Common structural features of the genes for two stable RNAs from *Halobacterium halobium*

A. Moritz, B. Lankat-Buttgereit, H. J. Gross and W. Goebel*

Institut für Genetik und Mikrobiologie, Röntgenring 11, 8700 Würzburg, FRG

Received 24 October 1984; Revised and Accepted 4 December 1984

**ABSTRACT**

The genes coding for the 5S rRNA and another stable RNA, termed 7S RNA, in *Halobacterium halobium* were isolated from a genomic library of this archaebacterium and their nucleotide sequences determined. Both genes are colinear with their transcripts (5S rRNA and 7S RNA), but 5S rRNA and possibly also 7S RNA isolated from other halobacteria carry additional nucleotides within the RNA transcript. Both genes are located in the G+C rich chromosomal fraction I of *H. halobium*. Comparison of the 3' non-coding regions of both genes shows a 20 bp sequence of high homology immediately at the 3' ends which is almost symmetrically flanked by two stem-loop structures, one being situated close to the 3' end but within the coding region and the other downstream of the common 20 bp sequence.

**INTRODUCTION**

As first shown by Woese (1), 16S ribosomal RNAs from archaebacteria differ considerably in their nucleotide sequence from those of normal eubacteria. More recent reports on the sequences of 5S rRNA from various archaebacteria, such as *Thermoplasma acidophilum*, *Halobacterium cutirubrum* and *Halococcus morrhuae* showed also substantial differences in their primary sequences as well as in their putative secondary structures compared to those of eubacteria and eukaryotes. Moreover it was demonstrated that the 5S rRNA of *H. morrhuae* contains an insert of 108 nucleotides within the mature form (2). This is of particular interest in the light of the recent finding that two rRNA genes from another archaebacterium *Sulfolobus solfataricus* contain putative introns (3) which have only been detected in eukaryotic but never in eubacterial genes. The 5S rRNA from *H. halobium* was previously shown (4) to be part of a single ribosomal RNA operon which has a similar gene arrangement (16S - 23S - 5S) as that of *Escherichia coli* (5). The transcription of this operon seems to start upstream of the 16S rRNA gene and terminates downstream of the 5S rRNA gene (4).

While analyzing the 5S rRNA of various halobacteria we observed in all strains the occurrence of another stable RNA which sediments in a sucrose
gradient at around 7S and is produced in all tested halobacteria in relatively large quantities. Its size (around 300 nucleotides) resembles that of the "7S RNA" found as core in the signal recognition particles ("SRP") which were first isolated and characterized from mammals (6).

We describe in this report the isolation of the genes for both, 5S rRNA and 7S RNA from H. halobium and some of the characteristic properties deduced from their nucleotide sequences.

MATERIALS AND METHODS

Bacterial strains. H. halobium NRC 817 was provided by D. Oesterhelt, H. trapanicum by W. Grant. Halococcus morrhuae was a purified strain, which we obtained from the Deutsche Stammssammlung für Mikroorganismen. For the cloning experiments the E. coli strains HB101 (H.Boyer), 5K and JM 103 were used.

Enzymes and chemicals. T4 polynucleotide kinase and T4 RNA ligase were products from P.L. Biochemicals and α- and γ-32 P-ATP were obtained from Amersham Buchler, Braunschweig. The RNases T1 and U2 were obtained from Calbiochem, Gießen, RNase A from Boehringer, Mannheim. Nuclease from Staphylococcus aureus was from Worthington, Freehold and Ribonuclease from chicken liver from BRL. Restriction enzymes were purchased from BioLabs, and E. coli DNA polymeraseI, Klenow fragment from Boehringer Mannheim.

Media and growth conditions. The halophilic strains were grown in salt medium, consisting of 4 M NaCl, 0.12 M MgSO4, 0.03 M KCl, 0.01 M trisodium citrate and 1 % peptone (Oxoid), pH 7.0, with shaking and illumination at 37°C, for 7 days. E. coli was grown in ENB medium (7) or YT medium (8)(JM 103) with shaking at 37°C.

Isolation of DNA and RNA. RNA was isolated as follows: cells (10 g) were collected by centrifugation at 10.000 rpm, lysed overnight in 83 ml 0.15 M NaCl, 0.1 M EDTA, pH 8.0 with 7 ml 25 % SDS-solution and heated to 75°C for 10 min. After adding 25 ml 5 M NaClO4 the solution was extracted twice with one volume of chloroform-isoamylalcohol (24:1) and the nucleic acids were precipitated with 2 volumes of ethanol (1 h, -20°C). The pellet was resolved in 30 ml 15 mM NaCl, 1.5 mM trisodium citrate, pH 7.0 with 3 ml 3M sodium acetate, 1 mM EDTA, pH 7.0. DNA was precipitated by adding of 17 ml isopropanol and vigorous stirring. The RNA was precipitated by adding two volumes of ethanol.

Isolation of DNA was carried out as previously described (9).

DNA and RNA separation. Restriction enzyme digested DNA was separated on 1
% agarose gels in Tris-phosphate-buffer (0.02 M NaH₂PO₄, 0.002 M EDTA, 0.018 M NaCl, 0.04 M tris-hydrochloride, pH 8.0). RNA was separated on 4% polyacrylamide gels containing 7 M urea according to Maniatis and Efstratiatis (10). Both gels were stained with ethidium bromide (0.5 μg/ml) and photographed under UV-light.

Transfer of DNA fragments to nitrocellulose filters and hybridization. Separated DNA fragments were transferred to nitrocellulose filters by the blotting technique of Southern (11). Radioactive labelling of DNA. ³²P-labelling of DNA was performed by nick translation as described by Rigby et al. (12).

Radioactive labelling of RNA. 7S RNA was partially hydrolysed by alkali treatment (50 mM Tris, pH 9.5, 10 min, 75° C). These fragments and the 5S RNA were ³²P-5'end-labelled by T4 polynucleotide kinase as described under RNA sequencing.

Cloning of total DNA of H. halobium. Partially Sau3A-cleaved DNA of H. halobium was ligated to BamHI-cleaved pHC79 DNA (13) and transformed into E. coli strain HB101 by in vitro packaging as described by Collins and Hohn (14).

Colony hybridization. Clones carrying the genes for 5S or 7S RNA were detected by colony-hybridisation according to R. Thayer (15).

DNA sequencing. DNA was sequenced by the technique of Sanger et al. (16, 17) using the vectors M13mp8 and mp9 (18). Subcloning was either done randomly by ligating HpaII-fragments into the AccI-cleaved vectors or HaeIII fragments into SmaI-cleaved vectors. In addition a non random strategy was used by eliminating subsequently a XmaI/XmaI fragment, a SacII/SmaI fragment, a XmaIII/SmaI fragment and a SphI/SmaI fragment. With the exception of the XmaI deletion the protruding ends were filled in with Klenow enzyme or T4 polymerase before ligating or removed with S1 nuclease or T4 polymerase.

Preparation of labelled RNA. 5'-³²P- and 3'-³²PpCp-labelling of RNA was performed with T4 polynucleotide kinase and with T4 RNA ligase (19), respectively. 3'-³²PpCp-labelling was followed by a partial enzymatic digest with RNase U2 (see below) due to purification problems with uncleaved RNA. The generated 3'-³²PpCp-labelled RNA fragments were purified and isolated by two-dimensional gel electrophoresis according to Lockard et al. (20) and Gross et al. (21).

Partial enzymatic digestion of RNA. Partial enzymatic digestion of RNA with RNase A, T1 and U2 was performed essentially as described (21, 22, 23).
Only the enzyme concentrations were changed in order to obtain fragments of the desired size. These fragments were directly 5'-\(^{32}\)P-labelled and isolated as above.

**Sequencing of labelled RNA.** Intact RNA as well as RNA fragments were sequenced by controlled enzymatic digestion of 5'-labelled RNA followed by polyacrylamide gel electrophoresis (19, 24). To determine the 5' and 3' terminal sequences of intact RNA and for short RNA fragments the two-dimensional mobility shift analysis according to Jay et al. (25), Gillum et al. (26) and Tyc et al. (27) was used.

**RESULTS**

**5S RNAs from Halobacteria.** RNA from several strains of halobacteria was isolated and separated on 4% polyacrylamide-urea gels. Most of 16S and 23S rRNA was removed with the DNA prior to the separation by precipitation with isopropanol. The 5S RNAs from most halobacteria tested possess the expected electrophoretic mobility (about 120 nucleotides). The only exception is the 5S RNA from *Halobacterium trapanicum* which bands at a position corresponding to a RNA with a size of about 231 nucleotides as shown in Fig. 1. A similar "5S RNA" was reported recently (2) for *Halococcus morrhuae* and the isolated RNA from this species indeed corresponds precisely in size to that of *H. trapanicum* (Fig. 1). For the "5S RNA" of *H. morrhuae* it was shown that an insert of 108 nucleotides is present in the mature form of the 5S RNA between positions 108 and 215.

**Isolation of the 5S RNA gene from *H. halobium.*** To test the hypothesis that the 108 b insert may represent an intron in all halobacterial 5S RNA genes which is properly excised by most halobacteria (e.g. *H. halobium*) but not by others (e.g. *H. trapanicum*) we have undertaken to isolate the 5S RNA gene from *H. halobium*. 5S RNA was isolated from the polyacrylamide gel, rebanded and electropheluted. This purified RNA was labelled at the 5' end with \(^{32}\)P and used as hybridization probe. DNA from *H. halobium* was separated into fractions I and II on bisacrylamide-malachite green columns as previously described (28). *PstI* cleaved DNA from both chromosomal fractions was hybridized with the \(^{32}\)P-labelled 5S RNA. Fig. 2 indicates that hybridization is obtained only in fraction I in a 1.1 kb *PstI* fragment. A cosmid library of fraction I DNA was constructed by ligating *BamHI*-cleaved *pHC79* with partially *Sau3A* cleaved fraction I DNA as described (13). Positive clones were detected by colony hybridization with the \(^{32}\)P-labelled 5S RNA probe in this cosmid library of *H. halobium*. Cosmid DNA from
Fig. 1. Total RNA of *H. halobium* (lane 1), *H. trapanicum* (lane 2) and *Halococcus morrhuae* (lane 3) separated on a 4% polyacrylamide-urea gel. The three lanes shown in this figure were cut out of the same slab gel. For the size determination of the 7S RNA from *H. halobium* and the 5S RNA from *H. trapanicum* the already determined 5S RNAs from *H. halobium* and *H. morrhuae* (123 b and 231 b, respectively) were used as internal markers.

Fig. 2. (A) Hybridization (lane 3 + 4) of $^{32}$P-labelled ribosomal RNA with PstI-digested DNA of *H. halobium* separated on a 1% agarose gel (lane 1 + 2). Lane 1 represents chromosomal DNA (fraction I) (68 mol % G+C), lane 2 a mixture of more A+T rich chromosomal and plasmid DNA (fraction II) (58 mol % G+C) of *H. halobium*. The two fractions were obtained as described by Ebert et al. (3). The size of the hybridizing PstI fragments were determined from the known sizes of the PstI fragments of pH1 seen in lane 2. The sizes of the two hybridizing PstI fragments fit well with those previously reported by Hofman et al. (4) and Neumann et al. (29).

(B) Hybridization of purified, $^{32}$P-labelled 5S rRNA with PstI-digested total DNA of *H. halobium* (lane 1 cleaved DNA; lane 4 hybridization), with a recombinant cosmid DNA, containing ribosomal gene fragments (lane 2 and 4) and with the M13mp9 subclone DNA, containing the small PstI fragment (lane 3 and 6).
Fig. 3. Sequence of the 5S rDNA (upper lane) and the 5S rRNA (lower lane) from *H. halobium* and putative 5'- and 3' non-coding regions of the 5S rRNA gene.

these clones was isolated and cut with PstI. All plasmid DNAs contained a 1.1 kb PstI fragment which hybridized exclusively with the 5S RNA probe. This fragment is adjacent to the large 30 kb PstI fragment, which carries the 23S RNA and the 16S RNA genes (29). The 1.1 kb fragment was isolated from the agarose gel and subcloned in both orientations into the M13mp8 vector (18).

**Sequencing of 5S RNA and 5S RNA gene.** The electrophoretically purified 5S RNA from *H. halobium* was sequenced using a strategy described before (19). The obtained RNA sequence (Fig. 3) is identical to that previously reported for another halobacterial species, *H. cutirubrum*, and confirms that of Mankin et al. (30). The "5S RNA" sequence from *Halococcus marrhuae* resembles closely that of *H. halobium* but shows the 108 nucleotide insert between positions 108 and 215 (2). Based on the size we assume that the "5S RNA" of *Halobacterium trapanicum* is similar to the 5S RNA of *H. marrhuae*.

The hybridization data shown in Fig. 2 indicated that the 5S RNA gene of *H. halobium* is located on the 1.1 bp PstI fragment. This fragment cloned into M13mp8 was sequenced from both ends using the dideoxynucleotide sequencing technique. Start and end of the 5S RNA gene was identified by comparing the obtained DNA sequence with the 5S RNA sequence. As shown in Fig. 3 the DNA sequence is perfectly colinear with the RNA sequence indicating that the 5S RNA gene of *H. halobium* does not contain an intron.

As recently described (31) and confirmed by our data, 5S rRNA and hence the coding region of the 5S rRNA gene of *H. halobium* shares about 50 percent sequence homology with corresponding sequences of lower eukaryotes (e.g.
Physarum) and eubacteria (e.g. E. coli and B. subtilis).

No sequence homology between the 3' and 5' non-coding regions of the 5S RNA genes of H. halobium and E. coli (and several other organisms) is observed although the arrangement of the ribosomal RNA genes in H. halobium is the same as that in E. coli, e.g. 16S - 23S - 5S (29) and the three genes form in both microorganisms a transcriptional unit (4) suggesting transcriptional termination signals at the 3' end of the 5S RNA genes. As will be described below there is, in fact, an interesting sequence similarity at the 3' ends of two genes from H. halobium, the 5S rRNA gene and the gene coding for another stable RNA, termed 7S RNA.

Identification of a 7S RNA in halobacteria and cloning of its gene from H. halobium. In addition to tRNA and 5S RNA which can be readily identified in RNA preparations on polyacrylamide urea gels, another major RNA species is visible in all RNA preparations from halobacteria which were tested. This RNA bands in polyacrylamide urea gels at a position corresponding to about 300 nucleotides and sediments in a sucrose gradient at around 7S. We will therefore term this RNA in the following "7S RNA". As indicated in Fig. 1, the 7S RNA differs slightly in size by about 30 nucleotides between various halobacterial species, which may again indicate the occurrence of an insert within this RNA in some strains.

7S RNA was isolated from polyacrylamide/urea gels, partially hydrolysed with alkali, 5' end-labelled with T4 polynucleotide kinase and used as probe for identifying the 7S RNA gene in a cosmid gene银行 constructed from total DNA of H. halobium using pH C79 as cosmid vector. The 7S RNA gene was identified on a 7.3 kb PstI fragment. The location of the 7S RNA gene could be further narrowed down on a SalI fragment of 850 bp (Fig. 4). This SalI fragment, also found in the fraction I of H. halobium chromosomal DNA (data not shown) was cloned into the SalI site of M13mp9 phage RF DNA and sequenced using again the dideoxynucleotide sequencing technique. Sequence of the 850 bp SalI fragment and determination of the coding region for 7S RNA. The 850 nucleotides of the SalI fragment were determined by sequencing it from both ends using a universal primer for mp9. The original SalI fragment was subsequently reduced in size by cleaving the SalI insert cloned in mp9 with restriction enzymes which uniquely cut within the insert (XmaI, SacI, XmaIII and SphI). The mp9 vector part was cleaved with SmaI or XmaI and the two ends were ligated either directly (XmaI/XmaI) or after filling in or removing protruding sticky ends of the other restriction site by T4 DNA polymerase as described (32).
Fig. 4. Restriction map of an 850 bp SalI fragment carrying the 7S RNA gene. This SalI fragment was subcloned from a 7.3 kb PstI fragment isolated from a recombinant pH79 cosmide. The position of the 7S RNA gene on this fragment (from coordinates 169 to 472) is given together with various restriction sites mentioned in the text.

In order to localize the gene for 7S RNA on this sequence we determined the RNA sequence of purified 7S RNA from the 3' and 5' ends. The 5' end was labelled with T4 polynucleotide kinase and the 3' end with T4 RNA ligase (19). Labelling at the 3' end was difficult to perform suggesting strong secondary structure at this end of the RNA molecule. Using conventional RNA sequencing techniques (19) 13 nucleotides could be clearly determined starting at the 5' end and 18 nucleotides starting from the 3' end. These
partial sequences of the 7S RNA at the 5' and 3' ends were sufficient, however, to detect start and end for the 7S RNA gene in the sequence of the 850 bp SalI fragment, assuming that the 7S RNA is the direct transcript of the gene. As indicated in Fig. 5 the 7S RNA gene starts at position 169 and ends at position 472. The length of the gene (304 bp) correlates well with the length of the RNA suggesting that the gene for the 7S RNA does not contain intron(s) although extremely small introns (less than 10 bp) can not be excluded by these data. This was further confirmed by sequencing fragments of 7S RNA obtained by partial digestion of this RNA with RNases U2, T1 and A (21, 22, 23).

These partial RNA sequences covered 92% of the total DNA sequence identified as the gene for the 7S RNA without any interruption. The RNA sequence gap at positions 359-369 is probably due to a structure exposing this region to nucleases. We never obtained an RNA fragment which contains this sequence; the other missing nucleotides in the RNA sequence derive from band compressions in the sequencing gels.

Analysis of the 3' and 5' non-coding regions of the 7S RNA gene and the 5S RNA gene. There is no indication that 7S RNA is processed from a larger precursor since hybridization of total RNA of H. halobium with an internal 7S RNA gene probe could not detect larger hybridizing RNA. Furthermore recent experiments have shown that 7S RNA is not part of isolated ribosomes of H. halobium (Moritz et al., manuscript in preparation). The 7S RNA is not part of the ribosomal RNA operon since the 7.3 kb PstI fragment which carries the 7S RNA gene is not overlapping or flanking the 30 kb and 1.1 kb PstI fragments carrying the ribosomal RNA operon (see above). We therefore assume that the 7S RNA is the direct transcript and the complementary DNA represents its structural gene. At the 5' end of this gene there is no sequence found which is reminiscent to any eubacterial or eukaryotic promoter consensus sequences (33). The most interesting feature is observed at the 3' end, where a G+C rich stem-loop structure is found 34 bp downstream of the putative transcriptional stop. A similar structure is likewise found at the 3' end of the 5S rRNA gene in about the same distance from the putative transcription stop. Moreover both genes share a highly homologous 20 bp sequence which is relatively A+T rich and is proximal to the 3' ends of the coding regions for the 5S rRNA and the 7S RNA, respectively. There is a partial overlapping of this common sequence and the last nucleotides of the coding regions of the two genes. Both genes possess in addition a second stem-loop structure few base pairs upstream of the

39
Fig. 6. Comparison of the 3' ends of the 5S rRNA and the 7S RNA genes. The ends of the mature transcripts are marked by (►). The common sequence is boxed in.

common 20 bp sequence within the structural gene. These three elements form a quite unique structure which is very similar in both genes as indicated in Fig. 6.

DISCUSSION

The gene for the 5S rRNA in *H. halobium* and probably in other halobacteria as well is part of a ribosomal RNA operon which has the same gene arrangement as in *Escherichia coli* (4). However, as first reported by Luehrsen et al. (31) rRNAs from archaeabacteria including *H. halobium* differ from 5S rRNAs of eubacteria in their primary sequences and in their putative secondary structures and resemble their eukaryotic counterparts in some features (34). The DNA sequence of the 5S rRNA gene from *H. halobium* reported here confirms these earlier observations for the structural part of the genes. The recent identification of putative introns in two tRNA genes from another archaeabacterium *Sulfolobus solfataricus* (3) has further strengthened the argument that archaeabacteria may exhibit both, procaryotic and eukaryotic, features in their genetic structures. The reported observation that the 5S rRNA from *Halococcus morrhuae* and - as shown here - from *Halobacterium trapanicum* contains a 108 nucleotide insert and that the 7S RNA, another stable RNA found in several halobacteria, is larger in some species than in others made us hypothesize that these genes may likewise contain introns which are properly excised in some halobacteria but not in others. The DNA sequences of both genes from *H. halobium*, however, are clearly colinear with their mature transcripts and indicate that the occurrence of additional nucleotides in some halobacterial 5S rRNA and 7S
RNA represents a peculiarity of the genes from these species. The origin of the inserts is as yet unknown but ongoing studies should reveal whether they are tolerated residues within the gene of one of the many insertion elements which are frequently occurring in halobacteria.

Little is known on signal structures essential for the expression of archaebacterial genes. It has been demonstrated (35, 36) that RNA polymerases from archaebacteria contain more subunits than the typical eubacterial RNA polymerase and that some subunits of archaebacterial RNA polymerases cross react with antibodies raised against the RNA polymerase II from yeast. This suggests that these RNA polymerases may require unique promoter and transcription termination structures. The 5S rRNA gene is localized at the end of the ribosomal RNA operon and hence should carry at its 3' end a transcription termination signal. The 7S gene is expected to carry at its 5' end a promoter sequence and at its 3' end a transcription termination signal since there is no indication that the 7S RNA is processed from a larger RNA precursor or is part of the ribosomal DNA operon. The comparison of the 3' non-coding regions from both genes revealed a striking similarity between a 21 bp sequence (7S RNA gene) and a 22 bp sequence (5S rRNA gene) following and partially overlapping the 3' ends of the coding regions.

This common sequence is surrounded by two stem-loop structures, one being still located within the structural gene and the other in an equal distance downstream of the 20 bp common sequence. Whereas the common sequence is relatively A+T rich, the stems are more G+C rich. The significance of these structures which could be involved in transcription termination is as yet unclear. This structure is, however, reminiscent of common sequences found at the 3' ends of the 5S RNA genes in the rrnC and rrnB operons of Escherichia coli (37). A similar conserved sequence is also found upstream of the T2 terminator structure of rrnB. Here, the conserved sequence includes also the last nucleotides of the 5S RNA, extends into the 3' non-coding region and overlaps partially with the terminator stem loop structure. These sequences are different from those recently identified at the 3' end of the opsin gene from H. halobium and claimed to represent termination signals for the transcription of this gene.

Promoter sequences for halobacteria or other archaebacteria have as yet not been clearly identified. Recently, DasSarma et al. (38) have pointed out some structural similarities between the 5' region of the opsin gene and that of several open reading frames identified in two insertion elements.
(ISH) from H. halobium. These include an A+T rich hexanucleotide (AAGTTA) 28 to 42 bp upstream from the putative transcriptional starts. The only common feature which we can observe in the 5' non-coding region of the 7S RNA gene is a AAGTA sequence but this is 52 bp upstream from the putative transcriptional start of the 7S RNA. Sequences resembling the consensus promoter sequences of many genes from E. coli (33) and other eubacteria are also not found in the 5' non-coding region of the 7S RNA gene.

ACKNOWLEDGEMENTS
The authors wish to thank K. Ebert, U. Rdest and B. Surek for advice and helpful discussion and E. Appel for editorial assistance. This work was supported by the Deutsche Forschungsgemeinschaft (Go 168/12-2 and Gr 576/4-5) and the Fonds der Chemischen Industrie.

*To whom correspondence should be sent

REFERENCES