BIOCHEMICAL STUDIES ON ARCHAEAL RIBONUCLEASE P REVEAL THEMATIC CONVERGENCE IN PROTEIN-FACILITATED RNA CATALYSIS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for

The Degree Doctor of Philosophy in the Graduate

School of The Ohio State University

By

Dileep Pulukkunat, M.S.

*****

The Ohio State University
2008

Dissertation Committee:

Dr. Venkat Gopalan, Adviser

Dr. Juan Alfonzo

Dr. Mark Foster

Dr. Kurt Fredrick

Approved by

Dr. Mark Foster

Adviser

The Ohio State Biochemistry Program
ABSTRACT

This study focuses on ribonuclease P (RNase P), an essential Mg\textsuperscript{2+}-dependent ribonucleoprotein enzyme (RNP enzyme) responsible for generating mature tRNAs from their precursors in all three domains of life. The bacterial RNase P holoenzyme is composed of a catalytic RNA subunit (RNase P RNA; RPR) and a protein cofactor (RNase P Protein; RPP). Eukaryal (nuclear) RNase P from yeast and human contain one RPR with nine and ten RPPs, respectively. Archaeal RNase P is associated with an RPR and at least four RPPs (designated as POP5, RPP30, RPP21, and RPP29). None of the archaeal RPPs are homologous to the bacterial RPP but have homologs in eukaryal RNase P. Interestingly, comparison of the secondary structures of the RPRs from all three domains of life indicates a closer relationship between the bacterial and archaeal RPRs rather than between the eukaryal and archaeal RPRs. Archaeal RNase P is thus a mosaic of its bacterial and eukaryal relatives. Therefore, we expected that detailed structure-function studies on archaeal RNase P, a simpler RNP complex with fewer (predicted) proteins, compared to its eukaryal counterpart, would be instructive in revealing insights about the evolution of one of the most ancient catalytic RNP complexes in nature and help address some fundamental queries regarding the role of multiple RPPs in archaeal/eukaryal RNase P holoenzymes.
Despite having a conserved catalytic core similar to the bacterial RPR, only a few archaeal/eukaryal RPRs display ptRNA processing activity. We reasoned that the failure to promote catalysis by most archaeal/eukaryal RPRs is likely due to structural defects in these RPRs that impair substrate binding. By constructing an enzyme-substrate conjugate in which the ptRNA substrate is attached to the active site of *Methanothermobacter jannaschii* (*Mja*) RPR, which was reported to be inactive under various conditions tested in vitro, we demonstrate that the *Mja* RPR alone is capable of accurately cleaving the leader sequence of the ptRNA with a rate comparable to that observed for the bacterial RPR.

Previous studies to designed dissect the individual role of RPPs in archaeal RNase P catalysis indicated that RPPs (designated as POP5, RPP30, RPP21 and RPP29) functioned as binary complexes (either POP5-RPP30 or RPP21-RPP29). A binary complex consisting of POP5-RPP30 enhances the $k_{cat}$ of the archaeal RPR to that observed with all four RPPs. Since $k_{cat}$ sets a lower limit for the rate of the steps subsequent to substrate binding in the catalytic cycle, POP5-RPP30 was inferred to play a role in cleavage or product release. In this study, we assessed the effect of each binary complex on the cleavage step by reconstituting the ptRNA-*Mja* RPR conjugate with POP5-RPP30 or RPP21-RPP29. After establishing that the observed rate of self-cleavage was reflective of the cleavage step, the rate of RPR-alone catalysis was compared to that of the RPP-assisted reaction to ascertain the role of individual RPPs in RPR-mediated catalysis. POP5-RPP30, but not RPP21-RPP29, enhances the catalytic potential of *Mja* RPR by increasing the rate of cleavage and affinity for Mg$^{2+}$. Consistent with these functional data, results from RNase T1 footprinting experiments with *Mja* RNase P
holoenzyme suggest that the binary complex consisting of POP5-RPP30 is proximal to the universally conserved nucleotides in the catalytic domain (C domain) of archaeal RPR. Moreover, a deletion derivative of archaeal RPR that retains only the C domain and is bereft of the putative specificity domain (S domain) is able to functionally assemble with POP5-RPP30. Since the C domains of bacterial and archaeal RPRs are similar, we examined whether archaeal POP5-RPP30 would heterologously reconstitute with bacterial RPR. The ability to reconstitute functional RNase P holoenzymes using various evolutionarily linked bacterial/archaeal/eukaryal (organellar) RPRs and archaeal RPPs revealed the presence of a conserved catalytic core in RNase P from different domains of life.

Taken together, this study provides insights into the interplay of RNA and protein subunits in RNase P catalysis and into the evolutionary transition from a simple RNP (in bacteria) to a more complex RNP (in archaea and eukarya).
Dedicated to my loving family
ACKNOWLEDGMENTS

First and foremost, I express my gratitude to my advisor Dr. Venkat Gopalan for his supervision, guidance, constant encouragement, and extreme patience throughout this work. I am also grateful to him for his help in cloning various constructs and assistance with protein purification that were integral to this dissertation.

I thank Drs. Juan Alfonzo, Mark Foster and Kurt Fredrick for their time in serving in my dissertation committee and providing valuable suggestions throughout the course of this work.

I am grateful to members of the Gopalan laboratory, past and present, for their support and comradeship, especially to Dr. Hsin-Yue Tsai, a former graduate student, for her valuable suggestions during the initial stages of this work. I thank Walter Woznick, a former undergraduate student in the laboratory, for his help in cloning and purifying the *Pyrococcus furiosus* (*Pfu*) RNase P protein subunits. I am grateful to Dr. Lien Lai for her valuable feedback and painstaking editing efforts on our manuscripts. I acknowledge Wen-Yi Chen and I-Ming Cho for their friendship and useful scientific conversations and Cecilia Go for her help in constructing two RNase P RNA-pterRNA conjugates. I also thank former members of the Gopalan laboratory, Drs. Pradip Biswas, Deepali Singh and Christopher Sudhahar for their friendship and assistance during the
course of this work. I am also thankful to members of the Foster laboratory (OSU), especially Ross Wilson and Yiren Xu, for helpful discussions on archaeal RNase P.

I am grateful to American Heart Association for a two-year predoctoral fellowship (2005-2007) and the Department of Biochemistry for a Burrell travel award.

Finally, I am thankful to my immediate family and friends for their unconditional love, support and encouragement provided throughout my graduate career.
VITA

January 13, 1973……………………Born – INDIA

1990-1993……………………………….B.Sc., Chemistry, Calicut University

1993-1995……………………………….M.Sc., Biochemistry, Poona University

1995-1997……………………………….Research Assistant, Department of
          Biochemistry, Indian Institute of Science,
          Bangalore, India.

1997-1999……………………………….Research Assistant, SPIC Science Foundation,
          Chennai, India.

2000-present…………………………….Graduate Teaching and Research Assistant,
          Department of Biochemistry,
          The Ohio State University

PUBLICATIONS

      diversity and systematic phylogeny from microbial metagenomes. Nucl. Acids Res.
      35:2283-94.

      Functional reconstitution and characterization of Pyrococcus furiosus RNase P.
      * The first two authors contributed equally to this work.

      Exploring the potential of plant RNase P as a functional genomics tool. Methods

      (2001) Cleavage of bipartite substrates by rice and maize ribonuclease P.

FIELD OF STUDY

Major Field: Biochemistry
# TABLE OF CONTENTS

Abstract ............................................................................................................................... ii  
Dedication ........................................................................................................................... v  
Acknowledgments .............................................................................................................. vi  
Vita................................................................................................................................... viii  
List of Tables ................................................................................................................... xiii 
List of Figures .................................................................................................................. xiv  
List of Abbreviations ....................................................................................................... xvi  

## Chapters

1. **Introduction**.............................................................................................................. 1  
   1.1 Bacterial RNase P ................................................................................................... 2  
      1.1.1 Bacterial RPR structure ................................................................. 2  
      1.1.2 Substrate recognition ........................................................................... 6  
      1.1.3 Mechanism of RNase P catalysis ......................................................... 7  
      1.1.4 Role of the protein subunit in bacterial RNase P catalysis ................. 9  
   1.2 Archaeal RNase P ............................................................................................... 10  
      1.2.1 Archaeal RNase P holoenzyme is made up of multiple RPPs that are homologous to eukaryal RPPs ............................................................. 12  
      1.2.2 Protein-protein interactions in archaeal RNase P holoenzyme ............ 13  
      1.2.3 In vitro reconstitution of archaeal RNase P using recombinant RPR and RPPs .................................................................................. 13  
   1.3 Eukaryal RNase P .............................................................................................. 15  
      1.3.1 Yeast RNase P ....................................................................................... 15  
      1.3.2 Human RNase P ................................................................................... 17  
      1.3.3 Organellar RNase P ............................................................................. 18  
   1.4 Research objectives and approaches ................................................................... 20  

2. **Reconstitution of Methanocaldococcus jannaschii (Mja) RNase P: further insights into the role of protein cofactors in archaeal RNase P catalysis** ............................... 36  
   2.1 Introduction ............................................................................................................ 36  
   2.2 Materials and methods .......................................................................................... 38  
      2.2.1 Cloning the genes encoding the archael type M (Afu, Mja and Mma) RPRs .................................................................................. 38  
      2.2.2 Construction of genes encoding ptRNA\textsubscript{Tyr-Mja} RPR conjugates with spacers of different length .................................................. 39  
      2.2.3 Construction of the gene encoding the EcoS-MjaC RPR chimera ............ 40  
      2.2.4 Cloning the genes encoding the Mja RPPs .............................................. 41
2.2.5 Construction of clones with two archaeal RPP ORFs in tandem ........................................................................... 41
2.2.6 Purification of archaeal RPPs ........................................................................................................................................ 42
2.2.7 RNase P activity assays .............................................................................................................................................. 44
2.2.8 Kinetic analysis .............................................................................................................................................................. 46
2.3 Results ..................................................................................................................................................................................... 47
2.3.1 Archaeal type M RPR supports ptRNA processing in the absence of its cognate RPPs ................................................................. 47
2.3.2 Reconstitution of an archaeal type M RNase P holoenzyme .............................................................................................. 51
2.3.3 Similarities between partially reconstituted type A and M RNase P holoenzymes ...................................................................................................................... 52
2.3.4 Effect of Mja RPPs on the first-order rate constant for ptRNA<sub>Tyr-S<sub>3</sub>-M RPR .................................................................................................................. 54
2.3.5 An Mja RPR-<i>ptRNA<sub>Tyr-S<sub>3</sub>-M RPR</i> conjugate lacking S domain also exhibits self-processing ................................................................................................................................. 57
2.4 Discussion .................................................................................................................................................................................... 58
2.4.1 RPR is the catalytic subunit in all three domains of life ........................................................................................................ 58
2.4.2 Role of RPPs in archaeal RNase P catalysis ........................................................................................................................ 60
2.4.3 Insights from studies on the C domain of type M RPR ........................................................................................................... 64

3. Identification of a minimal catalytic RNP core of archaeal RNase P ....................................................................................... 83
3.1 Introduction .................................................................................................................................................................................. 83
3.2 Materials and methods ............................................................................................................................................................... 85
3.2.1 Cloning of the RPR deletion mutants .................................................................................................................................. 85
3.2.2 Cloning, over-expression, and purification of <i>Pfu</i> RPPs and <i>Mja</i> RPPs ........................................................................................................ 86
3.2.3 Generation of RPRs using in vitro transcription .......................................................................................................................... 86
3.2.4 RNase P activity assay ....................................................................................................................................................... 87
3.2.5 Heterologous reconstitution of RNase P holoenzymes ......................................................................................................... 88
3.2.6 Use of RNase T1-based footprinting to map RNA-protein interactions in archaeal RNase P ................................................................................................................................. 89
3.3 Results ......................................................................................................................................................................................... 90
3.3.1 Identification of the catalytic domain of archaeal type A RPR ................................................................................................. 90
3.3.2 Identification of the catalytic domain of archaeal type M RPR .............................................................................................. 91
3.3.3 Mapping the binding sites of <i>Pfu</i> POP5-RPP30 complex on the C domain of <i>Pfu</i> RPR ................................................................................................................................. 93
3.3.4 Heterologous reconstitution of RNase P using RPR and RPPs ........................................................................................... 94
3.4 Discussion .................................................................................................................................................................................... 96
3.4.1 Identification of the catalytic domain of archaeal RPR ......................................................................................................... 96
3.4.2 Heterologous reconstitution of RNase P using RPR and RPPs ........................................................................................... 98

4. Discussion .................................................................................................................................................................................... 113
4.1 Introduction .................................................................................................................................................................................. 113
4.2 Biological catalysis - lessons from ribozymes .......................................................................................................................... 114
4.3 Need for protein cofactor: tertiary structure capture of the functional
conformation in the RNA?............................................................................................................. 116
4.4 Evolutionary path of RNase P from an RNA to an RNP enzyme ......................... 121
4.5 Why is the catalytic function retained in the RNA subunit of the catalytic RNP? Why has the catalytic role not been transferred to the protein subunit(s) as one might expect from the RNA world hypothesis? ......................................................... 122
4.6 Concluding remarks ........................................................................................................ 123

List of References .................................................................................................................. 129
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Oligonucleotides used in this study for cloning various archaeal RPRs or their derivatives</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>78</td>
</tr>
<tr>
<td>2</td>
<td>Oligonucleotides used in this study for cloning various archaeal RPPs</td>
<td>79</td>
</tr>
<tr>
<td>3</td>
<td>Characteristics of different archaeal RPPs used in this investigation</td>
<td>80</td>
</tr>
<tr>
<td>4</td>
<td>Effect of <em>Mja</em> RPPs on the rate of self-processing reaction catalyzed by <em>p</em>&lt;sup&gt;Tyr&lt;/sup&gt;-S&lt;sub&gt;3&lt;/sub&gt;-M RPR and <em>p</em>&lt;sup&gt;Tyr&lt;/sup&gt;-S&lt;sub&gt;3&lt;/sub&gt;-ΔS M RPR conjugates</td>
<td>81</td>
</tr>
<tr>
<td>5</td>
<td>Oligonucleotides used in this study for cloning various archaeal RPRs or their deletion derivatives</td>
<td>110</td>
</tr>
<tr>
<td>6</td>
<td>Heterologous reconstitution of <em>E. coli</em> RPR with archaeal RPPs</td>
<td>111</td>
</tr>
<tr>
<td>7</td>
<td>Heterologous reconstitution of <em>B. subtilis</em> RPR with archaeal RPPs</td>
<td>112</td>
</tr>
<tr>
<td>8</td>
<td>List of plasmid constructs used in this study that were made by other members of the Gopalan laboratory</td>
<td>128</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>24</td>
</tr>
<tr>
<td>1.2</td>
<td>25</td>
</tr>
<tr>
<td>1.3</td>
<td>26</td>
</tr>
<tr>
<td>1.4</td>
<td>27</td>
</tr>
<tr>
<td>1.5</td>
<td>29</td>
</tr>
<tr>
<td>1.6</td>
<td>31</td>
</tr>
<tr>
<td>1.7</td>
<td>32</td>
</tr>
<tr>
<td>1.8</td>
<td>33</td>
</tr>
<tr>
<td>1.9</td>
<td>34</td>
</tr>
<tr>
<td>1.10</td>
<td>35</td>
</tr>
<tr>
<td>2.1</td>
<td>66</td>
</tr>
<tr>
<td>2.2</td>
<td>67</td>
</tr>
<tr>
<td>2.3</td>
<td>68</td>
</tr>
<tr>
<td>2.4</td>
<td>70</td>
</tr>
<tr>
<td>2.5</td>
<td>71</td>
</tr>
</tbody>
</table>

The role of RNase P in tRNA maturation
Secondary structure representation of bacterial RPRs
Schematic representation of a phylogenetically conserved minimum bacterial RPR consensus
Tertiary structures of bacterial type A and B RPRs
Inter-domain interactions in bacterial RPR
Schematic summarizing the specific contacts between the bacterial RPR and a ptRNA substrate
The role of magnesium ions in the catalytic mechanism of RNase P
Similarity and variations in archaeal RPRs
Similarities in structures of bacterial RPP and archaeal POP5
Secondary structure representation of representative eukaryal RPRs
Schematic depicting secondary structures of RPRs from three domains of life
Bacterial (S)-archaeal (C) hybrid RPR is catalytically active and processes E. coli ptRNA\textsuperscript{Tyr} accurately
Design and optimization of reaction conditions for an active site-tethered ES conjugate
Secondary structure representation of ptRNA-\textit{Mja} RPR conjugate in which the S domain has been deleted (p\textsuperscript{Tyr}-S\textsubscript{3}-\Delta S RPR)
Comparison of self cleavage rate of pt\textsuperscript{Tyr}-S\textsubscript{3}-M RPR and pt\textsuperscript{Tyr}-S\textsubscript{3}-\Delta S M RPR
2.6 SDS-PAGE profiles depicting the purification of binary complexes of *Mja* RPP .................................................................72

2.7 Reconstitution of RNase P activity using *Mja* RPR and cognate RPPs purified as binary complexes .................................................................73

2.8 Effect of *Mja* RPPs on the rate of self-processing reaction catalyzed by pt$^{\text{Tyr}}$-S$_3$-M RPR ........................................................................74

2.9 Effect of *Mja* RPPs on the rate of self-processing reaction catalyzed by pt$^{\text{Tyr}}$-S$_3$-$\Delta$S M RPR ......................................................................76

2.10 Dependence of $k_{\text{obs}}$ on assay pH for the pt$^{\text{Tyr}}$-S$_3$-M RPR self-cleavage reaction in the absence and presence of RPPs ........................................77

3.1 Identification of a pared down, mini *Pfu* RNase P .....................................101

3.2 Identification of a minimal RNP core of *Mja* RNase P ................................102

3.3 Mapping the binding sites of the POP5-RPP30 complex on the catalytic domain of *Pfu* RPR .................................................................................103

3.4 Mapping the binding sites of POP5-RPP30 complex on the *Mja* RPR ........105

3.5 Heterologous reconstitution of RNase P holoenzymes using various archaeal RPRs and RPPs ........................................................................107

3.6 Heterologous reconstitution of RNase P holoenzyme using bacterial RPR and archaeal RPPs ........................................................................108

3.7 Heterologous reconstitution of RNase P using *Reclinomonas americana* mitochondrial RPR (Ram-mt RPR) and archaeal RPPs ..........................109

4.1 CYT-18 stabilizes the P4-P6 domain of LSU intron in a manner similar to the peripheral structural component P5abc of *Tetrahymena* group I intron ....125

4.2 Step-wise protein-induced changes in the telomerase RNA during the RNP assembly pathway .......................................................................126

4.3 Model depicting the differences in global stabilization of the RNA tertiary structure in bacterial and archaeal RPRs ........................................................127
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>°C</td>
<td>Degree Celsius</td>
</tr>
<tr>
<td>3-D</td>
<td>Three-Dimensional</td>
</tr>
<tr>
<td>A&lt;sub&gt;600&lt;/sub&gt;</td>
<td>Absorbance at 600 nm</td>
</tr>
<tr>
<td>α</td>
<td>alpha</td>
</tr>
<tr>
<td>Afu</td>
<td><em>Archaeoglobus fulgidus</em></td>
</tr>
<tr>
<td>B. subtilis</td>
<td><em>Bacillus subtilis</em></td>
</tr>
<tr>
<td>C-domain</td>
<td>Catalytic domain</td>
</tr>
<tr>
<td>C-terminal</td>
<td>Carboxy terminal</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>E. coli</td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>γ</td>
<td>gamma</td>
</tr>
<tr>
<td>g</td>
<td>gram(s)</td>
</tr>
<tr>
<td>h</td>
<td>hour(s)</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl-β-D-thiogalactopyranoside</td>
</tr>
<tr>
<td>kDa</td>
<td>kilodalton</td>
</tr>
<tr>
<td>L</td>
<td>Liter</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
</tr>
<tr>
<td>M</td>
<td>Molar</td>
</tr>
<tr>
<td>MES</td>
<td>2-[N-Morpholino]ethanesulfonic acid</td>
</tr>
<tr>
<td>mg</td>
<td>milligrams</td>
</tr>
<tr>
<td>min</td>
<td>minutes</td>
</tr>
<tr>
<td><em>Mja</em></td>
<td><em>Methanocaldococcus jannachii</em></td>
</tr>
<tr>
<td>ml</td>
<td>milliliter</td>
</tr>
<tr>
<td>mM</td>
<td>millimolar</td>
</tr>
<tr>
<td><em>Mth</em></td>
<td><em>Methanothermobacter thermautotrophicus</em></td>
</tr>
<tr>
<td>NAIM</td>
<td>Nucleotide analog interference mapping</td>
</tr>
<tr>
<td>ng</td>
<td>nanograms</td>
</tr>
<tr>
<td>nm</td>
<td>nanometer</td>
</tr>
<tr>
<td>nM</td>
<td>nanomolar</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td><em>Pfu</em></td>
<td><em>Pyrococcus furiosus</em></td>
</tr>
<tr>
<td><em>Pho</em></td>
<td><em>Pyrococcus horikoshii</em></td>
</tr>
<tr>
<td>ptRNA</td>
<td>Precursor transfer ribonucleic acid</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RNase P</td>
<td>Ribonuclease P</td>
</tr>
<tr>
<td>RNP</td>
<td>Ribonucleoprotein</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolutions per minute</td>
</tr>
</tbody>
</table>
RPP    RNase P protein
RPR    RNase P RNA
S domain Specificity domain
SDS-PAGE Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
*T. maritima* *Thermotoga maritima*
Tris    Tris-(hydroxymethyl) aminomethane
tRNA   Transfer ribonucleic acid
UV     Ultraviolet
μg     Microgram(s)
μl     Microliter(s)
μM     Micromolar
CHAPTER 1

INTRODUCTION

In all three domains of life, most transfer RNAs (tRNA) are transcribed as precursors with a 5' leader, mature domain, and 3' trailer sequence. Several complex steps of ordered enzymatic processing and post-transcriptional modifications are required to generate mature tRNAs from their primary transcripts. Ribonuclease P (RNase P) catalyzes a Mg$^{2+}$-dependent hydrolytic reaction that removes the leader sequence from precursor tRNAs (ptRNAs) to form mature tRNAs (mtRNAs; Fig. 1.1). RNase P activity has been identified and characterized from representative organisms encompassing all three domains of life (Bacteria, Archaea, and Eukarya) as well as organelles such as chloroplasts and mitochondria [1-5]. RNase P holoenzymes are ribonucleoprotein (RNP) complexes composed of an essential RNA subunit and one or more protein subunits depending on the source: at least one, four, and nine in Bacteria, Archaea, and Eukarya, respectively. In bacteria and yeast, genetic analyses have confirmed that RNase P activity is essential for cell viability [6-9]. As a preamble to the objective of this dissertation, a brief description of bacterial, archaeal, and eukaryal RNase P is provided below.
1.1 Bacterial RNase P

The bacterial RNase P holoenzyme is composed of a large RNA subunit (~400 nts) and a small protein cofactor (~120 amino acid residues). The RNA component alone is catalytically active in the absence of the protein subunit under *in vitro* conditions of high ionic strength and an appropriate divalent ion such as Mg$^{2+}$ (e.g., 50 mM Tris-HCl, pH 7.5, 100 mM NH$_4$Cl, and 60 mM MgCl$_2$) [10]. However, the protein cofactor is essential for RNase P function *in vivo* [7]. RNase P and parts of spliceosomes are considered as true cellular ribozymes which recognize and process their substrates in *trans* [10-12]. Because of its relative simplicity in terms of subunit make-up, the bacterial RNase P holoenzyme has been the subject of extensive biochemical and structural studies. Much of the current information regarding substrate recognition, catalytic mechanism, and tertiary structures of the individual RNA and protein subunits were obtained from studies on RNase P from *Escherichia coli* (*E. coli*) and *Bacillus subtilis* (*B. subtilis*). In the following text, RNase P RNA and Protein subunits are abbreviated RPR and RPP, respectively.

1.1.1 Bacterial RPR structure

Comparative sequence analysis of phylogenetically diverse bacterial RPRs was used to identify the universally conserved catalytic core in these RPRs and build a secondary structure model of the bacterial ribozyme (Fig. 1.2) [13-15]. Sequence analysis of many bacterial, archaeal, and eukaryal RPRs also revealed that universally conserved nucleotides, that are distal to each other in the primary sequence, are clustered in five conserved regions (CR I to V) in the secondary structure.
Based on the secondary structure, bacterial RPRs are classified into two structural types: A and B. Type A represents the ancestral class and includes most bacterial RPRs (e.g., *E. coli* RPR). However, only a subset of bacterial RPRs, especially those from low G+C gram positive bacteria, belong to type B (e.g., *B. subtilis*). At the secondary structure level, both type A and B RPRs show remarkable similarity in their highly conserved catalytic core consisting of helices (P1-P5 and P15) and their joining regions (J3/4, J5/15, J15/2 and J2/4), all of which play a role in catalysis (refer to Fig. 1.2 for an explanation of the abbreviations used to denote different secondary structure elements). However, type A and B RPRs differ in how various peripheral structural components, that do not directly take part in catalysis and are involved mainly in long-range interactions which stabilize the RPR, are arranged around the central catalytic core. These peripheral structural components are dispensable for catalysis and can be removed from the core structure to generate a minimal structure (Fig. 1.3). The mini-RPR thus generated is catalytically active but requires high concentrations of monovalent and divalent cations to stabilize its catalytic core [16, 17].

Both type A and B RPRs can be divided into two distinct domains that can fold independently: the catalytic (C) and specificity (S) domains [18, 19]. The C domain harbors phylogenetically conserved nucleotides that constitute the active site of the ribozyme. The S domain interacts with the precursor tRNA (ptRNA), specifically with the T-stem and loop and contributes significantly to the affinity of the RPR for ptRNA to generate the enzyme-substrate (RPR-ptRNA) complex [20]. Concomitant interactions between (i) the T-stem and loop of the ptRNA with the RPR’s S domain, and (ii) the
acceptor stem and the CCA sequence at the 3’ terminus (3’-CCA) of the ptRNA with the RPR’s C domain, play a pivotal role in positioning the scissile phosphodiester linkage into the active site of the ribozyme. The C domains of both structural types can be reconstituted with cognate RPP to generate a functional RNP complex [21, 22]. Moreover, although weak, the C domain of bacterial RPR alone is catalytically active under appropriate in vitro conditions (Kikovska and Kirsebom, personal communication) [22]. Taken together these results suggest that the C domain of the bacterial RPR harbors the active site of the ribozyme.

Substrate recognition and catalysis by enzymes depend on their three-dimensional (3-D) structures. Therefore, to derive mechanistic insights into the substrate specificity and catalytic activity of the bacterial RPR, many attempts were made to obtain their 3-D structures. Initially, distance constraints obtained from site-specific photo-crosslinking studies were used in molecular modeling studies to obtain insights into the global architecture of the bacterial RPR [23]. The 3-D structures of S domains from both type A and B RPRs were obtained subsequently. It is interesting to note that the secondary structures of the S domains of both type A and B RPRs are significantly different (Fig. 1.2). However, their 3-D structures reveal that both structures adopt a remarkably similar tertiary fold to generate a cleft that positions at similar locations the invariant nucleotides that are involved in T-stem and loop interaction [24-26].

The tertiary structures of representative type A and B full-length RPRs (i.e., with intact C and S domains) have been solved recently (Fig. 1.4) [27-29]. These structures highlight how the S and C domains are positioned and stabilized relative to each other.
This inter-domain interaction ensures that the ptRNA substrate simultaneously interacts with both domains and positions the scissile phosphodiester linkage at the active site of the enzyme. Highly conserved tertiary structural features that are common to both type A and B RPRs include: (i) the relative orientations of coaxial stacks P1/P4/P5, P2/P3, P8/P9, (ii) helix P15, (iii) the joining region between P3 and P4, and (iv) a unique structural module involving J11/12 and J12/11. Secondary structure elements that are not part of this common structural core play supporting functions that stabilize the overall structure through tertiary interactions and facilitate the intra/inter domain interactions necessary for function in each type. For instance, helix P18 of the C domain in type A RPR interacts with P8 through a tetraloop-tetraloop receptor interaction involving L18 and P8 (Fig. 1.5). This interaction represents one of the three main inter-domain contacts that brings together the C and S domains. Although P18 is absent in type B structures, helix P15.2 present at an analogous location in the type B RPR plays a similar role in establishing the tertiary interaction involving P8. Interestingly, archaeal and eukaryal RPRs lack P18 or its equivalent in their secondary structure; it is possible that the L18-P8 interaction found in bacterial RPRs has been replaced by RNA-protein interactions in archaeal and eukaryal RNase P holoenzymes.

In the absence of a co-crystal structure of the RPR-ptRNA complex, the high-resolution structures of bacterial RPRs alone do not provide information regarding the active site of the ribozyme or the catalytic mechanism. Nevertheless, the importance of these structures cannot be understated as these structures are informative in highlighting the importance of network of interactions between the bases in CR I-CR IV and CR V,
which although distant in the primary sequence come together to provide a unique structural platform necessary for coordinating catalytically important Mg\(^{2+}\) ions. Moreover, the structures also provided a foundation to generate models of the RPR-tRNA complex by docking the tRNA into the type A RPR structure [28]. To avoid bias in identifying the active site, the RPR-tRNA dock was created based on the available biochemical/cross-linking data on the interactions between T-stem loop of the ptRNA and the S domain of the type A RPR [28]. The RPR-tRNA complex thus obtained places the tRNA acceptor stem on the concave surface of the C domain in RPR, parallel to the P1/P4/P5 helical stack. Interestingly, this placement of tRNA also brings the 3’-CCA terminus proximal to L15 for base pairing interactions and the 5’-end close to the universally conserved P4 and the structural module consisting of CR I, CR IV and CR V. A detailed description of how substrate recognition is accomplished by the bacterial RPR is provided below.

1.1.2 Substrate recognition

Biochemical studies using a combination of approaches including cross-linking, chemical protection and modification interference have identified conserved nucleotides in the bacterial RPR that are critical for substrate binding. For example, cross-linking, and modification interference data demonstrated that conserved nucleotides in helices P11, P9 and P10 of the S domain of *E. coli* RPR are proximal to the T-stem and loop of the ptRNA and thus linked to substrate binding [30]. It has been demonstrated that specific 2’ OH groups in the T-stem loop of the ptRNA substrate are recognized by conserved adenosines in P10/P11 of the S domain [18]. Similarly, the involvement of the
universally conserved structural module composed of CR II-CR III (J11/12 and J12/11) in substrate binding has been confirmed by multiple methods such as direct cross-linking of the substrate to the RPR as well as chemical modification and protection of the RPR in the absence and presence of the substrate [18, 20, 31-34]. Interactions between the 3'-CCA of the ptRNA and the nucleotides in the loop region connecting P15-P16 (L15) of the RPR in the C domain have been demonstrated by various kinetic, mutagenesis, cross-linking and chemical protection studies using wild-type and mutant ptRNAs (Fig. 1.6) [30, 34-36]. The importance of functional groups of the nucleotides in J5/15 and J18/2 of the C domain in substrate binding and catalysis as revealed by nucleotide analogue interference mapping (NAIM) together with cross-linking studies indicate that J5/15 and J18/2 are proximal to the cleavage site [37-39]. Short-range cross-linking studies indicate close association of the P4 helix and the acceptor stem of the ptRNA; specifically, the universally conserved bulged uridine in P4 was crosslinked to +5 in the acceptor stem of ptRNA (cleavage takes place between -1 and +1 of the ptRNA) [40]. These available data on substrate recognition were used to generate the 3-D model of the enzyme-substrate complex in which the tRNA is docked on the bacterial RPR. It is interesting to note that the expected distance from the T-stem loop to the cleavage site in the free ptRNA (41 Å) is in good accordance with the distance between their respective contact sites in the S and C domains (46 Å) in the RPR’s crystal structure, attesting to the validity of the ES model.

1.1.3 Mechanism of RNase P catalysis

RNase P is a metalloenzyme. The rate of RPR-catalyzed hydrolysis of the phosphodiester bond is dependent on the Mg$^{2+}$ ion concentration [41]. A hydrated metal ion is believed
to be the source of the hydroxide nucleophile that attacks the scissile phosphorous center. Specific functional groups in the RPR and ptRNA could act as ligands for the hydrated metal ion which takes part in catalysis. Indeed such an expectation has been borne out by phosphorothioate substitutions of the non-bridging oxygen atoms at the cleavage site in the ptRNA or at strategic locations in the RPR. For example, the substitution of non-bridging phosphate oxygens (either at the R_p or S_p positions) of the scissile bond in ptRNA substrate resulted in a 10,000-fold decrease in the rate of cleavage (k_{obs}). Since inhibition of cleavage in the presence of Mg^{2+} due to phosphorothioate substitution at R_p was rescued by addition of a thiophilic metal much as Cd^{2+} in the reaction, a direct co-ordination of a metal ion to the non-bridging oxygen at R_p position of the scissile bond in ptRNA was deduced [42].

Based on the available biochemical and kinetic data, a two-metal ion scheme has been proposed for RNase P catalysis largely by extension from phosphoryl transfer reactions catalyzed by many protein enzymes and group I introns [43]. According to this mechanism, two metal ions are strategically positioned to co-ordinate to the pro-R_p oxygen of the reactive phosphorous center in the active site of the ribozyme (Fig. 1.7). This arrangement helps one of the metal ions to both position the water molecule serving as the nucleophile for an in-line attack at the phosphorous center and assist its deprotonation, while the second metal ion stabilizes the 3' leaving group in the transition state.

Since deprotonation to yield a hydroxide ion nucleophile is an essential step in the RNase P catalyzed phosphodiester bond hydrolysis reaction, the rate of cleavage
increases with increasing hydroxide ion concentration [41, 44, 45]. Therefore, a log-linear relationship between $k_{\text{obs}}$ and assay pH with slope near unity has been used as a criterion to establish that the observed rate ($k_{\text{obs}}$) indeed represents the rate of cleavage.

The catalytic cycle involves (i) rapid association of the ptRNA to the ribozyme, (ii) fast cleavage, and (iii) slow dissociation of the product. Consequently, multiple-turnover reactions are rate limited by the product release and are not informative in determining the rate of cleavage. However, kinetic measurements performed under single-turnover conditions (i.e., $[E] \gg [S]$) using transient kinetics or under altered reaction conditions such as low pH (see below) or substitution of Ca$^{2+}$ for Mg$^{2+}$ help slow down the reaction rate and ensure that cleavage is rate limiting [46].

1.1.4 Role of the protein subunit in bacterial RNase P catalysis

Since the presence of the sole bacterial protein cofactor significantly reduces the requirement for monovalent and divalent cations in the ribozyme-catalyzed ptRNA-processing reaction in vitro, it was originally thought that the protein cofactor acts as an electrostatic shield to overcome the charge repulsion expected between an RNA enzyme and its RNA substrate [10]. However, the tertiary structure of the bacterial RPP (determined both by NMR and X-ray crystallography) together with results from various kinetic, cross-linking and footprinting studies of the bacterial RNase P holoenzyme has led to a revision of this simplistic notion [47-51]. It has now been established that specific non-covalent interactions between the 5' leader sequence of the ptRNA and the protein cofactor probably accounts for the increased substrate affinity observed with the holoenzyme (i.e., catalytic RNA + protein cofactor [49]). A recent study suggests that the
RPP is responsible for normalizing the binding affinity and cleavage rate of the *E. coli* RNase P holoenzyme for all ptRNAs regardless of whether they possess consensus recognition sequences [52]. Interactions (as yet undefined) between the catalytic site in the RNA enzyme and the protein subunit likely contribute to the observed enhanced affinity for Mg$^{2+}$ and a higher rate of chemical cleavage with the holoenzyme [50, 53, 54]. A recent study demonstrated that binding of the RPP to its cognate RPR stabilizes the RPR by inducing local conformational changes at similar locations in both type A and B RPRs [55]. While computer-aided 3-D models of the *E. coli* RNase P holoenzyme are available [51, 56, 57], a high-resolution structure of the RNase P holoenzyme is needed to better understand the specific structural rearrangements in the RPR induced upon binding to RPP.

### 1.2 Archaeal RNase P

Archaea represent an evolutionarily distinct class of organisms more closely related to eukarya than to bacteria (Woese, et al., 1990). Archaeal RNase P holoenzymes have been partially purified from *Haloferax volcanii (Hvo)*, *Sulfolobus acidocaldarius (Sac)*, *Methanocaldococcus jannaschii (Mja)*, *Methanothermobacter thermautotrophicus (Mth)* and *Pyrococcus horikoshii (Pho)* [58-62]. Although initial studies confirmed the RNP nature of archaeal RNase P holoenzymes, disparities in their protein composition were evident based on their buoyant densities in CsCl or Cs$_2$SO$_4$ gradients. For example, while the buoyant density of *Hvo* RNase P holoenzyme was reported to be similar to that of bacterial enzyme suggesting that the enzyme is made up of mostly RNA, the buoyant
density of *Sac* RNase P indicated an enzyme with increased protein content akin to the eukaryal RNase P holoenzyme [60, 62].

On the basis of phylogenetic-comparative sequence analyses, secondary structure models for many archaeal RPRs have been proposed [15, 63]. Conserved structural elements (CR 1-CR V) similar to those found in the bacterial RPR have been identified in all archaeal counterparts, suggesting a common evolutionary origin for these RPRs. Based on the occurrence of secondary structure elements, most of the archaeal RPRs can be categorized into two well-defined structural classes: type A [represented by *Mth* RPR (Fig. 1.8A)] and type M [represented by *Mja* RPR (Fig. 1.8B)]. Archaeal type A RPR shows remarkable similarities to the bacterial type A RPR (especially to the catalytic domain) albeit lacking P13, P14 and P18 helices; the absence of these helices eliminates the possibility of tertiary interactions similar to those used for the stabilization of the bacterial RPR’s tertiary structure [28] (Figs. 1.2 and 1.5). In addition to P13, P14, and P18, the type M RPR also lacks the structural elements implicated in substrate binding in bacterial RPR especially P8 and L15 [20, 30, 64, 65]. Consistent with this observation, some type A archaeal RPRs (and not type M) are catalytically active in the absence of their protein subunits and in the presence of 4 M NH₄OAc and 300 mM MgCl₂ [66]. The high ionic strength used in these assays might have helped to stabilize the tertiary fold necessary for substrate binding and catalysis for at least a subset of type A RPRs. Since both type A and M archaeal RPRs harbor the highly conserved structural elements (CR I-CR IV) that constitute the catalytic core of the RPR, it is possible that the type M RPR is
the catalytic subunit but requires protein subunits for structural stabilization or substrate binding.

1.2.1 Archaeal RNase P holoenzyme is made up of multiple RPPs that are homologous to eukaryal RPPs

A chimeric holoenzyme consisting of the Mth (archaeal) RPR and the protein subunit of B. subtilis (bacterial) RNase P is weakly active [66]. This finding along with the similarity of the secondary structure of archaeal RPR and bacterial RPR might lead to the expectation that the protein subunit in the archaeal RNase P holoenzyme should be homologous to the protein subunit of bacterial RNase P. However, the sequences of various archaeal genomes have unambiguously demonstrated the presence of ORFs encoding homologs to the proteins associated with eukaryal RNase P and not bacterial RNase P. By using polyclonal antisera generated against four polypeptides of Mth (MTH 11, MTH 687, MTH 688 and MTH 1618, which were identified by database mining as homologs of yeast/human RNase P protein subunits RPP29, POP5, RPP30 and RPP21, respectively), Hall et al. (2002) immunoprecipitated ptRNA-processing RNase P activity from a partially purified Mth RNase P preparation [67]. Moreover, in vitro reconstitution of archaeal RNase P has been demonstrated using recombinant RPR and these four RPPs [58, 68, 69].

Tertiary structures of all four archaeal RPPs have been determined [68, 70-75]. Of relevance to this dissertation is the structure of POP5. The tertiary structure of POP5 reveals four antiparallel $\beta$-strands packed against $\alpha$-helices to form a compact $\alpha\beta$-sandwich fold of RNA recognition motif (RRM). Interestingly, the overall topology of
POP5 is strikingly similar to the bacterial RPP despite different secondary structure connectivities in these RPPs (Fig. 1.9). Structural homology between POP5 and bacterial RPP suggests similar functions for these protein subunits in their respective RNase P holoenzymes. Structural characterizations of these RPPs together with mutational analyses in reconstitution assays have been helpful in identifying regions of archaeal RPPs important for function [70, 72, 73].

A fifth protein subunit (homolog of human RPP38) has been found to increase the temperature optimum for the in vitro reconstituted Pho RNase P [76]. However, biochemical evidence that RPP38 is part of native Pho RNase P holoenzyme is lacking.

1.2.2 Protein-protein interactions in the archaeal RNase P holoenzyme

Yeast two-hybrid (Y2H) analyses have identified potential protein-protein interactions in two archaeal RNase P holoenzymes [77, 78]. Strong interactions involving POP5-RPP30 and RPP21-RPP29 are common to both Mth and Pho RNase P holoenzymes. Interactions identified by Y2H analysis were corroborated further using co-immunoprecipitation assay [78] and partial reconstitution of archaeal RNase P activity using recombinant RPR and pair-wise addition of RPPs. Among the six two-protein combinations, only POP5-RPP30 and RPP21-RPP29 were catalytically active when combined with the cognate RPR (see below).

1.2.3 In vitro reconstitution of archaeal RNase P activity using recombinant RPR and RPPs

Functional reconstitution of the RNase P activity has been demonstrated for archaeal RNase P holoenzyme using the recombinant RPR and the four RPPs [58, 68, 69].
Moreover, a functional RNase P can be reconstituted from RPR and either of the two interacting pairs of RPPs: POP5-RPP30 or RPP21-RPP29 [69]. Thus biochemical characterization of reconstituted \textit{Pfu} RNase P confirmed pair-wise interactions between protein subunits predicted earlier by Y2H studies [77, 78]. Steady-state kinetic analysis performed using the reconstituted \textit{Pfu} RNase P holoenzyme suggests that while \textit{Pfu} RPR alone is capable of multiple turnover, addition of all four RPPs to \textit{Pfu} RPR results in a decreased requirement for Mg$^{2+}$, a 25-fold increase in its $k_{\text{cat}}$ and a 170-fold decrease in $K_m$ [69]. These results parallel those reported with the sole bacterial RPP and lead to the inference that bacterial and archaeal RPPs might employ similar solutions to facilitate RNA catalysis.

Some inferences on the role of RPPs can be drawn from these kinetic studies. First, the overall catalytic efficiency of the RPR increases with protein complexity. Second, the POP5-RPP30 pair (but not RPP21-RPP29) enhances the $k_{\text{cat}}$ of RPR to that observed with all four RPPs indicating that the POP5-RPP30 pair must play a vital role in cleavage and/or product release. Third, although the POP5-RPP30 and RPP21-RPP29 pairs individually have the ability to decrease the $K_m$ value (~3- or 5-fold), there is a significant synergism (170-fold) when both pairs are present. However, a comprehensive understanding of how these RPPs promote catalysis requires a detailed map of the RNA-protein interactions in archaeal RNase P as well as kinetic studies to dissect their contribution to individual steps in the kinetic scheme.
1.3 Eukaryal RNase P

In eukaryotes, RNase P activity exists in the nucleus, mitochondrion and chloroplast [1]. Unlike bacterial RNase P, RNase P holoenzymes from eukaryal sources are enriched in protein (70% by mass compared to 45% and 10% for archaeal and bacterial RNase P holoenzymes, respectively). Yeast nuclear RNase P is associated with nine protein subunits [3]. At least ten protein subunits co-purify with human RNase P [79]. Although distinct organellar RNase P activity has been isolated and characterized from vertebrates, plants, and parasitic protozoa, the subunit composition of organellar RNase P is currently under debate.

1.3.1 Yeast RNase P

A highly homogenous preparation of nuclear RNase P from *S. cerevisiae* is associated with one RNA (RPR1) and nine protein subunits (designated as POP1, POP3, POP4, POP5, POP6, POP7, POP8, RPR2P and RPP1 [9]. Genetic manipulations of these subunits in yeast have demonstrated that all subunits are essential for cell viability [6, 8, 9]. Phylogenetic covariation analysis of RPR sequences from members of the *Saccharomycetaceae* family together with chemical and enzymatic secondary structure probing of *S. cerevisiae* (*Sc*) RPR have led to a secondary structure model of the yeast RPR (Fig. 1.10A) [80]. Conserved regions CR I - CR V in *Sc* RPR are present at identical positions to the bacterial and archaeal RPRs. Although the conserved four-way junction comprising of helices P7-P10/11 is present at a location in *Sc* RPRs similar to bacterial/archaeal RPRs, the nucleotide sequence is different and deviates from the consensus. Since these conserved nucleotides are implicated in interactions with the
substrate, Sc RPRs are likely to exhibit substrate-binding defects similar to archaeal type M RPRs. In addition, a notable deviation from the bacterial type A RPR structure is evident in helix eP15 of the Sc RPR. Although present at an equivalent location, eP15 lacks key structural features of the bacterial L15 involved in interaction with the substrate and Mg$^{2+}$ coordination. Therefore, it is not surprising that RNA alone activity has not been demonstrated for any of the yeast RPRs (however, see comment below on human RPR).

Binary interactions between yeast RPPs were revealed by Y2H analyses and in some cases further verified using glutathione-based in vitro affinity pull-down assays [81]. Interestingly, POP5-RPP1 (RPP30) and POP4 (RPP29)-RPR2P (RPP21) interactions are conserved in archaeal and eukaryal RNase P. Functional in vitro reconstitution of yeast RNase P from individual RPR and recombinant RPPs has not been reported. Nevertheless, in vivo genetic manipulation of the individual subunits in S. cerevisiae has identified a number of mutations in RPRs and RPPs that affect maturation and assembly of yeast nuclear RNase P [82-84]. For example, mutations in or around CR I, CR V and CR IV, that represent the universally conserved catalytic core of the RPR, resulted in mutant holoenzymes that exhibit a 10-fold reduction in catalytic efficiency (defect in catalysis rather than substrate binding) in a standard ptRNA cleavage assay in vitro. Mutations in CR II / CR III region in S domain, which constitute part of the Mg$^{2+}$-binding sites necessary for substrate binding show increase dependence on Mg$^{2+}$ for growth [85]. Similarly, functional characterization of many conditional mutations in the
highly conserved regions of POP1 has demonstrated defects in ptRNA processing as well as assembly of the RNase P holoenzyme \textit{in vivo} [82].

1.3.2 Human RNase P

Human nuclear RNase P consists of one RNA (H1RNA) and ten associated protein subunits (RPP14, RPP20, RPP21, RPP25, RPP29, RPP30, RPP38, RPP40, POP5 and POP1). Seven of these protein subunits have homologs in yeast; RPP14, RPP25 and RPP40 are unique to human RNase P [79].

Phylogenetic sequence analysis of many eukaryal RPRs revealed the universally conserved catalytic core in human RPR with conserved regions CR-I-CR-V at similar locations to the bacterial RPR (Fig. 1.10B). Human RPR, unlike yeast RPR, has been shown to be weakly active \( \left[ k_{\text{obs}}(\text{pH 6}) \right] = 10^{-5} \text{ min}^{-1} \) and can process ptRNA substrate and a model substrate \textit{in vitro} in the absence of its protein subunits [86].

The individual role(s) of human RPPs is currently unknown. Using gel-shift assays, human RPP21, RPP29 and RPP14 have been shown to bind ptRNA [87, 88]. Although RPP20 possesses weak ATPase activity, its significance remains unclear [89]. Since \textit{in vitro} cleavage of ptRNA using purified human RNase P does not require hydrolysis of ATP, the observed ATPase activity might play a role in rearranging the RPR structure during holoenzyme assembly \textit{in vivo}. RPP14, with its interacting partner OIP2, has been shown to process 3' to 5' exonuclease activity. The function of this activity in the context of holoenzyme is also not clear as highly purified human RNase P does not contain OIP2.
Although reconstitution of human RNase P using RPR and ten RPPs has proven difficult, partial reconstitution involving RPR and at least two RPPs (RPP21-RPP29) has been demonstrated [90]. The reconstituted activity was weak even under assay conditions that used excess enzyme over substrate. It is interesting to note that RPP29 is one of the four protein subunits evolutionarily conserved in archaea and eukarya and therefore likely to perform a similar function in the respective holoenzymes.

1.3.3 Organellar RNaseP

In addition to nuclear RNase P, eukaryotic cells also possess RNase P activity in their organelles such as mitochondrion and chloroplast. Perhaps the best characterized organellar RNase P is that from *S. cerevisiae* mitochondria where the holoenzyme was shown to be composed of mitochondrially-encoded RPR (Rpm1) and a single nuclear-encoded RPP (Rpm2) [91]. Biochemical and genetic analyses have demonstrated that both subunits are essential for cell viability. Database mining of mitochondrial genome sequences of many Ascomycete fungal lineage have identified many putative RPR sequences [92]. These RPRs exhibit large variations in size ranging from 140 (Saccharomyces fibuligera) to 423 nts (S. cerevisiae). Since Rpm2 does not have any significant homology to any of the bacterial or archaeal RPPs and there are large variations in mitochondrial RPR structure, mitochondrial RNase P might have diverged significantly from its bacterial or nuclear counterparts.

There is considerable debate over the subunit composition of RNase P from human mitochondria. RNase P isolated from human mitochondria, which is shown to have preference over ptRNAs of mitochondrial origin, was inferred to be made up of only
protein based on its insensitivity to nuclease treatment and protein-like buoyant density in a Cs$_2$SO$_4$ gradient [93, 94]. In an independent study, Puranam and Attardi (2001) reported that RNase P activity that is partially purified from enriched mitochondrial fraction contains an essential RPR that is identical in sequence to that of nuclear RPR [95].

Available bacterial and mitochondrial genome data support the notion that mitochondrial genome originated from a member of extant $\alpha$-proteobacteria. The mitochondrial genome most closely related to a bacterium was identified in a jakobid protist called *Reclinomonas americana* (*Ram*). Therefore, studying *Ram* mitochondrial RNase P is likely to provide valuable insights into mitochondrial RNase P evolution. While the genes encoding RPP(s) remain largely elusive, putative genes encoding RPRs have been identified in mitochondrial genomes of many jakobids based on their sequence homology to that of bacterial RPRs. Interestingly, *Ram* RPR is not catalytic in vitro, despite having the conserved catalytic core similar to the bacterial ribozyme [96].

Similar to human mitochondrial RNase P, the subunit composition of RNase P from spinach chloroplasts is also a subject of controversy. Based on insensitivity to nuclease treatment and a unique cleavage mechanism exhibited by partially purified RNase P preparation from spinach chloroplasts, it has been proposed that spinach chloroplast RNase P might be devoid of RPR [97-99]. In contrast to RNase P from chloroplast of higher plants, RNase P from the photosynthetic organelle (cyanelle) of primitive algae *Cyanophora paradoxa* has been shown to be an RNP complex based on its susceptibility to micrococcal nuclease and buoyant density measurement. Interestingly,
the cyanelle RPR structurally resembles its phylogenetically related cyanobacterial RPR [100]. However, unlike the cyanobacterial RNase P holoenzyme, the cyanelle enzyme is enriched in protein subunits whose identity is currently unknown [101, 102]. Therefore, the cyanelle RNase P holoenzyme represents an interesting scenario wherein the RPR is structurally similar to the bacterial RPR; however, unlike the bacterial RPR, the cyanelle RPR is associated with multiple RPPs as opposed to a single RPP in the bacterial RNase P holoenzyme.

1.4 Research objectives and approaches

Comparative analysis of primary sequence and secondary structures of many bacterial, archaeal and eukaryal RPRs reveal universally conserved nucleotides and a core structure that are likely to be the part of the active site of all RPRs. However, while all bacterial RPRs examined thus far are ribozymes, only a few representative archaeal and eukaryal RPRs are catalytically active without their cognate RPPs when tested under various conditions \textit{in vitro}. Moreover, the efficiency of ptRNA processing reaction, measured as \( k_{\text{obs}} \) (\( k_{\text{obs}} \) is the pseudo-first-order rate constant of ptRNA cleavage under single-turnover condition) catalyzed by the RPR alone decreases significantly from 10 min\(^{-1}\) for bacterial RPR to 10\(^{-2}\) min\(^{-1}\) and 10\(^{-5}\) min\(^{-1}\) for archaeal and eukaryal RPRs, respectively. This decrease in catalytic efficiencies of the archaeal/eukaryal RPR is accompanied by a concomitant increase in the number of RPPs that these RPRs associate with to form a functional RNase P holoenzyme. Although it is clear that the RPR is vital for RNase P catalysis in all three domains of life, it is unclear why there is an inverse correlation in the
catalytic efficiency of the archaeal/eukaryal RPRs and their number of RPPs. These observations provide the underpinning for the following queries central to this dissertation.

Is the sole RPR the catalytic moiety in all archaeal/eukaryal RNase P? If so, why do archaeal/eukaryal RNase P holoenzymes, which retain an RPR whose catalytic core is similar to the bacterial ribozyme, require multiple RPPs for function while a single RPP suffices for bacterial RNase P? Do archaeal/eukaryal RPPs merely play structural roles by replacing RNA-RNA interactions with RNA-protein or protein-protein interactions to create an RPR-based active site in archaeal/eukaryal RNase P holoenzymes? Or, do these RPPs participate directly in catalysis in archaeal/eukaryal RNase P holoenzymes? The overall objective of the present study is to address these queries using as a model system archaeal RNase P, an RNP complex that contains only four RPPs as opposed to nine in eukaryal RNase P.

To address whether RPR is the catalytic moiety in all archaeal RNase P holoenzyme, we focused on a subset of archaeal RPRs (type M RPRs) for which the RPR-alone activity has not been observed under various conditions tested \textit{in vitro}. We reasoned that the failure to promote catalysis by type M RPRs is likely due to structural defects in these RPRs that impair substrate binding. By constructing an enzyme-substrate conjugate in which the ptRNA substrate is attached to the \textit{Methanocaldococcus jannaschi} (\textit{Mja}) RPR, we demonstrated that the \textit{Mja} RPR is capable of accurately cleaving the leader sequence of the ptRNA even in the absence of cognate RPPs.
To address the specific role of RPPs in facilitating type M RPR catalysis, it is essential to reconstitute type M RNase P from individual RPR and RPPs. Towards this goal, we first optimized conditions for robust in vitro reconstitution of *Mja* RNase P using in vitro transcribed *Mja* RPR and recombinant RPPs (purified as binary complexes; either POP5-RPP30 or RPP21-RPP29). The effect of each binary complex on the rate of cleavage was then addressed by reconstituting ptRNA-*Mja* RPR conjugate with RPP21-RPP29 or POP5-RPP30. By comparing the RPR alone reaction rate to that of the protein-assisted reaction, we ascertained the role of individual RPPs in RPR-mediated catalysis.

To understand the role of multiple RPPs in archaeal/eukaryal RNase P, knowledge of the spatial organization of RPR and RPPs in the RNase P holoenzyme is necessary. Therefore, one of the main objectives of this study was to make inroads towards establishing the global architecture of the archaeal RNase P holoenzyme. Towards this aim, in vitro reconstituted archaeal RNase P holoenzymes were subjected to enzymatic footprinting using RNase T1. Results from such footprinting experiments show significant protection of the universally conserved nucleotides in the RPR in the presence of POP5-RPP30. Since these conserved nucleotides in the bacterial RPR have been shown to participate in Mg$^{2+}$ binding and interact with nucleotides at or around the cleavage site in the substrate, POP5-RPP30 is likely to be proximal to the cleavage site.

To delineate the contributions of various evolutionarily conserved structural elements of the archaeal RPR to substrate binding and catalysis, we constructed various deletion derivatives of archaeal RPR and assessed their ability to reconstitute with cognate RPPs. One such deletion derivative of archaeal RPR, bereft of the putative S
domain and retaining only the C domain, is able to functionally assemble with POP5-RPP30. Since bacterial and archaeal RPRs share a similar catalytic domain, we also examined whether archaeal POP5-RPP30 would heterologously reconstitute with bacterial RPRs. Such heterologous reconstitutions of functional RNase P holoenzymes using various evolutionarily related bacterial/archaeal/eukaryal (organellar) RPRs and archaeal RPPs reveal the presence of a conserved catalytic core in all three domains of life.
Figure 1.1 The role of RNase P in tRNA maturation. RNase P-mediated cleavage results in removal of the 5'-leader sequence from precursor tRNAs. The arrowhead indicates the site of RNase P cleavage in *E. coli* ptRNA<sup>Tyr</sup>. The leader sequence of ptRNA<sup>Tyr</sup> is indicated using outlined letters.
Figure 1.2 Secondary structure representation of bacterial RPRs. Panels A and B represent type A (E. coli) and type B (B. subtilis) bacterial RPRs, respectively. Both type A and B RPRs are composed of two independently folding domains: specificity domain (colored blue) and catalytic domain (colored black). Paired regions (P) are numbered sequentially from the 5' end. The nucleotide linker region joining helices P11 and P12 is denoted as J11/12. Universally conserved nucleotides are highlighted. CR I - V represent conserved regions. Figure reproduced from [4].
Figure 1.3 Schematic representation of a phylogenetically conserved minimum bacterial RPR consensus. Only those structural elements that are common to all bacterial RPRs are shown. Figure reproduced from [16].
Figure 1.4 Tertiary structures of the bacterial type A and B RPRs. Panel A depicts the tertiary structure of a type A RPR (Thermotoga maritima; T. maritima) while panel B illustrates the tertiary structure of type B RPR (Bacillus stearothermophilus; B. stearothermophilus). Similar secondary structures in both RPRs are represented using the same colors [28, 29].
Figure 1.5 Inter-domain interactions in the bacterial RPR. Panel A shows the tertiary interactions between L14 and L18 with helix P8 observed in the crystal structure of *T. maritima* RPR. Panel B is a representation of the secondary structure of bacterial RPR in which these tertiary interactions are depicted using dotted lines.
Figure 1.6 Schematic summarizing the specific contacts between the bacterial RPR and the ptRNA substrate. Nucleotides in the S domain interact with the T-stem loop of the ptRNA (blue circle). A specific interaction between an adenosine in the C domain and -1U in the ptRNA is indicated using a dotted line. Loop 15 (L15) of the bacterial RPR interacts with 3' CCA of the ptRNA (green circle) [103].
Figure 1.7 The role of magnesium ions in the catalytic mechanism of RNase P. The main features of the general two-metal ion mechanism proposed for protein- and RNA-catalyzed hydrolytic and phosphoryl transfer reactions are included. Figure reproduced from [42].
Figure 1.8 Similarity and variations in archaeal RPRs. Panels A and B represent secondary structure models of type A (Pfu) and M (Mja) archaeal RPRs, respectively. Secondary structures were derived by phylogenetic comparative sequence analysis [63]. Universally conserved nucleotides in all RPRs are highlighted. S domains (putative) in both types are colored in blue [1, 13].
Figure 1.9 Similarities in structures of bacterial RPP and archaeal POP5. Panels A and B represent high resolution structures of \textit{B. subtilis} RPP [48] and \textit{Pyrococcus furiosus} POP5 [74]. α-Helical regions are shown in red, β-strands in blue and loop regions in grey.
Figure 1.10 Secondary structure representation of eukaryal RPRs. Panels A and B represent secondary structure models of yeast and human RPRs, respectively. Universally conserved nucleotides in RPRs from all three domains are highlighted [1].
CHAPTER 2

RECONSTITUTION OF METHANOCALDOCOCCUS JANNASCHII (Mja) RNASE P: INSIGHTS INTO THE ROLE OF PROTEIN COFACTORS IN ARCHAEOAL RNASE P CATALYSIS

2.1 Introduction

Ribonuclease P (RNase P), an endoribonuclease responsible for 5'-maturation of tRNAs, is a ribonucleoprotein (RNP) complex containing an essential RNA (RNase P RNA; RPR) and one or more protein subunits (RNase P Protein; RPP) in all three domains of life: at least one, four, and nine RPPs in bacteria, archaea, and eukarya, respectively[1, 2, 4, 5, 67, 87]. The bacterial RPR is catalytically active in the absence of its RPP under in vitro conditions that include Mg$^{2+}$ (or another suitable divalent metal ion), which is essential for catalysis [10]. However, only a few representative archaeal and eukaryal RPRs are weakly active under various conditions tested in vitro [66, 86]. This result is surprising given the remarkable structural similarity of the putative catalytic core in all RPRs (revealed by comparative sequence analysis) and raises the possibility that not all RPRs are functionally equivalent (Fig. 2.1, [13]). Although unlikely, the
prospect remains that the archaeal/eukaryal RPRs do not participate directly in catalysis but act as a structural scaffold to create a protein-based active site in most archaeal/eukaryal RNase P holoenzymes. We investigate here a euryarchaeal RPR that was reported to be inactive and examine the basis for its inactivity as well as the role of its cognate RPPs in catalysis.

Based on the phylogenetically conserved structural elements of the RPR, euryarchaeal RNase P is classified into types A and M [63]. At the secondary structure level, type A RPRs resemble the ancestral bacterial RPRs while type M RPRs lack some of the structural elements implicated in substrate binding in bacterial RPRs (Fig. 1.6 and compare panel A to B in Fig. 1.8). Consistent with this observation, only some type A RPRs (and none of the type M RPRs) are catalytically active \textit{in vitro} in the absence of their RPPs [66]. In this study, we employ \textit{Methanocaldococcus jannaschii} (Mja) RPR as a prototype for type M RPRs. Using a unimolecular enzyme-substrate conjugate, expected to ameliorate substrate-binding defects in Mja RPR, we establish that Mja RPR (in the absence of RPPs) supports cleavage in cis, thus reaffirming the RPR’s pivotal catalytic role.

If the RPR is vital for RNase P catalysis in all three domains of life, what are the roles played by the respective RPPs? Why do archaeal/eukaryal RNase P holoenzymes require multiple RPPs for function \textit{in vivo} while bacterial RNase P employs at least one RPP to normalize the binding affinity and rate of cleavage of different ptRNAs by the RPR [50, 52, 54]? Biochemical characterization of a partially purified archaeal RNase P holoenzyme demonstrated that the RPR is associated with at least four RPPs (POP5,
RPP30, RPP21 and RPP29) which have eukaryal homologs [67, 87]. We recently reconstituted *Pyrococcus furiosus* (*Pfu*; type A) RNase P *in vitro* in an effort to develop archaeal RNase P as an experimental surrogate for the more complex eukaryal counterparts [69]. In this study, by examining the role of RPPs in aiding the self-processing of the type M RPR-ptRNA cis conjugate, we have uncovered functional parallels between the sole bacterial RPP and one pair of archaeal RPPs (POP5-RPP30) and also suggest how some bacterial RPR functions missing in archaeal RPR might have been substituted by the remaining pair (RPP21-RPP29) of archaeal RPPs.

2.2 Materials and methods

2.2.1 Cloning the genes encoding the archaeal type M (*Afu, Mja* and *Mma*) RPRs

The genes encoding the archaeal type M (*Afu, Mja* and *Mma*) RPRs were cloned by PCR using the respective genomic DNA as the template and the appropriate gene-specific primers (*AfuRPR*-F and *AfuRPR*-R; *MjaRPR*-F and *Mja-RPR*-R; *MmaRPR*-F and *MmaRPR*-R; Table 1). The PCR products containing the RPR genes were digested with restriction enzymes whose recognition site was included in the reverse primer (*HindIII* for *Afu* RPR and *BamHI* for both *Mja* and *Mma* RPRs) and then subcloned into pBT7 [104] that had been digested with *StuI* (which generates a blunt end) and either *HindIII* or *BamHI*. The resulting plasmids were named pBT7-*Afu* RPR, pBT7-*Mja* RPR, and pBT7-*Mma* RPR, respectively, and have the various RPR genes placed under the control of a T7 RNA polymerase promoter. The *Afu, Mja* and *Mma* RPRs were generated using
HindIII-linearized pBT7-Afu RPR, BsmAI-linearized pBT7-Mja RPR or BsmAI-linearized pBT7-Mma RPR as the template DNAs in separate in vitro transcription reactions.

Afu genomic DNA was a kind gift from Dr. F. Robert Tabita, Department of Microbiology, Ohio State University, Columbus, OH. Mja and Mma genomic DNAs were generously provided by Dr. Biswarup Mukhopadhyay, Virginia Bioinformatics Institute, Blacksburg, VA.

2.2.2 Construction of genes encoding ptRNA\textsuperscript{Ty}r-Mja RPR conjugates with spacers of different length

The gene encoding ptRNA\textsuperscript{Ty}r-S\textsubscript{10}-Mja RPR in which \textit{E. coli} ptRNA\textsuperscript{Ty}r is conjugated to the L15 of the Mja RPR with a 10-nt spacer was obtained by overlap-extension PCR. First, genes encoding the \textit{E. coli} ptRNA\textsuperscript{Ty}r, 5', and 3' fragments of Mja RPRs were obtained in three separate PCRs using the following primer sets: (i) MJATYR-F1 and MJATYR-R1, (ii) MJATYR-F2 and MJATYR-R2, and (iii) MJATYR-F3 and MJATYR-R3 (listed in Table 1); pUC19-ptRNA\textsuperscript{Ty}r [105] and pBT7-Mja RPR served as templates depending on whether the ptRNA or RPR was being amplified. Second, PCR products thus generated were combined and reamplified. Third, the extended product was subjected to another round of PCR using primers MJATYR-F1 and MJATYR-R3 and the PCR product thus obtained was digested with \textit{Bam} HI and ligated to pBT7 that had been digested with \textit{Stu} I and \textit{Bam} HI. The resulting plasmid was named as pBT7-ptRNA\textsuperscript{Ty}r-S\textsubscript{10}-Mja RPR.

To alter the spacer length, pBT7-ptRNA\textsuperscript{Ty}r-S\textsubscript{10}-Mja RPR was used as the template for PCR-based deletion mutagenesis. The primer pairs MjapTyrSp-F/MjapTyr0Sp-R and
MjapTyrSp-F/MjapTyr5Sp-R were used to construct pBT7-ptRNA\(^{Tyr}\)-S\(_0\)-Mja RPR and pBT7-ptRNA\(^{Tyr}\)-S\(_5\)-Mja RPR, respectively. Similarly, the primer pairs MjapTyrSp-F/MjapTyr3Sp-R and MjapTyrSp-F/MjapTyr4Sp-R were used to construct pBT7-ptRNA\(^{Tyr}\)-S\(_3\)-Mja RPR and pBT7-ptRNA\(^{Tyr}\)-S\(_4\)-Mja RPR, respectively.

The ptRNA\(^{Tyr}\)-Mja RPR RNA conjugates with spacers of 0, 3, 4, 5 and 10 nts were generated using BsmAI-linearized pBT7-ptRNA\(^{Tyr}\)-S\(_0\)-Mja RPR, pBT7-ptRNA\(^{Tyr}\)-S\(_3\)-Mja RPR, pBT7-ptRNA\(^{Tyr}\)-S\(_4\)-Mja RPR, ptRNA\(^{Tyr}\)-S\(_5\)-Mja RPR and pBT7-ptRNA\(^{Tyr}\)-S\(_{10}\)-Mja RPR, respectively, as the template DNAs in separate \textit{in vitro} transcription reactions.

\textbf{2.2.3 Construction of the gene encoding the EcoS-MjaC RPR chimera}

A hybrid RPR in which the S domain of the \textit{E. coli} RPR was fused with C domain of Mja RPR was obtained using overlap-extension PCR. The S domain of the \textit{E. coli} RPR was amplified by PCR using pJA2' [105] as the template and MJCECS-F and MJCES-R as primers. The 5'- or 3'-halves of the C domain of Mja RPR were amplified by PCRs using MJARPR-F and ECSMJC-R or ECSMJC-F and MJARPR-R, respectively, as the primers and pBT7-Mja RPR as the template. The PCR products thus obtained were mixed and extended in a thermal cycler using dNTPs and Taq DNA polymerase. The extended product was used as a template for the next round of PCR using primers MJARPR-F and MJARPR-R. The resulting PCR product was digested with \textit{EcoRI} and ligated to pBT7 that was digested with \textit{EcoRI} and \textit{StuI}. The resulting plasmid was named as pBT7-\textit{EcoS}\textit{MjaC} RPR and linearized with \textit{EcoRI} to generate the template DNA for \textit{in vitro} transcription to generate the \textit{EcoS-MjaC} hybrid RPR.
2.2.4 Cloning the genes encoding the *Mja* RPPs

The genes encoding *Mja* POP5, RPP21, RPP29 and RPP30 were amplified by PCR with *Mja* genomic DNA as the template and the respective gene-specific primers listed in Table 2; the names of the primers are self-explanatory. All PCR products were digested with *Bam*HI, whose recognition site was included in the various reverse primers. The *Mja* POP5 and RPP29-encoding ORFs were cloned individually into pLANT2b (Finkelstein et al., 2003) that had been (i) digested with *Nde*I, (ii) filled in with Klenow, and (iii) digested with *Bam*HI. The *Mja* RPP21 and RPP30-encoding ORFs were cloned into pET15b that had been (i) digested with *Nco*I, (ii) filled in with Klenow, and (iii) digested with *Bam*HI. These cloning approaches ensured that no additional amino acid residues were introduced in the four ORFs. The gene ID, molecular weight, and isoelectric point for each archaeal RPP used in this investigation are listed in Table 3.

2.2.5 Construction of clones with two archaeal RPP ORFs in tandem

Our overall strategy was to clone two archaeal RPP genes, each encoding one member of a binary RPP complex, into an expression vector such that each gene was under the control of its own T7 RNA polymerase promoter and control elements as originally conceived in the pET system; all these constructs have only one T7 terminator at the end of the tandem construct.

*MjaRPPTC1*: Digestion of pET15B-*Mja*RPP30 with *Bgl*II and *Bfl*I released the *Mja*RPP30 ORF together with its upstream T7 RNA polymerase promoter. This insert was then ligated to pLANT2b-*Mja*POP5 that had been digested with *Bam*HI and *Bfl*I to
generate pLANT2b-MjaRPPTC1, a construct in which *Mja* POP5 and *Mja* RPP30 ORFs are present in tandem.

**MjaRPPTC2**: Digestion of pLANT2b-MjaRPP29 with *Bgl*II and *Eco*RI released the *Mja*RPP29 ORF together with its upstream T7 RNA polymerase promoter and downstream T7 terminator as well as the RIL tRNA gene cluster. This insert was ligated to pET15B-MjaRPP21 that had been digested with *Bgl*II and *Eco*RI to generate pET15b-MjaRPPTC2, a construct in which *Mja* RPP21 and *Mja* RPP29 ORFs are present in tandem.

**Note**: The sequences of all the clones described above were confirmed by automated DNA sequencing at the OSU Plant-Microbe Genomics Facility. All the constructs used for protein expression and purification were made by Venkat Gopalan (Table 8).

### 2.2.6 Purification of archaeal RPPs

For co-overexpression of archaeal RPPs in *E. coli* BL21(DE3) cells, either individual ORFs subcloned in two compatible expression plasmids (i.e., pLANT2b and pET15b), or, in later stages of this study, tandem constructs that enabled expression of two ORFs from a single vector were used. The results were quite similar in both cases.

*E. coli* BL21(DE3) cells freshly transformed with pLANT2b-MjaRPPTC1 or pET15b-MjaRPPTC2 were inoculated into 5 mL Luria Broth (LB) media containing 35 μg/mL kanamycin (*Mja* POP5-RPP30) or 100 μg/mL carbenicillin (*Mja* RPP21-RPP29), respectively, and grown overnight at 37°C with shaking. These overnight cultures were used to inoculate 500 ml of fresh LB media containing the appropriate antibiotics. The cells were grown at 37°C with shaking until the cells reached an optical density at 600
Protein overexpression was induced by adding isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. To the RPP21-RPP29-expressing cultures, ZnCl₂ was also added to a final concentration of 1 mM at the time of addition of IPTG. The cells were then allowed to grow at 37°C for another 3 h (Mja RPP21-RPP29) or room temperature for 15 h (Mja POP5-RPP30) and harvested by centrifugation. The cell pellets were stored at -80°C until further use.

Frozen cells were thawed on ice and re-suspended in 20 mL of buffer A [25 mM Tris-HCl (pH 7.5), 5 mM DTT, 0.1 mM PMSF, 25 mM NaCl]. Five mL of buffer S [25 mM Tris-HCl (pH 7.5), 5 mM DTT, 0.1 mM PMSF and 5 M NaCl] was added to this re-suspension and the cells were lysed by sonication. Cell debris was pelleted by centrifugation (9,000 g, 15 min, 4°C). Polyethylenimine (PEI) was added to the cleared supernatant to a final concentration of 0.05% (v/v) and incubated on ice for 30 min. The PEI-precipitated nucleic acid was pelleted by centrifugation (9,000 g, 15 min, 4°C). To the resulting soluble fraction, finely powdered (NH₄)₂SO₄ (to a final saturation of 40%) was added slowly. The precipitation was performed over a period of 60 min with constant stirring on ice. The precipitated RPPs were recovered by centrifugation (9,000 g, 15 min, 4°C). The precipitate was dissolved 3 ml of buffer RS (1 M ammonium sulfate, 25 mM Tris-HCl (pH 7.5), 5mM DTT, 0.1mM PMSF, 25 mM NaCl) and diluted to 10 ml by adding buffer A. The solution was filtered through a 0.4 μm filter and loaded on a pre-equilibrated SP-Sepharose (GE Healthcare) column. The bound proteins were eluted with a 0 to 2 M NaCl gradient in buffer A. Fractions containing the archaeal RPPs were identified by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE)
followed by Coomassie Blue staining. Binary complexes typically eluted between 1.2 and 1.5 M NaCl. Using the extinction coefficients of the binary RPP complexes and their absorption at 280 nm, the concentration of each binary complex was calculated. The purified RPP complexes were dialyzed against buffer containing 50 mM Tris-acetate (pH 8), 0.8 M NH$_4$OAc, and either 10 mM Mg(OAc)$_2$ or 10 mM Ca(OAc)$_2$ and stored at room temperature until further use.

2.2.7 RNase P activity assays

For the cis cleavage reactions, the pt$_{\text{Tyr}}$-S$_x$-M RPR conjugates with spacers of different lengths were obtained by in vitro transcription of appropriate DNA templates. Subsequently, they were 5'-labeled with $\gamma$-[$^{32}$P]-ATP and T4 polynucleotide kinase and gel purified. These RNAs were then folded by first denaturing them at 50°C for 50 min in water followed by refolding at 37°C for 30 min in either buffer RF-2.5 [50 mM Tris-Acetate (pH 8), 100 mM Mg(OAc)$_2$, 2.5 M NH$_4$OAc] or RF-0.8 [50 mM Tris-Acetate (pH 8), 10 mM Mg(OAc)$_2$, 0.8 M NH$_4$OAc]. The self-cleavage reaction was initiated by addition of Mg(OAc)$_2$ to a final concentration of 500 mM. After incubation at 55°C for a defined time, the reactions were terminated by adding stop solution [10 M urea, 5 mM EDTA, 10% (v/v) phenol] and their contents subjected to ethanol precipitation. The pellets were resuspended in 5 μL of loading dye [10 M urea, 1 mM EDTA, 0.05% (w/v) xylene cyanol, 10% (v/v) phenol] prior to subjecting the contents to denaturing PAGE.

For the trans cleavage reaction involving the E. coli-Mja hybrid RPR (EcoS-MjaC), 100 nM RPR was incubated with 500 nM E. coli ptRNA$_{\text{Tyr}}$ (a portion of which was 5'-labeled). The reaction was performed at 50°C in 50 mM Tris-acetate (pH 8), 500
mM Mg(OAc)$_2$ and 2 M NH$_4$OAc for 1 h. The reactions were terminated by adding stop solution and and their contents subjected to ethanol precipitation. Cleaved products were visualized by autoradiography/PhosphorImaging.

When the pt$^{\text{Tyr}}$-S$_3$-M RPR conjugate was incubated in the presence of RPPs, the reactions were performed at pH 5.1 using two different pre-incubation strategies. In the first approach, 50 nM pt$^{\text{Tyr}}$-S$_3$-M RPR conjugate was mixed with 250 nM binary RPP complexes (either individually or together) in 50 mM MES (pH 5.1 at 55°C), 800 mM NH$_4$OAc, and 25 mM Ca(OAc)$_2$ and incubated at 37°C for 10 min and then at 55°C for 10 min. The reaction was initiated by adding equal volumes of 50 mM MES (pH 5.1 at 55°C), 800 mM NH$_4$OAc, and 200 mM Mg(OAc)$_2$ (final concentration, 100 mM). After incubation for a specified time at 55°C, the reactions were terminated by rapidly immersing the tube in liquid nitrogen. Equal volume of loading dye was added to the tube prior to subjecting its contents to denaturing PAGE. In the second approach, either 1 mM Mg(OAc)$_2$ (for POP5-RPP30) or 10 mM Mg(OAc)$_2$ (for RPP21-RPP29) was used in the pre-incubation step as these divalent concentrations permitted assembly without promoting significant cleavage. The reactions were initiated by adding 100 mM Mg(OAc)$_2$ and terminated as described above.

The dependence of rate on assay pH was established by obtaining rate of product formation at various pH values (at 55°C) ranging from 5.1 - 6.3 for the pt$^{\text{Tyr}}$-S$_3$-M RPR reaction and 4.65 - 5.1 for the pt$^{\text{Tyr}}$-S$_3$-M RPR + POP5-RPP30 complex using either 50 mM acetate (pH <5) or MES (pH>5) buffers. For each pH value, the experiments were repeated at least three times and the mean and standard deviation values are reported.
**Kinetic analysis:** For both cis and trans cleavage reactions (in the absence and presence of RPPs), the substrate and product RNAs were separated using either 8% or 10% (w/v) polyacrylamide/7 M urea gels and visualized by either autoradiography or PhosphorImaging (Molecular Dynamics). The extent of cleavage was quantitated using ImageQuant (Molecular Dynamics) software. The first-order rate constant for the self-cleavage reaction was calculated from a time-course analysis, i.e., determining the extent of processing as a function of time of incubation. The % product formed (P_t) at time (t) was fit using Kaleidagraph to P_t = P_\infty (1-e^{-kt}). If there was background cleavage (i.e., product formed at time t_0), then a correction was applied: P_t = P_0 + (P_\infty - P_0)(1-e^{-kt}) where P_0 refers to product formed at t_0.

*Mja* RNase P reconstitution experiments were initiated by pre-incubating folded *Mja* RPR (50 nM) with either all or a subset of four *Mja* RPPs (250 nM) in assay buffer [50 mM Tris·HCl (pH 7.5), varying amounts of NH_4OAc and MgCl_2] for 5 min at 37°C followed by 10 min at 55°C. The activity was assayed by adding a known concentration of *E. coli* ptRNA^{Tyr}, a trace amount of which was internally labeled with [\alpha-^{32}P]GTP. After a defined incubation time at 55°C, the reactions were quenched with urea-phenol dye [8 M urea, 0.04% (w/v) bromophenol blue, xylene cyanol, 0.8 mM EDTA, 20% (v/v) phenol] and analysed as described [69].
2.3 Results

2.3.1 Archaeal type M RPR supports ptRNA processing in the absence of its cognate RPPs

The recent demonstration that human RPR alone is weakly catalytic ($k_{\text{obs}} \approx 10^{-5} \text{ min}^{-1}$; $k_{\text{obs}}$ is the pseudo first-order rate constant of ptRNA cleavage under single-turnover conditions) under prolonged incubation at pH 6 in the presence of 0.2 M Mg$^{2+}$ and 0.8 M NH$_4^+$ prompted us to examine if the type M euryarchaeal RPRs would be active under similar conditions [86]. However, this was not the case when we examined *Archaeoglobus fulgidus* (*Afu*) and *Mja* RPRs (both type M; data not shown). No ptRNA processing activity was evident even under a variety of other assay conditions tested including different concentrations of monovalent and divalent ions (data not shown). We reasoned that if the inability of *Afu* and *Mja* RPR to promote catalysis is due to the absence of specific RNA structural elements (e.g., P8, L15; Fig. 2.1) that have been implicated in ptRNA binding in the bacterial RPR [Fig. 1.6] [18, 20, 34, 35, 106], alleviation of substrate-binding defects might render the type M RPR active.

Towards this objective, we constructed a bacterial-archaeal (type M) hybrid RPR with structural components from each RPR. The rationale for this experiment was based on three observations. First, biochemical studies delineated the presence of two independently folding modules in bacterial RPRs: a substrate-specificity (S) domain that has conserved nucleotides for recognition of the T stem-loop of the ptRNA; and a catalytic (C) domain which can (i) recognize the leader, acceptor stem, and the 3’ terminal CCA sequence, and (ii) cleave the ptRNA [18-20, 34, 35, 103, 106]. Second, we recently
demonstrated that in *Pfu* (type A) RPR, the C domain was capable of cleaving ptRNA in the presence of two of its cognate RPPs [69]. Lastly, mitochondrion-encoded RPRs in jakobid flagellates (e.g., *Reclinomonas americana*) became active upon fusion of their C domains with the S domain of *E. coli* (*Eco*) RPR [96]. Therefore, guided both by the high-resolution structures of bacterial RPRs [28, 29] and prior biochemical studies in which similar hybrids were generated [96], we constructed a hybrid RPR in which the S domain of *Eco* RPR is fused to the putative C domain of *Mja* RPR (*Eco*-*Mja* C RPR; Fig. 2.2A). Indeed, *Eco*-*Mja* C RPR accurately processed 5′-labeled ptRNA<sub>Tyr</sub> at 50°C in 50 mM Tris-acetate (pH 8), 500 mM Mg(OAc)<sub>2</sub> and 2 M NH<sub>4</sub>OAc (Fig. 2.2B). Therefore, *Mja* RPR supports trans cleavage of ptRNA<sub>Tyr</sub> if its S domain, which is probably defective in substrate binding, is replaced with the bacterial S domain that does not suffer from this limitation.

Since the above experiment revealed that the *Mja* RPR’s inability to promote catalysis in the absence of its cognate RPPs is due to defects in ptRNA recognition/binding, we inquired if an enzyme-substrate (ES) covalent conjugate in which a ptRNA is covalently tethered to the *Mja* RPR would facilitate its processing. Surprisingly, we did not observe any cis cleavage when we tested conjugates in which ptRNA<sub>Tyr</sub> was attached to the 3′ end of either *Afu* or *Mja* RPR [with spacers of either 50 nts (*Afu*) or 25 nts (*Mja*) to the cleavage site; data not shown].

As there was precedence for bacterial and archaeal type A RPRs supporting cis cleavage of a ptRNA<sub>Tyr</sub> attached to the L15 loop [66, 107], we made similar constructs with *Mja* RPR especially since L15 is part of the active site in bacterial RPR [28, 29, 34,
When we tested such conjugates with spacers of 0, 5, and 10 nts, only the 5- and 10-nt-spacer ES conjugates were cleaved. The accuracy of cleavage in ptRNA Tyr-S5-Mjä RPR (abbreviated as pt Tyr-S5-M RPR; Fig. 2.3A) was confirmed by comparing the migration of the 5' leader with that of size markers generated by partial digestion of 5'-labeled pt Tyr-S5-M RPR with RNase T1 (Fig. 2.3B). Since the 10-nt-spacer conjugate also supported some mis-cleavage, possibly due to formation of some secondary structure near the cleavage site, only the 5-nt-spacer conjugate was characterized further. Clearly, the length of the spacer is an important determinant for ensuring efficient cleavage and even influencing the rate-limiting step (see below).

To identify conditions that would permit efficient and accurate self-cleavage of the leader sequence from 5'-labeled pt Tyr-S5-M RPR, we varied different parameters. Either the Mg2+ concentration was held constant at 0.5 M while the NH4+ concentration was increased from 0 to 4 M or the NH4+ concentration was maintained at 2.5 M while the Mg2+ concentration was raised from 0 to 0.5 M (Fig. 2B). Such experiments together with others in which the assay pH or temperature was altered established the optimal condition as at 55°C in 100 mM Tris-acetate (pH 8), 500 mM Mg(OAc)2 and 2.5 M NH4OAc.

The extent of processing of 5'-labeled pt Tyr-S5-M RPR was determined by quantitating formation of the 5' leader. A single exponential function adequately describes the rate of product formation (data not shown; see also Fig. 2.4A). We next inquired if kobs corresponds to the rate of the cleavage step. Deprotonation of a hydrated Mg2+ ion in the RPR’s active site is believed to generate the hydroxide nucleophile for
attacking the scissile phosphodiester linkage in the ptRNA substrate [44, 108]. Therefore, evidence for cleavage being rate limiting is generally derived from correlating the rate of product formation with the hydroxide ion concentration [41]. A slope of ~1 in plots of log($k_{\text{obs}}$) vs. pH reflects cleavage being the sole rate-limiting determinant while values <<1 imply that other factors (e.g., conformational rearrangements, substrate docking) might also contribute to $k_{\text{obs}}$.

When we assayed pt$^{\text{Tyr}}$-$S_5$-M RPR at pH 5.5 and 7.5, we observed only a 2-fold increase in the rate of self-cleavage as opposed to the 100-fold enhancement that is expected if chemistry is rate limiting. Based on the premise that substrate docking in pt$^{\text{Tyr}}$-$S_5$-M RPR is probably suboptimal and contributes to the observed rate (in addition to cleavage), we inquired if alterations in the length of the spacer might yield a cis conjugate whose rate is determined solely by chemistry. Indeed, log($k_{\text{obs}}$) vs. pH for the self-cleavage of a 3-nt-spacer conjugate, pt$^{\text{Tyr}}$-$S_3$-M RPR, exhibited a slope of 0.95 in the pH range 5.1 to 6.3 (Fig. 2.10A). As might be expected for a first-order reaction, we found that the $k_{\text{obs}}$ values at different pt$^{\text{Tyr}}$-$S_3$-M RPR concentrations are nearly identical.

The $k_{\text{obs}}$ value of 0.15 min$^{-1}$ at pH 5.5 and 55°C for pt$^{\text{Tyr}}$-$S_3$-M RPR that we determined is comparable to the reported value of 0.75 min$^{-1}$ at pH 5.5 and 50°C for pt$^{\text{Asp}}$-$S_5$-Eco RPR [38]. Two inferences merit mention. First, although there are differences such as the ptRNA, site of conjugation, and assay temperature used in the two studies, the parallel in $k_{\text{obs}}$ is striking given that the type M RPR lacks the functional S domain present in E. coli RPR. Tethering of the substrate to the type M RPR appears to have rendered the S domain less important. Second, since Mja RPR is unable to process
ptRNA\textsuperscript{Tyr} in trans under these assay conditions where the cis construct is active, the covalently attached substrate in pt\textsuperscript{Tyr}-S\textsubscript{3} (and S\textsubscript{5})-M RPR must assist the RPR to attain a catalytically active conformation.

2.3.2 Reconstitution of a type M archaeal RNase P holoenzyme

To examine the contribution of \textit{Mja} RPPs to \textit{Mja} RPR catalysis, we cloned the genes encoding \textit{Mja} POP5, RPP21, RPP29, and RPP30 into vectors that allow T7 RNA polymerase-based overexpression in \textit{E. coli} BL21(DE3) cells; the \textit{Mja} RPP genes were easily identified based on their amino acid sequence homology to archaeal/eukaryal RPP homologs. By exploiting the high pi values of \textit{Mja} RPPs (Table 3), we expected to purify them by cation-exchange chromatography, a strategy that proved fruitful for \textit{Pfu} RPPs [69]. However, initial attempts to isolate the \textit{Mja} RPPs were unsuccessful; we either failed to overexpress the \textit{Mja} RPPs in \textit{E. coli} or purify them to homogeneity.

As an alternative, we explored the possibility of purifying \textit{Mja} RPPs as binary complexes of interacting partners after their co-overexpression in \textit{E. coli}. The rationale for this strategy was based on two observations. First, biochemical and genetic studies had already established pair-wise interactions between POP5-RPP30 and RPP21-RPP29 [77, 78]. Second, the realization that multi-component protein complexes are better reconstituted \textit{in vivo} than \textit{in vitro} (presumably due to protein folding) has led to the design of vectors that would permit co-overexpression of two or more polypeptides in a host such as \textit{E. coli} [109, 110]. The interacting pair of proteins could be expressed from either a bi-cistronic construct in one vector or from two separate (compatible) plasmids each of which harbors one ORF. We have successfully employed both routes. The
interacting proteins are overexpressed in *E. coli* and the binary complex generally isolated by affinity chromatography using an affinity handle that was engineered into one of the proteins. Rather than employ an affinity tag, we purified to homogeneity binary RPP complexes (types A and M) by employing heat denaturation, ammonium sulfate fractionation, and cation-exchange chromatography (Fig. 2.6). While this report was being prepared, Perederina et al. (2007) reported that two human RPPs (POP6 and POP7) that form a heterodimer could also be isolated as a recombinant, native complex without affinity tags.

When the *Mja* RPP binary complexes (i.e., POP5-RPP30 and RPP21-RPP29) were reconstituted with the *Mja* RPR in 50 mM Tris-acetate (pH 7.5), 30 mM Mg(OAc)$_2$ and 0.8 M NH$_4$OAc, the resulting *Mja* RNase P holoenzyme exhibited multiple turnover ($k_{cat} \sim 6-24 \text{ min}^{-1}$) at 55°C with *E. coli* ptRNA$^{\text{Tyr}}$ or *synechocystis* ptRNA$^{\text{Gln}}$ as the substrate. A $k_{cat}$ value of 34 min$^{-1}$ was reported for cleavage of *B. subtilis* ptRNA$^{\text{Asp}}$ by partially purified *Mja* RNase P holoenzyme at 50°C [59]. Neither the *Mja* RPR nor the binary RPP complexes alone promote cleavage of ptRNA$^{\text{Tyr}}$ under similar assay conditions (Fig. 2.7, lanes 3 and 4).

2.3.3 Similarities between partially reconstituted type A and M RNase P holoenzymes

The observations that *Pfu* (type A) RPR, which is capable of multiple turnover in the absence of RPPs, displays activity at lower substrate and Mg$^{2+}$ concentrations in the presence of either RPP21-RPP29 or POP5-RPP30 suggested that the type A RPR together with either RPP pair constitutes minimal functional complexes. The *Pfu* RPR +
RPP21-RPP29 and Pfu RPR+POP5-RPP30 complexes displayed 600- and 100-fold lower $k_{cat}/K_m$ values compared to the reaction with RPR+all four RPPs [69]. To test whether these findings are qualitatively applicable for a type M RNase P, we assayed at 55°C the partial RNP complexes generated from Mja RPR and either Mja RPP21-RPP29 or POP5-RPP30. Robust activity, under multiple turnover conditions, was observed upon reconstitution of Mja RPR with POP5-RPP30 (Fig. 2.7, lane 7). In contrast, under identical conditions using same amount of enzyme, we were not able to observe any activity with Mja RPR reconstituted with RPP21-RPP29 (Fig. 2.7, lane 6); however, after prolonged incubation, under single turnover conditions, we could observe weak activity with a partial complex involving Mja RPR + RPP21-RPP29 (Chen and Gopalan, unpublished observations). Since the Mja RPR alone fails to catalyze the trans cleavage reaction, it is remarkable that the RPP21-RPP29 pair is able to elicit even weak activity.

Before drawing the inference that the differences in behavior of the partially reconstituted Mja (type M) and Pfu (type A) RNase P complexes are attributable to fundamental differences in the structure and functioning of these two classes of archaeal RNase P, some other possibilities had to be ruled out. First, the Mja RPPs were purified as binary complexes while the Pfu RPPs used in the earlier study had been purified as individual RPPs [69]. We have now verified that the partial reconstitution results are not different when Pfu RPP binary complexes are used (Cho and Gopalan, unpublished observations). Second, there was the remote likelihood that Pfu RNase P might not have been an ideal type A representative. Therefore, we decided to purify RPPs from another member of the type A family [Methanothermobacter thermoautotrophicus (Mth)] and
repeat the partial reconstitution studies (Cho, Tsai and Gopalan, unpublished observations). The results obtained with *Mth* RNase P parallel those reported with *Pfu* RNase P [69]. Lastly, we confirmed that the results obtained with *Mja* RPR are valid even when a different type M RPR (e.g., *Afu*) was used in the reconstitution (data not shown). Hence, we conclude that the structural differences between type A and M RPRs translate into functional consequences during their assembly.

2.3.4 Effect of *Mja* RPPs on the rate of self-processing of ptRNA\textsubscript{Tyr}-S\textsubscript{3}-M RPR

Our previous studies on *in vitro* reconstituted *Pfu* RNase P revealed that the RPPs functioned as binary complexes (POP5-RPP30 and RPP21-RPP29). By measuring the $k_{\text{cat}}$ and $K_m$ values exhibited by partially and fully reconstituted *Pfu* RNase P, we recently determined that the POP5-RPP30 pair (but not RPP21-RPP29) enhances the $k_{\text{cat}}$ of the RPR by nearly 40-fold, in fact to the same extent observed with all four RPPs [69]. Since $k_{\text{cat}}$ sets a lower limit on the first-order rate constants for steps subsequent to substrate binding, the POP5-RPP30 pair was inferred to play a vital role in cleavage and/or product release. Although determining the rate constants for individual steps in the absence and presence of the RPP pairs would be ideal to fully understand the roles played by RPPs and formulate a comprehensive description of the kinetic scheme, we rationalized that studying the cis conjugate would allow us to focus solely on the cleavage step without any influence from substrate binding or product release.

Incubation of pt\textsubscript{Tyr}-S\textsubscript{3}-M RPR with *Mja* POP5-RPP30 or *Mja* RPP21-RPP29 revealed that both RPP pairs considerably decrease the concentration of monovalent and divalent ions required for cis cleavage (Table 4); for example, the Mg$^{2+}$ requirement
decreases from 500 to 100 mM in the presence of either binary complex. Presence of all
four RPPs caused a further decrease in the Mg$^{2+}$ requirement to 20 mM (Fig. 2.8A; Table
4).

*Mja* POP5-RPP30, but not RPP21-RPP29, enhances the rate of self-processing of
pt$^{\text{Tyr}-S_3}$-M RPR by nearly 100-fold (0.05 to 4.83 min$^{-1}$ at pH 5.1; Figs. 2.8A and B, Table
4). To verify that POP5-RPP30 actually contributes to the cleavage step, we determined
$k_{\text{obs}}$ at various pH values. A plot of log($k_{\text{obs}}$) vs pH for the pt$^{\text{Tyr}-S_3}$-M RPR + POP5-
RPP30 exhibited a slope of 0.92 proving that indeed cleavage is rate limiting for the
RPR-catalyzed reaction both in the absence and presence of POP5-RPP30 (Fig. 2.10A
and B).

We next proceeded to determine if all four RPPs together will enhance the rate to
a level greater than that observed in the presence of POP5-RPP30. However, technical
issues prevented us from doing this experiment under conditions employed for the binary
complex. For trans cleavage reactions involving RPR reconstituted with all four RPPs,
we have always employed a pre-incubation step (10 min at 37°C followed by 10 min at
55°C) to permit RNP complex formation prior to substrate addition. But such an approach
with the cis-conjugate triggers self-cleavage during the pre-incubation step thus
complicating rate calculations. Therefore, we had to devise an alternative strategy to
circumvent this problem.

Since kinetic and thermodynamic studies using bacterial RPRs and ptRNAs
revealed that the RPR adopts the correct tertiary fold for optimal enzyme-substrate (ES)
complex formation in a buffer containing Ca$^{2+}$, a divalent cation that promotes efficient
ptRNA binding but not processing [41] we pre-incubated pt\textsuperscript{Tyr-S}\textsubscript{3-M} RPR with either two RPP pairs (POP5-RPP30 + RPP21-RPP29) or POP5-RPP30 in a buffer containing 25 mM Ca\textsuperscript{2+} instead of Mg\textsuperscript{2+}; cleavage was initiated by adding an excess of Mg\textsuperscript{2+} to displace the Ca\textsuperscript{2+}. To determine the concentration of Mg\textsuperscript{2+} required for displacing the Ca\textsuperscript{2+} and promoting maximum cleavage, the rate of self-processing was determined at various concentrations of Mg\textsuperscript{2+}, added at the end of the pre-incubation step. Such a Mg\textsuperscript{2+}-titration analysis from 10 to 150 mM Mg\textsuperscript{2+} revealed that maximum cleavage was obtained when the pre-incubation mixture was supplemented with 100 mM Mg\textsuperscript{2+} during the assay.

Despite the viability of the above approach, it is conceivable that the calculated rate is influenced adversely by the pre-incubation with Ca\textsuperscript{2+} and its subsequent displacement by Mg\textsuperscript{2+}. Therefore, we compared the \(k_{\text{obs}}\) for self-processing of pt\textsuperscript{Tyr-S}\textsubscript{3-M-RPR + POP5-RPP30 under two conditions: assembled in a reconstitution buffer containing either 1 mM Mg\textsuperscript{2+} or 25 mM Ca\textsuperscript{2+} and subsequently assayed in the presence of 100 mM Mg\textsuperscript{2+}. In both cases, the result was nearly identical (4.8 versus 5.2 min\(^{-1}\)) thus validating the Ca\textsuperscript{2+}-based approach. Using this method, we determined that the rate of pt\textsuperscript{Tyr-S}\textsubscript{3-M-RPR self-cleavage in the presence of POP5-RPP30 (5.2 min\(^{-1}\)) was not appreciably enhanced by addition of RPP21-RPP29 (5.5 min\(^{-1}\); Table 4). This result suggests that the RPP21-RPP29 pair does not contribute to the rate of cleavage of the cis conjugate under the conditions employed (see Discussion).
2.3.5 An *Mja* RPR-ptRNA conjugate lacking the S domain also exhibits self-processing

Studies on type A bacterial (e.g., *Eco*) and archaean (e.g., *Pfu*) RPRs established that their C domains support catalysis albeit only in the presence of cognate RPPs [21, 69]. With the C domain of *B. subtilis* RPR (bacterial type B), the cleavage rate was 25,000-fold lower compared to the wild type [22]. These observations suggest that the S domain contributes to ptRNA recognition and helps position the ptRNA for optimal cleavage. Reasoning that failure to observe robust ptRNA processing with the C domains of bacterial/archaean RPRs was likely due to substrate-binding/positioning defects caused by removal of the S domain, we hypothesized that the *Mja* RPR-ptRNA cis conjugate should be active even without the S domain since the substrate is covalently tethered to the enzyme. Indeed, pt\(^{\text{Tyr}}\)-S\(^3\)-\(\Delta\)S *M* RPR, a cis conjugate lacking the S domain, is active (Fig. 2.9A, lane 1 and 2.9B). Interestingly, instead of the 25,000-fold defect observed in the trans cleavage scenario with the C domain of *B. subtilis* RPR, the \(k_{\text{obs}}\) of pt\(^{\text{Tyr}}\)-S\(^3\)-\(\Delta\)S *M* RPR is only 6-fold lower than wild type pt\(^{\text{Tyr}}\)-S\(^3\)-M RPR at pH 6.7 (0.24 ± 0.01 versus 1.5 ± 0.11 min\(^{-1}\); Fig. 2.9B). This difference is nullified in the presence of POP5-RPP30 (5.2 versus 4.9 min\(^{-1}\) at pH 5.1; Fig. 2.9A, Table 4). Although we cannot compare *Mja* RPR with and without its S domain for trans cleavage, the rate of self-processing of pt\(^{\text{Tyr}}\)-S\(^3\)-\(\Delta\)S *M* RPR suggests that covalent attachment of the ptRNA substrate results in a less adverse impact than what removal of the S domain might have caused in a trans cleavage reaction.

Although the RPP21-RPP29 complex does not increase the rate of self-cleavage of pt\(^{\text{Tyr}}\)-S\(^3\)-M RPR, it does lower the ionic requirement from 2.5 M to 0.8 M NH\(_4^+\) and
from 500 mM to 100 mM Mg\textsuperscript{2+} (Fig. 3C, Table 1). Addition of RPP21-RPP29 to pt\textsuperscript{Tyr}-S\textsubscript{3}-ΔS M RPR impacts neither the rate nor the ionic requirement suggesting that this binary complex might interact with the S domain (Fig. 3B, lane 3; data not shown). Indeed, ongoing footprinting analyses of \textit{in vitro} reconstituted \textit{Mja} RNase P holoenzyme indicate that RPP21-RPP29 binds predominantly to the S domain (Pulukkunat and Gopalan, unpublished observations).

2.4 Discussion

2.4.1 RPR is the catalytic subunit in all three domains of life

Although the bacterial RPR alone is catalytically active under \textit{in vitro} conditions of high ionic strength, RPRs from many archaeal and all eukaryal sources, despite sequence/structure similarities with their bacterial counterpart, were reported as incapable of catalysis in the absence of their cognate RPPs [10, 67, 111]. Structural defects that prevent generation of an active RNA tertiary fold are occasionally mitigated by high concentration of monovalent and divalent cations [112, 113]. Such an experimental approach has failed to render many archaeal and eukaryal RPRs active [114]. This has raised the possibility (if slim) that the decreased archaeal and eukaryal RPR activity might indicate a natural course towards a more complex RNP in which protein subunits have usurped catalytic roles. However, results from a recent study and the data reported here do not support such a premise. First, Kikovska et al. (2007) showed that the human RPR catalyzes processing of different ptRNAs and model substrates, albeit at a 10\textsuperscript{6}-fold lower rate than that of bacterial RPR. Second, we have demonstrated here that type M euryarchaeal RPR, thus far proven to be incapable of ptRNA processing, can catalyze the
ptRNA processing reaction if (i) its catalytic domain is fused to the S domain of bacterial RPR that is capable of interacting with the T-stem and loop of ptRNA, or (ii) the ptRNA substrate is provided in cis (Figs. 2.2 and 2.3).

Since the human and type M archaeal RPRs support ptRNA processing in the presence of 150 to 500 mM Mg\(^{2+}\), but in the absence of RPPs they must possess all the information necessary for generating the active site and confirm that the RPR is the catalytic moiety in all three domains of life. Although such a broad inference might seem inconsistent with various reports of failure to detect archaeal/eukaryal RPR alone-catalyzed ptRNA processing, these results might reflect either an inability to promote an active RPR fold \textit{in vitro} or a masking of the cleavage capability due to extremely weak ptRNA binding [114]. While we highlight the catalytic competence of the archaeal/eukaryal RPRs, we recognize that their cognate RPPs must play pivotal roles in promoting the RPR’s structural scaffold and contributing to interactions with the substrate and Mg\(^{2+}\) ions. Our ongoing footprinting and transient kinetic studies are expected to shed more light on how mutual interactions between the RPR and RPPs results in optimization of an RNA-based active site in the archaeal RNase P holoenzyme.

Extensive mutagenesis of the universally conserved P4 region in archaeal RPRs resulted in an RNase P holoenzyme for which the optimum Mg\(^{2+}\) requirement was elevated significantly [115] (Chen and Gopalan, unpublished work). Also, deletion of three nucleotides in the P4 helix of human RPR resulted in loss of activity [86]. Cumulatively, these data are again indicative of a common RNA-mediated catalytic mechanism in all RNase P holoenzymes and support the claim that the functionally
conserved P4 structure mediates binding of metal ions required for catalysis, although this might be indirect through substrate positioning [40].

2.4.2 Role of RPPs in archaeal RNase P catalysis

Despite the RPR being responsible for the cleavage step in all three domains of life, it is clear that the RPRs require more assistance from RPPs in archaea and eukarya compared to bacteria. Although the bacterial RPR is folded into a stable structure that is very active ($k_{\text{obs}} \sim 10 \text{ min}^{-1}$; [86, 116]), the sole RPP normalizes the binding affinity and rate of cleavage of different ptRNAs by the RPR [52]. In comparison, what are the roles played by the multiple archaeal/eukaryal RPPs? High-resolution structures of bacterial RPRs reveal how long-range intra-molecular RNA contacts provide the underpinning for precise positioning of the two independently folding S and C domains that in turn enable the remarkable inter-domain cooperation in substrate binding and catalysis [28, 29]. Are these tertiary structure-stabilizing intramolecular struts, which are missing in archaeal/eukaryal RPRs, replaced by RPPs? If so, could archaeal/eukaryal RPPs be distinguished by their structural (i.e., providing a substitute for the structural braces in bacterial RPRs) or catalytic roles in substrate binding and cleavage (akin to bacterial RPP)? Our studies on $Pfu$ and $Mja$ RNase P are providing insights into such a demarcation of roles.

Based on the similarities in the footprint of bacterial RPP and $Pfu$ POP5-RPP30 in the C domains of their respective RPRs and their ability to contribute to the cleavage step and or product release, we had speculated that the $Pfu$ POP5-RPP30 was proximal to the ptRNA cleavage site akin to the bacterial RPP [51, 56, 57, 69]. Our studies with $Mja$
RNase P now provide further support for this idea and specifically demonstrate that the POP5-RPP30 complex contributes to the cleavage. The strikingly similar three-dimensional structures of archaeal POP5 and bacterial RPP [74, 75] despite poor sequence homology and different secondary structure connectivities, lends significance to this functional parallel and suggests an instance of evolutionary convergence while optimizing the function of a catalytic RNP in the different domains of life.

The self-cleavage reaction of ptTyr-S3-M RPR, where we are monitoring only the rate of phosphodiester bond hydrolysis, is accelerated by POP5-RPP30 but not RPP21-RPP29 (Fig. 2.8). Similarly, only POP5-RPP30 promotes the trans cleavage of ptRNA by Mja RPR under multiple turnover conditions. The simplest interpretation of these data is that RPP21-RPP29 is unable to aid the cleavage step. However, addition of RPP21-RPP29 to the partial holoenzyme made up of ptTyr-S3-M RPR and POP5-RPP30 lowers the Mg2+ requirement for the cleavage reaction from 100 to 20 mM (Table 4). Therefore, Mja RPP21-RPP29 must be able to promote the RPR scaffold necessary for substrate binding and or catalysis but cannot expedite the chemical step directly as does POP5-RPP30. We interpret this as indicative of RPP21-RPP29 playing a structural role that does have catalytic consequences, although only when enabled by POP5-RPP30. In a similar vein, Pfu RPP21-RPP29 decreases the Km by five-fold while leaving the kcat of the Pfu RPR-catalyzed reaction unaltered; while both Pfu RPP21-RPP29 and POP5-RPP30 have the ability to decrease the Km value (~3- or 5-fold), there is a significant synergism (170-fold) when both pairs are present [69].
Although the cis conjugate has enabled us to validate the catalytic potential of type M RPRs and even dissect the role of RPPs in aiding RPR catalysis, some caution is required while interpreting these results. First, since we are yet to resolve technical issues which have thwarted the use of a rapid quench flow apparatus to perform transient kinetic assays at near-neutral pH and a temperature of 55°C, we resorted to using an assay pH of 5.1, which is lower than the physiological pH, but allowed us to manually determine the rate of self-cleavage of p\textsuperscript{Tyr}-S\textsubscript{3}-M RPR at 55°C. We recognize that protein–protein interactions in the archaeal RNase P holoenzyme at pH 5.1 might differ from those occurring at physiological pH thus requiring circumspection while drawing inferences on the influence of RPPs on the RPR-catalyzed reaction. Second, since the M\textit{ja} RPR fails to exhibit any trans cleavage activity in the absence of RPPs, we utilized the rate of self-processing of a cis conjugate as the baseline for calculating RPR’s catalytic potential and then determined how this rate was affected by RPPs. However, it is conceivable that this cis conjugate might yield a different pattern of activation by RPPs when compared to a trans cleavage reaction. For instance, if conjugation of the substrate in cis renders the S domain somewhat redundant, RPPs interacting with the S domain to make it functional might also be inferred as being unnecessary. Since RNase T1-based footprinting analysis of \textit{in vitro} reconstituted archaeal RNase P holoenzyme indicates that RPP21-RPP29 binds predominantly to the S domain of archaeal RPRs (Pulukkunat and Gopalan, unpublished observations), it is possible that the full contribution of RPP21-RPP29 might not be uncovered while using a cis substrate that might not require the S domain.
Despite the caveats discussed above, some inferences on the role of *Mja* RPPs can be drawn from our kinetic studies. Divalent metal ions play both structural and functional roles in RNA-mediated catalysis. Our results indicate that both *Mja* POP5-RPP30 and RPP21-RPP29 reduce the monovalent and divalent ion requirements for the self-cleavage of pt^{Tyr}-S_{3}-M RPR. Since the POP5-RPP30 pair (but nor RPP21-RPP29) enhances the rate of cleavage with a concomitant reduction in concentration of Mg^{2+} required for cleavage, we postulate that this binary complex functions primarily to enhance cleavage by promoting the affinity of the active site metal ions in the RPR (mimicking one of the roles of the bacterial RPP). On the other hand, the RPP21-RPP29 pair, which does not enhance the $k_{obs}$ for self-cleavage, is likely to play an indirect role in catalysis through stabilization of the RPR’s tertiary fold. We speculate that the RPP21-RPP29 pair substitutes for some of the RNA-RNA tertiary interactions that stabilize the core structure in bacterial RPRs.

The ability of RPP21-RPP29 to assist the RPR to adopt a catalytically-competent fold at lower concentrations of monovalent and divalent ions than the RPR-alone reaction parallels protein-facilitated self-splicing reactions of certain group I introns. Tyrosyl-tRNA synthetase (CYT-18), a bifunctional protein from *Neurospora crassa* (*N. crassa*), is both a group I intron protein cofactor and an amincoacyl synthetase. The observations that (i) most of the *N. crassa* group I introns, which are substrates of CYT-18, can form a tertiary structure only in the presence of CYT-18 despite possessing a conserved secondary structure, and (ii) CYT-18 can functionally replace the peripheral structural elements that stabilize the catalytic core of an unrelated *Tetrahymena* group I intron that
can generate an active tertiary fold at high Mg$_{2+}$ concentrations, suggested that CYT-18 might have taken over the role of certain RNA structural elements. Indeed, a recent co-crystal structure of CYT-18 bound to Neurospora group I intron elegantly illustrates how CYT-18 could promote the precise positioning of two different domains in the group I intron and thereby facilitate formation of a catalytically competent tertiary fold [117]. Analogously, we suggest that the RPP21-RPP29 pair has substituted for some intramolecular braces present in bacterial RPR and promotes an archaeal RPR tertiary structure required for optimal catalysis in the presence of POP5-RPP30 (see Fig. 4.3 for a comparison of models of the bacterial and archaeal RNase P holoenzymes).

### 2.4.3 Insights from studies on the C domain of type M RPR

Despite the identification of a conserved C domain in all RPRs, previous studies suggested that this domain was either not functional in the absence of cognate RPPs or that its activity was nearly 25,000-fold weaker than the wild type [22]. If the decreased or negligible cleavage observed in these instances is attributable to weak substrate binding and sub-optimal positioning of the substrate/metal ions, such a defect might be remedied when the substrate is covalently attached to the C domain. Indeed, our study shows that this is the case. Self-cleavage of pt$_{\text{Tyr-S3-ΔS}}$ M RPR is only 6-fold slower than pt$_{\text{Tyr-S3}}$-M RPR at pH 6.7; however, the decreased activity suggests that the S domain does play a role in substrate positioning (related to cleavage) even when the substrate is present in cis.

Interestingly, both pt$_{\text{Tyr-S3}}$-M RPR and pt$_{\text{Tyr-S3-ΔS}}$ M RPR when reconstituted with POP5-RPP30 display near identical rates of self-cleavage (Table 4). If POP5-RPP30 enhances the rate of cleavage by assisting the RPR in positioning the leader sequence or
the scissile linkage in the ptRNA substrate and such a function is redundant with the S domain, one can account for how the POP5-RPP30 binary complex compensates for the absence of the S domain.

Although pt\textsuperscript{Tyr}-S\textsubscript{3}-ΔS M RPR was used to study a cis cleavage event, ΔS \textit{Mja} RPR represents the smallest functional RPR yet identified (135 nts). We have now gathered evidence that this mini-RPR can be further trimmed to 110 nts and that it promotes robust trans cleavage activity after reconstitution with its four cognate RPPs (data not shown). These findings suggest that RNase P, historically classified as a large ribozyme, need not have begun its evolutionary odyssey as such in an RNA world.
Figure 2.1 Schematic depicting secondary structures of RPRs from three domains of life. Representative bacterial (A), archaeal (B and C), and eukaryal (D) RPRs are shown. Type A and M RPRs are illustrated in panels B and C, respectively. Nucleotides that are conserved in all three domains are indicated. Figure reproduced from [111]
Figure 2.2 Bacterial (S)-archaeal (C) hybrid RPR is catalytically active and can accurately processes *E. coli* ptRNA^{Tyr}. A. Secondary structure representation of a bacterial/archaeal hybrid RPR (*Eco*S-*Mja*C RPR). B. *EcoS-MjaC* RPR (2.5 μM) was tested for RNase P activity by incubating with 5' labeled *E. coli* ptRNA^{Tyr} (~1 pM) in the presence of various NH₄ and Mg²⁺ concentrations as indicated. M represents the size marker generated by processing of *E. coli* ptRNA^{Tyr} by *E. coli* RNase P.
Figure 2.3 Design and optimization of reaction conditions for an active site-tethered ES conjugate. A. Sequence and secondary structure representation of the ptRNA-Mja RPR conjugate. E. coli ptRNA\textsuperscript{Tyr} is attached to the L15 loop of Mja RPR with either a 3- or 5-nt-spacer sequence (abbreviated as pt\textsuperscript{Tyr}-S\textsubscript{3} or S\textsubscript{5}-M RPR in the text). Nucleotides that are not part of the Mja RPR or ptRNA are shown in lowercase letters. B. Titration of monovalent and divalent ions to identify the optimal conditions for maximum self-cleavage activity (see text for details).
Figure 2.4 Secondary structure representation of ptRNA-{\textit{Mja}} RPR conjugate in which the S domain has been deleted (p\textsuperscript{\textit{Tyr}}-{\textsubscript{S\textsubscript{3}}}Δ\textsubscript{S} M RPR).
Figure 2.5 Comparison of the self-cleavage rates of $pt^{Tyr-S_3-M}$ RPR and $pt^{Tyr-S_3-ΔS}$ M RPR. A single exponential function adequately describes the rate of product formation. Rates ($k_{obs}$) of product formation for $pt^{Tyr-S_3-M}$ RPR (A) and $pt^{Tyr-S_3-ΔS}$ M RPR (B) were found to be 1.5 and 0.24 min$^{-1}$, respectively. The rates were obtained by incubating the ptRNA-RPR conjugates at 55°C in 50 mM MES pH 6.7, 2.5 M NH$_4^+$, and 500 mM Mg$^{2+}$. 

71
Figure 2.6 SDS-PAGE profiles depicting the purification of binary complexes of *Mja* RPPs. Panels A and B denote the SDS-PAGE analysis to determine the homogeneity of the purified binary complexes RPP21-RPP29 and POP5-RPP30, respectively. UI and I denote the total crude lysate from the uninduced and induced *E. coli* BL21(DE3) cultures, respectively. IP represents the ammonium sulfate fractionated sample that was subjected to the SP-sepharose column chromatography and FT denotes the flow-through.
Figure 2.7 Reconstitution of RNase P activity using *Mja* RPR and cognate RPPs purified as binary complexes. See text for details.
Figure 2.8 Effect of *Mja* RPPs on the rate of self-processing reaction catalyzed by \( \text{pt}^{\text{Tyr-S}_3}\text{-M RPR} \). Panel A illustrates reconstitution of RNase P activity using 5' labeled \( \text{pt}^{\text{Tyr-S}_3}\text{-M RPR} \) and binary complexes of *Mja* RPPs at various concentrations of monovalent and divalent ions. The assays were performed in 50 mM MES (pH 5.1 at 55°C). B. In panel B, the rate of self-cleavage of the RPR alone (filled triangle) is compared to that of RPR + RPP21-RPP29 (filled circle) and RPR + POP5-RPP30 (filled square). The assays were performed in 50 mM MES pH 5.1 at 55°C.
Figure 2.9 Effect of *Mja* RPPs on the rate of self-processing reaction catalyzed by \( pt^{\text{Tyr-S}_3-\Delta S} \text{ M RPR} \). A. Reconstitution of RNase P activity using 5' labeled \( pt^{\text{Tyr-S}_3-\Delta S} \text{ M RPR} \) and binary complexes of *Mja* RPPs. B. Rate of self cleavage of the \( pt^{\text{Tyr-S}_3-\Delta S} \text{ M RPR + POP5-RPP30} \). The assays were performed in 50 mM MES pH 5.1 at 55°C.
Figure 2.10 Dependence of $k_{obs}$ on assay pH for the $\text{pt}^{\text{Tyr-S}_3}\text{-M RPR}$ self-cleavage reaction in the absence and presence of RPPs. Panels A and B represent plots of log $k_{obs}$ vs. pH for the self-cleavage reaction catalyzed by $\text{pt}^{\text{Tyr-S}_3}\text{-M RPR}$ and $\text{pt}^{\text{Tyr-S}_3}\text{-M RPR + POP5-RPP30}$, respectively. Average values from at least three independent experiments are provided. The slope values are 0.95 (for $\text{pt}^{\text{Tyr-S}_3}\text{-M RPR}$) and 0.92 (for $\text{pt}^{\text{Tyr-S}_3}\text{-M RPR + POP5-RPP30}$), respectively.
<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>AfuRPR-F</td>
<td>5'-TGGGAGCGGCGGCTGGG CGGC-3'</td>
</tr>
<tr>
<td>AfuRPR-R</td>
<td>5'-GGAAGCTTTGGGAGTGAGGCTGG-3'</td>
</tr>
<tr>
<td>MjaRPR-F</td>
<td>5'-GGGGGAATCCGTCTCGGGGCTATGGGGGC-3'</td>
</tr>
<tr>
<td>MjaRPR-R</td>
<td>5'-GAGGAGGAAAGGCCTGGGTGTAAGGCC-3'</td>
</tr>
<tr>
<td>MmaRPR-F</td>
<td>5'-GAGCAAGGGCAAGGGCTGGTGACTATC-3'</td>
</tr>
<tr>
<td>MmaRPR-R</td>
<td>5'-GCGAATTCGGCAGCTGGGCTATAGGCTG-3'</td>
</tr>
<tr>
<td>MJATYR-F1</td>
<td>5'-GAGCAGGCAGTAAAAAGCATTACC-3'</td>
</tr>
<tr>
<td>MJATYR-R1</td>
<td>5'-GCGCCGTACCTCTCTTGGGTGGGAAGGATTC-3'</td>
</tr>
<tr>
<td>MJATYR-F2</td>
<td>5'-CACCAACAGAAAGGTACGGCGCTTAGC-3'</td>
</tr>
<tr>
<td>MJATYR-2</td>
<td>5'-GGAATTCGGTGATGGGCCGAGAATGC-3'</td>
</tr>
<tr>
<td>MJATYR-F3</td>
<td>5'-GGAATTCGAGGCTAGGGTGGGCTG-3'</td>
</tr>
<tr>
<td>MJATYR-R3</td>
<td>5'-GGGATCCGTCTCGCTATGGAGCCGATG-3'</td>
</tr>
<tr>
<td>MJCECS-F</td>
<td>5'-CCCATTTAGGGTGCCAGGTAACGCTGGGGAAGA-3'</td>
</tr>
<tr>
<td>MJCECS-R</td>
<td>5'-CTTTGCAACCCGGGT GGAGTTTACGTC-3'</td>
</tr>
<tr>
<td>ECSMJC-R</td>
<td>5'-GGTAACCTGGGACCCCTTAAATGGGGGTGGCGGAACCTTC-3'</td>
</tr>
<tr>
<td>ECSMJC-F</td>
<td>5'-CTCCACCCGGTGCAAGGCCGCTTTCCGCTATAGC-3'</td>
</tr>
<tr>
<td>MjapTyrSp-F</td>
<td>5'-CGGCGCCTTAGCCGAAATGTC-3'</td>
</tr>
<tr>
<td>MjapTyr0Sp-R</td>
<td>5'-GATAGGTGGGGAAGGATTCGACC-3'</td>
</tr>
<tr>
<td>MjapTyr5Sp-R</td>
<td>5'-ATAGGTGGGGAAGGATTCGACC-3'</td>
</tr>
<tr>
<td>MjapTyr3Sp-R</td>
<td>5'-GATAGGTGGGGAAGGATTCGACC-3'</td>
</tr>
<tr>
<td>MjapTyr4Sp-R</td>
<td>5'-GATAGGTGGGGAAGGATTCGACC-3'</td>
</tr>
<tr>
<td>MjaD67-192-F</td>
<td>5'-GGTGCAAGCGGGCTTTCC-3'</td>
</tr>
<tr>
<td>MjaD67-192-R</td>
<td>5'-ATAAATGGGGTGGCGGAAC-3'</td>
</tr>
</tbody>
</table>

**Table 1 Oligonucleotides used in this study for cloning various archaeal RPRs or their derivatives**

*Note*: Wherever relevant, the italicized and underlined sequences indicate the restriction sites introduced for cloning and linearization of template (for *in vitro* transcription), respectively.
<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>MjaPOP5-F</td>
<td>5'-TGATAGAAATGCTGAAACACTGCCACCAACTTT AAGGG -3'</td>
</tr>
<tr>
<td>MjaPOP5-R</td>
<td>5'-GGGGATCCCTTATTTTTATTGTTTTTAGCTTTCAACCTTTTTC -3'</td>
</tr>
<tr>
<td>MjaRPP21-F</td>
<td>5'-AAAAAGTTCTTAGAAAAAAGCTAAAGAAGATAGC-3'</td>
</tr>
<tr>
<td>MjaRPP21-R</td>
<td>5'-GGGGATCCTTAAAGGTTTGTGAGTTTTGATTTTTGCTTTT AATCTTTCC-3'</td>
</tr>
<tr>
<td>MjaRPP29-F</td>
<td>5'-TGATCACTCCGCACAATATCCTGAGGATGAAGCTTATA GGGC -3'</td>
</tr>
<tr>
<td>MjaRPP29-R</td>
<td>5'-GGGGATCCCTAATACGGATACGATCTTTTATTTTTTTTCTTTT AATCTC-3'</td>
</tr>
<tr>
<td>MjaRPP30-F</td>
<td>5'-CGTATCGATATCAACAGAATAGAAAAAGGAAGAGG -3'</td>
</tr>
<tr>
<td>MjaRPP30-R</td>
<td>5'-GGGGATCCCTACTTTTTTTTTTCTTATAATTTCCA CTCC -3'</td>
</tr>
</tbody>
</table>

Table 2 Oligonucleotides used in this study for cloning the various archaeal RPPs

Note: The italicized sequences indicate the restriction sites introduced for cloning. All oligonucleotides were purchased from Integrated DNA Technologies.
Table 3 Characteristics of different archaeal RPPs used in this investigation

<table>
<thead>
<tr>
<th>Organism</th>
<th>RPP</th>
<th>Gene ID</th>
<th>Predicted mass, Da</th>
<th>Isoelectric point (pI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Pfu</em></td>
<td>POP5</td>
<td>PF1378</td>
<td>13,708</td>
<td>10.0</td>
</tr>
<tr>
<td></td>
<td>RPP21</td>
<td>PF1613</td>
<td>14,172</td>
<td>10.3</td>
</tr>
<tr>
<td></td>
<td>RPP29</td>
<td>PF1816</td>
<td>14,953</td>
<td>10.2</td>
</tr>
<tr>
<td></td>
<td>RPP30</td>
<td>PF1914</td>
<td>24,363</td>
<td>9.7</td>
</tr>
<tr>
<td><em>Mth</em></td>
<td>POP5</td>
<td>MTH0687</td>
<td>14,568</td>
<td>10.5</td>
</tr>
<tr>
<td></td>
<td>RPP21</td>
<td>MTH1618</td>
<td>16,936</td>
<td>11.1</td>
</tr>
<tr>
<td></td>
<td>RPP29</td>
<td>MTH0011</td>
<td>10,823</td>
<td>10.8</td>
</tr>
<tr>
<td></td>
<td>RPP30</td>
<td>MTH0688</td>
<td>27,662</td>
<td>8.2</td>
</tr>
<tr>
<td><em>Mja</em></td>
<td>POP5</td>
<td>MJA0494</td>
<td>15,941</td>
<td>10.0</td>
</tr>
<tr>
<td></td>
<td>RPP21</td>
<td>MJA0962</td>
<td>15,567</td>
<td>10.6</td>
</tr>
<tr>
<td></td>
<td>RPP29</td>
<td>MJA0464</td>
<td>10,888</td>
<td>9.7</td>
</tr>
<tr>
<td></td>
<td>RPP30</td>
<td>MJA1139</td>
<td>27,345</td>
<td>9.1</td>
</tr>
</tbody>
</table>
Table 4 Effect of Mja RPPs on the rate of self-processing reaction catalyzed by p^{Tyr}-S\textsubscript{3}-M RPR and p^{Tyr}-S\textsubscript{3}-ΔS M RPR conjugates. The optimum assay conditions (monovalent and divalent cations) for obtaining rate of self-cleavage of the p^{Tyr}-S\textsubscript{3}-M RPR and p^{Tyr}-S\textsubscript{3}-ΔS M RPR conjugate either in the absence or in the presence of RPPs are indicated. The cleavage assays were performed at pH 5.1, 55°C. B. To prevent self-cleavage during assembly, 25 mM Ca\textsuperscript{2+} was included in the pre-incubation step. The cleavage reaction was initiated by supplementing 100 mM Mg\textsuperscript{2+} during the assay.
**A**

<table>
<thead>
<tr>
<th>Catalytic entity tested</th>
<th>$k_{\text{obs}}, \text{min}^{-1}$</th>
<th>$\text{NH}_4^+$, M</th>
<th>$\text{Mg}^{2+}$, mM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p^{\text{Tyr}}$-$S_3$-$\text{M RPR}$ alone</td>
<td>$0.05 \pm 0.01$</td>
<td>2.5</td>
<td>500</td>
</tr>
<tr>
<td>with RPP21-RPP29</td>
<td>$0.06 \pm 0.01$</td>
<td>0.8</td>
<td>100</td>
</tr>
<tr>
<td>with POP5-RPP30</td>
<td>$4.83 \pm 0.02$</td>
<td>0.8</td>
<td>100</td>
</tr>
</tbody>
</table>

**B**

<table>
<thead>
<tr>
<th>Catalytic entity tested</th>
<th>$k_{\text{obs}}, \text{min}^{-1}$</th>
<th>$\text{NH}_4^+$, M</th>
<th>$\text{Mg}^{2+}$, mM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p^{\text{Tyr}}$-$S_3$-$\text{M RPR with all}$</td>
<td>$5.16 \pm 0.08$</td>
<td>0.8</td>
<td>100</td>
</tr>
<tr>
<td>$p^{\text{Tyr}}$-$S_3$-$\text{M RPR with all}$</td>
<td>$5.46 \pm 0.04$</td>
<td>0.8</td>
<td>100</td>
</tr>
<tr>
<td>four RPPs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p^{\text{Tyr}}$-$S_3$-$\Delta S$-$\text{M RPR}$</td>
<td>$4.86 \pm 0.06$</td>
<td>0.8</td>
<td>100</td>
</tr>
<tr>
<td>+ POP5-RPP30</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 3

IDENTIFICATION OF A MINIMAL CATALYTIC RNP CORE OF ARCHAEAL RNASE P

3.1 Introduction

Since recent studies on human and Mja RPRs (chapter 2) lead to the inference that the RPR is the catalytic subunit in all three domains of life, understanding the mechanism of RNase P-mediated catalysis requires identification of the RPR’s structural elements that are important for function. Earlier studies designed to understand structure-function relationships in the bacterial RPR were based on random deletion mutagenesis of the RPR [118]. However, the large size of bacterial RPR, together with lack of an experimentally determined tertiary structure, made it difficult to interpret results obtained from such mutagenesis studies seeking to establish a minimal structure that will sustain catalysis [16].

Secondary structures derived from phylogenetic covariation analysis of RPR sequences served as a good starting point in deciphering potentially dispensable sequences in bacterial RPRs. This is based on the assumption that sequences present in
all three domains of life are likely to be involved in catalytic function and
phylogenetically variable structures might not be essential for activity. To support this
notion, a minimal functional bacterial RPR was constructed by systematically deleting
various phylogenetically variable regions of the bacterial RPR, suggesting evolutionarily
conserved structures do relate to function [16].

Bacterial RPR is composed of two independently folding units with distinct
functions: substrate binding (S domain) and catalysis (C domain) [Fig.1.2]. Although
weak, the C domain alone is capable of independent catalysis in the presence of RPP [21].
Kinetic analysis of this minimal RNP complex indicated that it has a higher $K_m$ value
compared to the bacterial RNase P holoenzyme. The defect in substrate binding is
consistent with the data that conserved nucleotides in the S domain play a role in
substrate recognition and help precisely orient the substrate for cleavage by the C domain.

Secondary structure models based on phylogenetic covariation analyses have also
been proposed for archaeal and eukaryal RPRs. Archaeal RPRs can also be demarcated
into putative S and C domains (Figs. 1.8A and B). To test whether the C domain in
archaeal RPR supports catalysis, we constructed deletion derivatives of archaeal RPRs in
which the S domain is deleted while the highly conserved C domain is left intact (Figs.
3.1A and 3.2A). Our results demonstrate that the archaeal RPR’s C domain forms a
functional holoenzyme when reconstituted with the cognate RPPs. Interestingly, the C
domains of both type A and M RPRs are able to functionally assemble with the POP5-
RPP30 binary complex. This observation is consistent with the inference from our kinetic
studies that POP5-RPP30 (but not RPP21-RPP29) enhances the rate of the RPR-alone

84
reaction to that observed with all four RPPs. The C domain of the type M RPR represents the smallest functional RPR (143 nts) identified thus far which can functionally assemble with cognate RPPs.

Since the functional minimal core of a type M archaeal RPR is also conserved in RPRs from all three domains, we performed various heterologous reconstitutions using bacterial and eukaryal RPRs and archaeal RPPs. Results from these heterologous reconstitutions suggest that both bacterial and archaeal RPRs evolved from a common ancestor and archaeal RPPs are functionally interchangeable with the bacterial RPP in vitro.

3.2 Materials and methods

3.2.1 Cloning of the RPR deletion mutants

Cloning of cpPfu RPRΔ64-222. The gene encoding cpPfu RPRΔ64-222 was obtained using a two-step cloning approach. First, the gene encoding Pfu RPR was amplified by PCR using primers CP-F and CP-R (Table 5) and cloned into the EcoRI and PstI sites of pBT7-Pfu RPR to generate pBT7-(Pfu RPR)2, in which two copies of the RPR gene are present in tandem. The coding sequence for Pfu RPRΔ64-222 was amplified by PCR using pBT7-(Pfu RPR)2 as template and RPR 225-F and RPR 63-R as primers (Table 5). This PCR product was digested with BamHI (whose recognition site was included in RPR 63-R, the reverse primer) and ligated to pBT7 that had been digested with Stul (which generates a blunt end) and BamHI. The resulting plasmid, which we named pBT7-Pfu RPRΔ64-222, was linearized with FokI (whose recognition site was included
in RPR 63-R, the reverse primer) and used as the template in an *in vitro* transcription reaction to generate *Pfu* RPRΔ64-222.

**Cloning of *Mja* RPRΔ66-190.** The gene encoding *Mja* RPRΔ66-190 was generated using a PCR-based deletion mutagenesis approach using primers MjDSMnpT-F and MjDSMnpT-R (Table 5) and plasmid pBT7-*Mja* RPR as the template (Gopalan, unpublished work).

**Cloning of *Reclinomonas americana mitochondrial* (Ram-mt) RPR**

The gene encoding Ram-mt RPR was PCR amplified using primers RAM-F and RAM-R and plasmid (kind gift from Dr. Elias Seif, Franz Lang laboratory) as template. The PCR product was digested with *Hind*III and ligated to pBT7 to obtain pBT7-Ram-mt RPR (Gopalan, unpublished work; Table 8). Ram-mt RPR was generated by run off transcription reaction using T7 RNA polymerase and linearized pBT7-Ram-mt RPR as template.

The sequences of the clones described above was verified by using automated DNA sequencing at the Ohio State University Plant-Microbe Genomics Facility.

**3.2.2 Cloning, over-expression, and purification of *Pfu* RPPs and *Mja* RPPS**

Detailed description of the cloning of the genes encoding *Pfu* RPPs, over-expression of *Pfu* RPPs in *E. coli* and their purification are described elsewhere [69]. *Mja* RPPs were obtained as described in Chapter 2.

**3.2.3 Generation of RPRs using in vitro transcription**

The plasmids pBT7-*Pfu* RPR, pBT7-*Mja* RPR, pBT7-*Mth* RPR, pBT7-*Afu* RPR, pBT7-Ram-mt RPR, pBT7- *Pfu* RPRΔ64-222, and pBT7-*Mja* RPRΔ66-190 were linearized
with EcoRI, BsmAI, BamHI, HindIII, HindIII, FokI and Bsma I, respectively, and used as templates for T7 RNA polymerase-mediated run-off transcription using established protocols [69]. The RNA transcripts thus generated were subsequently subjected to dialysis to remove unincorporated rNTPs and their concentrations determined from A260 measurements. RPR (in water) was folded by incubating for 50 min at 50°C, 10 min at 37°C, and then for 30 min at 37°C in a Mg$^{2+}$-containing buffer [50 mM Tris·HCl (pH 7.5), 10 mM Mg(OAc)$_2$ and 800 mM NH$_4$OAc].

### 3.2.4 RNase P activity assays

Reconstitution experiments were initiated by pre-incubating folded Pfu RPRΔ64-222 RPR (500 nM) with either all or a subset of four RPPs (1 µM) in assay buffer [50 mM Tris·HCl (pH 7.5), 400 mM NH$_4$OAc and 120 mM Mg(OAc)$_2$] for 5 min at 37°C followed by 10 min at 55°C. The activity was assayed by adding 1 µM E. coli ptRNA$^{Tyr}$, a trace amount of which was internally labeled with [$\alpha$-$^{32}$P] GTP. After a 60 min incubation at 55°C, the reactions were quenched with urea-phenol dye [8 M urea, 0.04% (w/v) bromophenol blue, xylene cyanol, 0.8 mM EDTA, 20% (v/v) phenol].

Reconstitution experiments involving Mja RPRΔ66-190 and Mja RPPs were performed by combining Mja RPRΔ66-190 (500 nM) with either RPP21-RPP29 or POP5-RPP30 (1 µM) in assay buffer [50 mM Tris·HCl (pH 7.5), 400 mM NH$_4$OAc and 120 mM Mg(OAc)$_2$] followed by successive incubations (5 min each) at 37°C and 55°C. The activity was assayed by using 1 µM E. coli ptRNA$^{Tyr}$ as substrate. Mja RPRΔ66-190 was reconstituted with all four RPPs essentially as described above for the
binary complexes, except that the reconstitution buffer contained 30 mM Mg\(^{2+}\) instead of 120 mM.

3.2.5 Heterologous reconstitution of RNase P holoenzymes

Heterologous reconstitutions of RNase P holoenzymes using archaeal type A RPRs and type M RPPs were achieved by individually mixing *Pfu* and *Mth* RPRs (50 nM) with *Mja* RPPs (250 nM) in reconstitution buffer [50 mM Tris·HCl (pH 7.5), 800 mM NH\(_4\)OAc and 30 mM Mg(OAc)$_2$]. The RNP mixture was incubated for 5 min at 37°C followed by 10 min at 55°C. Similarly, the reconstitution of type M RPRs with type A RPPs were achieved by incubating *Afu* and *Mja* RPRs (50 nM) separately with *Pfu* RPPs (250 nM) in 50 mM Tris·HCl (pH 7.5), 800 mM NH\(_4\)OAc and 30 mM Mg(OAc)$_2$. RNase P assays were performed using 500 nM *E. coli* ptRNA\(^{Tyr}\) at 55°C 10 to 15 min.

Heterologous reconstitutions involving bacterial or organellar RPR and archaeal RPPs were performed by individually combining *E. coli*, *B. subtilis* RPR or *Ram* mt RPR (50 nM) and archaeal RPPs (500 nM) in 50 mM Tris·HCl (pH 7.5), 800 mM NH\(_4\)OAc and 30 mM Mg(OAc)$_2$. The mixture was pre-incubated at 37°C for 5 min followed by 5 min at 55°C. RNase P assays were performed at 55°C using 200 nM *E. coli* ptRNA\(^{Tyr}\) a trace amount of which was internally labeled with [\(\alpha^{32}\)P] GTP substrate. To prevent evaporation losses and aid in rapid temperature equilibration, assays were performed in 0.5-ml thin-walled PCR tubes and a thermal cycler (Eppendorf, Westbury, NY).

Uncleaved ptRNA\(^{Tyr}\) and products generated by RNase P were separated by 8% (w/v) polyacrylamide/7 M urea gel electrophoresis and bands were visualized by autoradiography.
3.2.6 Use of RNase T1-based footprinting to map RNA-protein interactions in archaeal RNase P

Each 20-µl footprinting reaction contained a mix of 5' labeled (100,000 dpm, ~500 pM) and unlabeled (500 nM) RPR either alone or complexed with RPP21-RPP29 or POP5-RPP30 in 50 mM Tris-acetate (pH 7.5), 120 mM Mg(OAc)$_2$, 400 mM NH$_4$(OAc). The reaction mix was preincubated for 10 min at 37°C followed by 10 min at 55°C. RNase T1 (0.002 U, Ambion) was added to the reconstitution mix and the incubation was continued at 55°C for an additional 4 min. The reaction was terminated by adding 10 µl of buffer-saturated phenol (pH 8) followed by extraction with phenol/chloroform. The RNA was precipitated by adding 2 vol of ethanol in the presence of 0.3 M sodium acetate (pH 5.2) and 20 µg/ml glycogen. The RNA sample was pelleted at 18,000g for 15 min in a microcentrifuge and the pellet was washed twice with 70% (v/v) ethanol. The air-dried RNA sample was resuspended in 10 µl loading dye [9 M urea/0.9 mM EDTA/0.05% (w/v) bromophenol blue/0.05% (w/v) xylene cyanol/10% (v/v) phenol] and separated by 8% (w/v) polyacrylamide/7M urea gel electrophoresis. The bands were visualized by using a PhosphorImager (Molecular Dynamics).

To map the RNA-protein interactions in *Mja* RNase P, *in vitro* reconstituted *Mja* RNase P was subjected to limited digestion with RNase T1 as described above. The reaction contents were precipitated, washed with 70% (v/v) ethanol, air-dried and resuspended in water. The oligonucleotide *Mja*RPR-R (Table 1), which is complementary to the 3’-end of *Mja* RPR was 5’-labeled with [γ-32P]-ATP and T4 polynucleotide kinase and used to prime the reverse transcription of the partial digestion
products of \textit{Mja} RPR. To minimize artificial stops caused by the secondary structure of the tRNA, the extension reactions were performed at 50°C using ThermoScript (Invitrogen) reverse transcriptase as specified by the supplier. The products of the reverse transcription reactions were separated in an 8% (w/vol) polyacrylamide/7M urea gel in parallel with a DNA sequencing ladder using MjaRPR-R as the primer and the plasmid pBT7-\textit{Mja} RPR as the template.

3.3 Results

3.3.1 Identification of the catalytic domain of archaeal RPR.

The secondary structure model obtained from phylogenetic covariation analysis of many archaeal RPR sequences identified the existence of conserved nucleotides that might constitute its catalytic core [5, 63, 119]. We decided to test whether this conserved core in archaeal RPR is sufficient to support catalysis by using a \textit{Pfu} RPR deletion derivative, whose design was based on inferences from bacterial RNase P studies [19, 21, 118].

Bacterial RPRs are composed of two independently folding units: the specificity (S) and catalytic (C) domains, with the latter capable of independent catalysis either alone or in the presence of the RPP [19, 21, 120]. The nucleotides in the S domain of bacterial RPR that recognize the T stem-loop of the ptRNA substrate were identified by an innovative circular permutation analysis coupled with dephosphorylation strategy [18] and by footprinting/photo-crosslinking studies [30]. Recent structural studies indicate that these interactions might play a role in precisely orienting the substrate for cleavage by the C domain [28, 29]. In the secondary structure model of \textit{Pfu} RPR, the P7-P10
four-way junction and adjacent loop regions (Fig. 1.2A and 1.8A) resemble the S
domain of bacterial RPR and suggests a similar function in ptRNA binding but not
cleavage. We therefore constructed \textit{Pfu} RPR Δ64-222 (Fig. 3.1A), a circularly permuted
deletion derivative of \textit{Pfu} RPR (i.e., Cp223-330 + 1-63), in which the S domain (from
nucleotides 64 to 222) is deleted while leaving intact the highly conserved C domain.
Indeed, \textit{Pfu} RPRΔ64-222, bereft of the S domain, forms a functional holoenzyme when
reconstituted with the \textit{Pfu} RPPs (Fig. 3.1B, lane 12), although its turnover number is 40-
fold lower than the holoenzyme reconstituted with \textit{Pfu} RPR (data not shown); \textit{Pfu}
RPRΔ64-222 was assayed for 40 min at 55°C in 50 mM Tris-acetate (pH 7.5), 120 mM
Mg(OAc)$_2$ and 400 mM NH$_4$OAc. This deletion derivative either alone or in
combination with a single RPP did not display any activity (data not shown). Because it
has been already established that the full-length \textit{Pfu} RPR could assemble with POP5-
RPP30 or RPP21-RPP29 to form a functional (although partial) RNase P holoenzyme
[69], we investigated how \textit{Pfu} RPRΔ64-222 would fare with these RPP pairs. We found
that \textit{Pfu} RPRΔ64-222 was able to functionally assemble only with the POP5-RPP30
combination (Fig.3.1B, lane 4), and that this activity remained unchanged upon addition
of RPP21-RPP29 (Fig. 3.1B, lane 4 versus 12).

3.3.2 Identification of the catalytic domain of a type M archaeal RPR

Based on the occurrence of secondary structure elements, most of the euryarchaeal
RPRs can be categorized into two well-defined structural classes: type A [represented by
\textit{Pfu} RPR (Fig.1.8A)] and type M [represented by \textit{Mja} RPR (Fig. 1.8B)]. Although the C
domain of type M RPR exhibits an apparent reduction in RNA structure, the C domains
from both type A and M RPRs show remarkable similarities to the C domain of the ancestral bacterial type A RPR with respect to the position and identity of the conserved nucleotides. Note that the highly conserved CR I-V and CR IV are retained in the C domains of both type A and M RPRs. However, helices P6, P16, and P17 and the loop regions connecting the P15 and P16 (L15), implicated in interaction with the 3’-ACCA of the ptRNA substrate, are absent in C domains of type M RPRs. Therefore, we investigated whether the C domain of a type M RPR can functionally assemble with cognate RPPs akin to C domain of a type A RPR (Pfu). To this end, we constructed a deletion derivative of Mja RPR (Mja RPRΔ66-190) in which the nucleotides 66 through 190 in the primary sequence are deleted without altering nucleotides in the putative C domain. Indeed, Mja RPRΔ66-190 was able to reconstitute with POP5-RPP30 (Fig 3.2B, lane 6). Interestingly, under identical assay conditions, we did not observe any noticeable RNase P activity when RPP21-RPP29 was incubated with Mja RPRΔ66-190 (Fig. 3.2B, lane 7). This result parallels that observed with Pfu RPRΔ64-222 + RPP21-RPP29 and is also consistent with the observation that POP5-RPP30 (and not RPP21-RPP29) was able to enhance the self-processing of the pTyr-S3-ΔS M RPR cis-conjugate (described in Chapter 2). However, unlike Pfu RPRΔ64-222, addition of RPP21-RPP29 to Mja RPRΔ66-190 + POP5-RPP30 resulted in significant enhancement in activity of Mja RPRΔ66-190 + POP5-RPP30 at 30 mM Mg2+ concentration. This observation is in agreement with the result obtained from the trans cleavage of ptRNA by the full length Mja RPR with all four RPPs, wherein addition of RPP21-RPP29 to Mja RPR + POP5-RPP30 enhanced its catalytic potential while reducing the Mg2+ requirement for optimal
activity (Chapter 2). By demonstrating that *Mja* RPRΔ66-190 can generate a functional RNase P holoenzyme with POP5-RPP30, we identified the smallest RPR (143 nts) that assemble with RPPs to generate a functional RNase P holoenzyme.

### 3.3.3 Mapping the binding sites of *Pfu* POP5-RPP30 complex on the C domain of *Pfu* RPR

To better understand how POP5-RPP30 supports catalysis by the archaeal RPR’s C domain, knowledge of the spatial organization of these RPPs on the cognate RPR is necessary. Because RNase T1 cleaves 3' to unpaired guanines, an examination of the RNase T1-digestion patterns of *Pfu* RPR in the absence and presence of the RPPs is expected to reveal at least a few RNA-protein interaction sites in *Pfu* RNase P. Although technical problems complicated footprinting experiments with the full-length *Pfu* RPR, the smaller C-domain-containing *Pfu* RPRΔ64-222 is better behaved in solution and was used in the mapping experiments reported here. Partial digestion of *Pfu* RPRΔ64-222 with RNase T1 was carried out at 55°C under conditions determined to be optimal for *Pfu* RNase P activity. Upon binding to POP5-RPP30, but not RPP21-RPP29, nucleotides in L3, P4, L15, J15/2 and J2/4 of *Pfu* RPR displayed decreased susceptibility to RNase T1 (Fig. 3.3, lanes 3 and 5). In contrast to the 3' strand of L15, Gs in the 5' strand are not protected. Although P6 seems to be protected, it is more likely that it is metastable in the absence of RPPs and that it becomes double-stranded in the presence of POP5-RPP30 (Fig. 3.3B, lanes 1 and 5). In the absence of RNase T1, we observed that certain regions in *Pfu* RPR become more susceptible to cleavage (likely Mg²⁺-induced) upon binding to POP5-RPP30 (Fig. 3.3, lane 6 versus lane 2). Interestingly, these positions are proximal
to the nucleotides that are protected from cleavage by RNase T1 (Fig. 3.3A and B). Because RPPs lower the Mg$^{2+}$ requirement of the RPR [69], changes in metal-ion binding at or around the RPP-binding sites are likely and these cleavages might reflect the same.

Results from phosphorothioate and hydroxyl radical-mediated footprinting experiments indicate that the bacterial RPP is proximal to P3, P4, J18/2, and J2/4 (located in conserved regions, CR I, IV and V) in its cognate RPR [51, 56] akin to the footprint of \textit{Pfu} POP5+RPP30 on \textit{Pfu} RPR. The conserved nucleotides in these regions have been shown by different biochemical approaches to participate in substrate and Mg$^{2+}$ binding [103]. In three-dimensional models of the bacterial RPR/RNase P-ptRNA complex [28, 29, 51, 57], these regions are proximal to the ptRNA acceptor stem and cleavage site.

Similar RNase T1 footprinting analysis of the \textit{Mja} RNase P holoenzyme indicates that binary complex consisting of POP5-RPP30 binds predominantly to the C domain (Fig.3.5A). Interestingly, positions in the \textit{Mja} and \textit{Pfu} C domains that are protected from RNase T1 digestion in the presence of POP5-RPP30 are remarkably similar in both cases suggesting that POP5-RPP30 interacts with conserved structural elements in both RPR.

**3.3.4 Heterologous reconstitution of RNase P using RPR and RPPs**

The observation that the C domain of \textit{Mja} RPR, which retains only the universally conserved catalytic core (that is also common to all bacterial and eukaryal RPRs) was able to form a functional complex with cognate RPPs prompted us to test whether archaeal RPPs will be able to heterologously reconstitute with bacterial and eukaryal
RPRs. Before we embarked on such heterologous reconstitutions involving archaeal RPPs and RPRs from different domains of life, we first tested whether we can establish heterologous reconstitution within archaea. Indeed we observed successful heterologous reconstitution, i.e., mixing type A RPR with type M RPPs and vice-versa, indicating the presence of a conserved RNP core in all archaeal RNase P holoenzymes (Fig. 3.5).
Since type A and type M RPRs reconstituted far more effectively with their cognate RPPs, the RPPs must have co-evolved with their RPRs as they display type-preferential reconstitution patterns.

Under optimal conditions for reconstitution, both *E. coli* and *B. subtilis* RPRs reconstituted with both types A and M archaeal RPPs to generate functional RNase P holoenzymes (Fig. 3.6 and Table 6 and 7). To examine whether all four protein subunits are needed for promoting bacterial RPR catalysis, we tested the functional reconstitution of bacterial RPR with either of the two binary complexes of archaeal RPPs (POP5-RPP30 or RPP21-RPP29). We observed that the binary complex consisting of archaeal POP5-RPP30 is sufficient to enhance the activity of *E. coli* RPR (bacterial type A RPR) and *B. subtilis* RPR (bacterial type B) by at least 6- and 10- fold, respectively (Table 6 and 7). The reason why both type A and M archaeal POP5-RPP30 enhanced the catalytic potential of of a type B bacterial RPR to a greater extent than than type A bacterial RPR is currently unclear. Note that in our heterologous reconstitution assays, POP5 alone was not sufficient to confer any significant increase in the bacterial RPR’s ability to promote catalysis (Fig. 3.6, lane 4). This observation is interesting considering the remarkable tertiary structure similarity of POP5 and bacterial RPP [74]. Since RPP21-RPP29 has no
significant effect on bacterial RPR’s activity, the results from our in vitro heterologous reconstitution involving bacterial RPR and archaeal RPPs suggest that POP5-RPP30 is functionally equivalent to the bacterial RPP.

To test whether the eukaryal RPR will be able to functionally assemble with archaeal RPPs, we examined whether heterologous reconstitutions involving yeast or human RPR and archaeal RPPs will result in the formation of a functional holoenzyme. However, these efforts were unsuccessful (data not shown). We then assessed whether archaeal RPPs could support organellar RPR catalysis. We focused mainly on the plastid (cyanelle)-encoded RPR of *Cyanophora paradoxa* (*Cpa*) and mitochondria (mt)-encoded RPR of *Reclinomonas americana* (*Ram*), a jakobid flagellate. While cyanelle RPR shows remarkable overall similarity to the bacterial RPRs, *Ram*-mt RPR represents a highly reduced version of the bacterial RPR lacking P6, P16, P17 and L15 (Fig. 3.7A).

Ribozyme activity has been demonstrated for cyanelle RPR. However, *Ram* mtRPR was found to be catalytically inactive under various conditions tested in vitro. Interestingly, archaeal RPPs heterologously reconstitute with *Ram* mtRPR but not with the *Cpa* RPR (Fig. 3.7B).

### 3.4 Discussion

#### 3.4.1 Identification of the catalytic domain of a type M archaeal RPR

Presence of conserved structural elements resembling that of the bacterial consensus in all archaeal counterparts suggests a common evolutionary origin for the RPR subunit and even the existence of a shared catalytic core in all RPRs [13, 63, 121]. Deletion mutants
of the *E. coli* RPR (e.g., Δ94-204 or Δ87-241; [21, 118]) that retain the ‘catalytic’ domain (P1 to P6 and P15 to P18) were catalytically active in the presence of the protein cofactor, albeit diminished in activity and with altered substrate specificity in some mutants. In a similar vein, *Pfu* RPRΔ64-222, a mutant derivative bereft of the ‘specificity’ domain is functional with POP5-RPP30 (Fig. 3.1B). This observation suggests that the active site resides in the postulated ‘catalytic’ domain of *Pfu* RNase P RNA and establishes an important, if largely anticipated, parallel between bacterial and archaeal RNase P.

We have also extended this finding to a type M RNase P. Indeed, similar to *Pfu* RPRΔ64-222, the C domain of *Mja* RPR, is able to support catalysis as borne out by the ability to reconstitute *Mja* RPRΔ66-190 with POP5-RPP30 but not with RPP21-RPP29 (Fig. 3.2). Interestingly, unlike *Pfu* RPRΔ64-222, addition of all RPPs to *Mja* RPRΔ66-190 resulted in the formation of an RNP complex with significantly lower requirement of Mg\(^{2+}\) for optimum activity (30 mM Mg\(^{2+}\) as compared to 120 mM Mg\(^{2+}\) for the *Mja* RPRΔ66-190 + POP5-RPP30). However, we did not observe significant protection of nucleotides in the C domain, when *Mja* RPR + RPP21-RPP29 was subjected to partial digestion with RNase T1. Furthermore, RPP21-RPP29 was not able to enhance the activity of a cis cleaving ptRNA-*Mja* RPR construct lacking S-domain in which ptRNA\(_{Tyr}\) is attached to the C domain of *Mja* RPR (pt\(_{Tyr}\)-S\(_3\)-ΔS M RPR). Based on these results, we speculate that interaction of RPP21-RPP29 with the *Mja* RPR’s C domain is likely to be indirect, i.e., via protein-protein interactions with POP5-RPP30.

Our finding that the C domain of *Mja* RPR is able to reconstitute with POP5-RPP30 led to identification of the smallest functional RPR after assembling with RPPs in
vitro (143 nts, Fig. 3.2A). Although many cis cleaving small ribozymes are known, this mini-\textit{Mja} RNase P holoenzyme can process ptRNA in \textit{trans} with robust multiple turnover. We believe that this new information on pared-down versions will aid high-resolution structural studies of archaeal RNase P.

An EDTA-Fe-based footprinting approach to map the RNA-protein contact sites in bacterial RNase P revealed that the C5 protein is proximal to P3 and the universally conserved nucleotides in J2/4, J3/4 [51]. The protection pattern observed with \textit{Pfu} Cp RPRΔ64-222 + POP5-RPP30 and \textit{Mja} RPR + POP5-RPP30 upon partial digestion with RNase T1 indicates that POP5-RPP30 is proximal to the universally conserved nucleotides near the active site in a manner similar to that of bacterial RNase P holoenzyme.

\textbf{3.4.2 Heterologous reconstitution of RNase P using RPR and RPPs}

Heterologous reconstitution studies using RPRs and RPPs from various bacterial clades demonstrates that bacterial RNase P is derived from a common ancestor since a functional catalytic core could be generated from unrelated RPRs and RPPs which differ in sequence and structure. Results from heterologous reconstitutions using archaeal RPRs and RPPs that differ significantly in sequence and structure (reconstitution of type A RPRs with type M RPPs and vice-versa) suggest that archaeal RPPs are also functionally interchangeable in vitro similar to their bacterial counterparts. Our findings that archaeal RPPs could facilitate bacterial RPR catalysis \textit{in vitro} albeit significantly weaker than the bacterial RPP, suggest that archaeal RPRs evolved from a common ancestor RNA which resembles the extant bacterial RPR. The observation that POP5-RPP30 and (not RPP21-
RPP29) enhances the catalytic potential of bacterial RPR once again highlights the role of these RPPs in catalysis as revealed from our kinetic studies (Chapter 2).

Based on a genetic complementation approach that utilizes the suppression of a conditional lethality of \textit{B. subtilis} RPP, Hartmann and coworkers recently concluded that both type A and M archaeal RPPs (expressed either as individual subunits or as RPP21-RPP29, POP5-RPP30 binary complexes) were not able to form a functional RNase P holoenzyme with \textit{B. subtilis} RPR \textit{in vivo} [122]. The results from similar \textit{in vitro} heterologous reconstitution performed in this study suggest that although archaeal RPPs could functionally replace the bacterial RPP, the resulting RNP complex exhibits at least a 60-fold decrease in activity compared to bacterial RNase P, perhaps accounting for their inability to support the ptRNA processing reaction \textit{in vivo}.

Although we have not exhaustively optimized reaction conditions and folding procedures for the RPR, two eukaryal nuclear RPRs (yeast and human) tested under many conditions were unable to form a functional holoenzyme with archaeal RPPs. On the other hand, \textit{Ram}-mt RPR with limited overall structural resemblance with archaeal RPR (except for the conserved catalytic core) reconstituted with archaeal RPPs suggesting a common evolutionary origin for these two RPRs.

Heterologous reconstitution results presented in this study establish a common evolutionary origin for bacterial and archaeal RPR and functional interchangeability of a pair of archaeal RPPs with the sole bacterial RPP. Although the secondary structure model derived from phylogenetic covariation analysis of many eukaryal RPRs identified a conserved structural core similar to archaeal RPR, failure to generate a functional
RNase P with archaeal RPPs suggest that eukaryal RPRs and their cognate RPPs might have diverged significantly from their bacterial and archaeal counterparts.
Figure 3.1 Identification of a pared down, mini-\textit{Pfu} RNase P. A. Secondary structure model of \textit{Pfu} RPRΔ64-222 based on comparative sequence analysis. The lowercase letters at the 5' and 3' ends represent nucleotides not present in the native \textit{Pfu} RPR. Paired helices (P1, P2, etc.) are numbered according to the \textit{E. coli} RPR nomenclature. B. Functional reconstitution of \textit{Pfu} RPRΔ64-222 with \textit{Pfu} RPPs.
Figure 3.2 Identification of a minimal catalytic RNP core of *Mja* RNase P. A.

Secondary structure model of *Mja* RPA66-190. B. Reconstitution of RNase P activity using *Mja* RPA66-190 and various combinations of *Mja* RPPs as indicated.
**Figure 3.3 Mapping the binding sites of POP5-RPP30 complex on the catalytic domain of *Pfu* RPR.**

A. Secondary structure representation of *Pfu* RPRΔ64-222. Unpaired Gs that are protected from RNase T1 are indicated with filled arrows. Asterisks indicate sites of Mg$^{2+}$-induced cleavage that occur upon binding to POP5-RPP30. B. RNase T1-footprinting analysis of *Pfu* RPRΔ64-222. 5'-end-labeled *Pfu* RPRΔ64-222 either alone, with RPP21-RPP29, or with POP5-RPP30 was incubated in the presence (lanes 1, 3, and 5, respectively) or absence (lanes 2, 4, and 6, respectively) of RNase T1. Alk. and T1 represent molecular size ladders generated by subjecting the 5'-end-labeled *Pfu* RPRΔ64-222 to alkaline hydrolysis and partial RNase T1 digestion under denaturing conditions, respectively.
Figure 3.4 Mapping the binding sites of POP5-RPP30 complex on the *Mja* RPR. A. *Mja* RPR alone (lane 2) or in the presence of POP5-RPP30 (lane 3) was subjected to partial digestion with RNase T1. Digestion products were subjected to primer extension analysis using 5′-end-labelled MjaRPR-R primer and reverse transcriptase (RT) and the cDNAs thus generated were separated in a sequencing gel. Lanes 1 and 3 represent the cDNA products generated from control experiments in which the RPR and RPR + POP5-RPP30 were incubated in the absence of RNase T1, respectively. G and C depict size markers in which pBT7-Mja RPR was subjected to DNA sequencing using the MjaRPR-R primer and dNTPs in which dideoxyguanine and dideoxycytidine, respectively, were used as chain terminating agents. B. Secondary structure representation of *Mja* RPR. Arrowheads indicate the locations where the band intensity is reduced in the presence of RPPs and indicates sites of RNA-protein interactions.
Figure 3.5 Heterologous reconstitution of RNase P holoenzymes using various archaeal RPRs and RPPs. Archaeal RPRs representing type A (Pfu and Mth) and type M (Mja and Afu) were reconstituted with type A (Pfu) RPPs (panel A) type M (Mja) RPPs (panel B).
Figure 3.6 Heterologous reconstitution of an RNase P holoenzyme using a bacterial RPR and archaeal RPPs. *E. coli* RPR was reconstituted with *Pfu* RPPs either singly or in certain combinations as indicated.
Figure 3.7 Heterologous reconstitution of an RNase P holoenzyme using *Reclinomonas americana* RPR (*Ram*-mt RPR) and archaeal RPPs. A. Secondary structure representation of *Ram* mitochondrial RPR. B. Reconstitution of RNase P using *Ram*-mt RPR and *Mja* RPPs.
<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP-F</td>
<td>5'-AAACTGCAGTAGGCAGGGGGCTGGGGGCTG-3'</td>
</tr>
<tr>
<td>CP-R</td>
<td>5'-GGGGATCCTAGGCGAGGGGCTATAGCCGC-3'</td>
</tr>
<tr>
<td>RPR 225-F</td>
<td>5'-TGCAAGGCGAGTTAGGGCCGATG-3'</td>
</tr>
<tr>
<td>RPR 63-R</td>
<td>5'-GGGGATCCCGGATGGACGTCTCTGCGTGTCGGTGGCCGGAAG-3'</td>
</tr>
<tr>
<td>MjDSMnpT-F</td>
<td>5'-AAA GGA GGA AGT TCC GCC CAC CCA TAG GGT GCA AGC CGA AAT AG-3'</td>
</tr>
<tr>
<td>MjDSMnpT-R</td>
<td>5'-CGG AAA GTC ACC AGC CCC CTC TCG GGG GCT ATA GCC CG-3'</td>
</tr>
<tr>
<td>RAM-F</td>
<td>5'-ATA AAA GTT TAT TGG ATG TCT GAT TAT AAT AAT GAT G-3'</td>
</tr>
<tr>
<td>RAM-R</td>
<td>5'-GCG AAT TCT TTA AAG TTT ATT TAT AAG CTG GAT TTT GTC C-3'</td>
</tr>
</tbody>
</table>

Table 5 Oligonucleotides used in this study for cloning various archaeal RPRs or their derivatives
Table 6 Heterologous reconstitution of *E. coli* RPR with archaeal RPPs.

Comparison of turnover numbers of chimeric enzymes in which *E. coli* RPR was reconstituted with either *E. coli* RPP or archaeal RPPs. Average values from three independent experiments are provided. The assays were performed at 55°C and pH 7.5.

<table>
<thead>
<tr>
<th>Composition</th>
<th>Turnover number (min⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> RPR</td>
<td>0.16 ± 0.02</td>
</tr>
<tr>
<td><em>E. coli</em> RPR + <em>E. coli</em> RPP</td>
<td>50 ± 5</td>
</tr>
<tr>
<td><em>E. coli</em> RPR + <em>Pfu</em> POP5-RPP30</td>
<td>1.3 ± 0.2</td>
</tr>
<tr>
<td><em>E. coli</em> RPR + all four <em>Pfu</em> RPPs</td>
<td>0.6 ± 0.23</td>
</tr>
<tr>
<td><em>E. coli</em> RPR + <em>Mja</em> POP5-RPP30</td>
<td>0.45 ± 0.15</td>
</tr>
<tr>
<td><em>E. coli</em> RPR + all four <em>Mja</em> RPPs</td>
<td>0.18 ± 0.04</td>
</tr>
<tr>
<td></td>
<td>Turnover number (min⁻¹)</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td><em>B. subtilis</em> RPR</td>
<td>0.06 ± 0.02</td>
</tr>
<tr>
<td><em>B. subtilis</em> RPR + <em>E.coli</em> RPP</td>
<td>48 ± 2</td>
</tr>
<tr>
<td><em>B. subtilis</em> RPR + <em>Pfu</em> POP5-RPP30</td>
<td>0.91 ± 0.18</td>
</tr>
<tr>
<td><em>B. subtilis</em> RPR + all four <em>Pfu</em> RPPs</td>
<td>0.58 ± 0.12</td>
</tr>
<tr>
<td><em>B. subtilis</em> RPR + <em>Mja</em> POP5-RPP30</td>
<td>0.62 ± 0.1</td>
</tr>
<tr>
<td><em>B. subtilis</em> RPR + all four <em>Mja</em> RPPs</td>
<td>0.26 ± 0.06</td>
</tr>
</tbody>
</table>

Table 7 Heterologous reconstitution of *B. subtilis* RPR with archaeal RPPs.

Comparison of turnover numbers of chimeric enzymes in which the *B. subtilis* RPR was reconstituted with either *E. coli* RPP or archaeal RPPs. Average values from three independent experiments are provided.
CHAPTER 4

DISCUSSION

4.1 Introduction

The RNA world refers to a hypothetical intermediate stage in the early evolution of life on earth when RNAs acted as both the genetic material (information storage) and the catalytic agents (enzymes). The extant DNA and protein world is postulated to have evolved from this primitive RNA world. DNA presumably took over the role of information storage because of its greater chemical stability while proteins became the specialized molecules suitable for biological catalysis due to their ability to employ a variety of amino acid functional groups ideal for catalysis at neutral pH. Catalytic RNAs (ribozymes) in extant life forms are evolutionary remnants of the RNA world and provide a unique opportunity to understand the evolution of macromolecular biological catalysis. Comparison of the mechanistic aspects of ribozyme and protein-catalyzed reactions will provide valuable insights into the limitations of RNA catalysis and the evolutionary pressure for the transition from RNA to the contemporary protein world.
Despite their catalytic potential, RNAs do not function as catalysts in the extant life forms but typically function together with protein cofactors for efficient function in vivo. In most cases, the specific contributions of individual RNA and protein subunits in assembly and function of such RNP complexes have not been fully understood. In the following section, I will compare the intrinsic catalytic prowess of RNA and protein enzymes to highlight the limitations of catalytic RNAs. I will then summarize the role of protein cofactor(s) in facilitating RNA catalysis using the knowledge gained from the most well characterized RNP complexes.

4.2 Biological catalysis - lessons from ribozymes and RNPzymes

Until twenty five years ago, it was generally accepted that biological catalysis was restricted to proteins. The discovery that RNAs could also catalyze chemical reactions resulted in a paradigm shift in biology [10, 123]. However, the lack of a wide repertoire of versatile functional groups in RNA, together with a seemingly less defined active site that might be expected from inefficient packing imposed by the high charge density and flexibility of the phosphodiester backbone, cast doubts on the likelihood of RNA serving as biological catalysts. Results from various biochemical and structural studies performed in the next two decades provided valuable mechanistic insights into ribozyme-catalyzed reactions [124]. Furthermore, in vitro selection of functional RNAs from a large pool of random sequences identified new RNA molecules with desired catalytic properties including nucleotide synthesis [125], RNA polymerization [126], aminoacylation of transfer RNA [127] and peptide bond formation [128]. Finally, the realization that an RNA-based active site for the peptidyl transfer reaction catalyzed by the ribosome
together with the recent discovery of many riboswitches that use an RNA-based mechanism to sense various metabolites to regulate gene expression (in response to changes in internal and external stimuli) have lead to a better appreciation of the remarkable structural and functional versatility of RNA.

Most naturally occurring ribozymes catalyze transesterification reactions involving the phosphodiester linkage. Like their protein counterparts, catalytic RNAs in principle could make use of a variety of strategies to accelerate a chemical reaction at the phosphoryl center. These include (i) general acid-base catalysis, (ii) use of intrinsic binding energy for catalysis, (iii) electrostatic destabilization of the ground state and stabilization of the negative charge build up in the transition state, and (iv) precise positioning and orientation of the reactive groups in the active site [129, 131].

It has been suggested that the limitations of RNAs as catalysts might stem from their inability to perturb the pKₐ of the nucleobase functional groups and utilize them effectively in general acid-base catalysis. However, catalytic RNAs are adept at utilizing metal ions as cofactors thus compensating at least in part for the lack of chemical versatility in their ribonucleobase building blocks. An RNA-based active site then could be imagined as consisting of a structural scaffold that appropriately positions the active site metal ions for use in one of the above mentioned catalytic strategies [130]. However, conformational rigidity of the active site is suboptimal for RNA because of the flexibility and high charge density of the phosphodiester backbone. Large ribozymes achieve rigidity by RNA-RNA interactions within the core and by employing long-range interactions with peripheral RNA structural elements [131]. One well characterized
example is the peripheral RNA structure, P5abc, which stabilizes the catalytic core of *Tetrahymena* group I intron. In addition, large RNA molecules fold into compact, ordered structures with precisely positioned functional groups serving as ligands for a multi-dentate metal ion to limit the active site flexibility. Therefore, it is not surprising that large ribozymes such as group I intron from *Tetrahymena* and the RNA subunit of bacterial RNase P can catalyze self splicing and ptRNA processing reactions, respectively, with a rate constant for the chemical cleavage step $\sim 200 \text{ min}^{-1}$. It should be emphasized that this rate represents a $10^{11}$-fold enhancement compared to the uncatalyzed reaction in solution and is comparable to the rate enhancement achieved by many protein enzymes although significantly less than the upper limit reported for protein enzymes ($10^{17}$-fold) [131].

4.3 Need for protein cofactors: tertiary structure capture of the functional conformation in the RNA?

Although a significant rate enhancement over the uncatalyzed reaction has been observed for many RNA-based enzymes, all *trans*-acting catalytic RNAs require protein cofactors for function *in vivo*. Extensive studies using simple catalytic RNPs/RNPzymes like the group I intron and bacterial RNase P have contributed significantly towards our understanding of how proteins modulate RNA function in an RNP complex. For example, although the terminal intron of yeast mitochondrial gene encoding cytochrome *b* (*COB*) pre-mRNA (*bI5, group I intron*) can self-splice *in vitro* at high concentrations of Mg$^{2+}$ [132], the protein encoded by the nuclear gene *CBP2* is essential for splicing at physiological concentrations of Mg$^{2+}$ [133]. In *Neurospora*, efficient splicing of
mitochondrial large ribosomal RNA (LSU) and other group I introns require mitochondrial tyrosyl-tRNA synthetase (mtTyrRS), also called as CYT-18 [112]. Both b15 and LSU are devoid of P5abc, the peripheral RNA element necessary for stabilizing the catalytic core of the prototype \textit{Tetrahymena} rRNA intron. CBP2 and CYT-18 were believed to compensate for the absence of this RNA structure by stabilizing the catalytically active structure of the intron core. This idea has been borne out by the recent co-crystal structure of CYT-18 bound to LSU [117] (Fig 4.1).

Another example of protein-facilitated capture of a functional RNA tertiary structure is provided by biochemical studies on telomerase, a catalytic RNP complex responsible for adding telomeric DNA to the 3’-end of the linear chromosomes and solving the end replication problem inherent to the linear chromosomes [134]. Although a reverse transcriptase (TERT) and an associated RNA (TER) that serves as template for DNA synthesis constitute a minimal functional core of telomerase, many proteins have been implicated in telomerase biogenesis and function. For example, biochemical characterization of a homogeneous preparation of \textit{Tetrahymena} telomerase revealed that in addition to TERT, at least four proteins copurify with TER [135]. TER on the other hand is remarkable for its species-specific size variation (150 nts in ciliates and more than 1500 nts in certain yeasts). Consequently, during telomerase biogenesis, TER could potentially traverse through several folding intermediates before attaining the active folded conformation. Several lines of evidence suggest that some of the telomerase-associated proteins could stabilize certain folding intermediates of TER and thereby assist the RNA through the correct folding pathway [136]. Direct visualization of such protein-
assisted tertiary structure capture by p65, one of the *Tetrahymena* telomerase associated proteins has been achieved by using a single molecule fluorescent resonance energy transfer (Sm FRET)-based approach. It has been demonstrated that p65 induces specific structural rearrangements in TER, which in turn directs the binding of the catalytic subunit, TERT to form the functional ternary complex (Fig. 4.2) [137].

In the two examples stated above, the presence of protein subunits ensures formation of the functionally relevant conformation in their cognate RNA. However, the protein subunit of RNase P provides an interesting scenario for protein facilitated-RNA catalysis, wherein the RNA subunit adopts the catalytic conformation even in the absence of the protein subunit. The RNA subunit of bacterial RNase P requires a small protein cofactor for function *in vivo*. Based on results from extensive biochemical and biophysical studies, many roles have been proposed for the bacterial RPP to facilitate RNA-based catalysis. These include (i) local stabilization of the RPR tertiary structure near its active site, (ii) increasing the affinity of the ribozyme for the catalytic metal ion, and (iii) preferential binding of the ptRNA substrate by the holoenzyme over the mature product due to specific recognition of the leader sequence by the protein. A relatively simpler system like bacterial RNase P is amenable for various biochemical experiments that help dissect the role of a single RNA and protein cofactor to catalysis. The problem is far more complicated when one attempts to ascertain contributions to catalysis from individual RNA and protein cofactors in complex catalytic RNP complexes such as the spliceosome or eukaryal RNase P which contains multiple protein subunits.
To address the individual role of RNA and protein subunits for assembly and catalysis in an RNP complex containing multiple protein subunits, we focused on archaeal RNase P as a model system. Archaeal RNase P is of intermediate composition in terms of its subunit make-up when compared to its bacterial and eukaryal relatives: one RNA and four protein subunits as opposed to one and ten in bacterial and eukaryal RNase P, respectively. Comparison of secondary structure models for bacterial, archaeal and eukaryal RPRs shows variations in RNA secondary structural elements outside the universally conserved catalytic core. While all bacterial RPRs tested so far are catalytically active \textit{in vitro} without their protein cofactor only a few archaeal and eukaryal RPRs are catalytic under certain \textit{in vitro} conditions suggesting that protein subunits play a direct role in catalysis in these RNase P holoenzymes. However, this dissertation provides evidence that despite defects in substrate binding, the catalytic center likely resides in the RNA moiety of all archaeal RNase P. If so, what are the individual roles of multiple RPPs in RPR-mediated catalysis?

We hypothesized that while bacterial RPRs take advantage of RNA-RNA interactions among the peripheral structural components for global structural stability of the RPR, which in turn ensure juxtaposition of S and C domains (a pre-requisite for catalysis), archaeal RPRs replace such RNA-RNA interactions with RNA-protein or protein-protein interactions. In this study, we have identified one such protein pair (RPP21-RPP29) which might be involved in structural stabilization of the archaeal RPR. This inference is based on the ability of RPP21-RPP29 to reduce the concentration of monovalent and divalent cations requirement for the RPR-alone reaction despite its
inability to enhance the rate of the RNA alone reaction. Moreover, footprinting analyses of partially and fully assembled \textit{Mja} RNase P suggest that the RPP21-RPP29 complex does not bind to the C domain but interacts predominantly with the S domain. Since RPP21-RPP29 complex does not contribute to cleavage, we propose that the role of RPP21-RPP29 in archaeal RNase P is structural stabilization of the RPR similar to the roles of CYT 18 or CBP2 proteins which stabilize the catalytic core of group I introns. On the other hand, the binary complex consisting of POP-RPP30 enhances the rate of cleavage by 100-fold. Consistent with this observation, footprinting analysis of the archaeal RNase P holoenzyme indicates that POP5-RPP30 interact with the C domain of the archaeal RPR. Moreover, a pared-down version of archaeal RNase P can be reconstituted from the catalytic domain of archaeal RPR and POP5-RPP30, once again highlighting the latter’s role in catalysis. In this regard, the function of archaeal POP5-RPP30 parallels the sole protein subunit of bacterial RNase P. Whether the POP5-RPP30 pair employs a mechanism similar to that employed by the bacterial RPP in enhancing the rate of cleavage of the RPR is not clear at this time but will be a subject of future investigation. Since POP5 shares tertiary structure similarity to the bacterial RPP [74], it is possible that POP5 could interact with the leader sequence in a manner similar to the bacterial RPP [49]. Interestingly, in heterologous reconstitutions involving bacterial RPR and archaeal RPPs, archaeal POP5-RPP30 was able to reduce the Mg\textsuperscript{2+} requirement of the RPR alone reaction similar to that reported for bacterial RPP.
4.4 Evolutionary path of RNase P from an RNA enzyme to an RNP enzyme

Since all RPRs share a similar catalytic domain, it appears that these RPRs evolved from a common ancestral ribozyme. Since there is no apparent sequence homology between the sole bacterial RPP and any of the archaeal/eukaryal RPPs, the ancestral ribozyme might be devoid of any protein cofactor. It is possible that the bacterial and archaeal lineages took independent evolutionary paths from a primordial RNA-alone enzyme to arrive at the extant RNP status of the RNase P holoenzyme. Bacteria retained an RNA-based solution to address the issue of stabilizing the RPR’s tertiary structure. To further enhance the catalytic potential of the RNA, bacteria might have recruited a single protein subunit to generate an RNase P holoenzyme while archaea recruited at least two protein subunits (POP5 and RPP30) to enhance the catalytic potential of their RPR. In addition, the peripheral RNA-RNA interactions that stabilize the bacterial RPR might have been replaced with protein-RNA or protein-protein interactions in archaea. It is likely that the binary complex RPP21-RPP29 was recruited to accomplish this task (Fig. 4.3). Further reduction in RNA structure is noticeable in eukaryal RPR as these RPRs lack helices P15, P16, P17 and the P6 pseudo-knot. Therefore, the protein contribution towards the overall stabilization of the RPR is likely to be even more pronounced in eukaryal RNase P holoenzymes as compared to their archaeal and bacterial counterparts.

In addition to structural stabilization of the RPR, multiple protein subunits might account for the higher substrate fidelity associated with archaeal and eukaryal RNase P holoenzymes. For example, when the substrate recognition properties of the bacterial, archaeal and plant RNase P holoenzymes are compared with atypical ptRNA substrates,
there is a direct correlation between protein content and higher fidelity of processing. The plant enzyme with multiple RPPs shows remarkable fidelity in cleavage site selection as opposed to the bacterial enzyme where significant mis-cleavage (as high as 60% with certain substrates) has been observed (Singh and Gopalan, unpublished observations). Although archaeal RNase P miscleaves various deletion derivatives, the extent of miscleavage is lower compared to the bacterial enzyme. This suggests that active site rigidity and plasticity are fine-tuned, presumably a pay off from the higher protein/RNA ratio in archaeal and eukaryal RNase P.

4.5 Why is the catalytic function retained in the RNA subunit of the catalytic RNP? Why has the catalytic role not been transferred to the protein subunit(s) as one might expect from the RNA world hypothesis?

Many recent reports suggest that in addition to ptRNAs, RNase P could cleave many cellular RNAs with highly-ordered structures that bear even weak resemblance to ptRNAs [138-141]. It is possible that a unique RNA-protein interface generated by the RPR and RPPs generates a tertiary fold that would recognize many RNA substrates with similar tertiary structure. This evolutionary scenario would preclude the need for generating many protein enzymes with diverse folds to recognize and process a diverse array of substrates. While it is conceivable that a protein-based catalyst equally versatile as the extant RNP could have evolved, failure to observe such an example in nature might attest to certain unique (unappreciated) aspects of RNP catalysis. For example, snoRNAs that direct the site-specific modification of ribosomal RNA (rRNA) confer broader substrate specificity to the SnoRNP complex by base-pairing interaction with target
sequence and hence rule out the need for evolving many protein enzymes that would modify the rRNA in a sequence-specific manner. Therefore, the need for an RNA subunit in SnoRNP is easily justified.

Currently, we do not have a complete list of substrates other than ptRNAs that are processed by RNase P. However, the transcriptomes of eukaryal genomes offer the possibility of many RNAs which could act as potential substrates for RNase P. Conversely, the highly reduced genome of certain organelles such as human mitochondria with few tRNA-encoding genes and mRNAs will be an ideal locale where the RPR might be dispensed with somewhat more readily. There is a recent report that human mitochondrial RNase P can be reconstituted from three mitochondrially-encoded polypeptides [142]. However, the catalytic mechanism or substrate specificity of the reconstituted activity is currently unknown.

4.6 Concluding remarks

Specific interactions between RNAs and proteins in RNP complexes play a central role in control of various cellular and developmental processes [143]. Although atomic resolution structures of some RNP complexes are available, the precise mechanisms by which proteins modulate RNA structure/function in an RNP complex are still a matter of intense investigation. Catalytic RNP complexes (ribosomes, spliceosomes, and RNase P) are excellent model systems for dissecting the synergistic interactions of RNA and protein components in RNP assembly and function. In this study, using archaeal RNase P as a model system, we have demonstrated the individual roles of RNA and protein subunits in RNP assembly and function. We have demonstrated that while the RNA
subunit of archaeal RNase P harbors the catalytic center, specific protein pairs contribute to stabilization of RNA structure and catalysis. By providing evidence that modest improvements in RNA structure and function can be attained by recruiting one of the two binary complexes of protein cofactors (RPP21-RPP29 or POP5-RPP30), we have gained some support for the idea that a gradual transition of an RNA to an RNP complex might have been accompanied by incremental functional gains.
Figure 4.1 CYT-18 stabilizes the P4-P6 domain of LSU intron in a manner similar to the peripheral structural component P5abc of *Tetrahymena* group I intron. A. Ribbon diagram representing the co-crystal structure of CYT-18/LSU intron. B. Ribbon diagram representing the tertiary structure of *Tetrahymena* group I intron. P5abc is shown in magenta. Figure reproduced from [117]
Figure 4.2. Step-wise protein-induced changes in the telomerase RNA during telomerase assembly. Interaction of p65 (red) with the RNA induces conformational rearrangements in the RNA, which results in binding of the catalytic subunit TERT (green) with concomitant changes in the RNA in the ternary complex. Figure reproduced from [137]
Figure 4.3 Model depicting the differences in global stabilization of the tertiary structure of the bacterial and archaeal RPRs. RNA-RNA interactions stabilize the tertiary structure in bacterial RPR and facilitate RPR-alone catalysis. A single RPP that interacts with the C-domain further enhances RPR’s catalytic potential. Lack of such RNA structural elements in archaeal RPR necessitates a protein-based stabilization (presumably RPP21-RPP29) of the RPR structure in archaeal RNase P. POP5-RPP30 enhances RPR-mediated catalysis similar to the sole bacterial RPP (Figure courtesy: Dr. Tim Eubank, OSU).
APPENDIX

Table 8. List of plasmid constructs used in this study that were made by other members of the Gopalan laboratory

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Obtained from</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBT7-ptRNA$^{Tyr}\cdot S_{50}^{Afu}$ RPR</td>
<td>Cecilia Go</td>
</tr>
<tr>
<td>pBT7-ptRNA$^{Tyr}\cdot S_{25}^{Mja}$ RPR</td>
<td>Cecilia Go</td>
</tr>
<tr>
<td>pBT7-$Ram$ mt-RPR</td>
<td>Venkat Gopalan</td>
</tr>
<tr>
<td>$Mja$ RPR $\Delta S66$-$190$</td>
<td>Venkat Gopalan</td>
</tr>
<tr>
<td>pBT7-ptRNA$^{Tyr}\cdot S_{0}^{Mja}$ RPR</td>
<td>Venkat Gopalan</td>
</tr>
<tr>
<td>pBT7-ptRNA$^{Tyr}\cdot S_{3}^{Mja}$ RPR</td>
<td>Venkat Gopalan</td>
</tr>
<tr>
<td>pBT7-ptRNA$^{Tyr}\cdot S_{4}^{Mja}$ RPR</td>
<td>Venkat Gopalan</td>
</tr>
<tr>
<td>pBT7-ptRNA$^{Tyr}\cdot S_{5}^{Mja}$ RPR</td>
<td>Venkat Gopalan</td>
</tr>
<tr>
<td>pBT7-ptRNA$^{Tyr}\cdot S_{8}^{Mja}$ RPR</td>
<td>Venkat Gopalan</td>
</tr>
<tr>
<td>pBT7-ptRNA$^{Tyr}\cdot S_{9}^{Mja}$ RPR</td>
<td>Venkat Gopalan</td>
</tr>
<tr>
<td>$Mja$ RPPTC1</td>
<td>Venkat Gopalan</td>
</tr>
<tr>
<td>$Mja$ RPPTC2</td>
<td>Venkat Gopalan</td>
</tr>
</tbody>
</table>
LIST OF REFERENCES


