J Mol Evol (1990) 31:25–32

Journal of Molecular Evolution

Secondary Structural Elements Exclusive to the Sequences Flanking Ribosomal RNAs Lend Support to the Monophyletic Nature of the Archaebacteria

Jørgen Kjems[1] and Roger A. Garrett[2]

[1] Biostructural Chemistry, Kemisk Institut, Aarhus University, DK-8000 Aarhus C, Denmark
[2] Institut for Biologisk Kemi B, Sølvgade 83, Copenhagen University, 1307 Copenhagen K, Denmark

Summary. Several sequences flanking the large rRNA genes of several transcripts from extreme thermophiles, extreme halophiles, and methanogens were aligned and analyzed for the presence of common primary and secondary structural features, which would bear on the concept of monophyletic archaebacteria. Few sequences were common to all the archaebacterial transcripts, and these were confined to short regions generally flanking putative double helices. At a secondary structural level, however, in addition to the previously characterized processing stems of the 16S and 23S RNAs, four helices were detected that were common to the archaebacterial transcripts: two in the 16S RNA leader sequence and two in the 16S–23S RNA spacer. Although all of these helices vary in size and form from organism to organism, three of them contain double helical segments that are strongly supported by compensating base changes among the three archaebacterial groups. Three extreme halophiles exhibited two additional helices in their relatively large spacers and a further helix preceding the 5S RNA, which are also supported by compensating base changes. Ribosomal RNA transcripts from eubacteria/chloroplasts and eukaryotes were also examined for secondary structural features with locations and forms corresponding to those of the archaebacteria, but none were detected. The analysis provides support for the monophyletic nature of the archaebacteria and reinforces their differences from eubacteria/chloroplasts and eukaryotes.

Offprint requests to: R.A. Garrett

Key words: Archaebacteria — rRNA operons — Secondary structure — Evolution

Introduction

The large rRNA sequences have been selected as evolutionary chronometers on account of their universality, slow rate of mutation, and conserved function. Moreover, phylogenetic trees derived from the aligned 16S RNA-like sequences provided the primary evidence for the monophyletic nature of archaebacteria, constituting the extreme thermophiles, the extreme halophiles, and the methanogens (Fox et al. 1980; Woese and Olsen 1986), and this evidence was reinforced by analyses of aligned 23S RNA-like sequences (Leffers et al. 1987). Nevertheless, this concept has remained contentious owing to the biochemical and genetic diversity among archaebacteria (Lake et al. 1986), including major differences in the organization of their rRNA operons (Lake 1988, 1989). Thus, the extreme halophile and methanogen rRNA operons exhibit eubacterial-like size and organization, in which the $tRNA^{Ile}$ + $tRNA^{Ala}$, or $tRNA^{Glu}$, in the 16S–23S RNA spacer is replaced by $tRNA^{Ala}$, whereas the extreme thermophiles have short spacers with no tRNA, a physically and transcriptionally uncoupled 5S RNA gene, and introns within the 23S RNA gene (Neumann et al. 1983; Larsen et al. 1986; Achenbach-Richter and Woese 1988).

The leader and spacer sequences are much less conserved in size and sequence than the rRNAs and

are not as amenable to alignment and secondary structure prediction as are the rRNAs. Therefore, as a basis for this study, we posed the question: Is there any evidence for structure, common and exclusive to the flanking sequences of the archaebacterial rRNAs, which would support the concept of monophyletic archaebacteria? To answer this, we undertook a comparative analysis of the sequence and secondary structure of these flanking regions for a range of archaebacteria and other organisms.

Materials and Methods

Sequence Analyses, Alignment, and Prediction of Secondary Structure. Whole or partial rRNA operon sequences from the following archaebacteria were examined: *Desulfurococcus mobilis* (Kjems and Garrett 1987), *Sulfolobus* B12 (Reiter et al. 1987), and *Sulfolobus solfataricus* (Olsen et al. 1985) (both partial), *Thermoproteus tenax* (Kjems et al. 1987; Wich et al. 1987), *Thermofilum pendens* (Kjems et al. 1990), *Thermoplasma acidophilum* (Ree et al. 1989), *Halobacterium cutirubrum* (Hui and Dennis 1985), *Halobacterium halobium* (Mankin and Kagramanova 1986) (the sequences of the latter two organisms are almost identical and we refer to them as *H. halobium* forthwith). *Halococcus morrhuae* (Larsen et al. 1986 and unpublished), *Methanococcus vannielii* (Jarsch and Böck 1985), *Methanobacterium thermoautotrophicum,* Marburg strain (Østergaard et al. 1987). We also examined the eubacteria/chloroplasts *Escherichia coli* rrnB (Brosius et al. 1981), *Bacillus subtilis* rrnB (Green et al. 1985), *Spinacia oleracea* (Massenet et al. 1987), and *Zea mays* (Kössel et al. 1982), and the eukaryotes *Saccharomyces cerevisiae* (Planta et al. 1980) and *Mus musculus* (Hassouna et al. 1984). Archaebacterial 16S–23S RNA spacer sequences were also examined from *Pyrodictium occultum, Thermococcus celer, Archaeoglobus fulgidus* VC-16, *Methanospirillum hungatei* (Achenbach-Richter and Woese 1988), *Halobacterium volcanii* (sequenced by R. Gupta) (GenBank accession numbers M19340–M19344), and the leader and spacer sequences from *Methanobacterium formicicum* (provided by A. Böck). All of these sequences were investigated for repeats, inverted repeats, and specific motifs using Staden's ANALYSEQ and DIAGON programs (Staden 1982). Regions of significant sequence similarity were aligned by a combination of visual inspection and editing using the program ALMA (S. Thirup and N. Larsen, unpublished results). Secondary structures were edited and plotted using the programs EDSTRUC and PLSTRUC, respectively (Leffers et al. 1987).

Results

Sequences flanking the rRNAs in the 15 available archaebacterial transcripts (seven complete and eight partial) were aligned and compared. Initially, this process was facilitated by aligning structural markers including tRNA in the 16S–23S RNA spacer, the previously characterized 16S and 23S RNA processing stems, and 5S RNA downstream from the 23S RNA. Sequence similarity was detected, especially among the extreme thermophiles and among the extreme halophiles (see below), although only short sequences were common to all archaebacterial transcripts (Fig. 1). On the basis of these sequence similarities, alignments were improved by hand. A search was then made for base pairing, initially within local sequence regions and, subsequently, over larger regions. Compensating changes between Watson–Crick pairings A–U, U–A, G–C, and C–G were considered as positive evidence; changes involving G.U were considered neutral; and other changes were considered negative.

By these criteria, all of the archaebacterial transcripts could generate common secondary structural elements in addition to the two long processing stems. Most could form two helices in the 16S RNA leader (A and B) and two in the 16S–23S RNA spacer (E and F). Moreover, the extreme halophiles could form two extra helices (C and D) in their relatively large spacers, and one (G) directly preceding the 5S RNA. These structures are presented for *T. pendens, H. halobium* (*H. cutirubrum*), and *M. thermoautotrophicum* in Fig. 1a–c, where sections of the helices that are supported by compensating base changes (CBCs) are indicated by vertical lines—thick lines for all archaebacteria and thin lines for one of the subgroups (extreme thermophiles, extreme halophiles, or methanogens). These helices are considered individually below.

Sequence alignments were further improved on the basis of the putative secondary structure, especially within the double helices, and nucleotide positions that are conserved among all the transcripts are underlined in Fig. 1, whereas less conserved positions, which are nevertheless generally conserved among the subgroups, are underlined with dots.

The archaebacterial sequences were also compared with those flanking the rRNAs of eubacteria (*E. coli, B. subtilis*), chloroplasts (maize and spinach), and eukaryotes (yeast, mouse). No evidence was found for conservation of the primary structure. Moreover, although many double helices can be formed (see Brosius et al. 1981), they do not correspond with those described above, except for the long processing stems in eubacteria/chloroplasts, which, however, lack the staggered bulged-loop structures that are boxed in Fig. 1.

Archaebacteria-Specific Helices

Of the four putative helices, three (B, E, and F) contain sections that are common to the archaebacterial transcripts. The appropriate sequences are aligned for the available transcripts in Fig. 2, where it can be seen that several base pairs in each helix are supported by CBCs. These base pairs are denoted by thick vertical lines in Fig. 1. Only the long helix A could not be subjected to a phylogenetic

analysis owing to difficulty in aligning the variable sequences.

Helix A

Helix A is irregular and can form in all the archaebacterial transcripts except for those of *D. mobilis* and *M. vannielii*. Although there is no formal phylogenetic support or conserved base pairing among all archaebacteria, the lower part of the helix was aligned for the three methanogens, and this yielded positive CBCs for 4 bp (data not shown), which are denoted by a thin vertical line (Fig. 1c).

Helix B

Helix B lies immediately upstream from the 16S RNA processing stem in all the transcripts. The first 6 bp of the lower part of the helix are supported by CBCs among all the transcripts (Fig. 2a). There was only one mismatch that occurred in the second base pair of *T. acidophilum* (Fig. 1d). In the extreme thermophiles and halophiles, and in *M. formicicum*, an additional variable internal loop and an extended upper helix can form, and the latter is supported by CBCs (Fig. 2a). The other methanogens only exhibit the lower part of the helix. Sequences directly flanking the helix are partially conserved as RUGNN–helix–YGA (Fig. 2a).

Helix E

Helix E can form in all transcripts except those of *T. tenax* and *P. occultum*, which both lack much of the spacer region between the two processing stems (Kjems et al. 1987; Achenbach-Richter and Woese 1988), and of *T. acidophilum*, which has no spacer (Ree et al. 1989). Among all the other archaebacterial transcripts, a helix with 9–12 bp can form where the first 9 bp are supported by multiple CBCs (Fig. 2b). In the extreme halophiles and methanogens, the helix exhibits an unpaired nucleotide in a fixed position on the 3′-side of the helix (except for *A. fulgidus* and *M. vannielii*) (Fig. 1b and c). The alignments (Fig. 2b) reveal the partially conserved flanking sequences RUGCA–helix E–GAAG.

Helix F

This helix directly precedes the 23S RNA processing stem and is subject to insertions/deletions in the apex of the hairpin. Almost all of the archaebacterial transcripts can generate at least 8 bp, and multiple CBCs occur for the first 4 with no mismatches. Extended helices can form in *T. pendens* (Fig. 1a), and in *S. solfataricus*, which forms a regular, A–U-rich, 20-bp helix. Among the extreme halophiles, the first 6 bp are supported by CBC's with no mismatches; then follows an unpaired nucleotide on the 5′-strand of the helix and a helix extension that is also supported by CBCs (Fig. 2b). Among the methanogens, helix F is variable in size; larger helices can form in *M. formicicum* (8 bp) and *M. hungatei* (5 bp), a small one in *M. thermoautotrophicum* (3 bp) (Fig. 1c), and the region is absent from *M. vannielii*. The helix location is defined by the bordering sequences GA and RUGA (Figs. 1 and 2b).

Extreme Halophile-Specific Helices

Three additional helices were discerned in the extreme halophile transcripts. Alignments are shown for the shorter helices C and G in Fig. 3 where the three CBCs occur for each helix. Helix C is located immediately upstream from the $tRNA^{Ala}$ and exhibits 6 bp except in the *H. morrhuae* transcript, which contains an additional 20 nucleotides at the apex, whereas helix G directly precedes 5S RNA [and has been proposed before without phylogenetic evidence (Larsen et al. 1986)]. The larger helix D is aligned for *H. halobium* and *H. morrhuae* in Fig. 4b where CBCs occur for 8 bp that are indicated by thin vertical lines in Fig. 1b. The remaining base pairs are either conserved or exhibit neutral changes (involving G.U pairs). In the *H. volcanii* transcript the helix is 3 bp longer. Helix D is irregular and its apex loop exhibits a conserved sequence corresponding to a promoter motif for 23S RNA transcription (Mankin et al. 1987; see Figs. 1b and 4b).

16S–23S RNA Processing Stems

Both processing stems also occur in eubacterial/chloroplast transcripts (Brosius et al. 1981). However, only the archaebacterial stems exhibit the two 3-nucleotide bulges, separated on opposite strands by 4 bp, which constitute a substrate for a processing enzyme (Mankin et al. 1984; Hui and Dennis 1985; Jarsch and Böck 1985; Kjems and Garrett 1987; Kjems et al. 1987). This feature is only absent from the 16S RNA transcript of *T. acidophilum* (Fig. 1d) which is transcribed independently of the 23S RNA (Ree et al. 1989). Comparative analyses revealed that the lengths and sequences of the archaebacterial stems are highly variable, even among closely related species, which renders a phylogenetic analysis of the secondary structures difficult.

Similarities between the Leader and Spacer Regions in the Extreme Halophiles

Superficially, the secondary structures flanking the 16S and 23S RNAs resemble one another. The clearest example is provided by the processing stems, but there are also similarities in position between helices A and B in the leader, and helices E and F, respectively, in the spacer (Fig. 1). Therefore, we examined

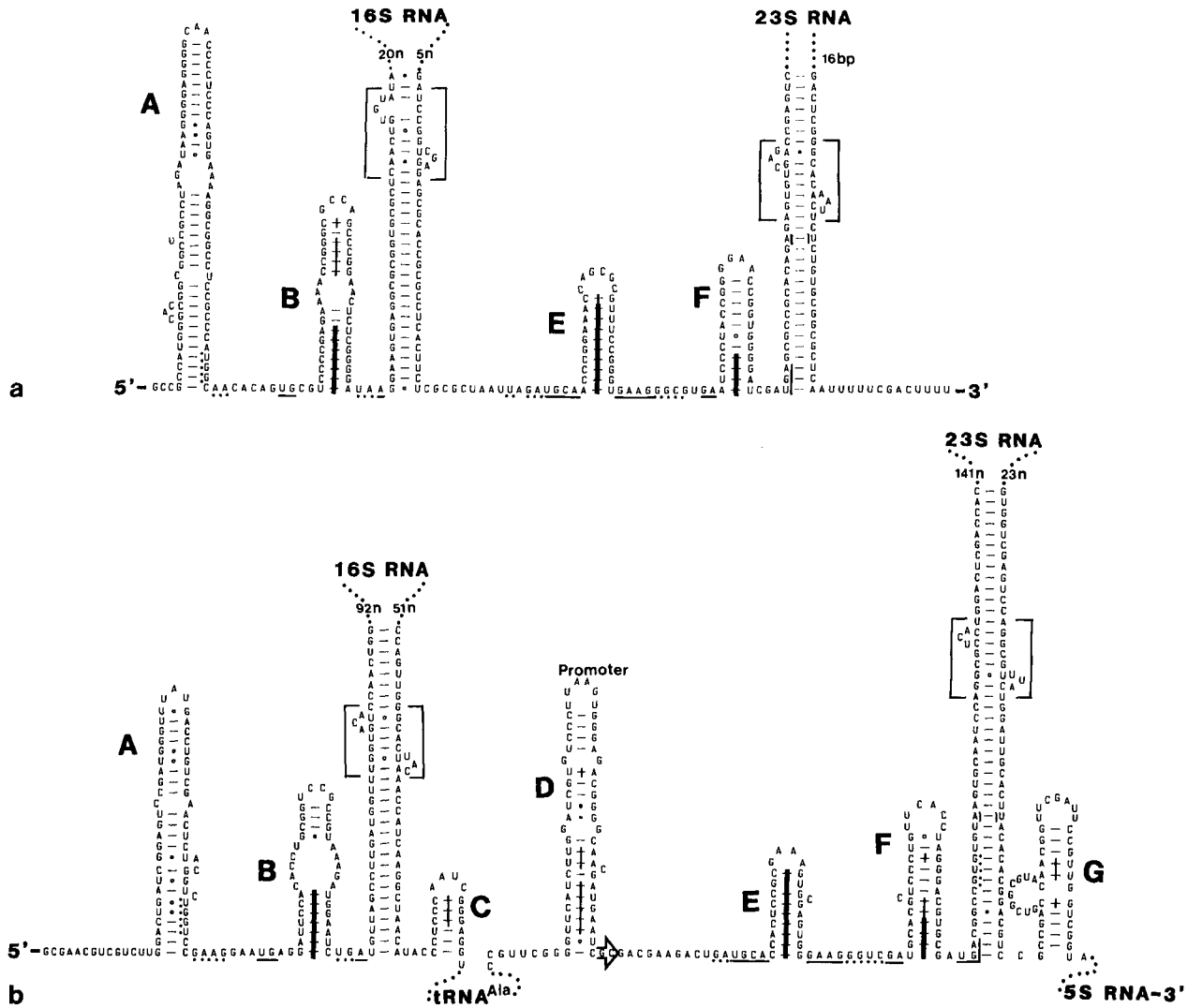

Fig. 1. Representative and putative secondary structures of the pre-rRNA transcripts are shown for **a** *T. pendens* (extreme thermophile), **b** *H. halobium* (extreme halophile), **c** *M. thermoautotrophicum*, Marburg strain (methanogen), and **d** *T. acidophilum* for which the 16S RNA is transcribed separately. Helices are labeled A–G, and shortened versions of the 16S and 23S RNA processing stems are shown. Base pairs that are supported by compensating base changes either among all the archaebacteria or within the subgroups are indicated by thick or thin vertical lines, respectively, between the paired bases. A few transcripts lack part or all of a double helical region and these are eliminated from the analysis. Processing sites are boxed. The transcripts are drawn from the putative initiation site closest to the 16S RNA and end at a putative termination site (in a and d) or at 5S RNA (in b and c). Nucleotides that are highly conserved among archaebacteria are underlined; less conserved positions, which are generally conserved within the subgroups, are underlined with dots. For the extreme halophiles, the putative promoter at the apex of helix D and the subsequent transcription initiation site (⇒) are indicated. Structures were edited and plotted using the programs EDSTRUC and PLSTRUC, respectively (Leffers et al. 1987).

the relationships between the leader and spacer sequences in more detail. Very few similarities were detected for the extreme thermophile and methanogen transcripts, but some were observed for those of the extreme halophiles, including that between the leader promoter and the spacer promoter that was first observed by Mankin et al. (1987).

In order to illustrate this, parts of the leader and spacer sequences of *H. halobium* and *H. morrhuae* are aligned together in Fig. 4. They reveal first, leader sequences that are highly similar to the 3′ one-third of the spacer $tRNA^{Ala}$s (Fig. 4a). These are preceded by similar sequences, extending over 40–50 nucleotides (Fig. 4b), that include helix D, promoters with the sequence TTAAGT, and transcription start sites (I). Finally, helices B and F are aligned (Fig. 4c) and both are flanked by conserved (boxed) sequences. This partial duplication of sequence/secondary structure between the leader and spacer sequences could reflect that the 16S and 23S RNA genes were transcriptionally uncoupled in ancestral cells, but it could also result from the high frequency

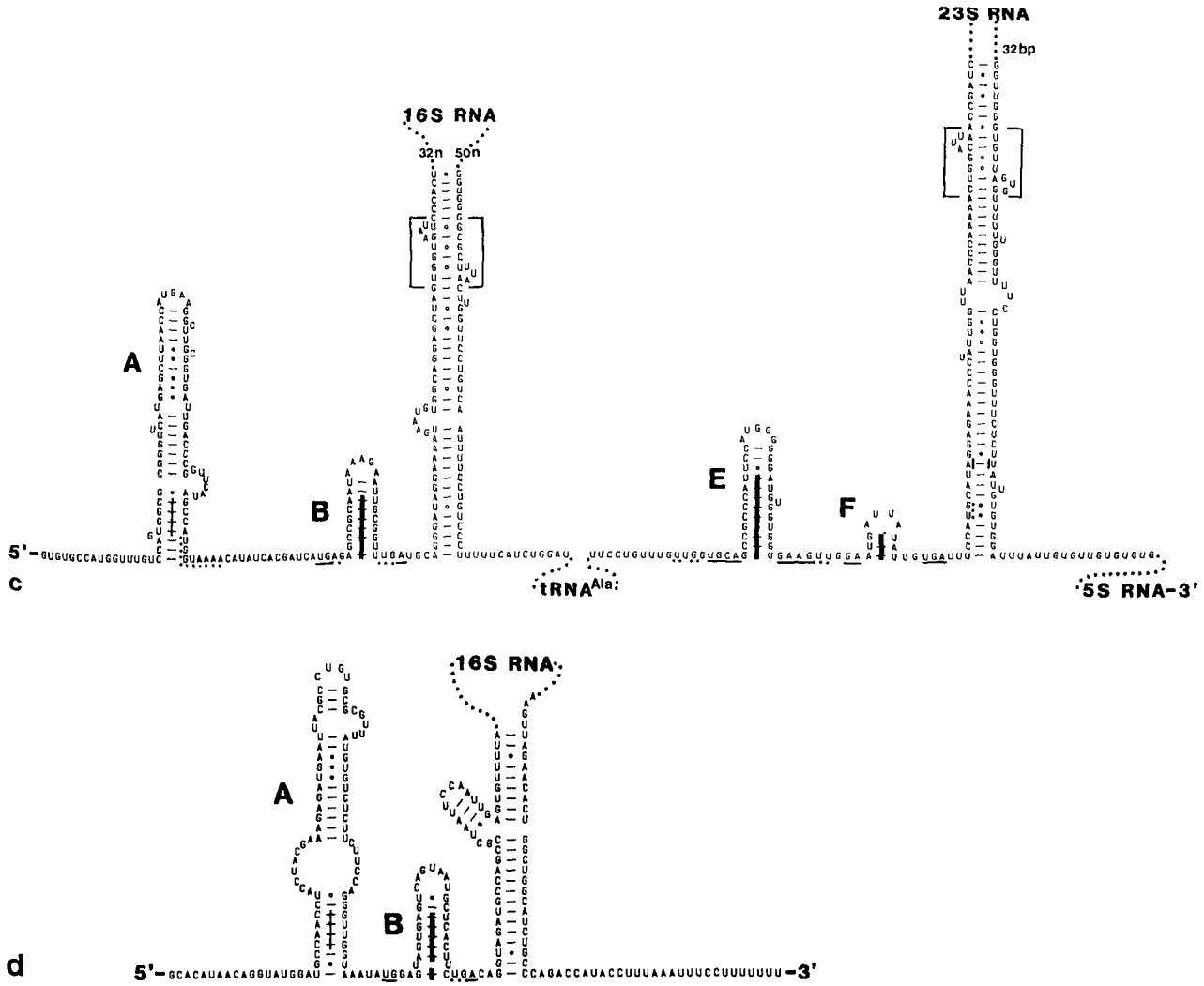

Fig. 1. Continued.

of recombination common among extreme halophiles.

Discussion

The leader, spacer, and downstream regions of the archaebacterial rRNA operons are diverse in size, organization, and sequence. The extreme thermophile operons, in particular, exhibit some eukaryotic-like characteristics including no tRNA in the spacer, uncoupled 5S RNA, and introns in the 23S RNA, and this was one important basis for the proposal of Lake and colleagues (Lake 1988, 1989, and references therein) that the extreme thermophiles are not monophyletic with the other archaebacteria but rather constitute a separate kingdom, the eocytes. This, in turn, led us to question whether the flanking sequences themselves yield useful evolutionary information bearing on the phylogenetic status of the archaebacteria.

Clearly, the primary sequences of the flanking regions are too variable for an extensive alignment, which precludes meaningful tree building. Nevertheless, we show that the archaebacterial transcripts can generate common secondary structures, and the results are summarized in Table 1. Clearly, the level of phylogenetic support does not generally reach the high standard observed for rRNA secondary structures. The main difference is in the degree of variation in size and complexity of some of the leader/spacer helices. Often one part of a helix is supported phylogenetically (for example the lower part of helix B), whereas the remainder is more variable. On the other hand, inserts in the apex loops of helices, as occur in helix C of *H. morrhuae,* are common in the rRNAs. This lower level of conservation of the

a

```
                                                                 16S stem
              <---------------------Helix B-------------------->   -->
              +++++++                                    +++++++
D.m.    AUGGG GAUUCAAGCACCUCGGGAGGCACGAUUACCUCCCGAGA---GCUUGAAUC CGACGA
S.s.    GUGGA GGGGCAAGAUCCCCGGGC---CCCUAA---GCCCGGGA--AGCUUGUCUC UGACAA
T.p.    GUGCG UCCCCGAGAAAACCGGGC----GCCA----GCCCGGAA--CUCUCGGGGA UAAGGA
T.t.    AUGUU CCCCCGACUAACCCCCCUGGG-UUUAG----CCCGGGCGGGCGUCGGGGG AGAAAA
H.h.    AUGAG GAUUCCACACCUGCGG------UCCG------CCGUAA--AGAUGGAAUC UGAUGU
H.m.    AUGAG GAUUCCGCCCCUGCGG------UCCG------CCGUCA--AGACGGGAUC UGACGU
T.a.    AUGGA GUAGUGAGU------------CAGUAAU-------------GCUCACUUC UGACAG
M.f.    AUGAG GGCCUAAAACUUUGC-------GACAA-----GCAAACUGCUUUUAGGUU UGAUGU
M.t.    AUGAG AGCCGCAAU-------------AAAGA--------------AUUGCGGUU UGAUGC
M.v.    GUGAU CGCCUGA--------------ACUAAAUCAC------------UCAGGUG UGAUGU
Cons.   RUG         A                                      U     YGA
```

b

```
                                                                              23S stem
            <---------Helix E---------->                <------Helix F------>    ---->
            +++++++++          +++ ++++++               ++++             ++++
D.m.  AUGCA ACACGGGG-----UUAU----CC-CCGUGU GAAGG---CGUGA ACCCCGUGGGAUCCACGGGGU --UGUGAGC
S.s.  AUGCA GUUCUCCUCU--UGGGCC-AGAG-GGGAAU GAAGUGC--CUAG GGCCCAUU-32n-UUAGGGCU -CAAUGAGG
T.p.  AUGCA ACCCGGAAAC--CAGCGC-GUUU-CCGGGU GAAGGGC--GUGA AUCCCUAC-10n-GUGGGGAU -CGAUGAGC
A.f.  GUGCA CCCGGCGAU----GAAA---GUC-GUCGGG GAAGGGU---UGA UGGGCAAG-----UUUGCCCA -GGGUGAAG
T.c.  GUGCA CAUCCCAAGC---CCC---GCUU-GGGAUG GAAGGGU--CCGA GACCCCGAA----CAGGGGUC ACGAUGAGG
H.h.  AUGCA CCACUCCGC----GAAA---GUGCGAGUGG GAAGGGU---CGA UGCACGCU-15n-GACGUGCG ---AUGACG
H.m.  AUGCA CCACCCCGU----GCAA---GCGUGGGUGG GAAGGGU---UGA GACGUGCC-12n-GGCGCGUC --CAUGAGA
H.v.  AUGCA CCAUCCCGC----GCAA---GCGAGGUUGG GAAGGGU---CGA AAACACGC-15n-GGGUGUUU --GAUGAAA
M.f.  GUGCA GCUAGUAUCAC--GGA--GUGAUAGCUAGU GAAGG------GA UAAAAAAA-22n-UUUUUUUG ---AUGAGU
M.h.  AUGCA CCCGGUCUG----GCGA---CAGAACCGGG GAAGGGCUUGAGA CCCCGUGAA-----AUCGGGG -CGAUGAUA
M.t.  GUGCA GCCGCCCAUUCCAUGGGGGGAUGUGGUGGU GAAG----UUGGA AUG------AUUA-----UAU --UGUGAUU
M.v.1 AUGCA GUCUAUGAUAU--UUG--AUAUU-AUAGAU CAAGGACC--AGA ---------UAUU-------- --AAUGAUG
M.v.2 GCCCA GCUCGUAAUU---UUUA--GAUU-ACGAUC UAGGGGCC--GGA ---------UAUU-------- --AAUGAUG
Cons. RUGCA                                GAAGGG     GA                          RUGA
```

Fig. 2. Alignment of the available sequences that generate **a** helix B in the leader and **b** helices E and F in the spacer. Complementary sequences, which generate the putative helices, are indicated by arrows over the sequences, and the limits of each helix are denoted by a free space in the alignment immediately before and after each helix. + indicate that putative complementary nucleotides in a helix are supported by CBCs. Thus, for helix B, the first 7 bp receive phylogenetic support. The approximate start points of the 16S and 23S RNA processing stems are indicated by arrows. Consensus sequences common to at least 80% of the organisms are given below. D.m.—*D. mobilis*, S.s.—*S. solfataricus*, T.f.—*T. pendens*, A.f.—*A. fulgidus*, T.c.—*T. celer*, H.h.—*H. halobium*, H.m.—*H. morrhuae*, H.v.—*H. volcanii*, T.a.—*T. acidophilum*, M.f.—*M. formicicum*, M.h.—*M. hungatei*, M.v.—*M. vannielii*.

```
Helix C:
           ----->                  <-----
          -ooo+++        apex      +++ooo-
H.h.      CCCUCCC        AAUC      GGGAGGU    -   tRNA

H.m.      GCCUUCC-20n-AAAC         GGAAGGC    -   tRNA

H.v.      ACCUAAG        AACUA     CUAGGGC -5n-   tRNA


Helix G:
           ---->             ---->              <---- <----
          -oooo+            -++ooo    apex      ooo++-+oooo-
H.h.      GCCGAC -loop-     CAACGG   UUCGAUU    CCGUUGGUCGGU

H.m.      ACCGAU  - C -     ACGCGG   UUCGAUU    CCGCGGAUCGGC
```

Fig. 3. Alignments of extreme halophile sequences for helices C and G. Complementary sequences, which generate the helices, are indicated by arrows. The signs over the aligned sequences designate the phylogenetic evidence (CBCs) for a putative base pair: + indicates positive evidence, 0 denotes neutral evidence, and − indicates negative evidence.

leader/spacer structures, relative to those of the rRNAs, almost certainly reflects a lower level of functional constraint during their evolution.

The leader/spacer structures may constitute regulatory or maturation signals. Thus, helices A and/or B could have an attenuator role in transcription. In the extreme halophiles, helices C and G could facilitate maturation of $tRNA^{Ala}$ and 5S RNA, respectively. Moreover, helix D, which contains the 23S RNA promoter in its apex loop, and its equivalents in the leader sequence, which are not shown, could be involved in transcriptional regulation at a DNA or RNA level. Clearly, however, these putative functional roles are not essential to all archaebacteria. For example, helix A, which is weakly supported phylogenetically, cannot form in *D. mobilis* and *M. vannielii.* Moreover, large parts of the spacers between the processing stems are absent from *T. tenax* (Kjems et al. 1987) and *P. occultum* (Achenbach-Richter and Woese 1988) without seriously impeding cellular growth. The most important helices functionally among the archaebacteria are probably helices B, E, and F, which are the most conserved and, also, the best supported phylogenetically (Table 1).

We conclude that there is sufficient conserved

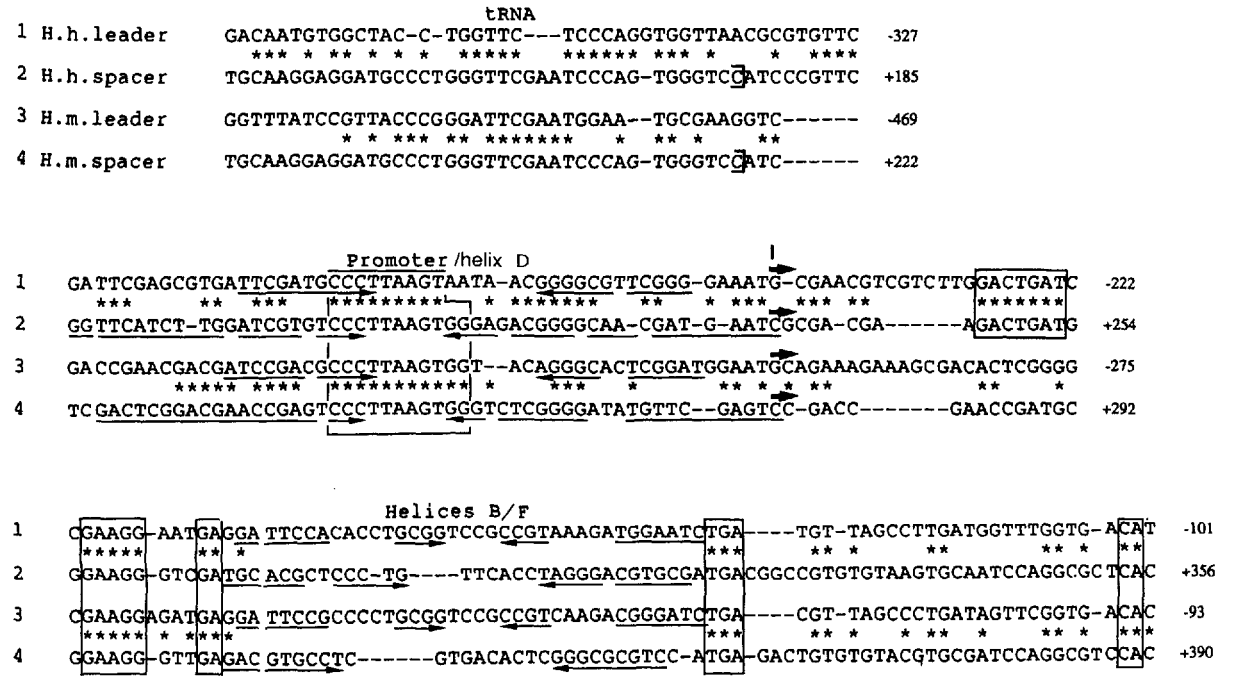

Fig. 4. Alignments of parts of the 16S RNA leader and spacer sequences for both *H. halobium* (H.h.—1 and 2) and *H morrhuae* (H.m.—3 and 4) on the basis of both primary and secondary structure. **a** The leader sequences are aligned with the spacer $tRNA^{Ala}$s. The 3'-ends of the tRNAs are indicated by squared brackets. **b** The promoter and transcription initiation regions (marked with an I and an arrow) in the leaders and spacers, respectively, are aligned. **c** Sequence regions containing helices B and F are aligned. Similar nucleotide positions in the leader and spacer regions of each organism are denoted by asterisks where the alignment is most reliable. Major conserved regions common to all four sequences are boxed. Complementary sequences generating the helices are underlined by arrows. For the leaders, nucleotide positions are numbered negatively from the 5'-ends of 16S RNA, and for the spacers they are numbered from the 3'-ends of 16S RNA. The alignment was edited using the program ALMA (S. Thirup and N. Larsen, unpublished).

Table 1. A summary of the secondary structural elements detected in the archaebacterial leader/spacer regions

Helix	Occurs in	Size	Base pairs supported by CBCs	
			Archaebacteria	Subgroups (t, h, m)
A	All −D.m., −M.v.	Large/variable	—	4 m
B	All	Variable	7	12 t, 7 h, 6 m
C	H.h., H.m., H.v.	6 bp	—	3 h
D	H.h., H.m., H.v.	20–24 bp	—	8 h
E	All −P.o., −T.a., −T.p.	9–12 bp	9	10 t, 9 h, 9 m
F	All −P.o., −T.a., −T.p., −M.v.	Variable	4	4 t, 6 h, 4 m
G	H.h., H.m., H.v.	11 bp	—	3 h

Individual organisms are designated as defined in the legend to Fig. 2. In the last column t denotes extreme thermophiles; h, extreme halophiles; and m, methanogens

structure in the leader and spacer regions of the rRNA operons to support the concept that the extreme thermophiles, extreme halophiles, and methanogens are closely related. The leaders and spacers of the eubacteria/chloroplasts and eukaryotes can also generate putative stem-loop structures (see, for example, Brosius et al. 1981), but by our criteria of location and form, none of those detected correlate with the archaebacterial helices. Clearly, we cannot eliminate the possibility that a helix with a conserved function occurs in an altered position or form in transcripts of eubacteria/chloroplasts or eukaryotes. However, our failure to detect corresponding structures, with the exception of the processing stems in eubacteria/chloroplasts, provides further support for the archaebacteria being monophyletic.

Acknowledgments. We thank L. Achenbach-Richter, Carl Woese, August Böck, and Ramesh Gupta for sharing unpublished sequences, and Niels Larsen, Søren Thirup, and Jacob Dalgaard for providing computing help. Alexander Mankin and the referees are thanked for their constructive criticism. Lisbeth Heilesen's help with the manuscript was appreciated. The Carlsberg Foundation and the Danish Natural Science Research Council provided grants to J.K. and for the research (to R.A.G.).

References

Achenbach-Richter L, Woese CR (1988) The ribosomal RNA gene spacer in archaebacteria. Syst Appl Microbiol 10:211–214

Brosius J, Dull TJ, Sleeter DD, Noller HF (1981) Gene organization and primary structure of a rRNA operon from *Escherichia coli*. J Mol Biol 148:107–127

Fox GE, Stackebrandt E, Hespell RB, Gibson J, Maniloff J, Dyer TA, Wolfe RS, Balch WE, Tanner R, Magrum L, Zablen LB, Blakemore R, Gupta R, Bonen L, Lewis BJ, Stahl DA, Luehrsen KR, Chen KN, Woese CR (1980) The phylogeny of prokaryotes. Science 209:457–463

Green CJ, Stewart GC, Hollis MA, Vold BS, Bott KF (1985) Nucleotide sequence of the *Bacillus subtilus* rRNA operon *rrn*B. Gene 37:261–266

Hassouna N, Michot B, Bachellerie JP (1984) The complete nucleotide sequence of mouse 28S rRNA gene. Implications for the process of size increase of the large subunit rRNA in higher eukaryotes. Nucleic Acids Res 12:3563–3583

Hui I, Dennis PP (1985) Characterization of the rRNA gene clusters in *Halobacterium cutirubrum*. J Biol Chem 260:899–906

Jarsch M, Böck A (1985) Sequence of the 23S rRNA gene from the archaebacterium *Methanococcus vannielii*: evolutionary and functional implications. Mol Gen Genet 200:305–312

Kjems J, Garrett RA (1987) Novel expression of the rRNA genes in the extreme thermophile and archaebacterium *Desulfurococcus mobilis*. EMBO J 6:3521–3530

Kjems J, Leffers H, Garrett RA, Wich G, Leinfelder W, Böck A (1987) Gene organization, transcription signals and processing of the single rRNA operon of the archaebacterium *Thermoproteus tenax*. Nucleic Acids Res 15:4821–4835

Kjems J, Leffers H, Olesen T, Holz I, Garrett RA (1990) Sequence, organization and transcription of the ribosomal RNA operon and the downstream tRNA and protein genes in the archaebacterium *Thermofilum pendens*. Syst Appl Microbiol (in press)

Kössel H, Edwards K, Koch W, Langridge P, Schiefermayer E, Schwarz Z, Strittmatter G, Zenke G (1982) Structural and functional analysis of an rRNA operon and its flanking tRNA genes from *Zea mays* chloroplasts. Nucleic Acids Res Symp 11:117–120

Lake JA (1988) Origin of the eukaryotic nucleus determined by rate-invariant analysis of rRNA sequences. Nature 331: 184–186

Lake JA (1989) Origin of the eukaryotic nucleus: eukaryotes and eocytes are genetically related. Can J Microbiol 35:109–118

Lake JA, Hendersen E, Clark MW, Scheinman A, Oakes MI (1986) Mapping evolution with three dimensional ribosome structure. Syst Appl Microbiol 7:131–136

Larsen N, Leffers H, Kjems J, Garrett RA(1986) Evolutionary divergence between the rRNA operons of *Halococcus morrhuae* and *Desulfurococcus mobilis*. Syst Appl Microbiol 7: 49–57

Leffers H, Kjems J, Østergaard L, Larsen N, Garrett RA (1987) Evolutionary relationships amongst archaebacteria: a comparative study of 23S rRNAs of a sulfur-dependent extreme thermophile, extreme halophile and a thermophilic methanogen. J Mol Biol 195:43–61

Mankin AS, Kagramanova VK (1986) Complete nucleotide sequence of the rRNA operon of *Halobacterium halobium*: a secondary structure of the archaebacterial 23S rRNA. Mol Gen Genet 202:152–161

Mankin AS, Teterina NL, Baratova LA, Kagramanova VK (1984) Putative promoter region of rRNA operon from archaebacterium *Halobacterium halobium*. Nucleic Acids Res 12:6537–6546

Mankin AS, Skripkin EA, Kagramanova VK (1987) A putative internal promoter in the 16S/23S intergenic spacer of the rRNA operons of archaebacteria and eubacteria. FEBS Lett 219:269–273

Massenet O, Martinez P, Seyer P, Briat J-F (1987) Sequence organization of the chloroplast ribosomal spacer of *Spinacia oleracea* including the 3'-end of the 16S RNA and the 5'-end of the 23S RNA. Plant Mol Biol 10:53–63

Neumann H, Gierl A, Tu J, Leibrock J, Staiger D, Zillig W (1983) Organization of the genes for ribosomal RNA in archaebacteria. Mol Gen Genet 192:66–72

Olsen GJ, Pace NR, Nuel M, Kaine BP, Gupta R, Woese CR (1985) Sequence of the 16S rRNA gene from the thermophilic archaebacterium *Sulfolobus solfataricus* and its evolutionary implications. J Mol Evol 22:301–307

Østergaard L, Larsen N, Leffers H, Kjems J, Garrett RA (1987) A rRNA operon and its flanking region from the archaebacterium *Methanobacterium thermoautotrophicum*, Marburg strain: transcription signals, rRNA structure, and evolutionary implications. Syst Appl Microbiol 9:199–209

Planta RJ, Klootwijk J, Raué HA, Brand RC, Veldman GM, Stiekema WJ, de Jonge P (1980) Transcription and processing of ribosomal RNA in yeast and bacilli. In: Osawa S, Ozeki H, Uchida H, Yura T (eds) Genetics and evolution of RNA polymerase tRNA ribosomes. Elsevier, Amsterdam, pp 451–470

Ree HK, Cas K, Thurlow D, Zimmermann R (1989) Structure and organization of the 16S rRNA gene from the archaebacterium *Thermoplasma acidophilum*. Can J Microbiol 35:124–133

Reiter WD, Palm P, Voos W, Kaniecki J, Grampp B, Schulz W, Zillig W (1987) Putative promoter elements for the rRNA genes of the thermoacidophilic archaebacterium *Sulfolobus* sp. strain B12. Nucleic Acids Res 15:5581–5595

Staden R (1982) An interactive graphics program for comparing and aligning nucleic acid and amino acid sequences (DIAGON). Nucleic Acids Res 10:2950–2696

Wich G, Leinfelder W, Böck A (1987) Genes for stable RNA in the extreme thermophile *Thermoproteus tenax*: introns and transcription signals. EMBO J 6:523–528

Woese CR, Olsen G (1986) Archaebacterial phylogeny: perspectives on the urkingdoms. Syst Appl Microbiol 7:161–177

Received July 18, 1989/Revised January 5, 1990