3

Genes for Stable RNAs and Their Expression in *Archaea*

Michael Thomm and Winfried Hausner

3.1 Introduction

The investigation of stable RNA sequences has led to the discovery of the *Archaebacteria* as a second procaryotic line of descent (Woese and Fox, 1977). Archaebacteria are no more related to typical bacteria than to the eucaryotic cytoplasms. Novel designations have been proposed to express this tripartite division of the living world, thus rejecting the conventional procaryote–eucaryote dichotomy (Woese et al., 1990). According to this proposal, life on earth can be seen as comprising three domains, called the *Bacteria* (formerly eubacteria), the *Archaea* (archaebacteria), and the *Eucarya* (eucaryotes). Two major phylogenetic branches (kingdoms) can be distinguished within the domain of *Archaea*: the coherent kingdom of *Crenarchaeota* (formerly sulfur-metabolizing thermophiles) and the phenotypically diverse second kingdom of *Euryarchaeota*, comprising the three orders of methanogens (Balch et al., 1979): the order *Thermococcales* (Zillig et al., 1987), the sulfate reducer *Archaeoglobus* (Achenbach-Richter et al., 1987), the extreme halophiles, and the genus *Thermoplasma* (Woese, 1987).

The genes for stable RNAs from representatives of all phylogenetic groups of *Archaea* have been cloned and their arrangements and sequences determined. Thus, enough structural data have been accumulated to allow a significant comparison of the molecular organization of tRNA and rRNA genes in bacterial, eucaryotic, and archaeal cells. This comparison will constitute the first part of this chapter.

Although DNA sequences upstream and downstream from numerous rRNA and tRNA genes of *Archaea* have been established, little was known about the mechanisms and transcription signals regulating the expression of stable RNA genes in *Archaea*. The major reason for this paucity of information has been the lack of genetic transfer procedures and in vitro transcription systems that would allow testing and refining the predictions inferred from structural studies in functional assays. However, a convincing proposal for two promoter motives of stable RNA genes in *Methanococcus* has been derived from sequence analyses and transcription mapping experiments (Wich et al., 1986a). These sequences comprise an AT-rich sequence, called box A, located between position −40 to −20 relative to the transcription start site, and a second motif at the transcription start site (box B). Footprinting (nuclease protection) experiments showed that some of these conserved sequences are bound by the purified RNA polymerase from *Methanococcus vannielii* (Thomm and Wich, 1988). From these footprinting experiments and from the sequence analyses of RNA genes from methanogens and the thermophile *Sulfolobus*, the octanucleotide TTTA T/A ATA (TATA box) has been inferred as a general promoter element for stable RNA genes in *Archaea* (Thomm and Wich, 1988; Reiter et al., 1988a).

A modified version of this motif was found

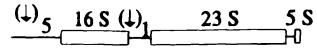
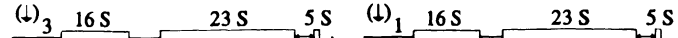
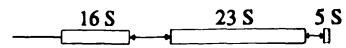
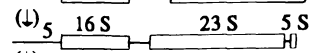
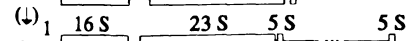
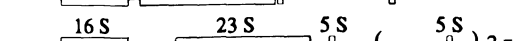
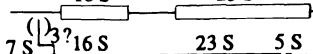
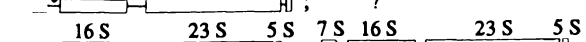
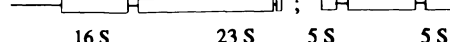
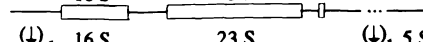
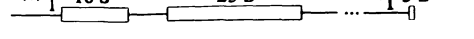
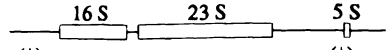
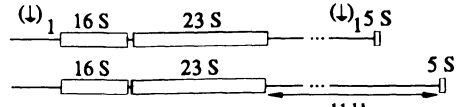
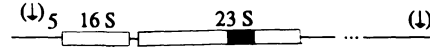
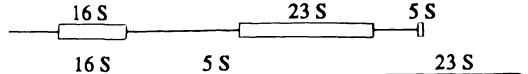
at the same location upstream from stable RNA genes of extreme halophiles (Mankin and Kagramanova, 1988; Thomm and Wich, 1988). Further footprinting experiments and sequence analyses supported the conclusion that this sequence also comprises a constituent of the promoters of archaeal protein-encoding genes (Brown et al., 1988; Reiter et al., 1988a; Thomm et al., 1988). Although the RNA polymerase of *Methanococcus* binds to the promoter, all purified RNA polymerases of *Archaea* are unable to initiate transcription at the correct site in vitro. However, cell-free transcription systems allowing the expression of tRNA genes of *Methanococcus* (Frey et al., 1990), rRNA genes of *Sulfolobus shibatae* (Hüdepohl et al., 1990), and a protein-encoding gene of *Methanobacterium thermoautotrophicum* (Knaub and Klein, 1990) have been described. The *Methanococcus* and *Sulfolobus* systems both initiate at the same site in vitro and in vivo. The *Methanobacterium* system starts transcription at a box B-like sequence 10 nucleotides upstream of the in vivo initiation site. The low efficiency and unusual start site of this in vitro system suggest that it lacks an essential component. The availability of specific transcription systems offers the opportunity to obtain biochemical evidence for the significance of conserved DNA sequences upstream and downstream from archaeal genes.

We summarize here some of our experiments defining the DNA sequences that promote and cause the termination of transcription in *Methanococcus vannielii*. These and other experiments (Hausner et al., 1991) demonstrate that the TATA box at −25 and a second signal at the transcription initiation site are indispensable for initiation of transcription. Accurate cell-free transcription in *Methanococcus* and *Sulfolobus* is mediated by soluble transcription factors (Frey et al., 1990; Hüdepohl et al., 1990); this is reminiscent of eucaryotic transcription systems. The similarities and differences of stable RNA transcription in *Archaea*, *Bacteria*, and *Eucarya* are discussed. Stable RNA genes are highly expressed in vivo, and because only a single type of RNA polymerase appears to be present in archaeal cells, the analysis of transcription in these cells may provide a basis for understanding the mechanism of archaeal transcription in general.

3.2 Organization of Stable RNA Genes

Within the *Euryarchaeota*, ribosomal RNA genes are arranged in the sequence 5′-16S-tRNA^{Ala}-23S-5S-3′ (Figure 3.1). The tRNA gene in the intercistronic spacer between 16S and 23S RNA is missing in all *Crenarchaeota* investigated so far. According to this feature, organisms that show a thermophilic phenotype, such as *Thermococcus* and *Archaeoglobus*, are also clearly linked with their phylogenetic relatives, the methanogens and halophiles (Achenbach-Richter and Woese, 1988). In *Desulfurococcus mobilis*, *Thermoproteus tenax*, and *Thermophilum pendens*, the 5S genes are unlinked from the 16S/23S genes and are transcribed from an independent promoter (Neumann et al., 1983; Kjems and Garrett, 1987; Kjems et al., 1990). Additional unlinked 5S genes have been observed in *Sulfolobus* strain B12 (Reiter et al., 1987; strain B12 has been described recently as a new species *Sulfolobus shibatae*; Grogan et al., 1990), *Thermococcus* (Neumann et al., 1983), and *Methanococcus* (Jarsch et al., 1983; Wich et al., 1987b). The additional 5S genes of *Methanococcus vannielii* and *M. voltae* are located within tRNA operons (Wich et al., 1984, 1987b). One rRNA operon of *Methanothermus fervidus* and *Methanobacterium thermoautotrophicum* is linked to a 7S gene (Haas et al., 1990; see Figure 3.1, and also following).

In *Thermoplasma*, the genes for 16S, 23S, and 5S rRNA are physically separated by 1.5 to 7.5 kb (Tu and Zillig, 1982) and transcribed from independent promoters (Ree and Zimmermann, 1990), indicating a unique arrangement of ribosomal RNA genes in this organism into three transcription units. The number of rRNA operons varies from one in the *Crenarchaeota* to four in *Methanococcus vannielii* (see Figure 3.1). The domain *Bacteria* shares with *Archaea* the general organization of rRNA genes. In the *Eucarya* the 5S gene is separated. However, a separate 5S rRNA gene has also been disco-

organism	gene structure	copy number
ARCHAEA		
Halobacterium halobium		1
Halobacterium marismortui		1 each
Halobacterium volcanii		1
Halococcus morrhuae		1
Methanococcus vannielii		4
Methanococcus voltae		1
Methanobacterium thermoautotrophicum		2*
Methanothermus fervidus		1 each
Sulfolobus acidocaldarius		1
Sulfolobus shibatae		1
Thermococcus celer		1
Thermofilum pendens		1
Thermoproteus tenax		1
Desulfurococcus mobilis		1
Desulfurococcus mucosus		1
Thermoplasma acidophilum		1

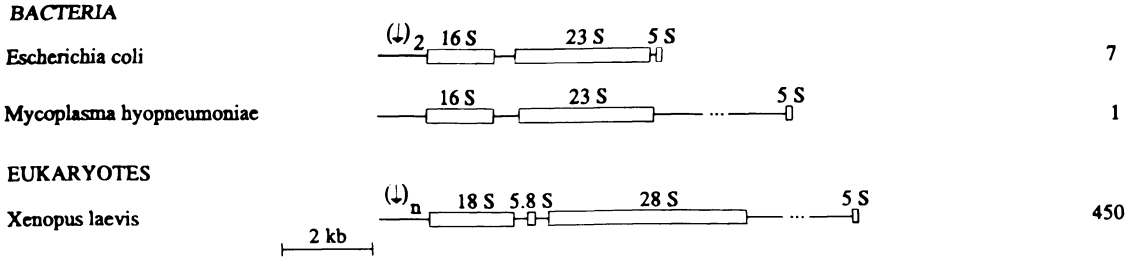


FIGURE 3.1 Arrangement and transcription of archaeal rRNA genes. Sequences encoding mature rRNA species are boxed. Dots indicate that distance of unlinked 5S gene to rRNA operon has not been precisely determined. Transcription start sites are indicated by (↓). Indices show number of mapped initiation sites; in *Xenopus*, Index n indicates multiple sites. Distance between 16S and 23S rRNA derived from mapping data has been corrected using DNA sequencing data provided by Achenbach-Richter and Woese (1988). Location of intron in 23S RNA of *Desulfurococcus mobilis* is indicated by black box.
*The organization of the second operon in *Methanobacterium thermoautotrophicum* has not yet been investigated, and interpretation of the mapping experiments for transcription start sites is difficult (see discussion in text).

vered in *Mycoplasma* (Taschke et al., 1986; Figure 3.1). Eucaryotic 18S, 5.8S, and 28S rRNA genes are cotranscribed as a large 40S or 45S precursor in lower and higher *Eucarya*, respectively. The 5.8S rRNA of *Eucarya* is homologous to the 5′ end of the 23S rRNA of *Bacteria* and *Archaea*.

Desulfurococcus mobilis shares with some lower *Eucarya* the presence of an intron within the 23S rRNA gene (Kjems and Garrett, 1985). The splicing process is similar to that of the class III introns of eucaryotic tRNAs. However, in contrast to these introns, which are linear and rapidly degraded after the splicing event, the excised *Desulfurococcus* intron circularizes in a manner similar to that of self-splicing introns (Kjems and Garrett, 1988). Because the *Desulfurococcus* intron is located in a "hot spot" intron site of eucaryotic rRNAs and is spliced by a mechanism similar to that of eucaryotic tRNA introns, both rRNA and tRNA introns of *Eucarya* may have a common origin in the introns of *Archaea*.

Transfer RNA genes in *Archaea* are organized in operons or single genes and may be located within rRNA operons, in the spacer region between 16S and 23S RNA and as trailers downstream from the 5S gene. This is the typical situation encountered in *Bacteria*. In *Eucarya*, tRNA genes are clustered, but neighboring genes are rarely cotranscribed into multimeric precursors. Transfer RNA genes hitherto have not been observed within an rRNA operon in *Eucarya*. The absence of a tRNA gene in the 16S/23S RNA spacer region of extreme thermophiles has been discussed as an important argument supporting the phylogenetic tree of Lake (1989). According to this proposal, these organisms (called eocytes by Lake) are more closely related to *Eucarya* than methanogens and halophiles, which cluster together with (eu)bacteria (Lake, 1988). However, the 16S/23S RNA spacer region of *Mycobacterium* (Suzuki et al., 1988a) and *Streptomyces* (Suzuki et al., 1988b) also does not contain a tRNA gene. Because this feature is also a common property of the *Actinomycetes*, an important group in the domain *Bacteria*, the presence or absence of a tRNA gene does not argue for a specific relationship of *Crenarchaeota* to *Eucarya*. No molecular feature of stable RNAs has yet been detected that is absolutely specific for one of the two archaeal kingdoms, and thus far, no phenotypic properties can be defined that clearly indicate a closer relationship of one of the two kingdoms of *Archaea* to *Eucarya* (see following). Also, almost all data accumulated thus far argue for the monophyletic nature of *Archaea*.

Some tRNA genes from *Sulfolobus solfataricus* (Kaine et al., 1983), *Haloferax volcanii* (Daniels et al., 1985), *Thermoproteus tenax* (Wich et al., 1987a), and *Thermofilum pendens* (Kjems et al., 1989) contain an intron. Most archaeal tRNA introns are located in the anticodon loop, as they are in eucaryotic nuclear tRNAs. The tRNA intron of *T. pendens* is located in the variable loop of the tRNA precursor. No intron has ever been detected at this position of the tRNA molecule. A $tRNA^{Trp}$ intron endonuclease has been purified from *Haloferax volcanii* (Thompson and Daniels, 1988). Unlike eucaryotic intron endonucleases, this enzyme appears to be sequence specific and does not require a complete mature tRNA structure for substrate recognition. In stable RNA genes of bacteria, no introns have yet been detected.

Most tRNA genes of *Archaea* lack the 3′ terminal CCA sequence, which must be added posttranscriptionally. These two features previously have been seen in eucaryotic tRNA genes (Melton et al., 1980), some tRNA genes of *E. coli* bacteriophages, and in some chromosomal tRNA genes of *Bacillus* (King et al., 1986).

A striking similarity to eucaryotes is the presence, in *Archaea*, of a gene encoding a stable RNA of about 300 nucleotides. This 7S RNA gene has been found in all *Archaea* examined (Moritz and Goebel, 1985; Haas et al., 1990; Kaine, 1990). The 7S RNAs of *Archaea* and *Eucarya* share a very similar secondary structure. However, the homology in the primary sequence is limited to a hairpin structure. This specific domain of about 40 nucleotides is also conserved in the 4.5S RNA (114 nucleotides) of *E. coli* and the small cytoplasmatic RNA (271 nucleotides) of *Bacillus subtilis* (Struck et al., 1988; Kaine and Merkel, 1989; Haas et al., 1990). Thus, these small bacterial RNAs and 7S RNAs might be evolutionary homologs. However, their size and secondary

structure result in a greater structural resemblance between archaeal and eucaryotic 7S RNA molecules. In eucaryotic cells, 7S RNA is a major constituent of the signal recognition particle involved in translocation of secretory proteins (Zwieb, 1989). The function of 7S RNA in archaebacterial cell metabolism remains an intriguing question.

There are so many excellent reviews and original papers about the structural features of archaeal stable RNAs and the phylogenetic trees based on these sequences (Böck et al., 1986; Leffers et al., 1987; Woese, 1987; Brown et al., 1988; Kjems and Garrett, 1990) that they cannot be discussed in detail here. This chapter therefore focuses on the following aspects of the expression of stable RNA genes.

3.3 Transcription of Stable RNA Genes

Transcription initiation sites

All living cells require specific mechanisms to synthesize the tremendous amount of RNA required to constitute the RNA component of ribosomes. In bacteria, three factors seem to contribute to the high expression rate of rRNA genes. First, there are multiple copies of the genes (seven in *E. coli*; Kenerley et al., 1977). Second, the genes are expressed from tandem promoters (Young and Steitz, 1979). Finally, the RNA polymerase shows a high affinity to these promoters, most likely because they show high homology to the eubacterial consensus sequence (Hawley and McClure, 1983).

Eucaryotic cells contain 50 to 500 identical repeat units of rRNA genes per haploid genome, which are clustered at a distinct site of the chromosome and separated by a nontranscribed spacer region. In *Xenopus laevis* and probably in other eucaryotes, the spacer is composed mainly of repeated DNA sequences, some of which contain promoter-like structures. These reduplicated promoter sites are bound by transcription factors and give rise to short transcripts upstream from the true rRNA promoter (Moss, 1983). Because RNA polymerase I does not detach after transcription of both the 40S rRNA precursor and the short DNA sequences upstream from the gene, the multiple promoter sites appear to deliver the RNA polymerase to the primary promoter, thus ensuring a high expression rate for rRNA genes (Sollner-Webb et al., 1987).

A similar initiation, termination, and reinitiation mechanism has been demonstrated to occur upstream from the 16S/23S operon of *Desulfurococcus mobilis* (Kjems and Garrett, 1987). Four transcription initiation sites seem to direct the RNA polymerase to the primary initiation site located 134 bp upstream from the mature 16S rRNA.

Multiple transcription initiation sites have also been mapped upstream from the RNA operon of *Halobacterium cutirubrum* (Dennis, 1985) and *H. halobium* (see Figure 3.1; Mankin and Kagramanova, 1986). These two organisms are closely related and should be subsumed into the species *H. salinarium* (Larsen and Grant, 1989). However, in contrast to *Desulfurococcus*, the transcripts initiating at distant sites of halophiles do not terminate upstream from the DNA region encoding the 16S rRNA. Thus, this readthrough from several upstream promoters to the terminator of the operon resembles the mechanism encountered in *bacteria* (Boros et al., 1983). Multiple transcription initiation sites have also been observed upstream from the rRNA operon of *Halococcus morrhuae* (Larsen et al., 1986) and from one of the two operons of *Halobacterium marismortui* (Mevarech et al., 1989). The three transcription start sites mapped upstream of one of the two rRNA operons of *Methanobacterium thermoautotrophicum* have not yet been clearly established (Ostergaard et al., 1987). The situation is further complicated by the presence of a 7S RNA gene immediately upstream of this operon that was only noticed later (Haas et al., 1990). Only a single transcription start site exists upstream from rRNA genes from *Methanococcus vannielii* (Wich et al., 1986a), *Sulfolobus shibatae* (Reiter et al., 1987), *Thermofilum pendens* (Kjems et al., 1990), *Thermoplasma acidophilum* (Ree and Zimmermann, 1990), and one operon of *Halobacterium marismortui* (Mevarech et al., 1989; see also Figure 3.1). Therefore, multiple start sites up-

stream from rRNA operons cannot be considered a general archaeal mechanism.

An additional putative transcription start site has been located in the spacer region upstream for the 23S RNA of *Halobacterium salinarium* by S1 nuclease mapping experiments (Mankin and Kagramanova, 1988). A consensus promoter sequence is located at the correct distance upstream from this nucleotide, suggesting that this S1 signal is caused by initiation of transcription and not by processing of the rRNA precursor. This additional promoter may help to adjust cellular levels of RNAs that are located far downstream of the primary promoter. These RNAs might otherwise be expressed at lower levels because of premature termination of transcription. It is unclear whether a similar mechanism operates in other organisms.

Only a few transcription start sites of archaeal tRNA genes and operons have been mapped (Wich et al., 1986a, 1987b; Kjems and Garrett, 1988). These data indicate the presence of a single initiation site. In *Bacteria*, tRNA genes can be expressed both from a single and from tandem promoters (Caillet et al., 1985). The transcription of eucaryotic tRNA and 5S rRNA genes is controlled by an intragenic promoter (see following).

Promoter sequences

The transcription initiation sites upstream from archaeal stable RNA genes have been located by S1 mapping or primer-extension experiments. A major disadvantage of the protocols used by most investigators was that they did not allow a distinctive discrimination between initiation start sites and processing sites (especially when applied to stable RNA genes). To determine the 5′ end of a primary transcript, it is important to demonstrate that: (1) the RNA initiates with a ribonucleoside triphosphate; and (2) the same initiation site is used when the RNA is synthesized in vitro in the absence of the processing machinery of the cell.

Wich et al. (1986a) were the first to locate the transcription start site of a primary transcript upstream from a rRNA operon using a guanosyl transferase capping experiment. They proposed two conserved DNA sequences as possible promoter signals for stable RNA genes of *Methanococcus*: the "box A" sequence ACCGAAA-TTTATATA-TA, extending from position 20–40 upstream from the 5′ end of the primary transcript, and the "box B" motif TGCAAGT, at the transcription start site. The footprint of the *Methanococcus* RNA polymerase extends from position −30 to +20 relative to the transcription start site (Thomm and Wich, 1988). Hence, the 3′ part of the "box A" of Wich et al. (1986a), the octanucleotide TTTA T/A ATA, is located within the RNA polymerase-binding site. This octanucleotide shows striking homology to the TATA box of eucaryotic promoters of protein-encoding genes in both location and sequence (Corden et al., 1980). A very similar sequence has been found upstream from primary transcripts of *Sulfolobus* genes (Reiter et al., 1988a). Thus, the DNA sequences upstream from transcription initiation sites appear highly conserved among phylogenetically distant *Archaea*. Sequence analyses of the DNA region upstream from stable RNA genes of a variety of further genera from both phylogenetic kingdoms of *Archaea* confirm this conclusion (Figures 3.2–3.4). However, although this strict conservation argues for the importance of these sequences in evolution, their function as promoter signals has not yet been demonstrated.

One way to investigate the significance of a conserved DNA sequence is to alter it in vitro and determine the effect of the mutation on the biological function. We used a cell-free transcription system (Frey et al., 1990) to define archaeal promoter sequences in a functional assay. As template for these experiments, we used the $tRNA^{Val}$ gene of *M. vannielii*, which has a −25 region with perfect homology to the consensus promoter sequence; at the 3′ end, it has an oligo-dT sequence, which has been proposed by Wich et al. (1986a) to be a potential terminator signal (Figure 3.5). Analysis of DNA deletion clones showed that the expression rate of this tRNA is not dramatically reduced after the DNA region from −590 to −35 is removed (Thomm et al., 1990; Hausner et al., 1991). However, removal of nucleotides extending into the TATA box or beyond leads to a complete inactivation of this template (Hausner

a)

TACCTAAAACAATACATATTACAACACGTTTTCATATTATgCAAATC	rRNA1	Mc.	vannielii	(Wich et al., 1986a)
TACCTAAAACAATACATATTACAACACGTTTTCATATTATgCAAATC	rRNA2		"	"
AACCGAAATATTTATATACTAGAATACCCTTCCTATACTATgCTCTT	tRNA-op.		"	(Wich et al., 1986b)
TACCGAAAACTTTATATATTATAACACTAGTATTCAGTATgCGAACA	5S/tRNA-op.		"	"
CACCGAAAAGTTTATATATCATGAATACTATGTTTAGTTTGCTCTCA	tRNA Val		"	"
CACCGAAAACTTTATATACTGTTTATTATGTATTTCATTTGGAAGTT	tRNA Thr		"	"
TACCGAAAACTTTATATAATATAATTTCAATCTAATTAATGCAAGTC	tRNA Gln		"	"
TGTAAAAAGGTTTATATAGTAGAATGTTAATTGTTATTGTGCGGTCA	tRNA-op.		"	"
TTACAAAAAGTATATATACTAAGAAAGATATGCTTAATTGTGGATTT	tRNA Phe		"	"
GATGGAAACATTTATATAGTATAATGGTGTTAGTCATTGTGCAAGTA	5S/tRNA-op.	Mc.	voltae	"
AATAGTAAACTATATAAGCTAGAACAAGTTATGTAATATTGGCTAGG	7S RNA		"	(Kaine and Merkel, 1989)
TATCAAAAAATTTAAATAAGATTGAAAAATAAAATATAAAATGGCAG	7S RNA	Mt.	fervidus	(Haas et al., 1990)
GTCCGAAAACTTTATATATGAAAAATTCAAAGGTAAATTATAGCTAA	tRNA-op.		"	(Haas et al., 1989)
AATCGAAAAATATAAATATGGTTTTATCTAATCTATCCATTAGCTTT	tRNA-op.		"	"
------AAGCTTTAAATAATAGAGCTGCCCTACAGTAATGG------	rRNA-op.	Mth.	soehng.	(Eggen et al., 1990)
TGCCATAACCTTTATAACTCACTGTGACAATACTTTATTTGGTGGGC	7S RNA	A.	fulgidus	(Kaine, 1990)
GTGCCAAAACTTTATATATCGACGGGGGAATAGAGTAACTGGCGGCG	7S RNA	Tc.	celer	"
CTTCGAAAGTTATATATACTGATTTGCTATTCTTTACTTtgCACATA	16S RNA	Tp.	acidophilum	(Ree and Zimmermann, 1990)
TCACGAAAATCTTATATAGATGTGTTCTATATAGTGTtCGGCAACG-	5S RNA		"	"
GATCAAAATGCTTATATCCCTCTTAATGATATAGTCCATaCACGCTT	23S RNA		"	"

b)

A_{17}	A_{20}	A_{19}	A_{12}	C_9	T_{15}	T_{14}	T_{20}	A_{20}	T_{15}	A_{20}	T_{18}	A_{17}
T_2		G_1	T_3	G_4	A_2	A_6			A_3		A_2	C_2
$^{-1}$			G_3	A_4	C_2				C_2			G_1
			C_2	T_3	G_1							
A	A	A	A	N	T	T	T	A	T	A	T	A

consensus methanogens
(usually low and intermediate GC)

FIGURES 3.2 to 3.4 The promoter sequences in *Archaea* are highly conserved. Figure 3.2 shows the results for methanogens and euryarchaeota excepting halophiles, Figure 3.3 for halophiles, and Figure 3.4 for the *Crenarchaeota*. **(A)** DNA sequences upstream from archaeal tRNA/rRNA and 7S genes have been aligned to yield maximal homology, with TATA box at −25. This promoter element has been identified by footprinting and cell-free transcription experiments using purified components from *Methanococcus* (Thomm and Wich, 1988; Thomm et al., 1990; Hausner et al., 1991). The consensus octanucleotide (methanogens and halophiles) and hexanucleotide (*Crenarchaeota*) are boxed. Lowercase letters indicate transcription start sites. *Euryarchaeota Archaeoglobus*, *Thermococcus*, and *Thermoplasma* are listed together with methanogens in Figure 3.2. In Figure 3.3, P_i indicates putative promoter in spacer region between 16S and 23S RNA in *Halobacterium salinarium*. **(B)** Conserved nucleotides in −25 region of archaeal promoters. Subscripts indicate base frequency at each position. Consensus derived is shown at bottom. DNA region that has been identified as being most important for cell-free transcription in *Methanococcus* is boxed (Figure 3.2). Corresponding DNA region in halophiles (Figure 3.3) and *Crenarchaeota* (Figure 3.4) is also boxed. Note modification of consensus sequence in halophiles, which can be correlated with their high GC content. (From Thomm et al., 1989.)

a)

```
CGCCGACATATTTATCCTCCGGCCTTGTGTTTGCATCCCaCGAAGAA   rRNA P1   H. salinarium  (Mankin and Kagramanova, 1988)
GGCGAAACTGCTTACAACGCCCCAACCCAACACGCACCCgCGTGGGT   rRNA P2        "                    "
TCGACGGTGTTTTATGTACCCCACCACTCGGATGAGATGCGAaCGAC   rRNA P3        "    (Dennis, 1985; Mankin et al., 1986)
GTCCGATGCCCTTAAGTACAACAGGGTACTTCGGTGGAATGCGAaCG   rRNA P4        "                    "
ATTCGATGCCCTTAAGTAATAACGGGTGTTCCGATGAGATGCGAaCG   rRNA P5        "                    "
ATTCGATGCCCTTAAGTAATAACGGGCGTTACGAGGAATTGCGaACG   rRNA P6        "                    "
ATTCGATGCCCTTAAGTAATAACGGGGCGTTCGGGGAAATGCGAaCG   rRNA P7        "                    "
GATCGTGTCCCTTAAGTGGGAGACGGGGCAACGATGAATCgCGACGA   rRNA Pi        "          (Mankin and Kagramanova, 1988)
CCGAAAGGCCCTTAAGAACGACCCGGGTAGGATGAGATGGACTAGGC   7S  RNA       "          (Moritz and Goebel, 1985)
TCGAAACCCCTTTAAGAAAAATCGCCATACGAGAGAGTGCAGACAGA   tRNA Trp   Hf. volcanii   (Daniels et al., 1985)
AAACAGAAGTCTTAACATAGCCAGACTCGTTTGTTGATCATG-----   tRNA Met       "          (Daniels et al., 1986)
AAAGGAAAGTCTATTTACCCACCGGCAGTACGAGAGATTGCAAGGG-   tRNA Lys       "                    "
ATCGAAACGGATTAAACTATCCGCGAGAGAGGCAACAATGGAAGCC-   tRNA Ser       "                    "
GATTCGAAAGCTTAAATTGTACCCGGACAACGGAGAGATGCGTCCGA   tRNA Val       "                    "
CACCGTCAGGCTTAAGCACAAGACCGGGATATCCAGTAACTGCGCC-   tRNA Cys       "                    "
-------AAGCTTAAATACAACCAGCACACAACAGGAAGTTGAGCCC   tRNA Met_m    "          (Datta et al., 1989)
TTCCGACGGGTTTATCCGTTACCCGGGATTCCGAATGGAAATGCGAA   rRNA P1   H. morrhuae    (Larsen et al., 1986)
ATCCGACGCCCTTAATTGGTACAGGGCACTCGGATGGAATGCAGAAA   rRNA P2        "                    "
CTTCGAAGGGTTTATACCCTCGAACGGTGTACGAAGAGATCCGAAGG   rRNA P3        "                    "
AGACCGTCCATTTATATACTTTTTTCCCATCGGATGTAATgCGAAGG   rRNA P    H. marismortui (Mevarech et al., 1989)
CTTCGACGGCGTTAAGTGTGGCTCACCCATCGGAATGAAATgCGAAC   rRNA P1        "                    "
TTCCGACGCCCTTAAGTGTAACAGGGCGTTCGGAATGAACgCAAAGG   rRNA P2        "                    "
ATCCGACGCCCTTAAGTGTAACAGGGTGCTCGGAATGAACgCGAACG   rRNA P3        "                    "
```

b)

```
G14 A16 C8  G11 C11 C11 C15 T23 T22 A22 A16 G12 T13 A10
A5  G4  A6  A6  G8  G7  T6      A1  T1  T6  A6  A5  G6
C3  T2  T5  C4  T2  T3  A1                  C1  C3  C5  T4
-1  -1  G3  T2  A2  A2  G1                      T2      C3
        -1

G   A   N   G   C   C   C   T   T   A   A   G   T   A
                G   G                                   consensus halophiles
                                                            (high GC)
```

FIGURE 3.3 See Figure 3.2 for caption.

a)

```
GCGAAAAATTTTTAATTTAGGGTGTTTTAGGATGGTCgCGCCTTAATTGTTTGT   rRNA          Tp. tenax   (Wich et al., 1987)
AGCGAAAAAATTTAAATCGGTGAGTAAGTACGCTCgGGGCCGGTAGTCTAGCGG   tRNA Ala           "              "
ACAAAAGCTTTTTAAATTCGCGCAAAGCTTAGACCTaGCGGGGTAGGCCAGCTA   tRNA Met           "              "
GGCTTAAAGCTTAAAATATCCTGTCATATAACGAGTTGGGCCGGTAGTCTAGCG   tRNA Ala           "              "
GCTGAAAAATTTTAAACTGAGCAGTTTATATCAGAGACGGCGGGGGTGCCCGAG   tRNA Leu           "              "
GAGAAAACATTTTAATCCTGAGGAGAAAATACTGGACAGGCGGGGGTGCCCGAG   tRNA Leu           "              "

ATGCTAAAGGTTTATTACCCAGGAAGTATTCCGGTCATgGGGGGTTACGAAGCC   tRNA Met    Tf. pendens  (Kjems et al., 1990)
CGGGAAAAGCTTTTAAGCATGCCTTTTTACTTCCTTCTAgAGGCTCAGCGGCCG   tRNA Gly           "              "
AAGCATAATATTCATATAACCCCCCGTTTACTAACTAGATTgCCGCCATGGGCA   rRNA               "              "
TAGTTAATTTTTTATATGTGTTATGAGTACTTAATTTTgCCCACCCGGCCACAG   5S rRNA     S. shibatae  (Reiter et al., 1987)
AGAAGTTAGATTTATATGGGATTTCAGAACAATATGTATAATgCGGATGCCCCC   rRNA               "              "
GGCATAACTTTTTAAAAGGTAACTATTTTATTATGTTATAGTGGGCCCGTAGCT   tRNA Met    S. solfataricus  (Kaine, 1987)
ACACGAAGAGTTTAAAAACGGGTAAAGATTAAACTATTAGAGAGGGCCCGTCGT   tRNA Val           "              "
ATCATTAAAGGTTAAATAGGCTTGAAAAAGATATTAATATTGCGGCCGTCGTCT   tRNA Gly           "              "
CCAATAAACCTATAAAGTCATATGTAAATAATAATAATGCCGCCGTAGCTCAGC   tRNA Phe           "              "

GCATAAAGTATATAAACCCTTATCGCATAGAGTAAGATTCCAGACGCTTACaGC   rRNA P1     D. mobilis  (Kjems and Garrett, 1988)
ACCCGTCATGATTAATACCCTTGGAGCAAATAGATTCATCaaGCCCGCGGCATT   rRNA P2           "              "
TAGTGAACGCTTTGAAAGCAGCTGGTGTTCCACGGAGTgAAGCACTCTACGTGG   rRNA P3           "              "
AGAGAACTGGTTCAAACACGTCAGGCTTTTCCCCGACGTCATCCCCGTgCTCAG   rRNA P4           "              "
-GAATTCATCCTTGAGGCAGTGGTGGGAACCGGGTTGAGCAGGGAGGATgCCGC   rRNA P5           "              "
AGAGTAAGGTTTTAAAACCCCAGTAATAGATTATGGGACTaCGGTGCCCGACCC   5S rRNA P1        "              "
CCTAACACACTATACAATATATTGATGCTCGCAATAGTGGTaGCCCTAATAGTC   5S rRNA P2        "              "
TAAGGAGATCTTTGAAAGCGCTGAGACAACACTgAAGTATCTTGAGAAAATCAT   5s rRNA P3        "              "
```

b)

A_{17}	A_{17}	A_{13}	T_9	C_7	T_{20}	T_{20}	T_{20}	A_{19}	A_{18}	A_{18}	T_8	C_8
T_5	C_3	C_5	G_7	T_7	A_1	A_3	C_2	G_3	T_4	T_4	A_8	T_5
C_1	G_2	G_3	A_6	G_5	G_1		A_1	T_1	C_1	G_1	C_4	A_5
	T_1	T_2	C_1	A_4	C_1						G_3	G_5
A	A	A	N	N	T	T	T	A	A	A		

consensus crenarchaeota
(low and intermediate GC)

FIGURE 3.4 See Figure 3.2 for caption.

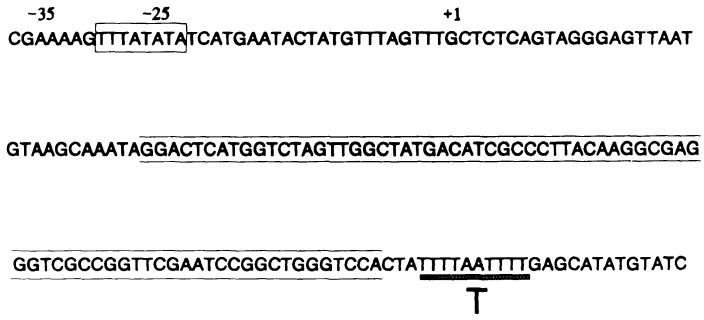


FIGURE 3.5 Genomic sequence of $tRNA^{Val}$ gene of *Methanococcus vannielii*. This gene was used as template to establish the DNA sequences promoting and terminating initiation of cell-free transcription. Recombinant plasmid shown harbors the wild-type upstream region to position −35; it contains archaeal consensus promoter sequence (Thomm and Wich, 1988; Reiter et al., 1988a) upstream (boxed octanucleotide) and also terminator signal proposed by Wich et al. (1986b) downstream from the gene (indicated by grey bar and the letter T below sequence). Sequence encoding mature $tRNA^{Val}$ is indicated by parallel lines above and below sequence.

et al., 1991). This finding supports the conclusion that the TATA box is necessary for initiation of transcription. To provide conclusive evidence for the significance of this sequence as a promoter signal, a series of point mutations has been introduced into the DNA region upstream from the $tRNA^{Val}$ gene (Hausner et al., 1991). The effects of some of these mutations on the rate of cell-free transcription are summarized in Figure 3.6. When the T in position 2, 5, and 7 of the consensus was replaced by G, the efficiency of transcription was dramatically reduced. In contrast, single-point mutations upstream and downstream from the TATA box did not significantly affect the expression rate of the tRNA gene (Figure 3.6). From these experiments, we concluded that the TATA box is a major constituent of an archaeal promoter. To assess the importance of the second box at the transcription start site, the initiator nucleotide was mutated to a T (Figure 3.6). Analysis of the in vitro transcripts from this template showed that this nucleotide is indispensable for initiation of transcription. Two further nucleotides of this box B sequence were required for a high rate of cell-free transcription (Figure 3.6; Hausner et al., 1991). Thus, a TATA box at a distance of about 20 nucleotides to an ATGC-like motif appears to be the minimal requirement of an archaeal promoter. The sequence AAAAG up-

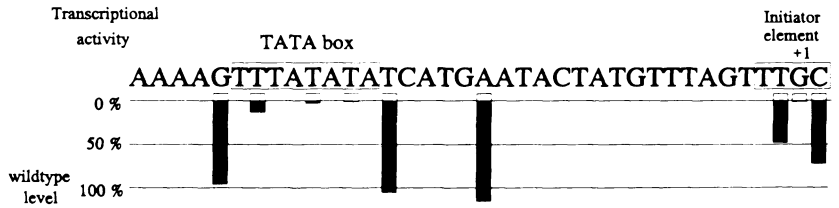


FIGURE 3.6 Analysis of effects of single-point mutations in the 5′ flanking region of the $tRNA^{Val}$ gene of *Methanococcus vannielii* on the rate of cell-free transcription. DNA sequence of 5′ flanking region of $tRNA^{Val}$ gene from position −35 to +2. Single-point mutations introduced into this region by in vitro mutagenesis are indicated by small open rectangles below sequence; T or A residue usually was replaced by G. G at position −31 was replaced by T; T and C residues at position −1 and +2 by A; and G at transcription start site by T. Black bars indicate template activity of plasmids containing single-point mutations. Note that promoter activity is dramatically reduced only by mutations in region of TATA box and initiator element.

stream of the TATA box (Brown et al., 1989; Figures 3.2–3.4) is not essential for initiation of transcription but appears to contribute to promoter function (Hausner et al., 1991). To avoid confusion with the well-established designations of A box and B box for the internal control regions of the RNA polymerase III promoter (see following), we suggest the designations TATA box and initiator element for the two structural elements of a typical archaeal promoter (Hausner et al., 1991). The archaeal TATA box appears homologous to the TATA box of eucaryotic polymerase II promoters at both the structural and the functional level (Hausner et al., 1991).

In contrast to organisms of the domains *Archaea* and *Bacteria*, eucaryotic cells have evolved different mechanism for the expression of rRNA and tRNA genes: These two classes of stable RNA genes are transcribed by two different types of RNA polymerase and do not share a common promoter sequence. In most eucaryotic systems, the DNA region from +10 to −40 constitutes the minimal RNA polymerase I promoter (Clos et al., 1986; Sollner-Webb et al., 1987). The sequences preceding rRNA genes from different eucaryotes do not show significant homologies. Thus, the polymerase I promoter appears to be species specific (Sommerville, 1984). In contrast to RNA polymerase I and II transcription systems, the promoter signals for eucaryotic tRNA and 5S genes reside downstream from the transcription initiation site (reviewed by Geiduschek and Tocchini-Valentini, 1988). Two sequences, A box and B box, corresponding to the region encoding the D and the TψC loops of mature tRNA, mediate initiation of transcription by RNA polymerase III. These sequences are highly conserved among all eucaryotes and are also found in the tRNA genes of bacteria. Moreover, owing to the presence of the A box and B box motives, tRNA genes of *E. coli* and the chloroplasts of *Euglena* are expressed with high efficiency by the polymerase III transcriptional machinery (Gruissem et al., 1982), although bacterial RNA polymerases initiate at upstream promoters. Inspection of the DNA sequences of archaeal tRNA genes revealed that the eucaryotic A box and B box sequences

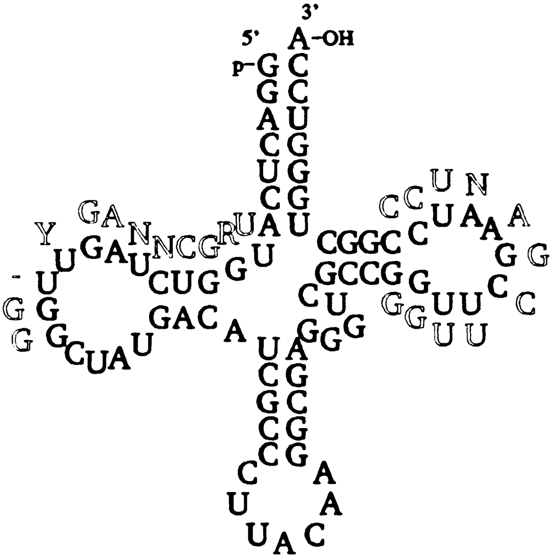


FIGURE 3.7 Archaeal tRNA sequence containing eucaryotic polIII promoter elements A box and B box: cloverleaf representation of tRNA^{Val} gene of *Methanococcus vannielii*. Intragenic polIII consensus promoter sequences are shown at corresponding position beside cloverleaf structure. Left: D loop, A box; right: TψC loop, B box. Abbreviations: R, purine; Y, pyrimidine; N, purine or pyrimidine.

are also highly conserved in the *Archaea* (Waldschmidt, 1989; Hausner, 1989; Figure 3.7). Deletion of internal sequences corresponding to the TψC loop did not abolish the template activity of the tRNA^{Val} gene (Thomm et al., 1990). However, a systematic analysis of the effects of internal deletions of an archaeal tRNA gene on the efficiency of cell-free transcription has not yet been performed. Thus, the function of internal DNA sequences in the initiation or termination of transcription remains to be elucidated.

Terminator sequences

Several DNA sequences downstream from archaeal genes have been proposed as possible terminator signals: transcripts from stable RNA genes of *Methanococcus vannielii* (Wich et al., 1986b), *Sulfolobus shibatae* (Reiter et al., 1988b), and *Thermofilum pendens* (Kjems et al., 1990) terminate within oligo-dT sequences strictly resembling the terminator sequences established

for polymerase III transcription systems (Geiduschek and Tocchini-Valentini, 1988). Similar sequences have been found at the 3′ end of protein-encoding genes from *Sulfolobus shibatae* (Reiter et al. 1988b). However, downstream from most protein-encoding genes of methanogens and extreme halophiles, sequences similar to rho-independent terminators of *E. coli* have been observed (see review by Brown et al., 1989). These structures can form hairpin-like structures, which are followed by an oligo-dT sequence. Transcripts from stable RNA genes from *Desulfurococcus mobilis*, *Methanobacterium thermoautotrophicum*, and *Thermofilum pendens* terminate at the end or after polypyrimidine sequences (Kjems and Garrett, 1987; Ostergaard et al., 1987; Kjems et al., 1990). However, transcription termination and 3′ processing sites can barely be distinguished by mapping the 3′ end of transcripts in vivo. Further, the function of these conserved oligo-dT and pyrimidine-rich sequences in the termination of transcription has not been established.

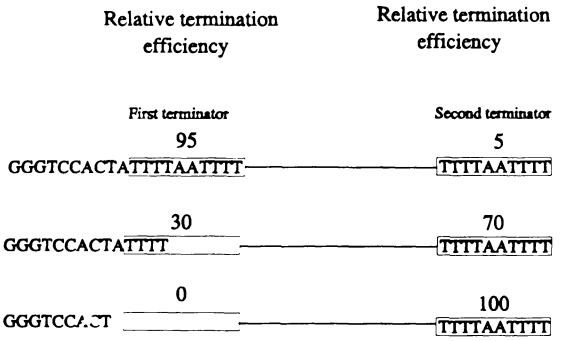


FIGURE 3.8 Oligo-dT sequences direct termination of transcription at tRNA genes of *Methanococcus vannielii*. DNA sequences of constructs containing putative terminator sequence TTTTAATTTT (Wich et al., 1986b) in tandem. DNA region between two terminator signals has been deleted by exonuclease III treatment. Efficiency of termination of transcription at first and second terminator was determined by measuring ratio of wild-type to elongated transcript. Relative amounts of corresponding transcripts are indicated above terminator signals (boxed).

To identify the DNA sequences necessary for termination of transcription in a member of the *Archaea*, we performed a mutational analysis of the DNA region downstream from the $tRNA^{Val}$ gene of *Methanococcus vannielii*. Various clones with deletions at the 3′ end of the gene were generated and ligated to a DNA fragment harboring the intact 3′ end of the $tRNA^{Val}$ gene (Figure 3.8). These constructs contain the two putative terminator sites TTTTAATTTT (Wich et al., 1986b) in tandem. When the RNA polymerase does not stop at the first (mutated) terminator, an additional longer transcript should be synthesized. The amount of this longer RNA product is inversely correlated with the efficiency of termination at the first terminator. Thus, the effect of a mutation can be quantitated twofold by measuring the ratio of wild-type to elongated transcript. When a construct containing two wild-type terminators in tandem was used as a template in an in vitro transcription experiment, the efficiency of termination at the first and second oligo-dT sequence was 95% and 5%, repectively (Figure 3.8, top row). When six nucleotides of the first terminator have been deleted, about two-thirds of the transcripts stop at the second and one-third at the first terminator (Figure 3.8, middle row). After the TTTT sequence has also been removed from the template (Figure 3.8, bottom row), transcription is terminated almost exclusively at the second terminator. These findings support the conclusion that the oligo-dT sequences downstream from stable RNA genes of M. *vannielii* are a major signal mediating termination of transcription.

3.4 Universal Features of Stable RNA Genes

No unique molecular property of stable RNA genes that is common to all the *Archaea* has been discovered so far. Some unique characteristics, such as unlinked expression of the three ribosomal RNAs (*Thermoplasma*), an intron in the variable loop of a transfer RNA (*Thermofilum pendens*), and a mixed-type splicing mechanism of a 23S rRNA intron (*Desulfurococcus mobilis*), appear to be restricted to a few genera or species.

Some features, such as the molecular orga-

nization of rRNA operons, may be *Eucarya*-like in one genus (*Thermoproteus*) and *Bacteria*-like in others (*Halobacterium*). However, both linked and unlinked expression of the 5S rRNA gene have been observed even within one genus (*Desulfurococcus mobilis* and *D. mucosus*). A universal feature of archaeal stable RNA genes that distinguishes *Archaea* from *Bacteria* is the presence of a 7S RNA.

Most similarities to *Eucarya* are found at the level of the transcriptional mechanisms. Both the polypeptides involved in the expression of stable RNA genes and the transcription signals appear to be very similar in the various genera of *Archaea*. All archaeal RNA polymerases show a multisubunit structure and genomic sequences resembling eucaryotic RNA polymerase (Huet et al., 1983; Schnabel et al., 1983; Pühler et al., 1989). The consensus promoter sequences are almost identical between *Methanococcus* and *Sulfolobus*, which represent the two phylogenetically separated kingdoms of the *Archaea* (see Figures 3.2 and 3.4). The archaeal TATA box is the major element determining the transcription start site (Hausner et al., 1991). This property, and the similarity in sequence and location to the eucaryotic TATA box, suggest homology of important parts of the archaeal and RNA polymerase II promoter. Minor differences in the consensus promoter sequences between extreme halophiles and methanogens (Figures 3.2 and 3.4) can be correlated with the high GC content of the cellular DNA of extreme halophiles (Thomm et al., 1989) and do not argue for a structural diversity of promoter structures in the *Euryarchaeota*. The existence of transcription factors provides a further similarity to eucaryotic gene transcription (Frey et al., 1990; Hüdepohl et al., 1990).

The finding that the purified RNA polymerase of *Archaea* binds to the promoter (Thomm et al., 1989) and shows semispecific initiation at initiator element-like sequences (Hüdepohl et al., 1990), unlike eucaryotic enzymes; but requires additional factors for correct initiation of transcription, as do eucaryotic RNA polymerases, suggests a novel function of the archaeal transcription factors. The further investigation of the biochemistry of archaeal transcription might contribute to a deeper understanding of the evolution of the transcription apparatus and lead to the discovery of new mechanisms of regulation of gene expression.

3.5 Summary

The sequential and structural organization of stable RNA genes from all major groups of *Archaea* has been analyzed. The transcription start sites located upstream of many rRNA/tRNA genes have been determined, and putative promoter and terminator sequences have been inferred from sequence comparisons. Footprinting and cell-free transcription experiments have been used to investigate the biological functions of these conserved DNA sequences, allowing the following conclusions.

Within the *Euryarchaeota*, operons encoding ribosomal RNAs show the bacteria-like organization: 5'-16S-tRNA-23S-5S-3'. *Crenarchaeota* do not contain a tRNA gene in the spacer between 16S and 23S rRNA and usually show unlinked 5S genes organized into a separate transcription unit. In *Methanococcus*, additional unlinked 5S genes exist; these are located within clusters of tRNAs. Some tRNA genes of *Archaea* and a 23S rRNA gene of *Desulfurococcus mucosus* contain an intron. However, this similarity to eucaryotes is restricted to a few tRNA species and has been found only in some genera of *Archaea*. Universal eucaryotic features are the presence of a gene encoding a 7S RNA of unknown function and the presence of eucaryotic promoter sequences. A TATA box octanucleotide at −25 and a second conserved sequence at the transcription start site are required for initiation of transcription. The archaeal TATA box element determines the transcription initiation site. Both in structure and function, this element closely resembles the TATA box of eucaryotic RNA polymerase II promoters. Oligo-dT sequences, which are conserved downstream from most archaeal genes, direct termination of transcription by the *Methanococcus* RNA polymerase. These sequences resemble the terminator signals recognized by RNA polymerase III of eucaryotes. The expression of archaeal stable

RNA genes is mediated by at least two transcription factors that might activate initiation of transcription by a hitherto unknown mechanism.

Acknowledgments. This work was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie. We thank Dr. Karl Stetter for supporting parts of this work by funds from the Leibniz Preis.

References

Achenbach-Richter, L. and C.R. Woese. 1988. The ribosomal gene spacer region in archaebacteria. *Syst. Appl. Microbiol.* **10**:211–214.

Achenbach-Richter, L., K.O. Stetter, and C.R. Woese. 1987. A possible biochemical missing link among archaebacteria. *Nature* (London) **327**:348–349.

Balch, W.E., G.E. Fox., L.J. Magrum, C.R. Woese, and R.S. Wolfe. 1979. Methanogens: reevaluation of a unique biological group. *Microbiol. Rev.* **1979**:260–296.

Böck, A., H. Hummel, M. Jarsch, and G. Wich. 1986. *In: Genes for Stable RNA in Methanogens: Phylogenetic and Functional Aspects. Biology of Anaerobic Bacteria,* H.C. Doubourguier, et al., eds. pp. 206–226. Amsterdam: Elsevier.

Boros, I., E. Csordas-Toth, A. Kiss, I. Kiss, I. Török, A. Udvardy, K. Udvardy, and P. Venetianer. 1983. Identification of two new promoters probably involved in the transcription of a ribosomal RNA gene of *Escherichia coli. Biochim. Biophys. Acta* **739**:173–180.

Brown, J.W., C.J. Daniels, and J.N. Reeve. 1989. Gene structure, organization and expression in Archaebacteria. *Crit. Rev. Microbiol.* **16**:287–338.

Brown, J.W., M. Thomm, G.S. Beckler, G. Frey, K.O. Stetter, and J.N. Reeve. 1988. An archaebacterial RNA polymerase binding site and transcription initiation of the *hisA* gene in *Methanococcus vannielii. Nucleic Acids Res.* **16**:135–150.

Caillet, J., J.A. Plumbridge, and M. Springer. 1985. Evidence that *pheV*, a gene for $tRNA^{Phe}$ of *E. coli* is transcribed from tandem promoters. *Nucleic Acids Res.* **13**:3699–3710.

Clos, J., A. Normann, A. Öhrlein, and I. Grummt. 1986. The core promoter of mouse rDNA consists of two functionally distinct domains. *Nucleic Acids Res.* **14**:7581–7595.

Corden, J., B. Wasylyk, A. Buchwalder, P. Sassone-Corsi, C. Kedinger, and P. Chambon. 1980. Promotor sequences of eukaryotic protein-coding genes. *Science* **209**:1406–1414.

Daniels, C.J., S.E. Douglas, and W.F. Doolittle. 1986. Genes for transfer RNAs in *Halobacterium volcanii. Syst. Appl. Microbiol.* **7**:26–29.

Daniels, C.J., R. Gupta, and W.F. Doolitle. 1985. Transcription and excision of a large intron in the $tRNA^{trp}$ gene of an archaebacterium, *Halobacterium volcanii. J. Biol. Chem.* **260**:3132–3134.

Datta, P.K., L.K. Hawkins, and R. Gupta. 1989. Presence of an intron in elongator methionine-tRNA of *Halobacterium volcanii. Can. J. Microbiol.* **35**:189–194.

Dennis, P.P. 1985. Multiple promoters for the transcription of the ribosomal RNA gene cluster in *Halobacterium cutirubrum. J. Mol. Biol.* **186**:457–461.

Eggen, R., H. Harmsen, and M. de Vos. 1990. Organization of a ribosomal RNA gene cluster from the archaebacterium *Methanothrix soehngenii. Nucleic Acids Res.* **18**:1306.

Frey, G., M. Thomm, B. Brüdigam, H.P. Gohl, and W. Hausner. 1990. An archaebacterial cell-free transcription system. The expression of tRNA genes from *Methanococcus vannielii* is mediated by a transcription factor. *Nucleic Acids Res.* **18**:1361–1367.

Geiduschek, P.E. and G.P. Tocchini-Valentini. 1988. Transcription by RNA polymerase III. *Annu. Rev. Biochem.* **57**:873–914.

Grogan, D., P. Palm, and W. Zillig. 1990. Isolate B12, which harbours a virus-like element, represents a new species of the archaebacterial genus *Sulfolobus, Sulfolobus shibatae,* sp. nov. *Arch. Microbiol.* **154**:594–599.

Gruissem, W., D. Prescott, B.M. Greenberg, and R.B. Hallick. 1982. Transcription of *E. coli* and *Euglena* chloroplast tRNA gene clusters and processing of polycistronic transcripts in a HeLa cell-free extract. *Cell* **30**:81–92.

Haas, E.S., C.J. Daniels, and J.N. Reeve. 1989. Genes encoding 5S rRNA and tRNAs in the extremely thermophilic archaebacterium *Methanothermus fervidus. Gene* **77**:253–263.

Haas, E.S., J.W. Brown, C.J. Daniels, and J.N. Reeve. 1990. Genes encoding the 7S RNA and a $tRNA^{Ser}$ are linked to one of the two rRNA operons in the genome of the extremely thermophilic archaebacterium *Methanothermus fervidus. Gene* **90**:51–58.

Hausner, W. 1989. In vitro Mutagenese von tRNA Genen Zur Ermittlung promotor-und terminatoraktiver DNA Sequenzen. Diploma Thesis, Universität Regensburg, FRG.

Hausner, W., G. Frey, and M. Thomm. 1991. Control regions of an archaeal gene. ATATA box and an initiator element promote cell-free transcription of the $tRNA^{Val}$ gene of *Methanococcus vanniellii, J. Mol. Biol.* L22:495–508.

Hawley, D.K. and R. McClure. 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res.* **11**:2237–2255.

Hüdepohl, U., W.D. Reiter, and W. Zillig. 1990. In

vitro transcription of two rRNA genes of the archaebacterium *Sulfolobus* sp. B12 indicates a factor requirement for specific transcription. *Proc. Nat. Acad. Sci. Natl. USA* **87**:5851–5855.

Huet, J., R. Schnabel, A. Sentenac, and W. Zillig, 1983. Archaebacteria and eukaryotes possess DNA-dependent RNA polymerases of a common type. *EMBO J.* **2**:1291–1294.

Jarsch, M., J. Altenbuchner, and A. Böck. 1983. Physical organization of the genes for ribosomal RNA in *Methanococcus vannielii*. *Mol. Gen. Genet.* **189**:41–47.

Kaine, B.P. 1987. Intron-containing tRNA genes of *Sulfolobus solfataricus*. *J. Mol. Evol.* **25**:248–254

Kaine, B.P. 1990. Structure of the archaebacterial 7S RNA molecule. *Mol. Gen. Genet.* **221**:315–321.

Kaine, B.P. and V.L. Merkel. 1989. Isolation and characterization of the 7S RNA gene from *Methanococcus voltae*. *J. Bacteriol.* **171**:4261–4266.

Kaine, B.P., R. Gupta, and C.R. Woese. 1983. Putative introns in tRNA genes of prokaryotes. *Proc. Nat. Acad. Sci. USA* **80**:3309–3312.

Kenerley, M.E., E.A. Morgan, L. Post, L. Lindahl, and M. Nomura. 1977. Characterization of hybrid plasmids carrying individual ribosomal RNA transcription units of *Escherichia coli*. *J. Bacteriol.* **132**:931–949.

King, T.C., R. Sirdegkmukh, and D. Schlesinger. 1986. Nucleolytic processing of ribonucleic acid transcripts in procaryotes. *Microbiol. Rev.* **50**:428–451.

Kjems, J. and R.A. Garrett. 1985. An intron in the 23S ribosomal RNA gene of the archaebacterium *Desulfurococcus mobilis*. *Nature* (London) **318**:675–677.

Kjems, J. and R.A. Garrett. 1987. Novel expression of the ribosomal RNA genes in the extreme thermophile and archaebacterium, *Desulfurococcus mobilis*. *EMBO J.* **6**:3521–3530.

Kjems, J. and R.A. Garrett. 1988. Novel splicing mechanism for the ribosomal RNA intron in the archaebacterium *Desulfurococcus mobilis*. *Cell* **54**:693–703.

Kjems, J. and R.A. Garrett. 1990. Secondary structural elements exclusive to the sequences flanking ribosomal RNAs lend support to the monophyletic nature of the archaebacteria. *J. Mol. Evol.* **31**:25–32.

Kjems, J., H. Leffers, T. Olesen, and R.A. Garrett. 1989. A unique tRNA Intron in the variable loop of the extreme thermophile *Thermofilum pendens* and its possible evolutionary implications. *J. Biol. Chem.* **264**:17834–17837.

Kjems, J., H. Leffers, T. Olesen, I. Holz, and R.A. Garrett. 1990. Sequence, organization and transcription of the ribosomal RNA operon and the downstream tRNA and protein genes in the archaebacterium *Thermofilum pendens*. *Syst. Appl. Microbiol.* **13**:117–127.

Knaub, S. and A. Klein. 1990. Specific transcription of cloned *Methanobacterium autotrophicum* transcription units by homologous RNA polymerase in vitro. *Nucleic Acids Res.* **18**:1441–1446.

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

Lake, J. 1989. Origin of the eukaryotic nucleus: eukaryotes and eocytes are genotypically related. *Can. J. Microbiol.* **35**:109–118.

Larsen, H. and W.D. Grant. 1989. Group III. Extremely halophilic archaeobacteria: Order Halobacteriales Ord. Nov. *In: Bergey's Manual of Systematic Bacteriology, Vol. 3.* J.T. Stanley, M.P. Bryant, N. Pfennig, and J.G. Holt, eds. pp. 2216–2224. Baltimore: Williams & Wilkins.

Larsen, N., H. Leffers, J. Kjems, and R.A. Garrett. 1986. Evolutionary divergence between the ribosomal RNA operons of *Halococcus morrhuae* and *Desulfurococcus mobilis*. *Syst. Appl. Microbiol.* **7**:49–57.

Leffers, H., J. Kjems, L. Ostergaard, N. Larsen, and R.A. Garrett. 1987. Evolutionary relationship amongst archaebacteria. A comparative study of 23S ribosomal RNAs of a sulphur-dependent extreme thermophile, an extreme halophile and a thermophilic methanogen. *J. Mol. Biol.* **195**:43–61.

Mankin, A.S. and V.K. Kagramanova. 1986. Complete nucleotide sequence of the single ribosomal RNA operon of *Halobacterium halobium*: secondary structure of the archaebacterial 23S rRNA. *Mol. Gen. Genet.* **202**:152–161.

Mankin, A.S. and V.K. Kagramanova. 1988. Complex promoter pattern of the single ribosomal RNA operon of an archaebacterium *Halobacterium haloblum*. *Nucleic Acids Res.* **16**:4679–4692.

Melton, D.A., E.M. De Robertis, and R. Cortese. 1980. Order and intracellular location of the events involved in the maturation of a spliced tRNA. *Nature* (London) **284**:143–148.

Mevarech, M., S. Hirsch-Twizer, S. Goldman, E. Yakobson, H. Eisenberg, and P.P. Dennis. 1989. Isolation and characterization of the rRNA gene clusters of *Halobacterium marismortui*. *J. Bacteriol.* **171**:3479–3485.

Moritz, A. and W. Goebel. 1985. Characterization of the 7S RNA and its gene from halobacteria. *Nucleic Acids Res.* **13**:6969–6979.

Moss, T. 1983. A transcriptional function for the repetitive ribosomal spacer in *Xenopus laevis*. *Nature* (London) **302**:223–228.

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

Ostergaard, L., N. Larsen, H. Leffers, J. Kjems, and R.A. Garrett. 1987. A ribosomal RNA operon and its flanking region from the archaebacterium *Methanobacterium thermoautotrophicum*,

Marburg strain: transcription signals, RNA structure and evolutionary implications. *Syst. Appl. Microbiol.* **9**:199–209.

Pühler, G., H. Leffers, F. Gropp, P. Palm, H.-P. Klenk, F. Lottspeich, R.A. Garrett, and W. Zillig. 1989. Archaebacterial DNA-dependent RNA polymerases testify to the evolution of the eukaryotic nuclear genome. *Proc. Nat. Acad. Sci. USA.* **86**:4569–4573.

Ree, H.K. and R.A. Zimmermann, 1990. Organization and expression of the 16S, 23S and 5S ribosomal RNA genes of the archaebacterium *Thermoplasma acidophilum*. *Nucleic Acids Res.* **18**:4471–4478.

Reiter, W., P. Palm, and W. Zillig. 1988a. Analysis of transcription in the archaebacterium *Sulfolobus* indicates that archaebacterial promoters and eukaryotic RNA pol II promoters are of the same type. *Nucleic Acids Res.* **16**:1–19.

Reiter, W., P. Palm, and W. Zillig. 1988b. Transcription termination in the archaebacterium *Sulfolobus*: signal structures and linkage to transcription initiation. *Nucleic Acids Res.* **16**:2445–2459.

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

Schnabel, R., M. Thomm, R. Gerardy-Schahn, W. Zillig, K.O. Stetter, and J. Huet. 1983. Structural homology between different archaebacterial DNA-dependent RNA polymerases analysed by immunological comparison of their components. *EMBO J.* **2**:751–755.

Sollner-Webb, B., J. Windle, S. Henderson, J. Tower, V. Culotta, S. Kass and N. Craig. 1987. Initiation and termination of ribosomal RNA transcription and processing of the primary rRNA transcript. *In*: *RNA Polymerase and the Regulation of Transcription*. W.S. Reznikoff, R.R. Burgess, J.E. Dahlberg, C.A. Gross, M.T. Record and M.P. Wickens, eds., pp. 187–194. New York: Elsevier.

Sommerville, J. 1984. RNA polymerase I promotors and transcription factors. *Nature* (London) **310**:189–190.

Struck, J.C.R., H.Y. Toschka, T. Specht, and V.A. Erdmann. 1988. Common structural features between eukaryotic 7SL RNAs, eubacterial 4.5S RNA and scRNA and archaebacterial 7S RNA. *Nucleic Acids Res.* **16**:7740.

Suzuki, Y., Y. Ono, A. Nagata, and T. Yamada. 1988a. Molecular cloning and characterization of an rRNA operon in *Streptomyces lividans* TK21. *J. Bacteriol.* **170**:1631–1636.

Suzuki, Y., A. Nagata, Y. Ono, and T. Yamada. 1988b. Complete nucleotide sequence of the 16S rRNA gene of *Mycobacterium bovis* BCG. *J. Bacteriol.* **170**:2886–2889.

Taschke, C., M. Klinkert, J. Wolters, and R. Herrmann. 1986. Organization of the ribosomal RNA genes in *Mycoplasma hyopneumoniae*: The 5S rRNA gene is separated from the 16S and 23S rRNA genes. *Mol. Gen. Genet.* **205**:428–433.

Thomm, M. and G. Wich. 1988. An archaebacterial promoter element for stable RNA genes with homology to the TATA box of higher eukaryotes. *Nucleic Acids Res.* **16**:151–163.

Thomm, M., B.A. Sherf, and J.N. Reeve. 1988. RNA polymerase-binding and transcription initiations sites upstream of the methyl reductase operon of *Methanococcus vannielii*. *J. Bacteriol.* **170**:1958–1961.

Thomm, M., G., Frey, W. Hausner, and B. Brüdigam. 1990. An archaebacterial in vitro transcription system. *In*: *Microbiology and Biochemistry of Strict Anaerobes Involved in Interspecies Hydrogen Transfer*, J.-P. Belaich, M. Bruschi, and J.-L. Garcia, eds., pp 305–312. New York: Plenum Press.

Thomm, M., G. Wich, J.W. Brown, G. Frey, B.A. Sherf, and G.S. Beckler. 1989. An archaebacterial promoter sequence assigned by RNA polymerase binding experiments. *Can. J. Microbiol.* **35**:30–35.

Thompson, L.D. and C.J. Daniels. 1988. A $tRNA^{Trp}$ intron endonuclease from *Halobacterium volcanicum*. Unique substrate recognition properties. *J. Biol. Chem. 263*:17951–17959.

Tu, J. and W. Zillig. 1982. Organization of rRNA structural genes in the archaebacterium *Thermoplasma acidophilum*. *Nucleic Acids Res.* **10**:7231–7245.

Waldschmidt, R. 1989. Transkriptionsfuktoren der RNA Polymerase III. Ph.D. Thesis, Phillips-Universität, Marburg, FRG.

Wich, G., M. Jarsch, and A. Böck. 1984. Apparent operon for a 5S ribosomal RNA gene in the archaebacterium *Methanococcus vannielii*. *Mol. Gen. Genet.* **196**:146–151.

Wich, G., H. Hummel, M. Jarsch, U. Bär, and A. Böck. 1986a. Transcription signals for stable RNA genes in *Methanococcus*. *Nucleic Acids Res.* **14**:2459–2479.

Wich, G., L. Sibold, and A. Böck. 1986b. Genes for tRNA and their putative expression signals in *Methanococcus*. *Syst. Appl. Microbiol.* **7**:18–25.

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

Wich, G., L. Sibold, and A. Böck. 1987b. Divergent evolution of 5S rRNA genes in *Methanococcus*. *Z. Naturforsch. Sect. C. Biosci.* **42**:373–380.

Woese, C.R. 1987. Bacterial evolution. *Microbiol. Rev.* **51**:221–271.

Woese, C.R., and G.L. Fox. 1977. Phylogenetic structure of the procaryotic domain: the primary kingdom. *Proc. the Natl. Acad. Sci. USA* **74**:5088–5090.

Woese, C.R., O. Kandler, and M. Wheelis. 1990. Towards a natural system of organisms. Proposal of the Domain Archaea, Bacteria and Eucarya. *Proc. Natl. Acad. Sci. USA* **87**:4576–4579.

Young, R.A. and J.A. Steitz. 1979. Tandem promoters direct *E. coli* ribosomal RNA synthesis. *Cell* **17**:225–234.

Zillig, W., I. Holz, H. Klenck, J. Trent, S. Wunderl, D. Janekovic, E. Imsel, and B. Haas. 1987. *Pyrococcus woesei* sp. nov., an ultra-thermophilic marine archaebacterium representing a novel order, Thermococcales. *Syst. Appl. Microbiol.* **9**:62–70.

Zwieb, C. 1989. Structure and function of signal recognition particle RNA. *Prog. Nucleic Acid Res. Mol. Biol.* **37**:207–234.