© 1990 Oxford University Press

Organization and expression of the 16S, 23S and 5S ribosomal RNA genes from the archaebacterium *Thermoplasma acidophilum*

Heesoo K.Ree and Robert A.Zimmermann*
Department of Biochemistry and Program in Molecular and Cellular Biology, University of Massachusetts, Amherst, MA 01003, USA

Received April 12, 1990; Revised and Accepted July 3, 1990 Genbank accession nos M20822, M32297, M32298

ABSTRACT

To elucidate the organization of the transcription units encoding the 16S, 23S and 5S rRNAs in the archaebacterium *Thermoplasma acidophilum*, the nucleotide sequences flanking the three rRNA genes were determined, and the 5′ and 3′ termini of the rRNA transcripts were mapped by primer extension and nuclease S1 protection. The results show that each of the rRNAs is transcribed separately, consistent with the lack of physical proximity among them in the *T. acidophilum* genome. The transcription initiation sites are preceded at an interval of approximately 25 base pairs by conserved A+T-rich sequences of the form CTTATATA, which strongly resemble the archaebacterial promoter consensus, TTTA$^{T}/_{A}$ATA. In all three cases, transcription termination occurs within T-rich tracts just downstream from inverted repeats which can be folded into relatively stable stem-loop structures. While no partially processed intermediates of the 16S or 5S rRNA transcripts were detected, the 23S rRNA transcript appears to be processed by a RNase III-like activity prior to final maturation. This is the only organism known in the prokaryotic world in which the 16S, 23S and 5S rRNAs are all expressed from separate transcription units.

INTRODUCTION

Thermoplasma acidophilum, a wall-less aerobic heterotroph which grows optimally at 60°C and pH 1−2, is one of the more unusual members of the archaebacterial kingdom, a group of organisms that share an evolutionary history distinct from eubacteria and eukaryotes (1,2). Its phylogenetic position among the archaebacteria is not entirely clear as it exhibits features characteristic of both main subdivisions, the methanogens and halophiles, on the one hand, and the sulfur-dependent thermoacidophiles, on the other. Analysis of the complete nucleotide sequence of *T. acidophilum* 16S rRNA, for instance, suggests that this species diverged deep within the methanogen/halophile lineage at about the same time that the methanogens and halophiles branched from one another (3−5). In contrast, a closer relationship with the sulfur-dependent thermoacidophiles has been inferred from the electrophoretic properties, immunological cross-reactivity and gene organization of its DNA-dependent RNA polymerase (6−8). The apparent lack of physical linkage among the single 16S, 23S and 5S rRNA genes of *T. acidophilum* (9) raises further uncertainties about the phylogenetic status of this organism since there is no other example of such an arrangement in the prokaryotic world. In methanogens and halophiles, as in nearly all eubacteria, the three rRNAs are co-transcribed from a polycistronic operon (10). In some thermoacidophilic archaebacteria, however, the 5S rRNA genes are transcribed independently of the 16S−23S rRNA operon while in at least two eubacteria, the 16S and 23S−5S rRNA genes comprise separate transcription units (10−12).

There is only a limited amount of information available on the transcriptional control signals utilized by archaebacteria (see ref. 10 for review). From studies of both stable RNA genes and protein-coding sequences, it has been suggested that archaebacterial promoters are composed of A+T-rich segments, centered approximately 25 nucleotides upstream from the site of transcription initiation (13−15). A consensus for transcription termination signals in archaebacteria has not yet been established. Transcription of some viral genes and a majority of the rRNA operons investigated thus far terminate at or near oligo(T) or oligopyrimidine tracts which are not in general preceded by inverted repeats, while a number of protein-coding genes contain sequences just upstream from the transcription termination sites that could fold into stable stem-loop structures (10). Most archaebacterial rRNA operons are expressed as long precursors which are later processed into individual rRNA molecules. Processing in these cases is thought to occur much as in eubacteria (16), since in a number of species the sequences immediately flanking the 16S and 23S rRNAs are capable of forming relatively stable base-paired stems in the middle of which are staggered bulge loops that may serve as recognition signals for processing enzymes (13,17−19).

To investigate the proposition that the 16S, 23S and 5S rRNA genes in *T. acidophilum* are separately transcribed, we have determined the nucleotide sequences flanking all three. These results, together with data from primer-extension and nuclease

* To whom correspondence should be addressed

S1 mapping, have revealed the presence of likely promoter and terminator sequences associated with each of the genes. We also discuss the possibility that the three rRNA transcripts are processed by different pathways and compare the transcriptional control and processing signals in *T. acidophilum* with those in other archaebacteria.

MATERIALS AND METHODS

Materials

Restriction endonucleases were purchased from Boehringer-Mannheim, International Biotechnologies, Inc., New England Biolabs or Promega Biotec, and used according to the manufacturers' instructions. *Escherichia coli* DNA polymerase I (large fragment) was obtained from New England Biolabs, modified T7 DNA polymerase (Sequenase) was from United States Biochemicals and AMV reverse transcriptase was from Life Sciences, Inc. T4 polynucleotide kinase and nuclease S1 were supplied by Boehringer-Mannheim. [α-^{32}P]dATP (800 Ci/mmol), [α-^{35}S]dATPαS (800 Ci/mmol) and [γ-^{32}P]ATP (6000 Ci/mmol) were the products of New England Nuclear, Amersham and ICN Biomedicals, respectively. The M13(−20) primer used in sequencing bacteriophage M13 recombinants and in certain mapping experiments was purchased from New England Biolabs, while specific oligodeoxynucleotides for primer-extension and nuclease S1 protection experiments were synthesized by Genetic Designs, Inc. *T. acidophilum* 122-1B2 cells were kindly provided by Dr. D. G. Searcy (University of Massachusetts, Amherst).

Plasmids and bacteriophages

The following vectors were used in establishing the sequence of the regions flanking the *T. acidophilum* rRNA genes. Bacteriophage mXX2-9, which contains a 1.5-kb *Xcy*I-*Xho*I fragment spanning the 5′ end of the 16S rRNA gene, and plasmid pTH2-2, which contains a 2.3-kb *Hin*dIII fragment covering the 3′ end of the 16S rRNA gene have been described previously (3). Plasmid pL8, a pBR322 derivative which harbors a 5.1-kb *Bam*HI-*Hin*dIII fragment encompassing the entire 23S rRNA gene and its flanking regions, was kindly provided by Drs. W.-D. Reiter and W. Zillig (Max-Planck-Institut für Biochemie, Martinsried, F.R.G.). Plasmid pTH1-1 was constructed by inserting a 2.7-kb *Hin*dIII fragment spanning the 3′ end of the 23S rRNA gene and its flanking region into pUC8. Plasmid pTH5PH23 carries a 0.7-kb *Pst*I-*Hin*dIII fragment covering the 5S rRNA gene and its flanking regions inserted into the corresponding sites of pUC18. To facilitate both sequencing and transcript mapping, subfragments from the above vectors were introduced into the polylinker regions of various bacteriophage M13 derivatives by the use of appropriate restriction-enzyme digests or by the deletion-subcloning method (20). Subclones were constructed in such a way that sequences could be read in both orientations. *E. coli* strains JM101 (21) and XL1-B (22) served as hosts for transformation and transfection throughout these experiments. Routine recombinant DNA procedures were carried out as described in ref. 23. DNA sequence analysis was performed by the dideoxynucleotide chain termination method (24) using either [α-^{32}P]dATP or [α-^{35}S]dATPαS in conjunction with DNA polymerase I or Sequenase (25).

Isolation of cellular RNA

T. acidophilum was grown for 1 or 2 days at 59°C in Darland's medium (26). After the culture was cooled to room temperature, cells were harvested by sedimentation at 4000 *g* for 2 min and resuspended in 0.02 M Na(OAc), pH 5.5, 0.5% SDS and 1 mM EDTA. An equal volume of phenol saturated with 0.02 M Na(OAc) was added to the cell suspension and the mixture was incubated at 60°C for 5 min with gentle shaking (27). The aqueous phase was re-extracted with phenol, and the RNA was precipitated several times with ethanol, dissolved in water and stored at −20°C.

Primer extension and nuclease S1 mapping

The sequences of the primers used for 5′-end mapping, as well as the coordinates of the complementary segments within the *T. acidophilum* rRNA genes, are listed below:

N16-1 d(TACTGACTCACTACTCC); −70 to −54, flanking 5′ end of 16S rRNA
N16-2 d(CCTGATAGCAGTGACCG); +21 to +37, within mature 16S rRNA
N23-1 d(AAGGTGCATTGTAAGCG); −89 to −73, flanking 5′ end of 23S rRNA
N23-2 d(ATCGCAGCTTGGCACGTC); +40 to +66, within mature 23S rRNA
N5 d(CCTCGCGGCATACAGTA); +77 to +93, within mature 5S rRNA

The M13(−20) primer was employed in all 3′-end mapping experiments.

To map the 5′ termini of rRNA transcripts by primer extension (28), 0.1 pmol of the appropriate 5′-^{32}P-labeled oligonucleotide was annealed to 10 μg of total *T. acidophilum* RNA in 10 μl of

```
16S rRNA

   -200   GAAAGTTATA TATACTGATT TGCTATTCTT TACTTTGCAC ATAACAGGTA
   -150   TGGATGCCAA CCTACCTACG AAAGAGATGA ATTACGCCTG TGCGCGTTTA
   -100   TGTGTCTCTT CTTCCAGGGT TGGTAAATAT GGAGTAGTGA GTCAGTAATG
    - 50  CTCACTTCTG ACAGGTAGAT GCCAGCCGCT AATTCCAATT GAGTGTTTTA

23S rRNA

   -200   AGTTAATTAA AATATATTTC AAATTCTGAA AAGTGAGTAT TATAAAATTA
   -150   TTGGATAATT GGGCTTCCGG ATCAAAATGC TTATATCCCT CTTAATGATA
   -100   TAGTCCATAC ACGCTTACAA TGCACCTTTT AATGCATCAC ATTAAAAGGG
    - 50  AAGATGTCGA ATTTTGATGA GACTAGATGA GAGATGGCAC AATGGATTGT

5S rRNA

   -200   CATGGGAAAA TGGGCATCTG CTGGAGACGA CTATGTACTT CATGCATCCA
   -150   TATTATCTTA TGATATTAAA ATGCTAAAGC GCATGATCGC TTCGGAAGGA
   -100   ATATCGACAG GGCGAGATTC GCTGTCTCAT AATCGCCCGG ATCTGCGCTG
    - 50  AAGCCGTAAA ATCACGAAAA TCTTATATAG ATGTGTTCTA TATAGTGTTC

                              (a)


16S rRNA

   1472   AAGTTAGAAC ACTGGCTGGC ATCTGCCCAG ACCATACCTT TAAATTTCCT
   1522   TTTTTTATAA TGTTTTGTTT CGGTGATTTG GTGTCAATTG AAGATAAGGC

23S rRNA

   2907   TCCATGCTAA ATTCTCGTCT AGTTGAAGTC ATAAGCGTGT TTGATTTTTG
   2957   TTTTTTATTG TTTTTGCCGT TGTAGTGTGA ACTATGATGA AAAATTATAT

5S rRNA

    124   TTTGAAATGA AAGTTTTTTA CATCTATCTT TAATTCAGTG GCAATCATTA
    174   ATCCTTGGTC ATCATCTGAT TTCTTCGACT ATGAGAGGCT AAAGAGAGGA

                              (b)
```

Figure 1. Nucleotide sequences flanking the 5′ and 3′ ends of the *T. acidophilum* rRNA genes. (a) 5′-flanking regions from the 16S, 23S and 5S rRNA genes. Putative promoter sequences are in italics and transcription initiation sites are underlined. (b) 3′-flanking regions from the 16S, 23S and 5S rRNA genes. Transcription termination sites are underlined. In the numbering system used here, +1 represents the first base of the mature rRNA sequence.

10 mM Tris-HCl, pH 7.5, and 40 mM KCl by heating at 95°C for 1 min and slowly cooling to room temperature. For cDNA synthesis, 20 μl of a buffer containing 70 mM Tris-HCl, pH 7.5, 12 mM $MgCl_2$, 7.5 mM DTT, 0.75 mM each of the four dNTPs and 75 μg/ml actinomycin D were added, and the mixture was incubated with 60 units of AMV reverse transcriptase at 42°C for 20 min. Following extraction with phenol, DNA was precipitated from the aqueous phase with ethanol.

Both the 5′ and 3′ termini of *T. acidophilum* rRNA transcripts were mapped by nuclease S1 protection using single-stranded DNA probes (23). To map the 5′ termini, probes were synthesized by extending 5′-^{32}P-labeled, rRNA-specific primers on recombinant M13 templates with DNA polymerase I (15). For mapping the 3′ termini, uniformly labeled probes were generated by annealing the M13(−20) primer to appropriate M13 recombinants and extending them with DNA polymerase I in the presence of ^{32}P-labeled dATP (29). In each case, the partially double-stranded product was cleaved with a restriction endonuclease to define the 3′ end of the probe, and labeled, single-stranded DNA was isolated by electrophoresis on a denaturing polyacrylamide gel (23). From 10^4 to 10^5 cpm of the probe were mixed with 5 μg of total cellular RNA in 10 μl of S1 hydridization buffer (40 mM PIPES, pH 6.5, 0.4 M NaCl, 1 mM EDTA and 80% formamide), and incubated at 85°C for 10 min; duplicate samples were then annealed for 3 hr at two different temperatures. After hybridization, 100 μl of S1 digestion buffer (50 mM Na(OAc), pH 4.6, 280 mM NaCl, 4.5 mM $ZnSO_4$ and 20 μg/ml denatured calf thymus DNA) containing 100 units of nuclease S1 were added to each sample. DNA:RNA hybrids were digested at 37°C for 30 min, extracted with phenol and precipitated with ethanol.

Extended cDNA (primer extension) or protected DNA probes

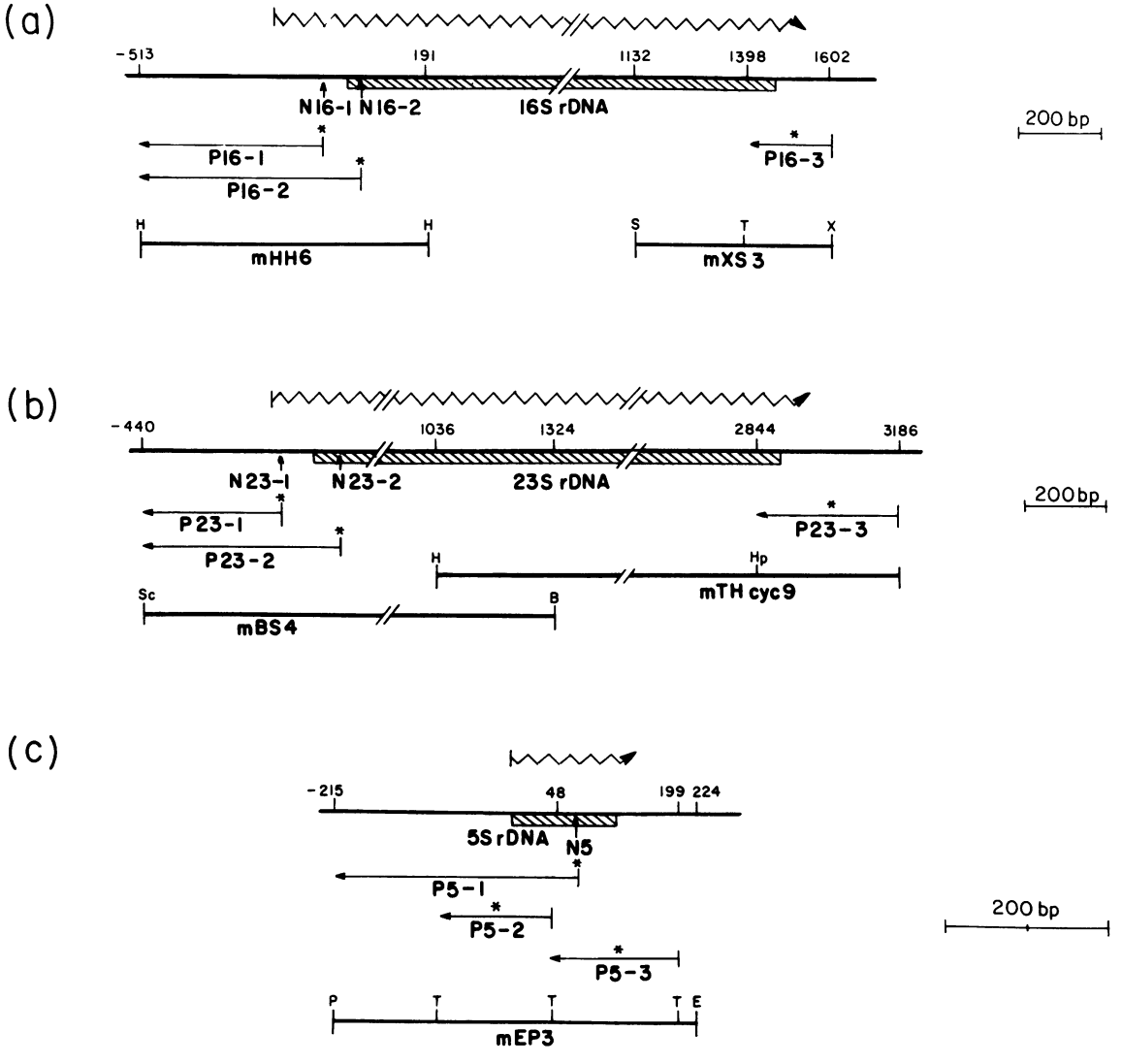

Figure 2. Mapping strategy for *T. acidophilum* rRNA transcripts. Sequences encoding the mature rRNAs are indicated by cross-hatched boxes and rRNA transcripts are represented by saw-toothed lines. (a) Oligonucleotides N16-1 and N16-2 were used for primer extension and for the synthesis of probes P16-1 and P16-2 from phage M13mHH6 templates. Probe 16-3 was synthesized from a phage M13mXS3 template using the M13(−20) primer. (b) Oligonucleotides N23-1 and N23-2 were employed for primer extension and for the synthesis of probes P23-1 and P23-2 from phage M13mBS4 templates. Phage M13mTHcyc9 was used to generate probe P23-3 with the aid of the M13(−20) primer. (c) Oligonucleotide N5 was used to synthesize probe P5-1, and the M13(−20) primer to synthesize probes P5-2 and P5-3, from phage M13mEP3 templates. Probes P5-2 and P5-3 were cleaved from the resulting double-stranded DNA by digestion with *TaqI* and purified by gel electrophoresis. P16-1, P16-2, P23-1, P23-2 and P5-1 were 5′-^{32}P-labeled, while P16-3, P23-3, P5-2 and P5-3 were uniformly ^{32}P-labeled. Restriction sites are designated as follows: *Hin*dIII, H; *Sal*I, S; *Taq*I, T; *Xmn*I, X; *Sac*I, Sc; *Bam*HI, B; *Hpa*II, Hp; *Pst*I, P; *Eco*RI, E.

(S1 mapping) were resolved by electrophoresis on 6–8% polyacrylamide gels containing 8 M urea. To locate the 5′ ends of the transcripts, sequence ladders were prepared from the same primers and the same M13 templates that had been used to synthesize the DNA probes and electrophoresed alongside the sample lanes. For 3′-end mapping, ladders produced from unrelated sequences were used to determine the lengths of the protected DNA probes.

RESULTS

Sequences flanking the 16S, 23S and 5S rRNA genes

The nucleotide sequences of the regions flanking the 5′ and 3′ termini of the 16S, 23S and 5S rRNA genes of *T. acidophilum* were determined by the dideoxynucleotide chain termination method (Fig. 1). The regions upstream and downstream from all three rRNA genes are far more A+T-rich than either the total genomic DNA of this organism (30) or the rRNA coding segments themselves (3,31). Although there are no sequences longer than six bases that are exactly conserved among the flanking regions, several features indicate that the 16S, 23S and 5S rRNA genes are encoded in separate transcription units. As illustrated in Figure 1(a), for instance, there are sequences within the upstream regions which exhibit the repeated TA motif found in many archaebacterial promoters (10), while the downstream regions, depicted in Figure 1(b), all include T-rich segments which characterize at least one class of archaebacterial terminators (10).

Mapping the 5′ and 3′ termini of the rRNA transcripts

To assess further the significance of the putative promoter and terminator sequences which adjoin each of the three rRNA genes, the 5′ and 3′ termini of cellular rRNA transcripts were mapped by primer extension and nuclease S1 protection according to the strategies outlined in Figure 2. Primers used for the synthesis of cDNA complemetary to the 5′ termini were designed to hybridize both upstream and downstream from the mature 5′ ends of the rRNAs, and single-stranded DNA probes for S1 mapping

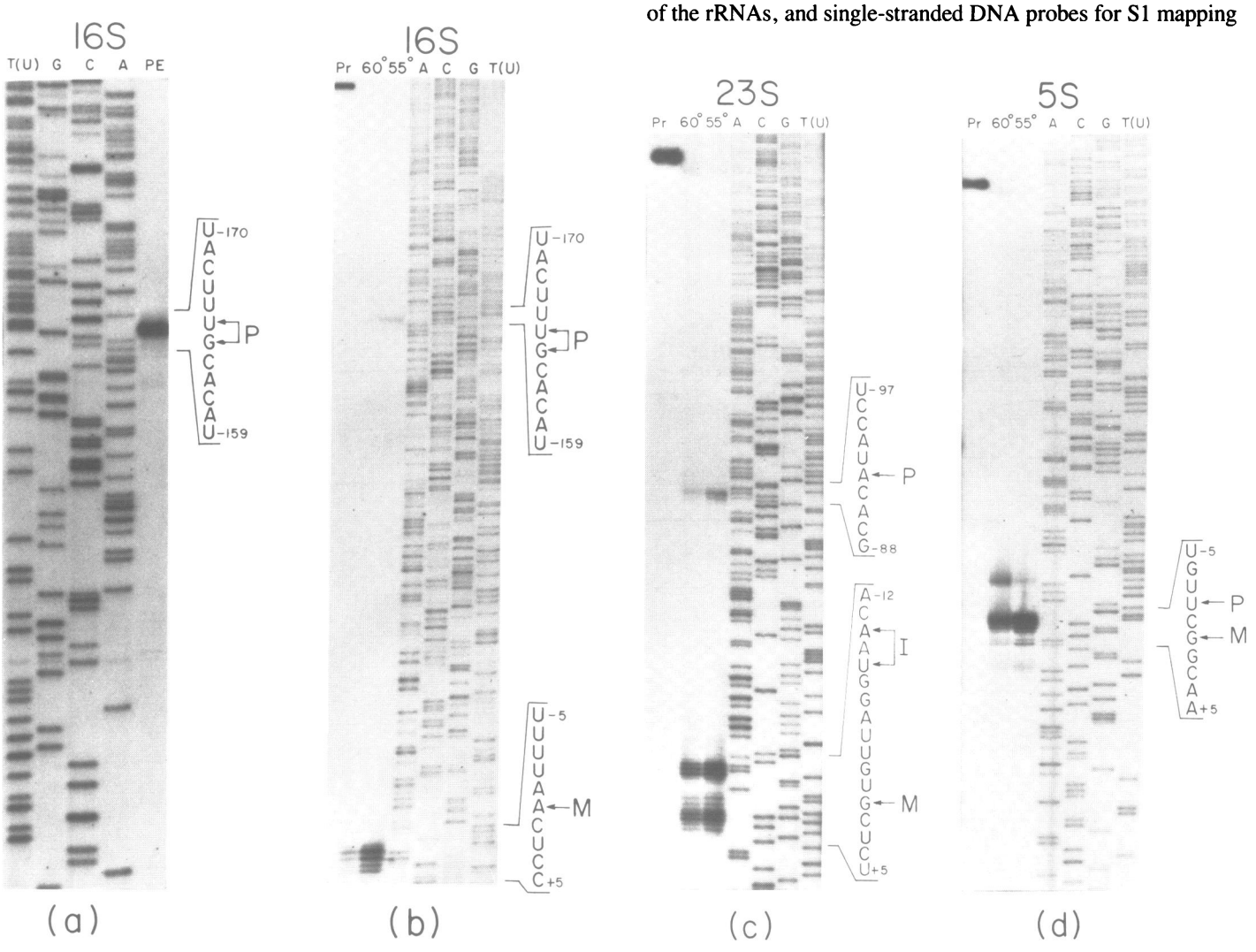

Figure 3. 5′-end mapping of *T. acidophilum* rRNA transcripts by primer extension and nuclease S1 protection. (a) Analysis of 16S rRNA transcripts by extension of oligonucleotide N16-1. T(U),G,C,A: DNA sequencing tracks obtained using N16-1 as primer; PE: primer-extension track. (b) Nuclease S1 mapping of 16S rRNA transcripts with probe P16-2. (c) Nuclease S1 mapping of 23S rRNA transcripts with probe P23-2. (d) Nuclease S1 mapping of 5S rRNA transcripts with probe P5-1. In (b)–(d), the 3′ ends of the probes were defined by the unique *Hin*dIII site in the polylinker of the M13 vectors. Pr: undigested probes; 60° and 55°: nuclease S1-digested probes from DNA-RNA hybrids annealed at 60° and 55°C, respectively; A,G,C,T(U): DNA sequencing tracks obtained using oligonucleotides N16-2(b), N23-2(c) and N5(d) as primers. Nucleotide position numbers correspond to those in Fig. 1(a). P,I,M refer to precursor, partially processed and mature forms of the rRNA molecules.

were extended from them. Probes for mapping the 3′ ends by nuclease S1 were synthesized from a common site in the vector, immediately adjacent to the cloned rRNA genes.

Transcription initiation sites

16S rRNA

Primer-extension analysis of transcripts encompassing the 16S rRNA was carried out with oligonucleotides N16-1 and N16-2 (Fig. 2(a)). The autoradiogram illustrated in Figure 3(a) indicates that transcription of the 16S rRNA gene begins at positions −165 and −164. Similar results were obtained by nuclease S1 mapping, using probes P16-1 and P16-2 (Fig. 2(a)). As shown in Figure 3(b), there are two bands, at positions −165 and −164, that represent the transcription initiation site(s), as well as a cluster of bands around position +1 that mark the 5′ end of mature 16S rRNA. No processing intermediates, whose 5′ ends fall between the initiation site(s) and the beginning of the 16S rRNA coding sequence, were detected by either method, even after prolonged exposure of the autoradiograms.

23S rRNA

The 5′ ends of the 23S rRNA transcripts could not be located conclusively by primer extension with oligonucleotides N23-1 and N23-2 (Fig. 2(b)), as the former hybridized too close to the site of transcription initiation and the latter primed cDNA synthesis at multiple sites within the cellular RNA. However, S1 mapping with probe P23-2 revealed that transcription of the 23S rRNA gene most likely initiates at position −92 (Fig. 3(c)). No significant bands upstream from this site were detected, even though there are several A+T-rich segments in this region that resemble known archaebacterial promoters. Interestingly, there are three consecutive bands located at positions −10 to −8 which appear to represent partially processed intermediates of the 23S rRNA transcript, in addition to a cluster of bands at position +1 that correspond to the mature 5′ end (Fig. 3(c)). The latter may reflect a small degree of heterogeneity in the site of final 5′ maturation.

5S rRNA

Oligonucleotide N5 and probe P5-1 were used to identify the 5′ terminus of 5S rRNA transcripts by primer extension and nuclease S1 mapping, respectively (Fig. 2(c)). Both approaches gave very similar results. S1 protection experiments with P5-1 indicated that transcription of the 5S rRNA gene begins at position −2 (Fig. 3(d)). Given the ±1-base error in this method, it is possible that the site of initiation actually coincides with the 5′ end of mature 5S rRNA as in two other thermoacidophilic archaebacteria, *Desulfurococcus mobilis* (18) and *Sulfolobus* B12 (28). Although another band often appeared a few bases upstream from the proposed initiation site, its intensity varied in different experiments, decreasing with increasing stringency of hybridization. When another probe, P5-2 (Fig. 2(c)), was used for nuclease S1 mapping, the upstream band was not observed. Thus, while the upstream band remains unexplained, its variability suggests that it is an artifact.

Transcription termination sites

16S rRNA

The 3′ end of the 16S rRNA gene was identified by nuclease S1 mapping with the uniformly labeled probe, P16-3 (Fig. 2(a)). From the autoradiogram presented in Figure 4(a), transcription can be seen to terminate within a T-rich sequence some 40−55 bases from the end of the 16S rRNA. As with the 5′ end, there was no evidence of fragments between the 3′ end of the mature 16S rRNA and the transcription termination sites that might indicate the presence of processing intermediates.

23S rRNA

Nuclease S1 mapping of the 3′ end of the 23S rRNA gene was carried out with probe P23-3 (Fig. 2(b)). As in the case of the 16S rRNA gene, transcription of the 23S rRNA gene terminates within a T-rich sequence roughly 45−60 nucleotides beyond the 3′ end of the coding sequence (Fig. 4(b)). Several strong bands located 9 to 11 bases downstream from the mature 3′ end corroborate the presence of partially processed 23S rRNA transcripts.

5S rRNA

For mapping the 3′ end of the 5S rRNA, probe P5-3 was utilized (Fig. 2(c)). As evident from Figure 4(c), termination of

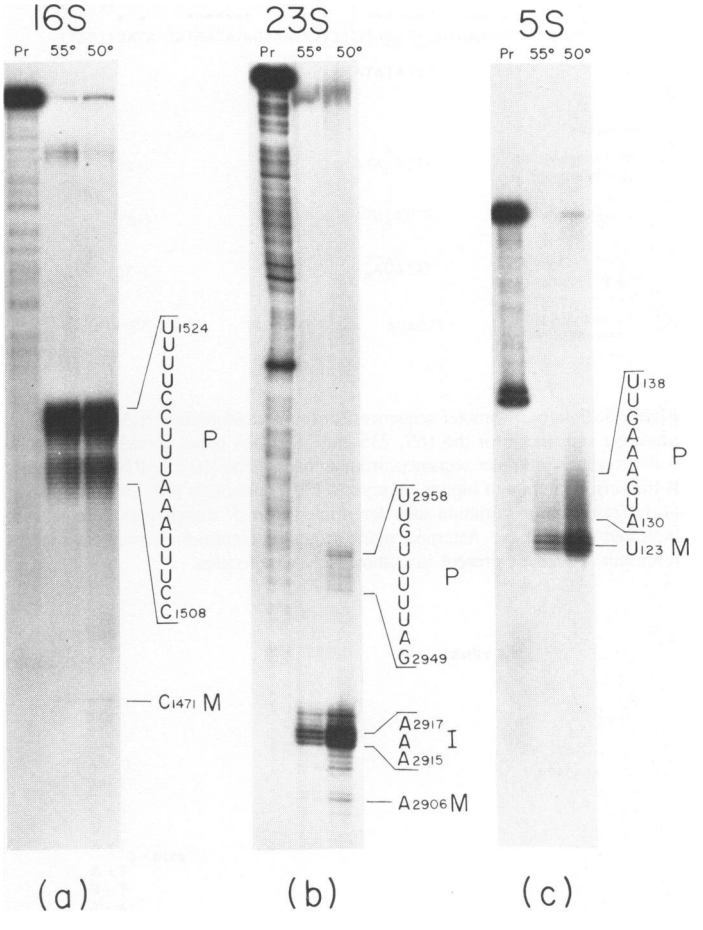

Figure 4. 3′-end mapping of *T. acidophilum* rRNA transcripts by nuclease S1 protection. (a) Mapping of 16S rRNA transcripts with probe P16-3. The 3′ end of the probe was defined by the *TaqI* site at position 1398 within the 16S rRNA coding sequence. (b) Mapping of 23S rRNA transcripts with probe P23-3. The 3′ end of the probe was defined by the *HpaII* site at position 2844 within the 23S rRNA coding sequence. (c) Mapping of 5S rRNA transcripts with probe P5-3. The 3′ end of the probe was defined by the *TaqI* site at position 48 within the 5S rRNA coding sequence. Pr: undigested probes; 55° and 50°: nuclease S1-digested probes from DNA-RNA hybrids annealed at 55° and 50°C, respectively. The lengths of the protected DNA probes were determined from unrelated sequence ladders on the same gels (not shown). Nucleotide position numbers correspond to those in Fig. 1(b). P,I,M refer to precursor, partially processed and mature forms of the rRNA molecules.

transcription of the 5S rRNA gene occurs just before a T-rich tract approximately 10–15 bases beyond the mature 3′ end of the 5S rRNA.

DISCUSSION

As anticipated from the lack of close physical proximity among the single 16S, 23S and 5S rRNA genes of *T. acidophilum* (9), transcript mapping by primer extension and nuclease S1 protection has confirmed that these genes do in fact comprise separate transcription units. Such a dispersal of rRNA genes within the genome is unique among living organisms studied to date.

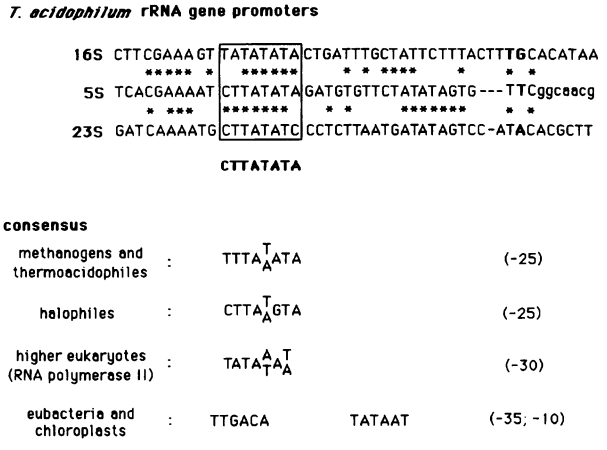

Figure 5. Putative promoter sequences for the *T. acidophilum* rRNA genes. The promoter sequences for the 16S, 23S and 5S rRNA genes (boxed) are aligned with consensus promoter sequences from archaebacteria (10,15), RNA polymerase II transcription units of higher eukaryotes (39), eubacteria (40) and chloroplasts (41). Transcription initiation sites determined from 5′ mapping experiments are indicated in boldface. Asterisks mark sequence identities. Lower case letters represent sequences present in mature rRNA molecules.

Promoters and transcription initiation

The transcription of archaebacterial genes generally begins with a purine residue, and the surrounding nucleotides are conserved to a limited extent (10). The most common sequence is **TRC**, where **R** represents the intiation site. Our mapping experiments showed that the *T. acidophilum* 16S rRNA gene is transcribed from either **T** or **G** in the sequence **TTGC**, while transcription of the 23S rRNA gene start at the **A** residue in the sequence **TAC**. Transcription of the 5S rRNA gene appears to be initiated at the **T** residue in the sequence **TTC***G*, where *G* is the first base of the mature rRNA molecule. The initiation sites are situated 164–165 and 92 bases upstream from the mature 5′ ends of the 16S and 23S rRNAs, respectively, while transcription of the 5S rRNA begins at or very near the mature 5′ end.

A+T-rich sequences some 25 bp upstream from the transcription initiation sites are likely candidates for promoters (Fig. 5). Although the three sequences are not identical, they are very similar both to one another and to the consensus for other archaebacterial genes (10,13–15,32). The consensus for the three putative rRNA gene promoters in *T. acidophilum* is CTTATATA, slighly closer to that of the methanogens and thermoacidophiles, $TTTA^{T}/_{A}ATA$, than to that of the halophiles, CTTAAGTA. As previously noted, these elements strongly resemble promoters recognized by RNA polymerase II from higher eukaryotes, both in sequence and in location relative to the transcription initiation sites (14,15,32). It is worthy of note that an exactly repeated dodecanucleotide sequence, GTTCCGCTTCGA, is found immediately upstream from the promoters for the 16S rRNA gene and for an open reading frame that precedes it (3). This sequence does not occur in the region flanking the 23S rRNA and 5S rRNA promoters, however, and it is not clear whether or not it plays a role in gene expression.

In the course of this work, we noted that a number of phage and plasmid recombinants containing long restriction fragments which encompass the promoters and other upstream elements from the three *T. acidophilum* rRNA genes either impaired cell growth, were not stably maintained, or could not be propagated. As these effects were observed with both pUC and M13 derivatives, incompatibility was not related to the nature of the

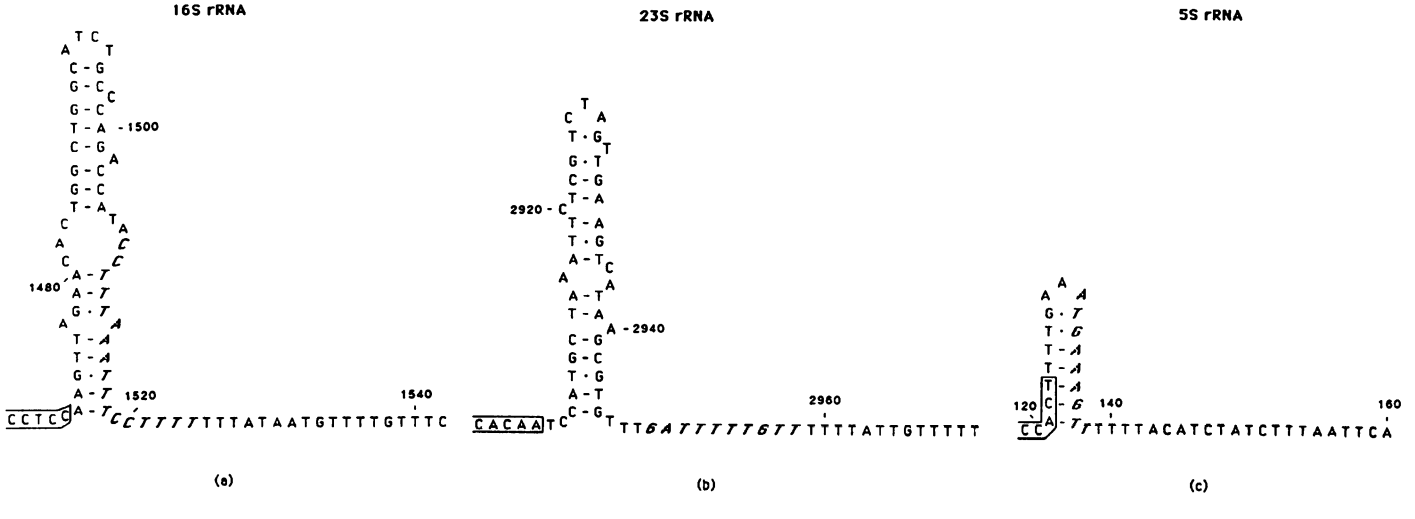

Figure 6. Putative transcription terminators for the *T. acidophilum* rRNA genes. Possible secondary structures for the immediate 3′ flanking regions of the (a) 16S, (b) 23S and (c) 5S rRNA genes are shown. The 3′ ends of the rRNA coding sequences are boxed. Transcription termination sites determined from 3′ mapping experiments are italicized.

parent vector, nor did it appear to be correlated with the location or orientation of the cloned archaebacterial promoters. One possible explanation for these findings is that the insert DNA—and perhaps the rRNA promoters themselves—bind and sequester RNA polymerase or other essential components of the transcription apparatus and deplete the supply available for host-cell gene expression. Alternatively, the *T. acidophilum* fragments may be transcribed, either from nearby promoters within the vector or from their own promoters. If the former were true, then the effects of the cloned sequences should be dependent upon their position and orientation relative to the vector promoters and this appears not to be so. While the expression of archaebacterial genes in *E. coli* has been documented (33,34), *in vitro* studies suggest that archaebacterial and eubacterial RNA polymerases do not utilize each other's promoters (15,35). It is nonetheless possible that *E. coli* RNA polymerase recognizes and makes use of archaebacterial sequences which resemble eubacterial promoters. In any case, transcription of the archaebacterial inserts would surely divert resources from normal cell growth and might even be toxic if the RNA products were able to interact with vital cell constituents such as ribosomal proteins.

Transcription termination

The nuclease S1 mapping experiments reported here revealed that the 3′ termini of all three rRNA transcripts from *T. acidophilum* are located within or near T-rich sequences (Fig. 4). The terminators are not identical, and transcription does not stop at a single, precise site but, rather, at any one of 5–10 positions within the T-rich segments. As depicted in Figure 6, the termination sites are preceded in each case by potential stem-loop structures. The transcription termination signals of the *T. acidophilum* rRNA genes thus resemble the rho-independent terminators of eubacteria (36), although the proposed secondary structures include a number of base bulges and mismatches which may limit the stability of the stems. Transcription of potential protein-coding genes from SSV1, the virus-like particle of *Sulfolobus* B12 (29), and of the rRNA operons of several other archaebacteria (10), also terminates within or near blocks of T residues or other pyrimidine-rich segments, but in these cases, there are no apparent secondary-structure elements preceding the termination sites. In contrast, termination of transcripts from the *H. halobium bop* gene has been mapped to a site just downstream from an inverted repeat consisting of ten G-C pairs but which contains no pyrimidine-rich sequences (37). From these examples, it can be concluded that several different structures may serve as recognition signals for termination in archaebacteria.

Processing of rRNA transcripts

The rRNAs of most prokaryotes are co-transcribed from large, polycistronic operons containing 16S, 23S and, usually, 5S rRNA genes. In *E. coli*, the regions immediately flanking the 5′ and 3′ ends of the 16S and 23S rRNA sequences in the precursor molecules form long base-paired stems that comprise sites for processing by RNase III (16). Because each rRNA gene in *T. acidophilum* is expressed from a separate transcription unit, the processing of its rRNA precursors must differ from that in eubacteria and most other archaebacteria where the mature rRNA species are excised from a single transcript (10,13,16–19). In particular, there is no need to distinguish sequences corresponding to each of the rRNAs since they are all synthesized separately.

Primer-extension and nuclease S1 mapping of 16S rRNA transcripts from *T. acidophilum* provided no evidence for processing intermediates, although it is possible that cleavage occurs very rapidly and that any intermediates formed were not detected in our experiments. Similarly, no partially processed 5S rRNA precursors were observed although here the transcription initiation site is close to or coincident with the mature 5′ end. By contrast, nuclease S1 mapping of 23S rRNA transcripts clearly demonstrated the presence of partially processed intermediates whose termini are centered 8–10 bases upstream and 9–11 bases downstream from the 5′ and 3′ ends of mature 23S rRNA, respectively (see Figs. 3 and 4). Figure 7 shows that a base-paired stem can be formed by sequences immediately adjacent to the *T. acidophilum* 23S rRNA gene. This structure contains two staggered bulge loops separated by four base pairs, a motif which occurs in several other archaebacterial rRNA precursors that are believed to be cleaved by RNase III-like activities (13,17–19). The scissions which define the putative 23S rRNA precursor take place within or close to the bulges, indicating that they may be the recognition sites for the processing nuclease. There is an interesting parallel here with rRNA processing in *E. coli* where cleavage by RNase III is not essential to the formation of mature 16S rRNA, whereas 23S rRNA requires processing by RNase III for its final maturation (38).

Stoichiometry and the regulation of rRNA synthesis

The assembly of ribosomes in all cells requires stoichiometric amounts of the constituent rRNAs. Transcription of the 16S, 23S and 5S rRNAs from polycistronic operons ensures that they are available in unit ratio. Although the three rRNA genes of *T. acidophilum* are located in separate transcription units, it is not unreasonable to assume that their expression is coordinated as

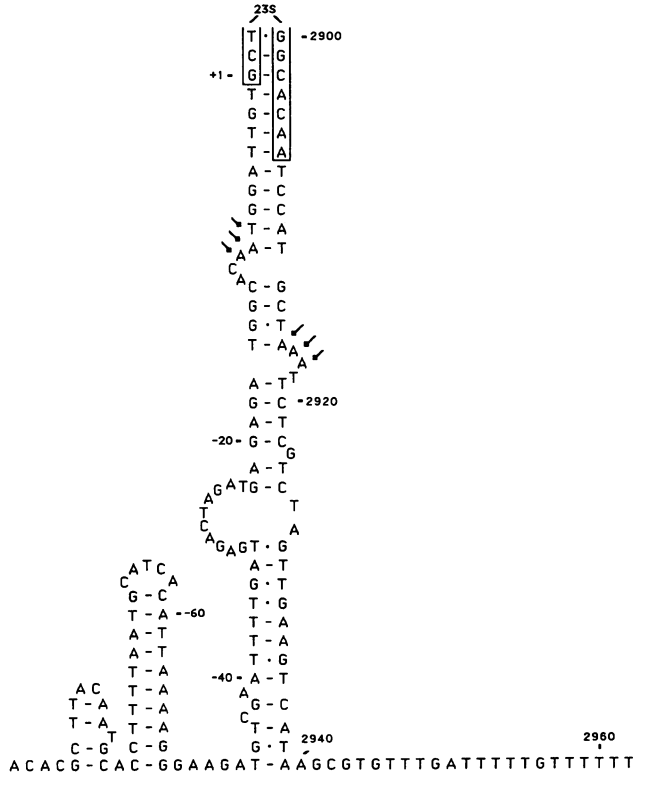

Figure 7. Proposed secondary structure for sequences flanking the *T. acidophilum* 23S rRNA coding sequence. Processing sites are indicated by arrows. The 5′ and 3′ ends of the rRNA coding sequence are boxed.

well and that they are subject to transcriptional controls of the same, or nearly the same strength. However, the promoter, initiation and termination sequences of the *T. acidophilum* rRNA genes, though similar, are not identical, sharing as much in common with those of other archaebacterial genes as they do with one another. The question of coordinate expression also arises in several sulfur-dependent thermoacidophiles where the 5S rRNA is transcribed separately from the 16S and 23S rRNAs (see ref. 10). In *Sulfolobus* B12, the two transcription units are apparently governed by the same promoter sequence, TTTATATG, while another conserved element, CTTATAT, which might play a regulatory role, is located about 50 base pairs upstream from each one (28). It is possible that in *T. acidophilum* other sequence elements serve to coordinate expression of the rRNA genes, but a search of the upstream flanking regions did not reveal any common sequences longer than a few nucleotides, although the CTTAT motif reminiscent of that preceding the *Sulfolobus* transcription units is found 55 bases upstream from the promoter for the *T. acidophilum* 16S rRNA gene. In fact, the three rRNAs may be synthesized at somewhat different rates and the surplus degraded. This would circumvent the need for a finely tuned regulatory mechanism. It can be argued that the unique organization of the rRNA genes in *T. acidophilum* represents a primitive stage in the evolution of rRNA transcription units. If so, the need for the three rRNA molecules in unit stoichiometry would have favored the evolution of the rRNA operons that are ubiquitous in the prokaryotic and eukaryotic worlds today.

ACKNOWLEDGEMENTS

We thank Drs. D. G. Searcy, W.-D. Reiter and W. Zillig for providing materials used in these experiments. This work was supported by National Science Foundation grant PCM84-10832.

REFERENCES

1. Woese,C.R. and Fox,G.E. (1977) *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 5088–5090.
2. Fox,G.E., Stackebrandt,E., Hespell,R.B., Gibson,J., Maniloff,J., Dyer,T.A., Wolfe,R.S., Balch,W.E., Tanner,R.S., Magrum,L.J., 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.
3. Ree,H.K., Cao,K., Thurlow,D.L. and Zimmermann,R.A. (1989) *Can. J. Microbiol.*, **35**, 124–133.
4. Woese,C.R. (1987) *Microbiol. Rev.*, **51**, 221–271.
5. Lake,J.A. (1989) *Can. J. Microbiol.*, **35**, 109–118.
6. Zillig,W., Schnabel,R., Tu,J. and Stetter,K.O. (1982) *Naturwissenschaften*, **69**, 197–204.
7. Schnabel,R., Thomm,M., Gerardy-Schahn,R., Zillig,W., Stetter,K.O. and Huet,J. (1983) *EMBO J.*, **2**, 751–755.
8. Zillig,W., Klenk,H., Palm,P., Pühler,G., Gropp,F., Garrett,R.A. and Leffers,H. (1989) *Can. J. Microbiol.*, **35**, 73–80.
9. Tu,J. and Zillig,W. (1982) *Nucleic Acids Res.*, **10**, 7231–7245.
10. Brown,J.W., Daniels,C.J. and Reeve,J.N. (1989) *CRC Critical Rev. Microbiol.*, **16**, 287–338.
11. Hartmann,R.K., Ulbrich,N. and Erdmann,V.A. (1987) *Biochimie*, **69**, 1097–1104.
12. Liesack,W. and Stackebrandt,E. (1989) *J. Bacteriol.*, **171**, 5025–5030.
13. Wich,G., Hummel,H., Jarsch,M., Bär,U. and Böck,A. (1986) *Nucleic Acids Res.*, **14**, 2459–2479.
14. Reiter,W.-D., Palm,P. and Zillig,W. (1988) *Nucleic Acids Res.*, **16**, 1–19.
15. Thomm,M. and Wich,G. (1988) *Nucleic Acids Res.*, **16**, 151–163.
16. King,T.C., Sirdeskmukh,R. and Schlessinger,D. (1986) *Microbiol. Rev.*, **50**, 428–451.
17. Chant,J. and Dennis,P. (1986) *EMBO J.*, **5**, 1091–1097.
18. Kjems,J. and Garrett,R.A. (1987) *EMBO J.*, **6**, 3521–3530.
19. Kjems,J., Leffers,H., Garrett,R.A., Wich,G., Leinfelder,W. and Böck,A. (1987) *Nucleic Acids Res.*, **15**, 4821–4835.
20. Dale,R.M.K., McClure,B.A. and Houchins,J.P. (1985) *Plasmid*, **13**, 31–40.
21. Messing,J. (1983) *Methods Enzymol.*, **101**, 20–78.
22. Bullock,W.O., Fernandez,J.M. and Short,J.M. (1987) *BioTechniques*, **5**, 376–378.
23. Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
24. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 5463–5467.
25. Tabor,S. and Richardson,C.C. (1987) *Proc. Natl. Acad. Sci. U.S.A.*, **84**, 4767–4771.
26. Darland,G., Brock,T.D., Samsonoff,W. and Conti,S.F. (1970) *Science*, **170**, 1416–1418.
27. Aiba,H., Adhya,S. and de Crombrugghe,B. (1981) *J. Biol. Chem.*, **256**, 11905–11910.
28. Reiter,W.-D., Palm,P., Voos,W., Kaniecki,J., Grampp,B., Schulz,W. and Zillig,W. (1987) *Nucleic Acids Res.*, **15**, 5581–5595.
29. Reiter,W.-D., Palm,P. and Zillig,W. (1988) *Nucleic Acids Res.*, **16**, 2445–2459.
30. Christiansen,C., Freundt,E.A. and Black,F.T. (1975) *Int. J. Syst. Bacteriol.*, **25**, 99–101.
31. Ree,H.K. and Zimmermann,R.A. (1990), submitted.
32. Zillig,W., Palm,P., Reiter,W., Gropp,F., Pühler,G. and Klenk,H. (1988) *Eur. J. Biochem.*, **173**, 473–482.
33. Wood,A.G., Redborg,A.H., Cue,D.R., Whitman,W.B. and Konisky,J. (1983) *J. Bacteriol.*, **156**, 19–29.
34. Morris,C.J. and Reeve,J.N. (1984) In Crawford,R.L. and Hanson,R.S. (eds.), Microbial Growth on C1 Compounds. American Society for Microbiology, Washington, D.C., pp. 205–209.
35. Brown,J.W., Thomm,M., Beckler,G.S., Frey,G., Stetter,K.O. and Reeve,J.N. (1988) *Nucleic Acids Res.*, **16**, 135–150.
36. Rosenberg,M. and Court,D. (1979) *Ann. Rev. Genet.*, **13**, 319–353.
37. DasSarma,S., RajBhandary,U.L. and Khorana,H.G. (1984) *Proc. Natl. Acad. Sci. U.S.A.*, 81, 125–129.
38. Srivastava,A.K. and Schlessinger,D. (1989) *Nucleic Acids Res.*, **17**, 1649–1663.
39. Corden,J., Waslyk,B., Buchwalder,A., Sassone-Corsi,P., Kedinger,C. and Chambon,P. (1980) *Science*, **209**, 1406–1414.
40. Hawley,D.K. and McClure,W.R. (1983) *Nucleic Acids Res.*, **11**, 2237-2256.
41. Kung,S.D. and Lin,C.M. (1985) *Nucleic Acids Res.*, **13**, 7543–7549.