Sequence of the 23S rRNA gene from the archaebacterium *Methanococcus vannielii*: evolutionary and functional implications

Michael Jarsch and August Böck
Lehrstuhl für Mikrobiologie der Universität München, Maria-Ward-Straße 1 a, D-8000 München 19, Federal Republic of Germany

Summary. The nucleotide sequence was determined of the genes coding for 23S and 5S rRNA within a putative 16S-23S-5S rRNA transcriptional unit from *Methanococcus vannielii*. The 23S rRNA gene is 2958 bases in length (compared to 2904 for the *E. coli* sequence). Determination of the homology of the primary structure with the *Escherichia coli* sequence gives a value similar to that previously found for the 16S rRNA sequences, indicating an identical phylogenetic drift for the two genes. The putative transcript of the 16S, 23S, and 5S rRNA genes plus flanking regions has been folded into a secondary structure following the comparative approach. The model provides evidence for processing signals of the putative primary transcript into pre-16S and pre-23S rRNA molecules. The 23S secondary structure when compared to those for eubacterial and eukaryotic large ribosomal RNA's reveals a mosaic architecture: It shows insertions at positions where they are also present in eu and eukaryotic large ribosomal RNA's; other parts of the molecule, however, are almost identical to eubacterial structures (e.g. the equivalent of the *E. coli* ribosomal protein L11 binding region). The 5S rRNA coding region is separated from the 23S rRNA gene by a spacer of 59 base pairs. This rRNA-operon-linked 5S gene differs by a total of 13 base substitutions from that of the tRNA-operon-linked 5S gene previously described. A correlation seems to exist between 23S rRNA structural domains and sensitivity or resistance to some 70S and 80S ribosome-targeted inhibitors: (i) The degree of sensitivity to the 80S ribosome inactivating RNase z-sarcin is correlated with the presence of the eukaryotic type cleavage site; (ii) inactivity of erythromycin as an inhibitor of *Methanococcus* protein synthesis may be a consequence of the presence of a G residue (instead of an “A” found in sensitive organisms) in position 2127 (the “peptidyl-transferase loop”) of the sequence; (iii) sensitivity of *Methanococcus* ribosomes to thiopeptin, on the other hand, is correlated with a highly conserved (compared to *E. coli*) binding domain for an L11 equivalent protein.

Materials and methods

Plasmids and strains. Plasmid pMV1 was described previously (Jarsch et al. 1983); it contains a DNA fragment of the *M. vannielii* genome coding for a complete rRNA operon. Plasmids were prepared by the method of Birnboim and Doly (1979) either from *E. coli* HB101 or from *E. coli* AB1157 (*dam*) to avoid inhibition of restriction endonuclease cleavage by *dam* modification (Dreiseikelmann et al. 1979).

Construction of deletion plasmids. Parts of the coding region of pMV1 were deleted by treatment with nuclease Bal31 after linearizing the plasmid with an appropriate restriction enzyme. A synthetic *BamHI* linker was attached to these shortened fragments and after secondary cleavage with restriction endonuclease *HindIII* the resulting fragments were subcloned into plasmid pBR322 by replacing a *BamHI*-*HindIII* fragment. The resulting subclones of pMV1 (pMV131a,b,d,e) are listed in Fig. 1. Alternatively, the Bal31 shortened plasmids were rejoined directly by means of a *BamHI* linker; the deleted parts of pMV1 in this case

Introduction

The genes for ribosomal RNA constitute a suitable model system for studying gene organization and expression in archaebacteria; their analysis, in addition, may give valuable information on the structural basis of many of the unusual functional properties of the archaebacterial protein synthesis system.

As a part of this project we have reported previously on the physical organization of the rRNA genes from *Methanococcus vannielii*, an organism belonging to the methanogen branch of archaebacteria (Jarsch et al. 1983). *Methanococcus* possesses four (putative) transcriptional units for 16S-rRNA-23S-5S RNA's and a single unlinked 5S rRNA gene possibly cotranscribed in a unique fashion with seven genes for tRNA's (Wich et al. 1984).

The nucleotide sequence of the 16S rRNA and intercistronic spacer region has been reported previously (Jarsch and Böck 1983, 1985); it confirmed and detailed the notion of a distinct relationship between methanogenic and extreme halophilic archaebacteria (Gupta et al. 1983), but also demonstrated that the methanogen 16S rRNA (like that of the halophile; Gupta et al. 1983) phylogenetically definitely is closer to the eubacterial type of small subunit ribosomal RNA than to the eukaryotic counterpart. In this communication we report on the primary structures of the 23S and 5S rRNA coding regions of one of the rRNA transcriptional units from *Methanococcus* and on phylogenetic and functional implications of the sequences.

Offprint requests to: A. Böck
are substituted by the linker. The subclones obtained in this way are listed as pMV132 a-d in Fig. 1.

DNA sequence analysis. DNA fragments were isolated and labeled at the 3' and 5' ends as described recently (Jarsch and Böck 1983). Sequencing reactions were carried out according to the chemical method of Maxam and Gilbert (1980), with the modification of the A+G reaction suggested by Gray et al. (1978). 6%, 8%, and 20% polyacrylamide gels containing 7 M urea were run at 60°C-70°C with the aid of thermostated plates.

Results and discussion

Physical map

The M. vannielii genome contains four copies of a transcriptional unit for 16S, 23S, and 5S rRNA and one additional physically unlinked 5S gene (Jarsch et al. 1983; Wich et al. 1984). Figure 1 gives the physical map for an EcoRI fragment of M. vannielii genomic DNA present in plasmid pMV1; it contains one of these rRNA transcriptional units. The 23S and 5S encoding segment is drawn out enlarged and supplied with a sequencing strategy. Parts of this sequence were derived from deletion plasmids as indicated. Their construction is detailed under Materials and methods.

Primary structure of the 23S rRNA gene

The 5' terminus of 23S rRNA from Methanococcus was determined by 5'-labeling of the rRNA with polynucleotide kinase and by sequencing labeled T1 ribonuclease digestion products (not shown). Two labeled oligonucleotides were obtained; a minor one corresponding to the sequence of the 5' terminus as given by Fig. 2 and a major one lacking the terminal U residue. Since it is plausible to consider the latter as a processing product, numbering of the primary sequence was started with U (Fig. 2). The 3' terminus was deduced by secondary structure considerations (Brosius et al. 1978; Noller 1984).

The M. vannielii 23S rRNA derived from the DNA sequence is 2958 bases in length, compared to 2904 bases of E. coli 23S rRNA. This difference in length is the result of a number of insertions and deletions, which are explained in more detail below.

The Methanococcus sequence was aligned (Fig. 2) to the 23S rRNA sequence of E. coli (Brosius et al. 1978). The sequences were first paired to optimal homology by introduction of gaps. Secondary structure features were taken into consideration in regions of unclear homology to correct this alignment. Homology values were determined according to McCarroll et al. (1983). The M. vannielii 23S rRNA shares a sequence homology of 60% with the E. coli sequence. This value is in good agreement with that reported for the 16S rRNA from these two organisms (Jarsch and Böck 1985). It indicates that the evolutionary drift of the genes for 16S and 23S rRNA – despite different functions of their products – was essentially identical.

Secondary structure of 23S rRNA

The M. vannielii 23S rRNA sequence derived from the gene sequence was folded into a secondary structure (Fig. 3) following the comparative approach (Noller 1984). The basic frames of the eubacterial and archaeabacterial versions of the molecule resemble each other quite closely. All of the six long range interactions proposed by Noller (1984) are also present in the archaeabacterial structure, dividing it into six structural domains.

The most striking difference from E. coli 23S rRNA