ANALYSES OF RNA IN *ESCHERICHIA COLI*: FROM GENOME-WIDE RNA PROCESSING, TO THE ANALYSIS OF INDIVIDUAL OPERONS, AND THE METHODS IN BETWEEN

by

MARK B. STEAD

(Under the Direction of Sidney R. Kushner)

ABSTRACT

The maintenance of homeostasis in *Escherichia coli* is complex. One aspect of control that is often overlooked is post-transcriptional regulation, which is often mediated by ribonucleases that can either destroy or stabilize an RNA species. The work described in this thesis examines the role of ribonucleases both in the control of gene expression and in the processing of ribosomal RNA. During the course of this work, it also became apparent that the RNA extraction methods currently being used for these analyses were not adequate. Accordingly a superior, novel method of extracting RNA was devised.

We analyzed the *in vivo* roles of two endoribonucleases (RNase E and RNase III) using tiling microarrays and demonstrated that both enzymes have considerably larger roles in gene expression than previously envisioned. In fact, we found that RNase E affects ~75% of small RNAs and RNase III affects greater than ten-fold more RNAs than previously thought. In addition, over 300

potentially novel genes were identified, along with the first demonstrated evidence of RNase III cleavage within a protein coding sequence in *E. coli.*

After working on transcriptome-wide gene expression analyses, we realized that no RNA extraction method resulted in the quantitative recovery of all RNA species present in the cell. In fact, we demonstrated that with the most commonly used commercial RNA extraction kits, RNA was selectively lost based on molecule size with some kits retaining only larger species, and some only small species. To circumvent this problem, we developed a very simple, cost-effective RNA extraction method that yields quantitative recovery of nearly 100% of RNA species in the cell, regardless of size, in a little as 15 minutes.

This work has also attempted to address long standing knowledge-gaps in the processing of ribosomal RNA, as the enzymes involved in many of the processing steps remain unidentified. While this work remains unfinished, we have constructed the tools required for the analysis of these processing steps and formed the foundation of future work which will resolve the processing pathway involved in ribosomal RNA maturation in *E. coli.*

INDEX WORDS: RNase E, RNase III, RNA processing, methods, *Escherichia coli*

ANALYSES OF RNA IN *ESCHERICHIA COLI*: FROM GENOME-WIDE RNA PROCESSING, TO THE ANALYSIS OF INDIVIDUAL OPERONS, AND THE METHODS IN BETWEEN

by

MARK B. STEAD

BS, Saint Martin's College, 2005

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2011

ANALYSES OF RNA IN *ESCHERICHIA COLI*: FROM GENOME-WIDE RNA PROCESSING, TO THE ANALYSIS OF INDIVIDUAL OPERONS, AND THE METHODS IN BETWEEN

by

MARK B. STEAD

Major Professor:	Sidney R. Kushner
Committee:	Richard B. Meagher Anna C. Glasgow Karls Michael J. McEachern Claiborne V. C. Glover

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2011

DEDICATION

For my wonderful family: I am incredibly grateful for, and blessed to have the love and support of my entire family. This work is dedicated to: my wife, children, parents, and grandparents, who have all made this endeavor possible.

ACKNOWLEDGEMENTS

My sincerest thanks go to Sidney R. Kushner, who forever changed the course of my life by inviting me to work and learn for a summer doing simple experiments in *Escherichia coli*. The advice, knowledge, and humor I have received from Sidney have been priceless and greatly appreciated. It has been a pleasure to have Sidney as my graduate advisor, and I am grateful for the opportunity to have worked under his direction.

Thank you to my committee members, who have been instrumental in guiding me over the course of my graduate studies. Thank you to Bijoy Mohanty for the expert advice, training, and friendship. The members of the Kushner laboratory have been wonderful to me, and I am truly appreciative of their humor, help, and camaraderie.

Finally, and most importantly, I would like to thank my entire family for their continuous encouragement; especially my wife, Jessica, whose unwavering love and support has made it possible for me to complete this work. My wonderful children Emma and Allison, who have been an endless source of joy and have reminded me that the important parts of life are not scientific. I would have never been able to pursue this endeavor without the love of my parents, Tim and Sherrill, who have always done everything in their power to support me.

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Starting in the 1950's, before the discovery and characterization of ribonucleases, scientists have attempted to elucidate the pathways and mechanisms involved with RNA metabolism in *Escherichia coli*, especially since it seemed that the bacteria had a very high content of ribonucleic acid (1-3). With the discovery of the first ribonuclease in *Escherichia coli*, it was determined that ribonucleases serve two basic functions within the cell: to turnover mRNA and recycle the ribonucleotides, and to process stable RNAs from their primary transcripts to the mature forms (4-5). Aside from the discovery in 1967 that RNase III was involved in ribosomal RNA maturation (6-7), throughout most of the first 30 years of research in the field, ribonucleases were viewed as little more than enzymes involved in a ribonucleotide recovery pathway.

While ribonucleases are primarily involved only in RNA processing and message decay, the field has begun to understand that many aspects of gene expression are directly or indirectly regulated by ribonucleases and that mRNA turnover and RNA processing are a means to effect cellular change in response

to environmental changes. According to an array-based chemical half-life estimation study, approximately 80% of *E. coli* mRNAs have half-lives between three and eight minutes, meaning most mRNAs are highly susceptible to RNase cleavages (8). Hence, ribonucleases are responsible, at least in part, for the amount of protein resulting from a single mRNA species. While there remains a great deal of mystery in the way *E. coli* regulates RNA stability and processing, work over the last 20 years has made it clear that ribonucleases play a vital role in the regulation of homeostasis and in the cellular response to environmental stimuli.

There are two types of ribonucleases, endoribonucleases, which cleave an RNA molecule internally, and exoribonucleases, which remove one nucleotide at a time from either the 5' or 3' end of an RNA molecule. For the purposes of this review, only endoribonucleases will be discussed as the work described in this thesis does not explore the contribution of exoribonucleases in RNA processing and gene expression. There are seven known endoribonucleases in *E. coli* at this time (RNase E, RNase G, RNase III, RNase P, RNase Z/RNase BN, RNase I, RNase LS), though there is certainly the possibility of additional enzymes, including *ybeY*, which has been recently reported as a potential endoribonuclease, but as of this date lacks a thorough enzymatic characterization (9). There are also gene members of the CRISPR/CAS genome defense system in *E. coli,* but according to a number of studies the system is inactive in our genetic background, and we currently have no evidence that the

two putative endoribonucleases within the system are expressed and active in the organism (10-12).

For the purposes of this work, three endoribonucleases are of the most importance: RNase E, RNase III, and RNase P. With regard to total endoribonuclease activity, the most active enzyme in *E. coli* is RNase E. Ribonuclease E was discovered, independently, as an essential gene involved with RNA turnover and processing by two groups, first as the *ams* gene (alteration of mRNA stability), aptly named because of the stabilization of a substantial number of mRNAs caused by inactivation of the protein using a temperature-sensitive allele (13-14), and subsequently as RNase E via biochemical methods to identify genes responsible for the processing of 5S rRNA (15-16). It was soon discovered that the *ams* and *rne* genes were actually the same gene, both encoding RNase E (17).

RNase E is involved in the processing or decay of a substantial fraction of the transcriptome (at least 40% of mRNA) (18) and is involved with every class of RNA, including the processing of tRNAs (19-21), rRNAs (16,22-27), non-coding RNAs (such as the RNA subunit of RNase P, the RNA quality control transcript *ssrA* otherwise known as tmRNA, and the antisense RNA plasmid replication ncRNA, RNA I) (28-31), and sRNAs (32-33). RNase E prefers to cleave single-stranded A-U rich regions of RNA and is somewhat inhibited by the presence of a 5′ triphosphate moiety, although RNase E can also cleave internally without regard to the status of the 5′ end on some substrates (34-49). RNase E-like

enzymes have also been found in many prokaryotes and the function of the ribonuclease seems to be highly conserved (50-58).

Additionally, the RNase E protein contains an assembly scaffold at the C-terminal end of the protein for the binding of RhlB (a DEAD-box RNA helicase), PNPase (a 3′-5′ exoribonuclease), and enolase (a glycolytic enzyme), which together have been termed the degradosome (23,59-69). Other proteins have been found to be associated with the degradosome, but at substoichiometric levels, such as polyphosphate kinase and the RNA binding protein required for most sRNA-mediated target degradation, Hfq (70-74). The degradosome is necessary for RNase E-mediated decay of some mRNAs, but does not significantly impact rRNA processing (66,75-77). RNase E exists *in vivo* as a homotetramer, and the quaternary structure is dependent upon a region known as the Zn-link, which is located after the N-terminal catalytic region, and before the C-terminal degradosome scaffolding region and uses coordinating zinc ions to form the structure (78-79). There exists some evidence that RNase E and the degradosome are spatially located near the cell membrane through an N-terminal membrane anchor region on RNase E (80-82), but this result is at odds with previous results suggesting a cytoskeletal filament organization of the degradosome (83).

Interestingly, RNase E is autoregulatory and can modulate the amount of enzyme relative to the amount of RNA substrate. The RNase E enzyme cleaves the 5' UTR of the *rne* transcript (the coding gene for RNase E), and the intact degradosome is required for autoregulation (40,46,66,84-86). Additionally, the

level of RNase E expression is also controlled by the use of three distinct promoters, which are required for autoregulation and the usage of which seem to affect translation efficiency of the *rne* transcript as the RNA level and protein level can be uncoupled (39). RNase E activity is also regulated in part by protein inhibitors, such as RraA and RraB, which bind the C-terminal domain of RNase E, both altering the composition of the degradosome and modulating the activity of RNase E itself (87-93). The ribosomal protein L4 has also been shown to interact directly with RNase E, in a region separate from the catalytic region, to stabilize some RNA substrates (94). In another level of the complex regulation of RNase E activity, the discovery that the 5' triphosphate, found on all primary transcripts in *E. coli,* inhibits RNase E cleavage on some substrates led to the discovery of an RNA pyrophosphohydrolase (RppH) which removes 5' terminal pyrophosphates from RNA (leaving behind a monophosphate residue) triggering RNase E-mediated degradation (95).

While RNase E is an essential enzyme for cell viability, studies over the last 20 years have utilized a variety of RNase E mutants, usually *rne-1,* a temperature-sensitive allele of RNase E, or truncation alleles such as *rneΔ645* which lack the C-terminal degradosome scaffolding region, in order to determine the effects of the enzyme *in vivo* (13-15,66,96-98). Recently a single amino acid change in RNase G (an endoribonuclease, which is closely related to RNase E, sharing approximately 50% amino acid sequence similarity to the catalytic region of RNase E, but lacking the degradosome scaffolding region) has been found that supports cell viability in the absence of RNase E, when the altered RNase G

protein (Rng-219) is over-expressed 12-fold relative to wild-type chromosomally encoded RNase G (99). This Rng-219 protein allows a much more physiologically relevant level of endonuclease expression to obtain complementation than a previous study, which required approximately a 174-fold increase of an extended form of RNase G to achieve viability in the absence of RNase E (18). Studies of *rng-219*, both *in vivo* and *in vitro,* suggest that the modified enzyme is less active than both wild-type RNase G and RNase E, yet with cleavage specificity that is more similar to RNase E than RNase G on most substrates (99). Unfortunately, however, it is not known what the essential function of RNase E activity is, and conversely, how exactly an RNase E deletion strain is complemented by Rng-219 (97,99-103). Future studies are being planned to elucidate the transcriptome-wide activity of Rng-219 versus both RNase E and wild-type RNase G, to better determine the true, total impact of RNase E activity on gene expression.

RNase III, while not nearly as prolific in activity as RNase E, was one of the first ribonucleases discovered in *E. coli* because of the enzyme's involvement in the processing of rRNA. It remains the only indentified active endoribonuclease in *E. coli* that specifically cleaves dsRNA (7). The relatively easy discovery of RNase III was allowed by the RNA phenotype of an RNase III-deficient mutant, as upon separation on an agarose gel, the primary rRNA transcript known as 30S pre-rRNA was clearly visible (5,104-106). Since the discovery of RNase III in *E. coli,* homologs have been found in many organisms, from prokaryotes to higher eukaryotes (such as mammals), and the enzyme has

been shown to process multiple classes of RNAs (107). While in some prokaryotes, including *Bacillus subtilis,* RNase III is essential (because of the processing of an essential sRNA), several *rnc* deletion mutants have been described in *E. coli* (104,107-110). The nuclease contains an N-terminal RNase H-like domain and a C-terminal dsRNA-binding domain, which as a functional homodimer preferentially cleaves hairpin structures that are at least 11 base pairs long (111-112). RNase III domains in higher eukaryotes are of special interest in the field of RNA silencing, as RNase III activity is responsible for the processing of microRNAs (113-116).

In bacteria, RNase III is primarily known for its role in rRNA processing (117), but has also been shown to be involved in the decay of a small number of mRNAs (approximately 50 transcripts), and in sRNA processing (104,107-108,118-120). Recent studies have demonstrated that some *E. coli* sRNAs regulate the stability and translation initiation efficiency of specific mRNAs through RNase III-dependent cleavages (121-123). RNase III itself is auto-regulatory in much the same manner as RNase E, as RNase III cleaves the 5' UTR of the polycistronic *rnc*-containing transcript (119,124). Additionally, RNase III activity is also modulated by a stress-responsive inhibitory protein, YmdB (125). While much remains to be investigated with regard to the role of RNase III in *E. coli,* the work described in Chapter 2 has increased the number of potential RNase III targets at least 10-fold, and demonstrates that RNase III may also be important in the regulation of a number of biological pathways (97).

RNase P is a highly conserved endoribonuclease found in all domains of life, and is a unique example of a ribozyme endoribonuclease, composed of both a catalytic RNA (M1) and a protein scaffold (C5). It is an essential enzyme in *E. coli* (126-128). RNase P in *E. coli* is primarily known for maturation of the 5' ends of tRNAs, but the ribozyme is also active on a limited number of other substrates (19,129-132). While RNase P may have some role in the control of gene expression and gene pathways, the role of the ribozyme in message decay is limited in comparison to RNase E and RNase III.

Control of biological pathways through ribonucleases

In order for the cell to adapt to a change in environment, there are two ways to change gene expression at the RNA level. The first is to synthesize new RNAs encoding needed proteins (or inhibit their degradation), and the second is to destroy unnecessary mRNAs leaving behind only those mRNAs needed for survival (or inhibit their synthesis). Creating new mRNAs requires a much greater amount of energy and resources than does destroying mRNAs, and when mRNAs are essentially in competition with each other for ribosomes, the destruction of unneeded mRNAs allows for a greater amount of protein production from the mRNAs left over. Ribonucleases therefore allow for a fast and low energy mechanism of cellular reprogramming. Because of this attribute of ribonuclease activity, some of the toxin/antitoxin systems in *E. coli* also utilize mRNA destruction to quickly reprogram the transcriptome in response to stress, such as nutrient starvation (133-135).

The MazE/F toxin/antitoxin system, for example, features an endoribonuclease (MazF) and an unstable antitoxin (MazE) which binds MazF and prevents activity. When environmental conditions arise that slow protein synthesis, the remaining MazE is quickly degraded by proteases and the MazF endoribonuclease is released to cleave RNAs specifically at ACA triplets (136-137). Recently, some evidence has suggested that MazF also cleaves mature 16S rRNA, generating a subpopulation of ribosomes that selectively translates leaderless mRNAs, wherein MazF has removed the 5′ UTR (including the normal ribosome binding site) (22). This selective translation system may allow the bacteria to more rapidly reprogram the transcriptome in response to a particular stress condition.

Ribonuclease cleavages of mRNAs can both destroy and stabilize the message. For example, in the case of the *gadX gadW* dicistronic operon, an RNase III cleavage, which is directed by the binding of the sRNA *gadY* in the intercistronic spacer region between *gadX* and *gadW,* actually stabilizes each mRNA coding sequence and is required for efficient translation (123). Most such examples of mRNAs, which are less abundant in the absence of an endoribonuclease, involve the role of a small RNA or functional ncRNA, which in themselves are often at least partially regulated by ribonucleases. These 40 – 500 nt sRNA/ncRNA species are known to both positively and negatively regulate gene expression in *E. coli* through both inhibition of translation of an mRNA by annealing to the ribosome binding site, and/or destabilization of the mRNA mediated by an endoribonuclease. sRNA activity is usually dependent

upon Hfq, an RNA binding protein required for the function of many sRNAs, and sometimes associated with the degradosome (138-141). Alternatively, a sRNA binding its target mRNA within the 5′ UTR may cause the ribosome binding site to become more available to ribosomes through the alleviation of secondary structure to enhance the translation efficiency (32,139-140,142-150).

Many sRNAs are regulated by endoribonucleases, both through processing of a pre-sRNA to a functional form and by degradation, and therefore gene expression control exerted by sRNAs many be considered as a secondary effect of ribonuclease cleavage (151). The activity of some sRNAs is also dependent upon the recruitment of an endoribonuclease to the sRNA-bound target, where the binding of the sRNA creates an efficient RNase cleavage site (33,122,152). Prior to the work described in Chapter Two, it was known that a number of sRNAs were affected in abundance by the absence of RNase E and RNase III, but it is clear that the role of RNases in the regulation of sRNA activity was significantly underestimated (97). In fact, it seems that the majority of known sRNAs are significantly affected in abundance in the absence of RNase E, but it remains unclear for many sRNAs whether endoribonucleolytic processing is required for activity.

Transcriptome-wide gene expression analysis has greatly expanded our ability to elucidate the mechanisms of control for biological pathways, and has proven to be a powerful tool for the analysis of ribonuclease mutant strains. The ability to characterize RNA abundance changes in gene families and gene networks has allowed the field to gain a top-down view of the role of

ribonucleases on gene expression. Hence transcriptome-wide analysis of ribonuclease mutants has been performed for a number of enzymes and organisms (76,97,153-161). For example, in 2002, Lee et al. (155) found that the abundance of 40% of mRNAs were affected in the absence of RNase E, kept viable with a large increase in the amount of an altered RNase G protein, and a significant number of entire genetic pathways seemed to be under post-transcriptional control. This result was followed up by a study from the same lab using ORF expression arrays to determine the role of the RNase E degradosome (76). However, there remained a number of questions after the 2002 RNase E expression array data were published as to the *in vivo* functions of RNase E, as the microarrays employed only used three oligoribonucleotide probes per coding sequence, and did not include probes for any other regions of the genome (8). Additionally, the strain used for the microarray analysis used an extended form of RNase G, which was over-expressed approximately 174-fold versus wild-type RNase G levels, in order to complement the loss of RNase E activity, which may have inherently complicated and influenced the data generated.

To address concerns with previous analyses, and to determine the effects of RNase E activity at the transcriptome-wide level, the work described in Chapter Two exploits the utility of tiling microarrays (at the resolution of 20 nt over the entire *E. coli* genome) to analyze changes in RNA abundance in an RNase E deletion mutant, kept viable by Rng-219 [the advantages of which were described previously (99)], versus wild-type *E. coli* (97). The results showed that RNase E has a dramatic effect on the abundance of the majority of annotated

sRNAs in *E. coli,* with 75% of sRNAs being affected either directly or indirectly by the absence of RNase E.

In addition to the transcriptome analysis of RNase E, duplicate experiments were also performed with RNA from an RNase III-deficient mutant versus wild-type, which represents the first published report of the effects of RNase III on the *E. coli* transcriptome (97). The RNase III results were surprising, as the number of mRNAs which were affected by RNase III was 10-fold more than previously published. In addition, nearly 12% of annotated sRNAs were also affected by the absence of RNase III. These results suggest that RNase III has a much more wide-spread role in the regulation of gene expression than previously envisioned. Using gene pathway analysis, we demonstrated that seven biological pathways were significantly affected in the absence of RNase III. Interestingly, the data also suggested a significant amount of overlap in activity between RNase E and RNase III, with 10% of coding sequences and 12% of ncRNAs being affected in the absence of either enzyme. The arrays were also used in the identification of potentially novel genes. As the vast majority of annotated ncRNAs are affected in the RNase E mutant, it stands to reason that at least some non-annotated ncRNAs are also affected. Genomic loci of unexplained changes in RNA abundance (which did not map to a known genome feature, without regard to the location of known ncRNAs) in the absence of RNase E versus wild-type were compiled, for a total of 402 loci, which included 39 annotated ncRNAs and one putative small open reading frame. While not all of the 362 loci of unexplained changes of RNA abundance will turn out to be

novel genes upon validation, even the discovery of one additional sRNA or small ORF could turn out to be a major discovery in the field.

The results described in Chapter Two draw attention to the significant complexity and impact of post-transcriptional regulation on gene expression in *E. coli*, and have highlighted the importance of ribonucleases in cellular homeostasis and environmental response. In addition, because of the significant changes in abundance of ncRNAs in the absence of RNase E or RNase III, it has become apparent that secondary effects of ribonuclease activity (such as those mediated by sRNAs) are likely major players in gene regulation, making the analysis of ribonuclease mutants much more complex.

RNA extraction methods

At the very core of molecular biology and specifically the study of RNA metabolism, is one of the most fundamental and overlooked steps, RNA isolation. Most scientists, who regularly work with RNA, understand that no method of RNA extraction is perfect, and in fact, far from it. In the early days of RNA biology, RNA was relatively hard to obtain, particularly RNA free of degradation, since RNA is significantly more susceptible to degradation than DNA. Early RNA extraction methods relied upon guanidium isothiocyanate to lyse cells and denature proteins, while the RNA was captured using a Cesium chloride cushion and ultracentrifugation (162). The RNA extraction method of choice up until the 1980s, which itself was a huge improvement in ease of use

versus Cesium gradients, involved using phenol, which was heated to lyse cells and inhibit RNases (163-164).

RNA extractions using hot phenol had significant problems both due to the toxicity of the phenol, and because the RNA obtained was not consistently of high-quality. Therefore, in the early to mid-1980's a protocol was developed to include guanidium isothiocyanate with phenol, which gave much more reproducible results when compared with previous methods (165-166). In the 1990s, several companies developed RNA isolation kits based upon the use of phenol, and later upon the use of silica/glass columns to bind the RNA, which led to higher quality RNA preparations and facilitated the use of RNA in more laboratories. The kits utilized several types of methods to lyse cells, remove proteins and DNA, and capture RNA, but certain RNA molecules, either due to size, secondary structure, or tightly-bound proteins, were not efficiently recovered with each method (see Chapter 3).

RNA isolation methods vary considerably with respect to two major aspects: cell lysis, and protein/ DNA removal. Cell lysis steps can either be mechanical, such as zirconium bead homogenization (often used with gram-positive bacteria or yeast), or chemical, such as heating cells in guanidine and detergents, or even a combination of the two. Additionally, the cell wall of some organisms—especially gram-positive bacteria—may require an enzymatic digestion step with lysozyme and/or proteinase K in order to achieve sufficient cell lysis and release of nucleic acid. The general problem faced with cell lysis for RNA extraction, is that RNA is highly unstable, so any and all ribonucleases

need to be very quickly inactivated upon exposure to the lysis solution. To accomplish denaturation of proteins including RNases, methods often rely on guanidium isothiocyanate and 2-mercaptoethanol, which in combination act to quickly inactivate ribonucleases and stabilize RNA.

Cell lysis steps may also include the use of sodium dodecyl sulfate (SDS) or other detergents to aid in cell disruption and protein inactivation. The use of detergents in cell lysis also led to the discovery of a cationic surfactant in the 1990's that is able both to aid in cells lysis and to bind RNA and DNA, allowing nucleic acid precipitation (167-169). DNA in the pellet is then subsequently removed by washing with LiCl, which takes the place of the detergent in interacting with RNA, but does not effectively precipitate DNA (167). The detergent-based RNA extraction procedure had the major advantage versus phenol-based extractions of being relatively non-toxic, and offered good yields of high-quality RNA. The trade name of the original detergent used for RNA extractions was Catrimox-14® (now owned by Qiagen), but its manufacture was soon discontinued. A replacement detergent, trimethyl(tetradecyl)ammonium bromide, was subsequently identified that had similar properties when compared with Catrimox-14®, which is now known as Catrimide (170).

In most RNA extraction methods, DNA and proteins are removed by selective precipitations (using LiCl or isopropanol precipitation methods) with RNA in water, or ethanol/low salt washes with silica columns; but selective removal of DNA and proteins often comes at a cost, as not all RNA molecules precipitate quantitatively or bind silica with the same efficiency. Depending upon

the method of selective RNA precipitation, some RNA isolation methods efficiently and quantitatively isolate large RNAs (greater than 200 nt) while effectively removing small RNAs from the RNA pool, while other methods have the opposite effect. Therefore, any method selected to extract RNA for a particular project could have large undesirable consequences on down-stream applications and data analysis. For example, if an RNA pool was practically devoid of RNA species less than 200 nt, the sample would not be suitable for the analysis of tRNAs and sRNAs.

While using a kit tailored for the needs of the experiment may appear to make sense, because of confusion in the field about the capabilities and limitations of RNA extraction kits, and the wide-spread use of transcriptome-wide gene expression analysis, many experiments have been conducted employing inappropriate RNA extraction methods. For some experiments, such as RNAseq which can analyze RNA of all sizes, no method exists that allows quantitative recovery of all RNA species without regard to RNA size or structure. With these problems in mind, the work described in Chapter Three first sought to determine quantitatively the recovery of a variety of specific RNA species using a number of commercial RNA extraction kits and the catrimide/LiCl method developed in the Kushner laboratory (170).

After some initial work, it was soon obvious that the true *in vivo* RNA profile of *E. coli* was unknown, as the method used for the RNA extractions upon which the RNA profile was determined was guanidine/CsCl gradients, a method that does not effectively capture small RNAs (1-3,171). Without knowing how

much RNA should be in a cell, and what the relative fractions should be of small species to large species, it was decided to attempt to make crude lysates of *E. coli,* in which all RNA species could be captured. The hope was that the crude lysate could be run directly on a gel, and the ratios of 23S, 16S, and 5S rRNAs could be compared with the ratios from commercial RNA isolation kits in order to determine if RNA was being quantitatively recovered. After unsuccessfully trying a variety of complicated methods to make an RNA-stabilized lysate, a rather unorthodox idea was attempted.

Cells from *E. coli* were suspended in formamide-based RNA extraction solution, heated for five minutes, and centrifuged. Subsequently the supernatant was run on a gel. The results were very surprising, as the procedure resulted in a relatively pure RNA sample. After a considerable amount of optimization, a fast, reproducible, and quantitative method for the isolation of high-quality RNA from *E. coli* was developed. This method, described in Chapter Three, is ideally suited for the analysis of RNA species of any size, and seems to be the only RNA extraction method to yield quantitative recovery of both RNAs as large as 5.7 kb and as small as 76 nt. Furthermore, the method recovers >99% of the RNA, takes 15 minutes to perform, and costs as little as three cents per sample. Furthermore, the RNA can be directly used for Northern analysis without any additional clean-up steps.

Ribosomal RNA processing

Ribosomal RNA in *E. coli* is transcribed as a large 30S precursor from seven polycistronic operons (*rrnA, rrnB, rrnC, rrnD, rrnE, rrnG, rrnH)*, each of which contains 16S, 23S, and 5S rRNAs and at least one tRNA; and each of the seven operons contribute equally to the rRNA pool (Fig. 1) (172-174). The 30S precursor must be initially cleaved by ribonucleases to generate the functional rRNAs required for ribosome assembly and translation. In exponentially growing wild-type cells, where rRNA transcription initiation rates are highest (172), the 30S rRNA precursor cannot be detected by Northern blotting (Stead and Kushner, unpublished data), which suggests 30S rRNA processing occurs very rapidly and efficiently. In wild-type cells, RNase III performs the initial four RNA cleavages of the 30S rRNA precursor at two physical locations (as the sequences 5' and 3' of both 23S and 16S rRNAs form dsRNA hairpin substrates suitable for RNase III cleavage), resulting in the separation of each of the pre-rRNA species from each other, generating 17S (containing 16S), 25S (containing 23S), and 9S (containing 5S rRNA) rRNA precursors (175-176). The RNase III cleavages take place 115 nt 5' to the mature 5' end of 16S rRNA, 33 nt 3' to the 3' mature end of 16S rRNA, 3 to 7 nt 5' to the mature 5' end of 23S (although data in Chapter Four demonstrates that RNase III may actually cleave either at the mature 5' end, or 1, 4, 5, 7, or 8 nt from the 5' end of 23S, and this pattern is not specific to a particular rRNA operon), and 7 to 9 nt 3' to the mature 3' end of 23S rRNA (24,105,163,177-178).

RNase III cleavage of the 30S rRNA transcript likely takes place after the binding of ribosomal proteins, as the ribosomal subunits confer RNase III specificity for 23S processing *in vitro,* and RNase III has been found to functionally interact with both 50S and 30S ribosomal subunits and more strongly with the 70S ribosome (179). This functional interaction between ribosomal proteins and RNase III is further supported by *in vitro* data suggesting 23S rRNA maturation is completed either in polysomes or in the 70S ribosome (177). Thus, it is important to remember that rRNA processing is closely linked to the transcription of 30S and to the immediate binding of ribosomal proteins, which likely initiate formation of 50S and 30S subunits before rRNA maturation even begins (177,179-180).

Following RNase III cleavage, 16S rRNA precursors are further processed at the 5' end by both RNase E (66 nt 5' of the mature 5' end of 16S) and RNase G (at the mature 5' end of 16S) (Fig. 2) (181). RNase E also cleaves the 9S rRNA precursor to release p5S with three nucleotides on each side of the mature sequence (27,182-183). The mature 3′ termini of 23S and 5S rRNAs are generated primarily by RNase T, a 3'-5' exoribonuclease, although other exoribonucleases may also be slightly involved (181,184-187). tRNAs within the 30S rRNA precursor are most likely processed by a combination of RNase P at the 5′ ends and various 3′ to 5′ exonucleases at the 3′ ends, but the processing of tRNAs within rRNA operons has not been specifically addressed (188).

It is not known which enzymes are responsible for the endonucleolytic cleavages resulting in the mature 5′ ends of 23S and 5S rRNAs, or the 3' end of

16S rRNA (181,183,189-190). RNase E has been shown to be involved with the 5′ maturation of 23S rRNA in α-*Proteobacteria* (*E. coli* is a γ-*Proteobacteria*) following RNase III cleavage, but experiments using the temperature-sensitive *rne-1* allele of RNase E have shown no effect on the 5' end of 23S in *E. coli* (see Chapter 4) (57). Furthermore, a group recently published that RNase G is responsible for the 5' maturation of 23S rRNA, but the experimental evidence lacks clarity, and in our hands, using a strain without RNase G, no build up of 23S 5' precursor products was observed upon primer extension analysis (See Chapter 4) (191).

Other factors are also involved with the maturation of rRNA in *E. coli,* but their roles in the pathway are much less understood than those of RNase III, RNase E, and RNase G. DEAD-box RNA helicases have been shown to be involved in the maturation of rRNAs, and most likely aid in the unwinding of secondary structure to allow single-stranded RNA-specific ribonucleases, such as RNase E, to process the rRNA (192). There are five DEAD-box helicases in *E. coli* (*deaD, srmB, rhlB, rhlE,* and *dbpA*); but two helicases, *deaD* and *srmB,* seem to have the most effect on ribosome formation, with *srmB* being implicated in 50S ribosome formation, and with total defects being especially pronounced in a *deaD srmB* double mutant (192-193). It has also been reported that *dbpA* binds to mature sequences of 23S rRNA in a number of locations, and while the exact role of interaction remains unclear, it is most likely involved with the unwinding of 23S helices for ribosome formation (194-195).

Another recent find was the involvement of a highly conserved gene, *ybeY,* in rRNA processing (9). *E. coli* strains deficient in YbeY have severe rRNA processing defects, especially with 16S processing at both the 5' and 3' ends in the MC4100 genetic background (9). Additionally, YbeY affects the processing of the 5' and 3' ends of both 23S and 5S rRNAs in MC4100 (a laboratory wild-type strain of *E. coli*), but the defect is not as dramatic as that of in 16S processing. It is not yet clear what role *ybeY* has in the processing of rRNA, or whether the effect is a direct interaction of the YbeY protein on the rRNA precursors, or whether the effect is caused by an unknown intermediary (9). Additionally, rRNA precursors accumulate in cells lacking one or more tRNA genes (190). It is currently unknown why tRNA levels affect rRNA maturation. Ribosomal RNA processing is also affected in temperature-sensitive mutants of RNase P at the non-permissive temperature, which is most likely due either to the processing role of RNase P on the 5′ end maturation in tRNAs or to a direct role in rRNA processing (190).

The cleavage of the 30S rRNA precursor is one of the best known functions of RNase III *in vivo,* and when steady-state RNA is examined from exponentially growing *E. coli* RNase III deficient strains, the 30S precursor can be visually identified on agarose gels (104). However, the amount of detectable 30S rRNA precursor in *rnc* mutants accounts for only a small fraction of the total processed rRNAs in the cell, and upon northern analysis with a probe to mature 16S rRNA, the signal from 30S accounts for less than 2% of the total signal (Stead and Kushner, unpublished results). This indicates that other

ribonucleases are able to process the 30S rRNA precursors very efficiently, even such that there is only a small impact on growth rate under standard laboratory conditions with *rnc*- strains (Stead and Kushner, unpublished data). In addition to a stabilization of the 30S precursor in an RNase III-deficient strain, several other stabilized rRNA precursors are easily visible after Northern blotting, indicating that the overall pathway of rRNA maturation may be altered in the absence of RNase III (Stead and Kushner, unpublished data). Interestingly, in some Enterobacteriaceae genera, such as Salmonella but not Escherichia, RNase III is also responsible for the excision of an intervening sequence within 23S rRNAs, which results in functional fragmented 23S rRNA species which are not re-ligated *in vivo* (196-199). While the enzyme responsible for rRNA fragmentation is not known for each organism and the physiological implications of this phenomenon are not clear, rRNA fragmentation is a common occurrence in many bacterial species and even eukaryotes (200).

In a 1984 paper, 23S rRNA was subjected to S1 nuclease protection assays to identify the 5′ and 3′ cleavages sites (201). Three unique cleavage sites near the 5′ end and one near the 3′ end of mature 23S rRNA were observed in an RNase III-deficient background (201). Since S1 nuclease protection assays are relatively primitive in comparison to more recent techniques such as primer extension analysis, all of these sites (with the exception of the site near the 3' end which was not tested) were more or less verified using primer extensions (Stead and Kushner, unpublished results). Furthermore, no mature 23S rRNA can be detected in RNase III mutant strains upon primer extension analysis, in

confirmation of previous studies. Instead all 23S rRNA species in RNase III mutant strains are either truncated at the 5' end (the primary truncated species being 1 and 3 nt shorter than the primary 5' end in wild-type *E. coli*) or significantly longer than the mature 5' end of 23S in wild-type (Stead and Kushner, unpublished results) and yet are still functional in ribosomes (178). 16S and 5S rRNAs also still function in ribosomes if incomplete processing occurs, indicating at least some ability of ribosomes in *E. coli* to tolerate aberrant processing for each rRNA species (181,183).

Processing of 17S pre-rRNA in the absence of RNase III was first analysed before the discovery of the roles of RNase E and RNase G on the 5′ end of p16S rRNA (106,202). Since RNase E and RNase G were shown to cleave in the presence of RNase III at two specific sites in p16S rRNA, and some of the cleavage sites shown in the absence of RNase III are in the same location, it can be reasonably assumed that RNase E and RNase G are able to produce mature 5′ ends of 16S rRNA without RNase III (106,181,202). The 3′ end of 16S rRNA is also processed in the absence of RNase III, but the responsible endoribonuclease(s) has not been identified (202). 5S rRNA processing is not affected in the absence of RNase III, but the method by which 9S rRNA is separated from the 30S precursor is not clear (106).

In summary, the wild-type 5' mature end of 23S rRNA is the only terminus that cannot be formed in the absence of RNase III, and rRNA processing without RNase III remains extremely efficient. The question now becomes, by what pathway does rRNA maturation take place in the absence of RNase III, and more

importantly, does that pathway also take place in wild-type cells as if it were in competition with an RNase III mediated pathway? In order to address this question a series of *E. coli* strains, lacking a variety of enzymes thought to be involved in rRNA maturation, were generated in a genetic background obtained from Cathy Squires laboratory (Stanford University) in which all chromosomally encoded rRNA operons were cleanly deleted and the strain was kept viable by plasmids encoding both a rRNA operon and the tRNAs originally contained within the chromosomally encoded ribosomal RNA operons. These strains were subjected (along with a series of other mutant strains constructed in a wild-type background) to primer extension analysis in order to determine which endoribonucleases are involved in the separation of 16S and 23S pre-rRNAs, and in the 5' maturation of 23S. While this story is not yet complete, the results obtained rule out a number of enzymes and will help guide this project to fruition.

Summary

The aim of my dissertation research—while the work can be easily separated into three distinct fields—has been to better understand the role of posttranscriptional regulation in *Escherichia coli.* To that end, in Chapter Two my research explores the *in vivo* roles of two endoribonucleases, RNase E and RNase III, on a transcriptome-wide scale. The data suggested that the role of posttranscriptional regulation of gene expression had been vastly underestimated, and that both endoribonucleases contribute both directly and indirectly to the regulation of a significant portion of the genome. In addition, the work in Chapter Two demonstrated that *E. coli* likely has many more

undiscovered genes (potentially more than 300), and highlighted the first evidence of an RNase III cleavage within a coding sequence of an mRNA in *E. coli.*

In Chapter Three, we addressed a methods problem which has plagued the world of RNA research since the discovery of RNA itself: how can high-quality RNA be easily and quantitatively isolated? After testing the RNA recovery of specific transcripts from RNA pools derived from a number of widely used RNA extraction methods, we determined that current methods did not yield quantitative recovery of all sized RNA species, and were inadequate for modern RNA analysis methods. We therefore sought to determine what the actual *in vivo* RNA pool consisted of, for the sake of comparison, and ended up inventing a novel method for the isolation of RNA. This method, which works in every bacterial species tested (both gram-positive and gram-negative), seems to be the only RNA extraction method to yield quantitative recovery of RNA without regard to the size of the RNA species. In addition, this method is easy enough for someone with very little laboratory experience to successfully perform, costs less than three cents a prep, and takes less than 15 minutes in most cases.

The research described in Chapter Four, while not complete, represents a solid beginning in the quest to determine the pathway responsible for both the maturation of the 5' end of 23S rRNA, and the separation of 16S and 23S rRNA precursors in *E. coli.* The enzyme responsible for the separation of precursor rRNAs in wild-type *E. coli* (RNase III) is surprisingly dispensable for maturation, and therefore an extremely efficient back-up pathway remains to be

characterized. Using a series of mutant strains and primer extension analyses, a large set of enzymes has been ruled out of involvement in rRNA processing, which will help to guide future studies towards the enzyme(s) responsible for rRNA maturation.

REFERENCES

1. Wade, H.E. (1955) Basophilia and high ribonucleic acid content of dividing *E. coli* cells. *Nature*, **176**, 310.
2. Manson, L.A. (1953) The metabolism of ribonucleic acid in normal and bacteriophage infected *Escherichia coli*. *J Bacteriol*, **66**, 703-711.
3. Borek, E., Rockenbach, J. and Ryan, A. (1956) Studies on a mutant of *Escherichia coli* with unbalanced ribonucleic acid synthesis. *J Bacteriol*, **71**, 318-323.
4. Singer, M.F. and Tolbert, G. (1965) Purification and properties of a potassium-activated phosphodiesterase (RNase II) from *Escherichia coli*. *Biochemistry*, **4**, 1319-1330.
5. Kindler, P., Keil, T.V. and Hofschneider, P.H. (1973) Isolation and characterization of an RNase III deficient mutant of *Escherichia coli*. *Mol. Gen. Genet.*, **126**, 53-69.
6. Robertson, H.D., Webster, R.E. and Zinder, N.D. (1967) A nuclease specific for double-stranded RNA. *Virology*, **12**, 718-719.
7. Robertson, H.D., Webster, R.E. and Zinder, N.D. (1968) Purification and properties of ribonuclease III from *Escherichia coli*. *J. Biol. Chem.*, **243**, 82-91.

8. Bernstein, J.A., Khodursky, A.B., Lin, P.-H., Lin-Chao, S. and Cohen, S.N. (2002) Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 9697-9702.

9. Davies, B.W., Kohrer, C., Jacob, A.I., Simmons, L.A., Zhu, J., Aleman, L.M., Rajbhandary, U.L. and Walker, G.C. (2010) Role of *Escherichia coli* YbeY, a highly conserved protein, in rRNA processing. *Mol Microbiol*, **78**, 506-518.

10. Pougach, K., Semenova, E., Bogdanova, E., Datsenko, K.A., Djordjevic, M., Wanner, B.L. and Severinov, K. (2010) Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. *Mol Microbiol*, **77**, 1367-1379.

11. Pul, U., Wurm, R., Arslan, Z., Geissen, R., Hofmann, N. and Wagner, R. (2010) Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS. *Mol Microbiol*, **75**, 1495-1512.

12. Westra, E.R., Pul, U., Heidrich, N., Jore, M.M., Lundgren, M., Stratmann, T., Wurm, R., Raine, A., Mescher, M., Van Heereveld, L. *et al.* (2010) H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. *Mol Microbiol*, **77**, 1380-1393.

13. Kuwano, M., Ono, M., Endo, H., Hori, K., Nakamura, K., Hirota, Y. and Ohnishi, Y. (1977) Gene affecting longevity of messenger RNA: a mutant

of *Escherichia coli* with altered mRNA stability. *Mol. Gen. Genet.*, **154**, 279-285.

14. Ono, M. and Kuwano, M. (1979) A conditional lethal mutation in an *Escherichia coli* strain with a longer chemical lifetime of mRNA. *Journal of Molecular Biology*, **129**, 343-357.

15. Misra, T.K. and Apirion, D. (1979) RNase E, an RNA processing enzyme from *Escherichia coli*. *Journal of Biological Chemistry*, **254**, 11154-11159.

16. Ghora, B.K. and Apirion, D. (1978) Structural analysis and *in vitro* processing to p5 rRNA of a 9S RNA molecule isolated from an *rne* mutant of *E. coli*. *Cell*, **15**, 1055-1066.

17. Babitzke, P. and Kushner, S.R. (1991) The Ams (altered mRNA stability) protein and ribonuclease E are encoded by the same structural gene of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, **88**, 1-5.

18. Lee, K., Bernstein, J.A. and Cohen, S.N. (2002) RNase G complementation of *rne* null mutation identified functional interrelationships with RNase E in *Escherichia coli*. *Molecular Microbiology*, **43**, 1445-1456.

19. Mohanty, B.K. and Kushner, S.R. (2008) Rho-independent transcription terminators inhibit RNase P processing of the *secG leuU* and *metT* tRNA polycistronic transcripts in *Escherichia coli*. *Nucleic Acids Research*, **36**, 364-375.

20. Ow, M.C. and Kushner, S.R. (2002) Initiation of tRNA maturation by RNase E is essential for cell viability in *Escherichia coli*. *Genes and Development*, **16**, 1102-1115.

21. Li, Z. and Deutscher, M.P. (2002) RNase E plays an essential role in the maturation of *Escherichia coli* tRNA precursors. *RNA*, **8**, 97-109.

22. Vesper, O., Amitai, S., Belitsky, M., Byrgazov, K., Kaberdina, A.C., Engelberg-Kulka, H. and Moll, I. (2011) Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in *Escherichia coli*. *Cell*, **147**, 147-157.

23. Lopez, P.J., Marchand, I., Joyce, S.A. and Dreyfus, M. (1999) The C-terminal half of RNase E, which organizes the *Escherichia coli* degradosome, participates in mRNA degradation but not rRNA processing *in vivo*. *Molecular Microbiology*, **33**, 188-199.

24. Li, Z., Pandit, S. and Deutscher, M.P. (1999) RNase G (CafA protein) and RNase E are both required for the 5' maturation of 16S ribosomal RNA. *EMBO J*, **18**, 2878-2885.

25. Srivastava, A.K. and Schlessinger, D. (1989) Processing pathway of *Escherichia coli* 16S precursor rRNA. *Nucleic Acids Research*, **17**, 1649-1663.

26. Roy, M.K., Singh, B., Ray, B.K. and Apirion, D. (1983) Maturation of 5S rRNA: Ribonuclease E cleavages and their dependence on precursor sequences. *European Journal of Biochemistry*, **131**, 119-127.

27. Ghora, B.K. and Apirion, D. (1979) Identification of a novel RNA molecule in a new RNA processing mutant of *Escherichia coli* which contains 5 S rRNA sequences. *J. Biol. Chem.*, **254**, 1951-1956.

28. Sim, S., Semi, K. and Lee, Y. (2001) Role of the sequence of the rne-dependent site in 3' processing of M1 RNA, the catalytic component of *Escherichia coli* RNase P. *FEBS letters*, **505**, 291-295.

29. Lin-Chao, S., Wei, C.-L. and Lin, Y.-T. (1999) RNase E is required for the maturation of *ssrA* and normal *ssrA* RNA peptide-tagging activity. *Proc. Natl. Acad. Sci. USA*, **96**, 12406-12411.

30. Xu, F., Lin-Chao, S. and Cohen, S.N. (1993) The *Escherichia coli pcnB* gene promotes adenylylation of antisense RNAI of ColE1-type plasmids *in vivo* and degradation of RNAI decay intermediates. *Proceedings of the National Academy of Sciences of the United States of America*, **90**, 6756-6760.

31. He, L., Soderbom, F., Wagner, E.G., Binnie, U., Binns, N. and Masters, M. (1993) PcnB is required for the rapid degradation of RNAI, the antisense RNA that controls the copy number of ColE1-related plasmids. *Molecular Microbiology*, **9**, 1131-1142.

32. Urban, J.H. and Vogel, J. (2007) Translational control and target recognition by *Escherichia coli* small RNAs in vivo. *Nucleic Acids Res*, **35**, 1018-1037.

33. Masse, R. and Gottesman, S. (2002) A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*.

Proceedings of the National Academy of Sciences of the United States of America, **99**, 4620-4625.

34. Kime, L., Jourdan, S.S., Stead, J.A., Hidalgo-Sastre, A. and McDowall, K.J. (2010) Rapid cleavage of RNA by RNase E in the absence of 5' monophosphate stimulation. *Mol Microbiol*, **76**, 590-604.

35. Lin-Chao, S., Wong, T.-T., McDowall, K.J. and Cohen, S.N. (1994) Effects of nucleotide sequence on the specificity of *rne*-dependent and RNase E-mediated cleavages of RNA I encoded by the pBR322 plasmid. *J. Biol. Chem.*, **269**, 10797-10803.

36. Yang, J., Jain, C. and Schesser, K. (2008) RNase E regulates the *Yersinia* Type 3 secretian system. *J. Bacteriol.*, **190**, 3774-3778.

37. Baker, K.E. and Mackie, G.A. (2003) Ectopic RNase E sites promote bypass of 5'-end-dependent mRNA decay in *Escherichia coli*. *Molecular Microbiol.*, **47**, 75-88.

38. Liu, Q., Ow, M.C. and Kushner, S.R. (2002) Analysis of RNase E cleavage specificity within the 5' UTR of the *rne* transcript of *Escherichia coli* K12. *Molec. Micro.*, in preparation.

39. Ow, M.C., Liu, Q., Mohanty, B.K., Andrew, M.E., Maples, V.F. and Kushner, S.R. (2002) RNase E levels in *Escherichia coli* are controlled by a complex regulatory system that involves transcription of the *rne* gene from three promoters. *Molecular Microbiol.*, **43**, 159-171.

40. Jiang, X., Diwa, A. and Belasco, J.G. (2000) Regions of RNase E important for 5'-end-dependent RNA cleavage and autoregulated synthesis. *Journal of Bacteriology*, **182**, 2468-2475.

41. Joyce, S.A. and Dreyfus, M. (1998) In the absence of translation, RNase E can bypass 5' mRNA stabilizers in *Escherichia coli*. *J. Molecular Biology*, **282**, 241-254.

42. Kaberdin, V.R., Chao, Y.H. and Lin-Chao, S. (1996) RNase E cleaves at multiple sites in bubble regions of RNA I stem loops yielding products that dissociate differentially from the enzyme. *J. Biological Chemistry*, **271**, 13103-13109.

43. McDowall, K.J., Kaberdin, V.R., Wu, S.-W., Cohen, S.N. and Lin-Chao, S. (1995) Site-specific RNase E cleavage of oligonucleotides and inhibition by stem-loops. *Nature*, **374**, 287-290.

44. Taraseviciene, L., Bjork, G.R. and Uhlin, B.E. (1995) Evidence for a RNA binding region in the *Escherichia coli* processing endoribonuclease RNase E. *J. Biol. Chem.*, **270**, 26391-26398.

45. McDowall, K.J., Lin-Chao, S. and Cohen, S.N. (1994) A + U content rather than a particular nucleotide order determines the specificity of RNase E cleavage. *J. Biol. Chem.*, **269**, 10790-10796.

46. Mudd, E.A. and Higgins, C.F. (1993) *Escherichia coli* endoribonuclease RNase E: autoregulation of expression and site-specific cleavage of mRNA. *Molecular Microbiol.*, **3**, 557-568.

47. Ehretsmann, C.P., Carpousis, A.J. and Krisch, H.M. (1992) Specificity of *Escherichia coli* endoribonuclease RNase E: *In vivo* and *in vitro* analysis of mutants in a bacteriophage T4 mRNA processing site. *Genes and Development*, **6**, 149-159.

48. Bouvet, P. and Belasco, J.G. (1992) Control of RNase E-mediated RNA degradation by 5' terminal base pairing in *E. coli*. *Nature*, **360**, 488-491.

49. Callaghan, A.J., Marcaida, M.J., Stead, J.A., McDowall, K.J., Scott, W.G. and Luisi, B.F. (2005) Structure of *Escherichia coli* RNase E catalytic domain and implications for RNA turnover. *Nature*, **437**, 1187-1191.

50. Lee, M., Yeom, J.H., Jeon, C.O. and Lee, K. (2011) Studies on a Vibrio vulnificus functional ortholog of *Escherichia coli* RNase E imply a conserved function of RNase E-like enzymes in bacteria. *Curr Microbiol*, **62**, 861-865.

51. Lehnik-Habrink, M., Newman, J., Rothe, F.M., Solovyova, A.S., Rodrigues, C., Herzberg, C., Commichau, F.M., Lewis, R.J. and Stulke, J. (2011) RNase Y in *Bacillus subtilis*: a Natively Disordered Protein That Is the Functional Equivalent of RNase E from *Escherichia coli*. *J Bacteriol*, **193**, 5431-5441.

52. Shin, E., Go, H., Yeom, J.H., Won, M., Bae, J., Han, S.H., Han, K., Lee, Y., Ha, N.C., Moore, C.J. *et al.* (2008) Identification of amino acid residues in the catalytic domain of RNase E essential for survival of *Escherichia coli*: functional analysis of DNase I subdomain. *Genetics*, **179**, 1871-1879.

53. Zeller, M.E., Csanadi, A., Miczak, A., Rose, T., Bizebard, T. and Kaberdin, V.R. (2007) Quaternary structure and biochemical properties of mycobacterial RNase E/G. *The Biochemical journal*, **403**, 207-215.

54. Kaberdin, V.R. and Bizebard, T. (2005) Characterization of *Aquifex aeolicus* RNase E/G. *Biochem. Biophys. Res. Comm.*, **327**, 382-392.

55. Even, S., Pellegrini, O., Zig, L., Labas, V., Vinh, J., Brechemmier-Baey, D. and Putzer, H. (2005) Ribonucleases J1 and J2: two novel endoribonucleases in *B.subtilis* with functional homology to *E.coli* RNase E. *Nucleic Acids Res*, **33**, 2141-2152.

56. Inagawa, T., Okamoto, S., Wachi, M. and Ochi, K. (2003) RNase ES of *Streptomyces coelicolor* A3(2) can complement the *rne* and *rng* mutations in *Escherichia coli*. *Biosci. Biotechn. and Biochem.*, **67**, 1767-1771.

57. Klein, F. and Evguenieva-Hackenberg, E. (2002) RNase E is involved in 5'-end 23S rRNA processing in *alpha-Proteobacteria*. *Biochem. Biophys. Res. Commun.*, **299**, 780-786.

58. Heck, C., Evguenieva-Hackenberg, E., Balzer, A. and Klug, G. (1999) RNase E enzymes from *Rhodobacter capsulatus* and *Escherichia coli* differ in context- and sequence-dependent *in vivo* cleavage within the polycistronic *puf* mRNA. *J Bacteriol*, **181**, 7621-7625.

59. Regonesi, M.E., Del Favero, M., Basilico, F., Briani, F., Benazzi, L., Tortora, P., Mauri, P. and Deho, G. (2006) Analysis of the *Escherichia coli* RNA degradosome composition by a proteomic approach. *Biochimie*, **88**, 151-161.

60. Chandran, V. and Luisi, B.F. (2006) Recognition of enolase in the *Escherichia coli* RNA degradosome. *J. Molecular Biology*, **358**, 8-15.

61. Callaghan, A.J., Aurikko, J.P., Ilag, L.L., Grossman, J.G., Chandran, V., Kuhnel, K., Poljak, L., Carpousis, A.J., Robinson, C.V., Symmons, M.F. *et al.* (2004) Studies on the RNA degradosome-organizing domain of the *Escherichia coli* ribonuclease RNase E. *J. Molecular Biology*, **340**, 965-979.

62. Liou, G.-G., Chang, H.-Y., Lin, C.-S. and Lin-Chao, S. (2002) DEAD box RhlB RNA helicase physically associates with exoribonuclease PNPase to degrade double-stranded RNA independent of the degradosome-assembling region of RNase E. *J. Biological Chemistry*, **277**, 41157-41162.

63. Carpousis, A.J. (2002) The *Escherichia coli* degradosome: Structure, function and relationship to other ribonucleolytic multienzyme complexes. *Biochemical Society Transactions*, **30**, 150-155.

64. Liou, G.-G., Jane, W.-N., Cohen, S.N., Lin, N.-S. and Lin-Chao, S. (2001) RNA degradosomes exist *in vivo* in *Escherichia coli* as multicomponent complexes associated with the cytoplasmic membrane via the N-terminal region of ribonuclease E. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 63-68.

65. Kuhnel, K. and Luisi, B.F. (2001) Crystal structure of the *Escherichia coli* RNA degradosome component enolase. *J. Molecular Biology*, **313**, 583-592.

66. Ow, M.C., Liu, Q. and Kushner, S.R. (2000) Analysis of mRNA decay and rRNA processing in *Escherichia coli* in the absence of RNase E-based degradosome assembly. *Molecular Microbiol.*, **38**, 854-866.

67. Coburn, G.A., Miao, X., Briant, D.J. and Mackie, G.A. (1999) Reconstitution of a minimal RNA degradosome demonstrates functional coordination between a 3' exonuclease and a DEAD-box RNA helicase. *Genes and Development*, **13**, 2594-2603.

68. Vanzo, N.F., Li, Y.S., Py, B., Blum, E., Higgins, C.F., Raynal, L.C., Krisch, H.M. and Carpousis, A.J. (1998) Ribonuclease E organizes the protein interactions in the *Escherichia coli* RNA degradosome. *Genes and Development*, **12**, 2770-2781.

69. Py, B., Higgins, C.F., Krisch, H.M. and Carpousis, A.J. (1996) A DEAD-box RNA helicase in the *Escherichia coli* RNA degradosome. *Nature*, **381**, 169-172.

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

71. Mackie, G.A., Coburn, G.A., Miao, X., Briant, D.J., Prud'homme-Genereux, A., Stickney, L.M. and Hankins, J.S. (2008) Preparation of the *Escherichia coli* RNase E protein and reconstitution of the RNA degradosome. *Methods Enzymol*, **447**, 199-213.

72. Masse, E., Escorcia, F.E. and Gottesman, S. (2003) Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes Dev*, **17**, 2374-2383.

73. Vytvytska, O., Jakobsen, J.S., Balcunaite, G., Andersen, J.S., Baccarini, M. and von Gabain, A. (1998) Host factor I, Hfq, binds to *Escherichia coli* ompA mRNA in a growth rate-dependent fashion and regulates its stability. *Proc Natl Acad Sci U S A*, **95**, 14118-14123.

74. Blum, E., Py, B., Carpousis, A.J. and Higgins, C.F. (1997) Polyphosphate kinase is a component of the *Escherichia coli* RNA degradosome. *Mol. Microbiol.*, **26**, 387-398.

75. Khemici, V. and Carpousis, A.J. (2004) The RNA degradosome and poly(A) polymerase of *Escherichia coli* are required *in vivo* for the degradation of small mRNA decay intermediates containing REP-stabilizers. *Molecular Microbiol.*, **51**, 777-790.

76. Bernstein, J.A., Lin, P.-H., Cohen, S.N. and Lin-Chao, S. (2004) Global analysis of *Escherichia coli* RNA degradosome function using DNA microarrays. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 2748-2763.

77. Spickler, C., Stronge, V. and Mackie, G.A. (2001) Preferential cleavage of degradative intermediates of *rpsT* mRNA by the *Escherichia coli* degradosome. *J. Bacteriol.*, **183**, 1106-1109.

78. Callaghan, A.J., Redko, Y.U., Murphy, L.M., Grossman, J.G., Yates, D., Garman, E., Ilag, L.L., Robinson, C.V., Symmons, M.F., McDowall, K.J. *et*

al. (2005) "Zn-Link": a metal sharing interface that organizes the quaternary structure and catalytic site of endoribonuclease, RNase E. *Biochemistry*, **44**, 4667-4775.

79. Callaghan, A.J., Grossman, J.G., Redko, Y.U., Ilag, L.L., Moncrieffe, M.C., Symmons, M.F., Robinson, C.V., McDowall, K.J. and Luisi, B.F. (2003) Quaternary structure and catalytic activity of the *Escherichia coli* ribonuclease E amino-terminal catalytic domain. *Biochemistry*, **42**, 12848-13855.

80. Liou, G.G., Jane, W.N., Cohen, S.N., Lin, N.S. and Lin-Chao, S. (2001) RNA degradosomes exist *in vivo* in *Escherichia coli* as multicomponent complexes associated with the cytoplasmic membrane via the N-terminal region of ribonuclease E. *Proc Natl Acad Sci U S A*, **98**, 63-68.

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

82. Taghbalout, A. and Rothfield, L. (2008) New insights into the cellular organization of the RNA processing and degradation machinery of *Escherichia coli*. *Mol Microbiol*, **70**, 780-782.

83. Taghbalout, A. and Rothfield, L. (2007) RNase E and the other constituents of the RNA degradosome are components of the bacterial cytoskeleton. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 1667-1672.

84. Andrade, J.M., Cairrao, F. and Arraiano, C.M. (2006) RNase R affects gene expression in stationary phase: regulation of *ompA*. *Molecular Microbiol.*, **60**, 219-228.

85. Sousa, S., Marchand, I. and Dreyfus, M. (2001) Autoregulation allows *Escherichia coli* RNase E to adjust continuously its synthesis to that of its substrates. *Molecular Microbiol.*, **42**, 867-878.

86. Schuck, A., Diwa, A. and Belasco, J.G. (2009) RNase E autoregulates its synthesis in *Escherichia coli* by binding directly to a stem-loop in the rne 5' untranslated region. *Mol Microbiol*, **72**, 470-478.

87. Lee, K., Zhan, X., Gao, J., Feng, Y., Meganathan, R., Cohen, S.N. and Georgiou, G. (2003) RraA: a protein inhibitor of RNase E activity that globally modulates RNA abundance in *E. coli*. *Cell*, **114**, 623-634.

88. Yeom, J.H. and Lee, K. (2006) RraA rescues *Escherichia coli* cells overproducing RNase E from growth arrest by modulating the ribonucleolytic activity. *Biochem. Biophys. Res. Comm.*, **345**, 1372-1376.

89. Zhao, M., Zhou, L., Kawaraski, Y. and Georgiou, G. (2006) Regulation of RraA, a protein inhibitor of RNase E-mediated mRNA decay. *J. Bacteriology*, **188**, 3257-3263.

90. Zhou, L., Zhao, M., Wolf, R.Z., Graham, D.E. and Georgiou, G. (2009) Transcriptional regulation of the *Escherichia coli* gene *rraB*, encoding a protein inhibitor of RNase E. *J Bacteriol*, **191**, 6665-6674.

91. Yeom, J.H., Shin, E., Go, H., Sim, S.H., Seong, M.J. and Lee, K. (2008) Functional implications of the conserved action of regulators of ribonuclease activity. *J Microbiol Biotechnol*, **18**, 1353-1356.

92. Yeom, J.H., Go, H., Shin, E., Kim, H.L., Han, S.H., Moore, C.J., Bae, J. and Lee, K. (2008) Inhibitory effects of RraA and RraB on RNAse E-related enzymes imply conserved functions in the regulated enzymatic cleavage of RNA. *FEMS Microbiol Lett*, **285**, 10-15.

93. Gao, J., Lee, K., Zhao, M., Qiu, J., Zhan, X., Saxena, A., Moore, C.J., Cohen, S.N. and Georgiou, G. (2006) Differential modulation of *E. coli* mRNA abundance by inhibitory proteins that alter the composition of the degradosome. *Mol Microbiol*, **61**, 394-406.

94. Singh, D., Chang, S.J., Lin, P.H., Averina, O.V., Kaberdin, V.R. and Lin-Chao, S. (2009) Regulation of ribonuclease E activity by the L4 ribosomal protein of Escherichia coli. *Proc Natl Acad Sci U S A*, **106**, 864-869.

95. Deana, A., Celesnik, H. and Belasco, J.G. (2008) The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature* **451**, 355-358.

96. McDowall, K.J., Hernandez, R.G., Lin-Chao, S. and Cohen, S.N. (1993) The *ams-1* and *rne-3071* temperature-sensitive mutations in the *ams* gene are in close proximity to each other and cause substitutions within a domain that resembles a product of the *Escherichia coli rne* locus. *J. Bacteriol.*, **175**, 4245-4249.

97. Stead, M.B., Marshburn, S., Mohanty, B.K., Mitra, J., Castillo, L.P., Ray, D., van Bakel, H., Hughes, T.R. and Kushner, S.R. (2011) Analysis of *Escherichia coli* RNase E and RNase III activity *in vivo* using tiling microarrays. *Nucleic Acids Res*, **39**, 3188-3203.

98. Apirion, D. and Lassar, A.B. (1978) Conditional lethal mutant of *Escherichia coli* which affects processing of ribosomal RNA. *J. Biological Chemistry*, **253**, 1738-1742.

99. Chung, D.H., Min, Z., Wang, B.C. and Kushner, S.R. (2010) Single amino acid changes in the predicted RNase H domain of *Escherichia coli* RNase G lead to complementation of RNase E deletion mutants. *RNA*, **16**, 1371-1385.

100. Kushner, S.R. (2007) In Böck, A., Curtis III., R., Gross, C. A., Kaper, J. B., Neidhardt, F. C., Nyström, T., Rudd, K. E. and L., S. C. (eds.), *Escherichia coli and Salmonella: cellular and molecular biology*. American Society for Microbiology Press, Washington, DC, pp. http://www.ecosal.org.

101. Kushner, S.R. (2004) In Higgins, P. (ed.), *The Bacterial Chromosome*. ASM Press, Washington, pp. 327-345.

102. Kushner, S.R. (2004) mRNA decay in prokaryotes and eukaryotes: Different approaches to a similar problem. *IUBMB Life*, **56**, 585-594.

103. Ow, M.C. and Kushner, S.R. (2002) Initiation of tRNA maturation by RNase E is essential for cell viability in *E. coli*. *Genes Dev*, **16**, 1102-1115.

104. Babitzke, P., Granger, L., Olszewski, J. and Kushner, S.R. (1993) Analysis of mRNA decay and rRNA processing in *Escherichia coli* multiple mutants carrying a deletion in RNase III. *J. Bacteriol.*, **175**, 229-239.

105. King, T.C., Sirdeshmukh, R. and Schlessinger, D. (1984) RNase III cleavage is obligate for maturation but not for function of *Escherichia coli* pre-23S rRNA. *Proceedings of the National Academy of Sciences of the United States of America*, **81**, 185-188.

106. Gegenheimer, P. and Apirion, D. (1980) Precursors to 16S and 23S ribosomal RNA from a ribonuclear III-strain of *Escherichia coli* contain intact RNase III processing sites. *Nucleic Acids Res.*, **8**, 1873-1891.

107. Drider, D. and Condon, C. (2004) The continuing story of endoribonuclease III. *J. Mol. Microbiol. Biotechnol.*, **8**, 195-200.

108. Nicholson, A.W. (1999) Function, mechanism and regulation of bacterial ribonucleases. *FEMS Microbiol. Rev.*, **23**, 371-390.

109. Takiff, H.E., Chen, S.M. and Court, D.L. (1989) Genetic analysis of the *rnc* operon of *Escherichia coli*. *J. Bacteriol.*, **171**, 2581-2590.

110. Herskovitz, M.A. and Bechhofer, D.H. (2000) Endoribonuclease RNase III is essential in *Bacillus subtilis*. *Mol. Microbiol.*, **38**, 1027-1033.

111. Pertzev, A.V. and Nicholson, A.W. (2006) Characterization of RNA sequence determinants and antideterminants of processing reactivity for a minimal substrate of *Escherichia coli* ribonuclease III. *Nucleic Acids Res.*, **34**, 3708-3721.

112. Blaszczyk, J., Tropea, J.E., Bubunenko, M., Routzahn, K.M., Waugh, D.S., Court, D.L. and Ji, X. (2001) Crystallographic and modeling studies of RNase III suggest a mechanism for double-stranded RNA cleavage. *Structure*, **9**, 1225-1236.

113. Meister, G. and Tuschl, T. (2004) Mechanisms of gene silencing by double-stranded RNA. *Nature*, **431**, 343-349.

114. Chen, P.Y. and Meister, G. (2005) microRNA-guided posttranscriptional gene regulation. *Biol. Chem.*, **386**, 1205-1218.

115. Bernstein, E., Caudy, A.A., Hammond, S.M. and Hannon, G.J. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, **409**, 363-366.

116. Carmell, M.A. and Hannon, G.J. (2004) RNase III enzymes and the initiation of gene silencing. *Nat. Struct. Mol. Biol.*, **11**, 214-218.

117. Dunn, J.J. and Studier, F.W. (1973) T7 early RNAs and *Escherichia coli* ribosomal RNAs are cut from large precursor RNAs *in vivo* by ribonuclease 3. *Proc. Natl. Acad. Sci. U S A*, **70**, 3296-3300.

118. Matsunaga, J., Simons, E.L. and Simons, R.W. (1996) RNase III autoregulation: structure and function of *rncO*, the posttranscriptional "operator". *RNA*, **2**, 1228-1240.

119. Bardwell, J.C., Regnier, P., Chen, S.M., Nakamura, Y., Grunberg-Manago, M. and Court, D.L. (1989) Autoregulation of RNase III operon by mRNA processing. *EMBO. J.*, **8**, 3401-3407.

120. Talkad, V., Achord, D. and Kennell, D. (1978) Altered mRNA metabolism in ribonuclease III-deficient strains of *Escherichia coli*. *J. Bacteriol.*, **135**, 528-541.

121. Afonyushkin, T., Vecerek, B., Moll, I., Blasi, U. and Kaberdin, V.R. (2005) Both RNase E and RNase III control the stability of *sodB* mRNA upon translational inhibition by the small regulatory RNA RyhB. *Nucleic Acids Res.*, **33**, 1678-1689.

122. Resch, A., Afonyushkin, T., Lombo, T.B., McDowall, K.J., Blasi, U. and Kaberdin, V.R. (2008) Translational activation by the noncoding RNA DsrA involves alternative RNase III processing in the rpoS 5'-leader. *RNA*, **14**, 454-459.

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

124. Matsunaga, J., Simons, E.L. and Simons, R.W. (1997) *Escherichia coli* RNase III (*rnc*) autoregulation occurs independently of *rnc* gene translation. *Molecular Microbiology*, **26**, 1125-1135.

125. Kim, K.S., Manasherob, R. and Cohen, S.N. (2008) YmdB: a stress-responsive ribonuclease-binding regulator of *E. coli* RNase III activity. *Genes Dev*, **22**, 3497-3508.

126. Stark, B.C., Kole, R., Bowman, E.J. and Altman, S. (1977) Ribonuclease P: an enzyme with an essential RNA component. *Proceedings of the*

National Academy of Sciences of the United States of America, **75**, 3719-3721.

127. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. and Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, **35**, 849-857.

128. Brown, J.W., Haas, E.S., James, B.D., Hunt, D.A., Liu, J.S. and Pace, N.R. (1991) Phylogenetic analysis and evolution of RNase P RNA in proteobacteria. *J Bacteriol*, **173**, 3855-3863.

129. Li, Y., Cole, K. and Altman, S. (2003) The effect of a single, temperature-sensitive mutation on global gene expression in *Escherichia coli*. *RNA*, **9**, 518-532.

130. Li, Y. and Altman, S. (2003) A specific endoribonuclease, RNase P, affects gene expression of polycistronic operon mRNAs. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 13213-13218.

131. Mohanty, B.K. and Kushner, S.R. (1999) Analysis of the function of *Escherichia coli* poly(A) polymerase I in RNA metabolism. *Mol. Microbiol.*, **34**, 1094-1108.

132. Mohanty, B.K. and Kushner, S.R. (2007) Ribonuclease P processes polycistronic tRNA transcripts in *Escherichia coli* independent of ribonuclease E. *Nucleic Acids Research*, **35**, 7614-7625.

133. Kamada, K. and Hanaoka, F. (2005) Conformational change in the catalytic site of the ribonuclease YoeB toxin by YefM antitoxin. *Molecular cell*, **19**, 497-509.

134. Pedersen, K., Zavialov, A.V., Pavlov, M.Y., Elf, J., Gerdes, K. and Ehrenberg, M. (2003) The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site. *Cell*, **112**, 131-140.

135. Sat, B., Hazan, R., Fisher, T., Khaner, H., Glaser, G. and Engelberg-Kulka, H. (2001) Programmed cell death in *Escherichia coli*: Some antibiotics can trigger *mazEF* lethality. *J. Bacteriol.*, **183**, 2041-2045.

136. Zhang, Y., Zhang, J., Hoeflich, K.P., Ikura, M., Qing, G. and Inouye, M. (2003) MazF cleaves cellular mRNAs specifically at ACA to block protein synthesis in *Escherichia coli*. *Molecular cell*, **12**, 913-923.

137. Zhang, Y., Zhang, J., Hara, H., Kato, I. and Inouye, M. (2005) Insights into the mRNA cleavage mechanism by MazF, an mRNA interferase. *J. Biological Chemistry*, **280**, 3143-3150.

138. Aiba, H. (2007) Mechanism of RNA silencing by Hfq-binding small RNAs. *Curr Opin Microbiol*, **10**, 134-139.

139. Udekwu, K.I., Darfeuille, F., Vogel, J., Reimegard, J., Holmqvist, E. and Wagner, E.G. (2005) Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. *Genes Dev*, **19**, 2355-2366.

140. Geismann, T.A. and Touati, D. (2004) Hfq, a new chaperoning role: binding to messenger RNA determines access for small RNA regulator. *Embo J*, **23**, 396-405.

141. Sledjeski, D.D., Whitman, C. and Zhang, A. (2001) Hfq is necessary for regulation by the untranslated RNA DsrA. *J. Bacteriology*, **183**, 1997-2005.

142. Reichenbach, B., Maes, A., Kalamorz, F., Hajnsdorf, E. and Gorke, B. (2008) 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.

143. Hershberg, R., Altuvia, S. and Margalit, H. (2003) A survey of small RNA-encoding genes in *Escherichia coli*. *Nucleic Acids Research*, **31**, 1813-1820.

144. Udekwu, K.I. (2010) Transcriptional and post-transcriptional regulation of the *Escherichia coli luxS* mRNA; involvement of the sRNA MicA. *PLoS One*, **5**, e13449.

145. Hankins, J.S., Denroche, H. and Mackie, G.A. (2010) Interactions of the RNA-binding protein Hfq with *cspA* mRNA, encoding the major cold shock protein. *J Bacteriol*, **192**, 2482-2490.

146. Vecerek, B., Beich-Frandsen, M., Resch, A. and Blasi, U. (2010) Translational activation of rpoS mRNA by the non-coding RNA DsrA and Hfq does not require ribosome binding. *Nucleic Acids Res*, **38**, 1284-1293.

147. Basineni, S.R., Madhugiri, R., Kolmsee, T., Hengge, R. and Klug, G. (2009) The influence of Hfq and ribonucleases on the stability of the small non-coding RNA OxyS and its target *rpoS* in *E. coli* is growth phase dependent. *RNA biology*, **6**, 584-594.

148. Updegrove, T., Wilf, N., Sun, X. and Wartell, R.M. (2008) Effect of Hfq on RprA-rpoS mRNA pairing: Hfq-RNA binding and the influence of the 5' rpoS mRNA leader region. *Biochemistry*, **47**, 11184-11195.

149. Brennan, R.G. and Link, T.M. (2007) Hfq structure, function and ligand binding. *Curr Opin Microbiol*, **10**, 125-133.

150. Zhang, A., Wassarman, K.M., Rosenow, C., Tjaden, B.C., Storz, G. and Gottesman, S. (2003) Global analysis of small RNA and mRNA targets of Hfq. *Molecular Microbiol.*, **50**, 1111-1124.

151. Afonyushkin, T., Vecerek, B., Moll, I., Blasi, U. and Kaberdin, V.R. (2005) Both RNase E and RNase III control the stability of *sodB* mRNA upon translational inhibition by the small regulatory RNA RyhB. *Nucleic Acids Research*, **33**, 1678-1689.

152. Masse, R., Escorcia, F.E. and Gottesman, S. (2003) Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes and Development*, **17**, 2374-2383.

153. Khodursky, A.B., Peter, B.J., Cozzarelli, N.R., Botstein, D., Brown, P.O., and Yanofsky, C. (2000) DNA microarray analysis of gene expression in response to physiological and genetic changes that affect tryptophan metabolism in *Escherichia coli*. *PNAS*, **97**, 12170-12175.

154. Bernstein, J.A., Khodursky, A.B., Lin, P.H., Lin-Chao, S. and Cohen, S.N. (2002) Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays. *Proc Natl Acad Sci U S A*, **99**, 9697-9702.

155. Lee, K., Bernstein, J.A. and Cohen, S.N. (2002) RNase G complementation of rne null mutation identifies functional interrelationships with RNase E in Escherichia coli. *Mol Microbiol*, **43**, 1445-1456.

156. Hatfield, G.W., Hung, S. and Baldi, P. (2003) Differential analysis of DNA microarray gene exprression data. *Molecular Microbiol.*, **47**, 871-877.

157. Conway, T. and Schoolnik, G.K. (2003) Microarray expression profiling:capturing a genome-wide portrait of the transcriptome. *Molecular Microbiol.*, **47**, 879-889.

158. Glynn, E.F., Megee, P.C., Yu, H.G., Mistrot, C., Unal, E., Koshland, D.E., DeRisi, J.L. and Gerton, J.L. (2004) Genome-wide mapping of the cohesin complex in the yeast *Saccharomyces cerevisiae*. *PLoS Biol*, **2**, E259.

159. Hundt, S., Zeaigler, A., Lange, C., Soppa, J. and Klug, G. (2007) Global analysis of mRNA decay in *Halobacterium salinarum NRC-1 at single-gene resolution using DNA microarrays*. *J. Bacteriol.*, **189**, 6936-6944.

160. Hughes, J.A., Brown, L.R. and Ferro, A.J. (1987) Nucleotide sequence and analysis of the coliphage T3 S-adenosylmethionine hydrolase gene and its surrounding ribonuclease III processing sites. *Nucleic Acids Res*, **15**, 717-729.

161. Hiley, S.L., Babak, T. and Hughes, T.R. (2005) Global analysis of yeast RNA processing indentifies new targets of RNase III and uncovers a link between tRNA 5' end processing and tRNA splicing. *Nucleic Acids Research*, **33**, 3048-3056.

162. Glisin, V., Crkvenjakov, R. and Byus, C. (1974) Ribonucleic acid isolated by cesium chloride centrifugation. *Biochemistry*, **13**, 2633-2637.

163. Babitzke, P., Granger, L. and Kushner, S.R. (1993) Analysis of mRNA decay and rRNA processing in *Escherichia coli* multiple mutants carrying a deletion in RNase III. *Journal of Bacteriology*, **175**, 229-239.

164. Sevag, M.G., Lackman, D.B., Smolens, J. (1938) The isolation of the components of streptococcal nucleoproteins in serologically active form *J Biol Chem*, **124**, 425-436.

165. Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem*, **162**, 156-159.

166. Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry*, **18**, 5294-5299.

167. Macfarlane, D.E. and Dahle, C.E. (1997) Isolating RNA from clinical samples with Catrimox-12TM and lithium chloride. *Journal of Clinical Laboratory Analysis*, **11**, 132-139.

168. Macfarlane, D.E. and Dahle, C.E. (1993) Isolating RNA from whole blood-The dawn of RNA-based diagnostics. *Nature* **362**, 186-188.

169. O'Hara, E., Wang, R.F., Meffen, F.V., Armstrong, K., Bargmann, C., Granger, L., Aldea, M.A., Babitzke, P. and Kushner, S.R. (1997) Identification of a new gene *(mrsA)* in *Escherichia coli* that is required for

cell viability and affects mRNA stability and cell division. *J. Bacteriol.*, submitted.

170. Mohanty, B.K., Giladi, H., Maples, V.F. and Kushner, S.R. (2008) Analysis of RNA decay, processing, and polyadenylation in *Escherichia coli* and other prokaryotes. *Methods Enzymol*, **447**, 3-29.

171. Gros, F., Hiatt, H., Gilbert, W., Kurland, C.B., Risebrough, R.W. and Watson, J.D. (1961) Unstable ribonucleic acid revealed by pulse labelling of *Escherichia coli*. *Nature*, **190**, 581-585.

172. Paul, B.J., Ross, W., Gaal, T. and Gourse, R.L. (2004) rRNA transcription in *Escherichia coli*. *Annu. Rev. Genet.*, **38**, 749-770.

173. Gourse, R.L., Gaal, T., Bartlett, M.S., Appleman, J.A. and Ross, W. (1996) rRNA transcription and growth rate-dependent regulation of ribosome synthesis in *Escherichia coli*. *Annual Review of Microbiology*, **50**, 645-678.

174. Hirvonen, C.A., Ross, W., Wozniak, C.E., Marasco, E., Anthony, J.R., Aiyar, S.E., Newburn, V.H. and Gourse, R.L. (2001) Contributions of UP elements and the transcription factor FIS to expression from the seven rrn P1 promoters in Escherichia coli. *J Bacteriol*, **183**, 6305-6314.

175. Nikolaev, N., Silengo, L. and Schlessinger, D. (1973) Synthesis of a large precursor to ribosomal RNA in a mutant of *Escherichia coli*. *Proc. Natl. Acad. Sci. U S A*, **70**, 3361-3365.

176. Nikolaev, N., Silengo, L. and Schlessinger, D. (1973) A role for ribonuclease 3 in processing of ribosomal ribonucleic acid and messenger

ribonucleic acid precursors in *Escherichia coli*. *J. Biol. Chem.*, **248**, 7967-7969.

177. Sirdeshmukh, R. and Schlessinger, D. (1985) Ordered processing of *Escherichia coli* 23S rRNA *in vitro*. *Nucleic Acids Res*, **13**, 5041-5054.

178. Sirdeshmukh, R., Krych, M. and Schlessinger, D. (1985) *Escherichia coli* 23S ribosomal RNA truncated at its 5' terminus. *Nucleic Acids Res*, **13**, 1185-1192.

179. Allas, U., Liiv, A. and Remme, J. (2003) Functional interaction between RNase III and the *Escherichia coli* ribosome. *BMC Mol Biol*, **4**, 8.

180. Lewicki, B.T., Margus, T., Remme, J. and Nierhaus, K.H. (1993) Coupling of rRNA transcription and ribosomal assembly *in vivo*. Formation of active ribosomal subunits in *Escherichia coli* requires transcription of rRNA genes by host RNA polymerase which cannot be replaced by bacteriophage T7 RNA polymerase. *J Mol Biol*, **231**, 581-593.

181. Li, Z., Pandit, S. and Deutscher, M.P. (1999) RNase G (CafA protein) and RNase E are both required for the 5' maturation of 16S ribosomal RNA. *EMBO J.*, **18**, 2878-2885.

182. Ghora, B.K. and Apirion, D. (1978) Structural analysis and *in vitro* processing to p5 rRNA of a 9S RNA molecule isolated from an *rne* mutant of *E. coli*. *Cell*, **15**, 1055-1066.

183. Li, Z. and Deutscher, M.P. (1995) The tRNA processing enzyme RNase T is essential for maturation of 5S RNA. *Proc. Natl. Acad. Sci. U S A*, **92**, 6883-6886.

184. Ow, M.C., Liu, Q. and Kushner, S.R. (2000) Analysis of mRNA decay and rRNA processing in *Escherichia coli* in the absence of RNase E-based degradosome assembly. *Mol. Microbiol.*, **38**, 854-866.

185. Li, Z., Pandit, S. and Deutscher, M.P. (1999) Maturation of 23S ribosomal RNA requires the exoribonuclease RNase T. *RNA*, **5**, 139-146.

186. Cheng, Z.F. and Deutscher, M.P. (2003) Quality control of ribosomal RNA mediated by polynucleotide phosphorylase and RNase R. *Proc. Natl. Acad. Sci. U S A*, **100**, 6388-6393.

187. Li, Z. and Deutscher, M.P. (1995) The tRNA processing enzyme RNase T is essential for maturation of 5S RNA. *Proc. Natl. Acad. Sci. USA*, **92**, 6883-6886.

188. Li, Z. and Deutscher, M.P. (2002) RNase E plays an essential role in the maturation of *Escherichia coli* tRNA precursors. *RNA*, **8**, 97-109.

189. Hayes, F. and Vasseur, M. (1976) Processing of the 17-S *Escherichia coli* precursor RNA in the 27-S pre-ribosomal particle. *Eur. J. Biochem.*, **61**, 433-442.

190. Slagter-Jager, J.G., Puzis, L., Gutgsell, N.S., Belfort, M. and Jain, C. (2007) Functional defects in transfer RNAs lead to the accumulation of ribosomal RNA precursors. *RNA*, **13**, 597-605.

191. Song, W.S., Lee, M. and Lee, K. (2011) RNase G participates in processing of the 5'-end of 23S ribosomal RNA. *J Microbiol*, **49**, 508-511.

192. Jagessar, K.L. and Jain, C. (2010) Functional and molecular analysis of *Escherichia coli* strains lacking multiple DEAD-box helicases. *RNA*, **16**, 1386-1392.

193. Charollais, J., Pflieger, D., Vinh, J., Dreyfus, M. and Iost, I. (2003) The DEAD-box RNA helicase SrmB is involved in the assembly of 50S ribosomal subunits in *Escherichia coli*. *Molecular Microbiol.*, **48**, 1253-1265.

194. Linden, M.H., Hartmann, R.K. and Klostermeier, D. (2008) The putative RNase P motif in the DEAD box helicase Hera is dispensable for efficient interaction with RNA and helicase activity. *Nucleic Acids Res*, **36**, 5800-5811.

195. Karginov, F.V. and Uhlenbeck, O.C. (2004) Interaction of Escherichia coli DbpA with 23S rRNA in different functional states of the enzyme. *Nucleic Acids Res*, **32**, 3028-3032.

196. Pabbaraju, K., Miller, W.L. and Sanderson, K.E. (2000) Distribution of intervening sequences in the genes for 23S rRNA and rRNA fragmentation among strains of the Salmonella reference collection B (SARB) and SARC sets. *J Bacteriol*, **182**, 1923-1929.

197. Miller, W.L., Pabbaraju, K. and Sanderson, K.E. (2000) Fragmentation of 23S rRNA in strains of Proteus and Providencia results from intervening sequences in the rrn (rRNA) genes. *J Bacteriol*, **182**, 1109-1117.

198. Mattatall, N.R. and Sanderson, K.E. (1998) RNase III deficient Salmonella typhimurium LT2 contains intervening sequences (IVSs) in its 23S rRNA. *FEMS Microbiol Lett*, **159**, 179-185.

199. Evguenieva-Hackenberg, E. and Klug, G. (2000) RNase III processing of intervening sequences found in helix 9 of 23S rRNA in the alpha subclass of Proteobacteria. *J Bacteriol*, **182**, 4719-4729.

200. Evguenieva-Hackenberg, E. (2005) Bacterial ribosomal RNA in pieces. *Mol Microbiol*, **57**, 318-325.

201. King, T.C., Sirdeshmukh, R. and Schlessinger, D. (1984) RNase III cleavage is obligate for maturation but not for function of *Escherichia coli* pre-23S rRNA. *Proc. Natl. Acad. Sci. U S A*, **81**, 185-188.

202. Srivastava, A.K. and Schlessinger, D. (1989) Processing pathway of *Escherichia coli* 16S precursor rRNA. *Nucleic Acids Res.*, **17**, 1649-1663.

Fig. 1. Model of *E. coli* rRNA processing.

This model is not drawn to scale. The 30S pre-rRNA transcript is first cleaved by RNase III at two physical hairpin locations, cleaving the RNA the hairpin structures formed by the spacer sequences adjacent to mature 16S and 23S rRNAs, generating 17S, 25S, and 9S pre-rRNAs (labeled in purple). 17S rRNA is cleaved first by RNase E and then by RNase G at the mature 5' end of 16S rRNA. RNase E also cleaves 9S three nt on each side of the mature termini of 5S rRNA forming p5S. RNase P cleaves at the mature 5' end of the tRNA. Exoribonucleases (primarily RNase T) are responsible for the 3' end maturation of the tRNA, 23S rRNA, and 5S rRNA, but the endoribonuclease(s) responsible for the 3' end maturation of 16S, the 5' end of 23S, and 5' end of 5S rRNAs remain unidentified (labeled in red).

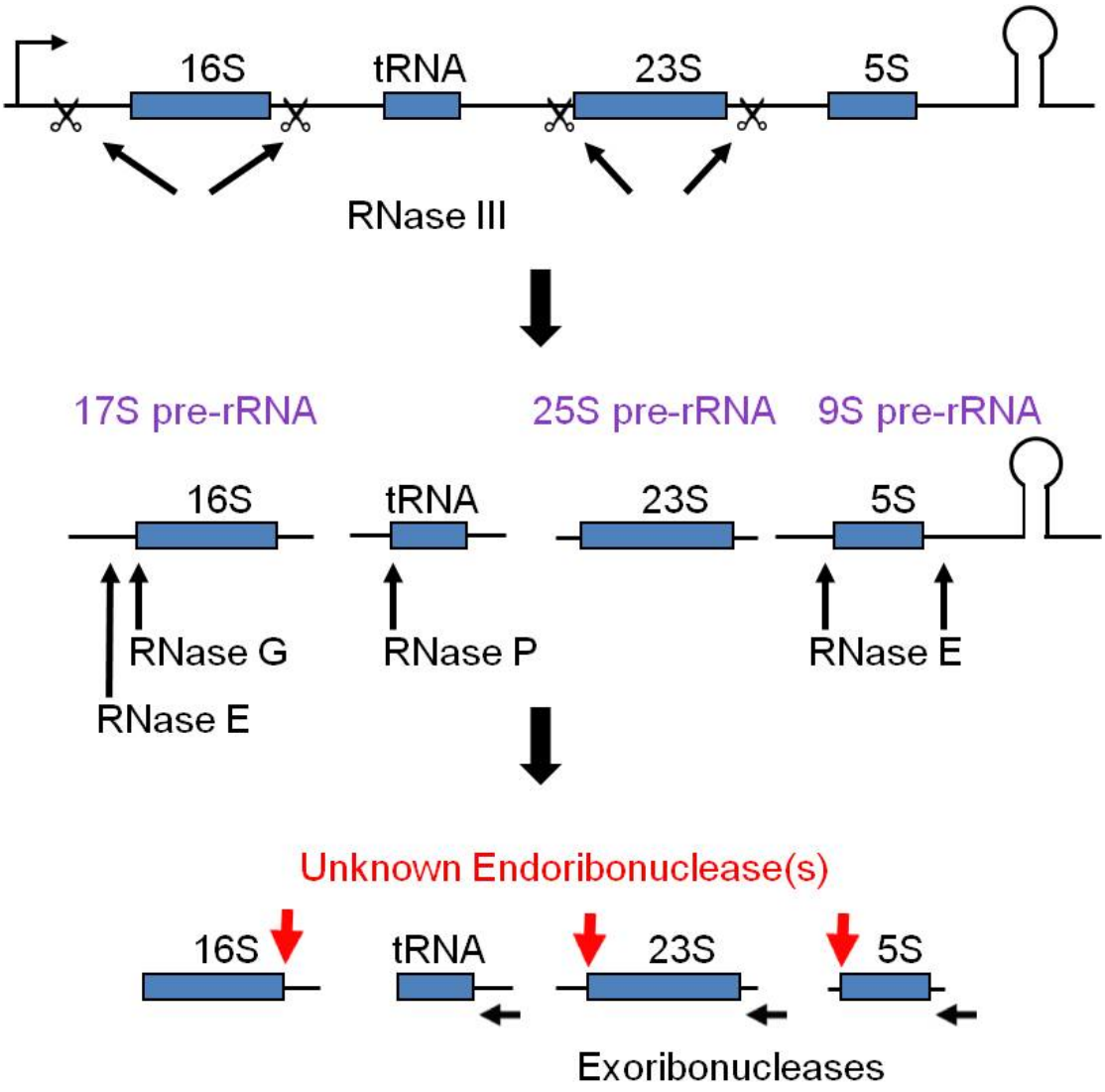
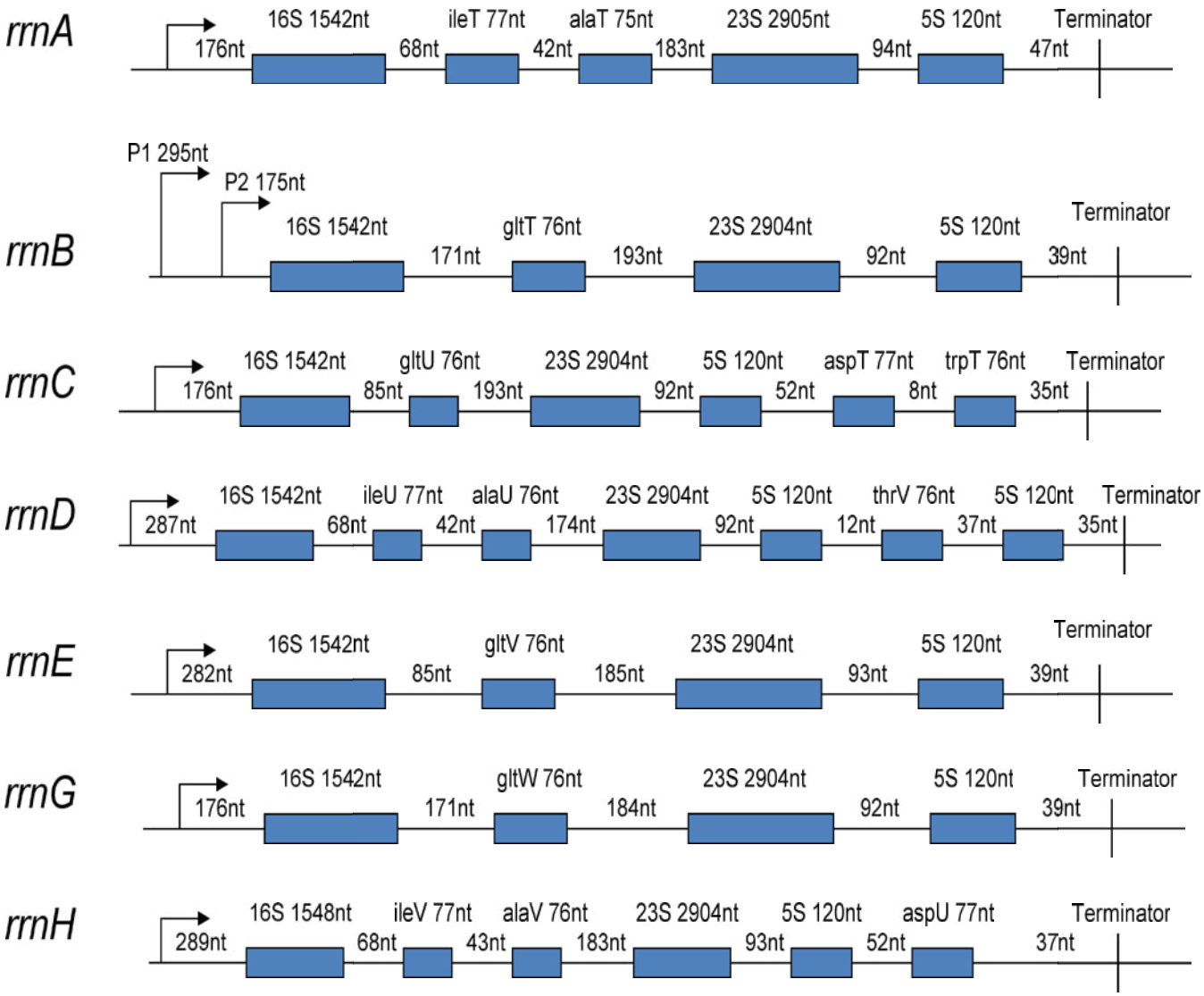

16S
tRNA
23S
5S
RNase III
17S pre-rRNA
25S pre-rRNA
9S pre-rRNA
16S
tRNA
23S
5S
RNase G
RNase E
RNase P
RNase E
Unknown Endoribonuclease(s)
16S
tRNA
23S
5S
Exoribonucleases

Fig. 2. Ribosomal RNA operons in *E. coli.*

This model is not drawn to scale. Each of the seven operons are identified on the left side by their respective *rrn* designations. Mature rRNAs or tRNAs are indicated by the light blue bars.


rrnA
16S 1542nt
ileT 77nt
alaT 75nt
23S 2905nt
5S 120nt
Terminator
176nt
68nt
42nt
183nt
94nt
47nt
rrnB
P1 295nt
P2 175nt
16S 1542nt
gltT 76nt
23S 2904nt
5S 120nt
Terminator
171nt
193nt
92nt
39nt
rrnC
16S 1542nt
gltU 76nt
23S 2904nt
5S 120nt
aspT 77nt
trpT 76nt
Terminator
176nt
85nt
193nt
92nt
52nt
8nt
35nt
rrnD
16S 1542nt
ileU 77nt
alaU 76nt
23S 2904nt
5S 120nt
thrV 76nt
5S 120nt
Terminator
287nt
68nt
42nt
174nt
92nt
12nt
37nt
35nt
rrnE
16S 1542nt
gltV 76nt
23S 2904nt
5S 120nt
Terminator
282nt
85nt
185nt
93nt
39nt
rrnG
16S 1542nt
gltW 76nt
23S 2904nt
5S 120nt
Terminator
176nt
171nt
184nt
92nt
39nt
rrnH
16S 1548nt
ileV 77nt
alaV 76nt
23S 2904nt
5S 120nt
aspU 77nt
Terminator
289nt
68nt
43nt
183nt
93nt
52nt
37nt

CHAPTER 2

ANALYSIS OF *ESCHERICHIA COLI* RNASE E AND RNASE III ACTIVITY *IN VIVO* USING TILING MICROARRAYS[1]

---

[1] Stead, M.B., Marshburn, S., Mohanty, B.K., Mitra, J., Castillo, L.P., Ray, D., van Bakel, H., Hughes, T.R. and S.R. Kushner. 2011. *Nucleic Acids Research*. 39:3188-3203.

Reprinted here with permission of the publisher.

ABSTRACT

Tiling microarrays have proven to be a valuable tool for gaining insights into the transcriptomes of microbial organisms grown under various nutritional or stress conditions. Here we describe the use of such an array, constructed at the level of 20 nucleotide resolution for the *Escherichia coli* MG1655 genome, to observe genome-wide changes in the steady-state RNA levels in mutants defective in either RNase E or RNase III. The array data were validated by comparison to previously published results for a variety of specific transcripts as well as independent Northern analysis of additional mRNAs and sRNAs. In the absence of RNase E, 60% of the annotated coding sequences showed either increases or decreases in their steady-state levels. In contrast, only 12% of the coding sequences were affected in the absence of RNase III. Unexpectedly, many coding sequences showed decreased abundance in the RNase E mutant, while more than half of the annotated sRNAs showed changes in abundance. Furthermore, the steady-state levels of many transcripts showed overlapping effects of both ribonucleases. Data is also presented demonstrating how the arrays were used to identify potential new genes, RNase III cleavage sites, and the direct or indirect control of specific biological pathways.

INTRODUCTION

The analysis of the post-transcriptional processing, maturation, and decay of RNA molecules in prokaryotes such as *Escherichia coli* has historically focused on individual classes of molecules such as rRNAs, tRNAs, mRNAs, and small RNAs (sRNAs). Thus considerable effort has been invested in understanding the maturation of 30S rRNA precursors into mature 16S, 23S and 5S rRNA species (1), the processing of tRNA precursors (2-7), the mechanisms of mRNA decay (8) and the processing and degradation of sRNAs (9-11).

Although it was originally thought that different ribonucleases might be involved in the processing, maturation and decay of particular classes of RNA molecules, work over the past fifteen years has clearly demonstrated that a limited set of ribonucleases mediate all aspects of RNA metabolism in *E. coli* (7,8,12) (http://ecosal.org). For example, the essential endoribonuclease RNase E, encoded by the *rne* gene, is involved in many aspects of RNA metabolism, including mRNA decay (13-17), sRNA processing and decay (10,18), tRNA processing (3,5,19), and rRNA maturation (20,21). In contrast, RNase G, a paralog of RNase E, appears to have a much more limited range of substrates, including some mRNAs, and 16S rRNA precursors (21-25). On the other hand, RNase III is primarily known for its role in rRNA maturation (26), but has also been shown to be involved, to a limited extent, in mRNA degradation and sRNA processing (27-32). In addition, recent studies have demonstrated that some *E. coli* sRNAs regulate the stability and translation initiation efficiency of specific mRNAs through RNase III-dependent cleavages (33,34).

Historically, the analysis of RNA transcripts has relied on either Northern blots or, to a lesser extent, qRT-PCR. Northern analysis is a particularly powerful method for studying RNA processing and decay, since it permits the visualization of both a full-length or mature transcript and its degradation or processing intermediates. However, both methods have significant limitations in that they are time consuming and cannot easily discern interactions among ribonucleases or provide an overview of general pathways of RNA processing and decay. Thus, despite years of research, many questions remain unanswered regarding the overall *in vivo* roles of ribonucleases such as RNase E and RNase III in *E. coli* RNA metabolism.

The development of DNA macro- and microarrays led to studies in *E. coli* that explored gene regulation in response to various stresses and growth conditions (35,36). A further application of macro- and microarrays has been to study the effect of nuclease mutations on overall mRNA abundance. For example, Mohanty and Kushner (37) used macroarrays to determine how the deletion of either polynucleotide phosphorylase (PNPase) or RNase II (both 3′ → 5′ exonucleases) affected the steady-state levels of all the *E. coli* open reading frames. In addition, Lee *et al.* (23) demonstrated that the steady-state levels of 40% of the coding sequences (CDSs) in a mutant containing an RNase E deletion, kept viable by a 174-fold increase in RNase G levels (25), changed in abundance compared to a wild-type control. However, these studies relied upon relatively low-resolution gene expression arrays, which only included information about mRNA abundances.

In contrast, tiling DNA microarrays provide the ability to study RNA processing on a transcriptome-wide scale. It is therefore now possible to simultaneously examine the role of any ribonuclease on all coding and non-coding RNAs in the transcriptome. Other potential advantages of the increased resolution associated with tiling microarrays are the ability to predict the approximate locations of RNase cleavage sites, the identification of potentially novel genes and small RNAs, and the examination of the effects of a particular ribonuclease on specific biological pathways.

In the work presented here, we have compared the transcriptomes, at twenty nucleotide resolution, of wild-type *E. coli* to both an RNase E deletion mutant (*rne*Δ*1018*), kept viable by an altered RNase G protein (*rng-219*) (25) and an RNase III null mutant (*rnc-14::*ΔTn*10*) (38). The RNase E deletion mutant was particularly useful because it contained only about a 12-fold increase in a mutant RNase G protein containing a single amino acid substitution in its RNase H domain (25) as compared to the strain employed by Lee *et. al.* (23), which contained a 174-fold increase in the level of an extended form of RNase G (25).

Our analysis of the tiling microarray data for the *rne* deletion strain showed that 1,520 CDSs (35% of the annotated CDSs) contained one or more regions (at least 100 nt, or two contiguous probes) that were increased in abundance compared to the wild-type control by at least 1.5-fold. In addition, 1,096 CDS (25% of CDSs) had at least one region of decreased abundance when compared to the wild-type control. We also determined that at least 47 annotated non-

coding RNAs (ncRNAs) were affected directly or indirectly by the absence of RNase E.

A second experiment employing total RNA isolated from an RNase III null mutant revealed that 391 CDSs (9%) contained one or more regions that were increased in abundance as compared to wild-type, while 120 CDSs (3%) exhibited decreased abundance. This is the first reported array study in *E. coli* using an RNase III null mutant and the data indicates that RNase III has a more widespread role in cellular RNA processing than previously envisioned. Furthermore, there appears to be a high level of redundancy and/or possible cooperation between RNase III/RNase E cleavage events for a number of transcripts. Of considerable interest was the number of changes in the abundance of small regulatory RNAs in the two mutants versus the wild-type control, which in and of itself could be responsible for a number of the observed changes throughout the transcriptome. The data have also permitted the identification of RNase III cleavage sites, including one within an mRNA coding sequence, and a large number of potentially novel transcripts.

MATERIALS AND METHODS

Construction of the *E. coli* tiling microarray

Both strands of the MG1655 genome (version U00096.2, GI:48994873) were tiled using 60 nucleotide-length probes containing 20 bp overlaps at each end. Sixty nucleotide long negative control probes were created by generating random sequences with a 50% GC content and selecting those with low cross-reactivity against *E. coli* probes using OligoPicker (39). The final array design

contained 231,984 *E. coli* probes, 6,415 duplicate *E. coli* probes and 3,000 random probes. Probe sequences were submitted to Agilent Technologies (Palo Alto, CA) for array manufacture (AMADID 015365). Probes were spotted randomly on the microarray slide to reduce any potential effects of background non-specific hybridization.

Bacterial strains and plasmids

All *E. coli* strains used in this study were derived from MG1693 (*thyA715 rph-1*) provided by the *E. coli* Genetic Stock Center, Yale University. SK3564 [*rne∆1018::bla thyA715 rph-1 recA56 srlD::*Tn*10* (Tc^r) /pDHK30(*rng-219* Sm^r/Sp^r)/pWSK129 (Km^r)] is an RNase E deletion strain in which cell viability is supported by a mutant RNase G (*rng-219*) protein synthesized from a single copy plasmid (25). SK4455 (*rnc-14::∆*Tn*10 thyA715 rph-1*) was constructed by moving the *rnc-14::∆*Tn*10* allele from HT115 (38) via P1 transduction into MG1693 and subsequent selection for Tc^r. The presence of the *rnc-14::∆*Tn*10* allele in SK4455 was confirmed by analysis of genomic DNA using PCR (data not shown) and by Western blot analysis with monoclonal anti-RNase III antibodies, which were a gift from D. Court (data not shown).

Growth of bacterial strains and isolation of total RNA

Strains were grown with shaking at 37°C in Luria broth supplemented with thymine (50 µg/ml) until a cell density of approximately 2.5 x 10^8 cfu/ml (60 Klett units above background, No. 42 green filter) was reached. RNA was extracted using the method described by O'Hara *et al.* (40), with the exception that 10%

trimethyl(tetra-decyl)ammonium bromide (Sigma) was used in place of Catrimox-14 (41). RNA was quantified on a NanoDrop™ (Thermo Scientific) apparatus. Five hundred ng of each RNA sample were run on a 1% Agarose-TAE gel and visualized with ethidium bromide to ensure accurate quantities and sufficient quality for further analysis. RNA to be used for microarray hybridization was further treated to remove any contaminating DNA with the DNA-free kit™ (Ambion) and was further analyzed for quality on a Bioanalyzer (Agilent Technologies).

Microarray analysis

Array hybridization, image processing, data normalization and visualization were performed as described in Hiley *et al.* (42). RNA samples were directly labeled using the Ulysses system from Invitrogen. The genome annotations used in this study were obtained from *Escherichia coli* strain MG1655, genome version NC_000913.2, GI:49175990. Small RNA annotation information contained in Table S2 was obtained from the EcoCyc database (43). Fold change numbers derived from the array data were calculated by averaging the normalized $\log_2$ ratios for all probes which either map within the genomic feature, or are within 100 nt of the feature to account for spacer regions. Experimental artifacts due to non-specific hybridization background were minimized in the averages by excluding the highest and lowest 10% of individual signal ratios as outliers. All fold change ratios reported in this study have been converted from $\log_2$ ratios to linear ratios. Greatest change fold ratios from the array data were obtained from PeakFinder (44) average binding ratios

representing the greatest change within a genomic feature. Values represented as negative fold changes in this study are derived by taking the negative inverse of the linear binding ratio to better describe the fold RNA abundance change between mutant and wild-type.

Gene Ontology analysis for changes with p-values ≤0.1 for SK4455 was performed using the Wallenius distribution method, described in Young *et al.* (45) and the gene ontology database (46). P-values were corrected for multiple testing using the Benjamini and Hochberg method (47).

Due to the large amounts of data generated in this study, all array data and electronic files necessary for the viewing and evaluation of the data are available online on the Kushner laboratory website (http://www.genetics.uga.edu/kushner/).

Northern analysis

Northern analysis was performed as described in O'Hara *et al.* (40). In the case of RNA species less than 500 nt in length, 10 µg of total RNA were run on 6% polyacrylamide-8.3 M urea gels. For RNA species larger than 500 nt, 10 µg of total RNA was run on 1.2% agarose/MOPS gels using the protocol of Vincze and Bowra (48). Northern blots were probed with ^{32}P- 5′ end-labeled oligonucleotides (41) and were scanned with a PhosphorImager (Storm^{TM} 840, GE Healthcare). The data were quantified using ImageQuant TL software (GE Healthcare). The Relative Quantities (RQ) reported were normalized based on the loading controls (either 5S or 23S rRNA) determined for each Northern blot.

Primer extension

Primer extension analysis for the *nirB* transcript was performed as described previously (4) with the exception of primers that were specifically designed for *nirB*.

Microarray data and oligonucleotide sequences

All microarray data and oligonucleotide sequences used in this study are available upon request. Oligonucleotide sequences used for the Northern blots in Fig. 4 were as follows: *cspE* (5′-GACGTATCTTACAGAGCGAT-3′); *dnaK* (5′-TCATGTGTTTCGGACCGGTCGCGTCTGCAGTGATGTAT-3′), *ompF* (5′-TCACCGTTACCCTTGGAAAAATAATGCAGACCAACAGCTTT-3′); *ryhB* (probe used was the same sequence as probe EM1 from Massé and Gottesman, 2002 (49) 5′-AAGTAATACTGGAAGCAATGTGAGCAATGTCGTGCTTTCAGGTTCTC-3′); 6S RNA (*ssrS*) (5′-ATATCGGCTCAGGGGACT-3′); 5S rRNA (5'-CTACTCAGGAGAGCGTTCACCG-3') (50); and, 23S rRNA (5'-CGTCCTTCATCGCCTCTGACT-3'). The oligonucleotide sequence for the primer extension reaction on the *nirB* mRNA is 5′-AGCGATGCGCGGTTCTTCAC-3′. The primers used to generate a PCR sequencing template for *nirB* are 5′-TCAGCCGTCACCGTCAGCAT-3′ and 5′-CGCGTTCGCCGACCAGAAC-3′.

RESULTS

Analysis of the *E. coli* transcriptome in the absence of either RNase E or RNase III

The construction of a tiling microarray at 20 nt resolution provided an opportunity to determine the impact of both RNase E and RNase III on every transcript generated from the entire genome. Total steady-state RNA from MG1693 (wild-type), SK3564 (*rne∆1018/rng-219*), and SK4455 (*rnc-14::∆*Tn*10*) was directly labeled with either Cy3 or Cy5 fluorescent dyes and hybridized to custom DNA tiling microarrays (See Materials and Methods). Duplicate arrays were performed with biological replicates and dye swaps to minimize experimental artifacts. Normalized $\log_2$ binding ratios were mapped to the genome and visualized using the Integrated Genome Browser (IGB) (51), while peaks associated with significant $\log_2$ ratios were identified using the PeakFinder program (44).

Initial validation of microarray data

For the initial validation of the microarray data, we examined the steady-state levels of the *rne* and *rng* mRNAs isolated from SK3564 (*rne∆1018/rng-219*). As expected, there was both a significant decrease in the *rne* mRNA along with a concomitant increase in the *rng* transcript (data not shown). As a second internal control, we examined the *rpsO pnp* operon that encodes ribosomal protein S15 and polynucleotide phosphorylase, respectively. It has previously been shown that the *pnp* transcript is significantly stabilized in the absence of RNase III (52-54), while the *rpsO* mRNA is very dependent on RNase E for its degradation

(13,17,22,25). As shown in Fig. 1, there was increased abundance for the *rpsO* mRNA in SK3564 but no significant change in SK4455. In contrast, in the absence of RNase III the *pnp* transcript showed a significant increase in abundance (Fig. 1). Finally, we observed that changes in the steady-state levels of six additional mRNAs (*rpsT, cspE, htpG, glpQ, raiA,* and *trxA*) were in excellent agreement with previously published results obtained from direct measurements of their half-lives in the RNase E mutants (Table 1).

It is important to note that changes in probe signal intensity between mutant and wild-type strains on a tiling microarray reflect a change in steady-state RNA levels and do not necessarily indicate alterations in RNA stability, since changes in transcription could also result in abundance differences. However, if changes in abundance are observed in externally or internally transcribed spacer regions flanking the mature sequences in non-coding RNAs (such as tRNAs, rRNAs and sRNAs), while the abundance of the mature sequences is unchanged, then transcription likely has not been affected and the resulting increases or decreases in RNA levels arise from either stabilization or destabilization of that region of an RNA transcript.

For example, for all seven rRNA operons in *E. coli*, which are known to undergo rapid initial maturation by RNase III to separate the 16S and 23S rRNAs from a larger 30S precursor (55,56), in the absence of RNase III we saw significant increases in the spacer regions of the rRNA operons but not in the abundance of the mature 16S and 23S species (Fig. 2, data not shown). Interestingly, small increases in the spacer regions were also observed in the

absence of RNase E (Fig. 2, data not shown), which may indicate a slight stabilization of the 30S rRNA precursor. Similar increases in the abundance of either intergenic regions, upstream or downstream sequences were also detected for a variety of tRNA transcripts (*pheU, glyW cysT leuZ, argX hisR leuT proM*) (data not shown) that have been shown to be dependent on RNase E for their initial processing (3,19).

Detailed analysis of specific transcripts

Although the preliminary analysis of the array data showed the expected changes in the steady-state levels for some specific mRNAs, we chose a number of additional transcripts to examine in more detail in order to better correlate the changes observed on the arrays with the actual alterations in steady-state RNA levels. For example, the *cspE* mRNA, which encodes a transcription antiterminator and RNA stability regulatory protein, has previously been shown to be more stable in strains carrying a temperature-sensitive allele of RNase E (57-59). Interestingly, our array data showed increased abundance of the *cspE* mRNA in both SK3564 (4.0-fold) and SK4455 (2.2-fold) versus the wild-type control (Fig. 3). When Northern analysis was used to verify the abundance of the *cspE* mRNA in wild-type, SK3564 and SK4455 strains, there were 3.4-fold and 1.3-fold increases in the full-length *cspE* mRNA in the SK3564 and SK4455, respectively (Fig. 4, Table 2).

The *dnaK* mRNA (60), which encodes heat shock protein 70, appeared to be significantly more abundant in both mutants compared to the wild-type control, while the downstream *dnaJ* transcript was more abundant in only the RNase III

strain (Fig. S1). The *dnaK* mRNA is transcribed from three separate promoters to produce a *dnaK dnaJ* dicistronic transcript, which also includes the putative *tpke11* sRNA between the two coding regions (61-63). Northern analysis showed that *dnaK* mRNA was 2.3-fold more abundant in SK3564 and 1.3-fold more abundant in SK4455 than in the wild-type control (Fig. 4, Table 2). The primary band detected on the Northern blot corresponded to a size of 2.1 kb, which contained only the coding region of *dnaK* with additional nucleotides at both the 5′ and 3′ ends. Additional mRNAs that were tested included *yncL* (Fig. S2) and *yncE* (Fig. S3). In both cases, the Northern data confirmed the changes observed on the microarrays (data not shown).

The *ompF* mRNA, which encodes the outer membrane porin F protein, has a single known promoter and is thought to be produced as a monocistronic transcript (64,65). The array data showed that in SK3564 the mRNA was actually less abundant with a fold change of -1.6 (the negative inverse of the linear ratio between mutant and wild type of 0.63. See Materials and Methods) as compared to the wild-type control (Fig. 5). In contrast, in the RNase III mutant, the *ompF* mRNA was 2.0-fold more abundant compared to the wild-type control (Fig. 5). Northern analysis of *ompF* confirmed the array data, showing that there was 2.7-fold more *ompF* in SK4455 and 1.3-fold less in SK3564 than in wild-type control (Fig. 4, Table 2). Although the decrease in the *ompF* mRNA was initially surprising, it has been shown that the *micF* sRNA is a negative regulator of *ompF* stability and is itself regulated by RNase E (66,67). In fact,

there was a significant increase (4.8-fold) in the amount of the *micF* sRNA in the absence of RNase E (Fig. 6).

The sRNA *rhyB* has been shown to be more abundant in both RNase E and RNase III mutant backgrounds (10). These results were confirmed by our array data and subsequent Northern blotting (Figs. S4, 4, Table 2). Also, 6S RNA, a sRNA involved in the regulation of stationary phase transcription, showed an increase in abundance at its 5′ end in SK3564 in both the array data and Northern analysis as compared to very little change in the SK4455 (Fig. S5, 4, Table 2), which was in agreement with previous 6S RNA biogenesis studies (18).

Genome-wide impact of RNase E and RNase III deletions

In order to determine the transcriptome-wide impact of eliminating either RNase E or RNase III, we established a threshold of significance based on our previous comparisons between the array data and published observations, in which the ratio of signal intensity between mutant and wild-type of at least two contiguous oligonucleotide probes (with respect to the genome coordinates, not physical location) had to change by at least ±1.5-fold (linear ratio comparing mutant versus wild-type, not a $\log_2$ ratio) in order for a difference between the mutant and wild-type strain to be considered significant. These criteria for significance, in the vast majority of cases, excluded both experimental noise (due to background non-specific hybridization on the microarray slides) and transcripts that were known to be unaffected in these mutants, from being included in the list of significant changes. We also ensured that our selection thresholds remained suitably sensitive to reliably include transcripts that have been shown to change

in abundance in these mutants, even if the abundance changes were relatively small.

Using the PeakFinder program (44) with the thresholds described above, we generated lists of genomic coordinates that had significantly different abundances of RNA when comparing our mutants to a wild-type control. The list generated using data from SK3564 included 1,520 CDSs (35%) that contained at least one positive peak, indicating higher RNA abundance in the RNase E deletion strain compared to the wild-type control (Figure 7, Table S1). Furthermore, 35 ncRNAs (Table S1) also showed increased in abundance in the RNase E deletion strain. In contrast, 1,096 CDSs (25 %) contained at least one negative peak, along with 12 ncRNAs (Table S1). Also of interest was the observation that in many messages, including the *clpX*, *lon,* and *ppiD* mRNAs, there were differential changes from the 5' to 3' end of each transcript (Fig. S6, data not shown).

Using the same criteria with data from SK4455, 391 CDSs (9%) containing at least one peak of higher abundance and 120 CDS (3%) containing at least one peak of lower abundance were detected (Fig. 7). Furthermore, eleven ncRNAs were more abundant in SK4455 than in the wild-type control, with four additional ncRNAs, showing decreased abundance.

While the overall number of RNA abundance changes between the two ribonuclease mutants was very different, there were a significant number of genomic features which were shared between both the SK3564 and SK4455 lists of affected genes (Table S1). The number of shared RNA abundance changes

within total genomic features included: 10% of the CDSs and 21% of annotated ncRNAs.

To obtain gene expression array profiling information about the overall abundance of a particular genomic feature, the RNA abundance ratios corresponding to each genomic feature that contained a significant change (as determined by our criteria described above) were averaged (see Material and Methods) (Table S1). We observed that many genomic features only had a higher or lower RNA abundance in the mutant versus wild-type in a small region of the feature, which may in fact still affect the functionality of the RNA species. Accordingly, in order to permit a more accurate numerical view of the changes in RNA abundance with these mutants, we also included in Table S1 a greatest fold change value, which was simply the greatest change in average binding ratio within the feature as detected and calculated by PeakFinder (44). The calculation of both the average change and greatest change (along with the location of the greatest change) allows perspective on which genomic features were affected uniformly (where the difference between the average and greatest values is minimal) or non-uniformly (where the difference between the average and greatest values is appreciable) and provides information about which area(s) of a transcript were most highly affected. Additionally, some genomic features contained both a peak of higher RNA abundance and a peak of lower abundance and are so noted in Table S1.

RNase E and RNase III do not significantly affect polycistronic mRNAs

A distinguishing feature of prokaryotic organisms is the presence of significant numbers of polycistronic transcripts. It is widely assumed that larger polycistronic transcripts are rapidly processed into smaller species by endonucleolytic cleavages. Accordingly, it was of considerable interest to determine if the absence of either RNase E or RNase III led to differential effects on individual CDSs within polycistronic transcripts. For example, as shown in Fig. 1 for the dicistronic *rpsO pnp* transcript, loss of RNase E activity led to the stabilization of the complete *rpsO* mRNA, while inactivation of RNase III increased the steady-state level of every oligonucleotide probe specific for *pnp* but not *rpsO*. In fact, for a number of short polycistronic operons (2-3 CDSs), we observed situations where inactivation of RNase E and/or RNase III each led to increased steady-state levels of both CDSs (*groS groL*, Table 3), decreased steady-state levels (*cysD cysN cysC*, Table 3), or as in the case of the *lac* operon, increased steady-state levels in the absence of RNase E but no change in the absence of RNase III (Table 3). However, for the vast majority of larger polycistronic transcripts examined (4-12 CDSs), there was little or no change in the steady-state levels of any of the CDSs within the larger transcripts or increased steady-state levels of the full-length transcripts (Table 3).

Identification of novel RNA transcripts

Microarrays have been used to identify novel sRNAs in microbial genomes by comparing total RNA harvested from wild-type cells grown at various stages of their growth cycle, grown under different stress or nutrient-starvation conditions

or RNA coimmunoprecipitated with antibodies raised against the RNA binding protein Hfq (68-71). Since it has been shown that RNase E is involved in the degradation of some microbial sRNAs (10), we hypothesized that the array using RNA from the RNase E deletion mutant could permit the detection of both annotated and potential new sRNAs as well as other novel transcripts. Consistent with this hypothesis was the fact that 47/63 (75%) of the annotated ncRNAs in the NCBI database (see Materials and Methods) showed at least a ±1.5-fold change in the RNase E deletion mutant relative to the wild-type control (Table S1). This observation highlighted the potential utility of tiling array analysis of ribonuclease mutant strains in screening for novel transcripts.

The array data also facilitated the detection of RNase E-dependent sRNAs under different cell growth conditions. For example, based on published data using wild-type cells, *ryhB* is only detected in either stationary phase or cells that have been grown in minimal medium (72). Yet in exponentially growing cells, the deletion of RNase E resulted in a >20-fold increase in the *ryhB* sRNA compared to the wild-type control (Figs. 4, S2, Table 2).

To extend this analysis, we manually mined the entire genome for unannotated regions that contained at least two adjacent oligonucleotide probes whose abundance was increased by ≥1.5-fold in the RNase E deletion strain compared to the wild-type control, without regard to the location of currently annotated ncRNAs. In this manner, we identified a total of 328 loci (Table S2). Within this list were 37 annotated sRNAs, none of which were placed on the list intentionally as the data was compiled without regard to the location of annotated

sRNAs. Of the remaining newly identified loci, 113 were located in intergenic regions, 156 were antisense to an annotated gene, 11 overlapped with the putative 5′ end of a gene, and 14 overlapped with the putative 3′ end of a gene. Additionally, we detected 74 loci whose abundance decreased in the absence of RNase E by ≤0.67 (the negative inverse of which gives a -1.5-fold change in SK3564 versus wild-type). Of these loci, 11 were located within intergenic regions, 62 were antisense to an annotated gene and none overlapped with either the putative 5′ or 3′ end of a gene. Furthermore, two annotated sRNAs, *isrA* and *fnrS*, fell into this category (Table S2).

Also of interest is candidate 62 in Table S2, which is a putative small coding sequence located on the opposite strand between genes *ompF* and *asnS* (Fig. 5) named C0240 by Tjaden *et al.* (69,73,74). Based on the identification of C0240 in our list of potentially novel transcripts, it is possible that other loci within Table S2 are also small ORFs, which remain to be validated. It should also be noted that the array data from the RNase E deletion mutant were utilized to assist in the development of a computational algorithm for *de novo* identification of non-coding RNAs in prokaryotes (75).

Identification of RNase III cleavage sites based on array data

RNase III cleavage sites, which are associated with certain stem-loop structures, are better characterized than those associated with many other ribonucleases (76). We hypothesized that based on the resolution of the tiling array data we might be able to predict the location of RNase III cleavage sites. For example, there is a sharp change in the RNA abundance ratios between the

rpsO and *pnp* transcripts (Fig. 1), which directly corresponds to the known RNase III cleavage site in the intercistronic region between these two genes (77). Accordingly, we scanned the array data derived from SK4455 for CDSs which showed either uniform increased or decreased abundance in the absence of RNase III. The 5′ UTRs for a number of selected CDSs were then folded using RNAstar (78) to determine if a hairpin could form that might be a suitable recognition sequence for RNase III. Subsequently, primer extension experiments were carried out to map potential RNase III cleavages. In the case of *nirB*, the gene encoding the large subunit of nitrite reductase, the steady-state level of the entire CDS was significantly lower in the absence of RNase III and it appeared that the decrease in abundance mapped closely to the start of translation (Fig. 8A). Surprisingly, primer extension analysis showed an RNase III dependent cleavage 61 nt downstream of the AUG translation start codon (Fig. 8B). The intensity of the extension product corresponding to the putative RNase III cleavage site was shown to be dependent on the amount of RNA loaded in each lane in both the wild-type and RNase E deletion strains (data not shown), indicating that this extension product was not an experimental artifact. The location of this RNase III cleavage site is unusual in that it is the first report of RNase III cutting within a coding sequence in *E. coli.*

RNase III affects the regulation of cysteine biosynthesis

As noted above, in both the RNase E and RNase III arrays there were a significant number of CDSs in which steady-state RNA levels were reduced in the absence of either ribonuclease. In the case of RNase III, a systematic

analysis of the CDSs identified twelve genes involved in cysteine metabolism, four associated with sulfate uptake (*cysP cysU cysW cysA*) and eight involved in *de novo* cysteine biosynthesis (*cysD cysN cysC cysJ cysI cysH cysK cysM*) (data not shown). The twelve genes are organized in three polycistronic and one monocistronic operons (*cysD cysN cysC, cysJ cysI cysH, cysP cysU cysW cysA cysM,* and *cysK*). However, bioinformatic analysis of the leader regions and the first CDS of each transcript did not reveal any significant regulatory sequence motif.

Gene Pathway Analysis

To better understand how inactivation of RNase E and RNase III impacted gene networks and pathways within the cell, we obtained the KEGG pathway annotations (79) for each genomic feature affected in both SK3564 and SK4455. Unfortunately, due to the relatively large number of genes that were affected in the strains, in addition to the large number of genes that remain unannotated in the KEGG pathway database, we were unable to draw any meaningful conclusions about changes in the metabolic and non-metabolic pathways in the cell (data not shown). However, gene ontology (GO) analysis (46) of the data derived from SK4455 revealed that seven GO terms were significantly affected (p value ≤0.1) in the absence of RNase III versus the wild-type control (Table 4). Among the seven GO terms that were changed was the assimilation of sulfate, which was also indicated by the changes in RNA abundance of 12 *cys* genes noted above. The pathway changes outlined in Table 4 also showed that in the absence of RNase III there were significant changes in the heat shock pathway,

iron transport, enterobactin production, the membrane fraction, cytosol, and unfolded protein binding activity. Changes in the membrane components may help to explain a rather odd but reproducible observation that cells deficient in RNase III tend to be more difficult to lyse than wild-type cells, or for that matter most other ribonuclease mutant strains (data not shown).

DISCUSSION

Even though microarrays have been previously used to examine changes in steady-state levels of *E. coli* ORFs (23,80), the data presented here demonstrate that tiling microarrays, based on their 20 nt resolution, provide a much more detailed perspective on the *in vivo* roles of ribonucleases on *E. coli* RNA metabolism. The array data were in excellent agreement with previously published results regarding changes in the steady-state levels of specific mRNAs in either RNase E or RNase III mutants (Figs. 1,3,6, S1 and Table 1). Furthermore, Northern analysis of additional CDSs and sRNAs demonstrated that changes in steady-state RNA levels, either up or down, relative to a wild-type control were accurately reflected by the array data (Figs. 3-5, S1, S4, S5, Table 1).

The RNA abundance changes in the CDSs of SK3564 were largely in agreement with previous CDS expression array studies using a different RNase E mutant (23). However, the mutant used in this study, along with the resolution of the arrays, facilitated a much more comprehensive view of the transcriptome changes in the absence of RNase E. This is particularly important as the transcriptome of RNase III deficient mutants in *E. coli* had not been previously

examined. Additionally, our data suggest that RNase III has a much greater role in the regulation of the transcriptome than previously thought. Furthermore, there also seems to be a considerable level of redundancy, and/or cooperation between RNase E and RNase III activity on a significant number of transcripts, as demonstrated by the 51 CDSs and 9 ncRNAs which were affected by at least ±1.5-fold in both mutants versus the wild-type control (Table S3). Interestingly, for 11 transcripts, their abundance was higher than wild-type in one mutant, but lower than wild-type in the other. However, additional experiments are needed to verify whether the endoribonucleases are involved in the same pathway.

It is important to note that the changes in abundance of many annotated ncRNAs (Tables S1-S2) highlight the complexity associated with interpreting some of the data derived from ribonuclease mutants. Although increases in RNA steady-state levels in the absence of a ribonuclease are easily explained by the loss of degradative capacity, we observed that many CDSs actually showed significantly reduced abundance in the RNase E and RNase III mutants compared to wild-type control. It is likely that some of these decreases in mRNA abundance were due to changes in the amount of regulatory RNAs in the mutant strains, as demonstrated by the decrease in abundance of *ompF* mRNA in SK3564 (Fig. 5). In this case, an explanation for the decrease in abundance in SK3564 versus wild-type may be related to the increase in the steady-state level of the *micF* sRNA (Fig. 6), which is known to regulate the stability of the *ompF* mRNA. Similarly, we hypothesize that for many other mRNAs changes in sRNA

abundance related to inactivation of either RNase E or RNase III indirectly leads to reduced steady-state levels of specific target mRNAs.

Thus, it seems likely that many of the changes observed, with both the RNase E and RNase III arrays, are secondary effects that can be associated with changes in the steady-state levels of various sRNAs or other unidentified regulatory factors. Since some of the annotated sRNAs have already been shown to have multiple targets (49), when coupled with the likelihood of additional sRNAs (Table S2), it seems quite probable that a significant portion of the *E. coli* transcriptome is affected by the abundance of sRNAs.

It is also important to note that there were some cases where increased abundance of a specific sRNA in either the RNase E or RNase III mutant was not accompanied by a decrease in the level of its known mRNA target. There could be a number of possible reasons for this type of result. The most likely explanation for this type of observation is that the abundance of the target mRNA is relatively low such that a significant increase in sRNA level would have little effect on its abundance. It is also possible that the sRNAs affected in either mutant are dependent on the missing ribonuclease for their processing. In the absence of processing, the full-length sRNA may not be functionally active. Additionally, it is possible that under conditions where there is a significant increase in the levels of a large number of sRNAs compared to wild type cells, that the amount of Hfq [an RNA binding protein required for the function of many sRNAs (70)] may be rate-limiting. If this were the case, changes in sRNA levels would not be expected to efficiently affect mRNA target levels.

Overall, our observations provide further support for the notion that the control of RNA stability is in fact a very important aspect of gene regulation in *E. coli*. Our results also re-emphasize that the changes in steady-state levels of a particular RNA species in the absence of a particular ribonuclease is not adequate proof that the enzyme is directly involved in its processing, maturation or decay. Although our data show that the analysis of ribonuclease mutants can be challenging, they also highlight the sophistication and complexity of post-transcriptional regulation with regard to gene expression in *E. coli* and most likely in other prokaryotes.

Another important finding regarding sRNAs in *E. coli* relates to the fact that even though the RNA used in this study was isolated from exponentially growing cells, we were able to detect changes in the steady-state levels of 75% of the annotated ncRNAs in SK3564, and 22% in SK4455 (Table S1). This fact is of considerable interest because many of these sRNA species have been shown to be present primarily in either early or late stationary phase cultures. Thus, the arrays will be useful for examining the effects of ribonuclease mutants on the transcriptome during all phases of growth. For example, manual analysis of the data indicated the possibility of 363 additional novel transcripts (excluding annotated sRNAs and a putative ORF, Table S2).

Another somewhat surprising observation was the failure to observe significant changes in the steady-state levels of large polycistronic mRNAs (Table 3). It has been assumed that RNase E is responsible for separating polycistronic operons into smaller but functional mRNAs that have differential

half-lives. While this notion may be true for a small fraction of polycistronic operons, the more likely scenario is that RNase E is not involved in establishing differential intra-operon half-lives. In fact, mRNAs within large operons may be grouped in such a way to guarantee equal mRNA copy numbers. Such a hypothesis is already widely accepted when dealing with mRNAs encoding ribosomal proteins, and may apply to other classes of mRNAs as well. This could be especially true for mRNAs which together encode multi-subunit enzymes which require the production of each subunit in stoichiometric amounts.

The analysis of the array data also produced some rather unexpected observations. For example, the array showed that the *nirB* mRNA was less abundant in SK4455 than the wild-type control (Fig. 8A). However, when we tested the transcript for an RNase III cleavage site because of a predicted hairpin within the 5′ UTR, we detected a strong RNase III cleavage site (Fig. 8B), but to our surprise the cleavage occurred at a site well within the coding sequence that would functionally inactivate the mRNA. Since functionally inactive mRNAs tend to have shorter half-lives because they are no longer being translated, it is not clear at this time why inactivating RNase III actually led to reduced abundance of the *nirB* message. It is possible that the *nirB* observations are another case of indirect effects, since there was also reduced abundance of the transcript in the absence of RNase E (Fig. 8A).

Another example of potential indirect effects associated with removing an important ribonuclease involved the decreased steady-state levels of the 12 genes associated with cysteine metabolism in the absence of RNase III. These

12 cysteine related genes are found in four different operons (three are polycistronic), and none of them seem to have an RNase III cleavage site within their 5′ or 3′ UTRs. It is thus likely in this case that the inactivation of RNase III leads to changes of a common factor (either a protein or perhaps a sRNA) that coordinately regulates the expression of these genes, rather than direct RNase III cleavages of the mRNAs themselves.

Additionally, while the mapping of promoter initiation sites in *E. coli* has been performed using a number of highly effective techniques (81,82), because of the plethora of transcriptome changes in the ribonuclease mutants, we suspect that tiling microarray data could allow for the detection of approximate locations of alternative transcription start sites within the transcriptome. By examining the array data near the 5′ ends of a few messages with experimentally verified transcription start sites (data not shown), we found that we could accurately predict the known start sites within a 60 nt window. We believe that our data may prove valuable as a preliminary screen for groups interested in genome-wide promoter usage determination.

It is already clear from the data presented here that the post-transcriptional regulation of gene expression in *E. coli* is far more complex than previously envisioned. As the transcriptomes of more ribonuclease mutants are analyzed, a more detailed picture of the complexity within transcriptome regulatory networks will begin to emerge and may begin to answer some of the questions which have plagued the prokaryotic RNA field since its inception. It should also be noted that in the time since our tiling microarray data sets were

generated, the RNAseq method of transcriptome analysis has been published and widely adopted. While RNAseq is a powerful method for transcriptome analysis (83), the associated costs are prohibitive for many researchers and therefore tiling microarray technology will continue to have a place in the analysis of many transcriptomes.

Finally, it is important to note that when we employed Peakfinder (40) to detect areas of significant changes, we used thresholds, which in the vast majority of cases, excluded both experimental noise (due to background non-specific hybridization on the microarray slides) and transcripts that were known to be unaffected in these mutants, from being included in the list of significant changes. We also ensured that our selection thresholds remained suitably sensitive to reliably include transcripts that have been shown to change in abundance in these mutants, even if the abundance changes were relatively small. Although it is possible that some of the 1.5-fold changes may not be biologically relevant, several such changes were validated using Northern blot analysis (Table 2 and data not shown). In fact, any change observed on a tiling array needs to be verified by an independent method of analysis. Thus, based on the nature of tiling array technology, no set of criteria will be perfect and the data reported here probably contain a very limited number of both false-positive and false-negative results. For example, regions of the genome that are repetitive such as related tRNA operons, may have some level of cross hybridization. Therefore the use of tiling array data for these types of genomic features has the potential to be misleading and has not been a focus of this

study. In spite of these caveats, the data from this study have allowed for a remarkably detailed overview of the transcriptome-wide impacts caused by the absence of either RNase E or RNase III, and will prove to be of value to anyone interested in these enzymes or the affects they have on specific transcripts.

FUNDING

This work was supported in part by grants from the National Institutes of Health (GM57220 and GM81554) to S.R.K. M.B.S. was supported in part by a NIH predoctoral traineeship from NIGMS (GM07106).

ACKNOWLEDGEMENTS

The authors would like to thank B. Mohanty for his technical expertise and D. Court for his gift of strains and antibodies.

REFERENCES

1. Srivastava, A.K. and Schlessinger, D. (1990) Mechanisms and regulation of bacterial ribosomal RNA processing. *Ann. Rev. Microbiol.*, **44**, 105-129.
2. Li, Z. and Deutscher, M.P. (1996) Maturation pathways for *E.coli* tRNA precursors: a random multienzyme process *in vivo*. *Cell*, **86**, 503-512.
3. Ow, M.C. and Kushner, S.R. (2002) Initiation of tRNA maturation by RNase E is essential for cell viability in *E. coli*. *Genes Dev.*, **16**, 1102-1115.
4. Mohanty, B.K. and Kushner, S.R. (2007) Ribonuclease P processes polycistronic tRNA transcripts in *Escherichia coli* independent of ribonuclease E. *Nucleic Acids Res.*, **35**, 7614-7625.

5. Mohanty, B.K. and Kushner, S.R. (2008) Rho-independent transcription terminators inhibit RNase P processing of the *secG leuU* and *metT* tRNA polycistronic transcripts in *Escherichia coli*. *Nucleic Acids Res.*, **36**, 364-375.

6. Mohanty, B.K. and Kushner, S.R. (2010) Processing of the *Escherichia coli leuX* tRNA transcript, encoding $tRNA^{leu5}$, requires either the 3'-5' exoribonuclease polynucleotide phosphorylase or RNase P to remove the Rho-independent transcription terminator. *Nucleic Acids Res.*, **38**, 5306-5318.

7. Deutscher, M.P. (2006) Degradation of RNA in bacteria: comparison of mRNA and stable RNA. *Nucleic Acids Res.*, **34**, 659-666.

8. Kushner, S.R. (2007) In Böck, A., Curtis III., R., Gross, C. A., Kaper, J. B., Neidhardt, F. C., Nyström, T., Rudd, K. E. and L., S. C. (eds.), *Escherichia coli and Salmonella: cellular and molecular biology*. American Society for Microbiology Press, Washington, DC.

9. Lundberg, U. and Altman, S. (1995) Processing of the precursor to the catalytic RNA subunit of RNase P from *Escherichia coli*. *RNA*, **1**, 327-334.

10. Masse, R., Escorcia, F.E. and Gottesman, S. (2003) Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes Dev.*, **17**, 2374-2383.

11. Li, Z., Pandit, S. and Deutscher, M.P. (1998) 3' Exoribonucleolytic trimming is a common feature of the maturation of small, stable RNAs in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA*, **95**, 2856-2861.

12. Carpousis, A.J. (2007) The RNA degradosome of *Escherichia coli*: An mRNA-dgrading machine assembled on RNase E. *Ann. Rev. Microbiol.*, **61**, 71-87.

13. Ow, M.C., Liu, Q. and Kushner, S.R. (2000) Analysis of mRNA decay and rRNA processing in *Escherichia coli* in the absence of RNase E-based degradosome assembly. *Mol. Microbiol.*, **38**, 854-866.

14. Ono, M. and Kuwano, M. (1979) A conditional lethal mutation in an *Escherichia coli* strain with a longer chemical lifetime of mRNA. *J. Mol. Biol.*, **129**, 343-357.

15. Arraiano, C.M., Yancey, S.D. and Kushner, S.R. (1988) Stabilization of discrete mRNA breakdown products in ams pnp rnb multiple mutants of *Escherichia coli* K-12. *J. Bacteriol.*, **170**, 4625-4633.

16. Mackie, G.A. (1991) Specific endonucleolytic cleavage of the mRNA for ribosomal protein S20 of *Escherichia coli* requires the products of the *ams* gene *in vivo* and *in vitro*. *J. Bacteriol.*, **173**, 2488-2497.

17. Regnier, P. and Hajnsdorf, E. (1991) Decay of mRNA encoding ribosomal protein S15 of *Escherichia coli* is initiated by an RNase E-dependent endonucleolytic cleavage that removes the 3' stabilizing stem and loop structure. *J. Mol. Biol.*, **187**, 23-32.

18. Kim, K.S. and Lee, Y. (2004) Regulation of 6S RNA biogenesis by switching utilization of both sigma factors and endoribonucleases. *Nucleic Acids Res.*, **32**, 6057-6068.

19. Li, Z. and Deutscher, M.P. (2002) RNase E plays an essential role in the maturation of *Escherichia coli* tRNA precursors. *RNA*, **8**, 97-109.

20. Ghora, B.K. and Apirion, D. (1978) Structural analysis and *in vitro* processing to p5 rRNA of a 9S RNA molecule isolated from an *rne* mutant of *E. coli*. *Cell*, **15**, 1055-1066.

21. Li, Z., Pandit, S. and Deutscher, M.P. (1999) RNase G (CafA protein) and RNase E are both required for the 5' maturation of 16S ribosomal RNA. *EMBO J.*, **18**, 2878-2885.

22. Ow, M.C., Perwez, T. and Kushner, S.R. (2003) RNase G of *Escherichia coli* exhibits only limited functional overlap with its essential homologue, RNase E. *Mol. Microbiol.*, **49**, 607-622.

23. Lee, K., Bernstein, J.A. and Cohen, S.N. (2002) RNase G complementation of *rne* null mutation identified functional interrelationships with RNase E in *Escherichia coli*. *Mol. Microbiol.*, **43**, 1445-1456.

24. Wachi, M., Umitsuki, G., Shimizu, M., Takada, A. and Nagai, K. (1999) *Escherichia coli cafA* gene encodes a novel RNase, designated as RNase G, involved in processing of the 5' end of 16S rRNA. *Biochem. Biophys. Res. Comm.*, **259**, 483-488.

25. Chung, D.-H., Min, Z., Wang, B.-C. and Kushner, S.R. (2010) Single amino acid changes in the predicted RNase H domain of *E. coli* RNase G lead to the complementation of RNase E mutants. *RNA*, **16**, 1371-1385.

26. Dunn, J.J. and Studier, F.W. (1973) T7 early RNAs and *Escherichia coli* ribosomal RNAs are cut from large precursor RNAs *in vivo* by ribonuclease III. *Proc. Natl. Acad. Sci. USA*, **70**, 3296-3300.

27. Matsunaga, J., Simons, E.L. and Simons, R.W. (1996) RNase III autoregulation: Structure and function of *rncO*, the posttranscriptional "operator". *RNA*, **2**, 1228-1240.

28. Drider, D. and Condon, C. (2004) The continuing story of endoribonuclease III. *J Mol. Microbiol. Biotechnol.*, **8**, 195-200.

29. Nicholson, A. (1999) Function, mechanism and regulation of bacterial ribonucleases. *FEMS Microbiol. Rev.*, **23**, 371-390.

30. Babitzke, P., Granger, L., Olszewski, J. and Kushner, S.R. (1993) Analysis of mRNA decay and rRNA processing in *Escherichia coli* multiple mutants carrying a deletion in RNase III. *J. Bacteriol.*, **175**, 229-239.

31. Bardwell, J.C.A., Regnier, P., Chen, S.-M., Nakamura, Y., Grunberg-Manago, M. and Court, D.L. (1989) Autoregulation of RNase III operon by mRNA processing. *EMBO J.*, **8**, 3401-3407.

32. Talkad, V., Achord, D. and Kennell, D. (1978) Altered mRNA metabolism in ribonuclease III-deficient strains of *Escherichia coli*. *J. Bacteriol.*, **135**, 528-541.

33. Afonyushkin, T., Vecerek, B., Moll, I., Blasi, U. and Kaberdin, V.R. (2005) Both RNase E and RNase III control the stability of *sodB* mRNA upon translational inhibition by the small regulatory RNA RyhB. *Nucleic Acids Res.*, **33**, 1678-1689.

34. Resch, A., Afonyushkin, T., Lombo, T.B., McDowall, K.J., Blasi, U. and Kaberdin, V.R. (2008) Translational activation by the noncoding RNA DsrA involves alternative RNase III processing in the rpoS 5'-leader. *RNA*, **14**, 454-459.

35. Khodursky, A.B., Peter, B.J., Cozzarelli, N.R., Botstein, D., Brown, P.O., and Yanofsky, C. (2000) DNA microarray analysis of gene expression in response to physiological and genetic changes that affect tryptophan metabolism in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA*, **97**, 12170-12175.

36. Zheng, M., Wang, X., Templeton, L.J., Smulski, D.R., LaRossa, R.A. and Storz, G. (2001) DNA microarray-mediated transcriptional profiling of the *Escherichia coli* response to hydrogen peroxide. *J. Bacteriol.*, **183**, 4562-4570.

37. Mohanty, B.K. and Kushner, S.R. (2003) Genomic analysis in *Escherichia coli* demonstrates differential roles for polynucleotide phosphorylase and RNase II in mRNA abundance and decay. *Mol. Microbiol.*, **50**, 645-658.

38. Takiff, H.E., Chen, S. and Court, D.L. (1989) Genetic analysis of the *rnc* operon of *Escherichia coli*. *J. Bacteriol.*, **171**, 2581-2590.

39. Wang, X. and Seed, B. (2003) Selection of oligonucleotide probes for protein coding sequences. *Bioinformatics*, **19**, 796-802.

40. O'Hara, E.B., Chekanova, J.A., Ingle, C.A., Kushner, Z.R., Peters, E. and Kushner, S.R. (1995) Polyadenylylation helps regulate mRNA decay in *Escherichia coli*. *Proc. Natl. Acad. Sci. U S A*, **92**, 1807-1811.

41. Mohanty, B.K., Giladi, H., Maples, V.F. and Kushner, S.R. (2008) Analysis of RNA decay, processing, and polyadenylation in *Escherichia coli* and other prokaryotes. *Meth. Enzymol.*, **447**, 3-29.

42. Hiley, S.L., Babak, T. and Hughes, T.R. (2005) Global analysis of yeast RNA processing indentifies new targets of RNase III and uncovers a link between tRNA 5' end processing and tRNA splicing. *Nucleic Acids. Res.*, **33**, 3048-3056.

43. Keseler, I.M., Bonavides-Martinez, C., Collado-Vides, J., Gama-Castro, S., Gunsalus, R.P., Johnson, D.A., Krummenacker, M., Nolan, L.M., Paley, S., Paulsen, I.T. *et al.* (2009) EcoCyc: a comprehensive view of *Escherichia coli* biology. *Nucleic Acids Res.*, **37**, D464-470.

44. Glynn, E.F., Megee, P.C., Yu, H.G., Mistrot, C., Unal, E., Koshland, D.E., DeRisi, J.L. and Gerton, J.L. (2004) Genome-wide mapping of the cohesin complex in the yeast *Saccharomyces cerevisiae*. *PLoS Biol*, **2**, E259.

45. Young, M.D., Wakefield, M.J., Smyth, G.K. and Oshlack, A. (2010) Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.*, **11**, R14.

46. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T. *et al.* (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.*, **25**, 25-29.

47. Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J. Royal Stat. Soc. Ser. B-Method*, **57**, 289-300.

48. Vincze, E. and Bowra, S. (2005) Northerns revisited: a protocol that eliminates formaldehyde from the gel while enhancing resolution and sensitivity. *Anal. Biochem.*, **342**, 356-357.

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

50. Babitzke, P. and Kushner, S.R. (1991) The Ams (altered mRNA stability) protein and ribonuclease E are encoded by the same structural gene of *Escherichia coli*. *Proc Natl Acad Sci USA*, **88**, 1-5.

51. Nicol, J.W., Helt, G.A., Blanchard, S.G., Jr., Raja, A. and Loraine, A.E. (2009) The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. *Bioinformatics*, **25**, 2730-2731.

52. Beran, R.K. and Simons, R.W. (2001) Cold-temperature induction of *Escherichia coli* polynucleotide phosphorylase by reversal of its autoregulation. *Mol. Microbiol.*, **39**, 112-125.

53. Robert-Le Meur, M. and Portier, C. (1994) Polynucleotide phosphorylase of *Escherichia coli* induces the degradation of its RNase III processed messenger by preventing its translation. *Nucleic Acids Res.*, **22**, 397-403.

54. Robert-Le Meur, M. and Portier, C. (1992) *Escherichia coli* polynucleotide phosphorylase expression is autoregulated through an RNase III-dependent mechanism. *EMBO J.*, **11**, 2633-2641.

55. King, T.C., Sirdeshmukh, R. and Schlessinger, D. (1984) RNase III cleavage is obligate for maturation but not for function of *Escherichia coli* pre-23S rRNA. *Proc. Natl. Acad. Sci. USA*, **81**, 185-188.

56. Srivastava, A.K. and Schlessinger, D. (1989) Processing pathway of *Escherichia coli* 16S precursor rRNA. *Nucleic Acids Res.*, **17**, 1649-1663.

57. Uppal, S., Akkipeddi, V.S. and Jawali, N. (2008) Posttranscriptional regulation of *cspE* in *Escherichia coli*: involvement of the short 5'-untranslated region. *FEMS Microbiol. Lett.*, **279**, 83-91.

58. Perwez, T. and Kushner, S.R. (2006) RNase Z in *Escherichia coli* plays a significant role in mRNA decay. *Mol. Microbiol.*, **60**, 723-737.

59. Bae, W.H., Xia, B., Inouye, M. and Severinov, K. (2000) *Escherichia coli* CspA-family RNA chaperones are transcription antiterminators. *Proc. Natl. Acad. Sci. USA*, **97**, 7784-7789.

60. Mogk, A., Deuerling, E., Vorderwulbecke, S., Vierling, E. and Bukau, B. (2003) Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation. *Mol. Microbiol.*, **50**, 585-595.

61. Wade, J.T., Roa, D.C., Grainger, D.C., Hurd, D., Busby, S.J., Struhl, K. and Nudler, E. (2006) Extensive functional overlap between sigma factors in *Escherichia coli*. *Nat. Struct. Mol. Biol.*, **13**, 806-814.

62. Zhou, Y.N., Kusukawa, N., Erickson, J.W., Gross, C.A. and Yura, T. (1988) Isolation and characterization of *Escherichia coli* mutants that lack the heat shock sigma factor sigma 32. *J. Bacteriol.*, **170**, 3640-3649.

63. Cowing, D.W., Bardwell, J.C., Craig, E.A., Woolford, C., Hendrix, R.W. and Gross, C.A. (1985) Consensus sequence for *Escherichia coli* heat shock gene promoters. *Proc. Natl. Acad. Sci. USA*, **82**, 2679-2683.

64. Cowan, S.W., Schirmer, T., Rummel, G., Steiert, M., Ghosh, R., Pauptit, R.A., Jansonius, J.N. and Rosenbusch, J.P. (1992) Crystal structures explain functional properties of two *E. coli* porins. *Nature*, **358**, 727-733.

65. Mitchell, J.E., Zheng, D., Busby, S.J. and Minchin, S.D. (2003) Identification and analysis of 'extended -10' promoters in *Escherichia coli*. *Nucleic Acids Res.*, **31**, 4689-4695.

66. Andersen, J., Forst, S.A., Zhao, K., Inouye, M. and Delihas, N. (1989) 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.

67. Schmidt, M. and Delihas, N. (1995) micF RNA is a substrate for RNase E. *FEMS Microbiol. Lett.*, **133**, 209-213.

68. Landt, S.G., Abeliuk, E., McGrath, P.T., Lesley, J.A., McAdams, H.H. and Shapiro, L. (2008) Small non-coding RNAs in *Caulobacter crescentus*. *Mol. Microbiol.*, **68**, 600-614.

69. Tjaden, B., Saxena, R.M., Stolyar, S., Haynor, D.R., Kolker, E. and Rosenow, C. (2002) Transcriptome analysis of *Escherichia coli* using high-density oligonucleotide probe arrays. *Nucleic Acids Res.*, **30**, 3732-3738.

70. Zhang, A., Wassarman, K.M., Rosenow, C., Tjaden, B.C., Storz, G. and Gottesman, S. (2003) Global analysis of small RNA and mRNA targets of Hfq. *Mol. Microbiol.*, **50**, 1111-1124.

71. Hu, Z., Zhang, A., Storz, G., Gottesman, S. and Leppla, S.H. (2006) An antibody-based microarray assay for small RNA detection. *Nucleic Acids Res.*, **34**, e52.

72. Argaman, L., Hershberg, R., Vogel, J., Bejerano, G., Wagner, E.G., Margalit, H. and Altuvia, S. (2001) Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr. Biol.*, **11**, 941-950.

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

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

75. Tran, T.T., Zhou, F., Marshburn, S., Stead, M., Kushner, S.R. and Xu, Y. (2009) De Novo Computational Prediction of Non-coding RNA Genes in Prokaryotic Genomes. *Bioinformatics*, **25**, 2897-2905.

76. Pertzev, A.V. and Nicholson, A.W. (2006) Characteization of RNA sequence determinants and antideterminants of processing reactivity for a minimal substrate of *Escherichia coli* ribonuclease III. *Nucleic Acids Res.*, **34**, 3708-3721.

77. Regnier, P. and Portier, C. (1986) Initiation attenuation and RNase III processing of transcripts from the *Escherichia coli* operon encoding ribosomal protein S15 and polynucleotide phosphorylase. *J. Mol. Biol.*, **187**, 23-32.

78. Gultyaev, A.P., van Batenburg, F.H.D. and Pleij, C.W.A. (1995) The computer simulation of RNA folding pathways using a genetic algorithm. *J. Mol. Biol.*, **250**, 37-51.

79. Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K.F., Itoh, M., Kawashima, S., Katayama, T., Araki, M. and Hirakawa, M. (2006) From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res*, **34**, D354-357.

80. Bernstein, J.A., Lin, P.-H., Cohen, S.N. and Lin-Chao, S. (2004) Global analysis of *Escherichia coli* RNA degradosome function using DNA microarrays. *Proc. Natl. Acad. Sci. USA*, **101**, 2748-2763.

81. Gordon, J.J., Towsey, M.W., Hogan, J.M., Mathews, S.A. and Timms, P. (2006) Improved prediction of bacterial transcription start sites. *Bioinformatics*, **22**, 142-148.

82. Mendoza-Vargas, A., Olvera, L., Olvera, M., Grande, R., Vega-Alvarado, L., Taboada, B., Jimenez-Jacinto, V., Salgado, H., Juarez, K., Contreras-Moreira, B. *et al.* (2009) Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in *E. coli. PLoS One*, **4**, e7526.

83. van Bakel, H., Nislow, C., Blencowe, B.J. and Hughes, T.R. (2010) Most "dark matter" transcripts are associated with known genes. *PLoS Biol*, **8**, e1000371.

84. Suzuki, K., Babitzke, P., Kushner, S.R. and Romeo, T. (2006) Identification of a novel regulatory protein (CsrD) that targets the global regulatory RNAs CsrB and CsrC for degradation by RNase E. *Genes Dev.*, **20**, 2605-2617.

Table 1. Comparison of array data with published experiments that employed various

RNase E mutants.

Gene/Operon	Published observation in absence of RNase E	Observed change in steady-state RNA levels on array	Reference
rpsT	Increased half-life	Increased abundance of transcript	(22,58)
rpsO	Increased half-life	Increased abundance of transcript	(13,22)
cspE	Increased half-life	Increased abundance of transcript	(58)
htpG	Increased half-life	Increased abundance of transcript	(58)
glpQ	Increased half-life	Increased abundance of transcript	(58)
raiA	No change in half-life	No change in transcript abundance	(22)
trxA	No change in half-life	No change in transcript abundance	(22)
csrC	Increased half-life	Increased abundance	(84)

		of transcript	
rrl rrs	Increased stability of spacer region between 23S and 5S sequences	Increased abundance of spacer region between mature 23S and 5S rRNA sequences	(22)
pheU	Increased abundance of full-length tRNA precursor	Increased abundance of terminator region	(3)
secG leuU	Increased half-life of *secG* mRNA and stabilization of *leuU* terminator region	Increased abundance of *secG* mRNA and *leuU* terminator region	(5)
glyW cysT leuZ	Increased abundance of polycistronic transcript	Increased abundance of spacer regions	(3,19)
tyrT tyrV tpr	Increased abundance of polycistronic transcript	Increased abundance of spacer regions	(3,19)
metT leuW glnU glnW metU glnV	Increased abundance of polycistronic transcript	Increased abundance of spacer regions	(5)

glnX			
argX hisR leuT proM	Increased abundance of polycistronic transcript	Increased abundance of spacer regions	(3,19)
valV valW	No change in abundance of polycistronic transcript	No change in abundance of polycistronic transcript	(4)
leuQ leuP leuV	No change in abundance of polycistronic transcript	No change in abundance of polycistronic transcript	(4)
metV metW metZ	No change in abundance of polycistronic transcript	No change in abundance of polycistronic transcript	(4)

In the case of mRNAs, published data relates to observed changes in individual half-lives. For tRNA transcripts, the published data relate to whether inactivation of RNase E led to changes in the levels of the full-length mono- or polycistronic transcripts.

Table 2. Comparison of transcript abundance between wild-type and mutant strains using array and Northern data.

Gene	Mutant/wild-type	Fold change from array	Fold change from Northern analysis
cspE	ΔRNase E	4.0	3.4
	ΔRNase III	2.2	1.3
dnaK	ΔRNase E	2.0	2.3
	ΔRNase III	2.0	1.3
ompF	ΔRNase E	-1.6	-1.3
	ΔRNase III	2.0	2.7
ryhB	ΔRNase E	7.6	20.4
	ΔRNase III	5.4	3.7
ssrS	ΔRNase E	ND	-1.6
	ΔRNase III	ND	1.1

Positive values indicate higher abundance in the mutant, while negative values indicate higher abundance in wild-type. Fold changes from the array data were calculated by averaging the ratios of every oligonucleotide probe across the genomic feature (see Materials and Methods). Fold change numbers from the Northern analyses were calculated as described in the Materials and Methods. ND, not determined due to the nature of the abundance change of *ssrS* (see Fig. S6).

Table 3. Changes in steady-state levels of individual ORFs within representative polycistronic transcripts in the absence of either RNase E[1] or RNase III[2].

Operon
↑ ↑ *groS groL* ↑ ↑
→ ↑ *nlpD rpoS* → ↑
↓ ↓ *tnaA tnaB* ↑ ↑
↑ ↑ ↑ *lacZ lacY lacA* → → →
↓ ↓ ↓ *cysD cysN cysC* ↓ ↓ ↓
→ → → → *bioB bioF bioC bioD* → → → →
→ → → → → *nsrR rnr rlmB yjfI yjfJ* → → → → →
→ → → → → *dgoR dgoK dgoA dgoD dgoT* → → → → →
→ → → → → *trpA trpB trpC trpD trpE* → → → → →
→ → → → → → → → → *flgB flgC flgD flgE flgF flgG flgH flgI flgJ* → → → → → → → → →
↑ → → → → → → → → *atpI atpB atpE atpF atpH atpA atpG atpD atpC* ↑ → → → → → → → →

→ → → → → → → → → → → *hyfA hyfB hyfC hyfD hyfE hyfF hyfG hyfH hyfI hyfJ hyfR* → → → → → → → → → → →
→ → → → → → → → → → → *rpsJ rplC rplD rplW rplB rpsS rplV rpsC rplP rpmC rpsQ* → → → → → → → → → → →
↑ ↑ ↑ ↑ ↑ ↑ ↑ ↑ ↑ → → *cydA cydB ygbT ygbE ygbC tolQ tolR tolA tolB pal ygbF* → → → → → → → → → → →

Arrows over the gene name indicate the steady-state abundance of the mRNA in the absence of RNase E versus wild-type, while arrows under the gene name indicate the abundance changes in the absence of RNase III versus wild-type. Upward arrows indicated increased steady-state levels in the mutant compared to the wild-type control. Downward arrows indicated decreased steady-state levels in the mutant versus the wild-type control. Horizontal arrows mean no significant change in the steady-state RNA levels between the mutant and wild-type control. In order to be included in this table, every oligonucleotide probe for a particular ORF had to be significantly increased, decreased or unchanged.

[1]*rneΔ1018/rng-219.* [2]*rnc-14::Δ*Tn*10*.

Table 4. Gene ontology pathways affected in the absence of RNase III. The following pathways are enriched for changes in SK4455 (*rnc-14*) versus wild-type at p-values ≤0.1 (see Materials and Methods).

Ontology Type	GO Term Affected	Definition of GO Term	P-Value	GO ID
Biological Process	Response to heat	A change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.)	0.0007	GO:0009408
Biological Process	Iron ion transport	The directed movement of iron (Fe) ions into, out of, within or between cells by means of some external agent such as a transporter	0.007	GO:0006826
Biological Process	Enterobactin biosynthetic process	The chemical reactions and pathways resulting in the formation of enterobactin, a catechol-derived siderochrome of Enterobacteria; 2,3-dihydroxy-N-benzoyl-L-serine and a product of the shikimate pathway.	0.03	GO:0009239
Cellular Component	Membrane fraction	That fraction of cells, prepared by disruptive biochemical methods,	0.03	GO:0005624

		that includes the plasma and other membranes.		
Cellular Component	Cytosol	The part of the cytoplasm that does not contain organelles but which does contain other particulate matter, such as protein	0.03	GO:0005829
Molecular Function	Unfolded protein binding	Interacting selectively and non-covalently with an unfolded protein.	0.07	GO:0051082
Molecular Function	Sulfate assimilation	The pathways by which inorganic sulfate is processed and incorporated into sulfated compounds.	0.08	GO:0000103

Fig. 1. Microarray data for the *rpsO pnp* operon.

Changes in the steady-state levels of the *rpsO* and *pnp* mRNAs in RNase E and RNase III deletion mutants. The image presented was obtained from a screen shot of the Integrated Genome Browser program (51). Labels for the data displayed are located on the far left of the image and identify all features within the same horizontal plane. Gene names appear above or below the horizontal bar indicating their location on the genome relative to actual nucleotide coordinates, which are displayed at the center of the graph. Objects above the genome coordinate line are on the forward strand, while objects below the coordinate line are on the reverse strand. Black arrows indicate the direction of transcription. CDSs are colored blue, while red indicates a sRNA. The actual array data are displayed as a series of vertical lines representing the $\log_2$ ratio of fluorescence between the mutant and wild-type strains along a horizontal that intersects the label on the far left designating the mutant strains and strand (+ or -). The horizontal line in the array data is equal to a $\log_2$ ratio of 0, with vertical lines going above or below the baseline representing changes in the $\log_2$ ratio of greater or less than 0 for each probe. Vertical lines above the baseline indicate higher RNA abundance in the mutant versus wild-type, while line extending below the baseline indicate lower RNA abundance in the mutant versus wild-type. Maximum and minimum peak heights displayed are equivalent to $\log_2$ ratios of $\pm$ 3.

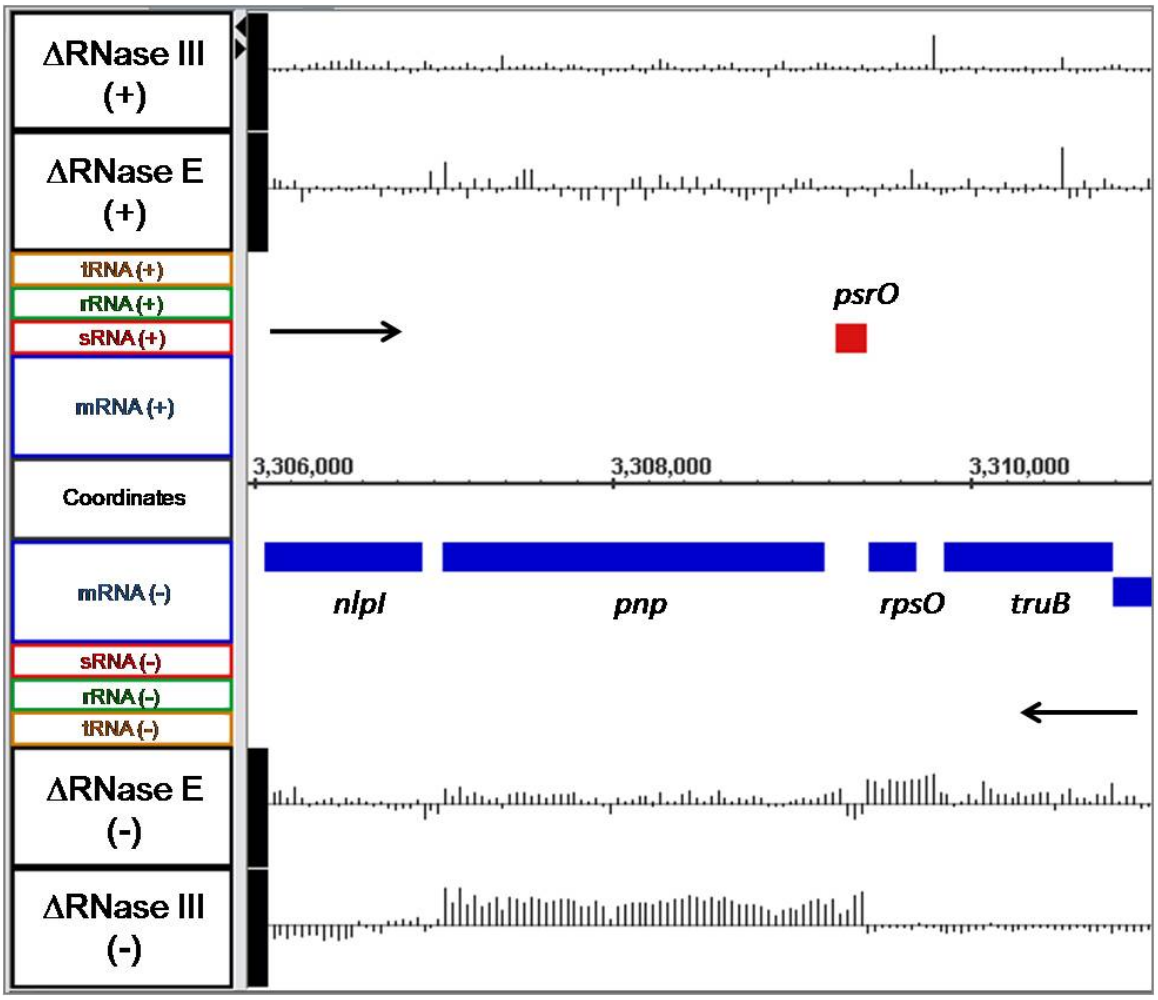

ΔRNase III (+)
ΔRNase E (+)
tRNA (+)
rRNA (+)
sRNA (+)
mRNA (+)
Coordinates
mRNA (-)
sRNA (-)
rRNA (-)
tRNA (-)
ΔRNase E (-)
ΔRNase III (-)
psrO
3,306,000
3,308,000
3,310,000
nlpI
pnp
rpsO
truB

Fig. 2. Microarray data for the *rrnE* ribosomal RNA operon.

Data is presented as described in Fig. 1. Green bars represent rRNA genes, while orange bars indicate tRNA genes.

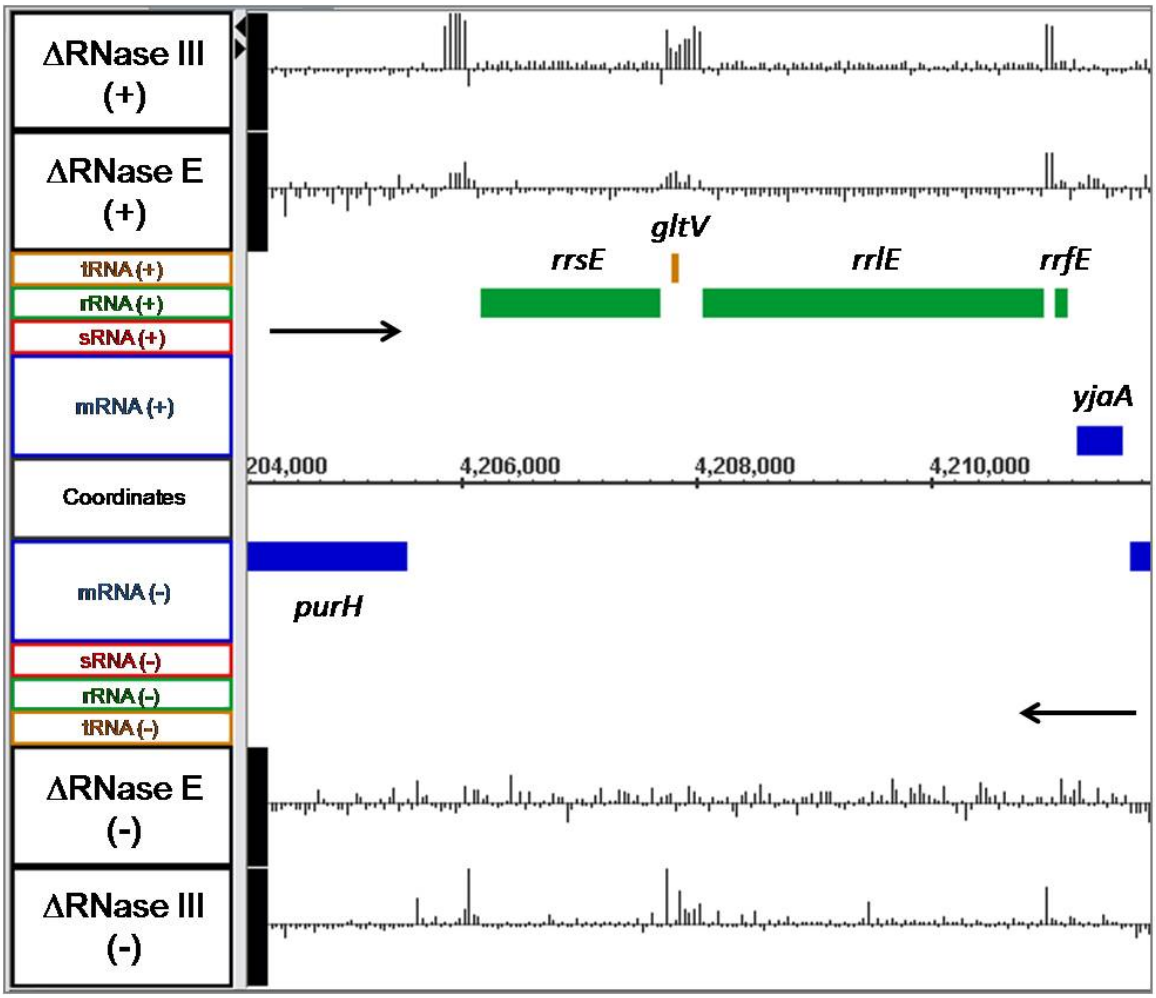

ΔRNase III (+)
ΔRNase E (+)
tRNA (+)
rRNA (+)
sRNA (+)
mRNA (+)
Coordinates
mRNA (-)
sRNA (-)
rRNA (-)
tRNA (-)
ΔRNase E (-)
ΔRNase III (-)
gltV
rrsE
rrlE
rrfE
yjaA
204,000
4,206,000
4,208,000
4,210,000
purH

Fig. 3. Microarray data for *cspE* and adjacent genes.

Data is presented as described in Fig. 1.

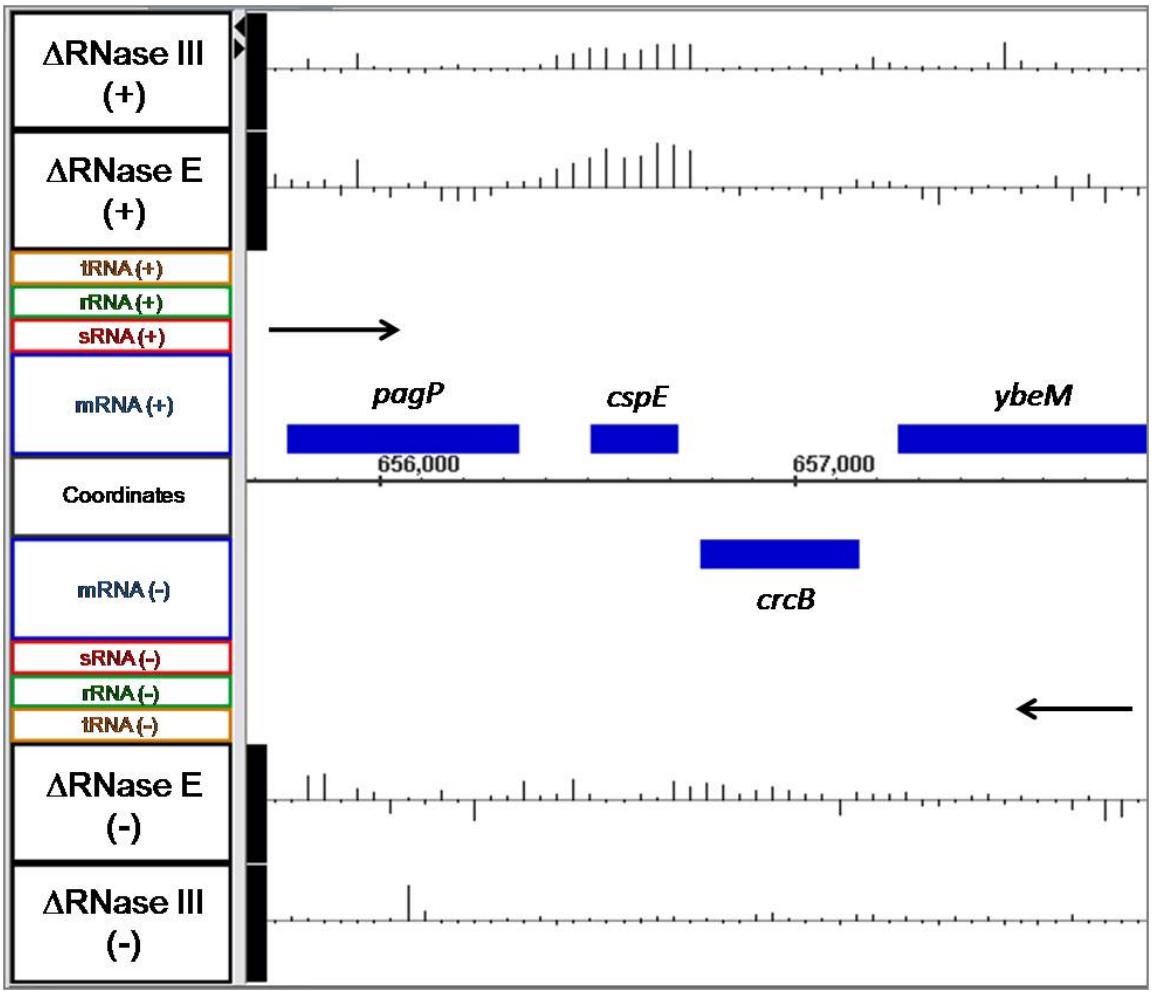

ΔRNase III (+)
ΔRNase E (+)
tRNA (+)
rRNA (+)
sRNA (+)
mRNA (+)
Coordinates
mRNA (-)
sRNA (-)
rRNA (-)
tRNA (-)
ΔRNase E (-)
ΔRNase III (-)
pagP
cspE
ybeM
656,000
657,000
crcB

Fig. 4. Northern analysis of specific transcripts.

Northern blots were conducted as described in Materials and Methods. For each blot with the exception of *dnaK*, the relative quantity values of total signal from the bands shown (RQ) represent an average of at least two independent experiments. The RQ values have been normalized based on an rRNA loading control appropriate to the gel used (each nylon membrane was stripped of the original oligonucleotide probe used as described in Materials and Methods, then hybridized with oligonucleotides specific for either 23S rRNA in the case of *dnaK* and *ompF*, and 5S rRNA for remaining blots). Standard deviations (±) are shown below the RQ values where appropriate. All Northern blots shown are representative of multiple independent replicates which have been performed (data not shown). Ribonucleotide size estimates are indicated with arrows for each band shown, with the exception of *ryhB* for which the annotated length is shown.

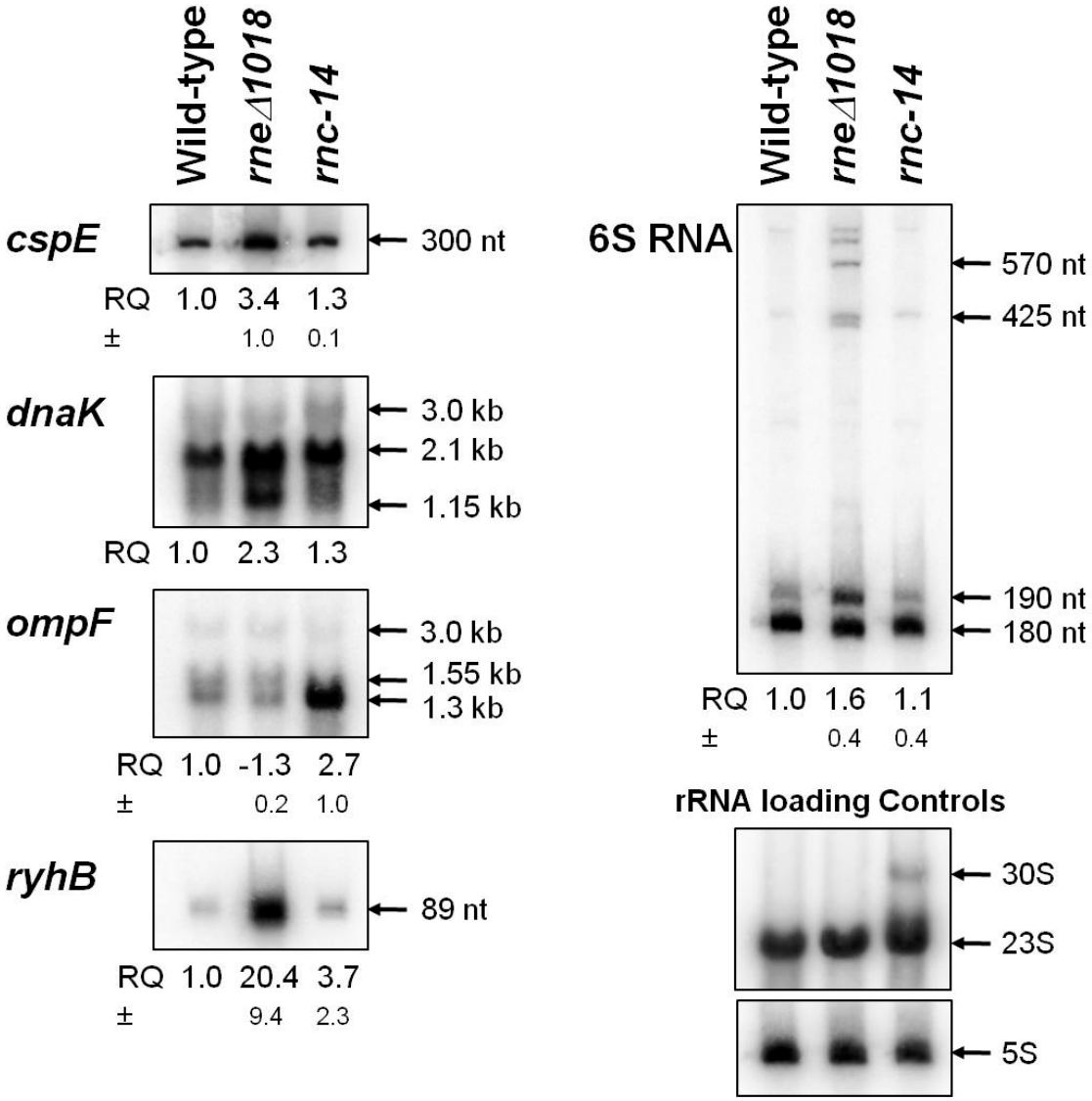

Wild-type
rne△1018
rnc-14
cspE
300 nt
RQ 1.0 3.4 1.3
± 1.0 0.1
dnaK
3.0 kb
2.1 kb
1.15 kb
RQ 1.0 2.3 1.3
ompF
3.0 kb
1.55 kb
1.3 kb
RQ 1.0 -1.3 2.7
± 0.2 1.0
ryhB
89 nt
RQ 1.0 20.4 3.7
± 9.4 2.3
6S RNA
570 nt
425 nt
190 nt
180 nt
RQ 1.0 1.6 1.1
± 0.4 0.4
rRNA loading Controls
30S
23S
5S

Fig. 5. Microarray data for *ompF* and adjacent genes.

Data is presented as described in Fig. 1.

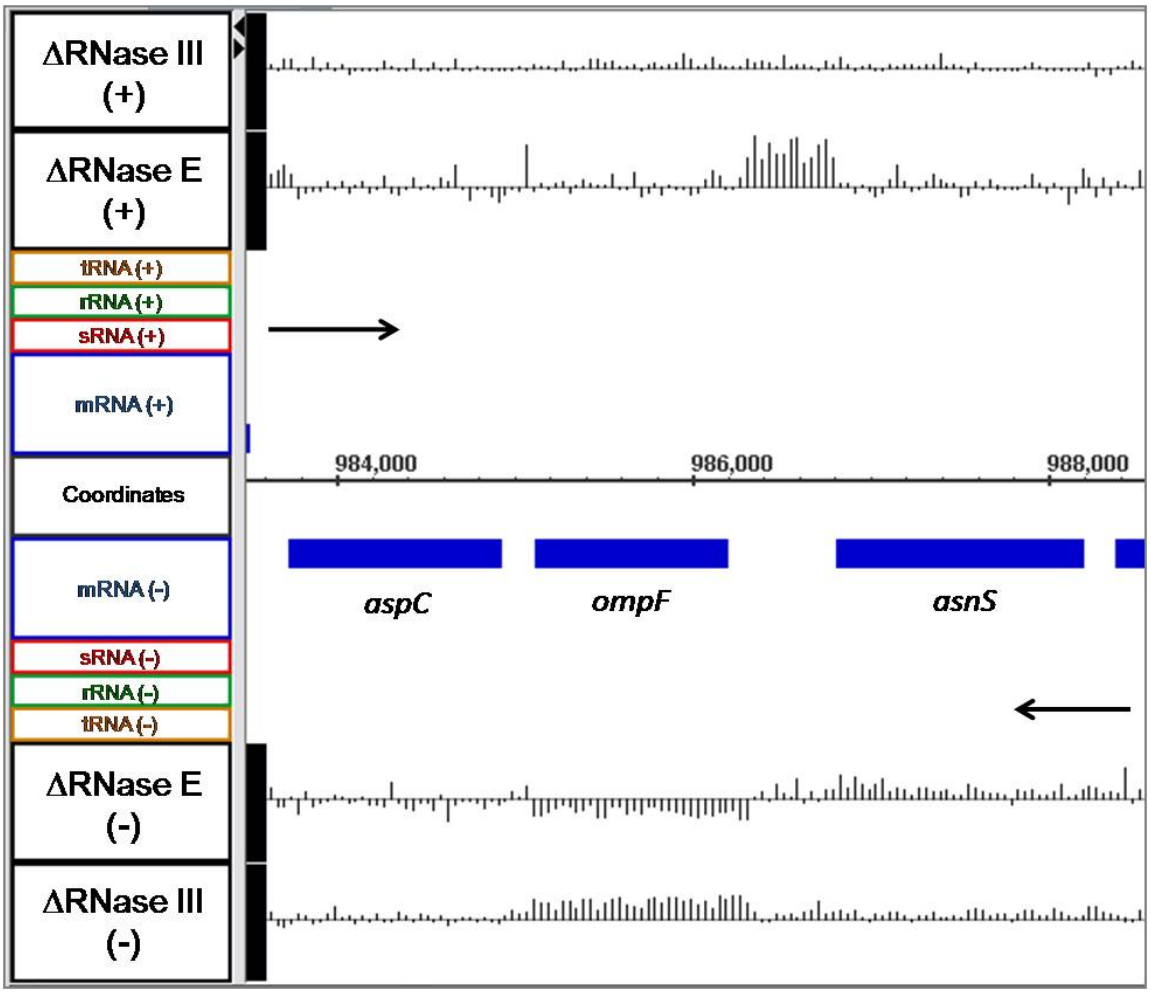

ΔRNase III (+)
ΔRNase E (+)
tRNA (+)
rRNA (+)
sRNA (+)
mRNA (+)
Coordinates
mRNA (-)
sRNA (-)
rRNA (-)
tRNA (-)
ΔRNase E (-)
ΔRNase III (-)
984,000
986,000
988,000
aspC
ompF
asnS

Fig. 6. Microarray data for the *micF* sRNA.

Data is presented as described in Fig. 1.

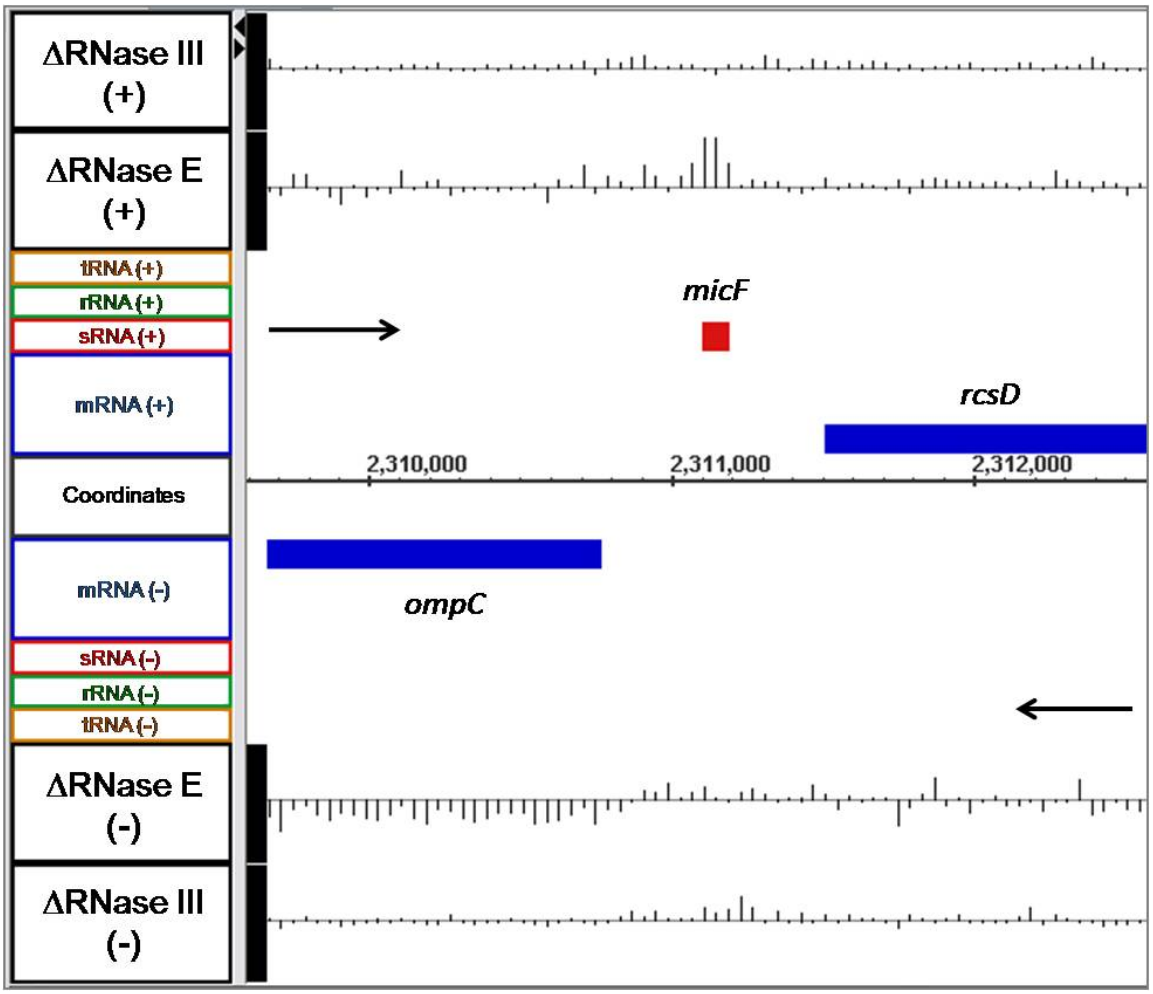

ΔRNase III (+)
ΔRNase E (+)
tRNA (+)
rRNA (+)
sRNA (+)
mRNA (+)
Coordinates
mRNA (-)
sRNA (-)
rRNA (-)
tRNA (-)
ΔRNase E (-)
ΔRNase III (-)
micF
rcsD
ompC
2,310,000
2,311,000
2,312,000

Fig. 7. Characterization of transcriptome changes in the absence of RNase E and RNase III.

The data within the graph is derived from the number of specific genomic features in each strain that contain a region of at least 100 nt changed by at least 1.5-fold, versus the total number of that gene feature annotated in the genome (see Materials and Methods). Dark grey bars that extend above the 0% X-axis indicate the number of a specific type of genomic feature which contain an area of increased abundance, versus the number of total features of that type. Light grey bars that extend below the 0% X-axis indicate the number of a specific type of genomic feature that contain an area of decreased abundance, versus the number of total features of that type. The list of genomic features affected in these strains can be found in Table S1. The two annotated tmRNA features are not included in this figure.

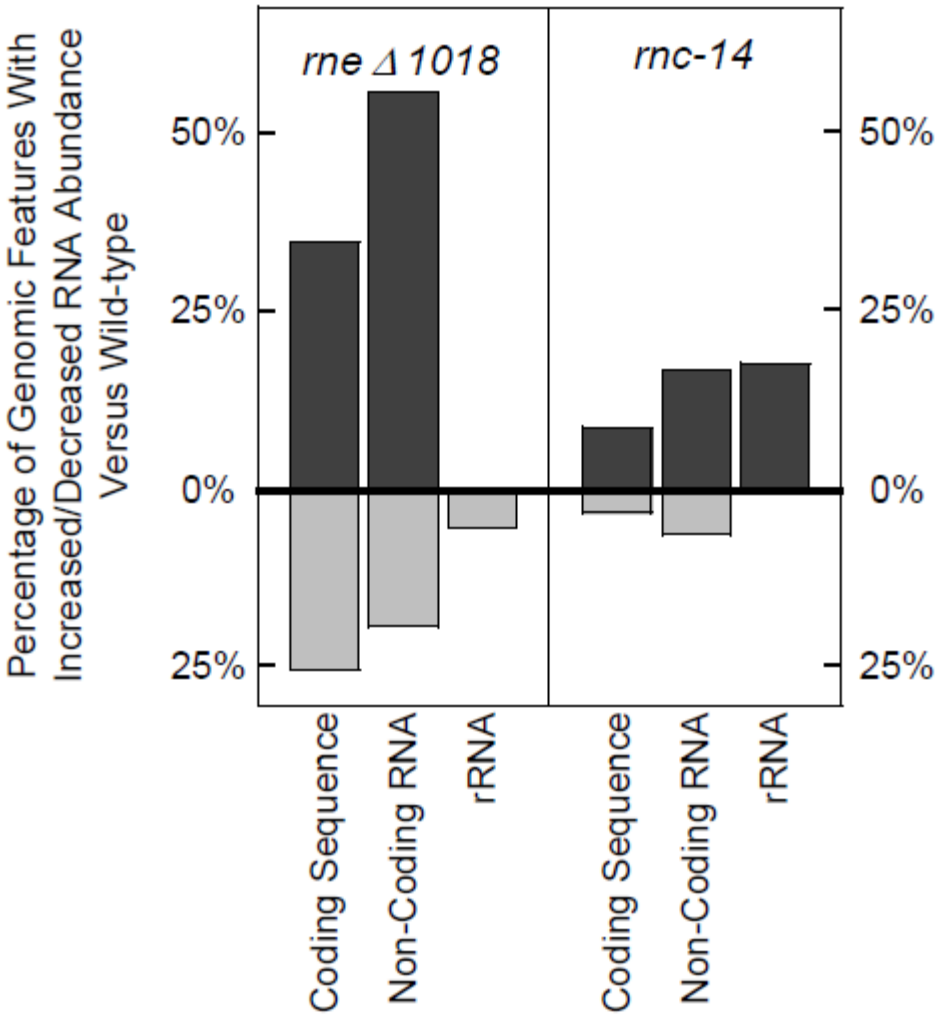
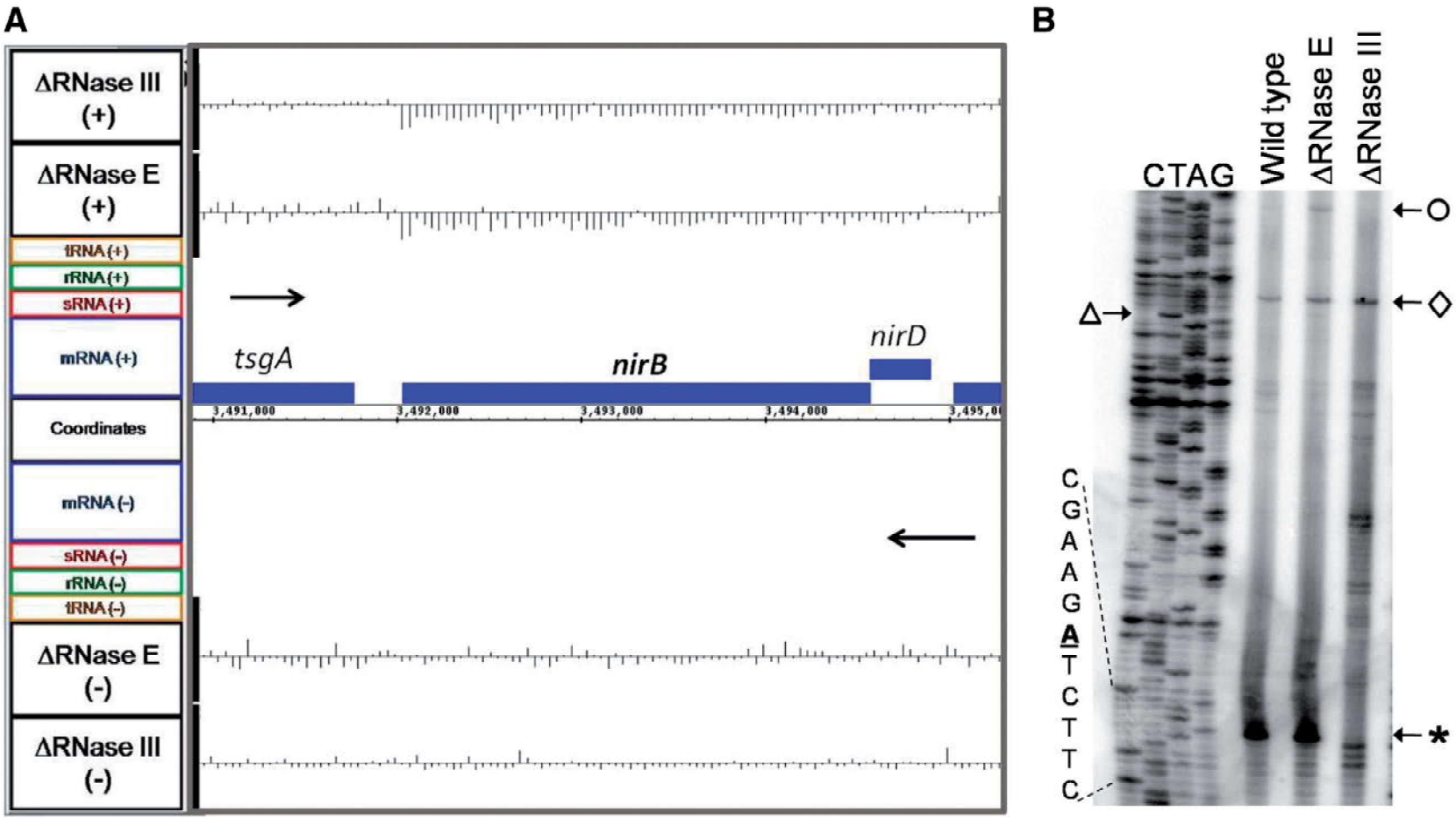

Percentage of Genomic Features With Increased/Decreased RNA Abundance Versus Wild-type
rne Δ 1018
rnc-14
50%
25%
0%
25%
50%
25%
0%
25%
Coding Sequence
Non-Coding RNA
rRNA
Coding Sequence
Non-Coding RNA
rRNA

Fig. 8. Transcriptional analysis of the *nirB* mRNA. (A) Microarray data for *nirB* and adjacent genes. Data is presented as described in Fig. 1. (B) Primer extension analysis of the *nirB* transcript. The primer extension was performed as described in the Materials and Methods. The transcription start site of the *nirB* operon transcript is indicated by an open circle. The diamond denotes a cleavage site found in all three strains located two nt upstream of the AUG translation start site, while the open triangle marks the start of the translation start codon in the DNA sequencing ladder. The asterisk (bottom right) indicates a major 5′ terminus, located within the *nirB* coding sequence that is not seen in the absence of RNase III. The putative RNase III cleavage site (in bold/underlined) is shown in the expanded sequence on the bottom left of the figure.


A
ΔRNase III (+)
ΔRNase E (+)
tRNA (+)
rRNA (+)
sRNA (+)
mRNA (+)
Coordinates
mRNA (-)
sRNA (-)
rRNA (-)
tRNA (-)
ΔRNase E (-)
ΔRNase III (-)
tsgA
nirB
nirD
3,491,000
3,492,000
3,493,000
3,494,000
B
CTAG
Wild type
ΔRNase E
ΔRNase III
Δ
C
G
A
A
G
A
T
C
T
T
C
○
◇
*

CHAPTER 3

RNA*EXPRESS*™: A QUANTITATIVE, INEXPENSIVE, AND RAPID METHOD FOR ISOLATING TOTAL RNA FROM BACTERIA[1]

---

[1] Stead, M.B., Agrawal, A., Mohanty, B.K., Nasir, R., Meagher, R.B., and S.R. Kushner. To be submitted to *Nature Methods*.

ABSTRACT

None of the currently available methods or commercial kits for isolating bacterial RNA are effective in recovering all species of intracellular RNAs (<50 - 5700 nt) with equal efficiency. Thus, we have developed a simple new procedure (**RNA*express*™**) that not only quantitatively recovers all intracellular RNA species (>99%) but is also faster (<15 min) and much less expensive (~3 cents) per sample than any currently available method. The **RNA*express*™** procedure yields ~60 µg of RNA from 10^8 *E. coli* cells that can be used directly for Northern analysis without any further purification. Northern analysis of specific transcripts ranging in size from from 76-5700 nt showed much smaller variations between biological and technical replicates compared to the commonly used RNA isolation kits or detergent-based procedures. An additional sodium acetate precipitation is required for the RNA to be used in enzymatic reactions such as reverse transcription or ligation. Significantly, the **RNA*express*™** method can be successfully used to isolate RNA from a wide range of Gram-negative and Gram-positive bacteria.

INTRODUCTION

Historically, working with RNA derived from bacterial cells has been technically difficult because of its highly labile nature and the procedures used for its isolation. Early RNA extractions relied on guanidium isothiocyanate to lyse cells and denature proteins, while the RNA was isolated using a cesium chloride cushion and ultracentrifugation(1). Subsequently, the use of hot phenol replaced cesium gradients(2-3). However, RNA extractions using hot phenol had significant problems both due to the toxicity of the phenol and because the RNA obtained was not consistently of high quality(4). Accordingly, in the mid 1980s a protocol was developed that included guanidium isothiocyanate and phenol, which yielded much more reproducible results compared to earlier methods(5-6).

As the interest in RNA metabolism in bacteria has grown, many companies developed kits making it easier for any laboratory to isolate RNA. These kits, which are relatively expensive, vary greatly in the chemistry and/or mechanics used to lyse the cells, denature and remove proteins, and to actually isolate the RNA. For example, with respect to cell lysis, the RNeasy® Protect Bacteria Mini Kit (Qiagen) uses an enzymatic digestion step for cell lysis of *Escherichia coli* cultures grown in rich medium. In contrast, the RiboPure™ kit (Ambion) uses mechanical lysis with zirconium beads within a phenol containing lysis buffer, while the TRIzol® Max™ kit (Invitrogen) employs heat in combination with guanidinium isothiocyanate. A quaternary amine-based detergent is used in each of the kits to help lyse the cells and stabilize the RNA by denaturing nucleases and other proteins (7).

For the actual isolation of RNA, the RNeasy® Protect Bacteria Mini Kit utilizes a silica column to capture RNA from the lysed sample, wash steps to remove DNA and proteins from the column, followed by elution of the RNA. For both the RiboPure™ Bacteria kit and TRIzol® Max™ Bacteria Kit, the RNA is separated by using a phenol/chloroform extraction step to aid in protein/DNA/RNA separation and subsequent RNA stability. Subsequently, the RiboPure™ Bacteria kit uses a glass filtration column followed by washes and RNA elution, which is similar to other column-based RNA extraction methods. In contrast, the TRIzol® Max™ Bacteria kit employs an isopropanol-based RNA precipitation to recover the RNA following the phenol/chloroform extraction.

The use of detergents to promote cell lysis led to the discovery of a cationic detergent (Catrimox-14®) that both aids cell lysis and captures RNA and DNA by precipitation(7-9). DNA in the pellet is subsequently removed by washing with LiCl, which takes the place of detergent in interacting with RNA, but does not effectively precipitate DNA(8). This method had the major advantage of not using phenol and provided good yields of high quality RNA. However, shortly after Qiagen acquired the patent rights to Catrimox-14® the detergent was withdrawn from the market. Subsequently, a variant of the Catrimox-14® isolation procedure was developed using a slightly different surfactant trimethyl(tetradecyl)ammonium bromide (subsequently called catrimide) that was a very effective and inexpensive substitute(10).

All current RNA isolation procedures contain multiple steps, leading to the possibility of reduced sample recovery. Furthermore, although each

manufacturer provides specifications for the yield and RNA quality resulting from their procedure, there is no published side-by-side comparison of the various methods in terms of total RNA yield, RNA quality, size distribution of the RNA molecules, time to carry out the procedure, and cost per sample. In fact, upon examination of the various RNA samples obtained using various kits and our own in-house experience with the catrimide/LiCl method, it was apparent that none of the current RNA isolation methods provide an accurate representation of various intracellular RNA pools, since each method appears to selectively enrich for either large or small RNAs relative to the levels of medium sized species. Thus, depending on the isolation method used, certain size classes of RNA will be either enriched or depleted relative to the total RNA population.

In contrast, the new **RNA*express***™ procedure described here quantitatively retains all RNA species. The isolation method is remarkably simple, rapid, reproducible, and inexpensive. With Gram-negative bacteria, it yields high quality RNA in less than 15 minutes that can be used directly for both polyacrylamide and agarose Northern analysis.

RESULTS

Development of a rapid RNA isolation method (RNA*express*™)

In most RNA isolation methods, the amount of RNA present is initially determined based on absorbance at 260 nm (A_{260}) or by total fluorescence of a nucleic acid-binding dye. Although these approaches provide an accurate estimate of the RNA present in a particular sample, the relative amounts of each RNA species

can vary widely depending on the particular isolation method employed. These variations are due to the biases towards either large (i.e., rRNA or large mRNAs) or small (tRNAs and sRNAs) RNAs associated with all of the current RNA isolation procedures. Thus, it is not possible to accurately assess the *in vivo* distribution of the various classes of RNA molecules in a particular sample by either A_{260} or fluorescent dyes. In order to help address the problem of representative and quantitative recovery, we sought to develop a one-step RNA extraction procedure that took place in a single tube in which total RNA was recovered in the supernatant and the bulk DNA and proteins were left in the pellet. We hypothesized that such an approach would provide a more accurate overview of the actual intracellular distribution of all RNA species, as losses are minimized by the simplicity of the procedure.

Surprisingly, we observed that exponentially growing *E. coli* cells were rapidly lysed, when suspended in a formamide-based RNA extraction solution and heated at 95^{o}C for seven minutes (see Methods). Following centrifugation for five minutes at 16,000 x g, the RNA was in the supernatant and the gelatinous pellet contained protein, cell debris, and the majority of the DNA. The RNA was quantified based on A_{260} by first blanking a spectrophotometer with the RNA extraction solution. It was important that the RNA extraction solution made on the same day to isolate the RNA sample was also used as the blank, since the A_{260} of the extraction solution itself changed over time after the addition of 2-mercaptoethanol. A one ml sample of an early exponential culture of *E. coli* (10^8 cells) yielded ~ 60 µg of total RNA with the entire procedure taking less than 15

minutes (Table 3). The quality of the RNA derived from this method was as good as or better than RNA obtained by the other methods tested in this study (Fig. 1, Table 1) and was suitable, without any further treatment, for Northern analysis (Fig 2). The minimal genomic DNA contamination in the **RNA*express*™** sample was comparable to was obtained with the other isolation methods (data not shown).

The RNA*express*™ method recovers >99% of all RNA species

Even though the **RNA*express*™** procedure was rapid and yielded more RNA per cell than any other method tested, it was important to determine how much RNA remained in the gelatinous pellet after extraction was carried out. Accordingly, we scaled up the isolation to 10 ml of culture (10^9 cells), and again carried out the protocol in a single tube. Following removal of the supernatant, the pellet was gently washed once with the extraction solution at room temperature. After a subsequent centrifugation, the pellet was resuspended in water and extracted using acidic phenol/chloroform (See Methods). The aqueous phase was precipitated with sodium acetate/ethanol and resuspended in water. Two hundred and fifty ng of RNA from both the pellet and supernatant were run on an agarose gel to confirm the presence, quality, and quantity of the RNA. In each of two replicates, approximately 2.7 µg of high quality RNA was recovered from the pellet (accounting for a 10% loss during acid phenol/chloroform extraction), while an average of 711.5 µg of RNA was found in the first supernatant, indicating that the efficiency of RNA recovery from *E. coli* using the **RNA*express*™** method was greater than 99% (data not shown).

Analysis of RNA*express*™ isolated RNA

In an attempt to test the quality and range of the transcripts isolated using the **RNA*express*™** method**,** we compared the RNA with RNA isolated by our previously optimized RNA isolation method (Catrimide/LiCl(10)) and three of the most widely used commercially available RNA isolation kits [TRIzol® Max™ Bacteria (Invitrogen), RNeasy® Protect Bacteria (Qiagen) and RiboPure™ Bacteria (Ambion)]. Each extraction method was tested using at least two independent biological replicates and at least four technical replicates per biological replicate. The quality of each RNA sample was assessed using three main parameters: purity as determined by a spectrophotometer ($A_{260/280}$ ratio), the 23S rRNA/16S rRNA ratio as determined by Bioanalyzer analysis (Agilent Technologies), and an RNA integrity score derived from Bioanalyzer analysis.

As shown in Fig. 1, all RNA preparations were of high quality based on both Bioanalyzer analysis (Agilent Technologies) (Fig. 1A) and the integrity of the 16S and 23S rRNA bands on agarose gels (Fig. 1B). The $A_{260/280}$ ratio of ~ 2.0 for all the RNA preparations (Table 1) indicated that all of the samples were relatively pure with the exception of the **RNA*express*™** samples without precipitation. Ideally, a $A_{260/280}$ ratio of 1.8-2 is indicative of highly purified RNA when resuspended in a buffered solution like TE (pH 8.0), since this ratio is highly dependent on pH (11). Thus, the low $A_{260/280}$ ratio of the **RNA*express*™** RNA samples may simply be due to a change in the absorption properties of RNA in the presence of formamide. In fact, resuspension of the RNA in RNase free water after a NaOAc/ethanol precipitation significantly improved the ratio (Table

1). Additionally, diluting the **RNA*express*™** RNA sample with RNase-free water also improved the $A_{260/280}$ ratio such that it was comparable to the other methods shown in Table 1 (data not shown). Interestingly, there were significant differences in terms of the amounts of the rRNAs and tRNAs present as well as RNA integrity scores (Table 1). The **RNA*express*™**, catrimide/LiCl, RNeasy® and RiboPure™ methods yielded comparable ratios of 23S/16S rRNAs, which were significantly higher than what was observed with the TRIzol® Max™ Bacteria method, with a ratio of approximately 2 being ideal (Table 1). In contrast, the TRIzol® Max™ Bacteria method yielded the highest concentrations of 5S rRNA and tRNAs, but the method had the second lowest total RNA yield of all the methods tested (Fig. 1A, 3).

Since there were obvious differences in the distribution of RNAs among of the most abundant RNA size classes obtained among the five RNA isolation methods, we attempted to quantitate the differences for specific RNA molecules ranging in size from 76-5700 nt using Northern analysis. The relative abundance of each of the transcripts (Fig. 2A) was calculated relative to catrimide/LiCl method (Table 2). The abundance of two out of three small RNA species (76-120 nt; *pheU/V* and 5S rRNA) in the RNA isolated by the **RNA*express*™** method was comparable to the TRIzol® Max™ Bacteria method (Table 2), while all the other methods had lower yields (Table 2). Furthermore, it should be noted that there were considerable variations among the independent isolations for both *pheU/V* and *rhyB* transcripts when either RNeasy® Protect Bacteria or

Ribopure™ Bacteria kits were used. These results were consistent with the manufacturer's stated limitations for species smaller than 200 nt.

Four out of five transcripts longer than 300 nt (except *ompF*) were present in the **RNA*express*™** isolated RNA in comparable levels compared to the catrimide/LiCl method, if not better than TRIzol® Max™, RNeasy®Protect Bacteria and Ribopure™ Bacteria methods. However, the 5700 nt *rpsJ* transcript was easily visualized in both the Catrimide/LiCl and **RNA*express*™** methods, and barely detected in either the TRIzol® Max or RNeasy® Protect Bacteria isolated RNA samples (Fig. 2, Table 2). While TRIzol® Max™, RNeasy® Protect and Ribopure™ isolated RNAs had limitations in detecting transcripts either larger than 3000 nt or smaller than 120 nt (Fig. 2, Table 2), all transcripts were detected in the catrimide/LiCl isolated RNA. In contrast, all of the transcripts were present with better representation in the **RNA*express*™** isolated RNA compared to the catrimide/LiCl method. The increased representation of transcripts in the range of 300-3000 nt in the Trizol® Max™, RNeasy® Protect and Ribopure™ kits compared to the **RNA*express*™** method may simply be due to an enrichment of these species at the expense of either smaller or larger species. It is important to note that since >99% of the cellular RNA was present in the **RNA*express*™** samples, data obtained from the Northern analysis of the **RNA*express*™** RNA (Fig. 2) probably most accurately represent the *in vivo* levels of each specific transcript.

Of the commonly used RNA isolation kits, TRIzol® Max™ was the best for isolating small RNAs, but it selectively lost larger RNA species (Table 2). In

contrast, RNeasy® RNA was comparable or better than the catrimide/LiCl isolated RNA, but also lost the larger *rpsJ* transcript (Table 2). In the size range between 300-3000 nt, all five isolation methods gave similar results (Fig. 2, Table 2).

Generality of RNA*express*™ RNA isolation method

Although the data shown in this study were generated using RNA isolated from exponentially growing cells, the **RNA*express*™** method worked equally well with late stationary phase cells (data not shown), unlike what has been observed with the catrimide/LiCl extraction procedure(10). In addition, the **RNA*express*™** procedure was easily and quantitatively scaled to handle 10 ml of culture (10^9 cells) for situations where larger amounts of RNA were needed.

Furthermore, the **RNA*express*™** RNA was used directly in polyacrylamide/urea gels. However, due to the nature of the RNA extraction solution, the RNA species (>1000 nt) reproducibly appeared larger on formaldehyde agarose gels than their actual size (as shown in Fig. 2B). However, once the RNA was precipitated out of the extraction solution and re-suspended in water (see Methods), the electrophoretic mobility of all RNA species were normal.

We have used the **RNA*express*™** method to successfully extract RNA from a number of unrelated Gram-negative bacteria including: *Alcalingenes faecalis* (ATCC 8750), *Serratia marcescens* (ATCC 14756), *Shigella flexneri* (ATCC 9199), *Pseudomonas aeruginosa* (ATCC 27853), *Salmonella enterica* (ATCC 29629), *Ruegeria pomeroyi* (ATCC 700808), and *Myxococcus xanthus* DK1622.

Additionally, using a slightly modified version of the **RNA*express*™** method in which zirconium bead homogenization was used for added lysis efficiency, high quality RNA was obtained from two Gram-positive bacteria, *Bacillus subtilis* (ATCC6633) and *Staphylococcus aureus* (ATCC 6538).

Overview of the five RNA isolation procedures

As shown in Table 3, the **RNA*express*™** method provided the highest total RNA yield of all five isolation procedures (1.7-4-fold higher). A comparison of the cost per sample, time to complete the RNA isolation, and the recommended size range for efficient recovery of specific RNA transcripts, as tested in this study, is shown in Table 3. Cost per sample ranged from over $8.00 for the RNeasy ® Protect Bacteria kit to approximately three cents for the **RNA*express*™** method.

DISCUSSION

We have described here a simple, rapid, and reproducible RNA isolation procedure (**RNA*express*™**) that yields RNA that can be used for Northern analysis without any further purification. Besides its simplicity, using eight transcripts with sizes ranging between 76 - 5700 nt, we have demonstrated that **RNA*express*™** gives little evidence of non-quantitative recovery or loss of any RNA species based on size. It quantitatively (>99%) recovered all RNA species from the cell, providing the most accurate representation of intracellular RNA pools compared to any of the other isolation methods tested. In addition, it works equally well with exponential and stationary phase cultures.

Surprisingly, RNA molecules larger than 1000 nt isolated using the **RNA*express*™** method exhibited a reproducible size-dependent decrease in electrophoretic mobility (Fig. 2) when used directly after isolation, although this phenomenon did not affect the quality of agarose Northerns (Fig. 2A). The change in electrophoretic mobility was eliminated via a sodium acetate/ethanol precipitation (data not shown). Following NaOAc precipitation, a water insoluble pellet remained even after the RNA was solubilized, indicating the removal of some contaminants. We have successfully used **RNA*express*™**-isolated RNA for primer extension analysis and RT-PCR following precipitation, DNase I treatment, and a further RNA precipitation step to remove digestion products. Overall, the quality and representative recovery offered by **RNA*express*™** method is unmatched by the other methods tested in this study and is uniquely suited for high-throughput gene expression analyses.

In lieu of the sodium acetate/ethanol precipitations, a faster but significantly more expensive option was the RNeasy® kit (or similar silica-column based extraction kit) or RiboPure™ kit, which can be used to recover the RNA from the formamide-based RNA extraction solution. Using either column-based method following the **RNA*express*™** extraction yielded extremely high-quality RNA suitable for any type of highly-sensitive RNA analysis (data not shown). However, the drawback to using a column, as demonstrated in this study (Table 2), was the non-quantitative recovery of RNA species depending on size and possible secondary structure of the RNA molecule.

With the advent of qRT-PCR, microarrays, and next generation sequencing, genome-wide expression profiling has become an indispensible tool to decipher biological systems. However, at the heart of the most robust and sophisticated gene expression analysis lies the quality and reproducibility of the extracted RNA pool. For example, if a research group were to use a column-based RNA extraction methodology, such as those tested in this study, to examine maturation of small RNAs less than 200 nt, the results of the study would be flawed due to non-quantitative recovery of RNA molecules less than 200 nt using the RNA extraction methods (Table 3). Alternatively, if a group were to examine the relative abundance of a 1,000 nt transcript compared with a 5,000 nt transcript, the ratio between the two abundances would vary considerably between various types of RNA extraction methodologies. More importantly, it is clear that no RNA isolation methodology (with the likely exception of the **RNA*express***™ method) is suitable for the study of all types and sizes of RNA molecules in the same experiment.

As such, the technical limitation of non-quantitative RNA recovery has arguably influenced many if not all gene expression studies published to date. The challenge of quantitative RNA isolation becomes more pressing in the field of RNA processing. For example, if using the catrimide/LiCl method for analysis of the maturation of tRNAs or other small RNAs, it may be possible for the method to effectively enrich for longer processing intermediates while losing some of the mature product which could lead to inaccurate assumptions about the efficiency or rate of RNA maturation. For any experiment involving RNA, our

study demonstrates that the method by which you isolate the RNA needs to be carefully considered with respect to the sizes of the RNAs of interest.

METHODS

Bacterial strains

E. coli strain MG1693 (*thyA715 rph-1*) (provided by the *E. coli* Genetic Stock Center, Yale University) was grown with shaking at 37°C in Luria broth supplemented with thymine (50 µg/ml) to exactly 50 Klett units above background (No. 42 green filter or OD_{600} 0.4), which is approximately 10^8 cfu/ml. Other strains were generously provided by the Departments of Microbiology and Marine Sciences at the University of Georgia.

RNA isolation methods

One ml of bacterial culture (10^8 cells) was used for each RNA isolation sample. Each RNA extraction method was performed with a minimum of two independent biological replicates and at least four technical replicates to measure reproducibility. For the **RNA*express*™** method, one ml of culture was centrifuged at 16,000 x g for 30 seconds and the supernatant was removed by aspiration. The cell pellet was stored in dry ice until ready for extraction. Cell pellets were then re-suspended in 100 µl of RNA extraction solution [18 mM EDTA, 0.025% SDS, 1% 2-mercaptoethanol, 95% Formamide (RNA grade)] by vortexing vigorously. The cells were lysed by incubating the sample at 95^{o}C in a sand bath for seven minutes. The cell debris was pelleted by centrifuging the warm sample at

16,000 x g for 5 minutes at room temperature. The supernatant was carefully transferred to a fresh tube without disturbing the clear gelatinous pellet.

The catrimide/LiCl method for RNA extraction used for these experiments was performed similarly to the method described by Mohanty *et al*.(10), but was modified for one ml samples. Briefly, one ml of bacterial culture was added to 500 µl of stop buffer, which was previously frozen horizontally in a 1.7 ml microcentrifuge tube. The cells were immediately mixed by vortexing vigorously, and then pelleted by centrifugation at 5,000 x g for 5 minutes at 4°C. The supernatant was carefully removed by aspiration, and the pellet was resuspended in 200 µl of lysis buffer by vortexing. The sample was then placed into a dry-ice ethanol slurry for 90 seconds, and followed by 90 seconds of incubation in a 37°C water bath. This freeze-thaw cycle was repeated four times in total. After the fourth 37°C incubation, the sample was transferred into the dry ice-ethanol slurry in order to refreeze the solution, and 35 µl of 20 mM acetic acid was then added to the frozen solution. The sample was then placed back into the 37°C water bath, followed by addition of 200 µl of 10% Catrimide [(trimethyl(tetradecyl)ammonium bromide)] when the sample was almost completely thawed. The sample was briefly vortexed and centrifuged at 16,000 x g for 10 minutes at 4°C. The supernatant was carefully removed by aspiration, and the pellet was suspended in 500 µl of 2 M LiCl in 35% ethanol by vortexing very vigorously. The sample was then incubated at room temperature for 5 minutes, followed by centrifugation at 16,000 x g for 10 minutes at 4°C. The supernatant was carefully removed by aspiration, and the pellet was

resuspended in 500 µl of 2 M LiCl in water followed by a repeat centrifugation. The pellet was briefly vortexed in 75% ethanol and centrifuged at 8,000 x g for 5 minutes at 4°C. The ethanol was removed by aspiration, and the tube was briefly centrifuged for a second time in order to collect and remove the remaining ethanol with a pipette. The pellet was allowed to air dry at room temperature for 10 minutes and subsequently hydrated by the addition of 100 µl of RNase-free water and incubated at room temperature for 10 minutes. The tube was vigorously vortexed, centrifuged at maximum force (21,000 x g) at room temperature for 1 minute to pellet cell debris, and the RNA containing supernatant was transferred to a new tube.

All other RNA extraction methods were done according to the manufacturer's recommendations and protocols specific for the number of *E. coli* cells and conditions in which they were grown. Any step described as optional, but that might improve the quality or yield of RNA was followed. No optional DNase I treatment was performed on any RNA sample used in this study. Every effort was made to ensure that the extracted RNA using each method met the manufacturer's guidelines in terms of overall RNA yield, A_{260}/A_{280} ratio, and RNA quality.

Determination of RNA quantity and quality

RNA quantity and A_{260}/A_{280} ratios were determined using a Nanodrop 2000c (Thermo Scientific). The amount of RNA in the **RNA*express***™ supernatants was determined by A_{260}, using the RNA extraction solution as a blank. RNA quality

was assessed by running 250 ng of each RNA sample, as determined by A_{260}, on a 1.2% agarose-0.5X TBE gel with Ethidium Bromide, run at 5 v/cm for 1 hour. RNA samples were denatured prior to loading by suspension in Gel Loading Buffer II (Ambion) and heating for 5 minutes at 95°C. Approximately 100 ng of each RNA sample was subsequently analyzed on a Bioanalyzer RNA chip (Agilent Technologies) using the manufacturer's recommendations.

Quantitive Determination of RNA recovery using the RNA*express*™ method

In order to estimate the amount of RNA remaining in the pellet, we performed an **RNA*express*™** extraction using 10 ml of *E. coli* cells (10^8 cells/ml) using 500 µl of RNA extraction solution. After the supernatant was recovered and placed into a separate tube, an additional 500 µl of room temperature RNA extraction solution was gently added to the gelatinous pellet in order to recover additional supernatant that could not be initially removed without disturbing the pellet. The tube was then spun at 16,000 x g for an additional five minutes, and the supernatant was again removed without disturbing the pellet. The pellet was then suspended in 100 µl of RNase-free water. Subsequently, another 100 µl of acidic phenol/chloroform (Ambion, 5:1 solution, pH 4.5) was added and the tube was vortexed vigorously for 30 seconds. The tube was then centrifuged at 16,000 x g for five minutes, and the aqueous phase was transferred to a fresh tube and sodium acetate/ethanol precipitated. The precipitated RNA was hydrated in 20 µl of RNase-free water. After the RNA was fully dissolved, the total amount of RNA was determined based on A_{260} and compared with the amount of RNA in the first 500 µl volume of RNA extraction solution recovered from the pellet.

Northern Analysis

Two types of Northern blots were performed in this study, a 6% polyacrlyamide/ 8.3 M Urea 1X TBE gel for small RNA species (*lpp, cspE,* 5S rRNA, *ryhB,* and *pheU/pheV*), and a 1.2% Agarose 1X MOPS gel for larger species (*rpsJ* operon, *adhE,* and *ompF*). Northern analysis was performed as described in Stead *et al.(12)* The RNA present in the supernatants obtained from the **RNA*express***™ method was used directly for polyacrylamide gels after dilution to the desired loading volume in a formamide-based RNA loading dye. For agarose Northerns, the RNA in extraction solution was brought up to a total volume of 10 µl with water. Subsequently, 4 µl of loading solution (3.8 µl of any formamide-based RNA loading dye along with 0.2 µl of 37% formaldehyde) was added. The samples were heated at 65°C for five minutes and placed on ice for one minute, followed by brief centrifugation before loading onto a 1.2% Agarose 1X MOPS gel, similar to the method of Vincze and Bowra(13).

The Northern membranes were then simultaneously probed such that the signal for *lpp,* 5S rRNA, and *pheU/V* were simultaneously visualized on a single membrane (similarly for *cspE/ryhB* and *adhE/ompF*). This approach helped determine if loading errors could account for differences in signal between the two replicates, as the percentage difference should be the same for each of those RNA species probed in the same lanes, unless the RNA extraction method used caused non-quantitative recovery of a particular RNA species. It was also technically possible that an error during the transfer of RNA from the gel to the nitrocellulose membrane accounted for a difference between replicates, but this

type of error is extraordinarily rare with polyacrylamide Northerns in our hands, and occurs infrequently with agarose Northerns.

RNA*express*™ RNA suspended in extraction solution was easily precipitated and resuspended in water using a sodium acetate/ethanol precipitation method in order to avoid slower electrophoretic mobility of larger transcripts during agarose Northern analysis. The **RNA*express*™** RNA sample was first diluted with two volumes of water followed by addition of 1/10 volume of 3 M sodium acetate, pH 5.2 and the sample was mixed by pipetting. Three volumes of 100% ethanol were then added, the sample mixed briefly by vortexing and incubated for at least 60 minutes at -80^{o}C. The tube was centrifuged at 16,000 x g for 30 min at 4^{o}C. The supernatant was carefully removed by aspiration, and the pellet was washed with 250 µl of 75% ethanol and centrifuged at 8,000 x g for five minutes at 4^{o}C. The supernatant was removed via aspiration, and the tube was briefly centrifuged again. Following the removal of any remaining ethanol, the pellet was air dried. The pellet was resuspended in water and centrifuged at 16,000 x g to pellet any remaining water insoluble proteins, and the RNA containing supernatant was transferred to a fresh tube.

ACKNOWLEDGEMENTS

This work was supported by a grant from the National Institutes of General Medical Sciences (GM81554) to S.R.K.

REFERENCES

1. Glisin, V., Crkvenjakov, R. and Byus, C. (1974) Ribonucleic acid isolated by cesium chloride centrifugation. *Biochemistry*, **13**, 2633-2637.
2. Sevag, M.G., Lackman, D.B. and Smolens, J. (1938) The isolation of the components of streptococcal nucleoproteins in serologically active form. *J. Biol. Chem.*, **124**, 425-436.
3. Bovre, K. and Szybalski, W. (1971) Multistep DNA-RNA hybridization techniques. *Methods Enzymol.*, **XXI**, 350-389.
4. Donovan, W.P. and Kushner, S.R. (1986) Polynucleotide phosphorylase and ribonuclease II are required for cell viability and mRNA turnover in Escherichia coli K-12. *Proc Natl Acad Sci U S A*, **83**, 120-124.
5. Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem*, **162**, 156-159.
6. Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry*, **18**, 5294-5299.
7. Macfarlane, D.E. and Dahle, C.E. (1993) Isolating RNA from whole blood-The dawn of RNA-based diagnostics. *Nature* **362**, 186-188.
8. Macfarlane, D.E. and Dahle, C.E. (1997) Isolating RNA from clinical samples with Catrimox-14 and lithium chloride. *Journal of clinical laboratory analysis*, **11**, 132-139.

9. O'Hara, E.B., Chekanova, J.A., Ingle, C.A., Kushner, Z.R., Peters, E. and Kushner, S.R. (1995) Polyadenylylation helps regulate mRNA decay in *Escherichia coli*. *Proc Natl Acad Sci U S A*, **92**, 1807-1811.

10. Mohanty, B.K., Giladi, H., Maples, V.F. and Kushner, S.R. (2008) Analysis of RNA decay, processing, and polyadenylation in *Escherichia coli* and other prokaryotes. *Methods Enzymol*, **447**, 3-29.

11. Wilfinger, W.W., Mackey, K. and Chomczynski, P. (1997) Effect of pH and ionic strength on the spectrophotometric assessment of nucleic acid purity. *Biotechniques*, **22**, 474-476, 478-481.

12. Stead, M.B., Marshburn, S., Mohanty, B.K., Mitra, J., P., C.L., Ray, D., Hughes, T. and Kushner, S.R. (2010) Analysis of *E. coli* RNase E and RNase III activity *in vivo* using tiling microarrays. *Nucleic Acids Res*, **39**, 3188-3203

13. Vincze, E. and Bowra, S. (2005) Northerns revisited: a protocol that eliminates formaldehyde from the gel while enhancing resolution and sensitivity. *Anal Biochem*, **342**, 356-357.

Table 1. RNA Quality Scores

	Catrimide/ LiCl	TRIzol® Max™ Bacteria	RNeasy® Protect Bacteria	RiboPure™ Bacteria	**RNA***express*™	**RNA***express*™ Precipitated*
$A_{260/280}$[1]	2.00 ± 0.01	1.97 ± 0.02	2.13 ± 0.01	2.12 ± 0.02	1.73 ± 0.01	1.92 ± 0.02
23S rRNA/ 16S rRNA ratio[2]	1.73 ± 0.15	1.21 ± 0.05	2.38 ± 0.55	2.05 ± 0.16	1.80 ± 0.01	1.21 ± 0.08
RNA integrity #[2]	9.4 ± 0.29	7.9 ± 0.17	9.0 ± 0.52	9.4 ± 0.21	9.5 ± 0.00	9.5 ± 0.35

[1] Obtained using a Nanodrop 2000c (see Methods)

[2] Obtained from Agilent Bioanalyzer analysis (see Methods). Each value is the average of at least two replicates.

*RNA from original extraction solution was precipitated using NaOAc/ethanol and resuspended in water (see Methods).

Table 2. Comparison of specific transcript levels in totalRNA preparations, using various methods.

		Relative RNA Abundance in total RNA isolated by each method				
RNA transcript	RNA ~size (nt)	Catrimide / LiCl	TRIzol® Max™ Bacteria	RNeasy® Protect Bacteria	RiboPure ™ Bacteria	**RNA*express*™**
pheU/V	76	1 ± 0.14	23.8 ± 3.8	2.0 ± 0.90	0.49 ± 0.42	29.6 ± 0.30
ryhB	90	1 ± 0.02	13.7 ± 2.9	10.9 ± 0.55	3.1 ± 0.68	0.89 ± 0.08
5S rRNA	120	1 ± 0.17	6.8 ± 1.0	2.1 ± 0.17	0.57 ± 0.18	4.9 ± 0.39
cspE	300	1 ± 0.04	4.6 ± 0.60	3.1 ± 0.03	5.1 ± 0.31	0.95 ± 0.10
lpp	330	1 ± 0.04	1.4 ± 0.20	1.3 ± 0.00	1.5 ± 0.03	1.0 ± 0.07
ompF	1000	1 ± 0.04	2.0 ± 0.14	2.4 ± 0.12	1.6 ± 0.03	0.67 ± 0.18
adhE	3000	1 ± 0.02	1.4 ± 0.28	2.1 ± 0.11	2.7 ± 0.16	0.78 ± 0.07
rpsJ operon	5700	1 ± 0.02	0.04 ± 0.00	0.03 ± 0.01	0.79 ± 0.06	0.99 ± 0.12

RNA was isolated using each of the methods listed as described in Methods. Each transcript was probed using a specific oligonucleotide probe (sequences available on request) labeled with ^{32}P. The data were obtained using a GE Storm PhosphorImager and quantified using ImageQuant TL software. The values obtained for the Catrimide RNA was set at 1 and used to normalize the other RNA samples. Each relative abundance value is the average of at least two independent replicates.

Table 3 . Overall comparison of the various RNA isolation procedures.

	Catrimide/ LiCl	TRIzol® Max™ Bacteria	RNeasy® Protect Bacteria	RiboPure ™ Bacteria	**RNAexpress ™**
Approx. cost/sample (U.S. dollars)	0.20	4.20	8.10	7.14	**0.03**
Approx. yield from 10^8 *E. coli* cells (µg)	35	27	35	15	**60**
Approx. duration of isolation (min)	60	60	40	40	**15**
RNA size range for efficient isolation (nt)	76-5700	76-3000	200-3000	300-5700	**76-5700**

Cost data based on the list price of chemicals or extraction kits. The approximate yield is the average yield of RNA isolated in this study based on 10^8 cells. All of these methods have the ability to handle more than 10^8 cells. For example, the RiboPure Bacteria kit recommends using 10^9 cells. The approximate duration of RNA isolation is based on the time required for each individual step using a small number of samples. The RNA size range data are based on the sizes of specific RNAs detected by Northern analysis (Fig. 2).

Fig. 1. RNA quality assessment of RNA samples from each method.

A. A representative composite bioanalyzer digital gel image using technical replicates of each of the RNA extraction methodologies tested (See Methods). B. A representative composite image of technical replicates of 250 ng of total RNA (based on A_{260}) from each RNA extraction method electrophoresed on a 1.2% agarose-0.5X TBE gel stained with ethidium bromide. The **RNA*express*™** samples denoted with an asterisk (*) were sodium acetate/ethanol precipitated prior to analysis.

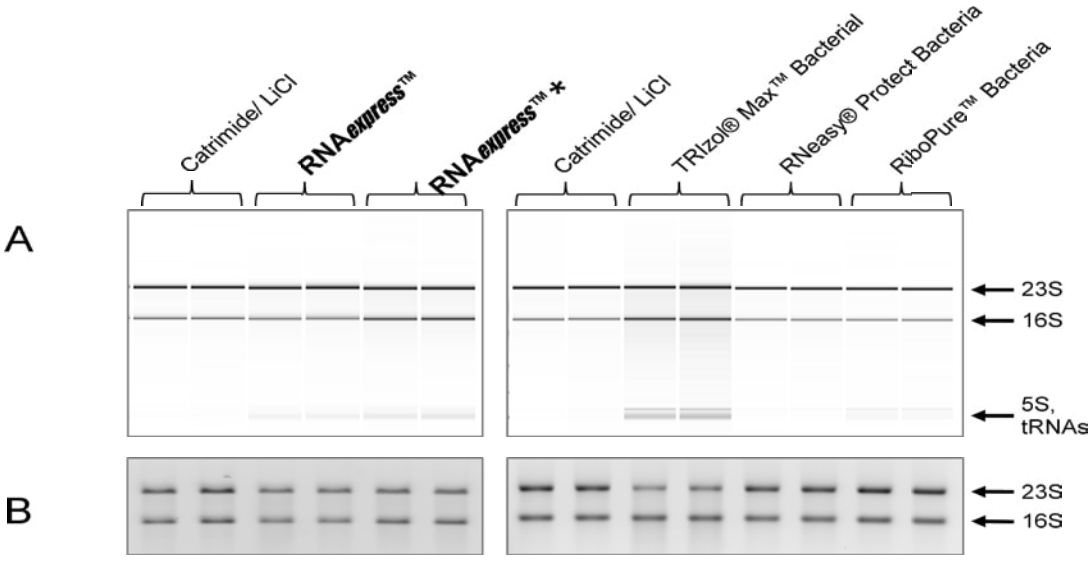
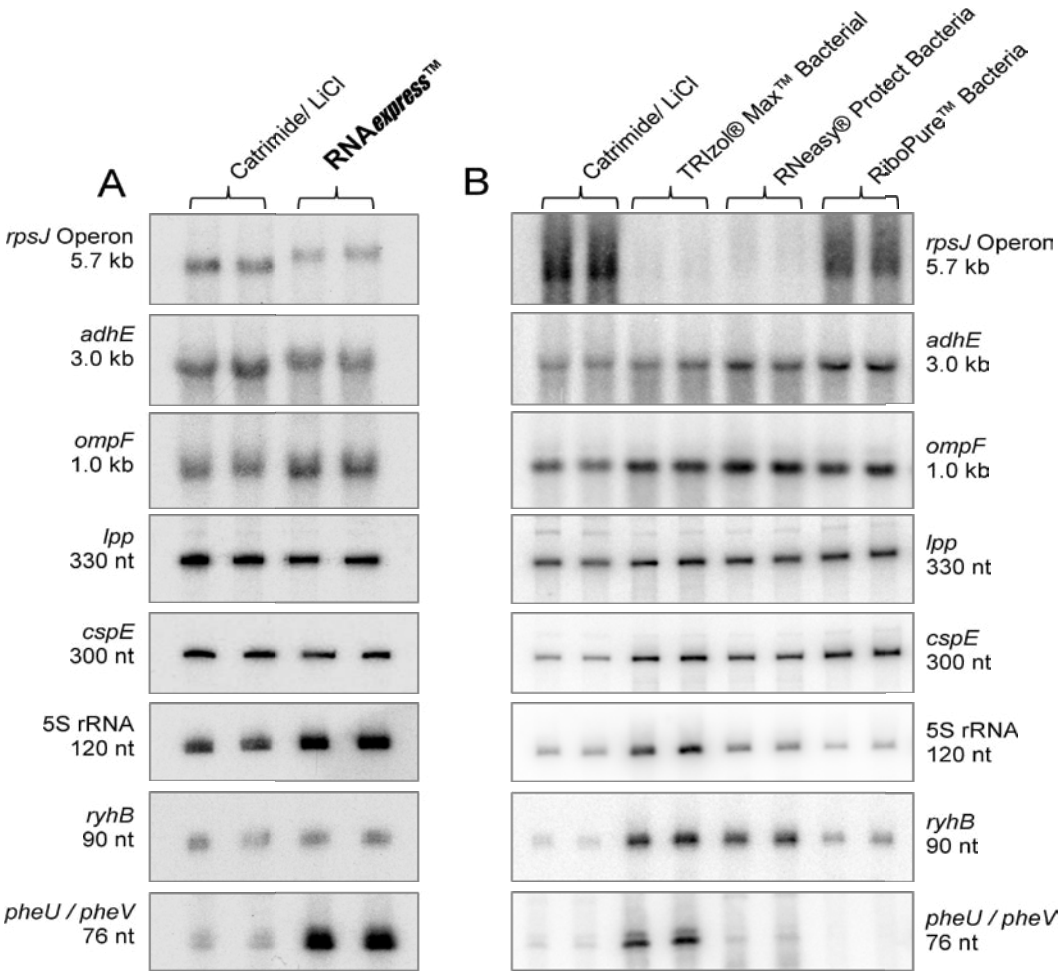

A
B
Catrimide/ LiCl
RNAexpress™
RNAexpress™ +
Catrimide/ LiCl
TRIzol® Max™ Bacterial
RNeasy® Protect Bacteria
RiboPure™ Bacteria
23S
16S
5S, tRNAs
23S
16S

Fig. 2. Northern analysis of specific RNA species using total RNA pool isolated by each RNA isolation method.

Five µg of total RNA (based on A_{260}) was used for Northern analysis on each of the eight specific RNAs listed on the sides of the autoradiograms along with the approximate size of each (see Methods). Two independent technical replicates are shown for each RNA isolation method and are loaded in adjacent lanes. A. Comparison of new **RNA*express*™** method versus Catrimide/ LiCl for specific RNAs. Each catrimide/LiCl and **RNA*express*™** RNA sample had equivalent amounts of extraction solution and water (See Methods). The *rpsJ* operon and *adhE* mRNA reproducibly exhibit a decrease in electrophoretic mobility in **RNA*express*™** samples (See Methods). B. Comparison of each common RNA extraction method for specific RNAs. RNA samples were treated as indicated in the Methods section.


A
Catrimide/ LiCl
RNAexpress™
rpsJ Operon
5.7 kb
adhE
3.0 kb
ompF
1.0 kb
lpp
330 nt
cspE
300 nt
5S rRNA
120 nt
ryhB
90 nt
pheU / pheV
76 nt
B
Catrimide/ LiCl
TRIzol® Max™ Bacterial
RNeasy® Protect Bacteria
RiboPure™ Bacteria
rpsJ Operon
5.7 kb
adhE
3.0 kb
ompF
1.0 kb
lpp
330 nt
cspE
300 nt
5S rRNA
120 nt
ryhB
90 nt
pheU / pheV
76 nt

CHAPTER 4

PROCESSING OF 30S rRNA AND 5' 23S rRNA MATURATION IN *ESCHERICHIA COLI*[1]

---

[1]Stead, M.B., Chaudhuri, G., Nasir, R., and S.R. Kushner. To be submitted to *Nucleic Acids Research.*

INTRODUCTION

Escherichia coli has seven distinct chromosomally encoded ribosomal RNA operons (*rrnA, rrnB, rrnC, rrnD, rrnE, rrnG, rrnH*). Each operon contains the sequence for 16S rRNA, 23S rRNA, and 5S rRNA, in addition to at least one tRNA (Fig. 1). Each of the seven rRNA operons is transcribed as a primary transcript, called 30S rRNA, which must be processed post-transcriptionally by ribonucleases to generate the mature, functional sequences that are incorporated into ribosomes. The pathway by which the 30S rRNA transcript is transcribed and processed into functional species has been studied in *E. coli* since the 1960's, but much of the pathway remains to be elucidated due to the complexity and essential nature of the system. For example, although many of the ribonuclease processing events in the maturation of 30S have been described, the endoribonuclease(s) responsible for the maturation of the 5' ends of both 23S rRNA and 5S rRNA, along with the 3' end of 16S rRNA, have not been identified (1-10).

It is known that following transcription of 30S rRNA and most likely after the binding of some ribosomal proteins, RNase III, an endoribonuclease specific for double-stranded RNA, cleaves the primary transcript at four locations to release 17S (pre-16S), 25S (pre-23S), and 9S (pre-5S) rRNA precursors (Fig. 2) (7-9,11-13). Following RNase III cleavage, RNase E and RNase G cleave successively to form the mature 5' end of 16S rRNA. In addition, RNase E is also responsible for the further processing of 5S rRNA by cleaving within 3 nt on each side of the mature sequence (Fig. 2) (2,10,14-16). The 3' ends of 23S and

5S rRNAs are matured primarily by the 3'-5' exoribonuclease RNase T (17-21). Surprisingly, the enzyme responsible for the first cleavage steps on the 30S rRNA transcript, RNase III, seems to be dispensable for rRNA processing, as efficient processing still occurs in the absence of RNase III, causing little change in the growth rate of the bacterium. Furthermore, only the 5' terminus of 23S rRNA is not properly matured in the absence of RNase III (8-9,15,22).

Since the loss of RNase III has only a very modest impact on the processing of the 30S rRNA precursor, what is complementing the loss of RNase III, and more importantly, what constitutes the non-RNase III-dependent pathway for rRNA maturation in wild-type *E. coli*? To answer these questions, we sought to determine the enzyme(s) responsible for the separation of 16S and 23S rRNA precursors and the maturation of the 5' end of 23S rRNA in the absence of RNase III. However, our experiments were complicated by the fact that failure to process 30S rRNA precursors would be lethal to the cell. In addition, it is known that cells are tolerant to improperly processed 23S rRNA. Thus, a mutant that resulted in aberrant 5' end processing of 23S rRNA could only be detected using primer extension analysis.

As an additional potential complication, each of the seven rRNA operons in *E. coli* are different from each other in terms of nucleotide sequence (of the exterior and interior transcribed spacer regions), length of the 30S transcript, rRNA and tRNA copy numbers, and tRNA identity (Fig. 1). The lack of homogeneity among the rRNA operons raised the question of whether they are

all processed by the same enzymes at the same locations, further complicating the analysis of processing by generating operon-dependent cleavage patterns.

A number of known enzymes could be potentially involved in the separation of 16S and 23S pre-rRNAs, including all of the single-stranded RNA specific endoribonucleases in *E. coli*: RNase E, RNase P, RNase Z, RNase G, RNase LS, and RNase I. Other enzymes, such as the five distinct DEAD-box RNA helicases, may also be involved as mutants lacking some combinations of these helicases have been shown to be defective in ribosome assembly and rRNA processing (11,23). Due to the hairpin formation by the sequences flanking both mature 16S and 23S rRNAs, RNA helicases may be required to unwind the secondary structure, allowing single-stranded RNA-specific endoribonucleases to carry out cleavage events.

Additionally, a highly conserved gene of relatively unknown function, *ybeY,* has been shown to significantly affect rRNA maturation, through a currently unidentified mechanism (24). It is possible that *ybeY* is itself either an rRNA binding protein, or an endoribonuclease, but a thorough characterization of the enzymatic activity of the protein has yet to be performed. It is thus possible that the effects of the *ybeY* deletion on rRNA processing could be either a direct or indirect function of the gene.

Because the endoribonuclease(s) involved in the maturation of the 5' end of 23S rRNA have not been identified, it is also possible that a previously uncharacterized gene is responsible. Alternatively a known endonuclease, which

is thought to be normally inactive in *E. coli*, such as the endoribonuclease members of the CRISPR/CAS genome defense system may be involved in processing (25-29). For example, the CRISPR/CAS system is repressed by a small RNA dependent upon RNase III for activity (30). Thus, it is formally possible that this system, while normally inactive, is functionally active in rRNA maturation in the absence of RNase III (25,27).

To help answer the question of which enzymes are involved in the separation of 30S rRNA and in the maturation of the 5' end of 23S rRNA, we used high-resolution agarose gel electrophoresis of total RNA and primer extension analysis in a search for multiple mutants which significantly altered the processing of 30S rRNA. In order to aid in the construction of multiple mutant strains, we generated low and single copy arabinose-inducible RNase III-controlled expression plasmids to facilitate the construction of mutants that might be inviable in the absence of RNase III. To determine if the processing of the seven rRNA operons—which are distinct in terms of both structure and spacer sequence identity—are processed identically, we made use of a strain of *E. coli* devoid of all chromosomally encoded rRNA operons, which allowed us to analyze the processing of individual rRNA operons. Our data have ruled out a number of known endoribonucleases in rRNA processing and lay the foundation for the elucidation of other enzymes involved in rRNA maturation in *Escherichia coli.*

RESULTS

Analysis of rRNA processing in strains containing all seven chromosomal rRNA operons

Although previous work has shown that some 30S rRNA is present in *rnc* mutants (3), we set out to construct an isogenic set of strains containing an *rnc* knockout as well as mutations in other known endoribonucleases. Initially we moved two distinct *rnc* mutations [*rncΔ38* (3) and *rnc-14* (31)] into our standard genetic background (MG1693, Table 1). Using a monoclonal antibody against RNase III obtained from Don Court, we showed that the *rnc-14* strain (SK4455) contained no detectable RNase III protein, while there was a small amount of a truncated RNase III protein in the *rncΔ38* strain (SK7622) (data not shown). In addition, it appeared that there was a slight increase in the level of 30S rRNA precursor in SK4455 compared to SK7622 (data not shown).

Based on these results, most of the multiple mutants were constructed, if possible, using the *rnc-14* allele. For the initial experiments, multiple mutants containing *rnc-14* and *rnpA49* (RNase P, SK10525) or *rnc-14* and *ΔybeY* (SK5372) were constructed using bacteriophage P1 transduction. In addition, a *rnpA49 ΔybeY* double mutant was also made (SK5374). RNA was extracted from these strains and the isogenic single mutants. As a further control, RNA was also isolated from a pentuple mutant containing mutations in RNase E, RNase P, RNase G, RNase Z, and RNase LS (SK4485). As expected, a faint band of 30S rRNA was observed in the *rnc-14* single mutant (SK4455) but not in

the RNase P strain (SK2525) or the wild-type control (MG1693) (Fig. 3, lanes 1-3). None of the other single mutant strains tested (*rne-1, Δrnz*) contained any visible 30S pre-rRNA species upon separation on agarose and staining with ethidium bromide (data not shown).

Construction of multiple mutant strains containing a single rRNA operon

Since we did not see any change in the level of 30S rRNA in the multiple mutants we initially tested (data not shown), we decided to try and simplify our analysis. Specifically, we sought to develop a system in which we could look at a single rRNA operon. We reasoned that such a system could potentially reduce the number of processing intermediates that might result from the presence of seven different rRNA operons, and it might also permit us to determine if all the rRNA operons were processed in the same manner.

Accordingly, we obtained an unpublished strain of *E. coli* (SQ2158, Table 2) that lacks all seven rRNA operons in the MG1655 genetic background. The mutant is kept viable through two plasmids, one a pSC101 derivative (6-8 copies per cell) that contains the *rrnB* operon, and one a p15A derivative (~15 copies per cell) that contains all of the tRNAs normally found within the deleted rRNA operons (Cathy Squires, unpublished results). Although the generation time of this strain was approximately 10 minutes longer versus a wild-type control in rich medium, the strain has proven to be more phenotypically robust than previously published strains that lacked all chromosomal rRNA operons (32) (data not shown). Unfortunately, the plasmid carrying the *rrnB* rRNA operon in the Squires

strain (pK4-16 in strain SQ2158, see Table 3), upon restriction digestion, did not match the restriction map we were provided and could not be easily displaced to study other *rrn* operons. As a result a new series of pSC101 origin plasmids was developed using pWSK29 derivatives (33) (pMSK2-4, Table 3).

Three rRNA operons were chosen for study to best address the variation amongst the seven operons: *rrnB, rrnG*, and *rrnD*. The *E. coli* rRNA operons tend to group into two main families, with *rrnB, rrnC, rrnE*, and *rrnG* being very similar, and *rrnA*, *rrnD*, and *rrnH* being very similar to each other (Fig. 1). The two operons representing the greatest divergence from each other are *rrnD* and *rrnB*. Therefore those two operons, along with *rrnG* (*rrnG* is very similar to *rrnB* and both operons should be similarly processed if nucleotide sequence is the primary determinant for cleavage) were cloned to generate pMSK2-4 (Table 3), respectively. Each plasmid contains a different selectable drug marker. Since *E. coli* will only maintain one type of pSC101 origin plasmid in the absence of selection, the introduction of a new *rrn* operon-containing plasmid with appropriate drug selection will displace an existing *rrn* operon plasmid. Plasmid displacement was carried out by a modification of the method of Ow et al. (34).

Analysis of rRNA processing in strains containing a single rRNA operon

RNA from mutant strains carrying only the *rrnG* operon were subjected to high-resolution agarose gel electrophoresis. The only strain which reproducibly seemed to contain more 30S rRNA than an *rnc-14* single mutant was the *ΔdeaD rnc-14* double mutant (SK5364, Table 2). However, this increase was less than

two-fold compared to the 30S rRNA species present in the *rnc-14* single mutant, and was still less than 5% of the total rRNA (Fig. 3, Lanes 5 and 12, data not shown).

Primer extension analysis of the 5' of 23S rRNA

Since none of the multiple mutants we examined showed any significant change in the amount of 30S rRNA, we decided to take a different approach to examining rRNA maturation. It had previously been shown that the 5' termini of 23S rRNA were significantly different in an *rnc* mutant compared to a wild-type control (7,9) (Fig. 4, lanes 1-4). Specifically, in wild-type *E. coli* the bulk of 23S rRNA has a single mature 5' terminus with small amounts of five slightly longer species (1, 3, 4, 7, and 8 nt longer) (Fig. 4, lanes 1-2). In contrast, in the *rnc-14* strain the mature species is completely missing accompanied by the appearance of many additional 5' termini that were either longer or shorter than the mature species (Fig. 4, lanes 3-4). Interestingly, inactivation of either RNase E (Fig. 4, lanes 5-6) or RNase P (Fig. 4, lanes 7-8) did not change the profile of 5' termini.

Even though the RNase E temperature-sensitive allele (*rne-1*)-containing strain upon shift to the non-permissive temperature did not cause any measurable changes in the processing of 23S (Fig. 4, lanes 5-6), the *rne-1* allele is known to have some residual activity and therefore any role of RNase E cannot currently be ruled out (35). Additionally, primer extension analysis of 23S rRNA from a multiple endoribonuclease deficient strain (missing RNase G, RNase Z, and RNase LS, with temperature-sensitive alleles of both RNase E and

RNase P) (SK4485, Table 1) showed no major defects in processing at the 5' end at either the permissive or non-permissive temperature (data not shown). This result was at odds with a recently published report of the effect of RNase G on the processing of the mature 5' end of 23S rRNA, because the multiple mutant is devoid of RNase G, and no build up of processing intermediates was present (data not shown) (36). However, RNA from the multiple endoribonuclease mutant strain (SK4485) did show a strong processing intermediate of 23S rRNA with the 5' end mapping near the 3' end of 16S rRNA, when shifted to the non-permissive temperature (data not shown). This result has not yet been repeated and the exact location of the cleavage event has not been mapped.

Primer extension results with the strains listed in Table 2 have indicated that, in fact, the processing steps of the *rrnB, rrnG,* and *rrnD* operons are most likely the same, with differences in cleavage pattern being attributable to the length of the spacer regions for each particular operon (data not shown). However, primer extensions using the single *rrn* operon strains have clearly indicated that RNase I, DeaD, SrmB, CasE, Cas2, and Cas3 by themselves (both in the presence and absence of RNase III, except in the case of SrmB where a double mutant strain with *rnc-14* could not be made) are likely not involved in the processing of the 5' end of 23S (data not shown).

The strongest defects in rRNA processing observed so far have been in the *ΔybeY rnpA49* double mutant (SK5374, Table 1). This strain produced a 16S precursor species approximately the size of 17S that was equally as intense upon ethidium bromide staining as mature 16S rRNA. This precursor abundance

is a stronger phenotype than previously reported, but still in agreement with the observations made in Davies et al. 2010 (24). The RNase P temperature-sensitive allele (*rnpA49*) by itself at both the permissive and non-permissive temperatures shows slight stabilization of 23S rRNA precursors 7-8 nt longer than the mature species at the 5' end (Fig. 4, Lanes 7,8). The stabilization of these same products was more pronounced in the *ΔybeY rnpA49* strain, also at both temperatures. Thus, the effect of both mutants may be additive, indicating that both YbeY and RNase P work in the same pathway to affect the processing of the 7 and 8 nt 5' 23S rRNA processing intermediates (Fig. 4, Lanes 13,14, data not shown).

Construction of controlled-expression RNase III plasmids

It is possible that in the absence of RNase III, a gene involved in the backup processing of the rRNA precursors could become essential thereby preventing the construction of a viable double mutant. For example, we have been unable to construct RNase Z, RNase III double mutants. This was an unexpected finding because *E. coli* RNase Z is thought to be primarily active on mRNAs (37). Therefore, in order to aid in the construction of potentially inviable multiple mutants, a set of controlled expression RNase III plasmids was constructed. The RNase III coding sequence was flanked, using multiple overlap PCR reactions, with the P_{BAD} arabinose-inducible promoter sequence from pKD46 (38) (the arabinose-inducible lambda *red* recombinase plasmid) and the Rho-independent transcription terminator from the *rrnB* operon. This PCR fragment was then inserted into two distinct vectors, one a 6-8 copy pSC101

origin plasmid (pMSK9, Table 3), and the other a single-copy mini-F origin plasmid (pMSK10, Table 3). These plasmids were able to express wild-type levels of functional RNase III with as little as 100 μM arabinose (data not shown). RNase III expression was effectively shut off in the presence of 10 mM glucose, as demonstrated by both Western blotting and by analysis of 30S rRNA amounts in chromosomal *rnc-14* strains containing either of the RNase III expression plasmids (data not shown). While the usefulness of these plasmids has yet to be utilized for the construction of synthetically lethal RNase III deficient strains, they will no doubt be a helpful tool for the elucidation of the rRNA processing pathway.

DISCUSSION

The data presented here has demonstrated that the currently accepted model for the processing of 30S rRNA precursors in *E. coli* by RNase III is only partially correct. As shown in Fig. 3, the inactivation of RNase III only led to the appearance of a very small amount of 30S rRNA (less than 5% of the total rRNA population). Thus, the bulk of the 30S rRNA is being processed by an alternative pathway(s) that involves some combination of additional endonucleases and possibly RNA helicases and/or RNA binding proteins. Based on our analysis of a variety of strains carrying mutations in known or putative endonucleases, it appears that inactivation of RNase E, RNase P, CasE, Cas2, Cas3, or RNase I plays no role in the processing of the 30S precursor in strains carrying RNase III (Fig. 3). There may have been a small effect of inactivation the DeaD RNA helicase in an RNase III mutant (Fig. 3, lane 12).

Furthermore, the data shown in Fig. 4 provides the first high resolution primer extension analysis of the 5' end of 23S rRNA. Interestingly, in wild-type cells there is a primary 5' terminus, representing the purported mature 5' end, along with five additional termini that are longer by 1, 3, 4, 7, and 8 nt. Since *E. coli* does not appear to contain any 5' → 3' exonucleases, it is not clear if these species are ever processed further, but rather are incorporated into 30S ribosomes with immature 5' termini. Strikingly, in the RNase III mutant, at both 37 °C and 42 °C, none of the 23S rRNA species has the mature 5' terminus, but rather there is a combination of over 20 different species that can be over 20 nt longer than the mature 23S rRNA (Fig. 4, lanes 3,4). Since the generation time of the *rnc-14* strain is only incrementally slower than the wild type control, it is clear that ribosomes containing these aberrantly processed 23S rRNA species are not significantly impaired in their ability to carry out protein synthesis.

Also of interest is the fact that inactivation of RNase E alone does not change the pattern of 5' termini compared to the wild-type control, at either 37 °C or 42 °C (Fig. 4, lanes 5,6). In contrast, inactivation of RNase P did seem to alter the ratio of the various longer species (Fig. 4, lanes 7,8). Similar alterations in the amounts of the longer species were also observed in the *ΔybeY rnpA49* double mutant (Fig. 4, lanes 13, 14).

A hypothesis that might explain some of the observations reported in this study center on the effects of the YbeY protein, which was first identified as a heat-shock protein and is required for the survival of *E. coli* grown at 42 °C or higher (39). At elevated temperatures, *ybeY* is three-fold more abundant than at

37 °C, and the YbeY protein has been shown to be required for efficient translation at any temperature, especially at elevated temperatures (40). Specifically, YbeY was shown to be required for the optimal activity of the 30S ribosomal subunit, which contains 16S rRNA, and is also the rRNA species most affected in the absence of YbeY (24,40). Because it is known that ribosomal proteins bind the 30S rRNA precursor before the first cleavage event and begin to initiate subunit formation (7,13), we hypothesize that YbeY actually promotes the binding of the ribosomal proteins to the 30S rRNA precursor—specifically the proteins involved in 30S subunit formation—which subsequently confer specificity to endoribonucleases, promoting the efficient processing of rRNA precursors. This binding would likely be less stable at higher temperatures, and thus may be the reason for the extreme temperature sensitivity exhibited in *ybeY* mutants (39). Upon shift to higher temperatures, the three-fold increase in the level of YbeY may increase the binding efficiency of the ribosomal proteins, to a level surpassing the efficiency at normal temperatures, and may be responsible for the decreased levels of rRNA processing intermediates at elevated temperatures.

The seemingly cumulative effects of the *rnpA49* and *ΔybeY* alleles on rRNA processing are intriguing, as RNase P has not previously been shown to affect the processing of 23S rRNA (Fig. 4). In light of our hypothesis on the activity and function of YbeY, it is possible that RNase P is responsible for the 5' end maturation of 23S rRNA, but is dependent upon the activity of YbeY to alter the secondary structure of the RNA, forming a suitable RNase P cleavage site. Testing this hypothesis will likely require the use of an RNase P expression

plasmid, both because the temperature-sensitive allele of RNase P may have residual activity, and because *ΔybeY* strains are highly temperature-sensitive themselves. Use of an RNase P expression plasmid should allow analysis of the system at physiological temperatures and without the additional complications from the heat-shock response or the stress response associated with reduced translation efficiency in *ΔybeY* strains.

The possibility also exists that the rRNA processing pathway has multiple endoribonucleases which function efficiently and redundantly in the system, allowing a normal rRNA processing phenotype until the right mutant combination is tested. Precedent for this hypothesis can be found in the processing of tRNAs, where multiple exoribonucleases need to be mutated or knocked out before maturation of the 3' end of tRNAs are significantly affected (41-42). Unfortunately, this possibility may require that a substantial number of mutant combinations be tested for rRNA processing defects by primer extension analysis. Additionally, while the cleavage patterns near the 5' end of 23S rRNA in both wild-type and *rnc-14* strains are essentially the same for each of the three operons, there are no data regarding the cleavage profile of other location on the 30S transcript. Also, the enzyme(s) responsible for other maturation steps may not behave identically with each operon, and therefore the cleavage pattern will be of interest once a gene (or a combination of genes) is identified which confers processing to the 5' end of 23S.

One of the major problems with the analysis of rRNA processing is the essential nature of the system, including the fact that two of the ribonucleases

tested (RNase E and RNase P) are essential for viability. In this study temperature-sensitive alleles were used for RNase E and RNase P and therefore bacteria had to be grown at either 30°C or 37°C and shifted to 42°C to cause loss of enzymatic activity. Furthermore, both the RNase P and RNase E temperature-sensitive alleles may not be fully inactivated at elevated temperatures (35,43). Also, in every strain examined (including wild-type) shifting to 42 °C from either 37 °C or 30 °C results in a reproducible loss of rRNA processing intermediates, which can either stem from a decrease in transcription of 30S, or increased processing of the rRNA precursors (Fig. 4). It is certainly possible that immediately upon shift to a higher temperature, the heat-shock response may temporarily decrease the transcription of rRNA, but as rRNA production is directly tied to growth rate, both wild-type bacteria and RNase III-deficient mutants grow faster at 44 °C than at 37 °C (unpublished observations). Additionally, because of the hairpin formation of the spacer sequences flanking the mature sequences of 16S and 23S rRNA, DEAD-box RNA helicases may be needed to unwind the secondary structure of the hairpin, allowing a suitable substrate for single-stranded RNA-specific endoribonucleases, and this activity is likely to be temperature-dependent as higher temperatures will de-stabilize the hairpin. These observations point to an increase of degradation capacity or efficiency rather than a decrease in rRNA transcription as the reason for the decrease in rRNA processing intermediates at elevated temperatures.

Though this project is not finished, a significant amount of information has been obtained from the data presented. It is now clear that RNase I, and

members of the CRISPR/CAS system, are not involved with either the separation of 30S pre-rRNA, or in the 5' end maturation of 23S rRNA, and that RNA helicases, individually, cannot account for the unknown processing activities. We have also identified a previously unreported stabilization of processing intermediates in the 5' maturation of 23S rRNA in the *ΔybeY rnpA49* double mutant. It remains unclear what role YbeY and RNase P play in the maturation of 23S rRNA, but we now have several hypotheses that may be tested. Furthermore, this work has developed a series of plasmids and strains which will aide in the discovery of the enzymes involved in these pathways and will allow a single-operon view of rRNA processing for the first time. Even with the many complexities involved with the elucidation of the rRNA processing pathway, the knowledge gained and the tools developed in this study will help to guide future work to fruition.

MATERIALS AND METHODS

Bacterial strains

Escherichia coli strain SQ2158 (MG1655, *ΔrrnABCDEGH /* pK4-16 pTRNA67) was a gift from C. Squires, and strain SK5390 (MG1655, *ΔybeY::cat*) was a gift from G. Walker. Mutant alleles of Cas2, CasE, Cas3, RNase I, SrmB, and DeaD were obtained from the *E. coli* Genetic Stock Center (Yale University) and were constructed as part of the Keio Collection (44-45) (See Table 1). All strains generated by the addition of chromosomal mutant alleles were constructed by P1 transduction and subsequent selection for the antibiotic

resistance marker associated with the mutant allele. For example, in the case of *ΔygbF756::kan,* P1 grown on JW5438 was used to transduce the allele into both SK5193 and SK5344. Transduction mixtures were spread on rich medium plates containing 25 μg/ml kanamycin, resulting in the isolation of strains SK5353 and SK5355, respectively. The genotypes of transductants were confirmed by PCR, or by DNA sequencing in the case of *rnpA49.* The lists of bacterial strains examined and constructed for this study are shown in Tables 1 and 2.

Plasmids

Plasmids carrying rRNA operons (pMSK2-4) were generated by PCR amplification of the rRNA operon using primers specific to each particular operon (*rrnD, rrnG,* and *rrnB,* respectively). The amplified fragments also contained restriction sites needed for digestion and ligation into pSC101 origin plasmids (pWSK29 for pMSK2 and pMSK4, and pVMK94 for pMSK3). The nucleotide sequences of the *rrn* operons in pMSK2-4 were confirmed as wild-type by DNA sequencing, and each plasmid contains the entire *rrn* operon including upstream transcription enhancer sequences (except in the case of pMSK3, in which the entire *rrnB* transcription enhancer region was not included by mistake, which requires that the plasmid be reconstructed at some point in the future, though the plasmid was not used in the generation of data presented in this study).

RNase III expression plasmids used in this study were constructed by PCR amplification of the P_{BAD} promoter region from pKD46 (38) (starting with the transcription terminator of the *araC* gene and ending one nucleotide before the

ATG of the gamma protein), the wild-type *rnc* gene from MG1693 (beginning with the ATG start codon and ending with the TGA stop codon), and the Rho-independent transcription terminator from the *rrnB* operon (starting at the 3' end of *rrfB* (5S rRNA) and ending after the second Rho-independent transcription terminator). The three PCR fragments were combined by overlap PCR reactions and ligated into the multiple cloning sites of both pWSK29 (6-8 copy, pMSK9) and pVMK194 (single-copy, pMSK10) using the *KpnI* and *SacII* restriction sites (which were included at the distal ends of the oligonucleotide primers used for the final PCR amplification of the P_{BAD}-*rnc-rrnB* terminator fragment). The nucleotide sequence of the RNase III coding region within pMSK9 and pMSK10 was confirmed as wild type by DNA sequencing. All plasmids generated in this study were electroporated into DH5α super-competent cells (Invitrogen) following ligation.

Oligonucleotide sequences used in this study are available upon request.

RNA isolation and analysis

RNA was isolated from strains grown in Luria broth supplemented with 0.05% thymine, with shaking, until 50 Klett units (No. 42 green filter, 10^8 cells/ml) above background was reached. RNA was extracted as described in Chapter Three, using the newly described **RNA*express***™ method. Primer extension analysis was performed as described in Stead et al. (46) with an oligonucleotide primer specific to the 5' end of mature 23S rRNA which is identical for each of the rRNA operons (5'-CGTCCTTCATCGCCTCTGACT-3'). Two-hundred and fifty ng

of total RNA was used for the reverse transcription reactions, with only half of the total product run on a gel. Total RNA (250 ng) from each strain was also analyzed by high-resolution agarose gels, where the RNA was separated on 0.7% agarose 0.9% synergel (Diversified Biotech) 0.5X TBE gels run at 5 V/cm for 5 hours and stained with ethidium bromide.

REFERENCES

1. Drider, D. and Condon, C. (2004) The continuing story of endoribonuclease III. *J Mol Microbiol Biotechnol*, **8**, 195-200.
2. Li, Z., Pandit, S. and Deutscher, M.P. (1999) RNase G (CafA protein) and RNase E are both required for the 5' maturation of 16S ribosomal RNA. *EMBO J*, **18**, 2878-2885.
3. Babitzke, P., Granger, L., Olszewski, J. and Kushner, S.R. (1993) Analysis of mRNA decay and rRNA processing in *Escherichia coli* multiple mutants carrying a deletion in RNase III. *J Bacteriol*, **175**, 229-239.
4. Vesper, O., Amitai, S., Belitsky, M., Byrgazov, K., Kaberdina, A.C., Engelberg-Kulka, H. and Moll, I. (2011) Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in *Escherichia coli*. *Cell*, **147**, 147-157.
5. Tock, M.R., Walsh, A.P., Carroll, G. and McDowall, K.J. (2000) The CafA protein required for the 5'-maturation of 16 S rRNA is a 5'- end-dependent ribonuclease that has context-dependent broad sequence specificity. *J. Biol. Chem.*, **275**, 8726-8732.

6. Srivastava, A.K. and Schlessinger, D. (1989) *Escherichia coli 1*6S rRNA 3'-end formation requires a distal transfer RNA sequence at a proper distance. *EMBO Journal*, **8**, 3159-3166.

7. Sirdeshmukh, R. and Schlessinger, D. (1985) Ordered processing of *Escherichia coli* 23S rRNA *in vitro*. *Nucleic Acids Res*, **13**, 5041-5054.

8. Sirdeshmukh, R., Krych, M. and Schlessinger, D. (1985) *Escherichia coli* 23S ribosomal RNA truncated at its 5' terminus. *Nucleic Acids Res*, **13**, 1185-1192.

9. King, T.C., Sirdeshmukh, R. and Schlessinger, D. (1984) RNase III cleavage is obligate for maturation but not for function of *Escherichia coli* pre-23S rRNA. *Proceedings of the National Academy of Sciences of the United States of America*, **81**, 185-188.

10. Roy, M.K., Singh, B., Ray, B.K. and Apirion, D. (1983) Maturation of 5S rRNA: Ribonuclease E cleavages and their dependence on precursor sequences. *European Journal of Biochemistry*, **131**, 119-127.

11. Gutgsell, N.S. and Jain, C. (2010) Coordinated regulation of 23S rRNA maturation in *Escherichia coli*. *J Bacteriol*, **192**, 1405-1409.

12. Kiyota, E., Okada, R., Kondo, N., Hiraguri, A., Moriyama, H. and Fukuhara, T. (2011) An Arabidopsis RNase III-like protein, AtRTL2, cleaves double-stranded RNA *in vitro*. *J Plant Res*, **124**, 405-414.

13. Allas, U., Liiv, A. and Remme, J. (2003) Functional interaction between RNase III and the *Escherichia coli* ribosome. *BMC Mol Biol*, **4**, 8.

14. Misra, T.K. and Apirion, D. (1979) RNase E, an RNA processing enzyme from *Escherichia coli*. *Journal of Biological Chemistry*, **254**, 11154-11159.

15. Hicks, L.D., Warrier, I., Raghavan, R. and Minnick, M.F. (2011) Ribozyme Stability, Exon Skipping, and a Potential Role for RNA Helicase in Group I Intron Splicing by *Coxiella burnetii*. *J Bacteriol*, **193**, 5292-5299.

16. Wachi, M., Umitsuki, G., Shimizu, M., Takada, A. and Nagai, K. (1999) *Escherichia coli cafA* gene encodes a novel RNase, designated as RNase G, involved in processing of the 5' end of 16S rRNA. *Biochemical and Biophysical Research Communications*, **259**, 483-488.

17. Ow, M.C., Liu, Q. and Kushner, S.R. (2000) Analysis of mRNA decay and rRNA processing in *Escherichia coli* in the absence of RNase E-based degradosome assembly. *Mol. Microbiol.*, **38**, 854-866.

18. Li, Z., Pandit, S. and Deutscher, M.P. (1999) RNase G (CafA protein) and RNase E are both required for the 5' maturation of 16S ribosomal RNA. *EMBO J.*, **18**, 2878-2885.

19. Li, Z., Pandit, S. and Deutscher, M.P. (1999) Maturation of 23S ribosomal RNA requires the exoribonuclease RNase T. *RNA*, **5**, 139-146.

20. Cheng, Z.F. and Deutscher, M.P. (2003) Quality control of ribosomal RNA mediated by polynucleotide phosphorylase and RNase R. *Proc. Natl. Acad. Sci. U S A*, **100**, 6388-6393.

21. Li, Z. and Deutscher, M.P. (1995) The tRNA processing enzyme RNase T is essential for maturation of 5S RNA. *Proc. Natl. Acad. Sci. USA*, **92**, 6883-6886.

22. Sirdeshmukh, R. and Schlessinger, D. (1985) Why is processing of 23 S ribosomal RNA in *Escherichia coli* not obligate for its function? *J Mol Biol*, **186**, 669-672.

23. Jagessar, K.L. and Jain, C. (2010) Functional and molecular analysis of *Escherichia coli* strains lacking multiple DEAD-box helicases. *RNA*, **16**, 1386-1392.

24. Davies, B.W., Kohrer, C., Jacob, A.I., Simmons, L.A., Zhu, J., Aleman, L.M., Rajbhandary, U.L. and Walker, G.C. (2010) Role of *Escherichia coli* YbeY, a highly conserved protein, in rRNA processing. *Mol Microbiol*, **78**, 506-518.

25. Westra, E.R., Pul, U., Heidrich, N., Jore, M.M., Lundgren, M., Stratmann, T., Wurm, R., Raine, A., Mescher, M., Van Heereveld, L. *et al.* (2010) H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. *Mol Microbiol*, **77**, 1380-1393.

26. Diez-Villasenor, C., Almendros, C., Garcia-Martinez, J. and Mojica, F.J. (2010) Diversity of CRISPR loci in *Escherichia coli*. *Microbiology*, **156**, 1351-1361.

27. Pul, U., Wurm, R., Arslan, Z., Geissen, R., Hofmann, N. and Wagner, R. (2010) Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS. *Mol Microbiol*, **75**, 1495-1512.

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

29. Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J. and Charpentier, E. (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, **471**, 602-607.

30. Pougach, K., Semenova, E., Bogdanova, E., Datsenko, K.A., Djordjevic, M., Wanner, B.L. and Severinov, K. (2010) Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. *Mol Microbiol*, **77**, 1367-1379.

31. Takiff, H.E., Chen, S. and Court, D.L. (1989) Genetic analysis of the *rnc* operon of *Escherichia coli*. *Journal of Bacteriology*, **171**, 2581-2590.

32. Asai, T., Zaporojets, D., Squires, C. and Squires, C.L. (1999) An *Escherichia coli* strain with all chromosomal rRNA operons inactivated: Complete exchange of rRNA genes between bacteria. *Proc. Natl. Acad. Sci. USA*, **96**, 1971-1976.

33. Wang, R.F. and Kushner, S.R. (1991) Construction of versatile low-copy-number vectors for cloning, sequencing and expression in *Escherichia coli*. *Gene*, **100**, 195-199.

34. Ow, M.C., Liu, Q. and Kushner, S.R. (2000) Analysis of mRNA decay and rRNA processing in *Escherichia coli* in the absence of RNase E-based degradosome assembly. *Molecular Microbiol.*, **38**, 854-866.

35. Chung, D.H., Min, Z., Wang, B.C. and Kushner, S.R. (2010) Single amino acid changes in the predicted RNase H domain of *Escherichia coli* RNase

G lead to complementation of RNase E deletion mutants. *RNA*, **16**, 1371-1385.

36. Song, W.S., Lee, M. and Lee, K. (2011) RNase G participates in processing of the 5'-end of 23S ribosomal RNA. *J Microbiol*, **49**, 508-511.

37. Perwez, T. and Kushner, S.R. (2006) RNase Z in *Escherichia coli* plays a significant role in mRNA decay. *Molecular Microbiology*, **60**, 723-737.

38. Datsenko, K.A. and Wanner, B.A. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 6640-6645.

39. Rasouly, A., Schonbrun, M., Shenhar, Y. and Ron, E.Z. (2009) YbeY, a heat shock protein involved in translation in *Escherichia coli*. *J Bacteriol*, **191**, 2649-2655.

40. Rasouly, A., Davidovich, C. and Ron, E.Z. (2010) The heat shock protein YbeY is required for optimal activity of the 30S ribosomal subunit. *J Bacteriol*, **192**, 4592-4596.

41. Mohanty, B.K. and Kushner, S.R. (2010) Processing of the *Escherichia coli leuX* tRNA transcript, encoding tRNA(Leu5), requires either the 3'-5' exoribonuclease polynucleotide phosphorylase or RNase P to remove the Rho-independent transcription terminator. *Nucleic Acids Res*, **38**, 597-607.

42. Li, Z. and Deutscher, M.P. (1994) The role of individual exoribonucleases in processing at the 3' end of *Escherichia coli* tRNA precursors. *J. Biol. Chem.*, **269**, 6064-6071.

43. Li, Y., Cole, K. and Altman, S. (2003) The effect of a single, temperature-sensitive mutation on global gene expression in *Escherichia coli*. *RNA*, **9**, 518-532.

44. Yamamoto, N., Nakahigashi, K., Nakamichi, T., Yoshino, M., Takai, Y., Touda, Y., Furubayashi, A., Kinjyo, S., Dose, H., Hasegawa, M. *et al.* (2009) Update on the Keio collection of *Escherichia coli* single-gene deletion mutants. *Mol Syst Biol*, **5**, 335.

45. Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L. and Mori, H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol*, **2**, 2006 0008.

46. Stead, M.B., Marshburn, S., Mohanty, B.K., Mitra, J., Castillo, L.P., Ray, D., van Bakel, H., Hughes, T.R. and Kushner, S.R. (2011) Analysis of *Escherichia coli* RNase E and RNase III activity *in vivo* using tiling microarrays. *Nucleic Acids Res*, **39**, 3188-3203.

47. Ow, M.C. and Kushner, S.R. (2002) Initiation of tRNA maturation by RNase E is essential for cell viability in Escherichia coli. *Genes Dev*., **16**, 1102–1115.

48. Arraiano, C.M., Yancey, S.D., Kushner, S.R. (1988) Stabilization of discrete mRNA breakdown products in *ams pnp rnb* multiple mutants of *Escherichia coli* K-12. *J Bacteriol,* **170**, 4625-4633.

49. Zaporojets, D., French, S., Squires, C.L. (2003) Products transcribed from rearranged *rrn* genes of *Escherichia coli* can assemble to form functional ribosomes. *J Bacteriol,* **185**, 6921-6927

Table 1. A partial list of strains used in this study containing all rRNA operons

Strain Name	Genotype	Source
JW0603	*Δrna-749::kan**	(44,45)
JW2560	*ΔsrmB744::kan**	(44,45)
JW2726	*ΔygcH758::kan**	(44,45)
JW2731	*ΔygcB763::kan**	(44,45)
JW5438	*ΔygbF756::kan**	(44,45)
JW5531	*ΔdeaD774::kan**	(44,45)
MG1655	Wild type: F^- *rph-1* λ^-	*E. coli* genetic stock center (Yale)
MG1693	*thyA715*	*E. coli* genetic stock center (Yale)
SK2525	*rnpA49 thyA715*	(48)
SK4455	*rnc-14*::ΔTn*10 thyA715*	(46)
SK4477	*Δrnz::apr thyA715*	Unpublished, S.R. Kushner
SK4485	*rne-1 rnpA49 Δrnz::apr ΔrnlA rng::cat thyA715*	Unpublished, S.R. Kushner

SK5372	*ΔybeY::cat rnc-14*::*Δ*Tn*10*	This study
SK5374	*ΔybeY::cat rnpA49*	This study
SK5389	*rnc-14*::*Δ*Tn*10 thyA715* / pMSK9	This study
SK5390	*ΔybeY::cat*	(24)
SK5392	*rnc-14*::*Δ*Tn*10 thyA715* / pMSK10	This study
SK5665	*rne-1 thyA715*	(47)
SK7622	*rncΔ38 thyA715*	(3)
SK10525	*rnc-14*::*Δ*Tn*10 rnpA49 thyA715*	This study

*-also contains: *Δ(araD-araB)567 ΔlacZ4787*(::rrnB-3) *Δ(rhaD-rhaB)568 hsdR514*

All strains listed in Table 1 are constructed in the MG1655 wild-type genetic background.

Table 2. A partial list of single rRNA operon strains used in this study

Strain Name	Genotype	Source
SQ2158	pK4-16 (*rrnB* operon)	Unpublished, C. Squires
SK2767	*Δrna-749::kan* / pMSK3 (*rrnB* operon)	This study
SK2769	*Δrna-749::kan rnc-14*::*Δ*Tn*10* / pMSK3 (*rrnB* operon)	This study
SK2771	*ΔsrmB744::kan* / pMSK3 (*rrnB* operon)	This study
SK2773	*rnpA49 rnc-14*::*Δ*Tn*10* / pMSK3 (*rrnB* operon)	This study
SK2775	*ΔdeaD774::kan* / pMSK4 (*rrnG* operon)	This study
SK2776	*ΔdeaD774::kan* / pMSK3 (*rrnB* operon)	This study
SK5191	pMSK2 (*rrnD* operon)	This study
SK5192	pMSK3 (*rrnB* operon)	This study
SK5193	pMSK4 (*rrnG* operon)	This study
SK5199	*rnc-14*::*Δ*Tn*10* / pMSK3 (*rrnB* operon)	This study
SK5342	*rnpA49* / pMSK3 (*rrnB* operon)	This study
SK5344	*rnc-14*::*Δ*Tn*10* / pMSK4 (*rrnG* operon)	This study
SK5346	*rnpA49* / pMSK4 (*rrnG* operon)	This study
SK5348	*ΔygcH758::kan* / pMSK4 (*rrnG* operon)	This study
SK5350	*ΔygcH758::kan rnc-14*::*Δ*Tn*10* / pMSK4 (*rrnG* operon)	This study
SK5352	*ΔygcB763::kan* / pMSK4 (*rrnG* operon)	This study

SK5353	*ΔygbF756::kan* / pMSK4 (*rrnG* operon)	This study
SK5355	*ΔygbF756::kan rnc-14*::*Δ*Tn*10* / pMSK4 (*rrnG* operon)	This study
SK5357	*Δrna-749::kan* / pMSK4 (*rrnG* operon)	This study
SK5359	*Δrna-749::kan rnc-14*::*Δ*Tn*10* / pMSK4 (*rrnG* operon)	This study
SK5361	*ΔygcB763::kan rnc-14*::*Δ*Tn*10* / pMSK4 (*rrnG* operon)	This study
SK5362	*ΔsrmB744::kan* / pMSK4 (*rrnG* operon)	This study
SK5364	*ΔdeaD774::kan rnc-14*::*Δ*Tn*10* / pMSK4 (*rrnG* operon)	This study
SK5370	*rnpA49 rnc-14*::*Δ*Tn*10* / pMSK4 (*rrnG* operon)	This study
SK5376	*ΔybeY::cat* / pMSK4 (*rrnG* operon)	This study
SK5378	*ΔybeY::cat rnc-14*::*Δ*Tn*10* / pMSK4 (*rrnG* operon)	This study
SK5380	*Δrnz::apr* / pMSK4 (*rrnG* operon)	This study
SK5393	*rnc-14*::*Δ*Tn*10* / pMSK2 (*rrnD* operon)	This study
SK5398	pMSK2 (*rrnD* operon) pMSK10 (*rnc*+)	This study
SK5399	*rnc-14*::*Δ*Tn*10* / pMSK2 (*rrnD* operon) pMSK10 (*rnc*+)	This study

All strains in Table 2 were constructed in the MG1655 wild type genetic background (See Table 1) and additionally contain: *ΔrrnABCDEGH* / pTRNA67 (encoding all tRNAs from rRNA operons). See Table 3 for plasmid details.

Table 3. Plasmids used in work

Plasmid Name	Genotype	Source
pK4-16[1]	*rrnB* rRNA operon, Kan^R	Unpublished, C. Squires
ptRNA67[2]	Asp-1 Trp Ile-1 Ala-1B Thr-1 tRNAs, Str^R/Spc^R	(48)
pMSK2[1]	*rrnD* rRNA operon, Ap^R	This study
pMSK3[1]	*rrnB* rRNA operon, Apr^R	This study
pMSK4[1]	*rrnG* rRNA operon, Ap^R	This study
pMSK9[1]	*rnc+,* Cm^R (P_{BAD} RNase III expression vector)	This study
pMSK10[3]	*rnc+*, Ap^R (P_{BAD} RNase III expression vector)	This study

[1]- pSC101 replication origin, 6-8 copies per cell

[2]- p15A replication origin, ~15 copies per cell

[3]- Mini-F replication origin, 1-2 copies per cell

Fig. 1. Ribosomal RNA operons in *E. coli.*

This model is not drawn to scale. Each of the seven operons are identified on the left side by their respective *rrn* designations. Mature rRNAs or tRNAs are indicated by the light blue bars. (From Chapter One, Fig. 2)

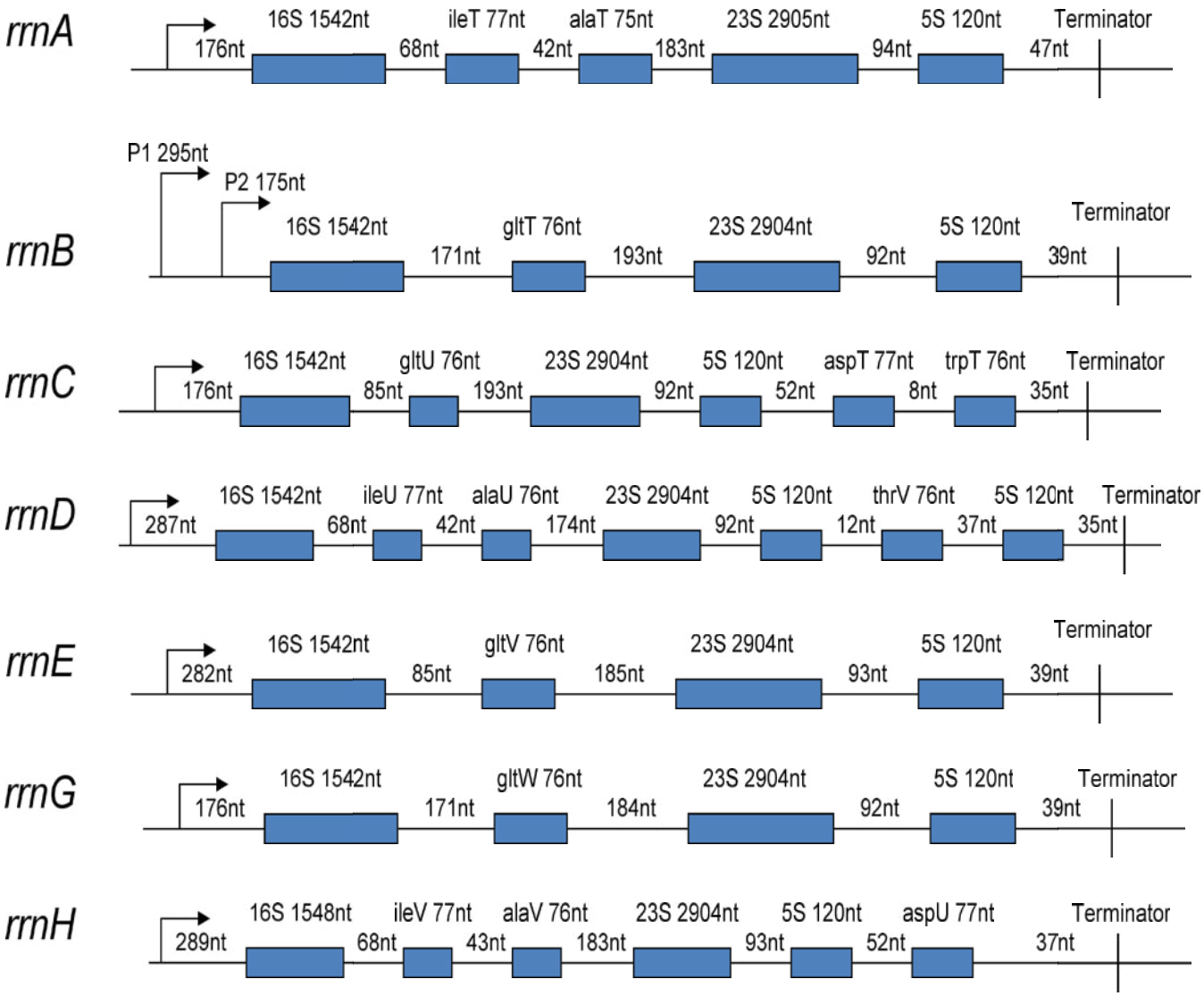

rrnA
176nt
16S 1542nt
68nt
ileT 77nt
42nt
alaT 75nt
183nt
23S 2905nt
94nt
5S 120nt
47nt
Terminator
rrnB
P1 295nt
P2 175nt
16S 1542nt
171nt
gltT 76nt
193nt
23S 2904nt
92nt
5S 120nt
39nt
Terminator
rrnC
176nt
16S 1542nt
85nt
gltU 76nt
193nt
23S 2904nt
92nt
5S 120nt
52nt
aspT 77nt
8nt
trpT 76nt
35nt
Terminator
rrnD
287nt
16S 1542nt
68nt
ileU 77nt
42nt
alaU 76nt
174nt
23S 2904nt
92nt
5S 120nt
12nt
thrV 76nt
37nt
5S 120nt
35nt
Terminator
rrnE
282nt
16S 1542nt
85nt
gltV 76nt
185nt
23S 2904nt
93nt
5S 120nt
39nt
Terminator
rrnG
176nt
16S 1542nt
171nt
gltW 76nt
184nt
23S 2904nt
92nt
5S 120nt
39nt
Terminator
rrnH
289nt
16S 1548nt
68nt
ileV 77nt
43nt
alaV 76nt
183nt
23S 2904nt
93nt
5S 120nt
52nt
aspU 77nt
37nt
Terminator

Fig. 2. Model of *E. coli* rRNA processing.

This model is not drawn to scale. The 30S pre-rRNA transcript is first cleaved by RNase III at two physical hairpin locations, cleaving the RNA the hairpin structures formed by the spacer sequences adjacent to mature 16S and 23S rRNAs, generating 17S, 25S, and 9S pre-rRNAs (labeled in purple). 17S rRNA is cleaved first by RNase E and then by RNase G at the mature 5' end of 16S rRNA. RNase E also cleaves 9S three nt on each side of the mature termini of 5S rRNA forming p5S. RNase P cleaves at the mature 5' end of the tRNA. Exoribonucleases (primarily RNase T) are responsible for the 3' end maturation of the tRNA, 23S rRNA, and 5S rRNA, but the endoribonuclease(s) responsible for the 3' end maturation of 16S, the 5' end of 23S, and 5' end of 5S rRNAs remain unidentified (labeled in red). (From Chapter One, Fig. 1)

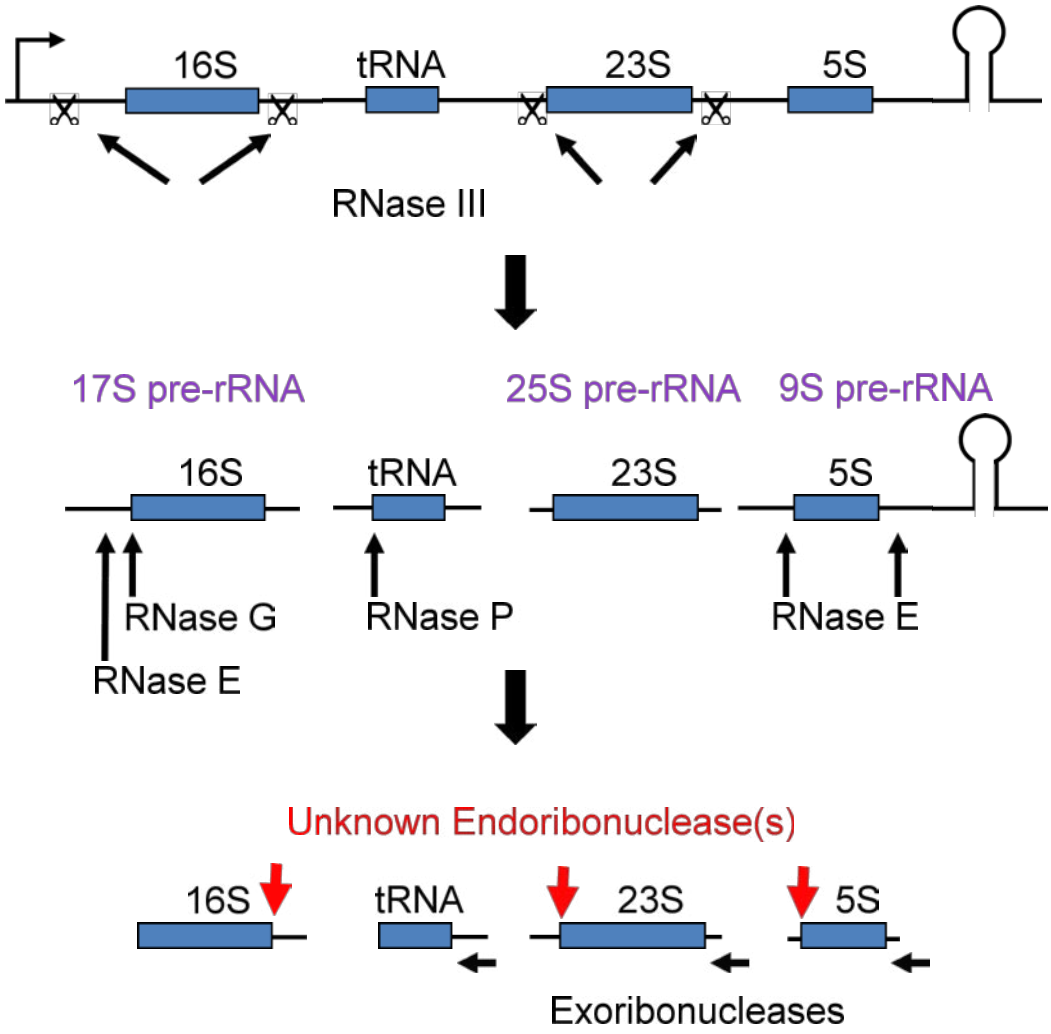

16S
tRNA
23S
5S
RNase III
17S pre-rRNA
25S pre-rRNA
9S pre-rRNA
16S
tRNA
23S
5S
RNase G
RNase E
RNase P
RNase E
Unknown Endoribonuclease(s)
16S
tRNA
23S
5S
Exoribonucleases

Fig. 3. rRNA processing profiles in a variety of mutant strains.

Two-hundred and fifty ng of total RNA from isogenic strains with mutations labeled upon each lane, was electrophoresed on a 0.7% agarose 0.9% Synergel 0.5X TBE high-resolution agarose gel and stained with ethidium bromide. The first three lanes contain total RNA from MG1693 (wild-type) derivative strains containing all seven rRNA operons, while the other lanes contain total RNA from SQ2158 (single *rrn* operon strain) derivative strains containing only the *rrnG* rRNA operon. Other mutations in the strains are as labeled. The locations of the 30S, 25S, and 17S pre-rRNAs, along with 23S and 16S rRNAs, are indicated on the right side of the gel image.

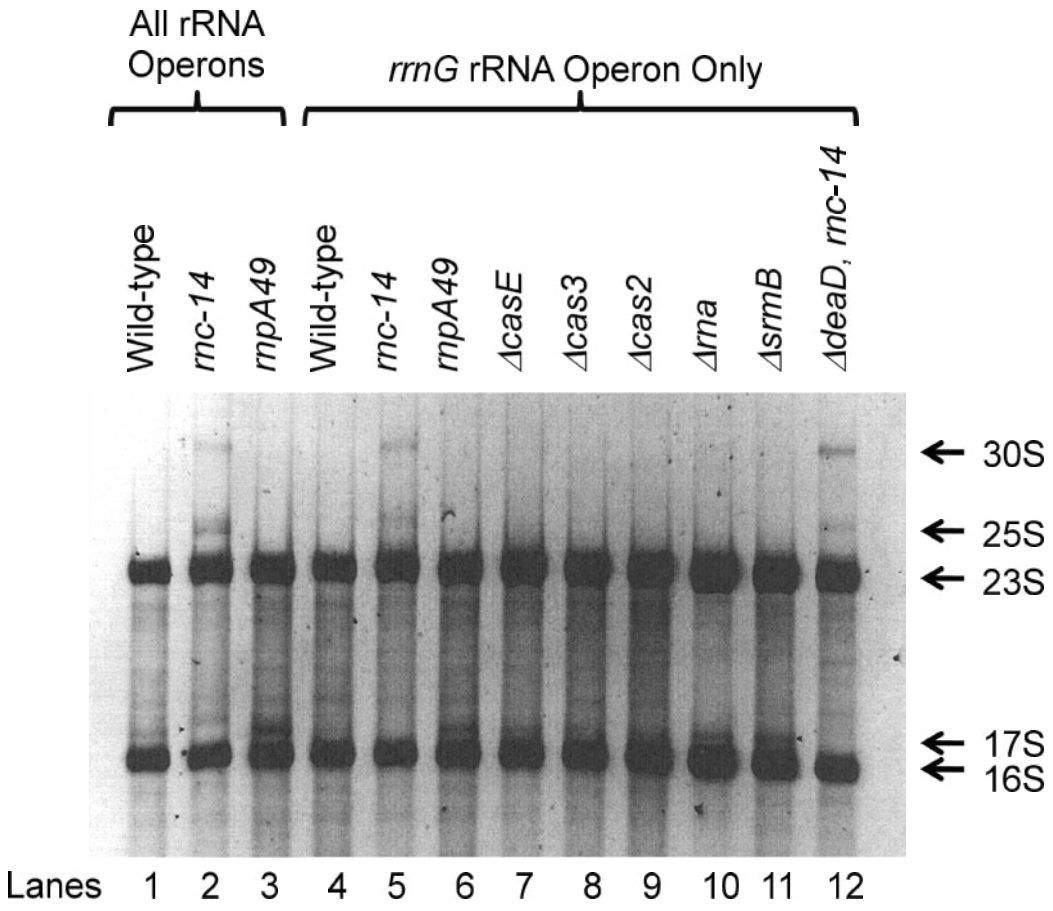
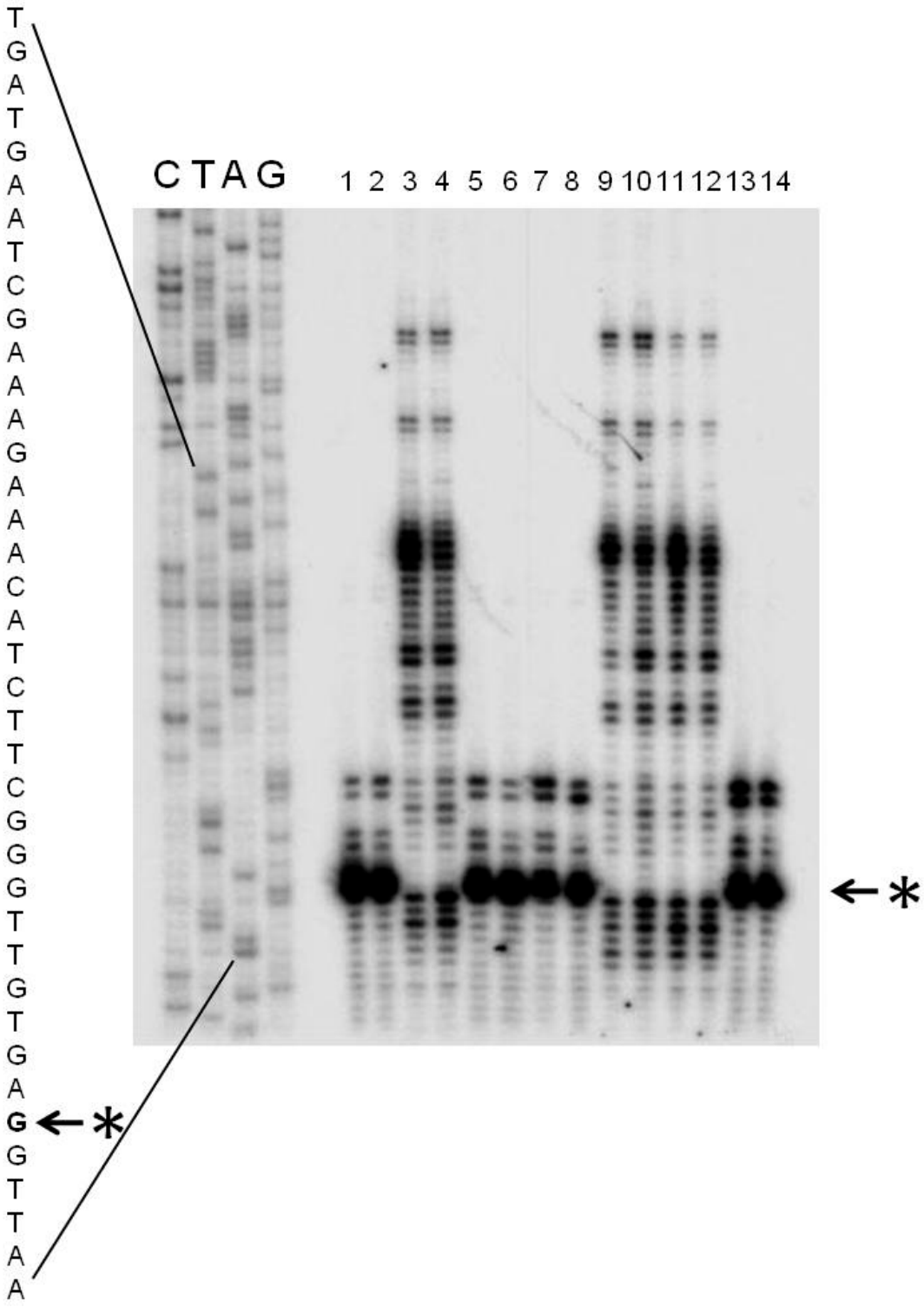

All rRNA Operons
rrnG rRNA Operon Only
Wild-type
rnc-14
rnpA49
Wild-type
rnc-14
rnpA49
ΔcasE
Δcas3
Δcas2
Δrna
ΔsmrB
ΔdeaD, rnc-14
30S
25S
23S
17S
16S
Lanes 1 2 3 4 5 6 7 8 9 10 11 12

Fig. 4. Primer extension analysis of the 5' end of 23S rRNA.

Primer extension analysis was conducted as described in Materials and Methods. The asterisks (*) indicate the location of the primary 5' mature end of 23S rRNA in wild type cells. The nucleotide sequence of the sequencing ladder used (as labeled CTAG on the top of the autoradiogram) was derived from the *rrnB* operon, and is expanded on the left side for clarity. Lanes 1,2, wild type; 3,4, *rnc-14*; 5,6, *rne-1;* 7,8, *rnpA49*; 9,10, *rnc-14 rnpA49*; 11,12, *rnc-14 ΔybeY;* 13,14, *rnpA49 ΔybeY.* Odd-numbered lanes contain extension products from RNA isolated from strains grown at 30°C, while even number lanes are shifted from 30°C to 42°C for 15 minutes. All bacterial strains used in the generation of this figure contain all seven rRNA operons.


T
G
A
T
G
A
A
T
C
G
A
A
A
A
G
A
A
A
C
A
T
C
T
T
T
C
G
G
G
G
T
T
G
T
G
A
G
G
T
T
A
A
C T A G
1 2 3 4 5 6 7 8 9 10 11 12 13 14
←*
←*

CHAPTER 5

CONCLUSIONS

My thesis research has explored the contributions of various endoribonucleases in post-transcriptional regulation of *Escherichia coli* gene expression. In Chapter 2 the roles of RNase E and RNase III in gene expression were examined using tiling microarrays which allowed an unprecedented view of the total transcriptome-wide activity of both enzymes *in vivo.* The work demonstrated that RNase E had much more wide-spread roles in the abundance of functional non-coding RNAs than previously envisioned. With approximately 75% of annotated small RNAs being affected by RNase E, and each sRNA having multiple mRNA targets, the implications for secondary effects of RNase E activity are potentially quite large. In addition, mining of the array data for areas of increased or decreased abundance, which were not attributable to known genes, allowed for the identification of more than 300 loci that are potentially novel genes. If even a small percentage of these loci (each of which are approximately the size of an average sRNA) are actually sRNAs or small ORFs, the potential impact would be large and would indicate an ever larger functional

role of RNase E activity on gene expression. RNase E was also thought to affect the stability of polycistronic mRNA transcripts such that some coding sequences could be less stable than others within the same primary transcript. However, the arrays demonstrated that for the vast majority of transcripts, RNase E did not play a significant role in differential stabilities, with most polycistronic mRNAs being uniformly affected by the absence of RNase E. However, it should be noted that the generation times of the wild type and RNase E deletion strains were significantly different.

Prior to the work described in Chapter 2, RNase III was thought to be active on only a relatively small set of substrates, primarily rRNAs, a few polycistronic mRNAs, and several sRNAs (1-4). In fact, our analysis has demonstrated that RNase III affects nearly 500 coding sequences (ten-fold more potential substrates than previously known), along with 12% of the annotated sRNAs. Interestingly, there also appears to be a large amount of overlap in activity between RNase E and RNase III, although it is unclear in the vast majority of cases whether both enzymes are functioning in the same pathway, or if they have different roles in the regulation of a single transcript. Furthermore, by taking advantage of the 20 nt resolution of the arrays, we were able to identify an RNase III cleavage site within an mRNA coding sequence, which is the first reported activity of its kind for RNase III in *E. coli*. Additionally, RNase III was also shown to significantly affect a number of biological pathways including heat response and iron transport, cysteine biosynthesis, and the protein composition of the membrane.

Although the work in Chapter 2 provides much needed detail in our understanding of the transcriptome-wide roles of RNase E and RNase III, it is by no means the end of the story. A number of very important questions remain to be addressed, such as whether any of the more than 300 loci of unexplained RNA abundance changes in the RNase E deletion strain are actually novel genes. Also, with the RNA transcripts, that are affected by both RNase E and RNase III, what roles do the enzymes play in the regulation of stability? In other words, is one enzyme required for efficient translation, while the other degrades the message? How many of the RNA abundance changes in each of the mutants are caused directly through cleavage by the endoribonucleases, and how many are secondary effects? Why are the majority of mRNAs which are targets of sRNAs not affected by alterations in the amount of sRNAs in the RNase E or RNase III mutant strains? Are genes unaffected by either RNase E or RNase III merely not transcribed during exponential phase, aerobic growth in rich medium? What role does the Rng-219 protein (the modified RNase G protein used in the RNase E deletion strain) perform that allows complementation of the loss of RNase E activity?

While many of these questions will likely remain unanswered for some time, the last question is particularly interesting as the essential function of RNase E remains to be identified (1,5). Future experiments to identify the activity of RNase E or Rng-219 will likely involve tiling array analysis of RNA from strains depleted of either RNase E or Rng-219, through the use of controlled-expression plasmids expressing the endoribonuclease of interest. For example, if Rng-219

were put onto a single-copy arabinose-inducible plasmid, such as the vector constructed for RNase III described in Chapter 4, the strain could be grown in arabinose, an aliquot of bacteria harvested, then the culture spiked with glucose to stop transcription of *rng-219*. After growth and viability curves are used to establish an appropriate harvest point following the addition of glucose, the RNA could be compared with the sample prior to glucose addition, and tiling array analysis should be able to determine which transcript levels are altered by the loss of Rng-219 activity. This experiment, while not perfect, should aid our search for the essential function of RNase E, and allow a more comprehensive analysis of the data presented in Chapter 2, since we could essentially control for the presence of Rng-219.

Other questions, such as whether any of the potentially novel genes are actually real, could be simply addressed by Northern analysis to search for a stable, discreet transcript originating from the specific location on the genome, followed by screening for open reading frames to determine if the RNA is a small ORF, or a potential ncRNA. Validation of these potential new genes is important, and could have a significant impact on the sRNA field.

In addition to questions regarding RNase E and RNase III, many other endoribonucleases in *E. coli* have yet to be examined for their transcriptome-wide impacts. Just as the array analyses have opened new avenues of experimentation on RNase E and RNase III, the same will likely happen for the other endoribonucleases (RNase I, RNase P, RNase Z, RNase G, and RNase LS). It will be particularly interesting to determine if different endoribonucleases

control different biological systems. For example, RNase III was shown to affect the response to heat but not response to cold. Therefore, it would be interesting if another endoribonuclease had the opposite effect, affecting cold response, but not heat. If so, endoribonucleases could be viewed as a type regulatory switch to effect gene expression changes in response to the environment.

Although transcriptome-wide gene expression data are useful, there are still a number of limitations dealing with the use of tiling microarray analysis to determine the effects of ribonucleases. One major problem, for which I do not have a simple answer, is simply the way the bacteria are grown. In aerobic, exponential-phase growth in rich medium not all of the genes in *E. coli* are expressed, as many are required only for very specific conditions. So how is it possible to gain a true understanding of ribonuclease activity on a whole without exploring its role in every stress and growth condition imaginable? Obviously, such open-ended experiments are not feasible, but the fact remains that only looking at a very particular phase of growth in a very particular growth condition is probably not the best way to gather the total transcriptome impact of an enzyme. Additionally, because of the chemistry employed in array hybridizations, a number of RNA abundance changes may actually be nothing more than experimental artifacts due to non-specific hybridization, even though a number of important steps are taken during an array experiment to reduce such effects. Even if the array technology used in Chapter 2 is completely supplanted by that of RNAseq, unfortunately many issues still remain, such as a way to

determine primary versus secondary effects of an endoribonuclease on a transcriptome-wide scale.

Even considering the limitations of transcriptome-wide RNA analysis, and consequently the many questions that remain unaddressed with regard to the role of RNase E and RNase III in *Escherichia coli*, the data in Chapter 2 represent a significant step forward in our understanding of the role of ribonucleases on gene expression. Once similar analyses are conducted with other endoribonuclease mutant strains, it is my expectation that the data will suggest a master regulatory role of endoribonucleases in pathways which are required to change rapidly in response to stimuli, especially in stress response pathways where *de novo* RNA synthesis alone may not be rapid enough to avoid cell death. However, it is clear that the role of endoribonucleases in gene expression was underestimated prior to the work described in Chapter 2.

In all transcriptome-wide gene expression experiments, the first step is to isolate high-quality RNA from the organism of interest. However, no commercial RNA extraction kit is perfect. The work in Chapter 3 shows that each of the commonly used commercial RNA extraction kits fail to provide quantitative recovery of all RNA species in *E. coli,* which can cause misleading results in any RNA-based experiment. Furthermore, while each of the kits represents an improvement over the traditional RNA extraction methods (i.e., hot phenol, or CsCl gradients) the methods can still be cumbersome to perform, and each isolation can cost as much as $8 per sample. Because of these limitations, we developed a novel method for the isolation of high-quality total RNA from all

bacteria, which takes less than 15 minutes to perform, costs as little as three cents per sample, and offers quantitative recovery of RNA regardless of size (Chapter 3).

The RNA extraction method described in Chapter 3 has the potential to change the way microbiology labs work with RNA, as no method can compare to its simplicity, ease of use, low cost, and quantitative yields. All methods have potential drawbacks, however. The weak link in our new method is the precipitation of the RNA from the RNA extraction solution, such that the RNA can be suspended in water for further enzymatic reactions. Sodium acetate/ethanol precipitations take time (about 2 hours), and may not always be quantitative depending on the skill of the experimenter. We have demonstrated that the RNA can be recovered from the extraction solution with columns, but silica columns tend to lose RNA molecules less than 200 nt. While these problems are not insurmountable, they need to be addressed in order for the technique to be accessible to a large number of laboratories that may not work with RNA on a regular basis. Future work with this method should also attempt RNA extractions from organisms outside of bacteria, such as eukaryotes and more specifically, human cells.

Gene expression analysis of cancer cells, for example, has become a hot topic in oncology, and like all RNA extraction methods the ones utilized for this most important task fall short of expectations (6). Adaption of the methods described in Chapter 3 to permit the RNA isolation from tissue samples could be of a large clinical benefit, as well as a boon to research in eukaryotes as well as

prokaryotes. The RNA extraction method described in Chapter 3 has also lent itself to the study of rRNA processing described in Chapter 4, as the 30S pre-rRNA transcript (being much larger in size than mature RNAs) was being non-quantitatively recovered in relation to smaller RNA species using the Catrimide/ LiCl RNA extraction method.

Ribosomal RNA was one of the first transcripts examined in *E. coli,* yet many aspects of its endoribonucleolytic remain to be elucidated. The work described in Chapter 4 lays the foundation for an innovative approach towards the determination of the processing steps involved in the maturation of the 5' end of 23S rRNA, and in the separation of 16S and 23S rRNA precursors in the absence of RNase III. While the work is not yet complete, a number of interesting hypotheses, and many more questions, have emerged from the data thus far. One of the most interesting questions is: What is the role of YbeY in rRNA processing? Is it an endoribonuclease? How does it affect the processing of every termini of 16S, 23S and 5S rRNAs (7)? Why does an RNase P mutant exacerbate the effects of a YbeY mutant?

One hypothesis posited in Chapter 4 involves the role of YbeY as a potential recruiter of ribosomal proteins to aid in the binding of the 30S rRNA transcript before processing occurs. This idea has merit, but does raise more questions. For instance, why does the YbeY protein not pull-down with ribosomes (7-9)? I think that the data presented may not be sufficient in eliminating a ribosome-YbeY interaction, though I am not convinced that a stable interaction is necessary for the recruitment of proteins to the rRNA transcript. At

this point in time, there is simply not enough known about rRNA processing, especially in relation to the ribosomal proteins which bind it, to predict exactly the processing pathways involved, but I think that the lines of experimentation started in Chapter 4 are bringing the field closer to an explanation.

In my estimation there will be no single “silver bullet” mutation that will cause massive defects in rRNA processing. Instead I believe we are now looking for a variety of small cumulative effects from a large number of enzymes, probably including RNase P, YbeY, DeaD, and SrmB. Unfortunately, the effects of multiple enzymes on a single processing step are going to be difficult to elucidate experimentally, but the use of the tools developed in Chapter 4 should be of use in the determination of the pathway.

While my thesis has covered three distinct areas (transcriptome analysis, methods, and the processing of rRNA), each aspect of the work has helped to shape our view of post-transcriptional regulation in *Escherichia coli.* Like many research projects, however, the work has generated more questions than it has answered, as the role of ribonucleases in the regulation of homeostasis seems to become more complex with each new finding. The data contained in this thesis has demonstrated that the transcriptome-wide activity of at least two endoribonucleases has been significantly underestimated, the methods all labs have relied upon for the basis of transcriptome-wide analysis have been inadequate, and that we still do not understand the processing of the single RNA molecule which has been studied for the longest period of time. And though these problems were clearly demonstrated in the work, a number of solutions

and answers were also revealed. While we do not understand all of the data yet, both RNase E and RNase III are clearly involved in the regulation of a significant portion of the genome, we presented a brand new method for RNA extraction that is superior to every previously described method, and we have set up the tools necessary for the discovery of the elusive processing pathway of ribosomal RNA. These findings will aid future work in the field, and help demonstrate how important post-transcriptional regulation truly is in the maintenance of biological systems.

REFERENCES

1. Stead, M.B., Marshburn, S., Mohanty, B.K., Mitra, J., Castillo, L.P., Ray, D., van Bakel, H., Hughes, T.R. and Kushner, S.R. (2011) Analysis of *Escherichia coli* RNase E and RNase III activity *in vivo* using tiling microarrays. *Nucleic Acids Res*, **39**, 3188-3203.
2. Opdyke, J.A., Fozo, E.M., Hemm, M.R. and Storz, G. (2011) RNase III participates in GadY-dependent cleavage of the *gadX-gadW* mRNA. *J Mol Biol*, **406**, 29-43.
3. Kim, K.S., Manasherob, R. and Cohen, S.N. (2008) YmdB: a stress-responsive ribonuclease-binding regulator of *E. coli* RNase III activity. *Genes Dev*, **22**, 3497-3508.
4. Drider, D. and Condon, C. (2004) The continuing story of endoribonuclease III. *J Mol Microbiol Biotechnol*, **8**, 195-200.
5. Ow, M.C. and Kushner, S.R. (2002) Initiation of tRNA maturation by RNase E is essential for cell viability in *E. coli*. *Genes Dev*, **16**, 1102-1115.

6. Fassan, M., Sachsenmeir, K., Rugge, M. and Baffa, R. (2011) Role of miRNA in distinguishing primary brain tumors from secondary tumors metastatic to the brain. *Front Biosci (Schol Ed)*, **3**, 970-979.

7. Davies, B.W., Kohrer, C., Jacob, A.I., Simmons, L.A., Zhu, J., Aleman, L.M., Rajbhandary, U.L. and Walker, G.C. (2010) Role of *Escherichia coli* YbeY, a highly conserved protein, in rRNA processing. *Mol Microbiol*, **78**, 506-518.

8. Rasouly, A., Davidovich, C. and Ron, E.Z. (2010) The heat shock protein YbeY is required for optimal activity of the 30S ribosomal subunit. *J Bacteriol*, **192**, 4592-4596.

9. Rasouly, A., Schonbrun, M., Shenhar, Y. and Ron, E.Z. (2009) YbeY, a heat shock protein involved in translation in *Escherichia coli*. *J Bacteriol*, **191**, 2649-2655.