INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Substrate recognition properties of the tRNA\textsuperscript{Trp} intron endonuclease from the archaeabacterium \textit{Halobacterium volcanii}

Thompson, Leo Douglas, Ph.D.

The Ohio State University, 1990
SUBSTRATE RECOGNITION PROPERTIES OF THE tRNATrp INTRON ENDONUCLEASE FROM THE ARCHAEBACTERIUM HALOBACTERIUM VOLCANII

DISSERTATION

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

By

Leo Douglas Thompson, B.S.

* * * * *

The Ohio State University
1990

Dissertation Committee:
Dr. C. Daniels
Dr. A. Darzins
Dr. D. Galloway
Dr. J. Reeve

Approved by
Charles Daniels
Advisor
Department of Microbiology
For Opa and Oma

I know they would be proud.
ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my advisor Chuck Daniels for not only being a wonderful mentor but also for being a good friend; without his support and encouragement this work would not have been possible. Many thanks also must go to Terri; her friendship and support helped my wife to accept the trials and tribulations a graduate student must at times endure.

There are many individuals to whom I am indebted for their friendship and assistance. The Reeve laboratory has been a continuing source of information and helpful discussion; thanks especially to Jim and Elizabeth Haas-Brown for being there! To the many individuals who have passed through our laboratory I express thanks for the help and good luck in future endeavors. A special appreciation goes to Dan Nieuwlandt for being a great labmate and all around good guy.

On a more personal note I wish to acknowledge Matisse, Fuzzbutt, and Boots for endless hours of stress reduction. For mom and dad, I owe you so much for your support and
unwavering faith in my desire to further my education and myself.

Last but most certainly not least I give my thanks and my love to Cindi. It has been long and tenuous road but we made it! I only hope I can support and assist you in your dreams as you have done for me during my studies.
VITA

March 10, 1959 ................. Born - Marietta, Ohio

1983 .......................... B.S., Virginia Polytechnic
             Institute and State
             University, Blacksburg,
             Virginia

1982-1985  .................... Research Associate, Genex
             Corp. Gaithersburg,
             Maryland

1985-Present  ................... Graduate Research
             Associate, Department of
             Microbiology, The Ohio
             State University

PUBLICATIONS

Thompson, L.D., Nieuwlandt, D., Brandon, L. and C.J.
Daniels. 1989. "Transfer RNA Processing in the
Halophilic Archaebacteria." Canadian J. Micro.,
35: 36-42.

Endonuclease from Halobacterium volcanii. Unique
Substrate Recognition Properties." J. Biol. Chem.,
263: 17951-17959.

of Subtilisin to Staphylococcal Protein A and its


FIELDS OF STUDY

Major Field: Microbiology

Studies in Archaebacterial Molecular Biology

vi
TABLE OF CONTENTS

ACKNOWLEDGEMENTS .................................................. iii
VITA ........................................................................ v
LIST OF TABLES ................................................... x
LIST OF FIGURES .................................................. xi
LIST OF ABBREVIATIONS ........................................ xlv

CHAPTER PAGE

I. INTRODUCTION ..................................................... 1

II. LITERATURE REVIEW ........................................ 4
   Archaebacteria ....................................................... 5
      Overview ......................................................... 5
      Gene structure and organization ......................... 11
      Archaebacterial introns ...................................... 18
   Intron processing .................................................. 18
      Nuclear mRNA (Class I) introns ......................... 19
      Group I/II (Class II) introns .............................. 23
      Chloroplast-encoded tRNA introns ....................... 27
      Nuclear-encoded (Class III) introns .................... 27
      Archaebacterial tRNA introns ............................. 43

III. MATERIALS AND METHODS ..................................... 48
   Growth of bacterial strains ................................. 49
   Halobacterium volcanii endonuclease purification .......... 49
   Sulfolobus solfataricus endonuclease preparation .......... 51
   Thermoplasma acidophilum endonuclease preparation ........ 51
   Yeast crude endonuclease extract preparation ............. 52
   tRNA runoff transcription .................................... 53
   Synthesis of HC23S and HC16S rRNA and Minsub RNA .......... 54
   Endonuclease cleavage assay .................................. 55
   Yeast endonuclease assay ..................................... 56
   Electrophoresis of nucleic acids ............................ 56
   UV shadowing of nucleic acids .............................. 58
   Gel purification of RNA ....................................... 59
   Nucleic acid precipitation .................................... 59
Large scale isolation of plasmid DNA from E. coli .................................. 60
Isolation of RF DNA from M13 infected E. coli .................. 61
Small scale isolation of plasmid DNA from E. coli .................. 62
Exonuclease III-directed deletions of DNA .......................... 63
DNA restriction and ligation .................................................. 64
Preparation of single-stranded template DNA ..................... 65
Construction of large-intron substrates .......................... 66
Reconstruction of tRNA\textsuperscript{Trp} ............................ 68
Construction of the minimal MinSub endonuclease substrate ........ 69
Oligonucleotide preparation ............................................. 69
Phosphorylation of oligonucleotides ................................. 70
Phosphorylation of DNA and RNA ..................................... 70
Oligonucleotide-directed mutagenesis ................................. 71
Recloning M13-based substrates into pUC18 DNA ...................... 73
Preparation of competent E. coli ....................................... 74
Transformation of competent E. coli ................................... 75
Transfection of competent E. coli ..................................... 76
Analysis of 5' end nucleotides ........................................... 76
RNA sequence analysis ..................................................... 78
Analysis of 3' end nucleotides from excised intron .............. 79
Synthesis of pCp .............................................................. 80
3' pCp end-labeling of RNA ............................................... 81
M13 Sanger dideoxy DNA sequencing .................................. 81
Kinetics assays .............................................................. 82
Structure probing analysis ............................................... 83
Protein precipitation for SDS-PAGE .................................... 84
Determination of protein concentration ............................... 85
SDS-PAGE .............................................................. 85

IV. RESULTS ........................................................................... 86
Development of an endonuclease assay ................................. 86
Purification of endonuclease .............................................. 91
Characteristics of the cleavage reaction ............................... 96
Fidelity of cleavage .......................................................... 96
Biochemical properties ...................................................... 103
Substrate specificity ......................................................... 111
Inability to cleave a nonhalophilic precursor tRNA .................. 111
Requirement for intact precursor ....................................... 114
Role of intron ................................................................. 120
Identification of a conserved structural motif ....................... 133
Reconstruction of o16\textsuperscript{7} and o16 ....................... 136
Analysis of the conserved structure at the exon-intron boundary ........ 139
Identification of cleavage sites by the endonuclease ............... 148
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Introns in chloroplast tRNA genes</td>
<td>30</td>
</tr>
<tr>
<td>2. Introns in nuclear tRNA genes</td>
<td>31</td>
</tr>
<tr>
<td>3. Introns in archaeabacterial genes</td>
<td>44</td>
</tr>
<tr>
<td>4. Purification of <em>Halobacterium volcanii</em> tRNA&lt;sub&gt;Trp&lt;/sub&gt; endonuclease</td>
<td>92</td>
</tr>
<tr>
<td>5. Comparison of yeast and halophilic tRNA intron endonucleases</td>
<td>181</td>
</tr>
<tr>
<td>6. List of oligonucleotides</td>
<td>206</td>
</tr>
<tr>
<td>7. Bacterial Strains</td>
<td>208</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Phylogenetic tree</td>
<td>8</td>
</tr>
<tr>
<td>2. Splicing mechanisms of Class I and Class II introns</td>
<td>21</td>
</tr>
<tr>
<td>3. Locations of preRNA introns</td>
<td>28</td>
</tr>
<tr>
<td>4. Cleavage pathways for Class III introns</td>
<td>33</td>
</tr>
<tr>
<td>5. Construction of the <em>H. mediterranei</em> tRNA&lt;sub&gt;Trp&lt;/sub&gt; encoding plasmid pT72HM&lt;sup&gt;16&lt;/sup&gt;</td>
<td>87</td>
</tr>
<tr>
<td>6. Structure and cleavage assay of <em>H. mediterranei</em> tRNA&lt;sub&gt;Trp&lt;/sub&gt; precursor</td>
<td>89</td>
</tr>
<tr>
<td>7. Hydroxylapatite column chromatography of <em>H. volcanii</em> tRNA endonuclease</td>
<td>94</td>
</tr>
<tr>
<td>8. RNA sequence analysis of the endonuclease products</td>
<td>97</td>
</tr>
<tr>
<td>9. Determination of the 3' terminal nucleotides</td>
<td>100</td>
</tr>
<tr>
<td>10. Time dependence of substrate cleavage with <em>H. volcanii</em> endonuclease</td>
<td>104</td>
</tr>
<tr>
<td>11. Determination of temperature optima for <em>H. volcanii</em> and <em>T. acidophilum</em> intron endonucleases</td>
<td>106</td>
</tr>
<tr>
<td>12. Ion requirements of <em>H. volcanii</em> tRNA endonuclease</td>
<td>109</td>
</tr>
<tr>
<td>13. Structure and cleavage assays for ^16 and yeast preRNA&lt;sub&gt;Phe&lt;/sub&gt; RNAs</td>
<td>112</td>
</tr>
</tbody>
</table>
14. Structure and cleavage assays of \(^{16},^{13},\) and \(^{13\text{II}}\) RNAs ........................................... 115

15. Structure and cleavage assays of \(^{16},^{16\text{M}},\) and \(^{13\text{II}}-5\) RNAs ........................................... 118

16. 3' end group analysis of the products from cleavage of \(^{16},^{13},\) and \(^{13\text{II}}\) RNAs ......................... 121

17. Competition assay of \(^{16}\) RNA with competing mature yeast tRNA ........................................... 123

18. Sequence and structural conservation of a potential stem and loop within the \(^{16}\) intron and the anticodon stem and loops of tRNA\(^{\text{Tyr}}\) and tRNA\(^{\text{Ile}}\) ........................................... 126

19. Structure and cleavage assays of \(^{13},^{13\text{II}-10},\) and \(^{13\text{II}-24}\) RNAs ........................................... 128

20. Summary of exon and intron deletions in \(^{16}\) RNA ........................................... 131

21. Conservation of exon-intron boundary structures among archaebacterial intron-containing precursor RNAs ........................................... 134

22. Reconstruction of the \(H. \text{mediterranei}\) intron-containing tRNA\(^{\text{Trp}}\) gene ........................................... 137

23. Structure and cleavage assays of \(o_{16}^{\text{M}}\) and \(o_{16}\) RNAs ........................................... 140

24. Analysis of structural alterations at the exon-intron boundary junctions of \(o_{16}\) ..................... 143

25. Structure probing of \(o_{16}\) and \(^{\text{GGAG}}\) substrate RNAs ........................................... 145

26. Structure and cleavage assay of the minimal intron endonuclease substrate Minsub ..................... 149

27. RNA sequence of end labeled intron derived from cleavage of Minsub RNA ................................. 151

28. Identification of cleavage sites by the halophilic endonuclease ........................................... 154

29. 5' end group analysis of products from cleavage site identification mutants ............................. 156

31. Analysis of RNase III cleavage site of the 16S and 23S rRNAs from H. cutirubrum ................. 163

32. Cleavage of o16 RNA by methanogenic and thermoacidophilic endonuclease extracts .............. 166

33. Cleavage of yeast pretrNA^Phe and o16 RNAs by yeast tRNA intron endonuclease .................. 169

34. 5' end group analysis of the products from cleavage of o16 RNA by Thermoplasma endonuclease ....................... 171

35. Cleavage of ^16, ^167, and ^1311-5 RNAs by Thermoplasma tRNA intron endonuclease ............ 173
ABBREVIATIONS

α  alpha
^  deletion substrate, deletion
β  beta
bp  basepair
BSA  bovine serum albumin (RNase free)
Cm  chloramphenicol
DNA  deoxyribonucleic acid
ddNTP  dideoxynucleotide triphosphate
DMSO  dimethylsulfoxide
dNTPs  dATP, dCTP, dGTP, TTP
DTT  dithiothreitol
EDTA  ethylenediaminoacetic acid, pH 8.0
EtBr  ethidium bromide
γ  gamma
g  gram
h  height
hr  hour
i  insertion
IPTG  isopropylthiogalactoside

xiv
μl  microliter
w  width
X-gal  5-bromo-4-chloro-3-indolyl-β-D-galatopyranoside
Chapter I
Introduction

The archaebacteria were proposed in 1977 to represent an evolutionarily distinct third primary kingdom, equally distant from the eukaryotes as from the eubacteria (244). Although they possess numerous unique biochemical properties, the archaebacteria share many characteristics with the eukaryotes and eubacteria. One feature shared by the eukaryotes and archaebacteria is the presence of introns in some genes. Of particular interest are the introns which interrupt tRNA genes, which must be spliced from the pre-tRNA before it can participate in protein biosynthesis. The study of tRNA intron structure and processing is complicated by the observation that there are three different types of tRNA introns. These intron types are the chloroplast encoded tRNA introns, the nuclear encoded tRNA introns, and the archaebacterial tRNA introns.

Chloroplast encoded tRNA introns can be found in several different locations within precursor tRNAs. Because
of their conserved structural elements, these introns are related to the Class II (group I and II) introns; many of these introns are self-splicing, and the chloroplast tRNA introns may be processed in a similar fashion. Nuclear encoded tRNA introns are restricted to one conserved location, between nucleotides 37 and 38, in the mature tRNA and are dependent upon protein-mediated excision for splicing. The archaebacterial tRNA introns are an enigma. Although, like the chloroplast encoded tRNA introns, they are found in many locations within pre-tRNAs, archaebacterial tRNA introns do not share the conserved structural elements of Class II introns and, like the nuclear encoded tRNA introns, are dependent upon protein-mediated intron cleavage.

This quandary raises the question of what relationship, if any, archaebacterial tRNA introns and their processing enzymes have to the eukaryotic chloroplast and nuclear encoded tRNA introns processing systems. Because of the presumably self-splicing nature of the chloroplast encoded tRNA introns, it is likely that these introns are processed by a mechanism different from both the nuclear and archaebacterial tRNA introns. Due to the conserved location of nuclear encoded tRNA introns, the mechanism employed by the nuclear encoded tRNA intron endonuclease to identify its cleavage sites must be different from the mechanism used by the archaebacterial tRNA intron endonuclease, since in the
latter intron location varies among pretRNAs. An examination of the mechanism and substrate specificity of the archaebacterial tRNA intron endonuclease may provide insight into the origin and relationship of these very different tRNA intron processing systems.
Chapter II

Literature Review

In 1977 two significant but unrelated publications appeared in the literature. These communications dealt with the postulation of a third primary kingdom, the archaebacteria, which was evolutionarily as distant from the eukaryotes as from the eubacteria (244), and the first report of an intervening sequence, or intron, in a coding gene (66). Considerable research has been conducted in these two fields of study and these investigations have led to a greater understanding of the evolution of gene structure, gene expression, organismal evolution, and development. This work reflects the merging of both field of study: tRNA intron processing in the archaebacteria. While the emphasis of this research is RNA processing, a brief overview of archaebacterial molecular biology and gene structure is appropriate.
Archaebacteria

Overview

The archaebacteria were proposed by Woese and Fox (244) in 1977 to represent an entirely new and evolutionarily distinct primary kingdom. Thus, organisms were classified not by the presence or absence of a nucleus (eukaryote and prokaryote, respectively), but on phylogenetic rather than morphological criteria.

As first suggested by Zuckerkandl and Pauling (251), certain types of biological molecules, called semantophoretic molecules or semantides, can serve as "information molecules" (molecules that carry information of genes or are a transcript thereof) as a measure of relatedness amongst organisms. Semantides can be used as an evolutionary molecular chronometer if they fulfill three criteria. First, they must exhibit clocklike behavior. That is, changes in the sequence must occur as randomly as possible with time. A sequence which has no selective constraints cannot be used as it mutates too quickly. Secondly, phylogenetic range must be suitable as to include all members examined. Lastly, a large molecule is needed for good statistics; the number of alterations must be statistically significant. Furthermore, one needs a molecule which is composed of a large number of domains which are somewhat independent of one another in an evolutionary state. Genes are primary semantides and best
fulfill these criteria. To the extent that two genes for the same function in different organisms have related sequences, the organisms are related. The extent to which two such sequences differ measures the time since the organisms diverged from a common ancestor. The molecule chosen by Woese and colleagues was 16/18S rRNA. A small (16S and 18S) ribosomal RNA exists in all organisms examined and serves the same function. It is a relatively large molecule (=1500 nucleotides) and is composed of several domains; an alteration in one domain, then, will not significantly perturb the remainder of the molecule. Digestion of 16S or 18S rRNA with ribonuclease T1 generated a catalog of RNA oligonucleotides which was characteristic for each organism from which the rRNA was isolated. A comparison of RNase T1 catalogs from two organisms allowed these investigators to formulate a number, termed the $S_{AB}$ value, which was a measure of the relatedness, evolutionarily, of the two organisms. This analysis was applied to the methanogens, a group of unusual obligately anaerobic prokaryotic organisms, and showed that these organisms had $S_{AB}$ values significantly different from more "conventional" prokaryotes, while amongst themselves they had very similar values. Further research indicated that these methanogenic organisms were also unrelated to eukaryotes species. Based upon these findings Woese and Fox proposed that the methanogens comprised a third primary
kingdom, the archaebacteria, which appeared to be equally
distant from the eukaryotes as from the other prokaryotes
(59), or eubacteria, as they are now called. Further work
has supported the contention that archaebacteria are a
distinct group. While sharing some traits with both the
eukaryotes and eubacteria, these organisms possess a number
of unique biochemical properties which are quite different
from the members of the two other primary kingdoms (54).
The phylogenetic relationship of these three kingdoms is
shown in Figure 1.

The archaebacteria as a kingdom are composed of three
subclasses: the methanogens, the halophiles, and the sulfur-
dependent acid-thermophiles (thermoacidophiles). The
methanogenic archaebacteria are obligate anaerobes and are
typically found in sediment, sludge digesters, cattle rumen,
and the large bowel of man and animals. They utilize C1
carbon sources and generate methane as an end product of
metabolism. While many methanogens are mesophilic, several
are thermophilic (230) and there is one report of a
methanogenic species capable of growth at 110°C (85). The
halophiles are noted for their ability to habitate
saturating (2.5 to 5 molar) salt environments. This is not
an adaptive process as these organisms require high salt for
maintainence of an intact cell envelope (116) and have a
very high internal potassium concentration which exceeds 2
molar (116, 117). Optimal growth temperature for these
Figure 1. Phylogenetic tree. Unrooted phylogenetic tree constructed from complete sequences of 16S or 18S rRNA. An approximate indication of evolutionary time is provided by the lengths of the lines which indicate sequence distances (the bar corresponds to 0.1 mutational events per sequence position). Taken from (244).
Archaebacteria

Eukaryotes

Eubacteria

Figure 1
organisms is 40-45°C. Some members of the halophiles have as part of their cell membrane the protein bacteriorhodopsin (10), which converts light energy to ATP. Halophiles attack salted foods and animal hides and are frequently found in salt lakes. The thermoacidophiles are noteworthy for their existence in perhaps the most inhospitable of environments (199). Low pH (pH 1-2 typically) and high temperature (ranging from 58°C to over 105°C) are typical growth conditions for these organisms. Thermoacidophiles are often isolated from acidic hot springs, volcanic areas, and deep sea hydrothermal vents.

The archaebacteria have many unusual biochemical properties. They lack peptidoglycan in their cell envelope and muramic acid is replaced by pseudomurine. Another component often found in their envelope, N-acetyl-talosamuric acid, a derivative of the sugar talose, is found nowhere else in nature (95). The archaebacteria exhibit four different types of cell envelopes. Type I, the most common cell wall found, is present in gram (+) archaebacteria and resembles a gram (+) eubacterial cell wall. Type II is only found in gram (+) M. fervidus and is covered with a surface layer of protein subunits. Type III is a typical gram (-) archaebacterial cell wall and is found in all thermoacidophiles and many halophiles and methanogens. Type IV is found in Methanospirillum and is the most complex cell envelope. Each cell is surrounded by
a layer of proteinaceous material; several cells are held together by a sheath of protein fibrils.

Archaebacterial cell membrane lipids are distinguished by being isoprenoid and having isoprenol glycerol ether lipids, rather than the straight-chain and fatty acid ester glycerol lipids which are characteristic of the membrane lipids found in eubacteria and eukaryotes (120).

Histone-like DNA binding proteins have been identified in *Thermoplasma* (187), *Methanothermus fervidus* (J. Krzycki, personal communication), and *Sulfolobus acidocaldarus* (70, 75). The archaebacteria have a complex DNA dependent RNA polymerase (249) which is insensitive to α-amanitin, an antibiotic inhibitory to DNA dependent RNA polymerases II and III of the eukaryotes. Rifampicins which inhibit all eubacterial polymerases at concentrations below 1 μg/ml are completely inactive against the halophiles, methanogens, and thermophiles. Most eubacterial protein synthesis inhibitors are inactive in all archaebacteria, however anisomycin, which affects eukaryotic peptidyl transfer in the 80S ribosome, does act against many archaebacteria (14).

**Gene structure and organization**

The genomes of archaebacteria appear to be structurally similar in size to those of the eubacteria. Genome sizes range from $8.4 \times 10^8$ daltons for *Thermoplasma acidophilum* (189) to $2.3 \times 10^9$ daltons for *Halobacterium halobium* and
Halococcus morrhuae (147). These genomes are similar in size to Escherichia coli, which has a genome of 2.5 X 10^9 daltons (147). The mole % G+C content of archaebacterial chromosomes varies from 25.8 in Methanosphaera stadtmaniae (144) to 68 in Halobacterium sodomense (145). No correlation between high G+C content and growth temperature is apparent. In fact, some methanogens which grow at higher temperatures have relatively low G+C content DNA; M. fervidus, with a growth optima above 85°C, has only 33% G+C in the DNA. High internal salt concentrations and DNA-binding histone-like proteins appear to assist in protection of double-stranded DNA from thermal denaturation (194).

A characteristic feature of several species of halophilic archaebacteria are the presence of A+T rich regions within the genome (38, 50, 51, 146, 147). One 70 kilobase pair A+T rich region is a site at which insertion elements (ISH1, ISH2, ISH26, ISH27, and ISH56) and uncharacterized repetitive sequences are clustered (168). The presence of insertion sequences typically found within the A+T rich regions of the halobacterial genomes is in part responsible for the high frequency of genomic rearrangement (181, 182) as most of these rearrangements appear to occur within the A+T-rich regions. In general, halobacterial insertion elements are similar in size and structure to eubacterial insertion elements, ranging from 520 to 1895 basepairs in length. They contain terminal inverted repeats
of 8 to 29 basepairs and their transposition causes a duplication of the target DNA (21).

Archaeabacteria also have extrachromosomal genetic elements. Plasmids have been identified in members of all three branches of the archaeabacteria (50, 192); the halobacteria have the largest reported number of plasmids. Plasmid pH1, a 150 kb plasmid of Halobacterium halobium, carries information coding for gas vacuole formation (41, 50) and bears representatives of all the insertion elements, except for ISh1, found in Halobacterium halobium (50). This plasmid has also been shown to integrate into the genome of H. halobium (41, 50). A number of small plasmids have been described in the halobacteria and methanogens. The plasmid pL, a derivative of the halobacterial phage φH, confers resistance to phage φH infection (185). Plasmid pVH2 of H. volcanii has been sequenced and found to contain four potential open reading frames (28). Constructs derived from this plasmid have been used to develop the first system for gene cloning in the archaeabacteria (28). Sulfolobus B12 carries the plasmid pSB12, which exists either integrated into the chromosome or as an extrachromosomal element. Amplification following ultraviolet radiation suggests that this plasmid may be a prophage (133). Methanobacterium thermoautotrophicum strain Marburg contains a plasmid, pME2001, derivatives of which have been constructed to replicate in E. coli, Staphylococcus aureus, and
*Saccharomyces cerevisiae* (139). *Methanolobus vulcani*
contains a plasmid, pMP1, which integrates into the host
genome and may also be a prophage (211).

Several archaebacterial viruses and virus-like particles are known. The temperate phage \( \phi H \) of *H. halobium*
is circularly permuted and terminally redundant, which suggests it may employ a headfull packaging method (183, 184). Virus-like particles from *Sulfolobus*, such as SSV1, are lemon shaped with a short tail, but it is not known if these particles function as viruses (133). Rod shaped virus-like particles have been detected in *Thermoproteus tenax* and contain linear double stranded DNA (88).

Gene organization in the archaebacteria resembles the eubacteria. Many archaebacterial operons or operon-like structures have been identified. Polycistronic mRNAs have been detected from transcripts isolated from *M. thermoautotrophicum* strain Marburg, *M. voltae*, *M. vannielii*, *F. formicicum*, and *Sulfolobus* (15, 111, 163, 171). An example of a well-characterized operon-like unit are the genes which encode methyl coenzyme M reductase, a multisubunit enzyme found in all methanogens. These genes have been cloned and sequenced from five methanogens and shown to be highly similar in sequence and gene organization (15, 16, 34, 106). The *mcrBDCGA* genes are tightly linked; preceding every gene is a sequence complimentary to the 3' terminus of 16S rRNA which presumably functions as a
ribosomal binding site, analogous to the eubacterial Shine/Delgarno sequence (190). A binding site for RNA polymerase and the site of transcription initiation have been identified upstream of the mcrBDCGA gene cluster from M. vannielii (212), and transcription termination sites have been mapped downstream of mcrA in M. vannielii, M. thermoautotrophicum strain Marburg, M. voltae, and M. fervidus (15, 16, 228).

Archaebacterial promoters have been identified primarily by comparison of conserved sequences near transcription initiation sites, and have been supported by DNA dependent RNA polymerase binding analysis (22, 200, 210, 212). Putative promoter regions have been identified from polypeptide and stable RNA genes. Consensus promoter sequences for methanogens, AAANNTTTATATA, and sulfur-dependent thermophiles, AAANNTTTATATA, are similar in sequence but differ significantly from the consensus halophilic promoter sequence, GANGCCCYTTAAGTA (Y any pyrimidine, N any residue) (21). Transcription termination is not well understood in archaebacteria but potential hairpin-loop structures, which resemble eubacterial rho-independent termination sequences, are found downstream of many archaebacterial genes (6, 13).

Because of their abundance and ease of isolation, stable (ribosomal, transfer, and 7S) RNAs have been extensively studied, providing much information on gene
organization and structure. Stable RNA genes from all three branches of the archaebacteria have been identified and sequenced (27, 39, 83, 86, 126, 132, 211, 232, 233). Archaebacterial ribosomes contain three ribosomal RNAs, 5S, 16S, and 23S rRNA. Complete rRNA operon sequences from *M. thermoautotrophicum* (158), *H. halobium* (86), *H. cutirubrum* (134), *T. tenax* (101), and *Desulfurococcus mobilis* (102) indicate that rRNA gene organization is similar to eubacterial rRNA operons: 16S-23S-5S (21); a tRNA gene is present in the 16S-23S spacer region and a tRNA gene is often found downstream of the 5S rRNA gene (21, 78). As in the eubacteria, tRNA\(^{\text{Ala}}\) is found in almost all 16S-23S spacer regions in the halophiles and methanogens. *H. morrhuae*, *H. halobium*, and one of two *H. volcanii* operons contain a tRNA\(^{\text{Cys}}\) downstream from the 5S rRNA (39, 121, 126). *M. vannielli* and *M. voltae* contain an unlinked 5S rRNA gene found within a tRNA gene coding cluster (231, 233), while *Sulfolobus acidocaldarius* and *Thermococcus celer* contain an unlinked 5S rRNA gene not in association with a tRNA gene cluster (150). In thermoacidophiles, the rRNA genes exist as single copies and do not contain tRNA genes in the spacer regions between the rRNA genes (159, 157). *D. mobilis* and *T. tenax* have only a linked 16S-23S rRNA operon (102). No linkage exists between the 16S, 23S, and 5S rRNAs in *Thermoplasma acidophilum* (215). *D. mobilis* is unusual in
having an intron of 622 nucleotides interrupting the 23S rRNA gene (102).

All archaebacteria contain an abundant stable RNA, the 7S RNA (78, 126). 7S RNA varies in size from 325 to 375 nucleotides and is not associated with ribosomes. Sequence comparisons of *H. halobium* 7S RNA show a high degree of sequence and structural similarity to eukaryotic 7S RNA and eubacterial 4.5S RNA (191, 214). The eukaryotic 7S RNA is a component of the signal recognition particle and this suggests that the archaebacterial 7S RNAs have a related function (78, 191).

tRNA genes have been identified and sequenced from tRNA clustered genes, genes in individual transcription units, and multigene transcription units. Within the methanogens (78, 231, 232) the largest tRNA operon, from *H. vannielli*, contains seven tRNA genes and a single 5S rRNA gene. A second operon from this organism contains five tRNA genes. A number of tRNA genes have been characterized from *H. volcanii* and related halophiles (38). The sequences of 41 tRNAs from *H. volcanii* have been published (76) and indicate that, overall, archaebacterial tRNAs agree in structure to eubacterial and eukaryotic tRNAs. However, archaebacterial tRNAs resemble eukaryotic tRNAs in that, with few exceptions, the 3' terminal CCA nucleotides found in eubacterial tRNAs are not present in archaebacterial tRNAs; the exceptions include tRNA^{Pro}_{UGG}, tRNA^{Asn}_{Guu}, and
tRNA\text{His}^{\text{GUG}} \text{from \textit{M. vannielii}} \text{and tRNA}\text{Pro}^{\text{UGG}} \text{from \textit{M. voltae}} (231, 232). Another similarity to eukaryotic tRNA genes is the presence of introns within some precursor tRNA transcripts (40, 93).

**Archaebacterial introns**

The discovery of introns interrupting the stable RNA genes within the members of some archaebacteria was an unexpected finding; introns were considered to be a characteristic of eukaryotic genomes only. As yet, no introns have been found to interrupt archaebacterial messenger RNAs, but introns have been found in transfer and ribosomal RNAs (38, 42, 93, 101, 102). All archaebacterial introns have been found in members of the halophiles and thermoacidophiles. While introns have not been found in methanogen genes, their presence would not be unexpected due to the close evolutionary relationship they share with the halophiles. The number of genes sequenced from these organisms is still rather limited, so the likelihood of discovering methanogen-encoded introns is good.

**Intron processing**

Intervening sequences, or introns, have been grouped into three classes, depending upon the structure of the intron and the mechanism employed to splice the intron (26): the Class I introns are represented by the nuclear mRNA
Introns, the Class II introns include the Group I and Group II introns, several members of which are self splicing, and the Class III introns are the nuclear encoded tRNA introns.

As a group, tRNA genes are diverse in their intron content; in the nucleus tRNA genes are interrupted by the small unstructured Class III introns, whereas the chloroplast tRNA genes are interrupted solely by Class II introns, primarily Group II. The archaebacterial tRNA introns, while small in size, are found in several positions in the mature domain, similar to that observed for chloroplast tRNA introns. Whether these introns were derived from a single common ancestral tRNA intron is not clear, however the structure and location of the tRNA introns play an important role in intron processing.

**Nuclear mRNA (Class I) Introns**

Nuclear premRNA (Class I) introns were the first introns to be discovered (30). These introns average in size from 200-500 bp but can be as large as 50,000 bp (53, 98, 167). The conserved elements of nuclear premRNA introns are limited to the exon-intron boundary regions (237). The 5' and 3' ends of a typical nuclear mRNA intron are defined by the consensus dinucleotides GU and AG, respectively (26, 161). Early studies demonstrated that introns from one organism can often be spliced correctly when introduced into another organism (143, 235), reflecting the general nature
of the splicing mechanism and the conserved nature of the splice junctions. There are variations in the efficiencies of heterologous processing reactions, and generally difficulties arise with the 3' splice site selection (236). Another essential and characteristic element of mRNA splicing is the sequence which acts as the internal branch site during processing (119).

Most knowledge of mRNA splicing has come from work done with yeast and HeLa cell splicing extracts (58, 161). Biochemical and genetic analysis has demonstrated a dependence upon trans-acting factors (snRNPs, small nuclear ribonucleoproteins) for splicing (128). These factors interact to form complexes known as spliceosomes (19, 68). After assembly of the spliceosome, splicing is initiated with a transesterification reaction (phosphodiester rearrangement) (189). This transesterification reaction involves the transfer of the 5' phosphate from the first nucleotide of the intron (G) to the 2'-OH of the ribose of the A residue at the branch site, near the 3' end of the intron (Figure 2). This reaction generates a free 5' exon with a 3'-OH and the intron-3' exon in the form of a branched circular intermediate, termed a lariat (160); these intermediates remain in the spliceosome. The second splicing step involves a transesterification reaction between the 5' phosphate preceding the first nucleotide of the downstream exon and the 3'-OH of the first exon to form
Figure 2. Splicing mechanisms of Class I and Class II introns. Left figure illustrates the mechanism employed to splice Class I (nuclear mRNA) introns. The reaction is initiated by the transfer of the 5' phosphate from the first nucleotide of the intron to the 2'-OH of the A residue at the branch site. The 5' phosphate of the 3' exon then attacks the 3'-OH of the 5' exon, releasing from the spliceosome the spliced exons and a lariat intron. Right figure illustrates the processing of group I and II introns. Splicing of the Group I intron is initiated by a guanosine moiety at the 5' splice site, releasing the 5' exon plus linear intron-3' exon, which has the added G residue at the 5' end. Next, the 3'-OH of the 5' exon reacts with the 3' splice site junction, releasing the exons and the intron as a linear product with the G residue at the 5' end. For Group II intron splicing, the 2'-OH of an A residue within the intron attacks the 5' splice junction to release the 5' exon and a lariat intron-3' exon intermediate. The released 5' exon attacks the 3' splice junction to produce the spliced exons plus an intron lariat.
Figure 2.
the spliced exon product. The phosphate at the splice junction originates from the 5' phosphate of the first nucleotide of the 3' exon. The excised intron terminates with a 3'-OH and the phosphate that forms the 2'-5' bond at the branch site is derived from the 5' phosphate of the first intron nucleotide (110). The spliceosome then disassembles, releasing the spliced mRNA. There is no net change in the number of phosphodiester bonds in the RNA substrate.

The transacting factors, snRNPs, are essential to the splicing reaction and are known to bind to the premRNA (23). snRNPs may function to hold the precursor in proper configuration for the transesterification reactions to occur.

**Group I/II (Class II) Introns**

Self-splicing prerRNA was discovered by T. Cech and his colleagues (112) while studying the large rRNA gene of *Tetrahymena thermophila*; this work was the first evidence that splicing could be entirely RNA mediated. Subsequent work showed that these self splicing introns, termed Group I introns, existed in a large number of organellar DNAs of fungi and plants. Examples of Group I introns include, but are not limited to, the cob gene of *Neurospora crassa* (63), the yeast cox I gene (221), bacteriophage T4, T2, T6 (31, 32), the large rRNA gene of nuclear RNA of *Physarum* (148),
and some chloroplast-encoded tRNA introns (18). Group I introns have a conserved core structure of at least seven loops and helices. These introns can be distinguished from Group II or nuclear mRNA introns by comparing the exon-intron boundary sequences; in the case of the Group I introns the last nucleotide of the 5' exon is U and the last nucleotide of the intron is a G. An internal guide sequence is present which is thought to align the 5' and 3' splice sites. Finally, some introns contain an ORF which often codes for a maturase or endonuclease which may participate in processing or transposition of the intron.

The self-splicing of Group I introns is mediated by the folded structure of the intron RNA itself and proceeds via two transesterification reactions (Figure 2). The first transesterification reaction (25) involves a guanosine or a guanosine nucleotide cofactor as the attacking nucleophile to release as intermediates the 5' exon with a 3'-OH and a linear intron-3' exon with the G added to the intron 5' end. In the second transesterification reaction the 3'-OH of the first exon reacts with the 3' splice junction to form the spliced exons and a linear intron product with the extra G still at the 5' end. The phosphate at the splice junction is derived from the first nucleotide of the 3' exon; the G cofactor is attached via a phosphate derived from the first nucleotide of the intron. The last three to seven nucleotides of the 5' exon are essential in defining the 5'
splice site by pairing with the internal guide sequence in the intron (25, 227). The excised intron of Tetrahymena can participate in further splicing reactions (25). These reactions include cyclization of the intron with release of an oligonucleotide from the 5' end that contains the added G; in addition the cyclic intron can reopen itself at the cyclization junction and act on external RNAs to promote sequence specific G addition, nucleotidyl transfer, and phosphotransfer (247, 248).

Several Group I introns from fungi require proteins for efficient splicing of the intron (2, 122). These proteins are referred to as maturases and are encoded by an open reading frame in the intron. It is believed the maturase functions in vivo by binding to sites on the premRNA and participating in the folding of the RNA (43).

Group II introns are found in fungal organellar mRNAs and chloroplast tRNAs (25, 149). Only a minority of Group II intron contain open reading frames, although some chloroplast tRNA introns do (26). These intron do not have the distinguishing features of Group I introns; rather, they have a different set of conserved sequences, or domains. These include short boundary sequences and a longer sequence located near the 3' end of the intron (25, 206). The 5' conserved sequence has the consensus GUGCG and the 3' conserved sequence is the dinucleotide AU or AC, with no known exceptions. Like the Group I introns, self-splicing
Group II introns also undergo a series of transesterification reactions (Figure 2) (25). The initiating nucleophile is the 2'-OH of an adenosine residue. The first transesterification step is the attack by this hydroxyl on the 5' splice junction, releasing the 5' exon with a 3'-OH end and a lariat intron-3' exon intermediate. The released 5' exon attacks the 3' splice junction to produce the spliced exons plus excised intron lariat. The phosphate at the splice junction is derived from the first nucleotide of the second exon, while the 2'-5' phosphodiester is from the first nucleotide of the intron. As with the Group I intron transesterification reactions, there is no net change in the number of phosphodiester bonds present, but unlike Group I splicing, this reaction does not depend upon a guanosine cofactor.

Group II introns do not have the internal guide sequence present in Group I introns, but Jacquier and Michel (87) have demonstrated two sequences located in domain 1 of yeast αl5g intron that basepairs with adjacent sequences near the 3' end of the 5' exon. As with some Group I introns, Group II introns may also participate in additional reactions (167). The release of the Group II intron as a lariat suggests a relationship between this mechanism and that of nuclear mRNA introns, as does the similarity of the exon-intron boundary junction sequences (26).
Chloroplast-encoded tRNA introns

Transfer RNA genes of chloroplast DNA are interrupted by Group I or, most often, Group II introns (18, 44, 45, 129, 149, 195, 202, 245). These introns are located within the anticodon and the DHU stem (Figure 3) and range in size from 450 to 2500 nucleotides (Table 1; (193)). Little is known about how these introns are spliced. Since they have the conserved structural elements of Group I and II introns, they presumably utilize the same splicing mechanisms employed by the non-chloroplast Group I and Group II introns.

Nuclear-encoded tRNA (Class III) Introns

Among the first eukaryotic genes reported to be interrupted by intervening sequences were tRNA genes of yeast (66, 219, 220), and the first indication that genes interrupted by introns were functional was the yeast SUP4 nonsense suppressor (66). Of the 360 yeast tRNA genes sequenced, ≈10% have introns (77, 97, 156, 220). Nine yeast tRNA gene families are interrupted by introns (155); these interrupted gene families are Tyr, Phe, LysUUU, ProUGC, Trp, SerCGA, LeuUGA, LeuCAA, and IleUAU. The introns range in size from 14 to 60 nucleotides. Nuclear encoded tRNA introns have also been found in tRNA genes from humans (130), Zea mays (112), Drosophila (178) and Nicotina rashica (200). Table 2 (193) lists the known nuclear encoded tRNA
Figure 3. Locations of pre-tRNA introns. Locations of introns within precursor tRNA encoded by eukaryotic nuclei, chloroplasts, and archaebacteria are indicated.
Figure 3.
Table 1

Introns in Chloroplast tRNA Genes

<table>
<thead>
<tr>
<th>Organism</th>
<th>Gene</th>
<th>Intron Size</th>
<th>Intron Location</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Cyanophora paradoxa</em></td>
<td>Leu</td>
<td>232</td>
<td>34</td>
<td>I</td>
</tr>
<tr>
<td><em>Glycine max</em></td>
<td>Ala</td>
<td>ND</td>
<td>38</td>
<td>&quot;II&quot;</td>
</tr>
<tr>
<td></td>
<td>Ile</td>
<td>ND</td>
<td>38</td>
<td>&quot;II&quot;</td>
</tr>
<tr>
<td><em>Hordeum vulgare</em></td>
<td>Gly</td>
<td>677</td>
<td>24</td>
<td>IIB'</td>
</tr>
<tr>
<td></td>
<td>Val</td>
<td>597</td>
<td>36-38</td>
<td>IIA'</td>
</tr>
<tr>
<td><em>Marchantia polymorpha</em></td>
<td>Ala</td>
<td>768</td>
<td>38</td>
<td>IIA</td>
</tr>
<tr>
<td></td>
<td>Gly</td>
<td>593</td>
<td>24</td>
<td>IIB</td>
</tr>
<tr>
<td></td>
<td>Ile</td>
<td>886</td>
<td>38</td>
<td>IIA</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>315</td>
<td>34</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>Lys</td>
<td>2111</td>
<td>38</td>
<td>IIA</td>
</tr>
<tr>
<td></td>
<td>Val</td>
<td>530</td>
<td>38</td>
<td>IIA</td>
</tr>
<tr>
<td><em>Nicotiana tabacum</em></td>
<td>Ala</td>
<td>710</td>
<td>37</td>
<td>IIA</td>
</tr>
<tr>
<td></td>
<td>Gly</td>
<td>691</td>
<td>24</td>
<td>IIB</td>
</tr>
<tr>
<td></td>
<td>Ile</td>
<td>707</td>
<td>37</td>
<td>IIA</td>
</tr>
<tr>
<td></td>
<td>Ile</td>
<td>729</td>
<td>37</td>
<td>IIA</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>503</td>
<td>34</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>Lys</td>
<td>2526</td>
<td>38</td>
<td>IIA</td>
</tr>
<tr>
<td></td>
<td>Val</td>
<td>571</td>
<td>36, 37</td>
<td>IIA</td>
</tr>
<tr>
<td><em>Pisum sativum</em></td>
<td>Lys</td>
<td>2447</td>
<td>38</td>
<td>IIA</td>
</tr>
<tr>
<td><em>Sinapis alba</em></td>
<td>Lys</td>
<td>2547</td>
<td>38</td>
<td>IIA</td>
</tr>
<tr>
<td><em>Sorghum bicolor</em></td>
<td>Leu</td>
<td>450</td>
<td>34</td>
<td>I</td>
</tr>
<tr>
<td><em>Triticum aestivum</em></td>
<td>Gly</td>
<td>676</td>
<td>24</td>
<td>IIB</td>
</tr>
<tr>
<td><em>Vicia faba</em></td>
<td>Leu</td>
<td>451</td>
<td>34</td>
<td>I</td>
</tr>
<tr>
<td><em>Zea mays</em></td>
<td>Ala</td>
<td>806</td>
<td>36-38</td>
<td>IIA</td>
</tr>
<tr>
<td></td>
<td>Ile</td>
<td>949</td>
<td>38</td>
<td>IIA</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>458</td>
<td>34</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>Val</td>
<td>603</td>
<td>36-38</td>
<td>IIA</td>
</tr>
</tbody>
</table>

1 branch A nucleotide at position 7
1' branch A nucleotide at position 8
### Table 2

<table>
<thead>
<tr>
<th>Organism</th>
<th>Gene</th>
<th>Intron Size</th>
<th>Intron Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabidopsis thaliana¹</td>
<td>Tyr GU</td>
<td>12</td>
<td>37</td>
</tr>
<tr>
<td>Dictyostelium discoideum</td>
<td>Trp CCA</td>
<td>13</td>
<td>37</td>
</tr>
<tr>
<td>Drosophila melanogaster</td>
<td>Leu CAA</td>
<td>38, 45</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Tyr GUA</td>
<td>105</td>
<td>37</td>
</tr>
<tr>
<td>Glycine max¹</td>
<td>Met CAU</td>
<td>11</td>
<td>37</td>
</tr>
<tr>
<td>Homo sapiens</td>
<td>Tyr GUA</td>
<td>20, 21, 22</td>
<td>37</td>
</tr>
<tr>
<td>Neurospora crassa</td>
<td>Leu AAG</td>
<td>27</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Phe GAA</td>
<td>16</td>
<td>37</td>
</tr>
<tr>
<td>Nicotiana rustica¹</td>
<td>Tyr GUA</td>
<td>13</td>
<td>37</td>
</tr>
<tr>
<td>Podospora anserina</td>
<td>Ser UGA</td>
<td>18</td>
<td>37</td>
</tr>
<tr>
<td>Saccharomyces cerevisiae</td>
<td>Ile UAU</td>
<td>60</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Leu CAA</td>
<td>32, 33</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Leu UAG</td>
<td>19, 19</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Lys UUU</td>
<td>23</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Phe GAA</td>
<td>18, 19</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Pro UGG</td>
<td>28, 31, 31</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Ser CGA</td>
<td>19</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Ser GCU</td>
<td>19</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Trp CAA</td>
<td>34</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Tyr GUA</td>
<td>14, 14</td>
<td>37</td>
</tr>
<tr>
<td>Schizosach. pombe</td>
<td>Leu UAA</td>
<td>16</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Lys CUU</td>
<td>8</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Ser CGA</td>
<td>16</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Ser UGA</td>
<td>15</td>
<td>37</td>
</tr>
<tr>
<td>Xenopus laevis</td>
<td>Tyr GUA</td>
<td>12, 13</td>
<td>37</td>
</tr>
</tbody>
</table>

¹Plant Nuclear
introns. Within a family of tRNA genes, either all members will contain introns, or none will. There is sequence conservation of introns within a tRNA family but not between tRNA families and all introns are located in the same position, one nucleotide 3' to the anticodon. Before these pretRNA transcripts can be functional in protein biosynthesis, the intron must be removed and the exon halves ligated. Knapp et al. (108), using cell free extracts prepared from *Saccharomyces cerevisiae*, were the first to demonstrate removal of introns *in vitro* from pretRNAs, indicating that introns were enzymatically removed at the RNA level. Subsequent to this splicing extracts have been prepared from *S. cerevisiae* (108), *Neurospora crassa* (218), *Xenopus laevis* oocytes (159), HeLa cells (122), wheat germ (198), mouse L cell nuclei (242), and *Chlamydomonas* (217). The ability of the yeast system to process all yeast intron-containing pretRNAs and heterologous intron-containing pretRNAs suggested that a single enzyme system acted on all precursors and that this system may be highly conserved in eukaryotic nuclei.

The process of nuclear-encoded tRNA splicing has been extensively studied in the yeast *S. cerevisiae* and a cleavage and ligation mechanism has been proposed (72, 155; Figure 4). These steps occur by two separate protein catalyzed reactions involving endonucleolytic excision of the intron followed by ligation of the exon halves (166).
Figure 4. Cleavage pathways for Class III introns. tRNA splicing occurs in two stages, endonuclease cleavage and ligation. See text for details of the cleavage mechanism. In yeast and wheat, ligation consumes two ATP molecules, one of which donates the phosphate that joins the half-tRNAs. In *Xenopus* and HeLa ligation one ATP molecule is consumed. The phosphate that joins the half-tRNAs is derived from the pretRNA. These phosphates are enclosed by a circle or diamond shape so their origins can be traced.
Class III

Figure 4.
In this system the endonuclease cleaves the pretRNA at the 5' and 3' splice sites to produce a 5' exon with a 2',3' cyclic phosphate terminus, a 3' exon with a 5'-OH terminus, and a free intron with both a 2',3' cyclic phosphate terminus and a 5'-OH terminus. Since mature tRNA structure is required for this reaction, the 5' and 3' exon halves are presumed to remain annealed during intron excision. A cyclic phosphodiesterase activity cleaves the 2',3' cyclic phosphate terminus of the 5' tRNA half to produce a 2' phosphate terminus, and a kinase activity phosphorylates the 5'-OH terminus of the 3' tRNA half by transfer of the Y-phosphate of an ATP cofactor, producing a 5' phosphate end. The ligase protein is adenylylated, with the release of pyrophosphate. The adenylate residue, consisting of the α-phosphate and adenosine, is transferred to the 5' phosphate terminus of the 3' tRNA half to produce an activated 3' half tRNA. This activated species is ligated to the 5' tRNA half, releasing the previously added adenyl residue, to form a mature tRNA which possesses a 2' phosphate at the splice junction. The phosphate at the splice junction forming the 3',5'-phosphodiester bond is derived from the Y position of the ATP cofactor. Finally, the 2'-phosphate is removed from the splice junction by a separate 2' phosphatase activity.

All nuclear tRNA intron endonucleases appear to generate similar cleavage termini, but some variation exists in the mechanisms of the ligases. Wheat germ ligase
activity can join tRNA halves in a manner similar to that of yeast, although the presence of a 3' terminal phosphate does not appear to be essential (65, 186). The vertebrate (Xenopus laevis and HeLa) ligases utilize a different mechanism; these enzymes catalyze the attack of the 2',3' cyclic phosphate of exon 1 by the 5' hydroxyl of exon 2, preserving the phosphate (56, 151). This is in contrast to plant, fungi, and T4 RNA ligase which incorporate phosphate from ATP into the splice junction (11, 60, 71, 72).

The yeast endonuclease exhibits properties of an integral membrane protein, and is solubilizable with nonionic detergents such as Triton X-100 (33). Ligase is soluble in the nucleus, but cytological evidence by immunoelectron microscopy has localized the ligase to the nuclear periphery (33). The ligase enzyme is a multifunctional protein which contains all the activities required for ligation: cyclic phosphodiesterase, kinase, and RNA ligase (55, 57, 72). Of the two splicing components, ligase is the best characterized; its gene has been cloned and sequenced (169). The ligase gene codes for a single 95.4 kdal protein which, when expressed in E. coli, exhibits all three activities which copurified with ligase from yeast extracts (229). Biochemical and enzymatic evidence has suggested that the endonuclease and ligase, while distinct proteins, actconcertedly in vivo in a tRNA splicing complex or spliceosome. This was suggested when T4 RNA ligase did not
compete for the tRNA half molecules produced in a coupled endonuclease-ligase assay (71). The yeast ligase had preferential access to the endonuclease products. It is generally believed that the endonuclease and ligase recognize common structural features in pre-tRNAs but evidence exists that each enzyme has its own unique recognition elements, as certain mutations affect only intron excision but not ligation, and some mutations affect excision and ligation differently (74). Endonuclease and ligase of *S. cerevisiae* will splice all intron-containing pre-tRNAs from *S. cerevisiae*, suggesting that these enzymes are exclusively responsible for tRNA splicing in yeast (74, 165). Greer et al. (74) have proposed a tRNA shuttle model wherein ligase initially interacts with end-matured pre-tRNA to form a presplicing complex, which in turn interacts with the membrane bound endonuclease, where splicing takes place, and is perhaps coupled to tRNA translocation into the cytoplasm.

Tertiary tRNA structure, rather than nucleotide specific recognition, has been implicated in substrate recognition by the endonuclease. X-ray crystallographic data (99, 205) has verified the cloverleaf structure proposed for mature tRNA (175), and the tertiary structure has an "L" shaped configuration stabilized by extensive base stacking interactions (77, 170, 204). Analysis of pre-tRNA structures in solution (47, 124, 206) showed that tRNA precursors are
folded into a mature configuration with the intron extending the anticodon helix. In many, but not all, pretRNAs pairing occurs between the anticodon and the intron. Mutations causing defects in splicing have been isolated in vivo as defective alleles of various intron-containing tRNA suppressors of *S. cerevisiae* and *S. pombe* (151, 164, 239). These point mutations were most often located in the mature domain of the precursor, although extensive deletions in the intron could cause defects in splicing, suggesting a minimum intron length may be required for splicing. Early studies demonstrated that insertions into the intron of yeast tRNA\textsubscript{Leu\textsubscript{3}} did not affect processing in vitro with Xenopus or yeast extracts, indicating that the structure of the intron was not important for recognition by the endonuclease (90). Introns of pretRNAs from *S. pombe* are not base paired to the anticodon helix as with *S. cerevisiae* pretRNA introns, although transcripts are spliced by *S. cerevisiae* extracts, indicating intron secondary structure is not required for splicing in yeast (74, 239). However, some structures at the exon-intron boundaries have a major effect on cleavage; these changes usually limit or block processing. Base pairing of the 3' cleavage site of yeast SUP32 resulted in a loss of cleavage at both sites by the yeast endonuclease (207), and removal of the 3' cleavage site of yeast tRNA\textsubscript{Phe} (135), or changes in the base pairing between the cleavage sites of yeast tRNA\textsubscript{3Leu} (4) decreased cleavage by the
Xenopus endonuclease. Mutations in exons of S. pombe suppressor pretrRNAs can lead to reduced suppressor efficiencies in vivo, and reduced splicing both in vivo and in vitro (65). Integrity of the helical structure of the D stem appears to be required for 5' end maturation and intron excision (5, 62, 135). Finally, Winey et al. (241) have inserted a 13 nucleotide intron into the naturally intronless single copy tRNA\(^{Gly}\) of S. cerevisiae and shown that it is functional in vivo, thereby supporting the idea that the splicing enzymes recognize the tertiary, rather than primary, structure of the precursor tRNA.

Since all nuclear encoded tRNA introns are located in the same position, the 5' and 3' cleavage sites are always eleven and six nucleotides away, respectively, from the topmost basepair of the anticodon stem (156). To study substrate recognition and to examine if these fixed distances were important for splice site selection by the yeast nuclear intron endonuclease (173), Reyes and Abelson designed a synthetic pretrRNA\(^{Phe}\) gene which was amenable to mutagenesis (172). With this gene, Reyes and Abelson were able to show that the introduction of a base pair in the anticodon stem did not block cleavage, however the excised intron from this altered substrate was observed to be two nucleotides larger than the wild type intron. Cleavage took place one base 5' and one base 3', respectively, from the wild type 5' and 3' cleavage sites. From this data they
concluded that the endonuclease maintained the distance from the central region of the mature body of the pretRNA (possibly the top of the anticodon stem) to the two splice sites and selected the splice sites by counting a fixed distance down the anticodon helix from this fixed point to the cleavage sites. When a two base pair insertion was introduced into the anticodon helix the excised intron was extended four nucleotides; the 5' and 3' splice sites were shifted two nucleotides 5' and 3' from the wild type sites, as predicted by the measurement model. Deletion of U₃₃, preceding the anticodon, generated an intron one nucleotide shorter than wild type; the 5' splice site had been shifted one base in the 3' direction, maintaining the fixed distance, while the 3' site was unaffected. Selection of the splice sites, then, was independent of one another; therefore, the endonuclease may have two distinct catalytic sites. Insertions into the amino acceptor stem, to localize the origin of the measurement mechanism, had no effect upon splice site selection, indicating that measurement originates somewhere within the center of the mature portion of the pretRNA. Experiments have also shown that the endonuclease has two probable contact points with the mature portion of the tRNA at U₈ and C₅₆; mutations which altered these nucleotides but did not affect tertiary structure of the precursor were not cleaved by the endonuclease. The highly conserved purine 3' to the anticodon, at the exon-
Intron boundary, may also be essential for intron excision. The intron itself appears to play a passive role in recognition; the tRNA\textsuperscript{Phe} intron could be replaced by poly U, provided G was present at the cleavage site, without loss in cleavage ability. Mattoccia et al. (136) have conducted similar experiments with the Xenopus endonuclease and have presented similar conclusions concerning endonuclease substrate specificity. However, a deletion of nucleotides 49-56, adjacent to the 3' splice site, in pretRNA\textsuperscript{Phe} drastically reduced cleavage of this substrate by the endonuclease. They suggest that the Xenopus enzyme may require the 3' splice junction to be single stranded for processing to occur, and a single-stranded loop at the 3' cleavage site may be a signal to the endonuclease that an intron is present in the molecule. Differences have been noted in the overall efficiency with which an endonuclease acts on heterologous substrates; some S. pombe pretRNAs are poor substrates for the S. cerevisiae endonuclease (61, 203). The wheat germ nuclear endonuclease is highly specific for only wheat germ pretRNAs and is incapable of cleaving S. cerevisiae, X. laevis, and human pretRNAs (199).

van Tol et al. (224) have presented data suggesting the possibility of self cleaving excision of the intron in pretRNAs. In control experiments with human pretRNA\textsuperscript{Tyr}, a minor amount of specific cleavage was observed at the 3' splice site, and to a lesser extent at the 5' splice site.
These data were interpreted as evidence for self cleavage and they propose that the role of the endonuclease is to maintain the structure of the precursor in an appropriate conformation for self cleavage at the 3' cleavage site, while the endonuclease cleaves at the 5' splice site. This does not appear to be a universal property of all precursors; there has not been evidence for self cleavage by other intron-containing tRNA precursors.

The location and temporal order of events in tRNA maturation have been studied in X. laevis oocytes with microinjection of pretRNAs or cloned tRNA genes, and this process is believed to be similar for most eukaryotes (71, 140). All processes involving size alteration of the transcript (end maturation, 3' CCA addition, intron excision and exon ligation) are always detected in the nucleoplasm, and never in the cytoplasm (46). Some modifications take place in the nucleus (m\(^5\)C) while others take place in the cytoplasm (G\(^\text{14}\) in tRNA\(^\text{TYR}\)) (90).

The lack of introns in most tRNA genes suggests that they do not have a general function in tRNA maturation, however the presence of an intron in some precursor tRNAs is essential for proper tRNA base modification. The intron of the tyrosine inserting suppressor SUP6 has been shown to be required for correct modification of the mature tRNA (91). When the intron was deleted from SUP6 the resulting gene was found to be less efficient as a suppressor than the intron-
containing homolog (91); the suppressor tRNA encoded by the intron deficient gene was found to lack the nucleotide 35 pseudouridine modification in the anticodon, which is found in the wild type tRNA. SUP53, a leucine inserting suppressor tRNA, also requires the intron for the m^5C^34 (5-methyl-cytidine) base modification in the anticodon of both the mature and suppressor tRNAs (200). There are several reports of base modification of tRNATyr which are intron dependent (29, 114, 200, 222). Introns are not, however, universally required for proper anticodon modification. Winey et al. (240) showed that the proline tRNA intron is not required for proper anticodon modification of the mature tRNA. Introns do not appear to shield tRNAs from improper modifications either. New modifications were not detected in any of four tRNAs encoded by intron deleted tRNA genes (90, 239).

**Archaebacterial tRNA introns**

Many archaebacterial tRNA genes have been found to be interrupted by introns (Table 3; (193)). The majority of the reported tRNA introns have been found in members of the thermophilic branch of the archaebacteria. Introns have been found in the genes encoding tRNA^Ser, tRNA^Leu, tRNA^Phe, tRNA^Gly, and tRNA^Met in *S. solfataricus*. These introns range in size from 15-25 nucleotides (94) and are located
<table>
<thead>
<tr>
<th>Organism</th>
<th>Gene</th>
<th>Intron size</th>
<th>Intron location</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Halobacterium volcanii</em></td>
<td>tRNA&lt;sub&gt;Trp&lt;/sub&gt;</td>
<td>106</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>tRNA&lt;sub&gt;Met&lt;/sub&gt;</td>
<td>75</td>
<td>37</td>
</tr>
<tr>
<td><em>Halobacterium cutirubrum</em></td>
<td>tRNA&lt;sub&gt;Trp&lt;/sub&gt;</td>
<td>105</td>
<td>37</td>
</tr>
<tr>
<td><em>Halobacterium mediterranei</em></td>
<td>tRNA&lt;sub&gt;Trp&lt;/sub&gt;</td>
<td>104</td>
<td>37</td>
</tr>
<tr>
<td><em>Solfolobus solfataricus</em></td>
<td>tRNA&lt;sub&gt;Gly&lt;/sub&gt;</td>
<td>15</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>tRNA&lt;sub&gt;Met&lt;/sub&gt;</td>
<td>18</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>tRNA&lt;sub&gt;Leu&lt;/sub&gt;</td>
<td>15</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>tRNA&lt;sub&gt;Phe&lt;/sub&gt;</td>
<td>18</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>tRNA&lt;sub&gt;Ser&lt;/sub&gt;</td>
<td>25</td>
<td>37</td>
</tr>
<tr>
<td><em>Thermoproteus tenax</em></td>
<td>tRNA&lt;sub&gt;Ala&lt;/sub&gt;</td>
<td>16</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>tRNA&lt;sub&gt;Leu&lt;/sub&gt;</td>
<td>18</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>tRNA&lt;sub&gt;iMet&lt;/sub&gt;</td>
<td>20</td>
<td>37</td>
</tr>
<tr>
<td><em>Thermofilum pendens</em></td>
<td>tRNA&lt;sub&gt;Gly&lt;/sub&gt;</td>
<td>18</td>
<td>46</td>
</tr>
<tr>
<td><em>Desulfurococcus mobilis</em></td>
<td>tRNA&lt;sub&gt;Met&lt;/sub&gt;</td>
<td>39</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>23S rRNA</td>
<td>622</td>
<td>2124</td>
</tr>
</tbody>
</table>
one nucleotide 3' to the anticodon. In addition to the intron in the 23S rRNA gene, a 39 nucleotide intron has been found in the tRNA^Met gene of D. mobilis, one nucleotide 3' to the anticodon (100). T. tenax genes encoding tRNA^Met, tRNA^Leu, and tRNA^Ala have been found to be interrupted by introns (234). These introns are small (16 to 20 nucleotides), however only tRNA^Met has an intron one nucleotide 3' to the anticodon. The intron of tRNA^Leu is located within the anticodon while the intron of tRNA^Ala interrupts the anticodon stem 5' to the anticodon. tRNA introns are not limited to the anticodon stem and loop region; tRNA^Gly from Thermophilum pendens contains an 18 nucleotide intron located within the extra arm sequence (105). Of the halophilic archaeabacteria, only two tRNA genes, tRNA^Trp and tRNA^Met, from H. volcanii and related halophiles, have been found to contain introns. These introns, 104-6 and 75 nucleotides, respectively, also found one nucleotide 3' to the anticodon (40, 42), are the largest of the archaeabacterial tRNA introns, and with the exception of pretRNA^Tyr from Drosophila (105 nucleotide), are also larger than any reported eukaryotic nuclear encoded tRNA introns.

Little is known about intron excision from archaeabacterial tRNAs. tRNA^Trp from H. volcanii is present as a single copy gene and northern analysis has shown the transcript to be processed in vivo into mature tRNA^Trp (37,
similar data exists for the interrupted tRNA genes of T. tenax (234), D. mobilis (100), and T. pendens (105). A mechanism does therefore exist to splice precursor tRNAs into mature tRNAs. Because of the variability of intron location within archaebacterial pretRNAs (Figure 3), it is unlikely that the measurement and substrate recognition mechanism employed by S. cerevisiae endonuclease to identify splice sites in precursor tRNA can occur with archaebacterial pretRNAs. Two possibilities can be proposed: either a different endonuclease exists for each intron-containing archaebacterial precursor tRNA, similar to charging enzymes, or a presently unknown substrate recognition pattern exists for the endonuclease such that one intron endonuclease recognizes and cleaves all pretRNAs independent of intron location. There are few conserved features shared by archaebacterial tRNA introns; one common feature is a presence of short helix flanked by two loops at the intron-exon boundary junctions (37, 94, 105, 213). It may be that structural or sequence features at the cleavage sites play a significant role in cleavage site recognition by the archaebacterial enzymes.

My research has focused on the tRNA \textsuperscript{T\textsubscript{TP}} intron endonuclease from Halobacterium volcanii and its substrate recognition properties. I have chosen to characterize the biochemical and substrate recognition properties of this endonuclease in an attempt to define the specificity of
substrate cleavage and identification of splice site selection in the archaeobacteria.
Chapter III
Materials and Methods

Restriction endonucleases, T4 DNA ligase, DNA polymerase large fragment (Klenow), exonuclease III, RNA ligase, and ribonucleases used in enzymatic sequence analysis of RNA were obtained from Bethesda Research Laboratories. Polyethylene glycol 8000, S1 nuclease, nuclease P1, cellulose thin layer plates with indicator (254 nM), zymolyase, and monophosphate nucleotide markers were obtained from Sigma. T7 RNA polymerase, phenol (molecular biology grade), and T4 polynucleotide kinase were obtained from U.S.Biochemicals. The radionucleotides \( \alpha-^{32}\text{P}\)ATP (3000 Ci/mmol), \( \alpha-^{32}\text{P}\)dATP (3000 Ci/mmol), and \( \gamma-^{32}\text{P}\)ATP (7000 Ci/mmol) were obtained from ICN Pharmaceuticals, Inc. X-ray film (X-OMAT AR, 20.3 x 25.4 cm and X-OMAT RP, 35 x 42 cm), developer (GBX), and fixer (GBX) were obtained from Kodak. Oligonucleotides for site directed mutagenesis and gene reconstruction were obtained from the Ohio State University Biochemistry Instruments Center (Table 6).
Growth of bacterial strains. *Escherichia coli* strains MV1190, CJ236, TB1, and JM101 (Table 7) were grown in YT media (8% tryptone, 5% yeast extract, 5% NaCl in H₂O) at 37°C. *Halobacterium volcanii* strain DS2 was grown in *H. volcanii* media containing, per liter, 125 g NaCl, 45 g MgCl₂, 10 g MgSO₄, 10 g KCl, 1.34 g CaCl₂, 3 g yeast extract, and 5 g tryptone with constant shaking at 37°C. *Saccharomyces cerevisiae* strain EJ101 was grown in YPD media (1% yeast extract, 2% peptone, and 2% glucose in H₂O) with constant shaking at 30°C. *Sulfolobus solfataricus* was grown in *S. solfataricus* media containing, per liter, 1 g yeast extract, 1 g casamino acids, 3.1 g KH₂PO₄, 2.5 g (NH₄)₂SO₄, 0.2 g MgSO₄, 0.25 g CaCl₂, 1.8 mg MnCl₂, 4.5 mg Na₂B₄O₇, 0.22 mg ZnSO₄, 0.05 mg CuCl₂, 0.03 mg Na₂MoO₄, 0.03 mg VOSO₄, 0.01 mg CaSO₄ per 1 liter H₂O, pH 4.0-4.5, with periodic agitation at 75°C. *Thermoplasma acidophilum* was grown in *T. acidophilum* media containing, per liter, 3 g KH₂PO₄, 0.5 g MgSO₄, 0.25 g CaCl₂, 0.2 g (NH₄)₂SO₄, 1 g yeast extract, 1% glucose (w/v), pH 1.5, with periodic agitation at 58°C.

*Halobacterium volcanii* endonuclease purification.

*Halobacterium volcanii* cells (10.5 g) were resuspended in 13 ml TMGK (40 mM Tris-HCl pH 7.5, 20 mM MgCl₂, 10% glycerol, 50 mM KCl) and passed twice through a French pressure cell at 10,000 psi. The lysate was centrifuged 15 min, 10,000
rpm, at 4°C in a Sorvall SS-34 rotor to remove unbroken cells. The supernatant was centrifuged for 1 hr, 45,000 rpm at 4°C in a Beckmann Ti60 rotor to remove particulate matter; this was the "S-100" fraction. The S-100 fraction was transferred to a beaker and PEG 8000 was slowly added with stirring to a final concentration of 10% (w/v). After the PEG 8000 was fully dissolved the lysate was kept on ice for 30 min and then centrifuged 20 min, 10,000 rpm, at 4°C to remove precipitated material. The PEG supernatant was dialyzed overnight against a 50-100 fold volume excess of TMGK to rid lysate of residual PEG 8000. A portion of the dialyzed lysate (10 ml) was applied to a 30 x 2 cm (h x w) DEAE Sephacel column, previously equilibrated with TMGK, and the column was washed with five column volumes of TMGK. Proteins were eluted from the DEAE Sephacel column with a linear 200 ml, 50-600 mM KCl, gradient in TMG (10% glycerol); fractions were collected in 4 ml aliquots. To identify fractions containing endonuclease activity all fractions were assayed; endonuclease activity eluted from the DEAE Sephacel column at 350 mM KCl. The active fractions were pooled, dialyzed against TMGK as described above, reapplied to the DEAE Sephacel column, and eluted under the same conditions. Fractions active for endonuclease activity (350 mM KCl) were pooled and dialyzed against 50-100 fold volume excess of 10 mM sodium phosphate pH 6.8. A portion of the dialysate (5 ml) was applied to a
16 x 1.3 cm (h x w) hydroxylapatite column, previously equilibrated with 10 mM sodium phosphate pH 6.8, and washed with five column volumes of 10 mM sodium phosphate pH 6.8. Protein was eluted with a linear 30 ml 10-400 mM sodium phosphate gradient and 1 ml fractions were collected and assayed for endonuclease activity; activity eluted from the column at 200 mM phosphate. Fractions which had endonuclease activity were pooled, dialyzed against TMG (2% glycerol), and stored as 50 µl aliquots at -70°C. Under these conditions endonuclease activity remained stable for three months.

*Sulfolobus solfataricus* endonuclease preparation. *S. solfataricus* cells (2 g) were resuspended in 10 ml TMGK. The cell suspension was passed twice through a French pressure cell at 10,000 psi. The cell lysate was centrifuged at 10,000 rpm for 15 min to pellet unlysed cells. The cleared cell lysate was used immediately for endonuclease cleavage assays. Assays were performed at 75°C. *S. solfataricus* endonuclease was unstable and activity was lost after storage for 24 hr at -70°C, -20°C, or 4°C.

*Thermoplasma acidophilum* endonuclease preparation. *T. acidophilum* cells (0.1 g) were placed in a 1.5 ml Eppendorf tube and resuspended in 200 µl TMG containing 0.1%
Triton X-100 to lyse the cells. After a 10-15 min incubation at room temperature, the cells were centrifuged for 15 min in an Eppendorf microfuge to pellet unlysed cells. The cleared lysate was transferred to a 1.5 ml Eppendorf tube and the lysate was stored at -70°C; the lysate retained endonuclease for several weeks. Extract was assayed at 58°C.

**Yeast crude endonuclease extract preparation.** *Saccharomyces cerevisiae* strain EJ101 was grown at 30°C in 40 ml YPD media (1% yeast extract, 2% peptone, and 2% glucose) to a density of $A_{600} = 2$. Cells were harvested at 4°C, resuspended in 40 ml H$_2$O, and pelleted. The cell pellet was resuspended in 0.5 ml Tris-DTT buffer solution containing 0.1 M Tris-HCl pH 7.4 and 10 mM DTT, transferred to a 1.5 ml Eppendorf tube, and incubated for 15 min at 23°C. Cells were isolated by centrifugation in the Eppendorf microfuge for 20 sec and resuspended in 200 μl zymolyase buffer containing 1.2 M sorbitol and 20 mM KPO$_4$ pH 7.5. The resuspended cells were again centrifuged 20 sec and resuspended in 200 μl zymolyase buffer containing 2 mg/ml zymolyase. The cell suspension was incubated at 23°C for 5 min to allow spheroblasts to form. Spheroblasts were isolated by centrifugation, washed twice with 1 ml of zymolyase buffer, then resuspended in 25 μl lysis buffer per 1 $A_{600}$ unit of the original cell culture. Lysis buffer contained 20 mM KPO$_4$ pH 7.5, 0.2 M
NaCl, 2 mM EDTA, 10 mM DTT, 1 mM PMSF, 10% glycerol, and 0.5% Triton X-100. Lysates were stored at -20°C.

tRNA runoff transcription (179). To generate runoff transcripts 1 μg of plasmid DNA, containing the T7 RNA polymerase promoter sequence upstream of the gene of interest, was digested with EcoRl if a ^16 derivative, HindIII if a ^13 derivative, or with BstNl for all other genes, in a total reaction volume of 25 μl in the appropriate restriction buffer. Digestion was conducted for 1 hr at 37°C for EcoRl and HindIII, or 58°C for BstNl. Transcription of the cleaved DNA was initiated by the addition of 10 μl TMG buffer, 50 μg BSA, DTT (10 mM final), ATP (0.1 mM final for synthesis of radioactively labeled RNA or 1 mM final for synthesis of nonradioactively labeled RNA), CTP, GTP, UTP (1 mM each, final) 100 μCi [α-^32P]ATP (3000 Ci/mmol, if radioactively labeling RNA), 80 units of T7 RNA polymerase, and H₂O to a total volume of 100 μl. For synthesis of 5' GMP-terminated RNA, the nucleotide concentrations were adjusted as follows: 2 mM final ATP, CTP, GTP, and UTP, 6 mM GMP; no [α-^32P]ATP. The reaction was incubated for 1 hr at 37°C; transcription was terminated by extraction of the synthesized RNA with 100 μl of phenol. The RNA, contained in the aqueous upper phase, was ethanol precipitated and resuspended in 10 μl of RNA loading buffer.
RNAs were separated by electrophoresis in denaturing polyacrylamide gels and gel purified.

**Synthesis of HC23S and HC16S rRNA and Minsub RNA.** The synthesis of these rRNA substrates took place with partially double stranded DNA oligonucleotides. The first oligonucleotide, either HC23S or HC16S, consisted of an 18 nucleotide sequence, complementary to the T7 RNA polymerase promoter sequence, fused to a sequence complementary to the RNA sequence of interest (the RNase III cleavage site for *H. cutirubrum* 16S or 23S rRNA). The second oligonucleotide ("T7Top") consisted of the 18 nucleotide sequence of the T7 RNA polymerase promoter. When annealed together, a partially double stranded molecule was formed which consisted of a double stranded T7 RNA polymerase promoter and a single stranded region encoding the complementary sequence of the RNA of interest. This method takes advantage of the ability of T7 RNA polymerase to transcribe single stranded DNA, if the T7 RNA polymerase promoter for the gene is double stranded.

For RNA synthesis, 1 µg of "T7Top" and 2 µg of the RNA-encoding oligonucleotide (HC23S or HC16S) were combined with 10 µl 10x TMG buffer and 50 µl H₂O, heated at 68°C for 7 min, and slow cooled at 23°C for 7 min to allow the oligonucleotides to anneal. To the annealed oligonucleotides were added 50 µg BSA, DTT (10 mM final), 2
mM final each of ATP, CTP, GTP, and UTP, 6 mM GMP, and 80 units of T7 RNA polymerase in a total volume of 100 μl. Transcription was carried out at 37°C for 1 hr, the reaction was terminated by extraction with an equal volume of phenol, and the RNA was gel purified as described. The RNA was resuspended in 20 μl of 10 mM Tris-HCl pH 7.5 containing 2 units bacterial alkaline phosphatase and incubated at 42°C for 1 hr to remove the 5' phosphate group. The dephosphorylated RNA was extracted with an equal volume of phenol, ethanol precipitated, and 5' phosphorylated with [γ-32P]ATP as described in "Phosphorylation of DNA and RNA".

A similar approach was used for Minsub RNA synthesis, however the two oligonucleotides encoding this RNA were completely complementary. When annealed, a 57 bp, double stranded DNA was formed. This DNA then served as a template for T7 RNA polymerase directed transcription. Minsub RNA was also derived from the cloned version of this gene; see below.

Endonuclease cleavage assay. The standard assay contained 10,000 cpm Cerenkov of substrate RNA if radiolabeled or 2 ng if not radiolabeled, endonuclease extract (200 ng total protein), and TMG or TSG buffer in a final volume of 50 μl. This mixture was incubated for 60 min at 37°C and the reaction was terminated by the addition of 50 μl phenol. Following centrifugation to separate the two phases, RNAs in
the upper aqueous phase were ethanol precipitated and separated by electrophoresis on a denaturing polyacrylamide gel as described below. Cleavage products were visualized via autoradiography or UV shadowing.

**Yeast endonuclease assay.** Assay buffer contained 20 mM Hepes pH 7.5, 5 mM MgCl₂, 2.5 mM spermidine, 0.1 mM DTT, 0.4% Triton X-100, 10% glycerol, substrate RNA (5 ng), and yeast endonuclease (1.5 μl) in a total volume of 10 μl. Assays were incubated at 30°C for 10 min, terminated by the addition of 1 μl of stop solution containing 2% SDS, 100 mM EDTA, 2 mg/ml proteinase K, and incubated an additional 10 min at 50°C. After the addition of 150 μl TE, the assay was extracted with an equal volume of phenol and the RNAs recovered by ethanol precipitation. RNA products were then separated by electrophoresis through 6% denaturing polyacrylamide gels and visualized by autoradiography. Combined endonuclease-ligase reactions were carried out under the same conditions with the addition of 1 mM ATP to the reaction buffer.

**Electrophoresis of nucleic acids** (48). Gel plates were assembled to receive acrylamide solution using either 0.5 or 0.3 mm spacers, coated with a light covering of Vaseline to prevent leakage. For a small gel (20 x 20 cm), 30 ml of 8.3M urea and 6%, 10%, or 20% polyacrylamide solution, 35 μl
of 10% (w/v) of ammonium persulfate, and 30 μl of TEMED were mixed with gentle swirling and quickly poured into the gel assembly. For a large gel (32 x 40 cm) 85 ml of 8.3M urea and 6%, 10%, or 20% polyacrylamide solution was mixed with 400 μl of 10% (w/v) ammonium persulfate and 400 μl TEMED. The well-forming comb was inserted and clamped in place until the acrylamide had polymerized, 15-30 min. The bottom spacer and comb were removed, and excess Vaseline and unpolymerized acrylamide were rinsed away with hot tap water. The gel assembly was placed into the lower buffer chamber of the electrophoresis apparatus, previously filled with 1x TBE electrophoresis running buffer. The upper chamber was filled with electrophoresis buffer and buffer was used to rinse and fill the top of the gel and the gel wells. A wick of Whatmann 1M chromatography paper was inserted between the gel plates to rest on the surface of the gel, and the other end was placed in the buffer of the upper tank. Saran wrap was used to cover the upper tank and electrode leads were connected to the power supply; if denaturing polyacrylamide was used, the gel was prerun at 400 volts (small gel) or 1000 volts (large gel) for 30 min. After rinsing excess urea from the wells, nucleic acid samples (5-25 μl in loading buffer) were loaded and were subjected to electrophoresis at 600 volts (small gel) or 1400 volts (large gel). After electrophoresis, the gel assembly was removed from the electrophoresis apparatus and
disassembled, leaving the gel on one glass plate. If autoradiography of the samples was to be performed, the gel was transferred to a sheet of used, developed X-ray film, covered with Saran wrap, and placed in an autoradiography cassette containing two Kodak X-omatic intensifying screens. In the darkroom a sheet of unexposed X-ray film was placed on top of the gel, the cassette was closed, and the film was allowed to expose at -70°C. The film was developed by immersion in Kodak developer and fixer as described by the manufacturer. If RNA bands were to be excised from the gel, the film or glass plate supporting the gel was placed on top of the developed autoradiogram such that any orienting marks were aligned, the corresponding gel slices containing the RNA bands were cut out of the gel, and the RNA was eluted, as described.

**UV shadowing of nucleic acids.** To visualize nucleic acid samples by UV illumination, the gel was placed between two sheets of Saran wrap and placed on a TLC plate containing a fluorescent (254 nm) indicator. The gel was illuminated from above with a long wave (300 nm) UV light source. The outline of the RNA, visualized as dark shadows on a green background due to the absorbance of UV light by the nucleic acid, was marked with a pen. The RNA was cut out of the gel and eluted, as described.
Gel purification of RNA. RNA to be gel purified was resuspended in 10 µl RNA loading buffer and loaded onto an 8.3M urea 6, 10, or 20% denaturing polyacrylamide gel and subjected to electrophoresis. After electrophoresis the RNA was visualized by UV shadowing or autoradiography and the location of the RNA band(s) was marked on the gel plate. Using a clean razor blade, the gel slice containing the RNA of interest was cut out of the polyacrylamide gel, the gel slice was placed into a 1.5 ml Eppendorf tube, and 350 µl each of RNA elution buffer and phenol were added to the Eppendorf tube. The tube was closed, sealed with parafilm, and placed on a rocking platform to rock overnight. The next morning, the tube was removed from the platform and the upper aqueous layer, containing the eluted RNA, was transferred to a siliconized Eppendorf tube. The RNA was precipitated as described.

Nucleic acid precipitation. To precipitate nucleic acids, one-tenth volume of 3M sodium acetate pH 4.5, and two and one-half volumes cold 95% ethanol were added to the solution. After vortexing, the solution was chilled 10-15 min in a dry ice-ethanol bath or 30 min at -70°C, and centrifuged for 10 min in an Eppendorf microfuge at room temperature. The supernatant was discarded, the pellet was dried under vacuum, and the nucleic acid was resuspended in an appropriate volume of the desired buffer.
Large scale isolation of plasmid DNA from *E. coli* (131). To prepare plasmid DNA, one liter of *E. coli* cells bearing plasmid was grown for 14 hr in YT media containing 60 mg/ml ampicillin. The cells were harvested by centrifugation for 10 min at 10,000 rpm in a Sorvall GSA rotor, at 4°C. The supernatant was discarded, the cell pellet was then resuspended in 10 ml sucrose buffer (10% sucrose, 20 mM Tris-HCl pH 7.5, and 5 mM EDTA pH 8.0); this was followed by the addition of 2 ml of 10 mg/ml lysozyme in sucrose buffer, and the cell suspension was incubated on ice for 5 min. To the cell suspension was added 16 ml of lysis buffer (2% Triton X-100, 50 mM Tris-HCl pH 8.0, 5 mM EDTA, pH 8.0). The cell suspension was rapidly mixed and the mixture was kept on ice for 20 min with occasional mixing. The cell lysate was cleared of cell debris by centrifuging for 1 hr at 45,000 rpm in a Beckmann Ti60 fixed angle rotor. After measuring the supernatant volume, the supernatant was transferred to 30 ml Corex tubes and solid CsCl was added at a concentration of 1.05 gm CsCl per 1 ml supernatant. After CsCl had completely dissolved, 1/20 volume of 5 mg/ml EtBr was mixed with the supernatant and the tubes were kept on ice for 30 min; this was followed by centrifugation for 10 min at 8,000 rpm in a Sorvall SS-34 rotor to allow the protein-CsCl-EtBr "pad" to collect at the surface of the supernatant. The supernatant was loaded into Beckmann quickseal VT165 polypropylene centrifuge tubes, the tubes
were balanced to within 0.1 g per pair, heat sealed, placed into a Beckmann VTi65 ultracentrifuge rotor, and centrifuged at 20°C for 12-14 hr at 60,000 rpm. Plasmid DNA was visualized with longwave (300 nm) ultraviolet illumination. To collect the plasmid DNA, a 20 gauge syringe needle was inserted into the top of the centrifuge tube to equalize the air pressure and a second 20 gauge needle, attached to a 3 ml syringe, was inserted into the tube below the plasmid DNA band. The plasmid-containing solution was withdrawn into the syringe and transferred to a 15 ml snap-cap tube. Ethidium bromide was removed via three extractions with an equal volume of 20xSSC-saturated IsOH; EtBr preferentially partitioned in the upper IsOH phase. To the plasmid solution was added 2.5 volumes of TE and 2.5 volumes of 95% EtOH. After mixing, the solution was stored at -20°C for 1-2 hr to precipitate plasmid DNA, and the precipitate was recovered by centrifugation in a Sorvall SS-34 rotor at 8,000 rpm for 10 min. The supernatant was decanted and the DNA pellet was dried under vacuum. The plasmid DNA pellet was resuspended in 400 μl TE.

Isolation of RF DNA from M13 infected E. coli. Two 5 ml cultures were established; one with the host strain, JM101 or MV1190, and a second with the cells bearing the M13 phage. Both cultures were incubated overnight at 37°C with shaking. The next morning 0.5 ml of the culture containing
the host alone was used to inoculate 50 ml of YT in a 125 ml flask and incubated for 2 hr at 37°C with shaking. From this culture, 2 ml were used to inoculate 250 ml of YT media in a sterile flask and 2.5 ml of the overnight culture containing M13 phage. This culture was incubated with shaking at 37°C for 2.5-3 hr. Following this, the cells were harvested and RF DNA was isolated as described under "Large scale isolation of plasmid DNA from E. coli".

Small scale isolation of plasmid DNA from E. coli (84). Cells from a 5 ml overnight culture of plasmid bearing E. coli were resuspended in 550 μl STET and the resuspended pellet was transferred to a 1.5 ml Eppendorf tube. Next, 50 μl of 10 mg/ml lysozyme in STET was added, the cell suspension was vortexed, and quickly placed in a boiling water bath for 90 sec. After boiling the tube was centrifuged in an Eppendorf microfuge for 10 min, the cell pellet debris was removed from the lysate with a toothpick, and the cleared lysate was extracted with an equal volume of phenol. The aqueous upper phase was transferred to an Eppendorf tube, 400 μl of Is0H was added, and the solution was mixed by vortexing. The tube was stored at -20°C for 20 min to precipitate the plasmid-rich nucleic acid fraction. The nucleic acids were pelleted by centrifugation for 10 min in an Eppendorf microfuge. The supernatant was discarded,
the nucleic acid pellet was dried under vacuum, and resuspended in 40 μl TE.

**Exonuclease III-directed deletions of DNA** (213). For unidirectional deletions into the 3' end of the tRNA<sub>Trp</sub> gene, 16 DNA (5 μg) was restricted with PstI and BamHI. For deletions within the intron of 16, the plasmid DNA was first propagated in the dam<sup>-</sup>, dcm<sup>-</sup>, hsdS<sup>-</sup> E. coli strain, GM272 (CGSC #6478), and the plasmid (5 μg) was cleaved with BclI. The restricted DNAs were resuspended in a solution containing 88 μl H<sub>2</sub>O, 10 μl 10x exo III buffer (660 mM Tris-HCl pH 7.6, 6.6 mM MgCl<sub>2</sub>, 10 mM β-mercaptoethanol), and 2 μl exonuclease III (150 units); digestion reactions were carried out at 23°C. Immediately upon addition of exonuclease III, 50 μl was removed into a tube containing 50 μl of 2x S1 buffer (2 M NaCl, 0.5 M NaCH<sub>3</sub>COOH pH 4.5, 10 mM ZnSO<sub>4</sub>, 100 mM MgCl<sub>2</sub>, and 5% glycerol) which terminated the exonuclease III reaction. Ten seconds later the remaining 50 μl reaction volume was added to 50 μl of 2x S1 buffer. Nuclease S1 (10 units) was then added to each of the tubes and these tubes were incubated for 15 min at 23°C to create blunt DNA termini. The S1 reactions were terminated by extraction with an equivalent volume of phenol. Following S1 treatment the DNAs were ethanol precipitated and ligated. The ligated DNAs were transformed into competent E. coli TB1 and individual colonies were inoculated into 5 ml YT media
containing 60 μg/ml ampicillin for overnight growth. Plasmid DNA was isolated via the small-scale plasmid isolation procedure. The size of DNA deletions were first determined by determining the sizes of EcoRl-HindIII restriction fragments in 8% nondenaturing polyacrylamide gels. DNA bands were visualized by ethidium bromide staining. To precisely define the extent of the deletions, the EcoRl-HindIII DNA fragments were ligated into EcoRl-HindIII digested M13mp19 vector and subjected to dideoxy DNA sequence analysis.

**DNA restriction and ligation** (131). A typical restriction reaction contained 2.5 µl of DNA (1-4 µg DNA), 2.5 µl of 10x restriction buffer, 1-2 µl of enzyme (10-20 units) and water to give a final volume of 25 µl. The mixture was incubated at the required temperature for 1-2 hr. Digestions were terminated by extraction with an equal volume of phenol, and the cleaved DNA was recovered by ethanol precipitation. Ligation of digested DNA was carried out in a 25 µl reaction volume in 1x ligase buffer (10 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 1 mM ATP, 10 mM DTT) with 0.2 units T4 DNA ligase per μg DNA at 14°C for 2 to 16 hr. Fragments to be cloned were present at an approximate four fold excess (moles of ends) to the vector.
Preparation of single-stranded template DNA (250). Insert-containing M13mpl8 or M13mpl9 (141) was transfected into competent MV1190 or JM101 cells, as described under "Transfection of competent E. coli". Individual plaques were stabbed with a sterile needle and inoculated into 2 ml YT media containing 40 µl of an overnight culture of MV1190 or JM101. The culture was incubated at 37°C with shaking for 5 hr. At the end of the incubation 1 ml of the culture was transferred to a 1.5 ml Eppendorf tube, and the cells were pelleted by centrifugation in an Eppendorf microfuge for 5 min. The supernatant was transferred to a 1.5 ml Eppendorf tube and 200 µl of 2.5M NaCl-20% PEG 8000 (w/v) was added. The tube was vortexed for 5 sec and incubated 30 min at room temperature. The tube was centrifuged for 5 min in an Eppendorf microfuge and the supernatant discarded. The M13 phage pellet was resuspended in 100 µl TE and extracted once with 100 µl phenol. The upper aqueous phase was transferred to a 1.5 ml Eppendorf tube and the single-stranded DNA was ethanol precipitated overnight at -20°C. The phage DNA was pelleted by centrifugation in an Eppendorf microfuge for 10 min, the supernatant was decanted, and the DNA pellet was dried under vacuum and resuspended in 10 µl TE. For large scale template preparation, 250 ml of YT (containing 30 µg/ml Cm if E. coli CJ236 was used) was inoculated with a 5 ml stationary-phase M13 culture. After an overnight incubation with shaking at 37°C the cells were
pelleted in a Sorvall GSA rotor, 5000 rpm for 10 min. The supernatant was collected and processed for single stranded template DNA as described; volumes for phage precipitation were adjusted accordingly. Phage pellet was resuspended in 2 ml TE and phenol extracted as described. After resuspension of template DNA, the concentration of single-stranded DNA was determined spectrophotometrically (1 A_{260} unit ≈ 20 μg/ml single-stranded DNA). Template DNA was stored in 10 μl aliquots at -70°C, used once for an oligonucleotide directed mutagenesis experiment or sequence analysis, and discarded.

Construction of large-intron substrates. DNA carrying the small (22 nucleotide) intron (10 μg) was cleaved with EcoRI and HindIII and the cleavage products were separated on an 8% nondenaturing polyacrylamide gel. The gel slice containing the DNA gene-specific fragment was removed from the gel and the restriction fragment was eluted overnight in RNA elution buffer. The DNA was precipitated, resuspended in 1x HinPl restriction buffer, and cleaved with HinPl for 1 hr to generate the two exon partial-intron fragments. The reaction was terminated by phenol extraction, and the products were recovered by ethanol precipitation and finally resuspended in 10 μl of H₂O. To isolate the 82 bp HinPl intron fragment, absent in the ω167 constructions, 10 μg of ω16 plasmid DNA was cleaved with HinPl and the restriction
fragments were separated on an 8% nondenaturing polyacrylamide gel. After visualization by EtBr staining, the gel slice containing the 82 bp HinPl intron band was cut out of the gel and eluted overnight, as described for RNA extraction. The intron fragment was ethanol precipitated, resuspended in 10 μl of H2O, and mixed with 2 μl of the exon partial-intron fragments and 250 ng of EcoRI-HindIII cleaved M13mp19 DNA. These DNA fragments were ligated in a final volume of 25 μl for 14 hr. An aliquot of the ligation mix (5 μl) was used to transfect competent E. coli MV1190. White plaques were picked into 2 ml YT media plus 40 μl stationary MV1190 in order to generate single stranded phage. To identify those plaques containing a full size intron, second-stranded synthesis analysis was conducted. Template DNA (4 μl) was mixed with 1 μl 10x "S" buffer (200 mM Tris-HCl pH 8.0, 100 mM MgCl2), 1 μl M13 universal primer (5 ng), 1 μl [α-32P] dATP (3000 Ci/mmol), 1 μl H2O and heated at 68°C for 7 min. This mixture was transferred to 23°C to allow primer-template annealing. To the annealed DNAs were added 2 μl of second-strand synthesis nucleotide mix (50 μM dGTP, dCTP, and TTP, plus 5 μM dATP) and 0.5 μl of Klenow fragment (2U). The reaction was incubated at 23°C for 15 min. To this reaction mixture was added 1.5 μl 10x restriction buffer, 12 μl H2O, 1 μl EcoRI (10 U), and 1 μl (10 U) HindIII. Double restriction digestion by EcoRI and HindIII releases the newly formed double stranded DNA. The
reaction was incubated 30 min at 37°C, DNA loading buffer was added, and the samples were subjected to electrophoresis on an 8% nondenaturing polyacrylamide gel. After electrophoresis was completed, the top glass plate was removed and the gel was covered with Saran wrap. In the darkroom, a sheet of X-ray film was placed on the gel for 2-5 min and then developed. To identify the correctly sized construct, the sizes of the cleaved intron-containing substrates were compared to standards also present on the gel, "16 or 016 plasmid DNA and "167 or 0167 plasmid DNA cleaved with EcoRl and HindIII. To determine the orientation of the intron HinP1 fragment and verify correct ligation, the templates of the putative positives were screened by dideoxy DNA sequence analysis.

Reconstruction of tRNA<sup>Trp</sup> (173). To reconstruct the gene encoding 0167, 6 pmol each of oligonucleotides "Trp2"-"Trp8" were phosphorylated, mixed (1.0 ml total volume) and heated to 60°C. After 5 minutes, 6 pmol each of unphosphorylated "Trp1" and "Trp9" were added (1.1 ml total volume) and the mixture was transferred to 90°C and incubated for 5 min. After heating, the mixture was slowly cooled to 23°C. The annealed oligonucleotides were ethanol precipitated, resuspended in 50 μl of 1x ligation buffer containing 50 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 5 mM DTT, 1 mM spermidine, 1 mM EDTA, and 1 mM ATP and ligated for 14 hr at 14°C. An
aliquot (5 μl) was removed and cleaved with BamH1 in a 100 μl reaction volume to generate the monomeric gene. The monomeric gene unit was separated by gel electrophoresis, purified, and ligated to BamH1-cleaved M13mp18 RF DNA. Correct gene assembly was verified by Sanger dideoxy DNA sequence analysis. To construct the intact (104 nucleotide) intron, an 82 bp HinP1 fragment, missing from the ol67 gene, was reintroduced into the constructed gene as described under "Construction of large-intron substrates". Proper reconstruction was verified by DNA sequence analysis and the complete gene, ol6, was subcloned into pUC18 for maintenance and propagation of the plasmid.

Construction of the minimal MinSub endonuclease substrate.
To construct the gene for Minsub, 32 pmol each of phosphorylated oligonucleotides "MinsubT" and "MinsubB" were combined with 2 μl TE (10 μl total volume), heated to 70°C for 10 min, and slowly cooled to 23°C. The annealed oligonucleotides were ligated into 1 μg of Sall-Smal digested pUC18 and transformed into E. coli TB1 for propagation of the plasmid DNA. Minsub was ligated into EcoR1-HindIII cleaved M13mp18 for confirmation of correct gene assembly by Sanger dideoxy DNA sequence analysis.

Oligonucleotide preparation. The oligonucleotides obtained from the OSU Biochemical Instrumentation Center were
debloked by heating at 58°C overnight. After deblocking, the oligonucleotide was placed on ice to cool and distributed into 1.5 ml Eppendorf tubes, 200 µl per tube. The oligonucleotides were dried in a Sorvall Speed-Vac, washed once with 100 µl H2O, and resuspended in 100 µl TE. The final DNA concentration was determined spectrophotometrically and the oligonucleotides were stored at -20°C.

Phosphorylation of oligonucleotides (250). To a 1.5 ml Eppendorf tube were added 240 pmol of oligonucleotide, 3 µl 10x T4 polynucleotide kinase buffer, 3 µl 10 mM ATP, 1 µl polynucleotide kinase (2 units), and H2O to a final reaction volume of 30 µl. This mixture was incubated for 30 min at 37°C. After incubation, 60 µl of H2O was added to bring the final concentration of the oligonucleotide to 2 pmol/µl. The phosphorylated oligonucleotide was stored at -20°C.

Phosphorylation of DNA and RNA (48, 131). DNA or RNA to be 5' phosphorylated was resuspended in a solution containing 25 µl T4 polynucleotide kinase buffer, 1 µl [γ-32P]ATP (100 µCi, 7000 Ci/mmol) if radiolabeling RNA or DNA, and T4 polynucleotide kinase (10 units). The reaction was carried out for 30 min at 37°C and terminated by extraction with an equal volume of phenol. The nucleic acid was ethanol
precipitated and, if RNA had been 5' end-labeled, gel purified as described.

Oligonucleotide-directed mutagenesis (115). Oligonucleotide directed mutagenesis reactions were prepared by labeling two 500 μl Eppendorf tubes as "+" and "-". To each tube was added, in a final volume of 10 μl, 0.1 pmol of template DNA, prepared from the dut-, unq- E. coli strain CJ236, 10 μg RNase A, 1 μl 10x annealing buffer (200 mM Tris-HCl pH 7.4, 20 mM MgCl₂, and 500 mM NaCl) and, for the reaction marked "+", 2 pmol of the phosphorylated mutagenic oligonucleotide. Both tubes were first incubated at 37°C for 10 min to degrade endogenous RNA which might function as a primer for nonspecific DNA synthesis, then incubated at 68°C for 7 min and slow cooled to 23°C to allow mutagenic primer to anneal to the template. To each tube was added 3 μl of the synthesis reaction mix containing 1 μl 10x synthesis buffer (10 mM ATP, 50 mM MgCl₂, 20 mM DTT, and 5 mM each dATP, dCTP, dGTP, TTP), 1 unit of T4 DNA ligase, and 1 unit of DNA polymerase large fragment. The tubes were incubated at 14°C for 5 hr, after which 90 μl of TE buffer was added to stop the synthesis reaction. An aliquot (20 μl) from each tube was used to transfect competent E. coli MV1190. If successful and specific second strand synthesis of the template had occurred, the transfection plates for the "+" reaction contained 100 fold greater number of plaques than
from the "-" reaction. From the "+" transfection plates eight plaques were picked into 2 ml YT media plus log phase E. coli MV1190 cells, and single stranded M13 phage were prepared as described. The presence of correct mutations were confirmed by Sanger dideoxy DNA sequence analysis.

Mutations utilizing oligonucleotides ^GGAG and C29:1 were introduced into o16. The altered substrate C29:1 was further mutated with oligonucleotide G146:1 to create the altered substrate C29:1-G146:1. The remaining mutations were first introduced into o16 7. The complete intron versions of these mutants were constructed by adding back the 82 bp HinPl fragment missing from the o16 7 gene. The HindIII-HinPl and EcoRl-HinPl fragments from the mutant o16 7 gene, which contained exon 1-partial intron and exon 2-partial intron respectively, were combining in a ligation mixture with the 82 bp HinPl fragment isolated from the wild type tRNATrp gene present in the ^16 clone, as described under "Construction of large-intron substrates". The complete mutated gene was then subcloned into pUC18 for maintenance and propagation of the plasmid (see below). C36:1 was created by combining mutant Exon 1 with wild-type Exon 2, while C36:1-G54:1 was generated by combining mutant Exon 1 with mutant Exon 2. Due to an eight of nine nucleotide match, the oligonucleotides C29:1, ^GGAG, and C36:1-G54:1 annealed to either the 5' exon or T7 RNA polymerase promoter sequences. To alleviate the formation
of deletion mutants, the following modification was employed during mutagenesis. A second 18 nucleotide oligonucleotide, "T7Bottom", was designed to be complementary, and anneal to, the T7 RNA polymerase promoter sequence, protecting this region from hybridizing with the mutagenic oligonucleotide. This oligonucleotide was included, 2 pmol, in these mutagenesis reactions.

Recloning M13-based substrates into pUC18 DNA. To subclone genes which were constructed in M13 vectors into plasmid pUC18 for propagation and maintenance, M13 template DNA (4 μl) was mixed with 2 μl H₂O, 1 μl 10x "S" buffer (200 mM Tris-HCl pH 8.0 and 100 mM MgCl₂), and 1 μl M13 universal primer (5 ng). The mixture was heated at 68°C for 7 min, slow cooled at 23°C for 7 min, 2 μl of 50 μM dNTPs and 0.5 μl Klenow fragment were then added, and the reaction was incubated for 15 min at 23°C. To this mixture was added 1.5 μl of 10x restriction buffer, 12 μl H₂O, 1 μl EcoR1 (10 U), and 1 μl HindIII (10 U) and incubation was continued for 45 min at 37°C. This procedure generated a double stranded fragment encompassing the gene region. The reaction was terminated by phenol extracted and the products were isolated by ethanol precipitation. The EcoR1-HindIII gene fragment was resuspended in 10 μl H₂O and added to a solution containing 0.5 μg EcoR1 and HindIII digested pUC18, 5 μl 5x T4 DNA ligase buffer, 1 μl T4 DNA ligase, and
sufficient H₂O to achieve a final volume of 25 µl. Ligation proceeded for 14 hr at 14°C. An aliquot of the ligation mix (10 µl) was used to transform competent *E. coli* TB1 cells. DNA from insert containing colonies were screened by restriction digest for the correct size inserts.

Preparation of competent *E. coli* (250). An isolated colony of *E. coli* from a minimal media plate was inoculated into 5 ml of YT media and shaken overnight at 37°C. An aliquot (500 µl) of this culture was used to inoculate 50 ml of YT media in a 125 ml or 250 ml sterile flask. The culture was shaken at 37°C until A₂₆₀ ≈ 0.5, about 2-3 hours. The cells were harvested by centrifugation in a sterile Oakridge tube in a Sorvall SS-34 rotor, 5000 rpm for 10 min at 4°C. The supernatant was discarded and the cell pellet was gently resuspended in an equal culture volume of ice-cold 10 mM MgSO₄. The cells were kept on ice for 30 min and then pelleted as described. After the supernatant was discarded, the pellet was resuspended in one-half volume of the original culture volume of ice-cold 50 mM CaCl₂ and kept on ice for an additional 30 min. The cells were harvested once more, and the cell pellet was resuspended in 0.5 ml of ice-cold 50 mM CaCl₂. The cells were then competent for transformation or transfection, as appropriate for the *E. coli* strain used, and stored at 4°C for up to 24 hr until needed.
Transformation of competent *E. coli*. DNA (0.1-1.0 μg) was combined with TE to a final volume of 50 μl in a sterile tube and placed on ice. Competent *E. coli* cells (100 μl) were added to the contents of the tube with gentle mixing, the reaction mixture was incubated on ice for 30 min, and then heat-shocked at 37°C for 2 min. YT media (900 μl) was added and the culture was incubated with shaking at 37°C for 1 hr to allow for the expression of antibiotic resistance. Aliquots (100 μl) were plated onto YT plates containing antibiotic (ampicillin at 60-75 μg/ml) and incubated inverted at 37°C. If the *E. coli* strain TB1 was transformed with a pUC-based plasmid vector and colonies were to be screened for inserts by α-complementation, cells were plated onto plates containing 60 μg/ml ampicillin and 0.4 mg/ml X-gal. pUC18 carries a portion of the lacZ operon interrupted by the in-frame multiple cloning block. TB1 also has a portion of the lacZ operon incorporated into its chromosome, such that the proteins resulting from these two operons can undergo α-complementation to produce a functional β-galactosidase enzyme. β-galactosidase can cleave X-gal, which when cleaved turns from white to blue. Any DNA cloned into the multiple cloning site of pUC18 will disrupt the reading frame, and the ensuing gene product will not be able to α-complement the gene product from the chromosomal lacZ product. In this situation X-gal will not be cleaved and the colony and will therefore remain white. Those cells
carrying only wild-type pUC18 will appear blue on YT plates containing X-gal.

**Transfection of competent E. coli.** Transfection of competent *E. coli* JM101, MV1190, or CJ236 was performed as described for the transformation procedure with the following modifications. After heat-shock at 37°C for 2 min, 50 and 100 μl of transformed cells were transferred to tubes containing 200 μl log phase recipient cells. These cells were required to ensure the formation of a lawn on the transfection plate. Molten YT top agar (YT media containing 0.7% agar; 3 ml) was added to the cell mixture and poured onto YT plates. After the top agar had hardened, the plates were incubated at 37°C until plaques were observed, generally in 8-16 hr. If color selection of plaques was desired, as described for pUC18-based vectors, 50 μl of 2% X-gal in dimethylformamide and 10 μl of 100 mM IPTG (to induce lacZ) were added to the 3 ml of molten top agar before plating cells. White plaques represented those *E. coli* colonies transfected with phage bearing an insertion in the M13 cloning site, while blue plaques bore no insert.

**Analysis of 5' end nucleotides** (109). Unlabeled RNA was cleaved by endonuclease and the cleavage products separated by electrophoresis through an 8.3M urea denaturing polyacrylamide gel (6% or 20%). The RNA was visualized by
UV shadowing and eluted from gel slices as previously described. The resulting RNA was phosphorylated at its 5' terminus in 1x T4 polynucleotide kinase buffer with 125 µCi [γ-32P]ATP and T4 polynucleotide kinase (2 units) for 30 min at 37°C. The reaction was terminated by phenol extraction and the labeled RNAs were recovered by precipitation and subjected to electrophoresis. The 5' end-labeled RNA, visualized by autoradiography, was purified from the gel as described. For recovery of small RNAs 20 µg carrier tRNA was added to the RNA elution buffer. Labeled RNA was resuspended in 8 µl of Nuclease P1 buffer containing 5 mM NH₄COOH₃ pH 5.0, plus 2 µl (2 units) of Nuclease P1, and the reaction was incubated for 4 hr at 37°C. Nuclease P1-digested RNA was spotted onto a cellulose TLC plate (Sigmacell Type 100 with 254 nm fluorescent indicator). Unlabeled nucleotide markers (5 µg each of pA, pC, pG, pU) were spotted in an adjacent lane for product identification. The TLC plate was placed in a development tank at 25°C containing 100 ml of chromatographic buffer, either Solvent System I (66:1:33 isobutyric acid/concentrated NH₄OH/H₂O) or Solvent System II (80:18:2 saturated ammonium sulfate/1 M sodium acetate/isopropanol). Plates were allowed to develop until the solvent front reached the top of the TLC plate. The plate was removed from the tank, allowed to air dry, wrapped in Saran wrap, and subjected to autoradiography to
identify labeled nucleotides. The mononucleotide standards were visualized by fluorescence under UV light.

**RNA sequence analysis** (48, 49, 127). To determine the RNA sequence of 5'-end labeled RNA, five duplicate reactions were set up. Into tubes 1-5 were added 5'-end labeled RNA (10,000 cpm) and 5 μg carrier tRNA. To tubes labeled 1-3 were added 2 μl of solution A (0.25 M sodium citrate pH 5), tube 4 received 2 μl of solution B (0.25 M sodium citrate pH 3.5), and tube 5 received 2 μl of Solution C (0.1 M NaPO₄ pH 6.5, 1 mM EDTA). To tubes 1, 2, and 4 were added 7 μl of solution D (10 M urea, 1.5 mM EDTA, and 0.5% each xylene cyanol and bromphenol blue) and sufficient water for a total volume of 9 μl. Tubes 2 and 5 received sufficient water for a final volume of 4 μl. All RNases were diluted to 0.5 and 2 U/μl in 1x of their reaction buffer before use (T₁, PhyM, and B. cereus in solution A, U₂ in solution B, and CL₃ in solution C). To the first set of five tubes were added 1 μl of enzyme diluted to 0.5 U/μl, while the second set received enzyme diluted to 2 U/μl. Tube 1 received RNase T₁, tube 2 received RNase PhyM, tube 3 received RNase B. cereus, tube 4 received RNase U₂, and tube 5 received RNase CL₃. All tubes were incubated for 15 minutes; tubes 1-4 were incubated at 55°C and tube 5 at 37°C. The reactions in tubes 1, 2, and 4 were terminated by chilling in a dry ice-ethanol bath. To terminate to reaction in tubes 3 and 5, 10 μl of solution D
was added to each mixture and the tubes were chilled in a dry ice-ethanol bath. To generate an alkaline hydrolysis RNA sizing ladder, end labeled RNA (40,000 cpm) was combined with 5 µg yeast tRNA, 1 µl of a 0.5 M sodium bicarbonate/carbonate pH 9.2, and sufficient water for a final volume of 10 µl. The mixture was incubated at 90°C for 7 min and then chilled in a dry ice-ethanol bath. An aliquot from each reaction (5 µl) was loaded onto a large 20% denaturing polyacrylamide gel and the cleavage products were separated by electrophoresis. Electrophoresis was terminated when the bromphenol blue dye had migrated 6 cm from the origin. The samples were visualized by autoradiography.

**Analysis of 3' end nucleotides from excised intron (109).** The ^16 [α-32P]ATP-labeled RNA was cleaved in a standard cleavage assay and the products were isolated by denaturing gel electrophoresis as described above. The radiolabeled intron band was visualized by autoradiography and the RNA was purified as described. The [α-32P]ATP-labeled intron RNA was resuspended in a solution containing 7 µl of 10 mM Tris-HCl pH 8.0, plus 2 µl of 5 mg/ml carrier tRNA. This mixture was heated at 100°C for 1 min, quick chilled on ice, and 1 µl (5 mg/ml) of RNase A was added. The RNA was digested for 5 hr at 37°C. At the end of the incubation period 10 µl of Solution D (BRL RNA sequencing kit; 10 M
urea, 1.5 mM EDTA, 0.05% each xylene cyanol and bromphenol blue) was added, and the RNase A-digested RNA was subjected to electrophoresis in an 8.3 M urea 20% polyacrylamide denaturing gel. The gel slice containing the labeled RNase A oligonucleotide (GpGpApGpApUp), corresponding to the 3' terminus, was cut out of the gel after visualization by autoradiography, and the RNA was precipitated, as previously described. The hexanucleotide was resuspended in a solution containing 2.3 µl H2O, 0.5 µl 0.25 M sodium citrate buffer pH 5.0, 1 µl carrier tRNA (5 mg/ml), and 1.2 µl Ribonuclease T2 (10 units). The oligonucleotide was digested for 2 hr at 37°C and then spotted onto a cellulose TLC plate (Sigmacell Type 100 with fluorescent indicator), previously prerun in saturated ammonium sulfate. Mononucleotide markers were run in an adjacent track for nucleotide identification. The TLC plate was developed in a tank containing 100 ml of chromatography buffer (80:18:2 saturated (NH4)2SO4/1 M NaCOOCH3/CH3CH2OH; (68)); development continued until the solvent front reached the top of the TLC plate. The plate was removed from the tank, allowed to air dry, wrapped in Saran wrap, and autoradiographed at -70°C to visualize radiolabeled nucleotides. Mononucleotide markers were visualized by fluorescence under UV light.

**Synthesis of pCp.** The following reagents were mixed together in a 1.5 ml Eppendorf tube: 10 µl 750 µM Cp, 5 µl
10x T4 polynucleotide kinase buffer, 6 μl (γ-\(^{32}\)P)ATP (7000 Ci/mmol), 28 μl H₂O, and 1 μl T4 polynucleotide kinase (5 units). This mixture was incubated overnight at 37°C. The reaction was terminated by boiling for 3 min, and denatured protein was removed by centrifugation for 5 min. The resultant supernatant was transferred to an Eppendorf microfuge tube and the radiolabeled pCp was stored at -20°C.

3' pCp end-labeling of RNA (173). The following reagents were combined in a 1.5 ml Eppendorf tube: 13 μl RNA, 25 μl 2x "partial mix", 4 μl 100 mM DTT, 5 μl radiolabeled pCp, 1 μl 10 mM ATP, 1 μl 1 mg/ml BSA, and 1 μl T4 RNA ligase (0.5 μg). This mixture was incubated overnight at 4°C. Labeled RNAs were recovered by ethanol precipitation, separated by gel electrophoresis, and purified as described.

M13 Sanger dideoxy DNA sequencing (141, 142). Sequencing of DNA was done by the dideoxynucleotide chain termination method of Sanger, et al. To prepare the sequencing reactions, four 500 μl Eppendorf tubes were labeled as A, G, C, or T and placed at 23°C. To these tubes were added, respectively, 3 μl of A, G, C, and T nucleotide mix. In a separate 500 μl Eppendorf tube single stranded template DNA (1 μg) was combined with 1 μl of annealing buffer (200 mM Tris-HCl pH 8.0 and 100 mM MgCl₂), 1 μl M13 primer (2 ng), 2 μl [α-\(^{32}\)P]dATP (3000 Ci/mmol), and sufficient water for a
final volume of 10 µl. The mixture was heated at 68°C for 7 min and allowed to slowly cool to 23°C to allow template and primer to anneal. To this mixture 2 µl of T4 DNA polymerase large fragment (1 U/µl) was added and 2 µl aliquots were added to each of the four tubes containing the nucleotide mixes. After the contents were gently mixed the reactions were incubated at 23°C for 15 min and then 1 µl of chase solution (50 µm dNTPs) was added. Incubation continued for an additional 15 min, the reactions were terminated by the addition of 10 µl of DNA sequencing-loading buffer (95% formamide and 0.1% w/v each of xylene cyanol and bromphenol blue in 1x TBE), and 5 µl aliquots from each reaction were loaded onto a large 6% denaturing polyacrylamide gel. Sequencing products were separated by electrophoresis and visualized by autoradiography.

**Kinetics assays.** The procedures for determining the \( K_m \) value for o16 and o16\(^7 \) were identical. Radiolabeled and nonradiolabeled RNA transcripts were generated from o16 and o16\(^7 \). The \( A_{260} \) absorbance value of the nonradiolabeled substrate was determined and the RNA concentration was calculated; the calculated extinction coefficient, based on nucleotide content, for o16 was \( 1.94 \times 10^6 (M^{-1}cm^{-1}) \) and the extinction coefficient for o16\(^7 \) was \( 1.05 \times 10^6 (M^{-1}cm^{-1}) \). A series of reactions were set up containing 5, 10, 50, 100 or 250 mM of nonlabeled RNA in a final volume of 45 µl of TMG.
Sufficient radiolabeled RNA was added to yield 100,000 cpm labeled RNA/nM intact unlabeled RNA at the 5 and 10 nM assay substrate concentrations, and 50,000 cpm/nM for the 25-250 nM assay substrate concentrations. The radiolabeled RNA was estimated to have a specific activity of 100,000 cpm per nM. DEAE-purified endonuclease (200 ng) was added to the reaction mix and 10 µl aliquots were removed into 10 µl of stop buffer (10 M urea, 1.5 mM EDTA, 0.05% xylene cyanol, 0.05% bromphenol blue) at time 0, 2, 4, and 6 min. The samples were loaded onto an 8.3 M urea 6% (016) or 10% (0167) denaturing polyacrylamide gel and subjected to electrophoresis. The cleavage products were visualized by autoradiography, the gel slices containing the bands corresponding to the intron were cut out of the gel, and Cerenkov radiation was determined by liquid scintillation. Nmol of intron released were determined and plotted versus time to obtain the velocity of intron release. k_m values were determined using Lineweaver-Burke analysis.

Structure probing analysis (124, 206). Unlabeled, 5'-GMP terminated RNA was synthesized and gel purified as described. The RNA was resuspended in 20 µl 10 mM Tris-HCl pH 7.5 containing 2 units bacterial alkaline phosphatase and incubated at 42°C for 1 hr to remove the 5' phosphate group. The dephosphorylated RNA was extracted with an equal volume of phenol, ethanol precipitated, and 5' phosphorylated with
[\gamma^{32}P]ATP as described in "Phosphorylation of DNA and RNA". Labeled RNA was resuspended in 30 µl H₂O. For S1 cleavage, 50 µl of S1 buffer, 10 µg yeast carrier tRNA, and sufficient S1 nuclease to give a ratio of 0, 0.1, 0.5, and 2 units/µg yeast tRNA was added to four separate tubes. Each tube then received 100,000 cpm of end labeled RNA and the tubes were incubated for 10 min at 37°C. The reactions were terminated with the addition of an equal volume of phenol, and the RNA ethanol precipitated. The RNA samples were resuspended in 10 µl RNA loading buffer and loaded onto a large (32 x 40 cm) denaturing 6% or 10% polyacrylamide gel. Uniformly [\alpha^{32}P]ATP-labeled 0I6 or oI67 RNA, previously cleaved with endonuclease, were also run as size markers. After electrophoresis the RNA bands were visualized by autoradiography and the sizes of the 5' end labeled RNAs, resulting from cleavage of the full length RNA by nuclease S1, were determined by comparison of their mobilities to the mobilities of the RNA bands generated from the endonuclease-cleaved RNA.

Protein precipitation for SDS-PAGE. To precipitate protein in solution (100-500 µl) an equal volume of 80% ice-cold TCA was added to the solution. After mixing, the solution was incubated on ice for 30 min and centrifuged for 10 min in an Eppendorf microfuge at room temperature to pellet proteins. The supernatant was discarded and the protein pellet was
dried under vacuum. The protein pellet was resuspended in SDS protein buffer and the proteins were separated by SDS-PAGE.

**Determination of protein concentration.** Total protein concentration was determined by reacting known volumes of extract with 200 µl BioRad Protein Assay Dye Reagent concentrate and water in a total volume of 1 ml, according to the manufacturer's directions. A595 values of unknown protein samples were determined and compared to a standard curve generated with known amounts of BSA; the protein concentration of the unknown samples were interpolated from the standard curve.

**SDS-PAGE.** SDS-PAGE was performed according to the method of U.K. Laemmli using 10% or 12% polyacrylamide (118). The gel was fixed and stained for 30 min in a solution containing 50% methanol, 10% acetic acid, and 0.2% Coomassie Brilliant Blue R, followed by overnight destaining in 10% acetic acid.
Chapter IV

Results

Development of an endonuclease assay.

To develop an assay for intron endonuclease, a suitable intron-containing tRNA substrate was needed. Previous work by Daniels et al. described the plasmid vector pT72HM^16 (40). This vector was constructed by subcloning a TaqI partial digest from pHMT9, a genomic clone from H. mediterranei coding for pretRNA^TrP, into the T7 RNA polymerase promoter transcription expression vector, pT72 (Figure 5). This subclone was further modified by deletion of some exogenous DNA sequences with exonuclease III to yield pT72HM^16, hereafter referred to as ^16. After cleavage at the 3' terminus of the gene with EcoR1, ^16 is transcribed by T7 RNA polymerase to yield a 219 nucleotide pretRNA^TrP transcript (Figure 6). In addition to the coding region this transcript contains a 34 nucleotide 5' leader region of vector sequence and two nucleotides at the 3' terminus. This transcript is similar in size to the predicted in vivo transcript of 185 nucleotides (40).
Figure 5. Construction of the *H. mediterranei* tRNA<sub>Trp</sub> encoding plasmid pT72HM<sup>16</sup>. The plasmid pHMT9 (37) was partially digested with TaqI and the tRNA<sub>Trp</sub> gene was ligated into TaqI-digested pT72 plasmid to generate the clone pT72HM. This plasmid was cleaved with BamHI and SstI, unidirectionally digested with exonuclease III to remove extragenic 3' sequences, and religated to create the plasmid pT72HM<sup>16</sup>. This plasmid was linearized with EcoRI and transcribed by T7 RNA polymerase with all four nucleotides to produce the pretRNA<sub>Trp</sub> substrate RNA used in endonuclease assays.
Figure 5.
Figure 6. Structure and cleavage assay of H. mediterranei tRNA^Trp precursor. A, primary sequence of the T7 RNA polymerase transcript (°16) produced from linearized pT72HM°16 plasmid. This plasmid contains the intact H. mediterranei tRNA^Trp gene. The five nucleotide changes which occur in the H. volcanii tRNA^Trp gene are indicated. Cleavage sites for endonuclease are designated by arrows, and the anticodon is enclosed in the box. The sequences which correspond to restriction sites used in the construction of the substrate clones pT72HM°167 (HinPI), pT72HV°13 (HaeIII), and the pT72HM°13 derivatives (BcII) are indicated. B, cleavage of [α-°32p]ATP labeled °16 substrate by the H. volcanii endonuclease. Products were analyzed by electrophoresis in 6% denaturing polyacrylamide gels. The sizes and identities of the products are indicated.
Figure 6.
Initial endonuclease assay conditions were established using a crude cell extract from *Halobacterium volcanii*, prepared in *H. volcanii* lysis buffer. When ~16 RNA was incubated in the presence of crude *H. volcanii* cell lysate, three products were generated which correspond in size to the intron (104 nucleotides), the leader plus 5' exon (75 nucleotides), and the 3' exon plus 3' flanking sequence (43 nucleotides; Figure 6). The identities of these cleavage products were confirmed by Northern blot analysis (unpublished). This established that the ~16 RNA could be used as a substrate for the characterization of the halobacterial tRNA intron endonuclease.

Purification of endonuclease.

Based on the ~16 RNA assay a partial purification procedure for endonuclease activity was developed (Table 4). Endonuclease activity was followed at each purification step by assaying fractions for endonuclease activity and retaining those fractions which generated the expected products in a cleavage assay. To quantitate the recovery of endonuclease, each fraction was incubated for 60 min with an equivalent amount of [α-32P]ATP-labeled ~16 substrate and the cleavage products were separated on a 6% denaturing polyacrylamide gel. Gel slices corresponding to the released intron were cut out of the gel and the radiolabel was quantitated by measuring the Cerenkov radiation. One
Table 4

Purification of *Halobacterium volcanii* tRNA<sup>Trp</sup> Endonuclease

<table>
<thead>
<tr>
<th>Step</th>
<th>Protein (mg)</th>
<th>Total Activity (units)</th>
<th>Specific Activity (units/mg)</th>
<th>Purification Fold</th>
<th>Yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>French Press Lysate</td>
<td>800.00</td>
<td>2228</td>
<td>2.8</td>
<td>1.0</td>
<td>100.0</td>
</tr>
<tr>
<td>100,000g Supernatant</td>
<td>685.00</td>
<td>1014</td>
<td>1.5</td>
<td>0.5</td>
<td>45.5</td>
</tr>
<tr>
<td>10% PEG Supernatant</td>
<td>104.00</td>
<td>693</td>
<td>6.7</td>
<td>2.4</td>
<td>31.1</td>
</tr>
<tr>
<td>DEAE Sephacel I</td>
<td>13.40</td>
<td>990</td>
<td>73.9</td>
<td>26.6</td>
<td>44.4</td>
</tr>
<tr>
<td>DEAE Sephacel II</td>
<td>4.80</td>
<td>574</td>
<td>119.6</td>
<td>43.0</td>
<td>25.8</td>
</tr>
<tr>
<td>Hydroxylapatite</td>
<td>0.24</td>
<td>159</td>
<td>651.0</td>
<td>234.3</td>
<td>7.1</td>
</tr>
</tbody>
</table>
unit of endonuclease activity was defined as the amount of endonuclease necessary to cleave 5 ng of ^16 substrate RNA to the maximum possible extent (47% of total counts are represented in the intron of [α-32p]ATP labeled intron) in 60 min. Specific activity and fold purification were then calculated at each step of the purification (Table 4).

Both selective precipitation and column chromatography procedures were utilized in the purification of endonuclease. A PEG precipitation step was used to separate the intron endonuclease activity from RNase P, whose activity (removal of 5' leader from pretrRNAs) would complicate product analysis. Using DEAE Sephacel chromatography, the endonuclease activity eluted at low KCl concentrations, prior to the elution of the majority of the nonactive protein. The active fractions from the DEAE Sephacel column were passed over DEAE Sephacel again to remove any remaining endogenous RNAs; endogenous RNAs carrying through in the purification result in a high background when nonradioactive products were end labeled prior to RNA sequence analysis. A second chromatography step, using hydroxylapatite, provided additional resolution of the active protein (Figure 7). SDS polyacrylamide gel electrophoresis of proteins from each of the purification steps indicated a decrease in the number of proteins present and the increase in the amounts of a few protein bands with
Figure 7. Hydroxylapatite column chromatography of *H. volcanii* tRNA endonuclease. Dialyzed enzyme from DEAE Sephadex chromatography was applied to the column and eluted with a linear gradient of $\text{PO}_4$. Assays were performed with $^{16}$RNA substrate.
Figure 7.
masses of 54 kdal, 38.5 kdal, 29.5 kdal, and 20 kdal (data not shown).

Several additional chromatography matrices were tested but did not provide additional resolution for protein purification. These matrices included QAE Sephadex (a strong anion exchanger), SP Sephadex (a strong cation exchanger), heparin agarose, Affigel Blue (a chromatographic matrix often utilized in the purification of DNA binding proteins), phenyl sepharose, and FPLC with the matrix Mono S (a strong cation exchanger).

Characteristics of the cleavage reaction.

Fidelity and mechanism of cleavage.

To confirm that cleavage had occurred accurately and precisely at the sites predicted from the mature tRNA_Trp sequence (Figure 6), unlabeled 16 pretRNA was cleaved by the endonuclease and the intron and 3' exon were isolated from a denaturing polyacrylamide gel. These RNAs were labeled at their 5' termini and subjected to enzymatic RNA sequence analysis. The RNA sequences obtained from this analysis (Figure 8) agreed with the sequence predicted for accurate cleavage of the precursor at both exon-intron boundaries. These data, and the absence of other cleavage products, demonstrate that the cleavage of 16 pretRNA by the endonuclease is both specific and accurate.
Figure 8. RNA sequence analysis of the endonuclease products. Intron and 3' exon molecules produced from the cleavage of unlabeled 16 substrate were isolated by gel electrophoresis, purified, and end labeled at their 5' termini with [γ-32P]ATP and polynucleotide kinase. Their sequences were determined by the enzymatic sequencing method, using the ribonucleases T1 (G), U2 (A), PhyM (A+U), Bacillus cereus (C+U), and CL3 (C). Products were analyzed by electrophoresis on 20% denaturing polyacrylamide gels.
Figure 8.
Since the 5' terminus of the intron and 3' exon products could be end labeled without phosphatase treatment, a 5'-OH must have been available for phosphorylation. This is consistent with the mechanism of phosphodiester bond cleavage proposed for the yeast nuclear tRNA intron endonuclease. As shown in Figure 4, cleavage of the phosphodiester bonds at the exon-intron boundaries by the yeast endonuclease results in a nearest neighbor transfer of the phosphate to produce 5'-OH and a 2',3' cyclic phosphate termini. Since a 5'-OH terminus was present on the intron and exon 2 after cleavage with the halophilic endonuclease, it was likely that nearest neighbor transfer had occurred and that the 2',3' cyclic phosphate termini would also be present.

A phosphate transfer was first investigated for the cleavage at the intron-exon 2 boundary. The cleavage mechanism predicted a nearest neighbor transfer of the α-phosphate from the 5' end of exon 2, A, to the 3' terminus of the intron, U, during the cleavage reaction (Figure 4). To assay for the transfer of a phosphate to the labeled 3' U of the intron, uniformly [α-32P]ATP-labeled ^16 transcript was cleaved by the endonuclease and the intron was isolated. This RNA was digested with Ribonuclease A, which cleaves specifically on the 3' side of U and C residues to generate a mixture of labeled RNA oligonucleotides (Figure 9). The largest oligonucleotide, representing the hexanucleotide
Figure 9. Determination of the 3' terminal nucleotides. A, isolated intron from endonuclease-cleaved [α-32P]ATP-labeled ^16 was subjected to digestion by RNase A. RNase A products were separated by electrophoresis on 20% denaturing polyacrylamide gels. The unique hexanucleotide derived from the 3' terminus is shown. Labeled phosphates originating from the input label and through transfer during the cleavage reaction are indicated by dots. B, purified hexanucleotide was digested with RNase T2, and the products were separated by thin layer chromatography. Mononucleotide markers are shown. C, 5' exon (lane 1) and intron (lane 2) products from [α-32P]ATP-labeled ^1311-5 substrate (Figure 15) were digested with nuclease P1. The products, pA derived from the input label and the unique pNP mononucleotides from the 3' termini, were separated by thin layer chromatography.
Figure 9.
GpGp*ApGp*ApUp* from the 3' terminus of the intron, was isolated and digested with Ribonuclease T2. This enzyme cleaves without nucleotide specificity to release 3' phosphate mononucleotides (Np). The resulting ribonuclease T2 products were separated by thin layer chromatography (Figure 9). The autoradiogram of the TLC plate indicated the existence of a radiolabeled U nucleotide, and that the expected labeled nucleotides were present in the predicted stoichiometric ratio of one Up* to two Gp*.

Since RNase T2 is known to cleave 2',3' cyclic phosphates to the 2' or 3' derivatives, this analysis could not establish whether a cyclic phosphate was present at the 3' termini of the intron or the 5' exon. To detect these phosphate groups, uniformly [α-32P]ATP labeled ~1311-5 (Figure 15) was cleaved with the endonuclease. The intron and 3' exon RNAs were isolated and digested with nuclease P1. Nuclease P1 cleaves without nucleotide specificity to yield 5' phosphate products, pN, and does not open 2',3' cyclic phosphates. Cleavage of the endonuclease products would produce p*A and a single unique pNp derivative from the 3' terminus. Analysis of the mononucleotide products by TLC demonstrated the presence of a 2',3' cyclic G phosphate derived from the 3' terminus of the 5' exon, and a 2',3' cyclic U phosphate derived from the 3' terminus of the intron (Figure 9). The mechanism of substrate cleavage by the halobacterial intron endonuclease appears to be very
similar, if not identical, to that employed by its eukaryotic counterpart.

**Biochemical properties.**

Nuclear mRNA intron processing requires the formation of a spliceosome before splicing of the precursor can occur. To determine the rate of cleavage and to test for assembly of a cleavage complex, a time-course for substrate cleavage by the endonuclease was performed. Beginning at time 0 min, when the endonuclease was added to a cleavage reaction, aliquots were removed at time 1, 2.5, 5, 10, and 30 min into an equal volume of phenol to terminate the cleavage reaction. The RNAs were precipitated and separated by electrophoresis through a 6% denaturing polyacrylamide gel; the autoradiogram of this assay is shown in Figure 10. No lag in the onset of cleavage was observed; cleavage began within one minute of incubation with the endonuclease and remained linear through the 30 minute reaction period.

A pH optimum curve was determined for the endonuclease at half-pH intervals from pH 5.5 to pH 9.5. This data (not shown) indicated the endonuclease has a broad pH optima, with maximum activity between pH 7.0 and pH 8.0.

A temperature optimum curve was also determined for the endonuclease. Cleavage assays were conducted at 25°C, 37°C, 42°C, 58°C, 68°C, and 75°C. As shown in Figure 11, the
Figure 10. Time dependence of substrate cleavage with *H. volcanii* endonuclease. Autoradiogram of time-dependent substrate cleavage with ^16 substrate RNA. Aliquots were removed from the reaction mixture at the indicated times and the reaction were terminated by phenol extraction of the RNA products. Identities of the RNA products are indicated.
Time (min)

1  2.5  5  10  30

pretRNA

intron

Exon 1

Exon 2

Figure 10.
Figure 11. Determination of temperature optima for H. volcanii and T. acidophilum intron endonucleases. Autoradiogram of cleavage assays for H. volcanii (left panel) and T. acidophilum (right panel) endonucleases. Assays were performed with ol6 RNA substrate. Assay temperatures and identities of RNA products are indicated.
Figure 11.
temperature optimum for the endonuclease is 42°C, which corresponds with its optimal growth temperature.

Mg$^{++}$ ions, which assist in the maintenance of nucleic acid tertiary structure (178), are required in the reactions for many RNA processing enzymes. To investigate the dependence of the endonuclease on the presence of Mg$^{++}$ ions, an assay was performed in which the concentration of MgCl$_2$ was varied from 0.5 mM to 50 mM (Figure 12). The endonuclease required a minimum concentration of 5 mM Mg$^{++}$ ions for efficient substrate cleavage; maximum activity was observed with 20 mM MgCl$_2$. Other divalent metal ions, including Ca$^{++}$ and Mn$^{++}$, were also evaluated; both Ca$^{++}$ and Mn$^{++}$ could substitute for Mg$^{++}$ in a cleavage assay. While the apparent intron release by the endonuclease was highest in the presence of Ca$^{++}$ ions, the nonstoichiometric recovery of exons suggested additional cleavage and Ca$^{++}$ was therefore not used as the divalent ion in the standard assay buffer. Spermidine, which stabilizes tRNA secondary structure (178), was also tested as a suitable polycation for the endonuclease. Spermidine, at concentrations as low as 0.5 mM, could replace 5 mM MgCl$_2$ in the cleavage assay (Figure 12) and was used in all subsequent cleavage assays, except as noted.

Because the internal monovalent ion concentrations of the halobacteria are in excess of 2 M (117), it was possible that added monovalent ions would increase the activity of
Figure 12. Ion requirements of *H. volcanii* tRNA endonuclease. Assays were performed with $^{16}$RNA substrate.
Figure 12.
the endonuclease. An assay was performed to evaluate the effect of monovalent ions on substrate cleavage. In the absence of Mg^{++} ions KCl did not support activity at concentrations up to 100 mM (Figure 12); in addition, activity could not be detected in extracts prepared in 2 M KCl (data not shown). When added to the Mg^{++}-containing TMG assay buffer, K^{+} ion concentrations greater than 100 mM significantly inhibited the activity of the endonuclease, and at 250 mM the activity was zero (Figure 12). Similar results have been obtained with Na^{+} ions (data not shown).

Substrate Specificity

Inability to cleave a nonhalophilic precursor tRNA.

All yeast nuclear pre-tRNAs are recognized and cleaved by one intron endonuclease (74, 165). Because the yeast and halophilic endonuclease cleavage mechanisms were very similar, it was possible that their substrate recognition properties would also be similar. To determine whether the halophilic tRNA intron endonuclease was capable of cleaving heterologous intron-containing tRNA precursors, a Saccharomyces cerevisiae intron-containing pre-tRNA, tRNA^{Phe}, was evaluated as a substrate for the enzyme (Figure 13). This RNA contains a 19 nucleotide intron located one nucleotide 3' from the anticodon and has been used to characterize the substrate recognition properties of the yeast nuclear-encoded tRNA intron endonuclease (172). When
Figure 13. Structure and cleavage assays for $^\cap 16$ and yeast pretRNA$^{Phe}$ RNAs. The left panel are the proposed secondary structures of the T7 RNA polymerase generated transcripts derived from the $^\cap 16$ and yeast pretRNA$^{Phe}$ genes. Cleavage sites for the endonuclease are indicated by arrows. The right panel is the autoradiogram of the cleavage assays with these substrate RNAs. Lane 1, $^\cap 16$ RNA without endonuclease; lane 2, $^\cap 16$ RNA with halophilic endonuclease; lane 3, pretRNA$^{Phe}$ with halophilic endonuclease; lane 4, pretRNA$^{Phe}$ without endonuclease. Sizes of RNA products are indicated.
Figure 13.
radiolabeled tRNA^{Phe} and \(^{16}\) RNAs were incubated with the halophilic endonuclease the \(^{16}\) RNA was readily cleaved, while pre-tRNA^{Phe} RNA was not, either specifically or nonspecifically (Figure 13). Since the endonuclease was capable of discriminating between these two intron-containing precursors, the halophilic precursor must have sequence or structural features which are lacking in the yeast precursor.

Requirement for intact precursor.

To determine if intact tRNA-like structure, an essential requirement for the eukaryotic tRNA intron endonuclease (62, 144), was required for substrate recognition, a gene encoding an exon-deletion derivative of \(^{16}\) was constructed. A HaeIII restriction endonuclease fragment from the plasmid pVT9, which encoded the \(H.\) volcanii tRNA^{Trp} homolog of \(^{16}\) (106 nucleotide intron rather than 104 nucleotides), was subcloned into the T7 RNA polymerase promoter expression vector pT71 to create the deletion substrate \(^{13}\). The transcript from this gene lacked the first 10 nucleotides of the 5' exon and the last five nucleotides of the 3' exon; while some aspects of the tRNA-like conformation could presumably be formed by the RNA transcript, it would lack the aminoacyl acceptor stem (Figure 14). The \(^{13}\) RNA was cleaved by the endonuclease in a cleavage assay (Figure 14). An exonuclease III digestion
Figure 14. Structure and cleavage assays of ^16, ^13, and ^13\textsuperscript{11} RNAs. The top panel is the proposed secondary structures of the T7 RNA polymerase generated transcripts derived from the ^16, ^13, and ^13\textsuperscript{11} genes. Cleavage sites for the endonuclease are indicated by arrows. The bottom panel is an autoradiogram of cleavage assays performed with ^16, ^13, and ^13\textsuperscript{11} substrate RNAs. Identities of RNA products are shown. Nonspecific cleavage products are attributed to nonspecific degradation of RNA due to contaminating ribonucleases.
Figure 14.
from the 3' end of \(^\text{13}\) resulted in the creation of \(^\text{13}^{11}\), which retained only the first seven nucleotides of the 3' exon. The transcript from this gene, in addition to lacking the aminoacyl acceptor stem, could not form the T\(\Psi\)C stem and loop of a mature tRNA (Figure 14), and therefore would not assume conventional tRNA tertiary structure. When \(^\text{13}^{11}\) RNA was assayed in a cleavage assay, discrete cleavage products of the predicted size were observed (Figure 14), indicating cleavage had occurred at the wild type cleavage sites. A more extensive exonuclease III generated deletion substrate, \(^\text{13}^{13}\), was also examined. The \(^\text{13}^{13}\) gene sustained a deletion up to the 3' exon-intron junction such that the resulting transcript had no 3' exon. The substrate RNA from this gene was not cleaved by the endonuclease (data not shown).

The requirement for an intact DHU stem and loop was next examined. An oligonucleotide-directed deletion, using oligonucleotide "\(^\text{3}'\text{Arm}\)", was introduced into \(^\text{13}^{11}\) such that the resulting deletion substrate, designated \(^\text{13}^{11-5}\), lacked the entire DHU and T\(\Psi\)C stem and loop structures, and retained only those exon sequences associated with the anticodon stem (Figure 15). The RNA transcript from this extensively deleted gene was cleaved by the endonuclease (Figure 15).

To demonstrate that the RNA products generated from the cleavage of \(^\text{13}\) and \(^\text{13}^{11}\) substrates were cleaved correctly,
Figure 15. Structure and cleavage assays of 16, 167, and 1311-5 RNAs. The left panel is the proposed secondary structures of the T7 RNA polymerase generated transcripts derived from the 16, 167, and 1311-5 genes. Cleavage sites for the endonuclease are indicated by arrows. The right panel is an autoradiogram of cleavage assays performed with 16, 167, and 1311-5 substrate RNAs. Identities of RNA products are shown.
Figure 15.
the isolated introns from \(^{13}\) and \(^{13_{11}}\) RNAs were subjected to 5' end group analysis. TLC analysis demonstrated an identical chromatographic pattern for these introns and that from \(^{16}\) RNA (Figure 16), both in identity and relative quantity. This data, in addition to RNA sequence obtained from the 5' ends of the respective introns (data not shown) indicated that \(^{13}\) and \(^{13_{11}}\) RNAs were cleaved accurately and precisely by the endonuclease. Unlike the eukaryotic enzyme, the halophilic endonuclease exhibited no dependence upon mature tRNA-like structure for substrate splice-site selection and cleavage.

Consistent with the lack of a requirement for mature-tRNA domains in the precursor, the halophilic endonuclease was not inhibited by mature tRNA. Varying amounts of heterologous mature yeast tRNA were mixed with \(^{16}\) RNA in a cleavage assay (Figure 17). At levels of up to 10 \(\mu\)g of yeast tRNA per 5 ng \(^{16}\) RNA per cleavage reaction (2000 fold excess yeast tRNA by weight; \(\approx5000\) fold molar excess) there was no significant inhibition in the cleavage of \(^{16}\) substrate.

**Role of intron.**

An examination of the RNA sequence of the intron from \(^{16}\) revealed a potential secondary structure within the intron (Figure 6), which resembled an anticodon stem-loop structure. The sequence of this putative stem and loop was
Figure 16. 3' end group analysis of the products from cleavage of 16, 13, and 131 RNAs. The radiolabeled intron product of 16, 13, and 131 RNAs were isolated, digested with nuclease T2, and the resulting mononucleotides separated by thin layer chromatography. Identities of radiolabeled nucleotides were determined by comparing their mobilities to known mononucleotide standards.
**RNase T2 Digestion**


**Up**

**Ap**

**Gp**  **Cp**

Δ16  Δ13  Δ1311

*Figure 16.*
Figure 17. Competition assay of &superscript;16 RNA with competing mature yeast tRNA. Equal amounts of [α-32P]ATP-labeled &superscript;16 substrate RNA were combined with increasing concentrations of unlabeled mature yeast tRNA in cleavage assays. The autoradiogram of the cleavage assays is shown. Concentrations of unlabeled yeast tRNA and the identities of RNA products are indicated.
Figure 17.
compared to known H. volcanii tRNA sequences and a 72% and 74% sequence similarity, respectively, was found to tRNA^{Ile} and tRNA^{TYE}, of H. volcanii (Figure 18). Akins and Lambowitz (2) reported that a tyrosyl-tRNA synthetase was required for splicing in N. crassa of the cyt-18 gene intron. The S elements of this and a number of other Group I introns can be folded into a small stem and loop structure containing the tyrosine anticodon. It was conceivable that the putative stem and loop structure, which resemble an anticodon stem and loop within the intron of the transcript from ^16, might interact with a synthetase-like enzyme, possibly associated with the endonuclease. To investigate this possibility, exonuclease III deletions were created at the BcII site within the intron of ^13^{11}, which was located within the putative anticodon loop. Two deletion substrates were selected for analysis. The RNA ^13^{11}-24 had a 29 nucleotide deletion, which removed the entire putative stem and loop sequence from the intron (Figure 19). A second RNA, ^13^{11}-10, had a 41 nucleotide deletion which removed half of the stem and loop, and half of the intron sequence which could basepair to the anticodon (Figure 19). These substrate RNAs were tested in an assay and were found to be cleaved by the endonuclease. The observed cleavage products from these RNAs agreed in size to the predicted cleavage products if cleavage at the wild type sites had occurred (Figure 19). Although these substrates were not cleaved to
Figure 18. Sequence and structural conservation of a potential stem and loop within the 16 intron and the anticodon stem and loops of tRNA$^{\text{Yr}}$ and tRNA$^{\text{Ile}}$. Dots indicate conserved nucleotides between the intron of 16 RNA and the anticodon stem and loop of H. volcanii tRNA$^{\text{Yr}}$ and tRNA$^{\text{Ile}}$. Arrows indicate inverted repeat sequences which can form a stem and loop structure. The anticodon nucleotides are underlined.
Figure 18.
Figure 19. Structure and cleavage assays of $^{13}$, $^{13}{^{11}-10}$, and $^{13}{^{11}-24}$ RNAs. The top panel shows the proposed secondary structures of the T7 RNA polymerase generated transcripts derived from the $^{13}$ gene. Deletions introduced into the gene encoding $^{13}$ and their identities are indicated. Arrows indicate the predicted cleavage sites for the endonuclease. Bottom panel is an autoradiogram of the cleavage assays for these substrate RNAs. Identities of the RNA products are shown.
Figure 19.
the same extent as the ^13 substrate, the putative structures and sequences removed from the 3' half of the intron were not essential for proper substrate cleavage. Therefore, an interaction of a "synthetase-like" protein, as described for N. crassa cyt-18, was not required by the halophilic endonuclease activity.

To further examine the role of structure or sequence in the intron, a more extensive intron-deletion substrate was constructed. The 219 nucleotide EcoRI-HindIII DNA restriction fragment coding for pretRNA^Trp was isolated from ^16 and digested with HinP1. The resulting EcoRl-HinPl and HindIII-HinPl DNA fragments, having the 5' and 3' exons, respectively, and a small portion of the intron, were ligated into pT72. The resulting tRNA gene, ^16^7 (Figure 15), retained only 22 nucleotides of the original intron sequence, including those sequences found at the exon-intron boundaries. The transcript from this gene was cleaved by the endonuclease in an assay containing 20 mM MgCl₂, but at a lower rate than for ^16 RNA. If, however, ^16^7 RNA was assayed in buffer containing 2.5 mM spermidine rather than 20 mM MgCl₂, substrate turnover increased significantly. Cleavage of this RNA indicates that complete intron was not a requirement for precursor cleavage and that the necessary sequences or structures must be present in the exon-intron boundary regions. Figure 20 summarizes the deletion substrates constructed.
Figure 20. Summary of exon and intron deletions in \(^{16}\) RNA. Secondary structure of \(^{16}\) substrate RNA is shown. Stippled areas indicate those regions of \(^{16}\) which can be deleted without affecting viability of the RNA as an active endonuclease substrate. Designations of the genes encoding these deletion genes are given. On the right is shown the common region shared by \(^{16}\), \(^{16^7}\), and \(^{1311-5}\) RNAs. This region consists of two, three nucleotide loops separated by four base pairs. Cleavage sites of the endonuclease are indicated by arrows.
Figure 20.
Identification of a conserved structural motif.

Although \(^{16}\), \(^{16'}\), and \(^{13}_{11-5}\) vary considerably in structure, their transcripts were all recognized and accurately cleaved by the halophilic endonuclease. They must all therefore share a common sequence or structural element recognized by the endonuclease. A comparison of \(^{16}\), \(^{16'}\), and \(^{13}_{11-5}\) structures disclosed only one region common to all three substrates (Figure 20), the exon-intron boundary junction of the anticodon helix. In \(^{16'}\) this region can be folded into a structure composed of two, three nucleotide loops flanking a four basepair stem, with cleavage sites occurring within the two loops. A compilation of archaeabacterial intron-containing precursors indicated that this structural motif was conserved at the exon-intron boundary junctions among many of the intron-containing precursor tRNAs (Figure 21). RNA structure-probing analysis indicated that the intron was extensively basepaired through the anticodon helix of \(^{16}\) (unpublished) and computer modeling of \(^{16}\) (provided by J. Brown, U. of Indiana, IN) suggested that the structure of \(^{16}\) was more accurately represented by the structure shown in Figure 15. Because the structural motif was highly conserved in intron-containing precursors, it was likely to be a recognition element for the endonuclease; this region was therefore chosen as a target for further study.
Figure 21. Conservation of exon-intron boundary structures among archaebacterial intron-containing precursor RNAs. Compilation of exon-intron boundary structures of all known archaebacterial intron-containing precursor RNAs. Identities of precursor RNAs are shown and arrows indicate sites of cleavage for the endonuclease.
Figure 21.
Reconstruction of o\textsuperscript{16} and o\textsubscript{16}.

One disadvantage of using RNA substrates derived from the \textsuperscript{16} gene was the possibility that the 5' leader sequence could contribute to alternative folding patterns in the mutant RNA which are unrelated to the gene encoding region. It was therefore desirable to reconstruct the pretRNA\textsuperscript{TrP} gene such that no exogenous RNA sequences would exist in the precursor RNA transcript. Due to the large size of the intron, the tRNA\textsuperscript{TrP} gene was constructed in two steps. First a set of nine overlapping oligonucleotides, encoding the \textit{H. mediterranei} \textsuperscript{16} tRNA\textsuperscript{TrP} gene, were annealed together and cloned into the plasmid pUC18 (Figure 22) to yield the gene o\textsubscript{16}\textsuperscript{7}. In addition to the gene sequences, this reconstructed gene contained the sequence for the T7 RNA polymerase promoter. The promoter was arranged such that the first nucleotide of the transcript was the same as 5' terminus of the pretRNA. The transcript of this gene ("o\textsubscript{16}^7") lacked the 5' leader sequences and had a 3' CCA terminus, which may be required for exon ligation. This substrate closely resembled the natural \textit{in vivo} precursor molecule in structure, except that it lacked base modifications. In addition, the transcript could be synthesized with a 5' monophosphate terminus, which allowed the 5' end to be easily radiolabeled. The full length \textsuperscript{16}-oligo tRNA\textsuperscript{TrP} ("o\textsubscript{16}") gene was constructed from the
Figure 22. Reconstruction of the *Halobacterium mediterranei* intron-containing tRNA\(^{Trp}\) gene. A, schematic of the partial intron-containing tRNA\(^{Trp}\) gene, o16\(^7\), showing the assembly of overlapping oligonucleotides; sequences of oligonucleotides are given in Table 6. Indicated below is the expected *in vitro*, end-matured, T7 RNA polymerase RNA transcript produced from this gene. The transcript can be synthesized with a 5' monophosphate or triphosphate termini. B, construction of the full length intron-containing tRNA\(^{Trp}\) gene, o16. The 82 base pair HindIII intron fragment, absent in o16\(^7\), was isolated from the ^16 tRNA\(^{Trp}\) gene of pT72HM^16 and reintroduced into the HindIII site of o16\(^7\). The *in vitro* transcript of this gene, o16 RNA, is shown below.
Figure 22.
synthetic gene by reassembling the exon 1-partial intron and exon 2-partial intron fragments of o167 with the 82 basepair HinP1 intron fragment from T16. Correct orientation of the HinP1 intron fragment was confirmed by DNA sequence analysis. Both o16 and o167 RNAs derived from the synthetic genes were cleaved accurately by the endonuclease (Figure 23), although o167 was somewhat less efficiently cleaved. This difference in cleavage efficiency was reflected in the apparent k_m's for these RNAs, 25-40 nM and 85-115 nM for o16 and o167 respectively. The difference in k_m values for o16 and o167 suggests that the endonuclease has a higher affinity for o16 than o167.

**Analysis of the conserved structure at the exon-intron boundary.**

Three goals were proposed to elucidate substrate specificity, defined by the conserved structure at the exon-intron boundaries. These goals were: 1) to determine whether removal or alteration of the proposed structure affected cleavage; 2) to determine whether the two cleavage sites were recognized independently; 3) to determine how the endonuclease identified the sites for cleavage. To determine whether removal or alteration of the proposed structure affected cleavage, the four intron nucleotides (5'-GGAG-3'), which participated in the base pairing between
Figure 23. Structure and cleavage assays of ol6\textsuperscript{7} and ol6 RNAs. The left panel is the proposed secondary structures of the T7 RNA polymerase generated transcripts derived from the ol6\textsuperscript{7} and ol6 genes. Cleavage sites for the endonuclease are indicated by arrows. The right panel is an autoradiogram of cleavage assays performed with ol6\textsuperscript{7} substrate RNA (lane 1) and ol6 substrate RNA (lane 2). Identities of RNA products are shown. Mutations in the ol6 gene used in the analysis of substrate recognition by the halophilic endonuclease are indicated (see text for details). ▽ represents an insertion, while △ represents a deletion.
Figure 23.
the two loops, were deleted by oligonucleotide-directed
mutagenesis (Figure 24). The resulting RNA transcript from
this gene (\^GGAG) retained both cleavage sequences, but the
cleavage sites were present in an internal loop. If only
sequence recognition was required for cleavage, this RNA
should remain an active substrate, however when this RNA was
assayed no cleavage was observed (Figure 24). It appears
that base pairing of these nucleotides and the formation of
the proposed structure is important for recognition, and
sequence in the cleavage site loops alone is not sufficient.
The possibility could not be excluded that those nucleotides
removed from the intron were primary sites for sequence
recognition. It was also possible that removal of the four
nucleotides might alter the structure of tRNA\textsuperscript{Trip} by allowing
alternate folding patterns to occur. To test this
possibility, 5' end labeled o16 and \^GGAG RNA were subjected
to nuclease S1 structure probing analysis. This analysis
(Figure 25) indicated that o16 and \^GGAG RNAs have identical
nuclease S1-sensitive cleavage patterns with the exception
of a new S1-sensitive site localized at the position of the
bulge created by the four nucleotide deletion. This four
basepair deletion did not appear to alter the overall
structure of the RNA. To test whether 5' and 3' cleavage
sites were cleaved independently, as occurs in the yeast
tRNA processing system, the nucleotides comprising the 3'
Figure 24. Analysis of structural alterations at the exon-intron boundary junctions of o16. Top panels are secondary structures of intron boundary junctions of substrate RNAs generated from the wild type and mutated o16 genes. Bold face type indicates nucleotide insertions and arrows indicate predicted cleavage sites for the endonuclease. Products from the cleavage assays for these RNA substrate are shown in the panels below. Identities of RNA products are indicated. The - and + designations indicate the absence or presence of endonuclease in the assay.
Figure 24.
Figure 25. S1 structure probing of 016 and ^GGAG substrate RNAs. Autoradiogram of nuclease S1 structure probing of 5' end labeled 016 and ^GGAG substrate RNAs. These RNAs were mixed with 10 µg unlabeled carrier yeast tRNA and digested with nuclease S1 for 10 min. Units of nuclease S1/µg unlabeled tRNA is indicated. Uniformly [α-32p]ATP labeled 016 was cleaved with endonuclease to generate size markers and loaded in the lane designated M; identities of cleavage products are shown. The common nuclease S1 digestion product is indicated on the right as S1. The * indicates the new nuclease S1 cleavage product from ^GGAG RNA.
Figure 25.
loop (AUA) were deleted (Figure 24). This deletion would be expected to increase the length of the anticodon helix while leaving the loop at the 5' cleavage site unchanged (Figure 24). This substrate (AUA) was not cleaved at the 5' cleavage site when assayed with the endonuclease (Figure 24). Similarly, a mutant was constructed where nucleotides complementary to the 5' cleavage loop (UCU) were introduced into the intron, extending the helix to include the 5' loop nucleotides (Figure 24). As observed with the 3' loop deletion, this substrate (UCU) was not cleaved by the endonuclease (Figure 24). These results indicated that the cleavage sites were not independently recognized and that both loops must be present in the substrate to be cleaved by the endonuclease.

It was possible that the endonuclease specifically required the nucleotides at the 5' cleavage site to be in a bulge loop structure. To test for this possibility, a second mutation was introduced into o16 in which the nucleotides AGA were inserted into the intron opposite the AGA of the 5' loop (Figure 24). This substrate RNA (AGA) had an intact 3' bulge loop and retained an unpaired 5' loop sequence, but the 5' cleavage site was localized in an internal loop rather than a bulge loop. When assayed, this substrate was not cleaved (Figure 24), in spite of retaining unpaired nucleotides in both loops. Apparently, specific structural constraints defined by the two, three nucleotide
loops separated by the four base pairs were required for cleavage by the endonuclease.

To verify that all of the necessary structural and sequence elements were present at the exon-intron boundaries, a model substrate containing only these proposed features was constructed (Figure 26). This model substrate, designated Minsub, consisted of a five nucleotide stem, the conserved structure of two, three nucleotide loops separated by four basepairs, and an intron of 17 nucleotides. When assayed, this RNA substrate was cleaved precisely to give four products (Figure 26). RNA sequence analysis of the excised intron (Figure 27) demonstrated that Minsub RNA was cleaved accurately by the endonuclease. Structure and sequences present in this molecule were therefore sufficient for recognition by the endonuclease.

Identification of cleavage sites by the endonuclease.

The eukaryotic endonuclease identifies its cleavage sites by measuring the distance from the top of the anticodon stem to the cleavage sites. This was demonstrated when addition of a base pair to the anticodon stem of tRNA^Phe caused the cleavage sites to shift by one nucleotide, maintaining a fixed distance from the top of the anticodon stem (134, 170). It was possible that the halophilic enzyme, while requiring a specific structure for binding, might also measure the cleavage sites from the top
Figure 26. Structure and cleavage assay of the minimal intron endonuclease substrate Minsub. The left panel is the proposed secondary structures of the T7 RNA polymerase generated transcripts derived from the gene encoding Minsub RNA. Arrows indicate the predicted sites of cleavage by the halophilic endonuclease. The right panel is an autoradiogram of the cleavage products generated with Minsub RNA. The - and + designations indicate the absence or presence of endonuclease in the assay. Identities of RNA products are indicated. Although not visible, location of the 3' exon is indicated.
Minimum Substrate

<table>
<thead>
<tr>
<th>preRNA</th>
<th>Exon 1 + Intron</th>
<th>Intron</th>
<th>Exon 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Exon 2)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 26.
Figure 27. RNA sequence of end labeled intron derived from cleavage of Minsub RNA. Unlabeled intron from cleavage of Minsub RNA was isolated, end labeled with T4 polynucleotide kinase and \([Y-^{32}P]ATP\), and subjected to enzymatic RNA sequence analysis. The identity of the 5' nucleotide was confirmed by end group analysis (data not shown).
Figure 27.
of the anticodon stem. To test this possibility, a single C residue was inserted into the 5' exon stem (C29:1, Figure 28), creating a one nucleotide bulge in the anticodon stem. This substrate was cleaved by the endonuclease, generating a 5' exon one nucleotide larger than wildtype 5' exon (Figure 28). Unlike the yeast intron endonuclease, which was incapable of cleaving a substrate carrying a one nucleotide insertion in the anticodon helix above the 3' cleavage site, any structural alteration in this stem as a result of a one nucleotide insertion did not affect the ability of the halophilic enzyme to cleave the RNA. Addition of a G residue in exon 2, adjacent to C29:1, resulted in the formation of a G-C nucleotide pair in the anticodon stem above the presumptive recognition structure (Figure 28). Assays of o16 and the base pair insertion C29:1-G146:1 demonstrated that the mutant substrate was cleaved by the endonuclease. The exons originating from C29:1-G146:1 were one nucleotide larger than those of o16, while the introns from each were the same size. This was consistent with cleavage at the original sites, as opposed to the nucleotide shift observed with the eukaryotic endonuclease where a base pair was added to the anticodon stem. Cleavage at these sites was verified by end group analysis. As predicted for o16, both the 3' exon and the intron terminated in A residues (Figure 29). Therefore, the halophilic endonuclease did not measure down the anticodon stem to the
Figure 28. Identification of cleavage sites by the halophilic endonuclease. Top panel shows the secondary structures at the exon-intron boundary junctions of wild type and mutant 016 substrate RNAs. Inserted nucleotides are indicated by bold face type and arrows indicate wild type cleavage sites. The lower panels are autoradiograms of cleavage assays from 016 (lane 1), single nucleotide insertions (lane 2), and double insertions (lane 3). Identities of RNA products are shown.
Figure 28.
Figure 29. 5' end group analysis of products from cleavage site identification mutants. The unlabeled 3' exon and intron products of ol6, C29:1-G146:1, C36:1, and C36:1-G136:1 RNAs (Figure 27) were 5' end labeled with T4 polynucleotide kinase and [γ-32P]ATP, digested with nuclease P1, and the resulting mononucleotides separated by thin layer chromatography using solvent system 2. Identities of radiolabeled nucleotides were determined by comparing their mobilities to known mononucleotide markers.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>EC2</td>
<td>I</td>
<td>E2 I</td>
<td>E2 I</td>
<td>E2 I</td>
</tr>
</tbody>
</table>

- PC __
- PU __
- PG __
- PA __

origin __

Figure 29.
cleavage sites; its site recognition was independent of anticodon stem length.

To test whether identification of cleavage sites involved a measurement between the loops, a single C residue or a CG basepair (C36:1 and C36:1-G136:1, respectively) were inserted into the helix region separating the two loops (Figure 28). The extra CG basepair would extend the length of the helix between the two loops by one basepair. The single C nucleotide insertion would result in the formation of a four, rather than three, nucleotide loop at the 5' cleavage site, while leaving the distance between the two loops at four basepairs. When assayed, both of these altered substrate RNAs were cleaved by the endonuclease (Figure 28), but at a greatly reduced efficiency (5-15% of the wild type substrate). Despite the reduced cleavage, end group analysis of the intron and 3' exon isolated from cleavage products of both RNAs was performed. Analysis of intron and exon 2 termini from the cleavage products of C36:1-G136:1 RNA indicated that cleavage occurred in the same location as in wild type substrate RNA (Figure 29). Product analysis of C36:1 indicated that cleavage occurred predominantly at the predicted sites, however a minor amount of the intron product (<5%-10%) had a 5' G terminus, indicative of a one nucleotide shift in the cleavage. Since cleavage is greatly reduced in these mutations, the distance between the cleavage sites must play some role in
recognition; however the fact that correct cleavage occurs suggests that other factors, such as interaction with the helices flanking the structure or specific nucleotides may also be involved.

The possibility of interactions with the helix below the cleavage sites was indicated when the analogous C36:1- G136:1 mutation was introduced into o16^7. In contrast to C36:1-G136:1, when this RNA was assayed, it was not cleaved by the endonuclease (Figure 30). The only difference between C36:1-G136:1 and C35:1-G56:1 is the presence of the large intron below the 5' cleavage site in o16. Furthermore, the o16^7 C35:1-G56:1 mutant differs from o16^7, a functional substrate, by having an additional base pair between the two cleavage site loops. In its worst effect, this mutation would rotate one of the cleavage sites by approximately 360° with respect to o16^7. It would appear that the enzyme makes multiple contacts with the substrate, possibly contacting the helices above and below the exon-intron boundaries.

To investigate the possibility of specific nucleotide requirements for the halophilic endonuclease, the C35:G56 basepair in the four basepair helix between the two loops was replaced with a U35:A56 basepair in o16^7 (Figure 30). When the substrate RNA was assayed it was cleaved only to a minor extent by the endonuclease (Figure 30). One interpretation for the loss of cleavage was that the UA
Figure 30.
pair, replacing the CG pair, destabilized the proposed structure; free energy calculations (216) predicted that the exon-intron boundary structure of the mutant RNA would be ≈1 kcal (35%) less stable than the wild type boundary structure. If structure alone was required, replacing the UA pair with a GC pair should return the substrate to its original stability and recover its activity as a substrate. This mutation was created in o167. When the G35:C56 RNA was assayed, it was cleaved only slightly better than the U35:A56 RNA (Figure 30). It is possible that these nucleotides may be points of contact with the endonuclease.

Since mature RNA structures and sequences were not required for substrate recognition and cleavage by the endonuclease, this enzyme could potentially cleave non-tRNA molecules. It is interesting to note that the RNase III processing sites of 16S and 23S rRNAs from several thermophiles and methanogens contain the conserved exon-intron boundary structure. This suggested that the endonuclease may function as a general stable RNA maturase. To determine whether these structures could be cleaved by the endonuclease, substrates mimicking the 16S and 23S rRNA processing sites of H. cutirubrum (Figure 31) were synthesized in the following manner. An 18 nucleotide oligonucleotide, which consisted of the promoter sequence for T7 RNA polymerase, was annealed to a second oligonucleotide. This second oligonucleotide consisted of
Figure 31. Analysis of RNase III cleavage site of the 16S and 23S rRNAs from H. cutirubrum. Top panel shows the organization (above) of the rRNA operon from H. cutirubrum and the location of the RNase III cleavage sites for the 16S and 23S rRNAs (below). Bottom panel shows the secondary structure of the T7 RNA polymerase generated transcripts produced from the oligonucleotides encoding the RNase III cleavage sites for 16S and 23S rRNA. Transcripts were synthesized with a radiolabeled 5' monophosphate guanosine, designated by *. Arrows indicate the predicted cleavage sites for the endonuclease. On the right is shown the autoradiogram of the cleavage assays for these substrate RNAs. Identities of RNA products are indicated. The designations - and + indicates absence or presence of endonuclease.
Figure 31.
18 nucleotides of complementary T7 RNA polymerase promoter sequence and the coding strand of either the 16S or 23S RNA RNase III sequences. These annealed oligonucleotides were then transcribed by T7 RNA polymerase, resulting in an RNA which, when self-annealed, would mimic the RNase III cleavage site of the in vivo ribosomal RNA. When 5' end labeled RNAs were assayed with the halobacterial endonuclease, products were produced which were consistent with cleavage in the bulge loop region (Figure 31).

Presence of endonuclease activity in other archaeabacteria.

The structure at the exon-intron boundary is highly conserved among archaeabacterial intron containing precursors, suggesting the possibility of a universal recognition mechanism. To test the generality of the endonuclease recognition pattern, extracts from organisms representing the methanogen and thermophilic branches of the archaeabacteria were evaluated. Using ol6 RNA as an assay substrate, cellular extracts from Halobacterium cutirubrum (data not shown), Sulfolobus solfataricus, Thermoplasma acidophilum, Methanosarcina barkeri, and Methanothermus fervidus were assayed for endonuclease activity. In all cases, the extracts were capable of cleaving ol6 at the physiological temperature of the host organism (Figure 32). It is likely that these organisms either contain interrupted genes or have structures resembling the conserved
Figure 32. Cleavage of o16 RNA by methanogenic and thermoacidophilic endonuclease extracts. Endonuclease extracts were prepared from *M. barkeri*, *M. fervidus*, *S. solfataricus*, and *T. acidophilum* and incubated with radiolabeled o16 substrate RNA at the temperatures indicated. An *E. coli* extract was assayed as a negative endonuclease control. The autoradiogram of these assays is shown. Identities of RNA products are indicated. The designations - and + indicate absence or presence of 1 mM ATP, which might be required for exon ligation.
Figure 32.
recognition structure. Yeast (Figure 33) and E. coli (Figure 32) extracts were not capable of cleaving ol6.

The endonuclease activity from Thermoplasma acidophilum was examined more closely. End group analysis of the intron and 3' exon isolated from a cleavage assay with ol6 RNA and T. acidophilum endonuclease (Figure 34) demonstrated the presence of a pA residue at the 5' termini of these cleavage products, indicating that the thermophilic endonuclease cleaved ol6 RNA at the correct position. This endonuclease also cleaved ol6^ and ~1311-5 RNAs (Figure 35), thus exhibiting the same substrate range as H. volcanii endonuclease. The temperature profile of the thermophilic endonuclease correlated with the optimal growth temperature of the organism (Figure 11). Because the intron and 3' exon could be 5' end labeled by T4 polynucleotide kinase with [Y-32p]ATP, these cleavage products must have 5'-OH group available for phosphorylation. The cleavage mechanism of T. acidophilum is likely to be identical to that of its halophilic counterpart.
Figure 33. Cleavage of yeast pre-tRNAPhe and 016 RNAs by yeast tRNA intron endonuclease. Autoradiogram pre-tRNAPhe and 016 substrate RNAs assayed with yeast tRNA intron endonuclease. Identities of RNA products are indicated. The designations - and + indicate the absence or presence of 1 mM ATP, which is required for exon ligation in the yeast extract.
Figure 33.
Figure 34. 5' end group analysis of the products from cleavage of o16 RNA by Thermoplasma endonuclease. The unlabeled 3' exon and intron products of o16 substrate RNA were 5' end labeled with T4 polynucleotide kinase and [γ-32P]ATP, digested with nuclease P1, and the resulting mononucleotides were separated by thin layer chromatography using solvent system 1. Identities of radiolabeled nucleotides were determined by comparing their mobilities to known mononucleotide standards.
Figure 34.
Figure 35. Cleavage of \(^{16}, {^{16}}\textsuperscript{7}\) and \(^{13}\textsuperscript{11}-5\) RNAs by *Thermoplasma* tRNA intron endonuclease. Radiolabeled \(^{16}, {^{16}}\textsuperscript{7}\), and \(^{13}\textsuperscript{11}-5\) substrate RNAs were incubated with *T. acidophilum* endonuclease at 58°C or *H. volcanii* endonuclease at 37°C. The autoradiogram of these cleavage assays is shown. Lane 1 is substrate RNA with halophilic endonuclease and lane 2 is substrate RNA with *T. acidophilum* endonuclease. Identities of RNA products are indicated.
Figure 35.
Partial purification of tRNA\textsubscript{Trp} intron endonuclease from \textit{Halobacterium volcanii}.

Since intron-containing pre\textit{tRNAs} are present \textit{in vivo} in very low concentrations, it was necessary to develop a method to produce pre\textit{tRNA}\textsubscript{Trp} \textit{in vitro}. A genomic subclone of tRNA\textsubscript{Trp} from \textit{H. mediterranei} was placed under the transcription control of a T7 RNA polymerase promoter such that the gene could be transcribed \textit{in vitro} by T7 RNA polymerase to generate biochemically useful amounts of pre\textit{tRNA}\textsubscript{Trp}. We used \textit{in vitro} transcripts of this construct (\textasciitilde16) to generate substrate RNA for tRNA intron endonuclease. An assay was then developed for \textit{H. volcanii} tRNA\textsubscript{Trp} intron endonuclease which was based on the ability of tRNA\textsubscript{Trp} intron endonuclease to cleave T7 RNA polymerase run-off transcripts from the plasmid \textasciitilde16. Cleavage of this RNA by the endonuclease gave three products which corresponded in size to the 5' exon plus leader, intron, and 3' exon. Using this assay, which allowed rapid
identification of endonuclease activity, a purification procedure was developed. This purification procedure employed lysis of halobacterial cells by passage of the cell suspension through a French pressure cell followed by PEG 8000 precipitation of nonactive proteins and RNase P activity. Soluble proteins were separated by passage twice over DEAE Sephacel followed by separation through hydroxylapatite. Implementation of this purification procedure resulted in a 230 fold purification of the endonuclease (Table 4). Analysis of protein fractions by SDS polyacrylamide gel electrophoresis indicated a reduction in the number of total proteins present in the extract, and also an enhancement in the relative quantity of a few specific proteins. These proteins had apparent molecular masses of 54 kdal, 38.5 kdal, 29.5 kdal, and 20 kdal; greater than 20 protein bands were still visible in the most purified extract. The endonuclease was not resolved by additional chromatographic separation methods. No further purification procedures were employed as the partial purification procedure gave an enzymatic activity which cleaved the substrate accurately and efficiently, and was suitable for examination of substrate specificity of the endonuclease, the main focus of this research.
Fidelity and mechanism of the cleavage reaction.

In order to demonstrate that the cleavage of tRNA\textsuperscript{Trp} was not due to the presence of a nonspecific endonuclease, it was necessary to show that cleavage was accurate and precise. To achieve this, it was necessary to purify the endonuclease to a level at which no endogenous RNA could be labeled with [\(\gamma\text{-}^3\text{P}\)]ATP by polynucleotide kinase; these RNAs would interfere with product analysis. Endonuclease passed twice through DEAE Sephacel was sufficiently free of contaminating RNA (demonstrated by pCp end labeling of extracts; data not shown) to be suitable for this analysis.

The cleavage mechanism for yeast tRNA intron endonuclease predicted the existence of 5' hydroxyl termini on the intron and 3' exon products. The ability to 5' end label cleavage products of the halophilic endonuclease with [\(\gamma\text{-}^3\text{P}\)]ATP and T4 polynucleotide kinase confirmed the presence of an available 5'-OH group. This property was exploited to determine the accuracy of the cleavage reaction by 5' end group analysis, as well as direct enzymatic RNA sequence analysis of released intron and 3' exon. End group analysis of these products demonstrated the existence of a pA nucleotide at the 5' terminus of both cleavage products; this was the nucleotide predicted to be present as the terminal 5' nucleotide if cleavage occurred at the expected locations. The RNA sequence obtained by direct enzymatic RNA sequencing from the 5' end of the intron and 3' exon
agreed with the predicted sequences of these products. The end group analysis and RNA sequence obtained from these products were unambiguous and showed that the cleavage products had discrete, homogeneous termini, indicating that cleavage by the endonuclease was precise. Analysis of the RNase A oligonucleotide isolated from the 3' end of uniformly $[\alpha^{-32P}]$ATP labeled intron demonstrated the presence of a labeled Up* (* representing a radiolabeled phosphate). This could only occur if a nearest neighbor phosphate transfer at the intron-exon junction (intron-pApUp*/A-exon 2) took place during cleavage. End group analysis of the 3' terminus of the excised intron and 5' exon revealed the presence of a 2',3' cyclic G residue at the 3' terminus of the 5' exon and a 2',3' cyclic U residue at the 3' terminus of the intron. Again, since the substrate RNAs were labeled with $[\alpha^{-32P}]$ATP the labeled phosphate must have been derived from a nearest neighbor phosphate transfer at the cleavage junction.

The 5' hydroxyl and 2',3' cyclic phosphate groups produced as part of the archaeabacterial cleavage reaction are the same reaction termini produced by the yeast nuclear-encoded tRNA intron endonuclease during cleavage of the yeast pre-tRNA. The biochemical nature of the 5' and 3' termini of the cleavage products suggest that the phosphodiester bond cleavage mechanism of the
archaebacterial and the eukaryotic nuclear-encoded tRNA intron endonuclease are the same.

Biochemical characterization of the endonuclease.

The time-dependent endonuclease cleavage assay demonstrated that endonuclease cleavage proceeded linearly with time up to 30 min with no apparent lag. There was no evidence for the assembly of a cleavage complex.

The endonuclease has a broad pH range, with optimal activity between pH 7 and pH 8. The temperature optimum of the halophilic tRNA intron endonuclease corresponded to the optimal growth temperature of the host organism (42°C), which was also observed with the intron endonuclease from *T. acidophilum*, which had a maximum at 58°C, corresponding to its optimal growth temperature.

The halophilic endonuclease exhibited a requirement for divalent ions. Although Ca⁺⁺ ions appeared to give highest substrate turnover by the endonuclease (Figure 12), the stoichiometry of the cleavage products was inconsistent with the expected distribution of label in the cleavage products. Ca⁺⁺ ions appeared to alter the specificity of the cleavage reaction. Both 20 mM MgCl₂ and 5 mM spermidine gave accurate and efficient substrate turnover and were used interchangeably in the standard assay. A requirement for Mg⁺⁺ ions or spermidine was not unexpected, as both MgCl₂ and spermidine are known to bind to mature tRNA and
participate in the maintenance of tRNA structure. tRNA molecules have binding sites for two spermidine molecules and six Mg$^{++}$ ions (178); Mg$^{++}$ ions stabilize tRNA structure through ionic bridges between Mg$^{++}$ ions and the negatively charged phosphate groups.

Monovalent ions were a potent inhibitor of the halophilic endonuclease, unlike the eukaryotic nuclear tRNA endonuclease (165). The basis for this inhibition is not known and was unexpected since the intracellular halophilic environment consists of high (2 M) concentrations of monovalent ions (117). This property was not unique to the halophilic endonuclease as similar properties have been observed for other halophilic proteins (82, 162). The possibility cannot be excluded that some additional component, which permits full activity or provides electrostatic shielding of the endonuclease from the high concentration of monovalent ions within the cell, is lacking in the partially purified endonuclease extract. Table 5 compares the biochemical properties of H. volcanii endonuclease and compares them with those of the Saccharomyces cerevisiae nuclear tRNA intron endonuclease.

Substrate Recognition.

Since a single yeast nuclear tRNA intron endonuclease cleaves all nuclear-encoded intron-containing tRNA precursors, and these introns are located in the same
<table>
<thead>
<tr>
<th>Property</th>
<th>Yeast Nuclear</th>
<th>Halophilic</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH optima</td>
<td>pH 7-8</td>
<td>pH 6-9</td>
</tr>
<tr>
<td>Ion requirement</td>
<td>Spermidine&gt;Mg^{++} Monovalent inhibitory at 120 mM</td>
<td>Spermidine&gt;Ca^{++}&gt;Mg^{++} Monovalent inhibitory at 50 mM</td>
</tr>
<tr>
<td>Cleavage termini</td>
<td>5' hydroxyl 2',3' cyclic phosphate</td>
<td>5' hydroxyl 2',3' cyclic phosphate</td>
</tr>
<tr>
<td>km</td>
<td>25-30 nM</td>
<td>35-42 nM</td>
</tr>
<tr>
<td>Substrate recognition</td>
<td>Mature tRNA structure in precursor</td>
<td>Exon-intron boundary regions</td>
</tr>
<tr>
<td>Substrate range</td>
<td>All yeast pretRNAs Doesn't cleave</td>
<td>pretRNA_{Trp} &amp; prerRNAs Doesn't cleave</td>
</tr>
<tr>
<td></td>
<td>H. volcanii tRNA_{Trp}</td>
<td>yeast tRNA_{Phe}</td>
</tr>
<tr>
<td>Subcellular location</td>
<td>Membrane</td>
<td>Soluble (and membrane)</td>
</tr>
<tr>
<td>Associated ligase activity</td>
<td>Yes, in crude preparation</td>
<td>Not detectable</td>
</tr>
</tbody>
</table>
position as the halophilic intron, it was possible that the halophilic endonuclease used a mechanism similar as the yeast enzyme for substrate recognition. To determine whether *H. volcanii* endonuclease was capable of cleaving a nonarchaebacterial precursor tRNA, a *S. cerevisiae* pretRNA\(^{\text{Phe}}\) was examined as a substrate for the halophilic enzyme. This precursor RNA has been shown to be an active substrate for the yeast nuclear tRNA endonuclease. The primary transcript of pretRNA\(^{\text{Phe}}\) has a 19 nucleotide intron located one nucleotide 3' to the anticodon, the same location as the intron interrupting *H. volcanii* tRNA\(^{\text{Trp}}\). When examined with the halophilic endonuclease, pretRNA\(^{\text{Phe}}\) RNA was not cleaved. This indicated that intron location was not sufficient for recognition by the halobacterial endonuclease and that its recognition properties differed from the yeast enzyme. The converse experiment showed that yeast nuclear endonuclease will cleave tRNA\(^{\text{Phe}}\) RNA *in vitro* but will not cleave \(^{16}\)RNA, again indicating a difference in these two enzyme systems. The halophilic endonuclease was also not inhibited by mature tRNA, therefore mature tRNA-like structure was not likely to be a primary recognition target for the archaebacterial endonuclease.

To define which regions of pretRNA\(^{\text{Trp}}\) were required for substrate recognition by the endonuclease, a series of deletion derivatives of tRNA\(^{\text{Trp}}\) genes were constructed. The first set of altered substrates were designed to examine the
requirement for exon structure on substrate recognition. These substrates lacked the acceptor stem ($^\text{13}$), acceptor stem plus $\Psi$C stem and loop ($^\text{13}_1^1$), and acceptor stem, $\Psi$C stem and loop, and DHU stem and loop ($^\text{13}_1^1-5$). In each case, the halophilic endonuclease cleaved the modified RNA substrate accurately and precisely. This definitely established that the halophilic endonuclease was not dependent upon mature tRNA-like structure in the precursor, an absolute requirement of the yeast endonuclease. Some exon sequences were required; a deletion which removed the entire exon 2 region, $^\text{13}_1^3$, was not cleaved at the 5' cleavage site by the endonuclease. Later experiments demonstrated a requirement for an intact exon-intron boundary junction.

A role for complete intron sequence and structure were also examined. The possible existence of a stem and loop structure within the intron (Figure 6), which had a high sequence similarity to the anticodon loops of tRNA$^{\text{TYR}}$ and tRNA$^{\text{ILE}}$, was investigated as a potential recognition element. Exonuclease III deletions originating from the BclI restriction site within the intron removed the putative stem and loop ($^\text{13}_2^4$) or removed the 3' half of the stem and loop plus half of a potential loop near the 3' intron-exon junction ($^\text{13}_1^0$). These substrate RNAs were cleaved by the endonuclease, indicating that the 3' half of the intron, including any potential internal structure, was not required
for cleavage. A larger deletion was introduced into \(^{16}\) to test the role of the 5' half of the intron. Taking advantage of two HinP1 sites located almost equidistantly from their respective exon-intron junctions, an 82 bp HinP1 intron fragment was deleted from \(^{16}\); this minimal-intron substrate, \(^{16^7}\), had an intron of 22 nucleotides. This substrate was cleaved by endonuclease, although less efficiently than the parental substrate \(^{16}\). Intact intron, then, was not an essential requirement for substrate recognition. These results indicated that recognition was directed to the exon-intron boundary regions. Moreover, the archaebacterial pre-tRNAs must possess specific sequences or structural elements since the halophilic enzyme could distinguish between the yeast tRNA\(^{\text{Phe}}\) and the \(^{16^7}\) substrates.

Nuclease S1 structure probing of \(^{16}\) RNA indicated that the structure of \(^{16}\) was best represented by Figure 15, where the intron was base paired into an extended helix. This structure also predicted the existence of a structural feature which consisted of two, three nucleotide loops separated by four base pairs, where cleavage by the endonuclease occurred within the two loops. An examination of the archaebacterial exon-intron boundaries from intron-containing RNA precursors (both tRNA and rRNA), revealed the potential to form similar structural motifs at the exon-intron junctions of all intron-containing precursors.
Because this motif was highly conserved, it seemed likely that its presence was significant and might be a signal for substrate recognition. An analysis of the possible role of this conserved structure was therefore undertaken.

To avoid potential mispairing after mutagenesis and to generate a substrate that closely resembled the predicted in vivo precursor, the tRNA\textsuperscript{Trp} gene was reconstructed. This gene was reconstructed such that transcripts would begin with the first nucleotide of exon 1 and terminate with CCA residues. Reconstruction was done in two steps, first constructing a synthetic gene for the small intron precursor \textsuperscript{oI6}, followed by the addition of the remaining intron sequence to give the complete gene \textsuperscript{oI6}. Both \textsuperscript{oI6} and \textsuperscript{oI6}, like the original genes \textsuperscript{~16} and \textsuperscript{~16}, produced transcripts which were accurately cleaved by the endonuclease. These synthetic genes were then used in subsequent mutagenesis protocols.

Initial analysis was directed towards establishing whether integrity of this conserved structure was essential for substrate recognition. For this analysis, the four nucleotides between the two loops at the 3' side of the intron were deleted. This mutation resulted in the alteration of the exon-intron boundary structure, while retaining the loop nucleotides in an internal loop. If nucleotide recognition alone was necessary for cleavage to occur, this altered substrate would be an acceptable
endonuclease substrate. The RNA from this gene, \(^{^\text{GGAG}}\), was not cleaved by the endonuclease. This argued that substrate recognition was not solely dependent on the sequences at the exon-intron boundaries or the distances from the cleavage sites to the top of the anticodon stem. It appeared that a specific structure, defined by the two bulge loops, was required for cleavage.

Studies on bulge loops in DNA helices have shown that these structures cause sharp bends or kinks in the molecule; the extent of the bend is determined by the size of the loop (9, 10, 174). In addition, studies with model RNAs indicate that bulge loops in RNA helices also result in kink formation. The occurrence of kinks or bends in a RNA molecule is associated with a distinct change in the mobility of the molecule in nondenaturing gel electrophoresis. Computer modeling of the conserved structural region of \(o_16\) with RNASE (64), an RNA structure analysis computer program, predicted a kinked or zigzag-like structure for the anticodon helices of \(o_16^7\) and \(o_16\) RNAs. \(o_16\) and \(^{^\text{GGAG}}\) RNA do in fact have differing migration patterns through 8% nondenaturing polyacrylamide gels (data not shown), which is consistent with a change in the shape of these molecules. Mutational analysis of this region was then initiated to determine if a defined structure was required at the exon-intron boundaries and if so, were the two cleavage sites recognized independently.
Mutations were introduced into o16 to test the requirement for the loop structures. The first alteration removed the three nucleotides which comprised the 3' loop, AUA. The resulting extended anticodon helix would only have a single loop at the 5' cleavage site. This RNA substrate was not cleaved by the endonuclease, either nonspecifically or specifically at the 5' cleavage sites. Two alterations were introduced at the 5' cleavage site. Rather than deleting the 5' cleavage loop nucleotides as was done for the 3' loop, the nucleotides UCU were introduced into the intron adjacent to the 5' cleavage loop such that the 5' loop nucleotides would be base paired as a part of an extended helix. The anticodon helix would have only a single loop at the 3' cleavage site. This substrate was also incapable of being cleaved by the endonuclease. Since neither single loop structure was a substrate for endonuclease, it appeared that a two bend or zigzag structure was necessary for cleavage. To distinguish whether this was a specific requirement for a structure defined by two bulge loops, three nucleotides (AGA) were added to the intron. The resulting RNA substrate had a six nucleotide internal loop, rather than a three nucleotide bulge loop, at the 5' cleavage site. The AGA nucleotides at the 5' cleavage site were accessible to the endonuclease but present in a slightly different structural context. When
this substrate was assayed with the endonuclease no cleavage was evident.

These experiments, probing structural elements within the conserved motif, demonstrated that a specific structure was absolutely required for endonuclease cleavage. It is likely that the proposed kinked structure, maintained by the two, three nucleotide bulge loops, is an essential element in substrate recognition by the halophilic endonuclease. Substrates having only one bend were not cleaved and may not be bound by the endonuclease. Although AGA, with the six nucleotide bulge, has intact cleavage sites, it must not have the correct structure to be cleaved by the endonuclease. These results may also explain the inability of the endonuclease to cleave yeast pretRNA. Even though the intron was located in the same position as the tRNA$^{\text{Tip}}$ intron, the yeast substrate lacked the required structural elements needed for cleavage.

Identification of the cleavage sites by the endonuclease was also addressed. The eukaryotic nuclear tRNA intron endonuclease has been shown to identify its cleavage sites via a measurement mechanism. Yeast and Xenopus endonucleases count a fixed number of nucleotides down the anticodon helix, from an anchoring point within the mature portion of the precursor, to the cleavage sites (136, 172). This was established when insertion of a nucleotide pair in the anticodon stem above the cleavage sites shifted
the cleavage sites such that the excised intron was one nucleotide larger than the wild type intron at each terminus. In addition, insertion of a nucleotide pair between the two cleavage sites resulted in correct cleavage at the 3' cut site, releasing a correctly sized 3' exon and an intron two nucleotides larger (136, 172). This was again consistent with a measurement, where the two cleavage sites are identified by their distance from the top of the anticodon stem.

Two sets of experiments were designed to determine whether this type of measurement mechanism was utilized by the halophilic endonuclease to identify its cleavage sites. Testing for measurement down the anticodon helix was performed by inserting either a single C or a GC pair in the 5' exon, above the bulge loop region. These mutations had little effect on the efficiency of cleavage. The single insertion substrate, C29:1, had an excised 5' exon one nucleotide larger than wild type, and the correct sized intron and exon 2. The second substrate, C29:1-G146:1, had excised exon 1 and 2 one nucleotide larger than wild type exon 1 and 2, and the correct sized intron. These data indicated that, unlike its eukaryotic counterpart, the endonuclease does not measure the length of the anticodon helix from a fixed point within the mature portion of the precursor to identify the cleavage sites.
In a second experiment a single C residue (C36:1) or a CG pair (C36:1-G136:1) were inserted into o16 between the two loops to investigate whether the identification of cleavage sites involved measurement between the two loops. Although substrate specificity and turnover were significantly reduced, end group analysis of the excised intron and 3' exon of these altered substrates indicated that the majority of cleavage took place in correct positions. In the case of the single C addition (C36:1), in which the 5' cleavage site loop was increased from three to four nucleotides, a minor amount of "slippage" was detected.

Identification of the two cleavage sites does not appear to utilize a simple measurement mechanism, however the distance between the two loops or loop size has an effect on cleavage efficiency. The ability of the endonuclease to find the correct cleavage sites in these mutants suggests that recognition of specific nucleotides may also be involved. It is possible that the two bulge loops orient the substrate on the surface of the enzyme so that specific nucleotide contacts at the cleavage sites can be made. Binding of substrate may also involve helical regions above and below the cleavage sites. Interestingly, o16^, bearing the C36:1-G55:1 mutation, is not cleaved by the halophilic endonuclease (Figure 30). The only difference between o16 C36:1-G136:1 and o16^ C36:1-G55:1 is the size of the intron, and the extent of helix below the 5'
exon-intron boundary junction. It is possible that the extended helix in 016 helps to stabilize the substrate-endonuclease interaction. This interaction may allow 016 C36:1-G136:1 to be "locked" into the proper orientation for cleavage, while 0167 C36:1-G55:1, lacking this extended helix, cannot be securely held in place, and the cleavage sites are then inaccessible to the endonuclease. The binding of the extended intron, and possibly the anticodon helix, would also tend to suppress the rotation of the cleavage sites (approximately 36° per base pair) imposed by the base pair addition. The interaction with the helical regions bordering the cleavage loops may also explain why 016 is more efficiently cleaved than 0167.

A role for sequence specific recognition was also suggested by the behavior of mutants with base pair changes between the two cleavage loops of 0167. When the base pair C35:A57, between the two cleavage loops, was replaced with U35:A57 in 0167 the resulting RNA was cleaved at a much lower efficiency. The initial interpretation was that the UA pair replacing the CG pair destabilized the base pairing between the two cleavage loops resulting in an open structure. Free energy calculations (216) predicted that the U35:A57 structure would be ~1 kcal less stable than the wild type structure. If this were occurring, then replacing the U35:A57 pair with a GC pair should restore wildtype stability to this region. However, when this mutation,
192

G35:C57, was created it was cleaved only slightly better than the U35:A57 substrate. It is possible that these nucleotides may be required for sequence-specific substrate recognition by the endonuclease. Other possibilities for sequence specific contacts might include the highly conserved C and U residues of the anticodon loop, the loop nucleotides, and those of the exon and intron helices.

Based upon the predicted structural and nucleotide specific recognition features required by the endonuclease, a model tRNA intron endonuclease substrate was constructed. This RNA substrate, "Minsub", consisted of a five nucleotide stem, two, three nucleotide loops separated by a four nucleotide helix, and an intron of 17 nucleotides (Figure 26). This RNA also had the sequences present at the exon-intron boundary junctions of the ol67 substrate. When assayed, this substrate was cleaved into discrete products by the endonuclease. RNA sequence analysis of the excised intron-sized fragment indicated that Minsub RNA was cleaved accurately and precisely at the predicted exon-intron boundaries. This data supports the proposal that the conserved structure and sequences at the exon-intron boundaries are sufficient for substrate recognition and cleavage. This model predicts cleavage of any intron-containing precursor tRNA, by the endonuclease, regardless of intron location, as long as this cleavage motif is present. This also implies that the endonuclease's
substrate range could include non-tRNA molecules, provided they had the necessary recognition elements.

**Cleavage of nontRNAs by the halophilic endonuclease.**

Since the endonuclease does not distinguish whether its substrate is a tRNA molecule, it was possible that other RNAs may be substrates for cleavage by this enzyme. The structural features required by the endonuclease, two loops of three nucleotides separated by four base pairs, most closely resembles the structure of eubacterial RNase III cleavage sites (176). RNase III cleaves at discontinuities, or loop regions, in the large hairpins which flank the 16S and 23S rRNAs in the rRNA polycistronic transcript. A survey of halobacterial RNase III-like cleavage sites revealed a striking similarity in structure to the predicted endonuclease cleavage site (87, 125, 132). Preliminary data using model substrates indicated that the halophilic endonuclease could cleave RNA substrates mimicking RNase III-like cleavage sites of 16S and 23S rRNA from *H. cutirubrum*. Thus, the tRNA intron endonuclease may act as an RNase III-like activity, releasing the individual rRNAs from the primary transcript. Whether a separate RNase III activity is present in the halophiles is not known. However, if such an activity were present, and it was similar to the eubacterial enzyme, it would differ from the tRNA endonuclease in its mechanism of phosphodiester bond
cleavage. RNase III from *E. coli* cleaves the phosphodiester bond to give 5' phosphates and 3' hydroxy termini, whereas the endonuclease cleaves to give 5' phosphates and 2',3' cyclic phosphate termini. The occurrence of the structural elements in the RNase III cleavage sites of methanogen (78) and *Thermoplasma* (R. Zimmermann, personal communication) rRNA primary transcripts suggests that the multifunctional property of this enzyme may extend beyond the halophiles to other members of the methanogen-halophile branch of the archaebacteria. RNA cleavage by the endonuclease may not be limited to intron-containing pre-tRNAs and rRNA primary transcripts. The only non-tRNA archaebacterial intron, the 23S rRNA intron of *D. mobilis*, has a very similar structure at its exon-intron boundaries, and it has been suggested that a tRNA intron endonuclease with recognition properties similar to the halobacterial enzyme removes this intron (102). The archaebacterial tRNA intron endonuclease might function in a more global role as a general stable RNA maturase, maturing rRNAs and removing introns from pre-tRNAs and rRNAs. Further experiments will be necessary to determine the extent to which endonuclease is involved in rRNA maturation and the relationship, if any, which exists between the various endonucleases.
Presence of endonuclease activity in other archaebacteria.

Since there was structural relatedness between the RNase III-like cleavage sites and the tRNA intron endonuclease sites in the halobacteria, the search for tRNA<sub>Trp</sub> cleavage activity was extended to other archaebacteria. Examination of RNase III sites from the methanogens and Thermoplasma revealed the conserved structural element. It was possible, even though interrupted genes had not been reported for these organisms, that an intron endonuclease activity like the H. volcanii enzyme could be present. Extracts from H. cutirubrum, S. solfataricus, T. acidophilum, Methanosarcina barkeri, and Methanothermus fervidus were assayed for endonuclease activity capable of cleaving 016 RNA. In each case endonuclease activity could be detected. These data established that an intron endonuclease activity was present in these organisms regardless of whether interrupted tRNA genes were present. It may be too premature to speculate on whether methanogens and Thermoplasma have interrupted tRNA genes as the number of genes characterized is still small. S. solfataricus, however, has been shown to have interrupted tRNA genes and its pre-tRNAs can assume structures similar to the halophilic pre-tRNA structure.

To better understand the relationship of these activities, the Thermoplasma enzyme was examined in more detail. As expected, the temperature profile of this
thermophilic endonuclease was different from that of the halophilic endonuclease; its maxima was 58°C, consistent with its higher optimal growth temperature. The thermophilic endonuclease did show the same substrate range as the halophilic endonuclease; it was capable of cleaving not only intact o16 RNA, but also the reduced intron o16^7 and minimum exon ^13-11-5 RNAs. End group analysis of the released intron and 3' exon also showed that cleavage of o16 RNA by T. acidophilum endonuclease was accurate and precise, and did not represent the activity of a general ribonuclease. The T. acidophilum endonuclease exhibited an apparent difference in nucleotide specificity when compared to the halophilic enzyme. The thermophilic endonuclease was capable of cleaving the o16^7 U35:A57 RNA more efficiently than the endonuclease from H. volcanii (data not shown). A differing nucleotide specificity for endonucleases originating from different host organisms may represent a "fine-tuning" to its particular substrate. In this organism, the endonuclease appears similar in substrate recognition and cleavage properties to the halophilic enzyme.

Archaebacterial ligase.

The reaction which occurs after tRNA intron excision is the ligation of the half-tRNAs. At no time, however, was mature tRNA^Trp detected in an H. volcanii endonuclease.
cleavage assay containing ATP. To eliminate the possibility that the ligase may only function under high salt conditions, the monovalent ion concentration of the assay reaction was raised to 2 M KCl following cleavage of 016 RNA by the endonuclease. No ligated exons were observed after incubation of the assay under these conditions. A similar observation was made using crude H. volcanii extracts prepared in TMG buffer containing 2 M KCl (data not shown). In addition, no ligase activity was detected in cleavage assays preformed with any methanogenic or thermophilic extracts. Combined endonuclease-ligase assays with the yeast extract indicated that, although 016 was inefficiently cleaved by the yeast endonuclease as compared to tRNA\textsuperscript{Phe}, half-tRNA\textsubscript{Trp} molecules could be ligated into mature tRNA\textsubscript{Trp} (data not shown). This data, in addition to the observation that mature tRNA\textsubscript{Trp} is detected in Northern blot analysis of H. volcanii RNA, indicates that these molecules are substrates for ligation and this event must occur \textit{in vivo}. Appropriate conditions for the ligase activity have yet to be established \textit{in vitro}.

It has been proposed that the yeast ligase makes the initial contact with the precursor tRNA, binding to the molecule and transporting it to the endonuclease which is membrane associated (71). It is at this splicing complex where the intron is excised, the half-tRNAs ligated, and the matured tRNA translocated across the nuclear membrane into
the cytoplasm. Recognition of the pretRNA by the ligase is probably mediated by structural queues rather than nucleotide specificity, likely a result of the conserved location of the introns in precursor tRNAs. Recognition of archaebacterial precursor tRNAs by the tRNA ligase probably occurs by a different recognition mechanism, which is reflected in the variable location of archaebacterial introns. For example, the tertiary structure of pretRNA<sub>Gly</sub> from *T. pendens*, whose intron is located within the variable arm of the pretRNA, would be expected to be much different from tRNA<sub>Trp</sub> of *H. volcanii*, whose intron is found in the anticodon stem. Even within an organism the intron location can vary, as is evident in the precursor tRNAs from *T. tenax* and whose intron are located within both the anticodon and the anticodon stem. Here, tRNA recognition by the ligase may be directed to the structure of the annealed half-tRNAs, after the intron has been excised, since once the intron has been removed all pretRNAs would have the same tertiary structure. Some level of nucleotide specificity by the ligase would not be unexpected, since ligated or non-intron containing tRNAs may otherwise compete with cleaved tRNA for the ligase. In this case the anticodon or the aminoacceptor stem might be a good candidate for nucleotide specific recognition, a recognition mechanism employed by many tRNA synthetases. One could envision in archaebacteria that initial substrate recognition and binding would occur by the
Intron endonuclease. After cleavage by this enzyme the substrate is released and interacts with the ligase, which binds to the annealed half-tRNAs and ligates the molecule.

Evolution of tRNA processing systems.

The relationship between the archaebacterial tRNA introns and their eukaryotic counterparts is an interesting question. Chloroplast and nuclear tRNA introns are readily distinguished from each other by their complexity and processing mechanisms. The archaebacterial tRNA introns appear to share some properties with both groups. Their relatively small size is similar to the nuclear encoded tRNA introns, however their variability is more like that observed for the chloroplast introns. Since archaebacterial tRNA introns are found in various locations, their processing enzymes, by necessity, must use a different mechanism for identifying the cleavage sites. The requirement for a defined structure at the exon-intron boundaries would provide the flexibility needed to identify introns at different locations, provided the conserved structure was present. The presence of tRNA<sub>Trp</sub> intron endonuclease activity in other archaebacteria supports the likelihood of a conserved intron endonuclease. Variability in intron location also suggests that the introns in the archaebacterial tRNA genes were not derived from a single ancestral intron-containing tRNA species; more likely these
introns were inserted, possibly as self-cleaving mobile introns, into preexisting tRNA genes. It has been proposed (82) that tRNA introns were originally capable of self-cleavage and that modern tRNA introns are the result of independent solutions to the transition from self-cleaving introns to protein assisted or protein mediated processing. The properties of the halophilic endonuclease, and the structural properties of archaebacterial tRNA introns, support this proposal. Further studies on the archaebacterial endonuclease proteins may reveal the origin of the protein components of these processes.
Chapter VI
Conclusion

Using an *in vitro* cleavage assay, the halophilic tRNA\textsubscript{Trp} intron endonuclease has been partially purified and biochemically characterized. Analysis of the cleavage mechanism reveals similarity to that of the yeast nuclear tRNA intron endonuclease. Cleavage of tRNA\textsubscript{Trp} by the halophilic endonuclease is accurate and precise.

An examination of substrate requirements demonstrates that the archaebacterial intron endonuclease exhibits no dependence upon mature tRNA-like structure for substrate recognition and cleavage, unlike its eukaryotic counterpart. Furthermore, the halophilic endonuclease does not rely upon a measurement mechanism from the top of the anticodon stem to the exon-intron boundaries to identify its cleavage sites, but rather recognizes a conserved structural element, present in many archaebacterial intron containing RNA precursors, and cleaves the substrate within the two, three nucleotide loops.
Based on cleavage requirements, a minimum endonuclease substrate was designed and synthesized. This substrate consists of two, three nucleotide loops flanking a four basepair stem and a small intron sequence. This substrate was cleavable by the endonuclease, demonstrating that this structural element is all that is necessary for recognition and cleavage by the endonuclease.

The ability of the endonuclease to cleave substrates mimicking RNase III-like cleavage sites of rRNAs suggests that the endonuclease may function in a more global role as a general stable RNA maturase. Endonuclease activity was detected in all archaebacterial strains examined, suggesting these organisms contain interrupted genes or the conserved recognition structure. Since no introns have been found to interrupt methanogen genes, the presence of endonuclease in these organisms supports the possible role of the endonuclease as an archaebacterial RNA maturase.

The ability of the halophilic endonuclease to recognize and cleave its substrate at the conserved recognition element is an elegant solution to the problem of recognizing and properly cleaving precursor tRNAs bearing introns in widely different and unrelated locations.
Commonly used buffers and reagents.

20% denaturing polyacrylamide- 19.0% acrylamide (w/v), 1.0% bis-acrylamide (w/v), 50% urea (w/v), 20% 5x TBE (v/v); filtered through Whatmann 3M filter paper.

20xSSC- 175.3 g NaCl, 88.2 g sodium citrate per liter H₂O, pH 7.5.

2x "partial mix"- 100 mM HEPES pH 7.2, 30 mM MgCl₂, 20% DMSO.

5x DNA loading buffer- 100 mM EDTA pH 8.0, 10% Ficol, 0.075% Bromphenol Blue.

6% denaturing polyacrylamide- 5.7% acrylamide (w/v), 0.3% bis-acrylamide (w/v), 50% urea (w/v), 20% 5x TBE (v/v); filtered through Whatmann 3M filter paper.

8% non-denaturing acrylamide- 18.6 ml of 29:1 (w/v) acrylamide:bis-acrylamide solution in H₂O, 42 ml 5x TBE, 107.7 ml H₂O.

A nucleotide mix- 50 mM NaCl, 10 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 1 mM DTT, 60 μM ddATP, 1.7 μM dATP, 30 μM dGTP, 30 μM dCTP, and 30 μM TTP.
G nucleotide mix- 50 mM NaCl, 10 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 1 mM DTT, 100 µM ddCTP, 1.7 µM dATP, 4.2 µM dCTP, 40 µM dGTP, and 40 µM TTP.

G nucleotide mix- 50 mM NaCl, 10 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 1 mM DTT, 150 µM ddGTP, 1.7 µM dATP, 4.2 µM dGTP, 40 µM dCTP, and 40 µM TTP.

H. volcanii lysis buffer- 10 mM Tris-HCl pH 7.4, 0.5% SDS.

Phenol- phenol was equilibrated against a solution containing 10 mM Tris-HCl pH 8.0, 1% 8-hydroxyquinoline, and 1 mM β-mercaptoethanol.

RNA elution buffer- 500 mM NH₄COOCH₃, 10 mM MgCOOCH₃, 100 mM EDTA pH 8.0, 0.1% SDS.

RNA loading buffer- 20 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0, 8 M urea, 0.05% Bromphenol Blue, 0.05% Xylene Cyanol.

STET- 60 ml 25% sucrose (w/v), 9 ml 1M Tris-HCl pH 8.0, 36 ml 0.25M EDTA pH 8.0, 9 ml Triton X-100, and 66 ml H₂O.

T nucleotide mix- 50 mM NaCl, 10 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 1 mM DTT, 120 µM ddTTP, 1.7 µM dATP, 40 µM dCTP, 40 µM dGTP, and 1.7 µM TTP.

T4 polynucleotide kinase buffer- 10x stock solution contains 0.5 M Tris-HCl pH 7.6, 0.1 M MgCl₂, 50 mM DTT, 1 mM spermidine-HCl, 1 mM EDTA pH 8.0.

TBE- 89 mM Trizma base, 89 mM boric acid, 2.5 mM EDTA.

TE- 10 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0.

TMG- 40 mM Tris-HCl pH 7.5, 20 mM MgCl₂, and 2% (v/v) glycerol.
TSG - 40 mM Tris-HCl pH 7.5, 5 mM spermidine, and 2% (v/v) glycerol.

YPD - 1% yeast extract, 2% peptone, and 2% glucose per liter of H₂O.

YT - 8 g tryptone, 5 g yeast extract, 5 g NaCl per 1000 ml H₂O. YT plates also contained 1.5% Oxoid or Difco Bacto agar.
Table 6

**List of Oligonucleotides**

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5' to 3')</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δ3'Arm</td>
<td>CTCTGGAGTCAGTCGCTCGTAATTCGTAATC</td>
<td>1</td>
</tr>
<tr>
<td>ΔGGAG</td>
<td>CAGCTTGATATGGTGTGCG</td>
<td>2</td>
</tr>
<tr>
<td>ΔAUA</td>
<td>CGATCGACTGACTCCGGGTGTGCG</td>
<td>3</td>
</tr>
<tr>
<td>UCU</td>
<td>CTGATATCTCTCTGTTGTGCG</td>
<td>4</td>
</tr>
<tr>
<td>AGA</td>
<td>CTGATATCTCCAGGTGTGCG</td>
<td>5</td>
</tr>
<tr>
<td>C29:1</td>
<td>CTGGAGTCAGTCGCCATG</td>
<td>6</td>
</tr>
<tr>
<td>G146:1</td>
<td>CCGATCGACCTGATATCTCC</td>
<td>7</td>
</tr>
<tr>
<td>C36:1-G136:1</td>
<td>CTGATATCTCCGGGTGTGCGCCAAGCCTCTTGGGAGT</td>
<td>8</td>
</tr>
<tr>
<td>C:36:U36</td>
<td>CCAAGCCTCTGAAGTCAGTCCG</td>
<td>9</td>
</tr>
<tr>
<td>G55:A55</td>
<td>CTGATATCTCCGTTGTGCGCC</td>
<td>10</td>
</tr>
<tr>
<td>C36G55:G36C55</td>
<td>CTGATATCTCCGGGTGTGCGCCAAGCCTCTTGGGAGT</td>
<td>11</td>
</tr>
<tr>
<td>MinsubB</td>
<td>ATAGTGAAGTCGATATAG</td>
<td>12</td>
</tr>
<tr>
<td>MinsubT</td>
<td>TCGACTAATACGACTCACTATAGGGAAGTCACTCCAGAG</td>
<td>12</td>
</tr>
<tr>
<td>T7Bottom</td>
<td>CTATAGTGATGTCGTATTA</td>
<td>13</td>
</tr>
<tr>
<td>T7Top</td>
<td>TAATACGACTATATAG</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>TCACCAAACCTATAGTATGCTGATAT</td>
<td></td>
</tr>
<tr>
<td>HC16S</td>
<td>GGTCTTGATATGCTCGCCACTGCCAGTGAGTTCG</td>
<td>15</td>
</tr>
<tr>
<td>HC23S</td>
<td>GCCGTCAGTCGCCGCCGAGGCTGAGTGAGTTCG</td>
<td>16</td>
</tr>
<tr>
<td>Trp1</td>
<td>GATCTGTATACGACTCAGTATAGGCGAGATAG</td>
<td>17</td>
</tr>
<tr>
<td>Trp2</td>
<td>CTTGGCCACAGCCCTATAGTGCTGATAT</td>
<td>17</td>
</tr>
<tr>
<td>Trp3</td>
<td>TGCCGAGGCCAGCCTGAGTGAGT</td>
<td>17</td>
</tr>
<tr>
<td>Trp4</td>
<td>TGGATGTCAGTGCGGCTCCAGCCAG</td>
<td>17</td>
</tr>
<tr>
<td>Trp5</td>
<td>ACTGACTTCCAGAGGGCTTGGGCCACACCGGAGATAG</td>
<td>17</td>
</tr>
<tr>
<td>Trp6</td>
<td>CTCCGGTAGTGCCAGCGCCAGCCCT</td>
<td>17</td>
</tr>
<tr>
<td>Trp7</td>
<td>CCCCGATCGAGATGAGTAG</td>
<td>17</td>
</tr>
<tr>
<td>Trp8</td>
<td>CGATCGGGGGTTTCAAATCCCTCCGGCCACACCAG</td>
<td>17</td>
</tr>
<tr>
<td>Trp9</td>
<td>GATCCCTGCTGGGGCCGGGAGGGGATTTGAA</td>
<td>17</td>
</tr>
<tr>
<td>Number</td>
<td>Purpose</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Deletion of DHU stem and loop in A16 gene</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Deletion of GAGG nucleotides between two loops in o16 gene</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Deletion of AUA nucleotides of 3' loop in o16 gene</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Insertion of UCU nucleotides in o16 gene to basepair with 5' loop nucleotides</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Insertion of AGA nucleotides opposite 5' loop nucleotides</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Insertion of C nucleotide in o16 gene to create C29:1 gene</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Insertion of G nucleotide into C29:1 gene to create C29:1-G146:1 gene</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Nucleotide conversion of C36 to U36 in o16\textsuperscript{7} gene</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Nucleotide conversion of G55 to A55 in o16\textsuperscript{7} gene</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Nucleotide conversion of C36:G55 to G36:C55 in o16\textsuperscript{7} gene</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Construction of Minsub gene</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Protection of T7 RNA polymerase promoter sequence from binding oligonucleotides during oligonucleotide directed mutagenesis</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Annealed to HC16S and HC23S oligonucleotides for synthesis of RNA</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Oligonucleotide coding for HC16S RNA</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Oligonucleotide coding for HC23S RNA</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Reconstruction of o16\textsuperscript{7} gene</td>
<td></td>
</tr>
</tbody>
</table>
Table 7

**Bacterial Strains**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Escherichia coli</em> MV1190</td>
<td>Δ(lac-pro AB), thi, supE, Δ(srl-recA) 306::Tn10(tet') [F':tra D36, pro AB, lacI^q ZΔM15]</td>
<td>Biorad</td>
</tr>
<tr>
<td><em>Escherichia coli</em> CJ236</td>
<td>dut, ung, thi, relA; pCJ105(Cm^F)</td>
<td>Biorad</td>
</tr>
<tr>
<td><em>Escherichia coli</em> TJ1</td>
<td>endA1, hadR (r_k^-, m_k^+), supE, sbc, thi, strA, Δlac-pro), l^−, [F':tra D36, pro AB, lacI^q ZΔM15]</td>
<td>BRL</td>
</tr>
<tr>
<td><em>Escherichia coli</em> JM101</td>
<td>Δ(lac-pro AB), rpsL, F80, ΔlacZΔM15, hsdR17(r_k^−, m_k^+)</td>
<td>BRL</td>
</tr>
<tr>
<td><em>Halobacterium volcanii</em> DS2</td>
<td>ATCC 29605, type strain</td>
<td>W.D. Doolittle</td>
</tr>
<tr>
<td><em>Saccharomyces cerevisiae</em> EJ101</td>
<td>trp 1, pro1-126, prb1-112 pep4-3, pcr1-126</td>
<td>J. Abelson</td>
</tr>
<tr>
<td><em>Sulfolobus solfataricus</em> 1616</td>
<td>ATCC 35091, type strain</td>
<td>ATCC</td>
</tr>
<tr>
<td><em>Thermoplasma acidophilum</em> 122-1B2</td>
<td>ATCC 25905, type strain</td>
<td>T. Langworthy</td>
</tr>
</tbody>
</table>


