POLYADENYLATION AND PROCESSING OF THE
POLYCISTRONIC RPSJ MESSENGER RNA TRANSCRIPT
IN ESCHERICHIA COLI K-12

by

KRISTEN BROWN MILDENHALL
(Under the Direction of Sidney R. Kushner)

ABSTRACT

The processing or decay of RNA in *Escherichia coli* involves an array of endonucleases and exonucleases. In addition, the decay of RNA transcripts also involves the post-transcriptional addition of 3’ poly(A) tails to mRNAs, rRNAs, and pre-tRNAs by poly(A) polymerase I (PAP I) and polynucleotide phosphorylase (PNPase). The addition of poly(A) tails helps target the RNAs for more rapid degradation. In this work, the processing and/or breakdown along with polyadenylation of the large polycistronic *rpsJ* mRNA transcript has been examined.

Our data suggest that RNase E is the major enzyme involved in the initial processing of the *rpsJ* transcript and is responsible for the removal of the 3’ Rho-independent transcription terminator. The results also suggest that the 5’ untranslated region of the *rpsJ* transcript remains intact in wild-type cells. Additionally, the processing pattern of the *rpsJ* transcript indicates that other endonucleases generate secondary cleavages. Interestingly, the cell maintains constant ratios of the individual ORF transcripts of the polycistronic *rpsJ* operon, despite changes in the processing of the
full-length transcript in RNase E mutants. These results suggest a greater flexibility in controlling ribosomal protein transcript levels than previously observed.

Furthermore, through our experiments involving the polyadenylation of the *rpsJ* operon transcript, we show that polyadenylation of mRNAs by PAP I is directly linked to an intact RNase E-based degradosome. The polyadenylation profiles of individual mRNAs encoded within the *rpsJ* operon (*rpsJ* and *rpsQ*) and a control transcript (*lpp*) changed dramatically both in the location of the tails and in their nucleotide composition when the scaffold region of the RNase E protein was deleted. Specifically, in the absence of degradosome assembly, the majority of the tails for these specific mRNA transcripts were heteropolymeric and were added by PNPase. Our data suggest the existence of a much larger polyadenylation complex and a more direct link between mRNA degradation and polyadenylation than previously envisioned.

INDEX WORDS: Polyadenylation, mRNA processing, RNase E, PAP I, PNPase, degradosome, *rpsJ*, *Escherichia coli*
POLYADENYLATION AND PROCESSING OF THE POLYCISTRONIC RPSJ MESSENGER RNA TRANSCRIPT IN ESCHERICHIA COLI K-12

by

KRISTEN BROWN MILDENHALL
B.S., Brigham Young University, 2007.

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
2013
POLYADENYLATION AND PROCESSING OF THE
POLYCISTRONIC RPSJ MESSENGER RNA TRANSCRIPT
IN ESCHERICHIA COLI K-12

by

KRISTEN BROWN MILDENHALL

Major Professor: Sidney R. Kushner
Committee: Lawrence Shimkets
Anna Karls
Robert Maier

Electronic Version Approved:
Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2013
DEDICATION

To the two most important people in my life—my husband, Taylor, and my son, Clark. And for my father-in-law, Tom—I feel privileged to vicariously carry on the family scientific PhD legacy; I only wish you had lived to see it to its fulfillment.
ACKNOWLEDGEMENTS

I would like to express my gratitude to my major advisor, Dr. Sidney R. Kushner, for his guidance, support and encouragement. I would also like to thank the other members of my advisory committee for their direction and suggestions. I also wish to thank the past and present members of the Kushner lab for their guidance, answers to questions, commiseration, encouragement, camaraderie, and support through the past few years.

Many thanks to my dear family and friends, who have supported and encouraged me throughout this process. Without you all, I would have never accomplished what I have been able to do.
# TABLE OF CONTENTS

**Page**

ACKNOWLEDGEMENTS.........................................................................................................................v

LIST OF TABLES........................................................................................................................................ix

LIST OF FIGURES .....................................................................................................................................x

### CHAPTER

1  INTRODUCTION AND LITERATURE REVIEW .................................................................................1
   INTRODUCTION .....................................................................................................................................1
   RIBONUCLEASES .................................................................................................................................1
   THE RNASE E-BASED DEGRADOSOME .............................................................................................9
   POLYADENYLATION IN *E. coli* ..........................................................................................................14
   THE *RPSJ* OPERON IN *E. coli*........................................................................................................22
   FIGURES ...............................................................................................................................................27
   REFERENCES .........................................................................................................................................30

2  RNASE E IS THE MAJOR ENDONUCLEASE RESPONSIBLE FOR PROCESSING THE POLYCISTRONIC *RPSJ* MESSENGER RNA
   ABSTRACT ............................................................................................................................................54
   INTRODUCTION ....................................................................................................................................55
   MATERIALS AND METHODS .............................................................................................................56
   RESULTS ..............................................................................................................................................60
3 POLYADENYLATION OF MESSENGER RNAs BY POLY(A) POLYMERASE I IN ESCHERICHIA COLI IS LINKED TO THE PRESENCE OF AN INTACT RNASE E-BASED DEGRADOSOME ........ 96

ABSTRACT ................................................................................................................. 97

INTRODUCTION ........................................................................................................... 98

MATERIALS AND METHODS .................................................................................... 101

RESULTS .................................................................................................................... 106

DISCUSSION ............................................................................................................... 113

TABLES AND FIGURES ........................................................................................... 120

REFERENCES ............................................................................................................. 134

4 CONCLUSIONS AND FUTURE DIRECTIONS ....................................................... 140

PROCESSING OF THE RPSJ OPERON ........................................................................ 141

POLYADENYLATION IN E. COLI ................................................................................... 141

FINAL CONCLUSION .................................................................................................. 144

REFERENCES ............................................................................................................. 145

APPENDICES ............................................................................................................. 146

A DELETION MUTANTS OF TRANSFER RNA_{LEU}^{2} ARE VIABLE AND EXHIBIT FLEXIBILITY IN CODON-ANTICODON RECOGNITION IN ESCHERICHIA COLI ........................................................................................................... 146

ABSTRACT ................................................................................................................. 147
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td>148</td>
</tr>
<tr>
<td>MATERIALS AND METHODS</td>
<td>149</td>
</tr>
<tr>
<td>RESULTS</td>
<td>152</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>154</td>
</tr>
<tr>
<td>TABLES AND FIGURES</td>
<td>156</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>165</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 2.1: List of bacterial strains and plasmids used in this study.................................74
Table 2.2: Oligonucleotide probes used in this study........................................................75
Table 2.3: Quantification of full-length transcript levels for agarose northern blots
   in various RNase E mutants.........................................................................................81
Table 3.1: List of bacterial strains and plasmids used in this study.................................121
Table 3.2: Previously published activities of PNPase point mutants compared to
   a wild-type control ........................................................................................................122
Table 3.3: Comparison of the polyadenylation profiles of degradosome mutants...........123
Table 3.4: Comparison of the polyadenylation profiles of PNPase mutants....................129
Table 3.5: Comparison of the polyadenylation profiles of strains carrying the
   pnp-7 allele.....................................................................................................................131
Table A.1: List of bacterial strains and plasmids used in this study ..............................156
Table A.2: Growth of strains on Luria broth plates after 48 hrs.....................................158
Table A.3: Cold sensitivity of SK10473..........................................................................158
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Model of mRNA decay pathways in <em>E. coli</em></td>
<td>27</td>
</tr>
<tr>
<td>1.2</td>
<td>Model of rRNA processing pathways in <em>E. coli</em></td>
<td>28</td>
</tr>
<tr>
<td>1.3</td>
<td>Schematic of wild-type RNase E protein</td>
<td>29</td>
</tr>
<tr>
<td>1.4</td>
<td>The <em>rpsJ</em> transcript</td>
<td>29</td>
</tr>
<tr>
<td>2.1</td>
<td>The <em>rpsJ</em> transcript</td>
<td>72</td>
</tr>
<tr>
<td>2.2</td>
<td>The 5′ UTR and 3′ Rho-independent transcription terminator of the <em>rpsJ</em> operon</td>
<td>73</td>
</tr>
<tr>
<td>2.3</td>
<td>Agarose northern blot of endonuclease mutants probed with 3′ <em>rpsJ</em> probe</td>
<td>76</td>
</tr>
<tr>
<td>2.4</td>
<td>RNase III is not involved in <em>rpsJ</em> processing</td>
<td>77</td>
</tr>
<tr>
<td>2.5</td>
<td>Agarose and polyacrylamide northern blots of various RNase E mutants probed with 5′ <em>rpsJ</em> UTR probe</td>
<td>78</td>
</tr>
<tr>
<td>2.6</td>
<td>Agarose and polyacrylamide northern blots of various RNase E mutants probed with the 3′ Rho-independent transcription terminator probe</td>
<td>79</td>
</tr>
<tr>
<td>2.7</td>
<td>Agarose northern blots of endonuclease or endo/exonuclease mutants probed for <em>rpsJ</em>, <em>rpsS</em>, <em>rpsC</em>, <em>rplP</em>, and <em>rpsQ</em></td>
<td>80</td>
</tr>
<tr>
<td>2.8</td>
<td>Polyacrylamide northern blots of various RNase E mutants probed for <em>rpsJ</em>, <em>rpsS</em>, <em>rplP</em>, and <em>rpsQ</em></td>
<td>82</td>
</tr>
</tbody>
</table>
Figure 2.9: Agarose northern blots of wild-type and rne-1 strains probed for rplC, rplD, rplW, rplB, rplV, and rpmC.................................................................................................83

Figure 2.10: Agarose northern blot of various RNase E mutants in the presence or absence of poly(A) polymerase I (PAP I) ........................................................................84

Figure 2.11: Differences in the rpsJ operon and the 16S rRNA in the rne-1 array ....85

Figure 2.12: Microarray comparison for the rpsJ ribosomal operon when strains were grown at comparable growth rates .................................................................87

Figure 2.13: Analysis of rpsJ transcript from strains grown at comparable growth rates in minimal medium .................................................................................................88

Figure 2.14: Agarose northern blots of wild-type and rneΔ1018/rng-219 strains grown in M9 minimal medium and probed for several intercistronic regions ........................................................................................................89

Figure 2.15: Chemical half-lives of the full-length rpsJ transcript .................................................................................................................................90

Figure 2.16: Chemical half-lives of the rpsJ operon transcript, the rpnN operon transcript, and adh transcript in SK4161 .................................................................................91

Figure 3.1: Pictorial representation of the rpsJ operon, also known as the S10 operon ......................................................................................................................120

Figure 3.2: Description of the RNase E degradosome mutants used this study ....122

Figure 3.3: Pictorial representation of tails found on the rpsQ transcript ............124

Figure 3.4: Pictorial representation of tails found on the lpp transcript ...............125

Figure 3.5: Percent composition of the post-transcriptionally added tails in various degradosome mutants ..........................................................................................126

Figure 3.6: Growth curves for various PNPase mutants .....................................127
Figure 3.7: Immunoprecipitation of PNPase and scaffold mutants ..................128

Figure 3.8: Shift in location of homopolymeric tails on rpsQ, lpp, and rpsJ transcripts for wild-type versus PNPase mutants ..................130

Figure 3.9: Shift in location of homopolymeric tails on rpsQ, lpp, and rpsJ transcripts for wild-type versus other PNPase mutants ..................132

Figure 3.10: Poly(A) sizing assay for various mutants ........................................133

Figure A.1: The presence of the leuU deletion in transductants was confirmed by Northern blotting ........................................157

Figure A.2: Growth curve and viability curve of SK10456 and SK10457 ΔsecG leuU mutants versus the MG1693 wild-type control ..................159

Figure A.3: Growth curve and viability curve of SK10475 ΔsecG leuU mutant versus the W3110 wild-type control ..................160

Figure A.4: Growth curve and viability curve of SK10473 ΔsecG leuU mutant versus the C600 wild-type control ..................161

Figure A.5: Comparative morphology of wild-type versus ΔsecG leuU deletion strain, as seen under bright field microscopy ..................162

Figure A.6: MG1693 ΔsecG leuU mutants with complementation in the absence of IPTG induction ..................163

Figure A.7: W3110 ΔsecG leuU mutant with complementation in the presence or absence of IPTG induction ..................164
CHAPTER 1
INTRODUCTION AND LITERATURE REVIEW

INTRODUCTION

Survival of a bacterial cell requires the adjustment of metabolic and physiological processes to changing environmental conditions. RNA turnover/decay is key to this rapid adaptation. RNA turnover is controlled by a combination of endonucleases and exonucleases, which leads to messenger RNA (mRNA) half-lives that vary anywhere from approximately 10 seconds to 20 minutes (1). The processing and breakdown of mRNAs directly affects steady-state transcript levels, leading to alterations in protein levels and the ability of the cell to rapidly adapt to changes in its growth environment.

Understanding the processes by which RNA processing, degradation, and polyadenylation occur in Escherichia coli is one key to the understanding the basics of cellular metabolism and regulation. Here we examine in detail the extensive work that has been done previously with regard to ribonucleases involved in the complex processes that control mRNA decay.

RIBONUCLEASES

The processing of RNA transcripts into mature forms along with the degradation of mRNAs, tRNAs, rRNAs, and sRNAs in E. coli is performed by a set of ribonucleases (RNases) consisting of endonucleases, which cleave at internal locations, and
exonucleases, which act in either a 5’→3’ or 3’→5’ manner to remove nucleotides from the extremities of RNA transcripts (Fig. 1.1, 1.2). [See (2) for an extensive comparison of the overlap in the functional activities of ribonucleases involved in mRNA decay and RNA processing.] To date, all the exonucleases in *E. coli* proceed in the 3’→5’ direction. It has been shown that both endo- and exonucleolytic action is required for the cell’s proper function and viability (3). These processing or degradative pathways are often complex, with multiple enzymes acting on a single substrate at varying rates and efficiencies causing a complex intertwined network of initial, secondary and tertiary cleavages.

**Endoribonucleases involved in *E. coli* RNA metabolism**

**A. RNase E**

The endonuclease, ribonuclease E (RNase E), encoded by the *rne* gene, prefers A/U rich single-stranded RNA regions as cleavage sites (4,5), especially those near stem-loops (6). RNase E is involved in many initial cleavage events and is regarded as the main endonuclease involved in RNA turnover (Fig 1.1). Tiling arrays have shown that more than half the RNA transcripts are targeted for RNase E cleavages (7). As such, RNase E has been shown experimentally to be involved in both rRNA and mRNA processing and the rapid degradation of mRNAs and sRNAs, including the processing of 5S rRNA precursors (8,9), 16S rRNA (10), tRNAs (11), and tmRNA (12).

Structurally, RNase E forms a 260kDa homotetramer organized as a dimer-of-dimers with each dimer resembling the blade and handles of an open pair of scissors (13-15). RNase E has a highly structured and essential catalytic region from amino acid
residues 1-529 (16), followed by a largely unstructured C-terminal scaffold domain from residues 530-1061 (17). Scaffold enzyme associations occur between residues 701-1061 (16-18) (Fig.1.3). The scaffold region provides binding sites that lead to the formation of a multimeric complex called the degradosome, which will be discussed in greater detail later in this chapter. The catalytic region is composed of five main domains, which include an S1 domain, a 5' sensor pocket, a non-catalytic RNase H domain, a DNase I domain, and a dimer-dimer interaction region (13-15).

RNase E is a 5'-end-dependent endonuclease that has a preference for 5' monophosphate substrates (19). This affinity is due to the physical constraints of the 5' sensing pocket of RNase E (13,14), whereby RNase E interacts with the 5' phosphate of the RNA transcript, causing the S1 domain to clamp down and properly orient the RNA for RNase E cleavage (Fig 1.1a).

The removal of the 5' triphosphate from primary RNA transcripts, which then allows RNase E to rapidly process or degrade the resulting 5' monophosphorylated RNA transcript (20), is performed by RppH, an RNA pyrophosphohydrolase. However, RNase E has also been shown to work on some substrates in a 5'-independent manner referred to as “internal entry” (Fig 1.1b), but which occurs at a slower rate than 5'-dependent processing by RNase E (21,22). Additionally, microarray analysis of rppH mutants suggests that the majority of transcripts processed by RNase E are likely processed in the slower 5'-independent manner (23).
B. RNase G

RNase G, encoded by the *rng* gene, is a parologue of the RNase E catalytic domain (10,24,25). It has no comparable C-terminal scaffold domain like RNase E and forms a homodimer rather than a dimer of dimers like RNase E (26). Additionally, unlike RNase E, RNase G is not essential for cell viability (10,25), and is found at lower levels in the cell (27). If overproduced at very high levels, RNase G has been shown to partially overcome the temperature sensitivity of *rne* mutants (27-29). Recently a single amino acid mutation in RNase G has been shown to allow RNase G to substitute for *rne* even at normal levels (30). In addition to substituting for or being a backup enzyme to RNase E, especially in 5S rRNA maturation, RNase G plays a role in its own right in the maturation of 16S rRNA (10,25), 6S RNA (31), some tRNAs (32,33), and mRNAs (28).

C. RNase III

After RNase E, RNase III is considered to be the second most important endonuclease involved in initial cleavage events. RNase III, encoded by the *rnc* gene, is an endonuclease with substrate specificity for double-stranded RNA, particularly double-stranded loops in intercistronic regions, which it cleaves at the base of the stem (34,35). Its phosphodiesterase cleavages generate a 2-nucleotide overhang on the 3’ end and a 5’-monophosphate and a 3’-hydroxyl terminus (36). It is the primary enzyme responsible for the separation of the pre-16S rRNA and pre-23S rRNA species from a 30S rRNA primary transcript (Fig. 1.2) (34,37). It also separates some mRNA transcripts between protein coding regions, such as *rpsO pnp* (35). Tiling arrays of *rnc* mutants have
indicated its role in processing or degradation of approximately 12% of all genes in the
*E. coli* genome (7).

**D. RNase P**

RNase P, encoded by the *rnpA* and *rnpB* genes, is a ribonucleoprotein, made up of a
protein subunit (M9) and a catalytic RNA subunit. It is an essential endonuclease and is
responsible for the 5′ maturation of tRNAs (Fig. 1.2). In addition, RNase P processes a
number of polycistronic tRNAs and/or mRNA transcripts (32,33,38).

**E. RNase Z/BN**

In some eukaryotes and prokaryotes, RNase Z is essential for the maturation of tRNA
precursors lacking encoded 3′ terminal CCA sequences. However, since all *E. coli*
tRNAs have chromosomally encoded CCA sequences (39), the function of RNase Z in
*E. coli* was initially unclear. RNase Z has since been shown to be an endonuclease that
plays a role in mRNA decay, including cleavage of the *rpsT* transcript at distinct sites
separate from its known RNase E cleavages (39). Other work has shown that RNase Z
cleaves endonucleolytically, producing a 3′ RNA byproduct that can be as small as 4
nucleotides (40).

RNase Z was initially discovered as a 3′→5′ exonuclease and was called RNase
BN (41). Its role as a potential exonuclease involved in RNA maturation will be
discussed in greater detail later in this chapter. The enzyme is most commonly referred
to RNase Z when acting as an endonuclease and RNase BN when acting as an
exonuclease.
F. Other Endoribonucleases

Ribonuclease H (RNase H) is an endonuclease found in both in prokaryotes and eukaryotes and cleaves RNA/DNA hybrids, including the RNA primers involved in DNA replication (42). RNase LS is an endonuclease that was thought to serve as an antagonist of infection by bacteriophage T4 (43), but has recently been suggested to encode the toxin of a novel toxin/antitoxin system (44). RNase I, and its altered form RNase M (45), is a nonspecific endonuclease that is found primarily in the periplasmic space and whose role in mRNA decay, if any, is still largely unknown.

Exoribonucleases involved in E. coli RNA metabolism

The major exoribonucleases in E. coli are discussed below. It is interesting to note that only PNPase, RNase II, RNase R and oligoribonuclease (Orn) are involved in mRNA decay (Fig. 1.1d). The other exonucleases are primarily involved in the 3’ maturation of precursor tRNAs (46).

A. PNPase

Polynucleotide phosphorylase (PNPase), encoded by the pnp gene, is a homotrimeric protein (47,48) that functions as a 3’→5’ exonuclease. PNPase has also been shown to be both a member of the degradosome complex (18,49) and to work biosynthetically as a poly(A) polymerase (50), both of which topics will be discussed in greater detail later in this chapter.

PNPase has been shown to be essential only if RNase II or RNase R are missing (3,51). PNPase is a phosphorylytic exonuclease that is inhibited by secondary structure,
stalling 6-8 nucleotides downstream of secondary structures. It is also ineffective on short oligonucleotides (52,53). PNPase’s enzymatic catalysis occurs in the central channel of the protein (54,55), into which the RNA molecules travel for cleavage. However, PNPase degrades secondary structures in conjunction with RhlB, an RNA helicase also found as part of the RNase-E based degradosome (56,57).

Recently PNPase has been shown to have a role in sRNA metabolism. PNPase can protect sRNAs from decay in exponentially growing cells (58), but in stationary phase PNPase actively degrades sRNAs, especially those not bound to Hfq (59).

B. RNase II

RNase II is a hydrolytic exonuclease that degrades single-stranded RNA, but is inhibited by secondary structures (53,60). However, RNase II degrades RNA transcripts that contain poly(A) tails (53,60). In fact, overproduction of RNase II can suppress the toxicity of overproduction of the poly(A) polymerase PAP I (60). RNase II also plays a minor role in 3’ maturation of tRNAs (61).

RNase II requires a single-stranded region of 10 nucleotides or more to bind effectively, as only single-stranded RNA can fit into the catalytic site due to steric hindrance (53,62). Interestingly, RNase II can also bind DNA but it does not degrade it; thus DNA can act as an inhibitor (63).

C. Other Exoribonucleases

Other exoribonucleases, including RNase R, oligoribonuclease, RNase T, RNase D, RNase PH, RNase BN, have been shown to play a role in RNA processing or decay.
Many of the functions of RNase T, RNase D, RNase PH, RNase II and RNase BN exonucleases are overlapping. A cell is viable with a wild-type copy of any one of these genes, even when the other four enzymes are missing (64), with RNase T, RNase PH, RNase D, RNase II, then RNase BN, respectively, being the most to least effective at supporting growth (64).

RNase R is a hydrolytic exonuclease with an ability to degrade structured single-stranded RNA, but requires single-stranded regions longer than 7 nucleotides to bind, leaving 2-5 nucleotide fragment byproducts (65). RNase R degrades defective tRNAs or rRNAs, mRNAs, especially those with poly(A) tails, and also in conjunction with YbeY to work on 70S rRNA degradation (66-71). Additionally, levels of RNase R have been shown to dramatically increase under cold shock conditions (72-75).

Oligoribonuclease (Orn) is an essential exonuclease (76) responsible for degradation into mononucleotides of the short 2-5 mer single-stranded RNA oligonucleotides that remain from the activity of other nucleases (77-79).

RNase T is a single-stranded specific hydrolytic exonuclease involved in the 3’ maturation of rRNAs and tRNAs (Fig. 1.2) (80,81). RNase T has also been shown to have DNA exonuclease activity (81).

RNase D is another hydrolytic exonuclease shown to act on denatured or damaged tRNAs, tRNA precursors, or other small structured RNAs (82-85).

RNase PH is a phosphorylytic exonuclease that works predominately on 3’ maturation of tRNA precursors (64,86-88) and is a paralogue of PNPase. It has also been shown to be involved in the modification of the 3’ end of sRNAs (89-91). It is
essential only in the absence of PNPase and RNase T (61,64,92), suggesting its function overlaps with the function of those two more prominent exonucleases.

RNase Z/BN is an endonuclease that has been shown to have in vitro exonuclease activity (40), and has also been suggested to function exonucleolytically in vivo as well (93). It has been hypothesized that the substrate range of the E. coli RNase Z/BN, as opposed to RNase Z family enzymes in other organisms, such as Bacillus subtilis, which have only endonucleolytic activity, is due to its narrower and more rigid channel downstream of its catalytic site (93). As an exonuclease, it is proposed to be active on both single-stranded and double-stranded RNA, but with double-stranded or duplexed RNA preferred (40). The exonuclease activity of RNase BN is inhibited by a run of C residues or the presence of either a phosphoryl group or CCA sequence on the 3' RNA terminus (40).

**THE RNASE E-BASED DEGRADOSOME**

The RNase E-based degradosome is a key complex involved in mRNA degradation. [See (94,95) for extensive reviews on the subject.] It has been shown to be involved in the majority of mRNA degradation in the cell (7), including the degradation of its own transcript (96-98). RNase E was first shown to be involved the degradation of small non-coding RNAs in 2003 by Massé et al. (99). Since then, it has been shown to be involved in the degradation of numerous sRNAs. [See (100) for a review of the past ten years of research on RNase E-dependent degradation of sRNAs.]
**Enzymes involved in the RNase E-based Degradosome**

The degradosome is composed of the RNase E endonuclease, along with PNPase, the RhlB RNA helicase, and enolase (49,56,101,102) which associate with the C-terminal scaffold domain of RNase E. Additionally, other proteins have been suggested to have the ability to interact under other physiological conditions, such as cold shock (103). It is also important to note that all the components of the degradosome are not present in equimolar amounts in the cell. Data has shown that PNPase and enolase are present in large excess relative to RNase E (5-10 fold during exponential growth), while RhlB is present in almost equimolar amounts (104).

**A. RNase E**

As discussed earlier in this chapter, RNase E is involved in rRNA and mRNA processing (4) and the rapid degradation of mRNAs and sRNAs. The protein has a highly structured catalytic region from amino acid residues 1-529, followed by a largely unstructured scaffold C-terminal domain from residues 530-1061 (17), with scaffold enzyme associations occurring between residues 701-1061 (16-18). It has been shown that mutants which contain a wild-type catalytic region of RNase E, even if missing the scaffold region, function normally with regards to mRNA decay and tRNA maturation (105).

Callaghan *et al.* (17) found that the scaffold region is highly unstructured under native conditions and was likely to remain unfolded even within the degradosome. The only exceptions to the unstructured nature of the scaffold are four 10-80 residue long regions of increased structural propensity (17) that correspond to the membrane binding
site of RNase E (106), the protein–RNA interaction site, and the interaction sites for enolase and PNPase (18), respectively. The arginine-rich RNA binding site (ARRBS) is located at amino acids 601-700 and is proposed to somehow enhance RNase E mRNA degradation (105, 107). Deletion experiments have demonstrated that association regions for RhlB RNA helicase, enolase and PNPase are located at amino acids 734-738, 739-845 and 844-1045 of the RNase E scaffold, respectively (18). Furthermore, it has been argued that Hfq, an abundant 11.2 kDa RNA-binding protein, can also associate with RNase E, competing with the RhlB RNA helicase for binding in the region of amino acids 734-738 (108).

There is strong evidence that the majority of RNase E binds to the inner cytoplasmic membrane (104, 106, 109, 110). This interaction occurs via amino acid residues 568-582. The phenylalanine at amino acid residue 575 in this α-helix region is key to this binding (106). RNase E proteins lacking this region no longer localize to the membrane, but are found in the cytoplasm (104, 106). This same region was previously identified as a short α-helix structured region and proposed to be involved in self-association (17).

Additionally, it has been suggested that RNase E needs to be localized to the inner membrane for normal levels of activity, since protein levels in mutants lacking residues 529-592 are increased compared with strains containing that region (106). Recent work has suggested that membrane binding of RNase E leads to secondary protein structural changes, which cause enzymatic activation, stabilization of folded RNase E, and an increased substrate binding affinity (111).
B. PNPase

PNPase is a 3’-to-5’ exonuclease and is a key component of the mRNA degradosome (3,49). PNPase associates with residues 844-1045 of the RNase E scaffold (18). More specifically, scaffold region residues 1021–1061 have been shown to bind PNPase by non-dissociating mass spectrometry, and pull-down assay analysis (17). Although RNase E is localized to the inner membrane, the majority of PNPase molecules have been found in the cytoplasm (104), a result that differs from other degradosome components. These data suggest that only a small fraction of cellular PNPase is bound to RNase E, an unsurprising conclusion due to the 5-10 fold excess of PNPase compared to RNase E in exponentially growing cells (104).

Additionally, PNPase binds directly to the RhlB helicase (112) in addition to its binding to the scaffold domain, suggesting that multiple associations occur to form or strengthen the degradosome complex. PNPase has also been shown to be a member of the polyadenylation complex through binding to both PAP I and Hfq (113), which will be discussed in greater detail later in this chapter.

C. RhlB

RhlB is a DEAD box ATP-dependent RNA helicase (56,114-116) that associates with the RNase E-based degradosome through an interaction with amino acid residues 734-738 of the C-terminal scaffold of RNase E (17,18,104). RhlB unwinds double-stranded RNA by employing ATP hydrolysis (114). It has been suggested that RhlB works in the degradosome to unwind RNA structures that inhibit PNPase exonuclease activity (56). Purified RhlB has been shown in vitro to have significantly increased ATPase activity
when complexed with RNase E, compared to when RhlB is on its own (17,18,115), suggesting that the interaction with RNase E stimulates RhlB activity. The majority of RhlB has also been shown to localize to the cellular membrane when the C-terminus of RNase E is present (104), confirming the interaction with the RNase E scaffold. RhlB has also been suggested to physically interact with PNPase in addition to the RNase E scaffold (117).

D. Enolase
Enolase is a glycolytic enzyme associated with the RNase E-based degradosome (56) through an interaction with the residues 739-845 of the RNase E scaffold (17,18). Its exact role in mRNA degradation is unknown. The majority of enolase has been found localized to the membrane when RNase E is membrane bound (104), confirming the physical interaction of RhlB with the RNase E scaffold.

E. Hfq
Hfq is a ubiquitous and abundant 11.2 kDa RNA-binding protein that exists \textit{in vivo} as a hexamer. It has been argued that Hfq can also associate with RNase E, competing with the RhlB RNA helicase for binding in the region of amino acids 734-738 (108). Hfq plays a major role in sRNA degradation by RNase E (99,118). Hfq has also been shown to be part of a polyadenylation complex with PNPase and PAP I (113), a role which will be discussed below.
POLYADENYLATION IN E. COLI

Polyadenylation at the 3’ terminus of a transcript is a key mechanism in mRNA metabolism in both Eukaryota and Bacteria, but it appears to function very differently in the two domains. In Eukaryota, the addition of a poly(A) tail provides stability to prevent 3’-to-5’ degradation and promotes more efficient translation (119-121). In contrast, bacterial polyadenylation has been shown to be involved in marking transcripts for degradation (122-126).

It has been hypothesized that since E. coli does not have a 5’→3’ exonuclease, nor can most of its exonucleases degrade through secondary structures, that the addition of poly(A) tails by PAP I or PNPase to the 3’ end of the mRNAs aids in the ability to degrade the secondary structure of mRNAs (53). In addition, it is unknown, but proposed that multiple rounds of polyadenylation and exonucleolytic degradation may be necessary before these inhibitory secondary structures are completely degraded (127).

Interestingly, although mRNA transcripts were originally believed to be the only substrates for polyadenylation, in E. coli rRNAs, tRNAs, sRNAs, and defective tRNAs can all be substrates for polyadenylation as well (69,126,128,129). It has been observed that the polyadenylation of tRNAs and sRNAs generates shorter tails than those found on mRNA transcripts.

**Enzymes involved in E. coli mRNA Polyadenylation**

Polyadenylation in E. coli is carried out primarily by poly(A) polymerase I (PAP I), but also involves the action of polynucleotide phosphorylase (PNPase) as the second
poly(A) polymerase. While there is much known about these two enzymes and their roles in polyadenylation, the exact mechanism by which PAP I and PNPase identify and select their RNA substrates for polyadenylation is still not completely understood.

A. PAP I

PAP I was first purified from *E. coli* in 1962 (130), but it was not until thirty years later that it was shown to be responsible for the addition of poly(A) tails onto the *lpp* transcript, the first direct evidence of PAP I function *in vivo* (131). At that time *pcnB*, the structural gene for PAP I, was also identified (132). Since then, PAP I has been shown to add poly(A) tails *in vivo* that exclusively contain adenosine residues. These PAP I-generated poly(A) tails are normally between 10 and 40 nucleotides in length (125). However, tail lengths are reduced by 3'→5' exonucleases such as RNase II and PNPase (60,125,133,134). *In vivo*, the homopolymeric tails are found on mRNA decay products or on full-length transcripts after Rho-independent transcription terminators (60,113,135,136), suggesting that Rho-independent transcription terminators, with their structure of a stable hairpin loop followed by a short U-rich single-stranded region (137), likely function as *in vivo* polyadenylation signals. In fact, over 80% of poly(A) tails are normally found after Rho-independent transcription terminators (113).

Deletion of the *pcnB* gene leads to a 90% decrease in intracellular poly(A) tail levels (136). Taken together, these data suggest that PAP I is responsible for the majority of all polyadenylation within the cell. Surprisingly, the PAP I overproduction was shown to be toxic (136). Recently, it has been demonstrated that the toxicity of
overproduction of PAP I is related to the polyadenylation of mature tRNAs and the concomitant inhibition of protein synthesis (129).

PAP I has also been shown to be associated with or localized to the inner membrane during exponential growth, but it appears to be released from the inner membrane as cells enter stationary phase (138,139). This release upon entry into stationary phase has been shown to be dependent on the adapter protein SprE (RssB) (139). Data showing that a PAP I leader-GFP fusion that contained the first 24 amino acid residues of PAP I localized to the inner membrane (138) suggests that the PAP I interaction with the inner membrane is dependent on some feature of the PAP I protein leader region. Additionally, evidence from immunoprecipitation of PAP I-GFP suggests that RNase E and the degradosome components are physically associated with PAP I, and that this interaction is SprE-dependent during stationary phase (140). However, to date, there has been no direct evidence of any interaction between the functions of PAP I or RNase E-based degradosome, other than that known association of polyadenylation aiding in mRNA degradation. In addition, it should be noted that PAP I is present in the cell at very low levels (30-50 molecules per cell) (113). An examination of a link between PAP I function and the presence of the RNase E-based degradosome will be discussed in Chapter 3.

**B. PNPase**

The evidence that a PAP I deletion mutant was not completely defective in polyadenylation led to the search for a second poly(A) polymerase (141). The f310 gene was initially identified as the structural gene for a second poly(A) polymerase (142), but
later work disproved this hypothesis (143). Subsequently, PNPase was identified as the second poly(A) polymerase (50). Although PNPase is a 3'→5' exonuclease and is a component of the mRNA degradosome (3,49), depending on the inorganic phosphate levels, PNPase can work either exonucleolytically or biosynthetically in vitro and in vivo (50,144). Subsequently, PNPase has also been identified in several other prokaryotes as the primary poly(A) polymerase in cases where PAP I is not present or the PAP I paralogue lacks poly(A) polymerase activity (145,146).

When PNPase functions as a poly(A) polymerase, it adds adenosine-rich heteropolymeric tails (50). These tails contain all four nucleotides, with compositions of more than 50% adenosine residues (50). While it was suggested that the presence of non-A nucleosides resulted from occasional PAP I misincorporation (147,148), heteropolymeric tails were eliminated in PNPase mutants of *E. coli*, suggesting that PNPase was actually responsible for these tails (50). In agreement with this data, strains with higher levels of PNPase have higher levels of heteropolymeric tails (60). It was also observed that PNPase, unlike PAP I, does not add tails after secondary structures such as Rho-independent transcription terminators but rather adds tails near the 5' termini of mRNA transcripts (60).

C. Hfq

Another protein involved in polyadenylation in *E. coli* is Hfq, an abundant 11.2 kDa RNA-binding protein. Hfq has been shown to be extensively involved in the function of sRNAs, stabilizing the sRNAs and aiding in their base pairing with target RNAs (149-152). However, it is still unclear exactly how Hfq brings a sRNA and its target RNA
together. Additionally, Hfq has been shown to have an affinity for single-stranded A/U rich RNA regions (153), similar to RNase E, suggesting that Hfq can compete with and block cleavages by RNase E (154).

Mutation of the hfq gene reduces the ability of PAP I to add poly(A) tails to transcripts after Rho-independent transcription terminators, even though intracellular PAP I protein levels remain unchanged (113). Total poly(A) levels are lower in Hfq mutants and the poly(A) tails observed are shorter, even when PNPase, RNase II or RNase E are also missing, demonstrating that the shortened tails are not the result nucleolytic degradation (113).

**The Polyadenylation Complex**

The existence of a polyadenylation complex containing PAP I, Hfq, and PNPase was proposed based on experimental evidence demonstrating physical interactions among PAP I, PNPase and Hfq (113). Co-purification of both PNPase and Hfq occurs under native conditions with histidine-tagged PAP I, as well as the co-immunoprecipitation of PAP I, PNPase, and Hfq under native conditions using PNPase or Hfq antibodies (113). This data suggests the existence of a multiprotein complex that is required for the polyadenylation of mRNAs with Rho-independent transcription terminators.

On a related note, it was observed that in a pnp mutant, PAP I adds poly(A) tails primarily to breakdown products rather than full-length transcripts with Rho-independent transcription terminators (113), suggesting a possible interaction between PNPase and PAP I function. Since the primary role of PNPase in *E. coli* is as a member of the
degradosome and as a 3′→5′ exonuclease (3,49), this potential PNPase/PAP I interaction could imply a link between PAP I mediated-polyadenylation and the functions of the degradosome. This question will be examined further in Chapter 3.

**Polyadenylation levels in *E. coli***

In 1992, Cao and Sarkar (131) found that 1.3% of the total RNA was polyadenylated in exponentially growing *E. coli* by measuring pulse-labeled RNA bound to oligo(dT)-cellulose. However, it was later shown that ~90% of the ORFs in exponentially growing cells were polyadenylated to some degree by PAP I employing genome-wide macroarray analysis of polyadenylated transcripts in wild-type and *pcnB* strains of *E. coli* (135). This result demonstrated a more global role for PAP I in the cell than previously thought. It is now hypothesized that only small percentage of total RNA transcripts are polyadenylated at any given time, but that the majority of transcripts are able to be substrates for polyadenylation.

Furthermore, there also seems to be a direct correlation between the nature of the transcription terminator and the enzyme responsible for the addition of poly(A) tails. It was observed from the macroarray analysis that the majority of mRNAs polyadenylated by PAP I were terminated in a Rho-independent manner (135). In contrast, the macroarray analysis suggested that transcripts lacking a Rho-independent transcription terminator were not substrates for PAP I (135).

Transcripts deemed primarily substrates of either PAP I or PNPase are found to have at least 60% homopolymeric tails or heteropolymeric tails, respectively (113,135). The composition varies among transcripts, but is reproducible for a specific transcript.
For example, *ompA* has been shown to have around 77% homopolymeric tails (135), while the *trxA* mRNA tails have been shown to be 100% heteropolymeric (113).

**Areas of polyadenylation needing further study**

Although much is known about polyadenylation and its role in mRNA degradation in *E. coli*, there are still many questions that need to be answered. How polyadenylation is signaled, what is the timing of polyadenylation in RNA turnover, what functional or physical interactions occur with the major processing endonuclease RNase E, and are the ORFs within a polycistronic operon polyadenylated in a similar manner are just a few of the questions unanswered with regards to polyadenylation in *E. coli*.

**A. Polyadenylation Signals**

It remains unclear precisely what signals PAP I or PNPase to add a tail to a particular substrate. The presence or absence of a Rho-independent transcription terminator could determine which polyadenylation enzyme is used (60, 113, 135, 136), but it is still unknown what the exact requirements, if any, are for that signal and how the signal is recognized. It is also still unclear if there is any physical interaction between PAP I and the secondary structure of the Rho-independent transcription terminator.

**B. Timing of Polyadenylation**

It also remains unclear at what point after transcription polyadenylation occurs. It was proposed that poly(A) tails mark the transcript for immediate degradation (124). However, this idea needs to be reexamined more fully in light of evidence that suggests
a larger role than previously thought for PAP I and polyadenylation in the cell. The possibility of polyadenylation occurring simultaneously with processing raises the question of how quickly a poly(A) tail signals that a particular transcript should be degraded. Is stability of polyadenylated transcripts consistent among the full-length transcripts, processed transcripts, and breakdown products? *In vivo* evidence suggests that PAP I-added tails are only found on mRNA decay products or after Rho-independent transcription terminators (60,113,136). These results could suggest that poly(A) tails have two slightly different functions and may be recognized by different enzymes at different stages of the decay pathway. The idea of two functions for PAP I could be supported by the evidence that in *pnp* mutants, PAP I adds poly(A) tails primarily to breakdown products rather than full-length mRNAs containing Rho-independent transcription terminators (113). However, it has yet to be determined precisely what percentage of the poly(A) tails added by PAP I are added to breakdown products versus full-length transcripts under wild-type conditions. Perhaps PAP I requires interaction with PNPase and the degradosome to function effectively as a poly(A) polymerase for full-length transcripts.

**C. Interaction with the RNase E-based degradosome**

It also remains to be determined what are the exact physical interactions of PAP I or the polyadenylation complex with RNase E or the components of the RNase E-based degradosome. The answer to this question could explain the preference of PAP I for polyadenylation after Rho-independent transcription terminators, since it has been shown that RNase E removes such substrates from a variety of primary tRNA
transcripts (33). The question of an PAP I interaction with the degradosome will be addressed in Chapter 3.

D. Polyadenylation of polycistronic transcripts

Experimental polyadenylation studies, until now, have only examined the polyadenylation of mono- or dicistronic transcripts, but have not looked at the polyadenylation of large polycistronic transcripts. It was observed from macroarray analysis that many ORFs within a polycistronic transcript showed polyadenylation patterns that were directly related to the transcription terminator of the operon (135), suggesting that the transcription terminator of a polycistronic transcript may be a polyadenylation signal for an entire operon. Thus, in a polycistronic transcript terminated in Rho-independent fashion, processed transcripts of the operon were predominantly substrates for PAP I, while processed units of a polycistronic transcript terminating in Rho-dependent manner were predominantly substrates for PNPase (135). However, this observation has yet to be examined in-depth to determine if the polyadenylation profiles of the processed ORF transcripts are actually similar to each other. This question will be addressed in Chapter 3 through the examination of the polyadenylation of the large polycistronic rpsJ operon.

THE RPSJ OPERON IN E. COLI

The E. coli ribosome is composed of the 5S, 16S, and 23S rRNAs together with more than 50 ribosomal proteins (r-proteins) (155). The r-proteins S1 to S21 assemble onto the 16S rRNA to makeup the 30S small ribosomal subunit, while the L1 through L36 r-
proteins assemble together with the 5S and 23S rRNAs to form the 50S large ribosomal subunit. Interestingly, 26 of the genes for the more than 50 r-proteins are transcribed by three transcriptional units clustered together—the \textit{rpsJ} (or \textit{S10}) operon, the \textit{rplN} (or \textit{spc}) operon, and the \textit{rpsM} (or \textit{a}) operon. Chapters 2 and 3 will include an in-depth examination of the processing and polyadenylation, respectively, of the \textit{rpsJ} operon.

**The \textit{rpsJ} operon**

The \textit{rpsJ} operon is a 5.2 kilobase polycistronic transcript made up of 11 genes (Fig. 1.4) (156,157). The genes and their encoded ribosomal proteins are as follows: \textit{rpsJ} (S10), \textit{rplC} (L3), \textit{rplD} (L4), \textit{rplW} (L23), \textit{rplB} (L2), \textit{rpsS} (S19), \textit{rplV} (L22), \textit{rpsC} (S3), \textit{rplP} (L16), \textit{rpmC} (L29), and \textit{rpsQ} (S17). All 11 genes encode 30S or 50S ribosomal subunit proteins. Additionally, the \textit{rpsJ} transcript has a highly structured 5' untranslated region (UTR) and a 3' Rho-independent transcription terminator.

**Cellular Control of Ribosomal Operons**

The cell uses a vast amount of energy in producing and maintaining ribosome content, so it has a vast network of regulatory mechanisms to tightly control ribosome production. The foundation of cellular control of rRNA and r-proteins is to maintain a balance of one-to-one ratio of rRNA to each r-protein to make functional ribosomes and also to not use excess cellular energy making more ribosomes than necessary. Overproduction of one of any components would be a great waste to the cell, while underproduction of any of these components would result in many incomplete and nonfunctional ribosomes. As such, the cell has been shown to tightly regulate the
transcription and translation of its r-protein operons. [See (158-160) for extensive reviews on the topic of ribosomal operons and their autogenous control.] In this dissertation, for the sake of brevity, only the autogenous control of the $rpsJ$ operon and its proteins will be discussed.

**Cellular Control of the $rpsJ$ Operon**

The $rpsJ$ operon has been well-studied over the past 30 years, and has been shown to have complex auto-regulation feedback loops for both of its transcription and translation through direct interactions of the L4 protein ($rplD$), the third gene in the $rpsJ$ operon, with the 5’ UTR of the polycistronic transcript (161-166). The 5’ UTR has been identified as containing two hairpin regions, HD and HE, which are responsible for transcriptional and translational control (161-164,166,167). The HD loop is necessary for L4/NusA activity and thus essential for transcriptional control, while the HE loop, which has homology to 23S rRNA, is where termination occurs and is essential to both transcriptional and translational control. Different nucleotides of HE hairpin are necessary for translational and transcriptional control (166), in that mutations of the region have been shown to abolish one without the loss of the other.

**A. Direct Control of $rpsJ$ transcript levels via the L4 protein**

One mechanism that helps control $rpsJ$ transcript levels is through L4 protein levels (168,169). In this mechanism, excess L4 protein, together with the NusA protein, will recognize and bind specific nucleotides in the hairpin loops of the 5’-UTR of $rpsJ$ causing transcriptional termination, which leads to decreased protein concentrations of
all the operon’s proteins in the cell (161-164,166). Additionally, L4 interaction with the hairpin is thought to sequester the Shine-Delgarno sequence and the ribosomal entry site, preventing translation (160,170).

B. Indirect Control of rpsJ transcript levels via the S10 protein

The second mechanism by which the rpsJ operon is controlled is through S10 protein concentrations. S10 (formerly known as NusE), is encoded by the rpsJ gene and is essential to the cellular viability. In this mechanism, which has yet to be fully elucidated, excess S10, together with NusB and other antitermination factors (171-173), promote antitermination of the seven rrn operons, allowing more rRNA to be transcribed. These rRNAs, together with the r-proteins, will then form mature ribosomes, directly leading to a decrease in excess r-proteins in the cell (172-174).

Processing of the rpsJ Operon

To date, processing of the large polycistronic rpsJ operon has not been examined. It is generally assumed that the rpsJ operon transcript it is processed into smaller, more manageable pieces, like other large polycistronic transcripts, but this hypothesis has never been directly tested. This hypothesis will be examined and discussed more fully in Chapter 2.

Polyadenylation of the rpsJ Operon

Macroarray data has suggested that at least some of the transcripts in the rpsJ operon are polyadenylated (135). However, this phenomenon has not been examined in detail,
nor has it been established if all ORFs within a polycistronic transcript are polyadenylated in the same manner. This question will be examined and discussed in Chapter 3 of this dissertation.
Figure 1.1. Model of mRNA decay pathways in *E. coli*. This figure shows the currently proposed pathways of mRNA decay and is modified from an extensive review on the topic (175).
Figure 1.2. Model of rRNA processing pathways in *E. coli*. This figure shows the currently proposed pathways of rRNA processing and is from an extensive review on the topic (175).
Figure 1.3. Schematic of wild-type RNase E protein. The areas labeled on the degradosome scaffold region indicate the associated binding site of the various proteins listed in the boxes. The S1 RNA binding domain and Arginine-rich RNA binding site (ARRBS) are also shown.

Figure 1.4. The rpsJ transcript. Two of the eleven encoded proteins are shown. Nucleotide lengths of the 5' UTR, each gene, intercistronic regions and the stem-loop terminator are shown above.
REFERENCES


not the C-terminal half, which is sufficient for degradosome assembly. *Proc. Natl. Acad. Sci. USA, 95*, 11637-11642.


175. Mackie, G.A. (2013) RNase E: at the interface of bacterial RNA processing and
CHAPTER 2

RNASE E IS THE MAJOR ENDONUCLEASE RESPONSIBLE FOR
PROCESSING THE POLYCISTRONIC RPSJ MESSENGER RNA TRANSCRIPT
IN ESCHERICHIA COLI\textsuperscript{1}

\textsuperscript{1} Kristen B. Mildenhall, Bijoy K. Mohanty, and Sidney R. Kushner. To be submitted to J. Bacteriol.
ABSTRACT

The processing and decay of RNA species in *Escherichia coli* involves the complex interaction of endonucleases and exonucleases. As such, most studies have only dealt with the analysis of either mono- or dicistronic mRNAs. In this study, the processing/decay of the large polycistronic *rpsJ* transcript (containing 11 ORFs) has been examined. Our data suggest that RNase E is the major enzyme involved in the initial processing of the *rpsJ* transcript and is responsible for the removal of the 3' Rho-independent transcription terminator. Our results also indicate that the 5' untranslated region of the *rpsJ* transcript remains intact in wild-type cells. Additionally, the processing pattern of the *rpsJ* transcript suggests a role for additional endonucleases in secondary cleavages. Interestingly, the cell appears to maintain fixed ratios of each ORF transcript within the polycistronic *rpsJ* operon, despite changes in the processing of the full-length transcript in RNase E mutants. These results indicate greater flexibility for the cell to control the ribosomal protein levels at the transcript level than previously observed.
INTRODUCTION

The processing or decay of mRNAs in *Escherichia coli* is performed by a combination of endonucleases and 3’→5’ exonucleases. The processing pathways are often a complex combination of primary and secondary endonucleolytic cleavages, along with exonucleolytic degradation of initial cleavage products. To date, other than a preliminary analysis of the *in vivo* turnover of the *trp* operon (1), no one has systematically examined the processing of a large polycistronic mRNA operon, despite the general assumption that these large transcripts are processed into smaller, more manageable species. Addressing the issue of processing of the large mRNA transcripts is especially pertinent for the ribosomal protein transcripts, as they are predominately found in large polycistronic operons. One such large polycistronic transcript is the *rpsJ* transcript (Fig. 2.1).

The *rpsJ* transcript is a 5.2 kb polycistronic transcript made up of 11 genes, all of which encode either 30S or 50S ribosomal subunit proteins (2,3). The *rpsJ* transcript has a highly structured 5’ untranslated region (UTR) and a 3’ Rho-independent transcription terminator (Fig. 2.2). The *rpsJ* operon has been well-studied and has been shown to have complex auto-regulation feedback of its transcription and translation through direct interaction of the L4 protein (*rplD*), the third gene in the *rpsJ* operon with the 5’ UTR of the polycistronic transcript (4-9).

Ribosomal proteins (*r*-proteins) are encoded in a one-to-one ratio to each other and to the 16S, 23S, and 5S rRNAs. The foundation of the cell’s control of its rRNA and *r*-proteins is to maintain one-to-one balance of rRNA to each *r*-protein such that only
mature ribosomes are formed. Therefore the cell tightly regulates the transcription and translation of many of its r-protein operons (10-12).

Preliminary polyadenylation data of the rpsJ transcript suggested that this large polycistronic transcript was highly processed. Since RNase E is the major enzyme involved in initiating RNA processing and decay in E. coli (13), we hypothesized that this enzyme was responsible for the initial processing of the rpsJ transcript. To test this hypothesis, the processing patterns of the rpsJ transcript in various endonuclease and exonuclease mutants were examined and compared with a wild-type control using northern blots. These data demonstrated that RNase E was the primary endonuclease responsible for rpsJ transcript processing and the removal of the 3’ Rho-independent transcription terminator. Furthermore, our results also indicate that other endonucleases are involved in a series of secondary cleavages.

**MATERIALS AND METHODS**

**Bacterial Strains and Growth**

All the strains and plasmids used in this study are listed in Table 2.1. They are all isogenic derivatives of MG1693. The rne-1 allele encodes a temperature-sensitive mutation that functionally inactivates RNase E at 44°C (14), while the rnpA49 allele encodes a temperature-sensitive mutation that functionally inactivates the protein subunit of RNase P at 44°C (15).

Bacterial strains were grown under one of three conditions: For experiments with temperature-sensitive mutations (rne-1 or rnpA49), all strains were grown under similar conditions at 30°C with shaking at 225 rpm in Luria broth supplemented with thymine
(50 µg/ml) and streptomycin (20 mg/mL), chloramphenicol (20 mg/mL) or kanamycin (25 mg/mL), where appropriate, until they reached a cell density of approximately $1 \times 10^8$/ml (50 Klett units above background, no. 42 green filter). Subsequently, the cultures were shifted to the non-permissive temperature of 44°C for 1 hour. Cultures were diluted with pre-warmed growth medium as needed to maintain them in exponential growth (Klett 50-90) and were subsequently harvested at approximately 60 Klett units above background.

For experiments not involving strains with temperature-sensitive mutations, strains were grown at 37°C with shaking at 225 rpm in Luria broth supplemented with thymine (50 µg/ml) and antibiotics, where appropriate, until they reached a cell density of approximately 60 Klett units above background.

For experiments employing minimal medium, MG1693 and SK3564 (rneΔ1018/rng-219) were grown at 37°C with shaking at 225 rpm in M9 medium supplemented with thymine (50 µg/ml) and additional supplements of 0.1% glucose for MG1693 and 0.75% glucose and 1% casamino acids for SK3564, until they reached a cell density of approximately 60 Klett units above background.

**Optimization of M9 Minimal Medium**

To obtain more comparable growth rates between MG1693 (wild-type) and SK3564 (rneΔ1018/rng-219), various concentrations of glucose and casmino acid supplements were tested as additions to the M9 medium. For these experiments, 5 mL aliquots of M9 medium supplemented with various concentrations of glucose were inoculated with 100uL of SK3564 and grown overnight at 37°C with shaking. Subsequently, readings
were taken to determine which concentration of glucose led to the highest cell density. Once the optimal glucose concentration had been determined, the experiment was repeated using the optimal concentration of glucose and varying concentration of casamino acids. Optimal concentrations for SK3564 were determined to be 0.75% glucose and 1% casamino acids (data not shown).

**Bacterial Growth Curves in Rich and Minimal Media**

To determine the differences in growth rates between MG1693 (wild-type) and SK3564 (rmeΔ1018/rng-219) in the various media, the strains were grown in Luria broth or M9 minimal medium as described above. Measurements under both growth conditions were taken every 30 minutes. Cells were diluted with pre-warmed growth medium as needed to maintain in exponential growth (Klett 50-90). Growth curves were replicated in triplicate and values averaged and then plotted.

**RNA Isolation and Northern Analysis**

Total RNA was isolated from various exponentially growing strains as described previously (16,17) under the growth conditions described above. The quality of RNA samples was checked by agarose gel electrophoresis and then quantified using a Nanodrop® ND-2000c spectrophotometer. Polyacrylamide northern analysis was carried out as described previously (18).

Glyoxyl agarose northern analysis was conducted using samples prepared with NorthernMax®-Gly sample loading dye (Ambion) and electrophoresis in a 1.5% agarose gel in BPTE buffer run at 80 volts for approximately 4.5 hours. Gel transfer and
subsequent steps were conducted as previously described (18). Formaldehyde agarose northern analysis was conducted as described previously (19). To analyze the processing pattern of the \textit{rpsJ} transcript, the northern blots were probed with $\gamma$-P$^{32}$-end-labeled oligonucleotides labeled using T4 PNK (NEB) that were specific for different regions of interest of the \textit{rpsJ} operon. The various oligonucleotide probes used are listed in Table 2.2. Quantification of band intensities was calculated using ImageQuant TL 7.0 software (GE Healthcare).

\textbf{\textit{rpsJ} mRNA Half-lives}

To determine the chemical half-lives of the \textit{rpsJ} transcript, rifampicin/naladixic acid was added to exponentially growing cells at the time of the temperature shift to 44°C as described previously (16,17). Time points were taken at 0, 1, 2, 4, 8, and 16 minutes. RNA was isolated and run on a 1.5% glyoxyl agarose northern as described above. Blots were probed and examined as described above.

\textbf{RNA Isolation for Tiling Arrays}

Strains were grown under conditions described above. Total RNA was isolated as described previously (16,17). The quality of RNA samples was checked by agarose gel electrophoresis and then quantified using a Nanodrop® ND-2000c spectrophotometer. The RNA was then labeled, fragmented, and hybridized to the tiling array, as described previously (13). These MG1655 based genome arrays were tiled using 60 nucleotide length probes containing 20 base pair overlaps at each end with both strands completely represented and manufactured by Agilent Technologies (Palo Alto, CA,
USA) as described previously (13). Results were viewed using the Integrated Genome Browser (20).

**RESULTS**

**Northern Analysis of rpsJ transcript processing**

To identify the full-length transcript of the *rpsJ* operon and examine its processing *in vivo*, polyacrylamide and agarose northern blots of RNA isolated from isogenic strains lacking various endonucleases were compared using a probe that hybridized to the 3' terminus of the *rpsJ* ORF, the first gene of the operon (Fig. 2.3). A transcript ~5.4 kb was present in every strain. Accumulation of the full-length transcript was seen in the RNase E mutant (Fig. 2.3, lane 4), with a 3.8-fold increase in full-length transcript compared to the wild-type control. In contrast, there was no increase in either the full-length transcript or decay intermediates in the RNase Z or RNase G deletion mutants (Fig. 2.3, lanes 2, 5). The wild-type strain appeared to have more processed fragments 3kb and smaller than the RNase E mutant (Fig. 2.3, lanes 1, 4).

This experiment was subsequently repeated using a RNase E deletion mutant (*rne* Δ1018/*rng*-219) and a RNase III deletion mutant (*rnc*-14). As shown in Fig. 2.4, there was less of the full-length *rpsJ* transcript in the RNase III mutant than the wild-type control. The increase of the full-length *rpsJ* transcript in the *rne* Δ1018/*rng*-219 mutant was comparable to what was seen with the *rne*-1 allele (Fig. 2.3, lane 4).
**Analysis of the rpsJ mRNA in various RNase E mutants.**

Since it was clear that RNase E played a major role in the processing or decay of the full-length rpsJ transcript (Figs. 2.3, 2.4), the fate of the transcript in strains carrying different rne alleles was analyzed. In this experiment, an RNase E truncation mutation lacking the degradosome scaffold region (rneΔ645), a newly isolated temperature-sensitive allele (rne-117), and a multiple mutant carrying rne-1 along with mutations in RNase II and polynucleotide phosphorylase were examined (Fig. 2.5). A probe that hybridized to the 5' UTR of the rpsJ transcript was utilized. The largest increase in the full-length transcript occurred in the rne-1 pnp-7 rnb-500 triple mutant (Fig. 2.15, lane 5).

Surprisingly, in all of the mutants, there were a series of decay intermediates that had retained the 5' UTR but had been shortened from the 3' terminus (Fig. 2.5). When the experiment was repeated with a probe for the 3' Rho-independent transcription terminator, increased levels of the full-length transcript were again observed in three of the four RNase E mutants. Unexpectedly, the amount of the full-length transcript present decreased in the rne-1 pnp-7 rnb-500 mutant (Fig. 2.6, lane 5). Some additional decay intermediates were observed, particularly in the rneΔ1018/rng-219 strain (Fig. 2.6). In addition, there was a very large increase in a ~200bp fragment that represented the Rho-independent transcription terminator (Fig. 2.6).

Interestingly, only the rne-1 pnp-7 rnb-500 strain showed stabilization of the full-length transcript compared to wild-type when the 5' UTR region was present (Fig. 2.5, lane 5), whereas transcripts containing the 3' stem-loop did not show any stabilization compared to wild-type for the rne-1 pnp-7 rnb-500 (Fig. 2.6, lane 5). The rneΔ645 scaffold mutant strain did not have any significant effect on full-length transcript
accumulation but did appear to have a difference on the fragment pattern of fragments less than 3kb (Figs. 2.5, 2.6, lane 2).

**Analysis of rpsJ transcript fragments using additional hybridization probes**

In order to determine if there was a distinct cleavage pattern of the *rpsJ* transcript, we initially examined polyacrylamide and agarose northern blots in various RNase E mutant strains using probes for five genes within the operon (Figs. 2.7 and 2.8). Subsequently, agarose northern analysis of wild-type versus the *rne-1* strain was also performed with probes for the remaining six genes of the operon (Fig. 2.9).

As expected, every probe showed an accumulation of the full-length *rpsJ* operon transcript in the *rne-1* and *rneΔ1018/mg-219* mutants (Table 2.3), in agreement with the results seen for the terminal regions (Figs. 2.5, 2.6). It is also important to note that the stabilization of the full-length transcript seen in the *rne-1 pnp-7 rnb-500* strain with the 5’ UTR probe was not seen for any of the other regions probed, including the *rpsJ* probe just downstream of the 5’ UTR region (Fig. 2.7A).

The RNase E mutant strains showed a build-up of various fragments for each region probed, ranging in size from ~5kb to ~500bp (Figs. 2.7, 2.9). Each region probed had a different and complex fragment pattern. In contrast, the only region affected in the polyacrylamide northern (which separates only ~700 bp transcripts or less), was for the *rpsQ* probed regions (Fig. 2.8D), which showed build-up of smaller size fragments, similar to what was seem for the 3’ transcription terminator probe.
**Effect of PAP I on *rpsJ* transcript processing**

To determine if polyadenylation had any effect on the processing and/or decay of the full-length *rpsJ* operon transcript or its 3’ terminus, the transcript profiles in strains carrying various combinations of RNase E and PAP I mutations were compared using a probe specific for the Rho-independent transcription terminator. The data obtained (Fig. 2.10) demonstrated that PAP I did not have a significant effect on the stability of the full-length transcript, as none of the double mutant strains defective in RNase E and PAP I showed any dramatic difference from the single RNase E mutants. There was some stabilization of smaller fragments in the *rne-1 ∆pcnB* strain compared to the *rne-1* single mutant (Fig. 2.10). The deletion of *pcnB* in the *rne-1 pnp-7 rnb-500* strain appeared to significantly stabilize a number of smaller fragments present (Fig 2.10, lane 8).

**Developing growth conditions to obtain comparable generation times for MG1693 and SK3564**

It is well established that the ribosome content of cells is directly related to growth rate (21). Therefore, to eliminate any growth dependent effects on ribosome concentrations, and thus r-protein concentrations, we attempted to obtain comparable doubling times between the MG1693 (wild-type control) and SK3564 (*rneΔ1018/mg-219*) mutant strain employing M9 minimal medium with glucose as the carbon source. It was determined through OD$_{600}$ measurements of a series of overnight cultures that M9 medium supplemented with 0.75% glucose and 1% casamino acids was optimal for SK3564 growth (data not shown). The wild-type MG1693 was grown at a lower concentration of the same carbon source (0.1% glucose) without any casamino acids to obtain a
comparable growth rate. When grown in M9 minimal medium supplemented with thymine, glucose, and casamino acids, MG1693 had a doubling time of 70-75 minutes, while SK3564 had a doubling time of 100-115 minutes. The doubling times of the strains grown in Luria broth supplemented with thymine were approximately 30 minutes for MG1693 versus 98-102 minutes for SK3564.

**Tiling Array Analysis for rne-1**

To examine the effects of RNase E on the *rpsJ* operon transcript in its entirety, a tiling array was performed for the *rne-1* strain versus the MG1693 wild-type control. In contrast to northern analysis, the array did not show any significant difference in the transcript levels of the *rpsJ* operon under the different conditions (Fig. 2.11A). This is in agreement, however, with the *rne∆1018/rng-219* arrays previously published (13), which also demonstrated no steady-state RNA level changes occurred for the *rpsJ* operon. The *rne∆1018/rng-219* strain was grown at 37°C and has been proposed to be a good substitute for the *rne-1* allele (22). Our results showed that the profile of the *rne-1* versus *rne∆1018/rng-219* array did not differ dramatically, which confirmed the idea that the *rne∆1018/rng-219* strain may be a good alternative to the temperature-sensitive *rne-1* allele.

Interestingly, however, one obvious difference between the *rne-1* and the *rne∆1018/rng-219* strains related to decreased levels of the 5’ and 3’ regions of 16S rRNA in all seven *rrn* operons (Fig. 2.11B). This result was in contrast to what was observed in the *rnc-14* and *rne∆1018/rng-219* arrays previously published (Fig.
However, the array pattern at the 3’ end of the 23S rRNA was similar in all the strains tested (Fig. 2.11B).

Additionally, as opposed to the rpsJ results, the rplN and rpsM ribosomal protein operons that are directly downstream of the rpsJ operon did show an increase in steady-state levels in the me-1 and meΔ1018/rng-219 arrays (Fig. 2.11A).

**Tiling Array Analysis for meΔ1018/rng-219 in Minimal Medium**

To eliminate any growth dependent effects from the tiling arrays on the ribosomal concentrations and r-protein concentrations, the meΔ1018/rng-219 tiling array was repeated using RNA isolated from the strain grown in minimal medium as described in the previous section. We were particularly interested in carrying out this experiment because both our me-1 array results and the previously published meΔ1018/rng-219 array results from our laboratory (13) did not show any significant difference in rpsJ transcript levels in MG1693 versus SK3564 when the strains were grown in Luria broth. These result were in striking contrast to the northern analysis data presented in Figs. 2.3-2.7, 2.9.

However, even when MG1693 and SK3564 were grown in minimal medium with comparable generation times there was not any significant difference in the rpsJ polycistronic transcript levels between the two strains (Fig. 2.12). We did observe increased levels of the downstream rplN and rpsM ribosomal protein operons (Fig. 2.12).
Northern Blot analysis of MG1693 AND SK3564 grown in minimal medium

Northern blot analysis was performed for the wild-type and \textit{rne}△1018/rng-219 strains grown in minimal medium to determine if accumulation of the full-length transcript was still seen in the RNase E mutant when the strains were grown with comparable generation times. Agarose northern blots of wild-type and \textit{rne}△1018/rng-219 strains were compared using probes for 5' UTR and the \textit{rpsS} gene (Fig. 2.13) and various intercistronic regions of the \textit{rpsJ} operon transcript (Fig. 2.14). Accumulation of the full-length transcript was still observed in the \textit{rne}△1018/rng-219 mutant (Fig. 2.13), with a 2.1-fold and 3.5-fold increase in full-length transcript compared to wild-type when probed with the \textit{rpsS} and 5' UTR probes, respectively. Similar results were obtained with the intercistronic probes (Fig. 2.14).

Chemical half-lives of the 5' \textit{rpsJ} UTR

To determine the stability of the \textit{rpsJ} full-length transcript in the \textit{rne}-1 mutant compared to the wild-type control, the chemical half-life of the full-length transcript was examined. Blots were probed for the 5' UTR of the \textit{rpsJ} transcript (Fig. 2.15). The half-life of the full-length \textit{rpsJ} transcript in SK5665 (\textit{rne}-1) was 3.86 ± 0.89 minutes, and 2.2 ± 0.56 minutes for MG1693 (wild-type). Surprisingly, full-length \textit{rpsJ} transcript levels increased up until 2 minutes for the wild-type control and 4 minutes for \textit{rne}-1 strain prior to decreasing (Fig. 2.15).

To determine if the increase in the amount of the \textit{rpsJ} transcript was some type of experimental artifact, we used previously isolated RNA from an \textit{rne}-1 strain that had been used for other half-life experiments that had not shown this behavior. In this
experiment an agarose northern blot and was probed for regions of the *rpsJ* and *rplN* operons and the monocistronic *adhE* transcript (Fig. 2.16). As we had observed in Fig. 2.15, the amount of the full-length *rpsJ* transcript increased until ~ 4 minutes before decreasing. This increase for several minutes before decreasing was also seen for the large polycistronic *rplN* polycistronic transcript (Fig. 2.16) and *rpsM* polycistronic transcript (data not shown), whereas the monocistronic *adhE* transcript did not display this effect (Fig. 2.16).

**DISCUSSION**

**RNase E processes the full-length *rpsJ* transcript**

The results presented here demonstrate that the *rpsJ* operon is a large polycistronic operon that is processed into smaller discrete fragments (Figs. 2.3-2.10, 2.13, 2.14). Accumulation of full-length and processed products in various RNase E mutants suggests that RNase E is the primary endonuclease responsible for this processing. Even when growth rates were adjusted to be comparable using minimal medium, inactivation of RNase E still was observed to stabilize the full-length transcript (Fig. 2.12).

Additionally, the analysis of the processing pattern of the RNase E scaffold mutant (*rne*Δ645) indicated a potential role for the scaffold domain, mainly in processing medium-sized fragments (less than 2kb). This result also suggested that only the catalytic domain of RNase E is critical for the initial cleavages or removal of the Rho-independent transcription terminator from the *rpsJ* transcript (Figs. 2.5-2.7, lane 2).
The lack of significant build-up of full-length or intermediates in RNase Z and RNase G mutants (Fig. 2.3) or RNase III mutants (Fig. 2.4) eliminated these endonucleases as primary processing enzymes.

**RNase E is likely responsible for the removal of the 3’ Rho-independent transcription terminator**

The 5’ UTR region was present on many processing fragments in the wild-type background (Fig. 2.5), while the 3’ stem-loop was predominately found on the full-length transcript (Fig. 2.6). This observation strongly suggested that the removal of the 3’ stem-loop as an early step in normal rpsJ processing. Additionally, in the *rne-1* mutant the presence of the 3’ stem-loop on fragments of varying sizes (Fig. 2.6 lane 4), with a strong accumulation of the ~200 bp fragment, suggested that RNase E was the primary enzyme responsible for stem-loop removal and its subsequent degradation, and that perhaps stem-loop removal is an early or rate-limiting step. The fact that the *rne-1 pnp-7 rnb-500* strain does not seem to have any change in the accumulation of the ~200 bp fragment compared to the *rne-1* single mutant, suggested that these two exonucleases do not play a major role in stem-loop degradation.

On the other hand, the 5’ UTR pattern in the *rne-1* strain did not seem to indicate any significant role of RNase E in removal of the 5’ UTR (Fig. 2.5). Failure to remove the 5’ UTR may be important because the 5’ UTR exerts both transcriptional and translational control on the entire operon (4,23).
Loss of RNase E stabilizes the full-length rpsJ transcript

Analysis of the chemical half-life of full-length rpsJ transcript (Fig. 2.15) suggested that not only does RNase E process the rpsJ full-length transcript, but that the lack of RNase E also stabilized the full-length rpsJ transcript. The observation that the abundance of full-length transcript continued to increase for 2-4 minutes after the equilibration of the rifampicin/naladixic acid seen for rpsJ operon transcripts (Figs. 2.15A, 2.16A) is likely due to the completion of transcription of the large polycistronic transcript that took place between the addition and equilibration of the rifampicin/naladixic acid. This is normally not seen in shorter transcripts, such as adhE (Fig. 2.16C), but this increase was also observed when the same northern blot was striped and reprobed for the rplN operon transcript (Fig. 2.16B) and the rpsM operon transcript (data not shown), the other large polycistronic ribosomal protein operons in the cell. It could also be a phenomenon that is specifically related to ribosomal protein transcripts.

Individual transcript ratios for the rpsJ operon transcript are maintained despite changes in its processing

The tiling array analysis of the rme-1 mutant and the rmeΔ1018/rng-219 grown in rich or minimal medium conditions both showed that the rpsJ operon steady-state levels were unchanged, even when conditions were adjusted to allow for comparable growth rates (Figs. 2.11, 2.12). At first glance, this result conflicts with the northern blot analysis that demonstrates a strong stabilization of the full-length transcript under the same conditions (Figs. 2.7 lane 3, 2.13). However, since the RNA is fragmented during the array procedure, the array inherently lacks the ability to detect differences if the cellular
concentration of each mRNA fragment is maintained; in other words, the tiling arrays cannot detect processing without breakdown or decay of a fragment more than 40 nucleotides in length.

Therefore, our result strongly suggested that the cell maintains the individual transcript ratios of each gene in the polycistronic *rpsJ* operon, despite changes in the processing of the full-length transcript in RNase E mutants. This result was supported by the lack of the strong banding pattern around 3kb in the *rne* mutants that is observed in the wild-type strain (Figs. 2.4 lane 2, 2.6 lane 3, 2.10AB). Our hypothesis is in agreement with the known tight cellular control of the r-protein operons and indicates greater flexibility and adaptation of the cell in controlling the ribosomal protein levels at the transcript level than previously observed. However, the maintenance of steady-state levels was not observed with any of the arrays for the *rplN* and *rpsM* operons that are directly downstream of the *rpsJ* operon (Figs. 2.11A, 2.12), a result that is still unclear.

**The *rpsJ* operon transcript is a substrate for other endonucleolytic cleavages when RNase E is missing**

The lack of complete build-up of full-length product or disappearance of intermediates in the RNase E mutants (Figs. 2.3-2.10, 2.13, 2.14) strongly suggested secondary cleavages by other endonucleases. The processing fragment pattern is very complex even when probed for other gene regions or intercistronic regions of the polycistronic transcript (Figs. 2.7, 2.10, 2.14), suggesting several other secondary players.

The processing pattern in the *rne-1 pnp-7 rnb-500* mutant showed a slight build-up of processed fragments (Figs. 2.5-2.7, lane 5), but the pattern was largely
unchanged from the *rne-1* single mutant (Table 2.3). These results indicated that these two major exonucleases only have a minor degradative role in the overall processing of the *rpsJ* operon transcript.

The overall lack of change in the full-length transcript profile of double mutant strains defective in *rne-1 ΔpcnB* strain compared to the *rne-1* single mutant suggested that polyadenylation is not playing a significant role in the absence of RNase E. This is not surprising, as RNase E is the major endonuclease which acts on polyadenylated substrates. The stabilization of smaller fragments in the *rne-1 ΔpcnB* strain and the *rne-1 pnp-7 rnb-500 ΔpcnB* compared to the *rne-1* single mutant (Fig. 2.10), suggested that polyadenylation may still be playing a role in the stability of the smaller fragments even when RNase E is missing.
Figure 2.1. The *rpsJ* transcript. Two of the eleven encoded proteins are shown.

Nucleotide lengths of the 5' UTR, each gene, intercistronic regions and the stem-loop terminator are shown above.
A. 5' Untranslated Region (UTR)

B. 3' Rho-independent transcription terminator

Figure 2.2. The 5' UTR and 3' Rho-independent transcription terminator of the *rpsJ* operon. The 5' UTR and 3' Rho-independent transcription terminator sequences were folded using the m-fold program (24).
Table 2.1. List of bacterial strains and plasmids used in this study. RNase = ribonuclease, PAP I = poly(A)polymerase I, PNPase = polynucleotide phosphorylase.

<table>
<thead>
<tr>
<th>Strains</th>
<th>Genotype</th>
<th>Endonuclease(s) missing</th>
<th>Exonuclease(s) or other missing</th>
<th>Reference Or Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>MG1693</td>
<td>thyA715 rph-1</td>
<td>none</td>
<td>none</td>
<td>E. coli Genetic Stock Center</td>
</tr>
<tr>
<td>SK2538</td>
<td>rng::cat thyA715 rph-1</td>
<td>RNase G</td>
<td>none</td>
<td>(25)</td>
</tr>
<tr>
<td>SK2685</td>
<td>rneΔ1018::bla srID300::Tn10 recA56 thyA715 rph-1 pDHK6 (rneΔ645 Sp/Sm&lt;sup&gt;R&lt;/sup&gt;)</td>
<td>RNase E scaffold domain</td>
<td>none</td>
<td>(26)</td>
</tr>
<tr>
<td>SK3543</td>
<td>rneΔ1018::bla srID300::Tn10 recA56 thyA715 rph-1 pDHK28 (mga219 low copy Km&lt;sup&gt;R&lt;/sup&gt;)</td>
<td>RNase E (rne-219 present in its place)</td>
<td>none</td>
<td>(22)</td>
</tr>
<tr>
<td>SK3564</td>
<td>rneΔ1018::bla ΔpcnB::apr srID300::Tn10 recA56 thyA715 rph-1 pDHK30 (mga219 single copy Sp/Sm&lt;sup&gt;R&lt;/sup&gt;)</td>
<td>RNase E (rne-219 present in its place)</td>
<td>None</td>
<td>Other: PAP I</td>
</tr>
<tr>
<td>SK3642</td>
<td>rneΔ1018::bla ΔpcnB::apr srID300::Tn10 recA56 thyA715 rph-1 pMOK13 (rne-1 low copy Cm&lt;sup&gt;R&lt;/sup&gt;)</td>
<td>RNase E (ts)</td>
<td>Other: PAP I</td>
<td>DuBose and Kushner, unpublished results.</td>
</tr>
<tr>
<td>SK3645</td>
<td>rneΔ1018::bla srID300::Tn10 recA56 thyA715 rph-1 pDHK30 (mga219 single copy Sp/Sm&lt;sup&gt;R&lt;/sup&gt;)</td>
<td>RNase E (rne-219 present in its place)</td>
<td>None</td>
<td>DuBose and Kushner, unpublished results.</td>
</tr>
<tr>
<td>SK3750</td>
<td>rneΔ1018::bla srID300::Tn10 recA56 thyA715 rph-1 pUGK49 (rne-117 low copy Km&lt;sup&gt;R&lt;/sup&gt;)</td>
<td>RNase E (ts)</td>
<td>none</td>
<td>Reyes-Darius, Perwez, and Kushner, unpublished results.</td>
</tr>
<tr>
<td>SK4161</td>
<td>rne-1 Δrnz::apr ΔmazEFG::kan thyA715 rph-1</td>
<td>RNase E (ts) RNase Z MazEFG (toxin-antitoxin system)</td>
<td>none</td>
<td>Richter-Roche and Kushner, unpublished results.</td>
</tr>
<tr>
<td>SK4455</td>
<td>rnc-14::ΔTn10 thyA715 rph-1</td>
<td>RNase III</td>
<td>none</td>
<td>(13)</td>
</tr>
<tr>
<td>SK4484</td>
<td>rne-1 rng::cat rnlA2::kan rnzΔ::apr rpsD296::Tn10 rnpA49 thyA715 rph-1</td>
<td>RNase E (ts) RNase G RNase Z RNase LS RNase P (ts)</td>
<td>none</td>
<td>Maples and Kushner, unpublished results.</td>
</tr>
<tr>
<td>Name</td>
<td>Sequence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>----------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rpsJ UTR</td>
<td>5’-CTCCTCAGACCCATTACGATTG-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3’ rpsJ</td>
<td>5’-AACCCAGGCTGATCTGCACG-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5’ rplC</td>
<td>5’-GTTACTGGGATAGAAACGCCGTC-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rplC-rplD intercistronic</td>
<td>5’-TCCATTGCTATCTCCTTACG-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5’ rplD</td>
<td>5’-AATCACGACCGAAGGTAGTTTCG-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5’ rplW</td>
<td>5’-AGAAACGTGCAGGTGCACGCA-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rplW-rplB intercistronic</td>
<td>5’-CATTGTATTACTCCTCACG-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5’ rplB</td>
<td>5’-AATCACGACCGAAGGTAGTTTCG-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rplB-rpsS intercistronic</td>
<td>5’-CATGGCTTATCCTCTAAATT-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3’ rpsS</td>
<td>5’-CTTCGCTTATCCTCTTACG-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rpsS-rplV intercistronic</td>
<td>5’-CATCTCTCCCTCTACCTTAAAATT-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5’ rplV</td>
<td>5’-ACCTCTGACGAGAAAGACGACG-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rplV-rpsC intercistronic</td>
<td>5’-ATTGCTAGTCTCCAGAGTCT-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5’ rpsC</td>
<td>5’-AATCGAGTTTACCTGAGTGAC-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rpsC-rplP intercistronic</td>
<td>5’-ACATCAGCGACGCTCCTTAT-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3’ rplP</td>
<td>5’-CATCAACGAGTTAATTGAGAGAG-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5’ rmcC</td>
<td>5’-ACTTGCAGCCTGATCTGACG-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5’ rpsQ</td>
<td>5’-ACCGAGACCCTGACAGTGAC-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rpsQ stem-loop</td>
<td>5’-CCGTTTATTGCGATGACGAG-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5’ rplN</td>
<td>5’-GTTCAATAGTGCTTTTCTTTG-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adhE</td>
<td>5’-CTTCACCGATCAATCTTG-3’</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 2.3. Agarose northern blot of endonuclease mutants probed with 3′ rpsJ probe. Cells were initially grown at 30°C prior to shifting to 44°C for 1hr. RNA was run on a 1.5% glyoxyl agarose gel. Strains were as follows: MG1693 (wild-type), SK2538 (∆rng), SK4484 (rne-1 ∆rng mzΔ500 mPA49 miaA2), SK5665 (rne-1), and SK9795 (rnzΔ500) in Lanes 1-5, respectively. Quantification of the full-length transcript was calculated with wild-type set at 1.0 and is shown in each lane. The sizes of the various transcripts based on a sizing ladder (RiboRuler™, Fermentas) are shown to the left of the blot.
**Figure 2.4. RNase III is not involved in rpsJ processing**

Cells were grown at 37°C. RNA was run on a 1.2% formaldehyde agarose gel. Strains were as follows: MG1693 (wild-type), SK3564 (rneΔ1018/rng-219), and SK4455 (mc-14) in Lanes 1-3, respectively.
Figure 2.5. Agarose and polyacrylamide northern blots of various RNase E mutants probed with 5’ rpsJ UTR probe. Cells were initially grown at 30°C prior to shifting to 44°C for 1hr. RNA was run on a 1.5% glyoxyl agarose gel (top) or a 6% urea polyacrylamide gel (bottom). Strains were as follows: MG1693 (wild-type), SK2685 (rneΔ645), SK3750 (rne-117), SK5665 (rne-1), and SK5704 (rne-1 pnp-7 rnb-500) in Lanes 1-5, respectively. Quantification of the full-length transcript was calculated with wild-type set at 1.0 and is shown in each lane. Predicted structures for the various rpsJ transcripts are shown to the right of the blot. The sizes of the various transcripts based on a sizing ladder (RiboRuler™, Fermentas) are shown to the left of the blot.
Figure 2.6. Agarose and polyacrylamide northern blots of various RNase E mutants probed with the 3’ Rho-independent transcription terminator probe. Cells were initially grown at 30°C prior to shifting to 44°C for 1 hr. RNA was run on a 1.5% glyoxyl agarose gel (top) or a 6% urea polyacrylamide gel (bottom). Strains were as follows: MG1693 (wild-type), SK2685 (rneΔ645), SK3564 (rneΔ1018/rng-219), SK5665 (rne-1), and SK5704 (rne-1 pnp-7 mb-500) in Lanes 1-5, respectively. Quantifications of the full-length transcript and the ~200 bp fragment were calculated with wild-type set at 1.0 and are shown in each lane. Predicted structures for the various rpsJ transcripts are shown to the right of the figure. The sizes of the various transcripts based on a sizing ladder (RiboRuler™, Fermentas) are shown to the left of the blot.
Figure 2.7. Agarose northern blots of endonuclease or endo/exonuclease mutants probed for \textit{rpsJ}, \textit{rpsS}, \textit{rpsC}, \textit{rplP}, and \textit{rpsQ}. Cells were initially grown at 30°C prior to shifting to 44°C for 1hr. RNA was run on a 1.5% glyoxyl agarose gel. Strains were as follows: MG1693 (wild-type), SK2685 (\textit{rne\Delta645}), SK3564 (\textit{rne\Delta1018/rng-219}), SK5665 (\textit{rne-1}), and SK5704 (\textit{rne-1 pnp-7 mnb-500}) in Lanes 1-5, respectively.
Table 2.3. Quantification of full-length transcript levels for agarose northern blots in various RNase E mutants. Quantification of the full-length transcript of blots shown in Figs. 2.5, 2.6 and 2.7 with wild-type set at 1.0.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Wild-type</th>
<th>rne△645</th>
<th>rne△1018/ rng-219</th>
<th>rne-1</th>
<th>rne-1 pnp-7 rnb-500</th>
</tr>
</thead>
<tbody>
<tr>
<td>5' UTR</td>
<td>1.0</td>
<td>1.2</td>
<td>ND</td>
<td>2.0</td>
<td>3.6</td>
</tr>
<tr>
<td>rpsJ</td>
<td>1.0</td>
<td>1.0</td>
<td>2.8</td>
<td>2.2</td>
<td>0.6</td>
</tr>
<tr>
<td>rpsS</td>
<td>1.0</td>
<td>0.9</td>
<td>3.3</td>
<td>2.8</td>
<td>0.6</td>
</tr>
<tr>
<td>rpsC</td>
<td>1.0</td>
<td>0.7</td>
<td>2.9</td>
<td>2.3</td>
<td>0.5</td>
</tr>
<tr>
<td>rplP</td>
<td>1.0</td>
<td>1.4</td>
<td>4.0</td>
<td>2.8</td>
<td>0.8</td>
</tr>
<tr>
<td>rpsQ</td>
<td>1.0</td>
<td>1.2</td>
<td>3.4</td>
<td>2.4</td>
<td>0.7</td>
</tr>
<tr>
<td>stem-loop</td>
<td>1.0</td>
<td>1.2</td>
<td>3.2</td>
<td>3.0</td>
<td>0.6</td>
</tr>
</tbody>
</table>
Figure 2.8. Polyacrylamide northern blots of various RNase E mutants probed for *rpsJ*, *rpsS*, *rplP*, and *rpsQ*. Cells were initially grown at 30°C prior to shifting to 44°C for 1hr. RNA was run on a 6% urea polyacrylamide gel. Strains were as follows: MG1693 (wild-type), SK2685 (rneΔ645), SK3564 (rneΔ1018/rng-219), SK5665 (rne-1), and SK5704 (rne-1 pnp-7 mb-500) in Lanes 1-5, respectively.
Figure 2.9. Agarose northern blots of wild-type and rne-1 strains probed for rplC, rplD, rplW, rplB, rplV, and rpmC. Cells were initially grown at 30°C prior to shifting to 44°C for 1hr. RNA was run on a 1.5% glyoxyl agarose gel. Strains were MG1693 (wild-type) and SK5665 (rne-1) in Lanes 1 and 2, respectively.
Figure 2.10. Agarose northern blot of various RNase E mutants in the presence or absence of poly(A) polymerase I (PAP I). Cells were initially grown at 30°C with shift to 44°C for 1 hr. RNA was run on a 1.5% glyoxyl agarose gel and the northern blot was probed with the oligonucleotide used in Fig. 2.6. Strains were as follows: MG1693 (wild-type), SK3543 (rneΔ1018/rng-219), SK3645 (rneΔ1018/rng-219 ΔpcnB), SK9704 (rne+), SK9937(rne-1), SK3642 (rne-1 ΔpcnB), SK5704 (rne-1 pnp-7 rnb-500), SK8901 (rne-1 pnp-7 rnb-500 ΔpcnB) in Lanes 1-8, respectively. Quantification of the full-length transcript was calculated with wild-type set at 1.0 and is shown in each lane.
A.

Figure 2.11. Differences in the \textit{rpsJ} operon and the 16S rRNA in the \textit{rne-1} array.

Changes in the steady-state levels of the genes that are involved in the \textit{rpsJ} (panel A) and \textit{rrn} (panel B) operons. The \textit{rnc-14} and \textit{rne}1018/\textit{rng}-219 array (bottom in A, top in B) were previously published (13). The \textit{rne}1018/\textit{rng}-219 (top in A, bottom in B) was
previously analyzed in this laboratory (Bowden and Kushner, unpublished results.) The rne-1 array was performed as part of this study. The image presented was obtained from a screen shot of the Integrated Genome Browser program (20). Gene or operon names appear above or below the operons that encode them, which indicate their location in the genome relative to nucleotide coordinates, as displayed in the center of the graph. Black arrows indicate the direction of transcription. The array data is displayed as vertical lines representing the log2 ratio of fluorescence between the mutant and wild-type strains.
Figure 2.12. Microarray comparison for the rpsJ ribosomal operon when strains were grown at comparable growth rates. Changes in the steady-state levels of the genes that are involved in the rpsJ operon. The rne-1 array was performed as part of this study and is also shown in Fig. 2.11. The rne\(\Delta\)1018/rng-219 was previously analyzed in this laboratory (Bowden and Kushner, unpublished results.) The rne\(\Delta\)1018/rng-219 in M9 medium array was performed as part of this study. The image presented was obtained from a screen shot of the Integrated Genome Browser program (20). Operon names appear above or below the operons that encode them, which indicate their location in the genome relative to nucleotide coordinates, as displayed in the center of the graph. Black arrows indicate the direction of transcription. The array data is displayed as vertical lines representing the log2 ratio of fluorescence between the mutant and wild-type strains.
Figure 2.13. Analysis of *rpsJ* transcript from strains grown at comparable growth rates in minimal medium. Cells were grown at 37°C. RNA was run on a 1.5% glyoxyl agarose gel. Strains were MG1693 (wild-type) and SK3564 (*rneΔ1018/rng-219*) in Lanes 1 and 2, respectively.
Figure 2.14. Agarose northern blots of wild-type and *rneΔ1018/rng-219* strains grown in M9 minimal medium and probed for several intercistronic regions. Cells were grown at 37°C. RNA was run on a 1.5% glyoxyl agarose gel. Strains were MG1693 (wild-type) and SK3564 (*rneΔ1018/rng-219*) in Lanes 1 and 2, respectively.
Figure 2.15. Chemical half-lives of the full-length $rpsJ$ transcript. The calculated half-life of the full-length $rpsJ$ transcript was $2.2 \pm 0.56$ minutes for MG1693 (wild-type) and $3.86 \pm 0.89$ minutes for SK5665 ($rne-1$). Cells were initially grown at 30°C with shift to 44°C for 1hr. RNA was run on a 1.5% glyoxyl agarose gel.
Figure 2.16. Chemical half-lives of the rpsJ operon transcript, the rplN operon transcript, and adh transcript in SK4161. Cells were initially grown at 30°C with shift to 44°C for 1hr. RNA from SK4161 (rne-1 Δrnz::apr ΔmazEFG) was run on a 1.5% glyoxyl agarose gel. Arrows indicate the full-length transcript of 5.2 kb, 5.6 kb, and 3.0 kb for rpsJ, rplN, and adhE, respectively.
REFERENCES


CHAPTER 3

POLYADENYLATION OF MESSENGER RNAs BY POLY(A) POLYMERASE I

IN *ESCHERICHIA COLI* IS LINKED TO THE PRESENCE OF AN

INTACT RNASE E-BASED DEGRADOSOME¹

ABSTRACT

In *Escherichia coli*, poly(A) polymerase I (PAP I) has been shown to account for 90% of the poly(A) tails that are added to mRNAs, rRNAs, and pre-tRNAs. Here we show that the polyadenylation of mRNAs by PAP I is directly linked to an intact RNase E-based degradosome. The polyadenylation profiles of specific mRNA transcripts (*lpp, rpsJ* and *rpsQ*) changed dramatically both in the location of the tails and in their nucleotide composition when the scaffold region of the RNase E protein was deleted. Specifically, in the absence of degradosome assembly, the majority of the tails for specific mRNA transcripts were heteropolymeric and were added by polynucleotide phosphorylase (PNPase). The polyadenylation profiles of PNPase mutants encoding either a non-functional PNPase, or enzymes with either reduced degradative activity or severely reduced biosynthetic activity were also examined to determine the possible role of PNPase in PAP I-mediated polyadenylation. Our data suggest the existence of a more extensive polyadenylation complex and a more direct link between mRNA degradation and polyadenylation than previously envisioned. Additionally, since RNase E and PAP I have both been previously demonstrated to be associated with the inner membrane during exponential growth, it appears that polyadenylation of mRNAs by PAP I likely occurs at the inner membrane.
INTRODUCTION

Polyadenylation at the 3' termini of mRNAs is an important post-transcriptional event in both eukaryotes and prokaryotes. In eukaryotes, the addition of poly(A) tails increases mRNA stability by preventing both 5'→3' and 3'→5' degradation and promotes more efficient translation (1-3). In contrast, polyadenylation of bacterial mRNAs, as demonstrated in *Escherichia coli*, leads to relatively short tails that appear to target the mRNAs for more rapid degradation (4-9).

In *E. coli*, less than 2% of total RNA is polyadenylated at one time, however, more than 80% of the mRNA transcripts can be substrates for polyadenylation (10,11). It has been hypothesized that since *E. coli* does not have a 5'→3' exonuclease, nor can most of its exonucleases degrade though secondary structures, that the addition of the poly(A) tails to the 3' end of mRNAs aid in their ability to degrade the secondary structures, particularly Rho-independent transcription terminators that are associated with mRNAs.

Polyadenylation in *E. coli* is carried out by two enzymes, poly(A) polymerase I (PAP I) and polynucleotide phosphorylase (PNPase). PAP I, encoded by *pcnB* (12), is responsible for 90% of polyadenylation and synthesizes short poly(A) tails (1-40 nt) that exclusively contain adenosine residues (7,9,13). However, tail lengths are reduced by 3'→5' exonucleases such as RNase II and PNPase (7,14-16). *In vivo*, poly(A) tails are found on both mRNA decay products and full-length transcripts after Rho-independent transcription terminators (9,11,15,17). In fact, over 80% of poly(A) tails are normally located after Rho-independent transcription terminators (17). Over production of PAP I causes cell death, since mature tRNAs become substrates for PAP I, causing a loss of
protein synthesis and cell death (18). However, in exponentially growing cells there are no significant changes in the growth properties of ΔpcnB strains (7,9,13).

The second enzyme that can post-transcriptionally add nucleotides onto the 3’ ends of mRNAs is PNPase. This enzyme functions both as a 3’→5’ exonuclease that degrades RNA using a phosphorolytic mechanism (19) and as a poly(A)-type polymerase that can biosynthetically synthesize long heteropolymeric tails that contain all four nucleotides (20). PNPase adds tails at locations within the mRNA transcripts that are significantly different than those synthesized by PAP I (17,20). Of particular interest is the fact that PNPase has been shown to form complexes with both PAP I (17) and RNase E (21). However, the exact mechanism by which PNPase functions in the polyadenylation of mRNAs is not clear.

PNPase is also an integral part of the RNase E-based degradosome, a key complex involved in mRNA degradation, which contains PNPase, the RhlB RNA helicase and enolase (21-23) associated with the RNase E scaffold domain. Deletion experiments have demonstrated that association regions for the RhlB RNA helicase, enolase and PNPase are located at amino acids 734-738, 739-845 and 844-1045 of the RNase E scaffold, respectively (24). Furthermore, it has been argued that Hfq, an abundant 11.2 kDa RNA-binding protein, can also associate with RNase E, competing with the RhlB RNA helicase for binding in the region of amino acids 734-738 (25).

Polyadenylation of mRNA transcripts in E. coli has also been shown to involve interaction of additional proteins. One protein that has been shown to be involved in PAP I polyadenylation is Hfq. Inactivation of Hfq dramatically reduces the ability of PAP I to add poly(A) tails to transcripts after Rho-independent transcription terminators, even
though intracellular PAP I protein levels remain constant (17). Additionally, total polyadenylation levels are dramatically lower and poly(A) tails observed in an *hfq* mutant are shorter, even when PNPase, RNase II or RNase E are also missing, demonstrating that the shortened tails are not the result of nucleolytic degradation (17). Pull-down and immuno-precipitation experiments have demonstrated physical interactions among PAP I, PNPase and Hfq, suggesting the existence of a multiprotein complex that is required for the polyadenylation of mRNAs with Rho-independent transcription terminators. However, it is important to remember that these proteins are not present in the cell in similar concentrations. PNPase has been shown to be present in large excess relative to RNase E (5-10 fold during exponential phase) (26), while PAP I has only been estimated to have ~32-50 molecules per cell (17).

Although the biological function of the heteropolymeric tails is currently unknown, polyadenylation of mRNAs by PNPase has been observed in a number of other bacteria (27-30). In fact, *E. coli* strains with higher levels of PNPase were observed to have higher levels of heteropolymeric tails (20). It was also shown that PNPase, unlike PAP I, usually does not add tails after secondary structures such as Rho-independent transcription terminators (17,20), although the reasons for the distinction are currently unclear.

Therefore, it was surprising when macroarray analysis of the *E. coli* genome revealed that multiple ORFs within a polycistronic mRNA were targets for polyadenylation (11). For example, with the *rpsJ* polycistronic mRNA, which is over 5,100 nt in length and contains 11 ORFs (Fig. 3.1), polyadenylation of many of its ORFs was detected with the array (11). When initial experiments confirmed that both the first
ORF in the operon \((rpsJ)\) and the last ORF \((rpsQ)\) were polyadenylated, we decided to examine this phenomenon in greater detail.

Since both Hfq and PNPase, which interact with PAP I, are also associated with the RNase E-based degradosome, we wondered if there was a link between PAP I mediated polyadenylation of mRNAs and the RNase E-based degradosome. Furthermore, if there were a link between polyadenylation and the RNase E-based degradosome, we wanted to determine what components of the degradosome were essential for their interaction. In order to delineate the relationship of these two complexes, if any, the polyadenylation profiles of three ORFs within the polycistronic \(rpsJ\) operon \((rpsJ, rplV, rpsQ)\) and the transcript for the monocistronic \(lpp\) mRNA were examined in \(E. coli\) mutants defective for degradosome assembly or PNPase function along with a wild-type control. Here we show that for transcripts primarily polyadenylated by PAP I, the main components of the degradosome are required for polyadenylation by PAP I. In the absence of an intact degradosome, the majority of the post-transcriptionally added tails were synthesized by PNPase.

**MATERIALS AND METHODS**

**Bacterial Strains and Plasmids**

All the strains and plasmids used are listed in Table 1. RNase E partial scaffold deletions were constructed in MG1693 \((thyA715 rph-1)\) by P1-mediated transduction using either AC-24 \((rne\Delta10 zce-726::Tn10 Tc^R)\) or AC-26 \((rne\Delta18 zce-726::Tn10 Tc^R)\) (31) as donor strains. The \(rne\Delta10\) and \(rne\Delta18\) mutations are missing the RNase E scaffold from amino acids 844-1045 and 728-845, respectively, such that PNPase fails
to bind in the \textit{rne}∆10 strain while enolase and the RhlB RNA helicase fail to bind in \textit{rne}∆18. AC-24 and AC-26 were the generous gifts of the Carpousis laboratory. MG1693 was provided by the \textit{E. coli} Genetic Stock Center (Yale University). Transductants were confirmed by PCR and DNA sequencing to verify the locations of scaffold deletions.

SK2683 was constructed using a plasmid displacement method. The \textit{rne}∆374 fragment from the 6-8 copy plasmid pMOK16 (32) was transferred into the single copy plasmid pMOK40 (Sm\textsuperscript{R}/Sp\textsuperscript{R}) (33) at the EcoRI-NotI sites and subsequently transformed into SK9714 (32). The covering plasmid (pSBK1) carrying \textit{rne}\textsuperscript{+} was displaced by pWSK129 (34), which had the same origin of DNA replication. Figures of each of the RNase E mutants are shown in Figure 3.2.

Strains were created that carried mutations in the \textit{pnp} gene in which single amino acid substitutions led to alterations in either the 3'\textendash5' exonuclease or polymerization activities, but retained the normal trimeric structure of the PNPase protein \textit{in vivo} (35). The R100D allele lacks both the biosynthetic and degradative activity of PNPase, the N435D has normal biosynthetic activity but severely reduced degradative activity, while the D490N allele has almost normal degradative activity but severely reduced biosynthetic activity (35) (Table 3.2).

Plasmids carrying \textit{pnp} point mutations were constructed in a 6-8 copy vector and under the control of the native \textit{rpsO pnp} promoter as follows: pKMK12 (\textit{rpsO}\textsuperscript{+} \textit{pnp}\textsuperscript{+}) was first created by \textit{HindIII/KpnI} digestion of pKAK7 (36) to excise the native \textit{rpsO}\textit{pnp} fragment which was subsequently ligated into the \textit{HindIII/KpnI} sites of pWSK29 (34). As both pWSK29 and pKAK7 are Ap\textsuperscript{R}, the ligated plasmid was digested with \textit{SalI} prior to transformation into DH5\textalpha to disrupt any possible pKAK7-derivatives. The pKMK12
plasmid was sequenced using M13R and M13F universal primers to ensure correct junctions.

Next, site directed mutagenesis was performed to create R100D, N435D, or D490N plasmid point mutations using Q5 Site-Directed Mutagenesis Kit (NEB) and the following primer sets to amplify pKMK12 \((rpsO^+ pnp^-)\): “R100Dpnp_Fwd2” [CCGCCCGATTGACCCGCTGTTC], “R100Dpnp_Rev2” [TCAATCAGACGCGCGATC], “N435Dpnp_Fwd2” [CACTGAATCCGACGGTTCCTC], “N435Dpnp_Rev2” [ATTTCAGACAACAACGTAC], “D490Npnp_Fwd2” [TCACCTGGGCAATATGGACTTC], and “D490Npnp_Rev2” [TCTTCGTCCGCCAAAATG]. The PCR reactions were digested with KLD enzyme mix as directed and subsequently transformed into NEB 5α. Plasmids were isolated from the transformants, sequenced for confirmation, and named pKMK13 \((rpsO^+ pnp^-R100D)\), pKMK14 \((rpsO^+ pnp^-N435D)\), and pKMK15 \((rpsO^+ pnp^-D490N)\).

Following plasmid construction, SK10019 \((pnp^-1683)\) (20) was transformed with either pKMK12 \((rpsO^+ pnp^-)\), pKMK13 \((rpsO^+ pnp^-R100D)\), pKMK14 \((rpsO^+ pnp^-N435D)\), pKMK15 \((rpsO^+ pnp^-D490N)\), or pWSK29 (vector) to create SK4436, SK4437, SK4438, SK4439, and SK4443, respectively.

The \(rne-1\) allele encodes a temperature-sensitive mutation that functionally inactivates RNase E at 44°C (37). The \(hfq-1\) mutation is a deletion/insertion mutation that produces no Hfq protein (17). The \(pnp^-7\) and \(pnp^-1683\) mutations are a point mutation and chromosomal deletion insertion, respectively, that have been described previously (20,38).
Bacterial growth and doubling times

To obtain growth curves, bacterial strains were grown at 37°C with shaking at 255 rpm in Luria broth supplemented with thymine (50 µg/mL) and streptomycin (20 mg/mL) or ampicillin (200 mg/mL), where appropriate. Measurements were taken every 30 minutes using a Klett-Summerson colorimeter with a green filter (No. 42). Cells were diluted with pre-warmed growth medium as needed to maintain the cultures in exponential growth. Growth curves were done in triplicate and values averaged before the data was plotted.

Pull-down analysis of the RNase E-based degradosome in various mutants

To determine if any of the pnp mutations led to alterations in the compositions of the RNase E-based degradosome complex, pull-down assays were performed. Strains were grown at 37°C with shaking at 255 rpm to Klett 80 in Luria broth supplemented with thymine (50 µg/mL) and antibiotic, where appropriate. Cell cultures (50 mL) were collected for protein isolation and quantified as described previously (17). Immuno-precipitations were performed at 4°C by mixing two mg of total protein and PNPase-antibody cross-linked to Pierce Protein A/G magnetic beads (Thermo Scientific; 25 µl) for 1 h in a shaker. The protein–bead complexes were washed three times with RIPA buffer (150mM NaCl, 50 mM Tris-HCl pH 8, 1.0% Nonidet P-40, 0.5% DOC, 0.1% SDS). The immunocomplexes were subsequently eluted from the beads in 40 µl of the SDS-PAGE loading buffer after heating at 55°C for 15 min, followed by boiling for 5 min. For PNPase (2 µl diluted 1:10) and for RNase E (10 µl) immunoprecipitation samples were separated on 7.5% (for RNase E) and 10% (for PNPase) SDS-PAGE gels and
probed with their respective antibodies at a 1:10,000 and 1:2,500 for PNPase (anti-rabbit from A.J. Carpousis laboratory) and RNase E (32), respectively.

**Mapping of post-transcriptionally added poly(A) tails**

Total RNA was isolated from various exponentially growing strains as described previously (7,39) and treated with DNase I to remove any residual DNA contamination present. The quality of RNA samples was checked by agarose gel electrophoresis and then quantified using a Nanodrop® ND-2000c spectrophotometer. Subsequently, oligo(dT)-dependent cDNAs were obtained for rpsJ, rplV, rpsQ, and lpp mRNAs using reverse transcription-PCR with 5' gene-specific and 3' oligo(dT)$_{17}$ primers, as described previously (15,39). The 5' gene-specific primers were as follows: “rpsJ 5' GS” [GTCTGAT CGATCAAGCAACC], “rpsQ 5' GS” [TAATGATCGATAAAATCCGTACTTCTGC], “rplV 5' GS” [AAACTATCGATAAACATCGCCATGCTC], and “lpp366” [GCTACATGGAGATTAACTCAAGCTTGAGGG]. All gene-specific primers contained a ClaI site, except for the lpp gene-specific primer which contained a HindIII site, and were designed such that the entire coding sequence of lpp and rpsQ ORFs and the final 260 and 329 nucleotides of the coding sequence of rpsJ and rplV, respectively, would be amplified. The cDNAs were all cloned into the ClaI/XbaI or HindIII/XbaI sites of pWSK29 (34), as described previously (20,39). Preliminary DNA sequencing was performed for MG1693-derived tails using BigDye® Terminator v3.1 Cycle Sequencing kit and an automated sequencer (Applied Biosystems 3730x1 DNA analyzer). All other sequencing was performed by MacrogenUSA using an M13 Reverse primer.

Sequences were then analyzed and compared to each specific gene using the online
BLAST blastn program (NCBI) selected for *E. coli* strain MG1655 (taxid:511145) in the NCBI nucleotide collection. Sequence results with tails starting less than three nucleotides following the 5' gene-specific primer sequence were not included to eliminate any mispriming events.

**Poly(A) Sizing Assay**

Total RNA was isolated from various exponentially growing strains as described previously (7,39). The RNA was treated with DNase I to remove any residual DNA contamination present. The quality of RNA samples was checked by agarose gel electrophoresis and then quantified using a Nanodrop® ND-2000c spectrophotometer. Subsequently, total RNA was 3'-end-labeled with [³²P]-pCp and then digested with RNase A and RNase T1 and then separated on a 20% PAGE, as described previously (7,39). A 5'-[³²P]-end-labeled oligo(dA) ladder was also run, as described previously (7,39). The gel was exposed to X-ray film overnight.

**RESULTS**

**Polyadenylation profiles of the *rpsJ*, *rplV*, and *rpsQ* transcripts in wild-type *E. coli***

In the macroarray analysis of the *E. coli* transcriptome, it was reported that many of the ORFs within the polycistronic *rpsJ* transcript were subject to polyadenylation (11). In order to confirm that the various ORFs within the *rpsJ* operon were in fact polyadenylated in wild-type *E. coli*, we chose three representative genes from the beginning (*rpsJ*), middle (*rplV*), and end (*rpsQ*) of the polycistronic transcript (Fig. 3.1).
for direct analysis. cDNAs from each ORF were obtained and cloned as described in the Materials and Methods.

As shown in Table 3.3, the majority of the tails associated with the \textit{rpsQ} transcript were homopolymeric indicating that they had been synthesized by PAP I. Additionally, as expected, the majority of the homopolymeric tails had been added after the Rho-independent transcription terminator (Fig. 3.3A), in agreement with previously reported results for the \textit{lpp} mRNA (Fig. 3.4A) (9,20). In the case of the \textit{rpsJ} ORF transcript, almost 50% of the tails associated were homopolymeric (Table 3.3). In contrast, only 20% of the tails associated with the \textit{rplV} transcript were homopolymeric (Table 3.3).

**Polyadenylation profiles in strains defective in functional degradosome assembly**

The presence of homopolymeric tails on multiple ORFs within the polycistronic \textit{rpsJ} transcript raised the possibility of an association between the polyadenylation reaction of PAP I and the processing of the full-length \textit{rpsJ} transcript by the RNase E-based degradosome. To examine a possible link between PAP I and the RNase E-based degradosome, the polyadenylation profiles of \textit{rpsJ}, \textit{rplV}, and \textit{rpsQ} transcripts were also analyzed in strains defective in functional degradosome assembly. As a positive control, the polyadenylation profile of the \textit{lpp} transcript was also examined, as it has been well-characterized previously (9,20). For these experiments, we employed a complete scaffold mutant (\textit{rne\Delta374}) which has previously been shown to prevent degradosome formation (32), and two partial scaffold mutants in which either the PNPase scaffold binding region was deleted (\textit{rne\Delta10}) or the RhlB and Enolase scaffold binding regions
were missing (rneΔ18) (Fig. 3.2). All three mutants contain a wild-type catalytic region of RNase E, and thus function normally with regards to mRNA decay and tRNA maturation (32).

As shown in Table 3.3, the majority of the tails associated with rpsJ, rpsQ and lpp were heteropolymeric in all three degradosome mutants. It is important to note that the deletion of the entire scaffold region (rneΔ374) had the most dramatic effect on the polyadenylation profiles with the percentage of homopolymeric tails dropping 2.6-fold for rpsJ, 6.5-fold for rpsQ, and 3-fold for lpp (Table 3.3). Surprisingly, even the prevention of the association of RhlB and enolase with the degradosome led to a significant reduction in the fraction of homopolymeric tails (Table 3.3). Not only was the number of homopolymeric tails dramatically reduced in the absence of degradosome assembly, but the physical location of the tails changed significantly for the rpsQ and lpp mRNAs in the absence of degradosome assembly (Figs. 3.3, 3.4). In the case of the lpp mRNA, the absence of a functional degradosome led to all of the tails being added well upstream of the Rho-independent transcription terminator. However, when we tested the rplV transcript in the rneΔ374 scaffold mutant, it polyadenylation profile was comparable to that observed in the wild-type control (Table 3.3, Fig. 3.5). It should also be noted that the heteropolymeric tails were always significantly longer than the homopolymeric tails (Table 3.3).

**Polyadenylation profiles for strains defective in Hfq**

Since the Hfq protein has been shown to play a role in polyadenylation of the lpp mRNA (17) and it has been suggested that it can also associate with the degradosome (25),
we examined the polyadenylation profiles of the *rpsJ* and *rpsQ* mRNAs in an *hfq-1* mutant. As shown in Table 3.3, the majority of the tails associated with *rpsJ*, *rpsQ* and *lpp* mRNAs were heteropolymeric in the *hfq-1* mutant, indicating that they had been synthesized by PNPase. These results were similar to the result seen in the degradosome mutants and seen previously for the *lpp* mRNA (17).

**Analysis of strains carrying point mutations in the gene for PNPase**

PNPase has been shown to play an integral role in both the functional degradosome (21) and the polyadenylation complex (17). It was therefore of interest to determine how important the presence of an intact PNPase trimer was to the polyadenylation of various mRNAs. To carry out these experiments, we took advantage of three specific point mutations that had previously been isolated by Jarriage *et al.* (35). The *pnp*-R100D allele encodes a PNPase with normal trimeric structure *in vivo* but no 3′→5′ exonuclease and polymerization activities, the N435D-*pnp* allele encodes PNPase with normal trimeric structure *in vivo* and almost biosynthetic normal activity but severely reduced degradative activity, and the D490N-*pnp* allele encodes PNPase with its normal trimeric structure *in vivo* and almost normal degradative activity but severely reduced biosynthetic activity (Table 3.2) (35).

In order to determine if any of these mutants played a role in polyadenylation we constructed a series of low-copy number plasmids carrying these alleles under the control of their native *rpsO* *pnp* promoter and transformed them into a *pnp*-Δ683 (20) deletion strain. Strains carrying *pnp*+, *pnp*-N435D, and *pnp*-D490N all grew with doubling times of 29.5 ± 2 minutes (Fig. 3.6). This result was comparable to the
doubling time of MG1693 (data not shown). The strains with the pnp-R100D allele or empty vector grew with doubling times of 52.5 ± 2 minutes and 45 ± 2 minutes, respectively. These results were comparable to the growth and doubling time of the pnpΔ683 strain with no plasmid present (Fig. 3.6).

**Pull-down analysis of the RNase E-based degradosome in various mutants**

Although it has been shown that all three PNPase point mutants make stable protein in vivo (35), we wanted to confirm that these proteins were incorporated in the RNase E-based degradosome. Accordingly, to determine if any mutant strain used prevented proper interaction of PNPase with the RNase E-based degradosome complex, pull-down assays were performed for the wild-type strain, the degradosome mutants (rneΔ374, rneΔ10, rneΔ18), the PNPase mutants (pnp-R100D, pnp-N435D, pnp-D490N), and an hfq-1 mutant strain. The rneΔ374 degradosome mutant and pnpΔ683 deletion mutant were used as internal controls, as they have been previously determined to lack interaction of PNPase with RNase E (32)(Mohanty and Kushner, unpublished results). The pull-down analysis demonstrated that all of the PNPase point mutants and the rneΔ18 mutant retained the interactions between RNase E and PNPase, as expected. Furthermore, the meΔ10 mutant strain, which lacks the PNPase binding region of degradosome was defective in degradosome assembly (Fig. 3.7).

**Polyadenylation profiles for strains with altered PNPase proteins**

To determine if alterations in the PNPase protein affected PAP I polyadenylation, the rpsJ, rpsQ and lpp transcripts from various PNPase mutants were analyzed. As shown
in Table 3.4, almost all of the tails associated with all three transcripts in the \textit{pnp\_A683} deletion strain were homopolymeric indicating that they had been synthesized by PAP I, but the physical location of the tails changed significantly for the \textit{lpp} and \textit{rpsQ} mRNAs in a strain lacking PNPase compared to the wild-type control (Figs. 3.3, 3.4). In the case of both the \textit{rpsQ} and \textit{lpp} mRNAs, the lack of a functional PNPase protein led to all of the tails being added upstream of the Rho-independent transcription terminator (Figs. 3.3C, 3.4C, 3.8AB). This shift was also seen in mutants defective for PNPase without a functional RNase E gene, and both with and without the deletion of the exonuclease RNase II (Fig. 3.9, Table 3.5).

Surprisingly, the absence of a functional protein did not alter the locations of the poly(A) tails for the \textit{rpsJ} transcript (Fig. 3.8C). Similar polyadenylation profiles were observed in strains carrying the \textit{pnp-7} allele as well as mutants in either RNase E (\textit{rne-1}) or the 3'\rightarrow5' exonuclease, RNase II (\textit{rnb500}) (Table 3.5). The locations of their tails were comparable with the data shown in Table 3.4 and Fig. 3.9, namely tail locations were only shifted if the polyadenylation was directly associated with a Rho-independent transcription terminator.

**Polyadenylation profiles for strains defective in PNPase activities**

To determine if the physical presence of PNPase in the degradosome complex or partial PNPase enzymatic activity was sufficient for proper PAP I polyadenylation, the \textit{rpsJ}, \textit{rpsQ} and \textit{lpp} transcripts from mutants with an altered PNPase were analyzed. A strain encoding PNPase that retains its normal trimeric structure \textit{in vivo} but lacks both 3'\rightarrow5' exonuclease and polymerization activities, a strain encoding PNPase which retains its
normal trimeric structure \textit{in vivo} and almost biosynthetic normal activity but has severely reduced degradative activity, and a strain encoding PNPase which retains its normal trimeric structure \textit{in vivo} and almost normal degradative activity but has severely reduced biosynthetic activity (35) were analyzed.

As shown in Table 3.4, almost all of the tails associated with the transcripts examined in both the \textit{pnp}\textbackslash{}A683 deletion mutant and the inactive \textit{pnp}\textbackslash{}R100D mutant were homopolymeric, indicating that they had been synthesized by PAP I. In the case of the \textit{lpp} and \textit{rpsQ} transcripts, the majority of the homopolymeric tails in the PNPase mutants were added within the ORF before the Rho-independent transcription terminator (Fig. 3.8). The profiles of \textit{pnp}\textbackslash{}N435D looked more similar to the wild-type control (Table 3.4), but still contained a higher percentage of heteropolymeric tails than observed in wild-type conditions and the tails were located throughout the coding sequence (Fig. 3.8), as opposed to primarily after the Rho-independent transcription terminator in wild-type conditions. Additionally, the profile of the tails added in the \textit{pnp}\textbackslash{}D490N mutant strain was predominately heteropolymeric, more similar to the degradosome scaffold mutant than to the wild-type control.

\textbf{Total poly(A) profiles in various \textit{E. coli} strains}

Although the data obtained with the \textit{rpsJ}, \textit{rpsQ} and \textit{lpp} mRNAs demonstrated an important role for degradosome assembly and Hfq in the polyadenylation of specific mRNAs, we wanted to confirm its importance in total polyadenylation in the cell. Accordingly, we examined total poly(A) levels using a poly(A) sizing assay for strains defective in either functional degradosome assembly or defects in PNPase activity.
Total RNA was 3'-end-labeled with $[^{32}\text{P}]-\text{pCp}$ and then digested with RNase A and RNase T1. Since RNase A cleaves after U and C residues and RNase T1 cleaves after G residues, 3’ poly(A) tails are protected from cleavage and are separated by PAGE, as described in the Materials and Methods. As shown in Fig. 3.10, failure to assemble the degradosome led to a significant reduction in the length and total amount of poly(A) tails in the cell.

**DISCUSSION**

**PAP I polyadenylation depends on an intact degradosome**

Examination of specific mRNAs (rpsJ, rpsQ, lpp) that are primarily polyadenylated by PAP I in wild-type strains showed that the percentage of homopolymeric tails decreased significantly in the absence of either full degradosome assembly or the Hfq protein (Fig. 3.3). No change was seen in the rplV transcript, which was primarily polyadenylated by PNPase I in the wild-type strain (Table 3.3, Fig. 3.5). The significant shift in polyadenylation profiles in the various degradosome mutants (Fig. 3.5), from primarily homopolymeric to primarily heteropolymeric, suggested that PAP I-polyadenylated transcripts require an intact degradosome, since the catalytic function of RNase E was intact in these degradosome mutants (32).

**Polyadenylation by PNPase can function independently of an intact degradosome**

Examination of the rplV transcript, which was primarily polyadenylated by PNPase in the wild-type strain (Table 3.3), showed that the percentage of heteropolymeric remained unchanged in the absence of either full degradosome assembly (Fig. 3.3). This result
strongly suggests that unlike PAP I, the biosynthetic activity of PNPase does not require an intact degradosome and is likely independent of the degradosome/polyadenylation complex interaction.

**PAP I added tails are shifted in location in the absence of an intact degradosome**

As shown in the contrasting results of Figs. 3.3B and 3.4B compared to Figs. 3A and 4A, failure to assemble an intact degradosome led to a dramatic shift in both the composition and location of the post-transcriptionally PAP I-added homopolymeric tails in the *rpsQ* and *lpp* transcripts. In the wild-type control, the homopolymeric tails of these transcripts were predominately located after the Rho-independent transcription terminator (Figs. 3.3A, 3.4A), while in the degradosome mutant strains (Figs. 3.3B, 3.4B) the homopolymeric tails were located throughout the *lpp* coding sequence.

In fact, it appeared that when the degradosome was missing, PAP I was no longer able to interact or recognize the 3’ terminus of the Rho-independent transcription terminator (Fig. 3.4A versus Fig. 3.4B), its preferred substrate in wild-type cells (17). This result was consistent with previously published data for the *lpp* mRNA in an *hfq-1* mutant (17), suggesting Hfq also plays a role in a polyadenylation complex-degradosome interaction.

**The physical presence of PNPase in the degradosome is not sufficient for proper PAP I polyadenylation**

Since the profile of the *pnpΔ683* deletion mutant and the inactive *pnp-R100D* mutant were experimentally similar (Table 3.4, Fig. 3.8), it would appear that the physical
presence of PNPase in its native trimeric form was not sufficient for wild-type PAP I polyadenylation of mRNAs. The profiles of pnp-N435D and pnp-D490N mutant strains seen in Table 3.4 were not comparable to what was seen in the wild-type control, especially with regards to the location of the tails (Fig. 3.8), even though PNPase retained partial function \textit{in vivo}. These results suggest that PNPase must not only be physically present in its native trimeric form, but also must have both normal exonucleolytic and biosynthetic activity for wild-type PAP I polyadenylation of mRNAs. Additionally, immunoprecipitation that showed all the PNPase point mutants properly interacted with the degradosome complex (Fig. 3.7) confirms that the effects on polyadenylation in these strains was due to the loss of normal function of PNPase and not any loss of PNPase/degradosome interaction.

**PAP I added tails are also shifted in location in the absence of functional PNPase**

As expected, in the \textit{pnp}Δ683 and \textit{pnp}R100D mutant strains, there were only a small number of tails that were not homopolymeric (Table 3.4). The few that did appear were dramatically shorter in length (Table 3.4), suggestive of either experimental artifacts or a rare misincorporation of a single nucleotide by PAP I. The elimination of heteropolymeric tails was consistent with previous work showing that deletion or inactivation of PNPase eliminated heteropolymeric tailing (15,17,20).

Furthermore, there was a dramatic shift in location of the homopolymeric tails present in all the PNPase mutant strains compared to the wild-type control (Figs. 3.3C, 3.4C, 3.8), particularly for transcripts directly associated with Rho-independent transcription terminators. These results were similar to what was observed for the
degradosome mutants and previously seen in *hfq-1* mutants (17). However, only a slight shift was seen for the *rpsJ* transcript (Fig. 3.8C), suggesting the shift was tied to the presence of a Rho-independent transcription terminator. These results suggest PNPase must be both present and functionally active for PAP I to polyadenylate mRNAs after Rho-independent transcription terminators and that when PNPase is absent, PAP I acts in a different manner to interact or act on its preferred mRNA substrates to post-transcriptionally add tails.

The shifts in location of tails raises the question of which interactions are key to PAP I’s dependency on a functional degradosome, and if the effect of PNPase on PAP I polyadenylation is due to the direct interaction of PNPase with the degradosome. Further examination of the manner by which PNPase affects PAP I substrate affinity is necessary.

**Processed ORF transcripts within a polycistronic transcript can have different polyadenylation profiles**

Previous polyadenylation studies had not examined the polyadenylation of large polycistronic transcripts. Thus, the observation from macroarray analysis that many ORFs within a polycistronic transcript showed polyadenylation patterns that were directly related to the transcription terminator of the operon (11) suggested that the transcription terminator of a polycistronic transcript may be a polyadenylation signal for an entire operon. Our results demonstrate that this is not the case, but rather that individual processed ORF transcripts within a polycistronic transcript can have significantly different polyadenylation profiles. Of the three genes chosen here for
analysis, the majority of the tails associated in the wild-type control with the \( rpsQ \) and \( rpsJ \) ORF transcripts had been synthesized by PAP I, while the majority of the tails associated with \( rplV \) had been synthesized by PNPase (Table 3.3). This was also confirmed by the changes in the polyadenylation profiles of \( rpsJ \) and \( rpsQ \) degradosome scaffold mutant (\(rne\Delta374\)), while \( rplV \) transcript profile remained unchanged compared to wild-type. These results suggest that the transcript terminator does not act as a polyadenylation signal for the entire polycistronic transcript.

**Total PAP I polyadenylation levels are reduced in strains lacking either an intact degradosome or Hfq**

When a poly(A) sizing assay on a degradosome scaffold mutant (\(rne\Delta374\)) and \(hfq-1\) mutant was performed, a dramatic drop in the number and length of poly(A) tails was observed compared to the wild-type strain (Fig. 3.10), although the drop was not as dramatic as previously seen for a \(\Delta pcnB\) deletion strain (17). These results were in agreement with previously published \(hfq-1\) poly(A) sizing assays (17).

This observed decrease in total polyadenylation levels would suggest that the shift for transcripts that are primarily polyadenylated by PAP I in wild-type strains to being primarily polyadenylated by PNPase in degradosome or \(hfq-1\) mutants (Fig. 3.3) was due to a decrease in overall PAP I polyadenylation as opposed to an increase in PNPase polyadenylation.
Proper PAP I polyadenylation of mRNAs is likely localized to the inner membrane

We hypothesize that the mechanism of polyadenylation of mRNAs is directly related to the interaction of the polyadenylation complex with the membrane-bound degradosome (40,41). This interaction is likely mediated through the action of the Hfq protein, which has been shown to interact with both the polyadenylation complex (17) and the degradosome scaffold region (25). The interaction of the two complexes brings PAP I in closer proximity to its mRNA substrates, which is important since PAP I is present in such low amounts in the cell (17). It may also explain why the polyadenylation of mRNAs requires Hfq, while the polyadenylation of tRNAs does not (13).

Conclusions

Taken together our results provide important insights into the role of an intact degradosome in PAP I polyadenylation in *E. coli*. Our data strongly suggest the presence of a more extensive polyadenylation complex and its need for interaction with an intact degradosome for normal PAP I polyadenylation. Our data also indicate that the polyadenylation complex interacts with the degradosome through some interaction, most likely with Hfq and possibly PNPase. Since previous studies have shown 90% of mRNA polyadenylation is done by PAP I under wild-type conditions (9,17), a degradosome-PAP I interaction could suggest a major role for the degradosome in proper mRNA polyadenylation in *E. coli*. The demonstration of a large polyadenylation complex is a novel concept in bacteria and will require further analysis to understand the precise role of PNPase and the degradosome in polyadenylation on PAP I polyadenylation.
Acknowledgements

Special thanks to Dr. A.J. Carpousis for his gift of AC-24 (rneΔ10), AC-26 (rneΔ18), and PNPase anti-rabbit antibodies. This work was supported by NIH grants GM57220 and GM81554 to S.R.K.
Figure 3.1. Pictorial representation of the *rpsJ* operon, also known as the S10 operon. The operon contains 11 genes, all of which encode ribosomal proteins. The full-length transcript is approximately 5.2 kb in length and contains a Rho-independent transcription terminator.
Table 3.1. List of bacterial strains and plasmids used in this study.

<table>
<thead>
<tr>
<th>Strains</th>
<th>Genotype</th>
<th>Reference or Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC-24</td>
<td>MC1061 _rne_10 (aaΔ 844-1045) _zce_726::Tn10 Tc&lt;sup&gt;R&lt;/sup&gt;</td>
<td>(31)</td>
</tr>
<tr>
<td>AC-26</td>
<td>MC1061 _rne_18 (aaΔ 728-845) _zce_726::Tn10 Tc&lt;sup&gt;R&lt;/sup&gt;</td>
<td>(31)</td>
</tr>
<tr>
<td>MG1693</td>
<td><em>thyA715 rph-1</em></td>
<td><em>E. coli</em> Genetic Stock Center</td>
</tr>
<tr>
<td>SK2683</td>
<td>_rne_1018::bla <em>thyA715 rph-1</em></td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>pDHK3(Sp/Sm&lt;sup&gt;R&lt;/sup&gt; _rne_374 single copy)</td>
<td></td>
</tr>
<tr>
<td>SK4433</td>
<td>_rne_10 (aaΔ 844-1045) _zce_726::Tn10 _thyA715 rph-1 Tc&lt;sup&gt;R&lt;/sup&gt;</td>
<td>This Study</td>
</tr>
<tr>
<td>SK4434</td>
<td>_rne_18 (aaΔ 728-845) _zce_726::Tn10 _thyA715 rph-1 Tc&lt;sup&gt;R&lt;/sup&gt;</td>
<td>This Study</td>
</tr>
<tr>
<td>SK4436</td>
<td>_pnp_683::spc/str <em>thyA715 rph-1</em></td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>pKMK12 ( <em>rpsO</em> pnp&lt;sup&gt;+&lt;/sup&gt;, Ap&lt;sup&gt;R&lt;/sup&gt; low copy)</td>
<td></td>
</tr>
<tr>
<td>SK4437</td>
<td>_pnp_683::spc/str <em>thyA715 rph-1</em></td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>pKMK13 ( <em>rpsO</em> pnp-R100D, Ap&lt;sup&gt;R&lt;/sup&gt; low copy)</td>
<td></td>
</tr>
<tr>
<td>SK4438</td>
<td>_pnp_683::spc/str <em>thyA715 rph-1</em></td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>pKMK14 ( <em>rpsO</em> pnp-N435D, Ap&lt;sup&gt;R&lt;/sup&gt; low copy)</td>
<td></td>
</tr>
<tr>
<td>SK4439</td>
<td>_pnp_683::spc/str <em>thyA715 rph-1</em></td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>pKMK15 ( <em>rpsO</em> pnp-D490N, Ap&lt;sup&gt;R&lt;/sup&gt; low copy)</td>
<td></td>
</tr>
<tr>
<td>SK4443</td>
<td>_pnp_683::spc/str <em>thyA715 rph-1</em></td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>pWSK29 (empty vector, Ap&lt;sup&gt;R&lt;/sup&gt; low copy)</td>
<td></td>
</tr>
<tr>
<td>SK5704</td>
<td>_rne-1 pnp-7 rnb-500 <em>thyA715 rph-1</em></td>
<td>(38)</td>
</tr>
<tr>
<td>SK5726</td>
<td>_pnp-7 <em>thyA715 rph-1</em></td>
<td>(38)</td>
</tr>
<tr>
<td></td>
<td>pDK39 ( _Cm&lt;sup&gt;R&lt;/sup&gt; rnb-500)</td>
<td></td>
</tr>
<tr>
<td>SK7988</td>
<td>_ΔpcnB::kan <em>thyA715 rph-1</em> Km&lt;sup&gt;R&lt;/sup&gt;</td>
<td>(7)</td>
</tr>
<tr>
<td>SK9714</td>
<td>_rne_1018::bla <em>thyA715 rph-1</em></td>
<td>(32)</td>
</tr>
<tr>
<td></td>
<td>pSBK1(_rne+, _Cm&lt;sup&gt;R&lt;/sup&gt;)</td>
<td></td>
</tr>
<tr>
<td>SK10019</td>
<td>_pnp_683::spc/str <em>thyA715 rph-1</em></td>
<td>(20)</td>
</tr>
<tr>
<td>SK10023</td>
<td>_hfq-1::Ω <em>thyA715 rph-1</em> Km&lt;sup&gt;R&lt;/sup&gt;</td>
<td>(17)</td>
</tr>
</tbody>
</table>
Figure 3.2. Description of the RNase E degradosome mutants used this study.

The label on the degradosome scaffold region indicates the associated binding site of the various proteins. The S1 domain is associated with RNA binding. The function of the arginine-rich RNA binding site (ARRBS) is not clear.

Table 3.2. Previously published activities of PNPase point mutants compared to a wild-type control (35).

<table>
<thead>
<tr>
<th>PNPase Mutants</th>
<th>Exonuclease Activity</th>
<th>Biosynthetic Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>R100D</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>N435D</td>
<td>26%</td>
<td>120%</td>
</tr>
<tr>
<td>D490N</td>
<td>70%</td>
<td>25%</td>
</tr>
</tbody>
</table>
Table 3.3. Comparison of the polyadenylation profiles of degradosome mutants. Poly(A) tails were compared with regards to percent of total tails in each tail type (homo-polymeric or heteropolymeric) and average length of each type of tail. Previously published polyadenylation profiles of *lpp* for wild-type and *hfq-1* (17) are included for comparison and are indicated by asterisks (*). N.D.= not determined.

<table>
<thead>
<tr>
<th></th>
<th><em>rpsJ</em></th>
<th></th>
<th><em>rpsQ</em></th>
<th></th>
<th><em>lpp</em></th>
<th></th>
<th><em>rpIV</em></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>wt</td>
<td>meΔ374</td>
<td>meΔ10</td>
<td>meΔ18</td>
<td>hq-1</td>
<td>wt</td>
<td>meΔ374</td>
<td>meΔ10</td>
</tr>
<tr>
<td>n =</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>11</td>
<td>23</td>
<td>17</td>
<td>15</td>
<td>25</td>
<td>24</td>
<td>22</td>
</tr>
<tr>
<td>% of homo-polymeric Tails (nt)</td>
<td>46</td>
<td>18</td>
<td>26</td>
<td>35</td>
<td>7</td>
<td>52</td>
<td>8</td>
<td>41</td>
</tr>
<tr>
<td>Avg. homo-polymeric tail length (nt)</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>15</td>
<td>9</td>
<td>20</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>% of hetero-polymeric tails</td>
<td>54</td>
<td>82</td>
<td>74</td>
<td>65</td>
<td>93</td>
<td>48</td>
<td>92</td>
<td>59</td>
</tr>
<tr>
<td>Avg. hetero-polymeric tail length (nt)</td>
<td>58</td>
<td>67</td>
<td>63</td>
<td>40</td>
<td>69</td>
<td>55</td>
<td>115</td>
<td>70</td>
</tr>
</tbody>
</table>

123
Figure 3.3. Pictorial representation of tails found on the rpsQ transcript. The number in parentheses following the tail is the number of that type of tail located in that region. Only the longest heteropolymeric tails are shown.
Figure 3.4. Pictorial representation of tails found on the *lpp* transcript. The number in parentheses following the tail is the number of that type of tail located in that region. Only the longest heteropolymeric tails are shown. The previously published polyadenylation profiles of *lpp* for wild-type (17) is included for comparison.
Figure 3.5. Percent composition of the post-transcriptionally added tails in various degradosome mutants. The complete data set is shown in Table 3.3.

Previously published polyadenylation profiles of lpp for wild-type and hfq-1 (17) are included for comparison and are indicated by asterisks (*).
Figure 3.6. Growth curves for various PNPase mutants. Strains carrying the \(pnp^+\), \(pnp\)-N435D, and \(pnp\)-D490N alleles grew with doubling times of 29.5 ± 2 minutes. The strains carrying the \(pnp\)-R100D allele and or an empty vector grew with doubling times of 52.5 ± 2 minutes and 45 ± 2 minutes, respectively.
Figure 3.7. Immunoprecipitation of PNPase and scaffold mutants. Proteins were immunoprecipitated using PNP-tagged A/G beads. Subsequent pull-down elutions were run on SDS-PAGE and blotted to PVDF. Western blots were probed with either PNPase or RNase E.
Table 3.4. Comparison of the polyadenylation profiles of PNPase mutants.

Poly(A) tails were compared with regards to percent of total tails in each tail type (homopolymeric or heteropolymeric) and average length of each type of tail. Previously published polyadenylation profiles of \textit{lpp} for wild-type (17) are included for comparison and are indicated by asterisks (*). N.D.= not determined.

<table>
<thead>
<tr>
<th></th>
<th>\textit{rpsJ}</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>\textit{wt}</td>
<td>\textit{pnpΔ683}</td>
<td>\textit{pnp-R100D}</td>
<td>\textit{pnp-N435D}</td>
<td>\textit{pnp-D490N}</td>
<td>\textit{wt}</td>
<td>\textit{pnpΔ683}</td>
<td>\textit{pnp-R100D}</td>
<td>\textit{pnp-N435D}</td>
<td>\textit{pnp-D490N}</td>
<td>\textit{wt*}</td>
<td>\textit{pnpΔ683}</td>
</tr>
<tr>
<td>(n=)</td>
<td>26</td>
<td>22</td>
<td>20</td>
<td>18</td>
<td>25</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>23</td>
<td>21</td>
<td>27*</td>
<td>28</td>
</tr>
<tr>
<td>% of homopolymeric Tails (nt)</td>
<td>46</td>
<td>91</td>
<td>85</td>
<td>33</td>
<td>12</td>
<td>52</td>
<td>85</td>
<td>96</td>
<td>43</td>
<td>29</td>
<td>70*</td>
<td>93</td>
</tr>
<tr>
<td>Avg. homopolymeric tail length (nt)</td>
<td>17</td>
<td>17</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>20</td>
<td>16</td>
<td>17</td>
<td>20</td>
<td>17</td>
<td>ND</td>
<td>16</td>
</tr>
<tr>
<td>% of heteropolymeric tails</td>
<td>54</td>
<td>9</td>
<td>15</td>
<td>67</td>
<td>88</td>
<td>48</td>
<td>15</td>
<td>4</td>
<td>57</td>
<td>71</td>
<td>30*</td>
<td>7</td>
</tr>
<tr>
<td>Avg. heteropolymeric tail length (nt)</td>
<td>58</td>
<td>15</td>
<td>23</td>
<td>57</td>
<td>64</td>
<td>55</td>
<td>20</td>
<td>49</td>
<td>29</td>
<td>41</td>
<td>ND</td>
<td>18</td>
</tr>
</tbody>
</table>
Figure 3.8. Shift in location of homopolymeric tails on \textit{rpsQ}, \textit{lpp}, and \textit{rpsJ} transcripts for wild-type versus PNPase mutants. The size of the dot is representative of the number of tails at particular location. Heteropolymeric tails are not shown. The \textit{pnpD490N} mutants are not shown as there are not enough homopolymeric tails present. (See Table 3.4 for summary of heteropolymeric tail data).

(A) \textit{rpsQ}

(B) \textit{lpp}

(C) \textit{rpsJ}
Table 3.5. Comparison of the polyadenylation profiles of strains carrying the pnp-7 allele. Poly(A) tails were compared with regards to percent of total tails in each tail type (homopolymeric or heteropolymeric) and average length of each type of tail. Previously published polyadenylation profiles of lpp for wild-type (17) are included in this work for comparison and are indicated by asterisks (*). N.D. = not determined.

<table>
<thead>
<tr>
<th></th>
<th>rpsJ</th>
<th>rpsQ</th>
<th>lpp</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rne-1 pnp-7 mb-500</td>
<td>rne-1 pnp-7 mb-500</td>
<td>rne-1 pnp-7 mb-500</td>
</tr>
<tr>
<td></td>
<td>pnp-7 mb-500</td>
<td>pnp-7 mb-500</td>
<td>pnp-7 mb-500</td>
</tr>
<tr>
<td></td>
<td>wt</td>
<td>wt</td>
<td>wt</td>
</tr>
<tr>
<td></td>
<td>n =</td>
<td>n =</td>
<td>n =</td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>25</td>
<td>27*</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>23</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>% of homopolymeric Tails (nt)</td>
<td>% of homopolymeric Tails (nt)</td>
<td>% of homopolymeric Tails (nt)</td>
</tr>
<tr>
<td></td>
<td>46</td>
<td>52</td>
<td>70*</td>
</tr>
<tr>
<td></td>
<td>85</td>
<td>96</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>96</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>Avg. homopolymeric tail length (nt)</td>
<td>Avg. homopolymeric tail length (nt)</td>
<td>Avg. homopolymeric tail length (nt)</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>20</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>% of heteropolymeric tails</td>
<td>% of heteropolymeric tails</td>
<td>% of heteropolymeric tails</td>
</tr>
<tr>
<td></td>
<td>54</td>
<td>48</td>
<td>30*</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Avg. heteropolymeric tail length (nt)</td>
<td>Avg. heteropolymeric tail length (nt)</td>
<td>Avg. heteropolymeric tail length (nt)</td>
</tr>
<tr>
<td></td>
<td>58</td>
<td>55</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>27</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>14</td>
<td>19</td>
</tr>
</tbody>
</table>
Figure 3.9. Shift in location of homopolymeric tails on rpsQ, lpp, and rpsJ transcripts for wild-type versus other PNPase mutants. The size of the dot is representative of the number of tails at particular location. Heteropolymeric tails are not shown. (See Table 3.5 for summary of heteropolymeric tail data).

(A) rpsQ

(B) lpp

(C) rpsJ
Figure 3.10. Poly(A) sizing assay for various mutants. Total RNA from MG1693, SK7988, SK10023, SK3564, SK4433, and SK4434 was 3'-end-labeled with $[^{32}P]$-pCp, digested, and then separated on a 20% polyacrylamide gel. A 5'-$[^{32}P]$-end-labeled oligo(dA) ladder is shown in lane 1.
REFERENCES


CHAPTER 4

CONCLUSIONS AND FUTURE DIRECTIONS

PROCESSING OF THE RPSJ OPERON

Understanding the mechanisms of RNA processing, degradation, and polyadenylation in *Escherichia coli* is key to understanding the basics of cellular metabolism and post-transcriptional regulation. In this work, we examined the processing and/or breakdown along with the polyadenylation profiles of the large polycistronic *rpsJ* mRNA transcript. Our data from Chapter 2 demonstrated that RNase E was the major ribonuclease involved in the initial processing of the *rpsJ* transcript and is responsible for the removal of the 3’ Rho-independent transcription terminator, confirming experimentally the general assumption that large polycistronic transcripts are processed into smaller, more manageable pieces.

As shown in Chapter 2, the cellular ratios of the individual ORF transcripts of the polycistronic *rpsJ* operon remained relatively constant, despite changes in the processing of the full-length transcript in RNase E mutants. This result suggested that the flexibility and adaptation potential to control the amounts of ribosomal proteins at the transcript level was much greater than previously observed or envisioned.
POLYADENYLATION IN *E. coli*

**The Role of the RNase E based degradosome in Polyadenylation**

Through the examination of the *rpsJ* operon transcript, our studies demonstrated the importance of the RNase E-based degradosome to maintain proper PAP I polyadenylation. In Chapter 3, we have shown that the presence of Hfq, a fully functional PNPase, and an intact degradosome are required for normal PAP I polyadenylation, particularly of transcripts containing a Rho-independent transcription terminator. When the scaffold region of the RNase E protein was deleted, the polyadenylation profiles of individual mRNAs encoded within the *rpsJ* operon (*rpsJ* and *rpsQ*) and a control transcript (*lpp*) changed dramatically both in the location of the post-transcriptionally added tails and in their nucleotide composition. The absence of PNPase or lack of functional PNPase resulted in similar observations. Our results provide important insight into the role of an intact degradosome in PAP I polyadenylation in *E. coli*, especially the need for interaction of the polyadenylation complex with an intact degradosome for normal PAP I activity.

**The Polyadenylation Complex**

As also shown in Chapter 3, our data suggested the existence of a much larger polyadenylation complex than previously proposed (1). We propose a complex composed of the RNase E-based degradosome, PNPase, PAP I, and Hfq. Our data also suggest a more direct link between mRNA degradation and polyadenylation than previously envisioned, with the polyadenylation complex interacting with the
degradosome via interactions with PNPase, Hfq, or both. The demonstration of a large polyadenylation complex is a novel concept in bacteria and requires further analysis to understand the precise role of PNPase and the degradosome on PAP I polyadenylation. Further work needs to be done to examine the actual interaction of PNPase, Hfq, and the degradosome scaffold region. It is still unknown what key physical interactions, if any, are necessary. Additionally, as there is strong evidence that the majority of RNase E binds to the inner cytoplasmic membrane (2-5) and that PAP I is associated with or localized to the inner membrane during exponential growth (6,7), it needs to be determined if localization to the inner membrane is an important component of normal polyadenylation of mRNAs. Further examination is also necessary to understand if PAP I is localized to the inner membrane independently, or due to its interaction with the RNase E-based degradosome. These experiments are important because it does not appear that the PAP I polyadenylation of tRNAs requires Hfq.

PAP I Dependence on PNPase

While there is much known about the polyadenylation enzymes in *E. coli*, the exact mechanism by which PAP I and PNPase identify and select their RNA substrates is still not completely understood. The dependence of PAP I on PNPase needs to be examined more fully. Our work in Chapter 3 demonstrated that PNPase-mediated polyadenylation is independent of PAP I function and is not affected by the absence of an intact degradosome, but that PAP I polyadenylation is dependent on the function of PNPase, the presence of intact degradosome, and Hfq. It remains to be understood
why PAP I requires interaction with PNPase, Hfq, and the degradosome to function effectively as a poly(A) polymerase, while PNPase function is independent.

**The Relationship of Polyadenylation and mRNA Processing and Decay**

As the RNase E-based degradosome plays a major role in both processing and polyadenylation, the extent of the link between the polyadenylation and processing functions needs to be examined further. The link between the RNase E-based degradosome and normal PAP I polyadenylation suggests the possibility of a closer interaction between processing or breakdown/decay and polyadenylation than previously understood. It also remains to be seen if the poly(A) tails after a Rho-independent transcription terminator function differently compared to the tails added to processed transcripts. Furthermore, it is not known if such tails are recognized by different enzymes at different stages of the decay pathway.

**Polyadenylation Signals**

It remains unclear precisely what signals PAP I or PNPase to add a tail to a particular substrate, although our data in Chapter 2 has demonstrated that polyadenylation of polycistronic transcripts occurs even when separated from the transcription terminator. Additionally, since the processed ORFs of the *rpsJ* polycistronic transcript are not all polyadenylated in the same manner, there appears to be some distinct polyadenylation signaling that is unique to each ORF transcript fragment.

**FINAL CONCLUSION**
Overall, our work has provided insight into important aspects of processing and turnover/decay of mRNA transcripts that are essential for the cell’s ability to rapidly adapt to changing environmental conditions. The understanding of these RNA processing, degradation, and polyadenylation activities in *Escherichia coli* is key to gaining a better understanding of the basics of cellular metabolism and regulation. Additionally, our work has strengthened the possibility of a closer tie between mRNA processing and turnover/decay than previously demonstrated, suggesting a larger control of these mechanisms by the involved enzymes.
REFERENCES


APPENDIX A

DELETION MUTANTS OF TRANSFER RNA_{LEU}^{2} ARE VAILABLE AND EXHIBIT FLEXIBILITY IN CODON-ANTICODON RECOGNITION IN ESCHERICHIA COLI\(^1\)

\(^1\) Kristen B. Mildenhall, Bijoy K. Mohanty, and Sidney R. Kushner. To be submitted to Nucleic Acids Research.
ABSTRACT

The process of translation requires that the codons within an mRNA be accurately recognized by appropriately charged tRNAs. The leuU tRNA (encoding tRNA^{Leu2}) of *E. coli* is unique in several aspects. It is one of the few tRNAs encoded in a polycistronic operon that also contains an mRNA, being co-transcribed with secG, a membrane translocation protein. More importantly, of the eight leucine tRNAs in *E. coli*, tRNA^{Leu2} is the only one that recognizes the CUC and CUU leucine codons in exponentially growing cells.

Here we report that *leuU* is not essential for cell viability in the MG1655, W3110, and C600 genetic backgrounds of *E. coli* K12. The ΔsecG leuU strains are viable at 30°C and 37°C but not at 44°C. Additionally, the cell morphology of the ΔleuU mutants was dramatically altered, with cells being significantly smaller in all dimensions compared to wild-type controls and the presence of elongated chains. Additionally, it was verified that the deletion of tRNA^{Leu2} was the cause of decreased growth and viability, temperature sensitivity, and altered cellular morphology by using complementation plasmids carrying either secG+ leuU+ or just leuU+.

We hypothesize that the cell is utilizing a more complex mechanism of alternative base-pairing to compensate for the inability to recognize the CUC and CUU codons, and that the CUC codon is also read by tRNA^{Leu3} through a U:C wobble base pairing.
INTRODUCTION

The leuU tRNA (encoding tRNA\textsuperscript{Leu\textsubscript{2}}) of \textit{E. coli} is unique in that it is one of the few tRNAs in a polycistronic operon that also contains an mRNA. It is co-transcribed with secG, a membrane translocation protein, which is directly involved in protein translocation across the membrane and is considered vital at low temperatures. A study done by Nishiyama \textit{et al.} (1) demonstrated that considerable amounts of a SecG precursor accumulated at 20°C but not at 37°C, suggesting the importance of this protein at cold temperatures due to the cell’s inability to survive in the absence of mature, functional protein.

More importantly, of the eight leucine tRNAs in \textit{E. coli}, tRNA\textsuperscript{Leu\textsubscript{2}} is the only one that recognizes the leucine CUU and CUC codons in exponentially growing cells. In fact, when Nishiyama and Tokuda (2) characterized the secG leuU transcript, they reported that a ΔsecG leuU mutant was not viable. Since secG could be deleted without effecting cell viability, they concluded that leuU was essential. Subsequently, Sorensen \textit{et al.} (3) also found that under starvation conditions, tRNA\textsuperscript{Leu\textsubscript{3}} (encoded by leuW) uses wobble from a uridine-5-oxyacetic acid modification to also recognize the CUU codon.

However, when experiments in our laboratory demonstrated that the leuU tRNA was initially processed from the secG leuU transcript by an RNase P cleavage at the mature 5’ terminus of leuU (4), the question arose as to whether the leuU tRNA was really essential for cell viability. Here we describe the construction and characterization of ΔsecG leuU mutants in three different genetic backgrounds. In all cases, the presence of a functional leuU tRNA was not essential for cell viability at either 30°C or 37°C.
Even though Sorensen et al. (3) found that under starvation conditions \( \text{tRNA}^{\text{Leu}3} \) (encoded by \( \text{leuW} \)) used wobble from a uridine-5-oxyacetic acid modification to also recognize the CUU codon, our results suggest a greater flexibility in codon-anticodon recognition, such that the cell can overcome the loss of this unique tRNA. We hypothesize that the cell is utilizing a more complex mechanism of alternative base-pairing to compensate for the inability to recognize the CUC and CUU codons, and that the CUC codon is also read by \( \text{tRNA}^{\text{Leu}3} \) through a U:C wobble base pairing.

**MATERIALS AND METHODS**

**Plasmids**

Plasmids were isolated using the QIAprep Spin Miniprep Kit (Qiagen) or E.Z.N.A.® Plasmid Mini Kit (OMEGA Bio-Tek). pAG3 (\( \text{secG}^+ \text{leuU}^+ \text{Ap}^R \)) was obtained from KN470/pAG3 strain (2) and contained \( \text{secG}^+ \text{leuU}^+ \) under the control of an arabinose-inducible promoter.

pBMK48 (\( \text{leuU}^+ \text{Cm}^R \)) was constructed by using overlap PCR to create a DNA fragment containing the coding sequence and Rho-independent transcription terminator of \( \text{leuU} \) with a 5’ \( \text{BamHI} \) and a 3’ \( \text{HindIII} \) restriction sites. The insert and vector pBMK11 DNA (5) were both digested with \( \text{BamHI} \) and \( \text{HindIII} \) and ligated using Fast-Link™ DNA Ligase (Epicentre®). Plasmid DNA was sequenced to ensure successful cloning of the insert. pBMK48 has \( \text{leuU}^+ \) under the control of an IPTG-inducible promoter and is a derivative of pAG702.
**Bacterial Strains**

Viable $\Delta$secG leuU deletions were constructed in the MG1693, W3110, and C600 genetic backgrounds of *E. coli* K-12 using a P1 transducing lysate obtained from a strain carrying a $\Delta$secG leuU::Sm$^\text{r}$/Sp$^\text{r}$ allele (K. Nishiyama laboratory). This mutation was a complete insertion/deletion of the dicistronic secG leuU operon. These three genetic backgrounds were selected, since C600 was used in the original work (2), while MG1693 (a MG1655 derivative) is the common strain used in our laboratory. W3110 was also included as its complete genome has also been sequenced. MG1693 (thyA715 rph-1) was provided by the *E. coli* Genetic Stock Center (Yale University).

Complementation plasmids were transformed into the desired strains by chemical transformation. Each parent wild-type strain and their four respective $\Delta$secG leuU derivatives strains were transformed with either pAG3 (secG$^+$ leuU$^+$) or pBMK48 (leu$^+$). Following transformation, the resulting strains were analyzed for temperature sensitivity and growth phenotypes. When indicated, 1mM IPTG or 1 mM arabinose was added to induce pBMK48 (leuU$^+$) and pAG3 (secG$^+$ leuU$^+$), respectively.

**PCR and Northern Analysis**

The presence of the leuU deletion in transductants was confirmed by both PCR and northern blotting. Total RNA was isolated as described previously (6). Subsequent quantification and normalization of RNA samples and northern analysis were then carried out as described previously (7). Northern blots were probed with a leuU probe, stripped, and reprobed with a glt probe. The four glt tRNA genes are located between the 23S and 16S rRNA subunits in four of the seven *E. coli* ribosomal RNA operons.
**Bacterial Growth and Viability**

To obtain growth curves, bacterial strains were grown in Luria broth supplemented with thymine (50 µg/mL) and streptomycin (20 mg/mL), chloramphenicol (20 mg/mL) or ampicillin (200 mg/mL), where appropriate, at 30°C with shaking until they reached a cell density of approximately 50 Klett units above background. The Klett units were measured using a Klett-Summerson colorimeter with a green filter (No. 42). Subsequently, the cultures were shifted to 44°C for 2-3 hours, as indicated. Measurements were taken at 30 minute intervals throughout growth.

Cell viability was determined by plating dilutions of cultures immediately before shift to 44°C and at 30 minute time points following the shift. Luria agar plates supplemented with thymine and antibiotic, as needed, were then grown at 30°C for 2-3 days and colonies were counted.

**Microscopy**

To examine cellular morphology, cultures were grown up with shaking, as described above, until they reached a cell density of 40-45 Klett units above background. Aliquots were removed and fixed with glutaraldehyde. The cultures were then shifted to 44°C. Aliquots were again taken at 150 min after shift and fixed with glutaraldehyde. Fixed cells were then placed in a solution of low-melting agarose, and imaged with a Leica DM6000B Light Microscope at 1000x.
RESULTS

Characterization, growth, and viability of ΔsecG leuU mutant strains

After the ΔsecG leuU deletions were confirmed by northern analysis (Fig. A.1), the growth properties for each ΔsecG leuU deletion strain were compared to their respective wild-type controls. All of the deletion strains were temperature sensitive at 44°C (Table A.2), except in the MG1693 genetic background, where both temperature-sensitive and temperature resistant transductants were obtained. Additionally, all of the transductants were cold-sensitive (Table A.3), which was consistent with previously identified properties of secG mutants (1).

To examine the temperature sensitive growth of these mutant strains, growth was plotted before and after shift to the non-permissive temperature. Aliquots were also taken at various points after temperature shift and plated at 30°C to examine the viability of the mutants. For the MG1693 derivatives, the temperature sensitive SK10457 (ΔsecG leuU) mutant both grew more slowly at 30°C and then growth leveled off around 120 minutes after temperature shift (Fig. A.2A). The temperature resistant mutant, SK10456 (ΔsecG leuU), grew slower than the MG1693 wild-type, but its growth was not affected by the temperature shift. The viability of the SK10457 (ΔsecG leuU) strain dropped at time points up to 90 minutes at which point it leveled off (Fig. A.2B).

For the W3110 derivative, the temperature sensitive SK10475 (ΔsecG leuU) mutant grew more slowly at 30°C and then growth leveled off around 120 minutes after shift to 44°C (Fig. A.3A). Viability of the SK10474 (ΔsecG leuU) mutant was lower than the wild-type control, but never decreased like what was observed with the MG1693-derivative mutant (Fig. A.3B).
For the C600 derivative, SK10473 (ΔsecG leuU) mutant grew more slowly at 30°C. Growth leveled off around 90 minutes after temperature shift after which time cell density actually decreased (Fig. A.4A). The loss of cell viability in SK10473 (ΔsecG leuU) was the most dramatic of the mutant strains, as it dropped immediately and never leveled off (Fig. A.4B).

**Characterization of strain morphology in ΔsecG leuU mutants**

To examine the effect of the ΔsecG leuU mutation on the morphology of the cell, both wild-type and mutants were examined using light microscopy. This showed that the cell morphology of the ΔsecG leuU mutants was dramatically altered. The mutant cells were significantly smaller in all dimensions from the wild-type control and also contained elongated chains (Fig. A.5).

**Complementation of ΔsecG leuU mutant strains**

Each parental wild-type strain and their respective ΔsecG leuU derivatives strains were transformed with either pAG3 (secG+ leuU*) and pBMK48 (leuU*) to verify by complementation that the leuU deletion was the cause of the growth defect at 44°C. The strains containing pBMK48 or pAG3 were tested for temperature sensitivity. Both the MG1693-derivative mutants and W3110 derivative mutants regained their growth (Figs. A.6, A.7) when complemented with either pAG3 (secG+ leuU*) or pBMK48 (leuU*). Of particular interest was the fact that this complementation occurred even when the leuU* plasmid was not induced (Figs. A.6, A.7).
DISCUSSION

The *leuU* gene is not essential for cell viability at 30°C or 37°C but leads to sensitive growth at 44°C in all three genetic backgrounds.

Here we show that the *leuU* gene is not essential for cell viability in the MG1655, W3110, or C600 *E. coli* K12 backgrounds. However, the *ΔsecG leuU* deletion in each genetic background demonstrated slower growth at the permissive temperature. Temperature-sensitive growth was seen in all three genetic backgrounds (Tables A.2, A.3). The *ΔsecG leuU* mutation exhibited the most dramatic effects in the C600 background, where we saw rapid loss in cell viability at 44°C.

This loss of growth ability was demonstrated through complementation to be caused by the *ΔsecG leuU* deletion, and was reversed by complementation of either a *secG*<sup>+</sup> *leuU*<sup>+</sup> or *leuU*<sup>+</sup> plasmid (Figs. A.6, A.7). Additionally, the complementation occurred even when the *leuU*<sup>+</sup> plasmid was present but not induced (Figs. A.6, A.7), suggesting that the minimal expression from the leakiness of the pBMK48 *lac* promoter was sufficient to restore the mutants to wild-type growth.

Another tRNA must be compensating for tRNA<sup>Leu<sub>2</sub></sup>

Since the *leuU* gene is not essential, but no tRNA is known to recognize the CUC codon, another tRNA must be compensating for the tRNA<sup>Leu<sub>2</sub></sup> through non-traditional wobble base pairing. Future peptidase sequencing and subsequent analysis of specific tRNA charging levels in the cell will help determine which amino acid is being inserted opposite the CUC codon and compensating for the tRNA<sup>Leu<sub>2</sub></sup>. Regardless, this tRNA
compensation suggests a greater flexibility in codon-anticodon recognition, such that the cell can overcome the loss of this unique tRNA.

Acknowledgments
Special thanks to Dr. K. Nishiyama for his gift of pAG3. Additional thanks to undergraduate students Wyatt Ramey, for his work on the viability curves and construction of the complementation strains and Mayson Callaway, for his work on complementation growth curves. Also thanks to Dr. Richard Meagher and Dr. M.K. Kandasamy for use and help with the Leica DM6000B Light Microscope. This work was supported by NIH grants GM57220 and GM81554 to S.R.K.
Table A.1. List of bacterial strains and plasmids used in this study. The \( \Delta secG \)\( \text{leuU::spc/str} \) is a complete deletion/insertion of the dicistronic \( secG \)\( \text{leuU} \) operon.

Unless indicated otherwise, \( \Delta secG \)\( \text{leuU} \) mutants are temperature sensitive at 44°C.

<table>
<thead>
<tr>
<th>Strains</th>
<th>Genotype</th>
<th>Plasmid</th>
<th>Reference or Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>MG1693</td>
<td>( \text{thyA715 rph-1} )</td>
<td>n/a</td>
<td>( E. \ coli ) Genetic</td>
</tr>
<tr>
<td></td>
<td>( \Delta secG )( \text{leuU::spc/str (non-ts)} )</td>
<td></td>
<td>Stock Center</td>
</tr>
<tr>
<td>SK10456</td>
<td>( \text{thyA715 rph-1} )</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>SK10457</td>
<td>( \Delta secG )( \text{leuU::spc/str thyA715 rph-1} )</td>
<td>n/a</td>
<td>This study</td>
</tr>
<tr>
<td>SK5270</td>
<td>( \text{thyA715 rph-1 leuU*} )</td>
<td>pBMK48 (leuU*)</td>
<td>This study</td>
</tr>
<tr>
<td>SK5271</td>
<td>( \Delta secG )( \text{leuU::spc/str (non-ts)} )</td>
<td>pBMK48 (leuU*)</td>
<td>This study</td>
</tr>
<tr>
<td>SK5273</td>
<td>( \Delta secG )( \text{leuU::spc/str thyA715 rph-1} )</td>
<td>pBMK48 (leuU*)</td>
<td>This study</td>
</tr>
<tr>
<td>SK5278</td>
<td>( \text{thyA715 rph-1 secG leuU*} )</td>
<td>pAG3 (secG(^+) leuU(^+))</td>
<td>This study</td>
</tr>
<tr>
<td>SK5279</td>
<td>( \Delta secG )( \text{leuU::spc/str (non-ts)} )</td>
<td>pAG3 (secG(^+) leuU(^+))</td>
<td>This study</td>
</tr>
<tr>
<td>SK5280</td>
<td>( \Delta secG )( \text{leuU::spc/str thyA715 rph-1} )</td>
<td>pAG3 (secG(^+) leuU(^+))</td>
<td>This study</td>
</tr>
<tr>
<td>W3110</td>
<td>( \text{rph-1} )</td>
<td>n/a</td>
<td>( E. \ coli ) Genetic</td>
</tr>
<tr>
<td>SK10475</td>
<td>( \Delta secG )( \text{leuU::spc/str rph-1} )</td>
<td>n/a</td>
<td>This study</td>
</tr>
<tr>
<td>SK5274</td>
<td>( \text{leuU*} )</td>
<td>pBMK48 (leuU*)</td>
<td></td>
</tr>
<tr>
<td>SK5275</td>
<td>( \Delta secG )( \text{leuU::spc/str rph-1} )</td>
<td>pBMK48 (leuU*)</td>
<td>This study</td>
</tr>
<tr>
<td>SK5281</td>
<td>( \text{secG leuU*} )</td>
<td>pAG3 (secG(^+) leuU(^+))</td>
<td>This study</td>
</tr>
<tr>
<td>SK5282</td>
<td>( \Delta secG )( \text{leuU::spc/str rph-1} )</td>
<td>pAG3 (secG(^+) leuU(^+))</td>
<td>This study</td>
</tr>
<tr>
<td>C600</td>
<td>( \text{leuB6 thr-1 thi-1 lacY1 glnV44} )( \text{fhuA21 cyn-101 glpR200} )</td>
<td>n/a</td>
<td>( E. \ coli ) Genetic</td>
</tr>
<tr>
<td>SK10473</td>
<td>( \Delta secG )( \text{leuU::spc/str} )( \text{leuB6 thr-1 thi-1 lacY1 glnV44} )( \text{fhuA21 cyn-101 glpR200} )</td>
<td>n/a</td>
<td>This study</td>
</tr>
<tr>
<td>SK5276</td>
<td>( \Delta secG )( \text{leuU::spc/str} )( \text{leuB6 thr-1 thi-1 lacY1 glnV44} )( \text{fhuA21 cyn-101 glpR200} )</td>
<td>pBMK48 (leuU*)</td>
<td>This study</td>
</tr>
<tr>
<td>SK5277</td>
<td>( \Delta secG )( \text{leuU::spc/str leuU*} )</td>
<td>pBMK48 (leuU*)</td>
<td>This study</td>
</tr>
</tbody>
</table>
SK5283
leuB6 thr-1 thi-1 lacY1 glnV44
fhuA21 cyn-101 glpR200
secG leuU
leuB6 thr-1 thi-1 lacY1 glnV44
pAG3 (secG' leuU')
This study

SK5284
fhuA21 cyn-101 glpR200
ΔsecG leuU::spc/str secG leuU
pAG3 (secG' leuU')
This study

| C600 wild-type | C600 ΔsecG leuU derivative (ts) |
| MG1693 wild-type | MG1693 ΔsecG leuU derivative (non-ts) |
| MG1693 ΔsecG leuU derivative (ts) | W3110 wild-type |
| W3110 ΔsecG leuU derivative (ts) |

Probed with leuU

Probed with glt

Figure A.1. The presence of the leuU deletion in transductants was confirmed by Northern blotting. Northern blots of all three wild-type strains and respective ΔsecG leuU mutants probed with leuU and glt RNA probes. Cells were grown at 30°C. RNA was run on 1.2% polyacrylamide northern. Strains were as follows: C600, SK10473 (ΔsecG leuU), MG1693, SK10456 (ΔsecG leuU), SK10457 (ΔsecG leuU), W3110, and SK10475 (ΔsecG leuU) in Lanes 1-7, respectively.
Table A.2. Growth of strains on Luria broth plates after 48 hours. Strains were as follows: MG1693, SK10456 (ΔsecG leuU), SK10457 (ΔsecG leuU), W3110, SK10475 (ΔsecG leuU), C600, and SK10473 (ΔsecG leuU).

<table>
<thead>
<tr>
<th></th>
<th>MG1693</th>
<th>SK10456</th>
<th>SK10457</th>
<th>W3110</th>
<th>SK10475</th>
<th>C600</th>
<th>SK10473</th>
</tr>
</thead>
<tbody>
<tr>
<td>30°C</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>44°C</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

Table A.3. Cold sensitivity of SK10473

<table>
<thead>
<tr>
<th>Viability of SK10473 at 4°C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Cell Viability</td>
</tr>
<tr>
<td>Viability after 3 days at 4°C</td>
</tr>
</tbody>
</table>
Figure A.2. Growth curve and viability curve of SK10456 and SK10457 ΔsecG leuU mutants versus the MG1693 wild-type control. Growth is shown in panel A at 30°C and after shift (indicated by the arrow) to 44°C. Cell viability is shown in panel B and is taken at 30 minute time intervals after shift to 44°C. Viability data is not shown for temperature resistant derivative SK10456.
Figure A.3. Growth curve and viability curve of SK10475 ΔsecG leuU mutant versus the W3110 wild-type control. Growth is shown in panel A at 30°C and after shift (indicated by the arrow) to 44°C. Cell viability is shown in panel B and is taken at 30 minute time intervals after shift to 44°C.
Figure A.4. Growth curve and viability curve of SK10473 ΔsecG leuU mutant versus the C600 wild-type control. Growth is shown in panel A at 30°C and after shift (indicated by the arrow) to 44°C. Cell viability is shown in panel B and is taken at 30 minute time intervals after shift to 44°C.
Figure A.5. Comparative morphology of wild-type versus ∆secG leuU deletion strain, as seen under bright field microscopy. The pictures of MG1693 and its ∆secG leuU derivative, SK10457 were taken prior to temperature shift, t = 0 (30°C) and again 150 minutes following the shift (44°C).
Figure A.6. MG1693 ∆secG leuU mutants with complementation in the absence of IPTG induction. Growth is shown at 30°C and after shift to 44°C at 210 minutes (panel A) and 120 minutes (panel B). Growth of SK5270, SK5271, SK5273 with pBMK48 (leuU) induction with 1mM IPTG (data not shown) was also plotted and does not show any differences from uninduced growth.
Figure A.7. **W3110 ΔsecG leuU** mutant with complementation in the presence or absence of IPTG induction. Growth is shown at 30°C and after shift to 44°C at 240 minutes. When indicated, 1mM IPTG was added to induce pBMK48 (*leuU*).
REFERENCES


