Maturation of tRNA in *Haloferax volcanii*.

Dissertation

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Richard Neil Nist BA, MS
Graduate Program in Microbiology

The Ohio State University
2011

Dissertation Committee:
Professor Charles J. Daniels, Advisor
Associate Professor Juan Alfonzo
Professor Venkat Gopalan
Abstract

Transfer RNA maturation in Archaea is a multi-faceted process requiring the activities of numerous enzymes in order to generate functional tRNAs from primary transcripts. Upon transcription, ribonuclease P cleaves the 5’ leader, tRNase Z cleaves the 3’ trailer, RNA-dependent and RNA-independent enzymes introduce modifications to bases and sugars, 3’ maturation is completed by the CCA-adding enzyme, and introns are removed. We examined the modification and intron removal stages of maturation for the *Halofex volcanii* intron-containing pre-tRNA.

*H. volcanii* tRNAMet possesses a 2’-O-methyl modification at position C34. A similar modification was characterized in pre-tRNATrp as being guided by a box C/D sRNP complex. Furthermore, just as sequence analysis and *in vitro* experiments have shown the intron-containing pre-tRNATrp to be a substrate for box C/D-guided modification, sequence analysis of intron-containing pre-tRNAMet suggests it is a substrate for box C/D-guided modification. We have determined not only that the C34 modification of tRNAMet is guided by a box C/D sRNP, but also that the intron-containing pre-tRNAMet appears to be a modification substrate *in vivo*.

By the action of intron endonuclease, tRNA introns are removed in Archaea. This is a requisite step for the maturation of tRNATrp, tRNAMet, and tRNAGlnUUG in *H. volcanii*. Sequence analysis and previous mapping experiments suggest that there are additional substrates for this enzyme, including the 16S rRNA and 23S rRNA flanking
sequences. Our data indicate that the 16S rRNA and 23S rRNA flanking sequences are cleaved by the intron endonuclease meaning there are substrates beyond the three tRNAs.

Ribosomal protein L7ae has been at the forefront or background of each study conducted in this research. This K-turn stabilizing protein is involved in several facets of tRNA maturation including as a box C/D sRNP complex component required for 2’-O-methylation of tRNA_{Trp} and tRNA_{Met}, as a component of archaeal RNase P, and as a box C/D sRNP complex component for the putative guide sRNA formed by the ligation of the 23S rRNA flanking sequences. Given the involvement of L7ae in these and other sRNP complexes, we attempted to capture L7ae-containing complexes from *H. volcanii*. We were able to demonstrate L7ae co-purification with each known and predicted class of L7ae-binding RNAs.
Dedication

To those who provided support/assistance for finishing the job, you know who you are.
Acknowledgements

This amount of work could never have been accomplished without considerable assistance from others. First, I would like to thank my advisor Dr. Chuck Daniels for allowing me to pick up and move this interesting project along. The back-and-forth discussions of data and suggestions of how to proceed next were greatly beneficial providing both insights I had not thought of and aiding me in seeing all sides of the problem. Also, the atmosphere of encouragement and continued interest in the science helped reinforce my own views.

Gratitude also goes out to my advisory and thesis committee, Dr. Juan Alfonzo, Dr. Kurt Fredrick, Dr. Joe Krzycki, and Dr. Venkat Gopalan for their willingness to share their insights into the research problems and lab protocols for similar procedures. The additional perspectives provided by them were significantly helpful in seeing likely explanations and problems that I otherwise would have missed.

The science may be the focus, but the people really make the lab. That said, thanks to my long-time lab-mates Dr. Anice Sabag-Daigle and Shoko Morimoto; Anice for being a willing sounding board for great discussions about our work and firming up protocols and Shoko for sharing protocols and keeping those buffer mixes handy that I always seemed to need. I would also be remiss if I didn’t thank them both (as well as past lab members Brian Gabel and Kara Batte) for putting up with me throughout these years, and making this a good working environment. I also thank past members of the lab whose contributions led to the project I was given.
I am appreciative of the assistance of our collaborators on the tRNAMet sRNA project provided by Dr. Jens Hemmingsen and his group at Capital University (Columbus, Ohio). In particular for the development of the RNase H assay which finally allowed us to “close the book” on the tRNAMet sRNA project. I am also appreciative of the understanding from the TA supervisors, allowing me to balance out teaching and research, and the friendly assistance of the other graduate students in the program at OSU, there are too many of you to name, but you have my thanks for our work and non-work related discussions.

Lastly, but hardly least, I wish to express my gratitude to my family, in particular to my parents Tim and Judy for their strong, constant support and interest throughout these last several years. Also, thanks to my brothers Jeff and Tom and sister Denise for their interest and attentive ears along the way, and to my grandparents Neil and Eris for their faith in me and support.
Vita

April 19, 1982 ................................................................. Born – Canton, Ohio

2004 ........................................................................ Bachelors of Art
                                                     Capital University
                                                     Columbus, Ohio

2007 ........................................................................ Masters of Science
                                                     The Ohio State University
                                                     Columbus, Ohio

2007-present ......................................................... Graduate Teaching and Research Associate
                                                     The Ohio State University
                                                     Columbus, Ohio

Field of Study

Major Field: Microbiology
# Table of Contents

Abstract .......................................................................................................................... ii
Dedication ......................................................................................................................... iv
Acknowledgements ........................................................................................................... v
Vita ....................................................................................................................................... vii
List of Tables ....................................................................................................................... xii
List of Figures ...................................................................................................................... xiii
List of Abbreviations ........................................................................................................... xvi

## Chapter 1: Archaea and Stable RNAs
- Domain Archaea .............................................................................................................. 1
- Stable RNAs ..................................................................................................................... 5
- Research Focus ................................................................................................................. 14

## Chapter 2: Materials and Methods
- Materials and Enzymes .................................................................................................. 16
  - Culturing of Bacteria and Archaea ................................................................................. 17
    - Culturing of *Escherichia coli* ...................................................................................... 17
    - Culturing of *Haloferax volcanii* ............................................................................... 21
    - Preparation of freezer stocks ...................................................................................... 21
    - Culturing of *H. volcanii* strains in competition ....................................................... 21
- Nucleic acid preparations ............................................................................................... 22
  - RNA isolation ................................................................................................................ 22
  - *H. volcanii* genomic DNA isolation ......................................................................... 24
  - Plasmid mini-preps ....................................................................................................... 25
  - Quantitation of DNA/RNA ........................................................................................... 26
Electrophoretic techniques for nucleic acids separation and purification ........................................26
  Denaturing polyacrylamide gel electrophoresis of DNA and RNA ........................................26
  Agarose gel electrophoresis of RNA .........................................................................................27
  Non-denaturing agarose gel electrophoresis of DNA .................................................................28
  Gel staining with ethidium bromide ..........................................................................................28
  Recovery of nucleic acids from agarose ....................................................................................29
  Recovery of nucleic acids from polyacrylamide ........................................................................29
  UV-shadowing ..........................................................................................................................30
Preparation of competent cells and transformation ........................................................................30
  E. coli competent cells preparations ..........................................................................................30
  Transformation of E. coli ............................................................................................................31
  Competent H. volcanii cell preparations ....................................................................................32
  Transformation of H. volcanii .....................................................................................................32
Recombinant DNA construction .....................................................................................................33
  Polymerase Chain Reaction .........................................................................................................33
  TA cloning ...................................................................................................................................34
  Site-directed mutagenesis ...........................................................................................................34
  Restriction enzyme digests .........................................................................................................35
  Ligation of DNA fragments ..........................................................................................................36
  Sequencing of plasmids ................................................................................................................36
Construction of strains and vectors ................................................................................................36
  Construction of the tRNA\textsubscript{Met} sRNA deletion strain ..................................................37
  Construction of the tRNA\textsubscript{Met} sRNA complementation strain ........................................37
  Construction of the inducible hexahistidine-tagged L7\textsubscript{ae} strain ........................................43
  Construction of the intron endonuclease triple mutant strain .....................................................44
Techniques for analysis of RNA .......................................................................................................45
  Radiolabeling of DNA probes .......................................................................................................45
  Northern analysis .........................................................................................................................45
  Quantitation of Northernns ..........................................................................................................47
  Stripping of Zeta-Probe\textsuperscript{TM} membranes ....................................................................47
  \textsuperscript{32}pCp synthesis and labeling ....................................................................................47
  Reverse transcription PCR ...........................................................................................................48
  \textit{In vitro} transcription .............................................................................................................49
  RNase H detection of 2’-O-methyl groups ..................................................................................50
Protein analysis and purification .....................................................................................................54
  Protein induction in \textit{H. volcanii} .................................................................................................54
  Preparation of whole-cell extracts from \textit{H. volcanii} ...............................................................54
  SDS-PAGE ....................................................................................................................................55
  Tricine-PAGE ..............................................................................................................................56
Western analysis.................................................................57
Dot-blotting.................................................................58
Protein A enrichment of antibodies........................................59
Affinity purification of hexahistidine-tagged proteins.................59

Chapter 3: Determining that modification at position C34 of H. volcanii tRNAMet elongator in vivo occurs via the box C/D mechanism
Introduction .....................................................................................61
Box C/D guided modification of archaeal tRNA .................................61
Functional significance of 2’-O-methyl groups .................................68
Results ...............................................................................................69
Identification and conservation of the tRNAMet sRNA gene ..........69
Detection of the transcript .................................................................75
Deletion of the tRNAMet sRNA gene from the chromosome .............82
Growth of the deletion strain ............................................................82
Complementation of the tRNAMet sRNA gene ...............................87
Impact of deleting the tRNAMet sRNA gene on 2’-O-methylation at position C34 of tRNAMet .........................................................87
Impact of tRNAMet sRNA gene deletion on the tRNAMet pool .........94
Discussion ..........................................................................................95
tRNAMet sRNA is not essential for viability or efficient growth .........95
Complementation of tRNAMet sRNA yields a new product ..............96
tRNAMet sRNA guides the C34 modification of H. volcanii tRNAMet in vivo ..............................................................................97
Modification at position C34 occurs in pre-tRNAMet, but is not required for efficient splicing in vivo .................................................102

Chapter 4: Capture and analysis of L7ae-containing complexes in H. volcanii
Introduction .....................................................................................103
Results ..............................................................................................118
Identification of the T2743 box H/ACA sRNA .................................118
Construction and verification of the affinity-tagged wild type L7ae overexpression strain DS70-ASD41-L6H .................................118
Discussion ..........................................................................................131
Discovery and capture of the box H/ACA sRNA T2743 .................131
Overexpression and purification of affinity-tagged L7ae-containing complexes ..............................................................................132
Chapter 5: Establishing the role of archaeal intron endonuclease in processing of the ribosomal RNA operon in vivo

Introduction .......................................................................................................................... 136
Results .................................................................................................................................. 144
Conservation of the ribosomal RNA operon including the Bulge-Helix-Bulge motif among Haloarchaea ........................................................................................................... 144
Construction of the intron endonuclease mutant strain ................................................. 149
Growth of the DS70-ASD41-endA* strain ....................................................................... 149
Induction of endA* ............................................................................................................. 152
Verification of intron endonuclease inhibition ................................................................. 157
Time-course accumulation of pre-tRNA ........................................................................ 160
Analysis of the impact of endA* expression on total RNA isolated from the mutant strain ........................................................................................................................................ 160
Impact of intron endonuclease inhibition on rRNA processing .................................... 163
Impact of intron endonuclease inhibition on 16S rRNA processing ............................ 163
Impact of intron endonuclease inhibition on tRNAAlaUGC processing .................... 173
Impact of intron endonuclease inhibition on 23S rRNA processing ............................ 178
Investigation of 16S rRNA and 23S rRNA flanking sequence ligation in H. volcanii ................................................................................................................................. 178
Discussion ............................................................................................................................ 188
Expression of endA* has a dominant negative impact on growth .............................. 188
Inactive monomer expression is sufficient to significantly inhibit Bulge-Helix-Bulge cleavage ......................................................................................................................... 190
There is an observable impact of endA* expression on total RNA ......................... 191
Intron endonuclease effects the release of the 16S rRNA from the primary transcript ................................................................................................................................. 192
Intron endonuclease acts prior to RNase P excision of tRNAAlaUGC ..................... 194
Intron endonuclease is responsible for excision of the 23S rRNA from the primary transcript ......................................................................................................................... 195
5S rRNA is unimpacted by intron endonuclease ......................................................... 197
Suggested model for rRNA primary transcript processing .................................... 197
Flanking sequences of the 23S rRNA released by intron endonuclease are ligated and may fulfill a box C/D role in H. volcanii ................................................................. 198

Perspectives .......................................................................................................................... 205
List of References .............................................................................................................. 207
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 List of halophiles whose genomes have been sequenced and published.</td>
<td>4</td>
</tr>
<tr>
<td>1.2 Known classes of non-coding RNAs and the domains in which they have been discovered so far</td>
<td>8</td>
</tr>
<tr>
<td>2.1 List of oligonucleotides used in the experiments presented in Chapter 3</td>
<td>18</td>
</tr>
<tr>
<td>2.2 List of oligonucleotides used in the experiments presented in Chapter 4</td>
<td>19</td>
</tr>
<tr>
<td>2.3 List of oligonucleotides used in the experiments presented in Chapter 5</td>
<td>20</td>
</tr>
<tr>
<td>2.4 List of <em>H. volcanii</em> strains and relevant features used for the experiments in this work and the strains generated by this work</td>
<td>38</td>
</tr>
<tr>
<td>3.1 Relative cleavage of intron-containing pre-tRNAMet and mature tRNAMet derived from the RNase H assay shown in Figure 3.10</td>
<td>93</td>
</tr>
</tbody>
</table>
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Tree of Life based on the analysis of Woese and colleagues depicting the approximate divisions in evolution among the three Domains and the two major divisions between Archaea</td>
<td>7</td>
</tr>
<tr>
<td>2.1 Relevant features of the plasmids used to generate the strains from this study described in Table 2.4</td>
<td>40</td>
</tr>
<tr>
<td>2.2 Schematic representation of the pop-in/pop-out genetics used to construct the tRNAMet sRNA deleted-strain</td>
<td>42</td>
</tr>
<tr>
<td>2.3 Schematic representation of the RNase H method for assessing whether or not a particular nucleotide is 2’-O-methylated</td>
<td>52</td>
</tr>
<tr>
<td>3.1 Representation of an archaeal canonical box C/D sRNA and its interaction with an RNA target</td>
<td>64</td>
</tr>
<tr>
<td>3.2 <em>H. volcanii</em> tRNAMet sRNA (DNA sequence) was aligned against BLASTN matches from the other halophilic Archaea using ClustalW</td>
<td>72</td>
</tr>
<tr>
<td>3.3 Alignment of haloarchaeal tRNAMet DNA sequences</td>
<td>74</td>
</tr>
<tr>
<td>3.4 Pairing of the box C/D sRNA with <em>H. volcanii</em> tRNAMet RNA</td>
<td>77</td>
</tr>
<tr>
<td>3.5 Gene organization of putative tRNAMet sRNAs in haloarchaeal genomes</td>
<td>79</td>
</tr>
<tr>
<td>3.6 Representation of the tRNAMet sRNA location, sequence, and evidence of expression</td>
<td>81</td>
</tr>
<tr>
<td>3.7 Sequence of the tRNAMet sRNA recovery module and expression of <em>H. volcanii</em> tRNAMet sRNA in wild type and mutant strains</td>
<td>84</td>
</tr>
</tbody>
</table>
3.8 Evidence of tRNAMet sRNA sequence deletion from the genome ..................86

3.9 PCR analysis of the competition between the wild type DS70-ASD40 (wt) and deletion DS70-ASD40-sno (Δ) strains.........................................................89

3.10 tRNAMet sRNA guides modification of both the intron-containing pre-tRNAMet (pre-met) and mature tRNAMet (Met) RNAs ..................................................92

3.11 Schematic model of tRNAMet modification and splicing..........................101

4.1 Predicted L7ae sRNP complexes in H. volcanii ......................................105

4.2 Potential box H/ACA RNA in H. volcanii .............................................110

4.3 Structure of RNase P RNA (RPR) from P. horikoshii and H. volcanii with probable L7ae binding sites indicated..................................................114

4.4 Conserved sequence elements of haloarchaeal 7S RNA and the potential binding sites for L7ae .................................................................116

4.5 Conservation of the second hairpin of the H. volcanii RNA originating from intergenic region T2743 .................................................................120

4.6 Expression of hexahistidine-tagged H. volcanii L7ae ................................123

4.7 Comparison of protein purification profiles of DS70-ASD41-L6H induced and uninduced cultures .................................................................126

4.8 RNAs associated with in vivo expressed affinity-tagged H. volcanii L7ae protein...128

4.9 Northern analysis of purified, tagged L7ae fractions for RNAs known and proposed to bind L7ae........................................................................130

5.1 Potential EndA substrates .......................................................................140

5.2 Conservation of the BHB elements surrounding the 16S rRNA in the haloarchaea .146

5.3 Conservation of the BHB elements surrounding the 23S rRNA in the haloarchaea .148

5.4 Mutation scheme for catalytically inactivating H. volcanii intron endonuclease......151

5.5 Growth of the DS70-ASD41-endA* strain.............................................154
5.6 Inducible expression of the endA* gene following urocanate addition ............... 156

5.7 Validation of intron endonuclease inhibition as measured by accumulation of intron-containing pre-tRNAs 4 hours post-induction of EndA* .................................................. 159

5.8 Time-course analysis of the accumulation of intron-containing pre-tRNA in DS70-ASD41-endA* under normal and under inhibitory conditions ......................... 162

5.9 Effect of EndA* on the population of small RNAs (<500 nt) .............................. 165

5.10 Schematic representation of the H. volcanii rRNA operon and positions of probes used in Northern analysis ................................................................. 167

5.11 Northern analysis of rRNAs from DS70-EndA* using the probes indicated beneath each panel .................................................................................. 169

5.12 Effect of EndA* on the processing of rRNA transcripts in the 5’ region of the 16S rRNA ........................................................................................................ 172

5.13 Effect of EndA* on the processing of rRNA transcripts in the 3’ region of the 16S rRNA ........................................................................................................ 175

5.14 Effect of EndA* on the processing of the 16S rRNA–tRNA region of the rRNA transcript .......................................................................................... 177

5.15 Effect of EndA* on the processing of rRNA transcripts in the 5’ region of the 23S rRNA ........................................................................................................ 180

5.16 Effect of EndA* on the processing of rRNA transcripts in the 3’ region of the 23S rRNA ........................................................................................................ 182

5.17 Sequence and location of the 23S flanking sequence ligated product ............ 184

5.18 Effect of EndA* on the 23S rRNA flanking sequence ligation product ........ 187

5.19 Schematic representation of H. volcanii16S rRNA and tRNAAla maturation .... 200

5.20 Schematic representation of H. volcanii16S rRNA and tRNAAla maturation .... 202
List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-FOA</td>
<td>5-flouroorotic acid</td>
</tr>
<tr>
<td>α</td>
<td>alpha</td>
</tr>
<tr>
<td>A</td>
<td>absorbance at the subscripted wavelength in nm</td>
</tr>
<tr>
<td>APS</td>
<td>ammonium persulfate</td>
</tr>
<tr>
<td>β</td>
<td>beta</td>
</tr>
<tr>
<td>BME</td>
<td>β-mercaptoethanol</td>
</tr>
<tr>
<td>BPB</td>
<td>bromophenol blue</td>
</tr>
<tr>
<td>°C</td>
<td>degrees Celsius</td>
</tr>
<tr>
<td>ChBD</td>
<td>chitin-binding domain</td>
</tr>
<tr>
<td>CDM</td>
<td>completely defined medium</td>
</tr>
<tr>
<td>CX</td>
<td>complex medium</td>
</tr>
<tr>
<td>dd</td>
<td>double distilled</td>
</tr>
<tr>
<td>Δ</td>
<td>delta</td>
</tr>
<tr>
<td>DMF</td>
<td>N,N-dimethylformamide</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethylsulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxynucleic acid</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>EtOH</td>
<td>ethanol</td>
</tr>
<tr>
<td>γ</td>
<td>gamma</td>
</tr>
<tr>
<td>g</td>
<td>relative centrifugal force</td>
</tr>
<tr>
<td>HGT</td>
<td>horizontal gene transfer</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl-β-D-thiogalactoside</td>
</tr>
<tr>
<td>kDa</td>
<td>kilodaltons</td>
</tr>
<tr>
<td>L</td>
<td>liter(s)</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
</tr>
<tr>
<td>m</td>
<td>milli</td>
</tr>
<tr>
<td>μ</td>
<td>micro</td>
</tr>
<tr>
<td>M</td>
<td>moles per liter</td>
</tr>
<tr>
<td>MCS</td>
<td>multiple cloning site</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>NaOH</td>
<td>sodium hydroxide</td>
</tr>
<tr>
<td>NEB</td>
<td>New England Biolabs</td>
</tr>
<tr>
<td>nm</td>
<td>nanometers</td>
</tr>
<tr>
<td>nt</td>
<td>nucleotide(s)</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PMGF</td>
<td>Plant Microbe Genetics Facility</td>
</tr>
<tr>
<td>PSI</td>
<td>pounds per square inch</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>rRNA</td>
<td>ribosomal RNA</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>sRNP</td>
<td>small ribonucleoprotein</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris-acetate</td>
</tr>
<tr>
<td>TBS</td>
<td>Tris-buffered saline</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris/boric acid/EDTA</td>
</tr>
<tr>
<td>TEMED</td>
<td>N,N,N&lt;sup&gt;1&lt;/sup&gt;, N&lt;sup&gt;1&lt;/sup&gt;-tetramethylethylenediamine</td>
</tr>
<tr>
<td>tRNA</td>
<td>transfer RNA</td>
</tr>
<tr>
<td>XC</td>
<td>xylene cyanol</td>
</tr>
<tr>
<td>V</td>
<td>volts</td>
</tr>
<tr>
<td>U</td>
<td>units of enzyme</td>
</tr>
<tr>
<td>w/v</td>
<td>weight per volume</td>
</tr>
<tr>
<td>v/v</td>
<td>volume per volume</td>
</tr>
</tbody>
</table>
Chapter 1: Archaea and Stable RNAs

Domain Archaea

In 1990, the three-Domain classification was proposed by Woese, Kandler, and Wheelis, and the development of the three Domain system has been reviewed recently (Cavicchioli, 2011; Woese et al., 1990). This proposal recommended a reordering of the taxonomic system to coincide with mounting molecular evidence toward three distinct evolutionary lineages of organisms: Archaea, Bacteria, and Eukarya (Winker and Woese, 1991; Woese et al., 1990). Consideration of the Archaea as separate from Bacteria was not a novel concept, but to this point and for some years after, many researchers resisted revision of the taxonomic structure.

Prior to the Woese proposal, Archaea were considered a subset of Bacteria based on their general phenotypic characteristics. However, these pre-existing views did not account for the distinct characteristics of Archaea at the molecular level including cell wall composition, lipid content, metabolism, 16S rRNA modification, 5S rRNA structure, and tRNA structure and modification (Woese et al., 1978). An absence of peptidoglycan, a well-established feature of bacterial cell walls, was documented in archaeal cell walls, and unique types of cell walls were discovered among the Archaea (Brock et al., 1972; Brown and Cho, 1970; Kandler and Hippe, 1977; Reistad, 1972; Steber and Schleifer, 1975; Woese et al., 1978). As the attributes of archaeal cell walls were inconsistent with what was known about Bacteria, there was a question of the validity of classifying
archaeal representatives as bacterial members.

Metabolic markers were another point of contention. Archaea had been found to contain a different major class of ether-linked lipids (Kandler and Konig, 1978; Langworthy et al., 1974; Steber and Schleifer, 1975; Tornabene et al., 1978; Woese et al., 1978). In conjunction with their niche, “extreme” habitats of high temperature or hypersaline conditions, some preliminary designations for Archaea were established (Cavicchioli, 2011; Woese et al., 1978).

Additional indicators at the molecular level were observed in the 16S rRNA and 5S rRNA. While the 16S rRNAs of Archaea contained modifications in similar locations to their bacterial counterparts, the modifications themselves were different (Magrum et al., 1978). Also, the “prokaryote loop” of the 5S rRNA in Archaea was unique relative to both Bacteria and Eukarya (Woese et al., 1978). Interestingly, these features were minimal when compared to archaean tRNAs. The “common arm” sequence (GTψCG) was absent and the tRNAs possessed a large number of modifications (Fox et al., 1977; Magrum et al., 1978; Woese et al., 1978). Some of these modifications are unique to the Archaea, but the most intriguing observation was that the modification patterns were much more akin to those observed in Eukarya. Collectively these observations and others led to the eventual suggestion that Archaea be recognized as a classification unto themselves.

Although the argument was now made to consider classifying Archaea separately, there was still debate as to which organisms constituted the Archaea. At first, methanogens and some thermophiles were thought to comprise Archaea, but analysis of
*Halobacterium halobium* (now known as *Halobacterium salinarum*) 16S rRNA modifications revealed that some halophilic “bacteria” were far more closely related to methanogens than previously thought. Enough similarity was observed so as to warrant their consideration among Archaea (Magrum et al., 1978). Numerous subsequent studies have confirmed that these halophiles are a definite subset of Domain Archaea on the basis of the same features mentioned above.

Presently, 95 archaeal genomes have been published in the Kyoto Encyclopedia of Genes and Genomes (KEGG); among these are 13 halophilic Archaea (Table 1.1), including the recently published *Haloferax volcanii* DS2 genome (Anderson et al., 2009; Baliga et al., 2004; Bolhuis et al., 2006; Falb et al., 2005; Goo et al., 2004; Hartman et al., 2010; Malfatti et al., 2009; Ng et al., 2000; Pfeiffer et al., 2008; Roh et al., 2010; Saunders et al., 2010; Tindall et al., 2009). At least two additional halophile genomes are currently in the assembly phase. Archaea are now an established third Domain of life due to the overwhelming molecular evidence segregating them from Bacteria and Eukarya (Cavicchioli, 2011; Winker and Woese, 1991; Woese, 2004). Studies have shown that they possess molecular machinery that had previously been thought exclusive to Eukarya. Among such machinery is that used to mature tRNAs, including sRNA-ribonucleoprotein complex (sRNPs) modification systems and intron removal and splicing enzymes; these will be discussed in detail later. Despite such characteristics, they retain physical properties that are distinctly bacterial such as the absence of nuclei and organelles. Therefore, Archaea remain a class of organisms that appear to be in-between the other
<table>
<thead>
<tr>
<th>Strain</th>
<th>Strain Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Halalkalicoccus jeotgali</em> B3</td>
<td>hje</td>
</tr>
<tr>
<td><em>Haloarcula marismortui</em> ATCC 43049</td>
<td>hma</td>
</tr>
<tr>
<td><em>Halobacterium salinarum</em> strains R1 and NRC-1</td>
<td>hal</td>
</tr>
<tr>
<td><em>Haloferax volcanii</em> DS2</td>
<td>hvo</td>
</tr>
<tr>
<td><em>Halogeometricum borinquense</em> DSM 11551</td>
<td>hbo</td>
</tr>
<tr>
<td><em>Halomicrobium mukohataei</em> DSM 12286</td>
<td>hmu</td>
</tr>
<tr>
<td><em>Haloquadratum walsbyi</em> DSM 16790</td>
<td>hwa</td>
</tr>
<tr>
<td><em>Halorhabdus utahensis</em> DSM 12940</td>
<td>hut</td>
</tr>
<tr>
<td><em>Halorubrum lacusprofundi</em> ATCC 49239</td>
<td>hla</td>
</tr>
<tr>
<td><em>Haloterrigena turkmenica</em> DSM 5511</td>
<td>htu</td>
</tr>
<tr>
<td><em>Natralba magadii</em> ATCC 43099</td>
<td>nma</td>
</tr>
<tr>
<td><em>Natronomonas pharaonis</em> DSM 2160</td>
<td>nph</td>
</tr>
<tr>
<td><strong>a</strong>Haladaptatus paucihalophilus** DX253</td>
<td>hpa</td>
</tr>
</tbody>
</table>

Table 1.1 List of halophiles whose genomes have been sequenced and published. Strain abbreviations used throughout the figures in this document are indicated in parentheses following the full name of the organism.

**a** This organism’s genome is not complete and is still in the assembly phase at the time of this writing.
two domains, but with their own unique characteristics that demand a separate taxonomic category (Figure 1.1).

The model organism, *Haloferax volcanii*, used in the following studies is classified among the *Euryarchaeota* in Domain Archaea. *H. volcanii* is moderately halophilic and can be cultured easily under aerobic conditions in either defined or complex media (Hartman et al., 2010). The organism was used in the elucidation of archaeal tRNA introns (Daniels et al., 1985; Thompson and Daniels, 1988). This organism also possesses the components necessary for sRNP complex-guided modification of tRNA and rRNA making it a suitable choice to study these processes in an archaeal system.

**Stable RNAs**

Our understanding of the diverse roles played by RNA in cellular processes has been greatly expanded in recent years with the discovery of several classes of novel non-coding RNAs noted in Table 1.2; many of these classes of RNAs are found in Archaea. The tmRNAs and microRNAs, regulatory RNAs for releasing stalled ribosomes and silencing transcripts, respectively, have been reviewed in the following references and will not be discussed in detail here as they have not been discovered in Archaea (Djuranovic et al., 2011; Wower et al., 2008). Among the stable RNAs Archaea have been shown to possess or have been implicated in possessing are rRNA, tRNA, riboswitches, ribonuclease P (RNase P), the 7S RNA component of Signal Recognition Particle (SRP), Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR),
Figure 1.1 Tree of Life based on the analysis of Woese and colleagues depicting the approximate divisions in evolution among the three Domains and the two major divisions within Archaea. The relative position of the model organism used in these studies, *Haloferax volcanii*, is indicated.
Figure 1.1 Tree of Life based on the analysis of Woese and colleagues depicting the approximate divisions in evolution among the three Domains and the two major divisions within Archaea.
Table 1.2 Known classes of non-coding RNAs and the domains in which they have been discovered so far. Domains are abbreviated by the first letter of the Domain (B: Bacteria, E: Eukarya, A: Archaea).

<table>
<thead>
<tr>
<th>Type of RNA</th>
<th>Function</th>
<th>Domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Riboswitch, RNAi</td>
<td>Regulation</td>
<td>B,E,(A?)</td>
</tr>
<tr>
<td>Spliceosomal (SNRP)s</td>
<td>Pre-mRNA splicing</td>
<td>E</td>
</tr>
<tr>
<td>Small Nucleolar RNAs (Box H/ACA and Box C/D)</td>
<td>RNA editing, RNA modification</td>
<td>E,A</td>
</tr>
<tr>
<td>tRNA</td>
<td>Translation</td>
<td>B,E,A</td>
</tr>
<tr>
<td>tRF</td>
<td>Regulation?</td>
<td>B,(A?)</td>
</tr>
<tr>
<td>Ribozymes (rRNA, RNase P)</td>
<td>Translation, tRNA 5’ maturation</td>
<td>B,E,A</td>
</tr>
<tr>
<td>SRP (7S RNA)</td>
<td>Protein Translocation</td>
<td>B,E,A</td>
</tr>
<tr>
<td>CRISPR</td>
<td>Phage Defense</td>
<td>B,A</td>
</tr>
</tbody>
</table>
tRNA-derived fragments (tRFs), small nucleolar (sno) RNAs, and unclassified regulatory RNAs apparently involved in adaptation to environmental conditions.

Substantial advances have occurred over the last decade in the understanding of ribosomal RNAs and transfer RNAs including the elucidation of the three-dimensional structure of the ribosome complexes from each Domain. There have also been improvements in the understanding of how these RNAs are modified and what function these modifications serve. Introduction of pseudouridines or 2’-O-methyl groups can be accomplished either by dedicated enzymes or by RNA guided complexes, and they affect the chemical properties of the RNA as well as improve stability of the molecule (Agris, 1996; Davis, 1995; Huang et al., 2011). Modifications to rRNA have been described in a recent review as having regulatory roles in the assembly of the large and small ribosomal subunits and in ribosome function given their proximity to the A, P, and E sites (Karijolich et al., 2010).

The function of modifications to tRNAs are at this point less understood, though no less prevalent. In fact, archaeal tRNAs are heavily modified relative to their bacterial counterparts; however, understanding the function of these modifications has lagged behind that of rRNA modifications. These modifications are now being demonstrated as important to survival in competitive environments and studies are underway to assign the actual functions of these modifications. The impact of 2’-O-methyl modifications and pseudouridinie modifications are discussed in greater detail in the following chapters: Chapter 3 for 2’-O-methylation, Chapter 4 for pseudouridine.
Riboswitches are regulatory leader sequences of genes that modulate expression of genes either through active permission of transcription or programmed transcription termination. Since their discovery, numerous examples have been identified in Bacteria. Some riboswitches are controlled by tRNA while others have been determined to be under control of specific metabolites (Smith et al., 2010; Zhang et al., 2010). Given the nature of coupled transcription and translation in Bacteria and Archaea, considerable speculation has suggested that riboswitches are present in Archaea, but no archaeal riboswitches have been demonstrated yet. Still, extensive structural motif searches have identified a riboswitch-like folding motif, \textit{crcB}, that is prevalent in Bacteria and Archaea (Weinberg et al., 2010). The folding of this sequence would allow a metabolite to control expression of the DNA damage stress control genes in these organisms. More work is necessary to determine if this RNA functions as a riboswitch in Archaea and to establish the controlling metabolite. Evidence that this is indeed the case may assist in the discovery of similar and new archaeal riboswitches.

Another well-characterized stable RNA is the RNA component of Ribonuclease P. Bacterial RNase P enzymes consist of the catalytic RNA and a single protein, this is in stark contrast to the eukaryal RNase P, which consists of the catalytic RNA and nine proteins (Jarrous and Gopalan, 2010). Archaea also possess an RNA-plus protein RNase P complex and the RNA component of RNase P has been demonstrated as the catalytic component of the complex in all three Domains of life (Guerrier-Takada et al., 1983; Kikovska et al., 2007; Pannucci et al., 1999). The archaeal ribonuclease P has four proteins: Pop5, Rpp21, Rpp29, Rpp30 (Hall and Brown, 2002). Recently, the
holoenzyme has been shown to have an additional protein, ribosomal protein L7ae (Cho et al., 2010; Fukuhara et al., 2006). This protein has been preliminarily mapped to bind in the P12 and P15-17 regions of the RNA component and it confers enhanced catalysis as well as an elevated maximum working temperature for RNase P.

The 7S RNA of the SRP is a stable RNA component of a multi-protein machine. The SRP is responsible for chaperoning proteins with specific signal tags from the ribosome to the SRP receptor FtsY by guiding the translating ribosome to the membrane. Bacterial SRP consists of a comparatively short RNA component with a single protein, SRP54 (Calo and Eichler, 2011). The archaeal SRP has two established protein components: SRP54 and SRP19; the eukaryal SRP has homologues to these proteins and has two additional proteins for which no archaeal homologue has been identified (Calo and Eichler, 2011; Zwieb and Bhuiyan, 2010). The archaeal complex may also contain the ribosomal protein L7ae found in box H/ACA and box C/D complexes. In a study attempting to isolate box C/D sRNAs through immuno-affinity chromatography targeted against L7ae, Nop5, and fibrillarin, it was noted that 7S RNA fragments co-purified. Binding studies and a modified toe-printing assay were used to determine the binding site of L7ae to 7S RNA (Zago et al., 2005). The results suggest that L7ae may be a bona fide component of the archaeal SRP.

CRISPR systems in bacterial and archaeal cells have been the subject of great interest in recent years. These elements consist of a series of CRISPR-associated (cas) proteins adjacent to long sequences of regularly spaced unique sequence elements that have been identified as being highly similar to known phage nucleic acid sequences.
Each of these sequences of 21-37 nt is separated from the next sequence by a similarly-sized repeat sequence (Jansen et al., 2002). CRISPR elements have been postulated to function as a defense mechanism against phages (Haft et al., 2005; Mojica et al., 2005). Work has shown that incorporation of phage DNA sequences into the element subsequently rendered the strain immune to infection by the phage source of the sequence (Barrangou et al., 2007). Interestingly, preliminary studies suggest that the CRISPR system in archaeal cells functions at the level of RNA degradation whereas the bacterial enzymes appear to target phage DNA (Hale et al., 2008; Marraffini and Sontheimer, 2008). Greater detail can be found in a recent review (Horvath and Barrangou, 2010).

A novel class of apparently regulatory RNAs, tRFs, has been identified in humans. These RNAs consist of specific, stable degradation products of tRNAs which serve largely undetermined functions as of yet. They are second in abundance only to micro RNAs, implying a significant role, and at least one example of a tRF RNA has been shown as necessary for cellular proliferation though the mechanism is unclear (Lee et al., 2009). Attempts to catalogue the sRNA population of Archaea has given hints that tRF RNAs may also develop and perform similar functions in Archaea. At present, tRF-like RNAs have been identified, but there is no evidence of their functional significance (Fischer et al., 2011).

Originally identified in Eukarya, the well-studied group of stable RNAs known as small nucleolar (sno) RNAs are responsible for guiding the modification of specific nucleotides in a target RNA (Ganot et al., 1997a; Ganot et al., 1997b; Kiss-Laszlo et al., 1996; Nicoloso et al., 1996). Archaeal homologues to these sno RNAs (referred to as
sRNAs in this dissertation) were discovered later (Omer et al., 2000). The two classes of sno RNAs known as box C/D and box H/ACA, function as part of an sRNP complex to guide 2’-O-methylation or pseudouridylation, respectively, of a target in rRNA (Archaea and Eukarya) or tRNA (Archaea only). The targets of these guide RNAs are in the rRNA of eukaryotes, and many sno RNAs in yeast have been identified and knocked-out in an extensive analysis of the essentiality and function of the modifications (Esguerra et al., 2008; King et al., 2003; Lowe and Eddy, 1999). These single-guide RNA knock-out strains proved viable and few displayed any growth deficiencies unless challenged by environmental stress, whereas the protein components proved to be essential (Tollervey et al., 1991). Deficiencies also became evident when multiple sno RNAs were inactivated in the same strain. Microarray and high throughput sequencing technologies have greatly improved the detection and identification of both types of sno RNAs within the Archaea, though discovering targets for them at times remains a challenge due to the small number of basepairs needed between the guide and target RNA. Improved detection of these RNAs for which targets have been identified has allowed for studies addressing the significance and necessity of some of these modifications in vivo. The composition and function of these complexes will be described in detail later on (box C/D: Chapter 3, box H/ACA: Chapter 4).

An interesting development is the emergence of unclassified regulatory RNAs. Extensive RNA cloning has led to cDNA libraries containing numerous apparently stable RNAs in Archaea. Knock-out analysis on at least two of these archaeal RNAs has resulted in sensitivity of the organism to temperature or salinity, implying that these
transcripts may account for a regulatory function under the respective conditions (Fischer et al., 2011; Straub et al., 2009). These studies have also suggested that there is an abundance of anti-sense transcripts present in the cell though no function has been substantiated for this at present. As the deluge of new data continues to emerge on the known categories of small regulatory RNAs (sRNAs), it seems likely that novel classes of sRNAs remain to be uncovered.

**Research Focus**

The research presented in this dissertation is related to the process of tRNA maturation. Specifically, it focuses on box C/D sRNA-guided modification, complexes containing ribosomal protein L7ae, and intron endonuclease.

The questions addressed with respect to box C/D sRNA-guided modification were first whether or not the 2’-O-methyl modification of C34 in tRNAMet (elongator) was introduced by a box C/D guide and whether or not the guide RNA was essential. We also inquired whether or not the intron-containing pre-tRNAMet was modified at position C34 in vivo, whether or not there was a growth advantage to cells possessing this modification, and whether or not there this modification had an impact on intron-containing pre-tRNAMet splicing.

Research investigations regarding L7ae centered on whether an affinity-tagged copy of L7ae be expressed and partially purified, and if the tagged protein be used to capture sRNP complexes associated with this protein.
Intron endonuclease research questions addressed in the following chapters consisted of whether an inactive monomer subunit could be conditionally expressed, whether expression of the inactive monomer would hinder intron endonuclease activity in vivo, and is the substrate range of intron endonuclease more extensive than simply three intron-containing tRNAs.
Chapter 2: Materials and Methods

Materials and Enzymes

T4 polynucleotide kinase, Superscript III RNase H reverse transcriptase, Trizol™ RNA isolation reagent, restriction endonucleases, Random Primer labeling kit, buffer saturated-phenol, and cloned Taq polymerase were purchased from Invitrogen (Carlsbad, CA). Ribozol™ was purchased from Amresco (Solon, OH). T4 DNA ligase, T4 RNA ligase 1, and DpnI were obtained from New England BioLabs (Ipswich, MA). The Wizard SV Miniprep DNA purification kit, Riboprobe™ In vitro Transcription System, DNase RQI, and pGEM-T Easy TA cloning Vector System were purchased from Promega (Madison, WI). Luria-Bertani Broth (LB), Tris-HCl, 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal), isopropylthio-β-D-galactoside (IPTG), Triton X-100, ampicillin, and 100 mM PCR nucleotide mix were purchased from USB (Cleveland, OH). Yeast extract, tryptone, and casamino acids were purchased from DifCo (Lawrence, KS). Sodium chloride, sodium dodecylsulfate, magnesium sulfate, tricine, and magnesium chloride were obtained from Fisher (Pittsburgh, PA). Potassium chloride was obtained from JT Baker (Phillipsburg, NJ). The tetra-his antibody and Q-Tip 100 columns were purchased from Qiagen (Valencia, CA). Agarose, 40% 1:29 bis-acrylamide, TEMED, and ammonium persulfate were purchased from National
Diagnostics (Atlanta, GA). HiTrap Protein A affinity and HisTrap Chelating columns were obtained from Pharmacia Biotech, Inc. (Piscataway, NJ). Hybond-P membrane was purchased from Amersham Biosciences (Piscataway, NJ). Zeta-Probe™ membrane, Freeze-N-Squeeze™, Dual-Color protein standards, and Criterion Electrophoresis System were purchased from Bio-Rad (Hercules, CA). Immobilon™ Western HRP Substrate was purchased from Millipore (Billerica, MA). Urocanic acid, PEG-600, chloroform, isopropanol, methanol, coomassie brilliant blue G, acetone, thiamine, biotin, potassium sulfate, and calcium chloride were obtained from Sigma-Aldrich (St. Louis, MO). Urocanic acid was also purchased from Acros Organics (Geel, Belgium). 5-Flouroorotic acid monohydrate was purchased from Toronto Research Chemicals (North York, ON Canada). Uracil was obtained from Kodak (Rochester, NY). Cloned Pfu polymerase was purchased from Stratagene (LaJolla, CA). Radioisotopes [γ-32P]ATP (100 μCi/μl, 7,000 Ci/mmol) and [α-32P]dGTP (20 μCi/μl, 6,000 Ci/mmol) were obtained from MP BioMedicals (Santa Ana, CA). X-ray film was purchased from Denville Scientific (Metuchen, NJ). Oligonucleotides in Tables 2.1-2.3 were synthesized by Integrated DNA Technologies (IDT) (Coralville, IA).

**Culturing of Bacteria and Archaea**

**Culturing of Escherichia coli**

*E. coli* DH5α and JM110 were routinely cultured in LB broth and supplemented with ampicillin (100 μg/ml) to select for pUC-19-based plasmids or pWL-based shuttle vectors. For solid media, 15 g of agar was added to 1 L of LB broth.
Table 2.1 List of oligonucleotides used in the experiments presented in Chapter 3.

*aThis oligonucleotide was a chimera composed of 2’-O-methylated NTPs (indicated by
the lowercase “m” following the base designation) and dNTPs indicated in bold.*

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence 5’ to 3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>metsno</td>
<td>CACCGAGTGCTAGTTTCGGTTCCGGAGCGGTGCC</td>
</tr>
<tr>
<td>9metRH</td>
<td>UmAmTGAGUmGmGmGmGmGmGmAmAmUmUmUm</td>
</tr>
<tr>
<td>metE1</td>
<td>AGTGCAGCGCTATGGCTCAAGCTACCCACCCGG</td>
</tr>
<tr>
<td>metIn</td>
<td>GGAGGCTTCCAAGGGCTTCCGACCCGATTGTG</td>
</tr>
<tr>
<td>metE2</td>
<td>GGGGTGGGCTCCGAACCCACAGATCTCCGACAT</td>
</tr>
<tr>
<td>MSD-F</td>
<td>CCTCCGGTGCGAGCGCTTCGGAATTGTGT</td>
</tr>
<tr>
<td>MSD-5’Lk</td>
<td>GGAGATCTAGCTAGCTAGCTAGGCTGCTACGTCTTTCTCCAAC</td>
</tr>
<tr>
<td>MSD-3’Lk</td>
<td>GGCCCTAGCTAGCTAGCTAGCTAGCTAGATTGATTATGTATGCG</td>
</tr>
<tr>
<td>MSD-R</td>
<td>CGTCTTCAAGATGCGGAATTGGAGTAGAC</td>
</tr>
<tr>
<td>MSDX-F</td>
<td>GACGACGACGCAGCCAGACCCAGAAAG</td>
</tr>
<tr>
<td>MSDX-R</td>
<td>AAGCCCTACGACACTACTAAGAGC</td>
</tr>
<tr>
<td>snorec-F</td>
<td>CCTCTAGATCGTGGTGGCGGCGGCGGACCGA</td>
</tr>
<tr>
<td>snorec-R</td>
<td>CGAATTCAAAAAACGCGGCGAGCAATACATCGGTTCCCTGTACCTGCTTTCTCCAAG</td>
</tr>
<tr>
<td>T7premet-F</td>
<td>TAATACGACTCACTATAGGGGCGGCGGTAGTGTTAGGAC</td>
</tr>
<tr>
<td>T7premet-R</td>
<td>CCTGGTTCCGGGAGGTTGGCTCCGGAACCGATCTCCG</td>
</tr>
<tr>
<td>T7met-F</td>
<td>GTAATACGACTCACTATAGGGGCGGCGGTAGTGTTAGGACATAGCGGGCAGCTCATAA</td>
</tr>
<tr>
<td>T7met-R</td>
<td>CCTGGTTCCGGGAGGTTGGCTCCGGAACCGATCTCCGCAATAGGATCGGGCGCTATAA</td>
</tr>
<tr>
<td>Primer Name</td>
<td>Primer Sequence 5' to 3'</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>L6H-F</td>
<td>CCGGATCCATGGCTGTACGAGACTTCG</td>
</tr>
<tr>
<td>L6H-R</td>
<td>CCTTAGATCAGTGTTGGGTTGTGTTGCTGAAGGTCTCCGACCTTAGTC</td>
</tr>
<tr>
<td>hindhut-F</td>
<td>CCGAAGCATTTTACAGTACGTACTGC</td>
</tr>
<tr>
<td>t2743-0</td>
<td>CCGGGAACCCCTCGTCGCGGGCGG</td>
</tr>
<tr>
<td>t2743-1</td>
<td>CCGAAGCGGGAACACTCGGGCGG</td>
</tr>
<tr>
<td>t2743-2</td>
<td>GCGGGCGGTCGGCTTTCCGGGGGC</td>
</tr>
<tr>
<td>unkCD</td>
<td>CTATCATCAAGGGGTGAGTGGG</td>
</tr>
<tr>
<td>trpE1</td>
<td>GAGTCAGTCGCATGCGGCTGCCCACAG</td>
</tr>
<tr>
<td>7S-R</td>
<td>GGACAGCGAGGTGGCCCGGCGTTCCG</td>
</tr>
<tr>
<td>5S-R</td>
<td>GAATAGGGCGGCGGAACTGGAATTTCCA</td>
</tr>
<tr>
<td>RPR</td>
<td>GGGTTAGCTCCCTCCCTTTACGGG</td>
</tr>
</tbody>
</table>

Table 2.2 List of oligonucleotides used in the experiments presented in Chapter 4.
<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence 5' to 3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>endA-F</td>
<td>GGGGATCCCAAGGACGGCTGAGCCGTCG</td>
</tr>
<tr>
<td>endA-R</td>
<td>GGTCTAGATCATGGTGTGAGCCGACTGACCCGAA</td>
</tr>
<tr>
<td>YF-F</td>
<td>TCGGACTTCGGCTTTCACCCAGATCAGTCG</td>
</tr>
<tr>
<td>YF-R</td>
<td>CGACTCGAATCTGGTGAAGAGCCGGAATCCG</td>
</tr>
<tr>
<td>HA-F</td>
<td>GTGAGACAGCTCTCGGCTCGAGATTCGTCG</td>
</tr>
<tr>
<td>HA-R</td>
<td>CACGAGAATCTCGAGGAGGCCGAGAGTCGTCG</td>
</tr>
<tr>
<td>KE-F</td>
<td>GCGCGGCGTGTCGCGCGAGCAATGTTTTCG</td>
</tr>
<tr>
<td>KE-R</td>
<td>CGCAAAAAACATTCGCTCGCGGACACCAGC</td>
</tr>
<tr>
<td>endAUS</td>
<td>GAGCCGTCCTTGGGATCCGTCGTCG</td>
</tr>
<tr>
<td>trpln</td>
<td>GGTCGGACAACACGCAGGATGATCGAGTCG</td>
</tr>
<tr>
<td>trpe2</td>
<td>GGCCCGAGGGATTGAGACCCCGATCGATC</td>
</tr>
<tr>
<td>glne1</td>
<td>CGCTATGATCTGCCTAACCCAG</td>
</tr>
<tr>
<td>glnlIntron</td>
<td>TCCTGGCAACGAAATGAACGTACCA</td>
</tr>
<tr>
<td>glnSpId</td>
<td>CCCTGACGGCCATCAAAGGCGC</td>
</tr>
<tr>
<td>t-asp A</td>
<td>CGTAGATGGGCTACTAC</td>
</tr>
<tr>
<td>t-asp B</td>
<td>GGTGAATGGCACAGTAATG</td>
</tr>
<tr>
<td>t-asp C</td>
<td>GGAAGATCGGTGTGAGCCAC</td>
</tr>
<tr>
<td>t-asp D</td>
<td>GTGTTAGCCCAAGTAGT</td>
</tr>
<tr>
<td>t-asp E</td>
<td>ACCCACAAGGTTCGAAACCTGGGC</td>
</tr>
<tr>
<td>t-asp F</td>
<td>GGTCATCTGTCTGTACGAATC</td>
</tr>
<tr>
<td>t-asp G</td>
<td>CGTGACGGCGAGAACGGG</td>
</tr>
<tr>
<td>t-asp H</td>
<td>CCCGTTCCTCGGGTTCAGTG</td>
</tr>
<tr>
<td>t-asp I</td>
<td>CGGTCGGGCGGCTTACCA</td>
</tr>
<tr>
<td>t-asp A'</td>
<td>CCTAGTATTTGTCTCGGACG</td>
</tr>
<tr>
<td>t-asp F'</td>
<td>GTTCAGTGGACGGCTGATGGTAG</td>
</tr>
<tr>
<td>Anti-F</td>
<td>GGATACGACGAGATGACCC</td>
</tr>
<tr>
<td>16S-R</td>
<td>GAGGTTATCGTCTCGAGGGTAG</td>
</tr>
<tr>
<td>16SRT-F</td>
<td>CCTGGAGTACGGCGCAAGACGG</td>
</tr>
<tr>
<td>23S-R</td>
<td>GGAATCCCATCGGAGATCCTCGT</td>
</tr>
<tr>
<td>75-F</td>
<td>GACCCGACACGGGCGGCGTTCGAGTCA</td>
</tr>
</tbody>
</table>

Table 2.3 List of oligonucleotides used in the experiments presented in Chapter 5.
Culturing of *H. volcanii*

*H. volcanii* DS70 was routinely cultured aerobically in baffled flasks with foam stoppers in completely defined (CDM) or complex (CX) media at 42°C with shaking at 220 rpm (Wendoloski et al., 2001). CDM contained the following components (per liter): 125 g NaCl, 50 g MgCl₂•6H₂O, 5 g K₂SO₄, and 0.26 g CaCl₂•2H₂O. After autoclaving, the following components were added to cooled media (basal salts) to final concentrations of: 0.45% (w/v) Na-succinate, 0.05% (v/v) glycerol, 10 mM NH₄Cl, 1 mM K₂HPO₄, 8 μg/ml thiamine, and 1 μg/ml biotin (Mevarech and Werczberger, 1985). Additionally, 1 ml Balch's trace elements was added (Gerhardt, 1994). CX contained the following components (per liter): 125 g NaCl, 45 g MgCl₂•6H₂O, 10 g MgSO₄•7H₂O, 10 g KCl, 1.34 g CaCl₂•2H₂O, 3 g of yeast extract and 5 g of tryptone. *H. volcanii* cells were grown at 42°C on 1.5 % (w/v) agar plates containing CDM or CX media. Media were supplemented with Casamino acids (5 g/L), 5-FOA (150-250 μg/ml) or uracil (50 μg/ml) as indicated.

**Preparation of freezer stocks**

All freezer stocked strains were frozen at -80°C in appropriate growth medium with a final concentration of 15% (v/v) glycerol.

**Culturing of *H. volcanii* strains in competition**

A starter culture of each strain was grown at 42°C in CX medium overnight. The absorbance at 550 nm was measured the next morning, and an equivalent amount of cells from both cultures was added to 100 ml at approximately a 1:1,000 dilution in fresh CX medium and incubated at 42°C.
To maintain cell cultures in exponential growth throughout the experiment, a subculture was made each successive morning for 28 days. The culture’s absorbance at 550 nm ($A_{550}$) was measured to verify that it remained in exponential phase prior to making the approximately 1:1,000 (adjusted based on the $A_{550}$ reading) subculture in fresh medium. Additionally, a 1 ml sample of the culture was taken and centrifuged for 5 minutes at 8,000 $g$ at 4°C. Supernatant was discarded and the cell pellet was frozen at -20°C.

A second form of the competition experiment in which the culture was allowed to reach stationary phase, $A_{550} \geq 1.0$, prior to subculturing was also run. A total of eight cell samples were taken. Equivalent amounts of cells were added to 100 ml of CX medium and incubated at 42°C for 72 hours. A 1:1,000 dilution into fresh CX medium was made while a 1 ml sample of the culture was centrifuged for 5 minutes at 8,000 $g$, 4°C. The supernatant was discarded and the cell pellet was stored at -20°C.

Analysis of strain populations was conducted using end-point PCR as described in Polymerase Chain Reaction. The cell pellets were resuspended in 100 μl (exponential samples) or 1 ml (stationary samples) of ddH$_2$O and cells were lysed by repeated pipetting of the suspension. Approximately 1-2 μl of cell lysate was sufficient to provide template for PCR.

**Nucleic acid preparations**

**RNA isolation**

Typically, RNA was harvested from *H. volcanii* cells using a modified form of the Trizol protocol. The volume of culture used per preparation was dependent on the
growth phase of the culture; for cells in exponential growth, 45 ml was used and for cells
in stationary phase 20 ml was used. Cultures were transferred into 50 ml Falcon tubes
and centrifuged at 3,700 g for 7 minutes at 4°C using a Beckman JA 25.50 rotor. The
cell pellet was then resuspended using 500 μl of spheroplasting solution (see Competent
*H. volcanii* Cell Preparations) and transferred to a 1.5 ml microcentrifuge tube. To lyse
the cells and extract total RNA, 500 μl of Trizol/Ribozol was added to the suspension and
the suspension was homogenized with a vortex mixer then allowed to stand at room
temperature for 5 minutes. Next, 200 μl of chloroform was added and the suspension was
homogenized with a vortex mixer for 30 seconds. Phases were separated by incubating
the suspension 3 minutes at room temperature before centrifugation at 4°C for 10 minutes
at 14,000 g. The top aqueous layer was transferred to a clean 1.5 ml microcentrifuge tube
and combined with 1 volume of cold isopropanol (99%) to precipitate the nucleic acid.
Precipitation was at room temperature for 10 minutes followed by centrifugation at 4°C
for 10 minutes at 14,000 g. The tube was then inverted on the bench top for about 15
minutes to allow air-drying of nucleic acid pellets. Pellets were resuspended in 100-200
μl of ddH2O, depending on pellet size. The RNA prep was repeated and resulting RNA
susensions were stored at -80°C.

The following description of small RNA isolation has been adapted from a
protocol developed by Qiagen. To isolate populations of RNA approximately 500 nt and
smaller, RNA was first prepared as described above. Qiagen Q-Tip 100 columns were
equilibrated with 15 ml of equilibration buffer [0.05 M MOPS (pH 7.0), 15% (v/v)
isopropanol, 1% (v/v) Triton X-100]. RNA suspensions in ddH2O were brought up to 0.1
M MOPS pH 7.0 and passed over the column 4 times. Columns were then washed twice with wash buffer [0.05 M MOPS (pH 7.0), 0.2 M NaCl]. Small RNAs were eluted by passing 15 ml of elution buffer [0.05 M MOPS pH 7.0, 0.75 M NaCl, 15% (v/v) EtOH] over the column. Fractions were collected manually in 1.5 ml microcentrifuge tubes. Generally, the first fraction contained no detectable RNA while the second and third fractions had the greatest amounts. Lesser amounts were detectable in subsequent fractions. To concentrate and combine the RNAs, fractions were precipitated at room temperature for 10 minutes with 1 volume of cold isopropanol (99%) followed by centrifugation at 14,000 g for 10 minutes at 4°C. RNA was resuspended in 25 μl ddH₂O per tube and combined in a single microcentrifuge tube prior to storage at -80°C.

**H. volcanii genomic DNA isolation**

Chromosomal DNA was isolated from 10 ml of *H. volcanii* culture in stationary-phase (A₅₅₀ ≥ 1.0). Culture was transferred into 1.5 ml microcentrifuge tubes and spun at 8,000 g for 3 minutes at 4°C to pellet the cells. The supernatant was discarded and the pellets were resuspended with 100 μl of TE pH 8.0 [10 mM Tris-HCl (pH 7.5), 1 mM EDTA (pH 8.0)] with 1% Triton X-100 (v/v). Cells were pipetted in this solution until completely lysed, then 200 μl of TE was added and the suspension was homogenized with a vortex mixer. Subsequently, 300 μl of phenol was added and the suspension was homogenized with a vortex mixer, then centrifuged for 5 minutes at 14,000 g at 4°C. The top aqueous layer was transferred to a clean microcentrifuge tube and 300 μl of chloroform was added. The suspension was homogenized with a vortex mixer and centrifuged for 5 minutes at 14,000 g at 4°C. Next, 300 μl of phenol : chloroform (50:50,
v/v) was added, the suspension was homogenized with a vortex mixer and centrifuged for 5 minutes at 14,000 g at 4°C. The top aqueous layer was transferred to a clean microcentrifuge tube and 300 μl of chloroform was added prior to homogenization with vortex mixer and centrifugation at 14,000 g for 5 minutes at 4°C. The top aqueous layer was transferred to a clean microcentrifuge tube and combined with 300 μl of cold 99% isopropanol to precipitate the nucleic acid. The tube was centrifuged 10 minutes at 14,000 g at 4°C; the liquid was discarded and the pellet was resuspended in 20 μl of TE buffer.

To eliminate RNA contamination, 2 μl of RNase A (200 ng, pre-boiled) was added to the DNA suspensions, and the tubes were incubated at 37°C for 30 minutes. RNase A was then inactivated and the DNA was extracted by adding 100 μl of TE and 300 μl of phenol. The suspension was homogenized with a vortex mixer and centrifuged at 14,000 g for 5 minutes at 4°C. The top aqueous layer was transferred to a clean microcentrifuge tube and combined with 300 μl of chloroform, homogenized, and centrifuged 5 minutes at 14,000 g at 4°C. The top aqueous layer was transferred to a clean microcentrifuge tube, and 300 μl of cold 99% isopropanol was added to precipitate the DNA. The tube was centrifuged at 14,000 g for 10 minutes at 4°C. Supernatant was discarded and the pellet was resuspended in 20 μl of ddH₂O. Preparations yielded approximately 1 μg/μl. The DNA was stored at -20°C until needed.

Plasmid mini-preps

Plasmids were recovered from E. coli routinely by using the Wizard SV Miniprep Kit (Promega). The manufacturer’s protocol was followed, and plasmids were eluted in
100 μl of nuclease-free water. Yields of DNA were at approximately 0.2-1 μg/μl concentration. Plasmid preparations were stored at -20°C until required.

Quantitation of DNA/RNA

Approximate quantitation of nucleic acids was achieved using a Spectronic™ Genesys 2 spectrophotometer. Dilutions of the nucleic acids in ddH₂O, usually 1:50, were placed in quartz cuvettes, and their absorbance at 260 nm was read. Concentrations of DNA or RNA were calculated with the following formula (Barker, 1998):

DNA concentration (μg/ml) = \( (A_{260}) \times \text{dilution factor} \times (50 \, \mu g/ml)/1 \, \text{OD}_{260} \) unit

RNA concentration (μg/ml) = \( (A_{260}) \times \text{dilution factor} \times (40 \, \mu g/ml)/1 \, \text{OD}_{260} \) unit

Electrophoretic techniques for nucleic acid separation and purification

Denaturing polyacrylamide gel electrophoresis of RNA and DNA

For separation of small RNAs, gels were prepared to a final concentration of 6% or 8% of 29:1 acrylamide : bis-acrylamide for Criterion™ (Bio-Rad) cassettes or 6% of 29:1 acrylamide : bis-acrylamide for sequencing gels (40 x 30 cm). For separation of DNA fragments ≤ 200 bp, only 6% gels were used. Per 100 ml of gel solution, 15 ml (6%) or 20 ml (8%) of AccuGel™ 29:1 (40%) acrylamide : bis-acrylamide solution (National Diagnostics) was combined with 20 ml of 5X TBE [prepared with the following per liter: 4.78 ml of 0.5 M EDTA (pH 8.0), 27.54 g of boric acid, and 59.2 g Tris base], 48 g of urea (final concentration of 8 M), and water to a final volume of 100 ml in a 250 ml erlenmeyer side-arm flask. A stir bar was added and the gel solution was heated using a hot plate to dissolve the urea. The gel solution was then allowed to cool to room temperature before addition of 1 ml of freshly mixed 10% (w/v) ammonium
persulfate and 20 μl of TEMED (National Diagnostics) to facilitate polymerization. Lastly, the solution was poured into an appropriate cassette, visible bubbles were removed, and an appropriate comb was inserted at the top of the gel. Criterion™ gels were allowed to polymerize for at least 1 hour while sequencing gels were covered at the top with damp paper towels and plastic wrap and allowed to polymerize overnight.

Before loading samples, gels were pre-run in 1X TBE buffer for 15 minutes (Criterion™) at 100 V or 1 hour (sequencing) at 1200 V. Wells were then rinsed with 1X TBE while samples were given 1 volume of denaturing loading dye [4 M urea, 10 mM EDTA (pH 8.0), 1 mM Tris-HCl, and a few grains each of xylene cyanol and bromophenol blue] and heated to 90°C for 2.5 minutes. Samples were loaded and electrophoresed at a constant voltage of 100 V (Criterion™) or 1200 V (sequencing) until the BPB approached the bottom of the gel. At this point, gels were either stained with ethidium bromide (see Gel Staining), transferred to Zeta-Probe™ membrane (see Northern Analysis), subjected to UV-shadowing (see UV-shadowing), or covered in plastic wrap and exposed to a storage PhosphorScreen (Amersham).

**Agarose gel electrophoresis of RNA**

For separation of larger RNAs, partially denaturing conditions were used for separation of RNA agarose gels [1.2% (w/v) agarose, 1X TBE]. The 65 ml gel mix was prepared in a 250 ml erlenmeyer flask and heated in a microwave to melt the agarose. After allowing the flask to cool, yet remain warm to touch, the gel mix was poured into a taped-off tray with an appropriate comb and allowed to cool completely. Samples were mixed with 1 volume of denaturing loading dye and heated to 90°C for 2.5 minutes just
before loading. The gel was covered with 1X TBE running buffer and samples were loaded and run at a constant voltage of 100 V until the BPB migrated three-fourths of the gel length. Gels were then either stained with ethidium bromide (see Gel staining) or transferred to Zeta-Probe™ membrane (see Northern analysis).

Non-denaturing agarose gel electrophoresis of DNA

To separate DNA fragments 200 bp and larger, 30 ml agarose gels [1% (w/v) agarose, 1X TAE] were prepared in a 250 ml erlenmeyer flask. The mix was heated in a microwave to melt the agarose. After allowing the flask to cool, yet remain warm to touch, the gel mix was poured into a tray with an appropriate comb and dams to either side and allowed to cool completely. The gel was then covered with 1X TAE [diluted from 50X stock: 242 g of Tris-base, 57.1 ml of glacial acetic acid, 100 ml of 0.5 M EDTA (pH 8.0)] running buffer. In general, 10 μl samples were combined with 2 μl of 6X non-denaturing loading dye [30% (v/v) glycerol, 0.25% BPB], loaded, and run at a constant voltage of 100 V until the BPB migrated three-fourths of the gel length. Gels were then subjected to ethidium bromide staining (see Gel staining).

Gel staining with ethidium bromide

For visualization of nucleic acid separation, both agarose and polyacrylamide gels were subjected to ethidium bromide staining. Trays were filled with enough water to cover the gel and 20 μl of ethidium bromide [10 mg/ml stock] were added. The solution was gently swirled to mix in the ethidium bromide. Gels were submerged in the trays and trays were set on a platform shaker to stain for 10 minutes. The ethidium bromide solution was then replaced with an equal volume of water was added to allow destaining.
for 10 minutes. Nucleic acids were visualized by placing gels on a UV light box and gels were documented with a digital camera.

**Recovery of nucleic acids from agarose**

Restriction enzyme digest products and PCR products were routinely stained with ethidium bromide for visualization and recovered using Freeze N Squeeze™ tubes (Bio-Rad), followed by precipitation with EtOH. Using a clean razor blade, bands were excised from the gel, chopped, and transferred into Freeze N Squeeze™ tubes. These tubes were incubated at -20°C for 5 minutes, and then centrifuged in a microcentrifuge at 13,000g for 3 minutes as per manufacturer’s instructions. The recovered aqueous solution containing the nucleic acid was transferred to a clean microcentrifuge tube. To precipitate and concentrate the nucleic acid, 0.1 volumes of 3 M sodium acetate and 2.5 volumes of 95% EtOH were added to the solution. The tube was placed at -80°C to facilitate precipitation for enough time to completely freeze the liquid (at least 45 minutes). Next, the tube was centrifuged in a microcentrifuge for 10 minutes at 14,000g at 4°C. Liquid was decanted and the nucleic acid pellet was resuspended in a volume of 20 μl of ddH₂O.

**Recovery of nucleic acids from polyacrylamide**

Recovery of DNA or RNA from polyacrylamide was accomplished by first visualizing the band of interest either through staining with ethidium bromide or UV-shadowing. The band was excised using a clean razor blade and the gel slice was placed inside a clean microcentrifuge tube. The entire slice was covered with Qiagen gel diffusion buffer [0.5 M ammonium acetate, 2 mM EDTA (pH 8.0), 0.1% SDS], and the
tube was placed at 4°C overnight. To recover and concentrate the nucleic acid, the gel diffusion buffer was transferred to a clean microcentrifuge tube and 1 volume of cold 99% isopropanol was added. Samples were mixed by inversion and incubated at room temperature for 10 minutes to precipitate the nucleic acids. The tube was then centrifuged for 10 minutes at 14,000g and the liquid was decanted. The nucleic acid pellet was resuspended in a suitable volume, usually 20 μl, and used immediately or stored at -20°C (DNA) or -80°C (RNA).

**UV-shadowing**

When it was necessary to avoid ethidium bromide intercalating in nucleic acids, UV-shadowing was used for visualization of nucleic acids electrophoresed on polyacrylamide gels. Once electrophoresis had finished, gels were placed on fluorescent indicator-coated TLC plates (Sigma). The gel and TLC plate were irradiated with UV light. Nucleic acids appeared dark against the fluorescent background; the bands were excised with a clean razor blade and transferred to clean microcentrifuge tubes for recovery of the nucleic acid (see Nucleic Acid Recovery from Polyacrylamide Gels).

**Preparation of competent cells and transformation**

*E. coli* competent cell preparations

Competent *E. coli* cells from an isolated colony of the desired strain (DH5α and JM110) were prepared by first incubating a 4 ml starter culture in LB medium overnight at 37°C with shaking at 220 rpm. The next day, the starter culture was used to inoculate a 400 ml LB culture (1/100 dilution), which was grown at 37°C with shaking until the culture was in the mid log phase (A₆₀₀ nm of 0.375-0.4). The culture was then aliquoted
into sterile, pre-chilled polypropylene tubes and placed on ice for 10 minutes. The cells were pelleted by centrifugation in a Beckman JA25.50 rotor at 1,100 g for 7 minutes at 4°C and then resuspended in 10 ml of ice-cold CaCl$_2$ solution [60 mM CaCl$_2$, 15% v/v glycerol, 10 mM PIPES pH 7.0] by gently inverting the tubes, which were then kept on ice. The centrifugation step was repeated at 750 g for 5 minutes at 4°C, the supernatant was decanted, and the pellets were gently resuspended again in 10 ml of the ice-cold CaCl$_2$ solution and incubated on ice for 30 minutes. After incubation the centrifugation step (750 g for 5 minutes at 4°C) was repeated. The pelleted cells were then resuspended with 2 ml of the ice-cold CaCl$_2$ solution per tube, and 400 μl aliquots were placed into pre-chilled microcentrifuge tubes then immediately frozen at -80°C.

**Transformation of E. coli**

All of the *E. coli* competent strains (DH5α and JM110) were transformed using the same transformation protocol differing only in the amount of DNA used for transformation, 1 μl (DH5α) or 10 μl (JM110) of plasmid DNA from a miniprep, or the total volume of a TA cloning reaction (10 μl) or ligation reaction (50 μl), was added to 100 μl of thawed competent cells in a microcentrifuge tube and was incubated on ice for 25 minutes. The cells were then heat shocked at 42°C for 60 seconds and then returned to ice for an additional 2 minutes. Each transformation tube received 1 ml of LB medium without antibiotics and the cells were incubated at 37°C for 1 hour with shaking at 220 rpm. After outgrowth, the cells were pelleted by centrifugation at 14,000 g for 3 minutes, the supernatant was removed and the pellet was resuspended in 100 μl of LB supplemented with 100 μg/ml of ampicillin for selection. The cells were then plated on
AXI [LB plates supplemented with X-gal (0.004%), IPTG (0.04 mM) and ampicillin (100 μg/ml)] and incubated overnight at 37°C.

**Competent *H. volcanii* cell preparations**

The following protocols for *H. volcanii* competent cells and transformation were adapted from Cline and co-workers (Cline et al., 1989). *H. volcanii* competent cells were freshly prepared for each transformation by growing the desired *H. volcanii* strain DS70-ASD40, DS70-ASD41, or DS70-ASD40-sno) in 30 ml of CX medium in a 50 ml Falcon tube at 42°C until the culture reached the mid-log phase (A<sub>550</sub> = 0.6-0.8). The culture was aliquoted into two sterile 50 ml Falcon tubes, and cells were pelleted by centrifugation with a Beckman JA 25.50 rotor at 3,700 g for 6 minutes at 4°C in sterile 50 ml Falcon tubes. The pellets were gently resuspended in 500 μl of spheroplasting solution [50 mM Tris pH 8.2; 0.8 M NaCl; 27 mM KCl; 15% (w/v) sucrose] in order to produce spheroplasts. These spheroplasts were competent for transformation and were used immediately.

**Transformation of *H. volcanii***

*H. volcanii* spheroplasts were transformed in 200 μl aliquots in 1.5 ml microcentrifuge tubes. First, 20 μl of 0.5 M EDTA (pH 8.0) was added to the spheroplasts. After gently mixing the suspension, 10 μl of DNA mix [3 μl of plasmid DNA miniprep, approximately 0.2 μg of DNA plus 7 μl spheroplast solution] was added to the spheroplasts and the suspension was incubated for 5 minutes at room temperature. Then 230 μl of 60% v/v PEG-600 [600 μl PEG-600; 400 μl spheroplast solution] was added to the spheroplasts, the suspension was mixed by gently inverting the tubes, and...
incubation was continued at room temperature for 15 minutes. Next, 1 ml of regeneration buffer [CDM; 10% (w/v) sucrose] was added to the cells, the suspension was gently mixed and the suspension was centrifuged at 8,000 g for 3 minutes at 4°C in a microcentrifuge. The supernatant was decanted, the pelleted cells were gently resuspended in 90 μl of regeneration buffer, and resuspensions were plated onto CDM plates. Incubation was at 42°C in a sealed plastic bag to prevent drying.

**Recombinant DNA construction**

**Polymerase chain reaction**

Unless stated otherwise, all polymerase chain reactions were performed using the following method. Oligonucleotide primers were purchased from Integrated DNA Technologies (IDT). Reactions were carried out using the My Cycler Personal Thermocycler (BioRad). Assembly of reaction components was done on ice in 200 μl PCR tubes. Components added to the tubes were 5 μl of 10X PCR Buffer supplied with the enzyme [200 mM Tris-HCl (pH 8.4), 500 mM KCl], 5 μl of DMSO, 3 μl of 10 μM forward primer, 3 μl of 10 μM reverse primer, 1-2 μl of template (approximately 1-2 μg of plasmid DNA or *H. volcanii* genomic DNA unless otherwise noted), 2 μl of 10 mM dNTPs, 1.5 μl of 50 mM MgCl₂, 5 units of *Taq* DNA Polymerase (Invitrogen), and water to a final volume of 50 μl. The thermocycler was programmed, unless specified otherwise, for 30 cycles with denaturation at 95°C for 30 seconds, primer annealing temperature varied depending on the Tₘ of the oligonucleotides used (typically 58°C-64°C), extension of the partial duplex at 72°C for one minute, followed by a 10 minute final extension step at 70°C after the 30 cycles.
Colony PCR was routinely used to screen *H. volcanii* transformant colonies to confirm the presence of a plasmid or to confirm that a region was deleted from the chromosome. Colony PCRs were assembled identically to standard PCR with one exception: a portion of an *H. volcanii* colony was added to the reaction mix to supply template DNA. Reactions were then run using the same conditions described above.

Nested PCR, as in the case of the tRNAMet sRNA deletion construct, was used to fuse together 5’ and 3’ regions, connected by a short linker sequence, for pop-in-pop-out deletion. The reaction was assembled as above using the agarose gel purified 5’ and 3’ fragments as primers. Standard PCR conditions were used for 10 cycles. At this point, 1 μl of 10 μM stocks of both forward and reverse primer were added and the reaction was continued for another 30 cycles under standard PCR conditions.

**TA cloning**

In general, PCR products generated with *Taq* polymerase were cloned by using the pGEM™- T Easy Vector System (Promega). The manufacturer protocol was followed for the ligation reactions in which 5 μ of 2X Rapid Ligation Buffer supplied with the kit, 1μl of pGEM™-T Easy Vector, 3 μl of PCR product, and 1 μl of T4 DNA ligase were combined on ice in a microcentrifuge tube, mixed, and incubated at 4°C overnight. The next day, the entire ligation reaction was used in the transformation of DH5α and plated on AXI (see Transformation of *E. coli*).

**Site-directed mutagenesis**

Three specific residues in the intron endonuclease were mutated using a method of site-directed mutagenesis that is derived from the QuikChange™ (Agilent
Technologies) protocol. Mutagenic primers were designed according to QuikChange™ specifications. Reactions were assembled on ice with 5 μl of DMSO, 5 μl of 10X reaction buffer provided with the enzyme [200 mM Tris-HCl (pH 8.8), 20 mM MgSO₄, 100 mM KCl, 100 mM (NH₄)₂SO₄, 1% Triton X-100, 1 mg/ml nuclease-free BSA], 5 to 50 ng of dsDNA template, 125 ng each of forward and reverse mutagenic primers, 1 μl of 10 mM dNTPs, and ddH₂O to a final volume of 50 μl. The reaction mix then received 2.5 units of Pfu polymerase.

The program for the mutagenic reaction in the MyCycler™ thermocycler (Bio-Rad) was based on the QuikChange™ program recommendations. First, a single 30 second step at 95°C, then multiple cycles of 95°C for 30 seconds, 55°C for 60 seconds, and 68°C for 60 seconds per Kb of plasmid length. The number of cycles was determined by the type of mutation: 12 cycles for single-nucleotide mutations or 16 cycles to mutate 2-3 adjacent nucleotides. Following the mutagenesis, reactions were cooled on ice for a minimum of 2 minutes before proceeding. The cooled reaction tubes received 10 units of DpnI and were incubated at 37°C for 1 hour to digest parental DNA. Finally, 1 μl of the reaction mix was used to transform E. coli DH5α cells (see Transformation of E. coli). Plasmids were recovered from transformants (see Plasmid mini-preps) and sequenced to verify the mutation.

**Restriction enzyme digests**

Unless otherwise specified, restriction enzyme digests were conducted by combining about 10 μg of plasmid DNA, 3 μl of appropriate 10X reaction buffer, 0.25 μl of each restriction enzyme, and water to a final volume of 30 μl in a microcentrifuge
tube. Digests were incubated at 37°C for 1 hour and separated by non-denaturing agarose gel electrophoresis or denaturing polyacrylamide gel electrophoresis depending on the size of the drop-out fragment.

**Ligation of DNA fragments**

Generally, DNA released from vectors by restriction enzyme digestion and appropriately digested vector DNA were combined after separation on and recovery from a non-denaturing agarose gel. Ligation reactions were prepared in microcentrifuge tubes on ice and carried out in 1X reaction buffer supplied with the enzyme [50 mM Tris-HCl, 10 mM MgCl₂, 10 mM DTT, 1 mM ATP, pH 7.5 at 25°C] using an approximate 3:1 ratio of purified fragment to purified linearized vector and 1 µl of T4 DNA ligase. The final volume of each reaction was brought up to 50 µl with water. Reactions were incubated at 4°C overnight and subsequently used in transformation of *E. coli* DH5α.

**Sequencing of plasmids**

Unless otherwise noted, all plasmid constructs were verified by DNA sequencing. Plasmids were submitted to the Ohio State University’s Plant Microbe Genomics Facility (PMGF) and sequenced by the 3730 DNA Analyzer (Applied Biosystems) with an appropriate primer.

**Construction of strains and vectors**

Strains generated from this work are summarized in Table 2.4. A summary of relevant features in the vectors used to construct these strains is given in Figure 2.1.
Construction of the tRNAMet sRNA deletion strain

A pop-in pop-out strategy was used to delete the putative tRNAMet sRNA gene from the *H. volcanii* chromosome (Figure 2.2). Briefly, sequences 5’ and 3’ of the tRNAMet sRNA region were amplified and joined using nested PCR (primers: MSD-F, MSD-5’Lk, MSD-3’Lk, and MSD-R). This DNA was cloned into the plasmid pBSK and introduced into *H. volcanii* DS70 ASD-40, a pyrF deletion strain. Plasmid pBSK does not replicate in *H. volcanii* and confers the Ura⁺ phenotype through homologous recombination of the tRNAMet sRNA flanking sequences; this is the pop-on stage. Single cross-over mutants are then challenged with 5-FOA and cells having undergone a second homologous recombination to remove the pyrF gene carried in the plasmid were recovered, the pop-out. Recombinants in which both pyrF and tRNAMet sRNA have been deleted were identified by PCR analysis with primers MSDX-F and MSDX-R. Products of PCR were separated with non-denaturing agarose gel electrophoresis and a strain, DS70-ASD40-sno, was found to have a product matching the size expected for deletion of tRNAMet sRNA. Additional confirmation of deletion was obtained by TA cloning the colony PCR product and sequencing plasmid DNA isolated from transformants. Northern analysis was also used to determine if the tRNAMet sRNA was detectable in the deletion strain. Figure 3.8 indicates that the hybridization signal observed in the *H. volcanii* DS70 ASD40 is absent in the deletion strain.
<table>
<thead>
<tr>
<th>Strain</th>
<th>Background</th>
<th>Genotype; Plasmid Conferred Phenotypes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Previously Generated Strains</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DS70</td>
<td>DS70</td>
<td>wildtype strain cured of pHV2</td>
<td>Wendoloski et al (2001)</td>
</tr>
<tr>
<td>DS70-ASD40</td>
<td>DS70</td>
<td>ΔpyrF; PyrF</td>
<td>Sabag-Daigle (2009)</td>
</tr>
<tr>
<td>DS70-ASD41</td>
<td>DS70-ASD40</td>
<td>ΔpyrFΔhutR; PyrF, HutR</td>
<td>Sabag-Daigle (2009)</td>
</tr>
<tr>
<td><strong>Strains from this study</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DS70-ASD40-msdC</td>
<td>DS70-ASD40</td>
<td>ΔpyrF; pBSK-metsnodel: PyrF⁺</td>
<td></td>
</tr>
<tr>
<td>DS70-ASD40-sno</td>
<td>DS70-ASD40</td>
<td>ΔpyrFΔtRNAmetsRNA; PyrF⁺, tRNAmetsRNA⁻</td>
<td></td>
</tr>
<tr>
<td>DS70-ASD40-snorec</td>
<td>DS70-ASD40-sno</td>
<td>ΔpyrFΔtRNAmetsRNA; pASDI-snorec: PyrF⁺, tRNAmetsRNA⁻</td>
<td></td>
</tr>
<tr>
<td>DS70-ASD41-L6H</td>
<td>DS70-ASD41</td>
<td>ΔpyrFΔhutR; pARS-L6H: PyrF⁺, HutR⁺, L6H⁺</td>
<td></td>
</tr>
<tr>
<td>DS70-ASD41-endA*</td>
<td>DS70-ASD41</td>
<td>ΔpyrFΔhutR; pARS-endA⁺: PyrF⁺, HutR⁺, endA⁺⁺</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.4 List of *H. volcanii* strains and relevant features used for the experiments in this work and strains generated by this work.
Figure 2.1 Relevant features of the plasmids used to generate the strains from this study described in Table 2.4. The U feature of the pARS plasmids indicates the unique start of the protein consisting of the first two HutU amino acids and those derived from the BamHI restriction site. The bolded T’s indicate transcription terminators. The 6H refers to the hexahistidine tag.
Figure 2.1 Relevant features of the plasmids used to generate the strains from this study described in Table 2.4.
Figure 2.2 Schematic representation of the pop-in/pop-out genetics used to construct the tRNAMet sRNA-deleted strain. Plasmid elements are shown in blue while chromosome elements are shown in white.
Figure 2.2 Schematic representation of the pop-in/pop-out genetics used to construct the tRNAMet sRNA-deleted strain.
Construction of the tRNAMet sRNA complementation strain

The region deleted in the DS70-ASD40-sno strain was amplified by PCR using DS70-ASD40 genomic DNA as a template (primers: snorec-F and snorec-R). The product was TA cloned, used in the transformation of DH5α, and sequenced. As the native promoter of this RNA is uncertain, the previously characterized *H. volcanii* tRNALys promoter was used to ensure transcription of the sRNA (Palmer and Daniels, 1995). Therefore, the tRNALys promoter was used to replace the tRNAProM reporter in the pASDII vector (Sabag-Daigle, 2009). In the resulting vector, pASDII-snorec, the tRNAMet sRNA gene is fused to the constitutive tRNALys promoter. This vector is non-replicative in *H. volcanii* and is designed to integrate into the chromosome at a site adjacent to the pyrF gene.

To complement tRNAMet sRNA, pASDII-snorec was introduced by transformation into DS70-ASD40-sno. PCR analysis was also used to verify that the tRNAMet sRNA encoding region was present at the site of recombination and the original deletion region remained unchanged. Northern analysis was used to confirm that the tRNAMet sRNA was transcribed in the complementation strain, DS70-ASD40-snorec.

Construction of the inducible hexhistidine-tagged L7ae strain

To study native L7ae-containing complexes and discover novel L7ae-containing complexes, a strain carrying an inducible carboxy-terminal-tagged L7ae was generated. Primers were designed to amplify the L7ae gene from DS70 genomic DNA while fusing a C-terminal hexahistidine tag to the protein (L6H-F and L6H-R). Due to the adjacent
TGA stop codon and XbaI restriction site (TCTAGA), the vector had to be passed through a Dam- strain as the methylase activity at GATC blocked XbaI activity. Therefore, the TA cloning vector containing the his-tagged L7ae (L6H) was used in transformation of the E. coli strain JM110. Purified L6H fragment was then used to replace the tRNAProM reporter in the vector pARS resulting in a fusion of L6H to the hutU promoter (Sabag-Daigle, 2009). This vector, pARS-L6H, was used for transformation of H. volcanii DS70-ASD41. The strain was designated DS70-ASD41-L6H. Presence of the plasmid in transformants was verified by colony PCR with primers hindhuF and L6HR.

Construction of the intron endonuclease triple mutant strain

To investigate the role of the essential intron endonuclease gene, endA, the inducible hutU promoter module was used to regulate the expression of a mutant form of the enzyme (Sabag-Daigle, 2009). First, the wild type intron endonuclease was amplified with primers endA-F and endA-R. Next, the enzyme was mutated to inactivate the catalytic core but still permit dimerization. This step was accomplished by using 3 rounds of site-directed mutagenesis to introduce mutations based on published work (Calvin et al., 2008; Kim et al., 2007). Mutations were Y274F, H285A, and K316E using primers: YF-F, YF-R, HA-F, HA-R, KE-F, and KE-R. Due to the adjacent TGA stop codon and XbaI restriction site (TCTAGA), the vector had to be passed through a Dam-strain as the methylase activity at GATC blocked XbaI activity. Therefore, the TA cloning vector containing the mutated intron endonuclease (endA*) was used in the transformation of the E. coli strain JM110. Purified endA* fragment was then used to
replace the tRNAProM reporter in the vector pARS resulting in a fusion of endA* to the,
hutU promoter. This vector, pARS-endA*, was used for transformation of H. volcanii,
DS70-ASD41. The strain was designated DS70-ASD41-endA*. Presence of the plasmid
in transformants was verified by colony PCR with primers hindhut-F and endA-R and by
performing a plasmid mini-prep on the strain. The mini-prep DNA was used for
transformation of E. coli DH5α, and plasmid recovered from transformants was analyzed
by restriction enzyme digest and sequencing.

**Techniques for analysis of RNA**

**Radiolabeling of DNA probes**

Approximately 10 picomoles of oligonucleotide probe was labeled at the 5’ end in
a reaction containing 1X Forward Reaction Buffer [70 mM Tris-HCl (pH 7.6), 10 mM
MgCl₂, 100 mM KCl, 1 mM 2-mercaptoethanol] supplied with the enzyme, 100 μCi of
[γ-³²P]ATP (100 μCi/μl, 7,000 Ci/mmol), and 5 units of T4 polynucleotide kinase. The
volume was increased to 25 μl with ddH₂O. Labeling reactions were incubated at 37°C
for 1 hour. The enzyme was inactivated by heating the reaction mixture to 65°C for 15
minutes. For Northern hybridizations, the entire 25 μl volume was added to
hybridization buffer [0.5 M Na₂HPO₄●7H₂O (pH 7.2), 7%SDS].

**Northern analysis**

RNA samples containing 5-15 μg of RNA were first separated by electrophoresis
in 6 or 8% (w/v) denaturing polyacrylamide or 1.2% agarose (w/v) as described in
acrylamide or agarose gel electrophoresis, respectively. Nucleic acid transfers to Zeta-
Probe™ membrane were accomplished by capillary transfer with the TurboBlotter™
system (Schleicher & Schuell). Membranes were equilibrated in transfer buffer [3 M NaCl, 8 mM NaOH]. Transfer stacks were assembled according to the manufacturer using a 3 cm stack of GB004 blotting paper cut to match membrane dimensions, Zeta-Probe™ membrane, the gel, an additional sheet of GB004 cut to match gel dimensions, a wick of GB004, and the wick cover. Gels were transferred 2.5 hours to overnight at room temperature. Membranes were fixed using UV irradiation at 125 mJoules/cm² with a GS Gene Linker™ from Bio-Rad.

Prehybridization of membranes consisted of placing membranes in a glass hybridization tube with 15 ml (150 μl/cm² of membrane) of hybridization buffer [0.5 M Na₂HPO₄•7H₂O (pH 7.2), 7%SDS (w/v)] and incubating at 65°C with rotation for 15 minutes. The prehybridization buffer was then exchanged for 15 ml of fresh hybridization buffer, and the labeled oligonucleotide probe(s) were added. Hybridization was carried out overnight with rotation in a hybridization oven at a temperature based on the Tₘ of the probe, typically between 52°C and 65°C.

The following day, the membrane was washed 3 times while allowing the oven to cool to room temperature. For the first 2 washes, hybridization buffer was replaced with 15 ml of wash buffer [45 mM Na₂HPO₄•7H₂O (pH 7.2), 5% SDS (w/v)] and washed with rotation for 20 minutes. The final wash was conducted with 15 ml of stringent buffer [0.8% SDS (w/v), 2X SSC] for 20 minutes with rotation. Membranes were then blotted with filter paper, covered in plastic wrap, and placed inside an exposure cassette with a bleached storage PhosphorScreen (Amersham).
Quantitation of Northern hybridizations was achieved by scanning storage PhosphorScreens (Amersham) using the PhosphorImager (Molecular Dynamics). The Volume Analysis Tool in the ImageQuaNT program (Molecular Dynamics) was then used to measure radioactive spots relative to each other, the background, or a control hybridization signal.

Stripping of Zeta-Probe™ membranes

In cases where a membrane required successive hybridizations, membranes were stripped of previous probes by placing them in glass hybridization tubes and adding 20 ml of stripping buffer [1% SDS, 0.1X SSC]. The tubes were placed in a hybridization oven about 30 minutes with rotation and the temperature was set to 95°C. Once the oven reached 95°C, the buffer was replaced with 20 ml of fresh stripping buffer, and the tubes were incubated an additional 45 minutes at 95°C with rotation. Buffer was then discarded and replaced with hybridization buffer and the oven temperature was lowered to 65°C for pre-hybridization of the membrane in preparation for the next radiolabeled probe.

$^{32}$pCp synthesis and labeling

For 3’ end-labeling, $^{32}$pCp synthesis was carried out in 1X Forward Reaction Buffer [70 mM Tris-HCl (pH 7.6), 10 mM MgCl$_2$, 100 mM KCl, 1 mM 2-mercaptoethanol] supplied with the enzyme, with the following components: 100 μCi of $[\gamma^{32}]$P]ATP (100 μCi/μl, 7,000 Ci/mmol), 5μl of 1.5 mM Cp, and 10 units of T4 polynucleotide kinase (Invitrogen). The remaining volume made up by ddH$_2$O to a final
volume of 50 μl. Synthesis reactions were incubated at 37°C overnight then subjected to 15 minutes at 65°C to heat-inactivate any remaining enzyme activity. The synthesized $^{32}$pCp was stored at -20°C.

All 3’ end-labeling reactions were set up with a master mix containing the following components per 10 μL: 1 μl $^{32}$pCp, 1 μl RNA, 10% DMSO, 1X reaction buffer [50 mM Tris-HCl, 10 mM MgCl$_2$, 10 mM DTT, 1 mM ATP, pH 7.8 at 25°C] supplied with the enzyme, and 5 units of T4 RNA Ligase 1 (New England Biolabs). Ligation reactions were incubated at 4°C overnight. Reactions were terminated by addition of 1 volume of denaturing loading dye (see Denaturing polyacrylamide gel electrophoresis of RNA) then stored at -20°C prior to denaturing gel electrophoresis.

For separation by denaturing gel electrophoresis on a 100 ml sequencing gel, a 5 μl sample was aliquoted into a 0.5 ml PCR tube and heated to 90°C for 2.5 minutes just before loading onto the gel. Electrophoresis continued at a constant voltage of 1200 V until the BPB was a few centimeters from the bottom of the gel. Spent x-ray film was used to pick up gels from the glass plates. Gels were covered in plastic wrap and placed in a cassette with a storage PhosphorScreen (Amersham).

**Reverse transcription PCR**

RT-PCR was performed using a suitable amount of RNA (5-15 μg) obtained through the method listed in Isolation of RNA. Samples were combined with 5 μl RQ1 RNase-Free DNase (Promega), 2.8 μl of 10X reaction buffer provided with the enzyme [400 mM Tris-HCl (pH 8.0 at 25°C), 100 mM MgSO$_4$, 10 mM CaCl$_2$], and ddH$_2$O for a total volume of 28 μl. Reactions were incubated at 37°C for 30 minutes to degrade any
remaining DNA in the samples. To terminate the reactions, 3 μl of stop solution provided with the enzyme [20 mM EGTA (pH 8.0 at 25°C)] was added to each tube. Samples were assessed for DNA contamination by running PCR with appropriate primers before proceeding.

Reverse transcription was conducted according to manufacturer’s instructions. An appropriate amount of DNase-treated RNA (0.5 μg – 5.0 μg) was combined with 1 μl of 10 mM dNTPs, 1 μl of 50 μM oligonucleotide primer, and ddH₂O to a final volume of 13 μl in 200 μl PCR tubes. Reactions were heated to 65°C for 5 minutes and then chilled on ice for 1 minute for primer annealing. To each cooled reaction, the following components were added to a total volume of 20 μl: 4 μl of 5X First-Strand Buffer provided with the enzyme [250 mM Tris-HCl (pH 8.3 at room temperature), 375 mM KCl, 15 mM MgCl₂], 1 μl of 0.1 M DTT, 40 units of RNaseOUT™, and 200 units of SuperScript™ III Reverse Transcriptase (Invitrogen). Tubes were incubated in a thermocycler at 55°C for 60-90 minutes for reverse transcription, and reactions were terminated by inactivating the reverse transcriptase at 70°C for 15 minutes.

PCR analysis of reverse transcription products was accomplished as described in PCR of DNA. The template DNA for these analyses consisted of 2 μl of reverse transcription product as per manufacturer’s recommendation.

**In vitro transcription**

T7 *in vitro* transcripts were synthesized using the Riboprobe™ System-T7 (Promega). The manufacturer’s protocol for Synthesis of Large Amounts of RNA was used to generate transcripts from PCR products that were purified by UV-shadowing (see
Recovery of nucleic acids from gels). Reaction components were added sequentially to a 200 μl PCR tube and assembled at room temperature as per manufacturer’s instructions. The mix consisted of 20 μl of Transcription Optimized 5X Buffer, 10 μl of 100 mM DTT, 100 units of Recombinant Rnasin™ Ribonuclease Inhibitor, 20 μl of nucleotide mixture [ATP, GTP, CTP, and UTP at 2.5 mM each], 2 μl purified PCR product template, 40 units of T7 RNA polymerase, and 43.5 μl of nuclease-free water. The tube was incubated at 37°C in a thermocycler for 2 hours. To terminate the reaction, 2 units of RQ1 RNase-Free DNase (Promega) were added and the tube was incubated an additional 15 minutes at 37°C. Newly synthesized transcripts were subsequently purified using either the Trizol method (RNA isolation) or UV shadowing following separation on a denaturing polyacrylamide gel (Recovery of Nucleic Acids from Gels).

**RNase H detection 2’-O-methyl groups**

This protocol is derived from an established method (Yu et al., 1997).

A chimeric oligonucleotide, metRH, consisting of a 2’-O-methyl backbone and four DNA bases was designed. This oligonucleotide was complementary to tRNAMet and the DNA bases spanned the site of methylation in the tRNA thereby targeting RNase H cleavage to the site of modification (Figure 2.3). The 5’ end nucleotide of the chimera pairs with A37, just before the splice site, allowing this oligonucleotide to pair with intron-containing pre-tRNAMet or mature tRNAMet.

In order to assay for 2’-O-methylation at C34, H. volcanii total RNA was isolated from the following strains DS70-ASD40-sno, DS70-ASD40, or DS70-ASD40--sno-rec during the exponential growth phase in CDM. Additionally, T7 transcripts of pre-
Figure 2.3 Schematic representation of the RNase H method for assessing whether or not a particular nucleotide is 2’-O-methylated. Sequences of tRNAMet and the targeting probe are given as indicated. The Cm and * represent a 2’-O-methylated nucleotide in the sequence or drawing, respectively. The DNA targeting sequence of the oligonucleotide is indicated in bold.
Figure 2.3 Schematic representation of the RNase H method for assessing whether or not a particular nucleotide is 2’-O-methylated.
tRNAMet and intronless tRNAMet were synthesized from PCR products of vectors containing these genes.

The reactions were carried out in 1.5 ml microcentrifuge tubes as follows. Approximately 10-15 μg of *H. volcanii* purified total RNA was precipitated with 1 volume of isopropanol and resuspended in 5 μl of ddH₂O. A total volume of 5 μl was also prepared from the T7 transcripts by combining 2.5 μl RNA with 2.5 μl ddH₂O. Duplicates of each sample were prepared to use in a no-enzyme control reaction. Next, 3 μl of chimeric oligonucleotide (50 μM) was added and the tubes were heated to 90-95°C for 3 minutes then placed in a 37°C water bath for 10 minutes to facilitate annealing.

During the annealing time, reaction master mix was prepared. Each reaction contained 2 μl of 10X reaction buffer [200 mM Tris-HCl (pH 7.5), 100 mM MgCl₂, and 1 M KCl], 5 μl of 20% sucrose (w/v), 2 μl of DTT (10 mM), 2 units of RNase Out (Sigma), and 1 unit of RNase H (Sigma) for a total volume of 12 μl. For the no-enzyme controls, mix in which ddH₂O was substituted for RNase H was also prepared.

Once the chimera/RNA mix was cooled, 12 μl of RNase H master mix or no-enzyme mix was added to the samples giving a final volume of 20 μl. The reactions were incubated at 37°C for 30 minutes. Termination of the reactions was achieved by adding 10 μl of denaturing loading dye (see Denaturing Polyacrylamide Gel Electrophoresis). Samples were heated to 90°C for 2.5 minutes then loaded onto 8% acrylamide Criterion™ (BioRad) gel. Electrophoresis was carried out as described in Denaturing Polyacrylamide Gel Electrophoresis. Gels were then transferred to (Zeta-Probe™)
membrane prior to hybridization with radio-labeled DNA oligonucleotides as detailed in the Northern analysis section and quantitated as described in Quantitation of Northerns.

**Protein analysis and purification**

**Protein induction in *H. volcanii***

Exploitation of the *hutU* promoter and HutR regulator to induce production of specific proteins in the presence of urocanic acid has been described previously (Sabag-Daigle, 2009). Strains of *H. volcanii* harboring plasmids for the controlled expression of mutant intron endonuclease (DS70-ASD41-*endA*) and affinity-tagged L7ae (DS70-ASD41-L6H) were grown to exponential phase (*A*<sub>550</sub>: 0.1 to 0.2) in CDM. At this point, cultures were split in half and the control half received a mock induction while the experimental half received freshly prepared urocanic acid (pH 7.3) to a final concentration of 10 mM. Culture samples were most commonly harvested at 2 to 4 hours post-induction by centrifugation for 10 minutes at 7,800 g 4°C using the Beckman JLA 10.500 rotor for whole-cell extracts (0.5 to 1 L cultures) or the same centrifugation conditions with the Beckman JA 25.50 rotor for RNA samples (45 ml aliquots of culture).

**Preparation of whole-cell extracts from *H. volcanii***

Total cell extract of *H. volcanii* was obtained through pressure lysis of cells. Cultures were grown to the desired optical density, as determined by *A*<sub>550</sub>, in appropriate growth medium for the strain, or samples of culture were taken at the desired timepoints post-induction. Cultures were transferred to sterile polypropylene bottles and cells were harvested by centrifugation at 7,800 g for 10 minutes at 4°C using the Beckman JLA 10.500 rotor. After decanting the spent media and removing remaining media with a
Pasteur pipette, the pellet from each bottle was gently resuspended on ice with 1 ml of TMK buffer [2.2 M KCl, 20 mM MgCl$_2$, 50 mM Tris-HCl (pH 7.5)]. Cell suspensions were combined by transferring them to a pre-chilled 50 ml Falcon tube on ice. Cells were then lysed by passing the cell suspension twice through a French Pressure Cell at 10,000 PSI. Lysates were cleared of insoluble material and unbroken cells by centrifugation at 7,700 g for 10 minutes at 4°C using a Beckman JA 25.50 rotor. Supernatant was transferred to a clean 15 ml Falcon tube without disturbing the pellet. Both the cleared lysate and the pellet were stored at -20°C.

**SDS-PAGE**

Separation of proteins ≥ 20 kDa was accomplished through SDS-PAGE. Unless otherwise noted, all gels were poured in Criterion™ cassettes. First 10% resolving gels were poured according to the manufacturer’s protocol. Next, a 3.7% stacking gel was poured and fitted with an appropriate comb.

Crude samples were prepared by adding 1 volume of 2X SDS loading buffer [0.125 M Tris-HCl (pH 6.8), 20% glycerol (v/v), 10% 2-mercaptoethanol (v/v), 4% SDS (w/v), 0.004% BPB (w/v)] and boiling the samples for 5 minutes immediately prior to loading on the gel. When greater precision was necessary, samples were precipitated with 2 volumes of cold acetone on ice for 30 minutes. The samples were centrifuged at 14,000 g for 10 minutes at 4°C and the liquid was decanted. Pellets were resuspended directly in 2X SDS loading buffer, boiled for 5 minutes, and immediately loaded onto the gel.
Gels were run at a constant voltage of 150 V in 1X SDS running buffer [192 mM glycine, 0.1% SDS (w/v), 25 mM Tris (pH 8.3)] until the BPB approached the bottom of the gel or suitable protein distribution was achieved as seen by the protein marker. Upon completion of the run, gels were either transferred to PVDF (see Western Analysis) or stained with coomassie blue.

To stain SDS PAGE gels, the gels were fixed in fixing solution [50% (v/v) EtOH, 10% (v/v) glacial acetic acid] for at least 20 minutes. Next, the fixing solution was replaced by coomassie blue staining solution [50% EtOH (v/v), 7.5% glacial acetic acid (v/v), 0.2% (w/v) coomassie brilliant blue G] for 1 hour to overnight. Destaining and storage of gels was in water.

**Tricine-PAGE**

Separation of proteins ≤ 20 KD was accomplished using the discontinuous buffer system and protocol for tricine gels (Schagger and von Jagow, 1987). All gels were poured in Criterion™ cassettes. Resolving gels were 16.1% acrylamide and contained 8.33 ml of 40% bis-acrylamide monomer solution, 6.66 ml of gel buffer [3 M Tris-HCl, 0.3% (w/v) SDS, pH 8.45], 2.67 g of glycerol, and 2.34 ml of water. The gel solution was stirred to attain homogeneity before adding 60 μl of freshly prepared 10% (w/v) of ammonium persulfate and 6 μl of TEMED to induce polymerization. Stacking gels of 4% and consisting of 1.25 ml of 40% bis-acrylamide monomer solution, 3.1 ml of gel buffer, and 8.15 ml of water were prepared and given 100 μl of 10% ammonium persulfate and 10 μl of TEMED to induce polymerization before adding the solution onto the already polymerized resolving gel.
Samples were prepared as described under SDS-PAGE and samples were run at a constant voltage of 125 V using the discontinuous buffer system of 1X cathode buffer [0.1 M Tris, 0.1% (w/v) SDS, pH 8.25] and 1X anode buffer [0.2 M Tris-HCl, pH 8.9] until suitable separation of proteins was achieved as determined by visual inspection of protein markers. Tricine gels were then subjected to staining (see SDS-PAGE) or transfer to PVDF (see Western Analysis).

**Western analysis**

Western blotting was used to confirm the presence of proteins in purifications. First, samples were prepared and run on either SDS PAGE or Tricine PAGE to obtain suitable 1 dimensional separation of proteins. Hybond™ PVDF membrane cut to gel-size were briefly soaked in 100% methanol as per manufacturer’s instructions, then equilibrated in transfer buffer [192 mM glycine, 25 mM Tris (pH 8.3), 10% (v/v) methanol] during the gel run. Gels were soaked in transfer buffer for 10 minutes to remove the SDS from the gel; during this time, fiber pads and filter paper were also soaked in transfer buffer.

Protein transfer was accomplished using the protocol for the Mini Trans-Blot Cell (Bio-Rad). Gel sandwiches were assembled, placed inside the tank, and covered with transfer buffer. Transfer proceeded for 2 to 2.5 hours at a constant voltage of 50 V with stirring. Cassettes were disassembled and membranes were briefly rinsed with ddH₂O, and then placed in blocking buffer [TBS (20 mM Tris-HCl, 137 mM NaCl, pH 7.7), 2% (w/v) milk powder] for 1 hour at room temperature or overnight at 4°C.
Primary antibodies used were either anti-endA (Kleman-Leyer et al., 1997) or anti-histidine (Qiagen tetra-his). To commence blotting, blocking buffer was replaced with antibody buffer [TBS, 1% (w/v) milk powder, 0.1% (v/v) Tween-20] with a 1:10,000 dilution (anti-endA) or a 1:25,000 dilution (anti-histidine) of primary antibody. Membranes were gently shaken at room temperature for 1 hour then washed 3 times with wash buffer [TBS, 0.1% (v/v) Tween-20] for 10 minutes per wash. Fresh antibody buffer containing a 1:50,000 dilution of secondary antibody was added and the membranes were shaken gently for 1 hour. Antibody buffer was discarded and the membranes were washed twice with wash buffer, then once with TBS at 10 minutes per wash. Membranes were drained of the buffer and placed face-up on clean plastic wrap. Immobilon™ Western Chemiluminescent HRP Substrate (Millipore) was used for detection. The manufacturer’s protocol for mixing and applying the reagents was followed. Reagents were then drained off the membranes, membranes were covered in clean plastic wrap, and exposed to x-ray film (Denville Scientific) for a suitable length of time prior to film development.

**Dot blotting**

For quick detection of fractions containing partially purified affinity-tagged proteins and complexes containing affinity-tagged proteins especially where UV absorbance was insufficient as a guide, dot blotting was used. Hybond™ PVDF membrane was cut to the size needed and wet in 100% methanol. The membrane was placed on a stack of filter papers cut to membrane size that were pre-soaked in transfer buffer (see Western analysis). Next, 5 μl of each sample and control was spotted onto the
membrane in an asymmetric grid and allowed to soak into the membrane by capillary action. After all samples had been absorbed, the membrane was placed in blocking buffer (see Western analysis) and treated identically to a Western blot with one exception: a brief rinse of the membrane with TBS was done just before adding primary antibody buffer.

**Protein A enrichment of antibodies**

The anti-EndA antibody was partially purified from whole sera (Kleman-Leyer et al., 1997) and enriched using a 1 ml HiTrap™ Protein A column (Pharmacia). The column was connected to a BioLogic™ LP system with fraction collector (Bio-Rad). To protect the antibody and neutralize the pH, 100 μl of 1 M Tris (pH 9) was placed in each 1.5 ml collection tube prior to antibody purification. The column was washed with 10 column volumes of binding buffer [20 mM sodium phosphate, pH 7.0]. Cleared sera was then applied to the column and followed by 10 column volumes of washing with binding buffer. The antibody was eluted in a step of 5 column volumes of elution buffer [0.1 M citric acid, pH 4]. Antibody elution was confirmed by separation with SDS-PAGE followed by staining using the UV absorbance reading from the BioLogic™ as a guide to identify protein-containing fractions. Purified antibody was stored at -20°C until needed.

**Affinity purification of hexahistidine-tagged proteins**

Proteins and complexes containing a protein with N or C terminal hexahistidine affinity tags were partially purified using a Pharmacia Biotech 5 ml HisTrap™ chelating column on a BioLogic™ LP system with fraction collector (Bio-Rad). The column was prepared according to the manufacturer’s instructions by first washing with 10 column
volumes of stripping buffer [20 mM NaH$_2$PO$_4$$\cdot$H$_2$O, 0.5 M NaCl, 50 mM EDTA (pH 8.0)]. Next, the column was washed with 10 column volumes of binding buffer [2.2 M KCl, 20 mM MgCl$_2$, 50 mM Tris-HCl (pH 7.5), 20 mM imidazole] followed by 10 column volumes of dH$_2$O. To recharge the column, 2.5 ml of 0.1 M NiSO$_4$$\cdot$6 H$_2$O was run through the column. The metal ion solution was removed by 5 column volumes of dH$_2$O and finally 5 column volumes of binding buffer.

Samples loaded on the HisTrap™ column consisted of cleared lysates (see Preparation of whole-cell extract) in TMK buffer. Lysates were centrifuged before loading on-column at 7,700 g for 10 minutes at 4°C using a Beckman JA 25.50 rotor. Supernatant was transferred to a clean 15 ml Falcon tube and brought up to about 10 ml final volume with binding buffer to reduce nonspecific binding. The lysate was applied to the column and washed by at least 10 column volumes of binding buffer. For gradient purifications, a linear gradient from 0 to 100% elution buffer 1 [2.2 M KCl, 20 mM MgCl$_2$, 50 mM Tris-HCl (pH 7.5), 0.5 M imidazole] was applied and 1 ml fractions were collected. The gradient was followed by an additional 10 ml of elution buffer which was also collected. For step purifications, the column washing was followed by 10 column volumes of elution buffer 1 or elution buffer 2 [2.2 M KCl, 20 mM MgCl$_2$, 50 mM Tris-HCl (pH 7.5), 80 mM imidazole]; 1 ml fractions were collected. Captured proteins were identified through staining of proteins separated by SDS-PAGE or Tricine-PAGE with coomassie blue and dot blot analysis.
Chapter 3: Determining that modification at position C34 of _H. volcanii_ tRNAMet elongator _in vivo_ occurs via the box C/D mechanism

Introduction

**Box C/D sRNA guided modification of archaeal tRNA**

Initial descriptions of box C/D complexes were derived from studies conducted on Eukarya, but a few years later, pull-down experiments in _Sulfolobus acidocaldarius_ established the presence of box C/D-guided methylation in the Archaea as well (Amiri, 1994; Kiss-Laszlo et al., 1996; Nicoloso et al., 1996; Omer et al., 2000; Tycowski et al., 1996). The first indications of sno-like proteins in Archaea came from the methanogens where fibrillarin-like proteins were noted in the Archaea (Amiri, 1994). Later, Omer et al. (2000) were able to show that there are homologues of both the protein and RNA components of box C/D guide sRNPs in Archaea. Subsequently, many examples of box C/D guide RNAs have been discovered in both major Kingdoms of Archaea. At present, only two specific box C/D guide sRNAs have been experimentally established in
the halophilic archaea, though our analysis and that of other groups has shown that there appear to be at least three (Bortolin et al., 2003; Singh et al., 2004; Weisel et al., 2010).

Box C/D sRNP complexes in Eukarya consist of four proteins and one snoRNA and this complex introduces a 2′-O-methylation in the target. The eukaryal proteins consist of the 15.5 kDa protein, Nop56, Nop58, and fibrillarin. In Archaea, each box C/D sRNP is composed of a single sRNA and two sets of three proteins: ribosomal protein L7ae, Nop5, and fibrillarin. The 15.5 kDa and L7ae proteins bind the RNA to stabilize the K-turn or K-loop structure and recruit the remaining proteins. The Nop5 protein in Archaea replaces the eukaryal Nop56/58 proteins. The Nop proteins interact with the K-turn complex and fibrillarin to form the complete sRNP complex. In both Eukarya and Archaea, fibrillarin functions as the methylase to introduce the 2′-O-methyl group to the target. These complexes in Archaea are capable of guiding the modification of two independent targets (Aittaleb et al., 2003; Reichow et al., 2007; Tran et al., 2003).

The basic structure of a box C/D guide RNA and its interaction with a target RNA is presented in Figure 3.1. These topics have been recently reviewed (Dennis and Omer, 2005; Reichow et al., 2007). The following discussion will focus on the archael systems. Briefly, the highly conserved C/C’ and D/D’ sequences provide the recognition for binding of ribosomal protein L7ae. These sequences are more loosely conserved among halophilic archaea than in Eukarya, allowing C to be substituted for U in the canonical sequence of UGAUGA (C/C’ box) and CUGA (D/D’ box); the substitution of C for U is likely the result of compensation for the relatively high G/C content of halophiles. Between the C box and D box is a sequence of approximately 9-12
Figure 3.1 Representation of an archaeal canonical box C/D sRNA and its interaction with an RNA target. Basic structure and orientation of the C/D and C’/D’ elements is given (left). The guide RNA basepairs with the target and the 2’-O-methyl group is introduced invariably five nt upstream of the D/D’ box (right).
Figure 3.1 Representation of an archaeal canonical box C/D sRNA and its interaction with an RNA target.
nucleotides which can form basepairs with the target RNA, thus facilitating the guide function of the sRNA. Invariably, the 2’-O-methyl group is introduced by fibrillarin into the target nucleotide five nucleotides upstream of the D/D’ box.

Assembly of archaeal box C/D sRNP complexes has been established through *in vitro* work. First, one L7ae protein binds the sRNA followed by another L7ae protein at the second site (Bortolin et al., 2003; Omer et al., 2002; Singh et al., 2004). These proteins stabilize the characteristic K-turn or K-loop in the guide RNA and facilitate the subsequent sequential binding of two Nop5 proteins and two fibrillarins. These complexes are sufficient to introduce 2’-O-methylations in the presence of the target and S-adenosyl methionine (SAM). Interestingly, extensive biochemical mapping has shown that a single copy of L7ae can recruit Nop5 and fibrillarin to both the C/D and C’/D’ regions of the guide RNA (Singh et al., 2008). This work has also demonstrated that both C/D and C’/D’ sequence regions need not be functional for guided methylation, though they can act sequentially to introduce modifications in a controlled order in the case of *H. volcanii* tRNATrp.

Alanine scanning and crystallography procedures have been used to map the interactions between the Nop5 and L7ae complex as well as the Nop5 and fibrillarin complex (Aittaleb et al., 2003; Hardin et al., 2009). Results from these studies support the model wherein the Nop5 component of a Nop5/fibrillarin heterodimer joins the sRNA already in complex with L7ae. The long-held model, which was reaffirmed by crystallography with a box C/D sRNA half-mer, contends that each sRNP complex consists of a single RNA with two copies of each protein (Ye et al., 2009). This view has
been challenged by single particle analysis of electron microscopy and size-based purification of *in vitro* assembled complexes, which indicate that archaeal box C/D complexes are dimeric (Bleichert et al., 2009). In the new model, two sRNAs are bridged together by the Nop5 proteins resulting in the sRNP complex consisting of two sRNAs and four copies of each protein. A crystal structure containing a bipartite C/D, C’/D’ sRNA bound by all three proteins has since been generated (Lin et al., 2011). Monomeric box C/D complexes were crystallized as well as dimeric, and both were found to be catalytically active. Additional analysis is necessary to assess how the dimeric form may function *in vivo*, but the current evidence is consistent with both models.

Among the best studied box C/D sRNAs in Archaea is the unusual case of the *H. volcanii* tRNATrp. This unique tRNA contains a large intron of over 100 nt that must be removed by the intron endonuclease for tRNA maturation (Daniels et al., 1985). It was established that the mature tRNATrp RNA had 2’-O-methylations at positions C32, C34, and U39 in the anticodon-stem loop region (Gupta, 1984). With the identification of sno RNA-mediated 2’-O-methylation of rRNAs in eukaryal cells, we, and others, proposed that sequences within the tRNATrp intron could function as guides for the 2’-O-methylations at positions C34 and U39 of this RNA (Armbruster, 1998; Bortolin et al., 2003). Additionally, basepairing between the target sequences and the guide sequences was noted as being more favorable in the intron-containing pre-tRNA suggesting that intron-containing tRNA may be the preferred substrate. Such a situation resulted in speculation that intron-containing pre-tRNATrp may be able to self-modify sequentially.
in cis following formation of the box C/D complex on the intron or that the free intron could form the sRNP complex and guide the modification in trans (Bortolin et al., 2003; Daniels et al., unpublished results). Other groups argued based on kinetic data that the C34 and U39 modifications are introduced sequentially, but only in trans by box C/D sRNP complexes formed with intron-containing pre-tRNATrp, the free tRNATrp intron, or the free, circularized tRNATrp intron (Singh et al., 2004). The mechanism of tRNATrp guide-target interaction has not been indisputably determined as of yet, though the C34 and U39 2’-O-methylations of either intron-containing pre-tRNATrp RNA and mature tRNATrp RNA have been conclusively demonstrated as being guided by the tRNATrp intron box C/D sRNP complex both in vitro and in vivo (Bortolin et al., 2003; Joardar et al., 2008; Singh et al., 2004). This presented an interesting question regarding the order of intron removal and 2’-O-methyl modification.

The H. volcanii tRNAMet (unless otherwise noted, tRNAMet in this document refers to the elongator tRNAMet of H. volcanii, not the initiator tRNAMet) presents a similar, yet simpler case as this tRNA has a 2’-O-methyl modification at position C34 and the tRNA is transcribed with an intron, but a box C/D guide sRNA is not encoded within the pre-tRNAMet intron (Gupta, 1984). Therefore, tRNAMet could only be modified by a box C/D sRNP complex in trans. As described below, an extended interaction between the intron-containing tRNAMet RNA and its cognate sRNP was also possible, providing further support for the model where modification could occur prior to intron removal. While this is not the only mechanism by which the 2’-O-methyl group could be introduced, few guide RNA-independent methylases have been identified in the
halophilic archaea. One such example would be the \(a\)Trm56 methylase which is responsible for the universal 2’-O-methylation of position C56 in archaeal tRNAs (Renalier et al., 2005). Based on the tRNATrp precedent, we hypothesized that the modification at position C34 is guided by a box C/D sRNP complex in \(H.\) \(v\)olcanii intron-containing pre-tRNAMet.

**Functional significance of 2’-O-methyl groups**

To this point, 2’-O-methyl modifications, either guided by sRNP complexes or introduced by independent methylases, have not been shown to be essential in any of the three domains (Benitez-Paez et al., 2010; Esguerra et al., 2008; Joardar et al., 2008). However, the loss of modifications has not always been without a discernible effect. In yeast and \(E.\) \(c\)oli, there is often not an observable growth defect in the strain unless the organism is challenged with stressful conditions or placed into competition with a modification-competent strain (Benitez-Paez et al., 2010; Esguerra et al., 2008). Ribosomal RNA 2’-O-methyl modifications have been implicated in proper folding of rRNA and ribosome function (Decatur and Fournier, 2002). In yeast strains depleted for some box C/D sno RNAs, deficiencies were noted in the -1 programmed ribosome frameshifting, but the strains were hyperaccurate at recognizing translation termination codons (Baxter-Roshek et al., 2007; Harger and Dinman, 2003; Karijolich et al., 2010). Other examples of the functional importance of 2’-O-methyl groups include providing a molecular marker of self versus non-self in eukaryal mRNA (Zust et al., 2011). These studies showed that a mouse coronavirus mutant lacking a 2’-O-methyltransferase was attenuated in wild type macrophages. Still, these modifications are to rRNA, and do not
sufficiently address the purpose behind introducing 2’-O-methyl groups to tRNA in the anticodon such as with *H. volcanii* tRNATrp and tRNAMet. Interestingly, a parallel study to the following work was recently conducted on the 2’-O-methylation of *E. coli* C34 of tRNALeuCAA (Benitez-Paez et al., 2010). While Bacteria lack the box C/D guide system, Benitez-Paez et al. (2010) identified the methylase, YibK, responsible for the modification and generated a knock-out strain. They observed, as we did in our study (see below), that there was no obvious defect in the position 34 modification depleted strain unless it was placed in competition with its non-mutant counterpart.

In the following work, we present evidence that 2’-O-methylation of C34 in tRNAMet of *H. volcanii* is guided by a box C/D sRNP complex. Deletion and complementation of the putative tRNAMet box C/D sRNA guide allowed for the predicted guide RNA’s presence/absence to be correlated with modification at position C34 in both mature tRNAMet and intron-containing pre-tRNAMet. We also show that the presence of this modification imparts an advantage to the cells during competition with the mutant strain.

**Results**

**Identification and conservation of the tRNAMet sRNA gene**

Modification patterns for a number of *H. volcanii* tRNAs have been established (Gupta, 1984). Similar to *H. volcanii* tRNATrp, the C34 of tRNAMet has a 2’-O-methyl modification, but unlike tRNATrp, no box C/D sRNA was apparent in the sequence of the tRNAMet intron. Therefore, if this modification is also guided by the box C/D mechanism, the tRNAMet sRNA must be located elsewhere within the genome.
A potential box C/D sRNA candidate was identified by our group and others. Using the pattern search tool fuzznuc (Rice et al., 2000), we located a candidate sRNA approximately 63 nucleotides from the 3’ end of the hypothetical gene Hvo_2294 in the intergenic region designated T2331 (Hartman et al., 2010). BLAST and ClustalW were used to identify and align the putative H. volcanii sRNA to sequences within the genomes of other halophiles (Altschul et al., 1990; Altschul et al., 1997; Chenna et al., 2003). This box C/D sRNA contained sequences complementary to the tRNAMet RNA. An 80 nucleotide region of the tRNAMet sRNA is conserved among all known halophile genomes with the exception of Natrialba magadii (Figure 3.2). Sequences related to the tRNAMet sRNA were not apparent in the N. magadii genome. This organism may lack 2’-O-methylation in tRNAMet at this position or may utilize a non-box C/D sRNA guided enzyme. Also, as shown in Figure 3.2, there is strong conservation of the box C/D motifs and absolute conservation of the basepairing to tRNAMet RNA. This property is most likely due to the considerable conservation of tRNAMet, itself (Figure 3.3). Interestingly, there is conservation of four additional residues at the exon-intron boundaries that would allow for extended pairing with the intron-containing pre-tRNAMet. The fourth nucleotide is the only variant, and is either an A or G in all identified tRNAMet sRNAs; however, in all halophile tRNAMets, the pairing nucleotide is uracil, thus maintaining the ability of the intron-containing pre-tRNA form an A-U or G-U pair. This observation suggests that the intron-containing pre-tRNA may be a substrate, if not the preferred substrate, for modification (Figure 3.4). Although
Figure 3.2 *H. volcanii* tRNAMet sRNA (DNA sequence) was aligned against BLASTN matches from the other halophilic Archaea using ClustalW. The C/C’ and D/D’ boxes are indicated as well as the conserved extended pairing (x-p) with the intron region of the pre-tRNAMet. Strain abbreviations are the same as those given in Table 1.1.
Figure 3.2 *H. volcanii* tRNA^Met^ sRNA (DNA sequence) was aligned against BLASTN matches from the other halophilic Archaea using ClustalW.
Figure 3.3 Alignment of haloarchaeal tRNAMet DNA sequences. *H. volcanii*

tRNAMet (DNA sequence) was aligned with other known halophilic Archaea tRNAMet RNAs using ClustalW. The universally conserved nucleotides in the anticodon stem and loop region, including the four additional base-pairs formed between the intron and the guide RNA are indicated by the bracket. The *H. volcanii* intron sequence is underlined. Strain abbreviations are the same as those given in Table 1.1.
Figure 3.3 Alignment of haloarcheal tRNAMet DNA sequences.
tRNAMet sRNA is conserved among the halophiles, its location in the genome is not conserved (Figure 3.5). We also noted that the tRNAMet sRNA has a second potential guide sequence element adjacent to the D’ region. These sequences are not highly conserved, and as of yet, a second target of this RNA has not been identified.

**Detection of the transcript**

As no obvious promoter was identified (see the sequence in Figure 3.6), we confirmed that the candidate tRNAMet sRNA was transcribed using Northern analysis, which was performed with total RNA isolated from *H. volcanii* using the metsno oligonucleotide as the probe. As seen in Figure 3.6, a band hybridizing to the tRNAMet sRNA is clearly detectable in the total RNA. To determine the approximate size of tRNAMet sRNA, the membrane was stripped of the metsno probe and subjected to hybridization with the metE1 probe. The metE1 probe resulted in signals from RNAs corresponding to intron-containing pre-tRNAMet, 148 nt, and mature tRNAMet, 77 nt. Size estimates based on these markers and others (not shown) suggest that tRNAMet sRNA is approximately 125 nt. Our microarray studies (described in Sabag-Daigle 2009) agree that the transcript is detectable since the region corresponding to the tRNAMet sRNA T2331 showed expression. A related transcript was also detectable, though smaller in size, in total RNA isolated from *H.*-NRC-1 using the same metsno probe (Figure 3.7).
Figure 3.4 Pairing of the box C/D sRNA with *H. volcanii* tRNA\textsubscript{Met} RNA. tRNA\textsubscript{Met} sRNA is capable of base-pairing with both intron-containing pre-tRNA\textsubscript{Met} (left) and mature tRNA\textsubscript{Met} (right). The C and D box sequences are noted. The tRNA is shown in black while the guide RNA is in blue. Intron-containing pre-tRNA\textsubscript{Met} is capable of forming four more base-pairs with the guide RNA than mature tRNA\textsubscript{Met} due to the guide RNA pairing with the intron.
Figure 3.4 Pairing of the box C/D sRNA with *H. volcanii* tRNAMet RNA.
Figure 3.5 Gene organization of putative tRNAMet sRNAs in haloarchaeal genomes.

FASTA searches in KEGG were used to identify sequences related to the *H. volcanii* tRNAMet sRNA. Boundaries of the tRNAMet sRNA in each case were defined as shown in Figure 3.2. The sequence viewer tool from NCBI was used to construct the gene neighborhood of tRNAMet sRNA in each organism. Arrows indicate the direction of transcription. Strain abbreviations are the same as those given in Table 1.1.
Figure 3.5 Gene organization of putative tRNAMet sRNAs in haloarchaeal genomes.
Figure 3.6 Representation of the tRNAMet sRNA location, sequence, and evidence of expression.  A: Location of tRNAMet sRNA with its sequence given below. The region conserved among halophiles is underlined while the box C/D elements are in red type.  B: Box C/D structure of the *H. volcanii* tRNAMet sRNA. Dotted lines indicate possible G-A pairs formed in K-turn/K-loop structures.  C: Northern analysis of total RNA isolated from *H. volcanii* DS70-ASD40 using either metsno (left) or metE1 (right) as probes for hybridization.
Figure 3.6 Representation of the tRNAMet sRNA location, sequence, and evidence of expression.
Deletion of the tRNAMet sRNA gene from the chromosome

The pop-in pop-out strategy described in Chapter 2 (Figures 2.1 and 2.2) was used to delete the putative tRNAMet sRNA gene from the *H. volcanii* chromosome. Recombinants in which both *pyrF* and tRNAMet sRNA have been deleted were identified by PCR and Northern analysis (Figure 3.8). The strain deleted for the tRNAMet sRNA region was designated DS70-ASD40-sno.

Growth of the deletion strain

DS70-ASD40-sno was cultured in both CDM and CX medium in order to determine if growth properties of the mutant strain differed from the wild type strain. Under either culturing condition, there was no observable difference in the growth rate of the deletion strain compared to the wild type strain DS70-ASD40 (data not shown). Since little has been established regarding the function and significance of these modifications in tRNAs, we investigated whether there was any advantage to having this modification by competing the wild type and tRNAMet sRNA deletion strains against each other in CX broth culture. Complex medium was chosen for the cultures to provide the best possible growth conditions for both strains. Two experiments were maintained; they derived from the same starter cultures. The strains were combined in equal amounts of cells based on A550 and were either passed once daily to maintain the cells in exponential growth or once every 72 hours to investigate whether reaching stationary phase could reveal a growth advantage. Cell samples were taken at each culture passage and stored for use as template for PCR analysis (See Chapter 2, Culturing of *H. volcanii* strains in competition). The same primers used to verify the deletion (MSDX-F and
Figure 3.7 Sequence of the tRNAMet sRNA recovery module and expression of *H. volcanii* tRNAMet sRNA in wild type and mutant strains. A: Sequence of the tRNAMet sRNA recovery module. The *H. volcanii* tRNALys promoter sequence is given in blue type, the recovery sequence is in black with the conserved region underlined (box C/D elements are as indicated in Figure 3.6 A), and a synthetic transcription terminator is in green type. B: Northern hybridization using metsno as a probe against total RNA isolated from *H. volcanii* wild type and mutant strains grown in Complex Medium (CX) or Completely Defined Medium (CDM). Strains are indicated as follows: DS70-ASD40 (wt), deletion, DS70ΔpyrFΔsno, complementation, DS70-snorec, or *Halobacterium* sp. NRC-1 (hal).
Figure 3.7 Sequence of the tRNAMet sRNA recovery module and expression of \textit{H. volcanii} tRNAMet sRNA in wild type and mutant strains.
Figure 3.8 Evidence of tRNAMet sRNA sequence deletion from the genome.  

A: Schematic representation of the deleted region. The approximate locations of the primers used for PCR verification are designated as arrows. B: PCR from the wild type DS70-ASD40 (wt) and deletion DS70ΔpyrFΔsno (Δ) strains using primers MSDX-F and MSDX-R. C: Northern hybridizations against total RNA isolated from each strain using probe metsno (top) or 7S rev (bottom).
Figure 3.8 Evidence of tRNAMet sRNA sequence deletion from the genome.
MSDX-R) were used to determine the relative proportions of the two genomes in the cell mixture.

According to the results of PCR, the wild type strain out-competed the mutant (Figure 3.9). Over the course of 24 days of maintaining cells in exponential growth, the band corresponding to the mutant strain steadily decreased and was barely detectable at 24 days. It also appears that growth into stationary phase has a more pronounced preference for the wild type strain as the mutant was no longer readily detectable after only 15 days. While it remains unclear as to what sort of advantage tRNAMet sRNA confers, there is an impact on the cell’s success against a comparable strain related either to modification of tRNAMet at C34 or the unknown second target of the sRNA.

**Complementation of the tRNAMet sRNA gene**

The region deleted in the DS70-ASD40-sno strain was fused to the previously characterized *H. volcanii* tRNAlys promoter as described in Chapter 2 (Palmer and Daniels, 1995). Notably, tRNAMet sRNA is transcribed at much higher levels in DS70-ASD40-snoec, and the larger primary transcript resultant from the promoter fusion is readily observable (Figure 3.7). We also noted that a smaller species of approximately 125 nt was also present suggesting that the primary tRNAMet sRNA transcript from the complementation strain underwent processing.

**Impact of deleting the tRNAMet sRNA gene on 2’-O-methylation at position C34 of tRNAMet**

Having established tRNAMet sRNA deletion and complementation strains, we analyzed the impact on C34 of tRNAMet utilizing the RNase H method for detection of
Figure 3.9 PCR analysis of the competition between the wild type DS70-ASD40 (wt) and deletion DS70-ASD40-sno (Δ) strains. The number above the lane corresponds to the number of passages for the culture; the same zero point was used to start both cultures. A: culture was passed once daily to maintain exponential growth phase throughout the experiment. B: culture was passed every third day to ensure that it reached stationary phase.
Figure 3.9 PCR analysis of the competition between the wild type DS70-ASD40 (wt) and deletion DS70-ASD40- sno (Δ) strains.
2’-O-methyl groups (Yu et al., 1997). This method allows for precise targeting of the residue of interest in total RNA. Briefly, an oligonucleotide complimentary to the sequence of tRNAMet consisting of a 2’-O-methyl RNA backbone with four DNA bases was designed. The oligonucleotide is then hybridized to the target in a sample of total RNA. The DNA bases direct RNaseH cleavage to the residue of interest, but the enzyme can only cleave if there is no 2’-O-methyl group on the target residue. This oligonucleotide was designed in a manner such that the 5’ end of the oligonucleotide begins adjacent to the splice junction, which enabled us to simultaneously test whether or not the intron-containing pre-tRNAMet and the mature tRNAMet were methylated (Figure 2.3). Results were assessed with Northern analysis following the RNase H cleavage reaction. As evidenced by the cleavage products there is C34-unmodified spliced tRNA in total RNA isolated from the deletion mutant (Figure 3.10). Using the unmodified T7 transcripts, RNase H cleavage results in cleavage of both intron-containing pre-tRNAMet and tRNAMet. Estimates, based on the hybridization signals indicated that 95-96% and 79% of the intron-containing pre-tRNAMet and tRNAMet were cleaved, respectively (Table 3.1). Comparison of the cleavage patterns of total RNA isolated from the wild type (Figure 3.10 B, lane wt) and the deletion strain (Figure 3.10 B, lane ∆) indicate that the loss of the tRNAMet sRNA is accompanied by an increase in cleaved intron-containing pre-tRNAMet and mature tRNAMet RNAs. The relative amounts of the cleaved RNA species are decreased when the tRNAMet sRNA is re-introduced into the deletion strain (Figure 3.10 B, lane C). Previous studies were
Figure 3.10 tRNAMet sRNA guides modification of both the intron-containing pre-tRNAMet (pre-met) and mature tRNAMet (met) RNAs. A: The RNase H assay oligonucleotide consists of a 2'-O-methyl backbone with four DNA bases, shown in bold, that target the cleavage by RNaseH to the specific nucleotide of interest. B: The RNase H cleavage assay was performed with \textit{in vitro} synthesized RNA (T7 pre-met and T7 met) and total RNA isolated from either wild type DS70-ASD40 (wt), deletion DS70ΔpyrFΔsno (Δ), or tRNAMet sRNA complementation DS70-snorec (c) strain in the presence (left panel) or absence (right panel) of RNase H. The apparent size discrepancy between the \textit{in vitro} RNA and cellular RNA is due to the primer used to synthesize the T7 transcripts adding additional nucleotides. RNAs were detected by Northern analysis using metE1 as a probe.
Figure 3.10 tRNAMet sRNA guides modification of both the intron-containing pre-tRNAMet (pre-met) and mature tRNAMet (met) RNAs.
Table 3.1 Relative cleavage of intron-containing pre-tRNAMet and mature tRNAMet derived from the RNase H assay shown in Figure 3.10.

<table>
<thead>
<tr>
<th>Source of RNA</th>
<th>% Cleavage of pre-tRNAMet&lt;sup&gt;a&lt;/sup&gt;</th>
<th>% Cleavage of spliced tRNAMet</th>
</tr>
</thead>
<tbody>
<tr>
<td>T7 Transcription</td>
<td>95-96</td>
<td>79</td>
</tr>
<tr>
<td>DS70-ASD40</td>
<td>14-16</td>
<td>&lt;1</td>
</tr>
<tr>
<td>DS70-ASD40-sno</td>
<td>68-74</td>
<td>35</td>
</tr>
<tr>
<td>DS70-ASD40-snorec</td>
<td>6</td>
<td>9</td>
</tr>
</tbody>
</table>

<sup>a</sup> The range of values is provided for pre-tRNAMet because it was calculated from hybridizations with both metE2 and metIn used as probes. The metIn probe does not detect mature tRNAMet.
successful in confirming that the box C/D sRNA in the *H. volcanii* tRNATrp intron is responsible for introducing the modifications it appears to guide, but the system used was an *H. volcanii* mutant strain in which the native intron-containing tRNATrp gene had been replaced by an intron-less version thereby preventing *in vivo* verification that intron-containing pre-tRNATrp is a substrate for modification (Joardar et al., 2008). Limitations in the RNase H assay left us unable to determine if pre-tRNAMet is the preferred or only substrate for modification by a box C/D sRNP, though we demonstrate here for the first time *in vivo* that intron-containing *H. volcanii* pre-tRNAs are substrates of the box C/D sRNP complex.

**Impact of tRNAMet sRNA gene deletion on the tRNAMet pool**

Having observed that the precursor tRNA was modified, we attempted to determine whether there was an impact on splicing. However, visual inspection of Northern hybridizations showed no significant differences between the levels of pre-tRNAMet and spliced tRNAMet in total RNA isolated from the deletion mutant and the wild type. Efforts to quantify the ratio of pre-tRNAMet to spliced tRNAMet resulted in values ranging between 60-90% spliced tRNAMet versus 50-90% spliced tRNAMet for the deletion mutant (data not shown). The slight difference on the low end of the range is likely due to the considerable biological or sample variation as opposed to representing a deficiency in splicing efficiency. Modification of C34 of intron-containing pre-tRNAMet is not necessary for splicing, and there is no obvious impact on the pre-tRNA to spliced tRNA ratio.
Discussion

tRNAMet sRNA is not essential for viability or efficient growth

While the machinery for box C/D sRNP guided modifications is clearly present within the Archaea, there have are few examples of in vivo studies of guide sRNAs in the Archaea, and no studies have yet demonstrated that a single box C/D guide is essential for growth. Furthermore, the function of these modifications in tRNA anticodon loops is not well-understood.

In this study, a putative box C/D guide sRNA targeting the tRNAMet RNA was identified in the genome of *H. volcanii*. This RNA is transcribed, and based on its sequence and the established modification pattern of tRNAMet RNA, the RNA is predicted to act as a guide for 2’-O-methylation at position C34 of this tRNA. We were able to construct a mutant strain where the tRNAMet sRNA was deleted from the chromosome, establishing that this gene is not essential. Wild type and mutant strains exhibited similar growth patterns in CDM and CX media.

The ability of cells to grow without the tRNAMet sRNA was not unexpected. Previous reports established that the tRNATrp intron, which functions as a box C/D sRNA directing 2’-O-methylation of tRNATrp at C34 and U39 and the *E. coli* YibK enzyme responsible for 2’-O-methylation of C34 of tRNALeuCAA are also non-essential (Benitez-Paez et al., 2010; Joardar et al., 2008).

Curiously, when forced into competition against the wild type, the tRNAMet sRNA deletion mutant was lost from the population in the culture over the course of time.
regardless whether cells were maintained in balanced growth or allowed to reach stationary phase.

Under these conditions it is not known whether *H. volcanii* cells can exchange DNA with each other (Rosenshine et al., 1989), but irrespective of that uncertainty, the only survivors in the culture give a product corresponding to the wild type with the sRNA intact. Based on the PCR results, the deletion strain was more successful when the culture was maintained in exponential growth, lasting about 20 days. The pressures of stationary phase made the deficiency in the deletion strain manifest in about 15 days suggesting that the role of tRNAMet sRNA is more significant when the cells reach this phase.

Having the ability to modify tRNAMet at C34 confers a growth advantage, though why this is the case is not clear. The modification could be a stabilizing factor in the tRNA molecule to aid in coping with stressful conditions. It also remains possible that the undiscovered potential second target of the sRNA could be the impetus behind the success of the wild type against the deletion mutant. However, similar work in *E. coli* suggests that the loss of C34 modification alone may be the reason for the deletion strain’s inability to successfully compete with the wild type (Benitez-Paez et al., 2010). Regardless the reason, the advantage is observable, but small under culturing conditions used in this study as it was only realized when the strain was forced into competition.

**Complementation of tRNAMet sRNA yields a new product**

Expression of tRNAMet sRNA in the complementation strain yielded two distinct products, both at higher levels than the wild type level of tRNAMet sRNA. The higher
levels of expression can be attributed to the fusion of tRNAMet sRNA to the constitutive
*H. volcanii* tRNALys promoter rather than the unidentified native promoter. The
presence of two distinct products, both at higher than wild type levels, suggests that the
tRNAMet sRNA primary transcript from the tRNALys promoter is processed. An
alternative explanation would be that the native promoter of tRNAMet sRNA is included
in the recovery sequence and there is transcription from both the tRNALys promoter and
the native promoter hence resulting in two products. This seems unlikely as a sole
explanation though, as the vector containing the tRNAMet sRNA is non-replicative in *H.
volcanii* and has integrated into the chromosome. Such conditions would result in wild
type levels of tRNAMet sRNA at the ~125 nt level. Therefore, while it remains possible
that both promoters drive transcription of tRNAMet sRNA, there is a processing event
that trims the primary transcript of the tRNALys promoter to the predominant
approximately 125 nt RNA. The mechanism of this processing event is unknown at this
time, though we have determined that the intron endonuclease does not appear to have
any role in this process (data not shown). While transcripts were readily detectable,
attends to map the 5’ end by primer extension were unsuccessful. Additionally, our
collaborators at Capital University (Jens Hemminsen and research group) were unable to
resolve this question through primer-walking PCR or cloning the RNA and sequencing it;
the clones produced variable ends.

**tRNAMet sRNA guides the C34 modification of *H. volcanii* tRNAMet in vivo**

Considering *H. volcanii* tRNATrp as a precedent, it was expected that, like
tRNATrp which contains its own guide sequence within its intron, the C34 2’-O-
methyltransferase of tRNAMet, another intron-containing tRNA, would be guided by a box C/D sRNP complex despite the guide sequence being absent from the tRNAMet intron. Data from the RNase H assay support this proposal. In the RNase H assay cleavage of the intron-containing pre-tRNAMet and mature tRNAMet can only occur if the target residue is not 2'-O-methylated and, only intron-containing pre-tRNA and mature tRNA isolated from the deletion strain showed significant levels of cleavage. Intron-containing pre-tRNAMet and mature tRNAMet from the tRNAMet sRNA deletion strain were both cleaved at lower levels than the equivalent T7 transcripts. The mature tRNAMet in particular exhibited a much lower level of cleavage than the precursor (35% vs 68-74%) suggesting that the mature tRNA structure, other modifications in the mature tRNAMet inhibit the assay, or that there is a redundant mechanism to introduce a 2'-O-methyl group to C34. Should the latter exist, it appears to favor the mature tRNA as a substrate. The wild type and complementation strains both largely blocked cleavage in the presence of RNase H indicating that both intron-containing pre-tRNAMet and mature tRNAMet were 2'-O-methylated at position C34. Percentages of pre-tRNA and spliced tRNA cleavage were comparable to the negative controls with the exception of the pre-tRNAMet isolated from the wild type. Approximately 14-16% of this pre-tRNAMet appeared to cut; this suggests that modification is not directly tied to transcription and that there is a background pool of pre-tRNAMet that has not encountered the box C/D sRNP complex yet. Alternatively, these C34 unmodified intron-containing pre-tRNAMet RNAs may be modified post-splicing by the sRNP complex or an unidentified redundant pathway. Intron-containing pre-tRNAMet RNA cleavage levels for the complementation
strain remain consistent with the controls. This can be explained by tRNAMet sRNA
being present in the cell at much higher than normal levels since it is fused to the H.
volcanii tRNALys promoter in the complementation strain. More tRNAMet sRNA
transcript likely results in more box C/D sRNP complexes thus decreasing the likelihood
that intron-containing pre-tRNAMet will be unmodified prior to splicing. Based on T7
transcript controls used in this assay, there was not 100% efficiency of cleavage (95-96%
of T7 intron-containing pre-tRNAMet, 79% of mature tRNAMet) suggesting that there
may be some inherent structural barriers especially with the mature tRNA which may
explain the lower levels of cleavage in the deletion strain as discussed above. We cannot
exclude the possibility that there is a redundant mechanism responsible for modification
of C34 in tRNAMet, but based on the tRNATrp intron deletion mutant, a redundant
pathway may be unlikely since no 2’-O-methylation of tRNATrp was detectable in the
tRNATrp intron deletion strain (Joardar et al., 2008).

Through the use of this assay, we were able to indirectly show that not only is the
mature tRNAMet modified by tRNAMet sRNA, but intron-containing pre-tRNAMet is
also likely to be modified by it. Figure 3.10 clearly shows that there is an observable
cleavage of intron-containing pre-tRNAMet in total RNA isolated from the tRNAMet
sRNA deletion strain in the presence of RNase H. Unfortunately, we are limited to
stating that pre-tRNAMet is modified at C34 in vivo, but we cannot prove that pre-
tRNAMet must be modified by the box C/D sRNP complex, though the near complete
blockage of RNase H activity on the RNA isolated from the wild type and
Figure 3.11 Schematic model of tRNAMet modification and splicing. While either pathway for maturation is possible, the data presented in this study suggests that the upper path of pre-tRNA modification is the more common route as designated by the thicker arrows.
Figure 3.11 Schematic model of tRNAMet modification and splicing.
complementation strains does suggest that the vast majority of pre-tRNAMet is modified (Figures 3.10 B and 3.11).

**Modification at position C34 occurs in pre-tRNAMet, but is not required for efficient splicing *in vivo***

The occurrence of box C/D sRNP modification of tRNA has, to this point, involved primarily tRNAs containing introns. In the case of both tRNATrp and tRNAMet, the intron-containing pre-tRNA is a substrate for the sRNP and may be the preferred substrate.

In spite of extended base-pairing interaction with intron-containing pre-tRNAMet and intron-containing pre-tRNAMet modification detected *in vivo*, C34 modification is not required for splicing and does not have an observable impact on splicing. There appears to be nothing to prevent spliced tRNAMet from box C/D modification, but based on the high levels of pre-tRNAMet modification, spliced tRNAMet as the substrate is likely a secondary pathway for tRNAMet maturation.
Chapter 4: Capture and analysis of L7ae-containing complexes in *H. volcanii*

Introduction

Ribosomal protein L7ae is an integral component of archaeal box C/D sRNP complexes. In addition to its role in box C/D sRNP complexes and its function in the large ribosomal subunits, L7ae is a requisite component of other sRNP complexes as well. L7ae binds RNA directly at defined sequence motifs referred to as Kink or K-turns and K-loops (Klein et al., 2001; Rozhdestvensky et al., 2003). The signature amino acids involved in RNA binding have been mapped as five highly conserved amino acids (Gagnon et al., 2010). L7ae stabilizes a structural conformation in the RNA fold and as it requires no other protein to bind, L7ae can be a part of any sRNP complex wherein the RNA has a K-turn or K-loop motif. At present, three other archaeal sRNP complexes shown in Figure 4.1 have been described as containing or likely containing L7ae: box H/ACA sRNP complexes, RNase P holoenzyme, and Signal Recognition Particle (SRP) (Cho et al., 2010; Rozhdestvensky et al., 2003; Zago et al., 2005).

Structures of L7ae binding RNA alone or in a complex with other proteins have been solved using a variety of techniques (Bleichert et al., 2009; Duan et al., 2009; Liang et al., 2009; Lin et al., 2011; Suryadi et al., 2005; Turner et al., 2005; Ye et al., 2009).
Figure 4.1 Predicted L7ae sRNP complexes in *H. volcanii*. The molecular weight and pI of each protein component is given along with the length (in nt) of RNA components.
Figure 4.1 Predicted L7ae sRNP complexes in *H. volcanii.*
The conformation of the RNA when bound by L7ae allows the appropriate proteins to subsequently bind to the RNA scaffold. In the instance of box C/D sRNP complexes, L7ae binds to the RNA first; the presence of the K-turn and L7ae then recruits Nop5 or the Nop5/Fibrillarin heterodimer to the complex (Hardin et al., 2009). Therefore in this complex, and based on current evidence all other complexes it is a part of, L7ae is a structural component that stabilizes the conformation of the RNA to recruit other requisite proteins. In this manner, L7ae can be necessary for sRNP complex formation or can improve the rate of catalysis, as in the case for RNase P (Cho et al., 2010; Fukuhara et al., 2006), of an sRNP complex without having a direct role in catalysis or guide/target interaction.

As mentioned above, L7ae is not only necessary for the assembly of box C/D complexes, but also for box H/ACA sRNP complexes. Box H/ACA sRNP complexes comprise the second major category of RNA-guided modification machines in the Archaea. The box H/ACA sRNP complex uses an sRNA component to specifically target a uridine for conversion to pseudouridine (Baker et al., 2005; Charpentier et al., 2005; Ganot et al., 1997a; Ganot et al., 1997b). In other systems, such as yeast, some highly conserved pseudouridine modifications in rRNA are required for normal growth, especially under stress conditions (King et al., 2003). Pseudouridines have been discovered in locations such as in molecular bridges that allow large and small ribosome subunits to stack as well as in the tRNA-binding regions (Decatur and Fournier, 2002). The chemistry of pseudouridines results in a more effective way to stabilize and improve the interaction (Agris, 1996; Davis, 1995). Mutants deficient in some of these
pseudouridines were found to have problems with ribosome assembly, tRNA binding, and experienced a decreased rate of translation (Badis et al., 2003; Karijolich et al., 2010; Liang et al., 2007). Therefore, while box H/ACA guided modifications are individually dispensable, they are collectively necessary and provide considerable benefits to the organism.

RNA-guided modification is not the only way to introduce pseudouridines. Examples including the archaeal Pus10 protein provide instances in which an enzyme can catalyze modification independent of an RNA guide (Gurha and Gupta, 2008; Watanabe and Gray, 2000). The pseudouridylase component of box H/ACA sRNP complexes, Cbf5, has even been shown to function as an RNA guide-independent enzyme in addition to its role in box H/ACA guided modification (Gurha et al., 2007; Watanabe and Gray, 2000). Curiously, not all Archaea introduce conserved modifications with the same machinery. In a case where archaeal Pus7 appeared insufficient to introduce the tRNA\text{Tyr} pseudouridine at position 35, the organism was found to have a box H/ACA RNA capable of guiding the modification (Muller et al., 2009).

The sRNA component of box H/ACA sRNP complexes has relatively loose sequence conservation and is defined more in terms of its hairpin structure with an open pseudouridine pocket between two helices. Despite the less conserved sequence, there are two commonly found sequence motifs flanking the hairpin structure, the box H (ANANNA) on the 5’ side and the box ACA motif located at the 3’ end of the hairpin (Reichow et al., 2007). The components of the box H/ACA complex are the sRNA and one of each protein per hairpin in the sRNA: L7ae, Nop10, Gar1, and Cbf5. Furthermore,
Examples have been found in this study and elsewhere of box H/ACA sRNAs that contain multiple hairpin structures enabling the same sRNA to guide modification at multiple targets (Figure 4.2).

Considerable efforts in structure determination, primarily through crystallography, have resulted in several high-resolution structures of the full box H/ACA sRNP complex granting new insight into the assembly and role of each component (Duan et al., 2009; Li and Ye, 2006; Liang et al., 2009; Manival et al., 2006; Rashid et al., 2006). Coupled with extensive biochemical and fluorescence mapping experiments, structure and function of box H/ACA components are quite well-defined (Liang et al., 2008; Youssef et al., 2007). Ribosomal protein L7ae binds at the top of the hairpin and stabilizes a K-turn. The pseudouridine pocket contains the guide sequence that will basepair with the target region while leaving the target uridine and the nucleotide 3’ of it unpaired in the center of the pocket. Nop10, Gar1, and Cbf5 form a stable complex independently of the sRNA. The function of Nop10 appears to be stabilizing the catalytic core of Cbf5 while Gar1 has been implicated in controlling the ability of Cbf5 to bind the sRNA, although Cbf5 is capable of binding the RNA on its own (Kiss et al., 2010; Youssef et al., 2007). Box H/ACA complexes can still function at an efficient rate of catalysis in vitro in the absence of Gar1, provided the remaining three proteins and the sRNA are present, but optimal activity is achieved when all components are present (Charpentier et al., 2005).
**Figure 4.2 Potential box H/ACA RNA in *H. volcanii*.** A: Northern analysis demonstrating detection of the putative box H/ACA RNA with probe T2743-2. Markers are shown to the right. B: The sequence of the RNA allows for the formation of two independent H/ACA hairpins; the general locations of the box H/ACA sRNP complex proteins are given. Dotted lines indicate possible G-A pairs formed in K-turn/K-loop structures.
Figure 4.2 Potential box H/ACA RNA in *H. volcanii*
While sRNA guided modification complexes have long been known to require L7ae, recent work has shown that L7ae is a fifth protein component of archaeal RNase P (Cho et al., 2010; Fukuhara et al., 2006). RNase P is responsible for removing the 5’ leader sequence of tRNA transcripts in all Domains of life. At present, the only characterized organism to proliferate naturally in the absence of RNase P activity is the archaeon *Nanoarchaeum equitans* (Randau et al., 2008). *Aquifex aeolicus* and the genus *Pyrobaculum* are other organisms suggested to exist without RNase P; however, a shortened RNase P has been discovered in *Pyrobaculum* and there is RNase P-like activity detectable in *A. aeolicus* although the cellular machinery for this activity remains elusive (Lai et al., 2010). The archaeal RNase P ribozyme complex was defined as the RNase P RNA and four proteins: Pop5, Rpp21, Rpp29, and Rpp30 (Hall and Brown, 2002; Tsai et al., 2006). Studies had demonstrated improved activity of *Pyrococcus horikoshii* RNase P *in vitro* when L7ae was added to the reaction (Fukuhara et al., 2006). Subsequent work by Cho et al. (2010) has conclusively shown that L7ae specifically binds to RNase P RNA and has mapped to likely binding location in the P12 region of RNase P RNA (Figures 4.3). The experimental evidence was obtained by using ion-exchange chromatography to partially purify *M. maripaludis* RNase P activity from ribosome-depleted S100 fractions. Western blotting for L7ae on peak activity fractions revealed the co-elution of L7ae with peak RNase P activity; however, L7ae is not required for catalysis *in vitro*. This finding is not unexpected as multiple studies have now proven the RNase P RNA to be the catalytic component in all three Domains suggesting that L7ae’s role is simply to sustain a more favorable conformation for
catalysis or recruitment of unidentified proteins. By manipulating reaction conditions, the naked RNase P RNA of Bacteria, Archaea, and Eukarya have been proven to be sufficient for catalyzing the removal of the 5’ leader sequence of tRNA (Guerrier-Takada et al., 1983; Kikovska et al., 2007; Pannucci et al., 1999). Furthermore, prior to the results mentioned above, *in vitro* studies of RNase P have been conducted in the absence of L7ae thereby demonstrating L7ae to not be an absolute requirement for ribozyme activity.

Another complex implicated in a previous study and in this study as containing ribosomal protein L7ae is the archaeal Signal Recognition Particle (SRP) complex. Archaeal SRP is responsible for delivering translating ribosomes to the membrane so that proteins can be translocated across the membrane (Calo and Eichler, 2011). Archaeal SRP, which is closely related to the eukaryal complex, has been defined as consisting of 7S RNA and two proteins: SRP54 and SRP19; both of these proteins have homologues in the eukaryal complex (Bhuiyan et al., 2000; Diener and Wilson, 2000; Tozik et al., 2002). What has been noted is that proteins are absent in archaeal genomes and based on the eukaryal complex a portion of the RNA would remain protein-free in the archaeal complex. It seems likely that additional proteins are associated with the archaeal SRP complex. The *H. volcanii* SRP was reconstituted to show specific interactions between the 7S RNA, and the two proteins (Tozik et al., 2002); however, the native complex has not been purified from these cells. We have noted that the 7S RNA sequence shows a highly conserved region exhibiting a characteristic recognition sequence for binding of L7ae suggesting that L7ae may be a component of the SRP complex (Figure 4.4 B).
Figure 4.3 Structure of RNase P RNA (RPR) from *P. horikoshii* and *H. volcanii* with probable L7ae binding sites indicated. A: *P. horikoshii* RPR structure with the relative binding locations of L7ae indicated by grey ovals as determined experimentally by Fukuhara et al 2006. B: *H. volcanii* RPR structure with putative L7ae binding sites, based on the *P. horikoshii* data, indicated by arrows.
Figure 4.3 Structure of RNase P RNA (RPR) from *P. horikoshii* and *H. volcanii* with probable L7ae binding sites indicated.


4.4 Conserved sequence elements of haloarchaeal 7S RNA and the potential binding sites for L7ae. A: Partial sequence representation illustrating conserved sequences of helix 5 of *H. volcanii* 7S RNA. The structure is based on Tozik et al. (2002). The established archaeal SRP proteins are indicated at their relative positions based on Calo and Eichler (2011). The likely binding location of L7ae is indicated based on our sequence analysis and on the prediction from *Sulfolobus* (Zago et al. 2005). B: Conservation of the putative L7ae binding site among the haloarchaeal 7S RNAs (DNA sequence shown). The 5’ loop and 3’ loop sequences are underlined.
Figure 4.4 Conserved sequence elements of haloarchaeal 7S RNA and the potential binding sites for L7ae.
As the SRP proteins are known to bind elsewhere on the RNA (Figure 4.4 A), this sequence should be available for L7ae interaction. Efforts to pull-down box C/D complexes from Sulfolobus using antibodies targeted to the box C/D proteins resulted in the capture of 7S RNA leading others to verify that L7ae will bind 7S RNA (Zago et al., 2005). The region of L7ae binding to 7S RNA was mapped by Zago et al. (2005) to the conserved region as indicated in Figure 4.4.

Given L7ae’s already wide distribution among sRNP complexes, we were interested in determining if these and novel complexes could be isolated by overexpression of affinity-tagged wild type L7ae protein. Previous work has demonstrated that the H. volcanii hutU promoter can be used to induce expression of a reporter gene in the presence of 10 mM urocanate (Sabag-Daigle, 2009). We reasoned that the combination of the reported stability of L7ae-bound complexes and this expression system would allow us to isolate L7ae containing complexes using nickel affinity chromatography; this would then provide a tool to examine the pool of native L7ae-containing complexes assembled in vivo with sufficient stability to remain intact throughout the purification. Having captured such complexes, the diversity of RNAs and proteins associated with L7ae sRNPs could be explored. Not all native L7ae-containing complexes have been fully characterized; for instance, RNase P complexity increases across the Domains from having only one protein in Bacteria to at least five in Archaea and nine in Eukarya. This complexity, along with the lack of compartmentalization in Archaea, allows for the possibility that there may be a wider substrate range for RNase P
and that there may yet be additional proteins associated with the native complex. This work also provides a method for expression of archaeal proteins in the native host.

**Results**

**Identification of the T2743 box H/ACA sRNA**

Analysis of *H. volcanii* tiled microarray experiments described in (Sabag-Daigle, 2009) revealed detectable expression from an area in the genome deemed intergenic, T2743. Subsequent Northern analysis showed that a transcript derived from this intergenic region was readily detectable (Figure 4.2 A).

Close inspection of the sequence contained by this region led to the discovery of the hallmark box H/ACA motifs. This putative box H/ACA sRNA was found to possess two hairpins and it is conserved among the halophile genomes as seen in Figure 4.2 B (Altschul et al., 1990; Altschul et al., 1997; Bailey and Elkan, 1994). The program mFold was used to assist in drawing the probable box H/ACA structure (Markham and Zuker, 2008). A comparison of the second hairpin against the known pseudouridylation map of the *H. marismortui* 23S rRNA (Del Campo et al., 2005) reveals that it is capable of basepairing with the 23S rRNA and guiding both pseudouridylations at the equivalent positions of 1956 and 1958 with the single hairpin (Figure 4.5). No target has been determined yet for the first hairpin structure of T2743.

**Construction and verification of the affinity-tagged wild type L7ae overexpression strain DS70-ASD41-L6H**

Wild type *H. volcanii* L7ae was cloned with a C-terminal hexahistidine tag into the *H. volcanii* replicative expression vector pARS-L6H (see Figure 2.1). Expression of
Figure 4.5 Conservation of the second hairpin of the *H. volcanii* RNA originating from intergenic region T2743. Based on sequence analysis, the second hairpin appears to guide the pseudouridylation at both positions 1956 and 1958 (*H. volcanii* numbering) of the 23S rRNA. Logos demonstrate the conservation of the sequence elements in this hairpin among the sequenced haloarchaeal genomes. Nucleotides universally conserved among the haloarchaeal are given in blue while the 23S rRNA target is shown in red. Dotted lines indicate possible G-A pairs formed in K-turn/K-loop structures.
Figure 4.5 Conservation of the second hairpin of the *H. volcanii* RNA originating from intergenic region T2743.

*Universally conserved nucleotides among halophiles*
tagged L7ae protein is controlled by the *hutU* promoter and HutR regulatory protein. A strain was confirmed with primers (hindhut-F and L6H-R) demonstrating a gene organization that does not occur in the *H. volcanii* genome but does exist in the recombinant plasmid, pARS-L6H.

Following isolation, the DS70-ASD41-L6H strain was cultured in CDM and assayed for expression of tagged L7ae. A culture that had reached balanced growth (A$_{550}$ of approximately 0.1-0.2) was portioned into two cultures and 10 mM urocanate was added to one culture to induce expression of the tagged L7ae gene. In general, 2-3 hours was allowed for induction prior to harvesting the cells by centrifugation. Whole-cell extract was generated from both cultures; samples of extract were precipitated with acetone to reduce the salt concentration and therefore reduce salt interference with electrophoresis. Proteins were separated using Tricine-PAGE due to the small size of L7ae (runs at ~14 kDa), and the gel was stained with coomassie blue. As seen in Figure 4.6 (Top panel), a prominent band corresponding to the expected size of L7ae is visible in the lane containing protein from the induced culture. Western blot analysis using a tetrahistidine anti-histidine antibody confirmed that the stained band observed upon induction contained a histidine tag.

Having successfully detected expression of the tagged protein, we developed a partial purification scheme that would allow for the isolation of L7ae-containing complexes. Using nickel-affinity chromatography we first utilized a linear gradient of 20 mM imidazole to 0.5 M imidazole. Dot-blotting with an anti-histidine antibody was used to survey elution fractions for the tagged protein. Fractions giving signal in dot-blots
Figure 4.6 Expression of hexahistidine tagged *H. volcanii L7ae*. Tricine-PAGE analysis of partially purified affinity-tagged L7ae (Top panel) and Western analysis using an anti-tetrahistidine primary antibody to confirm the presence of the tagged protein (Bottom panel). Labels indicate that samples came from an uninduced culture (U), an induced culture (I; 10 mM urocanate), flowthrough from the purification (Flow), or a particular purification fraction (fractions 21-23). Migration of molecular weight markers are indicated (mkr).
Figure 4.6 Expression of hexahistidine tagged *H. volcanii* L7ae.
were subjected to Tricine-PAGE and stained with coomassie blue. L7ae eluted in the range of fractions between 20 and 30 ml and the tagged protein’s presence was verified by Western blot (Figure 4.6, bottom panel). Subsequent examination of the fractions collected revealed that several protein bands specifically co-eluted with L7ae (Figure 4.7, top panel). To ensure that L7ae-RNA complexes were captured rather than free protein, samples of eluate were subjected to $^{32}$pCp labeling and separated on a DNA sequencing gel (denaturing acrylamide gel). Multiple signals were observed with several species showing enriched presence in the L7ae-containing fractions; not all signals could be assigned to a likely RNA candidate (Figure 4.8). Similar analysis of whole-cell extract from an uninduced DS70-ASD41-L6H culture showed an absence of protein and RNA bands thus validating the above results were resultant from capture of the tagged protein (Figure 4.7 bottom panel). To specifically determine if L7ae-associated RNAs were present, RNA was isolated from the elution fractions, separated with denaturing polyacrylamide gel electrophoresis and transferred to ZetaProbe™ membrane for Northern analysis. Hybridization of the membrane with $^{32}$P-labeled oligonucleotides resulted in detection of all known classes of RNAs bound by L7ae thereby suggesting that L7ae-containing sRNP complexes were successfully captured (Figure 4.9).

A step-elution scheme was then developed that enabled the elution of L7ae without apparent contamination by the known nickel-binding *H. volcanii* protein PitA (Allers et al., 2010). It has been reported that hexahistidine tagged proteins will, in general, elute at approximately 100-125 mM imidazole, making 80 mM imidazole a convenient concentration for a reasonably stringent wash to remove non-specific
Figure 4.7 Comparison of protein purification profiles of DS70-ASD41-L6H induced and uninduced cultures. A: Captured proteins from a culture induced with 10 mM urocanate. Fraction numbers are indicated above the lanes. B: Captured proteins from an uninduced culture. Equivalent samples from each purification were separated by Tricine-PAGE and stained with coomassie blue. Two large proteins known to bind nickel are present in both purifications, however, several unique bands are observed to co-elute when the affinity-tagged L7ae is expressed.
Figure 4.7 Comparison of protein purification profiles of DS70-ASD41-L6H induced and uninduced cultures.
Figure 4.8 RNAs associated with in vivo expressed affinity-tagged H. volcanii L7ae protein. $^{32}$pCp end-labeled fractions from either DS70 total RNA, or from column fractions (nickel affinity chromatography) of hexahistidine-tagged L7ae. Several RNAs co-elute specifically with the tagged L7ae (indicated by lines). Fraction numbers are indicated above the lanes.
$^{32}$pCp end-labeled RNA from fractions

Figure 4.8 RNAs associated with *in vivo* expressed affinity-tagged *H. volcanii* L7ae protein.
Figure 4.9 Northern analysis of purified, tagged L7ae fractions for RNAs known and proposed to bind L7ae. Each class of RNAs was detected in the captured fractions. Probes used were 7S-R for 7S RNA, RPR for RNase P RNA, trpE1 for intron-containing pre-tRNATrp (largest band) and tRNATrp intron (smallest band), T2743-2 for the box H/ACA RNA, unkCD for the box C/D RNA identified by Weisel et al. (2010), and 5S-R for 5S rRNA indicative of ribosomal subunits. The lanes contain either DS70 total RNA or RNA from a specific capture fraction as indicated by the number.
Figure 4.9 Northern analysis of purified, tagged L7ae fractions for RNAs known and proposed to bind L7ae.
proteins. While attempting to follow the initial 20 mM imidazole wash with the more stringent wash of 80 mM imidazole, we noted that L7ae eluted at 80 mM while PitA, based on its apparent molecular mass, did not elute until the imidazole concentration reached approximately 100-125 mM. Therefore, S100 samples were applied to the nickel column, washed with 20 mM imidazole and then eluted with 80 mM imidazole. Fractions were combined, acetone precipitated, and run using SDS-PAGE. Gels were stained with coomassie blue or transferred to PVDF for Western analysis.

Discussion

Discovery and capture of the box H/ACA sRNA T2743

Ribosomal protein L7ae has been described as having essentially one function: to secure a K-turn or K-loop conformation in RNA. The versatility of this function is widespread as L7ae binding sequences are present in both major classes of RNA-guided modifications, RNase P RNA, and SRP. L7ae has not been shown to provide catalytic functions to any of the sRNP complexes, and it is proposed that L7ae binding stabilizes the folding of the RNA scaffold to recruit other necessary proteins to fulfill the function of the RNA.

One such example of an RNA bound by L7ae in this work is the box H/ACA guide sRNA T2743. Analysis of tiled-array expression data across intergenic regions revealed that the intergenic region identified as T2743 showed marked expression. Examination of the sequence in this region suggested that there was a potential box H/ACA RNA encoded here. Since the box H/ACA motif is a fairly loose motif consisting primarily of structure rather than sequence, there is not yet an effective
program to predict these putative RNAs. We therefore used the *H. volcanii* sequence as a template to see if the RNA was conserved among other halophiles, and the results indicated that all halophiles had a similar sequence (see the LOGOS in Figure 4.5). This RNA has two conserved hairpin motifs, one of which appears capable of guiding formation of the two highly conserved pseudouridines in the 23S rRNA (Figures 4.2 B and 4.5). The transcript is readily detectable by Northern analysis during balanced growth as would be expected since there is considerable ribosome production during this growth phase. The fact that this RNA was pulled out attached to L7ae provides some *in vivo* evidence that this RNA is a box H/ACA guide RNA, but additional studies are necessary to prove that it recruits the remaining proteins and is responsible for the suggested modifications.

Recent evidence that L7ae is a component of RNase P in Archaea (Cho et al., 2010; Fukuhara et al., 2006) led us to examine the *H. volcanii* RNase P RNA sequence for the possible binding site(s). Comparison of the *H. volcanii* RPR structure with the *P. horikoshii* structure (Figures 4.3) suggests that there may be two binding sites for L7ae in the P12 and P15-17 regions. Likewise, our independent examination of the 7S RNA sequence showed that the 7S RNA helix 5 region was highly conserved among halophiles (Figure 4.4), and this region exhibits the characteristics of a putative L7ae binding K-loop domain.

**Overexpression and purification of affinity-tagged L7ae containing complexes**

In the present study, a system was developed to isolate of L7ae-containing sRNP complexes. *H. volcanii* L7ae was cloned with a C-terminal hexahistidine tag and placed
under the control of an inducible promoter. Once expressed, the tagged protein is capable of incorporating into complexes by binding RNA. Free tagged L7ae protein and complexes containing tagged L7ae can be captured using nickel affinity chromatography; the capture fractions appear to be free of contamination by known nickel-binding H. volcanii proteins when eluted with 80 mM imidazole. Gradient elution is also viable and leads to an interesting result of the various classes of complexes eluting in a reproducible order (Figure 4.9). Whether this is a result of varying affinity of the histidine tag or varying affinity of L7ae for these RNAs when the protein is in complex is uncertain. Although the factors controlling elution are unknown, this may provide a means to enrich these individual complexes.

All known box C/D sRNAs in H. volcanii except for the tRNAMet sRNA were identified in the tagged L7ae eluate (Figure 4.9). Very little tRNAMet sRNA is ordinarily detectable and the system utilized does not preclude the possibility that native tRNAMet sRNA is likely inaccessible or its synthesis is too low to detect new complexes via Northern blot. The T2743 box H/ACA RNA identified independently by our group and others was also present as were RNase P RNA, 7S RNA, and 5S rRNA; however, 5S rRNA could result from ribosomes. Detection of the RNA alone is insufficient to argue that complexes were captured in their entirety; in order to confirm the presence of all proteins involved in any of these complexes, a scheme must be devised to largely deplete the eluate of ribosomes since the abundance of ribosomal proteins may mask other less abundant proteins. Alternatively, the eluate could be tested for activity of the sRNP
complexes to provide evidence that at least a minimally functional version of the complex is present in the eluate.

We have confirmed a previous, but largely ignored prediction that the 7S RNA of the SRP appears bound by L7ae as 7S RNA could be detected in RNA isolated with L7ae pull-downs (Zago et al., 2005). A functional reconstituted halophile SRP has not yet been established, and may require additional proteins beyond those currently known. L7ae’s purpose may be to bind 7S RNA and lock in a conformation necessary for the recruitment of undetermined proteins which would presumably bind regions in the left half of the RNA (see Tozik et al. 2002 for full 7S RNA structure). The apparently low levels of 7S RNA are likely due to the fact that SRP associates primarily with the cell membrane, which is an insoluble component lost during the clearing centrifugation steps of whole-cell extract preparations. More telling results may be achieved by specifically isolating membrane components to determine if there is enrichment of 7S RNA when L7ae is captured.

Overall, the scheme developed in this study provides a means to over-express and partially purify native or presumably foreign proteins in *H. volcanii*. The advantage of protein production in the native host avoids difficulties sometimes encountered when attempting to produce foreign proteins in organisms such as *E. coli*. Protein folding problems and inclusion bodies can be overcome by making proteins in the native host. The conditional expression of the desired protein allows for expression to be activated at any time enabling protein over-expression during various growth phases or expression of a protein under conditions that it is not normally produced should that be desirable. Such
an approach is useful for attempting for discovering protein-protein or protein-RNA interactions as the overexpressed protein can integrate into complexes, and assuming the complex is stable under purification conditions, native complexes can be isolated. While this system is akin to the established tnaA expression system (Allers et al., 2010), the hutU promoter system is advantageous due to its undetectable expression in the absence of inducer molecule and the ability to use it in a background that is deleted for the plasmid components of the HutR regulatory protein.
Chapter 5: Establishing the role of archaeal intron endonuclease in processing of the ribosomal RNA operon in vivo

Introduction

Initial characterization of archaeal tRNA splicing noted that some tRNAs appeared to contain introns that were not among the self-splicing class of introns, but rather, resembled the eukaryal tRNA introns meaning that enzymatic activity was required to remove during tRNA maturation (Ogden et al., 1984; Strathern et al., 1982; Thompson and Daniels, 1988). The endonuclease responsible for intron excision was identified by its ability to cleave *H. volcanii* tRNATrp at the exon boundaries. The archaeal intron endonuclease bears a considerable structural similarity to the eukaryal intron endonuclease based on insights gained from examination of amino acid sequence modeling though no structure has yet been solved for the eukaryal enzyme (Calvin and Li, 2008). In fact, tRNA intron endonucleases share a highly conserved active site and
holoenzyme structure across domains (Calvin et al., 2008; Kim et al., 2007). Eukaryal intron endonuclease is an (αβγδ) heterotetramer predicted to achieve the same basic structure as the three different archaeal enzymes that have been crystallized, the α₄ homotetramer, the α₂ homodimer, and the (αβ)₂ dimer of heterodimers (Li and Abelson, 2000; Li et al., 1998; Mitchell et al., 2009; Xue et al., 2006). The aforementioned conserved active site, as suggested by analysis of the crystal structures, consists of a tyrosine, lysine, and histidine. An interesting finding resulting from the (αβ)₂ intron endonuclease is that this form of the archaeal enzyme, found among the Crenarchaea, appears to have less stringent substrate requirements. A “relaxed” BHB structure has been defined for substrate recognition meaning that non-canonical BHB structures can still be cleaved by this form of the enzyme but not by the α₄ or α₂ forms (Mitchell et al., 2009). As relaxed BHB motifs are found more frequently among the Crenarchaea, this finding explains how such substrates are cleaved by the intron endonuclease. The proposed catalytic mechanism appears to be similar to that of RNase A; the tyrosine and histidine are responsible for catalysis while the lysine residue’s charge stabilizes the catalytic pocket (Calvin et al., 2008). Mutational analyses have reaffirmed the crystal structure predictions that these amino acids, in particular histidine and lysine, are crucial for catalysis (Calvin et al., 2008; Kim et al., 2007).

For eukaryal intron endonuclease to catalyze intron removal it is necessary for the substrate to have tRNA structure (Baldi et al., 1992; Bufardeci et al., 1993; Reyes and Abelson, 1988). By contrast, the archaeal enzymes have no such requirement and instead only recognize a short structure motif known as the Bulge-Helix-Bulge (BHB) motif seen
in Figure 5.1 (Thompson and Daniels, 1990). This simple structural recognition element of the archaeal enzymes, is consistent with their ability to cleave intron-containing tRNA where introns are located at various positions within the tRNA. These data also suggest that tRNAs are not the only substrates of the intron endonuclease in Archaea.

The *Haloferax volcanii* tRNA intron endonuclease (EndA) is a member of the $\alpha_2$ homodimer class and was identified by its ability to excise the intron from *H. volcanii* tRNA$^{T rp}$ in vitro (Thompson and Daniels, 1988). Genome sequence and early cloning studies identified two additional tRNA genes interrupted by introns: tRNA$^{Met}$ (elongator) and tRNA$^{GlnUUG}$ (Gupta, 1984). All three intron-containing pre-tRNA introns are capable of forming a BHB motif for recognition by the endonuclease (Figure 5.1). Substrate recognition studies with this enzyme determined that only the BHB structural motif is recognized and the mature tRNA structure is not required (Armbruster and Daniels, 1997; Kleman-Leyer et al., 1997; Thompson and Daniels, 1990). This led to the suggestion that non-tRNA RNAs may also be substrates for this enzyme. Defining the BHB recognition mechanism of intron endonuclease shed light on the difficult problem of the elusive archaeal RNase III-related enzyme. In Bacteria, RNase III is the enzyme responsible for releasing the 16S and 23S rRNAs from the primary transcript of the rRNA operon through recognition of extended helices formed from 5’ and 3’ sequences flanking the 16S and 23S rRNA regions (Deutscher, 2009). Despite the presence of RNase III-like structures in archaeal rRNA transcripts, no such enzyme could be identified in the archaeal genomes suggesting that either rRNA operon primary transcript processing did not involve these structures or a different enzyme was
Figure 5.1 Potential EndA substrates. The BHB motifs of the three intron-containing pre-tRNAs and the BHB motifs formed around 16S rRNA and 23S rRNA in *H. volcanii*. Intron endonuclease cutsites are indicated by the arrows.
Figure 5.1 Potential EndA substrates.
responsible for excision of the 16S and 23S rRNAs. Early work describing maturation of archaeal rRNA by means of S1 nuclease mapping identified several discrete processing intermediates. Among these were intermediates containing the 16S rRNA and the 23S rRNA with small flanking sequences still attached (Chant and Dennis, 1986). These remaining flanking sequences mapped to the RNase III-like recognition sites, suggesting that a cleavage event occurred in those regions. Reexamination of the rRNA primary transcripts revealed the potential for the sequences flanking the 16S and 23S rRNAs to fold into structures that could form a BHB that could be recognized by the tRNA intron endonuclease (Figure 5.1). It has since been shown that the flanking sequences still attached to the 16S and 23S rRNAs indeed map to the predicted BHB cutsite (Durovic and Dennis, 1994). However, the difficulty of cloning rRNAs for in vitro studies and the challenge of inactivating an essential enzyme in vivo have prevented conclusive demonstration of the intron endonuclease’s apparent role in rRNA maturation.

Taking the view that the 16S and 23S rRNAs are essentially introns has allowed for the consideration that the sequences flanking them may be treated as exons and thus ligated together following the intron endonuclease cut. Indeed, Northern evidence has shown that 16S rRNA and 23S rRNA flanking sequences are ligated in *Archaeoglobus fulgidus* to give a discrete product from the 16S rRNA flanking sequences and a variable product from the 23S rRNA flanking sequences (Tang et al., 2002). The same study also concluded that the 16S rRNA and 23S rRNA flanking sequences were ligated together into a single product following release of the rRNA in *Sulfolobus solfataricus*. If the recently described archaeal ligase does not require a tRNA structure, it too, may
participate in the maturation of rRNA primary transcripts (Englert et al., 2011; Tanaka and Shuman, 2011). This work provided indirect evidence for the intron endonuclease releasing the rRNA from the primary transcript. It also gives additional evidence, see Armbruster et al. (1997), that the archaeal tRNA ligase is not tRNA-specific in vivo.

As an aside, the abundant, ligated 16S flanking sequences of *A. fulgidus* contains the characteristic box C/D motifs and this RNA bound L7ae in vitro suggesting that these ligated products have a role in the cell post-rRNA excision (Tang et al., 2002). Despite containing the box C/D sRNA hallmarks, no target(s) has been identified. The presence of the processing intermediates also has been detected for the *H. NRC-1* rRNA transcript using S1 mapping and DNA array analysis (Bonneau et al., 2007; Chant and Dennis, 1986; Durovic and Dennis, 1994). It remains possible that the flanking sequences may be responsible for guiding modifications or fulfilling an unidentified regulatory function in the cell given their abundance.

Interestingly, there are some apparent exceptions among the haloarchaea to the requirement of intron endonuclease for excision of at least the 16S rRNA from the primary transcript (Durovic and Dennis, 1994). While generally present, sequence analysis has shown that one of the *Haloarcula marismortui* rRNA operons (Dennis et al., 1998) lacks the potential for a canonical BHB motif to form with the 16S rRNA flanking sequences. Biochemical evidence in the form of S1 nuclease mapping has demonstrated that the 16S rRNA in this organism is released by an enzyme other than the intron endonuclease apparently at the mature 5’ end of the 16S rRNA, though the enzyme responsible for this activity has not yet been identified (Dennis et al., 1998). In the case
of organisms such as *Nanoarchaeum equitans*, relaxed BHB motifs are still recognized by the \((αβ)_2\) form of the intron endonuclease suggesting the possibility that organisms with this form of intron endonuclease could still excise the 16S rRNA without the need for a different enzyme (Mitchell et al., 2009). However, the \((αβ)_2\) form of the intron endonuclease has only been identified among the *Crenarchaea*, and the halophiles, which are among the *Euryarchaea*, have the \(α_2\) and \(α_4\) versions which will not accept the relaxed BHB motif. Additionally, evidence presented in the following work suggests that there is a redundant mechanism in *H. volcanii* for excision of the 16S rRNA independent of the intron endonuclease.

In the following study, we chose to investigate the role of the *H. volcanii* intron endonuclease, EndA, in the *in vivo* maturation of tRNA and rRNA primary transcripts by employing an inducible promoter system for the over-expression of an inactive EndA subunit. Given that our microarray data indicate a low level of expression for intron endonuclease, we reasoned that the native, wild type intron endonuclease activity could be masked by over-expression of a non-catalytic enzyme. Since the *H. volcanii* intron endonuclease holoenzyme is a homodimer, we hypothesized that over-expression of an inactive subunit would compete with the wild type enzyme for substrate, or that inactive monomers may form mixed dimers. The mixed dimers would be inactive since the active site is generated by sequences of both subunits at the subunit interface (Xue et al., 2006). The results indicate that intron endonuclease inhibition was successful and thorough analysis of the inactive subunit over-expression strain’s RNA revealed a definite impact on tRNA and rRNA maturation.
Results

Conservation of the ribosomal RNA operons including the Bulge-Helix-Bulge motif among Haloarchaea

As before, sequences were collected through use of BLASTN, aligned with ClustalW, and motif Logos were generated with MEME (Altschul et al., 1990; Altschul et al., 1997; Bailey and Elkan, 1994; Chenna et al., 2003). While previous studies have shown that there is not universal conservation of a canonical BHB among archaeal 16S and 23S rRNAs, there is conservation of the motif among the halophilic archaea with the exception of some rRNA operons in H. mukohataei and H. marismortui (Figures 5.2 and 5.3). The H. marismortui 16S rRNA, like the Sulfolobus 16S, has been shown to rely on another mechanism for excision of the rRNA (Dennis et al., 1998; Durovic and Dennis, 1994). The sequence, itself, is not conserved, but base-pairing in the stem around the BHB and within the internal helix is maintained by compensatory sequence differences. In the H. volcanii rRNA transcript, the result of excision at the BHB sites would still leave 108 nt and 77 nt on the 5’ and 3’ ends of the 16S rRNA, and 24 nt and 36 nt on the 5’ and 3’ ends of the 23S rRNA, respectively meaning that additional endonuclease or exonuclease activity is required for full maturation of the rRNA. Sequence resulting from the ligation of 16S and 23S rRNA flanking sequences does not appear to be conserved, although a box C/D sRNA-like 23S rRNA flanking sequence ligation product is present in H. volcanii. The role of these ligated fragments in the cell remains undefined (Figure 5.17).
Figure 5.2 Conservation of the BHB elements surrounding the 16S rRNA in the haloarchaea. 16S rRNAs that do not appear to be released by intron endonuclease, such as Hma2, were omitted. Alignments were generated with ClustalW and Logos were generated with MEME. Abbreviations are the same as those shown in Table 1.1. Numbers following the abbreviation indicate that multiple rRNA operons with differences in the intergenic sequences were present in that strain. Identical copies of the rRNA operon, such as in the case of *H. volcanii* were omitted from this alignment.
Figure 5.2 Conservation of the BHB elements surrounding the 16S rRNA in the haloarchaea.
Figure 5.3 Conservation of the BHB elements surrounding the 23S rRNA in the haloarchaea. Alignments were generated with ClustalW and Logos were generated with MEME. Abbreviations are the same as those shown in Table 1.1. Numbers following the abbreviation indicate that multiple rRNA operons with differences in the intergenic sequences were present in that strain. Identical copies of the rRNA operon, such as in the case of *H. volcanii* were omitted from this alignment.
Figure 5.3 Conservation of the BHB elements surrounding the 23S rRNA in the haloarchaea.
Construction of the intron endonuclease mutant strain

To investigate the role of the essential intron endonuclease gene, endA, the inducible hutU promoter module was used to regulate the expression of a mutant form of the enzyme (Sabag-Daigle, 2009). Mutation was accomplished by using three rounds of site-directed mutagenesis based on published work that identified potential active site residues as shown in Figure 5.4 (Calvin et al., 2008; Kim et al., 2007). Mutations were Y274F, H285A, and K316E. The pARS-endA* vector allowed for expression of the catalytically inactive intron endonuclease monomer in the presence of 10 mM urocanic acid.

Growth of the DS70-ASD41-endA* strain

The strain carrying the inducible endA* vector was cultured in CDM. When grown in broth medium, the culture exhibits a doubling time of approximately 4 hours during exponential growth; DS70-ASD41-endA* grows inherently slower than the wild type strain which we have found to double in as little as 3 hours during exponential growth. Background expression of endA* is not detectable at the mRNA level, and PCR evidence suggests that there may be some small amount of recombination between endA* and the wild type copy in the chromosome (data not shown). These observations may explain the slower growth and higher innate levels of intron-containing pre-tRNA in total RNA isolated from the strain as discussed later. The metabolic burden of the multi-copy plasmid may also contribute to the slower growth.

Upon addition of urocanate to 10 mM, growth appears to proceed as normal for one doubling period before a slight, yet reproducible, lag begins to develop. The growth
Figure 5.4 Mutation scheme for catalytically inactivating *H. volcanii* intron endonuclease reproduced from Kim et al (2007) and Calvin and Li (2008). A: Conservation of the intron endonuclease catalytic core across domains. B: Catalytic residues of intron endonuclease with the amino acid mutations indicated.
Figure 5.4 Mutation scheme for catalytically inactivating *H. volcanii* intron endonuclease reproduced from Kim et al. (2007) and Calvin and Li (2008).
lag continues to increase over the course of time based on $A_{550}$ and cultures receiving the inducer molecule never reach the same final $A_{550}$ as wild type or uninduced strains (Figure 5.5). Uninduced cultures typically reach a maximum $A_{550}$ of 1.8-2.0 while cultures induced with 10 mM urocanate achieve only a maximum $A_{550}$ of 1.1-1.2.

**Induction of endA***

Once the strain had been cultured, DS70-ASD41-endA* was assessed for successful production of the mutant intron endonuclease subunit. First, total RNA was isolated from both induced and uninduced DS70-ASD41-endA* cultures at various time points post-induction. RNA was separated with partially denaturing agarose gel electrophoresis and transferred to ZetaProbe™ membrane for Northern hybridization. The oligonucleotide used, (endAUS), takes advantage of the additional sequence added to the 5’ end of the endA* transcript from the hutU start and the BamHI restriction enzyme recognition site. This probe spans from the unique start region into the endA* gene, which results in hybridization only with the mutant transcript. The endA* transcript was readily detectable only in total RNA isolated from cultures that had been induced with urocanate (Figure 5.6 A). No signal was observed in the absence of urocanate indicating that transcription was inducible.

To verify that the endA* transcript was also translated, whole-cell extract was generated from uninduced and induced DS70-ASD41-endA* cultures. Protein was precipitated from the extract and subjected to SDS-PAGE for separation. The gel was stained with coomassie blue. Relative to the lane containing protein from an uninduced culture, there was a new prominent band in the induced lane (Figure 5.6 B). This band
Figure 5.5 Growth of the DS70-ASD41-\textit{endA*} strain. Time zero on the graph indicates the point at which the cultures (grown overnight) were split with one half receiving a mock induction and the other half receiving the inducer molecule urocanate to a final concentration of 10 mM. Growth was monitored by measuring the absorbance at 550 nm.
Figure 5.5 Growth of the DS70-ASD41-*endA* strain.
Figure 5.6 Inducible expression of the endA* gene following urocanate addition. A: The unique transcript from the endA* gene detected at the mRNA level with probe endAUS. The 5S-R probe was used to hybridize with the 5S rRNA as a loading control. B: EndA* detected in the total protein content as seen with SDS-PAGE. The migration of molecular weight standards are indicated. C: EndA and EndA* detected by Western blot using an anti-EndA primary antibody. Lanes labels indicate that the sample was taken from either an uninduced culture (U) or a culture induced with 10 mM urocanate (I).
Figure 5.6 Inducible expression of the endA* gene following urocanate addition.
matched the purported size of the intron endonuclease, running at approximately 37 kDa.
Finally, verification that the new band corresponded to increased amounts of intron
endonuclease subunits was carried out by way of a Western blot. The primary antibody
was a polyclonal antibody against the wild type EndA protein (Kleman-Leyer et al.,
1997). Results of western analysis show that above normal levels of intron endonuclease
are readily detectable upon induction (Figure 5.6 C).

**Verification of intron endonuclease inhibition**

Having confirmed that protein was induced, total RNA was examined in order to
confirm that there was inhibition of intron removal activity. The three known substrates,
tRNATrp, tRNAMet (elongator), and tRNAGlnUUG were used as indicators of relative
activity. Total RNA isolated from the strain was transferred to ZetaProbe™ for Northern
analysis. Typically, total RNA yields from *H. volcanii* strains such as DS70-ASD40 will
have a low, but detectable level of pre-tRNA for both tRNATrp and tRNAMet. RNA
from DS70-ASD41-*endA* was compared to RNA isolated from a nearly identical strain
containing the same vector with the inactive intron endonuclease subunit replaced by the
reporter gene yeast tRNAProM. The addition of urocanate to the culture expressing the
reporter gene shows no significant impact on the ratio of intron-containing pre-tRNATrp
to mature tRNATrp, while the DS70-ASD41-*endA* strain shows the elevated
background pre-tRNA exacerbated by the addition of inducer molecule (Figure 5.7 A).
This is likely due to a low level of expression of the EndA* protein under non-inducing
conditions. However, in RNA isolated from induced cultures, the levels of intron-
containing pre-tRNA increased measurably for all three intron-containing tRNAs (Figure
Figure 5.7 Validation of intron endonuclease inhibition as measured by accumulation of intron-containing pre-tRNAs 4 hours post-induction of EndA* protein. A: Comparison of tRNATrp cleavage in the control strain carrying the yeast tRNAProM under the control of the HutR-inducible promoter (P) and the strain carrying the endA* gene under the control of the same promoter (E). Northern analysis of RNA taken four hours post-induction provides the comparison of tRNATrp cleavage. Lanes labels indicate that the sample was taken from either an uninduced culture (U) or a culture induced with 10 mM urocanate (I). The trpE1 probe was used for hybridization. B: The pre-tRNAs and mature tRNAs are indicated as well as the ratio of pre-tRNA to mature tRNA as determined by the ImageQuant software. The non-intron-containing tRNAAspGUC served as a loading control. Lanes labels indicate that the sample was taken from either an uninduced culture (U) or a culture induced with 10 mM urocanate (I). Probes used to detect the various tRNA species were trpE1, metE1, glnE1, and t-asp.
Figure 5.7 Validation of intron endonuclease inhibition as measured by accumulation of intron-containing pre-tRNAs 4 hours post-induction of EndA* protein.
5.7 B). The consistently higher levels of pre-tRNA observed in total RNA from induced cultures strongly suggested that there was inhibition of intron endonuclease activity as a result of EndA* production.

**Time-course accumulation of pre-tRNA**

As a method to assess the length of time necessary for detectable inhibition of intron endonuclease, the build-up of pre-tRNAGlnUUG was monitored. Since this tRNA shows very low levels of intron-containing precursor ordinarily, it was used as the time course indicator. As shown in Figure 5.8, slight, but observable accumulation of pre-tRNA is present at only one hour post-induction with an apparent maximal effect at four hours post-induction. Inhibition and detection of the mutant intron endonuclease subunit transcript are still detectable at least eight hours post-induction (data not shown), though at this point the culture has left balanced growth and may be responding to stress signals. We therefore confined our analysis, unless otherwise indicated, to four hours or less post-induction.

**Analysis of the impact of endA* expression on total RNA isolated from the mutant strain**

Previous studies in this laboratory using an intron prediction search tool have indicated that there are >100 possible substrates of intron endonuclease in the *H. volcanii* genome (Ray and Daniels, 2003). While it is unlikely that all or even most of these potential substrates are real substrates, the structure-based recognition of the enzyme allows for more than just tRNAs and rRNA. To observe the impact of intron endonuclease inhibition on the total RNA population, total RNA was isolated and
Figure 5.8 Time course analysis of the accumulation of intron-containing pre-tRNA in DS70-ASD41-\textit{endA*} under normal and under inhibitory conditions.

The pre-tRNAGlnUUG was used to monitor the time course. Northern analysis was conducted with a 1:1 mix of glnIntron and glnSpld probes. The number in parentheses indicates the number of hours post-induction of EndA* the RNA sample was taken. Lanes labels indicate that the sample was taken from either an uninduced culture (U) or a culture induced with 10 mM urocanate (I).
Figure 5.8 Time course analysis of the accumulation of intron-containing pre-tRNA in DS70-ASD41-endA* under normal and under inhibitory conditions.
3’-end-labeled with $^{32}$pCp. Over the course of time, several signals, indicated in Figure 5.9, are lost in the RNA isolated from cells grown under EndA* induction conditions, while some new species appear. We examined the 3’-end-labeling on the small (<500 nt) RNA populations, and we expect that RNAs produced by ribonuclease activity producing 3’ or 2’-3’ cyclic phosphate would be excluded.

**Impact of intron endonuclease inhibition on rRNA processing**

Having completed the test case of known substrate splicing inhibition, processing of the rRNA operon transcript (Figure 5.10) was investigated with a series of Northern hybridizations. Hybridizations against the 16S, 23S, and 5S rRNAs were conducted, but there was no observable impact of induction beyond a small smear just above the 16S rRNA for RNA isolated from induced cultures (Figure 5.11). This is likely due to already high levels of mature rRNAs in the cells as they were harvested during exponential phase. We therefore chose to examine the impact on the rRNA operon intergenic regions. All RNA examined for the effects on the rRNA operon is total RNA isolated from DS70-ASD41-EndA* under normal conditions (exponential growth in CDM) or inhibition conditions (exponential growth in CDM, but given urocanate to 10 mM for 1-4 hours). Oligonucleotide probes were designed to target each region of the rRNA operon (Figure 5.10). Splice junction probes, A’ and F’, were also designed to detect ligated flanking sequences.

**Impact of intron endonuclease inhibition on 16S rRNA processing**

Results from probes A-D are presented in this section. At first, total RNA was separated with agarose gel electrophoresis under partially denaturing conditions or with
Figure 5.9 Effect of EndA* on the population of small RNAs (<500 nt). Total RNA isolated from DS70-endA* cultures at 1 hour or 4 hours post-induction was labeled at the 3’ end with $^{32}$pCp. Lanes labels indicate that the sample was taken from either an uninduced culture (U) or a culture induced with 10 mM urocanate (I). Some of the changes (gains designated by + or loss designated by -) apparent in the RNA population are indicated by arrows.
Figure 5.9 Effect of EndA* on the population of small RNAs (<500 nt).
Figure 5.10 Schematic representation of the *H. volcanii* rRNA operon and positions of probes used in Northern analysis. Intron endonuclease cleavage sites are marked with dashed lines and the length of genes and the distance between genes and cutsites is listed in nucleotides. Labeled arrows below the operon represent the approximate locations of Northern hybridization probes. Due to the multiple possible promoters, the distance between the U3-like sequence and the transcription start site is uncertain and therefore labeled with a “?”. 
Figure 5.10 Schematic representation of the *H. volcanii* rRNA operon and positions of probes used in Northern analysis.
Figure 5.11 Northern analysis of rRNAs from DS70-EndA* using the probes indicated beneath each panel. Total RNA was isolated four hours post-induction and separated using partially denaturing agarose gel electrophoresis. Lanes labels indicate that the sample was taken from either an uninduced culture (U) or a culture induced with 10 mM urocanate (I).
Figure 5.11 Northern analysis of rRNAs from DS70-EndA* using the probes indicated beneath each panel.
denaturing PAGE prior to transfer to ZetaProbe™ membrane for Northern analysis. Clear differences were observable for RNA species detected by probes A, B, and D while those related to C show no obvious change under inhibition.

Figure 5.12 A shows an apparent decrease in a ~110 nt species with an increase in a 200 nt species for the region 5’ of the 5’ 16S rRNA intron endonuclease cutsite when the RNA is separated by denaturing PAGE. When examining the same probe with agarose gel electrophoresis separation, the same shift in the smaller species is detected as well as an apparent increase in the signal at ~1700 nt suggesting that there is an increase in the species where the flanking sequencing hybridizing with probe A remains attached to the 16S rRNA. The occurrence of an approximately 190-200 nt species under non-induction conditions also indicates that the processing at the 5’ end of the 16S rRNA can proceed without cleavage at the BHB site.

Located between the intron endonuclease cutsite and the start of the 16S rRNA, the 108 nt region that probe B targets shifts up slightly when intron endonuclease is inhibited (Figure 5.12 B). The intermediate increases in size by approximately 10 nt based on a comparison against the same membrane probed for 5S and 7S markers. When separated by agarose gel electrophoresis, the shift of the small lower band is still evident, though there is no obvious increase in the intensity of the 1600-1700 nt band meaning intron endonuclease is not requisite for 5’ maturation. A band of ~200 nt was observed under induction conditions and may represent an intermediate derived from endonuclease activity to release the mature 5’ end of the 16S rRNA. The ~200 nt species may consist
Figure 5.12 Effect of EndA* on the processing of rRNA transcripts in the 5’ region of the 16S rRNA. Total RNA was isolated four hours post-induction and separated using denaturing polyacrylamide gel electrophoresis or partially denaturing agarose gel electrophoresis as indicated. Lanes labels indicate that the sample was taken from either an uninduced culture (U) or a culture induced with 10 mM urocanate (I). Approximate sizes were extrapolated using a set of known size markers (23S rRNA, 16S rRNA, RNase P RNA, 7S RNA, and 5S RNA).
Figure 5.12 Effect of EndA* on the processing of rRNA transcripts in the 5’ region of the 16S rRNA.
of the full length of the 5’ flanking sequence that has not been cleaved by intron endonuclease.

Probe D also shows a discernible change when comparing normal growth to inhibition conditions (Figure 5.13). The signal shifts from an ordinarily detectable processing intermediate of ~80 nt with some amount still attached to the 16S rRNA to a strong signal attached to the 16S and little to no detectable small processing intermediate.

**Impact of intron endonuclease inhibition on tRNAAlaUGC processing**

While RNase P is considered responsible for releasing tRNAAlaUGC from the rRNA operon, it may still be dependent on intron endonuclease for primary release. As seen in Figure 5.14, under non-inducing conditions tRNAAlaUGC still attached to the 16S rRNA awaiting release by intron endonuclease or RNase P; it appears that there are two distinct species. Several RNAs in the size range of the mature tRNAAlaUGC are detectable. The slightly larger band appears to be the tRNA with the 3’ flanking sequence of the 16S rRNA (probe D region) still attached. Under inhibition conditions, the amount of tRNA sequestered with 16S rRNA visibly increases while smaller species decrease, and a new band replaces the tRNA plus leader sequence band. This larger band of ~110 nt is likely the tRNA plus the full 3’ sequence between the tRNA and 16S rRNA, which is expected since the intron endonuclease cut would not have occurred. This also suggests that while not as efficient as 5’ maturation, 3’ maturation of the 16S rRNA can still occur without intron endonuclease activity. A fourth band >110 nt is present in the uninduced RNA; unfortunately, we are unable to account for this species with the expected cleavage sites.
Figure 5.13 Effect of EndA* on the processing of rRNA transcripts in the 3’ region of the 16S rRNA. Total RNA was isolated four hours post-induction and separated using denaturing polyacrylamide gel electrophoresis or partially denaturing agarose gel electrophoresis as indicated. Lanes labels indicate that the sample was taken from either an uninduced culture (U) or a culture induced with 10 mM urocanate (I). Approximate sizes were extrapolated using a set of known size markers (23S rRNA, 16S rRNA, RNase P RNA, 7S RNA, and 5S RNA).
Figure 5.13 Effect of EndA* on the processing of rRNA transcripts in the 3’ region of the 16S rRNA.
Figure 5.14 Effect of EndA* on the processing of the 16S rRNA–tRNA region of the rRNA transcript. Total RNA was isolated four hours post-induction and separated using partially denaturing agarose gel electrophoresis as indicated. Lanes labels indicate that the sample was taken from either an uninduced culture (U) or a culture induced with 10 mM urocanate (I). Approximate sizes were extrapolated using a set of known size markers (23S rRNA, 16S rRNA, RNase P RNA, 7S RNA, and 5S RNA).
Figure 5.14 Effect of EndA* on the processing of the 16S rRNA–tRNA region of the rRNA transcript.
Impact of intron endonuclease inhibition on 23S rRNA processing

The probes covering the regions 5’ (F and G) and 3’ (H and I) of the 23S rRNA were used to probe processing of the 23S rRNA. Neither G nor H hybridized to small (<500 nt) RNAs (data not shown), but each resulted in an increase in a larger RNA that was as large or larger than the 23S rRNA in RNAs from induced cells (Figures 5.15 and 5.16). No small bands were observed when using these probes, suggesting that once the 23S rRNA is released by intron endonuclease, these pieces are rapidly degraded, possibly by exonucleolytic activity.

The remaining probes, F and I, provide evidence of the involvement of intron endonuclease in the maturation of the 23S rRNA. Under normal conditions, processing intermediates approximately the sizes of the intergenic regions surrounding the 23S rRNA are readily detectable by Northern hybridization. These signals are lost under inhibition conditions and larger signals can be seen on the membrane at a size that correlates with the 23S rRNA (Figures 5.15 and 5.16).

Investigation of 16S rRNA and 23S rRNA flanking sequence ligation in *H. volcanii*

A previous report has identified ligation of the 16S rRNA flanking sequences in *S. solfataricus* and *A. fulgidus* (Tang et al., 2002). They reasoned that if the rRNAs are released as introns, then the flanking sequences could be as exons and therefore substrates for ligation. Additional analysis indicated that there were box C/D-like sequences in the abundant ligated product (Figure 5.17).

To corroborate this finding in the *H. volcanii*, we also used Northern hybridizations with total RNA from cultures grown under normal or inhibition
Figure 5.15 Effect of EndA* on the processing of rRNA transcripts in the 5’ region of the 23S rRNA. Total RNA was isolated one (F) or four (G) hours post-induction and separated using denaturing polyacrylamide gel electrophoresis or partially denaturing agarose gel electrophoresis as indicated. Lanes labels indicate that the sample was taken from either an uninduced culture (U) or a culture induced with 10 mM urocanate (I). Approximate sizes were extrapolated using a set of known size markers (23S rRNA, 16S rRNA, RNase P RNA, 7S RNA, and 5S RNA).
Figure 5.15 Effect of EndA* on the processing of rRNA transcripts in the 5' region of the 23S rRNA.
Figure 5.16 Effect of EndA* on the processing of rRNA transcripts in the 3’ region of the 23S rRNA. Total RNA was isolated one (H) or four (I) hours post-induction and separated using denaturing polyacrylamide gel electrophoresis or partially denaturing agarose gel electrophoresis as indicated. Lanes labels indicate that the sample was taken from either an uninduced culture (U) or a culture induced with 10 mM urocanate (I). Approximate sizes were extrapolated using a set of known size markers (23S rRNA, 16S rRNA, RNase P RNA, 7S RNA, and 5S RNA).
Figure 5.16 Effect of EndA* on the processing of rRNA transcripts in the 3’ region of the 23S rRNA.
Figure 5.17 Sequence and location of the 23S flanking sequence ligated product.

The full sequence of the ligated product is given (left panel). The box C/D-like regions of the ligated product are underlined while the sequences involved in formation of the BHB structure (right panel) are indicated in red. The ligation site is given in all capital letters.
Contains Box C/D-like sequences

tggtaagcagctactccactccgaaacgatcgtgcccccttaagt
gtggcaagacgttttctgaatgaatggatat tgacga cagatgca
ccatccccgccagcagcgagggtggaaggtcgaacacgcct
cggtatctc cccaggggtgttt gatga aacccgtgtgtacgtgca
atccaggccgtCCAactggattgacgcagcacacattgtgtaagacgg
ccgacccgatttggtatggcaacggttcgattccgtttgatcgtgta

Figure 5.17 Sequence and location of the 23S flanking sequence ligated product.
conditions. Attempts to detect a ligated product from the 16S rRNA flanking sequences with probe A’ resulted in no hybridization signals suggesting that these sequences may not ligate post 16S excision. However, when using a splice junction probe, F’, for the 23S rRNA flanking sequences, a strong signal was observed. The band migrated at approximately the same size as the signal detected with F, but was larger than the probe I signal. The apparent ligation product was only detectable in total RNA isolated from DS70-ASD41-\textit{endA} \* cells grown under normal conditions, hence implicating intron endonuclease in the formation of this product (Figure 5.18).

While it seems the flanking sequences of the 23S rRNA do ligate, closer inspection of the splice junction primer revealed a direct repeat of the first 6 nt across the junction which could mean that the band observed is a false positive signal. Additionally, signals observed from probe I were always extrapolated as smaller than those from probe F; if the flanking sequences are ligated, then both probes would presumably give the same signal. This signal could be the faint band observed in Figure 5.16 I at ~200 nt. To resolve the inconsistencies, RT-PCR was used on total RNA isolated from cells grown under both conditions. A product corresponding to the ligated flanking sequences of the 23S rRNA was observed, but no product from the 16S rRNA flanking sequences was observed (data not shown). These results were consistent with the initial Northern hybridizations, but to confirm that a ligated product was indeed preset, the RT-PCR product was TA cloned and sequenced. Sequencing data indicates that the 23S rRNA flanking sequences are joined exactly at the site of intron endonuclease cleavage. Ends of the ligated product were not determined as they were limited by the RT-PCR primers.
Figure 5.18 Effect of EndA* on the 23S rRNA flanking sequence ligation product.

Northern analysis using a probe which bridges the splice junction, F’, is indicated by the dotted arrow shown below a schematic representation of the region. Lanes labels indicate that the sample was taken from either an uninduced culture (U) or a culture induced with 10 mM urocanate (I). Approximate sizes were extrapolated using a set of known size markers (23S rRNA, 16S rRNA, RNase P RNA, 7S RNA, and 5S RNA).
Figure 5.18 Effect of EndA* on the 23S rRNA flanking sequence ligation product.
Analysis of the halophilic rRNA operons shows that the 23S rRNA flanking regions share greater conservation than the 16S rRNA flanking sequences. However, the box C/D-like elements in the ligated product only seem to be present in some of the H. volcanii (Figure 5.17), though some box C/D-like elements are also present in the ligated 23S rRNA flanking sequences from H. mukohatei. Still, these elements could explain why this product remains at such high levels in the cell rather than the flanking sequences undergoing degradation immediately after the rRNA is released.

Discussion

Expression of endA* has a dominant negative impact on growth

tRNA intron endonuclease has an established role in the removal of introns from tRNA transcripts and is predicted to be an essential enzyme. Loss of this activity by inhibition or inactivation is expected to result in a decrease in growth following the turnover of essential elements of the translation machinery. Based on a structural analysis of the archaeal α2 tRNA intron endonuclease and in vitro biochemical studies of the mutant EndA enzymes, a catalytic mechanism for this enzyme was proposed. The active site was defined by three key residues Y274, H285, and K316.

We took advantage of a controlled expression system for H. volcanii and the knowledge of the requirements for activity of the α2 EndA enzyme to selectively overproduce a catalytically inactive endonuclease subunit, EndA*. The production of EndA* was expected to produce inactive α2 homodimers and EndA* monomers could form mixed dimers, EndA-EndA*, which would also lack activity. Formation of inactive EndA* and mixed complexes, coupled with the turnover of wild type (EndA)2
complexes, was expected to decrease intron processing and ultimately compromise the translation machinery.

Consistent with this model, induction of endA* gene expression led to a delayed, but measurable decrease in the growth rate of the culture (Fig 5.5). The induced culture continued growth but did not reach the same final cell density as the non-induced cultures. Cells carrying the endA* gene grown under non-inducing conditions also displayed a slight increase in doubling time when compared with the wild type strain. We reasoned that this growth defect most likely resulted from a low level of endA* expression, possibly due to inducer formation from endogenous histidine. The multi-copy plasmid (about 20 copies per cell) places an additional metabolic burden on the strain, but PCR evidence (data not shown) indicates that there may be some background recombination at the site of the wild type intron endonuclease. As there are multiple copies of the chromosome present in each cell (Soppa, 2011), recombination of the vector into one copy of the chromosome could result a low background expression of EndA. The effects of EndA* expression provide further evidence that the intron endonuclease plays an essential role in the cell.

In addition to the slower growth of the endA*-containing strain under non-inducing conditions, there is a considerable growth defect that develops over time upon addition of the inducer. The intron endonuclease gene is not expressed at high levels in *H. volcanii* based on microarray analysis, and it is likely that the EndA homodimer is fairly stable. This is consistent with the growth properties of the EndA* strain following induction where the growth lag is not even detectable until at least one doubling period
has occurred and the pool of functional intron endonuclease has been diluted. The low expression levels of the wild type gene coupled with the rapid high expression of the mutant endA* gene is likely to produce high levels of EndA* monomers and EndA* dimer complexes and activity inhibition is apparent in the RNA population even before the growth of the cells is impacted as suggested in Figure 5.7. Growth is not completely stopped, though, suggesting that there is still enough wild type enzyme present to permit some amount of tRNA intron removal and rRNA release. The initial four-hour doubling period is never recovered and the final outgrowth does not achieve the same maximal levels as when cultured under normal conditions. These observations provide further support that intron endonuclease is essential for normal growth of H. volcanii.

Inactive monomer expression is sufficient to significantly inhibit Bulge-Helix-Bulge cleavage

Known substrates in H. volcanii consist of three pre-tRNAs: Trp, Met (elongator) and GlnUUG. Among these pre-tRNAs, the intron length varies from 31 to over 100 nt, but they all have the ability to form the BHB motif for recognition by intron endonuclease. These established substrates provided a test-case for evaluating the reduced intracellular intron endonuclease activity.

All three pre-tRNAs showed an observable level of accumulation when EndA* was expressed. Although the increases in levels of pre-tRNA-Trp and pre-tRNA-Met are somewhat modest, pre-tRNA-Gln is nearly undetectable in the total RNA pool unless the inactive endonuclease subunit is induced. This is particularly striking as the DS70-ASD41-endA* strain already shows higher pre-tRNA levels than strains without the
vector containing endA* (Figure 5.7 A). Upon induction, pre-tRNA\textsubscript{GlnUUG} becomes readily detectable at much higher levels (Figure 5.7 B). This suggests that although pre-tRNA\textsubscript{GlnUUG} appears to have the weakest 4-base-pair helix with 3 G-U pairs, it is readily processed \textit{in vivo}. The small intron size, less than half the size of the tRNA\textsubscript{Met} intron and about one-fourth the size of the tRNA\textsubscript{Trp} intron, has less complexity and may fold into the BHB more readily. As pre-tRNA\textsubscript{Met} cleavage efficiency is between pre-tRNA\textsubscript{Trp} and pre-tRNA\textsubscript{GlnUUG}, there is some support to this idea at the pre-tRNA level. However, another explanation for the comparatively lesser accumulation of pre-tRNA\textsubscript{Trp} is that intron endonuclease cleavage may be inhibited already as a result of box C/D machinery binding the pre-tRNA and converting it into an sRNP complex incapable of intron excision.

Irrespective of the rationale for the uneven increase of pre-tRNAs in the total RNA pool, all three show measurable accumulation. This result serves as proof of principle that intron endonuclease activity can be significantly reduced by stimulating production of the inactive subunit. Continued growth of the strain post-induction suggests that some active homodimers remain, but that overall intracellular activity has been largely inhibited.

**There is an observable impact of endA* expression on total RNA**

Since the known substrates of intron endonuclease were small RNAs, we examined the effects of EndA* on this population of RNAs. RNAs were 3’-end-labeled with \textsuperscript{32}pCp and separated by denaturing gel electrophoresis. As none of the bands resulting from the \textsuperscript{32}pCp were sequenced or identified, only general conclusions can be
made from Figure 5.9. Some of the changing RNAs appear to correspond to the accumulating pre-tRNAs. However, there are well more than three band differences between the normal and inhibitory conditions. Interestingly, the loss of bands from inhibitory conditions is most notable. These lost bands could be intron endonuclease substrates that are now too large to be effectively separated by polyacrylamide gel electrophoresis. It remains possible that some of the changing RNA profile is due to the cells experiencing stress in response to inactive endonuclease subunit production, though this does not seem likely as the growth does not significantly lag until after a doubling period and RNA samples were taken prior to or just at one doubling period. Earlier predictions of numerous intron endonuclease substrates in *H. volcanii* may be correct given the results presented in this study, but additional work identifying the bands is necessary to address this point.

**Intron endonuclease effects the release of the 16S rRNA from the primary transcript**

Comparative sequence analysis predicted that BHB structures occur in the flanking regions of the 16S and 23S rRNAs in some Archaea. The absence of an identifiable RNase III enzyme in these organisms suggested that tRNA intron endonuclease may function in the maturation of rRNA primary transcripts. Northern hybridization evidence seen in Figures 5.12 and 5.13 shows that there is an observable impact at three of the four regions surrounding the intron endonuclease cutsites when EndA* is expressed. Inspection of the region 5’ of the 16S rRNA reveals that a small, ordinarily detectable intermediate (Figure 5.12 A) decreases and a larger species appears.
There is also an increase in a band larger than the 16S rRNA. These results show that intron endonuclease participates in 16S rRNA processing. Interestingly, the results of Figure 5.12 B suggest that there is a secondary endonucleotic cut at or near the mature 5’ end of the 16S rRNA that releases the 5’ flanking sequence intermediate. As hybridization to the flanking sequence 5’ of the intron endonuclease cleavage site shows a shift up above the 16S rRNA size as well, this implies that the intron endonuclease cut normally occurs first. This cleavage does not appear requisite for the secondary enzymatic activity for full 5’ maturation though since hybridization 3’ of the cleavage site shows that a small intermediate larger than the normally observed intermediate is detectable.

Analysis of results from hybridizations 3’ of the 16S rRNA reveals a different pattern from the 5’ end. Data from probe C (Figure 5.13) representing sequences immediately 3’ of the 16S coding region shows no substantial impact from intron endonuclease inhibition. A single obvious band of about 1500 nt is detectable. The band is invariant under normal or inhibition conditions. A smaller intermediate is barely detectable when using polyacrylamide gel electrophoresis for nucleic acid separation prior to transfer. This signal does not show measurable variation either. In stark contrast, results from probe D (Figure 5.13) indicate that this region is now associated with a larger RNA similar in size to the 16S rRNA. Blockage of intron endonuclease activity causes the flanking sequence 3’ of the 16S rRNA to remain attached with the rRNA. The implication of these results is that 3’ maturation of the 16S rRNA, like the 5’ end, relies on intron endonuclease for release from the primary transcript followed by a secondary
cleavage by an as of yet unidentified endonuclease for full maturation. Endonucleolytic activity seems more likely for full 16S rRNA maturation since discrete intermediates were detectable under normal conditions. This finding is in agreement with the previous studies suggesting the likelihood of a secondary, independent endonuclease, perhaps the same endonuclease responsible for release of the 16S rRNA in operons that lack canonical BHB motifs (Dennis et al., 1998).

**Intron endonuclease acts prior to RNase P excision of tRNAAlaUGC**

Data from probe D indicate that the 3’ flanking sequence was attached to the 16S rRNA under inhibitory conditions; we examined the total RNA populations to determine if tRNAAlaUGC was also still associated with the rRNA. Even under normal growth conditions there is some amount of the tRNA detectable attached to the 16S rRNA (Figure 5.14). Additionally, there are two small bands on the gel that appear to correspond to the mature tRNA and the tRNA with the 20 nt leader still attached. Interestingly, inhibition of intron endonuclease activity results in the tRNA shifting more prominently to larger 16S rRNA-size molecules. There is also the development of another band over time that appears to be tRNAAlaUGC with a longer leader sequence. This could be explained by the lack of intron endonuclease cleavage combined with secondary endonuclease activity to mature the 3’ side of the 16S rRNA. The end-result is that the tRNA now has a leader consisting of the entire intergenic region rather than just the region 3’ of the endonuclease cleavage site. Detection of these intermediates and tRNAAlaUGC remaining attached to the 16S rRNA further suggests that RNase P does not normally, or cannot, remove the tRNAAlaUGC leader until after intron endonuclease
or the unknown endonuclease cleaves the region 3’ of the 16S rRNA. Such a cut would then enable RNase P to remove the leader sequence.

Unlike 5’ maturation, 3’ maturation of tRNAAlaUGC seems independent of intron endonuclease activity. Regardless the conditions used to culture the strain, total RNA isolated from cultures did not yield a signal larger than about 1700 nt when probe E was hybridized with the RNA (Figure 5.14), thus excluding the prospect of the tRNA remaining attached to the 23S rRNA. Subsequent analysis of the 23S rRNA flanking sequences, discussed below, agrees as the hybridization patterns did not reveal a signal with the approximate size of the tRNA plus the 23S rRNA flanking sequences. These data agree with previous reports that although efficient 5’ maturation by RNase P requires intron endonuclease cleavage first, 3’ maturation via tRNase Z is independent of intron endonuclease (Chant and Dennis, 1986; Dennis et al., 1998; Holzle et al., 2008).

**Intron endonuclease is responsible for excision of the 23S rRNA from the primary transcript**

The impact of intron endonuclease inhibition on 23S rRNA maturation was examined in the same manner as described for 16S rRNA maturation. Hybridization data from probes F and G (Figure 5.15) offer some elucidation of the 5’ end of 23S rRNA. Microarray data from *H.-NRC1* and the Northern data of this study both show that during exponential growth there is a readily detectable processing intermediate consisting of the intergenic sequence between the tRNA and 23S rRNA (Figures 5.15 and 5.16). When intron endonuclease activity is largely depleted, the intermediate is lost from the total RNA population. Like with the 16S rRNA flanking sequences, the loss of the smaller
RNA is coincident with a signal at the size of the 23S rRNA. This result suggests that once again, the intron endonuclease cut occurs first, then subsequent nuclease activity matures the 5’ end. Signal was hardly detectable with probe G, but a single large band migrating at just above the 23S rRNA band is visible. Due to the faintness of the band and the fact that there are only 24 nt between the cutsite and the 23S rRNA 5’ end, the remaining flanking sequence is most likely degraded. No discrete small intermediate was observed when using probe G for Northern hybridizations suggesting that these 24 nucleotides may be degraded directly by an exonuclease or its small size may make it difficult to detect. Previous experiments using S1 nuclease identified processing intermediates that were argued to imply a secondary endonuclease to mature the 5’ and 3’ ends of the 23S rRNA (Chant and Dennis, 1986; Dennis et al., 1998) after the intron endonuclease cut, but we could not corroborate that finding.

The 3’ maturation of 23S rRNA shows a similar pattern to the 5’ end. Probe I (Figure 5.16) shows a pattern akin to that of F in that a small processing intermediate is normally detectable, but is lost and the appearance of a signal close in size to the 23S rRNA when intron endonuclease is inhibited. Fortuitously, hybridizations from probe H (Figure 5.16) offer more insight than those conducted with probe G. An approximately 3500 nt species is detectable in total RNA isolated from cells cultured under normal conditions. When cells are cultured under inhibitory conditions this species is present along with a slightly larger RNA. This 3’ flanking sequence consists of 36 nt between the 23S rRNA 3’ end and the intron endonuclease cleavage site; in the absence of an intron endonuclease cut, additional sequence between the endonuclease cut site and the 5S
rRNA would add an additional 77 nt. The 5S rRNA is released by RNase Z (Holzle et al., 2008) and since there is no evidence that sequences 5’ of the 5S rRNA are associated with this RNA (Figure 5.11 And see below), it is unlikely that this larger species also carries the 5S rRNA region. As with the 5’ flanking sequence inside the intron endonuclease cutsite, there is no small intermediate observed, suggesting either an exonuclease or an endonuclease followed by exonuclease activity. Intron endonuclease activity has a clear and observable impact on the maturation of the 5’ and 3’ ends of the 23S rRNA.

5S rRNA is unimpacted by intron endonuclease

Studies have demonstrated that the release of the 5S rRNA’s 5’ end is facilitated by tRNase Z (Holzle et al., 2008). These in vitro studies argue that 5S excision is less efficient than 3’ end maturation of tRNA; however, the work presented in this study suggests that, in vivo, 5S rRNA release is rapid and efficient. We did not observe a signal when using probes designed to hybridize against the 5S rRNA at any size range other than the 122 nt of the 5S (Figure 5.11). This could be a limitation of the system used in this study as cells growing exponentially will already have an abundance of rRNA which could mask any minor signals from intermediates. Still, our data indicate that release of the 5S rRNA by tRNase Z is independent of intron endonuclease activity.

Suggested model for rRNA primary transcript processing

Placing this work into the context of rRNA processing in H. volcanii, our data build on the model in which intron endonuclease would first release the 16S rRNA from the primary transcript. Consistent with the findings of previous studies, secondary
endonucleolytic activity would then mature the 5’ and 3’ ends of the rRNA, though the enzyme(s) responsible for this activity has not yet been discovered (Chant and Dennis, 1986; Durovic and Dennis, 1994). Our data also agree with these studies that once tRNAAlaUGC has been transcribed, tRNase Z can cleave to mature the tRNA’s 3’ end while the 5’ leader would remain in place until after intron endonuclease cleavage. The 5’ leader would then be removed by RNase P. Further along the primary transcript, 23S rRNA would be released by intron endonuclease and the subsequent nuclease(s) could cause maturation of the 5’ and 3’ ends. While it has been argued that the final maturation of the 23S rRNA occurs with endonucleolytic cuts, the data presented here suggests that 23S rRNA maturation may rely on exonucleases after the intron endonuclease cut since no intermediates were readily detectable. Independently of all else, once the 5S leader sequence, the requisite recognition element for tRNase Z, the 5S rRNA could be released (Holzle et al., 2008). This appears to primarily occur prior to intron endonuclease cleavage since the flanking sequence 3’ of the cutsite could still be found attached to the 23S rRNA in RNA isolated from cells cultured in normal conditions whereas the 5S rRNA could not be found attached to the 23S rRNA using Northern analysis (Figures 5.19 and 5.20).

**Flanking sequences of the 23S rRNA released by intron endonuclease are ligated and may fulfill a box C/D role in H. volcanii**

A previous report demonstrated that the flanking sequences presumed released by intron endonuclease were ligated afterward (Tang et al., 2002). It was also noted that the
5.19 Schematic representation of *H. volcanii* 16S rRNA and tRNAAla maturation.

The primary transcript surrounding the 16S rRNA and tRNAAla appears to have at least seven processing sites. The 16S rRNA may be released directly by the activity of the unknown enzyme(s) indicated by “?” at its mature 5’ and 3’ ends, or the 16S rRNA may first be released by intron endonuclease at the sites marked “IE” followed by the unknown enzyme(s). Based on evidence from Figure 5.12, there seems to be an additional processing site just 5’ of the intron endonuclease site, though the enzyme responsible is unknown. 3’ maturation of the tRNAAla may occur at any time once the tRNase Z recognition site forms, but the activity of RNase P for 5’ maturation appears to be dependent on intron endonuclease cleaving first for greatest efficiency.
5.19 Schematic representation of *H. volcanii* 16S rRNA and tRNAAla maturation.
5.20 Schematic representation of *H. volcanii* 23S rRNA and tRNAAla maturation.

Both 5’ and 3’ end maturation of the 5S rRNA appear to be independent of intron endonuclease. The 5’ end of 5S rRNA may be released at any time once the tRNase Z recognition site forms. The unknown enzyme responsible for 3’ end maturation also appears to act independently of intron endonuclease. Maturation of the 23S rRNA 5’ and 3’ ends is accomplished by intron endonuclease cleavage to release the 23S rRNA at the sites designated “IE” followed by unknown nuclease activity to fully mature the 5’ and 3’ ends. The flanking sequences of the 23S rRNA, once released by intron endonuclease, are ligated, although the ligation product appears to be smaller than the total length of the combined flanking sequences (represented by the dotted line arrow and the <200 nt line below).
5.20 Schematic representation of *H. volcanii* 23S rRNA and tRNAAla maturation.
ligated product contained box C/D-like elements and could bind L7ae *in vitro* though a target was not identified for this putative guide. With this in mind, we assumed that the archaeal RNA ligase was not-tRNA dependent and investigated *H. volcanii* for this phenomenon. Since ligation of the flanking sequences can only occur when there is substrate present resulting from intron endonuclease activity, we analyzed total RNA from cells cultured in normal and inhibitory conditions to establish the necessity of the endonuclease cut. When using Northern hybridization with splice-junction probes (A’ and F’), there was only a product observed for the 23S rRNA ligated flanking sequences and only when normal intron endonuclease activity was present. Close examination of the splice junction probe, F’, revealed that there was a 6 nt direct repeat across the junction suggesting a false positive result. The product also appeared to match the size of the band detected with the F probe, but not the I probe. However there is a faint signal from probe I that could correspond to the ligated product. Overall, we were unable to resolve the question of flanking sequence ligation with Northern analysis.

To answer the question, the more sensitive method of RT-PCR was used. The results of RT-PCR agreed with the Northern hybridizations in that there was an obvious product from the primers used around the 23S rRNA, but not from around the 16S rRNA. Cloning and sequencing of the product revealed that it was the ligated flanking sequences joined precisely at the intron endonuclease cutsite. Though not conserved among the halophiles, the *H. volcanii* ligated product does contain the box C/D-like elements (Figure 5.17). There is less overall conservation of the 16S rRNA flanking sequences outside of the U3-like element and a lack of recognizable sequence elements, box C/D or
otherwise. Interestingly, processing intermediates from the 16S rRNA flanking sequences are detectable, but we found no evidence that they are ligated in *H. volcanii*. Whether ligation results in meaningful products from these flanking sequences or if they are just the result of having ligase substrate recognition elements remains to be shown. Identification of a target for the box C/D-like 23S rRNA ligated flanking sequences would help explain the abundance of what would normally be considered little more than ancillary leftovers of processing (Figure 5.18).
Perspectives

The work presented in this thesis allows for some conclusions to be drawn about the research questions addressed, though new questions have developed based on the now-available tools generated in the above.

With respect to the C34 modification of tRNA\textsubscript{Met} discussed in Chapter 3, evidence now shows that the intron-containing pre-tRNA is a substrate for modification \textit{in vivo}. Furthermore, it seems to be the preferred substrate, but no evidence as of yet can explain why this would be the case. It has been shown here that the modification at this position does not impact splicing and is therefore not a method to regulate the intron removal phase. However, modifying this position in the precursor may still be used as a regulatory step toward other modifications or sensing metabolites. The strains developed give a baseline to begin addressing some of these questions and whether or not there is a redundant mechanism for tRNA\textsubscript{Met} C34 modification.

The L7\textsubscript{ae} protein provided an interesting test of whether proteins could be over-expressed natively and whether RNA-binding proteins could be purified from the cell with their RNAs bound. Known classes of L7\textsubscript{ae}-binding RNAs were co-eluted with L7\textsubscript{ae} protein, yet the investigation for novel RNAs has just begun. An extension of this system wherein purified L7\textsubscript{ae} could then be bound to the column and \textit{H. volcanii} total RNA could be passed over the bound protein. As the protein has been over-expressed, it
remains likely that there is a sizeable pool of L7ae that is not binding RNA and passage of total RNA over the “free protein” would enable the capture of potentially novel L7ae-binding RNAs which could be subsequently identified. This may result in the discovery of novel RNAs in established classes or in the discovery of novel RNA classes. Additionally, detailed analysis of the protein fraction co-eluting with L7ae would allow for verification of whether or not whole complexes could be captured with this system. This may allow for a method of pulling out complete complexes of RNase P or SRP and allow for the identification of novel associated proteins in addition to allowing for the discovery of novel L7ae-containing sRNP complexes at the protein level.

Generation of the intron endonuclease mutant strain enabled in vivo verification of the enzymes role in maturation of the intron-containing tRNAs and in the release of the 16S rRNA and 23S rRNA from the rRNA primary transcript. End-labeling with $^{32}$pCp showed a clear change in the total RNA population, but the changes in band patterns have not been specifically correlated to individual RNAs yet. Sequencing these bands could reveal additional RNAs that are dependent on the intron endonuclease for maturation such as the ligated product of the 23S rRNA flankers. The archaeal tRNA intron endonuclease may be shown to be far more versatile than previously thought implicating that its apparent essentiality is not solely resultant from its action on the maturation of tRNA and rRNA.
References


Benitez-Paez, A., Villarroya, M., Douthwaite, S., Gabaldon, T., and Armengod, M.E. (2010). YibK is the 2'-O-methyltransferase TrmL that modifies the wobble nucleotide in Escherichia coli tRNA(Leu) isoacceptors. RNA 16, 2131-2143.


