Abstract

Vucson, Beverly Midge. Phylogenetic Comparative Analysis of Ribonuclease P RNA from Euryarchaeae. (Under the direction of Dr. James W. Brown.)

The purpose of this study has been to examine the RNA component of ribonuclease P (RNase P) from the Euryarchaeae. RNase P is found in all organisms and is essential for processing precursor tRNA. RNase P has been extensively studied in Bacteria in which it is composed of both protein and RNA components. In vitro, the RNA component is the catalytic component in the enzyme. Archaeal RNase P also contains an RNA component. Although the archaean RNA is similar to bacterial RNase P RNA, it has not been shown to be catalytic in vitro. In order to understand the catalytic defect of this RNA, a comparative analysis was used to develop a model of the euryarchaean subgroup’s RNA for comparison to that of the catalytic bacterial RNase P RNA.
Phylogenetic Comparative Analysis of Ribonuclease P RNA from Euryarchaeia.

By

Beverly Midge Vucson

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

Department of Microbiology

Raleigh, North Carolina

1996

Approved By:

[Signatures]

Chair of Advisory Committee
For my best friend, Kim
Biography

1. 1985 Graduated from Edison Sr. High School, Lake Station, Indiana

2. 1986 Served a four year term in United States Air Force

3. 1992 Graduated from the State University of New York at Plattsburgh
   Degree: B.A. Environmental Science

4. 1993 Employed for 1 year at Bristol-Myers Squibb Pharmaceutical Co.

5. 1994 Entered graduate school at North Carolina State University

Acknowledgements

I would like to thank Dr. Jim Brown for his patience and direction given to me while accomplishing this work. I have learned a tremendous amount while in his laboratory and have the utmost respect for him. In addition, thanks goes to my colleagues and friends in the laboratory, Dr. Elizabeth Haas for her assistance getting started and answers to lots of questions, and Jim Panucci and Kirk Harris for help with the computer, discussion of ideas regarding this work, and for being my friends. I also thank my graduate committee for assisting me with the direction of this work. Much appreciation goes to Ms. Linda Rudd for assisting with the editing of this work. Finally, thanks to Kim Colson for her support and encouragement during my graduate school experience.
# Table of Contents

List of Tables vi

List of Figures vii

Preamble viii

Phylogenetic Comparative Analysis of Ribonuclease P RNA 1

Introduction 1

Section I 1

Materials and Methods 6
  Sources for natural populations genomic DNA 6
  Sources for pure cultures 7
  Genomic DNA extraction 7
  PCR amplification 9
  Cloning of RNase P RNA-encoding genes 9
  Sequence determination 9
  Sequence analysis 10
  Oligonucleotide Reconstitution Assay 10

Results and Discussion 13
  PCR amplification 13
  Cloning and sequencing 15
  Oligonucleotide reconstitution assay 16
  Comparative analysis of secondary structure 21
  Archaeal versus bacterial RNase P RNA 27

Introduction 4

Section II 4

Materials and Methods 10
  Genomic DNA preparation 10
  PCR and cloning of archaeal 16S ribosomal genes 12
  Sequence determination and sorting of 16S rRNAs 12
  Sequence analysis 13

Results and Discussion 30
  16S ribosomal analysis of wastewater sludge 30

References Cited 34

Appendices 37

  Appendix 1. Synechococcus sp. R2 37
## List of Tables

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Methods for extracting genomic DNA and PCR primers used</td>
<td>8</td>
</tr>
<tr>
<td>2.</td>
<td>Universal archaeal RNase P PCR Primers</td>
<td>11</td>
</tr>
<tr>
<td>3.</td>
<td>Grouping of 16S rRNA clones</td>
<td>32</td>
</tr>
</tbody>
</table>
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Universal phylogenetic tree</td>
<td>3</td>
</tr>
<tr>
<td>2.</td>
<td>Archaeal phylogenetic tree</td>
<td>5</td>
</tr>
<tr>
<td>3.</td>
<td>Archaeal <em>versus</em> bacterial RNase P RNA</td>
<td>14</td>
</tr>
<tr>
<td>4.</td>
<td>Partial RNase P RNA plus oligodeoxynucleotide</td>
<td>17</td>
</tr>
<tr>
<td>5.</td>
<td>Oligo reconstitution assay using bacterial RNAs and specific oligonucleotides</td>
<td>18</td>
</tr>
<tr>
<td>6.</td>
<td>Oligo reconstitution assay using a universal inosine oligonucleotide</td>
<td>19</td>
</tr>
<tr>
<td>7.</td>
<td>Oligo reconstitution assay using archaeal RNAs and specific oligonucleotides</td>
<td>20</td>
</tr>
<tr>
<td>8.</td>
<td>Alignment of archaeal RNase P RNA primary sequences using the SeqApp program</td>
<td>22</td>
</tr>
<tr>
<td>11.</td>
<td>Archaeal <em>versus</em> bacterial consensus RNase P RNA</td>
<td>28</td>
</tr>
<tr>
<td>12.</td>
<td>Archaeal RNase P RNA phylogenetic tree</td>
<td>29</td>
</tr>
<tr>
<td>13.</td>
<td>Archaeal 16S wastewater clones phylogenetic tree</td>
<td>33</td>
</tr>
<tr>
<td>14.</td>
<td><em>Synechococcus</em> strain R2</td>
<td>37</td>
</tr>
</tbody>
</table>
Preamble

The primary focus of this study was to obtain RNase P RNA sequences from populations of organisms containing euryarchaeal species. These RNAs were then used in phylogenetic comparative analysis to examine their structural characteristics. These characteristics were compared to those of the catalytically proficient bacterial RNase P RNA. Because many euryarchaeal species have not been cultured in the laboratory, PCR primers were used to obtain RNase P gene sequences from populations thought to be rich in these organisms. However, none of the species typically isolated from wastewater sludge such as Methanosarcina barkerii were being amplified. Therefore, a study was done using universal archaean 16S ribosomal primers to examine the inhabitants of wastewater sludge. The results indicated that the predominant species found included novel methanomicrobiales species and a halophilic Archaeon. This result was consistent with the information obtained when examining wastewater sludge using archaeal RNase P primers. This information is useful for determining those organisms that have important roles in wastewater degradation as well as aiding in the development of different primers for successful PCR amplification of RNase P RNA genes from this environment.
Introduction

Section I

Phylogenetic comparative analysis of ribonuclease P RNA

Ribonuclease P (RNase P) is an essential enzyme present in all cells and organelles that synthesize tRNA (28, 29). It is an endonuclease that cleaves leader sequences from precursor tRNA to produce the necessary mature 5' ends of the molecule. This enzyme has been examined in all Domains of life, the Archaea, Eubacteria, and Eukarya. RNase P consists of both an RNA and protein component in all organisms excluding chloroplasts from vascular plants (11, 12, 13, 28). This enzyme is particularly interesting because at least in Bacteria, it is the RNA component, rather than the protein component, that is the catalytic element of the enzyme. In fact, this RNA is the only known naturally occurring catalytic RNA in that it turns over many substrate molecules (13, 29, 40).

RNase P has been extensively studied in the Bacteria. The holoenzyme in Bacteria is composed of a large RNA (approximately 350-400 nucleotides) and a small protein (approximately 120 amino acids) (28). Under appropriate salt conditions, the bacterial RNA alone can effectively process precursor tRNA molecules in vitro (13). In the Archaea and Eukarya, however, the RNA has not been shown to be catalytic in vitro in the absence of its protein component (25, 28). Preliminary study of archaeal and eukaryal RNase P reveals that the RNA component from these organisms is required since the enzyme is inactivated by treatment with micrococcal nuclease (28). The exception to this is the Sulfolobus acidocaldarius RNA. RNase P in S. acidocaldarius has a low buoyant density in cesium sulfate and originally appeared to be composed largely or entirely of protein (24). However, the RNA is shielded from micrococcal nuclease, presumably by a large protein component. Archaeal RNase P, including S. acidocaldarius, has an RNA similar in size to the bacterial RNA (28). For example, RNase P from Haloferax volcanii appears to be composed mostly of RNA because it is sensitive to micrococcal nuclease and has a high
buoyant density. In the case of the eukaryotes, the RNase P holoenzyme appears to be comprised of mostly protein rather than RNA. The best studied eukaryotic RNase P is that of *Saccharomyces cerevisiae*. While the size of its RNA is similar to that of Bacteria, its nucleotide sequence is quite different (5, 9). The secondary structure of the eukaryotic RNase P RNA has not been well defined since there is considerable diversity in the primary sequences of RNAs from different organisms (28, 34). A member of a gene family may produce the required RNAs necessary for eukaryotic cellular RNA processing. This is evidenced by the structural similarity of RNase P RNA to RNase MRP RNA (33, 34). When comparing RNase P RNA from the three Domains, it is important to consider their evolutionary relatedness. While the Archaea are phenotypically and morphologically similar to Bacteria, they are actually more closely related to the Eukarya (42) (Figure 1). In terms of RNase P, the RNA components of Bacteria and Archaea are similar in primary and secondary structure, yet the RNAs from Archaea and Eukarya have not been shown to be catalytic *in vitro* without their protein components.

Because the archaeal RNase P RNA component has not been shown to be catalytic while being similar in sequence and secondary structure to its bacterial counterpart, it is a useful tool for studying the structure-to-function relationship of this unique catalytic RNA. Phylogenetic comparative analysis is the approach used in this investigation to develop a model for the secondary structure of the archaeal RNA. This model must be based on a wide variety of RNAs to determine the most conserved elements of the structure which are likely essential for catalysis. The model of the archaeal RNA secondary structure can then be compared in detail to that of the catalytically proficient bacterial RNA to identify features of the archaeal RNA that might lend to its *in vitro* catalytic defect. This comparison is important for understanding the essential structural elements of the catalytically proficient RNase P RNA component.
Figure 1: Universal phylogenetic tree based on 16S rRNA analysis (14)
The purpose of this study was to obtain a number of RNase P RNA sequences from a wide range of different organisms belonging to the Euryarchaeota to develop a detailed model of the archaeal RNA's secondary structure. The Euryarchaeota is the largest subgroup of the Archaea (Figure 2). By phylogenetic comparative analysis, the structural components of the different RNAs can be identified and compared to determine the elements that are important for function. Because most Archaea have not been isolated or cultivated in the laboratory (35), a “natural populations” approach was used. In this approach, primers targeting conserved sequences are used in polymerase chain reactions (PCR) with DNA extracted from samples taken from environments suspected of being rich in Archaea. The amplified RNase P RNA-encoding genes can then be cloned and sequenced. This approach is also useful for obtaining a large number of sequences from just a few PCR experiments. In addition, similar primers were used to amplify RNase P RNA genes from cultivated species. RNase P RNA genes from cultured species are an important component of the phylogenetic analysis of RNAs obtained from natural populations. Accumulated RNA sequences are then used to construct secondary structures for the different RNAs to be used in comparative analysis.

Section II

16S ribosomal analysis of wastewater sludge

This study examines the methanogenic inhabitants of wastewater sludge. Anaerobic digestion of wastewater sludge involves the conversion of organic material into acetate, formate, hydrogen, and CO₂ by a consortium of Bacteria (10). Archaeal methanogenic microorganisms then use these substrates for the production of methane. Methane generation is the least understood step in the process because little is known about the microorganisms involved. In order for the waste degradation processes to be optimized, determination of the types and abundance of methanogenic organisms is necessary.
Archaea

Figure 2: Phylogenetic tree based on 16S rRNA analysis (14)
A molecular phylogenetic survey of the Archaea present in wastewater sludge was used in this study. Previous methods used to study methanogen populations were based on the classic approach of isolation and cultivation. This classic approach is not particularly useful for studying inhabitants of the environment because less than 1% of existing microorganisms can be readily cultivated. This study uses specific primers in PCR reactions to amplify archaeal 16S rRNA genes from wastewater sludge (31). Phylogenetic analysis of the sequences was used to reveal the abundant archaeal groups in the wastewater sludge population.

Materials and Methods
Section I

Phylogenetic comparative analysis of ribonuclease P RNA

Sources for natural populations genomic deoxyribonucleic acid (DNA):
Hot springs samples Queens Laundry (QL), Electric Monk (EM), E1, and Greenfinger pool (GFP) were all obtained from Yellowstone National Park and given as gifts by S. Barus (Pace Lab). White Mat (WM), Orange Mat (OM), Peanut Hill (PH), Blue Slime (BL), and Calcite spring samples were all obtained from deep sea hydrothermal vents and given as gifts by A. Reysenbach (Rutgers). Obsidian Pool Red (OR) and Obsdidian Pool Black (OB) were obtained from Yellowstone National Park and given as gifts by A. Reysenbach (Rutgers). Sludge (SL) was obtained from the Cary wastewater treatment plant in North Carolina. Rumenal fluid (RF) was obtained from a fistulated cow at the North Carolina State University (NCSU) veterinary school. Termite (TM) sample came from a woodpile in Apex, North Carolina. Methanogen enrichment cultures (ER) were obtained from J. Brown (NCSU). In these enrichments, wastewater sludge was inoculated by undergraduate microbiology students into ER medium (18) to enrich for the growth of methanogens. Samples were grown at 30°, 37°, 65° and 80°C. Growth was seen at 30° and 65°C. These cultures were combined for genomic DNA extraction.
Sources of pure cultures:

*Methanococcus vannielii* was purchased from the American Type Culture Collection (ATCC). *M. vannielii* was grown in ATCC medium 1343 at 37°C without shaking under 40 psi of 60% H₂/40% CO₂. *Methanobacterium formicicum* was given as a gift from John Reeve (Ohio State Univ.). *M. formicicum* was grown in a 2.5L fermenter containing ER medium (18) under constant agitation and 20 psi of 60% H₂/40% CO₂ at 37°C.

Genomic DNA extraction

The following methods were used to extract and purify genomic DNA from cultured and environmental samples. See Table 1 for specific method used in each case. The natural populations’ genomic DNA was extracted by either the 4-aminosalicylate (7) or microwave method (41). In the 4-aminosalicylate method, cell pellets were resuspended in STE (10 mM Tris-Cl [pH 9], 100mM NaCl, 1mM EDTA) (15ml/g of cells [wet weight]), treating them with lysozyme (1mg/ml) for 5 min at 37°C, adding 4-aminosalicylate and sodium dodecyl sulfate(SDS) to 4.5 and .75% respectively. This was followed by 3 freeze-thaw cycles using an ethanol/dry ice bath and 65°C water bath for 10 minutes, respectively. Phenol extractions were performed, followed by the addition of 1/10th volume 3M NaOAc with subsequent 2-propanol and ethanol precipitations. DNA was resuspended in 10mM Tris [pH8] 1mM EDTA (TE). The microwave lysis method involves heating a cell pellet for 2 minutes on high heat in the microwave, followed by treatment with TE+3.3% SDS. Phenol extractions were then performed followed by the addition of 1/10th volume 3M NaOAc and ethanol precipitations. DNA was resuspended in TE pH8. *Methanococcus vannielii* genomic DNA was extracted by treatment with 3% SDS, phenol extractions, and ethanol precipitations (35). *Methanobacterium formicicum* genomic DNA was extracted using the 4-aminosalicylate method. Genomic DNA from *Synechococcus* strain R2 and *Methanobacterium thermoautotrophicum* strains Marburg and ΔH were obtained from E. Haas (Brown Lab). Genomic DNAs from wastewater sludge, rumenal fluid,
<table>
<thead>
<tr>
<th>Natural Populations: Genomic DNA's</th>
<th>Primers used</th>
<th>Amplification Obtained</th>
<th>DNA Isolation technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Queen's Dairy</td>
<td>A59F/A347R, 59F/347R</td>
<td>59F/347R</td>
<td>Gift</td>
</tr>
<tr>
<td>Mork</td>
<td>A59F/A347R</td>
<td>No Amplification</td>
<td>Gift</td>
</tr>
<tr>
<td>E1</td>
<td>A59F/A347R, 59F/347R</td>
<td>No Amplification</td>
<td>Gift</td>
</tr>
<tr>
<td>GFP</td>
<td>A59F/A347R</td>
<td>No Amplification</td>
<td>Gift</td>
</tr>
<tr>
<td>White Mist</td>
<td>A59F/A347R, A59FM/R/A347RMB</td>
<td>No Amplification</td>
<td>4-aminosalicylate</td>
</tr>
<tr>
<td>Orange Mat</td>
<td>A59F/A347R, A59FI/A347R, A59RM/A347R</td>
<td>No Amplification</td>
<td>Microwave lysis</td>
</tr>
<tr>
<td>Peanut Hill</td>
<td>A59F/A347R, A59FI/A347R, A59RM/A347R</td>
<td>No Amplification</td>
<td>Microwave lysis</td>
</tr>
<tr>
<td>Blue Slime</td>
<td>A59F/A347R, A59FI/A347R, A59RM/A347R</td>
<td>No Amplification</td>
<td>Microwave lysis</td>
</tr>
<tr>
<td>Calcite</td>
<td>A59F/A347R, A59FI/A347R, A59RM/A347R</td>
<td>No Amplification</td>
<td>Microwave lysis</td>
</tr>
<tr>
<td>Onsetian Pool (Red)</td>
<td>A59F/A347R, A59FI/A347R, A59RM/A347R</td>
<td>No Amplification</td>
<td>4-aminosalicylate</td>
</tr>
<tr>
<td>Obsidian Pool (Black)</td>
<td>A59F/A347R, A59FI/A347R, A59RM/A347R</td>
<td>No Amplification</td>
<td>Microwave lysis</td>
</tr>
<tr>
<td>Sludge</td>
<td>A59F/A347R, A59FI/A347R, A59RM/A347R</td>
<td>No Amplification</td>
<td>Microwave lysis</td>
</tr>
<tr>
<td>Ruminal Fluid</td>
<td>A59F/A347R, A59FI/A347R, A59RM/A347R</td>
<td>No Amplification</td>
<td>Microwave lysis</td>
</tr>
</tbody>
</table>

Cloned and Sequenced Archaeal RNase P RNA's

- Methanobacterium thermoautotrophicum strain ΔH
- Methanobacterium thermoautotrophicum strain Marburg
- Methanobacterium formicicum
- ER-E (Methanogen enrichment culture)
- ER-H (Methanogen enrichment culture)
- SL2D (wastewater sludge)
- E1B (wastewater sludge)
- SL-C (wastewater sludge)

Cloned and Sequenced Bacterial RNase P RNA's

- Syntrophococcus strain R2

PCR Primers Used
- ΔH5/ΔH3
- ΔH5/ΔH3
- ΔH5/ΔH3
- ΔH5/ΔH3
- A59F/A347R
- A59F/A347R
- A59F/A347R

Table 1: Genomic DNA extraction techniques and primers used in PCR reactions.
M. formicicum, and M. vannielli were further purified using cesium chloride density gradients (37).

**PCR amplification:**
Polymerase chain reactions (36) were performed using buffer containing 50mM KCl, 10 mM Tris Cl (pH 8.3), 1.5 mM each dGTP, dCTP, dATP, and dTTP, 0.05% Nonident P40, and 200ng each primer oligonucleotide for all template DNAs tested. Amplifications using Sludge, Rumenal Fluid, and termite DNAs included 400ng/ul of Bovine Serum Album (BSA) (23). All amplifications included an initial 2 minute 94°C incubation and 30-40 amplification cycles (92°C for 1.5 min, 50-55°C for 1.5 min, 72°C for 0.5 min each cycle) followed by a 7 min 72°C incubation and a final 4°C soak.

**Cloning of RNase P RNA-encoding genes:**
PCR products were digested with appropriate restriction enzymes and purified by electrophoresis in 3% low-melting agarose gels (NuSieve GTG agarose, FMC, Rockland, ME). Agarose plugs containing the DNA bands were excised, melted at 65°C, and used directly in ligation reactions containing restriction endonuclease-digested pBluescript KS+ DNA (Stratagene). The pBluescript KS+ vectors containing the ligated DNAs were then cloned into Escherichia coli DH5αF'. Colonies containing the targeted gene inserts were screened using LB plates containing ampicillin and IPTG/XGal.

**Sequence determination**
The nucleotide sequences encoding RNase P RNAs were determined from double-stranded plasmid DNAs by the dideoxy-chain termination method (38) with either Sequenase Version 2.0 (Amersham, Arlington Heights, IL) or the FIDELITY DNA Sequencing System (Oncor, Inc., Gaithersburg, MD). M13 universal, M13 reverse and appropriate PCR primers were used in sequencing reactions.
Sequence analysis

RNA secondary structures were drawn using the archaenal RNase P RNA minimum consensus primary and secondary structures as a model. Sequences were then aligned manually using SeqApp (Don Gilbert, Indiana University). Comparative analysis of secondary structure was performed as previously described (7, 8). Regions of the RNAs used to construct PCR primers and those that are unique to a particular RNA sequence (P12, P15/16, and P19) were excluded from the analysis.

Oligonucleotide reconstitution assay

Plasmids containing PCR generated gene sequences for RNase P RNA from Escherichia coli, Thermoleophillum album, Planctomyces maris, Chlamydia psittaci, Chlamydia trachomatis, Methanobacterium thermoautotrophicum strain ΔH, Methanosarcina barkeri, and Methanothermus fervidus were linearized with Eco RI and transcribed with T7 RNA polymerase. Oligodeoxynucleotides were designed to fill in missing secondary structure of partial RNase P RNA clones obtained by PCR (Table 2). Transcription products and oligodeoxynucleotides (approximately 1ug each) in RNase P assay buffer containing 0.5 to 3M NH₄Cl were incubated at 65°C for 2 minutes and slow cooled prior to addition of ³²P-labeled Bacillus subtilis tRNA Asp. Enzyme assays were performed at 37°C to 60°C for 15 min to 24 hrs. Reaction products were separated in 8% acrylamide/7M urea gels and detected by autoradiography.

Section II

16S ribosomal analysis of wastewater sludge

Genomic DNA preparation

DNA was prepared as described above.
### Table 2: PCR primers used for amplification of RNase P RNA genes and oligodeoxynucleotides

#### Universal Archaeal RNase P PCR Primers
1. **A59Fxba** (GCTCTAGAGGAAAGTCGGA/ac/gCC)
   A347Rbam (CGGGATCTAAGGCTTg/gCTTggTT)
   Designed by Brown and Haas based on an available archaeal RNase P RNAs from both the Euryarchaea and Crenarchaea groups.
2. **A59FLixba** (GCTCTAGAGGAAAGTCGGA/ac/gCC)
   Designed by extending A59Fxba primer based on consensus sequence noticed among Archaeal RNase P RNAs.
3. **A59FSxba** (GCTCTAGAGGAAAGTCGGA/ac/gCC)
   Designed by shortening the A59Fxba primer to eliminate possible mismatching of the primer.
4. **A59FIIxba** (GCTCTAGAGGAAAg/acCTc/ac/CCTCCT)
   A347Rbam (CGGGATCTAAGGCTTg/gCTTggTT)
   Designed by comparing euryarchaeal RNAs and placing redundant bases in the sequence where variation was noticed.

#### Specific Archaeal RNase P RNA PCR Primers
1. **A59F(mb)xba** (GCTCTAGAGGAAAGTCTac/gCC)
   A347R(mb)bam (CGGGATAGGAAAGTCTac/gCTTggTT)
   Designed based on the 5' and 347R regions of *Mehtanosarcina barkeri*
2. **ΔH5'bam** (CGGGATCCACGGGGCAAGGGGC)
   Designed based on the 5' and 3' ends of *Methanobacterium thermoautotrophicum* strain ΔH. The entire structure of the *M. thermoautotrophicum* strain ΔH was obtained from the *Methanobacterium* genome sequencing project.
3. **A59Fxba(M.Bac)** (GCTCTAGAGGAAACTCCACCC)
   A347Rbam(M.Bac) (CGGGATCTAACCCACCTTCTCCT)
   Designed based on the 59F and 347R regions of the archaeal RNase P RNAs belonging to the *Methanobacterium* group.

#### Oligodeoxynucleotides for Reconstitution Assay
1. **41756-1094** (GATATCGAATTCCAGCAGGCCCCGGATCC)
   *Thermoleophilum album*
2. **43904-1194** (GATATCGAATTCCAGCAGGCCCCGGATCC)
   *Escherichia coli*
3. **45351-1194** (GATATCGAATTCCAGGCCCCGGATCC)
   *Planctomyces maris*
4. **45352-1194** (GATATCGAATTCCAGGCCCCGGATCC)
   *Chlamydia psittaci and trachomatis*
5. **Bacterial 5' 3' Inosine** (GATATCGAATTCCAGGCCCCGGATCC)
6. **43905-1194** (CCCGGGGATCGGACGTTAGAGGCAGG)
   *Methanobacterium thermoautotrophicum* strains ΔH and Marburg and *Methanobacterium fervidus*
PCR and cloning of archaeal 16S ribosomal RNA genes

16S ribosomal genes were amplified using primers 8FAPL
(5’GGCTGCAGTCTAGATCCGGTGTGGATCTCTGCGG-3’) and 1492RPL
(5’GGCTCGAGCCGACCACCGTTACCTGTTACGACTT-3’). Reactions were
performed using buffer containing 50mM KCl, 10 mM Tris Cl (pH 8.3), 1.5 mM each
dGTP, dCTP, dATP, and dTTP, 0.05% Nonident P40, 5% acetamide, and 200ng each
oligonucleotide. PCR amplifications were performed as described in Section I. The
amplified DNA products were digested with NotI and PstI, and separated by electrophoresis
in 3% low-melting agarose gels (NuSieve GTG agarose, FMC, Rockland, ME). Agarose
plugs containing DNA bands were excised, melted (65°C) and used directly in ligation
reactions containing restriction endonuclease-digested pBluescript KS+ DNA (Stratagene).

Sequence determination and sorting of 16S rRNAs

The nucleotide sequences encoding 16S rRNA genes were determined from double-
stranded plasmid DNAs by the dideoxy-chain termination method (38) with Sequenase
Version 2.0 (Amersham, Arlington Heights, IL) using M13 reverse primer. Upon initial
sequencing using M13 reverse primer to verify that sequences were archaeal 16S rRNAs,
further sequencing was done using the same primer and only a single termination reaction
(ddTTP). Positions of nucleotide T were compared for each clone and placed into groups
based on similarity to each other. Final analysis resulted in 13 different groups of clones.
A representative clone from groups having more than one representative was chosen to be
sequenced to completion by the DNA Sequencing Facility at Iowa State University. The
primers used were Universal (M13) (5’-TGTAAACGACCCGCTT-3’), Reverse-
1(M13) (5’CAGGAAAACAGCTATGACC-3’), and 515FPL
(5’GGCGATCTCTAGACTGCAGTGCCAGGCAGCCGGTAA3’).
Sequence analysis

Sequences were aligned manually using SeqApp (previously described). Masks utilized in obtaining phylogenetic trees using TreeTool (7), included the entire sequence, or excluded the hypervariable regions of the rRNA. The sequences were also compared to the 16S rRNA database and analyzed to determine nearest relatives and tree positions.

Results and Discussion

Section I

Phylogenetic comparative analysis of ribonuclease P RNA

PCR amplification:

Although many different genomic DNAs from various natural populations were examined in PCR reactions, only those from wastewater sludge yielded PCR products with the available primers. Table 1 lists the different primers tested with each of the natural populations' genomic DNAs and the results of the PCR reaction. These primers were designed based on strong consensus sequences located at the regions of nucleotides 59 (5' end) and 347 (3' end) of known archaeal RNase P RNAs representing both the Crenarchaea and Euryarchaea (Figure 3). Table 2 describes the primers used to PCR amplify RNase P RNA genes from the Archaea. As Table 1 indicates, eight new archaeal RNase P RNA genes from wastewater sludge and specific Methanobacterium species were amplified from genomic DNA using archaeal specific primers. While a large number of bacterial RNase P RNA genes have been successfully PCR amplified using a small group of specific primers, that has not been the case with archaeal RNase P RNA genes. Archaea specific RNase P RNA primers A59F/A347R were designed based on all available sequences. At the time, however, that collection of sequences was small and contained primarily halophilic representatives. A number of other primers have been designed based on variations of A59F/A347R with little success. Information from the Methanobacterium thermoautotrophicum strain ΔH genome sequencing project aided in the design of the
Archaeal versus Bacterial RNase P RNA

*Figure 3*: Two RNase P RNA representatives from the Domains Archaea and Bacteria, respectively. Helices are numbered P1-P18.
5'ΔH/3'ΔH primers. These primers were used to successfully amplify RNase P RNA genes from the 5' and 3' ends of different methanobacteria that included *Methanobacterium formicicum* and methanogen enrichment cultures inoculated with wastewater sludge. These primers were valuable for allowing amplification of intact genes as opposed to partial genes amplified with primers based on the 59F/347R regions. *Methanobacterium* specific 59F/347R primers have also been successful in PCR amplification. PCR amplification of *Methanococcus vannielii* genomic DNA with various archaeal primers has not been successful. The recent publication of the *Methanococcus jannaschii* genome sequence reveals that the 59F and 347R regions of its RNase P RNA gene have considerable variation from those of other known archaeal RNase P RNA genes. The published sequences at the 59F/347R regions are 5’GAGGAAGUUCGCC3’ and 3’AUCGGGCAGAAC5’. Based on this sequence, new primers can be constructed that will likely be successful in PCR amplifying RNase P RNA genes from the *Methanococcus* group. As more archaeal RNase P RNA gene sequences are obtained from different groups of the Archaea, primers can be designed to obtain sequences from genes out of natural populations. At this time, the integrity of PCR primers is likely the limiting factor in quickly obtaining a large enough number of archaeal RNase P RNA genes for comparative analysis.

**Cloning and sequencing**

Five RNase P RNA genes were cloned and sequenced from natural populations. Three of these were obtained from wastewater sludge and two were obtained from genomic DNA extracted from methanogen enrichment cultures inoculated with wastewater sludge. The three clones from the wastewater sludge are E1B, SL2D, and SL-C. The two clones from the enrichment cultures are ER-E and ER-H. RNase P RNA genes were cloned and sequenced from *Methanobacterium formicicum*, and *M. thermoautotrophicum* strains ΔH
and Marburg. The RNase P RNA encoding-gene from *Synechococcus* strain R2 was cloned and sequenced using bacterial specific primers (Appendix 1).

**Oligonucleotide reconstitution RNase P assay**

Primers for PCR amplification were designed based on conserved sequences found within the RNase P RNA rather than flanking regions. This results in a partial RNA sequence that cannot be directly assessed for catalytic activity. An approach has been developed to restore the secondary structure of these partial RNAs using DNA oligonucleotides (Figure 4). These DNA oligonucleotides are base paired to native sequence and flanking polylinker-derived sequences, allowing the RNA/DNA hybrid to fold properly into its native structure. This approach has been successful and useful for allowing examination of the catalytic potential of partial RNAs obtained by PCR. Partial bacterial RNase P RNAs obtained by PCR amplification were catalytically proficient when an appropriate DNA oligonucleotide that restored secondary structure was annealed to the RNA (Figure 5).

This indicates that that region of the RNA is important for its secondary structure rather than primary sequence for catalysis. A universal oligonucleotide using inosine in places of variable nucleotides in helix P2 was designed so that one oligonucleotide could be used for any *59FBam/347REco* PCR generated RNase P RNAs cloned into pBluescript KS+.

Again, catalytic proficiency was restored to these RNAs in the presence of the inosine oligonucleotide (Figure 6). A specific DNA oligonucleotide was constructed for three *Methanobacterium* species. Under a battery of different temperature, salt, and time conditions, these RNAs were not found to be catalytically proficient in the presence of the DNA oligonucleotide (Figure 7). Intact RNA’s of the above were also tested and not found to be active (data not shown). The partial RNAs from two *Chlamydia* species were also not shown to be catalytic even in the presence of a specifically designed oligonucleotide (Figures 5 and 6). These RNAs are unique in that they are the only bacterial RNase P RNAs not shown to be catalytically proficient *in vitro*. This phenomenon is currently being
Figure 4: The *Thermoleophilum album* partial RNA and a specifically designed oligodeoxynucleotide that base pairs with the structure to restore native secondary structure to the RNA.
Oligo Reconstitution Assay Using Bacterial RNAs and Specific Oligonucleotides

Figure 5: Oligonucleotide reconstitution assay using partial bacterial RNase P RNAs with specifically designed oligonucleotides to restore catalytic activity.
Figure 6: Partial bacterial RNase P RNAs were utilized with a universal oligonucleotide containing inosines to restore native secondary structure.
Figure 7: Partial archaean RNase P RNAs were utilized with specifically designed oligodeoxynucleotides to restore native secondary structure.
investigated. This assay has proven to be a useful tool for assessing catalytic proficiency in partial RNase P RNAs because many RNase P RNAs have been, and will be, obtained by PCR amplification using 59F/347R primers.

**Comparative analysis of secondary structure**

The archaeal RNase P RNA sequences were analyzed by comparative analysis (7) using a preliminary model of the archaeal primary and secondary consensus structures. RNase P RNA primary sequences were manually aligned in SeqApp (Figure 8). Alignment is based on consensus sequences among all available archaeal RNase P RNAs. Helices P12 and the region linking P15 to P6 were excluded from the alignment because of their considerable variability in sequence and length among different RNAs which renders their alignment impossible. These regions were folded using the Mfold program (7). After sequences were aligned, secondary structures were constructed using previously determined structures as preliminary models (Figures 9 and 10). Sequence covariation has been observed at previous sites of conservation in helices P10, P12, P5, P7, P8, and P15. The generally accepted proof that a helix exists is the observation of at least two compensatory changes. Compensatory changes in at least two positions were found in all proposed archaeal RNase P RNA helices except for those of P5 and P10. In those particular helices, there is only one position in which a compensatory change has occurred. These compensatory changes are interesting in that they have been consistent among groups rather than individual RNAs. For example, all of the examined RNAs from *Methanobacteria* contain an A/U base pair (bp) in helix P10 rather than a G/C bp. This A/U bp is also found in place of G/C at position 3 in helix P5 for all of the *Methanobacteria*. These changes probably represent unique characteristics of RNase P RNA from each particular group. Based on the consensus sequences and covariations observed in all available archaeal RNase P RNAs, a refined model of the core primary and secondary structures has been
Archaeal RNase P RNA Sequence Alignment

<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sulfolobus acidocaldarius</td>
<td></td>
</tr>
<tr>
<td>Halobacula volantium</td>
<td></td>
</tr>
<tr>
<td>Haloarcula marismarina</td>
<td></td>
</tr>
<tr>
<td>Neocallimicrobium gregoryi</td>
<td></td>
</tr>
<tr>
<td>Halococcus tarsus</td>
<td></td>
</tr>
<tr>
<td>Marmoricola marinum</td>
<td></td>
</tr>
<tr>
<td>M. echinosulfuritum</td>
<td></td>
</tr>
<tr>
<td>M. thermocarchesium</td>
<td></td>
</tr>
<tr>
<td>M. thermocarchesium prodigii</td>
<td></td>
</tr>
<tr>
<td>M. thermocarchesium psychii</td>
<td></td>
</tr>
<tr>
<td>En-Tech.</td>
<td></td>
</tr>
<tr>
<td>E-B</td>
<td></td>
</tr>
<tr>
<td>E-C</td>
<td></td>
</tr>
<tr>
<td>S-2D</td>
<td></td>
</tr>
<tr>
<td>Thermococcus olear</td>
<td></td>
</tr>
</tbody>
</table>

Figure 8: Alignment of archaeal RNase P RNA sequences.

Alignment is continued on next 2 pages.
Figure 8 continued.
Figure 9: Secondary structures of *Methanobacterium thermoautotrophicum* strains Marburg and ΔH, E1-B, and SL-2D RNase P RNAs.
Figure 10: Secondary structures of SL-C, ER-E, ER-H, and *Methanobacterium formicicum* RNase P RNAs.
developed (Figure 11). A phylogenetic tree based on this alignment was constructed using the GDE program (7) (Figure 12). When constructing this tree, a mask was used that eliminated helices P1, P2, P3, and P12 from the analysis, since E1B, SL2D, and Sl-C are partial RNAs obtained by PCR amplification and lack these sequences. P12 was not included as this region is highly variable and cannot be reliably aligned at this time. The RNase P RNAs in this analysis have positions as predicted based on ssu-rRNA phylogenetic analysis. The RNAs obtained with Methanobacterium specific primers are among that group on the tree. E1B and SL2D are near the Methanomicrobiales which is consistent with their known presence in wastewater sludge, although both are only distantly related to Methanosarcina barkeri. The RNA sequences obtained in this study have allowed for the refinement of the archael RNAse P RNA model. This model will be instrumental in the investigation of the catalytic deficiency of the archael RNAse P RNA in vitro. Further study is required to obtain representation of all the archael groups to develop a more detailed model of the archael RNAse P RNA secondary structure.

Archaeal versus bacterial RNAse P RNA

Although the archael RNAse P RNA is not catalytically proficient in vitro like its bacterial counterpart, the two secondary structures closely resemble each other (Figures 3 and 11). In addition, highly conserved sequences in bacterial RNAse P RNA are also present in those of the Archaea. Despite this similarity, the catalytic defect of the archael RNA must lie in the subtle differences in sequence or higher order structure (4, 14, 27).

In terms of secondary structure, one major difference between the bacterial and archael RNAse P RNAs has been identified. The Archaea lack the two base pair helix at P11 (Figures 3 and 11). Little is known about the region adjacent to P11 in the Bacteria or the homologous region in the Archaea (14). Helix P18 is not present in any of the archael RNAs, but is present in all bacterial RNAse P RNAs except two species of ‘Green Sulfur
Archaeal versus Bacterial Consensus RNase P RNA

Archaeal Consensus

Bacterial Consensus

Figure 11: These core archaeal and bacterial RNase P RNAs were constructed based on minimum observed consensus primary and secondary structure.
Figure 12: Phylogenetic tree depicting Archaeal RNase P RNAs
Bacteria'. Despite this major difference in structure, P18 is most likely not involved in the catalytic deficiency of archaeal RNase P RNA because it has been shown to not be necessary for catalysis in Bacteria (4, 14, 15). The region of P15 and P16 is important for substrate recognition in bacterial RNase P RNA (17). This region is conserved among the Bacteria but varies in the Archaea. However, the RNase P RNAs of the Methanobacteriales have a similar P15/P16 region that conforms to the bacterial sequence consensus. Both contain P15 and P16 helices connected by an A and G rich internal loop. However, these RNAs are also catalytically deficient in vitro. Studies are underway to determine whether or not the catalytic defect lies in other differences that inhibit substrate binding, catalysis, or product release. As more archaeal RNase P RNA sequences become available, the structure can be further defined to provide insight into those features of the RNA that have been preserved among all organisms and those that have been changed by evolution. These features are important in understanding the function of this RNA enzyme.

Section II

16S ribosomal analysis of wastewater sludge

A total of 34 clones containing archaeal 16S small subunit rRNA genes was isolated. To phylogenetically characterize those organisms that predominate in wastewater sludge, one clone from each sequence class containing multiple clones was sequenced (Table 3). The secondary structures of these RNAs were then used to align the sequences. Phylogenetic trees were constructed using masks that included the entire available sequence or only conserved portions of the sequence (Figure 13). It appears that those methanogens such as Methanosarcina barkeri and Methanospirillum hungatei that have repeatedly been cultivated from wastewater sludge are not the predominant Archaea present in that environment. It is interesting that the predominant clone isolated in this study was 41-1 (32% of 34 clones examined) which branched out with halophilic organisms in the tree analysis. This suggests that halophilic organisms may represent an important element of the archaeal
population in wastewater sludge. Clones 39-2 (15%), 46-1 (11%), and 69-1 (8%) all clustered together in the analysis and do not appear to be close relatives of any group of methanogens that have ever been isolated. These clones represent a novel group of Archaea that are probably the methanogens predominating in the wastewater sludge enrichment. Previously unidentified microorganisms appear to be present in abundance and probably play a significant role in the waste degradation process in wastewater sludge. New methods of isolating and cultivating them must be explored before their roles can be assessed, and before consequential progress can be made in the optimization of the waste degradation process in which these microorganisms are involved.
Table 3: Grouping of sludge 16s small subunit rRNA DNA clones based on "T-tracts corresponding to the 8Fα primer region

<table>
<thead>
<tr>
<th>1-group</th>
<th>2-group</th>
<th>3-group</th>
<th>4-group</th>
<th>5-group</th>
<th>6-group</th>
<th>7-group</th>
<th>8-group</th>
<th>9-group</th>
<th>10-group</th>
<th>11-group</th>
<th>12-group</th>
<th>13-group</th>
</tr>
</thead>
<tbody>
<tr>
<td>16c-3</td>
<td>16c-4</td>
<td>16a-1</td>
<td>13-1</td>
<td>13-1*</td>
<td>39-2</td>
<td>2-2</td>
<td>12a-4</td>
<td>40b-1</td>
<td>44c-1</td>
<td>57-1</td>
<td>38b-1</td>
<td>72-1</td>
</tr>
<tr>
<td>38-1</td>
<td>5a-2</td>
<td>56-1</td>
<td>36-1</td>
<td>36-1</td>
<td></td>
<td></td>
<td></td>
<td>60b-1</td>
<td>69-1*</td>
<td>69-1*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(*) Represents clones sent to Iowa State University for complete sequencing.
Figure 13: Phylogenetic tree based on 16S rRNA analysis
References


Appendix 1:
*Synechococcus* strain R2

![Secondary structure of *Synechococcus* strain R2](image)

**Figure 14:** Secondary structure of *Synechococcus* strain R2
RNase P RNA.