The EMBO Journal vol.5 no.5 pp.1091-1097, 1986

# Archaebacteria: transcription and processing of ribosomal RNA sequences in *Halobacterium cutirubrum*

John Chant and Patrick Dennis

Department of Biochemistry, University of British Columbia, Vancouver, British Columbia, V6T 1W5 Canada

Communicated by R.A.Garrett

**The chromosome of *Halobacterium cutirubum* contains a single ribosomal RNA gene cluster. The 5′ to 3′ organization of genes within this 6-kbp region is: 16S, alanine tRNA, 23S, 5S, cysteine tRNA. The entire gene cluster is transcribed as a single long primary transcript; processing of mature RNA sequences from the 5′ region of the transcript begins prior to the completion of synthesis at the 3′ end. There are five conserved octanucleotide direct repeats (TGCGAACG) in the 900-bp 5′-flanking sequence in front of the 16S gene. The positions of these repeat sequences correspond to the different 5′ ends of the primary transcript and probably represent the RNA polymerase start sites. The 16S and 23S rRNA genes are surrounded by long nearly perfect inverted repeat sequences. These sequences probably form duplex structures in the primary transcript and are recognized by an RNaseIII-like endonuclease activity that carries out the initial excision of the precursor 16S and 23S rRNA sequences. These precursors are rapidly trimmed to the mature 16S and 23S molecules and assembled into ribosomal particles. The processing sites for 5S rRNA appear to be at or very near to the mature ends of the 5S molecule. The tRNA sequences are processed with reduced efficiency from the primary transcript. Nuclease cuts have been detected at the ends as well as in the middle of the cysteine tRNA sequence suggesting that there may be alternative processing pathways, one resulting in proper excision of the mature tRNA sequence and the other resulting in improper excision and degradation of the tRNA sequence. The transcription termination sequence is believed to be at or beyond an AT-rich sequence preceded by a GC-rich sequence located distal to the cysteine tRNA gene.**

*Key words:* archaebacteria/*Halobacterium*/rRNA processing

## Introduction

Living organisms can be divided into three distinct phylogenetic kingdoms — the traditional eubacterial and eukaryotic kingdoms and the newly recognized archaebacterial kingdom (Woese, 1982; Woese and Fox, 1977). Although archaebacteria exhibit prokaryotic cell structure and organization, they are evolutionary quite distinct from eubacteria. Some of these distinguishing and unifying features include (i) the sequence and structure of rRNAs, ribosomal proteins and tRNAs (Fox, 1985; Matheson, 1985), (ii) the presence of intervening sequences in some genes encoding rRNA, tRNA and possibly also protein (Daniels *et al.*, 1985; Kaine *et al.*, 1983; Larsen *et al.*, 1986), (iii) an RNA polymerase subunit complexity reminiscent of and immunologically related to eukaryotic RNA polymerases (Huet *et al.*, 1983), and (iv) a translation apparatus that utilizes 16S, 23S and 5S rRNAs, that is sensitive to many eukaryotic ribosome-targeted antibiotics and to ADP ribosylation by diphtheria toxin and that initiates protein synthesis with methionine rather than formyl methionine (Woese, 1982).

The rRNA genes of the archaebacterial species *Halobacterium cutirubrum* are present in single copy and confined to a 6-kbp region of genomic DNA (Hui and Dennis, 1985). The organization of this cluster is eubacteria-like and has been shown by DNA sequence analysis to be 5′-16S, alanine tRNA, 23S, 5S, cysteine tRNA -3′; none of the genes contain intervening sequences.

The clustering and physical organization of the rRNA genes suggests that they are co-transcribed to produce a long polycistronic primary transcript (Hofman *et al.*, 1979; Hui and Dennis, 1985). The 900-bp region preceding the 16S gene contains three highly conserved and two less highly conserved direct repeat sequences; it has been suggested that these sequences might be elements of the rRNA promoter (Hui and Dennis, 1985; Mankin *et al.*, 1984). The 16S and 23S genes are each surrounded by long, nearly perfect inverted repeat sequences; this suggests that precursor 16S and 23S rRNA might be excised from a primary transcript by a double strand-specific RNaseIII-like endonuclease activity.

The analysis of *in vivo* transcription and processing of the rRNA gene cluster is greatly simplified by the fact that all rRNA transcripts are derived from a single gene cluster. Our previous results have indicated that the 5′-flanking direct repeat sequences correspond to the 5′ ends of the primary transcripts (Dennis, 1985). Here we show that the entire gene cluster is co-transcribed and that the inverted repeat sequences surrounding the 16S and 23S genes are sites for endonuclease excision of precursor 16S and 23S rRNA from the long primary transcript. Additional processing sites and the time ordering of processing events has also been examined.

## Results and Discussion

Primary transcripts reading into the 16S rRNA gene of *H. cutirubrum* have been shown by S1 nuclease mapping to have multiple 5′ ends (Figure 1, Dennis, 1985). The positions of five of these ends correspond to an octanucleotide element of a longer sequence which is directly repeated five times at intervals of ~125 nucleotides in the region between −800 and −200 bp in front of the 16S rRNA gene (Hui and Dennis, 1985). This octanucleotide, TGCGAACG, appears to represent start sites for transcription rather than RNA processing sites since no 3′ ends have ever been detected at any of these positions and the abundance of transcripts derived from upstream sequences appears to be negligible. The abundance of the different 5′ transcript ends varies by >50-fold; sequences that might be responsible for these differences in promoter efficiency have been located at positions −45 and −65 relative to the octanucleotide repeat. These sequences show no resemblance to either eukaryotic or prokaryotic promoter consensus sequences (Figure 2).

An additional 5′ transcript end was located ~108 nucleotides in front of the 16S rRNA sequence. This end was apparently generated by endonuclease processing within the first half of the

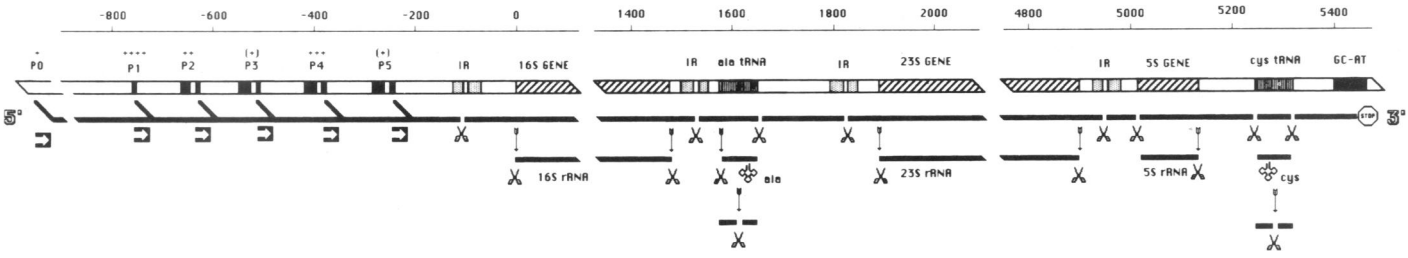


**Fig. 1.** Transcription and processing of the rRNA gene cluster of *H. cutirubrum*. The structure and organization of the rRNA gene cluster interrupted in the 16S and 23S genes is depicted by the bar portion of the graph. The nucleotide scale is given above and the primary transcript (heavy solid line) is given below. The direct repeat sequences (P1−P5) have been identified as putative transcription start sites. A sixth uncharacterized upstream start site (PO) is also included and the relative strengths of the different start sites are indicated [++++, +++, + and (+)]. The multiple 5′ ends of the primary transcripts are indicated by the branch points and the boxed arrows. The inverted repeat sequences (IR) surrounding the 16S and 23S genes are indicated. The initial processing sites within the primary transcript which have been identified by S1 mapping of *in vivo* RNA are indicated as scissored interruptions of the primary transcript. Subsequent processing events necessary to produce mature rRNA or tRNA sequences are also indicated. Termination of transcription possibly occurs in or beyond the tandem GC-rich−AT-rich sequence beyond the cysteine tRNA gene. Cutting within the transcribed tRNA sequence possibly represents an alternative processing pathway occurring within the primary transcript and results in degradation and loss of the alanine and cysteine tRNA sequences.

```
STRONG PROMOTERS

P1   -833 TGTCCCTGGTGTGGGCCGCCATCAC....13nt.....TTCGA...33nt...TGCGAACG -750
           • ●●● ●●●●●● ●● ● •  ●● ●            ●●●●●          ●●●●●●●●
P4   -452 TCTCCATGGTGTCGGTCTCACTCTC....11nt.....TTCGA...33nt...TGCGAACG -371

INTERMEDIATE PROMOTER

P2   -710 GCCCCCTCGTCCCGTTCGGACGGAA....15nt.....TCCGA...33nt...TGCGAACG -625
            ○ ●●○● ●●   ○● ○●○ ○○  ○                ● ●●●          ●●●●●●●●

WEAK PROMOTERS

P3   -585 ACTCCGCTGATCGGTTCGGCGTTCG....11nt.....TTCGA...33nt...TGCGAACG -504
            ○●●●   ●   ○● ○●○ ○ ●                  ●●●●●          ●●●●●●●●

P5   -323 CCGCCCTGGTGGGCAAACGTCACGC....15nt.....TTCGA...33nt...TGCGAACG -238
            ○ ●●○●●●●● ○        ○ ● ●              ●●●●●          ●●●●●●●●
```

**Fig. 2.** Putative rRNA promoter sequences. The nucleotide sequences upstream from the five transcript start sites have been compared. The distance upstream from the beginning of the 16S rRNA genes (nucleotide +1) is given at the start and end of each sequence. Transcripts start near the middle of the conserved octanucleotides at the end of each sequence. Bases conserved in both of the two strong putative Pl and P4 promoter sequences (●) and bases in P1 or P4 (○) are indicated. There is considerable sequence conservation in the 33-nucleotide spacer between the penta- and octanucleotide elements of P2, P3, P4 and P5 (Hui and Dennis, 1985) since the strong promoter P1 does not conserve these sequences, they are probably not essential for promoter activity.

long, nearly perfect inverted repeat that surrounds the 16S rRNA sequence. The corresponding 3′ end of the leader sequence released following this cut was easily detectable and more stable than the 5′ end of the excised precursor 16S rRNA (Dennis, 1985).

### *The 16S−23S intergenic space*

The 16S−23S intergenic space contains the second half of the 16S inverted repeat sequence, an alanine tRNA gene and the first half of the inverted repeat sequence that surrounds the 23S gene (Hui and Dennis, 1985). It was possible to demonstrate continuous transcription through the intergenic space by hybridizing total cellular RNA to the minus strand of a 752 nucleotide long *Msp*I fragment labeled either at the 3′ or 5′ end. In both cases continuous transcripts present in total cellular RNA protect the entire restriction fragment from S1 nuclease digestion (Figures 3 and 4).

In addition to full protection, a number of different partial protection products were observed. For the 3′ end-labeled fragment, partial protection products of 86 nucleotides, 152 nucleotides, 266 nucleotides and 435 nucleotides in length were observed (Figure 4A, 3′). These appear to correspond respectively to protection by (i) the 3′ end of mature 16S rRNA, (ii) the 3′ end of precursor 16S rRNA produced by excision within the 16S inverted repeat sequence, (iii) the 3′ end of an RNA produced by endonuclease cutting of the primary transcript at or near the 3′ end of the alanine tRNA sequence and (iv) the 3′ end of the long leader sequence (containing the 16S and alanine tRNA sequences) produced by cutting within the inverted repeat preceding the 23S gene.

Similarly, for the 5′ end-labeled *Msp*I fragment partial protection products of 155 nucleotides, 486 nucleotides and 600 nucleotides in length were observed (Figure 4A, 5′). These appear to correspond respectively to protection by (i) the 5′ end of mature 23S rRNA, (ii) the 5′ end of the trailer sequence liberated following cutting at or near the 3′ end of the alanine tRNA sequence and (iii) the 5′ end of the trailer sequence liberated by the excision of the precursor 16S rRNA sequence.

These S1 nuclease protection results suggest the following: first, that the 16S and 23S rRNA genes are co-transcribed; second, that the initial processing event within the 16S−23S spacer can occur either within the 16S inverted repeat processing signal, at or near the 3′ end of the alanine tRNA sequence or within the 23S inverted repeat processing signal; third, that 5′ processing of the alanine tRNA sequence occurs only after processing

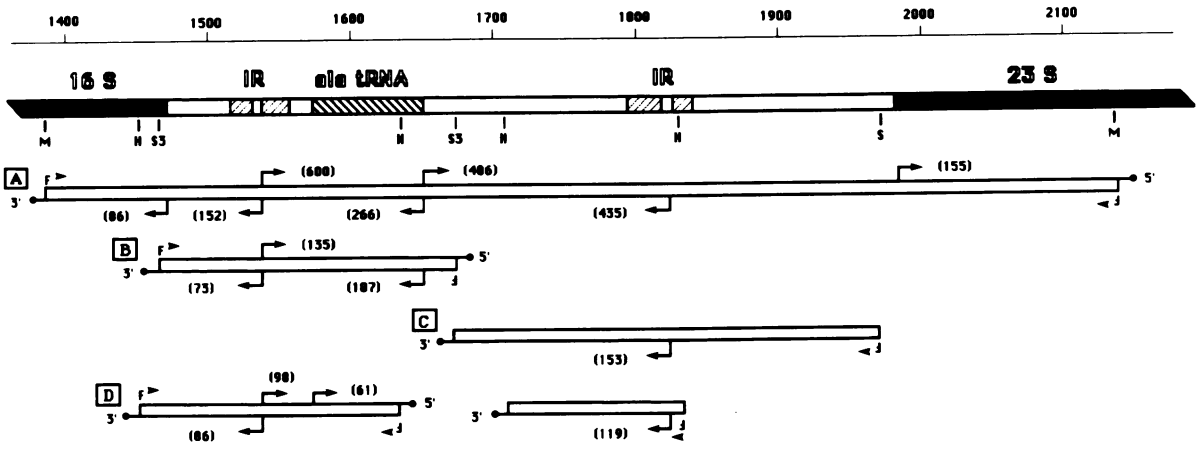


**Fig. 3.** The 16S−23S rRNA intergenic space: structure and processing of the primary transcript. The bar portion illustrates the structural organization of the 16S−23S rRNA intergenic space. The positions of the inverted repeats (IR) and the alanine tRNA gene are indicated. The nucleotide sequence scale is at the top. Restriction enzyme sites utilized in transcript mapping are: *Msp*I (M; 1384, 2136); *Hin*fI (H; 1459, 1633, 1705, 1830); *Sau*3AI (S3; 1462, 1670); *Sal*I (S; 1969). Some of the DNA fragments used for nuclease S1 protection are illustrated below. End labeling at the 3′ or 5′ positions, where utilized, are indicated respectively at the left or right end of each fragment. The size and position of the ends of transcripts which provide protection of end label are indicated; 5′ transcript ends are depicted above and 3′ transcript ends below each DNA fragment. Protection of the entire DNA fragment by an RNA transcript is indicated by an (F) at the end of the fragment. The autoradiograms of the four fragments labeled, A, B, C and D are illustrated in Figure 4.

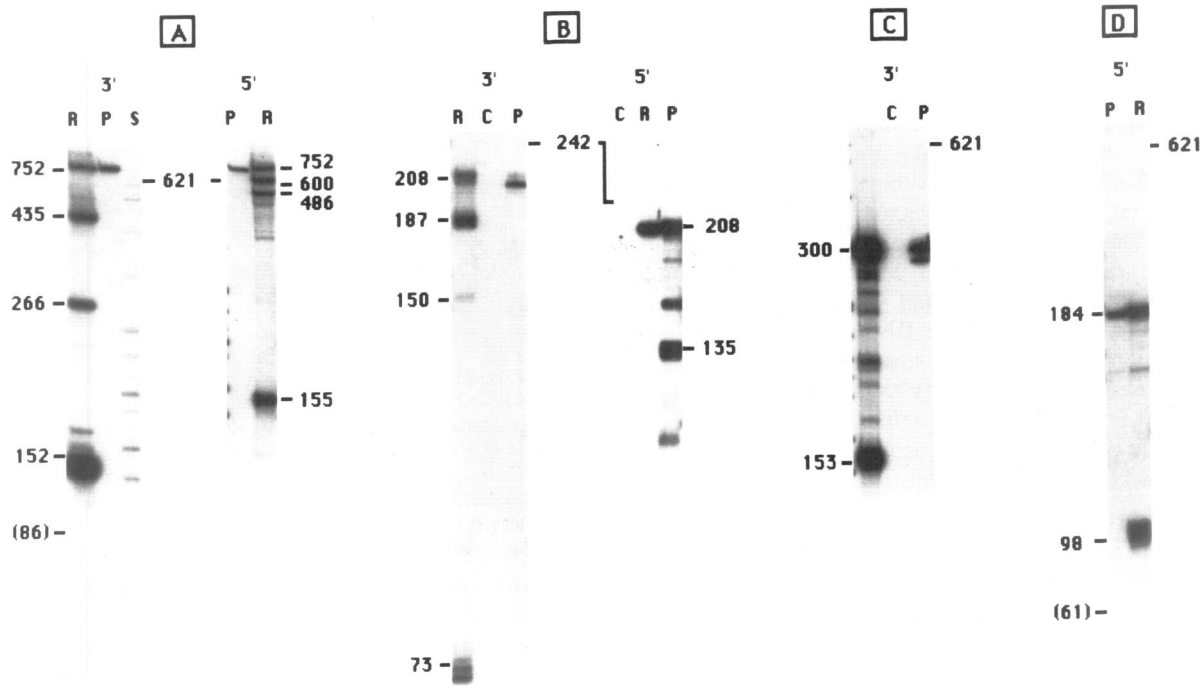


**Fig. 4.** Protection of end-labeled DNA fragments by rRNA primary transcripts and processed intermediates derived from the 16S−23S intergenic space. The four restriction fragments and their respective nucleotide positions of 3′ or 5′ labeling on the (−) DNA strand are: (**A**) *Msp*I, 3′ at nucleotide 1386 or 5′ at nucleotide 2138; (**B**) *Sau*3AI, 3′ at nucleotide 1465 or 5′ at nucleotide 2138; (**C**) *Sau*3AI-*Sal*I, 3′ at nucleotide 1673; (**D**) *Hin*fI, 5′ at nucleotide 1636. The lane designations are: S, molecular length standards; P, end-labeled untreated DNA probe; R, probe hybridized to 5μg of total RNA from *H. cutirubrum* and digested with nuclease S1; C, probe hybridised to 5μg of total RNA from *E. coli* and digested with nuclease S1. The molecular length standards were 3′ end-labeled *Msp*I fragments of pBR322. The lengths of these standards are 621, 527, 403, 309, 242, 238, 217, 202, 190, 180, 160, 147, 122, 110, 90, 76, 67, 34, 26, 15 and 9 nucleotides. For each autoradiogram the position of the 621 or 242 base standard is indicated. Estimates of the sizes of the protected fragments are accurate to within ~ ± 3%. Control experiments utilizing *E. coli* RNA showed little or no protection of the respective end-labeled probes and indicates that probe re-hybridization was negligible. The fragment lengths in parentheses and below the autoradiograms represent protection by the short ends of mature RNA sequences; these bands, although clearly visible on the autoradiograms, are not shown.

at the 3′ end has been initiated; fourth, that the 5′ end of the precursor 23S sequence liberated by cutting within the 23S inverted repeat is rapidly trimmed to the mature 5′ end of 23S rRNA whereas the 3′ end of the leader is transiently more stable.

In interpreting these results we have focussed almost exclusively on the major protection products and assumed that most (but not all) are produced by endonuclease cleavage events within the primary transcript. The 3′ and 5′ ended products of such an endonuclease event, when transiently stable, give rise to corresponding 3′ and 5′ protection products. For example, the 600

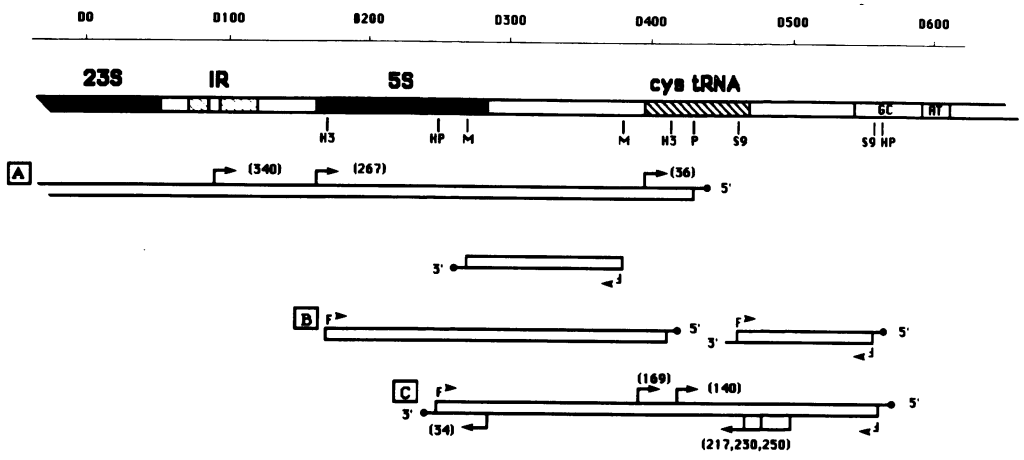


**Fig. 5.** The region distal to the 23S rRNA gene: structure and processing of the primary transcript. The bar portion illustrates the structural organization of the region distal to the 23S rRNA gene. The positions of the 23S inverted repeat (IR), the 5S gene, the cysteine tRNA gene and the GC-rich−AT-rich region are indicated. The nucleotide sequence scale is at the top. Restriction enzyme sites utilized in transcript mapping are: *Hae*III (H3; D169, D413); *Hin*PI (HP; D249, D561); *Msp*I (M, D268, D378); *Pst*I (P, D429) *Sau*961 (S9; D460, D556). Some of the DNA fragments used for nuclease S1 protection are illustrated below. Other details are as in the legend to Figure 2. The autoradiograms of the three fragments labeled A, B and C are illustrated in Figure 6.

nucleotide long 5′ protection fragment and the 152 nucleotide long 3′ protection fragment of the 752 nucleotide long *Msp*I DNA probe are generated by the products of an endonuclease cut within the 16S inverted repeat sequence within the primary transcript. If one or both of the primary products are rapidly trimmed or processed further, protection by the primary product is diminished or abolished. For example, the 5′ end of the precursor 23S rRNA produced by endonuclease excision within the 23S inverted repeat should result in a major 5′-protected fragment of ~317 nucleotides in length; only a minor band appears in the autoradiogram near this position suggesting that the precursor is rapidly trimmed to the mature 5′ end. Other minor bands could arise from contamination of probe DNAs or from transiently stable degradation intermediates of spacer RNA sequences present in or released from the primary transcript. Additional experiments with shorter restriction fragments from the 16S−23S intergenic space confirm the positions of these transcript ends (Figure 4, B,C,D and other unpublished results).

*The 23S distal region*

Because of the paucity of convenient restriction sites in the DNA, the abundance and concentration of potential processing sites and the apparent instability of some of the intermediates, processing in the distal region of the primary transcript has been difficult to study in detail. Our results suggest, however, that transcripts exiting the 23S gene read through the distal 5S and cysteine tRNA genes and terminate somewhere further downstream (Figure 5). Using the 1300 nucleotide long 5′ end-labeled *Pst*I fragment, partial protection products of 36 nucleotides, 267 nucleotides and 340 nucleotides were observed (Figure 6A). These appear to correspond respectively to protection by (i) the 5′ end of mature cysteine tRNA, (ii) the 5′ end of the trailer sequence produced by cutting at or in front of the mature 5S sequence and (iii) the 5′ end of the longer trailer sequence produced by excision within the 23S inverted repeat, releasing the upstream precursor 23S rRNA sequence from the primary transcript.

Using the more distal 312 nucleotide long *Hin*PI restriction fragment as probe, it was possible to detect full-length transcripts exiting the 5S gene and extending into the GC-rich cluster commencing ~75 nucleotides downstream from the cysteine tRNA gene (Figure 6C). In addition, a number of shorter transcripts having ends within or near the cysteine tRNA sequence were

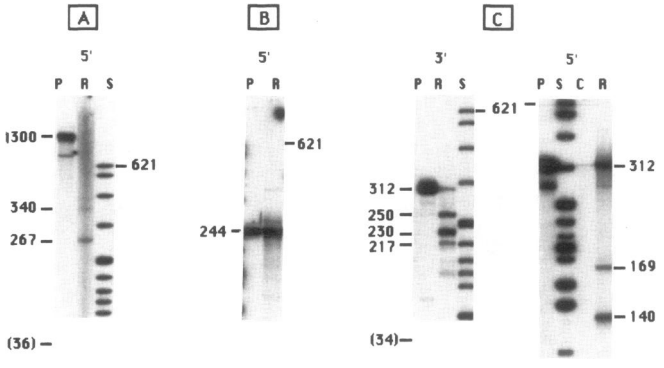


**Fig. 6.** Protection of end-labeled DNA fragments by RNA primary transcripts and processed intermediates derived from the region distal to the 23S rRNA gene. The three fragments and their respective nucleotide positions of 3′ or 5′ labeling on the (−) DNA strand are: (**A**) *Pst*I, 5′ at nucleotide D429; (**B**) *Hae*III, 5′ at nucleotide D414; (**C**) *Hin*PI, 3′ at nucleotide D251 and 5′ at nucleotide D563. Other details are as described in the legend to Figure 5.

detected. Using the 3′ end-labeled probe, partial protection products of 34, 217 and 230, and 250 nucleotides in length were observed. These correspond respectively to protection by (i) the 3′ end of mature 5S rRNA, (ii) the 3′ end of a leader transcript containing the entire tRNA sequence and terminating at its 3′ boundary, (iii) the 3′ ends of longer transcripts extending between 13 and 33 nucleotides beyond the 3′ tRNA boundary. With the 5′ end-labeled probe partial protection products of 140 and 169 nucleotides in length were observed. These appear to correspond respectively to protection by (i) the 5′ end of a trailer transcript containing the entire tRNA sequences beginning at the tRNA 5′ boundary and (ii) the 5′ end of a trailer transcript containing only the distal 46 nucleotides of the cysteine tRNA sequence.

A sequence which exhibits many of the features of a eubacterial factor independent transcription termination signal (i.e. a 13-bp inverted repeat followed by a $T_5$ sequence; Rosenberg and Court, 1979) is located in the 5S−cysteine tRNA intergenic space (see Figure 7). Termination at this site should have yielded a 3′ partial production product of ~81 nucleotides using the 312 nucleotide long *Hin*PI probe. The absence of a 3′ RNA terminus

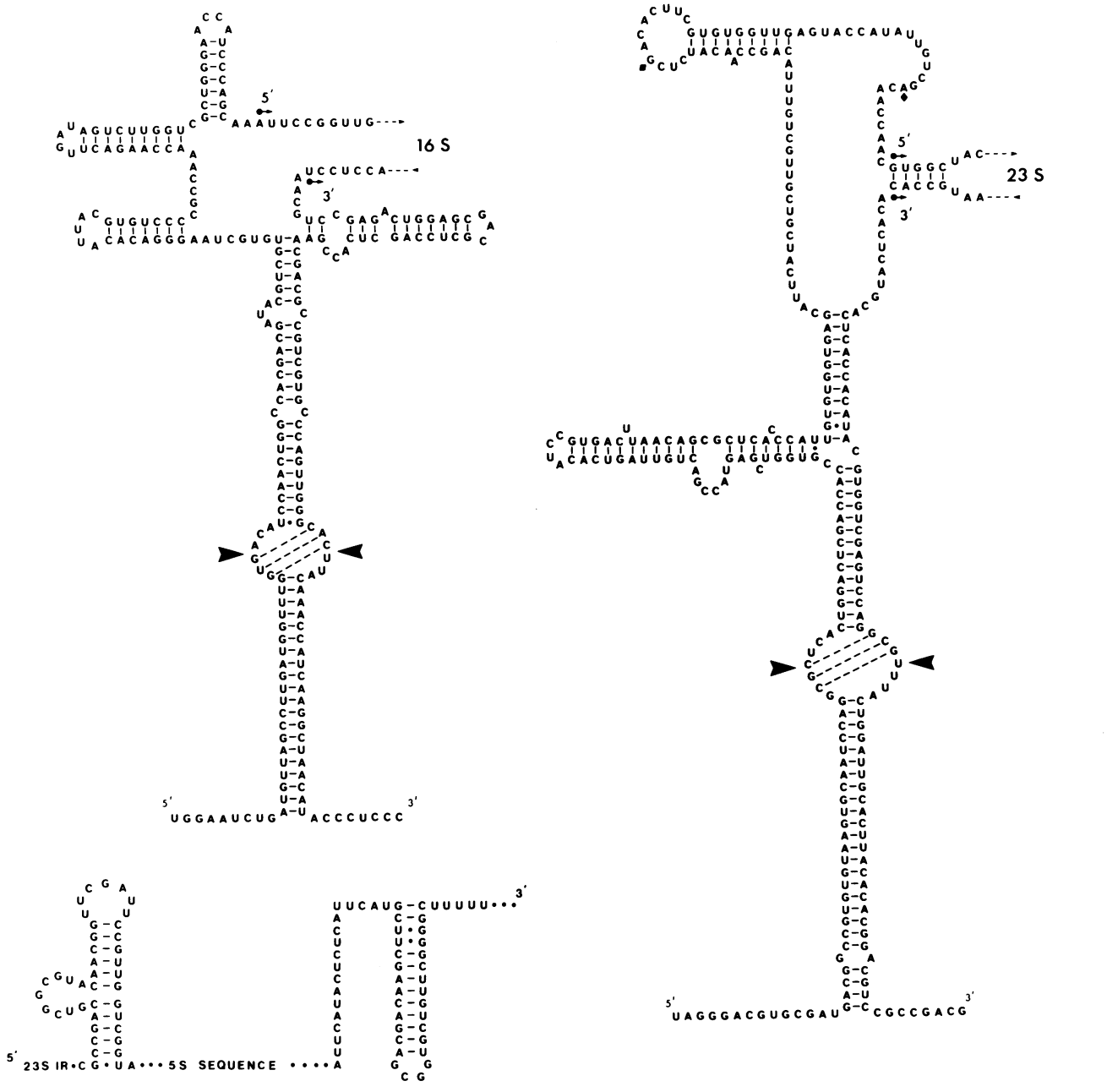


Fig. 7. Sequences surrounding the 16S, 23S and 5S rRNAs in the long primary transcript. The long, nearly perfect inverted repeat sequences flanking the 16S rRNA and 23S rRNA are illustrated in the upper left and right, respectively. Within these structures the 5′ and 3′ ends of the mature 16S and 23S molecules are indicated (●—►). The initial endonuclease cutting occurs on both sides of the helical stem near or within the major bulge (▶). Secondary structures are also possible in the 5′ and 3′ regions flanking the 5S rRNA sequence (lower left).

corresponding to this site was substantiated using the shorter 110 nucleotide long *Msp*I probe (Figure 5; autoradiogram not shown). Thus if transcription termination occurs at this site, the 3′ end must be rapidly and efficiently trimmed back to the 3′ boundary of mature 5S rRNA.

Co-transcription of the 5S and cysteine tRNA gene was also demonstrated using the 244 nucleotide long 5′ end-labeled *Hae*III probe (Figure 6B).

Taken together these results suggest the following: first, that transcripts exiting the 23S gene read through the 5S and cysteine tRNA gene and terminate somewhere beyond; second, that processing at the ends of the cysteine tRNA sequence is unordered and can occur first at either the 5′ or the 3′ boundary; third, that processing at the 3′ boundary of the 5S sequence does not commence until processing at the 5′ boundary of the cysteine has been initiated; fourth, that endonuclease cutting within the cysteine tRNA sequence may represent an alternate pathway for processing the primary transcript in which the cysteine tRNA sequence is not preserved.

The multiple 3′ transcript ends observed in the region immediately beyond the cysteine tRNA sequence could have been produced by transcription termination in this region or by endo- and exonuclease activity on a longer primary transcript. The intensities of the *Hin*PI full-length protection product and 169 and 140 nucleotide long partial protection products argue that many of the transcripts are continuous through this region (Figure 6C). The absence of detectable levels of a 5′-trailer transcript expected from an endonuclease event distal to the cysteine tRNA sequence

implies that the trailer sequence is unstable or alternatively that the observed 3′ ends are generated by transcription termination. We cannot rigorously exclude either of these possibilities.

*Putative activities required for RNA maturation*

In contrast to archaebacteria, much is known about the processing of the primary transcripts of the rRNA operons of *Escherichia coli* (for a review, see Gegenheimer and Apirion, 1981). Long, nearly perfect, inverted repeat sequences surrounding the *E. coli* 16S and 23S rRNA genes give rise to helical stems in the primary transcript with the 16S and 23S rRNA sequences protruding from the apical loops. These helical regions are substrates for RNase III; the enzyme introduces cuts at staggered positions on opposite sides of the stem, thereby releasing the precursor 16S and 23S sequences (Bram *et al.*, 1980; Gegenheimer and Apirion, 1981; Young and Steitz, 1978). The precise substrate characteristics of the enzyme have not been fully delineated although features such as primary sequence, unpaired bases, and non-Watson–Crick base interactions in addition to helicity are undoubtedly important. A number of additional activities are required to generate the mature 5′ and 3′ ends of 16S and 23S rRNA. The terminal step in 30S subunit assembly appears to be maturation of the 3′ end of the 16S rRNA.

In *H. cutirubrum*, excision of the precursor 16S and 23S rRNA appears also to occur by cutting on the opposite sides of an extended helical structure. The excision sites occur near or within a bulge in the helical stem. There is a degree of both primary sequence and secondary structure conservation between the two processing sites (Hui and Dennis, 1985). The sequences in the bulge regions are:

16S 5′–GUGACA . . . . CACUUA–3′
23S 5′–GCGCUCA . . . GCGUUUA–3′

where the underlined bases can potentially base pair across the bulge and the bold unpaired bases are conserved in the two structures (Figure 7). Similar sequence and potential structure have been noted in the inverted repeats surrounding the 16S and 23S rRNA genes of *Methanococcus vannielii, Halococcus morrhuae* and *Desulforococcus mobilis* and the 16S gene of *H. halobium* (Jarsch and Böck, 1985; Mankin *et al.*, 1984; Larsen *et al.*, 1986). Following excision of the precursors from the primary transcript, further processing at the 5′ and 3′ ends is required. Processing at the 5′ ends of the 16S and 23S precursor intermediates appears to be rapid whereas processing at the 3′ end of the 16S molecule proceeds more slowly, possibly reflecting a sequential 5′ to 3′ pathway of ribosome subunit assembly.

Secondary structural characteristics of 5S rRNA have been conserved throughout evolution (Fox, 1985). The precursor 5S rRNA in *E. coli* is excised from the primary transcript by RNase E which cuts first three nucleotides upstream and then three nucleotides downstream from the mature ends of the molecule (Gegenheimer and Apirion, 1981). The same processing order is followed in *H. cutirubrum* although there is little resemblance between primary sequence and secondary structure outside the coding sequence in the respective flanking regions. Considerable sequence homology exists between the flanking regions of the single *H. cutirubrum* (Daniels *et al.*, 1985; Hui and Dennis, 1985; Larsen *et al.*, 1986) and *H. morrhuae* 5S genes and the two *H. volcanii* 5S genes. Homology in the 5′ region preserves the secondary structure depicted in Figure 7. Homology in the 3′ regions is perfect for the first nine nucleotides and negligible thereafter. However, the non-homologous downstream regions in both *H. cutirubrum, H. volcanii* and *H. morrhuae* all exhibit inverted repeat symmetry followed by T-rich sequences; such sequences resemble the eubacterial *rho*-independent transcription termination signal. No 3′ transcript end associated with this structure was detectable in our S1 nuclease experiments. The precise role of these conserved secondary structural features in the 5′- and 3′-flanking regions of the 5S gene and their relationship to processing remain to be established.

In *E. coli* excision and maturation of the tRNA sequences from the rRNA primary transcript requires three additional enzymatic activities: RNase P, an endonuclease that cuts to produce the mature 5′ end of tRNA; RNase F, an endonuclease that cuts about five bases beyond the mature 3′ end of the tRNA sequence; RNase D, an endonuclease that trims the 3′ end of the precursor tRNA to the CCA terminus. Archaebacterial tRNA genes do not encode the CCA terminal sequence and therefore may utilize a somewhat different 3′ processing strategy. Multiple 3′ transcript ends have been detected in the 35 nucleotides downstream from the distal cysteine tRNA gene; the relationship of these ends to processing or termination remain to be established. Although the cysteine tRNA genes of *H. cutirubrum* and *H. volcanii* and the alanine tRNA genes of the *H. cutirubrum* and *M. vannielii* exhibit substantial sequence conservation, there is little homology in the 5′- and 3′-flanking sequences. This suggests that processing signals may be located within the tRNA coding sequence.

Two-dimensional gel analysis of total 5S tRNA of *H. cutirubrum* indicates that there are about six tRNAs per ribosome and that the amounts of none of the tRNA species are stoichiometrically equivalent to 5S RNA (Chant *et al.*, 1986 and unpublished results). This suggests that the processing efficiency of the alanine and cysteine tRNAs may be reduced compared with rRNA. A trailer RNA species with a 5′ end located ~29 nucleotides inside the cysteine tRNA sequence (140 nucleotide long protection product in Figure 6C) has been detected and may represent an intermediate in an alternate processing pathway in which the tRNA sequence is not preserved. Weak protection by an apparent leader RNA with a 3′ end within alanine tRNA sequence has also been observed (150 nucleotide long product in Figure 4C, 3′).

In summary, our results show that rRNA gene cluster of *H. cutirubrum* is transcribed as a long primary transcript. The direct repeat sequences in the 5′ region in front of the 16S gene correspond to the 5′ ends of the primary transcripts and probably represent RNA polymerase initiation sites. The inverted repeats surrounding the 16S and 23S rRNA sequences are utilized as processing sites for excision of precursor sequences from the primary transcript. Similar co-transcription and processing pathways probably exist in other archaebacterial species where the 16S, 23S and 5S genes are linked. In the sulfur-dependent thermoacidophiles the 5S gene is unlinked to the 16S and 23S genes; transcription and processing in these species has not yet been examined.

## Materials and methods

*H. cutirubrum* was cultured in defined amino acid medium supplemented with glycerol (0.4%) at 37°C in a rotary water bath shaker (Chant *et al.*, 1986). Total cellular RNA, including nascent rRNA, was isolated from exponential phase cells by phenol extraction of an SDS lysate (Dennis, 1985). Control RNA was isolated from exponential phase glucose grown *E. coli* C600 (Dennis, 1977).

Restriction enzyme fragments utilized as probes in Sl nuclease protection experiments were purified by electroelution from acrylamide gels (Maniatis *et al.*, 1982). The source of DNA was plasmid p4W, a pBR322 derivative, carrying a 7-kb *KpnI-BglII* fragment from phage λHc4 inserted between the *Eco*RI and *Bam*HI sites of pBR322 (Hui and Dennis, 1985). Restriction fragments were 5′ end-labeled with T4 polynucleotide kinase and [$\gamma$-$^{32}$P]ATP and 3′ end-labeled

with the Klenow fragment of DNA polymerase I and the appropriate [$\alpha$-$^{32}$P]dNTP (Maniatis *et al.*, 1985).

Nuclease Sl transcript mapping experiments were carried out as described previously (Dennis, 1985). Briefly, between $10^4$ and $10^5$ d.p.m. of end-labeled DNA fragment was co-precipitated with 5 $\mu$g of either *H. cutirubrum* or *E. coli* RNA. The dried pellet was resuspended in 20 $\mu$l of 80% formamide hybridization solution, denatured for 15 min at 80°C and allowed to hybridize for >3 h at 60°C. The hybridization products were digested with Sl nuclease (200–400 units/ml, P.L. Biochemicals) at 20 or 30°C for 30 min in the presence of single-stranded M13 DNA (20 $\mu$g/ml). The S1-resistant molecules were precipitated with isopropanol and electrophoresed on an 8% DNA sequencing gel. The gels were dried on filter papers and subjected to autoradiography. Control experiments with *E. coli* RNA indicate that there was negligible re-hybridization of the DNA probes. Molecular length standards were *Msp*I-cut pBR322 end-labeled with Klenow enzyme and [$\alpha$-$^{32}$P]dCTP.

## Acknowledgements

This work was supported by a grant from the Medical Research Council of Canada. We thank D.De Jong-Wong for technical assistance.

## References

Bram,R., Young,R. and Steitz,J. (1980) *Cell*, **19**, 393–401.
Chant,J., Hui,I., De Jong-Wong,D., Shimmin,L. and Dennis,P.P. (1986) *System. Appl. Microbiol.*, **7**, 106–114.
Daniels,C., Hofman,J., MacWilliam,J., Doolittle,W.F., Woese,C., Luehrsen,K. and Fox,G. (1985) *Mol. Gen. Genet.*, **198**, 270–274.
Daniels,C., Gupta,R. and Doolittle,W.F. (1985) *J. Biol. Chem.*, **260**, 3132–3134.
Dennis,P. (1985) *J. Mol. Biol.*, **186**, 457–461.
Dennis,P. (1977) *J. Mol. Biol.*, **115**, 603–625.
Fox,G. (1985) In Woese,C. and Wolfe,R. (eds), *The Bacteria*. Academic Press, NY, Vol. 8, pp. 257–310.
Gegenheimer,P. and Apirion,D. (1981) *Microbiol. Rev.*, **45**, 502–541.
Hofman,J., Lou,R. and Doolittle,W.F. (1979) *Nucleic. Acids Res.*, **7**, 1321–1333.
Huet,J., Schnabel,R., Sentenac,A. and Zillig,W. (1983) *EMBO J.*, **2**, 1291–1294.
Hui,I. and Dennis,P. (1985) *J. Biol. Chem.*, **260**, 899–906.
Jarsch,M. and Böck,A. (1985) *Mol. Gen. Genet.*, **200**, 305–312.
Kaine,B., Gupta,R. and Woese,C. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3309–3312.
Larsen,N., Leffers,H., Kjems,J. and Garrett,R. (1986) *System Appl. Microbiol.*, in press.
Maniatis,T., Fritsch,E. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, NY.
Mankin,A., Teterina,N., Rubstov,P., Baratova,L. and Kagramanova,V. (1984) *Nucleic Acids Res.*,**12**, 6537–6546.
Matheson,A. (1985) In Woese,C. and Wolfe,R. (eds), *The Bacteria*. Academic Press, NY, Vol. 8, pp. 345–377.
Rosenberg,M. and Court,D. (1979) *Annu. Rev. Genet.*, **13**, 319–353.
Woese,C. (1982) *Zentralbl. Bakt. Hyg. I Abt Orig.*, **C3**, 1–17.
Woese,C. and Fox,G. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5088–5090.
Young,R. and Steitz,J. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 3593–3597.

*Received on 14 January 1986*