Volume 15 Number 12 1987 Nucleic Acids Research

Gene organization, transcription signals and processing of the single ribosomal RNA operon of the archaebacterium *Thermoproteus tenax*

Jørgen Kjems*, Henrik Leffers, Roger A.Garrett, Günter Wich[1], Walfred Leinfelder[1] and August Böck[1]

Biostructural Chemistry, Kemisk Institut, Aarhus Universitet, DK-8000 Aarhus C, Denmark and [1]Lehrstuhl für Mikrobiologie der Universität, D-8000 München 19, FRG

Received March 13, 1987; Revised and Accepted May 12, 1987 Accession no. Y00346

ABSTRACT

The single ribosomal RNA (rRNA) operon from the extreme thermophile and archaebacterium *Thermoproteus tenax* was sequenced. Sites of transcriptional initiation and termination were established and the processing sites on the primary transcript were mapped with nuclease S1. The operon contained genes coding for 16S and 23S RNAs but lacked those coding for tRNA and 5S RNA. Transcription initiates 175 bp upstream from the start of the 16S RNA gene (Wich *et al.*, EMBO J. 6, 523-528, 1987 [1]) and terminates 49 bp downstream from the 23S RNA gene within a long pyrimidine sequence. An open reading frame downstream from the rRNA operon is transcribed.

The sequences bordering both 16S and 23S RNA genes can form putative processing stems in the primary transcript that involve the whole of the 16S-23S RNA spacer. The stems contain irregular features that constitute processing signals and are conserved in other archaebacteria. The 16S RNA stem is cut prior to that of the 23S RNA and RNA maturation follows. An unusual 14 bp helix can form between the extremities of the transcript such that the whole transcript is highly structured and a fork-like structure is formed together with the processing stems.

The 23S RNA sequence was aligned with other available 23S-like RNA sequences (Leffers *et al.*, J. Mol. Biol. 195, in press [2]): a putative secondary structure exhibiting archaebacterial-specific features was deduced using comparative sequence analyses. A rooted phylogenetic tree was also derived for the archaebacteria that confirms their division into three major subgroups.

INTRODUCTION

The archaebacteria constitute a diverse group of organisms which includes the extreme thermophiles, the extreme halophiles and the methanogens [3]. Comparative sequence studies of both the 16S RNAs [4] and 23S RNAs [2] are consistent with the archaebacteria having evolved as a separate kingdom. Nevertheless, there are major phenotypic differences amongst the archaebacteria that include the subunit patterns of the RNA polymerases [5] and ribosome shapes [6] which have led Lake [6] to question whether all of the archaebacteria belong to one taxon. The gene organization of the rRNA operons also exhibit differences [8, 9]: thus, tRNA genes are absent from the sequenced extreme thermophile operons, and the *Desulfurococcus mobilis* operon lacks a coupled 5S RNA gene [9] and exhibits an intron in the 23S RNA gene [10].

rRNA genes are convenient systems for studying archaebacterial gene organization and transcription signals because of their relative ease of isolation. The availability of operon

sequences from various archaebacteria including *Halobacteria* [11-14], *Halococcus morrhuae* [2, 9], *Methanococcus vannielii* [15 and references therein] and *Methanobacterium thermoautotrophicum* [16] provides a basis for comparative studies. Here we develop this approach by determining and analysing the sequence of the single rRNA operon from the extreme thermophile *Thermoproteus tenax;* the 16S RNA leader and gene sequences were presented elsewhere [1, 17].

MATERIALS AND METHODS

Preparation of cells, cellular RNA and DNA; cloning and DNA sequencing

Cells from *T. tenax* were kindly provided by W. Zillig and grown under the conditions described by Zillig *et al.* [18]. Total cellular RNA was prepared using hot phenol extractions (65°C) essentially as described by Aiba *et al.* [19]. The phenol solution was frozen at -80°C for 2 min and thawed before separating the aqueous phase. Two phenol-chloroform extractions were performed before the RNA was precipitated with ethanol and stored at -80°C.

The procedures for DNA preparation and genomic cloning were described earlier (20). Subcloning procedures and the sequencing of the 16S RNA leader sequence and gene have also been described [1, 17]. The remaining sequence was obtained by digesting large DNA clones with four base pair-specific restriction enzymes. The mixtures were fractionated on 5% polyacrylamide gels and cloned into M13mp18 or mp19 in both directions [21]. Single-stranded DNA was prepared and used as a template for DNA sequencing by the dideoxynucleotide procedure [22] using [α^{35}S]dATP as radioactive substrate. The DNA sequence was determined on 4-6% polyacrylamide gels containing salt or wedge-shaped gradients [23].

S1 nuclease mapping

0.1 to 1 μg DNA restriction fragment was labelled at either the 5'- or 3'-end [24]. It was digested with another restriction enzyme and purified on agarose or polyacrylamide gels to ensure that only one labelled end was present. 0.1-0.5 μg labelled DNA fragment (~100,000 cpm) was co-precipitated with 50-100 μg total cellular RNA. It was then denatured at 85°C for 10 min and hybridized in 40 mM PIPES, pH 6.4, 1 mM EDTA, 0.4 M NaCl, 80% formamide at 2-3 different temperatures. Optimal hybridization temperatures were calculated according to Favaloro *et al.* [25] and samples were also hybridized 5°C above and below these optima. After 3 hr, 300 μl ice-cold buffer (50 mM sodium acetate, pH 4.6, 0.28 M NaCl, 4.5 mM $ZnSO_4$) was added, containing 20 μg/ml carrier ssDNA and 1 unit/μl S1 nuclease, and the solution was incubated for 1 hr at 30°C. The protected sites were determined by co-electrophoresing the sequence of the DNA fragment [26].

Determination of the 23S RNA secondary structure and phylogenetic analysis

All secondary structure predictions were performed using the sequence comparison approach that is based on evidence for compensating base changes. All of the available 23S-like RNAs were aligned for archaebacteria, eubacteria and the eukaryotic cytoplasm as described in

detail by Leffers *et al.* [2]. Phylogenetic analyses for the archaebacterial kingdom were also based on the aligned 23S RNA sequences from *M. vannielii* [15]; *H. halobium* [14]; *H. morrhuae, M. thermoautotrophicum* and *D. mobilis* [2]. Homology values and K_{nuc} values were calculated as described by Hori and Osawa [27] and the phylogenetic tree was derived using the FITCH program of Felsenstein [28].

RESULTS

Gene organization and primary structure of the operon

Only one rRNA operon was detected in the *T. tenax* chromosome using the Southern blotting procedure [29] and probing with radioactively labelled clones from within both the 16S and the 23S RNA gene (results not shown). This agrees with the conclusion of Neumann *et al.* [8]. The gene organization of this rRNA operon is illustrated in Fig. 1 and the nucleotide sequence is given in Fig. 2. The primary transcript contained neither tRNA genes nor introns. Moreover, the longest transcript terminated 49 bp after the 23S RNA gene (see below) and there was no coupled 5S RNA gene. A similar gene organization has also been reported for the extreme thermophile *D. mobilis* [9].

We sequenced 780 bp upstream and 2380 bp downstream from the primary transcript. An analysis of these sequences revealed neither rRNA nor tRNA genes but there was an open reading frame (ORF) downstream from the 23S RNA gene that coded for 388 amino acids (Fig. 1); the amino acid sequence is given in Fig. 2. Northern blotting experiments [24] were performed to establish whether the putative gene was transcribed. A *Dde*I restriction fragment extending from nucleotides 6397 to 7094 (Fig. 2) was labelled with [α-^{32}P]dATP on the

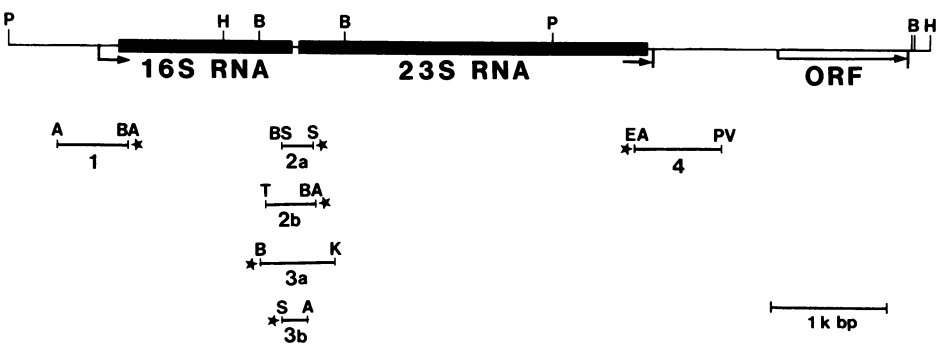


Figure 1: A map of the rRNA operon showing the arrangement of the rRNA genes, the location of the ORF and the major restriction sites. The limits of the rRNA transcript are indicated. The restriction fragments labelled 1 to 4 were used for S1 nuclease mapping. Asterisks denote the position of the end label on the DNA strand that was complementary to the RNA. Restriction enzyme abbreviations: A-*Alu*I, B-*Bam*HI, BA-*Ban*I, BS-*Bst*NI, EA-*Eag*I, H-*Hind*III, K-*Kpn*I, P-*Pst*I, PV-*Pvu*II and S-*Sau*96A.

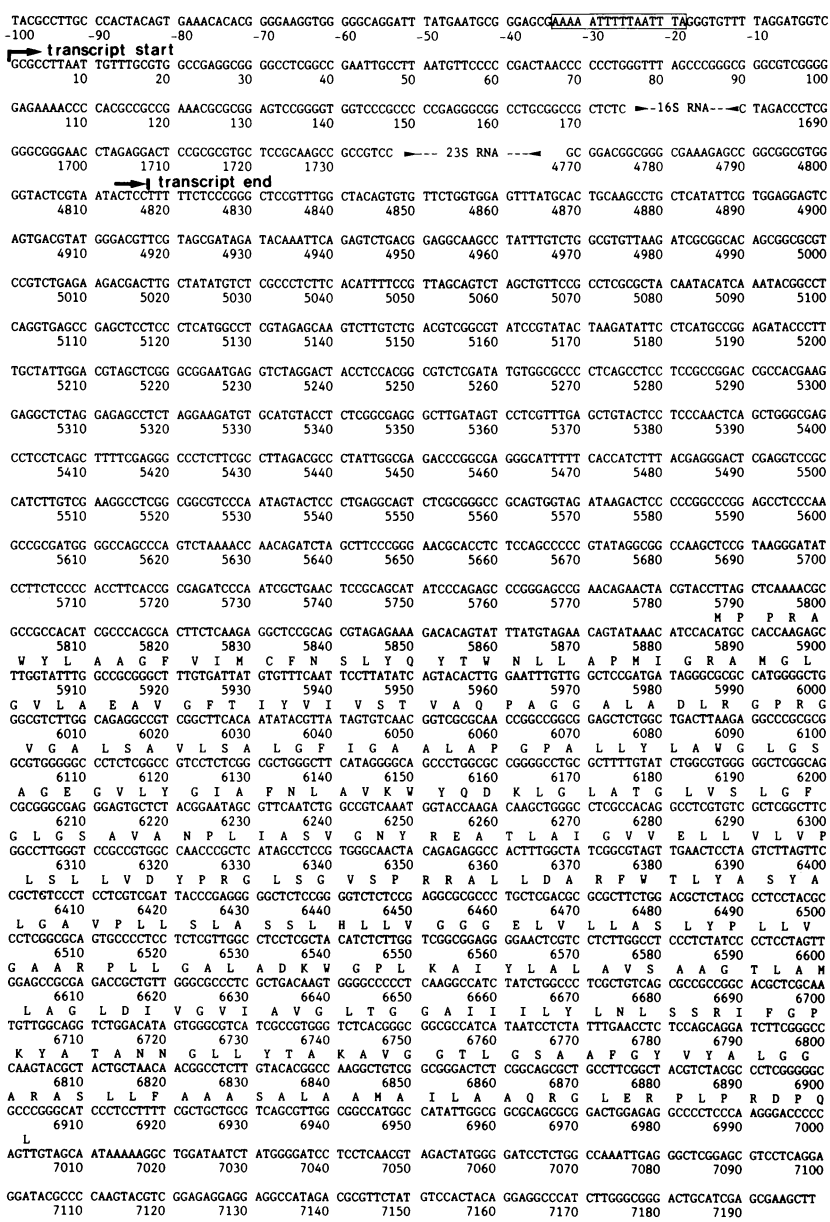


Figure 2: The DNA sequence corresponding to the rRNA primary transcript and the downstream ORF. The sequence of the 16S RNA was presented elsewhere [17] and the sequence of the 23S RNA is contained within the secondary structure shown in Fig. 5. The start of the primary transcript is indicated and a putative promoter sequence upstream it is boxed. The end of the transcript is also shown within a pyrimidine sequence. The amino acid sequence corresponding to the ORF sequence is included.

antisense strand and hybridized with total cellular RNA to reveal an band of about 4000 nucleotides (data not shown); we inferred, therefore, that the gene is expressed. The putative gene is preceded by an A-T rich sequences although none was unambiguously identified as a promoter (Fig. 2). Moreover, no Shine-Dalgarno sequence was detected. No homology was detected between this ORF and those flanking other archaebacterial rRNA operons (unpublished work).

Initiation and processing of the 16S RNA leader transcript

It has been shown previously by S1 nuclease mapping that initiation of rRNA transcription occurs at a single site 175 bp upstream from the 16S RNA gene [1]. A sequence comparison between the upstream region of this site and the upstream regions of tRNA primary transcripts revealed a common conserved A-T-rich sequence within a G-C-rich intercistronic spacer that could be a promoter motif [1]. This sequence, which is boxed in Fig. 2, differs from the putative promoter motifs found in the extreme halophiles [9, 11, 12] and methanogens [30].

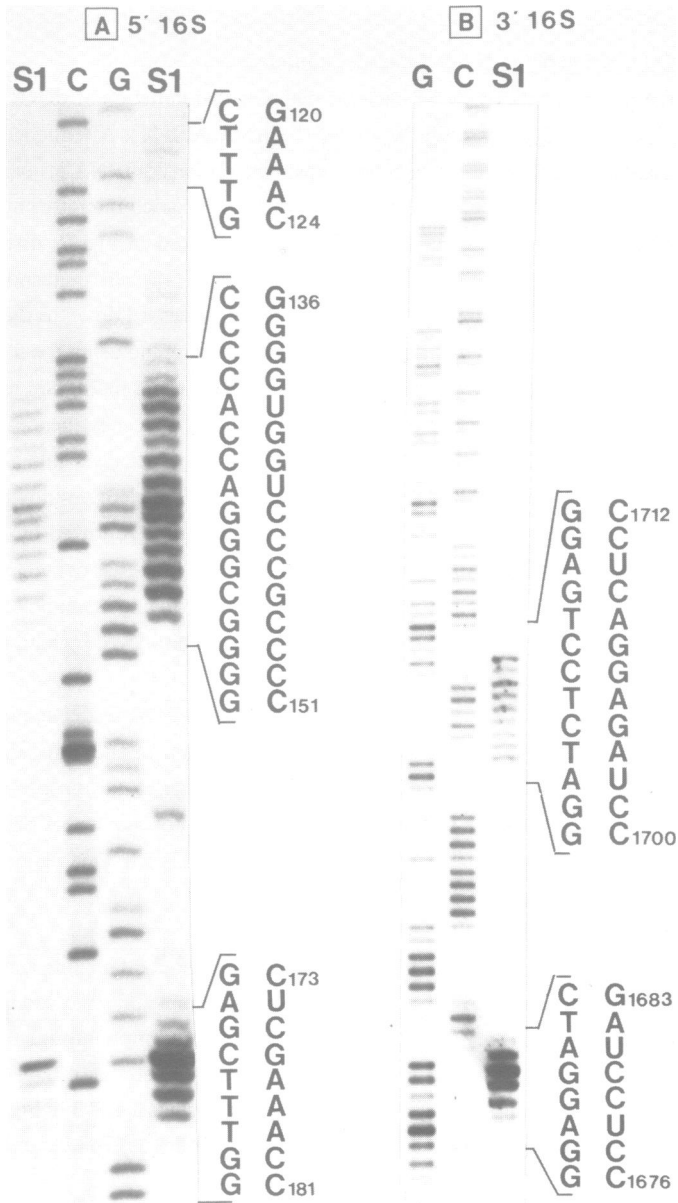
The present results reveal how the leader sequence is processed down to the mature 16S RNA. The processing sites were monitored by S1 nuclease mapping using the 5'-end labelled restriction fragment 1 (Fig. 1). The results are illustrated on the autoradiogram in Fig. 3A and presented on a putative secondary structure of the leader and spacer sequences in Fig. 4. A series of strong cuts were detected from position 139 (Fig. 3A), located within a 3-nucleotide bulge in the putative processing stem D, and extending about 9 nucleotides downstream (Fig. 4). This probably reflects an initial endonuclease cut at the bulge followed by exonuclease trimming but it could also result from a pyrimidine-specific endonuclease activity since the cuts occur in helix D where C-152 could migrate along the helix as an unpaired nucleotide (Fig. 4). Fairly strong processing cuts were also observed after A-122 and U-163 which lie in internal and terminal loops, respectively (Fig. 4).

Longer exposures revealed several weak bands in the 16S RNA leader sequence (Fig. 4). They probably result from unspecific degradation of the transcript and, thus, their exclusive occurrence in putative single-stranded regions lends support to the proposed secondary structure. Further support derives from the observation that all of the double helices in Fig. 4 are highly G-C-rich while the interhelical regions are A- and U-rich. One major 5'-end of the 16S RNA was detected (Figs. 3A and 4).

Processing of the 16S-23S RNA spacer

The spacer between the 16S and 23S RNA genes is only about 60 bp long and its sequence is complementary to those bordering both RNA genes; it is completely involved in the 16S and 23S RNA processing stems D and E (Fig. 4). Continuous transcription through the spacer was demonstrated by hybridizing total cellular RNA to the anti-sense strand of both 5'- and 3'-end labelled restriction fragments covering the entire spacer (fragments 2 and 3 in Fig. 1); the DNA fragments were protected against S1 nuclease digestion by continuous RNA transcripts (results not shown).

Evidence for partially protected DNA fragments was also observed and autoradiograms showing these cuts are given in Figs. 3B and C; the cuts are superimposed on the putative secondary structure in Fig. 4. 5'- and 3'-end labelled restriction fragments yielded different sets of data. 5'-end labelled fragments 2a and 2b (Fig. 1) revealed major cuts in the sequences



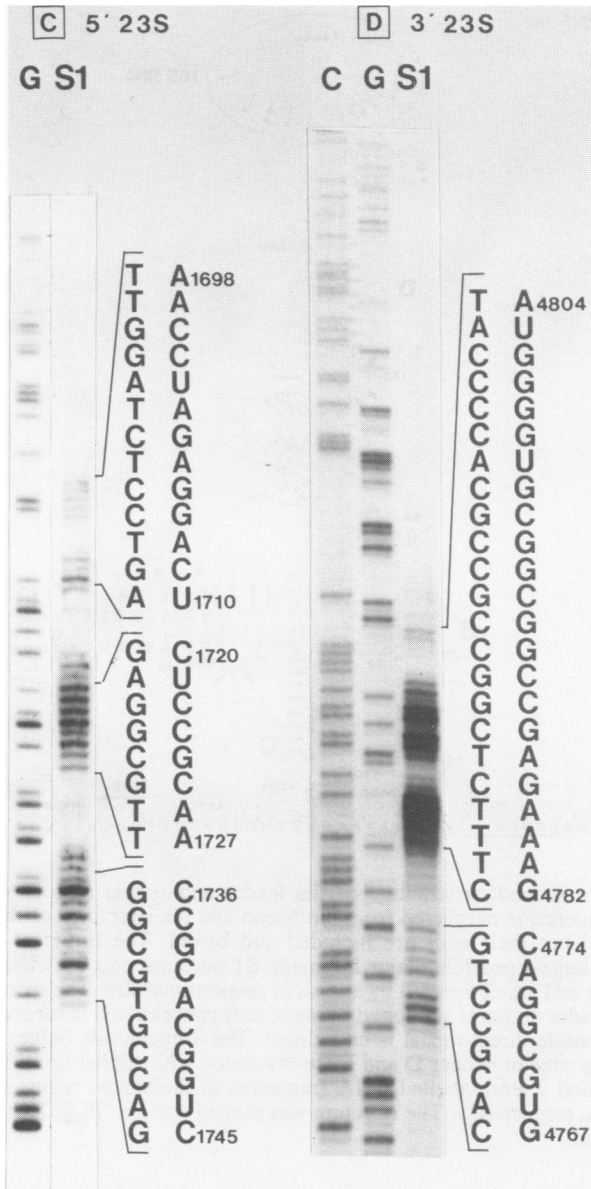


Figure 3: Autoradiograms showing examples of S1 nuclease mapping of the limits, and the processing sites, for the 16S and 23S RNAs. The limits of the 16S RNA were not determined experimentally earlier [17]. Experiments A to D were performed using fragments 1, 3b, 2a and 4, respectively, (see Fig. 1). The complementary RNA sequence is shown alongside the DNA sequence. C and G denote DNA sequencing tracks (Maxam-Gilbert). A correction was made for the one nucleotide difference beween the S1 nuclease tracks and the DNA sequencing tracks.

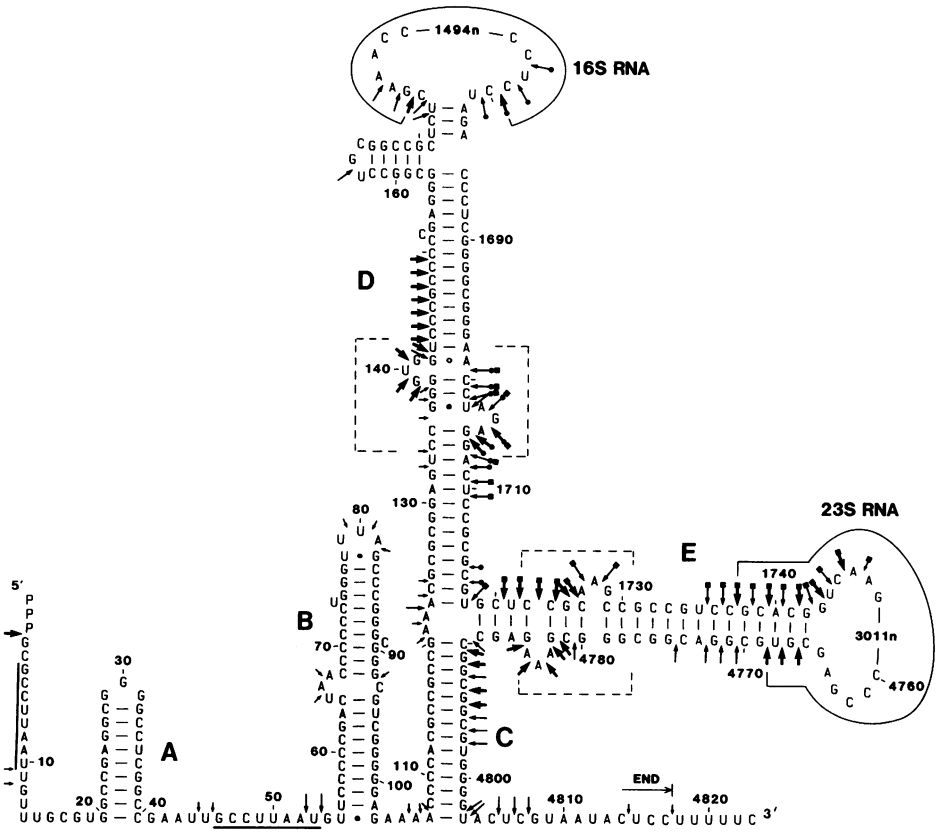


Figure 4: Putative secondary structure of the leader and spacer regions of the primary transcript. The sequence is numbered from the 5'-end and the structures of the 16S and 23S RNAs are omitted but their ends are included and boxed. The limits of the 23S RNA correspond to the largest possible major fragment. S1 nuclease cuts are classified as weak, medium and strong and are designated by arrows of proportional size. The secondary structure of the 16S RNA leader region is supported by weak cuts concentrated in unstructured regions, and an eight nucleotide direct repeat is underlined. The 3-nucleotide bulges that constitute primary processing sites in helices D and E are bracketed with stipled lines. Processing cuts detected with 5'- and 3'-end labelled DNA fragments in the spacer region are denoted by squares and circles, respectively. The structure was plotted using a "PLSTRUC" program (N. Larsen, unpublished).

C-1700 to G-1707, U-1721 to A-1728 and G-1738 to C-1745; these correspond to the 3-nucleotide loops in both 16S and 23S RNA processing stems and to the 5'-ends of mature 23S RNA; the heterogeneity at the 5'-end of 23S RNA was confirmed by primer extension experiments using reverse transcriptase which yielded essentially the same pattern of cuts (results not shown). 3'-end labelled fragments 3a and 3b (Fig. 1) revealed strong processing

cuts after the sequences C-1677 to C-1680 and A-1699 to A-1708 which correspond to the 3'-ends of mature 16S RNA (with a main end at C-1679) and the 3-nucleotide loop in processing stem D, respectively. These differences between the results obtained with 5'- and 3'-end labelled restriction fragments yield important information on the order of processing of the RNA transcript which is considered further in the Discussion.

Primary and secondary structure of the 23S rRNA

The 23S RNA contains approximately 3031 nucleotides and is larger than most archaebacterial 23S RNAs but shorter by about 46 nucleotides than that of *D. mobilis* [2]. The secondary structural model presented in Fig. 5 is based on a recent comparative sequence study of 23S-like RNA sequences from archaebacteria, eubacteria and the eukaryotic cytoplasm [2]. The format corresponds to that of the refined Santa Cruz/Urbana model for eubacteria [31]. Like the other extreme thermophile RNAs, the *T. tenax* RNA exhibits a very high G-C base pair content (79.5%) reflecting its stability at high temperature. The G·U content (9.9%) is slightly higher than the A-U content (8.9%), perhaps revealing the necessity to maintain unstable regions in the RNA structure that participate in cooperative effects or protein binding.

Most aspects of rRNA sequence and secondary structure that are typically archaebacterial [2] are present in the *T. tenax* structure and the secondary structural elements that are exclusive to archaebacteria are encircled in Fig. 5. The terminal helix joining the heterogeneous ends of the molecule is the maximum size (see Fig. 4) and may not form in the free RNA (see Discussion). There is an exceptional sequence in the region associated with peptidyl transfer in domain V where three highly conserved nucleotides are altered only in *T. tenax;* they are indicated by asterisks in Fig. 5. Two of these alterations have been shown to correlate with chloramphenicol resistance in mitochondrial ribosomes [32 and references therein].

Termination and processing of the transcript downstream from the 23S RNA

The termination site of the primary transcript and the processing sites distal to the 23S RNA were determined using the 750 bp restriction fragment 4 (Fig. 1). It contained about 120 bp of the 23S RNA gene and 630 bp downstream from the gene. It was hybridized to total cellular RNA and only very weak protection of the whole DNA fragment was observed; this was comparable to that observed in a control reaction containing total *E. coli* tRNA and probably resulted from DNA-DNA rehybridization. It was inferred, therefore, that the primary transcript terminates within this fragment. No single strong site was observed but a series of bands occurred in the sequences C-4790 to G-4797 and G-4802 to U-4809 downstream from the 23S RNA (see Figs. 3D and 4). The most downstream S1 nuclease cut occurred after C-4817 which lies in the long pyrimidine sequence: 5'-CTCC$_{4817}$TTTTTCTCCC-3'. There is no other T_5 sequence in the whole primary transcript. We could not infer from the S1 nuclease mapping results whether the heterogeneity at the 3'-end of the primary transcript reflects partially unspecific termination or exonuclease activity.

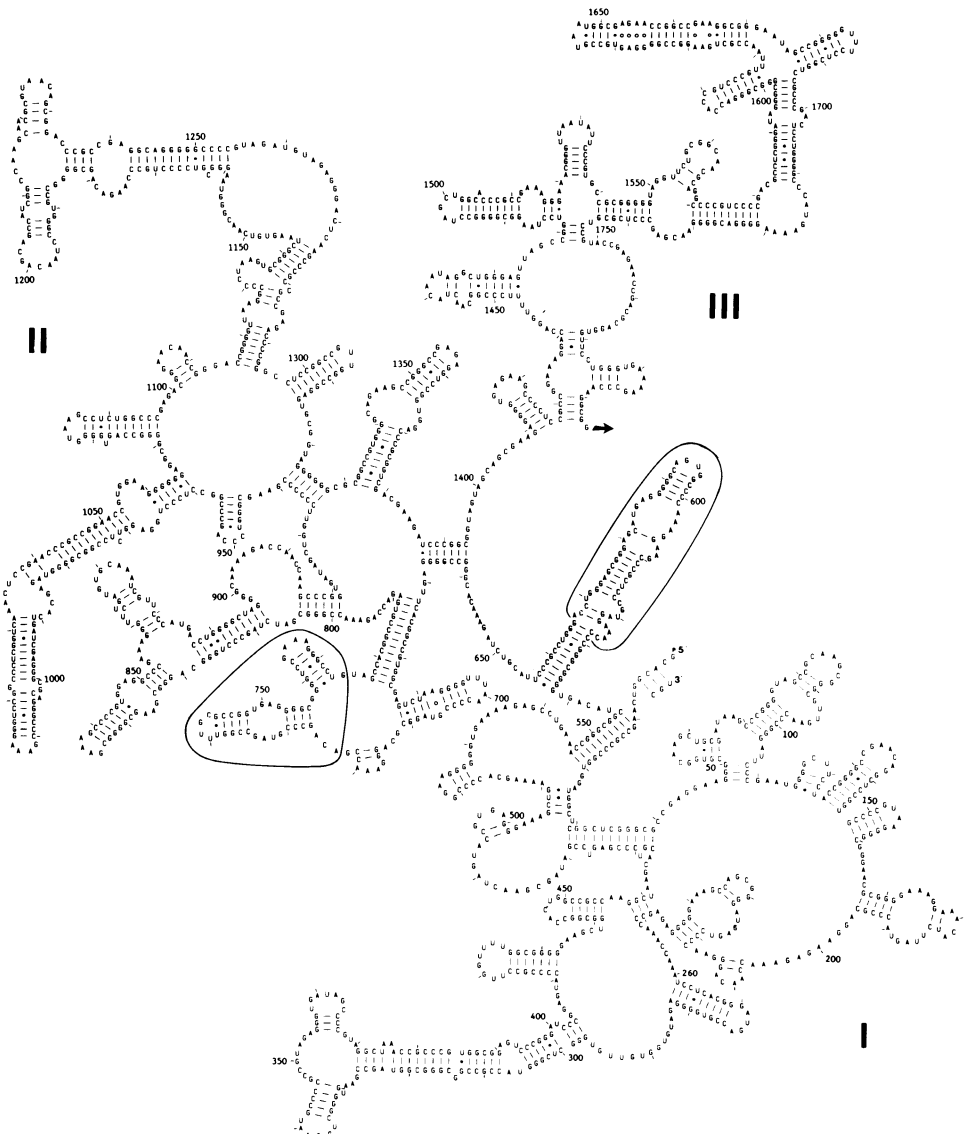
Several strong processing cuts occur between G-4782 and G-4786 which includes the

3-nucleotide loop in the 23S RNA processing stem E, in addition to the three strong cuts at C-4766 to U-4768 (Fig. 3D) which correspond to the 3'-ends of the mature 23S RNA.

DISCUSSION

Initiation signals: a promoter

A comparative sequence analysis of the 5'-flanking regions of the rRNA operon and four tRNA gene clusters from *T. tenax* revealed two common sequence features that could be



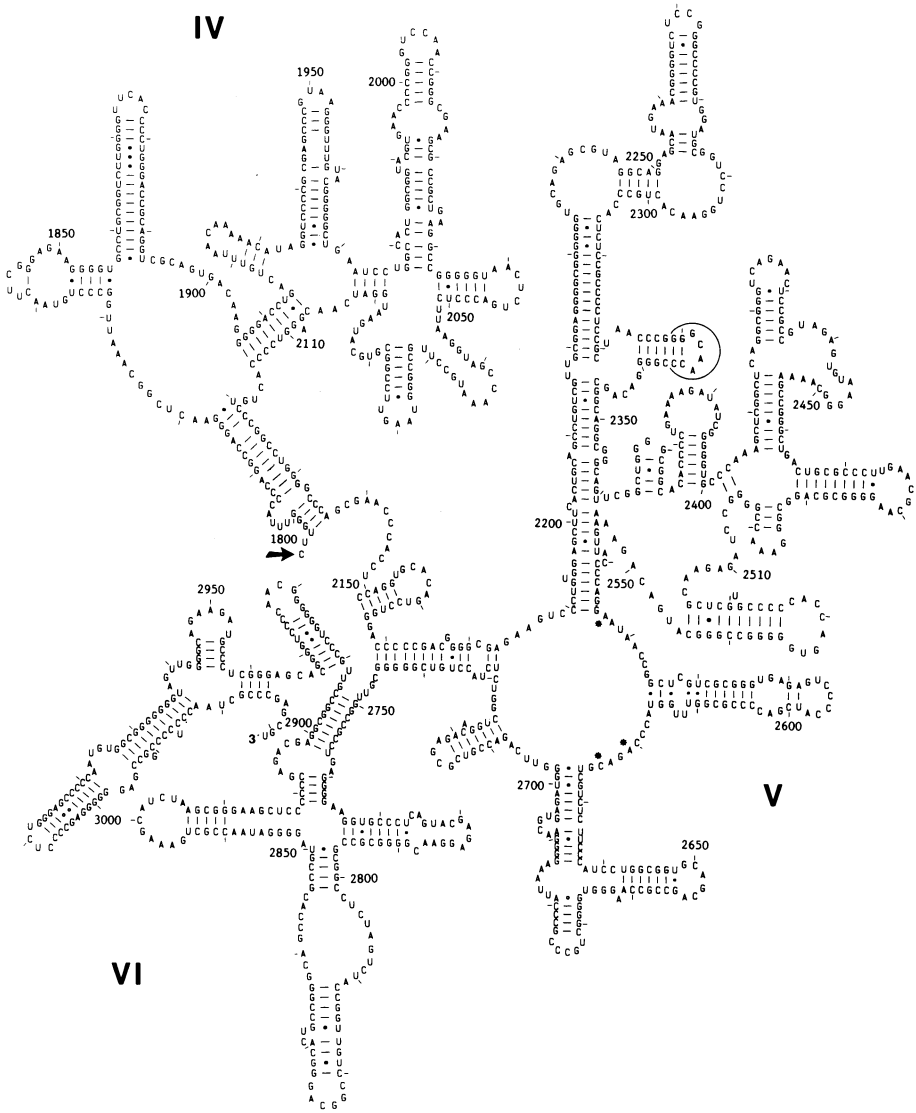


Figure 5: Putative secondary structure of the 23S RNA arranged as domains I to III (5'-half) and domains IV to VI (3'-half) based on a recently revised version of the Santa Cruz/Urbana model for *E. coli* [2]. Base pairs that are proven by sequence comparisons, or for which there is no negative evidence, are indicated by a line or dot between the nucleotides. Those which can form but for which there is conflicting evidence from sequence comparisons are juxtapositioned with no sign between them. Archaebacterial-specific features of the secondary structure are encircled. Three highly conserved nucleotides in the peptidyl transferase region of domain V that are only altered in *T. tenax* are denoted by asterisks. The structure was edited using an "EDSTRUC" program and plotted using a "PLSTRUC" program (N. Larsen, unpublished).

important for transcriptional initiation; one was at the site of transcription initiation and the other was upsteam and a putative promotor [1]. No homology was detected between the sequence at the initiation site and those found in *Methanococcus* and *Halobacteria* [11, 30]. Moreover, the conserved A-T-rich motif, and putative promoter, located about 18 bp upstream of the initiation site (Fig. 2) is absent from the corresponding sequences in other archaebacteria. This absence of common promoter and initiation motifs amongst the archaebacterial rRNA operons may relate to their adaptation to widely differing and extreme environments, and reflect the differing subunit structures of their RNA polymerases [33].

Secondary structure of the leader and spacer region

The putative secondary structure presented in Fig. 4 reveals an unusually high degree of secondary structure within the leader and spacer regions. Within the leader sequence the G-C-rich helices A and B are interspersed by A- and U-rich regions. Helix B has a counterpart in another extreme thermophile *D. mobilis*, and the base pairing scheme is supported by coordinated base changes (J.K., unpublished); such a helix would be a candidate for involvement in antitermination as has been found in eubacteria [34]. Other eubacterial and archaebacterial operons exhibit the long processing stems but none, so far, including the extreme thermophile *D. mobilis* [9] involve all of their spacer sequence in processing stems. This high level of structuring may yield some insight into the function of the 16S-23S RNA spacer. The structure presented offers two alternative coaxially stacked forms: either helix D on E or helix E on C (and the additional possibility of a functionally important exchange between them). Both forms will result in a large separation of the RNAs such that they can assemble without mutual interference and this may reflect the main function of all 16S-23S-like RNA spacer regions. Helix C could also act, simultaneously, as a termination signal (see below).

Processing signals and steps

It was proposed earlier that processing of the large rRNAs in *H. cutirubrum* and *M. vannielii* involved endonuclease cleavage of the primary transcript at or near the short discontinuities on opposite strands of the processing stems [12, 13, 15]. Both the processing stems and the short discontinuities are conserved in all archaebacterial RNA transcripts so far investigated [2 and references therein]. The processing sites may be comparable with the RNAse III recognition sites in eubacterial RNA precursors [35 and references therein]. Our results establish that a putative processing enzyme does, indeed, cut at the bulged loops that are staggered on opposite strands of both processing stems in *T. tenax* (Fig. 4). No conserved sequence motif is present at these sites but the purine content of the loops is generally high.

The S1 nuclease results suggest the following scheme of processing: Endonuclease digestion occurs at or near the 3-nucleotide loop on both strands of the 16S RNA processing stem. The 3'-end of the 16S RNA is then matured either by rapid exonuclease trimming, an endonuclease, or by a combination of both. The main 3'-end corresponds to C-1679 (Fig. 3B). This confirms an earlier estimate based on oligonucleotide analyses [36] and demonstrates that

the final cytidine of the putative Shine-Dalgarno sequence C-C-U-C-C is removed (Fig. 4). Processing of the 23S RNA appears to involve corresponding steps although here the heterogeneity of the mature ends suggests the involvement of an exonuclease. The high degree of heterogeneity contrasts with that found in both the 16S RNA of *T. tenax* and the 23S RNAs of other archaebacteria [2]. This result may be related to the exceptional fact that the terminal helix of the 23S RNA forms part of the processing stem in *T. tenax*. A consequence of these frayed ends, however, is that the terminal helix will exhibit a maximum of three base pairs (Fig. 5) and generally less (Fig. 4). This indicates, as for the 23S RNAs of *D. mobilis* and *M. thermoautotrophicum*, that it will generally be absent in the mature RNA [2].

By using both 5'- and 3'-end labelled DNA fragments covering the 16S-23S RNA spacer we were able to deduce the *order* of processing cuts. While cuts in the bulged loop on the 3'-side of the 16S RNA processing stem were detected using both 5'- and 3'-end labelled fragments those in the bulged loop of the 23S RNA processing stem were only detected using 5'-end labelled fragments (Fig. 4). This demonstrates that processing of the 16S RNA must occur prior to that of the 23S RNA. The S1 nuclease cuts obtained using 5'-end labelled fragments reveal, by a similar rationale, that processing of the 16S RNA occurs at the bulged loop prior to final maturation.

Finally, the S1 nuclease results (Fig. 3) show that many of the processing intermediates occur in high yield relative to the mature RNAs. To an unknown degree, the method gives an overestimate of these yields because DNA fragments preferentially select for larger RNA molecules. It is also likely, however, that the processing enzymes are less active at very high temperatures.

Termination signal: a terminator

The longest detectable transcript terminates within a 13-pyrimidine sequence immediately

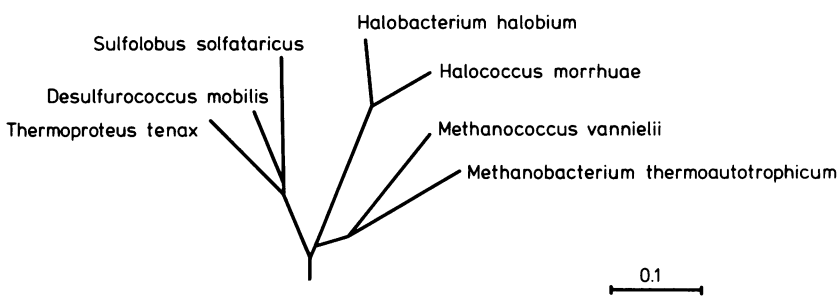


Figure 6: Unrooted phylogenetic tree for the archaebacteria derived from 23S ribosomal DNA sequences as described in Materials and Methods. The*S. solfataricus* 23S RNA sequence (C.R. Woese, pers. commun.) is included for comparison. The distance measure corresponds to 0.1 mutational events per sequence position.

downstream from helix C (Fig. 4). Other fragments resistant to S1 nuclease were detected with ends within this helix.

Our knowledge of other archaebacterial termination signals is limited. There is some similarity with a possible terminator in the rRNA operon of *D. mobilis* which consists of a T-rich sequence immediately following the 23S RNA processing stem [9]. On the other hand, although the primary transcripts of rRNA operons from both *M. vannielii* and *M. thermoautotrophicum* terminate within T-rich sequences they are followed by a short "hairpin" structure [16, 30] no such structure follows the T-rich sequence in *T. tenax*.

23S RNA structure and phylogenetic implications

The 23S RNA sequence was aligned with other available 23S-like RNA sequences on the basis of both conserved sequence regions and secondary structural elements as described by Leffers *et al.* [2]. Homology values and K_{nuc} values (the average number of mutational events per sequence position) were calculated. The latter give only an approximate measure of evolutionary distance [27] because variations occur in the evolutionary clocks (or mutational rates) of different species [37].

A phylogenetic tree was calculated from the K_{nuc} values for the archaebacteria and was rooted using a eubacterial 23S RNA (see Figure 6). It resembles the tree obtained with 16S RNA sequences [4] and demonstrates the three main subkingdoms of the archaebacteria. A similar tree was obtained by considering only the unstructured regions of the 23S-like RNAs (see Fig. 5) thus avoiding artefacts that could result from the high G-C content of the extreme thermophiles. As discussed earlier [38] such trees need to be tested by comparison with those based on other universal macromolecular sequences.

ACKNOWLEDGEMENTS

We thank Niels Larsen for his unstinting assistance with the computing and Heidi Hummel for her advice on culturing *T. tenax* cells. J.K. and the research were supported by the Danish Natural Science Research Council. J.K. also received an EMBO travel fellowship. H.L. was granted a candidate stipend from the Carlsberg Foundation. Lisbeth Heilesen is thanked for helping with the manuscript.

*To whom correspondence should be addressed

REFERENCES

1. Wich, G., Leinfelder, W. and Böck, A. (1987) EMBO J. 6, 523-528.
2. Leffers, H., Kjems, J., Østergaard, L., Larsen, N. and Garrett, R.A. (1987) J. Mol. Biol. 194, in press.
3. Fox, G.E., Stackebrandt, E., Hespell, R.B., Gibson, J., Maniloff, J., Dyer, T.A., Wolfe, R.S., Balch, W.E., Tanner, R., Magrum, L., Zablen, L.B., Blakemore, R., Gupta, R., Bonen, L., Lewis, B.J., Stahl, D.A., Luehrsen, K.R. Chen, K.N. and Woese, C.R. (1980) Science 209, 457-463.

4. Woese, C.R. and Olsen, G.J. (1986). System. Appl. Microbiol. 7, 161-177.
5. Zillig, W., Schnabel, R. and Stetter, K.O. (1985) Curr. Top. Microbiol. and Immunol. 33, 1-18.
6. Lake, J.A., Hendersen, E., Oakes, M. and Clark, M.W., (1984) Proc. Natl. Acad. U.S.A. 81, 3786-3790.
7. Lake, J.A. (1986) Nature 319, 626.
8. Neumann, H., Gierl, A., Tu, J., Leibrock, J., Staiger, D. and Zillig, W. (1983) Mol. Gen. Genet. 192, 66-72.
9. Larsen, N., Leffers, H., Kjems, J. and Garrett, R.A. (1986) System. Appl. Microbiol. 7, 49-57.
10. Kjems, J. and Garrett, R.A. (1985) Nature 318, 675-677.
11. Mankin, A.S., Teterina, N.L., Rubtsov, P.M., Baratova, L.A. and Kagramanova, V.K. (1984) Nucl. Acids Res. 12, 6537-6546.
12. Hui, I. and Dennis, P. (1985) J. Biol. Chem. 260, 899-906.
13. Chant, J. and Dennis, P. (1986) EMBO J. 5, 1091-1097.
14. Mankin, A.S. and Kagramanova, V.K. (1986) Mol. Gen. Genet. 202, 152-161.
15. Jarsch, M. and Böck, A. (1985) Mol. Gen. Genet. 200, 305-312.
16. Østergaard, L., Larsen, N., Leffers, H., Kjems, J. and Garrett, R.A. (1987). System. Appl. Microbiol. in press.
17. Leinfelder, W., Jarsch, M. and Böck, A. (1985) System. Appl. Microbiol. 6, 164-175.
18. Zillig, W., Stetter, K.O., Schäfer, W., Janekovic, D., Wunderl, S., Holz, I. and Palm, P. (1981) Zbl. Bakt. Hyg. I. Abt. Orig. C2, 205-227.
19. Aiba, H., Adhya, S. and de Crombrugghe, B. (1981) J. Biol. Chem. 256, 11905-11910.
20. Lechner, K., Wich, G. and Böck, A. (1985) System. Appl. Microbiol. 6, 164-170.
21. Norrander, J., Kempe, T. and Messing, J. (1983) Gene 26, 101-114.
22. Sanger, F., Nicklen, S. and Coulsen, A.R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463-5467.
23. Ansorge, W. and Labeit, S. (1984) J. Biochem. Biophys. Methods 9, 33-47.
24. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A laboratory manual. Cold Spring Harbor Press, N.Y.
25. Favaloro, J., Treisman, R. and Kamen, R. (1980) Methods in Enzymol. 65, 718-723.
26. Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 560-564.
27. Hori, H. and Osawa, S. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 381-385.
28. Felsenstein, J. (1978) J. Syst. Zool. 27, 401-410.
29. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
30. Wich, G., Hummel, H., Jarsch, M., Bär, U and Böck, A. (1986) Nucl. Acids Res. 14, 2459-2479.
31. Noller, H.F. (1984) Ann. Rev. Biochem. 53, 119-162.
32. Ettayebi, M., Prasad, S.M. and Morgan, E.A. (1985) J. Bact. 162, 551-557.
33. Gropp, F., Reiter, W.D., Sentenac, A., Zillig, W., Schnabel, R., Thomm, M. and Stetter, K.O. (1986) System. Appl. Microbiol. 7, 95-101.
34. Li, S.C., Squires, C.L. and Squires, C. (1984) Cell 38, 851-860.
35. Brosius, J., Dull, T.J., Sleeter, D.D. and Noller, H.F. (1981) J. Mol. Biol. 148, 107-127.
36. Woese, C.R., Gupta, R., Hahn, Ch..M., Zillig, W. and Tu, J. (1984) Syst. Appl. Microbiol. 5, 97-105.
37. Britten, R.J. (1986) Science 231, 1393-1398.
38. Woese, C.R., Pace, N.R. and Olsen, G.J. (1986) Nature 320, 401-402.