

# New insight into RNase P RNA structure from comparative analysis of the archaeal RNA

J. KIRK HARRIS,<sup>1,2</sup> ELIZABETH S. HAAS,<sup>1</sup> DANIEL WILLIAMS,<sup>1</sup> DANIEL N. FRANK,<sup>2</sup>  
and JAMES W. BROWN<sup>1</sup>

<sup>1</sup>Department of Microbiology, North Carolina State University, Raleigh, North Carolina 27695, USA

<sup>2</sup>Department of Molecular, Cellular and Developmental Biology, University of Colorado at Boulder,  
Boulder, Colorado 80309, USA

## ABSTRACT

A detailed comparative analysis of archaeal RNase P RNA structure and a comparison of the resulting structural information with that of the bacterial RNA reveals that the archaeal RNase P RNAs are strikingly similar to those of Bacteria. The differences between the secondary structure models of archaeal and bacterial RNase P RNA have largely disappeared, and even variation in the sequence and structure of the RNAs are similar in extent and type. The structure of the cruciform (P7–11) has been reevaluated on the basis of a total of 321 bacterial and archaeal sequences, leading to a model for the structure of this region of the RNA that includes an extension to P11 that consistently organizes the cruciform and adjacent highly-conserved sequences.

**Keywords:** Archaea; archaeobacteria; ribonuclease P; ribozyme; rnpB

## INTRODUCTION

Ribonuclease P is the endonuclease responsible for the removal of leader sequences from tRNA precursors (for review, see Pace & Brown, 1995; Karwan et al., 1995–1996; Frank & Pace, 1998). RNase P is required by, and present in, all cells and organelles that carry out tRNA biosynthesis. All RNase P enzymes in which constituents have been identified are ribonucleoproteins, and the RNA subunits from all systems are recognizably similar in sequence and secondary structure (Chen & Pace, 1997). The RNA is the catalytic subunit of the enzyme, and in Bacteria and some Archaea, the RNA is capable of catalyzing the cleavage of tRNA precursors in the absence of protein (Guerrier-Takada et al., 1983; Pannucci et al., 1999). Although the remaining RNase P RNAs (nuclear, mitochondrial, and those of some Archaea) are not by themselves catalytically active, there is evidence that the RNA is nevertheless the catalyst (Thomas et al., 2000).

Most of our knowledge on RNase P comes from investigation of the catalytically active bacterial RNA subunit. Comparative analysis of the bacterial RNase P RNA has provided a well-defined secondary structure

(in two related types, A and B), and several tertiary interactions have been identified (Brown et al., 1996; Haas et al., 1996; Massire et al., 1997; Haas & Brown, 1998). Models for the global three-dimensional structure of RNase P RNA have been proposed on the basis of the secondary structure and comparative data in combination with photoaffinity crosslinking constraints (Westhof & Altman, 1994; Harris et al., 1997; Chen et al., 1998; Massire et al., 1998). Comparative data have been used to identify discrete elements of structure in the RNA, and functional alternatives of these elements; this information in turn has been used to design mutant RNAs to assess the function of these elements (Darr et al., 1992; Haas et al., 1994; Siegel et al., 1996; Haas & Brown, 1998). Regions of the RNA that contribute to recognition of the substrate cleavage site, T-loop, and CCA tail are known (Burgin & Pace, 1990; Nolan et al., 1993; Harris et al., 1994; Kirsebom & Svård, 1994; Oh & Pace, 1994; Hardt et al., 1995; Harris & Pace, 1995; Oh et al., 1998; Christian et al., 2000). The functional “core” of the RNA has been identified on the basis of evolutionary conservation and confirmed by the analysis of synthetic, minimal RNase P RNAs (Waugh et al., 1989; Siegel et al., 1996).

The other RNase P RNAs are less well understood. The eukaryotic nuclear RNase P RNAs are similar in size to the bacterial RNAs, but the sequences are recognizably related only in five small conserved regions

Reprint requests to: James W. Brown, Department of Microbiology, 4521 Gardner Hall, North Carolina State University, Raleigh, North Carolina 27695, USA; e-mail: james\_brown@ncsu.edu.

(Chen & Pace, 1997). These RNAs contain much of the phylogenetically conserved secondary structure core of their bacterial homologs, but are extremely variable in both sequence and structural detail compared to the highly conservative bacterial RNAs.

The first archaeal RNase P RNA sequences to be determined were those of *Sulfolobus acidocaldarius* and *Haloferax volcanii* (Nieuwlandt et al., 1991; LaGrandeur et al., 1993). These RNAs are quite different in sequence, and could only generally be folded into structures similar to those of the bacterial RNA. Subsequent comparative analysis of 11 archaeal RNA sequences allowed the construction of the first model of archaeal RNase P RNA secondary structure without need to rely on the bacterial model for guidance (Brown & Haas, 1995–1996; Haas et al., 1996). Remarkably, the archaeal structures were shown to be even more similar to those of Bacteria than was originally thought. However, the archaeal RNAs were distinguished by the apparent lack of P11, P13/14, P17, and P18, and a host of more subtle differences in both sequence and secondary structure. Here we report a detailed comparative analysis of archaeal RNase P RNA structure on the basis of 37 sequences from a wide range of archaeal species, including both catalytically active and inactive RNAs, and a comparison of the resulting structural information with that of the bacterial RNA. The archaeal RNase P RNAs are strikingly similar to those of Bacteria, and provide variation in structure and sequence that contribute insight into the bacterial RNA.

## RESULTS AND DISCUSSION

### Sequences

A total of 37 archaeal RNase P RNA gene sequences are available. Seven new sequences are available from complete genome sequences [*Aeropyrum pernix* (Kawarabayasi et al., 1999), *Archaeoglobus fulgidus* (Klenk et al., 1997), *Methanobacterium thermoautotrophicum* strain  $\Delta$ H (Smith et al., 1997), *Methanococcus jannaschii* (Bult et al., 1996), *Pyrococcus horikoshii* (Kawarabayasi et al., 1998), *Pyrococcus abyssi* (unpublished, Genoscope, www.genoscope.cns.fr), and *Pyrococcus furiosus* (unpublished, Utah genome Center, www.genome.utah.edu)], and 18 new sequences have been determined in this study. Two of these sequences (*M. thermoautotrophicum* strains  $\Delta$ H and Marburg) are expansions of previously determined partial sequences. Ninety-nine additional bacterial RNase P RNA sequences have also been determined.

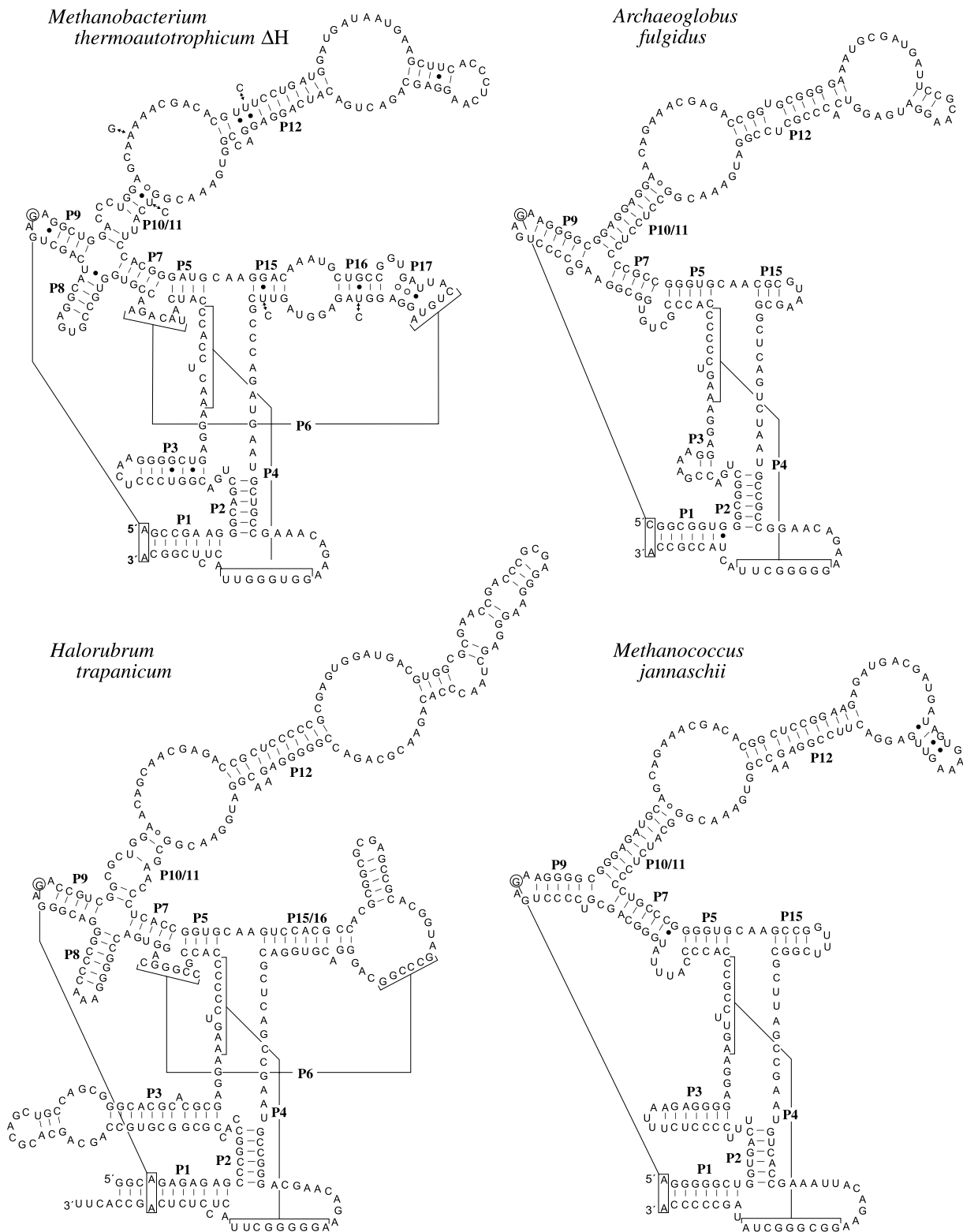
In three cases, *M. thermoautotrophicum*  $\Delta$ H, *P. furiosus*, and *Thermococcus celer* AL-1, more than one related sequence has been obtained from apparently pure cultures. It is unlikely that these different sequences represent errors introduced in the sequence

during the amplification process or contamination in the amplification reaction; the sequences were obtained repeatedly in entirely independent amplifications, are consistent with known sequence variation and the secondary structure of the RNA, and are specifically obtainable only from their respective sources. It is more likely that these differences represent allelic variation in nonclonal populations. These organisms are difficult to grow as isolated colonies on solid media, and even if the original isolate was clonal, all are routinely cultivated in liquid and passaged directly without clonal isolation; none have been grown from single colonies in recent history, if ever (see, e.g., Baross, 1995).

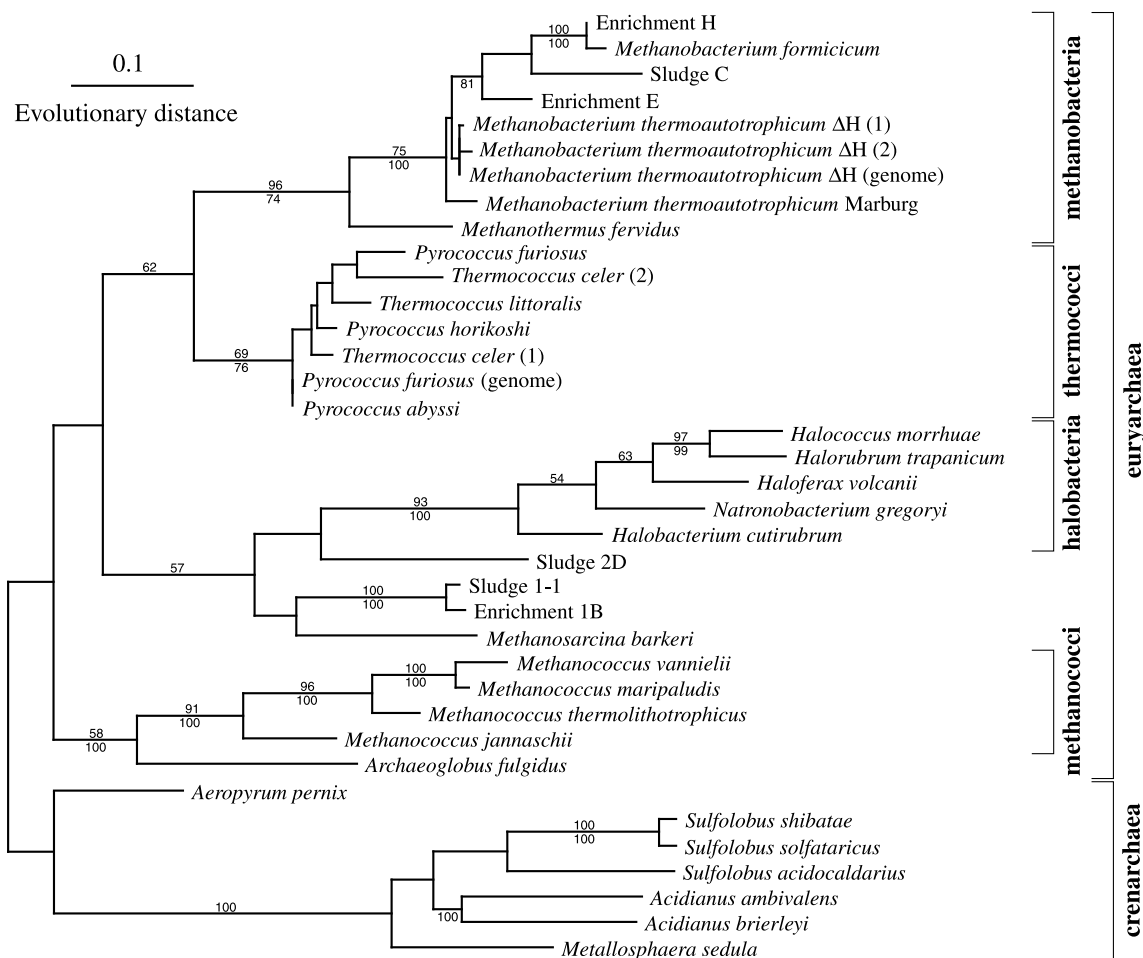
### Phylogenetic analysis

The sequences were rigorously aligned in a comparative analysis of secondary structure, providing an opportunity to compare phylogenetic relations derived from RNase P RNA sequences with those derived from small-subunit ribosomal RNA sequence from the same group of organisms. Phylogenetic trees based on the RNase P RNA alignment are in general agreement with those derived from ssu-rRNA analyses (Maidak et al., 2000) in regions of the tree that are well defined, as assessed by bootstrap analysis (Fig. 2). Each of the major groups of sequences form distinct, monophyletic branches, and the branching orders within these groups are similar to those of rRNA-based trees. The deeper branches connecting these major groups are not well-defined in the RNase P RNA tree, probably as a result of the higher rate of evolutionary change (approximately sixfold higher) in RNase P RNA relative to rRNA, and a broader range of evolutionary rates among sequences. The placement of *A. pernix* is especially volatile (this sequence frequently branches among the euryarchaea), and the placement of environmental sequences sludge 2D, sludge 1-1 and enrichment 1B with the halobacteria and *Methanosarcina barkeri* is uncertain; the environmental sequences frequently form a distinct, related clade.

Although the RNase P RNA sequences generally recreate trees similar to those based on rRNA, a significant exception is the placement of the sequence from *A. fulgidus*. In rRNA-based trees, this genus lies on a branch distinct from the other major euryarchaeal groups, separating from the other groups at approximately the bifurcation between methanobacteria and halobacteria/methanomicrobia (e.g., *Methanosarcina*; Maidak et al., 2000). The *A. fulgidus* RNase P RNA, however, is clearly related in structure (Fig. 1) and sequence (Fig. 2) to those of *Methanococcus*. Trees constructed using parsimony (DNAPARS) and maximum likelihood (DNAML) methods (Felsenstein, 1989) agree on the placement of this sequence as a relative of *Methanococcus*, and this placement is robust. In analyses using either the entire sequence alignment or only the



**FIGURE 1.** Representative archaeal RNase P RNA secondary structures. The *M. thermoautotrophicum*  $\Delta$ H and *Halorubrum trapanicum* represent the common archaeal type A RNAs, and the *M. jannaschii* and *A. fulgidus* RNAs represent type M RNAs. The *M. thermoautotrophicum*  $\Delta$ H sequence is the alternative sequence that differs from that of the genome sequence at five positions (sequence 2 in Fig. 2). The 5' and 3' ends of the *M. thermoautotrophicum*  $\Delta$ H and *M. jannaschii* RNAs have been determined (Pannucci et al., 1999); the ends of the *H. trapanicum* and *A. fulgidus* RNAs are predicted from the known locations of the ends of the *Halobacterium volcanii* (Nieuwlandt et al., 1991) and *Methanococcus jannaschii* RNA (Pannucci et al., 1999) RNAs. Helices are designated P1–P17 as previously defined for the bacterial RNA (Haas et al., 1994). P4 and P6, constituents of pseudoknots in the RNA, are connected by lines and brackets. The predicted tertiary interaction between the L9 and P1 (see text) is also shown.



**FIGURE 2.** Phylogenetic tree of archaeal RNase P RNA sequences. Sequences were aligned as described in the text, and the tree was generated using DNAML (Felsenstein, 1989). The tree is rooted between the crenarchaeal and euryarchaeal sequences. Multiple sequences obtained from a single strain are labeled parenthetically (see text). The scale bar relates estimated evolutionary distance (estimated fractional substitutions) to horizontal line length. Numbers above and below line segments are the percentage of trees containing that branch in 1,000 bootstrapped trees generated using DNAML (above) and DNAPARS (below) (Felsenstein, 1989). Bootstrap values below 50 are not shown, nor are they shown for very short branches regardless of their value.

nucleotides that comprise the phylogenetically conserved core of the RNA, 1,000 out of 1,000 bootstrap trees generated by parsimony place the *A. fulgidus* sequence as a specific relative of *Methanococcus*, without entering the phyletic range of this genus (i.e., *Methanococcus* remains monophyletic). Consensus trees generated from maximum likelihood bootstrapped core datasets also place *A. fulgidus* firmly with *Methanococcus*, in 579/1,000 trees. Placement of the *A. fulgidus* sequence was less robust (371/1,000) in the maximum likelihood trees based on all positions of the alignment, but no reasonable alternative placement occurred frequently (the next most preferred placement was with the crenarchaeote *A. pernix* in 213/1,000 trees).

The identification of conserved sequence insertions/deletions ("indels") is becoming an increasingly popular tool in molecular phylogenetic analysis (Gupta, 1998). In the present case, three major, specific deletions re-

late the *A. fulgidus* and *Methanococcus* RNase P RNAs: J5/7 (including the 5' strand of P6), P8, and P15/16/17. In addition, there are specific alterations of P7 and P10/11 shared by these RNAs (Fig. 1). The *Methanococcus* and *A. fulgidus* RNase P RNAs form a unique, derived structure class, "type M," that is compared in detail below to the type A (ancestral) RNA structure (Haas et al., 1996).

The most likely interpretations of the similarities between RNase P RNAs of *Methanococcus* and *A. fulgidus* are that either (1) the ribosomal RNA-based trees are for some reason misleading, and *A. fulgidus* is specifically related to the methanococci, or (2) that the gene encoding RNase P RNA has been transferred laterally from one group to another. Additional sequences of rRNAs and RNase P RNAs may allow the resolution of this problem. It may also be possible to distinguish between these possibilities once the protein

components of RNase P are identified in these organisms, especially if the critical missing elements of structure in the RNA are compensated for by additional polypeptide sequences (see below).

### Type A versus type M RNase P RNAs

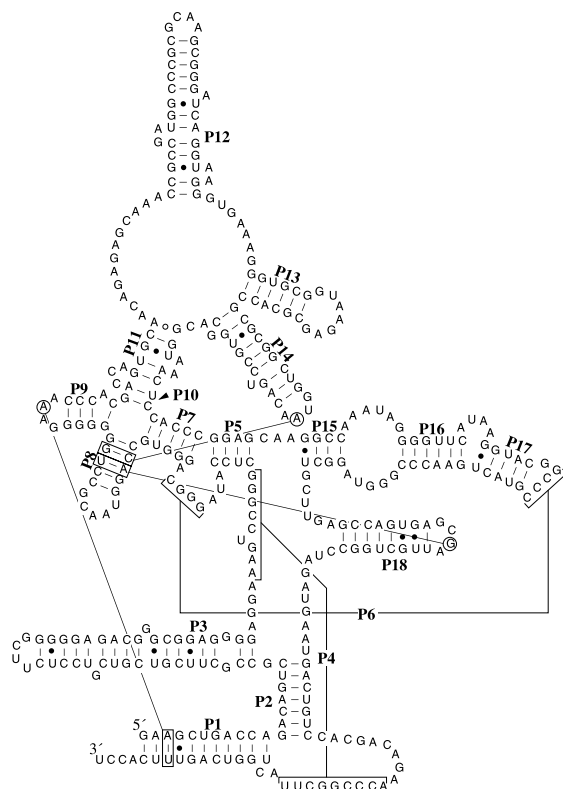
RNase P RNAs from Archaea fall into two distinct structural types, type A and type M. Type A is the common and apparently ancestral structure class, and is strikingly similar to the ancestral type A RNase P RNAs of Bacteria (Fig. 3). The archaeal type A RNAs are distinguished from those of most Bacteria primarily by the absence of P18 [also absent in green sulfur Bacteria (Haas et al., 1994)] and P13/14 [also absent in many  $\beta$ -purple Bacteria (Brown et al., 1991)], and the conserved, larger structure of P12 (see below). The type M structure class is a derived form of the secondary structure, found so far only in *Methanococcus* and *A. fulgidus* (Fig. 1). Type M RNAs lack P8 [known in Bacteria to be involved in substrate T-loop recognition (Nolan et al., 1993; Harris et al., 1994)] and everything distal to P15, including L15 [known to be involved in substrate 3'-CCA recognition (Kirsebom & Svård, 1994; Oh &

Pace, 1994; Hardt et al., 1995; Oh et al., 1998)]. The cruciform (P7–P11) of these RNAs has undergone a significant rearrangement in addition to the loss of P8, including lengthening of P9 and sometimes P7, and the lengthening and loss of extrahelical nucleotides from P10/11. The loss of L15/P16/P17 also eliminates P6 (one strand of which is in L17 in other RNase P RNAs). Bacterial RNase P RNAs that lack P16/17 (type B RNAs) retain L15, and apparently compensate for the lost structure with unique structural elements P5.1, P15.1, and P15.2. The type M archaeal RNase P RNAs do not contain additional structural elements that could compensate for the absence of the missing critical RNA structures. Type M RNase P RNAs are not capable of catalysis in the absence of protein (Pannucci et al., 1999).

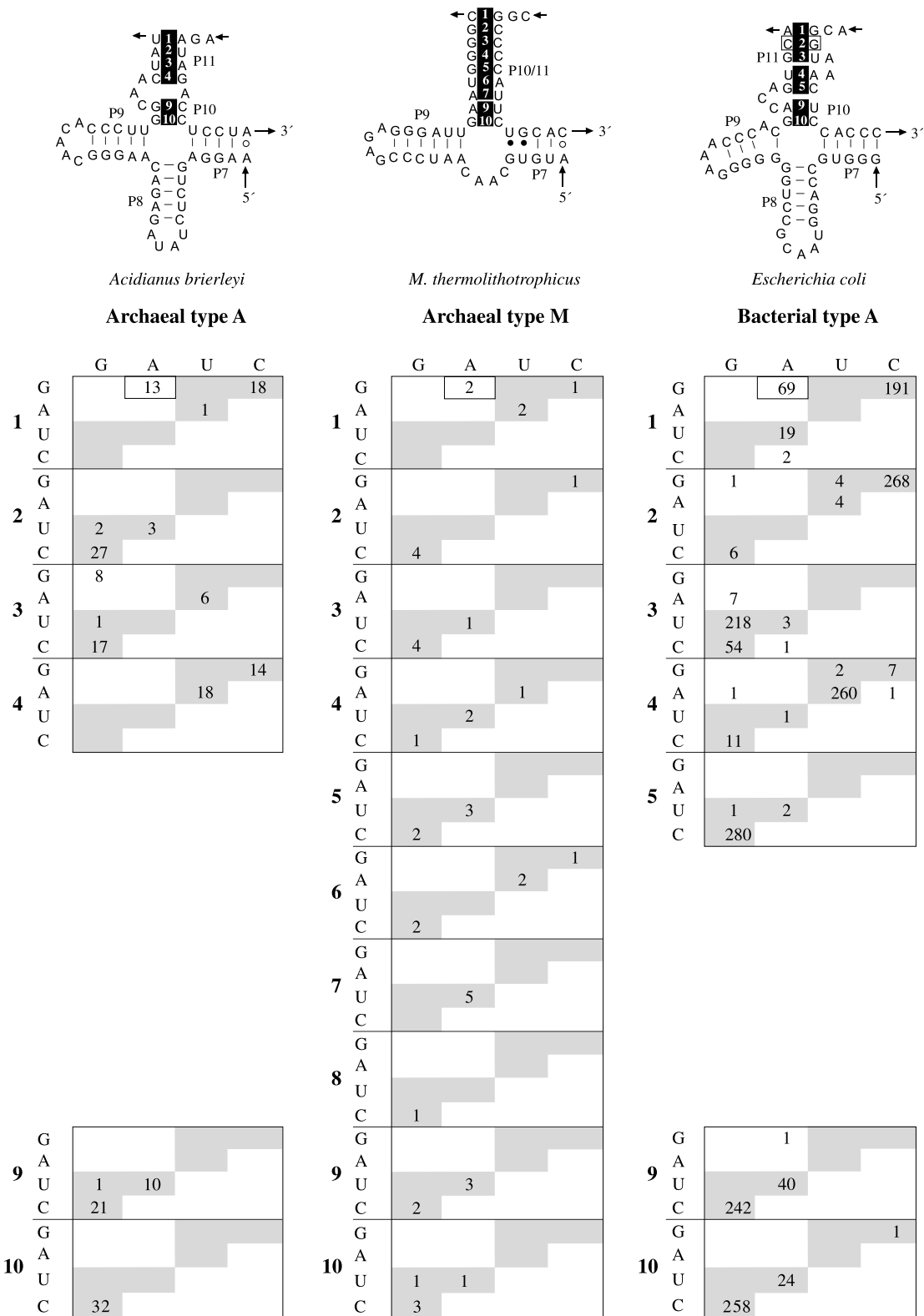
### P10/11 and the cruciform

Previous archaeal RNase P RNA secondary structures have not included P11, because convincing evidence for pairings in this region was not present in the few sequences available, and most of these sequences would not form a P11 using Watson–Crick or G•U pairs (Haas et al., 1996). The archaeal RNase P RNAs, however, do contain P11, which is now well-supported by sequence covariation (Fig. 4). This helix is 4 bp in length (usually with a G•G pair in the halophilic RNAs (e.g., *H. trapanicum* in Fig. 1) and frequently contains a terminal (distal) A•G pair (e.g., *M. thermoautotrophicum* and *A. fulgidus* in Fig. 1, *E. coli* in Figs. 3 and 4). These non-Watson–Crick alternatives covary with Watson–Crick pairings (Fig. 4). G•G pairs are common non-Watson–Crick pairs (e.g., the third base pair of P3 in the RNase P RNA of  $\gamma$ -purple Bacteria), and are thermodynamically favorable (Kierzek et al., 1999). The terminal A•G pair, where present, is reversed from the common helix:loop junction G•A configuration (Gautheret et al., 1994). In type M RNAs, P11 is an elongated, uninterrupted extension of P10. The unusual structure of P10/11 in these RNAs is part of the larger rearrangement of the cruciform, most notably involving the loss of P8, but also including lengthening of P7 and P9.

In archaeal RNase P RNAs, P11 ends directly adjacent to conserved sequences in the distal joining regions. This is consistent with previous proposals that an established base pair in the bacterial RNA, generally considered an isolated base pair, is part of a 3 bp distal extension of P11 (Tallsjo et al., 1993, Massire et al., 1998). Although this isolated base pair (C128:G230 in *E. coli*) has been confirmed genetically (Tallsjo et al., 1993), comparative support was previously weak for the extension of this base pair into an elongated P11. However, analysis of a larger alignment of bacterial sequences, including 99 new sequences determined in this study, support these additional base-pairs (Fig. 4). In bacterial RNAs, but not those of



**FIGURE 3.** The *E. coli* RNase P RNA secondary structure, with the extension of P11. Helices are designated P1–P18 as previously defined (Haas et al., 1994). P4 and P6, constituents of pseudoknots in the RNA, are connected by lines and brackets. Predicted tertiary interactions between the L9 and P1 (Massire et al., 1997), L14 and P8, and L18 and P8 (Brown et al., 1996), are also shown.



**FIGURE 4.** Cruciform structure in archaeal types A and M and bacterial type A RNAs. Secondary structures of representative archaeal type A, archaeal type M, and bacterial type A cruciforms are shown at the top. Conserved sequences flank the distal end of P11, except in those bacterial type A RNAs (such as that of *E. coli*) that contain the insertion of P13/14. Base pairs in P10/11 are numbered to correspond to the tables beneath, in which the number of RNAs that contain each possible base–base combination of that base pair are tabulated. The shaded regions of the table represent Watson–Crick and G•U pairs. The boxed numbers represent terminal A•G pairs (see text). The base pair boxed in the *E. coli* structure (bp 2) has been confirmed genetically (Tallsjo et al., 1993) and has previously been considered an isolated base pair. The extension of the base pair on either side, as shown, has been previously suggested (Tallsjo et al., 1993; Massire et al., 1998). The bacterial tables contain both types A and B sequences.

Archaea, the elongated P11 is usually interrupted by an extrahelical AA on the 3' strand. As in Archaea, the terminal base pair of this extension to P11 in Bacteria is frequently A•G. This brings the end of P10/11 to the same position in the flanking conserved sequences in both bacterial and archaeal RNase P RNAs, and consistently organizes length variation in this region of these molecules into an internal bulge and variation in the length of the extended P11.

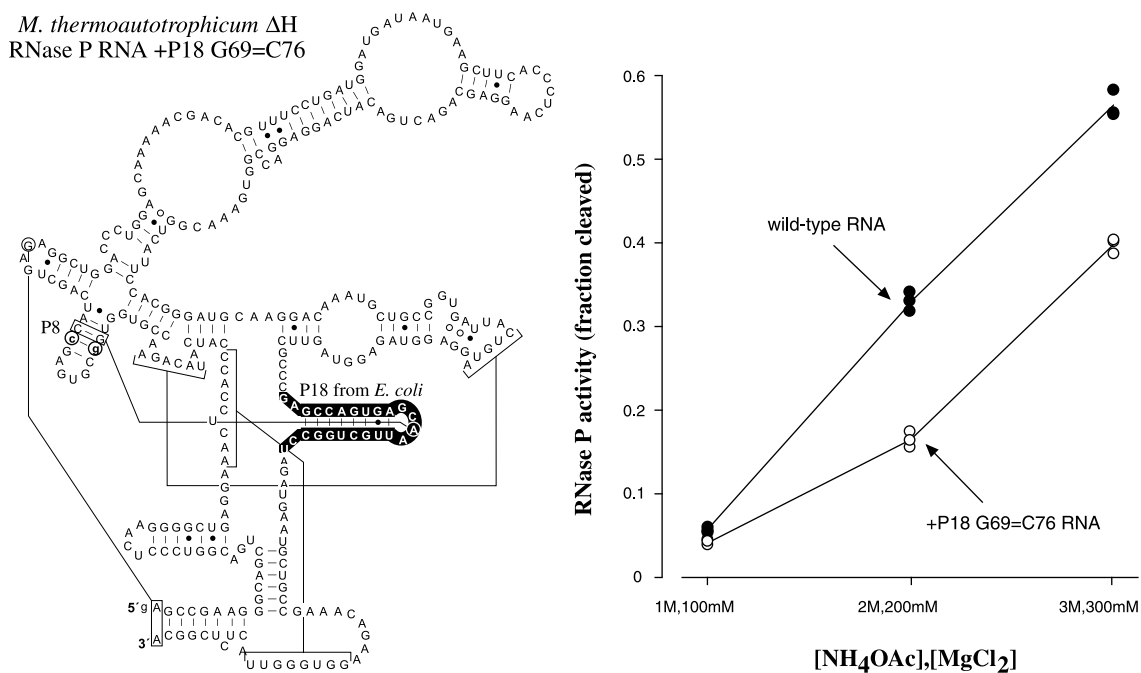
## P18

The most striking difference between archaeal and bacterial RNase P RNAs is the absence of P18 in the archaeal RNAs. In Bacteria, this helix is highly conserved in length (8 bp, flanked by an additional two bases on each side) and loop type (GNRA). This tetraloop interacts with nucleotides in P8 of the cruciform, in a now well-understood RNA:RNA tertiary motif (Brown et al., 1996; Abramovitz & Pyle, 1997). In type B RNAs (and that of *Thermomicrobium roseum*), P18 is replaced by a pair of helices and an extension of P8 that presumably compensate for the absence of the P8:L18 interaction, but *Chlorobium tepidum* and *Chlorobium limicola* are the only known bacterial RNAs that lack P18 entirely (Haas et al., 1994). It might be imagined, therefore, that the absence of stabilization of the cru-

ciform by P18 might be responsible for a large part of the catalytic deficiency of the archaeal RNAs. However, in this experiment, the insertion of P18 from *E. coli* into the *M. thermoautotrophicum* RNase P RNA, along with appropriate changes in P8 to allow the tertiary interaction between L18 and P8, did not improve either the overall activity of the archaeal RNase P RNA, nor did it relieve the activity from dependence on extremely high ionic strength and high  $Mg^{2+}$  concentration (Fig. 5).

## L9:P1 tertiary interaction

In bacterial RNase P RNAs, the GNRA loop of P9 in the cruciform forms a tertiary contact with base pairs in P1 (Harris et al., 1997; Massire et al., 1997). The loop of P9 in archaeal RNase P RNAs is very highly conserved, in this case GAGA, as is the length of the helix, consistent with a similar tertiary contact. The potential "docking" site for the L9 tetraloop in P1 is moved 1 bp proximally to a highly conserved G=C pair (G2:C291 in *M. thermoautotrophicum*) preceded by an A, appropriate for interaction with the third-position G of the L9 tetraloop (Michel & Westhof, 1991; Fig. 1). In *M. thermoautotrophicum* and *M. jannaschii*, this A residue is the first base of the RNA and is *not* base paired and internal to the helix, as is usually the case for this type



**FIGURE 5.** *M. thermoautotrophicum*  $\Delta H$  RNase P RNA + P18. The secondary structure of the *M. thermoautotrophicum*  $\Delta H$  RNase P RNA with P18 from *E. coli* inserted (highlighted in black). To maintain the possibility of the tertiary contact between L18 and P8, the third position of the L18 tetraloop was changed from the G present in the *E. coli* sequence to A (circled in white), to correspond to the third base pair (G=C) of P8 (boxed), and the next distal base pair in P8 was changed from C=G to G=C (lower-case bases circled in black). The nonnative additional G residue at the 5' end of the RNA (lower case) is required for efficient *in vitro* transcription. The graph at the right shows activity (assayed as described in the text) as a function of ammonium acetate and magnesium chloride concentration; both RNAs were at a concentration of 150  $\mu M$ .

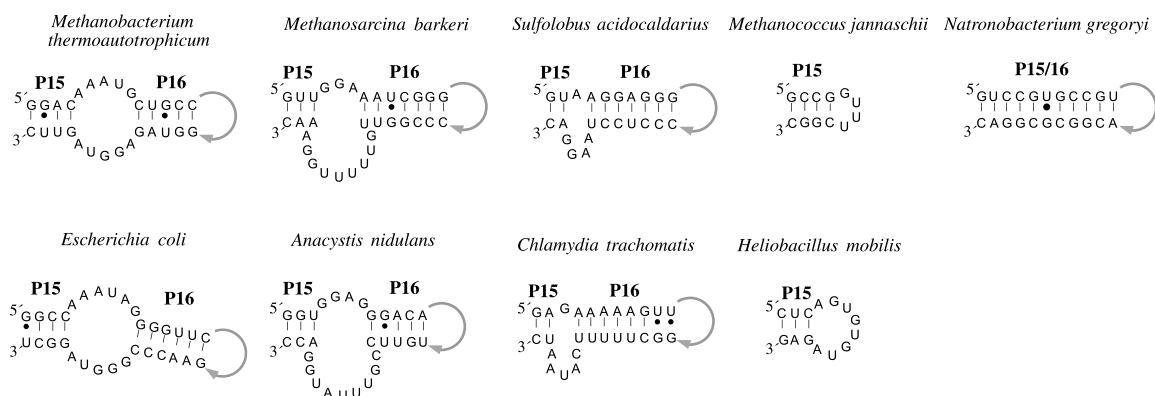
of interaction. Because the sequences upstream and downstream of P1 in these organisms maintain the potential to base pair (with sequence covariations), but are not present in the mature RNA, it seems likely that P1 is extended in an RNase P RNA precursor. The A•A mismatch in this helix may serve as a recognition site for processing as well as a docking site for interaction with L9. The RNase P RNAs of the extreme halophiles contain the same A•A mismatch, but the mature ends of the RNA have been mapped (in *H. volcanii*; Nieuwlandt et al., 1991) up- and downstream from this site, and so the mismatch is apparently retained internally in an extended P1 in these RNAs. In the Crenarchaea and thermococci, the putative docking site conforms to that found in Bacteria, with an A=U pair rather than an A•A pair terminal or internal to P1.

The conservation of the nucleotides involved in this interaction are very high, more so than for the homologous bases in the bacterial RNAs, suggesting a particularly important role for this interaction in the archaeal RNAs. P8 and P9 are thought to be stacked, and so the tertiary contact between L9 and P1 would stabilize both P9 and P8 (Harris et al., 1997; Massire et al., 1997). In the bacterial RNA, additional stabilization of P8 is attributed to similar tertiary interactions with L14 and L18 (Brown et al., 1996). P13/14 is a pair of stacked helices present in most bacterial RNase P RNAs that are rooted in conserved region III; L14 forms a tertiary contact with the proximal 2 bp of P8. P18 is a helix rooted in conserved sequences between P15 and P2; L18 forms a tertiary contact with the third and fourth base pairs of P8, directly adjacent to the tertiary contact with L14. Although P13/14 or P18 are absent in a few bacterial RNase P RNAs, no type A bacterial RNAs lack both of these elements. The archaeal RNAs lack stabilization of this region of the RNA, known in Bacteria to be involved in substrate recognition (Nolan et al., 1993; Harris et al., 1994). This is consistent with data sug-

gesting that even those archaeal RNAs with detectable activity in the absence of protein are structurally un-sound (Pannucci et al., 1999), and may explain the apparent high evolutionary pressure to maintain the remaining tertiary stabilization created by the L9:P1 contact.

## L15

Recognition of the 3'-NCCA tail of pre-tRNA by *E. coli* RNase P RNA is mediated by Watson-Crick pairing between the substrate NCC sequence (where N is usually a purine) and a highly conserved GGU sequence in the unpaired loop distal to P15. The structure of L15 (here defined as the loop of P15, whether a terminal loop as in archaeal type M RNAs and bacterial type B RNAs, or an internal loop between P15 and P16 as in other RNAs) is quite variable in archaeal RNase P RNAs. However, except in the halophiles and type M RNAs, this variation is consistent with the L15 motifs seen in the bacterial RNAs (Haas & Brown, 1998; Fig. 6). L15 structure in *Methanobacterium* matches the consensus *E. coli*-like bacterial L15 structure in every detail; L15 in *Thermococcus* and *Pyrococcus* is also similar, although consensus sequences are less well maintained. In crenarchaeal RNAs, L15 is typically a small, asymmetric loop similar to those of *Chlamydia* and relatives (Haas & Brown, 1998; Herrmann et al., 1996, 2000). L15 in *M. barkeri* is a larger asymmetric U-rich loop similar to that of many cyanobacteria (Vioque, 1992, 1997; Haas et al., 1996). Given the similarity in both sequence and structure in L15 between the archaeal and bacterial RNAs, it seems likely that this region in these archaeal RNAs participate in substrate 3'-NCCA recognition as in Bacteria. The halophilic RNAs, on the other hand, apparently lack L15 entirely; P15 and P16 form a single, uninterrupted helix. The continuity of P15 and P16 in halophilic Archaea has been used to sup-



**FIGURE 6.** Comparison of L15 structure in Archaea and Bacteria. Bacterial RNase P RNAs contain L15 that conform to one of four classes, represented beneath by *E. coli*, *Anacystis nidulans*, *Chlamydia trachomatis*, and *Heliobacillus mobilis* (Haas & Brown, 1998). L15 in archaeal RNase P RNAs also fall into these general classes, except in the cases of the extreme halophiles (e.g., *Natronobacterium gregoryi*), in which P15 and P16 are directly fused, creating a single uninterrupted helix.

port the reasonable hypothesis that the two “sides” of L15 form non-Watson–Crick pairs and remain generally A-form (Easterwood & Harvey, 1997; Harris et al., 1997; Massire et al., 1997). The length of P15/16 in halophilic Archaea, however, suggests that L15 is absent (deleted), rather than base paired (closed). Type M RNAs (those of *Methanococcus* and *Archaeoglobus*) contain an L15, but it is unlikely that these are functionally equivalent to those of other RNase P RNAs; these L15 sequences are small tetraloops or tri-loops, apparently simple closure loops for P15.

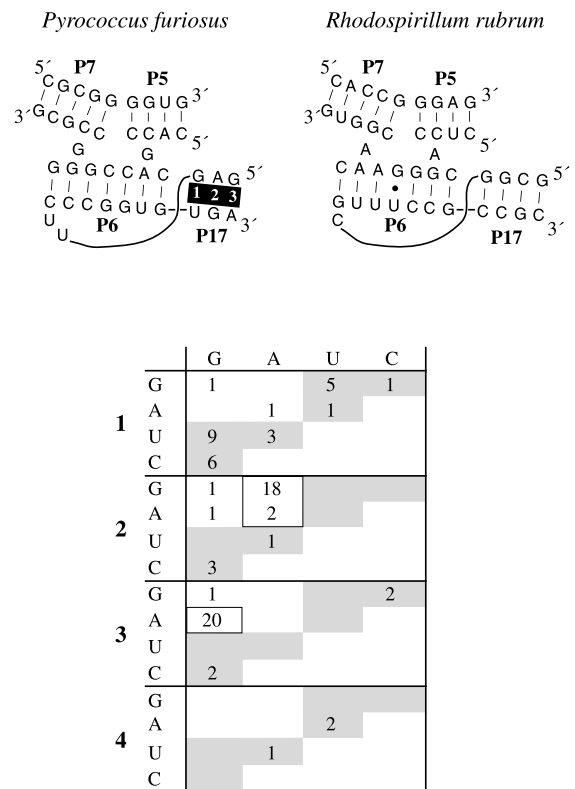
The sequences and structures of L15 in Archaea appear to be the best predictors of catalytic ability, or lack thereof, of these RNAs. The RNase P RNAs of *Methanobacterium* are the most catalytically proficient (Pannucci et al., 1999), and most closely resemble the bacterial consensus L15 sequence and structure, typified by that of *E. coli* (see Fig. 6). The RNAs of the thermococci are also active, although much less so. The remaining archaeal RNAs, which are not good matches to the standard motif/consensus L15, are not catalytically proficient in the absence of protein. The exception is the halophiles, the RNAs of which have detectable catalytic activity. All halophile RNase P RNAs contain the sequence “GGUA” immediately preceding P6 in L17; perhaps this replaces the same sequence in the consensus L15.

### P17

In Bacteria, P17 is one helix of a pseudoknot created by the pairing of J5/7 and L17 to form P6 (Haas et al., 1991). The lengths of P17 and P6 covary inversely in Bacteria, suggesting that they are stacked, which is consistent with molecular models of this region of the RNA (Haas et al., 1991). Previous models of archaeal RNase P RNA secondary structure have lacked P17; however, analysis of the sequences now available support the presence of P17 in the type A archaeal RNase P RNAs, except those of halophiles (Fig. 7). This helix was not previously identified because it is predominantly comprised of non-Watson–Crick pairs, but P17 in Bacteria also contains an unusually high frequency of non-Watson–Crick base pairs. The most common form of P17 in Archaea is a 3 bp helix formed by a G•A/A•R motif (Gautheret et al., 1994) followed by a Watson–Crick or G•U pair. Interestingly, P17 in *Methanothermus fervidus* is a single base pair. The spacing of P17 relative to P6 and P16 are highly conserved. The secondary structure in this region of the archaeal RNAs conforms to the P6/P17 motif present in  $\alpha$ -purple Bacteria (Haas & Brown, 1998).

### P12

In bacterial type A RNAs, P12 is extremely variable in sequence, sequence length, and structure distal to the



**FIGURE 7.** Comparison of P6 and P17 in Archaea and Bacteria. *P. furiosus* represents the archaeal RNAs, *Rhodospirillum rubrum* represents the bacterial RNAs. P6 and P17 are shown stacked, as proposed (Haas et al., 1991). P17 in the archaeal RNA is generally a 3 bp helix, most often with a GA:GA motif making up the proximal 2 bp (as in *P. furiosus*), with a 2 nt 5' connection between P17 and P6 (UU in *P. furiosus*). This structure class in the bacterial RNA (an alternative to that found in most bacterial RNase P RNAs) has been previously described (Haas & Brown, 1998). Base pairs in the archaeal P17 are numbered to correspond to the table below, in which the number of archaeal type A RNAs, excluding those of the halobacteria, that contain each possible base–base combination of that base pair are tabulated. The shaded regions of the table represent Watson–Crick and G•U pairs. Base pairs that conform to the frequently observed G•A/A•R motif (Gautheret et al., 1994), which covaries with Watson–Crick base pairs, are boxed.

first 5 bp [which are highly conserved in sequence and structure, and probably interact with P13 (Brown et al., 1991)]. P12 structure and sequence in Archaea is more conservative. In archaeal RNase P RNAs, P12 is 12 bp in length, contains an internal G + A-rich loop and a distal stem/loop. The RNAs of *Thermococcus* and *Pyrococcus* contain an additional internal helix and a second G + A-rich internal loop. These structures are also readily identifiable in the halophiles, despite the presence of large insertions in these internal loops. These internal loops could potentially be closed by the formation of primarily purine:purine pairs, but no specific arrangement of these pairings, or covariation among bases, is obvious. The nature of archaeal RNase P protein(s) are not known, but it seems likely that these conserved, internal loops in P12 are recognition sites for protein binding.

## CONCLUSION

The refinements to our understanding of archaeal RNase P sequence and structure described above primarily demonstrate that these RNAs are of the same "type" as bacterial RNase P RNAs. The differences between the secondary structure models of archaeal and bacterial RNase P RNA have largely disappeared, and even variation in the sequence and structure of the RNAs are similar in extent and type. The most conspicuous differences between bacterial type A and archaeal type A RNase P RNAs are the absence of P18 in the archaeal RNAs (although the addition of P18 to an archaeal RNase P RNA did not improve its activity, as might be expected) and minor differences in P10/11. It has already been shown that archaeal RNase P RNAs can be functionally reconstituted with the RNase P protein of *Bacillus subtilis* (Pannucci et al., 1999), although no protein similar to any bacterial RNase P protein is encoded in any available archaeal genome. In contrast, the type M RNase P RNAs of *Methanococcus* and *Archaeoglobus* are unique, lacking structural elements that are thought to be essential for substrate recognition. How the absence of these structures is overcome in the holoenzyme is not known.

## MATERIALS AND METHODS

### Cultures and preparation of DNA

Cultures of *Acidianus (Desulfurolobus) ambivalens* (DSM 3772), *Acidianus brierleyi* (DSM 1651), *Sulfolobus shibatae* (DSM 5389), and *Sulfolobus solfataricus* (DSM 1616) were obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ) and cultivated in the suggested media and conditions. Cell pastes of *Pyrococcus furiosus* (DSM 3638) were gifts from Chae Han and Robert Kelly (North Carolina State University). DNA from "*Thermococcus celer*" AL-1 and *Thermococcus littoralis* (ATCC 55233) were gifts from Anna-Louis Reysenbach (University of Portland). Cultures of *Methanococcus maripaludis* (DSM 2067), *M. thermolithotrophicus* (DSM 2095), and *Methanococcus vanniellii* (DSM 1224) were a gift from Dr. William Whitman at the University of Georgia, and were cultivated in ATCC media 1439 under 40 psi of 80% H<sub>2</sub>:20% CO<sub>2</sub> at 37 °C, 65 °C, and 37 °C, respectively. A culture of *M. thermoautotrophicum* strain ΔH and genomic DNA from strain Marburg (DSM 2133) were gifts from John Reeve (The Ohio State University). Strain ΔH was cultivated in ER media (Hook et al., 1984) at 65 °C under 40 psi of 80% H<sub>2</sub>:20% CO<sub>2</sub>. Frozen cells were lysed by grinding in dry ice in a mortar and pestle and resuspended in lysis buffer (Brown et al., 1991), except *M. vanniellii* and *M. maripaludis*, which were resuspended in lysis buffer and lysed by the addition of SDS to 0.5%. DNA was extracted as previously described (Brown et al., 1991).

Additional archaeal RNase P RNA gene sequences were obtained from DNA isolated from digested wastewater sludge (Cary South Municipal Wastewater Treatment Plant; "SL" sequences) and methanogen enrichment cultures (ER media

inoculated with this sludge, under 80% H<sub>2</sub>:20% CO<sub>2</sub>, incubated at 30–80 °C and pooled prior to DNA isolation; "E" sequences). DNA was isolated as described for cultivated species, followed by CsCl:ethidium bromide density gradient centrifugation (Sambrook et al., 1989).

A culture of *Haloanaerobium praevalens* (DSM 2228) was obtained from the DSMZ; DNA was extracted directly from this culture without further cultivation using the microwave lysis method (Reysenbach et al., 1994). Cell material of *Leptospirillum ferrooxidans* strains MK and CF12 were obtained from Barrie Johnson (University of Wales); DNA was extracted using the microwave method. Cultures of *Erysipelothrix rhusiopathiae* (ATCC 19414), *Plesiomonas shigelloides* (ATCC 51903), *Serratia liquefaciens* (ATCC 27592), *Pasteurella aerogenes* (ATCC 27883), *Providencia alcalifaciens* (ATCC 51902), *Leclercia adocarboxylata* (ATCC 23216), *Proteus vulgaris* (ATCC 881), *Streptococcus bovis* (ATCC 9809), and *Streptococcus equi* (ATCC 9528) were from Chrisope Technologies; *Streptococcus* spp. were cultivated on blood agar plates, the remainder in LB broth, and DNA was isolated using the Tissue Extraction kit (ClonTech) as recommended. DNA from *Fusobacterium mortiferum* (ATCC 9817) and *Thermodesulfobacterium commune* (ATCC 33708) was obtained from Brice Felden (University of Utah); DNA from *Verrucomicrobium spinosum* was obtained from Fred Rainey (DSMZ).

Additional bacterial RNase P RNA encoding sequences were obtained from DNA isolated from a salt marsh adjacent to the Berkeley Marina ("SM" sequences), obtained from Scott Dawson (University of California at Berkeley) and yard waste compost ("CP" sequences); DNA was isolated as previously described (Barnes et al., 1994). Sequences were also obtained from uncharacterized isolates, grown on LB plates, from a North Carolina State University Veterinary School horse barn drain ("VSDW") and drainage fluid ("EF"), and from an uncharacterized anaerobic purple photosynthetic isolate from West Lake, North Carolina ("purpleX"); DNA was isolated using the Tissue Extraction kit (ClonTech) as recommended.

### PCR amplification and sequencing

Polymerase chain reaction amplification was performed in 50 mL reactions containing 50 mM KCl, 10 mM Tris-Cl, pH 8.3, 1.5 mM each dGTP, dCTP, dATP, and dTTP, 200 ng each primer, and 30–60 ng of genomic DNA. Primers used to amplify the RNase P RNA gene from *A. ambivalens*, and *A. brierleyi* were A59FII*Xba* (GCTCTAGAGGAAASTCYMCCC) and A347RII*Bam* (CGGGATCCGTAASCYMCCTTCTGT). The *S. shibatae* and *S. solfataricus* RNase P RNA gene was amplified using SAC59F*Xba* (GCTCTAGAGGAAASTCCAGCC), and SAC347R*Bam* (CGGGATCCCTAASCAGCTTYTGT). The *Methanococcus* spp. RNase P RNA genes were amplified using Mja5'*Xba* (GCTCTAGAGGGTAAGGGGGCTGGTG) and Mja3'*Bam* (CGGGATCCGGTATGGGGGGCTATAGC). *Thermococcus* and *Pyrococcus* spp. RNase P RNA genes were amplified using Pyro5'*Xba* (GCTCTAGATAGGCGAGGGGGCTGGG) and Pyro3'*Bam* (CGGGATCCCTAGGCGACCCCCGTATAG). The *M. thermoautotrophicum* RNase P RNA genes were obtained using ΔH5'*Bam* (CGGGATCCACCGG GCAAGCCGAAGGGC) and ΔH3'*Xba* (GCTCTAGACCGGG CATGCCGAGAG). RNase P RNA gene sequences obtained from DNA from municipal wastewater sludge and methanogen enrichment cultures used primers ΔH5'*Bam* and ΔH3'*Xba*

(Enrichment E and Enrichment H) or A59FXba (GCTCTAG AGGAAAGTCCMSCC) and A347RBam (CGGGATCCTAAG CCMSCCTTYTGT) (Wastewater SL2D, Wastewater SL-C, Enrichment 1B and Sludge 1-1). The amplifications were performed with an initial incubation at 94 °C for 2 min, followed by 30 amplification cycles (92 °C for 1.5 min, 50 °C for 1.5 min, 72 °C for 0.5 min each cycle), and a final incubation of 72 °C for 7 min. The resulting PCR products were cloned and their nucleotide sequences determined. Bacterial sequences from both pure cultures or environmental DNAs were amplified using primers 59FBam with 347REco or 347RXba (Brown et al., 1991; Haas et al., 1996). The accession number for these sequences are AF192350–AF192365, AF295979–AF296074, AF084930, AF084931, AF160198, and AF160199.

### RNase P RNA genes in genome sequences

RNase P RNA-encoding sequences from *A. pernix* (Kawarabayasi et al., 1999), *A. fulgidus* (Klenk et al., 1997), *M. thermoautotrophicum* strain ΔH (Smith et al., 1997), *M. jannaschii* (Bult et al., 1996), *P. abyssi* (unpublished, Genoscope, www.genoscope.cns.fr), *P. furiosus* (unpublished, Utah genome Center, www.genome.utah.edu) and *P. horikoshii* (Kawarabayasi et al., 1998) were identified in their complete genome sequences using previously identified archaeal RNase P sequences to query GenBank using the NCBI BLAST server (Altschul et al., 1990). RNase P RNA encoding sequences were also identified in the 22 published bacterial genome sequences by the same method using the RNase P RNA sequence from their closest relative (predicted from ribosomal RNA-based phylogenetic trees). Potential RNase P RNA-encoding sequences were assessed by their potential to adopt the basic core of RNase P RNA secondary structure and correct placement of conservative residues.

### Alignment, phylogenetic, and comparative analysis

Sequences were aligned with those previously available using SeqApp (Don Gilbert, Indiana University), initially on the basis of previously identified secondary structure and sequence conservation (Haas et al., 1996). Alignments were refined as part of the comparative analysis for structural determination, primarily manually but also on the basis of mutual information using the software package Covariation (Brown, 1991), as previously described (Brown et al., 1996). Phylogenetic trees were generated using Phylip (Felsenstein, 1989). All sequences and secondary structures and alignments are available at the RNase P Database (<http://www.mbio.ncsu.edu/RNaseP>; Brown, 1999).

### *M. thermoautotrophicum* ΔH RNA +P18

Sequences corresponding to P18 were inserted into the plasmid-born RNase P RNA encoding gene from *M. thermoautotrophicum* strain ΔH by in vitro mutagenesis (Kunkel, 1985; Sambrook et al., 1989). The plasmid, based on pUC119, contains the RNase P encoding gene linked to a bacteriophage T7 promoter such that the in vitro transcription product

of BsaI-linearized plasmid contains a single extra 5' G, relative to the native molecule (required for efficient in vitro transcription) and the native 3' end (Pannucci et al., 1999). The P18 sequence, derived from that of *E. coli*, was inserted by in vitro mutagenesis using an oligonucleotide with the sequence CGGCAGCATTCATCTAGGCCAGCAATtGCTCACTGGCTCGGGCGGACTACCTC (the insertion is underlined). To maintain the possibility of the tertiary contact between L18 and P8 (Brown et al., 1996), the third position of the L18 tetraloop was changed from the G present in the *E. coli* sequence to A (lowercase t above), to correspond to the third base pair (G=C), and the next distal base pair in P8 was changed from C=G to G=C by in vitro mutagenesis using an oligonucleotide with the sequence CTCAGCTGATGgCTCA CGcCACCACGG (mutagenic nucleotides are lowercase). *M. thermoautotrophicum* ΔH RNase P RNAs, wild-type and +P18, were synthesized from these plasmids using T7 RNA polymerase, as described by the manufacturer (Promega).

RNase P RNA assays contained 1.5 nM <sup>32</sup>P-labeled *B. subtilis* pre-tRNA<sup>Asp</sup>, 50 mM Tris-pH 8, 100–300 mM MgCl<sub>2</sub>, 1–3 M ammonium acetate, 0.1% SDS, and 0.05% nonident P-40. Incubations were for 3 h at 50 °C. Reaction products were separated by electrophoresis on 8 or 12% polyacrylamide:7.5 M urea:TBE gels (SequaGel, National Diagnostics) and visualized by phosphorimager.

### ACKNOWLEDGMENTS

We thank all those who provided cultures and/or DNA used in this study: Scott Dawson, Brice Feldon, Chae Han, Barrie Johnson, Robert Kelly, John Reeve, Fred Rainey, Anna-Louis Reysenbach, and William Whitman. We also thank those who provided technical assistance by cloning RNase P RNA sequences during summer internships: Erika Richard, Alicia Lawson, Jessica Eaton, Julie Bright, Daniel West, Samar Seth, Jennifer Coughlan, and Ken MacGillivray. We also thank Bev Vucson, Julie Perez, and Jim Pannucci for cloning and determining RNase P RNA sequences. This work was supported by National Institutes of Health Grant GM52894.

Received September 1, 2000; returned for revision October 2, 2000; revised manuscript received October 30, 2000

### REFERENCES

- Abramovitz DL, Pyle AM. 1997. Remarkable morphological variability of a common RNA folding motif: The GNRA tetraloop-receptor interaction. *J Mol Biol* 266:493–506.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410.
- Barns SM, Fundyga RE, Jeffries MW, Pace NR. 1994. Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment. *Proc Natl Acad Sci USA* 91:1609–1613.
- Baross JA. 1995. Isolation, growth and maintenance of hyperthermophiles. In: Robb FT, Place AR, Sowers KR, Schreier HJ, Das-Sarma S, Fleischmann EM, eds. *Archaea: A laboratory manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 15–23.
- Brown JW. 1991. Phylogenetic comparative analysis of RNA structure on Macintosh computers. *Comput Appl Biosci* 7:391–393.
- Brown JW. 1999. The Ribonuclease P Database. *Nucleic Acids Res* 27:314

- Brown JW, Haas ES. 1995–1996. Ribonuclease P structure and function in Archaea. *Mol Biol Rep* 22:131–134.
- Brown JW, Haas ES, James BD, Hunt DA, Liu JS, Pace NR. 1991. Phylogenetic analysis and evolution of RNase P RNA in proteobacteria. *J Bacteriol* 173:3855–3863.
- Brown JW, Nolan JM, Haas ES, Rubio MA, Major F, Pace NR. 1996. Comparative analysis of ribonuclease P RNA using gene sequences from natural microbial populations reveals tertiary structural elements. *Proc Natl Acad Sci USA* 93:3001–3006.
- Bult CJ, White O, Olsen GJ, Zhou L, Fleischmann RD, Sutton GG, Blake JA, FitzGerald LM, Clayton RA, Gocayne JD, Kerlavage AR, Dougherty BA, Tomb JF, Adams MD, Reich CI, Overbeek R, Kirkness EF, Weinstock KG, Merrick JM, Glodek A, Scott JL, Geoghagen NSM, Venter JC. 1996. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* 273:1058–1073.
- Burgin AB, Pace NR. 1990. Mapping the active site of ribonuclease P RNA using a substrate containing a photoaffinity agent. *EMBO J* 9:4111–4118.
- Chen J-L, Nolan JM, Harris ME, Pace NR. 1998. Comparative photocross-linking analysis of the tertiary structures of *Escherichia coli* and *Bacillus subtilis* RNase P RNAs. *EMBO J* 17:1515–1525.
- Chen J-L, Pace NR. 1997. Identification of the universally conserved core of ribonuclease P RNA. *RNA* 3:557–560.
- Christian EL, Kaye NM, Harris ME. 2000. Helix P4 is a divalent metal ion binding site in the conserved core of the ribonuclease P ribozyme. *RNA* 6:511–519.
- Darr SC, Zito K, Smith D, Pace NR. 1992. Contributions of phylogenetically variable structural elements to the function of the ribozyme ribonuclease P. *Biochemistry* 31:328–333.
- Easterwood TR, Harvey SC. 1997. Ribonuclease P RNA: Models of the 15/16 bulge from *Escherichia coli* and the P15 stem loop of *Bacillus subtilis*. *RNA* 3:577–585.
- Felsenstein J. 1989. PHYLIP—Phylogeny Inference Package (Version 3.2). *Cladistics* 5:164–166.
- Frank DN, Pace NR. 1998. Ribonuclease P: Unity and diversity in a tRNA processing ribozyme. *Annu Rev Biochem* 67:153–180.
- Gautheret D, Konings D, Gutell RR. 1994. A major family of motifs involving G•A mismatches in ribosomal RNA. *J Mol Biol* 242:1–8.
- Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S. 1983. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* 35:849–857.
- Gupta RS. 1998. Protein phylogenies and signature sequences: A reappraisal of evolutionary relationships among archaeobacteria, eubacteria, and eukaryotes. *Microbiol Mol Biol Rev* 62:1435–1491.
- Haas ES, Armbruster DW, Vucson BM, Daniels CJ, Brown JW. 1996. Comparative analysis of ribonuclease P RNA structure in Archaea. *Nucleic Acids Res* 24:1252–1259.
- Haas ES, Brown JW. 1998. Evolutionary variation in bacterial RNase P RNAs. *Nucleic Acids Res* 26:4093–4099.
- Haas ES, Brown JW, Pitulle C, Pace NR. 1994. Further perspective on the catalytic core and secondary structure of ribonuclease P RNA. *Proc Natl Acad Sci USA* 91:2527–2531.
- Haas ES, Morse DP, Brown JW, Schmidt FJ, Pace NR. 1991. Long-range structure in ribonuclease P RNA. *Science* 254:853–856.
- Hardt WD, Schlegl J, Erdmann VA, Hartmann RK. 1995. Kinetics and thermodynamics of the RNase P RNA cleavage reaction: Analysis of tRNA 3'-end variants. *J Mol Biol* 247:161–172.
- Harris ME, Kazantsev AV, Chen J-L, Pace NR. 1997. Analysis of the tertiary structure of the ribonuclease P ribozyme-substrate complex by site-specific photoaffinity crosslinking. *RNA* 3:561–576.
- Harris ME, Nolan JM, Malhotra A, Brown JW, Harvey SC, Pace NR. 1994. Use of photoaffinity crosslinking and molecular modeling to analyze the global architecture of ribonuclease P RNA. *EMBO J* 13:3953–3963.
- Harris ME, Pace NR. 1995. Identification of phosphates involved in catalysis by the ribozyme RNase P RNA. *RNA* 1:210–218.
- Herrmann B, Pettersson B, Everett KD, Mikkelsen NE, Kirsebom LA. 2000. Characterization of the *mnpB* gene and RNase P RNA in the order *Chlamydiales*. *Int J Syst Evol Microbiol* 50:149–158.
- Herrmann B, Winqvist O, Mattsson JG, Kirsebom LA. 1996. Differentiation of *Chlamydia* spp. by sequence determination and restriction endonuclease cleavage of RNase P RNA genes. *J Clin Microbiol* 34:1897–1902.
- Hook LA, Corder RE, Hamilton PT, Frea JI, Reeve JN. 1984. Development of a plating system for genetic exchange studies in methanogens using a modified ultra-low oxygen chamber. In: Strohl WR, Tuovinen OH, eds. *Microbial chemoautotrophy*. Columbus, OH: The Ohio State University Press.
- Karwan R, Pluk H, van Venrooij WJ, eds. 1995–1996. *Special Issue: RNase MRP/RNase P Systems*. *Mol Biol Rep*. 22, Nos. 2/3.
- Kawarabayasi Y, Hino Y, Horikawa H, Yamazaki S, Haikawa Y, Jin-no K, Takahashi M, Sekine M, Baba S, Ankai A, Kosugi H, Hosoyama A, Fukui S, Nagai Y, Nishijima K, Nakazawa H, Takamiya M, Masuda S, Funahashi T, Tanaka T, Kudoh Y, Yamazaki J, Kushida N, Oguchi A, Aoki K, Kubota K, Nakamura Y, Nomura N, Sako Y, Kikuchi H. 1999. Complete genome sequence of an aerobic hyperthermophilic crenarchaeon, *Aeropyrum pernix* K1. *DNA Res* 6:83–101, 145–152.
- Kawarabayasi Y, Sawada M, Horikawa H, Haikawa Y, Hino Y, Yamamoto S, Sekine M, Baba S, Kosugi H, Hosoyama A, Nagai Y, Sakai M, Ogura K, Otsuka R, Nakazawa H, Takamiya M, Ohfuku Y, Funahashi T, Tanaka T, Kudoh Y, Yamazaki J, Kushida N, Oguchi A, Aoki K, Kikuchi H. 1998. Complete sequence and gene organization of the genome of a hyper-thermophilic archaeobacterium, *Pyrococcus horikoshii* OT3. *DNA Res* 5:55–76.
- Kierzek R, Burkard ME, Turner DH. 1999. Thermodynamics of single mismatches in RNA duplexes. *Biochemistry* 38:14214–14223.
- Kirsebom LA, Svärd SG. 1994. Base pairing between *Escherichia coli* RNase P RNA and its substrate. *EMBO J* 13:4870–4876.
- Klenk HP, Clayton RA, Tomb JF, White O, Nelson KE, Ketchum KA, Dodson RJ, Gwinn M, Hickey EK, Peterson JD, Richardson DL, Kerlavage AR, Graham DE, Kyrpides NC, Fleischmann RD, Quackenbush J, Lee NH, Sutton GG, Gill S, Kirkness EF, Dougherty BA, McKenney K, Adams MD, Loftus B, Peterson S, Reich CI, McNeil LK, Badger JH, Glodek A, Zhou L, Overbeek R, Gocayne JD, Weidman JF, McDonald L, Utterback T, Cotton MD, Spriggs T, Artiach P, Kaine BP, Sykes SM, Sadow PW, D'Andrea K, Bowman C, Fujii C, Garland SA, Mason TM, Olsen GJ, Fraser CM, Smith HO, Woese CR, Venter JC. 1997. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature* 390:364–370.
- Kunkel TA. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc Natl Acad Sci USA* 82:488–492.
- LaGrandeur TE, Darr SC, Haas ES, Pace NR. 1993. Characterization of the RNase P RNA of *Sulfolobus acidocaldarius*. *J Bacteriol* 175:5043–5048.
- Maidak BL, Cole JR, Lilburn TG, Parker CT, Saxman PR, Stredwick JM, Garrity GM, Li B, Olsen GJ, Pramanik S, Schmidt TM, Tiedje JM. 2000. The RDP (Ribosomal Database Project) continues. *Nucleic Acids Res* 28:173–174.
- Massire C, Jaeger L, Westhof E. 1997. Phylogenetic evidence for a new tertiary interaction in bacterial RNase P RNAs. *RNA* 3:553–556.
- Massire C, Jaeger L, Westhof E. 1998. Derivation of the three-dimensional architecture of bacterial ribonuclease P RNAs from comparative sequence analysis. *J Mol Biol* 279:773–793.
- Michel F, Westhof E. 1991. Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J Mol Biol* 216:585–610.
- Nieuwlandt DT, Haas ES, Daniels CJ. 1991. The RNA component of RNase P from the archaeobacterium *Haloferax volcanii*. *J Biol Chem* 266:5689–5695.
- Nolan JM, Burke DH, Pace NR. 1993. Circularly permuted tRNAs as specific photoaffinity probes of ribonuclease P RNA structure. *Science* 261:762–765.
- Oh B-K, Frank DN, Pace NR. 1998. Participation of the 3'-CCA of tRNA in the binding of catalytic Mg<sup>2+</sup> ions by ribonuclease P. *Biochemistry* 37:7277–7283.
- Oh B-K, Pace NR. 1994. Interaction of the 3'-end of tRNA with ribonuclease P RNA. *Nucleic Acids Res* 22:4087–4094.
- Pace NR, Brown JW. 1995. Evolutionary perspective on the structure and function of ribonuclease P, a ribozyme. *J Bacteriol* 177:1919–1928.
- Pannucci JA, Haas ES, Hall TA, Harris JK, Brown JW. 1999. RNase P RNAs from some Archaea are catalytically active. *Proc Natl Acad Sci USA* 96:7803–7808.
- Reysenbach AL, Wickham GS, Pace NR. 1994. Phylogenetic analysis of the hyperthermophilic pink filament community in Octopus

- Spring, Yellowstone National Park. *Appl Environ Microbiol* 60: 2113–2119.
- Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: A laboratory manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
- Siegel RW, Banta AB, Haas ES, Brown JW, Pace NR. 1996. *Mycoplasma fermentans* simplifies our view of the catalytic core of ribonuclease P RNA. *RNA* 2:452–462.
- Smith DR, Doucette-Stamm LA, Deloughery C, Lee H, Dubois J, Aldredge T, Bashirzadeh R, Blakely D, Cook R, Gilbert K, Harrison D, Hoang L, Keagle P, Lumm W, Pothier B, Qiu D, Spadafora R, Vicaire R, Wang Y, Wierzbowski J, Gibson R, Jiwani N, Caruso A, Bush D, Safer H, Patwell D, Prabhakar S, McDougall S, Shimer G, Goyal A, Pietrovski S, Church GM, Daniels CJ, Mao JI, Rice P, Nölling J, Reeve JN. 1997. Complete genome sequence of *Methanobacterium thermoautotrophicum*  $\Delta$ H: Functional analysis and comparative genomics. *J Bacteriol* 179:7135–7155.
- Tallsjö A, Svärd SG, Küfel J, Kirsebom LA. 1993. A novel tertiary interaction in M1 RNA, the catalytic subunit of *Escherichia coli* RNase P. *Nucleic Acids Res* 21:3927–3933.
- Thomas BC, Chamberlain J, Engelke DR, Gegenheimer P. 2000. Evidence for an RNA-based catalytic mechanism in eukaryotic nuclear ribonuclease P. *RNA* 6:554–562.
- Vioque A. 1992. Analysis of the gene encoding the RNA subunit of ribonuclease P from cyanobacteria. *Nucleic Acids Res* 20:6331–6337.
- Vioque A. 1997. The RNase P RNA from cyanobacteria: Short tandemly repeated repetitive (STRR) sequences are present within the RNase P RNA gene in heterocyst-forming cyanobacteria. *Nucleic Acids Res* 25:3471–3477.
- Waugh DS, Green CJ, Pace NR. 1989. The design and catalytic properties of a simplified ribonuclease P RNA. *Science* 244:1569–1571.
- Westhof E, Altman S. 1994. Three-dimensional working model of M1 RNA, the catalytic RNA subunit of ribonuclease P from *Escherichia coli*. *Proc Natl Acad Sci USA* 91:5133–5137.