with scramble RNA induction. 59
Figure S3.8. RT PCR of RNA transcribed during BONCAT assay analyzed on 10% native PAGE gels and stained with Ethidium Bromide. 59
Figure 4.1. Overview of ribosome isolation protocol. 62
Figure 4.2. Overview of RNA preparation workflow for next generation sequencing. 64
Figure 4.3. Induced stress conditions affect ribosomes and ribosomal subunit abundance and distributions. 68
Figure 4.4. Reads for the minimal medium condition are mostly ribosomal RNA. 70
Figure 4.5. Ribosomal RNA makes up the majority of reads for the hyperosmotic condition 70
Figure 4.6. Coding RNA reads are more abundant in the 30S region under tetracycline stress. 71
Figure 4.7. Increase in coding RNA and tRNA reads for the 50S tetracycline condition. 71
Figure 4.8. Less reads mapping to genes are identified in the minimal medium condition sample. 72
Figure 4.9. The hyperosmotic condition has a decrease in total reads mapping to genes. 73
Figure 4.10. Total number of genes do not vary between tetracycline and LB conditions in the 30S region. 73

Figure 4.11. A higher number of tRNA gene related reads are present in the 50S than the 70S. 74
Figure 4.12. ffs is more abundant in the 30S in minimal medium. 75
Figure 4.13. ncRNAs are predominantly lower in abundance for the minimal medium 30S compared to the LB 30S. 77
Figure 4.14. Five ncRNAs are significantly enriched in the 30S and 50S ribosomal subunits. 78
Figure 4.15. ncRNAs are significantly enriched with the hyperosmotic 70S when compared to the LB 70S. 79
Figure 4.16. More ncRNAs are enriched in the 30S2 peak than the 30S1 peak. 80
Figure 4.17. No ncRNAs are significantly more abundant or in low abundance for the 70S ribosome when compared to the LB 70S. 81
Figure 4.18. ncRNAs are in low abundance in the 30S1 peak and abundant in the 30S2 peak when compared to the LB 30S. 83
Figure 4.19. Majority of the ncRNAs are abundant in both the 50S and 70S when compared to the LB 50S and 70S ribosomal particles. 84
Figure 4.20. The majority of ncRNAs are shared between stress conditions but significantly enriched ncRNAs are condition specific. 85
Figure 4.21. Most ncRNAs are shared across all conditions while significantly enriched ncRNAs are condition specific. 86
Figure S4.1. RNA fractions extracted from *E. coli* MG1655 cells grown in LB medium. 97
Figure S4.2. rRNA operons are sequenced at specific positions. 98

LIST OF ABBREVIATIONS

5S	5 svedberg ribosomal RNA
16S	16 svedberg ribosomal RNA
23S	23 svedberg ribosomal RNA
30S	30 svedberg prokaryotic ribosomal subunit
40S	40 svedberg eukaryotic ribosomal subunit
50S	50 svedberg prokaryotic ribosomal subunit
60S	60 svedberg eukaryotic ribosomal subunit
70S	70 svedberg prokaryotic ribosome
AGO	Argonaute protein
ASO	Antisense oligonucleotide
ATP	Adenosine triphosphate
BAM	Binary alignment map
BONCAT	Biorthogonal non-canonical amino acid tagging
BP	Base pair
cDNA	Complementary DNA
CPP	Cell penetrating peptide
Csr	Carbon storage regulation
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DTT	Dithiothreitol
eCFP	Enhanced cyan fluorescent protein
HIV	Human immunodeficiency virus
IPTG	Isopropyl β-D-1-thiogalactopyranoside
K_D	Equilibrium dissociation constant
kDa	Kilodalton
L-AHA	L-Azidohomoalanine
LB	Luria broth
lncRNA	Long non-coding RNA
MIC	Minimum inhibitory concentration
miRNA	MicroRNA
mRNA	Messenger RNA
ncRNA	Non-coding RNA
NGS	Next generation sequencing
nt	Nucleotide
NTP	Nucleotide triphosphate
ORF	Open Reading Frame
OD	Optical density
Padj	Adjusted P value
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PNA	Peptide nucleic acid
rancRNA	Ribosome-associated non-coding RNA
RBS	Ribosome binding site
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RPKM	Reads per kilobase of transcript, per million mapped reads
RPM	Revolutions per minute
rRNA	Ribosomal RNA
RT PCR	Reverse transcription polymerase chain reaction
ScrRNA	Scramble RNA
SD	Shine-Dalgarno sequence
SDS	SEM

Sodium dodecyl sulfate	siRNA
Standard error of means	Small-interfering RNA
sncRNA	Small non-coding RNA
sRNA	Small RNA
SRP	Signal recognition particle
tmRNA	Transfer-messenger RNA
TOR	Target of rapamycin pathway
Trab	Transcription buffer
tRF	tRNA-derived fragment
tRNA	Transfer RNA
UTR	Untranslated region
UV	Ultraviolet
WHO	World Health Organization

CHAPTER 1: Introduction

Gene expression is a highly regulated process that occurs in all living cells. One of the later steps in gene expression is the process called translation. This step is ribosome-dependent and results in the decoding of messenger RNA (mRNA) transcripts into their corresponding protein. Previous work has identified several factors that aid in the regulation of translation, one of which being small non-coding RNAs (sncRNAs) (2). sncRNAs have been shown to interact with mRNA transcripts in multiple ways, resulting in mRNA degradation, translation initiation inhibition by base pairing, and recruitment of chaperones like Hfq to sequester mRNA (3-6). Additionally, sncRNAs have been found to aid in adaptation to stress conditions such as temperature, nutrient starvation, or osmotic effects through the regulation of translation (7, 8). While the majority of research has focused on sncRNA:mRNA regulation, tRNAs have also been recently found to be regulated by sncRNAs (9, 10). However, the regulation of the ribosome itself by sncRNAs is only poorly understood (11, 12).

In 2012, a new class of sncRNAs was identified by Norbert Polacek and called ribosome-associated non-coding RNAs (rancRNAs) (13, 14). These rancRNAs have been found to directly bind to the ribosome and regulate translation (12, 15). An overview of characterized rancRNAs is shown in Figure 1.1.

In *Haloferax volcanii*, a 5' $tRNA^{Val}$ fragment was demonstrated to compete with mRNA for binding in the mRNA channel under alkaline stress (14, 16). Interestingly, the tRNA fragment was shown to interfere with both eukaryotic and bacterial ribosome functions, suggesting conservation of this mechanism across the domains of life (16). Additionally, rancRNA_s194 was found in an intergenic region in *H. volcanii* that inhibits protein synthesis *in vitro,* but specifically regulates and represses translation of the mRNA *cstA* under xylose stress *in vivo* (17).

rancRNAs have also been identified in several eukaryotic organisms. In *Saccharomyces cerevisiae*, rancRNA_18, an mRNA-derived fragment from the *TRM10* gene, has been shown to be upregulated under hyperosmotic stress and to compete with tRNA binding to the ribosome (18, 19). Also in *S. cerevisiae,* several tRNA-derived fragments (3' $tRNA^{His}$, 3' $tRNA^{Leu}$, 3' $tRNA^{Gly}$, 3' $tRNA^{Ser}$, 3' $tRNA^{Thr}$, 5' $tRNA^{His}$) have been shown to interfere with global translation and can impact the ability of tRNA aminoacyl

synthetases to interact with the ribosome under heat shock, cold shock, acidic, basic, hyper/hypo conditions, UV exposure, and nutrient limitation (20, 21). In the eukaryotic protist *Trypanosoma brucei*, a 3' $tRNA^{Thr}$ half has been characterized that stimulates mRNA loading to the ribosome under nutrient starvation (22). Other small RNAs (sRNAs) have been found to simultaneously bind to mRNA and the ribosome in *T. brucei* as well (23). Furthermore, a proline 5' tRNA half, termed ProTiP, has been found to bind to 60S ribosomal subunits in several mammalian cell lines (CHO-K1, HeLa, HEK) and cause stalling and the accumulation of a peptidyl-tRNA (24). Lastly, 5' $tRNA^{Ala}$ fragments downregulate global translation *in vitro* in *Arabidopsis thaliana*, but the same effects could not be confirmed *in vivo* (25).

More recently, a rancRNA has been identified in bacteria for the first time. In *Staphylococcus aureus*, the 5' end of SprF1 RNA, a member of a toxin-antitoxin (TA) type I pair, can inhibit translation globally (26, 27). Preliminary work suggests that *Escherichia coli* sRNAs may be acting as rancRNAs in stationary phase (28). However, validation is required for these findings. In summary, rancRNA mechanisms are poorly understood and are weakly characterized, particularly in bacteria. Furthermore, although evidence exists for conservation of eukaryotic and archaeal rancRNAs in bacteria, no mechanistic information is available that confirms such a regulatory role (16, 18). Because the structure of the ribosome is highly conserved between the three domains of life, it is likely that there are evolutionary reasons for preserving this class of sncRNA (29).

To this end, we explored how rancRNAs may be functioning in bacteria, using *E. coli* as a model organism. Re-analyzing existing RNAseq data, we identified rancRNA candidates (arcZ, omrB, rprA, rttR, rybA, and tpke11) from existing annotated sRNAs for experimental validation (30-32). Primarily characterizing rttR, I demonstrate that rttR preferentially binds to 70S ribosomes with nanomolar affinity but does not bind to 50S or 30S ribosomal subunits. However, rttR only minimally effects protein synthesis *in vitro* and likely does not affect global translation *in vivo*. Further work is required to determine if rttR can be classified as a rancRNA as well as its mechanism of action.

In parallel, sRNAs were isolated and enriched from 70S ribosomes and 50S and 30S ribosomal subunits from cells grown under various stress conditions. Using DESeq2 and Rockhopper bioinformatic analyses, 14 ncRNAs and 67 ncRNAs were identified, respectively. Intriguingly, eight non-annotated

transcripts were detected using Rockhopper. Further *in silico* analysis will be required to determine candidate ncRNAs for experimental characterization.

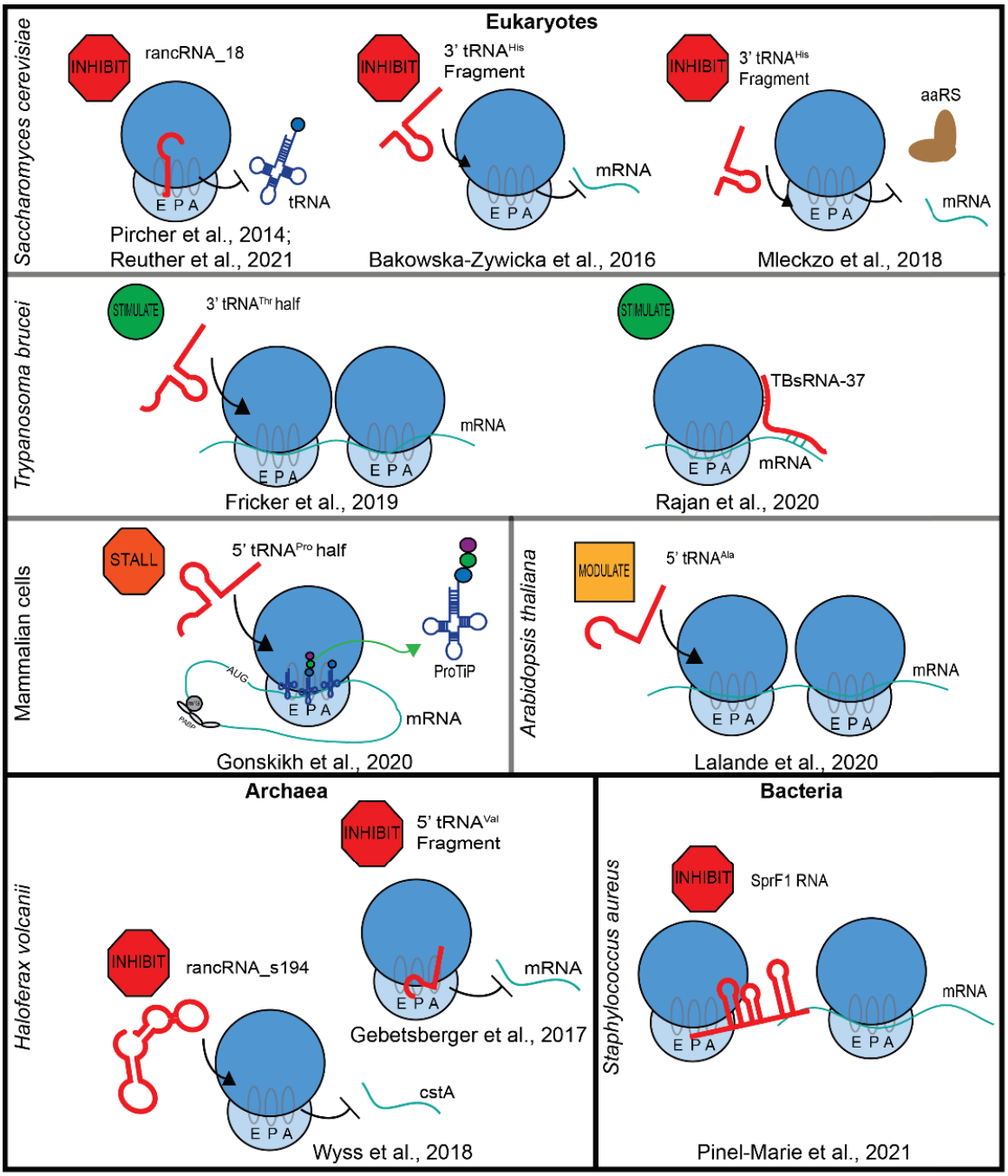


Figure 1.1. Overview of known rancRNA mechanisms. rancRNAs have been discovered in all three domains of life: eukaryotes, archaea, and bacteria. In eukaryotes and archaea, primarily mRNA and tRNA-derived fragments have been identified as rancRNAs and to downregulate translation by blocking a tRNA site (rancRNA_18, 5' $tRNA^{Pro}$ half) or the mRNA channel (3' $tRNA^{His}$ fragment, 5' $tRNA^{Ala}$ fragment, 5' $tRNA^{Val}$ fragment, rancRNA_s194). Positive effects on translation where mRNA loading is stimulated by rancRNAs has been observed in *Trypanosoma brucei*. In bacteria, SprF1 RNA, a toxin-antitoxin RNA, is the only known rancRNA where translation inhibition occurred under osmotic stress.

CHAPTER 2: Non-coding RNAs: What are we missing?

Reprinted with permission from “Non-coding RNAs: What are we missing?”

Carvalho Barbosa, C., Calhoun, S.H., and Wieden, H.-J. “Non-coding RNAs: What are we missing?” Biochem Cell Biol. 2020; 98(1): 23-30.

2.1: Abstract

Over the past two decades, the importance of small non-coding RNAs (sncRNAs) as regulatory molecules has become apparent in all three domains of life (archaea, bacteria, eukaryotes). In fact, sncRNAs play an important role in the control of gene expression at both the transcriptional and the post-transcriptional level, with crucial roles in fine-tuning cell responses during internal and external stress. Multiple pathways for sncRNA biogenesis and diverse mechanisms of regulation have been reported, and although biogenesis and mechanisms of sncRNAs in prokaryotes and eukaryotes are different, remarkable similarities exist. Here, we briefly review and compare the major sncRNA classes that act post-transcriptionally, and focus on recent discoveries regarding the ribosome as a target of regulation and the conservation of these mechanisms between prokaryotes and eukaryotes.

Key words: non-coding RNA, ribosome-associated non-coding RNA, translation regulation, bacteria.

2.2: Introduction

The term gene expression commonly refers to the process whereby the information encoded in DNA is converted into a messenger RNA (mRNA) copy, which is then translated into the corresponding protein (33). It is a complex, multi-level, and tightly regulated process fundamental to the development, homeostasis, and adaptation of all living cells (34, 35).

Gene regulation can be described as the processes that modulate the expression of a specific gene, or group of genes. It directs the efficient use of cellular resources and, in multicellular organisms, it defines the cell type, function, and structure (36). The regulation of gene expression

influences various biological processes and requires high fidelity and precise control. Even unicellular organisms use their genes selectively. It is well recognized that regulation of gene expression can occur by multiple mechanisms at several levels, and that misregulation is usually associated with physiological and developmental dysfunctions (33). For example, the misregulation of the transcriptional factor NF-KB has been linked to different diseases such as arthritis, asthma, and cancer (37).

Transcription, the synthesis of an RNA copy of the DNA, is the primary regulatory point in gene expression. The diversity, complexity, and variability of transcriptional regulators is vast, and they are present in all domains of life. Historically, transcriptional regulation has received the most attention (33, 36). However, post-transcriptional regulation, which includes translational regulation, adds additional points of control and complexity in the pathway from DNA to protein. Indeed, there seems to be a high level of coordination and interdependence between transcriptional and post-transcriptional control (36). For example, the activation of the conserved signaling target of rapamycin (TOR) pathway causes changes in transcription and subsequently affects the remainder of gene expression, activating metabolic pathways for cell growth, and coordinates ribosome and protein synthesis under nutrient and growth stress (33, 36, 38).

Translation, the synthesis of proteins using mRNA as a template, is one of the later steps during gene expression, and is a key component in its control. Protein synthesis is a highly energy-intensive process and, therefore, requires tight regulation (39). The consumption of energy by protein synthesis can reach 50% of the energy generated by rapidly growing *Escherichia coli* cells and 5% of the human caloric intake (40). The regulation of translation allows rapid and flexible changes in gene expression, and is used by cells and organisms to respond to intra- and extracellular signals, such as nutrient supply, hormones, and stress, thus playing an important role in modulating different biological processes (41).

Organisms have developed a variety of regulatory mechanisms to fine-tune protein synthesis (39). These include structural features, such as the addition of a 5’ 7-methyl guanosine cap and a

poly(A) tail in eukaryotes, or regulatory sequences including the 5' and 3' untranslated regions within the mRNA responsible for its translational fate, phosphorylation of initiation and elongation factors, and regulatory non-coding RNAs (ncRNAs) present in both eukaryotes and prokaryotes (41-43).

In addition to acting as intermediate carriers of genetic information, RNAs play key structural, catalytic, and regulatory roles in the cell. Over the past few decades, a fast-growing field of research regarding regulatory non-coding RNAs (ncRNAs) and their involvement in the control of gene expression, including translation, has developed. Since the first description of an mRNA-interfering complementary RNA in 1984 and microRNAs (miRNAs) in 1993, several lines of evidence, from bacteria to humans, have implicated riboregulators as mediators of gene expression at the post-transcriptional level, through both mRNA stability and translation regulation (4, 12, 44, 45). The prokaryotic genome is composed primarily of protein-coding sequences, with non-coding DNA reported in the range of 1% to 15% (46-48). In contrast, only 2% to 5% of the eukaryotic genome encodes proteins, while the remainder has been found to encode for non-coding RNAs (46). Regulatory ncRNAs are now recognized to be powerful and important translational regulators. A large number of studies has revealed an abundance of ncRNAs and several different mechanisms of regulation exerted by them. Nevertheless, a plethora of new fragments and potential mechanisms continue to be identified in all three domains of life (archaea, bacteria, and eukaryotes). It is likely that the variety and functions of these ncRNAs is underestimated, with mechanisms still unexplored. While similar ncRNA regulatory mechanisms have been identified in the three domains of life, there are noticeable gaps, especially in regard to the ribosome being the direct target of ncRNAs. Given that the ribosome is highly conserved amongst the three domains of life, we hypothesize that mechanisms that regulate expression at the ribosomal level exist in all domains of life, but might have gone unreported so far for some of them. In this review, we highlight the roles of ncRNAs involved in translational control in both prokaryotes and eukaryotes, and compare their mode of action focusing on recent discoveries in relation to the ribosome as the primary target of regulation.

2.3: Translational regulation by ncRNAs

Non-coding RNAs are a heterogeneous class of RNA molecules present in both prokaryotes and eukaryotes that do not encode for a protein, but that are functional within the organism. Although a variety of ncRNAs have been identified, characterization and systematic description of these ncRNA species is lagging behind. Among the ncRNAs, long non-coding RNAs (lncRNAs) are classified based on their length (greater than 200 to 300 nucleotides) (49). They have been identified in both eukaryotes and prokaryotes and have been shown to regulate post-transcriptionally by targeting the mRNA transcript (50, 51). However, lncRNAs have been found to be poorly conserved by comparison with other RNAs, and recent studies are exploring whether this ncRNA species is truly "non-coding" in all cases (52-54). Aside from lncRNAs, other ncRNA species have been identified and mainly classified based on their length, called small non-coding RNAs (sncRNAs). Figure 2.1 summarizes the classification and processing of sncRNAs. Short interfering RNAs (siRNAs) and microRNAs (miRNAs) in eukaryotes and small RNAs (sRNAs) in prokaryotes are the three most extensively studied classes of sncRNAs.

miRNAs and siRNAs are RNA transcripts of approximately 21 to 26 nucleotides (nt) in length and originate from larger RNA precursors. These sncRNAs are considered key regulators of gene expression (55). It has been suggested that, in higher order eukaryotes, more than 60% of all mRNAs are regulated by miRNAs (56-58), and humans express over 1000 different miRNAs. These short RNAs control gene expression by base-pairing with specific mRNAs and controlling their stability and translation. Although miRNAs and siRNAs are functionally equivalent, they can be distinguished by their mode of biogenesis and target complementarity (55, 59).

miRNAs are generated from an endogenous single-stranded precursor that forms a double-stranded stem loop and is then processed by two endonucleases, called Drosha and Dicer. The miRNA associated with Dicer and Argonaute (AGO) proteins forms an RNA-induced silencing complex (RISC) that guides the translational inhibition or cleavage of target mRNAs through the action of AGO proteins in a sequence-specific manner (56, 60). miRNAs show partial

complementarity to their targets, making them especially useful translational regulators, since a single miRNA can regulate different mRNA targets.

In contrast to miRNAs, siRNAs are usually exogenous molecules derived from viral particles, transposons, or transgenes, and are perfectly complementary to their mRNA target (58, 60, 61). They were originally discovered in plants, where they were shown to be involved in post-transcriptional gene silencing and predominantly in response to viral infection. Despite the differences, processing of both miRNAs and siRNAs is dependent on Dicer, and their regulatory function is exerted through the same RISC components previously identified for miRNAs (60). The presence of a foreign double-stranded RNA triggers a protein complex containing the endonuclease Dicer, which cleaves the RNA into short fragments. The short fragment is then loaded into RISC, which discards one strand and uses the remaining strand to target foreign mRNA molecules for degradation.

In contrast to their eukaryotic counterparts, prokaryotic sRNAs usually lack processing of primary transcripts and vary in size and structure (62). Typically, sRNAs are smaller than 300 nt in length and fulfill a variety of important regulatory functions, particularly under stress conditions such as nutrient starvation and oxidative stress (63, 64). The subdivision between perfect or imperfect base pairing with the target can also be applied to this class of sncRNAs. *Cis*- encoded RNAs are encoded on the antisense strand of the gene that they regulate, and show full complementarity with their targets, whereas *trans*-encoded RNAs are encoded at a distant genomic location and possess limited and imperfect base pairing with their targets (65-67). Like miRNAs, because of their partial complementarity to their targets, a single *trans*-encoded RNA can regulate different mRNA targets. Not unlike miRNAs and siRNAs, some sRNAs can be a part of a ribonucleoprotein complex, which in this case, contains an RNA chaperone such as Hfq or ProQ (62, 68). Recent reports have also indicated that RNase E interacts with the sRNA–Hfq complex for targeted degradation (69). Additionally, RNase E has been found to have a membrane-binding domain, and thus does not aid in the sRNA binding to its target (70, 71). However, given that RNA chaperones are not universally

conserved in prokaryotes, it has been speculated that complex formation is not essential for an sRNA to base pair with its target mRNA (72).

The number of mechanisms employed by sncRNAs to either up- or downregulate translation is large. A small number of sncRNAs can bind to and modulate protein activity in prokaryotes and eukaryotes. For example, in bacteria, CsrB and CsrC sncRNAs are known to bind to and regulate the activity of the translational regulator CsrA, which indirectly regulates the expression of a large set of genes (73). In the cell lines HEK293T, MDA-MB-231, and CN34, tRNA-derived fragments (tRFs) were found to antagonize YBX1, an RNA-binding protein that has been found to stabilize oncogenic transcripts (74, 75). However, the largest group of known sncRNAs activates or represses translation by base pairing with other nucleic acids, and the specificity of the action is related to the base pairing interactions (76). For example, the extent of the complementarity of miRNAs determines whether the target is going to be degraded or not. The range of regulatory mechanisms used by base-pairing sncRNAs are illustrated in Figure 2.2.

Even though sncRNAs differ in their mechanistic details between species or kingdoms, it is possible to draw parallels with respect to the overall design of their respective regulatory processes (62). For example, in prokaryotes, translation inhibition can occur by direct competition for the ribosome-binding site (RBS). In this case, the sRNA prevents the binding of the 30S subunit and translation initiation by base pairing with regions close to or overlapping with the Shine–Dalgarno sequence (SD) or start codon (77-79). Similarly, eukaryotic sncRNAs induce translational repression by preventing the binding of the 40S subunit and thus blocking initiation. However, in this case, translation initiation is inhibited by targeting the poly(A) tail (80). Although the mechanism of translation initiation in prokaryotes and eukaryotes differ, sncRNAs can induce translational

repression at the initiation step by preventing the binding of the small ribosomal subunit. Therefore, itis possible to suggest an evolutionary conservation of their regulatory processes and strategies.

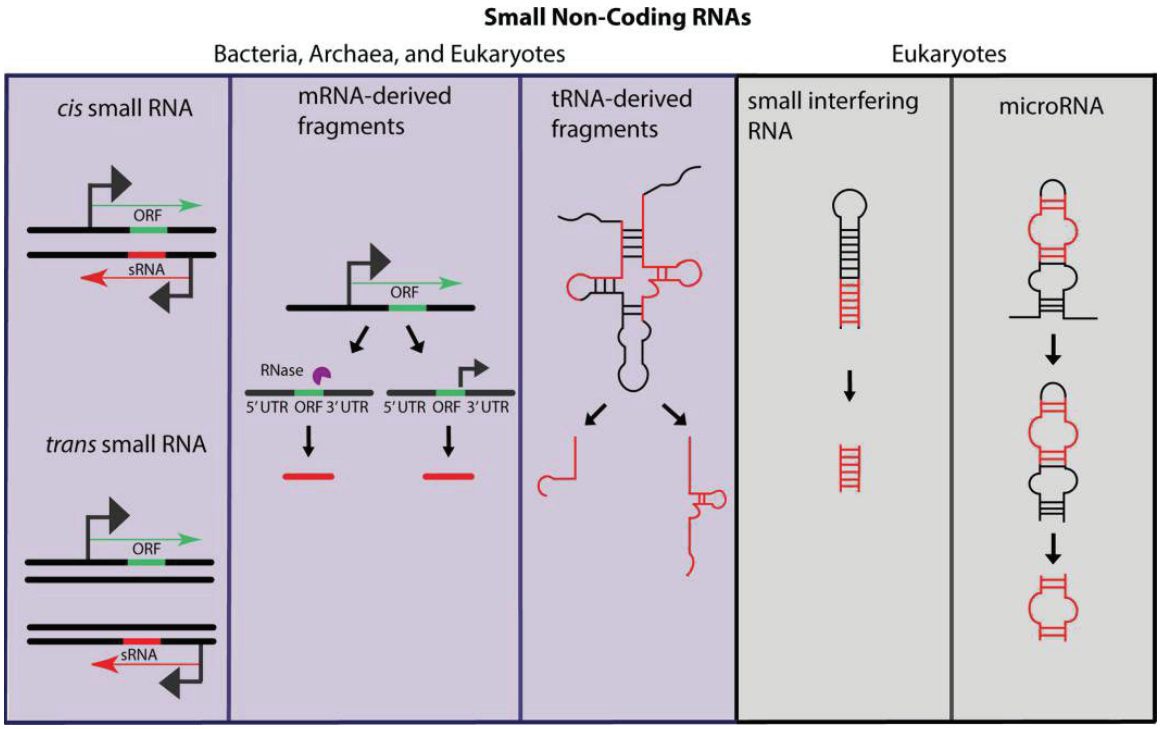


Figure 2.1. Biogenesis of small non-coding RNAs (sncRNAs) in prokaryotes and eukaryotes. In general, bacteria and archaea utilize processing pathways involving RNases such as E, Z, or P to generate sncRNAs. Eukaryotes have more complex pathways that involve multiple precursors and additional protein components like Drosha and Dicer for processing precursor molecules into a sncRNA.

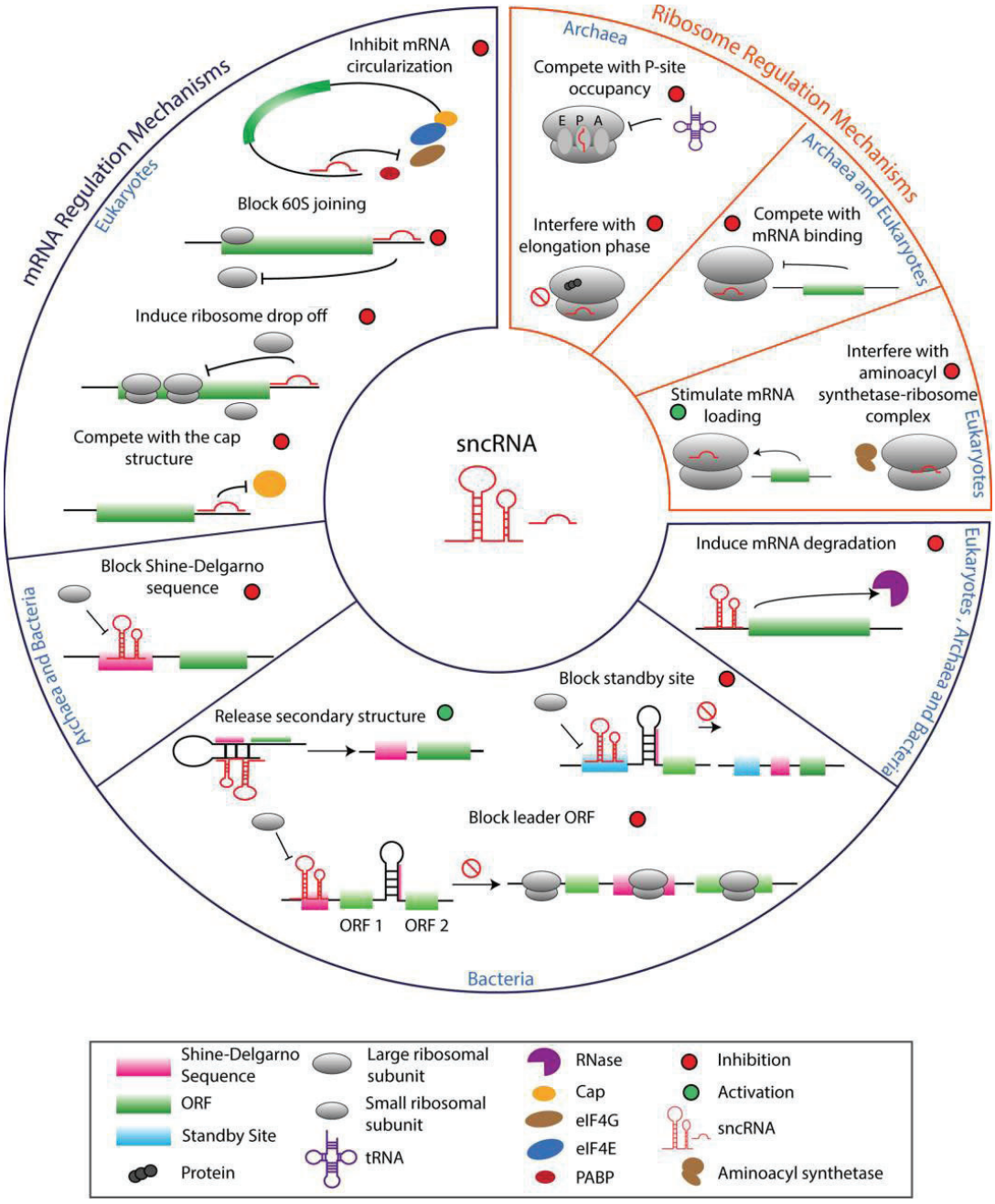


Figure 2.2. Mechanisms of base-pairing ncRNAs involved in modulation of protein synthesis. Base-pairing ncRNAs can be subdivided into two classes: mRNA-targeting and ribosome-targeting ncRNA. Non-coding RNAs are versatile and not always limited to one regulatory mechanism. For example, a single ncRNA is capable of regulating translation using two mechanisms, simultaneously blocking the ribosome from binding and inducing mRNA degradation.

2.4: ncRNAs derived from functional RNAs

As our understanding of gene expression regulatory targets and mechanisms progresses, the complexity of this field is increasingly evident. In the expanding repertoire of sncRNAs, recent transcriptomic studies in bacteria, archaea, and eukaryotes have provided a large amount of data that shed light on new sncRNAs generated by the processing of mature functional RNAs such as mRNAs and tRNAs (Fig. 2.1) (81, 82). Initially these fragments were considered random degradation products. However, accumulating evidence suggests that these are conserved, precisely processed fragments that can interact indirectly or directly during translation (9, 82).

In particular, sncRNAs generated from tRNAs have been identified in a wide range of organisms from all three domains of life (83). Besides their well-defined role in translation, it has been suggested that tRNAs exert regulatory function through the production of different sncRNAs derived from their transcripts. According to the relative size and biogenesis, the tRNA-derived fragments can be divided into two subclasses: tRNA halves, and tRFs. The first group is composed of transcripts of approximately 28 to 36 nt in length and usually are generated under stress conditions by anticodon nucleases such as ColicinE5 in bacteria, Rny1p in yeast, and angiogenin in humans (83-85). The latter group is composed of transcripts with an average length between 18 to 20 nt and are processed either by Dicer or RNase Z in eukaryotes and RNase E in *E. coli* (9, 83). Several studies have suggested that tRNA-derived fragments are involved in the regulation of protein synthesis in different organisms (86). Mechanistic studies have suggested that they function similarly to miRNAs, and their direct influence on protein biosynthesis has been revealed in archaea and in humans (14, 20, 87, 88).

Small ncRNAs derived from mRNA fragments have been reported in eukaryotes and bacteria (89-92). While mRNA fragments are often associated as products of mRNA degradation, cleaved RNA fragments may be used under certain circumstances to generate functional small regulatory RNAs (81, 82). Several studies have revealed sncRNAs that correspond to different known mRNA regions in both eukaryotes and bacteria (30, 79, 93). In addition, high-throughput sequencing of Hfq-bound and AGO-bound RNAs has revealed several mRNA-derived fragments associated with these proteins (79,

94). Although our current knowledge of their biological functions is limited, hints at the power of some of these mRNA-derived fragments in translation regulation were recently demonstrated in *Saccharomyces cerevisiae* and *Salmonella* (18, 92).

2.5: A new ncRNA target: the ribosome

The discoveries in sncRNA-mediated translational regulation are not limited to new molecules. New sncRNA targets and regulatory mechanisms have recently been discovered. As illustrated in Figure 2.2, the most extensively studied sncRNAs modulate protein synthesis through direct interactions with mRNA targets. However, increasing evidence suggests the ribosome as a common sncRNA target (18). This new group of sncRNAs, termed ribosome-associated non-coding RNA (rancRNA), was found to regulate translation through direct interactions with the ribosome in archaeal and eukaryotic cells (Fig. 2.2; Table 2.1) (18, 20). While the definition of a rancRNA is based on a direct interaction with the ribosome, it should be noted that there are examples where a rancRNA can target an mRNA as well as the ribosome. Firstly, transfer-messenger RNAs (tmRNAs) have been identified in bacteria, and function to monitor translation, relieving ribosome stalling on truncated mRNAs by facilitating tagging of the incomplete protein for degradation and ultimately recycling of the stalled ribosome (95, 96). Additionally, the RNA component of the universally conserved signal recognition particle (SRP) complex aids in localizing specific proteins to the endoplasmic reticulum in eukaryotes and the plasma membrane in prokaryotes (97, 98). Surprisingly, to our knowledge, no studies have been published on the presence of rancRNAs in bacteria where the ribosome is a specific target of the ncRNA. The presence of sncRNA transcripts on ribosome subunits is not new, but they were initially interpreted as biological noise. This view has changed after the direct association of specific individual rancRNA fragments to the ribosome has been experimentally confirmed (14, 18). Moreover, these associations were shown to change under different stress conditions and influence the process of protein synthesis.

The first evidence of ncRNAs targeting the ribosome was reported by the Polacek group in 2012 (14). In the archaebacteria *Haloferax volcanii*, Gebetsberger and colleagues found a 26 nt tRF that originated from the 5' region of valine tRNA (Val-tRF), which can downregulate global translation

under hyperosmotic stress, as well as the extent of peptide bond formation through direct interactions with the 30S subunit *in vivo* and *in vitro*. Further characterization provided evidence that the respective tRNA fragment binds to the small ribosomal subunit in the immediate vicinity of the mRNA channel, causing the dis- placement of mRNA transcripts from the initiation complex resulting in global translation attenuation (Fig. 2.2) (16). Furthermore, another study with the archaebacteria *H. volcanii* has identified an intergenic sRNA, designated rancRNA_194, that can inhibit protein synthesis of *cstA*, the mRNA transcript of a carbon starvation protein that was found to be involved when cells utilize xylose as the carbon source (17). Similarly to the Val-tRF, *in vitro* studies have shown that this rancRNA interacts with the large ribosomal subunit and inhibits peptide bond formation under specific stress conditions. However, in contrast to Val-tRF, rancRNA_194 is an mRNA specific regulator and does not influence global translation *in vivo* (17).

The existence of rancRNAs and their ability to regulate translation by direct interactions with the ribosome is not limited to the archaebacteria *H. volcanii.* It has also been demonstrated in the eukaryote *S. cerevisiae* and in the pathogenic protozoan *Trypanosoma brucei* (18, 20, 22). In 2014, Pircher and colleagues functionally characterized an ncRNA derived from an exon in the *TRM10* mRNA. This ncRNA can completely shut down global translation under hyperosmotic conditions through its direct interactions with polysomes (18). This rancRNA associates with 60S ribosomal subunits and possibly interferes with tRNA binding to the ribosomal P site (Fig. 2.2), suggesting a direct regulation of polysomes (18). In addition to the mRNA-derived fragment, tRNA-derived fragments from the 3' end of five tRNAs (tRNA-His, tRNA-Ser, tRNA-Gly, tRNA-Leu, tRNA-Thr) and from 5' end of one tRNA (tRNA-His) were also reported to directly interact with ribosome and induce translation inhibition in *S. cerevisiae* under stress conditions (20). A more recent study also suggests that tRNA-derived fragments can modulate translation via interactions with ribosome-associated aminoacyl-tRNA synthetases, which is an alternative to the direct regulation of ribosomal inhibition (21).

The most recent example of rancRNAs identifies a 3' tRNA half, derived from tRNA-Thr, that is capable of interacting with ribosomes and polysomes in *T. brucei* (22). Similar to the rancRNAs summarized so far, this respective tRNA fragment is produced during nutrient stress (starvation and stationary phase). Surprisingly, this rancRNA can stimulate, rather than inhibit, protein synthesis by promoting recruitment of mRNA during the initiation phase of translation (22).

Taken together, it seems that the rancRNAs identified so far are just a few examples of an unexplored class of translation modulators. More studies are necessary to investigate and better understand the principles and the outcomes behind the association of sncRNAs with the ribosome. However, it seems reasonable to assume that sncRNAs are capable of regulating translation by directly targeting the translational machinery.

Table 2.1. Summary of previously identified rancRNAs in different organisms and their mechanism of action.

rancRNA	Organism	Stress Condition	Mechanism	References
5' tRF-Val	*Haloferax volcanii*	Alkaline	Compete with mRNA binding	(14, 16)
18-mer RNA	*Saccharomyces cerevisiae*	Hyperosmotic	Compete for P-site occupancy	(18)
3' tRF-His 3' tRF-Leu 3' tRF-Gly 3' tRF-Ser 3' tRF-Thr 5' tRF-His	*Saccharomyces cerevisiae*	High pH, low pH, nutrient starvation, heat shock, cold shock, high salt, UV exposure, hypoosmotic, hyperosmotic, anaerobic	Ribosome-bound aminoacyl-transfer RNA synthetase association	(21)
rancRNA_194	*Haloferax volcanii*	Xylose carbon source	Unknown	(17)
3' tRF-Thr	*Trypanosoma brucei*	Nutrient starvation	Stimulate mRNA binding	(22)

2.6: What are we missing?

The ribosome is a universally conserved molecular machine with a crucial role for protein synthesis in all three domains of life (29, 99). Traditionally, the ribosome has been viewed as a dynamic entity with an integral, but passive, role during translation. This perspective overlooks recent studies

suggesting the ribosome itself as a novel hub in translation regulation (29, 100). Phosphorylation of ribosomal proteins, for example, may affect translational initiation. One eukaryotic example is the phosphorylation of the ribosomal protein S6, which stimulates translation of mRNAs by processing the 5' terminal oligopyrimidine sequence (101). Phosphorylation of ribosomal proteins also occurs in prokaryotes, with S7 and S11 being identified as important for readthrough of mRNAs (102). In addition, mounting evidence indicates the existence of distinct ribosomal subpopulations that exhibit heterogeneity in ribosomal composition in response to environmental signals. This provides a platform for increasing diversity in ribosomes leading to adaptation and functional specialization (103). For example, in *E. coli*, the stress-induced endoribonuclease MazF generates specialized ribosomes lacking the anti-Shine–Delgarno sequence through cleavage of the 3' end of the 16S ribosomal RNA (rRNA). These specialized ribosomes are selective for leaderless mRNAs that are also generated by MazF (104).

Given the important role that ribosomes play for protein synthesis and the recent evidence indicating that ribosomes are sncRNAs targets makes them a prime target for a highly sensitive and reversible regulation of gene expression. In addition, the small size of rancRNAs and their immediate availability provides them with an advantage as first-wave regulators during stress conditions. This raises the question as to whether a largely unexplored ancient and conserved mechanism of translation regulation exists. The fact that inhibition of protein synthesis by rancRNAs is strongly associated with stress conditions, and that rancRNAs are capable of inhibiting protein synthesis in both archaeal and eukaryotic cells, supports the idea of an evolutionarily conserved mechanism (Fig. 2.2) (14, 18, 20).

Surprisingly, this mechanism has not been shown in bacteria so far. Although bacteria, archaea, and eukaryotes have evolved different mechanisms of regulatory ncRNAs (Fig. 2.2), the overall design of their regulatory process is, however, intriguingly similar. Therefore, it is possible that because more complex organisms utilize this unique regulatory mechanism, bacteria might also use it in some form. Evidence to support such a conservation of this mechanism was reported by

Gebetsberger and colleagues in 2017. Using the archaeal Val-tRF in *S. cerevisiae* and *E. coli* model organisms, they demonstrated that inhibition still occurred across species, suggesting that this mechanism may be conserved across domains of life (16). Interestingly, studies on the ncRNA interactome in bacteria have identified sncRNAs to be enriched in polysomal and subpolysomal fractions, along with the identification of ncRNA–rRNA interactions (30-32).

Perhaps the best question to ask in this context is whether these sncRNAs participate in a conserved mechanism for translational control. Although some specific functions have been indicated, current research is still in the phase of identifying and validating ncRNAs with the potential to directly bind and regulate the ribosome, functional studies are limited so far. It will be an important challenge to investigate the biology of rancRNAs and provide insight into a completely novel layer of gene expression regulation. This is not only interesting from a fundamental perspective to understand the basic processes of life, but also from a synthetic biology and bioengineering approach. The rational design of gene expression circuits is at the heart of biotechnology, with the construction of artificial cells and increasing importance for commercial applications (105). Furthermore, the majority of clinically relevant antibiotics act during gene expression in one way or another, with approximately 50% of them targeting ribosome-dependent processes (106, 107). This highlights the potential for the development of novel antimicrobial strategies exploiting the rancRNA-dependent regulatory processes. This is of particular urgency, as the emergence of antibiotic resistance has been identified as a high priority by the World Health Organization (WHO), who endorsed a global action plan on antimicrobial resistance in May 2015. A part of their plan identifies the need to develop a sustainable plan to develop antimicrobials in which, we believe, rancRNAs may play a role.

2.7: Acknowledgements

This work was supported by funding from Natural Sciences and Engineering Research Council of Canada (NSERC) Undergraduate Student Research Award (to S.H.C.), and Alberta Innovates (to H.J.W.; Strategic Chairs Program, SC60-T2). We thank Emily Wilton for helpful comments on the manuscript, and Harland Brandon for helpful discussions.

CHAPTER 3: Investigation of annotated bacterial small RNAs functioning as ribosomal regulators

3.1: Introduction

The process of gene expression is regulated by various molecules. Focusing on translation, regulatory factors have been identified including proteins and nucleic acid (108, 109). Ribosomal proteins are known to use molecular mimicry, where two individual biological molecules share sequence and/or structural similarities, to repress translation of another ribosomal protein to maintain stoichiometric levels (110, 111). RNases also act to prevent the translation of faulty mRNAs. RNase R degrades RNA non-specifically at the 3' end, while RNase III typically degrades double-stranded RNA (112-115). sRNAs have been found to regulate mRNAs post-transcriptionally by base pairing to allow or inhibit translation or cause the mRNA to be degraded (67, 116, 117). Proteins and sRNAs can also work in conjunction with each other as regulatory RNA-protein complexes. For example, protein CsrA activity is regulated by the sRNAs csrB and csrC binding to high affinity sites on CsrA, where traditionally the sites would be bound by target mRNAs (118, 119). The binding of csrB and csrC by mimicking a target mRNA leads to sequestration of CsrA (118, 119).

Overall, RNA has been proposed to be advantageous as a regulator since RNA can have a tunable response to cellular or environmental signals (120, 121). In bacteria, RNA is approximately 125 times energetically cheaper for cells to produce compared to proteins (122, 123). RNA can also regulate gene expression at multiple levels. For example, sRNAs target numerous sites on mRNAs such as 5' and 3' UTRs, regions that transcription regulators do not target (124). Additionally, sRNAs can act at the transcriptional and post-transcriptional level to improve the speed of regulation. It has been shown that post-transcriptional down regulation is more efficient in response to external cues compared to signaling pathways and transcription regulators (125, 126).

Although sRNAs have been shown to be advantageous at various points during gene expression regulation, the molecular details of how sRNAs interact with the ribosome are poorly understood. More recent studies have identified ribosome-associated non-coding RNAs (rancRNAs) as direct regulators of

the ribosome (12). These rancRNAs have been identified in all domains of life but are only poorly characterized, especially in bacteria (16-26). Recently, SprF1 RNA has been classified as an attenuator of global translation in *Staphylococcus aureus*, but no other bacterial rancRNAs have been identified (26).

To better understand this class of sRNA in bacteria and to potentially identify previously overlooked rancRNAs, I have re-analyzed two previously published RNAseq datasets for sRNAs associated with ribosomes (30, 31). These datasets were obtained using two different techniques to isolate ribosome-associated RNAs for subsequent next generation sequencing. The first technique called GradSeq separated all RNA associated with ribosome and ribosomal subunits using glycerol density ultracentrifugation followed by next-generation sequencing (30). The second technique utilized *in vivo* UV cross-linking and proximity ligation to sequence RNA:RNA duplexes, where sRNA:rRNA interactions are identified (31). rancRNA candidates were identified based on their presence in both datasets after bioinformatic analysis and subsequently characterized *in silico* and experimentally. An overview of the characterization pipeline is shown in Figure 3.1.

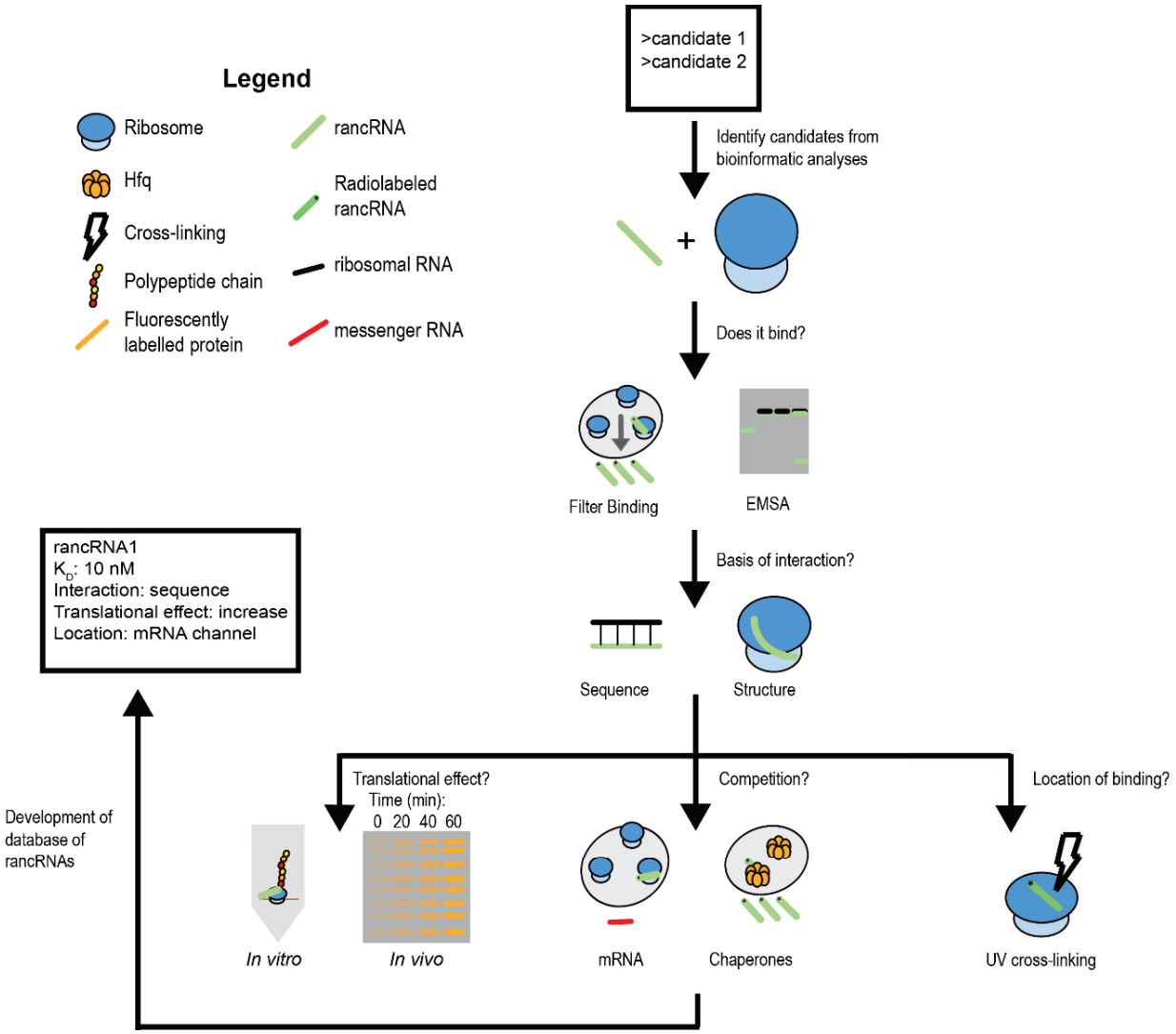


Figure 3.1. Overview of characterization methodology for rancRNA candidates *in silico* and experimentally. After candidate identification, binding to 70S ribosomes and/or 50S/30S ribosomal subunits *in vitro* will confirm its association with ribosomal particles. Subsequently, the candidate's mode of interaction will be predicted using a combination of *in silico* methods. Lastly, the mechanism of action will be investigated by examining the candidate's effect on translation, if the candidate binds to additional factors, and experimentally validating where the candidate binds on the ribosome. This experimental data will be compiled into a database together with functional information for all known rancRNAs.

3.2: Methods

3.2.1: Identification of rancRNA candidates

A literature search was conducted to identify bacterial RNA sequencing datasets in which ribosomes and/or ribosomal subunits were directly sequenced, or the experimental technique to obtain the RNA for sequencing resulted in an indirect association with rRNA. Initially, two datasets were identified for re-analysis with the first being a *Salmonella* Typhimurium Gradseq dataset (30). In brief, ribosomes and ribosomal subunits from *Salmonella* Typhimurium were separated using glycerol density ultracentrifugation and all ribosome-associated RNA contained in the respective fractions of the glycerol gradient were sequenced (30). The corresponding Fastq files were obtained from the NCBI GEO databank (accession number: GSE62988) and subsequently aligned to the *Salmonella* Typhimurium SL1344 genome (GenBank: FQ312003.1) using a Burrows Wheeler Alignment (127). Transcript abundance was quantified, and differential expression was computed using SeqMonk to identify rancRNA candidates.

The second dataset was generated using *E. coli* O157:H7 cells (31). Although the aim of this study was to identify sRNA:mRNA pairs using UV cross-linking and proximity ligation, sRNA:rRNA cross-links were also present in the sequencing dataset. sRNA:rRNA sequences were identified and extracted using Microsoft Excel (2016) for the second set of rancRNA candidates. Both sets of rancRNA candidates were subsequently compared and sRNAs were selected as rancRNA candidates if they were present in both lists. Subsequently, an additional study was found where RNA from ribosomes from *E. coli* O157:H7 cells were sequenced using RNAseq and RIBOseq providing additional information if the identified RNAs are translated by the ribosome (32). The original rancRNA candidate list was then compared to the sRNAs identified in the additional study to identify if any of the candidates contain open reading frames (ORFs) and may be translated.

3.2.2: *In silico* analysis of rancRNA candidates

rancRNA structure prediction:

rancRNA candidate secondary structure was predicted using RNAfold to obtain the dot bracket structure (settings: minimum free energy and partition function, avoid isolated base pairs) (128). The dot bracket structure and the respective sequence were entered into VARNA to generate the secondary structure representation (129). The tertiary structure was also predicted based on the RNA sequence and dot-bracket secondary structure predicted by RNAfold with the program SimRNA (default simulation steps (500), default % of the lowest energy frames taken to clustering (1) (130). Tertiary structures were visualized and further analysed using PyMOL (align function) (The PyMOL Molecular Graphics System, Version 2.0, Schrödinger, LLC).

Conservation of rancRNA candidates:

rancRNA candidate sequences were input into BLASTn with default settings (Standard databases, 100 max target sequences, expect threshold = 0.05, word size = 28, match/mismatch scores = 1-2, linear gap costs, filter low complexity regions, mask for lookup table only) (131). The resulting sequences were then aligned with Clustal Omega (settings: enter RNA sequence, ClustalW with character counts output) and a phylogeny tree was generated (132, 133).

Potential rRNA regions of interaction:

5S, 16S, and 23S rRNA sequences as well as the rancRNA candidate sequences were obtained from EcoCyc (134). To identify any possible base pairs between rRNA and rancRNA candidates, the rRNA sequence and the reverse strand of the rancRNA candidate were aligned using Clustal Omega (settings: enter RNA sequence, ClustalW with character counts output) (132, 133). sRNA:rRNA pairs identified in Waters *et al.* were also used to identify possible interaction regions based on the recovered sequences (31).

To further evaluate possible base pairing interactions, various software tools were utilized. IntaRNA (settings: number of interactions per RNA pair = 5, suboptimal interaction overlap – can overlap in query, no lonely base pairs, no GU at helix ends, minimum number of base pairs in seed = 7, ignore seeds with GU ends) and RNAUp (settings: avoid isolated base pairs, length of unstructured region = 4, maximal length of region of interaction = 25) were used to determine base pairing based on energy calculations (135-138). For both tools, default settings were used. The align function of PyMOL was used to generate a superimposition and RMSD calculation between the rRNA and rancRNA candidate (The PyMOL Molecular Graphics System, Version 2.0, Schrödinger, LLC).

3.2.3: RNA preparation and purification

Small RNA Synthesis:

The DNA sequences corresponding to the identified small RNAs were modified to include a T7 promoter and terminator for *in vitro* transcription and were synthesized and subcloned into pUC57 by Azenta/Genewiz (South Plainfield, New Jersey) (Table S3.3). Plasmid DNA was then transformed into *E. coli* DH5α cells and isolated using a BioBasic EZ-10 Spin Column Plasmid DNA Miniprep kit following the manufacturer's protocol.

Amplification of Small RNAs using Polymerase Chain Reaction:

Phusion polymerase in GC buffer (Thermo Scientific), dNTPs (BioBasic), and construct specific primers (IDT, Table S3.3) were used to amplify the region of interest from the respective pUC57 plasmid containing the small RNA sequence and a T7 promoter and terminator. Reaction conditions used for PCR are summarized in Table 3.1. Amplification was performed using a T Gradient Biometra thermocycler (Table 3.2) and reactions were subsequently restriction digested with DpnI (Thermo Scientific). To confirm the correct product size, PCR products were analyzed using 10% Native Polyacrylamide Gel Electrophoresis (PAGE), followed by staining with ethidium bromide (Thermo Scientific). All gels were visualized using the Amersham Imager 600 or Amersham Typhoon.

Table 3.1. PCR conditions for a 20 µL reaction using Phusion Polymerase.

Component	Final Concentration
5x Phusion GC Buffer	1x
10 mM dNTPS	200 µM (for each nucleotide)
Forward Primer	0.5 µM
Reverse Primer	0.5 µM
Template DNA	0.175 ng/µL
Phusion DNA Polymerase	0.02 U/µL
H_2O	

Table 3.2. PCR cycle conditions for rancRNA candidates.

Cycle Step	Temperature (°C)	Time (s)	Number of Cycles
Initial Denaturation	98	30	1
Denaturation	98	10	25
Annealing	49.7-60.1*	30	25
Extension	72	30	25
Final Extension	75	300	1

*Range of annealing temperatures varied for rancRNA candidates.

In vitro Transcription of PCR Product:

Correctly amplified PCR products were *in vitro* transcribed using T7 polymerase (in-house preparation), Transcription (Trab) buffer (200 mM Tris-Cl (pH 7.5), 75 mM $MgCl_2$, 50 mM NaCl, 10 mM sperimidine), DTT (Thermo Scientific), NTPs (Sigma Aldrich), GMP (Sigma Aldrich), iPPase (Sigma Aldrich), and RNase Inhibitor I (Thermo Scientific) for three hours at 37 °C. DNase I (Thermo Scientific) was added to the *in vitro* transcription reactions followed by a one-hour incubation at 37 °C. RNA was then purified using a BioBasic EZ-10 Spin Column RNA clean-up and concentration kit following manufacturer's instructions. To confirm the correct size of the RNA product, samples were analyzed on a 10% Urea PAGE, followed by staining with ethidium bromide (Thermo Scientific).

Table 3.3. *In vitro* transcription components for 100 µL reaction.

Component	Final Concentration
5x Trab Buffer	1x
100 mM DTT	10 mM
25 mM NTPs	3 mM
100 mM GMP	5 mM
0.5 U/µL iPPase	0.01 U/µL
T7 Polymerase	1.4 µM
40 U/µL RNase Inhibitor I	0.12 U/µL
DNA Template	10 ng/µL
H_2O	-

3.2.4: Nitrocellulose filter binding

RNA Radiolabeling:

Purified RNA product (300 nM to 3640 nM) was dephosphorylated using 1 U/µL shrimp alkaline phosphatase (New England BioLabs) for one hour at 37 °C and then heat deactivated for 10 minutes at 95 °C. 10 µCi of radiolabeled gamma ^{32}P-ATP (Perkin Elmer) was then added with 10 U/µL T_4 Polynucleotide Kinase (Thermo Scientific) for one hour at 37 °C. The resulting radiolabeled RNA was then purified using a BioBasic RNA clean-up kit and eluted using RNase-free water (New England BioLabs). Subsequently the RNA concentration was determined using UV Spectroscopy at 260 nm (Ultrospec 3000, Pharmacia Biotech) and the respective extinction coefficient calculated based on the respective RNA sequence (https://www.fechem.uzh.ch/MT/links/ext.html).

Nitrocellulose Filter Binding Assay:

2 000-4 000 nM stock Radiolabeled RNA (2 nM final concentration, previously eluted in RNase-free water) was heated at 95 °C for 10 minutes, followed by slow cooling to 37 °C. Concentrations of in-house purified 70S, 50S, 30S ribosomes and ribosomal subunits were determined with a BioDrop at 260 nm (MBI Lab Equipment) and their respective extinction coefficients (70S – 3.91×10^7 $M^{-1}cm^{-1}$, 50S – 2.55×10^7 $M^{-1}cm^{-1}$, 30S – 1.37×10^7 $M^{-1}cm^{-1}$) (139). Serial dilutions were made (0.034 µM to 2.4 or 3.2 µM, 20 µL volume) from the ribosome stock (10-20 µM) using 1x $TAKM_5$ buffer (50 mM Tris-HCl, 70 mM NH_4Cl, 30 mM Potassium Chloride, 5 mM Magnesium Chloride). 5 µL of 2 nM Radiolabeled RNA was then pipetted

into each ribosome dilution, followed by incubation for 15 minutes at 37 °C. After the incubation period, the RNA-ribosome mixture was immediately removed and pipetted (25 µL total volume) onto a 0.2 µm nitrocellulose membrane (0.2 µm, 25 mm diameter GE Healthcare Life Sciences), followed by washing of the membrane with 1 mL of ice cold 1x $TAKM_5$ buffer. The membrane was then transferred into a scintillation vial and dissolved in 10 mL of EcoLite scintillation cocktail (EcoLite (+), MP Biomedical). This procedure was repeated for all dilutions. Additionally, a scintillation vial containing 5 µL of radiolabeled RNA pipetted directly into 10 mL of scintillation cocktail was used to determine the maximum radioactivity for the RNA and a vial with no RNA to account for background radioactivity were analyzed for each trial. Vials were then vortexed for one minute and left at room temperature for approximately 30 minutes followed by a second vortexing (one minute) and the amount of RNA retained by binding to the ribosome was quantified by scintillation counting (Liquid Scintillation Analyzer Tri-Carb 2810 TR, Perkin Elmer). Prism GraphPad version 9 was used for visualisation and fitting with a one-site binding equation ($y = Bmax*x/(K_D+x)$) to extract the equilibrium dissociation constant (K_D) and the fraction of RNA bound to the ribosome (Bmax).

3.2.5: *In vitro* translation

Using the PURExpress® system (New England Biolabs), 100 ng/µL of pUC57 coding for enhanced cyan fluorescent protein (eCFP) was translated. To ensure all ribosomes in the PURExpress® system (ribosome content is 60 picomoles) would be targeted, 60 picomoles of the rancRNA candidates were added. The experiment was conducted under two conditions: (1) eCFP template and rancRNA were added at the same time and (2) ribosomes were pre-incubated with the rancRNA where pUC57 encoding for eCFP was added after the rancRNA candidate was incubated with the PURE components for 15 minutes at 37 °C. A negative control with no template and a control experiment containing only template and no rancRNA were included. The respective reaction mixtures were then incubated for 16 hours at 37 °C. The resulting eCFP fluorescence was analysed using a QuantaMaster Fluorimeter (Photon Technology International (Canada) Inc). eCFP was excited at 439 nm and the emission scan was measured from 454 to 554 nm. Obtained emission spectra were subsequently analyzed using Microsoft Excel (2016) after the negative control was subtracted from the respective experiment control and reactions containing the rancRNA

candidates. A one-tailed t-test was used to determine statistical significance using Prism GraphPad version 9.

To ensure the RNA added during the pre-incubation experiments remained intact with the additional incubation time, the PURE reactions containing rancRNA candidates were reverse transcribed using SuperScript IV (Invitrogen), following the manufacturer's protocol. In brief, the corresponding reverse primer for the rancRNA candidate, dNTPs, 500 ng of PURE reaction, and H_2O were combined in a microcentrifuge tube and heated at 65 °C for 5 minutes, and subsequently incubated on ice for 1 minute (Table 3.4). The remaining reaction components listed in Table 3.4 were added to the microcentrifuge tube. The final reaction was then incubated at 55 °C for 10 minutes, followed by an inactivation step with incubation at 80 °C for 10 minutes. cDNA product was amplified using Phusion polymerase (Thermo Scientific). The experiment was set up as detailed earlier and according to Table 3.1 and the PCR cycle as described in Table 3.2.

Table 3.4. Reverse transcription components for 20 µL reaction.

Component	Final Concentration
2 µM gene-specific reverse primer	0.1 µM
10 mM dNTPs	0.5 mM
100 mM DTT	5 mM
5x SSIV Buffer	1x
SuperScript® IV Reverse Transcriptase (200 U/µL)	10 U/µL
40 U/µL RNase Inhibitor I	2 U/µL
RNA Template	500 ng
H_2O	-

3.2.6: *In vivo* translation

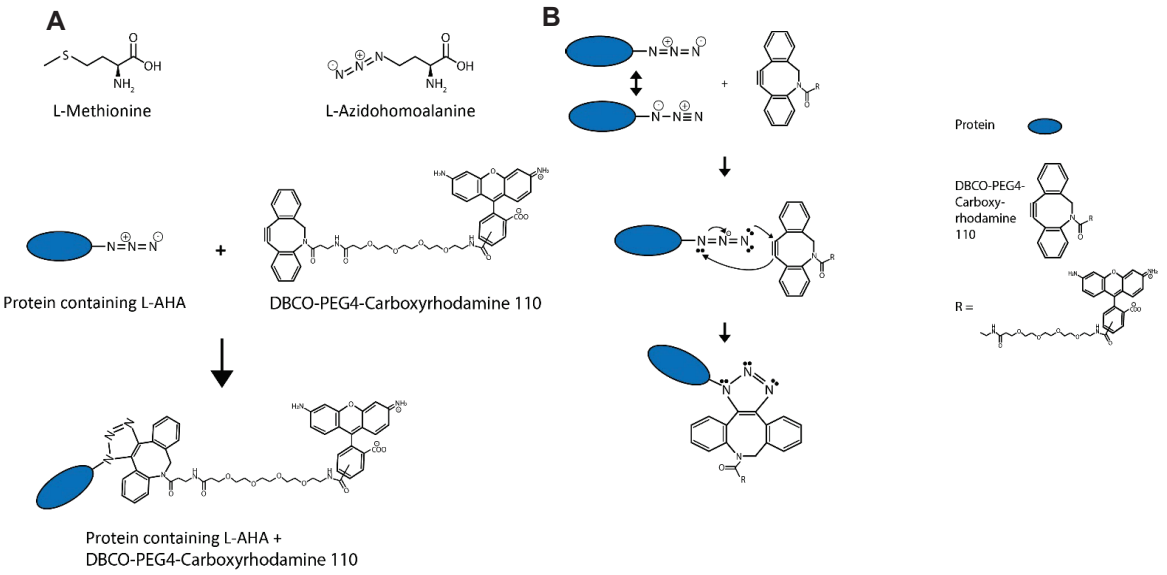


Figure 3.2. Overview of Biorthogonal non-canonical amino acid tagging (BONCAT) assay. **(A)** L-azidohomoalanine (L-AHA) is used in place of L-methionine and is incorporated into *E. coli* proteins. These proteins can then be labelled using dyes containing an alkyne bond. **(B)** Overview of the "Click chemistry" approach based on the reaction between the azide and alkyne bond used to incorporate the labelling dye (DBCO-PEG4-Carboxyrhodamine-110).

To measure *de novo* protein synthesis in live *E. coli* cells, an assay adapted from Hatzenpichler and colleagues (140) was used (Figure 3.2). In brief, a 50 mL LB culture of *E. coli* BL21(DE3) cells transformed with pACYCDuet-1 encoding eCFP tagged with the RNA Mango III aptamer and pUC57 encoding one of the rancRNA candidates were grown at 37 °C and 220 rpm overnight (Max^{Q} 4000 Barnstead Lab-Line). Cells were pelleted at 5 000xg for 15 minutes at 4 °C (Heraeus Multifuge X3R Centrifuge, Thermo Scientific) and subsequently washed once using M9 salts minimal medium (5x M9 salts, 1 M $MgSO_4$, 20% glucose (w/v), 1 M $CaCl_2$) at 5 000xg for 15 minutes at 4 °C (Heraeus Multifuge X3R Centrifuge, Thermo Scientific). After, cells were added to 50 mL of M9 salts minimal medium at an OD_{600} of 0.1 and grown in tandem at 37 °C with shaking at 220 rpm (Max^{Q} 4000 Barnstead Lab-Line). At an OD_{600} of 0.2, 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to one flask and a second flask was induced at an OD_{600} of 0.4. When all three flasks reached an OD_{600} of 0.4, 100 mM L-Azidohomoalanine (L-AHA) was added to a 10 mL subculture of the respective 50 mL starting culture. Two 0.5 OD_{600} samples

were removed from each subculture every 20 minutes up to 60 minutes and stored on ice. Cells were pelleted at 16 000xg for five minutes (Sorvall Legend Micro 17R, Thermo Scientific) and re-suspended in 1x PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4) with 3% (w/v) formaldehyde. The cells were subsequently fixed for 1 hour at room temperature, pelleted by centrifugation at 16 000xg for five minutes (Sorvall Legend Micro 17R, Thermo Scientific), and the supernatant was decanted. Following this, cells were washed three times with 1x PBS by centrifugation at 16 000xg for five minutes (Sorvall Legend Micro 17R, Thermo Scientific), and stored in 1x PBS containing 50% ethanol at -20 °C.

Cells were subsequently pelleted and re-suspended in the BONCAT protein extraction buffer (1% (w/v) sodium dodecyl sulfate, 50 mM Tris, pH 8.4, 150 mM NaCl, 100 mM EDTA, 1 mM $MgCl_2$) and boiled for 30 minutes, followed by snap cooling on ice for five minutes. Samples were then centrifuged at 16 000xg for five minutes (Sorvall Legend Micro 17R, Thermo Scientific). The supernatant was transferred to a new microfuge tube, and 100 mM 2-chloroacetamide (Sigma Aldrich) (50 µL of 1 M solution) was added, followed by incubation in the dark for one hour at room temperature with shaking. 10 mM DBCO-PEG4-Carboxyrhodamine 110 (Click Chemistry Tools) was added and incubated in the dark at room temperature with shaking for an additional 30 minutes. Proteins were then extracted using a 6:1.5:4 mixture of methanol:chloroform:water and washed three times with 100% methanol by centrifuging at 16 000xg for five minutes (Sorvall Legend Micro 17R, Thermo Scientific). The protein pellet was then dried using the Savant speedvac vacuum concentrator or air dried in a fumehood.

The resulting dried protein pellets were then re-suspended in 1 M DTT, 6x SDS loading buffer without dye, and 8 M urea. The resulting protein concentration was determined using A_{280} and A_{260} absorbance obtained by UV spectroscopy (Ultrospec 3000, Pharmacia Biotech). Samples were then analyzed on a 12% SDS PAGE at 80V for 20 minutes followed by 200V for 45 minutes in the dark with a prestained protein ladder (Frogga Bio). The labelled proteins were then visualized using an Amersham Typhoon (excitation at 488 nm and the Cy2 525BP20 emission filter) and subsequently stained with Coomassie Blue overnight. The gel was then scanned using the optical densitometry setting on an Amersham Typhoon. Densitometry analysis was performed for proteins in the range of 100 kDa and 35

kDa using ImageJ (141). Results were further analysed using Microsoft Excel (2016) and GraphPad Prism v9. Dunnett's test was used for statistical analysis (142).

To ensure the rancRNA candidate was transcribed during cell growth following induction by IPTG, total RNA was isolated using the EZ-10 Spin Column Total RNA Miniprep kit (BioBasic) according to manufacturer's instructions. Isolated RNA was subsequently digested with 0.5 U/µL DNAse I (Thermo Scientific) for 30 minutes at 37 °C, followed by the addition of 1 µL 50 mM EDTA and incubation at 65 °C for 10 minutes. This RNA was then reverse transcribed using SuperScript IV (Invitrogen), following the manufacturer's protocol as described earlier (Table 3.4). cDNA product was amplified using Phusion polymerase (Thermo Scientific). The experiment was set up as detailed earlier, according to Table 3.1, and the PCR cycle as described in Table 3.2.

3.3: Results

3.3.1: rancRNA candidate identification

rancRNA candidates were identified using the two previously published bacterial datasets (30, 31), The workflow approach is illustrated in Figure 3.3, A. Using data reported in Smirnov *et al.* (30), 54 sRNAs were identified in *Salmonella* Typhimurium that are also present in *E. coli* (Table S3.1). 51 sRNAs were identified by filtering for sRNA-rRNA interactions present in pathogenic *E. coli* from Waters *et al.* (Table S3.1) (31). Correlating the two lists with each other, the following was taken into consideration when selecting the rancRNA candidates: abundance in fractions 10-20 of the glycerol gradient where 70S ribosomes and 50S/30S ribosomal subunits were isolated (30), the number of unique hybrids that map to the rRNA:sRNA interaction (termed count total), the Gibbs free energy of the rRNA:sRNA interaction, and false discovery rate of rRNA:sRNA cross-links (31). A total of six candidates (arcZ, omrB, rprA, rttR, rybA, and tpke11) were identified in *E. coli* (Table 3.4). Of these six candidates, five were also present in the Neuhaus *et al.* data set (Figure 3.3, B) (32). rttR is processed from the tRNA *tyrT* operon, where the 5' end of rttR is directly adjacent to the 3' processing site for tyrV (Figure S3.2) (143, 144). As other rancRNAs are

known to be tRNA-derived fragments, we chose to focus on characterizing rttR because of the proximity of the processing site to $tRNA^{Tyr}$.

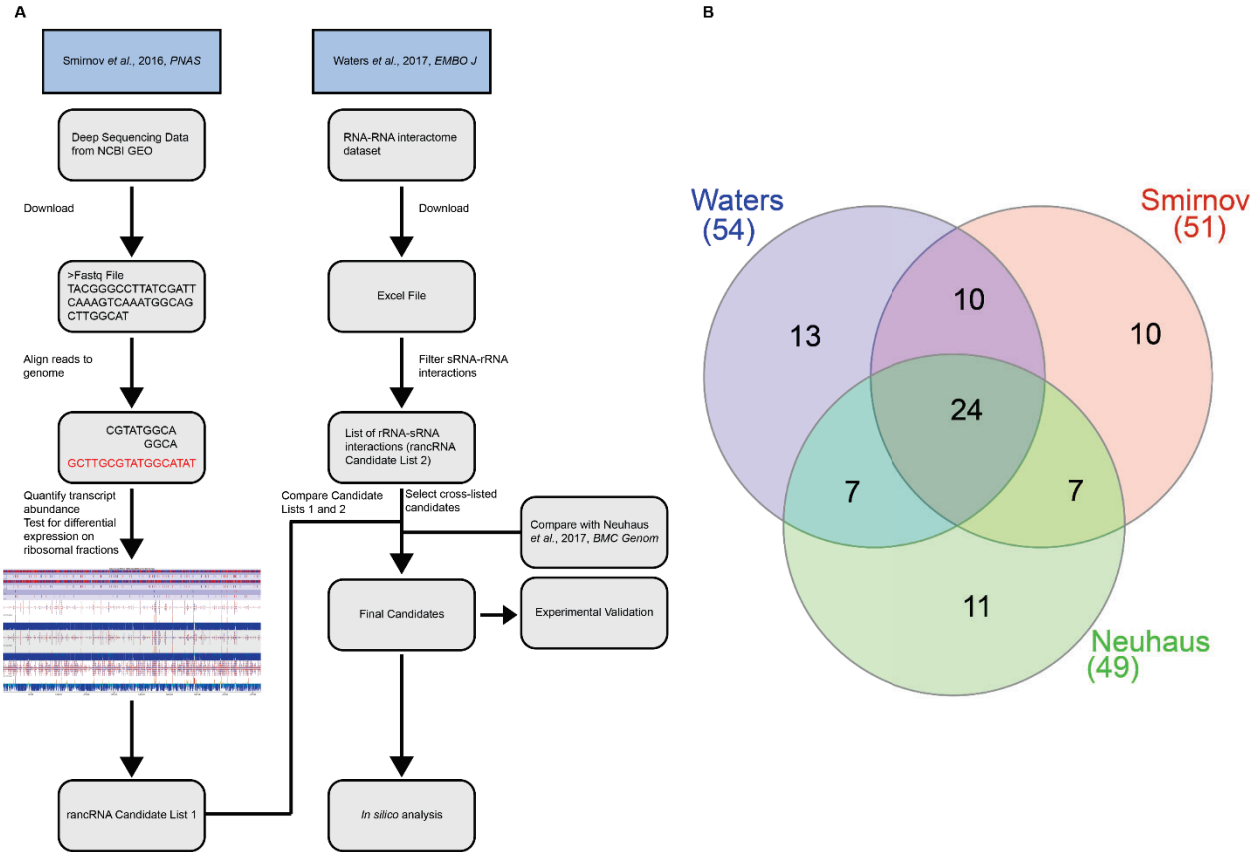


Figure 3.3. Identification of rancRNA candidates. **(A).** Workflow to identify candidates. Data from Smirnov *et al*. was aligned to the *Salmonella* Typhimurium genome and transcript abundance was quantified for sRNAs using SeqMonk (30). Data from Waters *et al*. was filtered for sRNA-rRNA interactions (31). Candidates were compared between the two datasets and ones present in both were selected for validation. This list was then compared to sRNAs identified in a study by Neuhaus *et al*. to assess if any of these contain ORFs and are translated by the ribosome(32). **(B)**. Overlap of sRNAs identified in the three studies analyzed (145).

Table 3.5. List of rancRNAs identified.

rancRNA Candidate	Length (nts)	Role in *E. coli*	References
arcZ	122	Increases translation of sigma factor 38 (RpoS); inhibits Rho-dependent transcription termination	(146-150)
omrB	84	Regulation of outer membrane protein composition	(151)
rprA	107	Produce RpoS during osmotic shock	(152, 153)
rttR	173	Processing product from *tyrT* operon	(143)
rybA	91	Regulation of aromatic acid synthesis under peroxide stress	(154)
tpke11	91	Phantom gene within heat shock protein locus	(32, 155)

3.3.2: *In silico* Analysis of rttR

To investigate if certain regions of rttR are likely to be of structural or functional importance, I focused on the sequence conservation of rttR among different rttR containing Enterobacteria using BLASTn and CLUSTAL omega (Figure 3.4) (131-133). These revealed that the rttR sequence is conserved amongst the Enterobacteriaceae family, primarily *Escherichia, Shigella, Salmonella, Klebsiella,* and *Citrobacter* (Figure 3.4, A). Percent identity for all alignments ranged from 89% to 100% (Figure 3.4, B). The most closely related sequences were found in the genera of *Escherichia*, *Shigella,* and *Salmonella* with sequence identities of 100% (Figure 3.4. B). Although a functional role of rttR is not well characterized in bacteria, the conservation of rttR's sequence amongst Enterobacteria shows promise that rttR is indeed a viable rancRNA candidate.

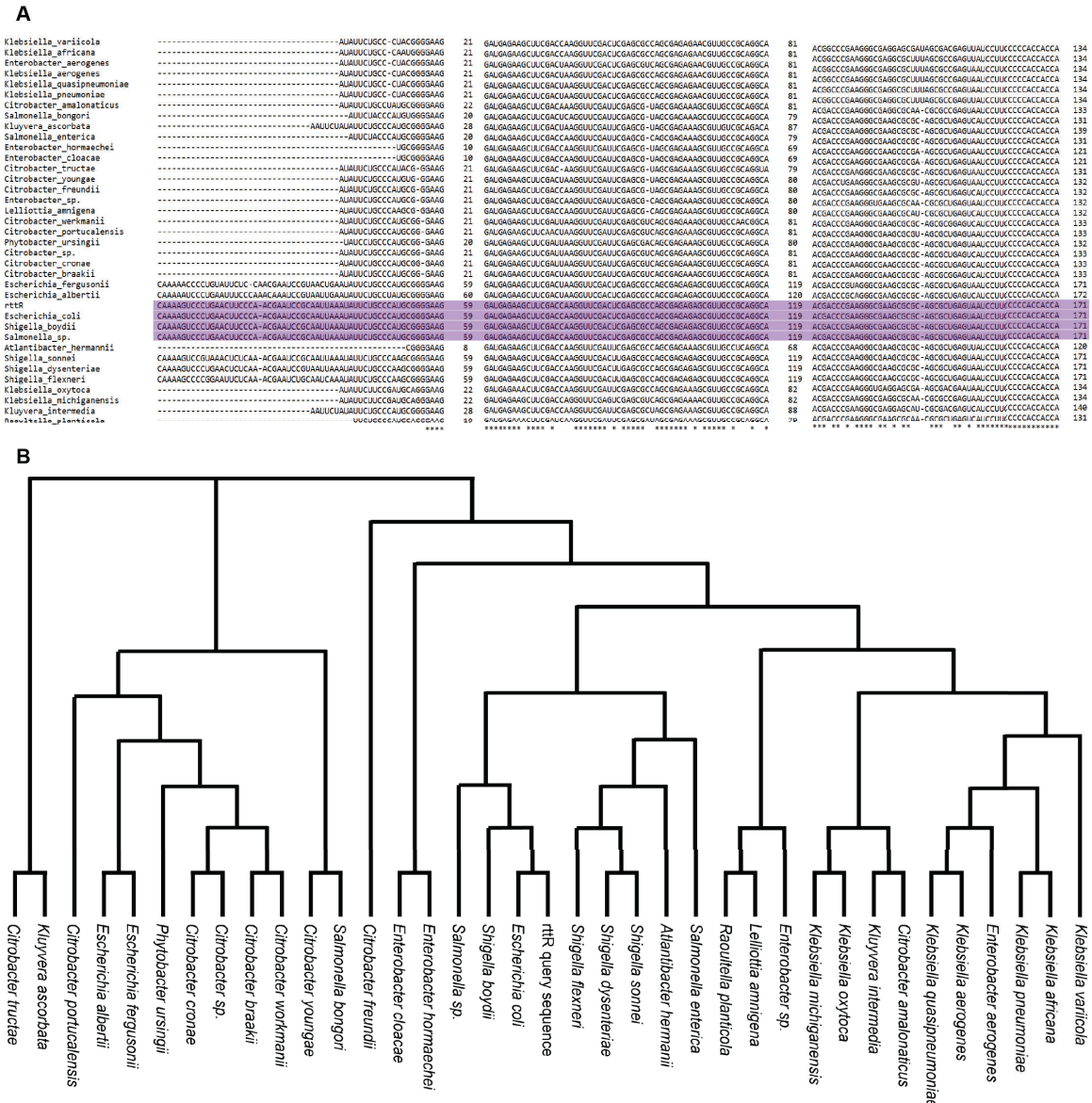
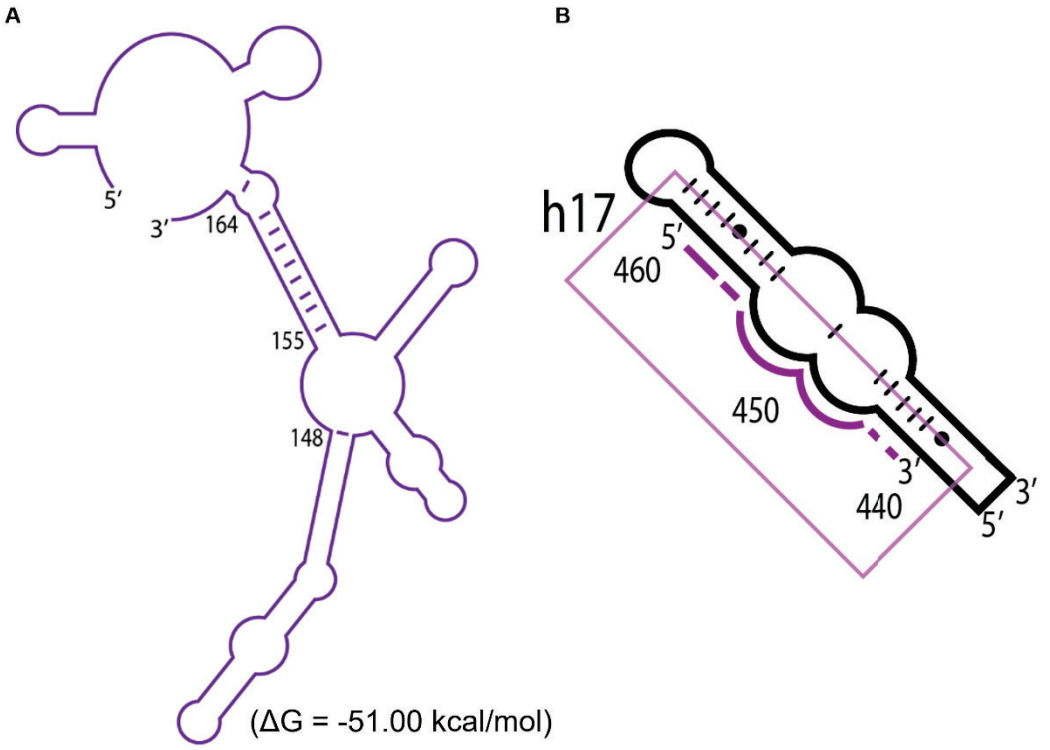


Figure 3.4. rttR is conserved in Enterobacteriaceae. **(A).** Clustal Omega alignment of sequences obtained through a BLASTn search reveals sequence conservation of rttR (131-133). Highlighted in purple are rttR sequences with 100% sequence identity. **(B).** Phylogenetic tree showing evolutionary relatedness of rttR (132, 133).

To better understand the structure of rttR, several secondary and tertiary structure predictions were completed (Figure 3.5). The predicted secondary structure using RNAfold (128) and VARNA (129) indicated that rttR is highly structured with over 90 nucleotides out of its 178 nts being involved in base pairing (Figure 3.5, A, Figure S3.1). I further explored if rttR is likely to bind to the ribosome by analyzing the regions of rRNA that rttR cross-linked to and where rttR is likely to align using Clustal Omega, IntaRNA, and RNAUp

(132, 133, 135-138). Based on the dataset from Waters *et al.*, one cross-linked region is helix 17 in the 16S rRNA 5' domain (Figure 3.5 B) (31). Manual alignment using Clustal Omega of the reverse sequence of rttR to helix 17 indicates the potential for base pairing, which was previously experimentally validated (Figure 3.5, B, C) (31). Interestingly, after visualizing this interaction using PyMOL, a structural alignment including all atoms was generated yielding an RMSD of 3.5 Å, indicating that rttR superimposes well on helix 17 (Figure 3.5, D). Additional *in silico* analyses are found in Table S3.2. Taken together, rttR may bind 16S rRNA helix 17.



C

rttR	3’	— 171	ACCACCA CCCCCUUCCUA AUGAGU	148 —	5’
16S rRNA	5’	— 441	AGCGGGGAGGAAGGGAGUAAAGU	463 —	3’

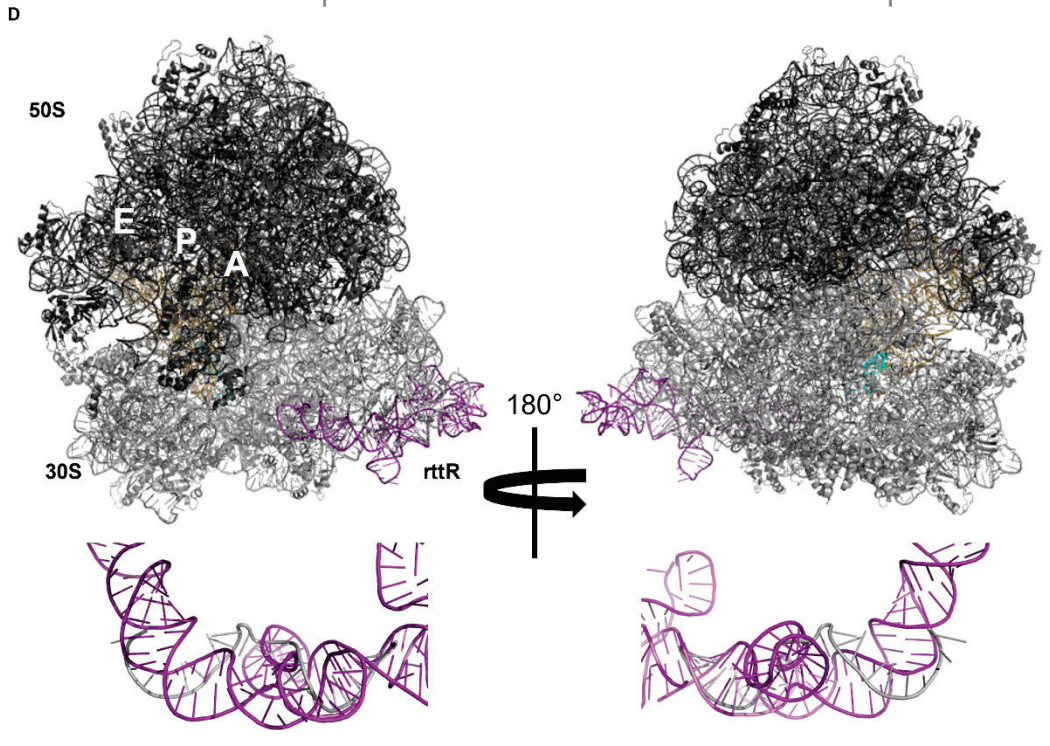


Figure 3.5. rttR is predicted to bind to 16S rRNA helix 17. **(A).** Secondary structure of rttR predicted by RNAfold and VARNA with nucleotides 148, 155-162, and 164 highlighted as predicted interacting nucleotides from Waters *et al.* (31, 128, 129). **(B).** 16S rRNA secondary adapted from Noller structures (http://rna.ucsc.edu/rnacenter/ribosome_images.html) with nucleotides 441-463 highlighted as predicted interacting nucleotides (31, 156). The purple box indicates the side of helix 17 and rttR interaction. Shown in purple are the nucleotides predicted to interact with rttR. **(C).** Alignment of rttR and helix 17 sequences predicted by Clustal Omega (132, 133). Predicted base pairs are shown by black lines between the respective nucleotides. **(D).** 70S ribosome (PDB: 6bu8) with tertiary structure of rttR aligned using PyMOL. Beneath the 70S is an enlarged view focused on the alignment of the interacting nucleotides of rttR (purple) and the 16S rRNA (light grey).

3.3.3: rancRNA candidates preferentially bind to ribosomes and ribosomal subunits

To determine if the rancRNA candidates selected indeed bind to ribosomes and/or ribosomal subunits, nitrocellulose filter binding was performed (Figure 3.6, A, *in vitro* transcribed RNA shown in Figure S3.3). As a binding control, $tRNA^{Phe}$ was used which binds 70S ribosomes with an affinity of 99 nM, consistent with the previously reported affinity of ~83 nM (157). As expected, no binding was observed with individual subunits as the tRNA binding sites are only fully formed when the subunits associate (158). The candidates arcZ, omrB, rprA, rttR, and rybA are capable of binding to the 70S with affinities ranging from 17 nM to 117 nM (Figure 3.6, B, Table 3.5). Interestingly, binding affinity differences were observed for some candidates between ribosomes and ribosomal subunits, but not all. ArcZ and rybA both have higher affinities for the 70S (93 nM and 101 nM, respectively) but weaker binding for the 30S ribosomal subunit for arcZ (273 nM) and 50S ribosomal subunit for rybA (195 nM) (Figure 3.6, Table 3.5). RprA also has a higher affinity for the 70S (28 nM), but all dissociation constants are below 100 nM (50S – 82 nM, 30S – 78 nM) (Table 3.5). Out of all the candidates rancRNAs, omrB is the only candidate that binds all ribosomal particles with high affinity (70S – 17 nM, 50S – 37 nM, and 30S – 19 nM) (Figure 3.6, Table 3.5). Interestingly, rttR binds to the 70S (117 nM) but neither ribosomal subunit (Figure 3.6). Here, I confirm that five of the six rancRNA candidates (arcZ, omrB, rprA, rttR, and rybA) are able to bind to 70S ribosomes and/or 50S/30S ribosomal subunits with nanomolar affinity that is comparable or higher than $tRNA^{Phe}$.

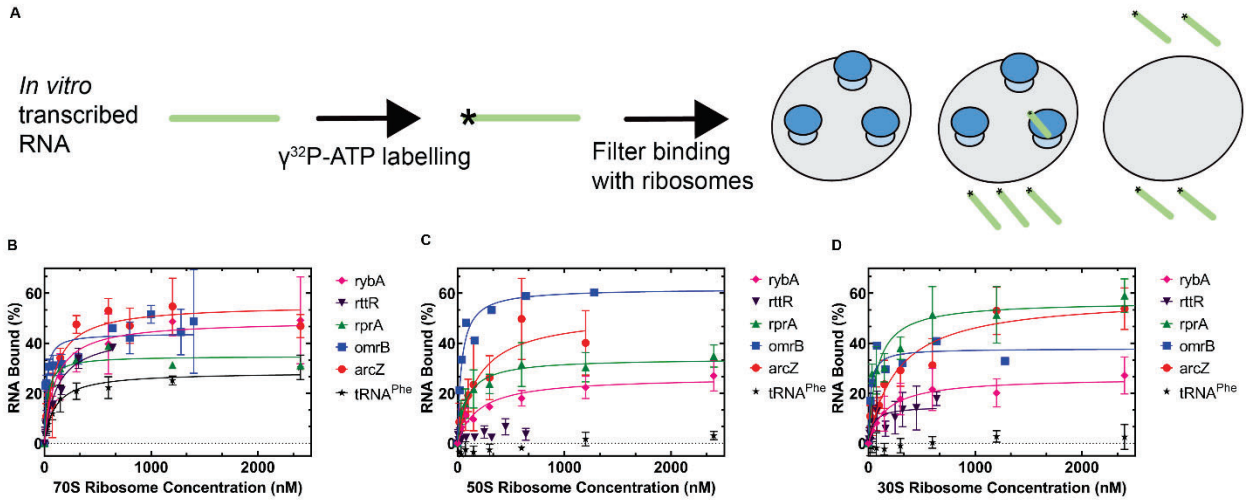


Figure 3.6. Binding of rancRNA candidates to ribosomes and ribosomal subunits. **(A)**. Nitrocellulose filter binding assay. *In vitro* transcribed RNA was labelled using γ^{32}P-ATP. A set concentration of radiolabelled RNA was incubated with varying concentrations of 70S, 50S, or 30S ribosomes,and subsequently pipetted onto a nitrocellulose membrane. Following a washing step, ribosomes and ribosomes with any bound RNA remain on top of the filter and free RNA is washed away. The retained radioactivity is then quantified using scintillation counting. **(B).** Binding of candidates to 70S ribosome. All candidates bind to the 70S with similar or tighter binding compared to the control. **(C).** Binding of candidates to the 50S ribosomal subunit. All candidates bind to the 50S except for rttR. **(D).** Binding of candidates to the 30S ribosomal subunit. All candidates bind to the 30S except for rttR. For all, tRNA^{Phe} was used as a binding control.

Table 3.6. Apparent dissociation constants determined for rancRNA candidates (n=3 ± SD, * denotes n=1).

Ribosome/Ribosomal Subunit Dissociation Constant (nM)	arcZ	omrB	rprA	rttR	rybA	tRNA^{Phe}
70S (nM)	93 ± 22	17 ± 5	28 ± 9*	117 ± 23	101 ± 29	99 ± 20
50S (nM)	188 ± 77	37 ± 8*	82 ± 21	No binding	195 ± 51	No binding
30S (nM)	273 ± 69	19 ± 8*	78 ± 21	43 ± 26	164 ± 55	No binding

3.3.4: Assessing rancRNA candidate effects on translation

In order to determine if any of the rancRNA candidates have a direct effect on translation and to minimize the influence of other factors present in the cellular environment, an *in vitro* coupled transcription-translation system (PURExpress®) was used (159) (Figure 3.7, A, *in vitro* transcribed RNA shown in Figure S3.3). Translation of eCFP encoded on pUC57 using the PURExpress® system allows for protein synthesis to be monitored by measuring the endpoint fluorescence of eCFP. A control of eCFP being produced by the PURExpress® system serves as the baseline fluorescence for eCFP (Figures 3.7, B, C, PURE). In the presence of a rancRNA candidate, any increases or decreases in fluorescence compared to the control

would indicate that a particular rancRNA candidate is capable affecting the production of eCFP. To determine if this effect is specific to the sequence of the used candidate RNA, a scrambled version of the respective RNA sequence was also included and had no effect on eCFP fluorescence (Figure 3.7, D, E). When incubating the template and candidate rancRNA concurrently with the rest of the PURE components, arcZ, rprA, rttR, and rybA did not influence eCFP fluorescence (Figure 3.7, B, C). However, addition of omrB resulted in increased fluorescence. Incubating the rancRNA candidates with the PURE components prior to adding the eCFP template, no significant increases or decreases on the fluorescence of eCFP could be observed (Figure 3.7, D, E, Figure S3.4). Therefore, the *in vitro* translation assay is functional. However, in the presence of the rancRNA candidates, there are minimal effects on protein production, with the exception of omrB.

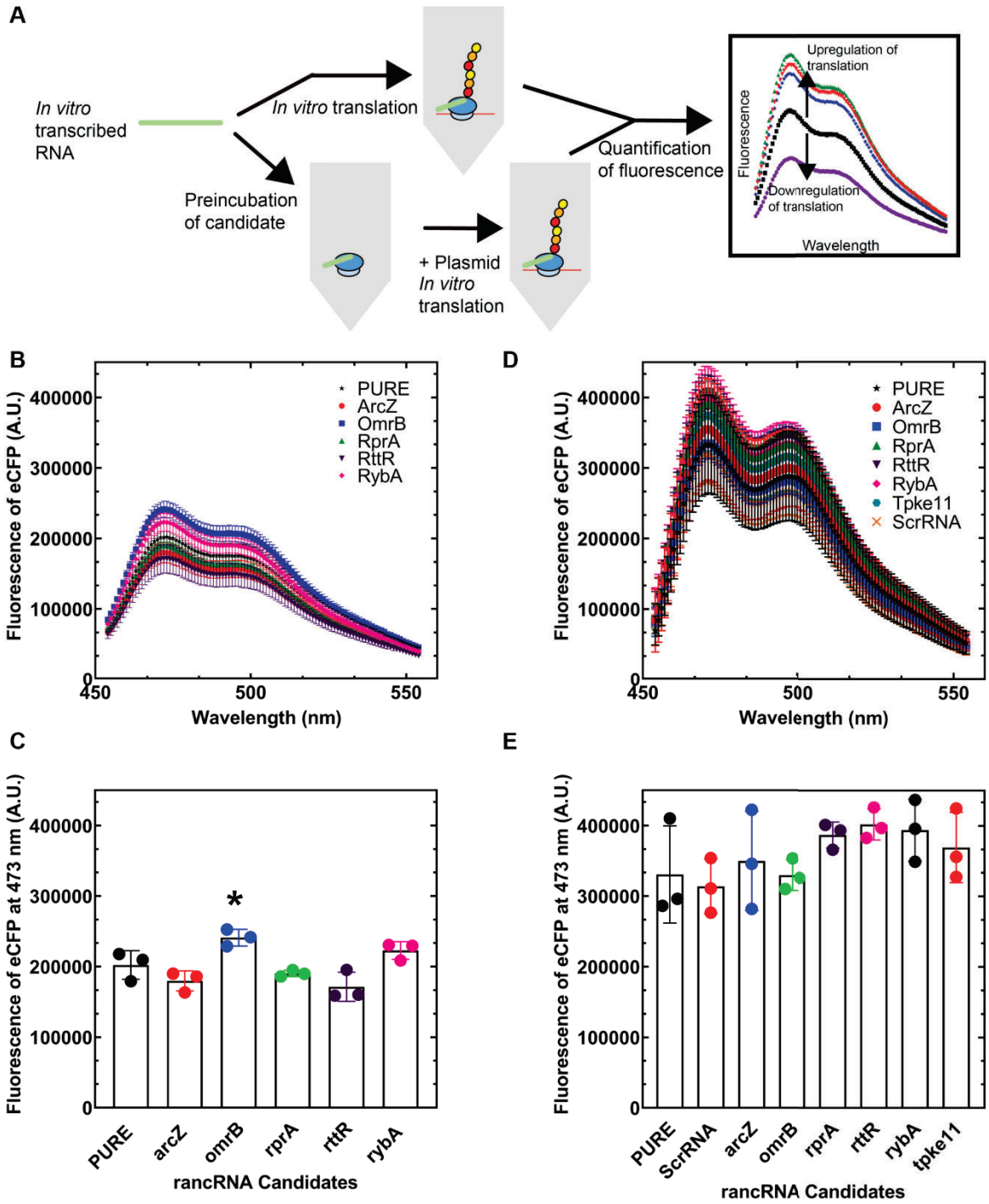


Figure 3.7. rancRNA candidates, except for omrB, have minimal effects on translation *in vitro.* **(A).** Overview of *in vitro* translation assay. Effects on the fluorescence of eCFP produced using the PURExpress system with the addition of a rancRNA candidate in a 1:1 ratio to ribosomes (2.4 μM). Candidates were either added at the same time as the plasmid encoding eCFP or pre-incubated for 15 minutes at 37 °C with all PURE components followed by addition of the plasmid. **(B, C).** Candidate rancRNAs were added at the same time as the eCFP template (n=3 ± SD, * - p = 0.0181) or **(D, E)** pre-incubated with the PURE components prior to adding the eCFP template (n=3 ± SD).

To determine if rttR's limited effects on translation *in vitro* can also be seen *in vivo*, I used BONCAT (140) (Figure 3.2). Here *de novo* protein synthesis can be monitored by incorporating the L-Methionine analogue L-AHA during cell growth and extracted proteins can be labelled with the fluorescent dye DBCO-carboxyrhodamine-110 through an alkene-alkyne reaction. By transforming *E. coli* BL21(DE3) cells with pUC57 encodes rttR, rttR transcription is controllable by IPTG induction. To mimic the *in vitro* translation assays, IPTG induction of the rancRNA candidates occurred at two different optical densities (OD_{600} of 0.2 and 0.4), while L-AHA in both cases was added at an OD_{600} of 0.4.(160). If rttR does indeed influence global translation, one would expect to see either an increase or decrease in the fluorescence of the labelled protein content when rttR transcription is induced. To determine if this effect is independent of the RNA sequence being transcribed, a control experiment utilizing a scramble RNA was conducted as well (Figures S3.6, S3.7, S3.8 B, Tables S3.6, S3.7). Overall, I do not observe any effects on translation under any of three conditions, except for the 40-minute time point without IPTG induction (Figure S3.6 D). However, one biological replicate appears to be skewing the average, resulting in significant changes at this point between conditions (Figure S3.6 D, Tables S3.6, S3.7). Thus, it is likely that the control is effective in showing that a randomized sequence of RNA is not impacting global translation.

Cellular growth during IPTG induction and L-AHA incorporation was monitored, revealing a similar rate of growth when no IPTG was added or if IPTG was added at an OD_{600} of 0.4 (Figure 3.8, C). However, there is a slight delay in growth when IPTG was added at an OD_{600} of 0.2, which is expected (Figure 3.8, C (160)). Samples for protein extraction and fluorescent labelling were taken just prior to adding L-AHA (0-minute time point) and every 20 minutes following the addition of L-AHA up to 60 minutes. The extracted and labelled proteins from the time course and three tested conditions were subsequently analyzed via SDS PAGE (Figures 3.8, A, B, Figure S3.5). The transcription of rttR was also confirmed by RT PCR (Figure S3.8 A). For all conditions, the 0-minute time point (no L-AHA added) is lower compared to the rest of the time points taken (Figure 3.8, A). This is expected as the chemical reaction between DBCO-Carboxyrhodamine-110 dye and L-AHA is unable to occur. Qualitatively, there are minimal differences when comparing the 20-, 40-, and 60-minute time points for the No IPTG and OD_{600} 0.4 conditions (Figure

3.8, A). For the OD_{600} 0.2 condition, the fluorescence intensity for the 40- and 60-minute time points increase compared to 20 minutes (Figure 3.8, A). Figure 3.8, B shows no significant differences in protein levels after Coomassie staining, indicating the amount of protein loaded was equal for all samples.

To assess differences more quantitatively between the tested conditions, band intensities at 100 kDa and 35 kDa were analyzed by densitometry (Figure 3.8, D). Comparing the differences following rttR transcription induction, a statistically significant decrease in intensity for both 20-minute time points compared to the 20-minute time point with no IPTG induction can be observed (Table S3.4). However, the time course in the absence of IPTG is inconclusive because the 0- and 60-minute time points show no significant difference (Table S3.4). No significant differences are found between any time points for either induction condition when compared to no IPTG induction (Table S3.4). Within each condition, there is a statistically significant increase for the 20-minute and 40-minute time points when compared to the 0-minute time point for the no IPTG condition and an increase for the 40-minute time point compared to the 0-minute time point when IPTG was added at 0.2 OD_{600} (Table S3.5). However, the standard error of means between the three biological replicates makes it difficult to determine if this increase suggests an upregulatory effect (161). No statistically significant changes are observed for when IPTG was added at 0.4 OD_{600} (Table S3.5). This indicates that rttR may not be able to be transcribed at a fast enough rate to determine any effects when IPTG and L-AHA are added simultaneously (160). Altogether, rttR likely does not impact global translation.

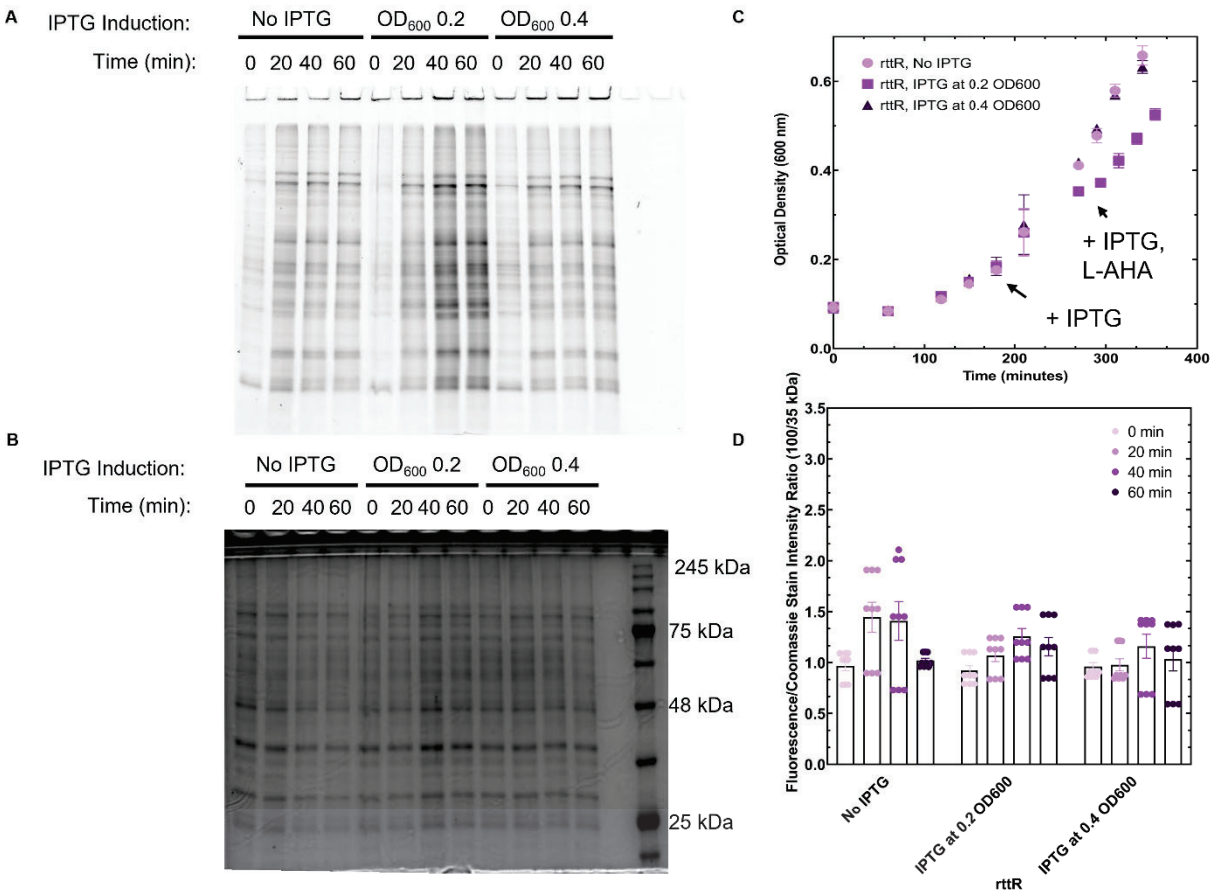


Figure 3.8. *In vivo* monitoring of global *de novo* protein synthesis. **(A).** 12% SDS PAGE of fluorescently labelled proteins. **(B).** Same SDS PAGE stained with Coomassie blue. **(C).** Growth curve of *E. coli* BL21(DE3) cells transformed with plasmid encoding rttR (n=3 ± SEM). **(D).** Normalized fluorescence of the 100 and 35 kDa bands identified in panel A and B (n=3 ± SEM).

3.4: Discussion

rancRNAs have been discovered in all domains of life (14, 16-26). To gain a better understanding of this RNA class in bacteria, I re-analyzed three reported RNAseq datasets to identify potential rancRNA candidates in *E. coli* (Table S3.1) (30-32). Overall, I identified six candidates (arcZ, omrB, rprA, rttR, rybA, and tpke11) (Figure 3.3, Table 3.4). From these, arcZ, omrB, and rprA have previously been reported to interact with Hfq, a known RNA chaperone that facilitates interactions been ncRNA and their target mRNA (147, 151, 152). While Hfq is known to be involved in 30S subunit maturation, it is unknown if there are any interactions where the ribosome, Hfq, sRNA, and target mRNA form a complex (162, 163). In *H. volcanii*, rancRNA_194 was reported to bind to the ribosome and regulate a specific mRNA involved in xylose

metabolism (17). Furthermore, TBsRNA37 is capable of simultaneously binding to the ribosome and target mRNA in *T. brucei* (23). This suggests that rancRNAs may be able to bind to multiple targets as part of their functional role, increasing the regulatory complexity of the mechanism of action.

Tpke11 is a putative ncRNA, that Neuhaus *et al.* have suggested is translated based on RIBOseq data (32). However, Pinel-Marie *et al*. showed that SprF1 RNA, which contains a putative open reading frame, can functions as a ncRNA in *S. aureus.* However, SprF1 RNA does not produce a protein when it regulates translation attenuation under osmotic stress (26). This indicates that sRNAs can behave differently depending on the environmental conditions, further supporting a highly complex regulatory network. Although rybA has been implicated in aromatic amino acid synthesis regulation, it is still relatively uncharacterized with respect to its cellular interactions (154).

Several rancRNAs identified in *Saccharomyces cerevisiae, Haloferax volcanii*, mammalian cells, and *Arabidopsis thaliana* have been identified as tRNA-derived fragments (14, 16, 20, 21, 24, 25). RttR was previously shown to be processed from the *tyrT* operon, encoding tyrosine tRNA tyrV (143). Given the prevalence of tRNA related byproducts in rancRNA studies, they may be conserved amongst the domains of life. To understand the sequence conservation of rttR, BLASTn searches demonstrated it is conserved in enterobacteria (Figure 3.4), which is consistent with the overall close evolutionary relationship between members of the Enterobacteriaceae family (164, 165). Previous experimental data revealed that rttR cross-links to helix 17 in the 5’ domain of the 16S rRNA ((31), Figure 3.5, Table S3.2). Little experimental data is available regarding the role of helix 17 during translation. Helix 17 is located on the periphery of the 30S ribosomal subunit and is unlikely to play a critical role during initiation or elongation (166). Interestingly, rttR superimposes well on helix 17 (RMSD = 3.5 Å), suggesting that rttR may be mimicking this section of 16S rRNA (Figure 3.5 D). Previous reports have identified similar behaviour for mRNA to aid in maintaining stoichiometry between rRNA and ribosomal proteins (167-169). Ribosomal protein S16 is shown to bind near the region where rttR resembles the structure of helix 17 (170, 171). S16 is a secondary binding protein, requiring ribosomal proteins S4 and S20 for binding to the 16S rRNA 5’ domain and that aids in suppressing excessive formation of 30S assembly intermediates (171, 172). I propose that rttR may

function during the formation of the 30S ribosomal subunit by blocking interactions between helix 17 and ribosomal protein S16, preventing progression of 30S ribosomal subunit maturation. This would limit the number of ribosomes available for translation, which occurs during times of cellular stress (173, 174). Many rancRNAs are known to function during cellular stress, such as osmotic shock (18, 26). The presence of rttR associated with stressed ribosomes would need to be validated in the future.

The reported filter binding data confirms that rttR binds specifically to the 70S ribosome, and neither the 50S or 30S ribosomal subunits (Figure 3.6, Table 3.5). Since helix 17 is located on the periphery of the 30S ribosomal subunit, binding to the 50S is not expected (166). To ensure that rttR is not binding in the mRNA channel when binding to the 70S ribosome, an mRNA competition assay will need to be completed (175). Interestingly, most of the candidate rancRNAs bind to 70S, 50S, and 30S ribosomes and subunits (Table 3.5). This raises the question if the candidates are binding specifically or non-specifically. It is well known that the components of the ribosome are able to interact with multiple RNA sequences, and contain several nucleic acid binding motifs (176). To gain a better understanding of the dissociation constants of RNA binding to the ribosome, further binding experiments with scramble RNAs and mRNA need to be conducted (177). However, the fact that the $tRNA^{Phe}$ control RNA does not bind to the 50S and 30S ribosomal subunits alone and the nM affinity for the 70S ribosome suggest that the identified rancRNA candidates are indeed specifically binding to the ribosome, which is consistent with their presence in the initial RNA sequencing datasets (30-32).

If rttR acts during 30S ribosomal subunit assembly, it is unlikely that the process of translation would be affected under conditions when primarily mature 30S ribosomal subunits are present. Consistent with this assumption, rttR did not affect eCFP production when tested in the PURExpress® system (Figure 3.7. Figure S3.4). Use of the PURExpress® system to test the other rancRNA candidates for effects on translation show that omrB has a positive effect on protein production when incubated concurrently with the PURExpress® components (Figure 3.7). A positive regulatory effect for omrB is surprising because it is known to be involved in negative regulatory pathways, although its mechanism of action is not completely understood (151, 178). Given that this interaction occurred in purified and reconstructed minimal gene

expression environment, omrB may have a previously unknown role in addition to its known targets (179). Interestingly, when the rancRNA candidates were incubated with the PURExpress® components prior to adding the eCFP template, neither increased nor decreased protein production. Given that the PURExpress® system is a purified reconstituted minimal transcription-translation system, it is possible that additional factors such as chaperones, may be necessary for the candidates to interact with the ribosome (180).

Interestingly, rttR likely does not impact global translation based on the *in vivo* protein synthesis assay using the BONCAT method (Figures 3.2, 3.8, S3.5, S3.6, S3.7, S3.8). Significant impacts on the 30S ribosomal subunit formation would have resulted in an overall decrease in translation as immature 30S ribosomal subunits enter translation initiation slowly and have lower translational fidelity (181, 182). Although rttR has a high affinity for the ribosome, it is unknown what the actual concentration of rttR is during the *in vivo* assay. While the number of ribosomes in *E. coli* fluctuates, approximately 6 800 ribosomes are present per cell during mid-log phase growth (183). Additionally, immature 30S ribosomal subunits are known to be utilized in translation during times of stress (182, 184). To better understand the state of the 30S ribosomal subunit, obtaining cryo-EM structures of rttR bound to the 70S ribosome would aid in understanding if ribosome biogenesis is impacted by rttR *in vivo* (19, 185). Furthermore, I cannot rule out that rttR may regulate specific mRNAs as has been observed with rancRNA_194 (17). To address this issue, mass spectrometry can be utilized to determine if specific proteins are up or downregulated (17, 161).

In conclusion, I have re-examined previously published data and identified and validated small RNAs that are capable of binding to the 70S ribosome and/or 50S/30S ribosomal subunits. With a particular focus on rttR, this work reports for the first-time equilibrium binding constants for five of the candidate rancRNAs to the 70S ribosome and the 50S / 30S ribosomal subunits. Additionally, *in vitro* and *in vivo* translation assays have been established to assess future rancRNA candidates. Although some of the rancRNA candidates can modulate translation *in vitro*, no *in vivo* effects could be detected. Further characterization is necessary to determine if rttR can indeed be classified as a rancRNA.

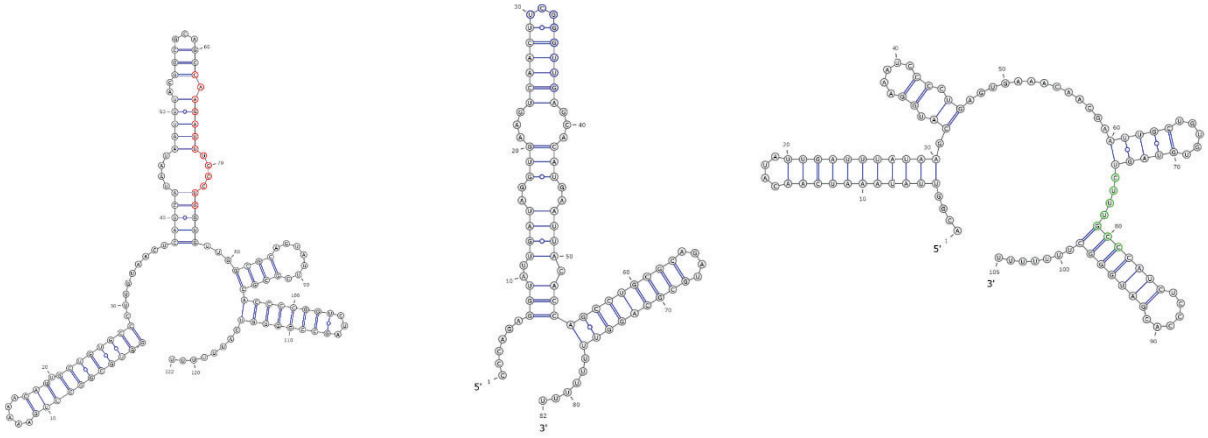
3.5: Supplementary Information

Table S3.1. List of identified sRNAs found to associate with ribosomes in previously published datasets (30-32). Highlighted in yellow are sRNAs found in all datasets indicated.

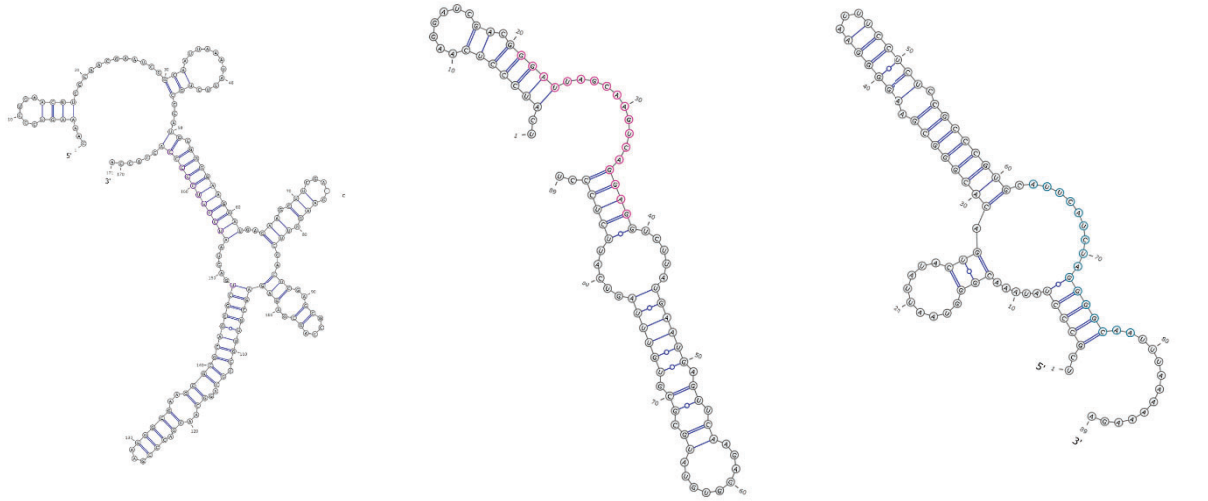
sRNA	Waters *et al.*, 2017	Smirnov *et al.*, 2016	Neuhaus *et al.*, 2017
arcZ	X	X	X
arrS			X
C0299	X		X
C0343	X		
C0362	X		
C0465			X
chiX		X	
cpxQ		X	
csrB	X	X	X
csrC		X	X
cyaR	X	X	X
dicF	X		X
dsrA	X	X	X
ffs		X	X
fnrS	X	X	X
gadY	X		X
gcvB	X	X	X
glmY	X	X	X
glmZ	X	X	X
isrA	X		
isrB	X	X	
istR			X
mgrR	X	X	
micA	X	X	X
micC		X	X
micF	X	X	X
micL		X	
micM	X		
omrA	X	X	X
omrB	X	X	X
orzP			X
oxyS	X		X
psrD	X		
rdlA	X		

rdlB	X		
rdlD	X	X	X
rnpB	X	X	X
rprA	X	X	X
rseX		X	X
rttR	X	X	X
rybA	X	X	
rybB	X	X	X
rybD		X	
rydB	X	X	X
rydC	X	X	X
ryeA			X
ryeB	X		X
ryfA		X	X
ryfD	X	X	
ryhB	X		X
ryjA	X		
ryjB		X	
sdsR		X	
sgrS	X	X	
sibA	X	X	
sibB	X		
sibC	X	X	
sibD		X	
sibE	X		
sokC	X		
sokX			X
spf	X	X	X
sraA			X
sraB		X	X
sraC		X	
sraF		X	
sraG			X
sraL		X	X
sroA		X	
sroB			X
sroC	X	X	X
sroD			X
sroE	X		X
sroG	X	X	
sroH	X		X

ssrA	X	X	
ssrS	X	X	X
symR	X	X	X
tff	X	X	
tpke11	X	X	X
tpke70	X		



arcZ (ΔG = -42.30 kcal/mol) **omrB (ΔG = -26.60 kcal/mol)** **rprA (ΔG = -23.90 kcal/mol)**



rttR (ΔG = -51.00 kcal/mol) **rybA (ΔG = -16.20 kcal/mol)** **tpke11 (ΔG = -32.70 kcal/mol)**

Figure S3.1. Predicted secondary structure of rancRNA candidates using RNAfold and generated with VARNA (128, 129).

Table S3.2. *In silico* analysis predicting base pairing interactions of rancRNA candidates and 23S rRNA or 16S rRNA (31, 130, 132, 133, 135-138).

Candidate	Cross-linked Region			
	23S rRNA	RMSD (Å) (atoms)	16S rRNA	RMSD (Å) (atoms)
ArcZ	170-199	23.131 (228/248)	608-669	23.167 (625/638)
122 nt	784-804	5.850 (145/146)	918-954	11.911 (406/421)
	1405-1469	23.695 (590/592)	1043-1079	16.184 (360/360)
	2232-2257	13.647 (275/275)	1392-1464	25.028 (675/675)
	2699-2740	16.663 (421/421)		
	Base Pairing Probabilty (IntaRNA)			
	23S rRNA	RMSD (Å) (atoms)	16S rRNA	RMSD (Å) (atoms)
	182-193	0.947 (65/81)	923-930	0.693 (71/84)
	785-803	5.625 (205/223)	1420-1429	5.070 (121/126)
	1408-1418	5.065 (101/104)		
	2233-2254	2.258 (105/120)		
	2733-2739	1.192 (42/42)		
	Alignment to rRNA			
	23S rRNA	RMSD (Å) (atoms)	16S rRNA	RMSD (Å) (atoms)
	2739-2742	0.148 (13/20)		
	RNAUp			
	23S	RMSD (Å) (atoms)	16S	RMSD (Å) (atoms)
	1248-1259	6.848 (109/109)	1408-1418	1.288 (122/124)
	Cross-linked Region			
	23S rRNA	RMSD (Å) (atoms)	16S rRNA	RMSD (Å) (atoms)
OmrB	749-788	13.136 (357/357)	517-592 bp	32.225 (598/598)
82 nt	887-943	29.813 (488/488)	614-672 bp	34.716 (603/603)
	1453-1497	32.084 (529/529)	1046-1067 bp	10.679 (210/210)
	2272-2321	29.425 (515/515)		
	Base Pairing Probabilty (IntaRNA)			
	23S rRNA	RMSD (Å) (atoms)	16S rRNA	RMSD (Å) (atoms)
	898-905	0.792 (40/43)	617-626	6.582 (143/143)
	Alignment to rRNA			
	23S rRNA	RMSD (Å) (atoms)	16S rRNA	RMSD (Å) (atoms)
	748-753	1.276 (16/18)	517-533	30.628 (201/201)
	751-753	0.311 (12/18)	557-568	5.843 (104/106)
	780-785	0.125 (12/18)	557-580	10.308 (242/242)

	876-880	0.410 (20/20)	570-577	5.397 (100/101)
	886-891	0.031 (13/20)	618-623	0.088 (13/20)
	930-932	0.899 (15/18)		
	940-942	0.639 (15/18)		
	1451-1461	3.005 (42/43)		
	1495-1497	0.270 (18/22)		
	2172-2175	0.128 (15/20)		
	2183-2185	0.262 (22/22)		
	2306-2311	3.251 (59/61)		
	Base Pairing Probability (IntaRNA)			
	23S rRNA	RMSD (Å) (atoms)	16S rRNA	RMSD (Å) (atoms)
			(1+)618-623	6.582 (143/143)
			618-623(+1)	0.823 (51/60)
	RNAUp			
	23S	RMSD (Å) (atoms)	16S	RMSD (Å) (atoms)
	1023-1026	0.090 (16/19)	1472-1482	2.752 (140/147)
	Cross-linked Region			
	23S rRNA	RMSD (Å) (atoms)	16S rRNA	RMSD (Å) (atoms)
RprA 105 nt	752-780	15.818 (379/380)	551-609	21.025 (481/481)
	859-891	14.394 (309/310)	1150-1249	20.511 (878/878)
	Base Pairing Probabilty (IntaRNA)			
	23S rRNA	RMSD (Å) (atoms)	16S rRNA	RMSD (Å) (atoms)
	774-780	0.076 (16/20)	578-601	6.859 (337/390)
	868-884	4.166 (184/184)		
	Alignment to rRNA			
	23S rRNA	RMSD (Å) (atoms)	16S rRNA	RMSD (Å) (atoms)
	751-753	0.284 (11/18)	557-580	13.911 (228/228)
	780-783	0.081 (12/18)	606-609	0.753 (15/18)
	876-880	0.218 (16/20)	1176-1183	0.212 (22/22)
	886-891	0.189 (15/20)	1187-1191	0.469 (18/20)
			1212-1214	0.121 (14/20)
			1221-1223	0.093 (23/23)
			1236-1241	0.133 (17/22)
	Base Pairing Probability (IntaRNA)			
	23S rRNA	RMSD (Å) (atoms)	16S rRNA	RMSD (Å) (atoms)
	(2+)876-880	0.528 (37/43)	1221-1223 (+4)	4.555 (60/65)
	RNAUp			

	23S	RMSD (Å) (atoms)	16S	RMSD (Å) (atoms)
	1022-1027	5.276 (96/96)	1219-1234	8.839 (129/129)
RttR 171 nt	Cross-linked Region			
	23S rRNA	RMSD (Å) (atoms)	16S rRNA	RMSD (Å) (atoms)
	1649-1693	14.268 (287/349)	441-463	3.504 (116/116)
			1166-1195	10.787 (169/179)
	Base Pairing Probabilty (IntaRNA)			
	23S rRNA	RMSD (Å) (atoms)	16S rRNA	RMSD (Å) (atoms)
	1655-1662	1.783 (137/149)	444-450	0.685 (15/18)
			1177-1187	0.123 (14/22)
	Alignment to rRNA			
	23S rRNA	RMSD (Å) (atoms)	16S rRNA	RMSD (Å) (atoms)
	1652-1654	1.032 (16/18)	447-455	0.108 (12/18)
	1671-1680	8.298 (120/120)	1166-1171	1.572 (32/42)
	1687-1691	0.268 (23/23)	1176-1183	0.110 (16/20)
			1187-1191	0.042 (18/22)
	RNAUp			
	23S	RMSD (Å) (atoms)	16S	RMSD (Å) (atoms)
	857-863	0.319 (19/20)	418-429	6.227 (185/194)
RybA 89 nt	Cross-linked Region			
	23S rRNA	RMSD (Å) (atoms)	16S rRNA	RMSD (Å) (atoms)
	426-450	15.165 (262/282)		
	961-1012	24.489 (421/421)		
	Base Pairing Probabilty (IntaRNA)			
	23S rRNA	RMSD (Å) (atoms)	16S rRNA	RMSD (Å) (atoms)
	433-445	1.187 (95/121)		
	Alignment to rRNA			
	23S rRNA	RMSD (Å) (atoms)	16S rRNA	RMSD (Å) (atoms)
	426-432	2.306 (42/42)		
	445-448	0.079 (15/20)		
	971-975	0.706 (15/18)		
	979-986	0.758 (15/18)		
	990-992	0.678 (15/18)		
	1005-1011	0.747 (19/23)		
	Base Pairing Probability (IntaRNA)			

	23S rRNA	RMSD (Å) (atoms)	16S rRNA	RMSD (Å) (atoms)
	426-432	2.661 (119/119)		
	RNAUp			
	23S	RMSD (Å) (atoms)	16S	RMSD (Å) (atoms)
	1429-1451	3.810 (322/358)	1221-1238	1.561 (217/271)
	Cross-linked Region			
	23S rRNA	RMSD (Å) (atoms)	16S rRNA	RMSD (Å) (atoms)
Tpke11 89 nt	2388-2407	11.535 (207/207)	441-479	13.377 (294/294)
			572-597	17.548 (265/265)
			1048-1078	15.499 (189/189)
	Alignment to rRNA			
	23S rRNA	RMSD (Å) (atoms)	16S rRNA	RMSD (Å) (atoms)
			447-455	1.178 (16/18)
			463-470	4.166 (61/62)
	RNAUp			
	23S	RMSD (Å) (atoms)	16S	RMSD (Å) (atoms)
	2768-2783	5.858 (236/236)	1487-1492	134.2 (83/83)

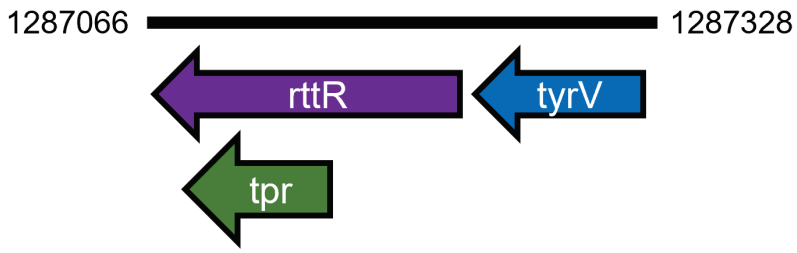


Figure S3.2. rttR genomic location in *E. coli* MG1655 (134). rttR is a part of the *tyrT* operon that encodes for the tyrosine tRNA tyrV and the protein tpr (143, 186). The black line indicates the position in the *E. coli* chromosome.

Table S3.3. rancRNA candidate sequences and primers.

rancRNA	rancRNA Construct	Forward Primer	Reverse Primer
arcZ	5'-TAATACGACTCACTATAGGGTGCGGCCTGAAAAACAGTGCTGTGCCCTTGTAACTCATCATAATAATTTACGGCGCAGCCAAGATTTCCCTGGTGTTGGCGCAGTATTCGCGCACCCCGGTCTAGCCGGGGTCATTTTTTGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATTCCGAGGGGACCGTCCCCTCGGTAATGGCGAATGGGACCTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG-3'	TAATACGACTCACTATAGGGTGCGGC	AAAAAATGACCCCGGCTAGACCG
omrB	5'-TAATACGACTCACTATAGGGCCCAGAGGTATTGATAGGTGAAGTCAACTTCGGGTTGAGCACATGAATTACACCAGCCTGCGCAGATGCGCAGGTTTTTTTTGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATTCCGAGGGGACCGTCCCCTCGGTAATGGCGAATGGGACCTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG-3'	TAATACGACTCACTATAGGGCCCAGA	AAAAACCTGCGCATCTGCGCA
rprA	5'-TAATACGACTCACTATAGGGACGGTTATAAATCAACATATTGATTTATAAGCATGGAAATCCCCTGAGTGAAACAACGAATTGCTGTGTGTAGTCTTTGCCCATCTCCCACGATGGGCTTTTTTTGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATTCCGAGGGGACCGTCCCCTCGGTAATGGCGAATGGGACCTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG-3'	TAATACGACTCACTATAGGGACGGTTA	AAAAAGCCCATCGTGGGAGA
rttR	5'-TAATACGACTCACTATAGGGCAAAAGTCCCTGAACTTCCCAACGAATCCGCAATTAAATATTCTGCCCATGCGGGGAAGGATGAGAAGCTTCGACCAAGGTTCGACTCGAGCGCCAGCGAGAGAGCGTTGCCGCAGGCAACGACCCGAAGGGCGAAGCGCGCAGCGCTGAGTAATCCTTCCCCCACCACCAGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATTCCGAGGGGACCGTCCCCTCGGTAATGGCGAATGGGACCTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG-3'	TAATACGACTCACTATAGGGCAAAAGT	TGGTGGTGGGGGAAGGATTACT
rybA	5'-TAATACGACTCACTATAGGGTCATCCCTCAAGGATCGACGGGATTAGCAAGTCAGGAGGTCTTATGAATGAGTTCAAGAGGTGTATGCGCGTGTTTAGTCATTCTCCCTGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATTCCGAGGGGACCGTCCCCTCGGTAATGGCGAATGGGACCTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG-3'	TAATACGACTCACTATAG	AGGGAGAATGACTAAACA

tpke11	5’- TAATACGACTCACTATAGGGTCGCCCTATAAA CGGGTAATTATACTGACACGGGCGAAGGGGA ATTTCCTCTCCGCCCGTGCATTCATCTAGGGG CAATTTAAAAAAGAGGCCGGCATGGTCCCAG CCTCCTCGCTGGCGCCGGCTGGGCAACATTC CGAGGGGACCGTCCCCTCGGTAATGGCGAAT GGGACCTAGCATAACCCCTTGGGGCCTCTAA ACGGGTCTTGAGGGGTTTTTTG-3’	TAATACGACTC ACTATAG	TTTTTAAATTGC CCCTAGA
Scramble RNA	5’- GAATTCTAATACGACTCACTATAGGGGAACTC GCAAGGGTCGTGAAGTCGGTTCCTTCAATGG TTAAAAAACAAAGGCTTACTGTGCAGACTGGA ACGCCCATCTAGCGGCTCGCGTCTTGAATGC TCGGTCCCCTTTGTCATTCCGGATAAATCCCT AGCATAACCCCTTGGGGCCTCTAAACGGGTC TTGAGGGGTTTTTTGAAGCTT-3’	TAATACGACTC ACTATAGGGG AACTCGC	CAAAAAACCCC TCAAGACCCGT TTAG

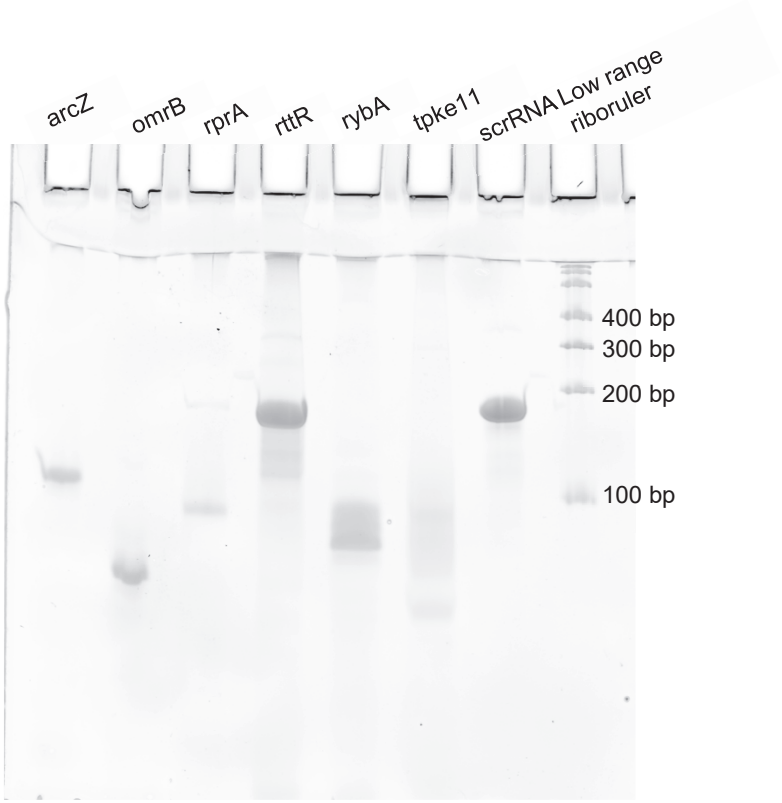


Figure S3.3. *In vitro* transcribed RNA of all rancRNA candidates and a scramble RNA control analyzed on a 10% urea PAGE and stained with ethidium bromide. The RNA size marker is a low range riboruler (Thermo Fisher).

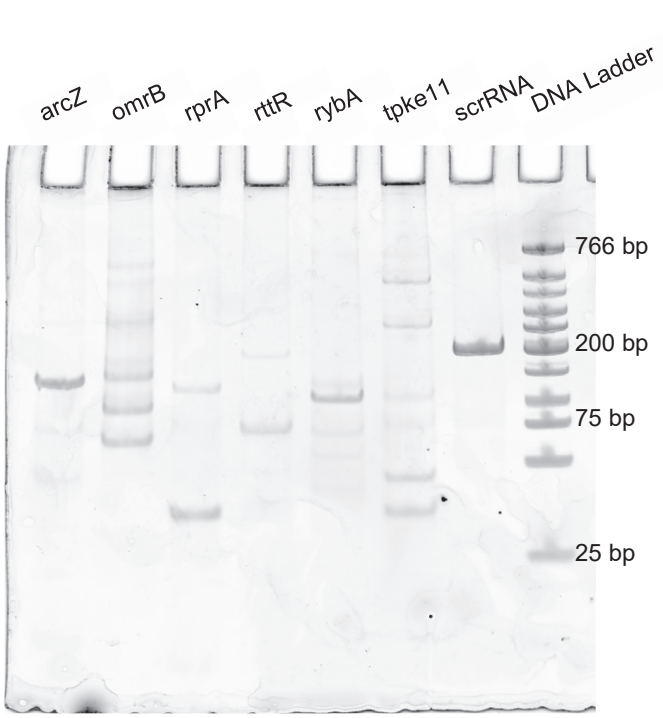


Figure S3.4. RT PCR of RNA added into PURExpress® system analyzed on a 10% native PAGE and stained with ethidium bromide. The DNA size marker is a DNA low molecular weight ladder (NEB).

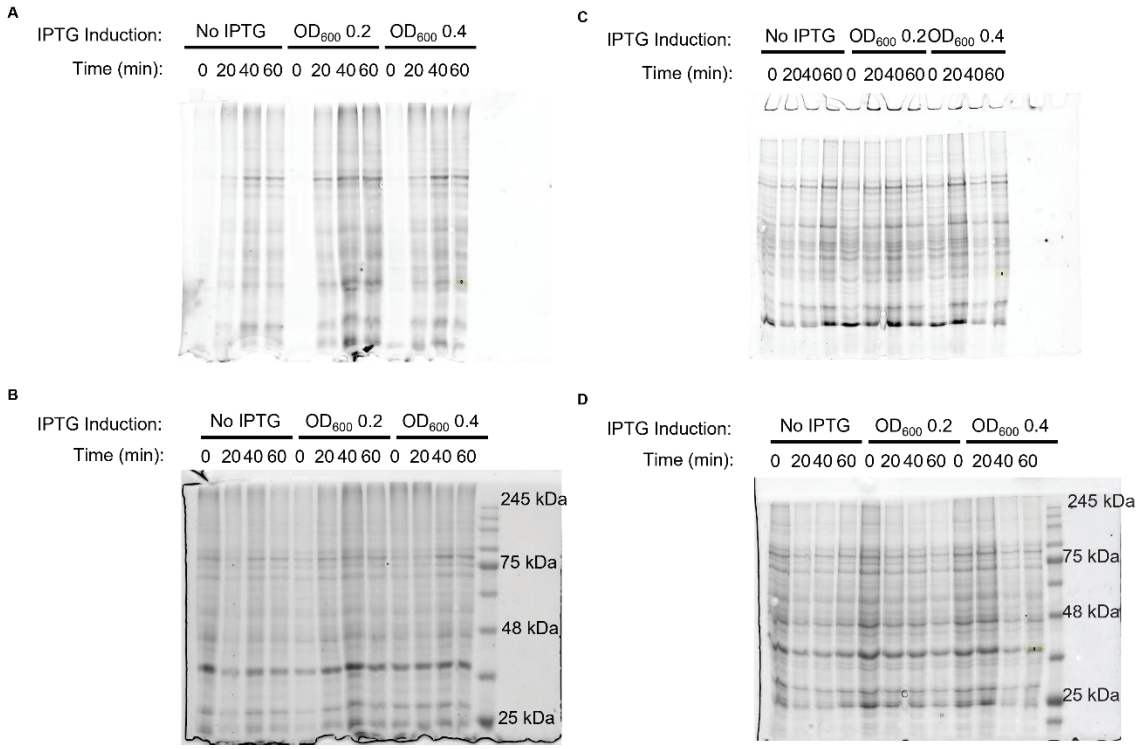


Figure S3.5. Biological replicates of extracted proteins from BONCAT assay with rttR induction. **(A).** 12% SDS PAGE of fluorescently labelled proteins (biological replicate 2). **(B).** Same SDS PAGE with Coomassie staining (biological replicate 2). **(C).** 12% SDS PAGE of fluorescently labelled proteins (biological replicate 3). **(D).** Same SDS PAGE with Coomassie staining (biological replicate 3).

Table S3.4. Dunnett's multiple comparison test for quantitative analysis of rttR's effect on *in vivo* translation between different induction times (* - $p < 0.05$, ** - $p < 0.01$, ns – not significant).

Effect of rttR Between Conditions	Padj (significance)
0 min	
No IPTG vs IPTG at 0.2 OD_{600}	0.9240 (ns)
No IPTG vs IPTG at 0.4 OD_{600}	0.9982 (ns)
20 min	
No IPTG vs IPTG at 0.2 OD_{600}	0.0141 (*)
No IPTG vs IPTG at 0.4 OD_{600}	0.0019 (**)
40 min	
No IPTG vs IPTG at 0.2 OD_{600}	0.4412 (ns)
No IPTG vs IPTG at 0.4 OD_{600}	0.1292 (ns)
60 min	
No IPTG vs IPTG at 0.2 OD_{600}	0.5094 (ns)
No IPTG vs IPTG at 0.4 OD_{600}	0.9925 (ns)

Table S3.5. Dunnett's multiple comparison test for quantitative analysis of rttR's effect on *in vivo* translation within each induction condition (* - p < 0.05, ** - p < 0.01, ns – not significant).

Effect of rttR Within Conditions	Padj (significance)
No IPTG	
0 min vs 20 min	0.0022 (**)
0 min vs 40 min	0.0050 (**)
0 min vs 60 min	0.9645 (ns)
IPTG at 0.2 OD_{600}	
0 min vs 20 min	0.5765 (ns)
0 min vs 40 min	0.0432 (*)
0 min vs 60 min	0.2183 (ns)
IPTG at 0.4 OD_{600}	
0 min vs 20 min	0.9987 (ns)
0 min vs 40 min	0.3367 (ns)
0 min vs 60 min	0.9119 (ns)

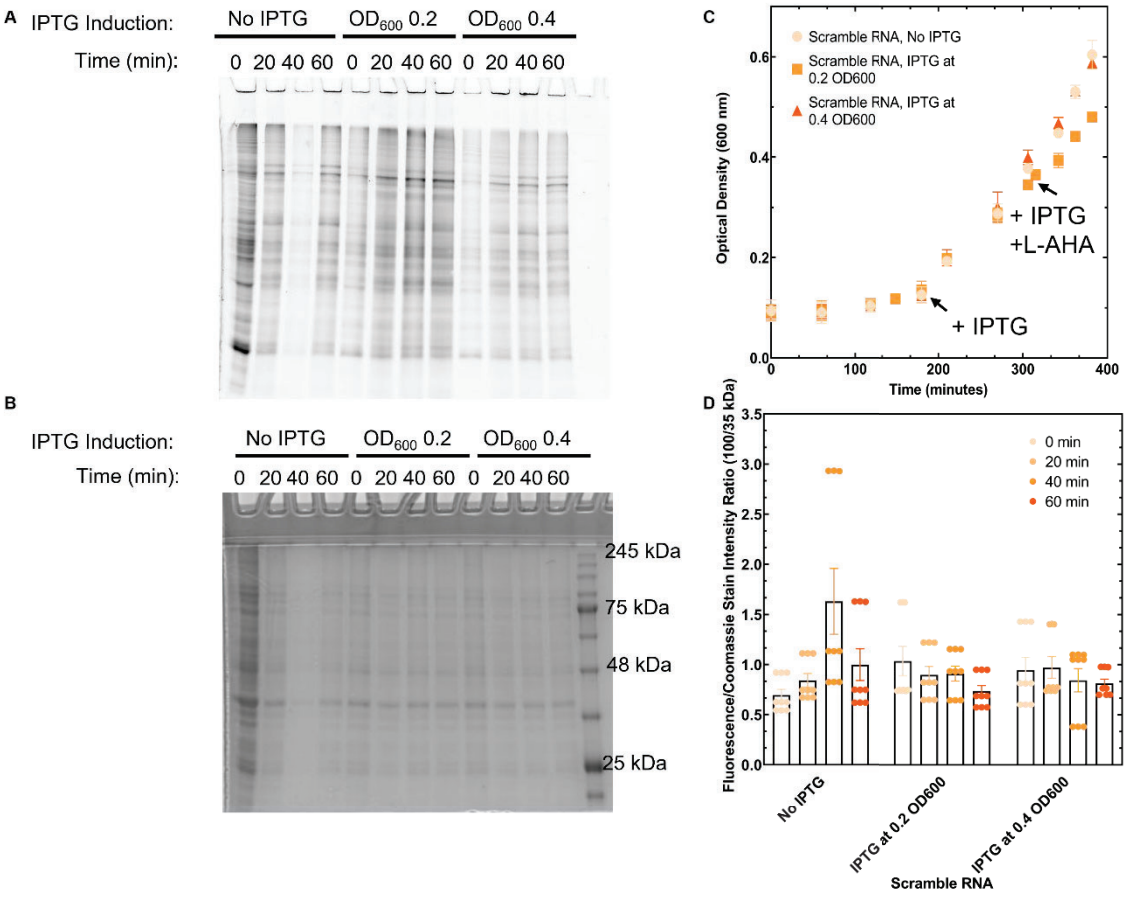


Figure S3.6. Control experiment of *in vivo* monitoring of global *de novo* protein synthesis with scramble RNA induction. **(A).** 12% SDS PAGE of fluorescently labelled proteins. **(B).** Same SDS PAGE with Coomassie blue staining. **(C).** Growth of *E. coli* BL21(DE3) cells transformed with plasmid encoding rttR (n=3 ± SEM). **(D).** Normalized fluorescence of the 100 and 35 kDa bands (n=3 ± SEM).

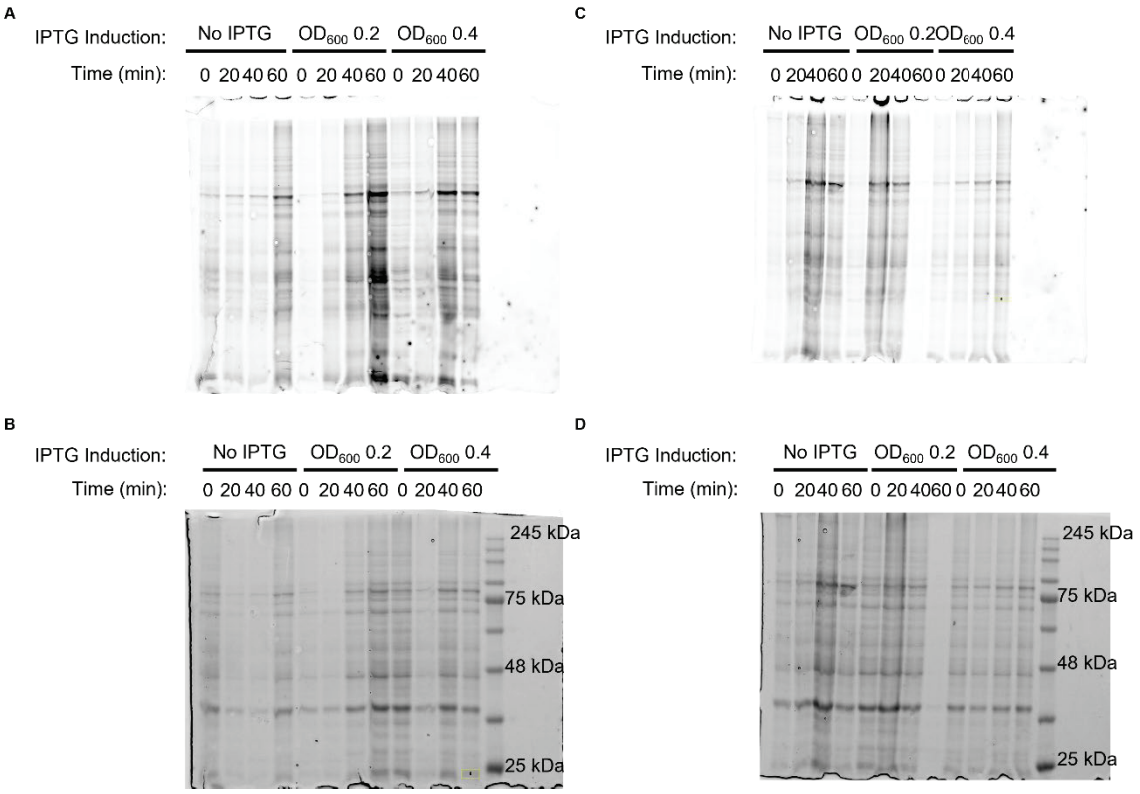


Figure S3.7. Biological replicates of extracted proteins from BONCAT assay with scramble RNA induction. **(A).** 12% SDS PAGE of fluorescently labelled proteins (biological replicate 2). **(B).** Same SDS PAGE with Coomassie staining (biological replicate 2). **(C).** 12% SDS PAGE of fluorescently labelled proteins (biological replicate 3). **(D).** Same SDS PAGE with Coomassie staining (biological replicate 3).

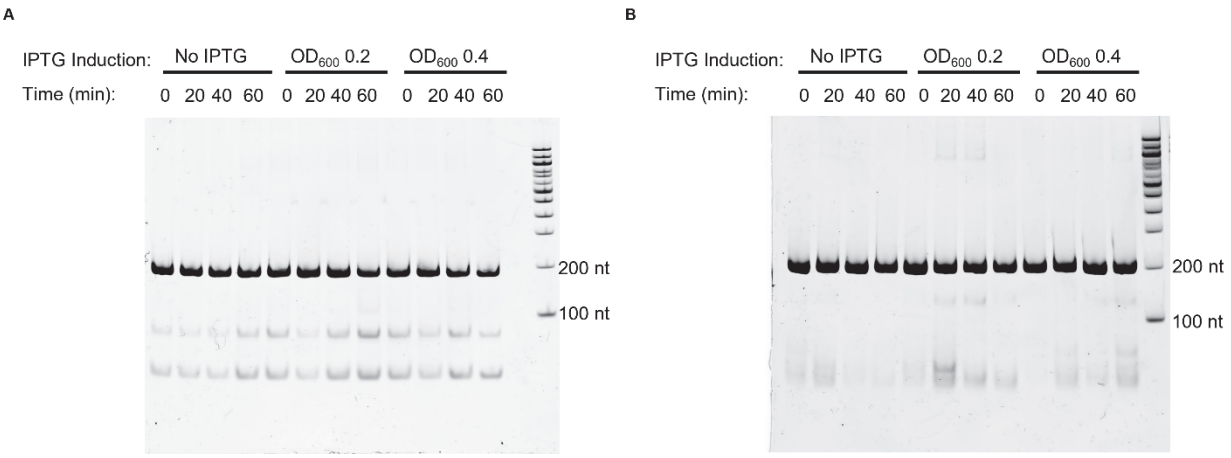


Figure S3.8. RT PCR of RNA transcribed during BONCAT assay analyzed on 10% native PAGE gels and stained with Ethidium Bromide (ladder – 100 base pair, Thermo Scientific). **(A).** cDNA of rttR (178 nt) from BONCAT assay. **(B).** cDNA of scramble RNA (198 nt) from BONCAT assay.

Table S3.6. Dunnett's multiple comparison test for quantitative analysis of the scramble RNA's effect on *in vivo* translation between induction conditions (*** - $p < 0.001$, ns – not significant).

Effect of rttR Between Conditions	Padj (significance)
0 min	
No IPTG vs IPTG at 0.2 OD_{600}	0.1387 (ns)
No IPTG vs IPTG at 0.4 OD_{600}	0.3186 (ns)
20 min	
No IPTG vs IPTG at 0.2 OD_{600}	0.9412 (ns)
No IPTG vs IPTG at 0.4 OD_{600}	0.7220 (ns)
40 min	
No IPTG vs IPTG at 0.2 OD_{600}	0.0006 (***)
No IPTG vs IPTG at 0.4 OD_{600}	0.0002 (***)
60 min	
No IPTG vs IPTG at 0.2 OD_{600}	0.2957 (ns)
No IPTG vs IPTG at 0.4 OD_{600}	0.5279 (ns)

Table S3.7. Dunnett's multiple comparison test for quantitative analysis of the scramble RNA's effect on *in vivo* translation within the induction condition (**** - $p < 0.0001$, ns – not significant).

Effect of rttR Within Conditions	Padj (significance)
No IPTG	
0 min vs 20 min	0.7836 (ns)
0 min vs 40 min	<0.0001 (****)
0 min vs 60 min	0.2706 (ns)
IPTG at 0.2 OD_{600}	
0 min vs 20 min	0.8102 (ns)
0 min vs 40 min	0.8511 (ns)
0 min vs 60 min	0.2800 (ns)
IPTG at 0.4 OD_{600}	
0 min vs 20 min	0.9985 (ns)
0 min vs 40 min	0.9068 (ns)
0 min vs 60 min	0.8247 (ns)

CHAPTER 4: Identification of potential rancRNA candidates under stress conditions in *Escherichia coli*

4.1: Introduction

Translation is highly dynamic during cellular stress as cells up and down regulate the expression of genes (187, 188). sRNAs have long been known to be an important aspect in signalling during environmental stress (7, 8, 189). Many studies provide evidence for sRNAs acting post-transcriptionally by base pairing with a target mRNA to modulate translation of the respective mRNA, to impact transcript stability, or act in association with protein partners like Hfq (2, 190, 191).

Nutrient stress impacts metabolic regulation of the cell (192, 193). Known regulators in *E. coli* are the RpoS pathway that has wide implications for gene regulation, as well as the carbon storage regulatory (Csr) system including the protein CsrA and two small RNAs csrB and csrC (194-198). In addition to csrB and csrC, only a few sRNAs have been characterized in *E. coli* that function during nutrient depravation including cyaR, gcvB, and sdsN, inhibiting translation (199-201). However, there is still a lack of detailed mechanistic understanding with respect to how the ribosome may be impacted by sRNAs (199, 202-204). Interestingly, several rancRNAs in *S. cerevisiae, T. brucei,* and *H. volcanii* have been found to regulate the ribosome during nutrient stress (17, 20, 22).

In several strains of *Salmonella*, the sRNAs rybB, cyaR, and invR are upregulated under hyperosmotic stress, though the exact mechanisms of action are uncertain (205). In *E. coli*, rprA has been shown to activate translation of the *rpoS* mRNA in response to osmotic stress (153). RancRNAs in *S. cerevisiae* and *S. aureus* have also been shown to negatively regulate translation by binding the ribosome during hyperosmotic stress (18, 20, 26).

Furthermore, sRNAs have been implicated in antibiotic resistance mechanisms (106, 206). Intriguingly, the only rancRNA identified so far in bacteria was found to induce persister cell formation, which promote antibiotic tolerance (26, 207). Interestingly, tetracyclines have been reported to increase the formation of persister cells (208).

As reviewed in Chapters 1 and 2, rancRNAs have been proposed to impact translation under stress (209). In bacteria, SprF1 RNA is the only sRNA that has been identified (*S. aureus)* to date (26). This suggests that indeed rancRNA mechanisms are exploited for translation regulating in bacteria but have been overlooked so far. I investigated if rancRNAs are present in *E. coli* under the following stress conditions: nutrient starvation (M9 salts minimal medium), hyperosmotic, and antibiotic stress (1/2 minimum inhibitory concentration (MIC) of tetracycline).

4.2: Methods

4.2.1: Ribosome isolation

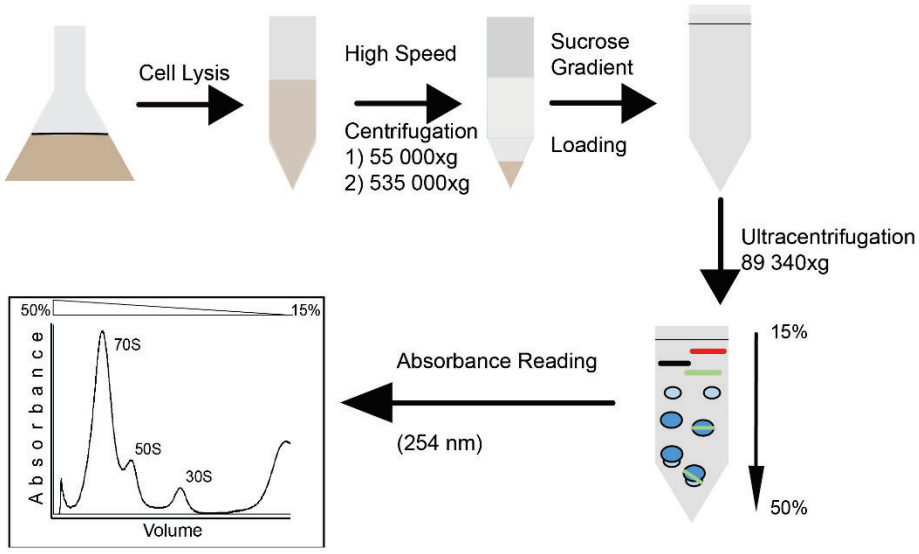


Figure 4.1. Overview of ribosome isolation protocol. *E. coli* MG1655 cells induced with a stress condition are lysed and subsequently centrifuged in two steps to collect ribosomal particles. The ribosomal particles are then separated by ultra centrifugation using a 15-50% sucrose gradient and subsequently fractionated to obtain the respective ribosomal particles.

Cell Growth:

Overnight cultures of *Escherichia coli* MG1655 were grown in 50 mL of Luria Burtani (1% Sodium Chloride (weight/volume), 1% Tryptone (w/v), 0.5% Yeast Extract (w/v), pH = 7) medium at 37 ºC and 220

rpm in a shaking incubator (Max^{Q} 4000 Barnstead Lab-Line). Cells were then centrifuged at 5 000xg for 10 minutes at 4 °C (Heraeus Multifuge X3R Centrifuge, Thermo Scientific), the cell pellet was resuspended in fresh LB medium and used to inoculate 500 mL of LB medium at a starting optical density of 600 nm (OD_{600}) at 0.100. Cells were grown until OD_{600} = 0.600 and harvested by centrifugation at 5 000xg for 15 minutes at 4 °C using a JLA 8.1000 rotor in an AvantiJ-26 XPI centrifuge (Beckman Coulter). The pellets were then resuspended in 500 mL of fresh LB medium, 0.7 M NaCl containing LB medium, minimal medium (5x M9 salts, 1 M $MgSO_4$, 20% glucose (w/v), 1 M $CaCl_2$), or ½ MIC (2 µg/mL) of tetracycline in LB medium. The resulting cultures were then incubated for another 20 minutes at 37 °C and subsequently collected by centrifugation using the same rotor and conditions as described above. After discarding the media, cells were washed in 1 mL of cell opening buffer (20 mM Tris-Hydrochloric acid, pH = 7.6 at 4 °C, 60 mM Ammonium Chloride, 5.25 mM Magnesium Acetate, 3 mM β-Mercaptoethanol) and transferred to a 15 mL falcon tube. The cells were collected by centrifugation at 5 000xg for 10 minutes at 4 °C (Heraeus Multifuge X3R Centrifuge, Thermo Scientific). The supernatant was discarded, and the resulting cell pellet (~0.7 g) was flash frozen using liquid nitrogen and stored at -80 ºC until further use.

Sucrose Density Gradient Ultracentrifugation:

An overview of the procedure for isolating ribosomes from *E. coli* MG1655 cells grown under stress is shown in figure 4.1. Herein, this portion of the study was initially designed based on Smirnov *et al.* (30) and thus focused on sequencing sRNAs isolated from 70S ribosomes and 50S/30S ribosomal subunits. Frozen cell pellets of approximately 0.7 g were thawed and resuspended in 5 mL of cell opening buffer, 1 mg/mL lysozyme (BioBasic), a few crystals of DNase I (Sigma), and 0.016 U/µL of RNase inhibitor I (Thermo Scientific). Cells were then lysed using a French pressure cell press at 4 000 psi (Sim-Aminco Spectronic Instruments) by passing the cell lysate through the machine three times. The cell lysate was then centrifuged at 55 000xg for 30 minutes at 4 ºC using an S140-AT rotor in a Sorvall MX 150 Micro-Ultracentrifuge (Thermo Scientific). Subsequently, the cleared lysate was transferred to new tubes and centrifuged at 535 000xg for two hours at 4 °C using the same rotor and centrifuge as in the previous step. Ribosome pellets were slowly re-suspended in cell opening buffer on ice for at least two hours by gentle

shaking. The absorbance at 260 nm of the dissolved pellet was determined using a BioDrop (MBI Lab Equipment) and the total optical density (OD) was calculated (60-120 OD_{260}, 1 OD = A_{260} *dilution factor*total volume*path length).

Sucrose gradients were prepared using 50% sucrose (20 mM Tris-HCl, 60 mM NH_4Cl, 5.25 mM $MgAc_2$, 3 mM βME, 50% sucrose (w/v), pH = 7.6 at 4 °C) and 15% sucrose (20 mM Tris-HCl, 60 mM NH_4Cl, 5.25 mM $MgAc_2$, 3 mM βME, 15% sucrose (w/v), pH = 7.6 at 4 °C). The gradients were formed using a Gradient Master 108 (BioComp) in 38.4 mL open top tubes (Beckman Coulter). Equivalent amounts (50 ODs) of ribosomes collected as described above were loaded on each individual gradient. The samples were then centrifuged at 89 340xg for 16 hours at 4 ºC using an SW28 rotor (Beckman Coulter) in a Beckman Coulter Optima XPN-100. Following centrifugation, the gradients were fractionated using an ÄktaPrime Plus (GE Healthcare Life Sciences) and the absorbance at 254 nm was recorded to generate an absorbance profile. The obtained absorbance profiles were subsequently analyzed using Microsoft Excel (2016) and Prism GraphPad version 9. RNA was isolated from fractions that contained minimal cross-contamination between 70S, 50S, and 30S ribosomal peaks.

4.2.2: Preparation of RNA for next generation sequencing

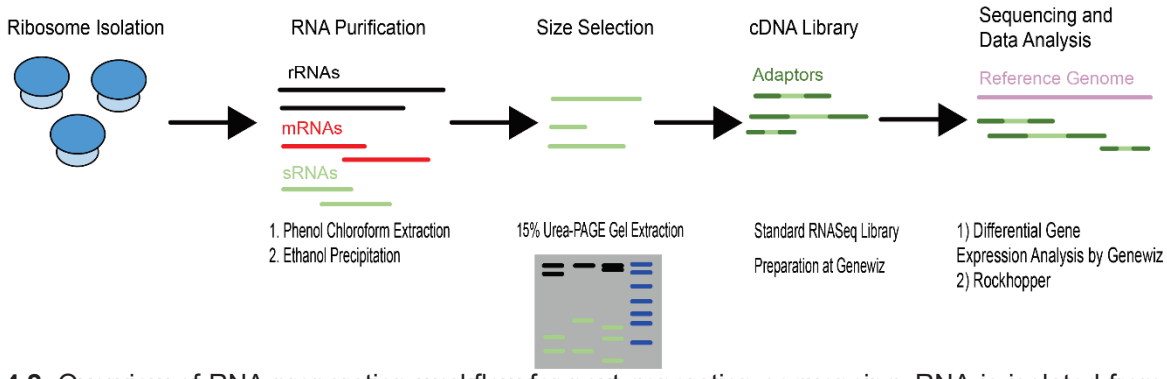


Figure 4.2. Overview of RNA preparation workflow for next generation sequencing. RNA is isolated from ribosomal particles by phenol chloroform extraction and ethanol precipitation to enrich sRNAs for subsequent gel extraction. A cDNA library is generated from the isolated sRNAs and sequenced using next generation sequencing followed by bioinformatic analysis.

Total RNA Purification:

The RNA purification workflow is summarized in figure 4.2. RNA from selected fractions corresponding to the 70S, 50S, and 30S ribosomal particles was purified using phenol-chloroform extraction in 2 mL phase-lock tubes according to the manufacturer's protocol (Quantabio). The aqueous phase was precipitated overnight at -20 °C in 100% isopropanol, 3 M NaOAc. pH 5.2, and glycoblue (Ambion) in 1.5 or 2 mL lo-bind tubes (Eppendorf). Samples were centrifuged at 13 000 rpm for 30 minutes at 4 °C (Sorvall Legend Micro 17R, Thermo Scientific). The supernatant was decanted, and the RNA pellet was washed with 70% ethanol and centrifuged using the same conditions as before. After discarding the supernatant, pellets were dried in a fume hood for approximately 20 minutes and the resulting dried pellets were resuspended in nuclease-free water (New England Biolabs).

Size Selection:

To eliminate ribosomal RNA and enrich for small RNAs, the purified samples and a low range riboruler (Thermo Fisher) were heated to 95 °C for 10 minutes. A 15% urea polyacrylamide gel was pre-heated at 300V for 30 minutes in hot 1x TBE buffer (130 mM Tris (pH 7.6), 45 mM borate, 2.5 mM EDTA). Subsequently, the purified RNA samples and size marker were separated on the pre-heated 15% urea polyacrylamide gel at 300V for 30 minutes. RNA was visualized via UV shadowing and the whole length of the lane below the 300-base pair marker band was extracted. The gel slices were then mixed with 2 volumes of gel extraction solution (0.3 M Sodium Acetate, 50 mM EDTA, pH = 7.0) in 2 mL lo-bind microcentrifuge tubes and shaken overnight at 4 °C with a rocker at speed/tilt 15 (Fisher Scentific). The tubes were then centrifuged at 5 000xg for one minute (Sorvall Legend Micro 17R, Thermo Scientific) and the resulting supernatant was transferred to a new 2 mL lo-bind microfuge tube. RNA was precipitated overnight at -20 °C in 100% isopropanol, 3 M NaOAc. pH 5.2, and glycoblue (Ambion) in 1.5 or 2 mL lo-bind tubes (Eppendorf). Samples were subsequently centrifuged at 13 000 rpm for 15 minutes at 4 °C (Sorvall Legend Micro 17R, Thermo Scientific) and the resulting supernatant was decanted. The RNA pellet was washed with 70% ethanol and centrifuged using the same conditions as before. After discarding the supernatant, pellets were dried in a fume hood for approximately 20 minutes. Dried pellets were resuspended in

nuclease-free water (NEB). Collected samples were used to generate a cDNA library by reverse transcription PCR, followed by adaptor ligation and subjected to NGS using an Illumina HiSeq (2x 150 bp paired end) (cDNA library preparation and NGS completed by Azenta/Genewiz).

4.2.3: Bioinformatic analysis

A two-pronged approach was used to analyze the obtained sequencing data (Azenta/Genewiz). The first analysis was completed by Azenta/Genewiz. After evaluating sequence quality, reads were trimmed using Trimmomatic v.0.36 (210) and mapped to the *E. coli* genome ((211), GenBank number: U00096.3) using Bowtie2 to generate Binary Alignment Map (BAM) files and mapping statistics (212). The unique gene hit counts were then extracted using featureCounts v.1.5.2 (213) and differential gene expression analysis comparing the LB 70S, 50S, and 30S ribosomes and ribosomal subunits to the ribosomes and ribosomal subunits from the stress conditions was completed with DESeq2 using gene hit counts (number of reads that map to the specific gene) greater than 10 (214). P-values were calculated using the Wald test and log2fold changes based on an estimated size factor (library size and composition) and dispersion (variability between biological replicates) for each gene relying on the assumption that most genes will not be differentially expressed. Adjusted p-values less than 0.05 and absolute log2fold changes greater than 1 were identified as being significant.

To categorize the RNA as non-coding RNA (ncRNA), rRNA, coding RNA, or tRNA, the list of genes from the DESeq2 analysis was input into PANTHER v16 (215) to determine which genes encode proteins. Genes that did not match to any genes within the PANTHER database were manually categorized using EcoCyc (134). The data was then further analyzed using Microsoft Excel (2016) and GraphPad Prism version 9.

The second analysis was used to identify non-annotated transcripts, using the Rockhopper computational analysis system (216-218). In brief, trimmed FastQ files from Azenta/Genewiz were used as input for Rockhopper along with the annotated reference genome for *E. coli* ((211), GenBank number: U00096.3). Rockhopper aligned the sequence reads to the reference genome using a Burrows-Wheeler

index, and the read counts were normalized with upper quartile normalization to allow samples to be compared (216). Transcript boundaries were identified, and novel transcripts were predicted in between these boundaries. Results were viewed using Integrative Genomics Viewer (219) and transcripts with a false discovery rate of less than 0.01 were deemed significant. Data was further analyzed using Microsoft Excel (2016) and GraphPad Prism version 9.

4.3 Results

4.3.1: sRNA isolation and sequencing

To obtain small RNAs associated with ribosomes for sequencing, 70S ribosomes, and the 50S/30S ribosomal subunits were separated using sucrose (15-50%) gradient density ultracentrifugation. The respective three peaks (70S/50S/30S) were observed for the ribosome pool collected from cells grown in LB, minimal medium, and hyperosmotic conditions (Figure 4.3). Interestingly, four peaks were observed for the cells grown in the presence of tetracycline, with an additional peak in the <30S region. Visual inspection of the ratio of the 50S and 30S peaks relative to the 70S peak revealed an increase in 50S and 30S ribosomal subunits for cells grown under the hyperosmotic condition. For the minimal medium condition, the lowest quantity of ribosomes and ribosomal subunits were collected although more free RNA was present.

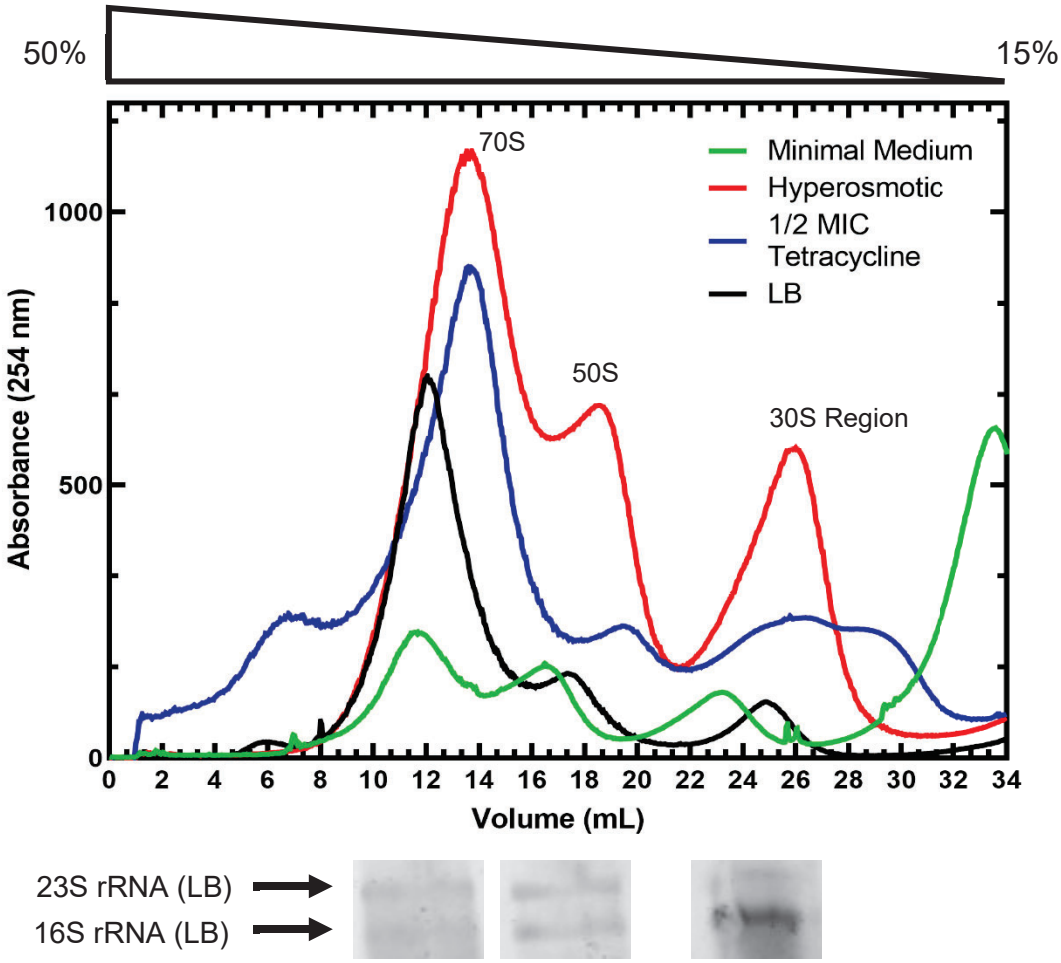


Figure 4.3. Induced stress conditions affect ribosomes and ribosomal subunit abundance and distributions. 50 OD_{260} of re-suspended ribosome pellets from *Escherichia coli* MG1655 cells grown with different stress conditions were loaded onto a 15-50% sucrose gradient and separated using ultracentrifugation. Gradients were fractionated and total RNA was extracted from select fractions (shown for LB condition below the gradient). Representative gradient profiles for each condition are shown.

To ensure that sRNAs would be present in the expected size range for bacteria (20-300 nt) and to reduce rRNA contamination, an enrichment step was necessary (63, 64). To this end, extracted RNA was separated using 10% (staining with Sybr Gold, only bands corresponding to rRNA were detectable (Figure S4.1)) or 15% urea PAGE and the part of each lane corresponding to below 300 nt was excised. In total, 39 samples (9 samples for the LB, hyperosmotic, and minimal medium conditions each and 12 samples for the tetracycline condition) were submitted for commercial cDNA library generation and Illumina sequencing.

4.3.2: Ribosomal RNA is present in all sequencing reads

Bioinformatic analyses for each condition was completed by comparing the LB control condition (LB 30S – LB30S, LB 50S – LB50S, LB 70S – LB70S) to the other three conditions (Minimal Medium 30S – MM30S, Minimal Medium 50S – MM50S, Minimal Medium 70S – MM70S, Hyperosmotic 30S – S30S, Hyperosmotic 50S – S50S, Hyperosmotic 70S – S70S, Tetracycline 30S peak1 – T30S1, Tetracycline 30S peak2 – T30S2, Tetracycline 50S – T50S, Tetracycline 70S – T70S). On average, 92.9% of reads across all samples mapped to the reference genome (Tables S4.1-4.4). At least two biological replicates for each sample had similar read counts with unique reads ranging from 239660 to 22256952 reads.

The majority of samples were found to have more reads compared to the LB control condition samples (MM70S – 7.4x more, MM50S – 35x more, MM30S – 28x more, S70S – 9x more, S50S – 13x more, S30S – 8.5x more) (Figure 4.4, 4.5). In contrast, the tetracycline condition yielded comparable or lower reads when compared to the control LB condition samples (T70S – 1.08x more, T50S – 2.68x less, T30S1 – 1.4x less, T30S2 – 2.2x less) (Figure 4.6, 4.7). Analysis of the number of reads with respect to the type of RNA (ncRNA, coding RNA, tRNA, or rRNA), revealed that rRNA related reads were the most abundant reads for all conditions and samples (MM – 98-99%, S – 98-99%, T – 89-96%, LB – 95-97%). Interestingly, the rRNA operons are shown to have reads enriched in specific locations of the rRNA sequence (Figure S4.2). The abundance of coding RNA and ncRNA reads is lower for all samples when compared to the LB control samples (MM: coding RNA - 0.05-0.49%, ncRNA – 0.05-0.57%, S: coding RNA – 0.25-0.64%, ncRNA – 0.1-0.43%, LB: coding RNA – 2.6-3.5%, ncRNA – 0.24-1.2%). All tetracycline samples, except for T70S, yielded more reads for coding and ncRNA when compared to the respective LB control sample (T: coding RNA – 3-9.6%, ncRNA – 0.17-0.83%, LB: coding RNA – 2.7-3.5%, ncRNA – 0.24-1.2%). The respective tRNA reads fractions are only higher for T30S1 (0.1%), T30S2 (0.28%), and T50S (0.1%) compared to the LB30S (0.02%) and LB50S (0.01%) samples.

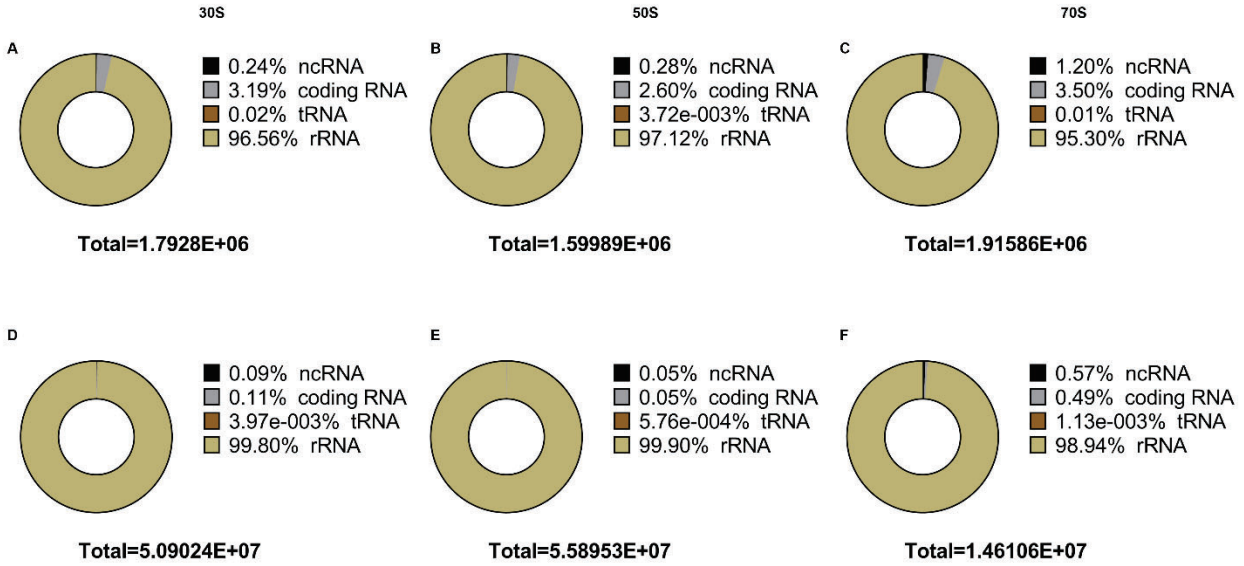


Figure 4.4. Reads for the minimal medium condition are mostly ribosomal RNA. The sum of the normalized hit counts for each classification (total indicated below each pie graph) were compared for the LB and minimal medium conditions using the average of three biological replicates after DESeq2 analysis. **(A-C).** Read counts distribution for LB condition for 30S, 50S, and 70S, respectively. **(D-F).** Read count distribution for minimal medium condition for 30S, 50S, and 70S, respectively.

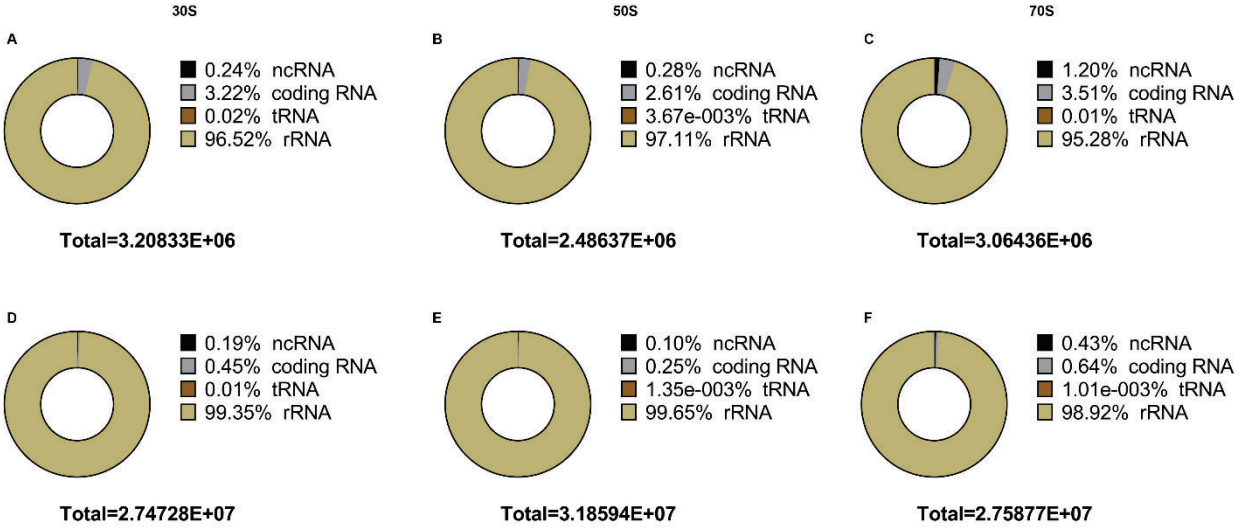


Figure 4.5. Ribosomal RNA makes up the majority of reads for the hyperosmotic condition. The sum of the normalized hit counts for each classification (total below each pie graph) were compared for the LB and hyperosmotic conditions using the average of three biological replicates after DESeq2 analysis. **(A-C).** Read count distribution for LB condition for 30S, 50S, and 70S, respectively. **(D-F).** Read count distribution for hyperosmotic condition for 30S, 50S, and 70S, respectively.

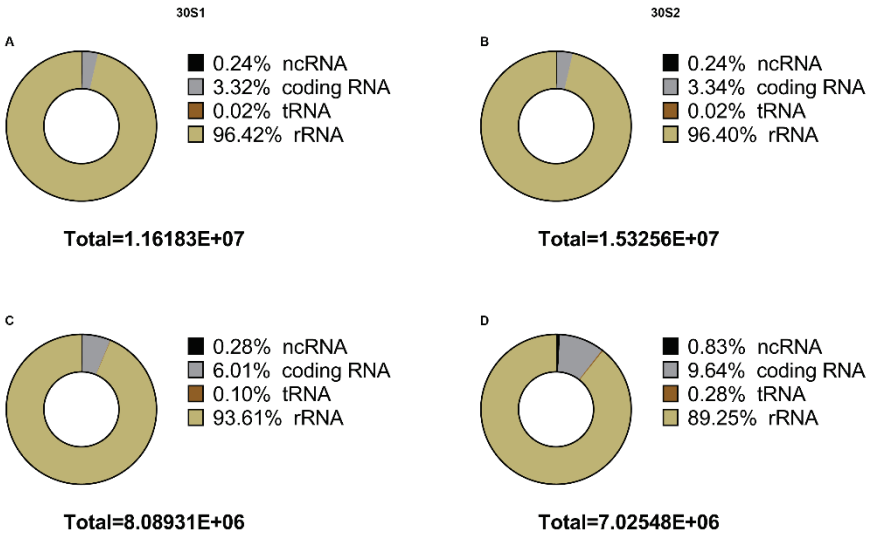


Figure 4.6. Coding RNA reads are more abundant in the 30S region during tetracycline stress. The sum of the normalized hit counts for each classification (total below each pie graph) were compared for the LB and tetracycline conditions using the average of three biological replicates after DESeq2 analysis. **(A, B).** Read count distribution for LB condition for the 30S. **(C, D).** Read count distribution for tetracycline condition for 30S1 and 30S2 peaks, respectively.

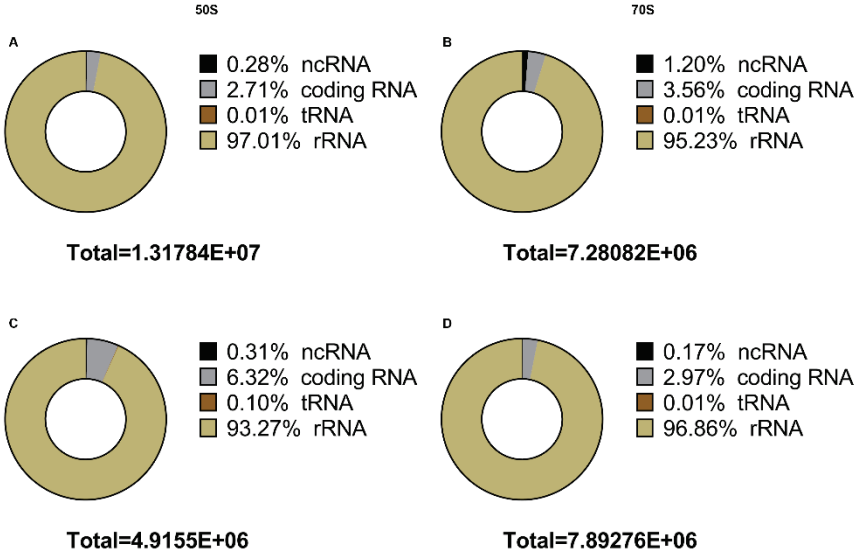


Figure 4.7. Increase in coding RNA and tRNA reads for the 50S tetracycline condition. The sum of the normalized hit counts for each classification (total below each pie graph) were compared for the LB and tetracycline conditions using the average of three biological replicates after DESeq2 analysis. **(A, B).** Read count distribution for LB condition for 50S and 70S, respectively. **(C, D).** Read count distribution for tetracycline condition for 50S and 70S, respectively.

4.3.3: Sequencing reads primarily map to Coding RNA genes

To determine what genes the obtained reads represent, a gene hit list for the different conditions were compared. Obtained reads primarily map to coding RNA (Figure 4.8, 4.9, 4.10, 4.11). Interestingly, the total number of hits are lower for the minimal medium condition when compared to the LB condition (MM70S – 19%, MM50S – 38%, and MM30S – 32%). However, there are only minimal differences between ncRNA, tRNA, and rRNA encoding genes for the minimal medium, hyperosmotic, and tetracycline conditions compared to the LB condition.

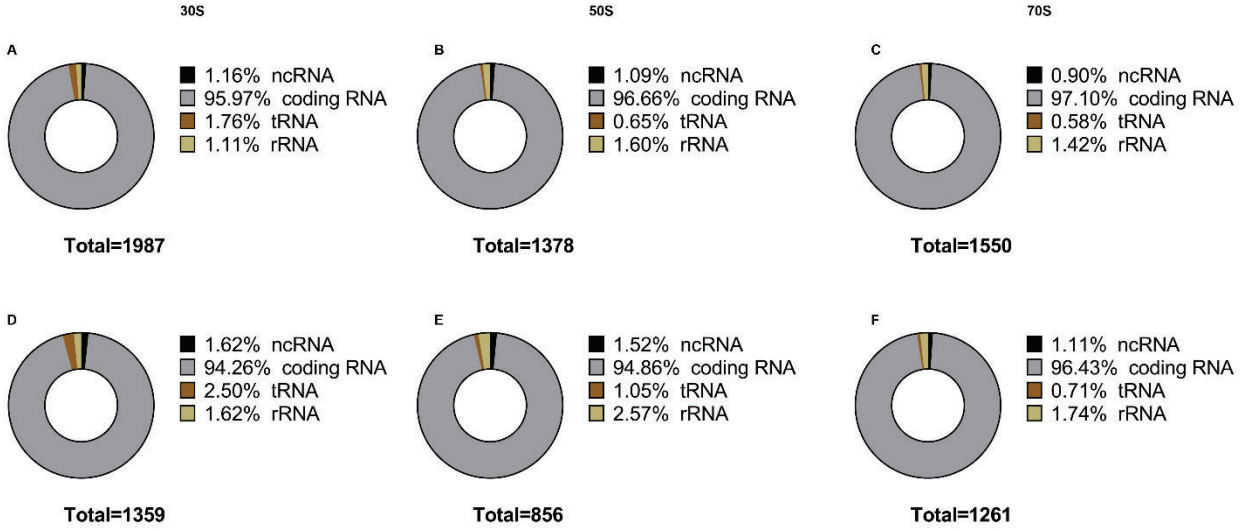


Figure 4.8. Less reads mapping to genes are identified in the minimal medium condition sample. The total number of genes identifies (total underneath each pie chart) identified in the DESeq2 analysis were compared between the LB and minimal medium conditions. Genes with 0 hit counts were not included in the total. **(A-C).** Fraction of reads mapping to genes present in LB condition for 30S, 50S, and 70S, respectively. **(D-F).** Fraction of reads mapping to genes present in minimal medium condition for 30S, 50S, and 70S, respectively.

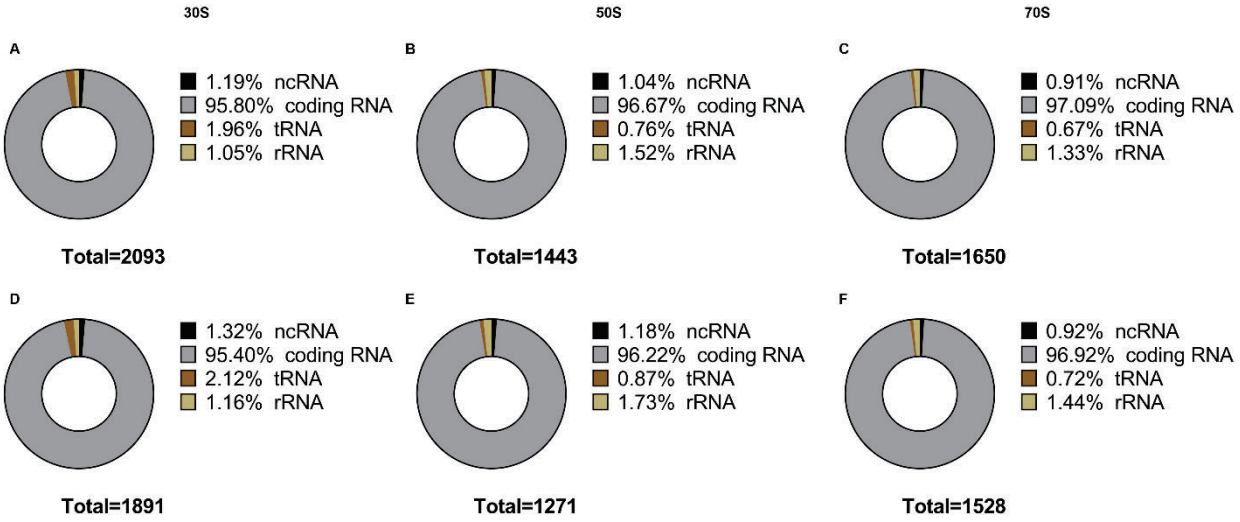


Figure 4.9. The hyperosmotic condition has a decrease in total reads mapping to genes. The total number of genes identified (total beneath each pie chart) from the DESeq2 analysis were compared for the LB and hyperosmotic conditions. Genes with 0 hit counts were not included in the total. **(A-C).** Fraction of reads mapping to genes in LB condition for 30S, 50S, and 70S, respectively. **(D-F).** Fraction of reads mapping to genes in hyperosmotic condition for 30S, 50S, and 70S, respectively.

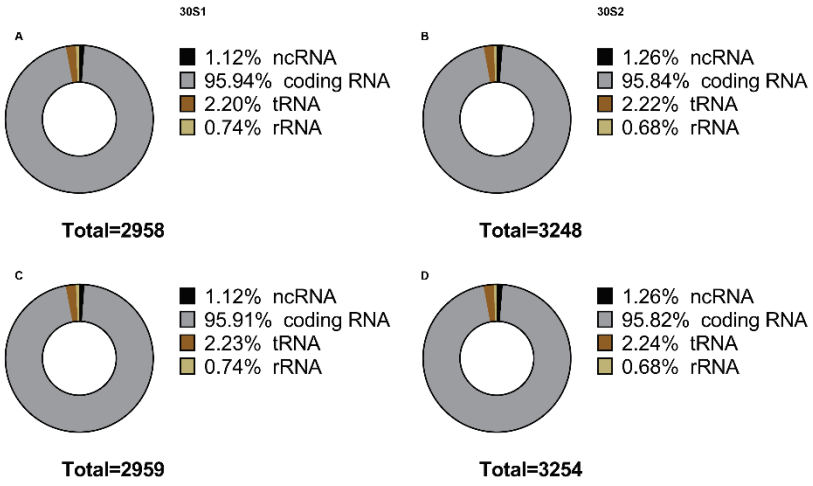


Figure 4.10. Total number of genes does not vary between tetracycline and LB conditions in the 30S region. The total number of genes (total underneath each pie chart) from the DESeq2 analysis were compared for the LB and tetracycline conditions. Genes with 0 hit counts were not included in the total. **(A, B).** Fraction of reads mapping to genes present in LB condition for the 30S. **(C, D).** Fraction of reads mapping to genes present in tetracycline condition for 30S1 and 30S2 peaks, respectively.

Figure 4.11. A higher number of tRNA gene related reads are present in the 50S than the 70S. The total number of genes (total underneath each pie chart) from the DESeq2 analysis were compared for the LB and tetracycline conditions. Genes with 0 hit counts were not included in the total. **(A, B).** Fraction of reads mapping to genes present in LB condition for the 50S and 70S, respectively. **(C, D).** Fraction of reads mapping to genes present in tetracycline condition for the 50S and 70S, respectively.

4.3.4: Few ncRNAs are present in ribosomal particles under minimal medium stress

To address if any ncRNAs present were significantly represented under minimal medium stress when compared to the LB control condition, the samples were analyzed using DESeq2 with a p-value of less than 0.05 being identified as significant. MM30S has the most ncRNAs identified with 23, while MM50S has 15 and MM70S has 14 (Figure 4.12). Overall, the proportion of normalized mean hit counts for all ncRNAs is higher under minimal medium condition (Figure 4.12, A-C, Table S4.5). However, the only significant hit identified in the minimal medium condition is ffs, which encodes the 4.5S RNA component of the signal recognition particle, in MM30S with a five-fold increase (p=0.040).

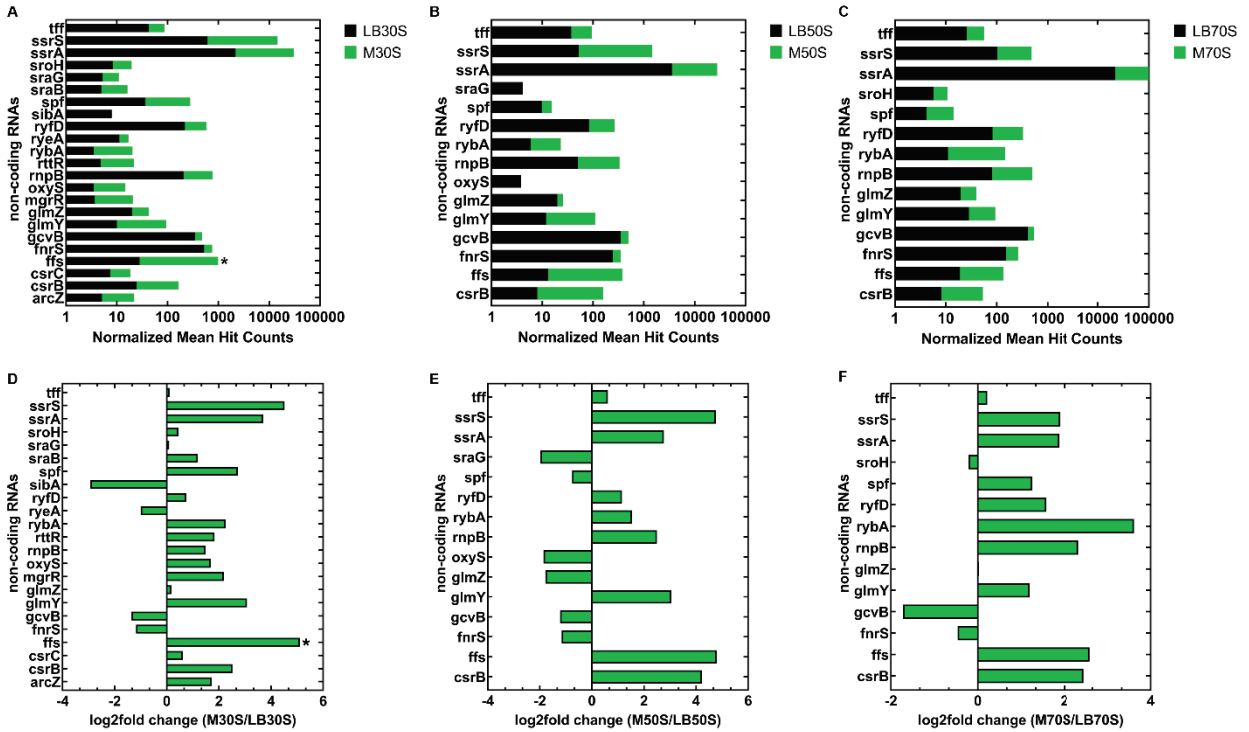


Figure 4.12. ffs is more abundant in the 30S in minimal medium (* = p < 0.05). The analysis was completed using DESeq2 (214). **(A-C).** The proportion of hit counts for ncRNAs between the LB and minimal medium conditions. The x-axis is on a logarithmic scale. **(D-F).** Fold changes of ncRNAs between the LB and minimal medium conditions.

Because the DESeq2 analysis is only able to identify annotated transcripts, additional analyses were conducted using Rockhopper to identify novel sRNAs in unannotated regions of the *E. coli* genome.

Strikingly, more ncRNAs were identified in this analysis for MM70S (45), MM50S (46), and MM30S (58) (Table S4.8). All ncRNAs identified in the DESeq2 analysis are also present in the Rockhopper analysis. Figure 4.13 summarizes the ncRNAs identified as having a significant differential abundance in comparison to the LB control condition ($q < 0.01$). Most ncRNAs are more abundant in MM70S (abundant ncRNAs: chiX – $q = 0.0001$, cpxQ – $q = 2.5 \times 10^{-5}$, gadY – $q = 7.55 \times 10^{-33}$, rydB – $q = 0.0039$, ryhB – $q = 0$, ryjB – $q = 6.58 \times 10^{-20}$, sdhX – $q = 9.00 \times 10^{-27}$; low abundance ncRNAs: gcvB – $q = 7.90 \times 10^{-16}$, ryeA – $q = 0.0023$, sroE – $q = 4.63 \times 10^{-77}$) and MM50S (abundant ncRNAs: rydB – $q = 1.75 \times 10^{-110}$, rydC – $q = 6.64 \times 10^{-99}$, ryhB – $q = 9.22 \times 10^{-136}$, ryjA – $q = 3.16 \times 10^{-19}$; low abundance ncRNA: gcvB – $q = 1.66 \times 10^{-47}$) compared to LB70S and LB50S, respectively. However, the majority of ncRNAs are in low abundance for MM30S (abundant ncRNAs: ffs – $q = 2.62 \times 10^{-29}$, rydB – $q = 2.05 \times 10^{-71}$, ryhB – $q = 0$; low abundance ncRNAs: agrA – $q = 3.31 \times 10^{-21}$, csrC – $q = 0.0046$, gcvB – $q = 0.0074$, micL – $q = 6.66 \times 10^{-41}$, rdlA – $q = 2.76 \times 10^{-12}$, rdlD – $q = 2.53 \times 10^{-11}$, ryeA – $q = 6.30 \times 10^{-9}$, ryfA – $q = 1.77 \times 10^{-5}$, sdsR – $q = 0.0037$, sibA – $q = 8.05 \times 10^{-87}$, sokC – $q = 1.37 \times 10^{-8}$, sroC – $q = 4.96 \times 10^{-91}$, symR – $q = 2.16 \times 10^{-12}$) when compared to LB30S. In agreement with both analysis techniques, fss is significantly abundant in MM30S. Interestingly, several ncRNAs are absent (transcript abundance = 0) in the minimal medium condition when compared to the LB condition (MM70S – 18, MM50S – 27, MM30S – 15) (Table S4.8).

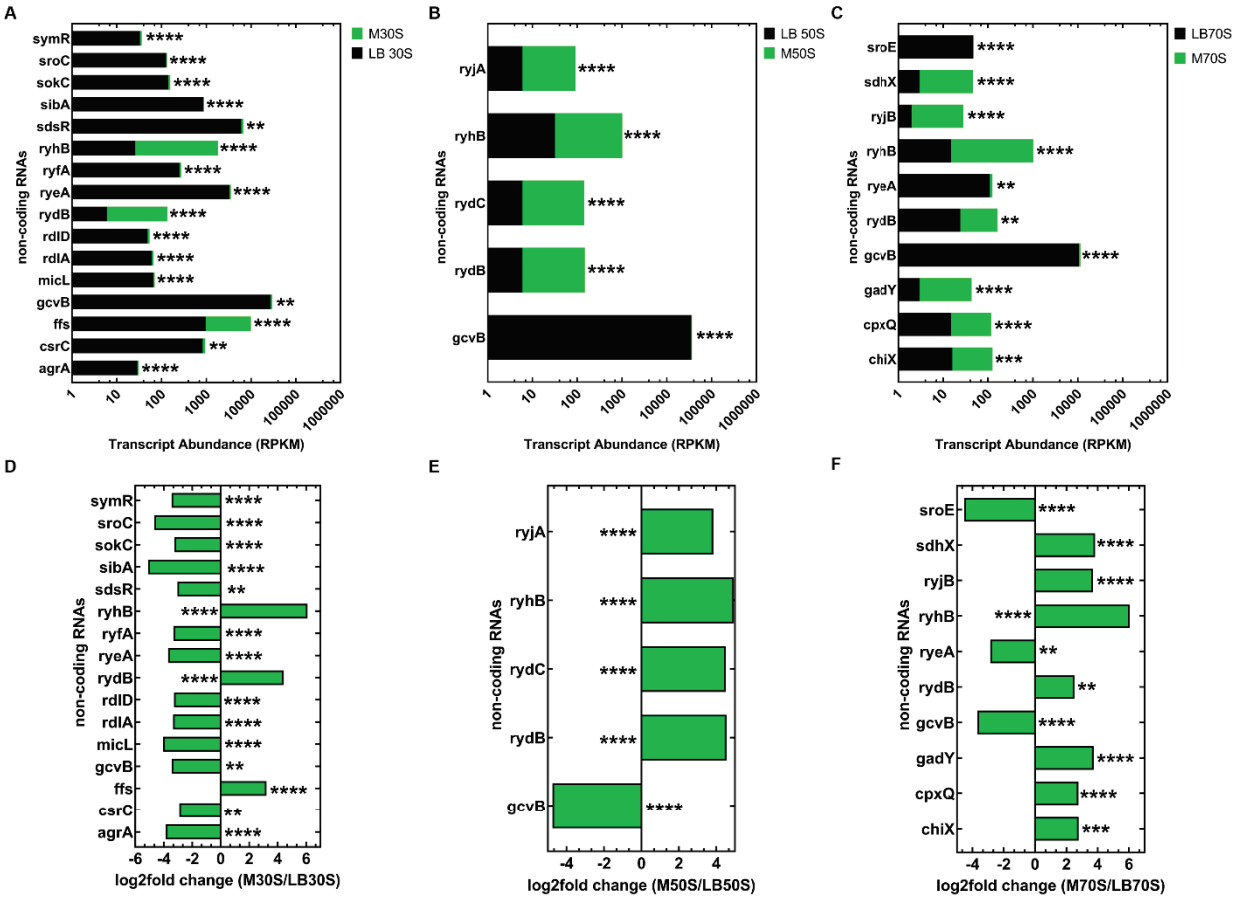


Figure 4.13. ncRNAs are predominantly lower in abundance for the minimal medium 30S compared to the LB 30S (** = q < 0.01, *** = q < 0.001, **** = q < 0.0001). The analysis was completed using Rockhopper (216-218). **(A-C).** Proportion of increased or reduced abundance transcripts when compared between the LB and minimal medium conditions. The x-axis is on a logarithmic scale. **(D-F).** Fold changes for high or low abundance ncRNAs.

4.3.5: ncRNAs are abundant with 50S and 30S ribosomal subunits under hyperosmotic stress

A comparable number of ncRNAs are identified for both the hyperosmotic condition (S70S – 15, S50S – 15, S30S – 25) and the minimal medium condition when compared to the corresponding LB control (Figure 4.14, Table S4.6). Three ncRNAs are identified for S30S (glmY – p = 0.007, ryeA – p = 0.037, and sibC – p = 0.047) and two ncRNAs for S50S (glmY – p = 0.006 and ssrS – p = 0.027).

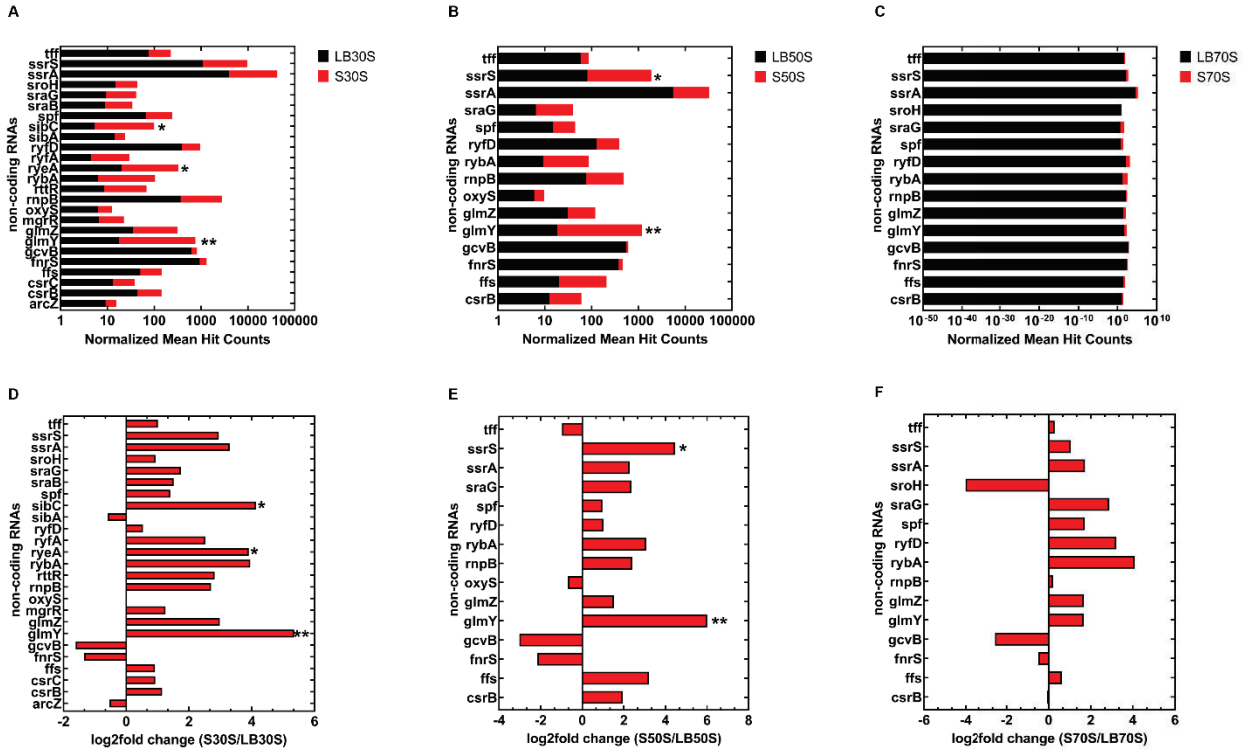


Figure 4.14. Five ncRNAs are significantly enriched in the 30S and 50S ribosomal subunits (* = p < 0.05, ** = p < 0.01). The analysis was completed using DESeq2 (214). **(A-C).** The proportion of hit counts for ncRNAs between the LB and hyperosmotic conditions. The x-axis is on a logarithmic scale. **(D-F).** Fold changes of ncRNAs between the LB and hyperosmotic conditions.

For the Rockhopper analysis, more ncRNAs are identified in S70S (43), S50S (52), and S30S (69) compared to the DESeq2 analysis (Table S4.9). A similar number of ncRNAs are present in high or low abundance for S50S (abundant ncRNAs: cpxQ – q = 0.0005, glmY – q = 8.74×10^{-30}, pspH – q = 0.0002, rdlB – q = 6.22×10^{-16}, rdlC – q = 1.24×10^{-41}, rttR – q = 4.66×10^{-25}, sroA – q = 1.17×10^{-11}; low abundance ncRNAs: arcZ – q = 2.70×10^{-25}, gcvB – q = 3.49×10^{-12}, micA – q = 4.81^{-36}, oxyS – q = 1.56×10^{-19}, ryjB – q =

0.0002, sibA – q = 1.41x10^{-11}, sibB – q = 0.0052) and S30S (abundant ncRNAs: cpxQ – q = 1.28x10^{-6}, micF – q = 3.14x10^{-9}, rdlC – q = 7.20x10^{-11}, rttR – q = 5.03x10^{-10}, rydB – q = 9.17x10^{-96}; low abundance ncRNAs: fnrS – q = 4.63x10^{-7}, gadY – q = 3.26x10^{-6}, gcvB – q = 0.0004, mgrR – q = 8.88x10^{-5}, rprA – q = 8.99x10^{-17}, sibA – q = 4.08x10^{-167}, sibB – q = 2.96x10^{-28}) when compared to LB50S and LB30S, respectively (Figure 4.15 A-D). In contrast, almost all ncRNAs are present with high abundance in S70S (abundant ncRNAs: 3'ETS-leuZ – q = 1.26x10^{-98}, cpxQ – q = 1.20x10^{-6}, glmY – q = 3.44x10^{-7}, istR – q = 3.17x10^{-11}, rttR – q = 0.0019, ryfD – q = 0.0016, sdhX – q = 3.65x10^{-9}; low abundance ncRNAs: gcvB – q = 0.0012) compared to LB70S (Figure 4.15 C, F). Interestingly, a predicted RNA (PR1S70S, q = 9.14x10^{-110}) is also identified for S70S. Additionally, some ncRNAs are absent in the hyperosmotic condition (S70S – 24, S50S – 19, S30S – 9) when compared to the LB control condition (Table S4.9).

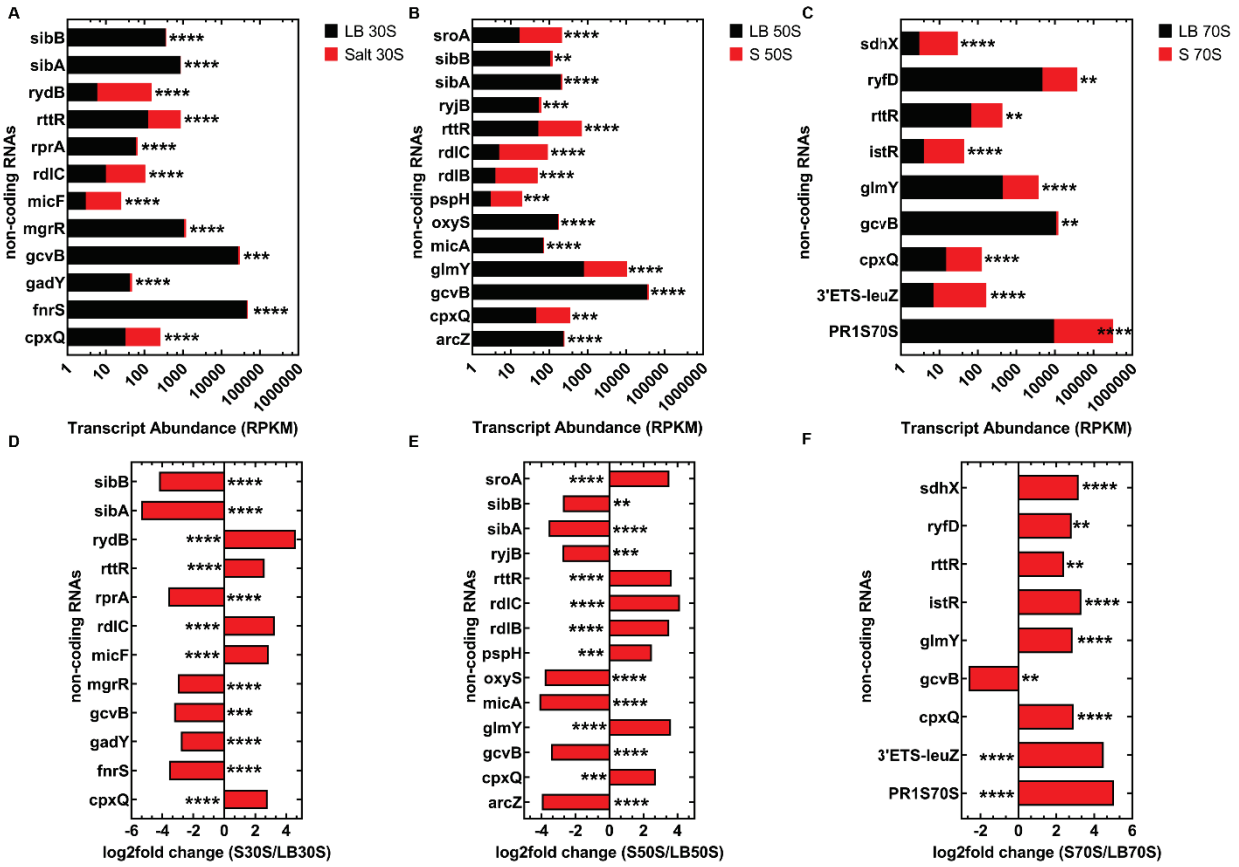


Figure 4.15. ncRNAs are significantly enriched in the hyperosmotic 70S when compared to the LB 70S (** = q < 0.01, *** = q < 0.001, **** = q < 0.0001). The analysis was completed using Rockhopper (216-218). **(A-C).** Proportion of significant transcripts enriched when comparing the LB and hyperosmotic conditions. The x-axis is on a logarithmic scale. **(D-F).** Fold changes for significant ncRNAs.

4.3.6: ncRNAs are enriched in the Tetracycline 30S2 peak

Of all the conditions, the T30S1 and T30S2 peaks have the most ncRNAs present (T70S – 20, T50S – 33, T30S1 – 33, T30S2 – 41) (Figure 4.16, Table S4.7). Remarkably, more ncRNAs are abundant in the 30S2 peak compared to LB30S. Additionally, both peaks have the largest number significantly enriched or depleted ncRNAs when compared to LB30S, with T30S1 having seven (abundant ncRNAs: cyaR – p = 0.008, mgrR – p = 0.032, oxyS – p = 0.008, rttR – p = 0.038, ryjA – p = 0.047, sraB – p = 0.038; low abundance ncRNA: ffs – p = 0.039) and T30S2 having eight (abundant ncRNAs: cyaR – p = 0.029, mgrR – p = 0.037, oxyS – 0.020, rttR – p = 0.004, ryfA – p = 0.043, ryjA – p = 0.020, sibD – p = 0.014, sraB – p = 0.022).

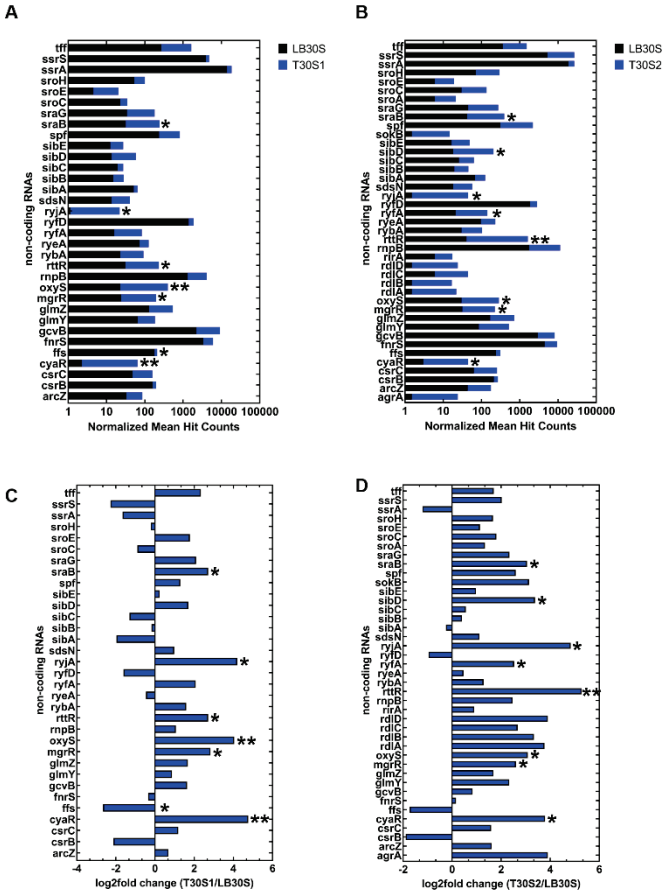


Figure 4.16. More ncRNAs are enriched in the 30S2 peak than in the 30S1 peak (* = p < 0.05, ** = p < 0.01). The analysis was completed using DESeq2 (214). **(A, B).** The proportion of hit counts for ncRNAs between the LB and tetracycline conditions. The x-axis is on a logarithmic scale. **(C, D).** Fold changes of ncRNAs when comparing between the LB and tetracycline conditions.

While T50S has a higher number of ncRNAs identified (33), T70S (20) is similar to the minimal medium and hyperosmotic conditions (Figure 4.17, Table S4.7). No ncRNAs are found to be significantly in high or low abundance for T70S when compared to LB70S. For T50S, three ncRNAs show significantly differential abundance (abundant ncRNAs: rttR – p = 0.010, ryjA – p = 0.049; low abundance ncRNA: ssrA – p = 0.046).

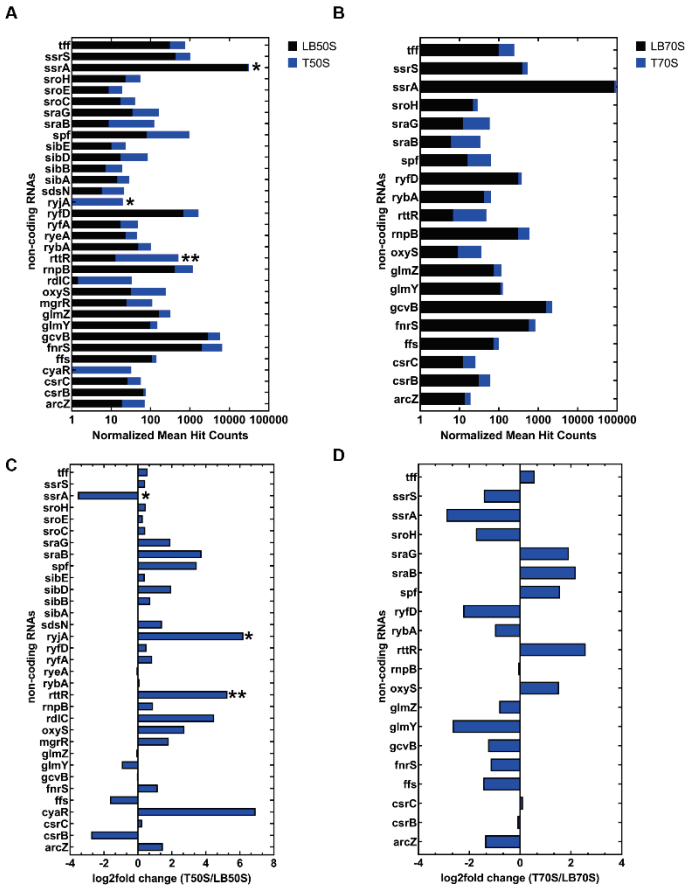


Figure 4.17. No ncRNAs are significantly more abundant or in lower abundance for the 70S ribosome when compared to the LB 70S (* = p > 0.05, ** = p > 0.01). The analysis was completed using DESeq2 (214). **(A, B).** The proportion of hit counts for ncRNAs between the LB and tetracycline conditions. The x-axis is on a logarithmic scale. **(C, D).** Fold changes of ncRNAs when comparing between the LB and tetracycline conditions.

For the Rockhopper analysis, more ncRNAs are present in the tetracycline condition when compared to the hyperosmotic and minimal medium conditions (T70S – 72, T50S – 70, T30S1 – 74, T30S2 – 78) (Table S4.10). Strikingly, more ncRNAs have significantly lower abundance in T30S1 when compared

to LB30S (abundant ncRNAs: cyaR – q = $2.63x10^{-32}$, oxyS – q = 0.0008, rdlB – q = $4.38x10^{-19}$, rdlC – q = $2.93x10^{-27}$, rttR – q = $1.17x10^{-58}$, rydC – q = 0.0022, sokE – q = $2.81x10^{-7}$; low abundance ncRNAs: csrB – q = $1.88x10^{-281}$, ffs – q = 0.0004, fnrS – q = 0.0056, glmY – q = $4.91x10^{-8}$, omrA – q = $1.18x10^{-5}$, omrB – q = $1.34x10^{-39}$, rprA – q = $1.67x10^{-5}$, ryeA – q = $4.11x10^{-123}$, ryfD – q = $1.08x10^{-7}$, sdsR – q = $2.22x10^{-85}$, sgrS – q = $1.06x10^{-6}$, sibA – q = $6.16x10^{-67}$, sibB – q = $2.49x10^{-14}$, sokB – q = $1.80x10^{-5}$, sroH – q = 0.0011, ssrS – q = $2.03x10^{-158}$) (Figure 4.18 A, C).

In T30S2, there is a proportional distribution of significantly high and low abundance ncRNAs in comparison to LB30S (abundant ncRNAs: 3'ETS-leuZ – q = $1.11x10^{-8}$, cyaR – q = 0.0003, rdlB – q = $6.69x10^{-18}$, rdlC – q = $1.76x10^{-62}$, rdlD – q = 0.0021, rttR – q = 0, rydC – q = 0.0056, ryfA – q = $2.50x10^{-5}$, ryjA – q = $6.42x10^{-6}$, sibD – q = 0.0002, sokE – q = $6.20x10^{-16}$; low abundance ncRNAs: cpxQ – q = $1.98x10^{-7}$, csrB – q = $4.96x10^{-91}$, omrB – q = $2.37x10^{-86}$, ryeA – q = $7.88x10^{-6}$, sdsR – q = $7.71x10^{-5}$, sgrS – q = 0.0014, sibA – q = $8.55x10^{-9}$, sibB – q = $6.93x10^{-6}$) (Figure 4.18 B, D). One predicted RNA was found for T30S1 (PR1T30S1 – q = $9.70x10^{-41}$) and four for T30S2 (PR1T30S2 – q = $1.43x10^{-16}$, PR2T30S2 – q = 0, PR3T30S2 – q = 0, PR4T30S2 – q = $3.14x10^{-13}$). Similarly, to the hyperosmotic and minimal medium conditions, some ncRNAs are absent in the tetracycline condition compared to the LB condition (T30S1 – 5, T30S2 – 4) (Table S4.10).

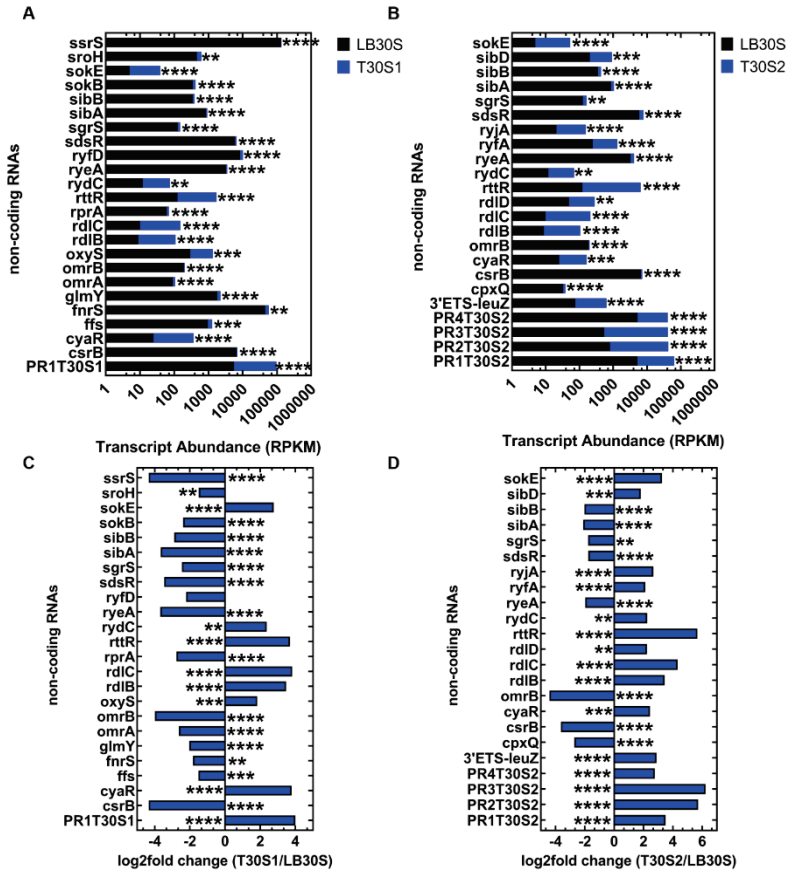


Figure 4.18. ncRNAs are in low abundance in the 30S1 peak and abundant in the 30S2 peak when compared to the LB 30S (** = q < 0.01, *** = q < 0.001, **** = q < 0.0001). The analysis was completed using Rockhopper (216-218). **(A, B).** Proportion of significant transcripts between the LB and tetracycline conditions. The x-axis is on a logarithmic scale. **(C, D).** Fold changes for significant ncRNAs.

The majority of ncRNAs are significantly abundant in both T50S (abundant ncRNAs: 3'ETS-leuZ – $q = 5.03 \times 10^{-7}$, agrB – $q = 3.83 \times 10^{-6}$, eyeA – $q = 8.26 \times 10^{-5}$, oxyS – $q = 8.26 \times 10^{-7}$, rdlB – $q = 3.16 \times 10^{-126}$, rdlC – $q = 0$, rdlD – $q = 8.14 \times 10^{-5}$, rirA – $q = 5.71 \times 10^{-9}$, rttR – $q = 0$, rydC – $q = 1.03 \times 10^{-10}$, ryfA – $q = 3.57 \times 10^{-7}$, ryjA – $q = 3.32 \times 10^{-41}$, spf – $q = 0.0053$, symR – $q = 6.35 \times 10^{-34}$; low abundance ncRNAs: cpxQ – $q = 9.88 \times 10^{-33}$, csrB – $q = 1.37 \times 10^{-19}$, ffs – $q = 7.64 \times 10^{-7}$, glmY – $q = 0.0068$, sgrS – $q = 3.62 \times 10^{-8}$, ssrA – $q = 7.24 \times 10^{-5}$, ssrS – $q = 0.0026$) and T70S (abundant ncRNAs: 3'ETS-leuZ – $q = 9.57 \times 10^{-92}$, cyaR – $q = 2.62 \times 10^{-7}$, eyeA – $q = 0.0020$, mcaS – $q = 0.0068$, rdlC – $q = 2.66 \times 10^{-37}$, rttR – $q = 1.42 \times 10^{-43}$, ryjB – $q = 1.80 \times 10^{-83}$, sdhX – $q = 5.09 \times 10^{-9}$, sdsN – $q = 4.92 \times 10^{-26}$, sibD – $q = 0.0007$, sraB – $q = 0.0040$, sroE – $q = 0.0062$; low abundance ncRNA: csrB – $q = 0.0054$) when compared to LB50S and LB70S, respectively (Figure 4.19).

Two predicted RNAs are also identified for the 50S subunit (PR1T50S – q = $8.58x10^{-9}$, PR2T50S – q = $1.77x10^{-36}$). As ncRNAs are absent in the 30S peaks, there are also some ncRNAs absent for T70S (4) and T50S (7) when compared to LB70S and LB50S, respectively (Table S4.10).

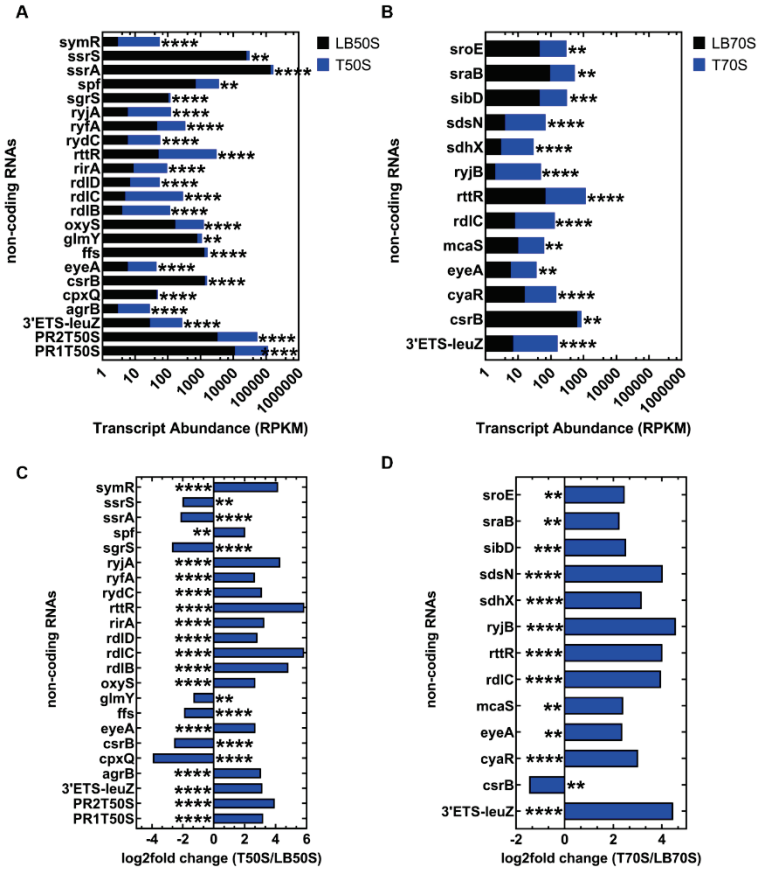


Figure 4.19. Majority of the ncRNAs are abundant in both the 50S and 70S when compared to the LB 50S and 70S ribosomal particles (** = q < 0.01, *** = q < 0.001, **** = q < 0.0001). The analysis was completed using Rockhopper (216-218). **(A, B).** Proportion of significant transcripts comparing the LB and tetracycline conditions. The x-axis is on a logarithmic scale. **(C, D).** Fold changes for significant ncRNAs.

4.3.7: Comparison between conditions reveals a common response to stress

Looking at the overlap of all ncRNAs identified from the DESeq2 analysis in all conditions, most ncRNAs are present in all stress conditions, although tetracycline did have some ncRNAs exclusively identified in all sequenced particles (Figure 4.20, A-C)). Additionally, the hyperosmotic and tetracycline conditions shared a couple ncRNAs (ryfA, sibC) while the minimal medium condition did not have any

identified exclusively. Looking at the ncRNAs identified as significantly abundant or in low abundance compared to the LB condition, all are found to be present in only one stress condition, except for ffs (Figure 4.20, D-E).

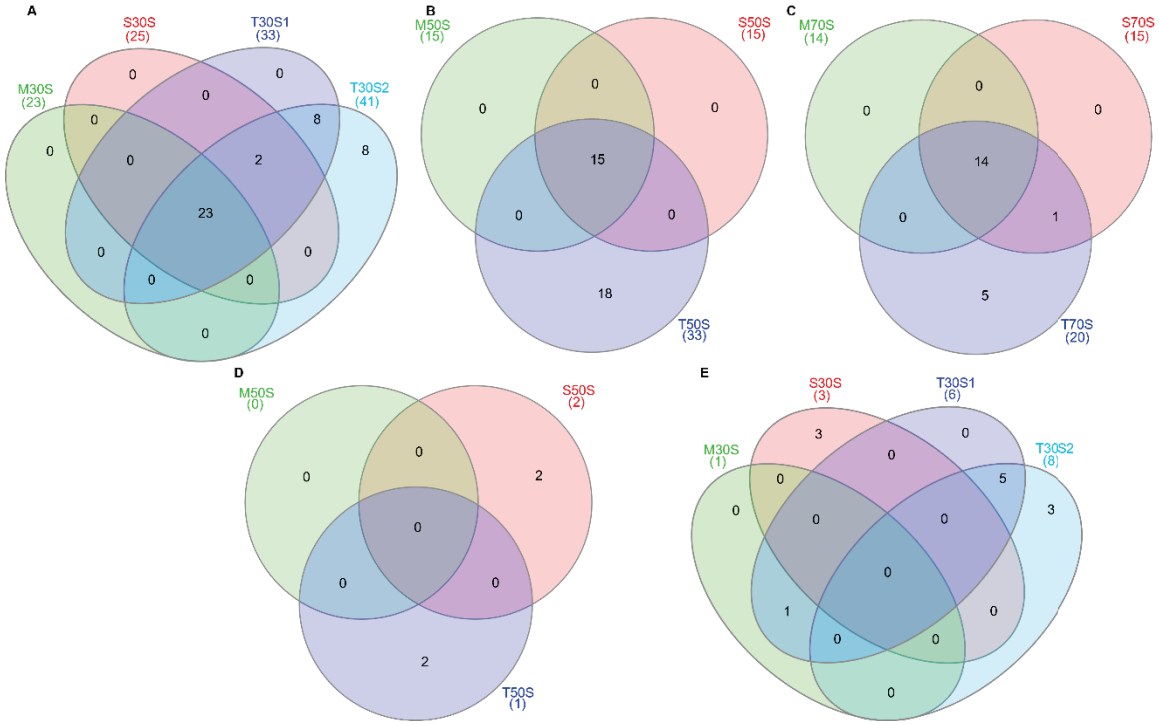


Figure 4.20. The majority of ncRNAs are shared between stress conditions but significantly enriched ncRNAs are condition specific (145). All ncRNAs identified in the **(A)** 30S, **(B)** 50S, and **(C)** 70S. Significantly enriched ncRNAs are shown in the **(D)** 50S and **(E)** 30S.

For all ncRNAs identified using Rockhopper, the majority are also shared between all conditions tested (Figure 4.21, A-C). No ncRNA unique to T30S1 and M70S are identified. For the significantly enriched or low abundance ncRNAs when compared to the LB condition identified, most are also specific to one stress condition (Figure 4.21, D-F). However, sibA is shared between all conditions in the 30S and sdhX is shared between all conditions for the 70S.

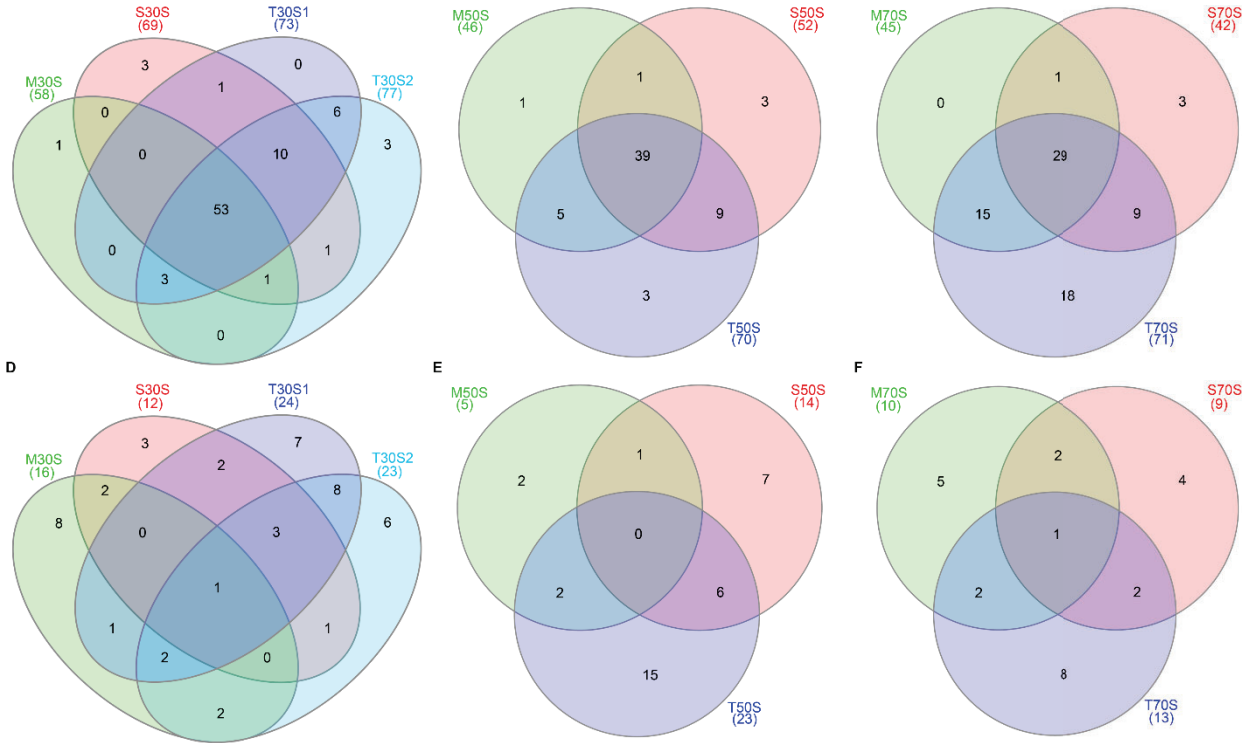


Figure 4.21. Most ncRNAs are shared across all conditions while significantly enriched ncRNAs are condition specific (145). All ncRNAs found in the **(A)** 30S, **(B)** 50S, and **(C)** 70S. Significantly enriched ncRNAs identified in the **(D)** 30S, **(E)** 50S, and **(F)** 70S.

The ncRNAs identified as significantly enriched or in low abundance when compared to the LB condition using DESeq2 analysis are listed in Table 4.1. In total, 14 ncRNAs are found to be significantly enriched or in low abundance across all conditions. All the ncRNAs have been annotated although their cellular functions are not well characterized.

Table 4.1. Significant (p > 0.05) ncRNAs identified in all conditions through DESeq2 analysis.

Condition	ncRNA	Size (nt)	Known Information	Reference
Minimal Medium (30S), Tetracycline (30S1)	ffs	114	Signal recognition particle 4.5S RNA, can bind to Elongation Factor G and ribosomes	(220, 221)
Hyperosmotic (30S, 50S)	glmY	184	Increases synthesis of GlmS, antagonizes glmZ,	(222)
Hyperosmotic (30S)	ryeA	272	Also known as sraC, antisense RNA	(223)
Hyperosmotic (30S)	sibC	141	Antitoxin with IbsC	(89, 224)
Hyperosmotic (50S)	ssrS	183	6S RNA, stationary phase regulation of transcription of sigma70	(225)
Tetracycline (30S1, 30S2)	cyaR	87	Promotes degradation of *ompX, yqaE, nadE, luxS, rpoS*	(89, 200, 202, 226)
Tetracycline (30S1, 30S2)	mgrR	98	May affect lipopolysaccharide composition	(227)
Tetracycline (30S1, 30S2)	oxyS	117	Oxidative stress response, antimutator effect	(228)
Tetracycline (30S1, 30S2, 50S)	rttR	171	May be released from *tyrT* operon during processing	(143)
Tetracycline (30S1, 30S2)	sraB	169	Usually expressed in late stationary phase	(146)
Tetracycline (30S2)	ryfA	304	Not much known, in *Salmonella*, encodes toxic protein TimP	(89, 229)
Tetracycline (30S2)	ryjA	141	Also known as sraL, expressed in stationary phase	(230)
Tetracycline (30S2)	sibD	145	Antitoxin RNA	(89, 224)
Tetracycline (50S)	ssrA	363	Transfer-messenger RNA, releases stalled ribosomes	(231)

Significantly enriched or low abundance ncRNAs identified using Rockhopper are shown in Table 4.2. Using this analysis, 67 ncRNAs are found to be significantly enriched or in low abundance when compared to the LB condition. The majority of the ncRNAs are annotated. However, eight predicted RNAs have been discovered in the hyperosmotic and tetracycline conditions.

Table 4.2. Significant (q > 0.01) ncRNAs identified in all conditions through Rockhopper analysis.

Condition	ncRNA	Size (nt)	Known Information	Reference
Minimal Medium (30S)	agrA	82	Encoded in the *arsR-gor* intergenic region	(232)
Minimal Medium (70S)	chiX	84	Negatively regulates DpiA/DpiB	(233)
Minimal Medium (70S), Hyperosmotic (all), Tetracycline (30S2, 50S)	cpxQ	58	Located in the 3' UTR of cpxP mRNA, reduces CpxP protein levels	(234)
Minimal Medium (30S)	csrC	246	Modulates regulation of carbon storage regulator CsrA	(196, 235)
Minimal Medium (30S), Tetracycline (30S1, 50S)	ffs	114	Signal recognition particle 4.5S RNA, can bind to Elongation Factor G and ribosomes	(220, 221)
Minimal Medium (70S), Hyperosmotic (30S)	gadY	106	Positively regulates *gadX* and *gadW*, increases the expression of several genes	(236, 237)
Minimal Medium (all), Hyperosmotic (all)	gcvB	206	Regulates expression of periplasmic transporter components DppA and OppA	(146, 201)
Minimal Medium (30S)	micL	80	Negatively regulates *lpp*	(238)
Minimal Medium (30S)	rdlA	67	Putative antisense regulatory RNA	(239)
Minimal Medium (30S), Tetracycline (30S2, 50S)	rdlD	64	Acts as a Type I antitoxin	(239)
Minimal Medium (all), Hyperosmotic (30S)	rydB	68	Overexpression decreases *rpoS*, highest levels in minimal medium	(89)
Minimal Medium (50S), Tetracycline (30S1, 30S2, 50S)	rydC	65	Activates expression of Cfa during acid shock	(240, 241)
Minimal Medium (30S, 70S), Tetracycline (30S1, 30S2)	ryeA	272	Also known as sraC, antisense RNA	(223)
Minimal Medium (30S), Tetracycline (30S2, 50S)	ryfA	304	Not much known, in *Salmonella*, encodes toxic protein TimP	(89, 229)
Minimal Medium (all)	ryhB	95	Reduces iron consumption under low-iron conditions	(5, 242, 243)
Minimal Medium (50S), Tetracycline (30S2)	ryjA	141	Also known as sraL, expressed in stationary phase	(230)
Minimal Medium (70S), Hyperosmotic (50S), Tetracycline (70S)	ryjB	90	May stabilize *rnb* mRNA, expression increases at higher concentrations of magnesium	(244, 245)
Minimal Medium (70S), Hyperosmotic (70S), Tetracycline (70S)	sdhX	101	Connects TCA cycle gene to acetate metabolism, represses AckA, FdoG, and KatG	(246, 247)
Minimal Medium (30S), Tetracycline (30S1, 30S2)	sdsR	104	Deletion decreases ampicillin-induced mutagenesis, targets *tolC* and *yhcB*	(248, 249)

Minimal Medium (30S), Hyperosmotic (30S, 50S), Tetracycline (30S1, 30S2)	sibA	144	Small RNA in the *yegL-yegM* intergenic region, may be in a toxin-antitoxin system with *ibsA*	(89, 224)
Minimal Medium (30S)	sokC	55	Downregulates expressin of HokC toxin	(250)
Minimal Medium (30S)	sroC	163	Destabilizes GcvB	(91, 251)
Minimal Medium (70S), Tetracycline (70S)	sroE	92	May interact with *ligB*	(151, 252)
Minimal Medium (30S), Tetracycline (50S)	symR	77	In a toxin-antitoxin pair with *symE*	(253)
Hyperosmotic (70S), Tetracycline (30S2, 50S, 70S)	3'ETS-leuZ	67	Acts as a sponge for transcriptional noise	(9)
Hyperosmotic (50S)	arcZ	122	Increases translation of sigma factor 38 (RpoS); inhibits Rho-dependent transcription termination	(146-150)
Hyperosmotic (30S), Tetracycline (30S1)	fnrS	122	Metabolism regulation during anaerobiosis, reported expression only under anaerobic conditions	(254, 255)
Hyperosmotic (50S, 70S), Tetracycline (30S1, 50S)	glmY	184	Increases synthesis of GlmS, antagonizes glmZ,	(222)
Hyperosmotic (70S)	istR	140	Part toxin-antitoxin system with TisB	(256)
Hyperosmotic (30S)	mgrR	98	May affect lipopolysaccharide composition	(227)
Hyperosmotic (50S)	micA	73	Downregulates OmpA expression	(77, 257)
Hyperosmotic (30S)	micF	93	Negatively regulates OmpF in high osmolarity	(258)
Hyperosmotic (50S), Tetracycline (30S1, 50S)	oxyS	117	Oxidative stress response, antimutator effect	(228)
Hyperosmotic (50S)	pspH	111	Able to base pair with Spot 42	(259)
Hyperosmotic (50S), Tetracycline (30S1, 30S2, 50S)	rdlB	66	Putative antisense regulatory RNA	(239)
Hyperosmotic (30S, 50S), Tetracycline (all)	rdlC	68	Putative antisense regulatory RNA	(239)
Hyperosmotic (70S), Tetracycline (30S1)	ryfD	143	With small RNA pair RaiZ, associates with ProQ	(260, 261)
Hyperosmotic (30S), Tetracycline (30S1)	rprA	107	Produce RpoS during osmotic shock	(152, 153)
Hyperosmotic (all), Tetracycline (all)	rttR	171	May be released from *tyrT* transcript during processing	(143)
Hyperosmotic (30S, 50S), Tetracycline (30S1, 30S2)	sibB	136	Small RNA in the *yegL-yegM* intergenic region, may be in a toxin-antitoxin system with *ibsB*	(89, 224)
Hyperosmotic (50S)	sroA	93	May be generated during attenuation of *thiBPQ*	(151)
Hyperosmotic (70S)	PR1S70S	99	May be in the 5' UTR of cold shock protein CspG	(262)

Tetracycline (50S)	agrB		82	A part of toxin-antitoxin system with DinQ	(232)
Tetracycline (all)	csrB	369	Modulates regulation of carbon storage regulator CsrA		(196, 235)
Tetracycline (30S1, 30S2, 70S)	cyaR	87	Promotes degradation of *ompX, yqaE, nadE, luxS, rpoS*		(89, 200, 202, 226)
Tetracycline (50S, 70S)	eyeA	75	Increases at higher concentrations of magnesium		(245)
Tetracycline (70S)	mcaS	94	Regulates expression of curli by inhibiting translation of *csgD* mRNA, part of a ribonucleoprotein complex with S1		(263, 264)
Tetracycline (30S1)	omrA	88	Regulates protein composition of the outer membrane		(89, 146, 151, 179)
Tetracycline (30S1, 30S2)	omrB	84	Regulation of outer membrane protein composition		(151)
Tetracycline (50S)	rirA	73	Interacts with the transcription antiterminator RfaH, overexpression decreases lipopolysaccharides		(265)
Tetracycline (70S)	sdsN		137	Negatively regulates NfsA nitric oxide dioxygenase, and NarP	(199)
Tetracycline (30S1, 30S2, 50S)	sgrS	227	Regulates glucose transporter PtsG		(266)
Tetracycline (30S2, 70S)	sibD		145	Antitoxin RNA	(89, 224)
Tetracycline (30S1)	sokB	56	Putative antisense RNA, may regulate MokB and HokB		(267)
Tetracycline (30S1, 30S2)	sokE		59	A part of toxin-antitoxin system with HokE	(267)
Tetracycline (50S)	spf		111	Regulates *galK,* can block 30S from binding	(5, 268)
Tetracycline (70S)	sraB		169	Usually expressed in late stationary phase	(146)
Tetracycline (30S1)	sroH		161	Interacts with *kbl* mRNA	(252)
Tetracycline (50S)	ssrA		363	Transfer-messenger RNA, releases stalled ribosomes	(231)
Tetracycline (30S1, 50S)	ssrS		183	6S RNA, stationary phase regulation of transcription of sigma70	(225)
Tetracycline (30S1)	PR1T30S1		100	May be in the 5’ UTR of cold shock protein CspG	(262)
Tetracycline (30S2)	PR1T30S2		126	May be in the 5’ UTR of cold shock protein CspG	(262)
Tetracycline (30S2)	PR2T30S2		85	185 nt downstream of cold shock protein CspB	(269, 270)
Tetracycline (30S2)	PR3T30S2		85	185 nt downstream of cold shock protein CspB	(269, 270)
Tetracycline (30S2)	PR4T30S2		36	Located in 5’ UTR of cold shock protein CspB	(269-271)
Tetracycline (50S)	PR1T50S		119	May be in the 5’ UTR of cold shock protein CspG	(262)
Tetracycline (50S)	PR2T50S		88	Located in 5’ UTR of cold shock protein CspB	(269-271)

4.4 Discussion

rancRNAs have been identified in all domains of life (14, 16-26). However, the cellular and molecular mechanisms of rancRNAs are still poorly understood, especially in bacteria. To address this issue, we investigated sRNAs that associate with 70S ribosomes and 50S/30S ribosomal subunits from *E. coli* MG1655 cells grown under different stress conditions (minimal medium, hyperosmotic, and tetracycline stress).

To obtain sRNAs associated with ribosomes and the respective subunits, *E. coli* MG1655 cells were grown under one of the selected stress conditions, ribosomes were isolated using sucrose density gradient ultracentrifugation (Figure 4.3). As expected, three peaks corresponding to the 70S ribosome and 50S/30S ribosomal subunits contained 23S rRNA and 16S rRNA (Figure 4.3). The presence of 23S rRNA in the 30S fraction and 16S rRNA in the 50S fraction is likely caused by cross contamination between the 70S ribosome or 50S subunit, which has been observed previously (272). An additional peak in the 30S region is present in the tetracycline condition, but no other condition (Figure 4.3). Previous reports have observed the presence of such a peak (273-275). However, no detailed characterization of its composition has been reported so far. It has been speculation that this peak results from ribosome biogenesis being compromised as RNAse III is inhibited by tetracycline *in vitro* (276). Further studies are required to determine the protein and RNA composition of the 30S region during tetracycline stress.

The analysis reported here reveals that more 50S and 30S ribosomal subunits are present in the hyperosmotic stress condition samples (Figure 4.3), consistent with reports by Hase *et al.* (277). Under the minimal medium stress condition, 70S ribosomes and 50S/30S ribosomal subunits are in much lower abundance compared to the other conditions (Figure 4.3). During nutrient starvation, *E. coli* is known to reduce ribosome biogenesis to maintain translation of stress factors that promote cell survival (278, 279). Additionally, there is evidence that rRNA is degraded during nutrient starvation (280-282). This would result in fewer ribosomes and an increase in the free RNA as observed here. Based on the reported behaviours of ribosome particles under the tested stress conditions, it is likely that the cellular conditions used in this study are sufficient to induce the desired stress effects.

Sequencing results reveal that the majority of the reads mapped to the *E. coli* genome (Tables S4.1-4.4). This is expected as the *E. coli* genome is well annotated (283). However, in all samples the sequencing reads predominantly map to rRNA (95-99%), although many reads still correspond to coding RNA, tRNA, and ncRNA (Figures 4.4, 4.8, 4.12, 4.13). rRNA typically constitutes over 90% of total RNA, which means a ribosomal depletion step is required to ensure other areas of the genome can still be sequenced (284). sRNAs between 20 and 300 nt were enriched for using a size selection step and rRNA was depleted during this step as well (Figure S4.1). However, rRNA derived fragments that migrate in the enriched size range would not have been excluded, resulting in the high abundance of rRNA reads (285). Interestingly, only specific locations of the respective rRNA operons are represented (Figure S4.2). This suggests the possibility that these rRNA derived fragments may be specific products that are produced selectively, as was observed in a sequencing study of *Amblyomma testudinarium* (286). Rajan and colleagues also identified sRNAs located within rRNA operons that bind to the ribosome and mRNA simultaneously (23). Confirmation of the stability of these products and if they have a function will need to be studied in the future.

The distribution of transcripts show that coding RNA (95-97%) is predominantly present (Figures 4.5, 4.9, 4.14, 4.15), which is consistent with previously published results from bacteria (28, 30). mRNA and tRNA fragments are known to appear in sequencing data (14, 18, 20, 28). Initial analysis of the generated sequencing data for tRNA and mRNA fragments has been inconclusive and a more stringent and specific analysis for identifying fragments based on Raad *et al.* is being explored (28).

The two data analyses with DESeq2 (Figures 4.12, 4.14, 4.16, 4.17, Tables S4.5-4.7) and Rockhopper (Figures 4.13, 4.15, 4.18, 4.19, Tables S4.8-4.10) differ in the number of ncRNAs identified. These differences can be attributed to how the data is processed. In the DESeq2 analysis performed by Azenta/Genewiz, any ncRNAs with hit counts less than 10 are excluded. Additionally, DESeq2 analysis relies on the assumption that most genes are not differentially expressed, and data is normalized based on the library size and variation between biological replicates (214, 287). This method is known to have a lower number of significantly enriched or low abundance ncRNAs compared to the control condition (214, 288).

The in-house performed Rockhopper analysis did not have an initial exclusion criterion, allowing all data to be analyzed. Furthermore, Rockhopper normalizes data using upper quartile normalization based on gene length and sequencing depth and quantifies results in Reads per Kilobase of Transcript, per Million mapped reads (RPKM) (216-218). This has been shown to result in a higher number of significantly enriched or low abundance ncRNAs being identified, but also increases the number of false positives (287). Thus, the differences in the number of reported ncRNAs is consistent with the described functions of the respective bioinformatic pipeline. Therefore, the sequencing and bioinformatic analysis pipelines described earlier perform as expected.

The majority of sRNAs identified by both sequence analysis pipelines are annotated in the *E. coli* MG1655 genome (Tables 4.1, 4.2). However, many of the sRNAs are not well characterized or listed as putative, where the reported sRNAs have not been validated as bonafide sRNAs. Looking at overlaps in the identified ncRNAs between each stress condition, both analyses pipelines reveal that the majority of the ncRNAs are common amongst all tested stress conditions (Figures 4.20 A-C, 4.21 A-C). More importantly, both analyses distinguish that ncRNAs found to be significantly enriched or in low abundance when compared to the LB condition are mostly present in only one stress condition (Figures 4.20 D-E, 4.21 D-F). Interestingly, no ncRNAs are enriched or in low abundance in the 70S ribosomes across all conditions using DESeq2, but some ncRNAs are identified using Rockhopper. This is most likely due to the difference in techniques used to normalize the data (214, 216-218). Consequently, it is possible that the Rockhopper analysis is providing an overestimate of ncRNAs abundant or in low abundance for the 70S ribosomes. Experimental validation by performing Northern blot analysis for specific ncRNAs with isolated ribosomal fractions would confirm which ribosome and/or ribosomal subunit a ncRNA is associating with (18).

There are at least 80 reported sRNAs in *E. coli* (134, 191). From the sRNAs reported in EcoCyc, 71 are present in the generated dataset and 15 are absent (C0293, chiZ, ftsO, ipeX, ispZ, malH, nc1, och5, orzP, raiZ, rbsZ, ryeG, small RNA, sraA, and timR) (Tables S4.5-4.10). Many of the sRNAs absent have been shown to be expressed during stationary phase, alternative stress conditions, are recognized as putative, or have low cellular copy numbers, which were not examined here (155, 245, 289, 290).

Additionally, some sRNAs are identified to act as sponges such as ispZ and ftsO (291, 292). They may be absent due to the higher abundance of their targets oxyS and rybB and cpxQ, respectively (291, 292). Most importantly, no sRNAs that perform critical roles are absent in the dataset, suggesting that the obtained ribosomal particles and associated RNA had been isolated from *E. coli* MG1655 cells without any functional defects. Additionally, the Rockhopper analysis identified that several ncRNAs are absent in the minimal medium and hyperosmotic stress conditions (Table S4.8, S4.9). This is most likely due to the cell prioritizing the production of cellular factors that will enable metabolism, which is critical for the cell to survive (293).

Interestingly, functional RNAs (fss – MM30S, T30S1, T50S; ssrS – S50S, T30S1, T50S, and ssrA – T50S) are significantly enriched or depleted under specific stress conditions (Figures 4.6, 4.10, 4.17, 4.7, 4.11, 4.19). While no studies have examined fss (4.5S RNA) specifically under nutrient starvation, it is known that the depletion of the SRP in *E. coli* results in a major global response that impacts the ribosome directly by inactivation (294, 295). Moreover, it is known that the 4.5S RNA can bind to Elongation Factor G and the 30S subunit (221, 296). An increase in 4.5S RNA with the 30S suggests a protective role under stress although further evidence is required to confirm the functional role of the 4.5S RNA in this context. However, the 4.5S RNA is less abundant in the ribosome fraction in the presence of tetracycline. Previously, it was shown that the 4.5S RNA cross-links to ribosomal protein S7, which is in close proximity to the tetracycline binding site 6 on the 30S ribosomal subunit (296, 297). This might suggest that tetracycline binding could prevent 4.5S RNA from accessing this site.

ssrS (6S RNA) is also enriched in the hyperosmotic condition, but depleted in tetracycline stress (Tables 4.1, 4.2). In the α-proteobacteria *Rhodobacter sphaeroides*, 6S RNA was found to play a role in hyperosmotic survival (298). Additionally, although the 6S RNA regulates σ^{70} RNA polymerase, depletion of it causes a decrease in translational machinery including EF-G and ribosomes (299). No previous research has shown that the 6S RNA binds to the ribosome or is involved in antibiotic resistance mechanisms (300). However, depletion of the 6S RNA was shown to impact persistence (301). This could suggest that 6S RNA may be sensitive to tetracycline.

ssrA (tmRNA) is also depleted in the 50S ribosomal subunit under tetracycline stress. Deletion of the ssrA gene has been shown to not have any significant effects on sensitization to tetracycline (302). However, other reports suggest the MIC is lowered (303, 304). As part of its functional mechanism, tmRNA binds to both the 30S and 50S and some of the binding sites overlap with reported tetracycline binding sites (297, 305). Therefore, tetracycline may prevent tmRNA from binding near the A-site; inhibiting stalled ribosomes from being released. Interestingly, the 50S subunit is where tmRNA is depleted. Although most tetracycline binding sites have been identified for the 30S, there is suggestive evidence that tetracycline may also bind to 23S rRNA (275, 297). Taken together for these three functional RNAs (fss, ssrS, and ssrA) in *E. coli*, they are functioning as expected based on published data. This suggests that the stress conditions are indeed working as was noted earlier based on Figure 4.3.

Many sRNAs identified as significantly abundant or in low abundance compared to the LB condition are known to be members of Type I Toxin-Antitoxin (TA) systems (agrB, glmY, istR, rdlA, rdlB, rdlC, rdlD, ryeA, sibA, sibB, sibC, sibD, sokB, sokC, symR) (Tables 4.1, 4.2). TA systems are known to be activated under stress conditions (306). Furthermore, TA systems have been implicated in antibiotic persistence, which results in bacteria surviving exposure to antibiotics (306, 307). Interestingly, all have the same mode of action by inhibiting protein synthesis via base pairing to the mRNA toxin partner (308). However, it has been reported that simple base pairing is not sufficient to prevent translation of the toxin. Hok/Sok and Ldr/Rdl partners were found that within the 5' UTRs of the mRNA toxins were the bindings sites for the sRNA antitoxins (309, 310). Intriguingly, the only identified rancRNA in bacteria is also involved in a TA system (26). SprF1 RNA was found to be bi-functional with its 5'-end impairing global translation and the 3'-end binding to the SprG1 toxin (26). rancRNAs in *T. brucei* and *H. volcanii* were found to interact with the ribosome and mRNA simultaneously (17, 23). This suggests further investigation into how the ribosome may be involved when antitoxins act as antitoxins or if antitoxins have an additional cellular role in addition to their involvement in a toxin-antitoxin pair.

The main reason for including the additional Rockhopper analysis was to identify non-annotated ncRNAs that associated with the ribosome and/or ribosomal subunits. This analysis has been successful

in identifying eight transcripts in non-annotated regions of the *E. coli* MG1655 genome from the hyperosmotic and tetracycline conditions (Table 4.2). PR1S70S, PR1T30S1, PR1T30S2, and PR1T50S are likely to be located in the 5' UTR of cold shock protein CspG (262). This suggests that these predicted RNAs may be the same RNA although PR1S70S and PR1T30S1 are approximately 20 nucleotides shorter than PR1T30S2 and PR1T50S. PR4T30S2 and PR250S are located in the 5' UTR of cold shock protein CspB, but these predicted RNAs also differ in size (269-271). Interestingly, PR2T30S2 and PR3T30S3 are located 185 nt downstream of cold shock protein CspB (269, 270). There has been no evidence to suggest a definitive 3' UTR for CspB. However, it is interesting that all predicted ncRNAs are within or in close proximity to cold shock proteins. Cold shock proteins are known to aid in the translation of mRNAs as chaperones during cold stress (311, 312). Additionally, RbfA, a 30S ribosomal binding factor, has previously been shown to be involved in the *E. coli* cold shock response (313). There has been little information on the role ncRNAs may play during cold shock with respect to translation (314). Further sequence analysis in this area of the genome for transcription termination sites or other RNA motifs can help in discerning if these sets of predicted RNAs for CspG and CspB are located on one RNA transcripts or in a 3' UTR.

Altogether, the reported data shows that the experimental approach for isolating ribosomes from cells under defined stress conditions works and that the two-pronged approach for bioinformatic analysis is able to identify annotated and non-annotated transcripts. Many known *E. coli* sRNAs that associate with the 70S ribosome and 50S and 30S subunits have been identified and are enriched or depleted under minimal medium, hyperosmotic, and tetracycline stress. Additionally, known functional RNAs are specifically impacted during stress conditions as expected. This indicates that they may have roles more directly related to the ribosome. TA systems are also in high abundance as was shown in a previous rancRNA study (26). This suggests there may be a conserved mechanism in bacteria to aid in antibiotic persistence. Lastly, eight non-annotated transcripts with limited experimental data published on the genomic regions these transcripts have been identified in. These non-annotated transcripts and antitoxin RNAs will be further analyzed *in silico* to determine their viability as a rancRNA candidate that will move forward for experimental validation.

4.5 Supplementary Information

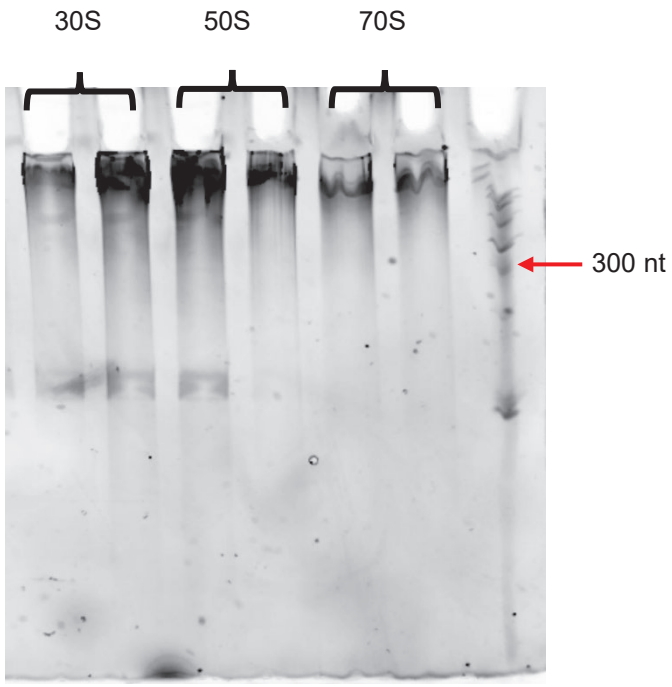


Figure S4.1. RNA fractions extracted from *E. coli* MG1655 cells grown in LB medium. Fractions were analyzed on a 10% Urea PAGE gel and stained with Sybr Gold. Lanes L-R: 30S fractions, 50S fractions, 70S fractions, and RiboRuler Low Range RNA Ladder (Thermo Scientific).

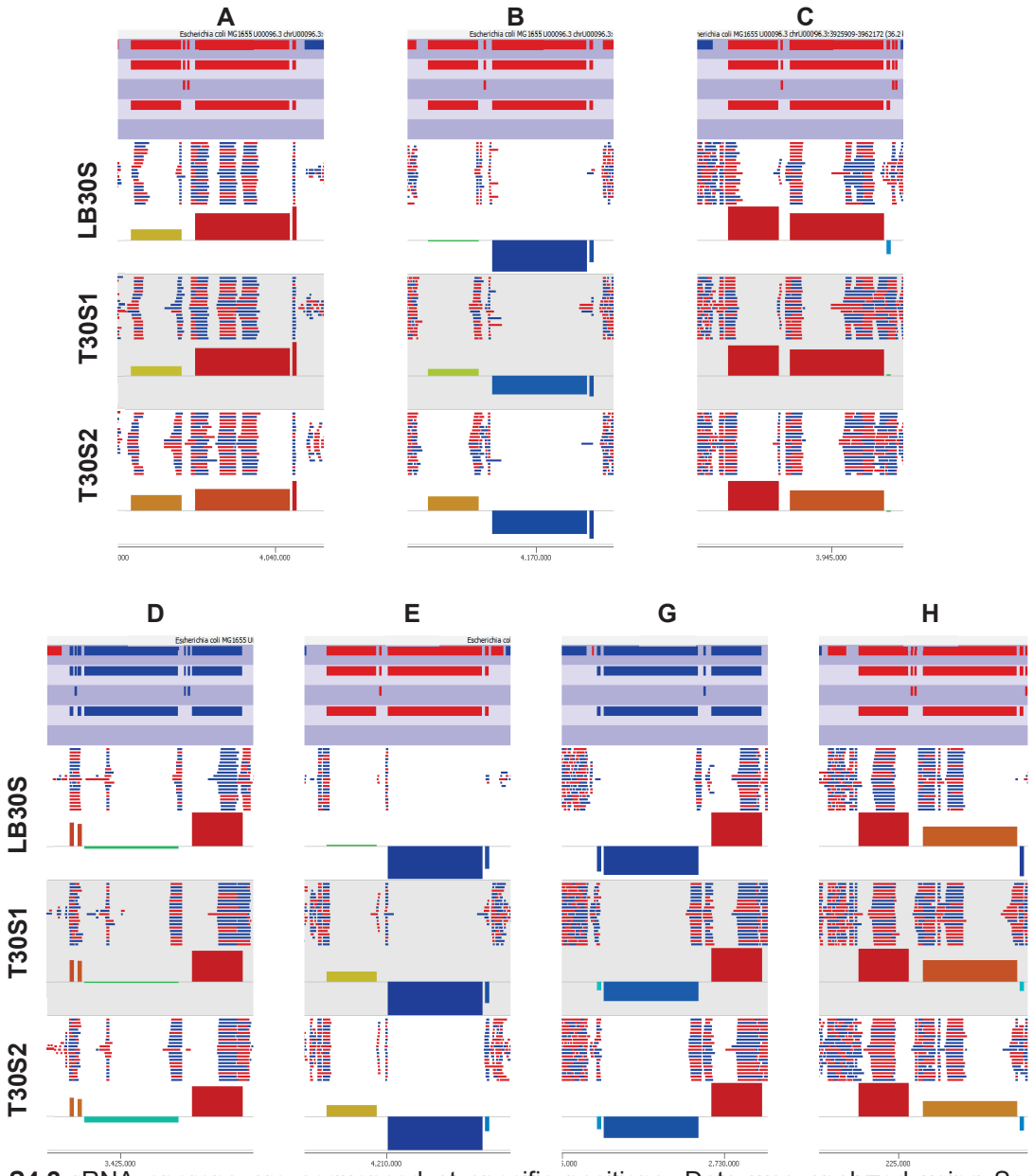


Figure S4.2 rRNA operons are sequenced at specific positions. Data was analyzed using SeqMonk (https://www.bioinformatics.babraham.ac.uk/index.html, version 1.46). Probes were generated for rRNA operons and quantified using the enrichment tool. Shown are the results for all seven rRNA operons (the letter on top of each graph is representative of each operon) for LB and tetracycline 30S region peaks.

Table S4.1. Sequencing statistics for 70S, 50S, and 30S ribosomes and subunits from LB condition.

Reads to -	Reads to +	Read-2	Read-1	mapq >= mapq_cut	mapq < mapq_cut	Unmapped reads	Total records		
781101	780320	780378	781043	1561421	19049370	105467	20716258	1	
5688226	5693420	5610977	5770669	11381646	373451194	1901256	386734096	2	**30S**
624974	624764	624653	625085	1249738	13933422	48498	15231658	3	
617155	616125	616362	616918	1233280	17667643	99171	19000094	1	
11131993	11124959	11056728	11200224	22256952	392930329	2267379	417454660	2	**50S**
331285	331057	331047	331295	662342	14421461	75023	15158826	3	
722065	720671	719541	723195	1442736	18268671	4986429	24697836	1	
6400709	6402584	6326899	6476394	12803293	403811619	2360864	418975776	2	**70S**
421876	421620	421727	421769	843496	17758020	78868	18680384	3	

Percent reads mapped to reference	Reads mapped in proper pairs
99.5	1555016
99.5	9942724
99.7	1245420
99.5	1228338
99.5	20413330
99.5	659418
79.8	1423832
99.4	11112638
99.6	840892

Table S4.2. Sequencing statistics for 70S, 50S, and 30S ribosomes and subunits from hyperosmotic condition.

mapq >= mapq_cut	mapq < mapq_cut	Unmapped reads	Total records		
1561421	19049370	105467	20716258	1	**30S**
11381646	373451194	1901256	386734096	2	
1249738	13933422	48498	15231658	3	
1233280	17667643	99171	19000094	1	**50S**
22256952	392930329	2267379	417454660	2	
662342	14421461	75023	15158826	3	
1442736	18268671	4986429	24697836	1	**70S**
12803293	403811619	2360864	418975776	2	
843496	17758020	78868	18680384	3	

Percent reads mapped to reference	Reads mapped in proper pairs	Reads to -	Reads to +	Read-2	Read-1
99.5	1555016	781101	780320	780378	781043
99.5	9942724	5688226	5693420	5610977	5770669
99.7	1245420	624974	624764	624653	625085
99.5	1228338	617155	616125	616362	616918
99.5	20413330	11131993	11124959	11056728	11200224
99.5	659418	331285	331057	331047	331295
79.8	1423832	722065	720671	719541	723195
99.4	11112638	6400709	6402584	6326899	6476394
99.6	840892	421876	421620	421727	421769

Table S4.3. Sequencing statistics for 70S, 50S, and 30S ribosomes and subunits from minimal medium condition.

Reads to -	Reads to +	Read-2	Read-1	mapq >= mapq_cut	mapq < mapq_cut	Unmapped reads	Total records		
385136	384660	384699	385097	769796	17226635	3499063	21495494	1	**30S**
459058	459000	458964	459094	918058	18474881	68705	19461644	2	
886175	885670	885286	886559	1771845	23448366	138971	25359182	3	
300552	300244	300333	300463	600796	18999010	912816	20512622	1	**50S**
322140	321568	321600	322108	643708	15313764	109472	16066944	2	
361785	361337	361407	361715	723122	14964196	86592	15773910	3	
120192	119468	119585	120075	239660	5660401	9379037	15279098	1	**70S**
306483	306168	306181	306470	612651	12517888	56793	13187332	2	
419469	418902	418909	419462	838371	17552217	115564	18506152	3	

Percent reads mapped to reference	Reads mapped in proper pairs
83.7	767236
99.6	915966
99.5	1753178
95.5	598638
99.3	640248
99.5	718744
38.6	237098
99.6	610258
99.4	832542

Table S4.4. Sequencing statistics for 70S, 50S, and 30S ribosomes and subunits from tetracycline condition.

mapq >= mapq_cut	mapq < mapq_cut	Unmapped reads	Total records		
1955225	17329805	98472	19383502	1	**30S1**
1976124	16222099	59371	18257594	2	
1757388	13785880	81870	15625138	3	
3124752	20151830	94906	23371488	1	**30S2**
2341132	20769160	62556	23172848	2	
2054651	15026104	74045	17154800	3	
1583970	13293214	59854	14937038	1	**50S**
1572001	13604808	116865	15293674	2	
1827798	16216672	93580	18138050	3	
813799	11239298	6165003	18218100	1	**70S**
2581822	17750558	75916	20408288	2	
3138175	23655490	96115	26889780	3	

Percent reads mapped to reference	Reads mapped in proper pairs	Reads to -	Reads to +	Read-2	Read-1
99.5	1940220	977969	977256	977119	978106
99.7	1967910	988113	988011	987936	988188
99.5	1743528	878951	878437	878318	879070
99.6	3110544	1562838	1561914	156205	1562699
99.7	2330994	1170756	1170376	117041	1170718
99.6	2037458	1027470	1027181	102695	1027696
99.6	1575138	792387	791583	791727	792243
99.2	1559362	786044	785957	785827	786174
99.5	1813684	914066	913732	913572	914226
66.2	809846	407268	406531	406706	407093
99.6	2563900	1291130	1290692	129052	1291296
99.6	3116802	1569414	1568761	156849	1569676

Table S4.5. List of ncRNAs identified in the minimal medium condition through DESeq2 analysis.

ncRNA	Norm. Mean Hit Counts (LB70S)	Norm. Mean Hit Counts (M70S)	Log2FoldChange	Pvalue
csrB	8.2	45	2.456224389	0.263038
ffs	19	115	2.597560489	0.195045
fnrS	151.6	110	-0.462753612	0.794944
gcvB	412	125	-1.720696447	0.34523
glmY	28.4	65	1.194553068	0.539904
glmZ	19.6	20	0.029155732	0.989533
rnpB	81.2	410	2.336075141	0.212044
rybA	11	135	3.617375544	0.085721

ryfD	82	245	1.57909111	0.385414
spf	4.2	10	1.251535921	0.666073
sroH	5.8	5	-0.214111483	0.947806
ssrA	22095.4	81050	1.875083109	0.292887
ssrS	101.8	380	1.900266574	0.298625
tff	25.8	30	0.217599811	0.916447
	Norm. Mean Hit Counts (LB50S)	**Norm. Mean Hit Counts (M50S)**		
csrB	8.007572174	149.3586288	4.221247514	0.15791
ffs	13.0558242	361.9074467	4.792828472	0.112908
fnrS	244.5791066	109.1466903	-1.163994328	0.672413
gcvB	348.5034672	149.3586288	-1.222351637	0.656457
glmY	11.8372806	97.65756499	3.04438336	0.293434
glmZ	19.84485278	5.744562647	-1.78845981	0.576842
oxyS	3.829708431	0	-1.857691033	0.64343
rnpB	49.96028726	281.4835697	2.49419955	0.3774
rybA	5.918640302	17.23368794	1.541890686	0.602179
ryfD	82.68688658	183.8260047	1.15262553	0.6745
spf	9.748348734	5.744562647	-0.762939327	0.812248
sraG	4.177863743	0	-1.983219912	0.618309
ssrA	3548.398939	24092.69574	2.763380862	0.329498
ssrS	52.39737444	1418.906974	4.759124423	0.113865
tff	37.42669603	57.44562647	0.61814412	0.82323
	Norm. Mean Hit Counts (LB30S)	**Norm. Mean Hit Counts (M30S)**		
arcZ	5.126524164	16.97056275	1.726976953	0.511671
csrB	24.74873734	141.4213562	2.514571979	0.269097
csrC	7.424621202	11.3137085	0.607687283	0.822364
ffs	28.10749455	967.3220767	5.104952951	0.03963
fnrS	524.6732316	231.9310242	-1.177697951	0.578915
gcvB	345.775216	135.764502	-1.348702828	0.530145
glmY	10.07627163	84.85281374	3.073992454	0.196012
glmZ	19.97576657	22.627417	0.179831163	0.93975
mgrR	3.712310601	16.97056275	2.192635429	0.41444
oxyS	3.535533906	11.3137085	1.678064751	0.554856
rnpB	205.0609665	571.3422792	1.478313075	0.48846
rttR	4.772970773	16.97056275	1.83006936	0.488776
rybA	3.535533906	16.97056275	2.263023939	0.401108
ryeA	11.3137085	5.656854249	-0.999979635	0.734634
ryfD	218.8495488	367.6955262	0.748585914	0.72128
sibA	7.954951288	0	-2.934489286	0.443638
spf	36.76955262	243.2447327	2.72582429	0.22954
sraB	4.949747468	11.3137085	1.192643635	0.667818
sraG	5.303300859	5.656854249	0.093117815	0.97622
sroH	8.308504679	11.3137085	0.44541744	0.86838
ssrA	2181.247643	28295.58496	3.697361502	0.112286
ssrS	614.4757929	14125.16506	4.522768371	0.06082
tff	42.07285348	45.254834	0.105194369	0.962258

Table S4.6. List of ncRNAs identified in the hyperosmotic condition through DESeq2 analysis.

ncRNA	Norm. Mean Hit Counts (LB70S)	Norm. Mean Hit Counts (S70S)	Log2FoldChange	Pvalue
csrB	13.11372838	12.50597811	-0.068449396	0.977654
ffs	30.38546819	46.89741791	0.626133787	0.750206
fnrS	242.4440514	171.957199	-0.495588975	0.775884
gcvB	658.8848891	112.553803	-2.54938605	0.177554
glmY	45.41827876	143.8187483	1.662910557	0.374481
glmZ	31.34500929	100.0478249	1.674385022	0.383317
rnpB	129.8578956	150.0717373	0.208728637	0.905487
rybA	17.59158684	300.1434746	4.092685945	0.050424
ryfD	131.1372838	1219.332866	3.216943333	0.0979
spf	6.716787704	21.88546169	1.704127289	0.47897
sraG	5.117552537	37.51793433	2.874045236	0.216477
sroH	9.275563973	0	-4.011497722	0.323888
ssrA	35335.74072	116649.5108	1.722998394	0.329336
ssrS	162.8021401	337.661409	1.052465659	0.548757
tff	41.26026733	50.02391244	0.277873295	0.88496
	Norm. Mean Hit Counts (LB50S)	**Norm. Mean Hit Counts (S50S)**		
csrB	12.44304096	48.05899153	1.949466468	0.358117
ffs	20.28756678	188.5391206	3.21619203	0.10434
fnrS	380.053751	85.02744655	-2.160182141	0.233577
gcvB	541.5427826	66.54321904	-3.024684894	0.111853
glmY	18.39406055	1182.990561	6.007036136	0.006003
glmZ	30.83710151	88.72429205	1.52466523	0.421193
oxyS	5.951019589	3.696845502	-0.686818537	0.850074
rnpB	77.63375554	410.3498507	2.402100848	0.18677
rybA	9.197030274	77.63375554	3.077437369	0.147074
ryfD	128.4879229	262.4760307	1.030560033	0.545929
spf	15.14804986	29.57476402	0.965238914	0.659267
sraG	6.49202137	33.27160952	2.357546288	0.313297
ssrA	5513.89015	26824.31096	2.282410763	0.192901
ssrS	81.42076801	1800.36376	4.466742545	0.026668
tff	58.15769144	29.57476402	-0.975588903	0.620624
	Norm. Mean Hit Counts (LB30S)	**Norm. Mean Hit Counts (S30S)**		
arcZ	9.170605214	6.32455532	-0.536035706	0.856414
csrB	44.27188724	98.03060747	1.146845755	0.49626
csrC	13.28156617	25.29822128	0.929614658	0.667335
ffs	50.2802148	94.86832981	0.915940884	0.582705
fnrS	938.5640095	369.9864862	-1.342968353	0.373373
gcvB	618.5415103	202.3857703	-1.611747641	0.299273
glmY	18.02498266	736.8106948	5.353211418	0.006979
glmZ	35.73373756	281.4427118	2.977481316	0.090938
mgrR	6.640783086	15.8113883	1.251539073	0.624885
oxyS	6.32455532	6.32455532	1.20E-05	0.999997
rnpB	366.8242086	2390.681911	2.704265633	0.100157
rttR	8.538149682	60.08327554	2.814962743	0.168938
rybA	6.32455532	98.03060747	3.95418526	0.054423

ryeA	20.23857703	303.5786554	3.906884535	0.036818
ryfA	4.427188724	25.29822128	2.514564449	0.302145
ryfD	391.4899743	566.0477012	0.531956259	0.715316
sibA	14.23024947	9.486832981	-0.5849475	0.818456
sibC	5.375872022	94.86832981	4.141343109	0.047472
spf	65.77537533	173.9252713	1.402852246	0.386321
sraB	8.854377448	25.29822128	1.514573179	0.503789
sraG	9.486832981	31.6227766	1.736964829	0.423771
sroH	14.862705	28.46049894	0.937268321	0.654944
ssrA	3901.934405	38402.69991	3.298954081	0.053274
ssrS	1099.207715	8493.877795	2.949965154	0.076742
tff	75.26220831	151.7893277	1.012078088	0.526092

Table S4.7. List of ncRNAs identified in the tetracycline condition through DESeq2 analysis.

ncRNA	**Norm. Mean Hit Counts (LB70S)**	**Norm. Mean Hit Counts (T70S)**	**Log2FoldChange**	**Pvalue**
arcZ	13.67198146	5.266244709	-1.376366795	0.427678
csrB	31.14173554	28.9643459	-0.10456782	0.928925
csrC	12.15287241	13.16561177	0.115480637	0.936658
ffs	72.1576799	26.33122354	-1.454371399	0.249569
fnrS	575.7423302	258.0459907	-1.157781021	0.459776
gcvB	1564.682322	653.0143439	-1.260666119	0.434497
glmY	107.8567426	17.1152953	-2.655746334	0.064368
glmZ	74.43634348	42.12995767	-0.821155144	0.495864
oxyS	9.114654304	26.33122354	1.530514948	0.270499
rnpB	308.3791373	293.5931425	-0.070878284	0.960535
rttR	6.835990728	40.81339649	2.577818155	0.074962
rybA	41.77549889	21.06497884	-0.987805373	0.41763
ryfD	311.4173554	65.82805886	-2.242060794	0.132015
spf	15.95064503	47.39620238	1.571157467	0.214659
sraB	6.076436203	27.64778472	2.185864775	0.142772
sraG	12.15287241	46.0796412	1.922832055	0.145021
sroH	22.02708123	6.582805886	-1.742495125	0.257575
ssrA	83913.30529	11244.74901	-2.899597073	0.176786
ssrS	386.6132534	143.5051683	-1.429775964	0.333126
tff	97.98253377	146.1382907	0.576741944	0.64269
	Norm. Mean Hit Counts (LB50S)	**Norm. Mean Hit Counts (T50S)**		
arcZ	18.6185724	51.66883795	1.472554162	0.394298
csrB	65.88110233	9.775185558	-2.752649315	0.136133
csrC	25.77956178	30.72201175	0.253051071	0.88086
cyaR	0	32.11846683	6.929429663	0.306032
ffs	107.4148408	34.21314945	-1.650556977	0.315516
fnrS	2012.238017	4532.893189	1.171641093	0.438802
gcvB	2867.260149	2875.301009	0.00405243	0.997786
glmY	97.38945561	50.27238287	-0.953988859	0.546271
glmZ	163.2705579	153.6100588	-0.087983454	0.953261
mgrR	24.3473639	85.8819874	1.81859164	0.284721
oxyS	31.50835329	210.1664895	2.737725852	0.116561
rdlC	1.432197877	32.11846683	4.487073554	0.077219

rnpB	411.0407906	761.0680184	0.888751634	0.553922
rttR	12.88978089	497.1380084	5.26933294	0.009765
rybA	48.69472781	52.36706549	0.104901093	0.947381
ryeA	22.91516603	21.64505374	-0.082258132	0.962276
ryfA	17.18637452	30.72201175	0.838011446	0.632908
ryfD	680.2939914	965.6486876	0.505349245	0.732222
ryjA	0	19.55037112	6.21351308	0.048672
sdsN	5.728791507	15.36100588	1.422969748	0.49847
sibA	14.32197877	14.66277834	0.033934489	0.985546
sibB	7.160989383	11.86986818	0.729076501	0.726823
sibD	17.18637452	67.02984382	1.963538915	0.262987
sibE	10.02538514	13.26632326	0.40411613	0.837615
spf	80.20308109	882.5596104	3.459963323	0.052776
sraB	8.59318726	115.2075441	3.744890197	0.055166
sraG	34.37274904	129.8703224	1.917736606	0.251085
sroC	17.18637452	23.04150881	0.422975326	0.811949
sroE	8.59318726	10.4734131	0.28547213	0.889862
sroH	22.91516603	31.42023929	0.455396856	0.78926
ssrA	29193.92152	2495.465227	-3.548252614	0.046087
ssrS	431.0915609	577.4341754	0.421670845	0.775571
tff	307.9225435	456.6408111	0.568500245	0.702703
	Norm. Mean Hit Counts (LB30S)	**Norm. Mean Hit Counts (T30S1)**		
arcZ	33.17686642	53.32028581	0.684509079	0.521512
csrB	160.1641827	36.71232794	-2.125206664	0.079771
csrC	48.04925482	109.2628808	1.185219427	0.279532
cyaR	2.288059753	62.06131627	4.761484722	0.008282
ffs	181.9007504	28.84540052	-2.656727812	0.039465
fnrS	3395.480674	2663.391981	-0.350339214	0.799592
gcvB	2237.722439	6976.216411	1.640418468	0.136732
glmY	65.20970297	118.8780143	0.866325855	0.414644
glmZ	129.2753761	412.5766378	1.674217444	0.172936
mgrR	24.02462741	171.324197	2.834142224	0.031779
oxyS	22.88059753	378.486619	4.048041749	0.007749
rnpB	1327.074657	2776.151274	1.064846415	0.488927
rttR	30.88880667	202.7919067	2.714843794	0.037602
rybA	22.88059753	69.05414064	1.593604262	0.168771
ryeA	73.21791211	53.32028581	-0.457508317	0.658479
ryfA	16.01641827	67.30593455	2.071182296	0.093509
ryfD	1416.308987	463.2746144	-1.612182549	0.259082
ryjA	1.144029877	20.97847311	4.196694985	0.046698
sdsN	13.72835852	27.09719443	0.980985986	0.434833
sibA	51.48134445	13.11154569	-1.973204611	0.116467
sibB	14.8723884	13.11154569	-0.181794939	0.888802
sibC	19.4485079	7.866927415	-1.305781012	0.342225
sibD	13.72835852	44.57925535	1.699213932	0.174819
sibE	12.58432864	14.85975178	0.239784606	0.854658
spf	237.9582143	584.7749378	1.2971774	0.299825
sraB	32.03283655	210.6588341	2.71728475	0.037657
sraG	34.3208963	146.8493117	2.097178181	0.08368
sroC	22.88059753	12.23744265	-0.902817583	0.479287

sroE	4.576119507	15.73385483	1.781674322	0.247802
sroH	53.7694042	46.32746144	-0.214914555	0.834929
ssrA	14116.18465	4472.785287	-1.658085643	0.26795
ssrS	3976.647851	824.2791725	-2.270325697	0.16616
tff	272.2791106	1373.215885	2.334404033	0.11538
	Norm. Mean Hit Counts (LB30S)	**Norm. Mean Hit Counts (T30S2)**		
agrA	1.508738872	22.53537747	3.900761996	0.057372
arcZ	43.7534273	133.8866544	1.613545757	0.144912
csrB	211.2234421	57.00124891	-1.889697101	0.097048
csrC	63.36703264	190.8879033	1.590923156	0.149937
cyaR	3.017477745	41.75672885	3.790585412	0.029265
ffs	239.8894807	72.24576896	-1.731377744	0.123204
fnrS	4477.936973	4984.295253	0.154565971	0.908374
gcvB	2951.093234	5249.417341	0.830917405	0.551473
glmY	85.99811573	432.8118086	2.331364406	0.063556
glmZ	170.4874926	546.8143064	1.681388138	0.16667
mgrR	31.68351632	192.2135138	2.600905467	0.037248
oxyS	30.17477745	255.8428149	3.083841829	0.01969
rdlA	1.508738872	20.54696181	3.767495991	0.066733
rdlB	1.508738872	15.24452006	3.336863241	0.107196
rdlC	6.03495549	38.44270275	2.6712923	0.072886
rdlD	1.508738872	22.53537747	3.900761996	0.057372
rirA	6.03495549	11.26768874	0.900777989	0.551158
rnpB	1750.137092	9646.46717	2.462539437	0.10335
rttR	40.73594955	1584.104475	5.281212748	0.003631
rybA	30.17477745	73.5713794	1.285803161	0.246285
ryeA	96.55928783	133.2238492	0.46436827	0.638496
ryfA	21.12234421	121.9561604	2.529520561	0.042865
ryfD	1867.818724	959.7419583	-0.960623629	0.489297
ryjA	1.508738872	43.08233929	4.835662676	0.019595
sdsN	18.10486647	39.10550797	1.110995993	0.356015
sibA	67.89324926	57.00124891	-0.252271282	0.7996
sibB	19.61360534	25.84940357	0.398278931	0.737127
sibC	25.64856083	37.77989753	0.558743422	0.623529
sibD	18.10486647	188.2366824	3.378096292	0.013553
sibE	16.5961276	32.47745577	0.968593825	0.425597
sokB	1.508738872	13.2561044	3.135230106	0.132253
spf	313.8176855	1898.27415	2.596693467	0.089722
sraB	42.24468843	351.2867665	3.055806924	0.022001
sraG	45.26216617	228.6678008	2.336876475	0.053661
sroA	6.03495549	15.24452006	1.336876151	0.367041
sroC	30.17477745	105.38603	1.804269679	0.114902
sroE	6.03495549	13.2561044	1.135242722	0.447289
sroH	70.910727	224.6909695	1.66386823	0.136674
ssrA	18616.32895	8076.944409	-1.204675467	0.308831
ssrS	5244.37632	21305.21099	2.022370069	0.115069
tff	359.0798516	1165.211576	1.698220052	0.210839

Table S4.8. List of ncRNAs identified in the minimal medium condition through Rockhopper analysis.

ncRNA	Transcript Abundance (LB70S)	Transcript Abundance (M70S)	Log2FoldChange	q value
3'ETS-leuZ	7	0		1
arcZ	100	28	-1.84	0.493873612
arrS	10	0		1
chiX	16	110	2.78	0.000146391
cpxQ	15	102	2.77	2.50E-05
csrB	631	518	-0.28	1
csrC	95	116	0.29	1
cyaR	16	0		1
dsrA	1	0		1
esrE	129	134	0.05	1
eyeA	6	0		1
ffs	518	2088	2.01	0.036319273
fnrS	7657	5645	-0.44	1
gadY	3	40	3.74	7.55E-33
gcvB	10582	820	-3.69	7.90E-16
glmY	447	509	0.19	1
glmZ	470	678	0.53	1
mcaS	10	24	1.26	1
mgrR	203	139	-0.55	1
micA	17	0		1
micC	3	0		1
micL	46	0		0
omrA	42	0		0
omrB	15	0		1
oxyS	111	0		0
pspH	1	0		1
ralA	0	13		1
rdlB	0	15		1
rdlC	8	29	1.86	0.128640188
rdlD	0	9		1
rirA	22	0		0
rnpB	2730	2775	0.02	1
rttR	68	116	0.77	1
rybA	362	1710	2.24	0.018460524
rybB	2	0		1
rydB	24	139	2.53	0.003915612
rydC	24	0		0
ryeA	110	15	-2.87	0.002276971
ryfA	42	121	1.53	0.394902437
ryfD	4753	13403	1.50	1
ryhB	15	999	6.06	0
ryjA	0	28		0
ryjB	2	26	3.70	6.58E-20
sdhX	3	43	3.84	9.00E-27
sdsN	4	6	0.58	1
sdsR	193	0		1

sgrS	20	45	1.17	1
sibA	89	0		0
sibB	17	0		1
sibD	46	165	1.84	0.186834716
sibE	46	210	2.19	0.031582662
sokB	34	25	-0.44	1
sokC	54	160	1.57	0.53722664
spf	162	375	1.21	0.963334284
sraB	94	76	-0.31	1
sraG	403	78	-2.37	0.093931613
sroA	10	5	-1.00	1
sroC	41	168	2.03	0.082127129
sroE	46	2	-4.52	4.63E-77
sroH	49	80	0.71	1
ssrA	409112	466134	0.19	1
ssrS	4199	9285	1.14	1
tff	1911	430	-2.15	0.133725532
	Transcript Abundance (LB50S)	**Transcript Abundance (M50S)**		
3'ETS-leuZ	28	34	0.28	1
agrA	48	13	-1.88	0.242091991
agrB	3	0		1
arcZ	235	74	-1.67	1
arrS	0	17		0
chiX	9	0		1
cpxQ	46	21	-1.13	1
csrB	1352	1528	0.18	1
csrC	345	124	-1.48	1
cyaR	84	0		0
dsrA	1	0		1
esrE	141	67	-1.07	1
eyeA	6	0		1
ffs	1302	4538	1.80	0.080236723
fnrS	23285	7082	-1.72	1
gadF	10	0		1
gadY	41	141	1.78	0.31107105
gcvB	35105	1305	-4.75	1.66E-47
glmY	790	1743	1.14	0.658431525
glmZ	1554	539	-1.53	1
isrC	18	0		1
istR	4	0		1
istR	4	0		1
mcaS	62	0		0
mgrR	260	211	-0.30	1
micA	69	59	-0.23	1
micF	10	0		1
micL	67	0		0
ohsC	2	0		1
omrA	42	0		0
omrB	53	0		0

oxyS	168	135	-0.32	1
pspH	3	10	1.74	1
rdlA	34	0		0
rdlB	4	0		1
rdlC	5	0		1
rdlD	7	0		1
rirA	9	0		1
rnpB	7291	3374	-1.11	1
rttR	52	135	1.38	0.544594589
rybA	368	591	0.68	1
rybB	16	45	1.49	1
rydB	6	141	4.55	1.75E-110
rydC	6	136	4.50	6.64E-99
ryeA	217	39	-2.48	0.082447967
ryfA	47	97	1.05	1
ryfD	4556	6546	0.52	1
ryhB	32	979	4.94	9.22E-136
ryjA	6	86	3.84	3.16E-19
ryjB	54	11	-2.30	0.0295834
sdhX	47	0		0
sdsN	40	15	-1.42	1
sdsR	462	93	-2.31	0.178742306
sgrS	105	0		0
sibA	202	147	-0.46	1
sibB	106	0		0
sibC	51	0		0
sibD	187	110	-0.77	1
sibE	127	317	1.32	0.479152459
sokB	96	0		0
sokC	40	113	1.50	1
sokX	7	0		1
spf	715	479	-0.58	1
sraB	195	69	-1.50	1
sraG	239	124	-0.95	1
sroA	17	19	0.16	1
sroC	33	84	1.35	1
sroE	125	113	-0.15	1
sroH	140	98	-0.51	1
ssrA	138117	232637	0.75	0.382203869
ssrS	25538	32522	0.35	1
symR	3	0		1
tff	2101	1053	-1.00	1
	Transcript Abundance (LB30S)	**Transcript Abundance (M30S)**		
3'ETS-leuZ	75	25	-1.58	1
agrA	29	2	-3.86	3.31E-21
agrB	8	2	-2.00	1
arcZ	454	205	-1.15	0.875682113
arrS	13	0		1
chiX	70	0		0

cpxQ	33	17	-0.96	1
csrB	6705	2188	-1.62	1
csrC	836	111	-2.91	0.004632215
cyaR	25	8	-1.64	0.37872031
dicF	6	0		1
dsrA	3	0		1
esrE	327	101	-1.69	0.926520054
eyeA	23	0		0
ffs	961	8901	3.21	2.62E-29
fnrS	45413	13853	-1.71	0.735647307
gadF	7	12	0.78	1
gadY	42	39	-0.11	0.944177154
gcvB	27296	2513	-3.44	0.007388945
glmY	1814	2183	0.27	1
glmZ	1528	383	-2.00	0.614509708
isrC	1	11	3.46	1
mcaS	39	67	0.78	1
mgrR	1078	173	-2.64	0.033606858
micA	76	45	-0.76	0.878997068
micF	3	0		1
micL	66	4	-4.04	6.66E-41
omrA	92	0		0
omrB	191	0		1
oxyS	296	214	-0.47	1
pspH	16	0		1
rdlA	61	6	-3.35	2.76E-12
rdlB	9	6	-0.58	1
rdlC	10	7	-0.51	1
rdlD	49	5	-3.29	2.53E-11
rirA	54	0		0
rnpB	29119	7099	-2.04	1
rseX	21	0		1
rttR	124	252	1.02	0.369481831
rybA	319	352	0.14	1
rybB	32	49	0.61	1
rydB	6	129	4.43	2.05E-71
rydC	12	19	0.66	1
ryeA	3284	253	-3.70	6.30E-09
ryfA	250	25	-3.32	1.77E-05
ryfD	8359	11767	0.49	1
ryhB	26	1780	6.10	0
ryjA	21	4	-2.39	1
ryjB	39	49	0.33	1
sdhX	20	41	1.04	1
sdsN	180	120	-0.58	1
sdsR	6022	728	-3.05	0.003696796
sgrS	126	40	-1.66	0.875682113
sibA	862	25	-5.11	8.05E-87
sibB	352	166	-1.08	0.875682113
sibC	398	142	-1.49	1
sibD	203	316	0.64	0.615888541

sibE	123	98	-0.33	0.999762311
sokB	354	0		1
sokC	142	15	-3.24	1.37E-08
sokE	5	0		1
sokX	12	0		1
spf	1745	3316	0.93	0.461362297
sraB	500	195	-1.36	1
sraG	304	121	-1.33	0.875682113
sroA	49	0		0
sroC	127	5	-4.67	4.96E-91
sroE	72	171	1.25	0.500765356
sroH	457	362	-0.34	1
ssrA	81412	156300	0.94	1
ssrS	130260	358175	1.46	1
symR	33	3	-3.46	2.16E-12
tff	3971	1922	-1.05	0.875682113

Table S4.9. List of ncRNAs identified in the hyperosmotic condition through Rockhopper analysis.

ncRNA	Transcript Abundance (LB70S)	Transcript Abundance (S70S)	Log2FoldChange	q value
PR1S70S	9498	309584	5.03	9.14E-110
PR2S70S	121516	329265	1.44	1
3'ETS-leuZ	7	158	4.50	1.26E-98
arcZ	100	79	-0.34	1
arrS	10	0		1
chiX	16	0		1
cpxQ	15	113	2.91	1.20E-06
csrB	631	402	-0.65	1
csrC	95	156	0.72	1
cyaR	16	62	1.95	0.141447812
dsrA	1	0		1
esrE	129	176	0.45	1
eyeA	6	0		1
ffs	518	644	0.31	1
fnrS	7657	3035	-1.34	1
gadY	3	0		1
gcvB	10582	1709	-2.63	0.001196766
glmY	447	3254	2.86	3.44E-07
glmZ	470	1462	1.64	0.249373798
isrC	3	0		1
istR	9	35	1.96	0.174252815
istR	4	40	3.32	3.17E-11
mcaS	10	0		1
mgrR	203	360	0.83	1
micA	17	77	2.18	0.067221134
micC	3	6	1.00	1
micL	46	0		0

omrA	42	22	-0.93	1
omrB	15	0		1
oxyS	111	124	0.16	1
pspH	1	0		1
rdlC	8	0		1
rdlD	0	131		1
rirA	22	0		0
rnpB	2730	2014	-0.44	1
rprA	1	0		1
rttR	68	362	2.41	0.001897204
rybA	362	1013	1.48	0.464370208
rybB	2	0		1
rydB	24	0		0
rydC	24	0		0
ryeA	110	43	-1.36	1
ryfA	42	33	-0.35	1
ryfD	4753	33013	2.80	0.00163752
ryhB	15	0		1
ryjB	2	0		1
sdhX	3	27	3.17	3.65E-09
sdsN	4	0		1
sdsR	193	108	-0.84	1
sgrS	20	16	-0.32	1
sibA	89	61	-0.54	1
sibB	17	0		1
sibC	0	6		1
sibD	46	176	1.94	0.129065682
sibE	46	0		0
sokB	34	0		0
sokC	54	0		0
spf	162	267	0.72	1
sraB	94	77	-0.29	1
sraG	403	354	-0.19	1
sroA	10	0		1
sroC	41	44	0.10	1
sroE	46	17	-1.44	1
sroH	49	20	-1.29	1
ssrA	409112	605252	0.57	1
ssrS	4199	3958	-0.09	1
tff	1911	1733	-0.14	1
	Transcript Abundance (LB50S)	**Transcript Abundance (S50S)**		
3'ETS-leuZ	28	60	1.10	1
agrA	48	0		0
agrB	3	0		1
arcZ	235	15	-3.97	2.70E-25
chiX	9	0		1
cpxQ	46	304	2.72	0.00053886
csrB	1352	2082	0.62	1
csrC	345	132	-1.39	1

cyaR	84	0		0
dsrA	1	4	2.00	1
esrE	141	133	-0.08	1
eyeA	6	0		1
ffs	1302	1348	0.05	1
fnrS	23285	5109	-2.19	0.075429434
gadF	10	0		1
gadY	41	0		0
gcvB	35105	3269	-3.42	3.49E-12
glmY	790	9706	3.62	8.74E-30
glmZ	1554	2924	0.91	1
isrC	18	0		1
istR	4	0		1
istR	4	0		1
mcaS	62	0		0
mgrR	260	79	-1.72	1
micA	69	4	-4.11	4.81E-36
micF	10	0		1
micL	67	0		0
omrA	42	61	0.54	1
omrB	53	53	0.00	1
oxyS	168	12	-3.81	1.56E-19
pspH	3	17	2.50	0.000237216
rdlA	34	0		0
rdlB	4	46	3.52	6.22E-16
rdlC	5	88	4.14	1.24E-41
rdlD	7	16	1.19	1
rirA	9	30	1.74	0.427331071
rnpB	7291	7008	-0.06	1
rttR	52	656	3.66	4.66E-25
rybA	368	322	-0.19	1
rybB	16	25	0.64	1
rydB	6	16	1.42	0.428269095
rydC	6	0		1
ryeA	217	229	0.08	1
ryfA	47	110	1.23	0.615562682
ryfD	4556	3916	-0.22	1
ryhB	32	52	0.70	1
ryjA	6	0		1
ryjB	54	8	-2.75	0.000228838
sdhX	47	52	0.15	1
sdsN	40	61	0.61	1
sdsR	462	229	-1.01	1
sgrS	105	93	-0.18	1
sibA	202	17	-3.57	1.41E-11
sibB	106	16	-2.73	0.005181689
sibC	51	31	-0.72	1
sibD	187	89	-1.07	1
sibE	127	0		0
sokB	96	47	-1.03	1
sokC	40	42	0.07	1

sokX	7	0		1
spf	715	446	-0.68	1
sraB	195	211	0.11	1
sraG	239	289	0.27	1
sroA	17	197	3.53	1.17E-11
sroC	33	55	0.74	1
sroE	125	50	-1.32	1
sroH	140	129	-0.12	1
ssrA	138117	161497	0.23	1
ssrS	25538	17487	-0.55	1
symR	3	0		1
tff	2101	1438	-0.55	1
	Transcript Abundance (LB30S)	**Transcript Abundance (S30S)**		
3'ETS-leuZ	75	159	1.08	0.825194599
agrA	29	28	-0.05	1
agrB	8	25	1.64	0.324467258
arcZ	454	91	-2.32	0.074396687
arrS	13	40	1.62	0.343621555
chiX	70	149	1.09	0.659099125
cpxQ	33	230	2.80	1.28E-06
csrB	6705	2936	-1.19	1
csrC	836	194	-2.11	0.155259333
cyaR	25	6	-2.06	0.075929421
dicF	6	14	1.22	1
dsrA	3	0		1
esrE	327	121	-1.43	0.623662182
eyeA	23	14	-0.72	1
ffs	961	1023	0.09	1
fnrS	45413	3819	-3.57	4.63E-07
gadF	7	0		1
gadY	42	6	-2.81	3.26E-06
gcvB	27296	2865	-3.25	0.00041589
glmY	1814	6462	1.83	0.0373216
glmZ	1528	3579	1.23	0.350447753
isrC	1	0		1
mcaS	39	23	-0.76	1
mgrR	1078	134	-3.01	8.88E-05
micA	76	36	-1.08	1
micC	0	11		1
micF	3	22	2.87	3.14E-09
micL	66	31	-1.09	1
ohsC	1	12	3.58	1
omrA	92	85	-0.11	1
omrB	191	87	-1.13	1
oxyS	296	60	-2.30	0.092589326
pspH	16	4	-2.00	1
rdlA	61	33	-0.89	1
rdlB	9	17	0.92	1
rdlC	10	96	3.26	7.20E-11

rdlD	49	45	-0.12	1
rirA	54	71	0.39	1
rnpB	29119	24653	-0.24	1
rprA	61	5	-3.61	8.99E-17
rrrD	10	0		1
rrrQ	13	8	-0.70	1
rseX	21	0		1
rttR	124	758	2.61	5.03E-10
rybA	319	282	-0.18	1
rybB	32	12	-1.42	0.573770009
rydB	6	147	4.61	9.17E-96
rydC	12	0		1
ryeA	3284	2155	-0.61	0.966307947
ryfA	250	107	-1.22	0.811104241
ryfD	8359	4869	-0.78	1
ryhB	26	67	1.37	0.440939854
ryjA	21	11	-0.93	1
ryjB	39	66	0.76	1
sdhX	20	11	-0.86	1
sdsN	180	113	-0.67	1
sdsR	6022	4434	-0.44	0.993208256
sgrS	126	99	-0.35	1
sibA	862	21	-5.36	4.08E-167
sibB	352	19	-4.21	2.96E-28
sibC	398	909	1.19	0.148980565
sibD	203	39	-2.38	0.082092772
sibE	123	37	-1.73	0.546607106
sokB	354	103	-1.78	0.470066758
sokC	142	80	-0.83	0.927095625
sokE	5	0		1
sokX	12	0		1
spf	1745	1037	-0.75	0.892556129
sraB	500	145	-1.79	0.371786575
sraG	304	409	0.43	0.732214974
sroA	49	13	-1.91	0.197831225
sroC	127	26	-2.29	0.172790755
sroE	72	83	0.21	1
sroH	457	206	-1.15	1
ssrA	81412	181250	1.15	1
ssrS	130260	74769	-0.80	1
symR	33	0		1
tff	3971	2761	-0.52	0.965166135

Table S4.10. List of ncRNAs identified in the tetracycline condition through Rockhopper analysis.

ncRNA	Transcript Abundance (LB70S)	Transcript Abundance (T70S)	Log2FoldChange	q value
3'ETS-leuZ	7	154	4.46	9.57E-92
agrA	0	36		1
agrB	0	4		1
arcZ	100	200	1.00	0.607567456
arrS	10	0		1
chiX	16	24	0.58	1
cpxQ	15	0		1
csrB	631	227	-1.47	0.005415804
csrC	95	206	1.12	0.766537973
cyaR	16	130	3.02	2.62E-07
dsrA	1	0		1
esrE	129	240	0.90	1
eyeA	6	31	2.37	0.002020499
ffs	518	238	-1.12	0.11181392
fnrS	7657	6788	-0.17	1
gadY	3	8	1.42	1
gcvB	10582	13399	0.34	0.831281729
glmY	447	349	-0.36	0.434923764
glmZ	470	808	0.78	1
isrC	3	11	1.87	1
istR	4	3	-0.42	1
istR	9	10	0.15	1
mcaS	10	53	2.41	0.006766374
mgrR	203	235	0.21	0.851761725
micA	17	33	0.96	1
micC	3	0		1
micF	0	7		1
micL	46	19	-1.28	0.454418658
ohsC	0	4		1
omrA	42	9	-2.22	0.010530681
omrB	15	5	-1.58	1
oxyS	111	310	1.48	0.233645419
pspH	1	10	3.32	1
ralA	0	4		1
rdlA	0	61		1
rdlB	0	44		1
rdlC	8	125	3.97	2.66E-37
rdlD	0	36		1
rirA	22	59	1.42	0.453484604
rnpB	2730	2297	-0.25	0.92608972
rprA	1	21	4.39	1
rttR	68	1104	4.02	1.42E-43
rybA	362	624	0.79	1
rybB	2	11	2.46	1
rydB	24	16	-0.58	1
rydC	24	48	1.00	1
ryeA	110	122	0.15	0.776736044

ryfA	42	187	2.15	0.012572412
ryfD	4753	3864	-0.30	1
ryhB	15	22	0.55	1
ryjA	0	15		1
ryjB	2	48	4.58	1.80E-83
sdhX	3	27	3.17	5.09E-09
sdsN	4	66	4.04	4.92E-26
sdsR	193	188	-0.04	0.779790448
sgrS	20	27	0.43	0.830128865
sibA	89	74	-0.27	1
sibB	17	36	1.08	1
sibC	0	60		1
sibD	46	266	2.53	0.00074587
sibE	46	76	0.72	1
sokB	34	37	0.12	0.888571938
sokC	54	56	0.05	0.88051718
sokE	0	10		1
sokX	0	30		1
spf	162	601	1.89	0.05225547
sraB	94	451	2.26	0.003996253
sraG	403	705	0.81	1
sroA	10	5	-1.00	1
sroC	41	19	-1.11	0.492747357
sroE	46	254	2.47	0.006194268
sroH	49	131	1.42	0.311517802
ssrA	409112	111531	-1.88	1
ssrS	4199	3570	-0.23	1
symR	0	22		1
tff	1911	4163	1.12	1
	Transcript Abundance (LB50S)	**Transcript Abundance (T50S)**		
PR1T50S	11245	103513	3.20	8.58E-09
PR2T50S	3320	51437	3.95	1.77E-36
3'ETS-leuZ	28	248	3.15	5.03E-07
agrA	48	52	0.12	0.828281002
agrB	3	25	3.06	3.83E-06
arcZ	235	410	0.80	1
chiX	9	18	1.00	1
cpxQ	46	3	-3.94	9.88E-33
csrB	1352	227	-2.57	1.37E-19
csrC	345	427	0.31	0.948393408
cyaR	84	211	1.33	0.336032396
dsrA	1	0		1
esrE	141	181	0.36	0.965578075
eyeA	6	39	2.70	8.26E-05
ffs	1302	343	-1.92	7.64E-07
fnrS	23285	24601	0.08	0.808052944
gadF	10	0		1
gadY	41	51	0.31	0.789027687
gcvB	35105	24732	-0.51	0.489198666

glmY	790	318	-1.31	0.006814023
glmZ	1554	1554	0.00	0.716332457
isrC	18	7	-1.36	1
istR	4	6	0.58	1
istR	4	0		1
mcaS	62	28	-1.15	0.34883989
mgrR	260	524	1.01	0.835856476
micA	69	78	0.18	0.925248696
micF	10	0		1
micL	67	50	-0.42	1
ohsC	2	5	1.32	1
omrA	42	0		1
omrB	53	0		1
oxyS	168	1085	2.69	8.26E-07
pspH	3	0		1
ralA	0	5		1
rdlA	34	82	1.27	0.762106744
rdlB	4	114	4.83	3.16E-126
rdlC	5	287	5.84	0
rdlD	7	50	2.84	8.14E-05
rirA	9	87	3.27	5.71E-09
rnpB	7291	5883	-0.31	0.860388391
rprA	0	11		1
rttR	52	3026	5.86	0
rybA	368	366	-0.01	0.508642709
rybB	16	32	1.00	1
rydB	6	21	1.81	1
rydC	6	52	3.12	1.03E-10
ryeA	217	129	-0.75	0.138136461
ryfA	47	299	2.67	3.57E-07
ryfD	4556	5807	0.35	1
ryhB	32	11	-1.54	0.257528161
ryjA	6	119	4.31	3.32E-41
ryjB	54	92	0.77	0.804506525
sdhX	47	15	-1.65	0.1221228
sdsN	40	142	1.83	0.093153599
sdsR	462	348	-0.41	0.307316534
sgrS	105	16	-2.71	3.62E-08
sibA	202	102	-0.99	0.152929744
sibB	106	38	-1.48	0.113500185
sibC	51	116	1.19	0.466210594
sibD	187	527	1.49	0.197141087
sibE	127	75	-0.76	0.341065226
sokB	96	53	-0.86	0.532957479
sokC	40	31	-0.37	1
sokE	0	33		0
sokX	7	24	1.78	0.09573145
spf	715	2923	2.03	0.005340886
sraB	195	735	1.91	0.021852226
sraG	239	560	1.23	0.415529533
sroA	17	32	0.91	1

sroC	33	102	1.63	0.181028216
sroE	125	173	0.47	1
sroH	140	160	0.19	0.712534113
ssrA	138117	31100	-2.15	7.24E-05
ssrS	25538	6277	-2.02	0.002617346
symR	3	54	4.17	6.35E-34
tff	2101	4405	1.07	0.490317389
	Transcript Abundance (LB30S)	**Transcript Abundance (T30S1)**		
PR1T30S1	5644	90261	4.00	9.70E-41
3'ETS-leuZ	75	248	1.73	0.0644502
agrA	29	72	1.31	0.281230584
agrB	8	30	1.91	0.079815368
arcZ	454	234	-0.96	0.053611494
chiX	70	35	-1.00	0.329482618
cpxQ	33	8	-2.04	0.01886393
csrB	6705	331	-4.34	1.88E-281
csrC	836	463	-0.85	0.105385549
cyaR	25	345	3.79	2.63E-32
dicF	6	0		1
dsrA	3	3	0.00	1
esrE	327	219	-0.58	0.166651286
eyeA	23	34	0.56	0.902821324
ffs	961	336	-1.52	0.000396232
fnrS	45413	12896	-1.82	0.00558617
gadF	7	0		1
gadY	42	75	0.84	0.166170026
gcvB	27296	29721	0.12	0.875772882
glmY	1814	444	-2.03	4.91E-08
glmZ	1528	1940	0.34	1
isrC	1	12	3.58	1
istR	0	15		1
istR	0	16		1
mcaS	39	70	0.84	0.515131359
mgrR	1078	509	-1.08	0.014954359
micA	76	57	-0.42	0.725118481
micC	0	4		1
micF	3	0		1
micL	66	45	-0.55	0.586799699
ohsC	1	4	2.00	1
omrA	92	15	-2.62	1.18E-05
omrB	191	12	-3.99	1.34E-39
oxyS	296	1051	1.83	0.000764631
pspH	16	33	1.04	0.447125221
ralA	0	6		1
rdlA	61	72	0.24	0.911389416
rdlB	9	100	3.47	4.38E-19
rdlC	10	141	3.82	2.93E-27
rdlD	49	140	1.51	0.15155205
rirA	54	79	0.55	0.669591391

rnpB	29119	10668	-1.45	0.126579806
rprA	61	9	-2.76	1.67E-05
rseX	21	0		1
rttR	124	1605	3.69	1.17E-58
rybA	319	423	0.41	1
rybB	32	24	-0.42	1
rydB	6	13	1.12	1
rydC	12	62	2.37	0.002229521
ryeA	3284	255	-3.69	4.11E-123
ryfA	250	573	1.20	0.115058889
ryfD	8359	1802	-2.21	1.08E-07
ryhB	26	0		0
ryjA	21	67	1.67	0.094538524
ryjB	39	52	0.42	0.750922338
sdhX	20	11	-0.86	1
sdsN	180	198	0.14	1
sdsR	6022	549	-3.46	2.22E-85
sgrS	126	23	-2.45	1.06E-06
sibA	862	68	-3.66	6.16E-67
sibB	352	47	-2.90	2.49E-14
sibC	398	182	-1.13	0.022754779
sibD	203	366	0.85	0.501707516
sibE	123	103	-0.26	0.573518287
sokB	354	68	-2.38	1.80E-05
sokC	142	104	-0.45	0.582038522
sokE	5	34	2.77	2.81E-07
sokX	12	32	1.42	0.252479244
spf	1745	1674	-0.06	0.92087401
sraB	500	975	0.96	0.31523039
sraG	304	518	0.77	0.479632959
sroA	49	41	-0.26	1
sroC	127	64	-0.99	0.126579806
sroE	72	117	0.70	1
sroH	457	162	-1.50	0.001106804
ssrA	81412	23780	-1.78	0.098493294
ssrS	130260	6447	-4.34	2.03E-158
symR	33	53	0.68	0.708945811
tff	3971	7270	0.87	0.542184005
	Transcript Abundance (LB30S)	**Transcript Abundance (T30S2)**		
PR1T30S2	5191	59302	3.51	1.43E-16
PR2T30S2	797	42257	5.73	0
PR3T30S2	543	40761	6.23	0
PR4T30S2	5256	35803	2.77	3.14E-13
3'ETS-leuZ	75	556	2.89	1.11E-08
agrA	29	101	1.80	0.057262649
agrB	8	28	1.81	0.098822296
arcZ	454	626	0.46	1
arrS	13	20	0.62	1
chiX	70	30	-1.22	0.188321052

cpxQ	33	5	-2.72	1.98E-07
csrB	6705	540	-3.63	4.96E-91
csrC	836	702	-0.25	0.511937988
cyaR	25	137	2.45	0.000278674
dicF	6	0		1
dsrA	3	8	1.42	1
esrE	327	196	-0.74	0.113977532
eyeA	23	24	0.06	1
ffs	961	456	-1.08	0.024547666
fnrS	45413	18838	-1.27	1
gadF	7	0		1
gadY	42	51	0.28	0.846645747
gcvB	27296	22595	-0.27	0.83360911
glmY	1814	1295	-0.49	0.31467202
glmZ	1528	2689	0.82	0.404673509
isrC	1	8	3.00	1
istR	0	17		1
istR	0	14		1
mcaS	39	82	1.07	0.372222408
mgrR	1078	481	-1.16	0.014103908
micA	76	38	-1.00	0.311917371
micC	0	5		1
micF	3	0		1
micL	66	111	0.75	0.717047259
ohsC	1	1	0.00	1
omrA	92	2	-5.52	0
omrB	191	9	-4.41	2.37E-86
oxyS	296	771	1.38	0.040076748
pspH	16	27	0.75	0.605772866
ralA	0	1		1
rdlA	61	98	0.68	0.721549582
rdlB	9	98	3.44	6.69E-18
rdlC	10	202	4.34	1.76E-62
rdlD	49	232	2.24	0.002141486
rirA	54	47	-0.20	0.761579055
rnpB	29119	38568	0.41	0.981697981
rseX	21	0		1
rttR	124	6364	5.68	0
rybA	319	397	0.32	1
rybB	32	24	-0.42	1
rydB	6	23	1.94	1
rydC	12	58	2.27	0.005597505
ryeA	3284	847	-1.96	7.88E-06
ryfA	250	1085	2.12	2.50E-05
ryfD	8359	3201	-1.38	0.079107831
ryhB	26	11	-1.24	0.298226082
ryjA	21	134	2.67	6.42E-06
ryjB	39	77	0.98	0.433041826
sdhX	20	25	0.32	1
sdsN	180	244	0.44	1
sdsR	6022	1776	-1.76	7.71E-05

sgrS	126	37	-1.77	0.001360036
sibA	862	200	-2.11	8.55E-09
sibB	352	87	-2.02	6.93E-06
sibC	398	174	-1.19	0.014174204
sibD	203	719	1.82	0.000215777
sibE	123	142	0.21	1
sokB	354	130	-1.45	0.028039149
sokC	142	109	-0.38	0.575376253
sokE	5	48	3.26	6.20E-16
sokX	12	7	-0.78	1
spf	1745	4469	1.36	0.056610913
sraB	500	1468	1.55	0.016646943
sraG	304	613	1.01	0.21086378
sroA	49	91	0.89	0.585956095
sroC	127	362	1.51	0.03093173
sroE	72	140	0.96	0.666485701
sroH	457	581	0.35	1
ssrA	81412	28387	-1.52	1
ssrS	130260	76427	-0.77	1
symR	33	38	0.20	1
tff	3971	5220	0.39	0.981270272

CHAPTER 5: Conclusions and Future Directions

In this thesis I studied the class of RNA called rancRNAs in *E. coli* with the aim to expand the fundamental knowledge of rancRNAs in bacteria. Firstly, I identified six previously uncharacterized rancRNA candidates from annotated sRNAs (arcZ, omrB, rprA, rttR, rybA, and tpke11) that associated with ribosomes (Figure 3.3, Table 3.4). Subsequently, I determined dissociation constants for five of the six candidates (arcZ, omrB, rprA, rttR, and rybA) to 70S ribosomes and 50S/30S ribosomal subunits, revealing that most bind with nanomolar affinity (Figure 3.6, Table 3.5). Using a reconstituted, coupled *in vitro* transcription-translation system, I observed that all candidates have minimal effects on protein synthesis, with the exception of omrB (Figure 3.7). rttR also does not appear to influence global translation *in vivo* (Figure 3.8). Using mass spectrometry can aid in determining if rttR regulates the translation of specific mRNAs. Additionally, determining a specific binding location for rttR on the ribosome biochemically by ultraviolet cross-linking and structurally using, for example, cryo electron microscopy will provide further insights into the mechanism of action for rttR.

I also conducted my own RNAseq study where I sequenced sRNAs isolated from 70S, 50S, and 30S ribosomes and subunits from cells grown under tetracycline, minimal medium, and hyperosmotic stress (Figure 4.3). I identified a total of 14 ncRNAs across all conditions and samples through DESeq2 analysis and 70 significantly enriched or depleted ncRNAs across all conditions and samples, including eight non-annotated transcripts, using Rockhopper (Tables 4.1, 4.2). Further *in silico* studies will be conducted on the non-annotated transcripts based on the pipeline established in chapter 3, followed by experimental validation to determine if these are true ncRNAs, specifically rancRNAs. Additionally, I observed a shoulder peak in the 30S region in the tetracycline stress condition that has not been characterized previously (Figure 4.3). My sequencing data provides some of the first insights into the composition of this peak. Investigating this is beyond the scope of my thesis.

REFERENCES

1. C. Carvalho Barbosa, S. H. Calhoun, H. J. Wieden, Non-coding RNAs: what are we missing? *Biochem Cell Biol* **98**, 23-30 (2020).

2. S. Gottesman, The small RNA regulators of Escherichia coli: roles and mechanisms*. *Annu Rev Microbiol* **58**, 303-328 (2004).

3. K. Prevost, G. Desnoyers, J. F. Jacques, F. Lavoie, E. Masse, Small RNA-induced mRNA degradation achieved through both translation block and activated cleavage. *Genes Dev* **25**, 385-396 (2011).

4. T. Mizuno, M. Y. Chou, M. Inouye, A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). *Proc Natl Acad Sci U S A* **81**, 1966-1970 (1984).

5. G. Desnoyers, E. Masse, Noncanonical repression of translation initiation through small RNA recruitment of the RNA chaperone Hfq. *Genes Dev* **26**, 726-739 (2012).

6. G. Desnoyers, M. P. Bouchard, E. Masse, New insights into small RNA-dependent translational regulation in prokaryotes. *Trends Genet* **29**, 92-98 (2013).

7. H. Luidalepp, S. Berger, O. Joss, T. Tenson, N. Polacek, Ribosome Shut-Down by 16S rRNA Fragmentation in Stationary-Phase Escherichia coli. *J Mol Biol* **428**, 2237-2247 (2016).

8. Z. Kang *et al.*, Small RNA regulators in bacteria: powerful tools for metabolic engineering and synthetic biology. *Appl Microbiol Biotechnol* **98**, 3413-3424 (2014).

9. D. Lalaouna *et al.*, A 3' external transcribed spacer in a tRNA transcript acts as a sponge for small RNAs to prevent transcriptional noise. *Mol Cell* **58**, 393-405 (2015).

10. S. P. Keam, A. Sobala, S. Ten Have, G. Hutvagner, tRNA-Derived RNA Fragments Associate with Human Multisynthetase Complex (MSC) and Modulate Ribosomal Protein Translation. *J Proteome Res* **16**, 413-420 (2017).

11. N. T. Ingolia, Ribosome profiling: new views of translation, from single codons to genome scale. *Nat Rev Genet* **15**, 205-213 (2014).

12. A. Pircher, J. Gebetsberger, N. Polacek, Ribosome-associated ncRNAs: an emerging class of translation regulators. *RNA Biol* **11**, 1335-1339 (2014).

13. M. Zywicki, K. Bakowska-Zywicka, N. Polacek, Revealing stable processing products from ribosome-associated small RNAs by deep-sequencing data analysis. *Nucleic Acids Res* **40**, 4013-4024 (2012).

14. J. Gebetsberger, M. Zywicki, A. Kunzi, N. Polacek, tRNA-derived fragments target the ribosome and function as regulatory non-coding RNA in Haloferax volcanii. *Archaea* **2012**, 260909 (2012).

15. V. Pecoraro, A. Rosina, N. Polacek, Ribosome-Associated ncRNAs (rancRNAs) Adjust Translation and Shape Proteomes. *Noncoding RNA* **8** (2022).

16. J. Gebetsberger, L. Wyss, A. M. Mleczko, J. Reuther, N. Polacek, A tRNA-derived fragment competes with mRNA for ribosome binding and regulates translation during stress. *RNA Biol* **14**, 1364-1373 (2017).

17. L. Wyss, M. Waser, J. Gebetsberger, M. Zywicki, N. Polacek, mRNA-specific translation regulation by a ribosome-associated ncRNA in Haloferax volcanii. *Sci Rep* **8**, 12502 (2018).

18. A. Pircher, K. Bakowska-Zywicka, L. Schneider, M. Zywicki, N. Polacek, An mRNA-derived noncoding RNA targets and regulates the ribosome. *Mol Cell* **54**, 147-155 (2014).

19. J. Reuther *et al.*, A small ribosome-associated ncRNA globally inhibits translation by restricting ribosome dynamics. *RNA Biol* **18**, 2617-2632 (2021).

20. K. Bakowska-Zywicka, M. Kasprzyk, T. Twardowski, tRNA-derived short RNAs bind to Saccharomyces cerevisiae ribosomes in a stress-dependent manner and inhibit protein synthesis in vitro. *FEMS Yeast Res* **16** (2016).

21. A. M. Mleczko, P. Celichowski, K. Bakowska-Zywicka, Transfer RNA-derived fragments target and regulate ribosome-associated aminoacyl-transfer RNA synthetases. *Biochim Biophys Acta Gene Regul Mech* 10.1016/j.bbagrm.2018.06.001 (2018).

22. R. Fricker *et al.*, A tRNA half modulates translation as stress response in Trypanosoma brucei. *Nat Commun* **10**, 118 (2019).

23. K. S. Rajan *et al.*, Developmentally Regulated Novel Non-coding Anti-sense Regulators of mRNA Translation in Trypanosoma b rucei. *iScience* **23**, 101780 (2020).

24. Y. Gonskikh *et al.*, Modulation of mammalian translation by a ribosome-associated tRNA half. *RNA Biol* **17**, 1125-1136 (2020).

25. S. Lalande, R. Merret, T. Salinas-Giege, L. Drouard, Arabidopsis tRNA-derived fragments as potential modulators of translation. *RNA Biol* **17**, 1137-1148 (2020).

26. M. L. Pinel-Marie *et al.*, RNA antitoxin SprF1 binds ribosomes to attenuate translation and promote persister cell formation in Staphylococcus aureus. *Nat Microbiol* **6**, 209-220 (2021).

27. C. Riffaud, M. L. Pinel-Marie, G. Pascreau, B. Felden, Functionality and cross-regulation of the four SprG/SprF type I toxin-antitoxin systems in Staphylococcus aureus. *Nucleic Acids Res* **47**, 1740-1758 (2019).

28. N. Raad, H. Luidalepp, M. Fasnacht, N. Polacek, Transcriptome-Wide Analysis of Stationary Phase Small ncRNAs in E. coli. *Int J Mol Sci* **22** (2021).

29. M. Sauert, H. Temmel, I. Moll, Heterogeneity of the translational machinery: Variations on a common theme. *Biochimie* **114**, 39-47 (2015).

30. A. Smirnov *et al.*, Grad-seq guides the discovery of ProQ as a major small RNA-binding protein. *Proc Natl Acad Sci U S A* **113**, 11591-11596 (2016).

31. S. A. Waters *et al.*, Small RNA interactome of pathogenic E. coli revealed through crosslinking of RNase E. *EMBO J* **36**, 374-387 (2017).

32. K. Neuhaus *et al.*, Differentiation of ncRNAs from small mRNAs in Escherichia coli O157:H7 EDL933 (EHEC) by combined RNAseq and RIBOseq - ryhB encodes the regulatory RNA RyhB and a peptide, RyhP. *BMC Genomics* **18**, 216 (2017).

33. J. Mata, S. Marguerat, J. Bahler, Post-transcriptional control of gene expression: a genome-wide perspective. *Trends Biochem Sci* **30**, 506-514 (2005).

34. E. J. a. H. Nestler, S.E., "Regulation of gene expression" in Neuropsychophrmacology: The Fifth Generation of Progress. (Philadelphia, Penn., 2002).

35. L. Lopez-Maury, S. Marguerat, J. Bahler, Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nat Rev Genet* **9**, 583-593 (2008).

36. S. Bhattacharjee, K. Renganaath, R. Mehrotra, S. Mehrotra, Combinatorial control of gene expression. *Biomed Res Int* **2013**, 407263 (2013).

37. G. Courtois, T. D. Gilmore, Mutations in the NF-kappaB signaling pathway: implications for human disease. *Oncogene* **25**, 6831-6843 (2006).

38. M. Shimobayashi, M. N. Hall, Making new contacts: the mTOR network in metabolism and signalling crosstalk. *Nat Rev Mol Cell Biol* **15**, 155-162 (2014).

39. J. W. Hershey, N. Sonenberg, M. B. Mathews, Principles of translational control: an overview. *Cold Spring Harb Perspect Biol* **4** (2012).

40. M. B. Mathews, Sonenberg, N., and Hershey, J. W. B., "Origins and principles of translational control" in Cold Spring Harbor Laboratory Press. (Cold Spring Harbor, N.Y., 2000).

41. F. Gebauer, M. W. Hentze, Molecular mechanisms of translational control. *Nat Rev Mol Cell Biol* **5**, 827-835 (2004).

42. T. Preiss, M. W. Hentze, Dual function of the messenger RNA cap structure in poly(A)-tail-promoted translation in yeast. *Nature* **392**, 516-520 (1998).

43. G. S. Wilkie, K. S. Dickson, N. K. Gray, Regulation of mRNA translation by 5'- and 3'-UTR-binding factors. *Trends Biochem Sci* **28**, 182-188 (2003).

44. R. C. Lee, R. L. Feinbaum, V. Ambros, The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843-854 (1993).

45. T. R. Cech, J. A. Steitz, The noncoding RNA revolution-trashing old rules to forge new ones. *Cell* **157**, 77-94 (2014).

46. F. Hube, C. Francastel, Coding and Non-coding RNAs, the Frontier Has Never Been So Blurred. *Front Genet* **9**, 140 (2018).

47. I. B. Rogozin *et al.*, Congruent evolution of different classes of non-coding DNA in prokaryotic genomes. *Nucleic Acids Res* **30**, 4264-4271 (2002).

48. J. S. Mattick, RNA regulation: a new genetics? *Nat Rev Genet* **5**, 316-323 (2004).

49. L. Ma, V. B. Bajic, Z. Zhang, On the classification of long non-coding RNAs. *RNA Biol* **10**, 925-933 (2013).

50. H. D. Bernstein, D. Zopf, D. M. Freymann, P. Walter, Functional substitution of the signal recognition particle 54-kDa subunit by its Escherichia coli homolog. *Proc Natl Acad Sci U S A* **90**, 5229-5233 (1993).

51. C. Carrieri *et al.*, Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature* **491**, 454-457 (2012).

52. T. R. Mercer, M. E. Dinger, J. S. Mattick, Long non-coding RNAs: insights into functions. *Nat Rev Genet* **10**, 155-159 (2009).

53. Z. Ji, R. Song, A. Regev, K. Struhl, Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. *Elife* **4**, e08890 (2015).

54. S. W. Choi, H. W. Kim, J. W. Nam, The small peptide world in long noncoding RNAs. *Brief Bioinform* **20**, 1853-1864 (2019).

55. P. D. Zamore, B. Haley, Ribo-gnome: the big world of small RNAs. *Science* **309**, 1519-1524 (2005).

56. M. A. Valencia-Sanchez, J. Liu, G. J. Hannon, R. Parker, Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes Dev* **20**, 515-524 (2006).

57. R. C. Friedman, K. K. Farh, C. B. Burge, D. P. Bartel, Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* **19**, 92-105 (2009).

58. M. R. Fabian, N. Sonenberg, W. Filipowicz, Regulation of mRNA translation and stability by microRNAs. *Annu Rev Biochem* **79**, 351-379 (2010).

59. Y. Tomari, P. D. Zamore, Perspective: machines for RNAi. *Genes Dev* **19**, 517-529 (2005).

60. R. W. Carthew, E. J. Sontheimer, Origins and Mechanisms of miRNAs and siRNAs. *Cell* **136**, 642-655 (2009).

61. S. M. Elbashir, W. Lendeckel, T. Tuschl, RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev* **15**, 188-200 (2001).

62. S. Gottesman, G. Storz, Bacterial small RNA regulators: versatile roles and rapidly evolving variations. *Cold Spring Harb Perspect Biol* **3** (2011).

63. C. H. Hoe, C. A. Raabe, T. S. Rozhdestvensky, T. H. Tang, Bacterial sRNAs: regulation in stress. *Int J Med Microbiol* **303**, 217-229 (2013).

64. D. Lalaouna, M. Simoneau-Roy, D. Lafontaine, E. Masse, Regulatory RNAs and target mRNA decay in prokaryotes. *Biochim Biophys Acta* **1829**, 742-747 (2013).

65. L. S. Waters, G. Storz, Regulatory RNAs in bacteria. *Cell* **136**, 615-628 (2009).

66. J. Georg, W. R. Hess, cis-antisense RNA, another level of gene regulation in bacteria. *Microbiol Mol Biol Rev* **75**, 286-300 (2011).

67. G. Storz, J. Vogel, K. M. Wassarman, Regulation by small RNAs in bacteria: expanding frontiers. *Mol Cell* **43**, 880-891 (2011).

68. A. Smirnov, C. Wang, L. L. Drewry, J. Vogel, Molecular mechanism of mRNA repression in trans by a ProQ-dependent small RNA. *EMBO J* **36**, 1029-1045 (2017).

69. Y. Chao *et al.*, In Vivo Cleavage Map Illuminates the Central Role of RNase E in Coding and Non-coding RNA Pathways. *Mol Cell* **65**, 39-51 (2017).

70. V. Khemici, L. Poljak, B. F. Luisi, A. J. Carpousis, The RNase E of Escherichia coli is a membrane-binding protein. *Mol Microbiol* **70**, 799-813 (2008).

71. H. Strahl *et al.*, Membrane recognition and dynamics of the RNA degradosome. *PLoS Genet* **11**, e1004961 (2015).

72. A. Jousselin, L. Metzinger, B. Felden, On the facultative requirement of the bacterial RNA chaperone, Hfq. *Trends Microbiol* **17**, 399-405 (2009).

73. J. Timmermans, L. Van Melderen, Post-transcriptional global regulation by CsrA in bacteria. *Cell Mol Life Sci* **67**, 2897-2908 (2010).

74. Y. Wu *et al.*, Strong YB-1 expression is associated with liver metastasis progression and predicts shorter disease-free survival in advanced gastric cancer. *J Surg Oncol* **105**, 724-730 (2012).

75. H. Goodarzi *et al.*, Endogenous tRNA-Derived Fragments Suppress Breast Cancer Progression via YBX1 Displacement. *Cell* **161**, 790-802 (2015).

76. G. Storz, S. Altuvia, K. M. Wassarman, An abundance of RNA regulators. *Annu Rev Biochem* **74**, 199-217 (2005).

77. K. I. Udekwu *et al.*, Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. *Genes Dev* **19**, 2355-2366 (2005).

78. H. Kawamoto, Y. Koide, T. Morita, H. Aiba, Base-pairing requirement for RNA silencing by a bacterial small RNA and acceleration of duplex formation by Hfq. *Mol Microbiol* **61**, 1013-1022 (2006).

79. Y. Chao, K. Papenfort, R. Reinhardt, C. M. Sharma, J. Vogel, An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs. *EMBO J* **31**, 4005-4019 (2012).

80. A. Valinezhad Orang, R. Safaralizadeh, M. Kazemzadeh-Bavili, Mechanisms of miRNA-Mediated Gene Regulation from Common Downregulation to mRNA-Specific Upregulation. *Int J Genomics* **2014**, 970607 (2014).

81. A. C. Tuck, D. Tollervey, RNA in pieces. *Trends Genet* **27**, 422-432 (2011).

82. N. R. De Lay, D. A. Garsin, The unmasking of 'junk' RNA reveals novel sRNAs: from processed RNA fragments to marooned riboswitches. *Curr Opin Microbiol* **30**, 16-21 (2016).

83. M. R. Garcia-Silva, F. Cabrera-Cabrera, M. C. Guida, A. Cayota, Hints of tRNA-Derived Small RNAs Role in RNA Silencing Mechanisms. *Genes (Basel)* **3**, 603-614 (2012).

84. T. Ogawa *et al.*, A cytotoxic ribonuclease targeting specific transfer RNA anticodons. *Science* **283**, 2097-2100 (1999).

85. D. M. Thompson, R. Parker, Stressing out over tRNA cleavage. *Cell* **138**, 215-219 (2009).

86. M. Shigematsu, S. Honda, Y. Kirino, Transfer RNA as a source of small functional RNA. *J Mol Biol Mol Imaging* **1** (2014).

87. R. L. Maute *et al.*, tRNA-derived microRNA modulates proliferation and the DNA damage response and is down-regulated in B cell lymphoma. *Proc Natl Acad Sci U S A* **110**, 1404-1409 (2013).

88. A. Sobala, G. Hutvagner, Small RNAs derived from the 5' end of tRNA can inhibit protein translation in human cells. *RNA Biol* **10**, 553-563 (2013).

89. K. M. Wassarman, F. Repoila, C. Rosenow, G. Storz, S. Gottesman, Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev* **15**, 1637-1651 (2001).

90. M. Ghildiyal *et al.*, Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. *Science* **320**, 1077-1081 (2008).

91. M. Miyakoshi, Y. Chao, J. Vogel, Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA. *EMBO J* **34**, 1478-1492 (2015).

92. Y. Chao, J. Vogel, A 3' UTR-Derived Small RNA Provides the Regulatory Noncoding Arm of the Inner Membrane Stress Response. *Mol Cell* **61**, 352-363 (2016).

93. T. R. Mercer *et al.*, Expression of distinct RNAs from 3' untranslated regions. *Nucleic Acids Res* **39**, 2393-2403 (2011).

94. A. M. Burroughs *et al.*, Deep-sequencing of human Argonaute-associated small RNAs provides insight into miRNA sorting and reveals Argonaute association with RNA fragments of diverse origin. *RNA Biol* **8**, 158-177 (2011).

95. B. D. Janssen, C. S. Hayes, The tmRNA ribosome-rescue system. *Adv Protein Chem Struct Biol* **86**, 151-191 (2012).

96. H. Himeno, D. Kurita, A. Muto, tmRNA-mediated trans-translation as the major ribosome rescue system in a bacterial cell. *Front Genet* **5**, 66 (2014).

97. D. Akopian, K. Shen, X. Zhang, S. O. Shan, Signal recognition particle: an essential protein-targeting machine. *Annu Rev Biochem* **82**, 693-721 (2013).

98. I. Saraogi, S. O. Shan, Co-translational protein targeting to the bacterial membrane. *Biochim Biophys Acta* **1843**, 1433-1441 (2014).

99. R. P. Perry, Balanced production of ribosomal proteins. *Gene* **401**, 1-3 (2007).

100. M. Barna, Ribosomes take control. *Proc Natl Acad Sci U S A* **110**, 9-10 (2013).

101. D. Barth-Baus *et al.*, S6 phosphorylation-independent pathways regulate translation of 5'-terminal oligopyrimidine tract-containing mRNAs in differentiating hematopoietic cells. *Nucleic Acids Res* **30**, 1919-1928 (2002).

102. G. Y. Soung, J. L. Miller, H. Koc, E. C. Koc, Comprehensive analysis of phosphorylated proteins of Escherichia coli ribosomes. *J Proteome Res* **8**, 3390-3402 (2009).

103. Z. Shi, M. Barna, Translating the genome in time and space: specialized ribosomes, RNA regulons, and RNA-binding proteins. *Annu Rev Cell Dev Biol* **31**, 31-54 (2015).

104. O. Vesper *et al.*, Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in Escherichia coli. *Cell* **147**, 147-157 (2011).

105. B. C. Buddingh, J. C. M. van Hest, Artificial Cells: Synthetic Compartments with Life-like Functionality and Adaptivity. *Acc Chem Res* **50**, 769-777 (2017).

106. D. N. Wilson, Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nat Rev Microbiol* **12**, 35-48 (2014).

107. P. Greulich, M. Scott, M. R. Evans, R. J. Allen, Growth-dependent bacterial susceptibility to ribosome-targeting antibiotics. *Mol Syst Biol* **11**, 796 (2015).

108. V. R. Kaberdin, U. Blasi, Translation initiation and the fate of bacterial mRNAs. *FEMS Microbiol Rev* **30**, 967-979 (2006).

109. N. Majdalani, C. K. Vanderpool, S. Gottesman, Bacterial small RNA regulators. *Crit Rev Biochem Mol Biol* **40**, 93-113 (2005).

110. P. Singer, M. Nomura, Stability of ribosomal protein mRNA and translational feedback regulation in Escherichia coli. *Mol Gen Genet* **199**, 543-546 (1985).

111. C. Ehresmann *et al.*, Molecular mimicry in translational regulation: the case of ribosomal protein S15. *RNA Biol* **1**, 66-73 (2004).

112. S. Domingues *et al.*, The role of RNase R in trans-translation and ribosomal quality control. *Biochimie* **114**, 113-118 (2015).

113. Z. F. Cheng, M. P. Deutscher, Purification and characterization of the Escherichia coli exoribonuclease RNase R. Comparison with RNase II. *J Biol Chem* **277**, 21624-21629 (2002).

114. D. L. Court *et al.*, RNase III: Genetics and function; structure and mechanism. *Annu Rev Genet* **47**, 405-431 (2013).

115. H. D. Robertson, R. E. Webster, N. D. Zinder, Purification and properties of ribonuclease III from Escherichia coli. *J Biol Chem* **243**, 82-91 (1968).

116. V. Pfeiffer, K. Papenfort, S. Lucchini, J. C. Hinton, J. Vogel, Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat Struct Mol Biol* **16**, 840-846 (2009).

117. K. Papenfort, M. Bouvier, F. Mika, C. M. Sharma, J. Vogel, Evidence for an autonomous 5' target recognition domain in an Hfq-associated small RNA. *Proc Natl Acad Sci U S A* **107**, 20435-20440 (2010).

118. A. H. Potts *et al.*, Global role of the bacterial post-transcriptional regulator CsrA revealed by integrated transcriptomics. *Nat Commun* **8**, 1596 (2017).

119. C. A. Vakulskas, A. H. Potts, P. Babitzke, B. M. Ahmer, T. Romeo, Regulation of bacterial virulence by Csr (Rsm) systems. *Microbiol Mol Biol Rev* **79**, 193-224 (2015).

120. J. Vazquez-Anderson, L. M. Contreras, Regulatory RNAs: charming gene management styles for synthetic biology applications. *RNA Biol* **10**, 1778-1797 (2013).

121. E. Levine, Z. Zhang, T. Kuhlman, T. Hwa, Quantitative characteristics of gene regulation by small RNA. *PLoS Biol* **5**, e229 (2007).

122. M. Lynch, G. K. Marinov, The bioenergetic costs of a gene. *Proc Natl Acad Sci U S A* **112**, 15690-15695 (2015).

123. N. Lane, W. Martin, The energetics of genome complexity. *Nature* **467**, 929-934 (2010).

124. C. L. Beisel, G. Storz, Base pairing small RNAs and their roles in global regulatory networks. *FEMS Microbiol Rev* **34**, 866-882 (2010).

125. M. Y. Sherman, S. B. Qian, Less is more: improving proteostasis by translation slow down. *Trends Biochem Sci* **38**, 585-591 (2013).

126. Y. Shimoni *et al.*, Regulation of gene expression by small non-coding RNAs: a quantitative view. *Mol Syst Biol* **3**, 138 (2007).

127. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).

128. R. Lorenz *et al.*, ViennaRNA Package 2.0. *Algorithms Mol Biol* **6**, 26 (2011).

129. K. Darty, A. Denise, Y. Ponty, VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics* **25**, 1974-1975 (2009).

130. M. J. Boniecki *et al.*, SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction. *Nucleic Acids Res* **44**, e63 (2016).

131. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J Mol Biol* **215**, 403-410 (1990).

132. M. Goujon *et al.*, A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res* **38**, W695-699 (2010).

133. H. McWilliam *et al.*, Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res* **41**, W597-600 (2013).

134. I. M. Keseler *et al.*, EcoCyc: a comprehensive database of Escherichia coli biology. *Nucleic Acids Res* **39**, D583-590 (2011).

135. A. Busch, A. S. Richter, R. Backofen, IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics* **24**, 2849-2856 (2008).

136. M. Mann, P. R. Wright, R. Backofen, IntaRNA 2.0: enhanced and customizable prediction of RNA-RNA interactions. *Nucleic Acids Res* **45**, W435-W439 (2017).

137. U. Muckstein *et al.*, Thermodynamics of RNA-RNA binding. *Bioinformatics* **22**, 1177-1182 (2006).

138. A. R. Gruber, R. Lorenz, S. H. Bernhart, R. Neubock, I. L. Hofacker, The Vienna RNA websuite. *Nucleic Acids Res* **36**, W70-74 (2008).

139. S. D. Moore, T. A. Baker, R. T. Sauer, Forced extraction of targeted components from complex macromolecular assemblies. *Proc Natl Acad Sci U S A* **105**, 11685-11690 (2008).

140. R. Hatzenpichler *et al.*, In situ visualization of newly synthesized proteins in environmental microbes using amino acid tagging and click chemistry. *Environ Microbiol* **16**, 2568-2590 (2014).

141. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**, 671-675 (2012).

142. S. Lee, D. K. Lee, What is the proper way to apply the multiple comparison test? *Korean J Anesthesiol* **71**, 353-360 (2018).

143. M. Bosl, H. Kersten, A novel RNA product of the tyrT operon of Escherichia coli. *Nucleic Acids Res* **19**, 5863-5870 (1991).

144. T. Sekiya *et al.*, Total synthesis of a tyrosine suppressor transfer RNA gene. XVI. Enzymatic joinings to form the total 207-base pair-long DNA. *J Biol Chem* **254**, 5787-5801 (1979).

145. H. Heberle, G. V. Meirelles, F. R. da Silva, G. P. Telles, R. Minghim, InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics* **16**, 169 (2015).

146. L. Argaman *et al.*, Novel small RNA-encoding genes in the intergenic regions of Escherichia coli. *Curr Biol* **11**, 941-950 (2001).

147. K. Papenfort *et al.*, Specific and pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-dependent small RNA. *Mol Microbiol* **74**, 139-158 (2009).

148. T. Soper, P. Mandin, N. Majdalani, S. Gottesman, S. A. Woodson, Positive regulation by small RNAs and the role of Hfq. *Proc Natl Acad Sci U S A* **107**, 9602-9607 (2010).

149. P. Mandin, S. Gottesman, Integrating anaerobic/aerobic sensing and the general stress response through the ArcZ small RNA. *EMBO J* **29**, 3094-3107 (2010).

150. N. Sedlyarova *et al.*, sRNA-Mediated Control of Transcription Termination in E. coli. *Cell* **167**, 111-121 e113 (2016).

151. J. Vogel *et al.*, RNomics in Escherichia coli detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic Acids Res* **31**, 6435-6443 (2003).

152. N. Majdalani, S. Chen, J. Murrow, K. St John, S. Gottesman, Regulation of RpoS by a novel small RNA: the characterization of RprA. *Mol Microbiol* **39**, 1382-1394 (2001).

153. R. Madhugiri, S. R. Basineni, G. Klug, Turn-over of the small non-coding RNA RprA in E. coli is influenced by osmolarity. *Mol Genet Genomics* **284**, 307-318 (2010).

154. K. Gerstle, K. Klatschke, U. Hahn, N. Piganeau, The small RNA RybA regulates key-genes in the biosynthesis of aromatic amino acids under peroxide stress in E. coli. *RNA Biol* **9**, 458-468 (2012).

155. E. Rivas, R. J. Klein, T. A. Jones, S. R. Eddy, Computational identification of noncoding RNAs in E. coli by comparative genomics. *Curr Biol* **11**, 1369-1373 (2001).

156. H. F. Noller, C. R. Woese, Secondary structure of 16S ribosomal RNA. *Science* **212**, 403-411 (1981).

157. A. Meskauskas, J. D. Dinman, Ribosomal protein L3: gatekeeper to the A site. *Mol Cell* **25**, 877-888 (2007).

158. S. C. Blanchard, H. D. Kim, R. L. Gonzalez, Jr., J. D. Puglisi, S. Chu, tRNA dynamics on the ribosome during translation. *Proc Natl Acad Sci U S A* **101**, 12893-12898 (2004).

159. Y. Shimizu *et al.*, Cell-free translation reconstituted with purified components. *Nat Biotechnol* **19**, 751-755 (2001).

160. L. Gomes, G. Monteiro, F. Mergulhao, The Impact of IPTG Induction on Plasmid Stability and Heterologous Protein Expression by Escherichia coli Biofilms. *Int J Mol Sci* **21** (2020).

161. J. D. Bagert *et al.*, Quantitative, time-resolved proteomic analysis by combining bioorthogonal noncanonical amino acid tagging and pulsed stable isotope labeling by amino acids in cell culture. *Mol Cell Proteomics* **13**, 1352-1358 (2014).

162. J. M. Andrade, R. F. Dos Santos, I. Chelysheva, Z. Ignatova, C. M. Arraiano, The RNA-binding protein Hfq is important for ribosome biogenesis and affects translation fidelity. *EMBO J* **37** (2018).

163. I. M. Sharma, A. Korman, S. A. Woodson, The Hfq chaperone helps the ribosome mature. *EMBO J* **37** (2018).

164. K. McLean *et al.*, Improved Species-Level Clinical Identification of Enterobacteriaceae through Broad-Range dnaJ PCR and Sequencing. *J Clin Microbiol* **57** (2019).

165. M. Adeolu, S. Alnajar, S. Naushad, S. G. R, Genome-based phylogeny and taxonomy of the 'Enterobacteriales': proposal for Enterobacterales ord. nov. divided into the families Enterobacteriaceae, Erwiniaceae fam. nov., Pectobacteriaceae fam. nov., Yersiniaceae fam. nov., Hafniaceae fam. nov., Morganellaceae fam. nov., and Budviciaceae fam. nov. *Int J Syst Evol Microbiol* **66**, 5575-5599 (2016).

166. M. O'Connor, A. E. Dahlberg, Enhancement of translation by the epsilon element is independent of the sequence of the 460 region of 16S rRNA. *Nucleic Acids Res* **29**, 1420-1425 (2001).

167. N. B. Leontis, E. Westhof, The annotation of RNA motifs. *Comp Funct Genomics* **3**, 518-524 (2002).

168. R. R. Gutell, N. Larsen, C. R. Woese, Lessons from an evolving rRNA: 16S and 23S rRNA structures from a comparative perspective. *Microbiol Rev* **58**, 10-26 (1994).

169. J. C. Wu, D. P. Gardner, S. Ozer, R. R. Gutell, P. Ren, Correlation of RNA secondary structure statistics with thermodynamic stability and applications to folding. *J Mol Biol* **391**, 769-783 (2009).

170. M. Bubunenko *et al.*, 30S ribosomal subunits can be assembled in vivo without primary binding ribosomal protein S15. *RNA* **12**, 1229-1239 (2006).

171. P. Ramaswamy, S. A. Woodson, S16 throws a conformational switch during assembly of 30S 5' domain. *Nat Struct Mol Biol* **16**, 438-445 (2009).

172. W. A. Held, B. Ballou, S. Mizushima, M. Nomura, Assembly mapping of 30 S ribosomal proteins from Escherichia coli. Further studies. *J Biol Chem* **249**, 3103-3111 (1974).

173. R. Kaplan, D. Apirion, The fate of ribosomes in Escherichia coli cells starved for a carbon source. *J Biol Chem* **250**, 1854-1863 (1975).

174. U. Maivali, A. Paier, T. Tenson, When stable RNA becomes unstable: the degradation of ribosomes in bacteria and beyond. *Biol Chem* **394**, 845-855 (2013).

175. G. Z. Yusupova, M. M. Yusupov, J. H. Cate, H. F. Noller, The path of messenger RNA through the ribosome. *Cell* **106**, 233-241 (2001).

176. D. E. Draper, L. P. Reynaldo, RNA binding strategies of ribosomal proteins. *Nucleic Acids Res* **27**, 381-388 (1999).

177. A. R. Ferre-D'Amare, RNA Binding: Getting Specific about Specificity. *Cell Chem Biol* **23**, 1177-1178 (2016).

178. M. Guillier, S. Gottesman, G. Storz, Modulating the outer membrane with small RNAs. *Genes Dev* **20**, 2338-2348 (2006).

179. M. Guillier, S. Gottesman, Remodelling of the Escherichia coli outer membrane by two small regulatory RNAs. *Mol Microbiol* **59**, 231-247 (2006).

180. V. H. Nagaraj, J. M. Greene, A. M. Sengupta, E. D. Sontag, Translation inhibition and resource balance in the TX-TL cell-free gene expression system. *Synth Biol (Oxf)* **2**, ysx005 (2017).

181. J. L. McGinnis *et al.*, In-cell SHAPE reveals that free 30S ribosome subunits are in the inactive state. *Proc Natl Acad Sci U S A* **112**, 2425-2430 (2015).

182. B. Roy-Chaudhuri, N. Kirthi, G. M. Culver, Appropriate maturation and folding of 16S rRNA during 30S subunit biogenesis are critical for translational fidelity. *Proc Natl Acad Sci U S A* **107**, 4567-4572 (2010).

183. F. Fegatella, J. Lim, S. Kjelleberg, R. Cavicchioli, Implications of rRNA operon copy number and ribosome content in the marine oligotrophic ultramicrobacterium Sphingomonas sp. strain RB2256. *Appl Environ Microbiol* **64**, 4433-4438 (1998).

184. S. F. Clatterbuck Soper, R. P. Dator, P. A. Limbach, S. A. Woodson, In vivo X-ray footprinting of pre-30S ribosomes reveals chaperone-dependent remodeling of late assembly intermediates. *Mol Cell* **52**, 506-516 (2013).

185. A. Jomaa *et al.*, Understanding ribosome assembly: the structure of in vivo assembled immature 30S subunits revealed by cryo-electron microscopy. *RNA* **17**, 697-709 (2011).

186. M. R. Bosl, Genetic map of the tyrT region of Escherichia coli from 27.1 to 27.7 minutes based exclusively on sequence data. *J Bacteriol* **175**, 7751-7753 (1993).

187. A. L. Starosta, J. Lassak, K. Jung, D. N. Wilson, The bacterial translation stress response. *FEMS Microbiol Rev* **38**, 1172-1201 (2014).

188. A. Bartholomaus *et al.*, Bacteria differently regulate mRNA abundance to specifically respond to various stresses. *Philos Trans A Math Phys Eng Sci* **374** (2016).

189. K. M. Wassarman, Small RNAs in bacteria: diverse regulators of gene expression in response to environmental changes. *Cell* **109**, 141-144 (2002).

190. S. Gottesman *et al.*, Small RNA regulators and the bacterial response to stress. *Cold Spring Harb Symp Quant Biol* **71**, 1-11 (2006).

191. E. C. Hobbs, J. L. Astarita, G. Storz, Small RNAs and small proteins involved in resistance to cell envelope stress and acid shock in Escherichia coli: analysis of a bar-coded mutant collection. *J Bacteriol* **192**, 59-67 (2010).

192. C. N. Peterson, M. J. Mandel, T. J. Silhavy, Escherichia coli starvation diets: essential nutrients weigh in distinctly. *J Bacteriol* **187**, 7549-7553 (2005).

193. M. J. Brauer *et al.*, Conservation of the metabolomic response to starvation across two divergent microbes. *Proc Natl Acad Sci U S A* **103**, 19302-19307 (2006).

194. C. L. Patten, M. G. Kirchhof, M. R. Schertzberg, R. A. Morton, H. E. Schellhorn, Microarray analysis of RpoS-mediated gene expression in Escherichia coli K-12. *Mol Genet Genomics* **272**, 580-591 (2004).

195. H. E. Schellhorn, Function, Evolution, and Composition of the RpoS Regulon in Escherichia coli. *Front Microbiol* **11**, 560099 (2020).

196. T. Weilbacher *et al.*, A novel sRNA component of the carbon storage regulatory system of Escherichia coli. *Mol Microbiol* **48**, 657-670 (2003).

197. T. Romeo, Global regulation by the small RNA-binding protein CsrA and the non-coding RNA molecule CsrB. *Mol Microbiol* **29**, 1321-1330 (1998).

198. K. Jonas, O. Melefors, The Escherichia coli CsrB and CsrC small RNAs are strongly induced during growth in nutrient-poor medium. *FEMS Microbiol Lett* **297**, 80-86 (2009).

199. Y. Hao, T. B. Updegrove, N. N. Livingston, G. Storz, Protection against deleterious nitrogen compounds: role of sigmaS-dependent small RNAs encoded adjacent to sdiA. *Nucleic Acids Res* **44**, 6935-6948 (2016).

200. W. Kim, Y. Lee, Mechanism for coordinate regulation of rpoS by sRNA-sRNA interaction in Escherichia coli. *RNA Biol* **17**, 176-187 (2020).

201. M. L. Urbanowski, L. T. Stauffer, G. V. Stauffer, The gcvB gene encodes a small untranslated RNA involved in expression of the dipeptide and oligopeptide transport systems in Escherichia coli. *Mol Microbiol* **37**, 856-868 (2000).

202. N. De Lay, S. Gottesman, The Crp-activated small noncoding regulatory RNA CyaR (RyeE) links nutritional status to group behavior. *J Bacteriol* **191**, 461-476 (2009).

203. C. M. Sharma, F. Darfeuille, T. H. Plantinga, J. Vogel, A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. *Genes Dev* **21**, 2804-2817 (2007).

204. C. M. Sharma *et al.*, Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA. *Mol Microbiol* **81**, 1144-1165 (2011).

205. H. Wang *et al.*, Resistance Profiles of Salmonella Isolates Exposed to Stresses and the Expression of Small Non-coding RNAs. *Front Microbiol* **11**, 130 (2020).

206. P. Dersch, M. A. Khan, S. Muhlen, B. Gorke, Roles of Regulatory RNAs for Antibiotic Resistance in Bacteria and Their Potential Value as Novel Drug Targets. *Front Microbiol* **8**, 803 (2017).

207. I. Levin-Reisman *et al.*, Antibiotic tolerance facilitates the evolution of resistance. *Science* **355**, 826-830 (2017).

208. B. W. Kwan, J. A. Valenta, M. J. Benedik, T. K. Wood, Arrested protein synthesis increases persister-like cell formation. *Antimicrob Agents Chemother* **57**, 1468-1473 (2013).

209. Y. Gonskikh, N. Polacek, Alterations of the translation apparatus during aging and stress response. *Mech Ageing Dev* **168**, 30-36 (2017).

210. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).

211. F. R. Blattner *et al.*, The complete genome sequence of Escherichia coli K-12. *Science* **277**, 1453-1462 (1997).

212. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359 (2012).

213. Y. Liao, G. K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014).

214. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).

215. H. Mi *et al.*, PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Res* **49**, D394-D403 (2021).

216. R. McClure *et al.*, Computational analysis of bacterial RNA-Seq data. *Nucleic Acids Res* **41**, e140 (2013).

217. B. Tjaden, De novo assembly of bacterial transcriptomes from RNA-seq data. *Genome Biol* **16**, 1 (2015).

218. B. Tjaden, A computational system for identifying operons based on RNA-seq data. *Methods* **176**, 62-70 (2020).

219. J. T. Robinson *et al.*, Integrative genomics viewer. *Nat Biotechnol* **29**, 24-26 (2011).

220. V. Ribes, K. Romisch, A. Giner, B. Dobberstein, D. Tollervey, E. coli 4.5S RNA is part of a ribonucleoprotein particle that has properties related to signal recognition particle. *Cell* **63**, 591-600 (1990).

221. K. Nakamura, Y. Fujii, T. Shibata, K. Yamane, Depletion of Escherichia coli 4.5S RNA leads to an increase in the amount of protein elongation factor EF-G associated with ribosomes. *Eur J Biochem* **259**, 543-550 (1999).

222. B. Reichenbach, A. Maes, F. Kalamorz, E. Hajnsdorf, B. Gorke, The small RNA GlmY acts upstream of the sRNA GlmZ in the activation of glmS expression and is subject to regulation by polyadenylation in Escherichia coli. *Nucleic Acids Res* **36**, 2570-2580 (2008).

223. A. K. Gupta, N. Siddiqui, D. Yadav, L. Arora, T. Dutta, Regulation of RyeA/SraC expression in Escherichia coli. *Biochem Biophys Res Commun* **516**, 661-665 (2019).

224. E. M. Fozo *et al.*, Repression of small toxic protein synthesis by the Sib and OhsC small RNAs. *Mol Microbiol* **70**, 1076-1093 (2008).

225. K. M. Wassarman, G. Storz, 6S RNA regulates E. coli RNA polymerase activity. *Cell* **101**, 613-623 (2000).

226. J. Johansen, M. Eriksen, B. Kallipolitis, P. Valentin-Hansen, Down-regulation of outer membrane proteins by noncoding RNAs: unraveling the cAMP-CRP- and sigmaE-dependent CyaR-ompX regulatory case. *J Mol Biol* **383**, 1-9 (2008).

227. X. Yin *et al.*, The small protein MgtS and small RNA MgrR modulate the PitA phosphate symporter to boost intracellular magnesium levels. *Mol Microbiol* **111**, 131-144 (2019).

228. H. Cho, K. S. Kim, Escherichia coli OxyS RNA triggers cephalothin resistance by modulating the expression of CRP-associated genes. *Biochem Biophys Res Commun* **506**, 66-72 (2018).

229. L. Andresen, Y. Martinez-Burgo, J. Nilsson Zangelin, A. Rizvanovic, E. Holmqvist, The Small Toxic Salmonella Protein TimP Targets the Cytoplasmic Membrane and Is Repressed by the Small RNA TimR. *mBio* **11** (2020).

230. I. J. Silva, A. D. Ortega, S. C. Viegas, F. Garcia-Del Portillo, C. M. Arraiano, An RpoS-dependent sRNA regulates the expression of a chaperone involved in protein folding. *RNA* **19**, 1253-1265 (2013).

231. A. Muto *et al.*, Structure and function of 10Sa RNA: trans-translation system. *Biochimie* **78**, 985-991 (1996).

232. R. Weel-Sneve *et al.*, Single transmembrane peptide DinQ modulates membrane-dependent activities. *PLoS Genet* **9**, e1003260 (2013).

233. P. Mandin, S. Gottesman, A genetic approach for finding small RNAs regulators of genes of interest identifies RybC as regulating the DpiA/DpiB two-component system. *Mol Microbiol* **72**, 551-565 (2009).

234. M. Grabowicz, D. Koren, T. J. Silhavy, The CpxQ sRNA Negatively Regulates Skp To Prevent Mistargeting of beta-Barrel Outer Membrane Proteins into the Cytoplasmic Membrane. *mBio* **7**, e00312-00316 (2016).

235. M. Y. Liu *et al.*, The RNA molecule CsrB binds to the global regulatory protein CsrA and antagonizes its activity in Escherichia coli. *J Biol Chem* **272**, 17502-17510 (1997).

236. J. A. Opdyke, E. M. Fozo, M. R. Hemm, G. Storz, RNase III participates in GadY-dependent cleavage of the gadX-gadW mRNA. *J Mol Biol* **406**, 29-43 (2011).

237. J. A. Opdyke, J. G. Kang, G. Storz, GadY, a small-RNA regulator of acid response genes in Escherichia coli. *J Bacteriol* **186**, 6698-6705 (2004).

238. M. S. Guo *et al.*, MicL, a new sigmaE-dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein. *Genes Dev* **28**, 1620-1634 (2014).

239. M. Kawano, T. Oshima, H. Kasai, H. Mori, Molecular characterization of long direct repeat (LDR) sequences expressing a stable mRNA encoding for a 35-amino-acid cell-killing peptide and a cis-encoded small antisense RNA in Escherichia coli. *Mol Microbiol* **45**, 333-349 (2002).

240. A. M. King, C. K. Vanderpool, P. H. Degnan, sRNA Target Prediction Organizing Tool (SPOT) Integrates Computational and Experimental Data To Facilitate Functional Characterization of Bacterial Small RNAs. *mSphere* **4** (2019).

241. C. M. Bianco, K. S. Frohlich, C. K. Vanderpool, Bacterial Cyclopropane Fatty Acid Synthase mRNA Is Targeted by Activating and Repressing Small RNAs. *J Bacteriol* **201** (2019).

242. E. Masse, S. Gottesman, A small RNA regulates the expression of genes involved in iron metabolism in Escherichia coli. *Proc Natl Acad Sci U S A* **99**, 4620-4625 (2002).

243. E. Masse, C. K. Vanderpool, S. Gottesman, Effect of RyhB small RNA on global iron use in Escherichia coli. *J Bacteriol* **187**, 6962-6971 (2005).

244. H. Chen, A. Previero, M. P. Deutscher, A novel mechanism of ribonuclease regulation: GcvB and Hfq stabilize the mRNA that encodes RNase BN/Z during exponential phase. *J Biol Chem* **294**, 19997-20008 (2019).

245. R. Raghavan, E. A. Groisman, H. Ochman, Genome-wide detection of novel regulatory RNAs in E. coli. *Genome Res* **21**, 1487-1497 (2011).

246. M. Miyakoshi, G. Matera, K. Maki, Y. Sone, J. Vogel, Functional expansion of a TCA cycle operon mRNA by a 3' end-derived small RNA. *Nucleic Acids Res* **47**, 2075-2088 (2019).

247. F. De Mets, L. Van Melderen, S. Gottesman, Regulation of acetate metabolism and coordination with the TCA cycle via a processed small RNA. *Proc Natl Acad Sci U S A* **116**, 1043-1052 (2019).

248. A. Gutierrez *et al.*, beta-Lactam antibiotics promote bacterial mutagenesis via an RpoS-mediated reduction in replication fidelity. *Nat Commun* **4**, 1610 (2013).

249. J. S. Choi *et al.*, The small RNA, SdsR, acts as a novel type of toxin in Escherichia coli. *RNA Biol* **15**, 1319-1335 (2018).

250. L. K. Poulsen, A. Refn, S. Molin, P. Andersson, The gef gene from Escherichia coli is regulated at the level of translation. *Mol Microbiol* **5**, 1639-1648 (1991).

251. D. Lalaouna, A. Eyraud, A. Devinck, K. Prevost, E. Masse, GcvB small RNA uses two distinct seed regions to regulate an extensive targetome. *Mol Microbiol* **111**, 473-486 (2019).

252. M. K. Mihailovic *et al.*, High-throughput in vivo mapping of RNA accessible interfaces to identify functional sRNA binding sites. *Nat Commun* **9**, 4084 (2018).

253. M. Kawano, L. Aravind, G. Storz, An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin. *Mol Microbiol* **64**, 738-754 (2007).

254. S. Durand, G. Storz, Reprogramming of anaerobic metabolism by the FnrS small RNA. *Mol Microbiol* **75**, 1215-1231 (2010).

255. A. Boysen, J. Moller-Jensen, B. Kallipolitis, P. Valentin-Hansen, M. Overgaard, Translational regulation of gene expression by an anaerobically induced small non-coding RNA in Escherichia coli. *J Biol Chem* **285**, 10690-10702 (2010).

256. J. Vogel, L. Argaman, E. G. Wagner, S. Altuvia, The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide. *Curr Biol* **14**, 2271-2276 (2004).

257. A. A. Rasmussen *et al.*, Regulation of ompA mRNA stability: the role of a small regulatory RNA in growth phase-dependent control. *Mol Microbiol* **58**, 1421-1429 (2005).

258. J. Andersen, S. A. Forst, K. Zhao, M. Inouye, N. Delihas, The function of micF RNA. micF RNA is a major factor in the thermal regulation of OmpF protein in Escherichia coli. *J Biol Chem* **264**, 17961-17970 (1989).

259. S. Melamed *et al.*, Global Mapping of Small RNA-Target Interactions in Bacteria. *Mol Cell* **63**, 884-897 (2016).

260. M. Kawano, A. A. Reynolds, J. Miranda-Rios, G. Storz, Detection of 5'- and 3'-UTR-derived small RNAs and cis-encoded antisense RNAs in Escherichia coli. *Nucleic Acids Res* **33**, 1040-1050 (2005).

261. S. Melamed, P. P. Adams, A. Zhang, H. Zhang, G. Storz, RNA-RNA Interactomes of ProQ and Hfq Reveal Overlapping and Competing Roles. *Mol Cell* **77**, 411-425 e417 (2020).

262. K. Nakashima, K. Kanamaru, T. Mizuno, K. Horikoshi, A novel member of the cspA family of genes that is induced by cold shock in Escherichia coli. *J Bacteriol* **178**, 2994-2997 (1996).

263. M. G. Jorgensen *et al.*, Small regulatory RNAs control the multi-cellular adhesive lifestyle of Escherichia coli. *Mol Microbiol* **84**, 36-50 (2012).

264. R. W. van Nues, D. Castro-Roa, Y. Yuzenkova, N. Zenkin, Ribonucleoprotein particles of bacterial small non-coding RNA IsrA (IS61 or McaS) and its interaction with RNA polymerase core may link transcription to mRNA fate. *Nucleic Acids Res* **44**, 2577-2592 (2016).

265. G. Klein *et al.*, Multiple Transcriptional Factors Regulate Transcription of the rpoE Gene in Escherichia coli under Different Growth Conditions and When the Lipopolysaccharide Biosynthesis Is Defective. *J Biol Chem* **291**, 22999-23019 (2016).

266. C. K. Vanderpool, S. Gottesman, Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Mol Microbiol* **54**, 1076-1089 (2004).

267. K. Pedersen, K. Gerdes, Multiple hok genes on the chromosome of Escherichia coli. *Mol Microbiol* **32**, 1090-1102 (1999).

268. X. Wang *et al.*, Expression of each cistron in the gal operon can be regulated by transcription termination and generation of a galk-specific mRNA, mK2. *J Bacteriol* **196**, 2598-2606 (2014).

269. J. P. Etchegaray, P. G. Jones, M. Inouye, Differential thermoregulation of two highly homologous cold-shock genes, cspA and cspB, of Escherichia coli. *Genes Cells* **1**, 171-178 (1996).

270. S. J. Lee *et al.*, Family of the major cold-shock protein, CspA (CS7.4), of Escherichia coli, whose members show a high sequence similarity with the eukaryotic Y-box binding proteins. *Mol Microbiol* **11**, 833-839 (1994).

271. L. Fang, Y. Hou, M. Inouye, Role of the cold-box region in the 5' untranslated region of the cspA mRNA in its transient expression at low temperature in Escherichia coli. *J Bacteriol* **180**, 90-95 (1998).

272. R. Brielle, M. L. Pinel-Marie, S. Chat, R. Gillet, B. Felden, Purification, identification, and functional analysis of polysomes from the human pathogen Staphylococcus aureus. *Methods* **117**, 59-66 (2017).

273. L. E. Day, Tetracycline inhibition of cell-free protein synthesis. I. Binding of tetracycline to components of the system. *J Bacteriol* **91**, 1917-1923 (1966).

274. B. Liu, C. Chen, Translation Elongation Factor 4 (LepA) Contributes to Tetracycline Susceptibility by Stalling Elongating Ribosomes. *Antimicrob Agents Chemother* **62** (2018).

275. C. U. Chukwudi, rRNA Binding Sites and the Molecular Mechanism of Action of the Tetracyclines. *Antimicrob Agents Chemother* **60**, 4433-4441 (2016).

276. C. U. Chukwudi, L. Good, Interaction of the tetracyclines with double-stranded RNAs of random base sequence: new perspectives on the target and mechanism of action. *J Antibiot (Tokyo)* **69**, 622-630 (2016).

277. Y. Hase, T. Tarusawa, A. Muto, H. Himeno, Impairment of ribosome maturation or function confers salt resistance on Escherichia coli cells. *PLoS One* **8**, e65747 (2013).

278. M. H. Pontes, J. Yeom, E. A. Groisman, Reducing Ribosome Biosynthesis Promotes Translation during Low Mg(2+) Stress. *Mol Cell* **64**, 480-492 (2016).

279. H. Cheng-Guang, C. O. Gualerzi, The Ribosome as a Switchboard for Bacterial Stress Response. *Front Microbiol* **11**, 619038 (2020).

280. M. Fessler, B. Gummesson, G. Charbon, S. L. Svenningsen, M. A. Sorensen, Short-term kinetics of rRNA degradation in Escherichia coli upon starvation for carbon, amino acid or phosphate. *Mol Microbiol* **113**, 951-963 (2020).

281. A. Jacobson, D. Gillespie, Metabolic events occurring during recovery from prolonged glucose starvation in Escherichia coli. *J Bacteriol* **95**, 1030-1039 (1968).

282. M. A. Zundel, G. N. Basturea, M. P. Deutscher, Initiation of ribosome degradation during starvation in Escherichia coli. *RNA* **15**, 977-983 (2009).

283. S. H. Yoon *et al.*, Comparative multi-omics systems analysis of Escherichia coli strains B and K-12. *Genome Biol* **13**, R37 (2012).

284. A. Conesa *et al.*, A survey of best practices for RNA-seq data analysis. *Genome Biol* **17**, 13 (2016).

285. K. C. McGrath *et al.*, Isolation and analysis of mRNA from environmental microbial communities. *J Microbiol Methods* **75**, 172-176 (2008).

286. Z. Chen *et al.*, Two featured series of rRNA-derived RNA fragments (rRFs) constitute a novel class of small RNAs. *PLoS One* **12**, e0176458 (2017).

287. M. A. Dillies *et al.*, A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Brief Bioinform* **14**, 671-683 (2013).

288. L. A. Corchete *et al.*, Systematic comparison and assessment of RNA-seq procedures for gene expression quantitative analysis. *Sci Rep* **10**, 19737 (2020).

289. J. Hor *et al.*, Grad-seq shines light on unrecognized RNA and protein complexes in the model bacterium Escherichia coli. *Nucleic Acids Res* **48**, 9301-9319 (2020).

290. A. Bar, L. Argaman, Y. Altuvia, H. Margalit, Prediction of Novel Bacterial Small RNAs From RIL-Seq RNA-RNA Interaction Data. *Front Microbiol* **12**, 635070 (2021).

291. P. P. Adams *et al.*, Regulatory roles of Escherichia coli 5' UTR and ORF-internal RNAs detected by 3' end mapping. *Elife* **10** (2021).

292. G. Storz *et al.*, Correction: Regulatory roles of Escherichia coli 5' UTR and ORF-internal RNAs detected by 3' end mapping. *Elife* **10** (2021).

293. M. Ziegler, J. Zieringer, R. Takors, Transcriptional profiling of the stringent response mutant strain E. coli SR reveals enhanced robustness to large-scale conditions. *Microb Biotechnol* **14**, 993-1010 (2021).

294. S. Brown, M. J. Fournier, The 4.5 S RNA gene of Escherichia coli is essential for cell growth. *J Mol Biol* **178**, 533-550 (1984).

295. D. Wickstrom *et al.*, Consequences of depletion of the signal recognition particle in Escherichia coli. *J Biol Chem* **286**, 4598-4609 (2011).

296. S. Q. Gu *et al.*, Conformation of 4.5S RNA in the signal recognition particle and on the 30S ribosomal subunit. *RNA* **11**, 1374-1384 (2005).

297. M. Pioletti *et al.*, Crystal structures of complexes of the small ribosomal subunit with tetracycline, edeine and IF3. *EMBO J* **20**, 1829-1839 (2001).

298. D. Elkina *et al.*, 6S RNA in Rhodobacter sphaeroides: 6S RNA and pRNA transcript levels peak in late exponential phase and gene deletion causes a high salt stress phenotype. *RNA Biol* **14**, 1627-1637 (2017).

299. T. Neusser, T. Polen, R. Geissen, R. Wagner, Depletion of the non-coding regulatory 6S RNA in E. coli causes a surprising reduction in the expression of the translation machinery. *BMC Genomics* **11**, 165 (2010).

300. D. Lalaouna, A. Eyraud, S. Chabelskaya, B. Felden, E. Masse, Regulatory RNAs involved in bacterial antibiotic resistance. *PLoS Pathog* **10**, e1004299 (2014).

301. A. E. Trotochaud, K. M. Wassarman, 6S RNA function enhances long-term cell survival. *J Bacteriol* **186**, 4978-4985 (2004).

302. J. de la Cruz, A. Vioque, Increased sensitivity to protein synthesis inhibitors in cells lacking tmRNA. *RNA* **7**, 1708-1716 (2001).

303. H. Luidalepp, M. Hallier, B. Felden, T. Tenson, tmRNA decreases the bactericidal activity of aminoglycosides and the susceptibility to inhibitors of cell wall synthesis. *RNA Biol* **2**, 70-74 (2005).

304. A. F. Brito, J. W. Pinney, Protein-Protein Interactions in Virus-Host Systems. *Front Microbiol* **8**, 1557 (2017).

305. F. Weis *et al.*, tmRNA-SmpB: a journey to the centre of the bacterial ribosome. *EMBO J* **29**, 3810-3818 (2010).

306. S. Ronneau, S. Helaine, Clarifying the Link between Toxin-Antitoxin Modules and Bacterial Persistence. *J Mol Biol* **431**, 3462-3471 (2019).

307. A. Brauner, O. Fridman, O. Gefen, N. Q. Balaban, Distinguishing between resistance, tolerance and persistence to antibiotic treatment. *Nat Rev Microbiol* **14**, 320-330 (2016).

308. J. Wen, E. M. Fozo, sRNA antitoxins: more than one way to repress a toxin. *Toxins (Basel)* **6**, 2310-2335 (2014).

309. T. Franch, A. P. Gultyaev, K. Gerdes, Programmed cell death by hok/sok of plasmid R1: processing at the hok mRNA 3'-end triggers structural rearrangements that allow translation and antisense RNA binding. *J Mol Biol* **273**, 38-51 (1997).

310. M. Kawano, Divergently overlapping cis-encoded antisense RNA regulating toxin-antitoxin systems from E. coli: hok/sok, ldr/rdl, symE/symR. *RNA Biol* **9**, 1520-1527 (2012).

311. W. M. El-Sharoud, P. L. Graumann, Cold shock proteins aid coupling of transcription and translation in bacteria. *Sci Prog* **90**, 15-27 (2007).

312. E. Rennella *et al.*, RNA binding and chaperone activity of the E. coli cold-shock protein CspA. *Nucleic Acids Res* **45**, 4255-4268 (2017).

313. P. G. Jones, M. Inouye, RbfA, a 30S ribosomal binding factor, is a cold-shock protein whose absence triggers the cold-shock response. *Mol Microbiol* **21**, 1207-1218 (1996).

314. M. de la Fuente, S. Valera, J. L. Martinez-Guitarte, ncRNAs and thermoregulation: a view in prokaryotes and eukaryotes. *FEBS Lett* **586**, 4061-4069 (2012).