RIBONUCLEASE P (RNase P) is a ribozyme (an enzyme composed of RNA) that is involved in tRNA biosynthesis. RNase P catalyzes the endonucleolytic removal of 5' leader sequences from precursors of tRNA (pre-tRNA), which generates the mature 5' ends of tRNAs. Every cell and every eukaryotic subcellular compartment that synthesizes tRNAs has been shown to contain RNase P or an RNase P-like activity, and the enzymes generally have been shown to contain an RNA subunit.

Further insight into the catalytic core of the RNase P RNA structure is now being sought by the acquisition of new RNase P RNA sequences from widely divergent organisms. In the following sections we describe variations in known examples of RNase P and comment on some of the recent developments in our understanding of its structure. We use the nomenclature for the primary phylogenetic domains proposed by Woese. ‘Bacteria’ refers to organisms previously known as eubacteria, ‘Archaea’ refers to the group previously called archaebacteria and ‘Eukarya’ refers to the eukaryotic nuclear lineage.

RNase P in bacteria

The bacterial RNase P, for instance that of Escherichia coli or Bacillus subtilis, is by far the best-characterized version of the enzyme. In vivo, the bacterial RNase P holoenzyme comprises both an RNA and at least one protein subunit. In vitro, at high ionic strength, the RNA is catalytically active in the absence of the protein. In contrast, none of the RNAs derived from the eukaryal, the archaeal or the organellar enzymes is active after deproteinization. Thus, only the bacterial RNase P RNA can be shown directly to be the catalytic subunit of the enzyme.

The detailed structure of the bacterial holoenzyme is as yet unknown, although footprinting analysis has suggested some regions of interaction between the RNA and the protein subunit. The role of the protein in the function of the enzyme in vivo is unknown. However, the RNase P protein does not appear to be involved in the specificity of substrate binding; the holoenzyme $K_m$ for substrate is the same as that displayed by the RNA subunit in the absence of the protein subunit at high ionic strength. Since the protein can be replaced by high ionic strength, it is likely that the function of the protein is coulombic, to reduce electrostatic repulsion, due to phosphates between enzyme and substrate RNAs or within the RNase P RNA superstructure. Several mutations in the RNA have been reported which render the RNA catalytically inactive except in the presence of the protein subunit. This indicates that the protein stabilizes the active conformation of the RNase P RNA, in a similar way to the role imagined for some ribosomal proteins; in those cases stabilizing the folded structure of RNA.

The bacterial RNase P RNA has been cloned and sequenced from representatives of five of the approximately 11 main bacterial lineages. An alignment of most of the known bacterial RNase P RNA (and protein) sequences has recently been published. The sizes of bacterial RNase P RNAs vary from approximately 338 to approximately 444 nucleotides (nts); however, even the homologous sequences within the term ‘homologous’ in its strict sense: homologous sequences have common ancestry.) In view of the wide distribution and essentiality of RNase P, it is surprising that there is so little sequence conservation in the RNA subunit. For instance, only 17% of the tuts are identical in all bacterial RNase P RNAs (compare the E. coli RNA to the consensus structure in Fig. 1). Thus, the essential features of the RNA enzyme probably lie in its conserved secondary and tertiary structure rather than in conserved sequence motifs. The only conserved contiguous sequence motif is an 11-nt segment, GAGGAaGUCc, which is present in all known bacterial sequences (with the exception of the lower case a, which is a C in one instance.) This sequence corresponds to nts 61–71 in the E. coli RNA (Fig. 1).

Secondary structures of the RNase P RNA from E. coli and B. subtilis, and a minimal consensus structure, are shown in Fig. 1. These structures are comprised of many local hairpin elements, as well as helices formed by pairing of widely distant portions of the sequence. These helices are well proven by instances of RNase P RNAs in which paired bases co-vary to maintain the complementarity. For the purposes of this review, we highlight the long-range base-pairings by labeling each with a letter (see Fig. 1 legend). These long-range interactions organize the RNA structure into a central core that has been modified through evolution by the addition of local helical features.
hairpin elements. The presence and absence of these hairpins are evident in the E. coli and B. subtilis structures in Fig. 1. Note that the relative position of the 11-nucleotide conserved sequence motif is maintained in both structures.

Comparison of the E. coli and B. subtilis structures in Fig. 1 highlights another important point. The E. coli RNA, and indeed most other bacterial RNase P RNAs, forms long-range helix E while the RNAs from members of the low G+C content Gram-positive bacterial group, such as B. subtilis, do not. The Bacillus sequences lack one of the pairing elements, the helix composed of nts 260–290 in the E. coli RNA. Helix E was recently identified by phylogenetic comparisons and site-directed mutagenesis. It is known that helix E is important because mutations which disrupt base pairs within the region are lethal. In addition, in vitro analysis has demonstrated that deletion of E. coli nts 260–290, which eliminates one of the pairing partners of helix E, reduces substrate binding affinity by 100-fold.

substrate-binding site. In contrast, the B. subtilis RNase P, which apparently does not contain helix E, exhibits wild-type affinity for substrate. The Bacillus RNase P RNAs apparently have developed a structural adaptation that compensates for the loss of helix E: the E. coli-type RNA, containing helix E, is clearly the ancestral version (below). It has been proposed that the short stem-loop at nts 62–85 in the B. subtilis RNA produces a structure that is similar to the E. coli helix E (Ref. 9). However, if this structure indeed replaces helix E, it cannot act alone. When nts 62–85 and 250–261 of the B. subtilis RNA were inserted into the E. coli RNase P RNA in place of the E. coli structures in those regions, wild-type affinity for substrate was not produced (S. Darr, K. Zito and N. Pace, unpublished). Other parts of the B. subtilis structure must contribute to the substrate-binding site.

The structure of the RNase P RNA from Thermotoga maritima, a representative of the most deeply branching bacterial phylum so far defined, is of the T. maritima RNA to that of E. coli (Fig. 1) is striking. Since T. maritima and E. coli represent the most extreme poles in the phylogenetic domain of Bacteria, that common structure must also have pertained to the ancestor of all Bacteria.

Archaeal RNase P

The Archaea are the second major phylogenetic division of the prokaryotes, as different evolutionarily from Bacteria as are Eukarya. The nature of the archaeal RNase P reflects this deep evolutionary divergence, but similarities between the archaeal and bacterial RNase P RNAs are evident. The archaeal RNase P has been characterized most extensively from two species: Halobacterium volcanii, a member of the halophilic group, and Sulfolobus solfataricus, a thermoacidophile. The isolation or characterization of RNase P from a member of the third major physiological group of Archaea, the methanogens, has not yet been reported. In both of the examined cases, the archaeal RNase
in the absence of protein\textsuperscript{11,12}. In other properties, however, the two archaeal enzymes seem to be quite different from each other. The \textit{H. volcanii} enzyme was reported to be sensitive to nuclease treatment and to display a density of 1.61 g ml\textsuperscript{-1} in cesium sulfate gradients\textsuperscript{33}, as does bacterial RNase P, while the \textit{S. solfatarius} enzyme is resistant to nuclease treatment and has a density of 1.27 g ml\textsuperscript{-1} in cesium sulfate\textsuperscript{12}. These results suggest a considerably greater component of protein in the \textit{S. solfatarius} holoenzyme than in the \textit{H. volcanii} enzyme. However, nothing is known about the protein components of either enzyme.

The sequence of the RNA subunit from \textit{H. volcanii} is shown in Fig. 3. The small RNA subunit of the RNAse P RNAs from \textit{T. maritima}, based on phylogenetic comparisons (Ref. 9; J. W. Brown and E. S. Haas, unpublished), Watson-Crick pairs are shown with lines (-), non-Watson-Crick pairs with filled circles (*). The locations of the 5' and 3' ends of the RNA have not been determined experimentally, but are based on the mature ends of the \textit{E. coli} RNase P RNA. \textit{T. maritima} represents the earliest divergence in the bacterial lineage; the similarity of RNase P RNA secondary structures from this organism and all other characterized RNase P RNAs (with the exception of \textit{B. subtillis}, see text) indicates that this is the ancestral form of the RNA in the Bacteria.

In contrast, the \textit{S. solfatarius} RNA is only 306 nts in length (T. LaGrandeur, S. Darr and N. Pace, unpublished). However, this size difference between the \textit{S. solfatarius} and the \textit{H. volcanii} RNAs does not account for the apparent differences in the density or nuclease sensitivity. The \textit{S. solfatarius} holoenzyme must contain other features, perhaps related to its thermophilic character, which distort the density of the enzyme and protect the RNA from nuclease action.

Since the archaeal RNA subunits of RNase P are not by themselves catalytic, there is no direct proof that they are bona fide RNase P RNAs. The major evidence that these RNAs are in fact the RNA subunit of an RNase P is twofold. First, the RNAs always co-purify with enzyme activity, even though the RNAs and holoenzymes from the two organisms inspected are different in size. Second, and most convincing, the sequences of these RNAs allow them to fold into much of the secondary structure of the bacterial RNA. The bacterial and the archaeal RNAs are clearly homologous over much, but not all, of their lengths. The \textit{H. volcanii} structure (Fig. 3) correctly positions, in the overall context, the conserved 11-nt consensus sequence GAGGAAUGCC, the long-range pairings A-F and other features. The presence of some of the long-range pairs, including helix E, in the \textit{H. volcanii} structure is supported by sequence covariation data (E. Haas and C. Daniels, pers. commun.).

Organellar RNase P

Both the chloroplasts and the mitochondria of eukaryotes contain RNase P-like activity. In every case so far examined, these organellar varieties of RNase P have been surprisingly different from the bacterial prototype from which they evolved. Chloroplasts derived from cyanoacteria, which contain a normal bacterial RNase P RNA, while mitochondria stemmed from the ‘purple bacteria’ (‘proteobacteria’), the same phylogenetic group as \textit{E. coli}.

The chloroplast enzyme from spinach has to some extent been characterized\textsuperscript{14}. The enzyme endonucleolytically cleaves precursor tRNAs, to produce the same products as the bacterial RNase P. Unlike the bacterial enzyme, however, the chloroplast enzyme is of low density (1.28 g ml\textsuperscript{-1}) in cesium sulfate gradients and it is resistant to treatment with micrococcal nuclease\textsuperscript{14}. These observations and the seeming preparations of the enzyme led Gegenheimer and colleagues to suggest that the chloroplast RNase P does not contain significant RNA. However, the observation of density and nuclease resistance is not conclusive proof regarding the presence or absence of an RNA subunit. The RNase P from \textit{Sulfolobus solfatarius}\textsuperscript{12} also exhibits low density and is resistant to nuclease, yet it co-purifies with an RNA molecule. Further characterization of the chloroplast RNase P is required before any conclusions can be made regarding its lack of an RNA subunit.

In contrast, the mitochondrial RNase P clearly contains an RNA element. Several instances of this enzyme, including that of human\textsuperscript{15}, have been
most information on the RNA moiety derives from studies with yeast. Martin and colleagues have demonstrated that a locus in the Saccharomyces cerevisiae mitochondrial genome is required for mitochondrial tRNA processing and encodes the RNA subunit of RNase P. This RNA has some remarkable features: it is more than 85% A-U and it can be fragmented during purification into several small pieces without significant loss of activity in vitro. More recently, additional yeast mitochondrial RNase P RNAs, from Candida glabrata and Saccharomyces fibuliger, have been identified. These RNAs also have a high A-U content and their variation in size is remarkable. Whereas the S. cerevisiae RNA is 490 nts, larger than any of the known bacterial RNase P RNAs, the C. glabrata and S. fibuliger RNAs are much smaller, 227 and 140 nts, respectively. These latter RNAs are the smallest RNase P-associated RNAs known.

All of the known mitochondrial RNase P RNAs seem to contain a homolog of the 11-nt conserved sequence element found in the bacterial RNase P RNAs. Both sides of helix C are preserved in the fragments of RNA remaining in the isolated S. cerevisiae mitochondrial RNase P RNA. The high A-U bias of these RNAs makes it difficult to recognize other structures.

Like the archaean enzymes described above, the yeast mitochondrial RNAs are not catalytic when protein is absent. This may not be surprising, considering the sequence of these RNAs: the highly AU-rich sequences may not be able to adopt stable structures required for activity without an external scaffold. It may be that some other component of the enzyme, possibly a protein subunit, provides structural information for the RNA. This notion is consistent with the observation that fragmented RNAs can remain assembled in an active enzyme; the scaffold does not require intact RNAs to maintain an active structure.

**RNase P in Eukarya**

RNase P has been characterized from several species of animals and yeast, which is, of course, only a limited sampling of Eukarya. The partially purified enzymes have all been demonstrated to be sensitive to nucleases, suggesting that they have nucleic acid subunits that are essential for activity. Two characteristics, however, make the eu-

teral versions. The eukaryal enzymes are light in cesium sulfate gradients and much less dense than most ribonucleoprotein complexes; in several cases it is the density of a protein enzyme. In addition, these holoenzymes are large, e.g., 450 kDa in Schizosaccharomyces. This is much larger than the bacterial enzyme, which is 150 kDa. These characteristics, while clearly different from the bacterial enzyme, are also shared by at least some archaean and organellar enzymes that have been characterized so far.

RNAs that co-purify with the eukaryal RNase P have now been identified in S. cerevisiae, Xenopus laevis, human and several species of Schizosaccharomyces. These RNAs are not active in the absence of protein, under a variety of conditions. Furthermore, they display limited sequence similarity to the bacterial RNA or even to each other, with the exception of the Schizosaccharomyces species, which are at least 80% identical to one another, and the Xenopus RNA, which is 60% identical to the human RNA.

In the absence of enzyme activity in the RNA subunit, and without extensive sequence similarities to the bacterial RNA, what is the evidence that these RNAs are RNase P RNAs? First, these RNAs always co-purify with RNase P activity. Careful attention has been paid to demonstrating that the RNAs co-purify at every step with the enzyme activity. However, this criterion is inherently unsatisfactory because the RNAs are often degraded during purification. Second, for the yeast RNAs, genetic evidence has been gathered to show that the gene encoding the isolated RNA is essential. If one copy of the gene is knocked out, or if the gene is replaced by a rearranged construction, processing of precursor-RNAs is reduced. Finally, for human RNase P, immunological data provide good evidence that the RNAs are part of RNase P. Antibodies to the protein subunit of the E. coli enzyme will immunoprecipitate RNase P activity from human extracts. Those immunoprecipitates contain the same RNA that co-purifies with RNase P activity.

A satisfactory secondary structure for the eukaryal RNase P RNAs has not yet been developed, though several have been proposed. In general, these structures do not bear up to phylogenetic comparisons; phylogenetic co-variation of base-paired nucleotides is the best support for a secondary structure. However, additional sequence information may yet provide crucial evidence.
lar deeply divergent species, is needed before co-varying nucleotides can be identified. A positive step in developing the structure of the eukaryal RNase P RNA is the possible identification of the long-range helices A, B, and C. In addition, the 11-nt region conserved in the bacteria is present, though somewhat modified, in the eukaryal RNAs (Table I). This observation is not completely satisfying because while there are some co-varying base pairs in the eukaryal sequences, there are also nts that do not co-vary in a complementary fashion. For instance, there is a G.C basepair at E. coli equivalent nts 68 and 358 that is U.A in S. cerevisiae, but G.A in human, X. laevis and S. pombe. Nevertheless, these sequences suggest that the eu-

karyal RNA contains a version of the most diagnostic structure in the RNase P RNA: the 11-nt conserved region overlapping with the long-range helix C.

Ribonuclease MRP

Evidence is growing that another putative ribozyme, ribonuclease MRP, shares some common features with RNase P. Ribonuclease MRP (RNase MRP) participates in mitochondrial DNA replication by cleaving the RNA primer complementary to the D-loop region of the mitochondrial genome. Like RNase P it is composed of both RNA and protein subunits. In both mouse and humans, the RNA and the protein subunits are encoded in the nuclear genome. All components of the enzyme must be imported into the mitochondria for function.

Both RNase MRP and RNase P are immunoprecipitated by a group of sera from patients with autoimmune diseases. Sera from 30 different patients that immunoprecipitated the RNase MRP RNA (also called Th RNA) also immunoprecipitated the human RNase P RNA. In addition, the prototype anti-Th serum also immunoprecipitated the Xenopus RNase P RNA. This immunoprecipitation is dependent upon protein; it no longer works in phenol-extracted samples. The simplest explanation for this observation is that RNase P and RNase MRP share a common protein epitope, perhaps even a common protein subunit. Comparison of the polypeptides immunoprecipitated by the autoimmune sera from both ribonucleases should identify any common polypeptides.

Sequence analysis indicates that limited regions of the RNase MRP RNA are similar to RNase P RNA. In particular, a region similar to the bacterial conserved 11-nt segment is also present in the MRP RNA (Table I). This region is-as similar to the bacterial sequence as any of the eukaryal RNase P RNA sequences. The MRP RNA can also be folded into a structure that contains long-range elements A, B and C. A structural analysis using chemical modification has been reported for the MRP RNA. The structure does not propose the pairings described by Forster and Altman; however, the chemical modification data do not rule out their possibility.

The significance of these similarities between the RNAs of RNase MRP and RNase P is unknown. It has been suggested that the RNAs originated from a common ancestor or share a common function. However, such conclusions await further analysis of the two ribonucleases and the extent of their commonality.

Conclusions

The analysis of the sequences of RNase P RNAs from Bacteria, Archaea and Eukarya is revealing common secondary structures among RNAs which share relatively little sequence identity. In addition, a conserved 11-nt sequence motif first identified in Bacteria appears to be present in all RNase P RNAs known to date. This motif is also present in the RNase MRP RNA; however, the relationship between RNase P and RNase MRP is otherwise unknown. Clearly, there are many varieties of RNase P. It is important, now, to identify their commonalities, for it is the common features that are likely to be responsible for the mechanism of action of this interesting and ancient enzyme.

References

7 Brown, J. W. and Pace, N. R. Nucleic Acids Res. (in press)
20 Knapp, G. et al. (1986) EMBO J. 5, 1697–1703