ABSTRACT

PANNUCCI, JAMES A. Archaeal RNase P RNAs are Catalytically Active. (Under the direction of James W. Brown.)

RNase P is the endoribonuclease responsible for processing the 5′ termini of precursor-tRNAs. In Bacteria, RNase P consists of two subunits, a catalytic RNA and a 14 kD protein. Physiological ionic conditions require the presence of both subunits for pre-tRNA processing, but under appropriate in vitro conditions (increased ionic strength), the RNA is catalytically active without the RNase P protein. The Archaea, although more closely related to eukaryotes, have RNase P RNAs that are structurally similar to Bacteria.

RNase P RNAs from certain archaeal groups, the methanobacteria, halobacteria, and thermococci, can process pre-tRNAs without RNase P protein when extremely high concentrations of monovalent and divalent cations are used in vitro. RNA-alone catalytic activity was confirmed by in vitro RNA synthesis, size fractionation, and selective RNase H degradation of RNase P RNA. Synthetic archaeal RNase P RNAs that were active in vitro were also active when RNA was extracted from cell lysates of the same archaeal species. RNase P RNAs that were not active after synthetic production, they were not active when extracted from cells lysates. Catalytic activity from RNA synthesized in vitro suggests that base modifications are not a catalytic requirement.

Kinetic properties of the archaeal RNase P RNA reaction were characterized using the RNase P RNA from Methanobacterium thermoautotrophicum strain ΔH. Poor substrate affinity and extreme salt requirements resemble the functional impairments of severe structurally defective bacterial RNase P RNA mutants. Structure probes of the M. thermoautotrophicum strain ΔH RNase P RNA demonstrate the presence of bacteria-like RNase P RNA tertiary structure, but also confirms structural instability.
ARCHAEOAL RNASE P RNAs ARE CATALYTICALLY ACTIVE

by

JAMES A. PANNUCCI

A thesis submitted to the Graduate Faculty of
the North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

MICROBIOLOGY

Raleigh
1998

APPROVED BY:

________________________

________________________
Chair of Advisory Committee
ACKNOWLEDGMENTS

Completion of my degree was made possible by the advice, assistance, and cooperation of many people. I wish to express my gratitude in a way that would acknowledge all contributions without rank because regardless of size, I consider every effort made on my behalf to be an essential part of my success at N.C. State. Thank you.

Jim
BIOGRAPHY

Graduated from Columbia High School, Maplewood, NJ 1988

B.S. Degree in Pathobiology, University of Connecticut, Storrs, CT 1992

Employed at the Schering-Plough Research Institute, Kenilworth, NJ 1992-1995

M.S. Degree in Biology, Seton Hall University, South Orange, NJ 1994

Ph.D. Degree in Microbiology, North Carolina State University, Raleigh, NC 1998
# TABLE OF CONTENTS

**LIST OF FIGURES** ...................................................................................... vi

**PREFACE** ........................................................................................................... viii

**A REVIEW OF THE RNASE P RNA** ................................................................. 1

- Structure and Function ............................................................................... 1
- Achievements .............................................................................................. 15
- Future Trends in RNase P Research ......................................................... 19
- Methods ....................................................................................................... 20
- Conclusions ................................................................................................. 30
- Figures .......................................................................................................... 31
- References .................................................................................................... 45

**RNASE P RNAS FROM CERTAIN ARCHAEAE ARE CATALYTICALLY ACTIVE** ........... 53

- Abstract ....................................................................................................... 53
- Introduction ................................................................................................. 53
- Materials and Methods ............................................................................ 55
- Results and Discussion ............................................................................ 59
- Figures .......................................................................................................... 64
- References .................................................................................................... 68

**BIOCHEMICAL CHARACTERIZATION OF M. THERMOAUTOTROPHICUM STRAIN ΔH RNase P RNA** .................................................................................. 71

- Abstract ....................................................................................................... 71
- Introduction ................................................................................................. 71
- Materials and Methods ............................................................................ 72
- Results .......................................................................................................... 75
- Discussion .................................................................................................... 79
- Figures .......................................................................................................... 84
- References .................................................................................................... 95

**APPENDIX 1**

- The effect of spermidine on the *M. thermoaototrophicum* strain ΔH RNase P RNA ................................................................................................................. 98

**APPENDIX 2**

- Site directed mutagenesis to repair the *M. barkeri* unimolecular PT construct ...... 101
TABLE OF CONTENTS (con’t)

APPENDIX 3

Characterization of the archaeal unimolecular constructs ............................ 107

APPENDIX 4

Figure A4. End-mapping of the *M. jannaschii* RNase P RNA ........................ 110
LIST OF FIGURES

A REVIEW OF THE RNase P RNA

Figure 1. *E. coli* and *B. subtilis* RNase P RNA secondary structures .......... 31
Figure 2. Examples of RNase P substage secondary structures .................. 32
Figure 3. Secondary structure of an *in vitro* active archael RNase P RNA ....... 33
Figure 4. Secondary structures of minimal RNase P RNAs ....................... 34
Figure 5. RNase P RNA consensus structures ...................................... 35
Figure 6. Secondary structures of RNase P RNA from organelles .................. 36
Figure 7. A three-domain tree of life .............................................. 37
Figure 8. Three-dimensional model of bacterial RNase P RNA ................... 38
Figure 9. External guide sequence RNA ............................................ 39
Figure 10. Tandem repeat method of cloning a circular permuted RNase P RNA .... 40
Figure 11. Two PCR method of cloning a circular permuted RNase P RNA ........ 41
Figure 12. An example of an RNase P RNA assay ................................ 42
Figure 13. Partial alignment of bacterial RNase P RNAs ........................... 43
Figure 14. Mutual information plot of the bacterial RNase P RNA alignment ..... 44

RNASE P RNAs FROM CERTAIN ARCHAEA ARE CATALYTICALLY ACTIVE

Figure 1. Comparison of archael and bacterial RNase P RNAs .................... 64
Figure 2. Catalytic activity of synthetic archael RNase P RNAs ................. 65
Figure 3. Oligonucleotide-directed RNase H depletion of RNase P RNAs ....... 66
Figure 4. Reconstitution of active chimeric holoenzymes ........................ 67

BIOCHEMICAL CHARACTERIZATION OF *M. THERMOTROPHICUM* STRAIN ΔH RNASE P RNA

Figure 1. RNase P RNA consensus structures ...................................... 84
Figure 2. A three-domain tree of life .............................................. 85
Figure 3. Secondary structures of *M. formicicum* unimolecular constructs .... 86
Figure 4. Secondary structures of *M. barkeri* unimolecular constructs ........ 87
Figure 5. Ionic optimization of *M. thermoautotrophicum* strain ΔH RNase P RNA catalytic activity ........................................ 88
Figure 6. Temperature and pH optimization of *M. thermoautotrophicum* strain ΔH RNase P RNA reaction ............................................ 89
Figure 7. Mn** versus Mg** graph of catalytic activity ............................ 90
Figure 8. Enzyme and substrate concentration dependence ....................... 91
Figure 9. Time course of catalytic activity of *M. thermoautotrophicum* strain ΔH RNase P RNA and the *M. formicicum* TP construct ............. 92
Figure 10. Secondary structures and metal ion cleavage sites of *E. coli* and *M. thermoautotrophicum* strain ΔH RNase P RNA ....................... 93
Figure 11. Pb cleavage of gel *E. coli*, *M. thermoautotrophicum* strain ΔH, and *M. sedula* RNase P RNAs ........................................... 94
LIST OF FIGURES (con’t)

APPENDIX 1

Figure A1. The effect of the spermidine on the *M. thermoautotrophicum* strain ΔH RNase P RNA ...................................................... 100

APPENDIX 2

Figure A2. Characterization of the *M. formicicum* unimolecular construct, TP ...... 106

APPENDIX 3

Figure A3. Sequence repair of the *M. barkeri* unimolecular PT construct .......... 109

APPENDIX 4

Figure A4. End-mapping of the *M. jannaschii* RNase P RNA ........................ 112
PREFACE

The body of this thesis is arranged in three parts - A REVIEW OF THE RNase P RNA, RNase P RNAs FROM CERTAIN ARCHAEA ARE CATALYTICALLY ACTIVE, and BIOCHEMICAL CHARACTERIZATION OF *M. THERMOAUTOTROPHICUM* STRAIN ΔH RNase P RNA. The review contains data from elsewhere in the thesis, which is not usually included in historical reviews for a thesis, but appears because the review will be published in the forthcoming book, *Ribozymes. Biology and Biotechnology* (BioTechniques Books, G. Krupp, ed.). The second and third chapters were combined into a manuscript that was submitted to the Proceedings of the National Academy of Science - USA. Since the second and third chapters represent manuscript material, they also contain experimental contributions from others. In the chapter, RNase P RNAs FROM CERTAIN ARCHAEA ARE CATALYTICALLY ACTIVE, E.S. Haas contributed cell lysate and extracted RNA experiments. T. Hall contributed bacterial RNase P protein reconstitution experiments. In the chapter, BIOCHEMICAL CHARACTERIZATION OF *M. THERMOAUTOTROPHICUM* STRAIN ΔH RNase P RNA, E.S. Haas contributed experiments with the *M. barkeri* unimolecular (TP) construct.
A REVIEW OF THE RNase P RNA

STRUCTURE AND FUNCTION

Background

Ribonuclease P is the enzyme responsible for the specific removal of 5' leader sequences from precursor transfer RNAs in all cells and organelles that synthesize tRNA (3, 41, 60). The subunit composition of RNase P is particularly interesting. RNase P has a highly structured RNA (250-450 nucleotides) that is complemented by a protein subunit that ranges from a single 14kD polypeptide in Bacteria to a large, multiple protein complex in the yeast nucleus (13, 17, 46, 58). Previous work with bacterial RNase P has demonstrated the RNA as the catalytic subunit. Physiological conditions require the presence of both subunits for efficient pre-tRNA processing, but under appropriate in vitro conditions (increased ionic strength), the RNase P RNAs from Bacteria and certain Archaea (this thesis) are catalytically active without the RNase P protein (32, 62, 65).

In vitro RNA-alone activity has not been demonstrated by RNase P RNAs from eukaryotic nuclei, organelles, and several Archaea, but the RNase P RNAs are still presumed to be the catalytic subunit based on a shared and highly conserved set of functional core components (11, 18, 27, 40). The tertiary structure of the RNase P RNA, formed by the interaction of conserved and variably present secondary structural elements, is undoubtedly the key to catalytic activity.

In cells, the RNase P RNA is bound by its protein subunit which is largely responsible for shielding the electrostatic repulsion of the RNA backbone and enabling RNA tertiary folding (65). RNA-alone catalysis requires structural elements to be arranged in a proper tertiary conformation so that the global architecture is sufficient to correctly position both
the enzyme and substrate, but even if all key structural features are present, incompatible global arrangements can prevent enzymatic function. The *in vitro* phenomenon of RNA-alone catalytic activity, and the conditions required to produce that activity, demonstrates the ability of an RNase P RNA to self-fold, and that it contains all the necessary structural features to bind pre-tRNAs and cleave 5' leaders (32).

In this chapter, the structure and function of the RNase P RNA will be reviewed with consideration towards the common RNase P structural core, its role in catalysis, and how it is effected by variably present structural elements and the RNase P protein subunit. *(Special note: Reference will be made to the RNase P RNA secondary structures from many species. A general structure figure is included (Fig. 1), but readers may wish to view specific secondary structures by visiting the RNase P Database web site, www.mbio.ncsu.edu/RNaseP)*

**Substrate recognition and cleavage site selection**

The highly-defined secondary and tertiary structure of tRNA molecules has been applied to understanding the substrate recognition and cleavage site selection requirements of RNase P. *In vivo*, RNase P processes all the pre-tRNAs of a given cell or organelle. *In vitro*, RNase P will process pre-tRNAs from any species (2, 22, 60). The RNase P interaction with substrate depends on the highly conserved features of the mature tRNA domain rather than aspects of the precursor that may vary with each specific form (16). Substrate interaction experiments have identified the acceptor stem and the TΨC stem-loop as the critical recognition structures (Fig. 2). The D and anticodon stem-loops serve as additional, but catalytically unnecessary, anchors (16, 79).

The acceptor stem of almost all pre-tRNAs contains a 7 base pair helix which appears critical for correct cleavage site selection (15). Analysis of RNase P cleavage of circularly
permuted substrates (Circularly permuted pre-tRNAs have 5' and 3' ends generated in unusual locations after circularization and hydrolysis of substrate RNA) shows that interruptions in the pre-tRNA sugar-phosphate backbone are not generally tolerated in the tRNA acceptor stem or in the 5'P3'C stem-loop (16). RNase P may measure the stem length to find the appropriate cleavage site (15). Using the RNase P from the *Xenopus laevis* oocyte nucleus, a pre-tRNA<sup>pre</sup> substrate with two base pairs inserted in the acceptor stem has cleavage shifted two base pairs 3' to maintain the seven base pair acceptor stem. In contrast, the *E. coli* RNase P (both the RNA-alone and the holoenzyme) cleaves a pre-tRNA<sup>Tyr</sup> derivative with an insertion of three base pairs in the acceptor stem at the same base cleaved in the wild type pre-tRNA<sup>Tyr</sup> leaving a 10 base pair acceptor stem (72). Removal of 4 base pairs from the acceptor stem of tRNA<sup>Tyr</sup> by changing nucleotide identity to leave single-stranded 5' and 3' ends will be cleaved by the *E. coli* RNase P RNA, but aberrant cleavage shows that the identity of the acceptor stem bases is important. (72).

Acceptor stem length was demonstrated as an important determinate of substrate interaction and RNase P cleavage site selection. Changes to the number of acceptor stem bases affects catalytic efficiency of substrate processing, (alters K<sub>m</sub> and k<sub>cat</sub>) but sequence as well as acceptor stem length is important for cleavage site selection (43, 72). A G nucleotide at the site of cleavage is an important primary structural feature. Substitutions at the cleavage site G have resulted in aberrant cleavage, usually 1 base in the 3' direction in the RNA-alone reaction; however, a G to A nucleotide substitution at the cleavage site of pre-tRNA<sup>Tyr</sup> will be correctly cleaved by the *E. coli* RNase P holoenzyme (43, 73). A tRNA<sup>Ser</sup> variant (tRNA<sup>Ser</sup> Su1) (73) is dependent on the identity of the nucleotide at position -2, but in parallel experiments tRNA<sup>Tyr</sup> is not. The CCA tail, present on all tRNAs (discussed below) also plays a significant role in cleavage site selection. For the tRNA<sup>Ser</sup>
Su1 variant, shifting the CCA tail along the 3' acceptor stem sequence causes a compensatory shift in cleavage site location.

Pre-tRNAs can be covalently crosslinked to various positions on the RNase P RNA (see methods). The crosslinks are dependent on the tertiary arrangement of both the enzyme and substrate RNAs, but structural conformation of the RNAs is not the only important consideration. There are specific Watson-Crick base pairs formed between the enzyme and substrate RNAs. Bases G292 and G293 of the P15/16 loop region of the E. coli RNase P RNA (Fig. 1) base pair with bases +75 and +74 (+76 and +75 of Fig. 2) of the substrate RNA located in the CCA tail (44). Also possible is the extension of the base pairing to include the RNase P RNA base U294 if there is a corresponding purine in the tRNA at position +73 (Fig 2 base +74). The presence of the base pair interactions have been supported by compensatory mutations (44). The tRNA -2, -1, +1: +72, +73, +74 region is thought to be a Mg** binding site and important for RNase P cleavage activity (72). The formation of base pairs near the site of RNase P cleavage on the substrate may position and expose the cleavage site by unwinding the acceptor stem helix (44).

CCA tails are present in all pre-tRNAs. E. coli and many others species have the CCA sequence encoded in most tRNA genes, but some have the CCA added by a separate enzymatic system. A CCA tail is not absolutely necessary for cleavage, but it does influence rate and cleavage site positioning (72, 73). E. coli RNase P RNAs with mutations in the G291-A295 sequence (the P15/16 loop region that base pairs to the tRNA CCA) exhibits aberrant cleavage (44). The P15/16 loop region of RNase P RNAs is not universally conserved. In the Archaea, many of the RNase P RNAs without the P15/16 loop GGU sequence are not catalytically active in vitro. The idea that the RNase P protein subunit compensates for the absence of CCA base pairing is appealing. However, there are examples of bacterial (e.g. cyanobacteria and chlamydia) as well as archaeal RNase P
RNAs (e.g. *Thermococcus celer*) that do not have P15/16 loop sequences complimentary to CCA, but perform *in vitro* catalysis accurately without a protein subunit (35, 75).

Full-length pre-tRNAs are not the only substrate recognized by RNase P in both the RNA and holoenzyme form. Bacterial RNase P will cleave an artificial minimal substrate that consists of just the coaxial stack of the acceptor stem and TΨC stem-loop connected directly to each other by deletion of the D and anticodon stem-loops (16, 79). *X. laevis* oocyte nuclear RNase P will not cleave the prokaryotic minimal substrate, but addition of at least five bases in the junction of the 5′ acceptor stem and TΨC stem-loop restores activity (16).

Artificial and naturally occurring alternate RNase P substrates have much in common. The 4.5S RNA and the SRP RNAs of Archaea and Eukaryotes are natural, non-pre-tRNA RNase P substrates that mimic the structure of the tRNA acceptor stem stacked on the TΨC stem-loop (Fig. 2) (45). The 4.5S RNase P reaction has a higher $K_m$ and lower $k_{cat}$ compared to pre-tRNAs; it is not an ideal substrate, but it is still sufficiently cleaved to satisfy cellular needs. The 10Sa RNA and certain plant viral RNA (i.e. the turnip yellow, eggplant, and cucumber mosaic viruses) have tRNA structural mimics near the ends of RNA sequences that are processed by RNase P (33, 45, 54). Two hypotheses of RNase P processing in plant viral RNAs have been suggested; one is that RNase P is associated with a host defense function, and the other suggests RNase P activity as part of multi-step RNA editing event. The 10Sa RNA tRNA domain is likely processed by RNase P to facilitate interaction with the ribosome and its role in abortive polypeptide decomposition.

A natural substrate for RNase P has been identified in *Drosophila* (42) after infection by a 5kb retrovirus-like RNA. In all other instances of retrovirus infection, full length, mature tRNAs are used as primers for minus strand reverse transcription, but in the *Drosophila* retrovirus-like infection, a fragment of the mature initiator tRNA$^{Met}$ corresponding to the 5′
39 bases is used to prime reverse transcription. Although convincing evidence is still forthcoming, tRNA$_{\text{Met}}$ is thought to assume an alternate conformation that is still an RNase P substrate. The alternate fold most likely arranges the acceptor stem and D stem loop in a coaxial stack recognized by RNase P; cleavage occurs between bases 39 and 40 of the mature tRNA$_{\text{Met}}$ after the initial RNase P maturation of the 5' end occurs. The 5' 39 bases, freed from the 3' section of the molecule, bind the Drosophila retrovirus-like RNA and prime reverse transcription.

**Catalytic mechanism**

RNase P catalyzes a direct hydrolysis rather then a transphosphorlyation reaction that occurs in some other ribozymes. Beebe and Fierke (6) proposed a four step model of RNase P RNA catalysis based on the RNase P RNA and pre-tRNA$^{\text{Ap}}$ from Bacillus subtilis. The cycle begins with rapid binding of pre-tRNA by the RNase P RNA followed cleavage of a phosphodiester bond. Dissociation of products occurs in a kinetically favored pathway; the 5' fragment leaves first followed by the rate limiting tRNA dissociation. The final step is the formation of a second conformer of RNase P RNA that is less stable and binds pre-tRNA significantly more slowly. The second conformation must convert prior to catalysis (20) and that most likely occurs after the formation of the initial enzyme:substrate collision complex (6). A conformational change comparable to the insertion or docking of the H-bonded substrate region into the catalytic site of Tetrahymena group I intron is a plausible example for many nucleic acid enzymes including RNase P (20).

The *E. coli* DNA Polymerase I exonuclease domain was proposed as a model for RNase P catalyzed phosphodiester cleavage (70). An essential Mg$^{2+}$ is coordinated directly to the pro-Rp phosphate group oxygen and indirectly to two other ligands near the scissile bond. One ligand is the upstream ribose 2' hydroxyl and the other is the downstream
purine N7. The catalytic Mg$^{++}$ most likely positions and deprotonates H$_2$O for in-line nucleophilic attack on the scissile bond phosphate (20, 70). Hydroxide oxygens attack the phosphate group immediately 5' to mature tRNA domain of substrate. The bond between the phosphate and O3' (3' oxygen of upstream ribose) is broken and the upstream moiety becomes the leaving group. The RNase P reaction leaves the mature tRNAs with 5' phosphates and leader sequences with 3' hydroxyls.

A photocrosslinking assay (described in methods section) was used to determine the effect of metal ions on RNase P substrate affinity (70). Combined with data from RNase P activity assays, crosslinking of bacterial RNase P RNA to various pre-tRNAs suggests that monovalent ions primarily reduce electrostatic repulsion and facilitate substrate binding. The ionic radius of different monovalent ions had an effect on the RNase P RNA. Cs$^+$, K$^+$, NH$_4^+$ and Rb$^+$ have ionic radii of 1.3-1.7 angstroms. Ions with radii outside that range do not have a simple correlation between binding and cleaving of substrate. Specifically, Na$^+$ has an ionic radius of 0.95 angstroms, and is three-fold less effective than NH$_4^+$ for binding substrate at a concentration of 1M, but at 1M Na$^+$ affords a three-fold faster cleavage rate (70).

Divalent ions are required for cleavage, but not formation of enzyme:substrate complexes. The role of divalent cations is primarily catalytic. Reaction velocity increases with Mg$^{++}$ concentration until affinity between product-tRNA and the RNase P RNA slows product release and lowers reaction rates in multiple turnover reactions. Ca$^{++}$ allows RNase P cleavage of substrate to occur, but at slower rates than Mg$^{++}$. Mn$^{++}$ is also able to catalyze RNase P reactions. Mn$^{++}$ has a greater nucleotide affinity then Mg$^{++}$ and is more effective at lower concentrations, but the denaturing effect caused tight binding of Mn$^{++}$ to nucleotide bases and phosphates reduces substrate binding and catalytic rate (70).
Various ions were also tested (this thesis) for their effect on the archaeal *in vitro* RNase P RNA reaction (62). The monovalent ions, K\(^+\), Na\(^+\), and NH\(_4^+\), stabilized the *M. thermoautotrophicum* strain ΔH RNase P RNA (Fig. 3 for secondary structure) and Mn\(^{++}\) could enable pre-tRNA processing comparably to Mg\(^{++}\). Like bacterial RNase P RNA, Mn\(^{++}\) was more effective in the archaeal RNase P RNA reaction at lower concentrations than Mg\(^{++}\), but it also increased RNA degradation. Polyamines, such as spermine and spermidine, were tested for their ability to substitute or aid monovalent ions in RNase P RNA reactions, but no increased activity was observed in either bacterial or archaeal RNase P RNA reactions (70, 62, this thesis). In general, the extreme salt requirements and high \(K_m\) of catalytically active archaeal RNase P RNAs resemble defective bacterial RNase P RNA mutants, but archaeal RNase P RNAs are clearly homologous to bacterial RNase P RNAs, share core structural components, and respond in a similar manner to various ions.

The role of monovalent ions is primarily shielding of electrostatic repulsion, but divalent cations have three apparent functions; to facilitate the deprotonization of a H\(_2\)O molecule, to polarize a phosphoryl oxygen in order to increase electropositivity of the corresponding phosphorous, and to coordinate that phosphoryl oxygen to stabilize the transition state (20). For RNase P RNA to bind divalent ions and cleave pre-tRNA its tertiary structure has to be stabilly arranged. Although the cleavage mechanism depends heavily on divalent ion chemistry, it is the RNase P RNA that binds and positions the divalent ions so cleavage can occur, and it is possible for non-catalytically active RNase P RNA to have regions of the molecule in a conformation able to bind Mg\(^{++}\), but with a global arrangement not compatible with catalytic activity. Enzymatic chemistry depends on the catalytically active structure of the RNase P RNA.

**Variable elements, conserved core, and consensus structures**
Although *in vitro* catalytic activity has only been demonstrated by bacterial and certain archaeal RNase P RNAs (this thesis), the RNA subunit is presumed to be the catalytic entity in all cases where it is present (a protein RNase P has been proposed for certain cases where an RNA subunit cannot be found (66, 67)). Structural conservation, in particular of the catalytic core of the molecule, suggests functional conservation (18).

For an RNase P RNA to be active *in vitro*, all components have to be arranged in a catalytically competent global configuration. *In vitro* RNA-alone reactions provide information about which structural components are necessary versus beneficial, or what structural role is accomplished by the RNase P RNA or RNase P protein. The presence or absence of non-core structural elements can alter the ability RNase P RNA to function *in vitro*. An RNase P RNA that is inherently unstable, and fails to be properly stabilized by salt, will have a low affinity for substrate with the possibility of poor or no *in vitro* catalysis (62).

In contrast to the core RNase P RNA structural components is the variety in primary structure and peripheral secondary structural features. Much of the structural identity of RNase P RNAs has been determined by comparative analysis, but confirmation has been provided by physical experiments designed to probe RNase P RNAs for structural features (see methods) (37). A survey of bacterial RNase P RNAs reveals 24 different helical elements that occur in two forms of the molecule. Type A RNase P RNAs, such as the RNase P RNA from *E. coli.*, are the most common and represent the ancestral form. The type B RNase P RNA occurs in low G+C Gram positive Bacteria such as *Bacillus subtilis* (Fig. 1) and were a late evolving structural form. Although no single RNase P RNA contains all the possible helical variations, each contains the same fundamental core elements. The minimal bacterial consensus structure contains the helices, P1-5, P7-11, and P15 (Fig. 4). Helices P12 and P18 are present in >98% of the known bacterial sequences.
Variable helical elements often protrude from regions of the RNase P RNA core and are therefore most highly conserved at their foundations. Terminal loop sequences are also conservative because they can serve as structural anchors via tertiary interactions with nucleotides elsewhere in the molecule (37). The helices P8, P9, P13, 14, P16/17, and P18 of bacterial RNase P RNAs are examples of helices with conserved foundations and terminal loops. P9 originates in the conserved cruciform (P7-P10) and docks with nucleotides in P1. P8 docks with P4 (55). P13 is known to interact with P12; P14 with P8 (separately from the P8:P4 docking). P16/17 is located distal to the P15 loop region (which is part of the minimal bacterial consensus structure) and interacts with sequence between P5 and P7 to form P6. P6 is present in many type A RNase P RNAs, but it is not part of the conserved functional core. In the type B RNase P RNAs, P6 and P17 are replaced by P5.1, which is also not part of the minimal consensus. P18 extends upstream from a highly conserved region adjacent to the central loop and P4, and is an important stabilizing structure (see below).

The nucleotide sequences of RNase P RNAs range from 100% to 24% identical. Variation is often seen as insertion or deletion of non-core structural features and in the sequence of the stems of helices (37). Non-essential structural features are, by definition, not required for catalysis, but are often important as elements of structural stability.

The native Chlorobium sp. RNase P RNAs have a $K_m$ similar to the native E. coli RNase P RNA under similar conditions. The Chlorobium RNase P RNAs do not have a P18, a structure that is highly conserved among Bacteria and when not present in the E. coli RNase P RNA (E. coliΔP18 deletion mutant) increases the monovalent salt requirement (34). Mutant bacterial RNase P RNAs and the catalytically active archaeal RNase P RNAs (this thesis) require increased ionic conditions for structural stabilization. The absence of P18 from the Chlorobium RNase P RNA means it is not required for
highly conserved region in Bacteria and demonstrated to directly base pair with the CCA 3' sequence of pre-tRNAs. Interestingly, many of the archaeal RNase P RNAs that have catalytic activity in vitro have the P15 loop with the GGU sequence present in Bacteria (this thesis). The archaeal consensus structure is likely to be refined in the future because there are considerably fewer archaeal sequences than bacterial (30 vs. 150) (7).

The collection of eukaryotic RNase P RNA sequences is steadily increasing, but a comprehensive consensus structure has not been reported. Based on a vertebrate RNase P RNA consensus structure constructed from an alignment of 17 sequences, eukaryotic RNase P RNAs do not have P15-P18, but they do have a large internal loop in P3 (Fig. 5) (63). The P15-P18 region is also absent in a recently reported secondary structure of the Schizosaccharomyces pombe RNase P RNA (18). The loss of P15/16 and associated loop removes a highly conserved Mg++ binding site (present in both the bacterial and archaeal consensus structures) and the location of the tRNA CCA base pairing. The eukaryal RNase P protein may compensate for the loss of RNA structure; it has been speculated that the loop in P3 may fulfill this role (18).

RNase P from organelles closely resembles the RNase P RNA structure of their prokaryotic ancestors, in certain instances (Fig. 6). For example, the RNase P RNA from photosynthetic organelles of Cyanophora paradoxa contains many of the helices commonly found in bacterial RNase P RNA; furthermore, the C. paradoxa cyanelle RNase P RNA has an elongated A+U rich P15 loop region like that of the cyanobacterial RNase P RNA (5). Other examples of prokaryotic similarity is observed in mitochondrial RNase P RNAs. The mitochondrial RNase P RNAs from Reclinomonas americana, Nephroselmis olivacea, and Jakoba libera (Lang, unpub.) are similar to bacterial A-form RNase P RNAs; they include many of the variable helical elements seen in all prokaryotic RNase P RNA plus unique
highly conserved region in Bacteria and demonstrated to directly base pair with the CCA 3’ sequence of pre-tRNAs. Interestingly, many of the archaeal RNase P RNAs that have catalytic activity *in vitro* have the P15 loop with the GGU sequence present in Bacteria (this thesis). The archaeal consensus structure is likely to be refined in the future because there are considerably fewer archaeal sequences than bacterial (30 vs. 150) (7).

The collection of eukaryotic RNase P RNA sequences is steadily increasing, but a comprehensive consensus structure has not been reported. Based on a vertebrate RNase P RNA consensus structure constructed from an alignment of 17 sequences, eukaryotic RNase P RNAs do not have P15-P18, but they do have a large internal loop in P3 (Fig. 5) (63). The P15-P18 region is also absent in a recently reported secondary structure of the *Schizosaccharomyces pombe* RNase P RNA (18). The loss of P15/16 and associated loop removes a highly conserved Mg$^{2+}$ binding site (present in both the bacterial and archaeal consensus structures) and the location of the tRNA CCA base pairing. The eukaryal RNase P protein may compensate for the loss of RNA structure; it has been speculated that the loop in P3 may fulfill this role (18).

RNase P from organelles closely resembles the RNase P RNA structure of their prokaryotic ancestors, in certain instances (Fig. 6). For example, the RNase P RNA from photosynthetic organelles of *Cyanophora paradoxa* contains many of the helices commonly found in bacterial RNase P RNA; furthermore, the *C. paradoxa* cyanelle RNase P RNA has an elongated A+U rich P15 loop region like that of the cyanobacterial RNase P RNA (5). Other examples of prokaryotic similarity is observed in mitochondrial RNase P RNAs. The mitochondrial RNase P RNAs from *Reclinomonas american*, *Nephroselmis olivacea*, and *Jakoba libera* (Lang, unpub.) are similar to bacterial A-form RNase P RNAs; they include many of the variable helical elements seen in all prokaryotic RNase P RNA plus unique
structural adaptations. The mitochondrial RNase P RNA sequence and structural similarity to Bacteria was cited as evidence of their evolutionary relatedness to α-proteobacteria (50).

Other mitochondrial RNase P RNAs pose an interesting challenge to the study of RNase P structure. The mitochondrial RNase P RNA from *Rhizopus stolonifer* has the basic set of RNase P RNA-type helices interrupted by the insertion of 575 nucleotides divided among two locations (P12 and P15 regions) (Lang, unpub.). The mitochondrial RNase P RNA from *Mucor mucedo* is missing the entire region between P15 and P18, and has 55 extra nucleotides at the distal end of P5.1 (Lang, unpub.). The *Aspergillus nidulans* mitochondrial RNase P RNAs is missing P3 (which is part of the bacterial minimum consensus) and the helices P15 through P18 (52, 53). Yeast mitochondrial RNase P RNAs are nearly unrecognizable and contain only a few typical RNase P helices, in addition to being extremely A+U rich (85%) (31). The *Saccharomyces cerevisiae* mitochondrial RNase P RNA has P1, P4, a helix analogous to P18, and 285 uninterpreted nucleotides (21).

Insertions, deletions, and unusual base composition make it difficult to determine a consensus for mitochondrial RNase P RNA. In addition, structural representations can be unreliable because organelle RNase P RNA sequences are usually forced into bacterial structural models without the benefit of meaningful comparative analysis. Organelle RNase P RNAs are equally difficult in physical experiments; their unique sequence and structural character make identification a challenge, which becomes no less difficult when confronted with the possibility of an all-protein RNase P, as postulated for the human mitochondria and vascular plant chloroplasts, where RNAs have not been co-purified with RNase P activity (66).

RNase P RNAs are evolutionarily homologous and are expected to have common structural foundations in the catalytic core of the molecule (18). Certain organelle RNase P RNAs may be worthy tests of this rule, but the lack of sequences and reliable structures
require postponement of judgment. An alignment of 10 evolutionarily diverse RNase P RNA sequences, with representatives from all three evolutionary domains and one mitochondrial sequence, revealed five conserved regions (CR I-V) in RNase P RNAs (18). The conserved regions are all located in the conserved core of the RNase P RNA, and are present in both the archaeal and bacterial minimal consensus structures. CR-I (GNAANNUC) is located in the 5′ portion of P4; CR-II (AGNRA) is in the universal internal loop preceding P12, then CR-III (UGNRA) is in the same loop but 3′ of P12; CR-IV (AGNNNNAU) is adjacent to the 5′ P4 sequence, just 3′ of P18, and finally, CR-V (ACNNRANNNNGNNUA) is the 3′ sequence of P4 (Fig. 5).

The RNase P protein subunit

The bacterial RNase P subunit is comprised of just one polypeptide, and it appears to be unrelated in sequence to the 10 polypeptides associated with the yeast nuclear RNase P protein (17, 65). The yeast proteins were isolated from Saccharomyces cerevisiae and several are shared by the mitochondrial ribosomal processing ribonuclease (RNase MRP). A crystal structure of a eukaryotic RNase P protein has not been reported so it is not known if there are motifs similar to the bacterial protein subunit. An archaeal RNase P protein subunit has not been purified, and comparison of the three completely sequenced archaeal genomes with the known bacterial RNase P proteins has not produced a homologous gene product (Hall & Brown unpub.).

The biochemical analysis of the function of the RNase P protein was an obvious curiosity after demonstration of the RNA as the catalytic subunit. Early experiments compared the B. subtilis RNase P RNA reaction (800mM NH₄OAc, 100mM MgCl₂) with the holoenzyme reaction (100mM NH₄OAc, 10mM MgCl₂) and demonstrated that the RNA and the holoenzyme had a similar Kₘ. The addition of RNase P protein increased kᵥat by 20
fold, which lead to the conclusion that the protein facilitated product release in addition to
shielding electrostatic repulsion (65). Later experiments with under different experimental
conditions (100mM NH₄OAc, 10mM MgCl₂ in the RNA alone reactions) and using the B.
subtilis RNase P RNA and holoenzyme offered an alternate conclusion. The B. subtilis
RNase P holoenzyme appears to have greater affinity for the precursor tRNA^Asp than it does
for the mature molecule. The increased affinity for substrate lowered K_m by 120-fold,
which caused catalytic efficiency (k_cat/K_m) to increase by 180 fold compared to RNA-alone
reactions that had a relatively similar k_cat (48).

ACHIEVEMENTS

RNA-alone catalysis by bacterial and archaeal RNase P RNAs.

The hallmark discovery in the study of RNase P was that bacterial RNase P RNAs
possess the necessary structural components to be catalytically active in vitro without
protein. The RNase P RNA, by itself, can perform each step in the RNase P reaction
cycle. RNA-alone activity has been observed for all tested bacterial RNase P RNAs, and
archaeal RNase P RNAs from the euryarchaeal groups: methanobacteria, halobacteria, and
thermococci (this thesis) (Fig. 7).

Catalytically active RNAs, such as the group I introns, hammerhead, hairpin, varkud
RNAs, and the hepatitis delta virus ribozymes can be engineered to work in trans, but
occur naturally as single-turnover, cis-acting elements. RNase P is unique because it is a
naturally occurring, multiple-turnover, trans-acting RNA enzyme. The demonstration of
catalytic activity by an RNase P RNA in the absence of protein is a significant achievement
because it shows that RNA can be truly enzymatic and it provides a simple model to
examine the relationship between structure and function in RNA.
Three-dimensional models

Structural arrangement is the critical determinant of catalytic activity. Three-dimensional models of the RNase P RNA are constructed to gain insight into global structure and propose how the molecule functions (Fig. 8). Three-dimensional models have been built and refined based on the combination of crosslinking data and phylogenetic analysis (see methods for details). Successful x-ray crystallography of an RNase P RNA has not been reported.

Accurate three-dimensional models represent an important achievement in the study of RNase P RNA because they allow visualization of the catalytic core of the molecule and approximate location of the structure-stabilizing peripheral elements. Construction of a tertiary model requires the identification of core versus peripheral components, but also, and more significantly, where the various structural features are located in space and in reference to one another.

Accumulation of sequences

Secondary structural refinement and tertiary model construction has been made possible by the continual collection of new RNase P RNA sequences. Sources of RNase P RNA genes are pure cultures, sequences of complete genomes, and isolation from environmental samples (discussed in methods). Obtaining additional sequences is important because the discovery of unique variations may represent evolutionary intermediates or reflect an evolutionary change not observed in RNase P RNAs from other species. In addition, the statistically based methods of comparative analysis decrease in uncertainty with a larger sample size.

A large and growing collection of RNase P RNA and RNase P protein sequences are available on the RNase P database (www.mbio.ncsu.edu/RNaseP). The RNase P database
is an excellent organizational tool; it puts the published RNase P RNA and protein sequences in one easily accessible location (7). In most cases, sequence data can be readily transported into various analytical software tools.

**Individual structural components**

Tertiary models of the RNase P RNA structure predicts the arrangement of structural features, but not necessarily their function. A notable achievement in the study of RNase P RNA has been the dissection of structure to determine the function of various individual structural features.

The stabilizing role of bacterial RNase P RNA helices P18 and P12 have been discussed above. Equally impressive work has been performed to elucidate Mg$^{2+}$ binding sites and the substrate tRNA CCA base pairing functions of the P15/16 loop of type A, and P15 loop of type B, bacterial RNase P RNAs. Understanding the function of individual structural features certainly increases knowledge of the molecule, but function is also valuable as a constraint in tertiary model construction. Computer generated models of the bacterial RNase P RNA P15/16 loop region (25), and the Mg$^{2+}$ binding P10/11-P12 region of yeast RNase P RNA (80) had to account for the function of particular structures when devising the global conformation. It is not enough to merely arrange structure if the final product is not compatible with known function.

**An RNase P protein crystal structure**

The recent report of the *B. subtilis* RNase P protein crystal structure (71) added another interesting feature to the already remarkable enzyme. The protein subunit of RNase P has an unusual left-handed $\beta$-$\alpha$-$\beta$ crossover connection rather than the right-handed crossover connection typically seen between $\beta$ strands. The unusual RNase P protein topology is
shared with the ribosomal protein S5 and the ribosomal translocase elongation factor G, and may represent evolution from a common RNA binding ancestor (71). The RNase P protein also has a large central cleft and a metal binding loop, which together with the left-handed βαβ, constitute three possible RNA binding sites. The discovery of a rare RNA binding motif shared by certain ribosome-associated proteins is a significant achievement that will contribute to unraveling the evolutionary history of RNase P.

**RNase P based gene therapy**

The most promising approach to RNase P based gene therapy is the design of an external guide sequence (EGS). The EGS is an oligoribonucleotide that is complimentary to the sequence of a target RNA (Fig 9.) (28). The annealing of the EGS to a target RNA (e.g. mRNA or viral RNA) results in the formation of an RNA structure that mimics RNase P substrate recognition elements, in particular the tRNA amino acid acceptor stem and the ΨΨC stem loop (77). EGSs require bases complimentary to the target RNA and a 3′ CCA for optimal cleavage activity (28). Target RNAs become the 5′ portion of an RNase P substrate and cleavage is directed at a particular base in the target RNA. Extra sequence of target RNAs are tolerated so long as the structural requirements of the an RNase P substrate 5′ and 3′ ends are met. Endogenous RNase P will recognize the trans-assembled substrate and cleave the target RNA, and ideally, render it inactive. By using EGSs, any RNA is a theoretical target provided that the EGS can bind the target RNA and cause the formation of an RNase P substrate (28).

The primary challenge of RNase P based, EGS mediated, gene therapy will be in delivery (1); however, design of successful EGSs is also a considerable obstacle. Experiments with the 2.1kb hepatitis B RNA have shown certain areas of the RNA are more suitable for EGS annealing (77, 78). Similar experiments with EGSs that targeted the
chloramphenicol acetyltransferase mRNA expressed in human cell cultures also demonstrated site-dependent optimization. Differences in binding efficiency are most likely due to the target RNA containing higher order structure at the less successful locations (79). Transfecting multiple EGSs appears to increase efficacy when separate EGSs were designed to target the polymerase and nucleocapsid mRNAs of human influenza virus.

The preliminary work with RNase P as an agent of gene therapy is a significant achievement. Although in vivo RNase Ps themselves are unchanged, and most of the experimental designs have involved creation of suitable substrates, it still requires intimate knowledge of the structure and function of the RNase P RNA to make these potential therapies effective. The development of RNase P as a gene therapy agent signifies the importance of the structure and function of RNA in the treatment of disease.

FUTURE TRENDS IN RNASE P RESEARCH

Future trends in RNase P research are likely to involve the increased study of the RNase P protein subunit, both alone and as part of the RNase P holoenzyme. The recent report of the crystal of the B. subtilis RNase P protein is the first look at the structure of an RNase P protein and how that structure relates to its function. The study of the relationship between structure and function of RNase P will certainly take advantage of research and technological advancements. Cryo-electron microscopy has already been applied to the structural study of ribosomal subunits with great success. Combined with increasingly sophisticated imaging methods, improved experimental techniques will produce high-resolution, detailed images of RNase P.

The details of structure will contribute to the understanding of function. Armed with this knowledge, RNase P applications, such as gene therapy, can be advanced to the point of predicting or designing function can be routinely accomplished based on knowledge of structure. The progression of RNase P based gene therapy is a good example of the initial
efforts in this direction. Understanding the structure of the RNA, in particular the RNA:RNA enzyme:substrate interaction, lead to the design of effective EGSs.

Future experiments may resolve issues such as the instances of RNase P where an RNA has not been identified. If it is true that RNase P activity can be accomplished by an all protein system, it would be interesting to compare the novel catalytic mechanisms to those understood for the catalytic RNase P RNA. A better understanding of the evolution of the RNase P RNA and RNase P protein, in particular the possibility of their co-evolution (Frank, unpub.) is also an anticipated future result; especially if there are examples of all protein RNase Ps to compare evolutionary origins.

METHODS

Cloning RNase P genes

The difficulty in cloning RNase P RNA genes has been greatly reduced by the accumulation of sequences from different organisms. Initially, hybridization of labeled probes to bacterial colonies with plasmids containing fragments of genomic DNA was used to identify RNase P RNA genes (36, 49). A collection of sequences, combined with comparative analysis, identified highly conserved regions of RNase P RNAs that were specific for a particular evolutionary domain: Bacteria, Archaea, or Eukarya (36, 63, 68). Based on conserved sequences, PCR primers sets were designed that would selectively target RNase P RNA genes from each evolutionary domain (see Oligonucleotides).

PCR amplification with domain-specific primers only yields partial RNase P RNA genes. The primer annealing sites roughly correspond to the RNase P RNA helix P4, and omits P1, the 5' portion of P2, and P3 from amplification. Sequencing upstream and downstream of the domain-specific primers would produce the missing sequence, but that
is not possible when genes are amplified directly from environmental samples, mixed cultures or any instance when exact genomic DNA is unavailable. Fortunately, sequence analysis demonstrated genus-level conservation in P1 at the 5’ and 3’ ends of RNase P RNAs (7, 35, 36). PCR primers can be designed that narrow the evolutionary scope of gene searches (for examples, see archaeal RNase P RNA gene genus-specific PCR primers below in Oligonucleotides).

PCR template DNA can be obtained from a variety of sources. Amplification from pure or enriched cultures is common, as is obtaining RNase P RNA genes by traditional methods from genomic DNA preparations. Isolating DNA from prokaryotic species requires relatively simple extraction and precipitation procedures (8). Specific protocols may be necessary to isolate RNase P RNA genes from eukaryotic nuclei and organelles (26, 17, 74). Complete genome sequences are a valuable source of RNase P RNA gene sequences. Database searches can be performed using the nucleotide sequences of PCR primers, or used to design primers specific for the amplification of the RNase P RNA gene from that organism.

PCR amplification from natural populations is a technique that eliminates the need to isolate or grow individual species (17). An environmental sample, such as wastewater sludge or pond water, is collected and DNA is extracted. PCR primers for RNase P RNA genes (or any gene of choice) are then used to amplify DNA. Amplification of novel genes from natural populations of organisms often produces sequences from unidentifiable species, but general evolutionary relatedness can be determined based on phylogenetic analysis. Natural populations amplification is advantageous because it generates many more sequences than could be obtained from growing pure cultures, especially since the majority of organisms in any particular sample are unculturable.
Circularly permuted RNAs

A circularly permuted RNA is a molecule with artificial 5' and 3' ends that has its native 5' and 3' ends incorporated into continuous sequence. Circularly permuted RNase P RNAs (or substrates) are used in crosslinking and unimolecular enzyme:substrate complexes experiments (19, 38, 57). An RNase P RNA or pre-tRNA with artificial ends allows the addition of a molecular tag such as the photoreactive compound, azidophenacyl bromide, (see crosslinking) to be placed at specific locations within the RNA sequence. A unimolecular complex has the substrate RNA cloned as part of the RNase P RNA gene. Substrate RNA genes are usually cloned into regions of the RNase P RNA that will position them in the substrate binding pocket of the RNase P RNA so that transcription creates an enzyme:substrate complex. Substrate bound to enzyme potentially removes that step from the catalytic cycle and allows direct evaluation of catalysis.

There are two general methods of generating a circularly permuted RNase P RNA. One method requires tandem RNase P RNA genes to be cloned into a plasmid (Fig. 10). The tandem RNase P RNA genes are then amplified by PCR primers designed to create the desired artificial 5' and 3' ends. Amplification through the junction of the two RNase P RNA genes provides the continuous sequence at the native ends of the molecule. The other method of circular permutation requires two rounds of PCR amplification (Fig. 11). In the first PCR amplification, primers corresponding to the native 5' and 3' are used to amplify an RNase P RNA gene. The amplification product is treated with DNA ligase to produce circular DNA. The second PCR amplification is performed using primers designed to anneal to the circularized RNase P RNA gene and produce the desired artificial 5' and 3' ends. Restriction endonuclease sites are often added to the end of the PCR primers to clone amplification products into plasmids, and phage promoters can also be added to the 5' primer for direct transcription.
RNase P activity assays

The RNase P RNA activity assay is used to test function (Fig. 12). Catalytic activity is an evaluation of structural mutations, kinetic parameters, and used as a reporter system in various nuclease or modification experiments. Monovalent and divalent cation concentrations are critical factors when designing an RNase P RNA activity assay. \( \text{NH}_4^+ \) and \( \text{Mg}^{++} \) are preferred, but activity has also been observed with other ions (62, 69). Activity assays can be attempted using an RNase P RNA from any species or organelle, but \textit{in vitro} RNA-alone cleavage of pre-tRNA has only been observed using bacterial and certain archaeal RNase P RNAs (this thesis) (62).

Bacterial RNase P RNAs are catalytically active under a wide range of ionic conditions. Minimally, 50-100 mM \( \text{NH}_4\text{OAc} \) and 10-15mM \( \text{MgCl}_2 \) is enough to stabilize the RNase P RNAs from \textit{E. coli} and \textit{B. subtilis}. For the same organisms, 800mM-1M \( \text{NH}_4\text{OAc} \) and 25-50mM \( \text{MgCl}_2 \) produces optimal catalysis \textit{in vitro}. However, some RNase P substrates are cleaved better in lower concentrations of \( \text{NH}_4^+ \). Mutant bacterial RNase P RNAs often require increased ionic concentrations (1-3M \( \text{NH}_4\text{OAc} \), 100-200mM \( \text{MgCl}_2 \)) to stabilize mis-formed RNase P RNA tertiary structures (58, 60, 62). High salt conditions have also been used to test the catalytic activity of bacterial RNase P RNAs at elevated temperatures (up to 75°C) where increased stability is needed to prevent thermal denaturation (9). The minimal ionic conditions needed to observe activity for the catalytic archaeal RNase P RNAs is approximately 1-1.5M \( \text{NH}_4\text{OAc} \) and 100mM \( \text{MgCl}_2 \), but optimal conditions are much higher: 3M \( \text{NH}_4\text{OAc} \) and 300mM \( \text{MgCl}_2 \) (this thesis).

For the bacterial and archaeal catalytically active RNase P RNAs, optimal pH is 8.0 and is achieved by including 50mM Tris-HCl pH8 in reactions. Assays are typically resolved using 8-12% denaturing polyacrylamide gel-electrophoresis of reactions that contained \( ^{32} \text{P} \)
labeled pre-tRNA substrate. If viewing gels by autoradiography, approximately five-
thousand CPM per reaction is recommended, but if using phosphorimagry, two-thousand
CPM per reaction produces satisfactory results.

RNase P activity assays are also performed using the RNase P holoenzyme. In such
cases, salt is held at a minimum (100mM NH₄OAc, 10-25mM MgCl₂) (48, 65). Since the
RNase P protein subunit stabilizes the RNase P RNA more effectively then salt, RNase P
holoenzymes can be assayed at higher temperatures in lower salt conditions then the RNA
alone.

A commonly used substrate pre-tRNA, such as the pre-tRNA\textsuperscript{Asp} from \textit{B. subtilis} (Fig.
2) (pDW128, 64) is 103 nucleotides in length after run-off transcription. Expected results
upon complete cleavage by RNA is a 77 bases mature tRNA and a 26 base leader (Fig. 12).
When using the \textit{E. coli} or \textit{B. subtilis} RNase P RNA, expected K\textsubscript{m} is in the range of 10-
100nM, but will vary with alteration of reaction conditions and sequence mutation.

**Purification of RNase P**

Purification of RNase P RNAs or holoenzymes is often a necessary step in the
completion of a particular experiment. The isolation and purification of RNase P was
essential in demonstrating that the RNase P RNA was the catalytic subunit. Purification is
also required to isolate RNase P proteins or RNAs from other cellular components, and
when RNase P RNAs are made \textit{in vitro}, a purification step is necessary to separate the full
length RNA from the various incomplete transcripts.

Purification of bacterial RNase P begins with the growth of a suitable amount of cells
(150-400g). \textit{E. coli} RNase P purification (47) starts with cell disruption and centrifugation
to remove cell debris. Passage of supernatent through a NH₄Cl gradient used on a DEAE-
Sephadex column is followed by a Sepharose-4B column of peak fractions. Peak
Sepharose column fractions were pooled, precipitated with ammonium sulfate, and applied
to a G200 sephadex column. Two n-octyl-sepharose columns are used to further purify the
*E. coli* RNase P followed by dialysis against 0.5M NH₄OAc, and addition of solid urea to
7M. The material is applied to a CM-sephadex column and eluted with NaCl from 0 to
0.5M.

The *E. coli* purification protocol described above is quite thorough. A more simple
procedure was reported for the purification of the *B. subtilis* RNase P (30). Cells are
passed twice through a French press, then centrifuged to remove cell debris. Supernatent
is incubated on ice for 30 minutes in 0.22M NH₄Cl then centrifuged at 45K rpm for 4hrs.
at 2°C. The second supernatent is dialyzed with 0.5M Tris-HCl pH7.3, 0.06M NH₄Cl,
0.01M MgCl₂, 0.1mM DTT, and 5% glycerol buffer that contained 100µg of DNase and
0.5mM PMSF. Dialysis is followed by application to a DEAE-cellulose column and
elution with 0.2-0.5M NH₄Cl in dialysis buffer.

Nuclear and organelle RNase P has been purified from several eukaryotic organisms
using procedures similar to those used in Bacteria (4 26, 29, 31, 63, 74). The same is true
for archaeal RNase Ps. Specific protocols should be consulted to insure critical details
pertaining to growth conditions and isolation methods of individual organisms are satisfied.

An interesting approach to RNase P purification is the use of affinity column purification
as reported for the purification of the RNase P from *Tetrahymena thermophila* (74).
Biotinylated tRNAs were bound to streptavidin-conjugated beads. Partially purified *T.
thermophila* RNase P preparation was mixed with the beads to allow binding with the
tRNA. Unbound material was separated, then highly purified RNase P was eluted from the
tRNA conjugated beads.

Previous purifications demonstrated that bacterial RNase P contains more RNA (120kD)
than protein (14kD), which is reflected in the buoyant density (1.55g/ml) of the *E. coli*
RNase P (51) in Cs₂SO₄. Eukaryotic RNase Ps have lower buoyant densities (1.34g/ml
for the wheat RNase P) which corresponds to a much larger protein:RNA ratio (4). For reference, RNase P RNA alone has a buoyant density of 1.7g/ml and RNase P protein, 1.2g/ml. Instead of a single polypeptide protein subunit as in Bacteria, the yeast, _Saccharomyces cerevisiae_, has a 10 polypeptide RNase P protein subunit (17). Although archaeal RNase Ps have been purified (23, 56), the number and identity of protein subunit polypeptides has not been determined. Buoyant density measurements indicate mixed possibilities. For crenarchaeal organisms, like _Solfolobus acidocaldarius_ (density = 1.27g/ml), there is a high protein:RNA ratio suggesting large or multiple polypeptides. Euryarchaeal organisms, such as _Haloflex volvanii_ (density = 1.61g/ml), have a higher buoyant density, that suggests a lower protein:RNA ratio and the possibility of one or small polypeptides in the RNase P protein.

Isolating RNase P RNA from cells in the form of total RNA is often a preparative step used in experiments such as RNA end-mapping or reverse transcription. A 20-30 ml cell culture is usually sufficient. Cultures are passed through a French press and centrifuged to remove cell debris. Organic extraction is used to remove proteins and is accomplished by extraction twice with phenol followed by extraction with phenol-chloroform (1:1), or some alternate, but equally effective, combination of organic solvents. RNase P RNA is more stable than many cellular RNAs, but RNA-handling precautions are still recommended.

RNase P RNA synthesized by _in vitro_ run-off transcription can be purified by denaturing polyacrylamide gel-electrophoresis. RNA is readily excised from gels using the UV shadow technique, ethidium bromide staining, or if the RNA is labeled with ^32_P, excision by autoradiogram positioning is possible. Elution of RNase P RNA from gel slices can be passive (overnight incubation in a suitable buffer at 4°C) or active (electro-elution).
Crosslinking

Physical analyses such as X-ray crystallography and NMR are used to generate tertiary structural models for protein molecules and small RNAs, but these techniques have not been successfully used with large RNAs like the RNase P RNA. Instead, tertiary models of RNase P RNAs are made by using constraints from intermolecular and intermolecular UV-light induced specific crosslinks (19, 38, 39, 57). Crosslinking is used to constrain elements of secondary structure in space. For example, crosslinks demonstrated that the loop at the end of E. coli RNase P RNA helix P9 is adjacent to sequence in P1 (19), so any tertiary model must account for the proximity of these helices. Crosslinking experiments that directly compared the tertiary arrangement of the E. coli and B. subtilis RNase P RNA revealed that the non-homologous structures P6/16/17 of E. coli RNase P RNA and P5.1-P15.1 interaction of B. subtilis RNase P RNA occupy the same space and probably serve an analogous function.

Experimental procedures are essentially the same whether an RNase P RNA is being crosslinked to itself (intramolecularly) or to a tRNA (intermolecularly). The RNA designated for addition of the photoagent is transcribed in the presence of guanosine monophosphorothioate (GMPS) will incorporate into nascent RNA only at the position +1. RNA polymerase requires nucleotide triphosphates for chain elongation, so the GMPS that will not be present at any other position. Circularly permuted RNAs (described above) are employed to position GMPS on nucleotides normally in the interior of RNA sequences.

The photoagent commonly used in RNase P crosslinking experiments is azidophenacyl bromide, which was specifically attached to the 5' thiol group of GMPS in reactions containing 40% methanol, 20mM sodium bicarbonate pH 9, 1% SDS that incubated for 1hr. at room temperature. Excess azidophenacyl bromide is removed by phenol extraction and tagged RNAs are collected by ethanol precipitation (14). After a short preincubation,
RNA(s) are exposed to UV light (302nm). The UV light was filtered through polystyrene to reduce shorter wavelength light that could cause photoagent-independent crosslinks. Crosslinked products separated from non-crosslinked material on denaturing polyacrylamide gels. The crosslinked positions are determined by primer extension from oligonucleotides annealed to the target RNA. Reverse transcriptase will elongate a DNA chain until it reached the crosslink site. Products are compared to size markers, and used to measure which nucleotide(s) in the target RNA sequence were crosslinked.

Crosslinking of a nucleotide to the azidophenacyl photoagent does not indicate base pairing or specific contact, but does demonstrate that the tagged nucleotide and the base it is crosslinked to are in proximity. The bond length of the photoagent is 5-10 angstroms. Several crosslinks to one tagged-RNA position are among the expected results (14). There is a reasonable probability that more than one nucleotide is within crosslinking distance of the photoagent. Multiple crosslinks are usually located near each other; however, when one tagged position has crosslinks in distant regions of the target RNA it suggests close tertiary structural positioning, which is information vital to tertiary model construction (Fig. 8).

Phylogenetic comparative analysis

The key to phylogenetic comparative analysis is conservation of higher-order structure despite changes in nucleotide sequence. Evolutionary changes of a nucleotide will correlate with compensatory changes elsewhere in the molecule indicating interactions between these bases. The structure of an RNA can be determined by comparing homologous positions in a collection of RNA sequences for compensatory base changes. When compensation occurs in homologous RNAs from different species it is referred to as covariation.

Covariation is studied starting with RNA sequence alignments (Fig. 13). RNA sequences are written in a single line with each sequence listed above the next. Homologous nucleotides are positioned vertically so the same nucleotide position in each
sequence appears in a column. Closely related sequences are easy to align because they will be similar in sequence length. When aligning more distantly related RNAs, it is necessary to insert numerous gaps in the alignment that correspond to the existence of bases in one RNA that are not present in others.

Once RNA sequences are properly aligned, the construction of structural models is based on results from analysis of covariation. Phylogenetic comparative analysis (7) was used to construct secondary structures of many RNase P RNAs, and was recently used for refining the secondary structure and generating a tertiary model of bacterial RNase P RNA (55). As mentioned above, phylogenetic comparative data is often combined with results of crosslinking experiments to produce a model that represents both comparative and physical constraints.

Oligonucleotides

Universal bacterial PCR primers:

59FBam: CGG GAT CCG IIG AGG AAA GTC CII GC
347REco: CGG AAT TCR TAA GCC GGR TTC TGT

Universal archaeal PCR primers:

A59FXba: GCT CTA GAG GAA AGT CCC/A C/GCC
A347RBam: CGG GAT CCT AAG CCC/A C/GCT TT/CT GT (35, 36)

Sequence analysis shows that there is less P4 conservation in eukaryotes, so PCR primers are not as universal as in Bacteria. The following are two general eukaryal PCR primers ((63) additional eukaryal primers also reported).

Euk-Z1: GCG GAA GGA AGC TCA CTG TAG AGG
Euk-A: CGG AAT TCC GGA GAG TGG TCT GAA TTG GGT T

Methanobacterial genus-specific PCR primers:
CONCLUSIONS

The study of structure and function encompasses nearly all aspects of RNase P. Catalytic requirements, kinetic parameters, secondary structures, and tertiary models are all dependent on the conservation of a structural core that is influenced and stabilized by variably present peripheral elements. The RNase P protein subunit is required in vivo, but is not necessary for bacterial and certain archaeal RNase P RNAs in vitro where increased ionic conditions can support catalytic activity. In vitro RNA-alone activity indicates that the RNase P RNA contains the essential elements for substrate recognition and catalysis, and they possess the ability to self-assemble into a catalytically competent tertiary structure. Where RNase P RNAs are not catalytically active in vitro, a more complete understanding of the RNase P protein, combined with increasingly detailed tertiary models, should clarify the discrepancy between lack of activity and the function of a conserved structural core.
Figure 1. RNase P RNA secondary structures with numbered helices.
Bacillus subtilis pre-tRNA(Asp)

Escherichia coli 4.5s RNA

Figure 2. Two RNase P Substrates. The pre-tRNA is from the clone pDW128 (64) and the upstream leader sequence is cloning vector polylinker sequence (45).
*Methanobacterium thermoautotrophicum*
strain ΔH RNase P RNA

Figure 3. Secondary structure of an archaial RNase P RNA with helices numbered (35, 62, this thesis).
Figure 4. Secondary structures of minimal RNase P RNAs. (A) The minimal bacterial consensus structure. Capital letters (G, A, U, C): 100% conserved bases, lower case letters (g, a, u, c): 80% conserved bases, closed circles: bases present in all RNase P RNAs, and open circles: bases present in 80% of RNase P RNAs. (B) Micro P. A minimal but catalytically active in vitro RNase P RNA constructed based on sequence information from *M. fermentans* and the bacterial consensus structure (68).
Figure 5. RNase P RNA consensus structures. Capital letters (GAUC) represent bases at least 80% conserved, closed circles represent bases 80% present. A comprehensive eukaryal consensus structure has not been reported; the vertebrate consensus was chosen as an example. The position of P12 is labeled, but because of sequence length and variation in structure is not drawn (36, 63, 68).
Figure 6. Secondary structures of selected organelle RNase P RNAs. The *Reclinomonas americana* mitochondrial RNase P RNA has helices numbered for reference (Lang, unpubl.).
Figure 7. A three-domain tree of life (35). The boxed Archaea are those with RNase P RNAs that are active in vitro without RNase P protein. All bacterial RNase P RNAs tested are active without RNase P protein. In vitro RNA-alone activity has not been reported for any eukaryotic or organellar RNase P RNA.
Figure 8. A three-dimensional model of the bacterial type A RNase P RNA proposed by Massire and Westhof (55). Colors represent sequence variation as determined by comparative analysis.
Figure 9. External guide sequence RNA shown base paired to a region of the chloramphenicol acetyltransferase mRNA to form an RNase P substrate. (28, 79)
Step 1. Clone tandem repeats of RNase P RNA genes.

Step 2. PCR amplification with primers (arrows) that will create artificial 5' and 3' ends while linking the native ends of the RNase P RNA gene.

Step 3. Proceed with transcription, end-labeling, addition of photo-reactive agent, or cloning into unimolecular construct.

Figure 10. Tandem repeat method of cloning circularly permuted RNase P RNAs.
**Step 1.** Amplification of RNase P RNA gene from the native 5’ and 3’ ends with complimentary primers that contain linker sequences.

**Step 2.** Treatment with restriction endonucleases and T4DNA ligase. Creates circular amplification fragment.

**Step 3.** PCR amplification with primers that create artificial 5’ and 3’ ends.

**Step 4.** Proceed with transcription, end-labeling, addition of photo-reactive agent, or cloning into unimolecular construct.

*Figure 11.* Two-PCR method of cloning circularly permuted RNase P RNAs.
Figure 12. An example of an RNase P assay.
Figure 13. A portion of a sequence alignment of bacterial RNase P RNAs (7). This is usually the first step in comparative analysis. Gaps are used to position homologous bases or blocks of sequence.
Figure 14. Mutual information plot of Bacterial RNase P RNA. Helical elements appear as short lines perpendicular to the diagonal.
REFERENCES


RNase P RNAs FROM CERTAIN ARCHAEA ARE CATALYTICALLY ACTIVE

ABSTRACT

Bacterial RNase P RNAs are catalytically active by themselves in vitro; i.e. they are ribozymes. The RNase P RNAs of Archaea and Eukarya are apparently catalytically inactive in the absence of protein; this similarity is consistent with their phylogenetic affiliation. However, a variety of archael RNase P RNAs were tested for activity in ionically extreme conditions, and although many remain inactive, those of the methanobacteria, thermococci, and halobacteria were catalytically active. In addition, chimeric holoenzymes reconstituted from the Methanobacterium RNase P RNA and the B. subtilis RNase P protein subunits were functional at moderate ionic strength.

INTRODUCTION

RNase P is the enzyme responsible for the endonucleolytic removal of 5’ leader sequences from transfer RNA precursors (1, 26, 17). RNase P is best understood in Bacteria, in which it consists of two subunits: a large (ca. 140kD, 400 nucleotide) RNA, and a small (ca. 14 kD, 120 amino acid) protein (25, 18). Both RNA and protein are required in vivo, and for optimal activity in vitro in reactions at moderate ionic strength (19, 29). The RNA is the catalytic subunit of the bacterial enzyme; at elevated ionic strength it is by itself capable of processing pre-tRNAs catalytically, i.e. it is a ribozyme (12). The RNase P enzymes of Archaea (formerly archaebacteria) and Eukarya (the nuclear/cytoplasmic portion of the eukaryotic cell) also contain RNA subunits similar to bacterial RNAs (7, 10), but these RNAs have not been shown to be catalytically active.
Although it seems likely that the catalytic function of the archaeal and eukaryal enzymes reside in the RNA, the expression of this activity apparently requires the presence of the protein subunits. The similarity of the archaeal and eukaryal RNase P RNAs, in terms of their catalytic deficiency, is consistent with the specific relationship of the Archaea and Eukarya, which are sister taxa relative to their bacterial cousins (3, 16).

RNase P enzymes have been characterized from only two archaeal species: Sulfolobus acidocaldarius (8) and Haloferax volcanii (24). The S. acidocaldarius RNase P is resistant to micrococal nuclease treatment, but contains a 315 nt RNA that persists after nuclease treatment. The enzyme is large (ca. 400kD apparent molecular weight) and has a low density in Cs₂SO₄ (1.27g/cm³), similar to those of the nuclear enzyme and implying a high protein:RNA content. The H. volcanii RNase P, on the other hand, resembles the bacterial enzyme density in Cs₂SO₄ gradients (1.61g/cm³) and nuclease sensitivity. The halophile enzyme is associated with a 435 nt RNA.

Proteins associated with the archaeal enzyme have not been identified, nor have sequences encoding polypeptides with recognizable similarity to eukaryotic or bacterial RNase P proteins been identified. The putative RNase P RNAs from S. acidocaldarius and H. volcanii are similar in size and to some extent sequence to those of Bacteria, but these RNAs by themselves could not be shown to be catalytically active following deproteinization or when synthesized in vitro (20, 24). This, and subsequent tests with the same result with RNase P RNAs from Methanosarcina barkeri and Methanococcus jannaschii (13, 15), led to the conclusion that archaeal RNase P RNAs, like those of eukaryotes, are not by themselves catalytically active. However, this report will show that in ionically extreme conditions, RNase P RNAs from the archaeal groups methanobacteria, thermococci, and halobacteria are, by themselves, catalytically active.
MATERIALS AND METHODS

Cloning RNase P RNA genes

RNase P RNA-encoding sequences were cloned using previously-described methods (4, 13, 15). Oligonucleotide primers for amplification from genomic DNA were ΔH5′Bam and ΔH3′Xba for obtaining sequences from Methanobacterium spp., Pyro5′Xba and Pyro3′Bam for obtaining sequences from Pyrococcus and Thermococcus spp., Mja5′Xba and Mja3′Bam for obtaining sequences from Methanococcus spp., and Sac5′Bam and Sac3′Xba for obtaining the sequence from M. sedula. See Oligonucleotides (below).

Preparation of archaenal cellular RNA

A 20-30 ml cell culture was used to isolate total RNA from cells. Cultures were passed through a French press and centrifuged to remove cell debris. Proteins were removed from archaenal cell lysates by two phenol extractions followed by one 1:1 phenol:chloroform extraction (27). Supernatants were ethanol precipitated and resuspended in 10mM Tris-HCl pH8, 1mM EDTA. RNase P RNA assays contained 100-400ng cellular RNA, 1.5nM uniformly $^{32}$P-labeled B. subtilis or M. thermoautotrophicum pre-tRNA$^{Asp}$, 50mM Tris-pH8, 25-300mM MgCl$_2$, and 1-5M NH$_4$OAc. Incubations ranged from 2-16 hours at 30 to 75°C. Reaction products were separated by electrophoresis on 8% polyacrylamide:8M urea gels and analyzed by phosphorimagy.
In vitro transcription of RNase P RNAs

RNase P RNAs were synthesized from cloned genes using T7 RNA polymerase. Transcription reaction products were separated by denaturing polyacrylamide gel electrophoresis, visualized after staining with ethidium bromide, or if the RNA is labeled with $^{32}$P, excision by autoradiogram positioning is possible. Gel lanes were fractionated and eluted in 10mM Tris-pH8:1mM EDTA:0.1% SDS. RNA concentration in each fraction was evaluated from absorbance at 260nm, and RNase P activity was assayed as described in Materials and Methods. In each case, RNase P activity coincided exactly with the predominant RNA band of the expected size.

RNase P RNA end mapping

The location of the *M. thermoautotrophicum* (strain ΔH) 5' end was mapped by primer extension using the oligonucleotide Mfo145R (8). Total RNA was harvested from *M. thermoautotrophicum* cultures as described. 5' $^{32}$P-end-labeled oligonucleotides were annealed to RNase P RNA by slow cooling to the hybridization temperature after a 10 min. 80°C incubation in 0.4M NaCl, 40 mM PIPES pH 6.4, 1mM EDTA, and 80% (v/v) formamide. The hybridization temperature was determined using the following formula: 29.3 + 0.41(%G+C of the oligonucleotide). Hybridization reactions were precipitated with ethanol and NH$_4$OAc and resuspended in 10mM PIPES pH 6.4, 0.4M NaCl. 5 units of AMV reverse transcriptase was incubated with annealed RNAs for 1 hr. at 42°C in 50mM Tris-HCl pH 8.2, 10mM DTT, 6mM MgCl$_2$, and 25 ug/ml actinomycin-D. Reactions were phenol-extracted and ethanol precipitated prior to being resolved by denaturing polyacrylamide gel (6%) electrophoresis.
The location of the 3' end was determined by S1 nuclease mapping using a probe DNA fragment that overlapped the 3' end of the *M. thermoautotrophicum* strain ΔH RNA and containing sequences downstream of the gene for the RNase P RNA, made by polymerase chain reaction amplification from genomic DNA using oligonucleotide primers MthΔH138F and MthΔH908RXba. The probe DNA was annealed to total cellular RNA in 0.4M NaCl, 40 mM PIPES pH 6.4, 1mM EDTA, and 80% (v/v) formamide. Hybridization reactions were incubated at 80°C for 10 min., then slow cooled to 45°C and incubated for 16hrs. The hybridization reaction was added to a 300μl S1 nuclease reaction that contained, in addition to S1 nuclease, 28mM NaCl, 5mM NaOAc pH 4.6, 0.45mM ZnSO₄, and 2μg salmon sperm DNA, and incubated for 45min at 37°C. Reactions were phenol-extracted and ethanol precipitated prior to being resolved by denaturing polyacrylamide gel (8%) electrophoresis.

**RNase H degradation assay**

RNase P RNAs were degraded by RNase H in 100mM KCl, 10mM MgCl₂, 20mM HEPES-KOH (pH 7.4) and 1 mM DTT. Reactions that contained 0.2μg RNase P RNA and 2μg of oligonucleotide that binds specifically to a single stranded region in *E. coli* or *M. thermoautotrophicum* RNase P RNA were heated to 80°C for three minutes then slow cooled to 30°C; 2 units RNase H (Life Technologies, Inc.) were added and incubated for 1 hr at 30°C. RNase P activity remaining was assessed in RNase P RNA activity assays (described below).
RNase P activity assays

RNase P RNA activity assays contained 50mM Tris-HCl (pH 8), 300 mM MgCl$_2$, 3M NH$_4$OAc and were incubated at 37°C. Cleavage products were analyzed by electrophoresis in 8% denaturing acrylamide gels followed by phosphorimagry. Reactions containing archaeal RNase P RNA and negative controls containing no RNase P RNA were incubated at 37°C for 16 hours. Reactions containing *E. coli* RNase P RNA were incubated at 37°C for only 5-10 minutes. The pre-tRNA$^{\text{A}^\text{p}}$ substrate from *B. subtilis* was synthesized and labeled with $^{32}$P-GTP using T7 RNA polymerase from the clone pDW128 (64).

Reconstitution of RNase P holoenzymes

*B. subtilis* RNase P protein was a gift from C. Fierke (Duke University). RNase P holoenzyme assays were conducted in 50 mM Tris-HCl (pH 8.0), 100 mM NH$_4$OAc, 25 mM MgCl$_2$ and ca. 1.5nM pre-tRNA$^{\text{A}^\text{p}}$. All reactions included 18 μg/ml antisense *M. thermoautotrophicum* RNase P RNA as a decoy for traces of nuclease activity in the *B. subtilis* RNase P protein preparation.

Oligonucleotides

deltaH5'Bam:  CGG GAT CCA CCG GGC AAG CCG AAG GGC
deltaH3'Xba:  GCT CTA GAC CGG GCA TGC CGA GAG
Pyro5'Xba:  GCT CTA GAT AGG CGA GGG GGC TGG GG
Pyro3'Bam:  CGG GAT CCT AGG CGA GGG GGC TAT AG
Mja5'Xba:  GCT CTA GAG GGT AAG GGG GCT GGT G
Mja3'Bam:  CGG GAT CCG GTA TGG GGG CTA TAG C
RESULTS AND DISCUSSION

A well-defined secondary structure model was recently constructed of RNase P RNAs from Archaea using comparative sequence analysis (Fig. 1) (4, 13, 15). Nearly all of the conserved sequences and secondary structures in the catalytically-active bacterial RNAs are also present in the archaeal RNAs. The structural distinctions between the bacterial and archaeal RNAs have largely disappeared as our understanding of their structures has improved, so the dramatic difference in their ability to catalyze pre-tRNA 5'-processing was puzzling. RNase P RNA-encoding sequences were obtained from three cultivated species of Methanobacterium (M. thermoautotrophicum strains Marburg and ΔH, and M. formicicum), and two related sequences cloned from a methanogen enrichment culture inoculated with municipal wastewater sludge. The archaeal RNAs of the methanobacterial group are particularly similar to their bacterial homologues, so they were tested for pre-tRNA processing activity in a wide variety of conditions, including ionic conditions outside the range usually considered.

In preliminary experiments, cell lysates from a variety of archaeal cultures were capable of processing either the B. subtilis or M. thermoautotrophicum pre-tRNA<sup>Ap</sup> (27). Extraction of proteins from these lysates in many cases eliminated catalytic activity, but those of the methanobacteria and the thermococci (Pyrococcus furiosus and Thermococcus
(celer) retained RNase P activity. Archaeal RNase P RNA activity was resistant to exhaustive phenol extraction and the inclusion of SDS (1%) in the reaction buffer, supporting the belief that this activity was independent of protein. RNAs purified by organic extraction from the extreme halophiles *Haloferax volcanii* and *Natronobacterium gregoryi* failed to process pre-tRNA, but the activity present in cell lysates was resistant to the inclusion of 1% SDS in the reaction buffer, suggesting that the RNA by itself was responsible for this activity.

The convincing proof that other ribozymes, including bacterial RNase P RNA, are catalytically proficient in the absence of associated proteins came from experiments using RNAs transcribed *in vitro* (11). Synthetic RNAs from cloned archaean RNase P RNA-encoding genes were therefore tested for pre-tRNA 5′-processing activity; those of the methanobacteria, the extreme halophiles, and thermococci, but not other Archaea, were found to contain RNase P activity (Fig. 2). The catalytic proficiency of the synthetic RNAs demonstrates that they are not dependent on protein for activity, nor are post-transcriptional modifications such as pseudouridylation or 2′-O-methylation essential for activity.

The extent of activity for all of the active archaean RNase P RNAs was quite low; assays included large molar excesses (in the range of 100:1) of RNase P RNA over the labeled substrate and incubation for several hours, but resulted only in partial processing of the substrate. The potential for contamination by catalytically-active bacterial RNase P RNAs was therefore a significant concern. Previous descriptions of catalytic activity by the *H. volcanii* RNA (24), and reconstitution of chimeric RNase P holoenzymes from archaean and bacterial components (21, 24), were previously discounted as the result of contamination, because the amount of activity recovered was very small and the results could not be reproduced (13). Three observations exclude the possibility that this catalytic activity might result from contamination by bacterial RNase P RNA rather than from the archaean RNA:
1., correlation of the presence or absence of activity from RNA extracted from cells and RNAs synthesized in vitro from genes cloned from the same species; 2., the unusual biochemical properties of the activity (Fig. 2), and 3., the electrophoretic co-migration of activity with the archaeal RNA. The direct experimental evidence that the observed activity resides in the archaeal RNase P RNA comes from specific oligonucleotide-directed RNase H depletion experiments using the *M. thermoautotrophicum* RNA (transcribed in vitro from the cloned gene) and oligonucleotides complementary to either the *M. thermoautotrophicum* or *E. coli* RNase P RNA (Fig. 3). Treatment of the *M. thermoautotrophicum* RNase P RNA sample with RNase H in the presence of the *M. thermoautotrophicum*-specific oligonucleotide, but not the *E. coli*-specific oligonucleotide, reduces RNase P activity, whereas the *E. coli*-specific oligonucleotide and RNase H reduced the activity of the *E. coli*, but not *M. thermoautotrophicum*, RNase P RNA.

The reaction conditions required by the archaeal RNase P RNA activity are very unusual. The *M. thermoautotrophicum* RNase P RNA synthesized in vitro was found to require 300mM MgCl₂ and 3M ammonium acetate for maximal activity (28), well above the requirements of any characterized bacterial RNase P RNA. The MgCl₂ requirement was not reduced by increases in ammonium acetate concentration, or *visa versa*, suggesting that the RNA is structurally defective (in the absence of the protein subunit[s] of the holoenzyme). The high ionic strength may be required for stabilization of the RNA structure, and the high Mg²⁺ for further stabilization and/or to overcome poor binding of catalytically-involved Mg²⁺. The RNase P RNA from *Haloferax volcanii* also required very high ionic strength for activity (4M ammonium acetate and 300mM MgCl₂) for maximal activity. The optimal temperatures for the thermophilic RNase P RNAs has not be determined because of the extent of non-specific hydrolysis in these reactions at elevated temperatures, but high
temperature alone did not allow optimal catalysis at lower ionic strength or MgCl₂ concentrations.

Although the catalytic proficiency of these archaeal RNase P RNAs is much lower than their bacterial homologues, the ability of these RNAs to process pre-tRNA at any rate implies that they contain all of the sequences and structures necessary for substrate recognition and catalysis. This is consistent with our understanding of the conserved features of sequence and secondary structure of the archaeal and bacterial RNase P RNAs; these RNAs are remarkably similar in both sequence and structure, and all of the invariably-present sequences and structures of the catalytically-active bacterial RNAs are present in some or all of the archaeal RNAs as well (Fig. 1).

The protein component(s) of the archaeal RNase P have yet to be identified despite the availability of several complete archaeal genome sequences and RNase P protein sequences from a wide variety of Bacteria and the yeast nucleus and mitochondrion (and some of the human sequences) (6, 9, 23, 30). Because of the similarity of the archaeal and bacterial RNase P RNAs, we attempted to reconstitute chimeric holoenzymes containing archaeal RNase P RNA and bacterial RNase P protein (Fig. 4). Catalytic activity by the Methanobacterium RNase P RNAs was reconstituted at moderate ionic strength and Mg²⁺ concentration using the B. subtilis RNase P protein, indicating the functional interaction of these heterologous subunits. We did not detect enhanced activity of the Metallosphaera sedula, Haloferax volcanii, Methanosarcina barkeri or Methanococcus vannielii RNAs in the presence of the bacterial protein under these conditions, although analogous reconstitution of H. volcanii RNA and B. subtilis protein has been reported (24). RNase P activity at moderate ionic strength was absolutely dependent on the presence of the protein; this is unlike the E. coli or B. subtilis RNase P RNAs, which can process this substrate at rates of ca. 20% that of the reconstituted enzymes in these conditions (19, 29).
It seems likely, therefore, that the structure and function of the bacterial and archaeal protein components are, at least in part, similar in the archaeal and bacterial enzymes.
Figure 1. Comparison of archaeal and bacterial RNase P RNAs. The secondary structure of the RNase P RNA from *Methanobacterium thermoautotrophicum* strain ΔH is shown in panel A (4, 13, 15). Helices are numbered P1-P16 as previously described (14). Helices P4 and P6 are shown with lines and brackets. The 5' and 3' termini of the RNA were determined in primer extension and nuclease S1 protection analyses. A consensus sequence/structure for type A bacterial RNase P RNAs is shown in panel B (2). Only nucleotide positions that are present in all instances of the RNA are shown, as letters (A, G, C or U) if invariant or dots if variable in identity. Grey lines in the background indicate the backbone of the *Escherichia coli* RNA for orientation.
Figure 2. Catalytic activity of synthetic archaeal RNase P RNAs. Assays contained 4M ammonium acetate, 300mM MgCl₂, 50mM Tris-Cl (pH 8), 0.1% SDS, 0.05% nonident P-40, 1.5nM uniformly-labeled Bacillus subtilis pre-tRNAAsp, and ca. 300nM RNase P RNA (synthesized from cloned genes by in vitro transcription), and were incubated for 3.5 hr at 45°C. The location of substrate and product bands are indicated. RNase P RNAs tested were: Escherichia coli (Eco), Metallosphaera sedula (Msc), Sulfolobus acidocaldarius (Sac), Pyrococcus furiosus (Pfu), Thermococcus celer (Tce), Methanococcus jannaschii (Mja), Methanococcus thermolithotrophicus (Mth), Methanococcus maripaludis (Mma), Methanococcus vannii (Mva), Methanobacterium formicicum (Mfo), Methanobacterium thermoautotrophicum ER-H (ERH), Methanobacterium thermoautotrophicum Marburg (MbM), Methanobacterium thermoautotrophicum ER-E (ERE), Methanobacterium thermoautotrophicum ΔH (MΔ), Methanosarcina barkeri (Mba), Haloferax volcanii (Hvo), and Natronobacterium gregoryi (Ngr) (27). Activity by the N. gregoryi RNA is evident but weak. The phylogenetic relationships between these organisms based on ssu-rRNA sequences (22) is shown as a cladogram above the gel. Placement of the ER-E and ER-H sequences, which were cloned from enrichment cultures inoculated with wastewater sludge, are based on similarity to the M. thermoautotrophicum strains ΔH and Marburg sequences.
Figure 3. Oligonucleotide-directed RNase H depletion of RNase P activity. RNase P RNAs were annealed with oligonucleotides complementary to either the *M. thermoautotrophicum* (MtΔ) or *E. coli* (Eco) RNase P RNA, then digested with ribonuclease H. Reagents included in each reaction are indicated above each lane. Activity by the *M. thermoautotrophicum* RNase P RNA was reduced only in reactions containing the *Methanobacterium*-specific oligonucleotide; reactions lacking the oligonucleotide or containing the *E.coli*-specific oligonucleotide were not inhibited. Activity by the *E. coli* RNase P RNA was reduced only in reactions containing the *E.coli*-specific oligonucleotide; reactions lacking the oligonucleotide or containing the *Methanobacterium*-specific oligonucleotide were not inhibited.
Figure 4. Reconstitution of active chimeric holoenzymes. RNase P RNAs from E. coli (Eco, 38μg/ml), M. thermoautotrophicum ΔH (MtΔ, 33μg/ml), M. formicicum (Mfo, 60μg/ml) or H. volcanii (Hvo, 170μg/ml) were assayed for RNase P activity in the presence of 0, 0.1, 1 or 10μg/ml B. subtilis RNase P protein (increasing protein concentration indicated by black wedges above the reaction lanes). Specific RNase P cleavage products are indicated by black arrows. The E. coli RNase P RNA is somewhat active by itself under these conditions, but activity is enhanced by the inclusion of the B. subtilis RNase P protein. The M. thermoautotrophicum and M. formicicum RNase P RNAs are not active in the absence of protein under these conditions, but were activated by inclusion of the B. subtilis RNase P protein. Correct processing by the H. volcanii RNase P RNA was not observed under these conditions either in the presence or absence of the B. subtilis RNase P protein. However, a specific inappropriate cleavage was generated at the highest concentration of protein; the nature of these products and the apparent mis-cleavage has not been determined.
REFERENCES


68


BIOCHEMICAL CHARACTERIZATION OF M. THERMOAUROTTROPHICUM STRAIN ΔH RNase P RNA

ABSTRACT

Archaeal RNase P RNAs from the methanobacteria, thermococci, and halobacteria, like the RNase P RNAs of Bacteria, are catalytically active in vitro without the RNase P protein subunit. The RNase P RNA from M. thermotrophicum strain ΔH was used to characterize in vitro processing of pre-tRNA. Even in optimal reaction conditions, the archaeal RNase P ribozyme substrate affinity is extremely poor. Metal ion probing of the M. thermotrophicum strain ΔH RNase P RNA indicates some discrete tertiary structure despite their functional impairment. RNA structural defects are only partially compensated for by the high salt concentrations used in archaeal RNase P activity assays.

INTRODUCTION

RNase P is the endoribonuclease responsible for processing the 5' end of precursor-tRNAs (1, 18, 23). In cells, RNase P is a ribonucleoprotein; it has a large (300 to 400 nucleotide) RNA and a protein subunit that ranges from a small, ca. 14kD single polypeptide in Bacteria to large, multiple polypeptide complexes recently identified in the yeast nucleus (4, 6, 17, 21). Previous work with bacterial RNase P demonstrated that the RNase P RNA is the catalytic subunit, and that with increased ionic strength the RNase P RNA can process pre-tRNAs in vitro without the RNase P protein (12). Simply stated, the bacterial RNase P RNA is a ribozyme. Catalytic activity by a eukaryal RNase P RNA has not been demonstrated, but it is presumed that the RNA is responsible for catalysis in the holoenzyme because it contains similar core sequence and structural components (Fig. 1).
RNase P RNA from certain Archaea (formerly archaebacteria) are, like their bacterial homologues, able to process pre-tRNAs in the absence of proteins (previous chapter). The Archaea, although more closely related to eukaryotes, have RNase P RNAs that are similar to Bacteria (Fig. 2). Sequence and structure analysis shows that the archaeal RNase P RNAs possess the same core structural components required for catalysis in Bacteria (13, 15, 26); however, the catalytic activity of archaeal RNase P RNAs *in vitro* is severely limited.

Purified archaeal RNase P enzymes and RNase P activity from archaeal cell lysates require moderate ionic concentrations (100mM NH$_4^+$, 25mM Mg$^{2+}$) similar to those of bacterial RNase P enzymes (11, 19, 20). Catalytically active synthetic archaeal RNase P RNAs and catalytically active RNA extracted from archaeal cultures (which would contain any post-transcriptional modifications) have a preference for high concentrations of both monovalent salt (e.g. 3M NH$_4^+$) and Mg$^{2+}$ (300mM). The need for extremely high salt concentrations suggests a structural impairment of the RNA. Biochemical characterization and metal ion structure probing of the archaeal RNase P RNA from *Methanobacterium thermoautotrophicum* strain ΔH reveals partially stable ribozyme with low affinity (high K$_m$) for pre-tRNA substrate.

**MATERIALS AND METHODS**

**Measurement of enzymatic activity**

RNase P RNA from *M. thermoautotrophicum* strain ΔH was synthesized from pHtrimmed#18 (22), using T7 RNA polymerase as described (11, 20). The pre-tRNA$^{32}$P from *B. subtilis* (in plasmid pDW128 (25)) was transcribed using T7 RNA polymerase in the presence of $25Cl/mmole$ α$^{32}$P-GTP. RNAs were purified by electrophoresis in
denaturing polyacrylamide gels. The appropriate bands were excised and eluted by diffusion for 16 hrs. in 50mM Tris-EDTA:0.1% SDS. After phenol extraction and ethanol precipitation, RNAs were renatured in 50mM Tris-HCl (pH 8) by slow cooling to 37°C after a three minute incubation at 80°C. RNase P RNA activity assays contained 50mM Tris-HCl (pH 8), 300 mM MgCl₂, 3M NH₄OAc and were incubated at 45°C unless otherwise indicated. Temperature optimum was determined in reactions incubated between 30°C and 75°C. Optimal pH was evaluated in activity assays that used 50 mM HEPES buffer. Reactions were resolved on 8% denaturing polyacrylamide gels; imaged and quantitated using a Molecular Dynamics Phosphorimager.

Cloning of unimolecular constructs

Unimolecular enzyme:substrate RNAs (10) were made for Methanobacterium formicicum (Fig. 3) and Methanosarcina barkeri (Fig. 4). RNase P RNA genes were amplified by PCR using the primers ΔH5'BAM and ΔH3’Xba (M. formicicum) and Msb5'Bam and Msb3'RBam (M. barkeri) which would anneal to the 5’ and 3’ ends of each RNase P RNA gene. Amplified DNA was phosphorylated with T4 polynucleotide kinase, and treated with T4 DNA ligase to generate circular genes. A second PCR amplification was performed using primers MF261FBglIIPT and MF260RHindTP (M. formicicum TP); MF274FECoPT and MF273RKpnPT (M. formicicum PT); MsbTP2505’ and MsbTP2503’ (M.barkeri TP); Msb267FPstPT and Msb266RKpnPT (M. barkeri PT) which are designed to anneal to the DNA sequence of the RNase P RNA substrate binding site (based on homology to bacterial RNase P RNA gene sequences) to create a circularly permuted (cp) RNase P RNA gene. The cpRNase P gene was cloned into the plasmid vector pDW128 immediately upstream (PT) or downstream (TP) of the gene for pre-tRNA^asp from B. subtilis (see Appendix 2). Radiolabeled RNAs were synthesized using T7 RNA polymerase.
in the presence of 25Ci/m mole α32P-GTP. Purification and activity assays were performed as described above.

**Lead cleavage**

RNase P RNA tertiary structure was probed using Pb++ and Mg++ cleavage. RNase P RNAs (E. coli, M. thermoautotrophicum strain ΔH, and M. sedula) were synthesized using T7 RNA polymerase, treated with alkaline phosphatase to remove 5' phosphates, and 5' end-labeled using γ32P-ATP and T4 polynucleotide kinase. Following purification on denaturing polyacrylamide gels, and elution (described above), RNAs were renatured in 50mM Tris-HCl (pH 8) by slow cooling to 37°C after a three minute incubation at 80°C. RNA cleavage reactions were assembled in 1M NH₄OAc, 25mM MgCl₂, 50mM Tris-HCl (pH 8). RNAs were preincubated at 45°C for 15 minutes. Pb(OAc)₂ was added to each reaction (0-250mM) and incubated for 30 minutes at 15°C. Reactions were stopped with addition of EDTA to 50mM and ethanol precipitation. The cleaved RNA was resuspended in RNA gel loading buffer, resolved over an 8% denaturing polyacrylamide gel, and analyzed by phosphorimagry.

**Oligonucleotides**

ΔH5'Bam:  
CGG GAT CCA CCG GGC AAG CCG AAG GGC

ΔH3’Xba:  
GCT CTA GAC CGG GCA TGC CGA GAG

Msb5’Bam:  
CGG GAT CCA TGC GAG AGA GGC TGG

Msb3’RBam:  
CGG GAT CCA TGC GAG TGA GGC ACG

MF261FBglIIITP:  
GAA GAT CTA GGT AAC TCG CAT AGA TG

MF260RHindTP:  
CCC AAG CTT CTG CCT CAT ACA GGA TTC

74
MF274FecoPT: CGG AAT TCA GAT GAA TGC TGC CAC C
MF273RKpnPT: GGG GTA CCA TGC GAG TTA CCT CTG CC
MsbTP2505': CGT CTA GAA ACG CAT AGC CGA ATG
MsbTP2503': AAC TGC AGA ACA ACC GGG AGA GTC CG
Msb267FPstPT: AAC TGC AGA ATG CCG TCA CTG CAA G
Msb266RKpnPT: GGG GTA CCG GCT ATG CGT TTC CAA AAA CAA C

RESULTS

Reaction optimization

Optimal reaction conditions were determined for the *M. thermoautotrophicum* strain ΔH RNase P RNA in a series of activity assays by co-titrating the concentrations of NH$_4$OAc (1-5M) and MgCl$_2$ (0-400mM) (Fig. 5). Optimal catalytic activity was reached at 3M NH$_4$OAc, 300mM MgCl$_2$. At the highest NH$_4$OAc and MgCl$_2$ concentrations tested, decreased catalysis (less substrate processing) was coupled with non-specific tRNA (and presumably RNase P RNA) decomposition (measured as loss of signal in the substrate and product gel bands). The temperature and pH optima were determined in separate experiments (Fig. 6) using optimal reaction conditions. Activity increased with temperature up to 45°C; at higher temperatures; non-specific RNA hydrolysis masked catalytic activity. The optimal pH of the reactions is 8.0; at pH > 8.5, reactions could not be assayed because of non-specific hydrolysis.

Other monovalent ions (1M, 2M, and 3M KCl or NaCl) were tested in the presence of 300mM MgCl$_2$ at 45°C for their ability to promote archael RNase P RNA activity. Each allowed some degree of catalysis, but the other monovalent ions increased the rate of non-specific hydrolytic degradation of substrate and probably the RNase P RNA (data not shown). Divalent cations (CaCl$_2$, MnCl$_2$, ZnCl$_2$, and CuCl$_2$) were tested at concentrations...
up to 300 mM in the presence of 3M NH₄OAc at 45°C; Mn²⁺ was the only divalent cation able to promote catalytic activity, but it also hastened the hydrolytic degradation of substrate. There was a four-fold greater degradation of substrate when activity peaked at 200 mM Mn²⁺ reactions then in peak Mg²⁺ reactions at 300 mM in parallel experiments with concentrations up to 500 mM (Fig. 7). Spermidine, a polyamine, was tested for its ability to enhance or aid monovalent ions in RNase P RNA reactions, but no increased activity was observed in either bacterial (previously reported, (3)) or archaean RNase P RNA reactions (see Appendix 1).

**Biochemical measurements**

Pre-tRNA cleavage by the *M. thermoautotrophicum* strain ΔH RNase P RNA is directly proportional to enzyme concentration (Fig. 8A) and time (Fig. 9). Both figures also demonstrate that *M. thermoautotrophicum* strain ΔH RNase P RNA performs multiple rounds of enzymatic turnover. Using optimal reaction conditions, substrate was titrated from 1 nM to 10 μM in an attempt to produce a maximum rate of catalysis and determine *Kₘ* (Fig. 8B). *Vₘₐₓ* could not be achieved; catalytic activity remained essentially linear throughout the range of substrate concentrations. Therefore, *Kₘ* is ≥ 10 μM in experiments where the maximum velocity obtained was approximately 3,000 nM per hour. The pre-tRNA<sup>Asp</sup> from *B. subtilis* was used as the substrate in the archaean RNase P reactions, and represents a possible source of abnormal catalytic activity, but exploratory experiments with the pre-tRNA<sup>Asp</sup> from *M. thermoautotrophicum* strain ΔH showed a similar ability to be cleaved by the catalytically active archaean RNase P RNAs, and archaean RNase P holoenzymes (22).
Unimolecular constructs

RNase P unimolecular constructs are single molecules that have the substrate RNA transcribed in the substrate binding site of the RNase P RNA (10); this creates a unimolecular enzyme:substrate complex. Unimolecular RNAs potentially bypass the substrate binding process, providing access to the RNase P catalytic step (in cis) and allowing measurement of the rate of enzymatic chemistry. Unimolecular constructs have been used extensively with bacterial RNase P RNAs. Experiments with a unimolecular construct (TP) made from *M. formicicum* (a catalytically active archaeal RNase P RNA and close relative to *M. thermoautotrophicum strain ΔH*) (Fig. 3) revealed a rate constant (k = 2.303 m, where m = the slope log[TP] vs. time) of 0.138 per minute, which is slow relative to the analogous bacterial system (TP292 made with the *E. coli* RNase P RNA (10)) which has an apparent rate constant of 1.2 per minute, a 9-fold difference (10). In addition, the *M. formicicum* TP requires 4M NH₄OAc and 400 mM MgCl₂ for optimal reactivity (see Appendix 3). Both slow activity and high ionic requirements relative to the *E. coli* TP292 are indications of global structural defects in the archaeal RNase P RNA.

Unimolecular constructs made from the RNase P RNA gene from *M. barkeri* (a non-catalytically active RNase P RNA) failed to cleave at all (data not shown) suggesting that substrate binding is not the only defect in this non-active archaeal RNase P RNA. Since the *M. barkeri* constructs and the PT construct made from *M. formicicum* did not cleave, it is possible the unimolecular RNAs are artificially defective, and that the lack of catalysis is not due to a problem with the native RNase P RNAs. Similar observations have been noted for various bacterial unimolecular constructs. The possibility of an inactive unimolecular construct is why both TP and PT constructs were made; however, unimolecular constructs exactly analogous to the *M. formicicum* TP and PT are catalytically active when based on the *E. coli* RNase P RNA (10).
**Lead cleavage**

Specific metal ion cleavage of RNase P RNAs was used as an indicator of global tertiary structure. RNase P is a metalloenzyme that requires the specific binding of Mg ions. However, certain other divalent ions can be substituted for Mg$^{2+}$. Pb$^{2+}$ is used for RNase P structural analysis because it will replace Mg$^{2+}$ and cause hydrolytic cleavages in the area of the RNase P RNA in which it is bound. 5'-end-labeled RNase P RNAs from *M. thermoautotrophicum* strain ΔH, *Metallosphaera sedula* (a non-catalytically active archaeal RNase P RNA), and *E. coli* were exposed to lead ions (0-250mM Pb(OAc)$_2$) in reactions containing 1M NH$_4^+$ and 25mM Mg$^{2+}$. Cleavages in the *E. coli* RNase P RNA generally matched previously observed sites (8, 28) of Pb$^{2+}$ induced specific hydrolysis (C123, C138, U303 as numbered in the *E. coli* secondary structure in Fig. 10). Differences in cleavage patterns were probably caused by differences in reaction conditions.

The *M. thermoautotrophicum* strain ΔH RNase P RNA had cleavages sites that roughly correspond to previously reported *E. coli* cleavage sites. There were also common cleavage sites in the *M. thermoautotrophicum* strain ΔH and *E. coli* RNase P RNAs in our experiments (see Fig. 10 for placement of cleavage sites on the *M. thermoautotrophicum* strain ΔH and *E. coli* secondary structures); some were in P9 and P12 while others were in the single-stranded loop region between P10 and P12. Cleavages were more distinct and specific in the *E. coli* RNase P RNA, suggesting a more discrete tertiary structure with well-defined metal binding sites. The *M. sedula* RNase P RNA produced mostly faint bands that were generally not Pb$^{2+}$ concentration dependent. However, there were some bands that corresponded to cleavages observed for the other RNAs tested (Fig. 11).
DISCUSSION

Biochemical properties

In vitro RNA-alone catalytic activity indicates archael RNase P RNAs have the necessary structural components to bind and cleave pre-tRNA. The necessary structural components are present and arranged in a discrete, stable, functional conformation. RNA-alone catalytic activity is a balance of the electrostatic repulsion of the negatively charged phosphate backbone and the folding energy that drives the formation of tertiary structure. If folding energy is decreased by deletion of structural components or removal of the RNase P protein subunit, causing the molecule becomes less stable, structure can be restored by decreasing electrostatic repulsion through the addition of monovalent cations. The balance must also be maintained when temperature is increased and RNA structure become less stable. Increased stability can be achieved by increasing folding energy with additional stabilizing structural components or by further shielding electrostatic repulsion with higher concentrations of salts. Structurally defective RNase P RNAs have been rescued in vitro by increasing the ionic strength or concentration of Mg^{2+} in activity assays (15, 26). Increased salt concentration has also been used to stabilize RNase P RNA at elevated temperatures (3).

The biochemical characterization of the M. thermoautotrophicum strain ΔH RNase P RNA demonstrated that cleavage of pre-tRNA was dependent on enzyme concentration and time. The high salt concentrations required for optimal M. thermoautotrophicum strain ΔH RNase P RNA activity indicates the need for extensive structural stabilization (electrostatic shielding). Even under optimal conditions, the M. thermoautotrophicum strain ΔH RNase P RNA has low affinity for substrate (extremely high $K_m$), which is most likely caused by
persistent structural defects. However, it is possible that the archaeal RNase P RNA structural defect only effects substrate binding and that catalysis is unaffected.

The construction of unimolecular enzyme:substrate molecules potentially removes substrate binding from the reaction cycle and allows for the direct measurement of catalysis. The increase in reaction rate observed for unimolecular constructs suggests that transcribing the substrate RNA at the site of catalysis improves the rate of activity. The *E. coli* unimolecular construct (TP292) (10), is structurally compromised (requires 3M salt) and has a rate constant that is ca. 100-fold faster the corresponding native reaction (at $K_m$) and is ca. 50-fold faster than the $V_{\text{max}}$ of the *E. coli* RNase P RNA.

The *M. formicicum* derived unimolecular construct (TP) is also structurally compromised. Optimal salt conditions for the TP construct are higher then the *M. thermoautotrophicum* strain ΔH RNase P RNA, but this was expected because of the increased ionic requirements of the bacterial unimolecular constructs. A plot of log[TP] vs. time (used to calculate the rate constant) is shown in the Figure 9 inset. The biphasic nature of the curve suggests two populations of TP molecules; one that is structurally competent and fast-acting, and one that is slow to form the active conformation. Half-life and rate constant were determined using the steep-lined phase that corresponds to the initially active population of unimolecular RNAs.

The archaeal unimolecular reaction cannot be directly compared the native *trans* reaction because of the inability of the *M. thermoautotrophicum* strain ΔH RNase P RNA to reach $V_{\text{max}}$. This prevents the calculation of various biochemical parameters. However, if the relationship between bacterial *cis* and *trans* reactions is the same for archaeal reactions, the apparent rate constant for the *M. thermoautotrophicum* strain ΔH RNase P RNA should be about 100-fold slower, ca. 0.0014 min$^{-1}$.
Tertiary structure

Metal ion cleavage of the *M. thermoautotrophicum* strain ΔH ribozyme was used to evaluate the state of the RNA tertiary structure compared to that of the RNase P RNAs from *E. coli* (presumed to have a stable tertiary structure) and *M. sedula* (a non-active RNase P RNA). There were distinct Pb**⁺** and Mg**⁺⁺** dependent cleavage bands for the *E. coli* RNase P RNA (some matched those reported in previous experiments); all of which reflected the state of a stable RNase P RNA tertiary structure. The metal cleavages of the *M. thermoautotrophicum* strain ΔH RNase P RNA were compared to those from the *E. coli*. Although the two RNAs are distantly related, previous experiments with distantly related bacterial RNase P RNAs showed that stable RNase P RNA tertiary structures will have some common cleavages (5). The *M. thermoautotrophicum* strain ΔH RNase P RNA produced a band pattern that was weaker than that of the *E. coli* RNase P RNA, but still indicated some degree of distinct tertiary structure. Metal binding sites were present in the *M. thermoautotrophicum* strain ΔH RNase P RNA tertiary structure and were mapped to the same general regions of the *E. coli* secondary structure. The cleavages of the *M. sedula* RNase P RNA were faint and few in number, and when compared to the results of *M. thermoautotrophicum* strain ΔH and *E. coli* suggested the lack of a stable, catalytically competent tertiary structure. Our interpretation of the metal ion cleavage data is that specific bands indicate stable tertiary structure, and the observation of such bands in the *M. thermoautotrophicum* strain ΔH RNase P RNA, but not *M. sedula*, suggests that the deranged tertiary structure of the *M. sedula* RNase P RNA is the reason it is not catalytically active.

Additional evidence for partial tertiary structure is that chimeric RNase P holoenzymes were reconstituted from certain archaeal RNAs and RNase P protein from *B. subtilis* (previous chapter in this thesis). The archaeal RNase P RNAs required ≥ 100-fold higher
protein concentration than the *E. coli* RNase P RNA, but even minimal reconstitution suggests that the archaean RNase P RNAs resemble the bacterial tertiary form.

**Similarity to bacterial mutants**

Variable helical appendages are thought to dock into the RNA superstructure and stabilize global structure (6, 16). The deletion of variable structural elements often results in destabilization (15, 26). The conserved core structural elements are present in RNase P RNAs from both bacterial and archaean evolutionary domains (13, 15). In fact, archaean and bacterial RNase P RNAs have secondary structures similar to the point that archaean RNase P RNAs more closely resemble those of the Bacteria than those of their sister taxa, the eukaryotes (Fig. 1) (7, 9, 24). In addition, the methanobacterial RNase P RNAs have a P15/P16 region (a region thought to be critical for catalytic activity) similar to that of Bacteria. The archaean RNase P RNAs are not missing any helical elements that are not variably present in Bacteria, but extreme ionic requirements for in vitro catalytic activity are strong evidence for structural defects.

The extremely high $K_m$ of the *M. thermoautotrophicum* strain ΔH RNase P RNA reaction suggests a need for structural stabilization. The preference of the *M. thermoautotrophicum* strain ΔH RNase P ribozyme for high monovalent and divalent salt concentrations resembles severely defective, mutant bacterial RNase P RNAs. The stabilization of the archaean structure, like the bacterial mutants, is achieved by increased cation concentration (10, 15, 26, 27).

The Min 1 mutant RNase P RNA has non-conserved portions of helices 3, 12a, 13, 14, 16 and 17 (see Fig. 10 for helix numbering) removed. *E. coli* ΔP18 has P18 replaced with a C nucleotide to resemble the sequence and structure of the homologous area in *Chlorobium* species. The *Chlorobium* RNase P RNA does not have a P18, a helix shown to have
specific tertiary structure stabilizing effects (the archaeal RNase P RNAs lack P18 (Fig. 1) (13)). In both Min 1 and \textit{E. coli} ΔP18 RNase P RNAs, optimal activity was observed at 3M NH4⁺.

The Min 1 and \textit{E. coli} ΔP18 RNase P RNAs (and the RNase P RNA from the \textit{Chlorobium}) function in moderate (15 to 100mM) Mg⁺⁺. The bacterial RNase P RNA mutant, Micro P, was constructed based on the sequence and structure of the RNase P RNA from \textit{Mycoplasma fermentans} (26). Compared to the \textit{E. coli} RNase P RNA, Micro P lacks the helices P6, P13, P14, P17, P18, and most significantly, P12, which has been found in all bacterial, archaeal, and eukaryal RNase P RNAs except \textit{M. fermentans}. The Micro P, like other mutants, requires high monovalent salt concentrations, but also needs 300mM Mg⁺⁺ for optimal catalytic activity. The high archaeal RNase P RNA Mg⁺⁺ optimum may indicate a persistent structural defect that, like the Micro P, requires this Mg⁺⁺ concentration to drive those ions into mis-shaped binding sites.

\textbf{Conclusion}

The \textit{M. thermoautotrophicum} strain ΔH RNase P RNA tertiary structure is stable enough to be catalytically active \textit{in vitro}, but retains structural defects that impair its ability to bind substrate. In cells, RNase P is the result of billions of years of evolutionary selection; it is assumed that the enzyme works as well \textit{in vivo} in Archaea as in any other organism. \textit{In vitro} experiments with reconstituted bacterial RNase P subunits (RNA and protein) reduced the need for salt from 1M to 0.1M and increased the catalytic efficiency of the RNase P reaction by improving the rate of product release (9). The uncharacterized archaeal RNase P protein(s) must also provide structural stability to the RNA, and compensate for the structural instability of the RNA observed \textit{in vitro}.
Figure 1. RNase P RNA consensus structures. Capital letters (GAUC) represent bases at least 80% conserved, closed circles represent bases 80% present. A comprehensive eukaryal consensus structure has not been reported; the vertebrate consensus was chosen as an example. The position of P12 is labeled, but because of variation in sequence length and structure, it is not shown (13, 24, 26).
Figure 2. A three-domain tree of life (35). The boxed Archaea are those with RNase P RNAs that are active in vitro without RNase P protein. All bacterial RNase P RNAs tested are active without RNase P protein. In vitro RNA-alone activity has not been reported for any eukaryotic or organellar RNase P RNA.
Figure 3. *M. formicicum* unimolecular constructs, TP and PT. RNase P cleavage sites are indicated by arrows. Lower-case letters indicate non-RNase P or tRNA sequence.
Figure 4. *M. barkeri* unimolecular constructs. TP and PT.RNase P cleavage sites are indicated by arrows. Lower-case letters indicate non-RNase P or tRNA sequence.
Figure 5. Optimization of ΔH RNase P RNA catalytic activity. All reactions contained 50mM Tris-HCl pH8. NH$_4^+$ and Mg$^{++}$ titration. MgCl$_2$ (0-500mM) was added to reactions containing NH$_4$OAc (1-5M) and RNase P RNA activity was plotted as percent of optimal activity.
Figure 6. The effect of temperature (A) and pH (B) on the *M. thermoautotrophicum* strain ΔH RNase P RNA. Reactions were performed in 3M NH$_4$OAc, 300mM MgCl$_2$. 
Figure 7. Mn^{++} versus Mg^{++} graph of catalytic activity. (A) Effect of each metal on catalytic activity. (B) Effect of each metal on substrate stability measured by decrease in total substrate counts. pH was the same in reactions for each metal.
Figure 8. Enzyme and substrate concentration dependence. Using optimized reaction conditions (3 M NH$_4$OAc, 300 mM MgCl$_2$, 50 mM Tris-Cl pH 8 at 45°C), (A) the *M. thermoautotrophicum* strain ΔH RNase P RNA was titrated (5-100 nM) to show activity was linearly dependent on enzyme concentrations. 1.5 nM substrate was used in 2 hr. reactions. Activity was plotted as percent of substrate processed. (B) The pre-tRNA$^{{\text{Asp}}}$ from *B. subtilis* was titrated up to 10 μM and velocity was plotted as nM cleaved/hr.
Figure 9. Time course of catalytic activity of *M. thermoautotrophicum* strain ΔH RNase P RNA and the TP unimolecular construct made from *M. formicicum*. Reactions contained 3M NH₄OAc, 300mM MgCl₂, 50mM tris-HCl at 45°C for 2 hrs. RNase P activity was plotted as percent of substrate processed. Inset: log[TP] vs. time plot to determine exponential decay rate constant.
Figure 10. Secondary structures of the *E. coli* and *M. thermoautotrophicum* strain ΔH RNase P RNAs. Arrows indicate sites of Pb** or Mg** dependent cleavage of each RNase P RNA. Large filled arrows represent strong Pb** sites; small filled arrows, weak Pb** sites. Large open arrows represent Mg**-only dependent sites; small open arrows, weak Mg**-only dependent sites.
E. coli  Mth H  Mse

PbOAc

230
190
163
136
123
110
97
80
65

Figure 11. Metal ion cleavage of RNase P RNA. E. coli, ΔH, and M. sedula RNase P RNAs were exposed to PbOAc (0-250mM) in 1M NH₄OAc, 25mM MgCl₂, 50mM Tris-HCl pH8.0.
REFERENCES


2. Beebe, J.A., J.C. Kurz, and C.A. Fierke. 1996. Magnesium ions are required by Bacillus subtilis RNase P RNA for both binding and cleaving precursor tRNA^{Asp}.


APPENDIX 1

THE EFFECT OF SPERMIDINE ON THE *M. THERM OAUTOTROPHICUM* STRAIN ΔH RNase P RNA.

Purpose:

Polyamines such as spermidine are generally thought to stabilize RNA and DNA structure. By binding to the sugar-phosphate backbone, polyamines shield phosphate repulsion and reduce ionic strength requirements; however, the bacterial RNase P RNA does not respond to the presence of polyamines (previous chapter, and reference 3). The archaeal RNase P RNA *in vitro* reaction depends on extreme ionic strength apparently for structural stability of the RNA. Spermidine was therefore tested for its ability to decrease the salt dependence of *M. thermoautotrophicum* strain ΔH RNase P RNA enzymatic activity.

Procedure:

RNase P RNA activity assays were performed in the presence and absence of 1mM spermidine. MgCl₂ concentration was held constant at 300mM, while NH₄OAc was titrated between 0 and 5M. 50mM Tris-HCl pH 8 was included in reactions that contained 100nM *M. thermoautotrophicum* strain ΔH RNase P RNA and 10nM ³²P-labeled *B. subtilis* pre-tRNA⁴⁴⁸ substrate. Reactions were incubated for 1.5hrs. at 45°C. Assays were resolved on an 8% polyacrylamide gel and quantitated by phosphorimagry.

Results and conclusions:

Enzymatic activity of reactions with and without spermidine, in a range of NH₄OAc concentrations, is shown Fig A1. Although the two curves do not match exactly, the presence of 1mM spermidine does not notably enhance the catalytic activity of the *M.*
thermoautotrophicum strain ΔH RNase P RNA or reduce its dependence on monovalent cations.
Figure A1. The effect of 1mM spermidine on the *M. thermosautrophicum* strain ΔH RNase P RNA in vitro reaction. MgCl₂ concentration was held constant at 300mM, while NH₄OAc was titrated between 0 and 5M. 50mM Tris-HCl pH 8 was included in reactions that contained 100nM ΔH RNase P RNA and 10nM *B. subtilis* pre-tRNAASP substrate. Reactions were incubated for 1.5hrs. at 45°C.
APPENDIX 2

SITE DIRECTED MUTAGENESIS TO REPAIR MUTANT M. BARKERI PT UNIMOLECULAR CONSTRUCTS.

Purpose:

Several attempts were made to create the M. barkeri PT unimolecular construct, but all initial cloning produced artifactual mutations, including sequence deletions. The decision was made to repair one of the mutant clones by site directed mutagenesis (SDM) rather than continuing to generate the desired unimolecular construct de novo.

Clone PT17 contained a point mutation (G41→A in the M. barkeri RNase P RNA) at a universally conserved position, and a cluster of substitutions at the site of one of the PCR primers used to generate the clone (Fig. 3A). Sequential SDM using the oligonucleotides Mbp267-41G to repair the G41→A and Msb267FPst-sdm to repair several G→C substitutions between M. barkeri positions 267-284 was successful in creating the desired sequence. (see below for oligonucleotide sequences)

Procedures:

The following protocol was taken from the lab procedure files and was used in all site directed mutagenesis experiments. Step-by-step instructions are listed to ensure reproducibility of this detailed procedure in the future.

OLIGONUCLEOTIDE-DIRECTED SITE SPECIFIC MUTAGENESIS

Introduction
In this procedure, a synthetic oligonucleotide with an internal mismatch is used to direct the generation of a mutation. The mutagenic primer is annealed to a uracil-containing single-stranded template prepared in dut', ung' E. coli CJ236 (Kunkel). The primer is extended with Klenow DNA polymerase and the resulting nicked circles ligated with T4 DNA ligase. The heteroduplex DNA containing the mismatch is then used to transform competent dut', ung' cells, wherein newly synthesized strand gives rise to progeny plasmid DNA. The mutants can then be identified by direct sequencing, restriction analysis, etc.
In order to increase the frequency of mutant isolates, template-strand progeny are eliminated by preparing the single-stranded template in *E. coli* dut-, ung* cells. These cells are deficient in production of dUTPase, an enzyme that hydrolyzes deoxyuridine triphosphates, and allows incorporation of one uracil residue about every 100 bases. These residues are normally removed by uracil-N-glycosylase, which is also lacking in this strain. The heteroduplex DNA is transformed in ung* cells, so the wild-type template strand is selectively degraded, thereby enhancing the recovery (to 50-70%) of progeny derived from the non template (mutant) strand.

**Preparation of U-containing Single Strand Template DNA**

1. Transform *E. coli* CJ236 (dut, ung*) with target plasmid DNA. Spread on 2X YT plates containing chloramphenicol (Cam) (25 ug/ml) and drug to select for target plasmid (usually Amp at 200 ug/ml). Cam is required for CJ236 cells.

2. Grow a 10 ml overnight culture of isolate in 2XYT + drugs (as above)

3. Inoculate 50 ml of 2XYT (no drugs!) with 1.0 ml of overnight culture
   - add uridine to 0.25 ug/ml.
   - shake @ 37C, 30 min. (250 ml flask)
   - add 200 ul helper phage M13KO7 (>1 x 10^11 pfu/ml)
   - shake @ 37C, 30 min.
   - add 70 ul kanamycin (kan) (50 mg/ml) and continue shaking overnight

4. Transfer cells to Oak Ridge tubes and pellet @ 12KRPM 15 min.
   - pour supernatant into a fresh tube and spin again 15 min.
   - transfer the supernatant to a 50 ml graduated dispo-tube
     - and add 1/4 volume 3.5M NH4OAc / 20% PEG (MW 6000)
   - invert to mix and pour into a fresh centrifuge tube
   - leave on ice 30 min.
   - centrifuge 15 min, 17,000g
   - thoroughly drain the supernatant and resuspend the pellet in 200 ul TE pH8 (10mM Tris-HCl pH8 / 1mM EDTA).
   - transfer to an Eppendorf tube and add 100 ul buffer saturated phenol and 100 ul chloroform
   - vortex for 1 full minute and spin 5 min. in a microfuge
   - remove the top (aqueous) phase and transfer to a fresh tube.
     (leave behind the large interface)
   - repeat the extraction until there is only a slight interface
     (5-6X)
   - add 200 ul chloroform, vortex, and spin for 2 min.
   - transfer the aqueous phase to a fresh tube

5. Precipitate ssDNA with 100 ul 7.5M NH4OAc and 600 ul ethanol
   - leave on ice for 30 min. Spin for 15 min at 4°C
   - drain the supernatant and carefully rinse the pellet with 95% EtOH
   - dry in speed vac (no heat)- dissolve pellet in 40 ul TE

6. Examine on a 1% minigel - load ca. 1ul per lane. The ssDNA should migrate 5-10% faster that the supercoiled dsDNA if the gel is run in TAE buffer.
7. Run the entire ssDNA sample in a ca. 50mm wide lane on a 1% preparative low-melting agarose gel. Load the sample with 5% glycerol ONLY added, rather than tracking dyes (run the dyes in an adjacent lane). Stain with ethidium bromide (or examine directly via UV-shadow) and carefully excise the ssDNA band with razorblade. The purpose of the gel-purification is to remove traces of dsDNA, so don't take any extra agarose; don't be greedy.

8. Weigh the gel slice, add 4 vol of TE and heat to 65°C to melt the gel.
   - add an equal volume of phenol, vortex, and spin for 2 min.
   - transfer the upper (aqueous) phase to a fresh tube
   - repeat the phenol extraction & collect the aqueous phase
   - add 1/2 vol 7.5M NH₄OAc and 3 vol EtOH - leave on ice 30 min.
   - Spin 15 min. (Sorval SS34 - 12KRPM), drain and dry the pellet.
   - dissolve the pellet in 250ul TE & transfer to a microfuge tube
   - add 125ul 7.5M NH4OAc and 1.125ml EtOH - leave on ice 30 min
   - spin 15 min, drain and dry the pellet - dissolve in 20ul TE

9. Quantitate ssDNA
   - dilute 2.5 ul of yield to 500 ul in H2O (1:200)
   - measure absorbance at 260nm
   - assume 1 A260/ml = 40 ug/ml ssDNA

**Preparation of oligonucleotide primer**

1. In most instances, the oligonucleotide primer will not require purification. If the oligo comes in ammonium hydroxide, simply dry the oligo in a speedvac, redissolve in an equal volume of H2O & redry, then resuspend in TE to a final conc. of ca. 2 mg/ml

2. Quantitate oligo
   - prepare a 300ul sample diluted to ca. 10ug/ml
   - measure Absorbance at 260nm
   - assume 1 A260/ml = 20 ug/ml oligomer

**5'-Phosphorylation of Oligonucleotide**

If the oligonucleotide was synthesized with 5' phosphate, skip this step.

Reaction:
10 ug oligonucleotide
5 ul 10X Kinase Buffer *
5 ul 10mM ATP
H2O to 49 ul
1 ul Polynucleotide Kinase (10 units)
----------
50ul final volume
- incubate 60 min., 37°C
- heat 65°C, 10 min, to inactivate kinase

* 10X Kinase Buffer:
  0.5 M Tris Cl (7.6)
  0.1 M MgCl2
  50 mM dithiothreitol
  1 mM spermidine
  1 mM EDTA

In vitro mutagenesis (as in Maniatis, Molecular Cloning, section 15.63)

1. Set up the following three reactions in 0.5ml microfuge tubes:

   (1)   (2)   (3)
   1 µg ssDNA 1 µg ssDNA 1 µg ssDNA
   1 µl oligo (20ng) 1 µl oligo ----
   5 µl 10X NTB* 5 µl 10X NTB 5 µl 10X NTB
   H2O to 25 µl H2O to 25 µl H2O to 25 µl

* 10X NTB (nick translation buffer)
  500mM Tris-Cl pH 7.2
  100mM MgSO4
  1mM DTT
  500ug/ml nuclease-free BSA

2. Incubate at 80°C for 2 min, slow cool to room temp by removing heat block from heater (about 30 min) to anneal primer to ssDNA

3. Transfer 12.5 µl of each reaction into fresh tubes, then add:

   (1)   (2)   (3)
   1 µl 2.5mM dNTPs 1 µl 2.5mM dNTPs 1 µl 2.5mM dNTPs
   2.5µl 10mM ATP 2.5µl 10mM ATP 2.5µl 10mM ATP
   6.5µl H2O 9 µl H2O 6.5µl H2O
   2.5µl Klenow (12u) ---- 2.5µl Klenow
   0.5µl ligase (4u) ---- 0.5µl ligase

4. Incubate overnight at 15°C.

5. Transform 12.5µl of each reaction into 100 µl of competent E. coli DH5αF'. Plate 1/10 of each transformation onto Xgal/Amp plates, and the remaining 9/10 onto another plate. Incubate overnight at 37°C.

6. Run the remaining reaction volumes on a 1% agarose gel, along with samples of untreated ds & ssDNA. Reaction (1) should be smeared from the original ssDNA band to another band migrating about 1/3 slower (i.e relaxed dsDNA). The other reactions should look like the untreated ssDNA.
7. Reaction (1) should yield numerous colonies (numbers would depend on the competency of the cells). Reaction (3) should yield no more than 1/3 of the number of colonies in (1) from random-priming, and reaction (2) should yield few or no colonies (from dsDNA in the ssDNA prep).

The method used to screen clones for the mutation will depend on the design of the mutagenic primers. If the mutation creates or removes a restriction site, restriction digests of miniprep DNAs can be used to screen the clones. Otherwise, the mutation can be detected by sequence analysis or PCR (if the mutation is a large insertion or deletion). We routinely screen 6-10 clones per experiment, and 30-70% of the clones generally are found to contain the mutation.

References


**Oligonucleotides:**

Mbcp267-41G: GGA GAC TTT CCT CAG ACC CGA AG

Msb267FPst-sdm: CCT GAG GAC TCT TGC AGT GAC GGC ATT CTG CAG TTA TTC CCT ATA G

**Results & conclusions:**

In each round of site directed mutagenesis, the desired alteration was obtained. Figure A2 shows the correct secondary structure and sequence of the *M. barkeri* PT unimolecular construct with the original mutations highlighted, and the location of the repair oligonucleotides designated.
Figure A2. Characterization of the *M. formicicum* unimolecular construct, TP. (a) Reaction optimization (b) determination of concentration independence. (c) time course of the unimolecular reaction. Reactions were performed as described in procedures and above.
APPENDIX 3

CHARACTERIZATION OF THE ARCHAEOAL RNase P RNA UNIMOLECULAR ENZYME:SUBSTRATE COMPLEXES.

Purpose:

Unimolecular enzyme:substrate RNAs are designed to examine catalysis in cis. Substrate and enzyme are forced into close proximity which essentially eliminates the substrate binding step from the catalytic cycle. Unimolecular constructs of bacterial RNase P RNAs have cleavage rates roughly 100 fold faster the corresponding trans reaction, which provides a close approximation to the rate of enzymatic chemistry.

Both PT and TP unimolecular constructs were created using the *M. formicicum* and *M. barkeri* RNase P RNAs to examine catalysis and to address the possibility that by providing the stable tRNA substrate in approximately the right position it might serve as a scaffold for folding of the RNase P RNA. The following experiments were performed to test and characterize the enzyme:substrate RNAs.

Results & conclusions:

The *M. barkeri* unimolecular RNAs (both PT and TP) and the *M. formicicum* PT RNA lack catalytic activity; cleavage was never observed in reactions that were tested under a variety of conditions (data not shown). *E. coli* RNase P RNA was added to inactive unimolecular reactions to cleave the pre-tRNA from the unimolecular constructs to ensure the substrate was folded into a recognizable structure that could be removed by RNase P.

The *M. formicicum* enzyme:substrate complex was catalytically active in vitro. Optimal ionic conditions were determined in a series of NH₄OAc and MgCl₂ titrations (Fig. A3a). The *M. formicicum* TP construct preferred 4M NH₄OAc and 400mM MgCl₂, which is
higher than the optima for the *M. thermoautotrophicum* strain ΔH RNase P RNA (3M NH₄OAc and 300mM MgCl₂) and suggestive of poor structural stability.

The *M. formicicum* TP unimolecular construct was tested for concentration dependence to ensure that the observed reaction was in cis. At each concentration of the *M. formicicum* TP tested cleavage activity was constant, which confirms molecules are self-cleaving (Fig. A3b). The rate of the cis reaction was initially studied in a time course reaction (0-250 min.) (Fig. A3c). Like the *M. thermoautotrophicum* strain ΔH RNase P RNA reaction, the unimolecular cis cleavage is slow compared to bacterial systems, which reflects the general structural instability of the archaeal RNase P RNA.
Methanosarcina barkeri cp267 RNase P RNA

Built from two site-directed mutagens. The final product called PT-sdm-Pst.

Bacillus subtilis
pre-tRNA(Asp)

Figure A3. Sequence repair of the M. barkeri unimolecular PT construct by site-directed mutagenesis. Non-native sequences are shown as lower case letters. The large arrow indicates the RNase P cleavage site.
APPENDIX 4

DETERMINATION OF THE 5' AND 3' ENDS OF THE METHANCOCOCUS JANNASCHII RNase P RNA

Purpose:

Sequence alignment and phylogenetic structural analysis demonstrated that there are highly conserved RNase P RNA structural elements missing from the M. jannaschii RNase P RNA. To test for structural elements 5' and 3' of the P1 region that might compensate for the missing structural components, end-mapping of cellular RNase P RNA was performed.

Procedures:

End mapping was performed using the methods described in reference 56 of the first thesis chapter. The location of the M. jannaschii RNase P RNA 5' end was mapped by primer extension using the oligonucleotides Mfo145R and Mjan3'Bam (see Oligonucleotides). The location of the 3' end was determined by S1 nuclease mapping using a probe DNA fragment that overlapped the 3' end of the M. jannaschii RNase P RNA and containing sequences downstream of the gene for the RNase P RNA, made by polymerase chain reaction amplification from genomic DNA using oligonucleotide primers Mja155(F)EcoRI and Mja3'(231)XbaReverse.

Oligonucleotides:
Mfo145R: CGT GTC GTT TCT GCT CC
Mjan3'Bam: CGG GAT CCG GTA TGG GGG CTA TAG C
Mja155(F)EcoRI: GGA ATT CCG GAG AAC CGG TGA AAC G
Mja3'(231)XbaReverse: GCT CTA GAC ACT ATC CAT TTT CTC CC
Results and conclusions:

The 5' and 3' ends of the M. jannaschii RNase P RNA were determined to be identical (relative to secondary structure) to those of M. thermoautotrophicum strain ΔH RNase P RNA. The ends of the RNA occur at an A:A mismatch in what was originally assumed to be the middle of helix P1; this may be a recognition site for an uncharacterized RNA processing enzyme. Two oligonucleotides were used to confirm the 5' end because of weak signals. See figure A4 for oligonucleotide binding sites and secondary structure of the M. jannaschii RNase P RNA.
Figure A4. The *Methanococcus jannaschii* secondary structure. 5' and 3' ends were determined by primer extension and S1 nuclease, respectively. The P1 extension represents a continuation of base pairing that is presumably present in RNase P primary transcript, but not in the native RNA. A similar observation was made for the native ends of the ΔH RNase P RNA.