JOURNAL OF BACTERIOLOGY, Jan. 1988, p. 478–481
0021-9193/88/010478-04$02.00/0

Vol. 170, No. 1

Identification of a Transcript and Its Promoter Region on the Archaebacterial Plasmid pME2001

LEO MEILE, JERZY MADON, AND THOMAS LEISINGER*

Mikrobiologisches Institut, Eidgenössische Technische Hochschule, ETH-Zentrum, CH-8092 Zurich, Switzerland

Received 10 March 1987/Accepted 23 October 1987

The cryptic multicopy plasmid pME2001 of *Methanobacterium thermoautotrophicum* Marburg encodes a 611-base-pair transcript containing several consecutive, short open reading frames. Scrutiny of the 5'-flanking region did not reveal homology to putative archaebacterial consensus promoter sequences. However, 28 base pairs upstream of the transcription start point, there was a sequence with strong homology to a sequence preceding the *purE* gene of *M. thermoautotrophicum*.

The thermophilic methanogen *Methanobacterium thermoautotrophicum* Marburg contains a 4.5-kilobase high-copy-number plasmid (pME2001) of unknown function (12). To develop cloning vectors based on pME2001 (13), information about the location of essential plasmid functions on pME2001 is needed. We therefore have examined crude RNA preparations of the host strain by Northern (RNA) hybridization analysis for pME2001-encoded transcripts and have detected a prominent plasmid-encoded RNA. Since information on the structure of mRNA molecules and on transcription start points of methanogenic bacteria is scarce (3), we have determined the nucleotide sequence encoding this RNA and its 5'- and 3'-flanking regions.

M. thermoautotrophicum (DSM2133) was grown as described previously (12). Total DNA and plasmid DNA was isolated after mechanical (12) or enzymatic (9) lysis of the cells. Total RNA was prepared by mixing 1 g of wet cells with 7 ml of 50 mM Tris hydrochloride (pH 8.5) containing 1% tri-isopropyl-naphthalenesulphonate, 4% 4-aminosalicylate, and 6% phenol-chloroform (1:1 [vol/vol]). The mixture was vigorously shaken for 2 min with 15 g of glass beads (3-mm diameter). Three extractions by phenol-chloroform (1:1 [vol/vol]) were followed by isopropanol precipitation (10). The pellet was suspended in 0.5 ml of 50 mM Tris hydrochloride (pH 7.5)–5 mM $MgCl_2$–2 mM dithiothreitol containing 100 U of RNasin (Boehringer Mannheim Biochemicals), treated with 5 μg of DNase I for 1 h at 30°C, and extracted with phenol. The RNA was precipitated with isopropanol and suspended at 1 mg/ml in 20 mM Tris hydrochloride (pH 8.2) and 2 mM dithiothreitol containing RNasin (100 U/ml). DNA/RNA hybridization and digestion with nuclease S1 were performed as described elsewhere (5, 15). ^{32}P-end-labeled DNA fragments (75,000 cpm; 0.1 to 1.0 μg) and 50 μg of RNA from *M. thermoautotrophicum* were routinely used in a volume of 30 μl at a hybridization temperature of 49°C. Both strands of the indicated region of plasmid pME2001 (Fig. 1) were sequenced by the chemical degradation method (11).

A single strong band was observed on Northern blots of total RNA from *M. thermoautotrophicum* hybridized to the plasmid pME2001 probe (Fig. 2). When the blots were hybridized to fragments I, II, and III (Fig. 1) from pME2001, the same signal, corresponding to an RNA of approximately 0.6 kilobases, was obtained with fragment I but not with fragments II or III (Fig. 1 and 2). To map the transcription initiation site and to determine the direction of transcription on plasmid fragment I, the 0.43-kilobase *Bcl*I-*Bst*EII fragment Ia and the 0.6-kilobase *Bcl*I-*Xho*I fragment Ib were 5'end labeled with [γ-^{32}P]ATP and polynucleotide kinase at the *Bcl*I restriction site (Fig. 1). The sense strand was determined from the data shown in Fig. 3A. The data show that the RNA protected the 5' end of fragment Ia but not of fragment Ib. When fragments Ia and Ib were 3' end labeled with [α-^{32}P]dGTP and Klenow fragment of DNA polymerase I at the *Bcl*I site, hybridized RNA afforded protection of the label from nuclease S1 digestion in fragment Ib and not in fragment Ia (not shown). Thus, initiation of transcription is on fragment Ia, and RNA synthesis proceeds towards the *Xho*I site (Fig. 1).

The transcription start point was located 93 nucleotides upstream of the *Bcl*I site, close to the *Kpn*I restriction site

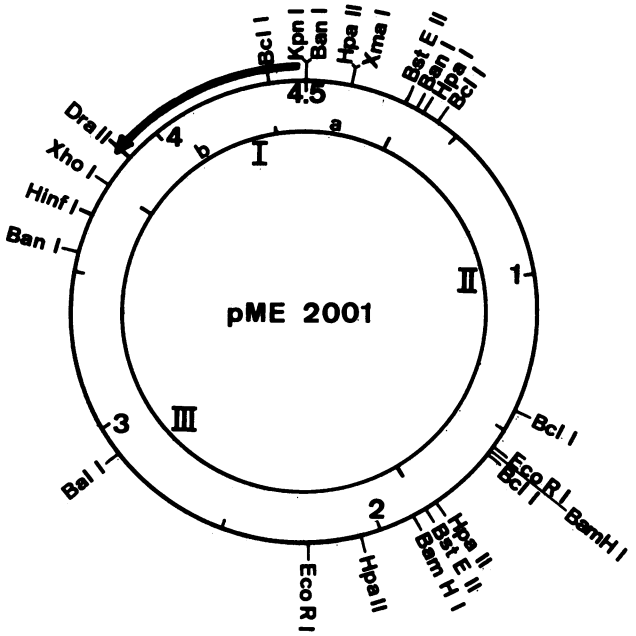

FIG. 1. Restriction map of plasmid pME2001. The restriction map was constructed by analyses of electrophoresed restriction fragments and from the sequencing data. Only one each of several *Dra*II and *Hin*fI sites is indicated. The transcript is represented by a thick arrow. I, II, and III are the designations of the restriction fragments used for Northern blot analysis.

*Corresponding author.

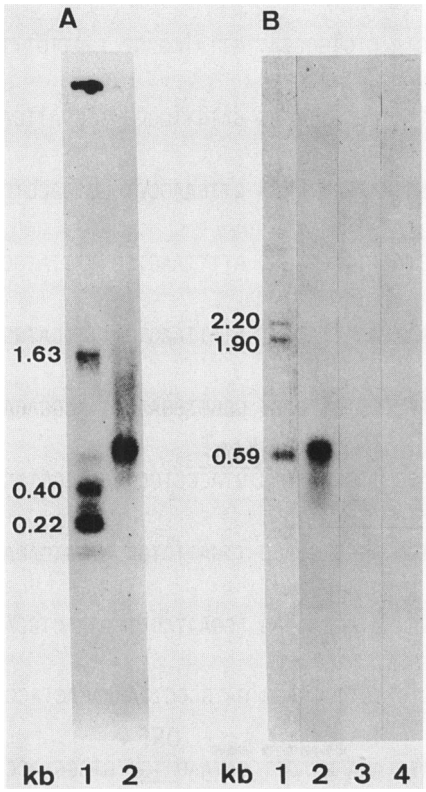

FIG. 2. Northern blot analysis of total RNA from *M. thermoautotrophicum*. (A) RNA probed with undigested, nick-translated ^{32}P-labeled pME2001 DNA (lane 2). (B) RNA probed with nick-translated fragments I (lane 2), II (lane 3), and III (lane 4), which are indicated in Fig. 1. The size markers were glyoxal-treated, *Hin*fI-restricted, ^{32}P-end-labeled pBR322 DNA (panel A, lane 1) and glyoxal-treated, *Hin*dIII-restricted, ^{32}P-end-labeled λ DNA (panel B, lane 1).

(Fig. 3A and 4). Transcription starts from a G, which is followed by CCCTG. The relative intensities of the four bands observed in the S1 mapping experiment changed with an increase in the concentration of nuclease S1. This change indicates that the apparently staggered initiation of transcription is probably due to imprecise digestion with nuclease S1. We cannot exclude the possibility that the transcript is formed by processing of a larger primary transcript. However, in two archaebacterial mRNAs whose 5′ termini have been examined, there was no evidence for posttranscriptional processing (2, 4).

The results of a nuclease S1 protection experiment with the 255-base-pair (bp) *Hin*fI-*Dra*II fragment at coordinate 3.7 (Fig. 1), which was 3′ end labeled at the *Dra*II site, are presented in Fig. 3B. They showed that the plasmid-encoded transcript ends closely downstream of the *Dra*II site, at nucleotide 611 (Fig. 4). Unlike other archaebacterial transcription termination sites (4, 14, 16), the putative transcription termination region does not exhibit an inverted repeat or an oligo(T) track.

The DNA sequence coding for the pME2001 transcript and its flanking regions is presented in Fig. 4. The pME2001 transcript started at position +1, in a region of strong dyad symmetry. In the RNA, this region could form an 8-nucleotide hairpin loop with a ΔG (at 25°C) value of −19.8 kcal/mol (23 hydrogen bonds). The nucleotide sequence encoding the transcript contained four possible open reading frames (ORFs) that ranged between 90 and 234 bp in size. Three of them were preceded by at least three consecutive bases of the methanogen consensus ribosome-binding sequence 5′-AGGTGA-3′ (7). In the absence of expression studies, it is not possible to predict whether any of these ORFs give rise to polypeptides in *M. thermoautotrophicum*. However, the complementarity in 9 out of 10 bases to the 3′ end of the 16S rRNA of *M. thermoautotrophicum* (1) is striking for the nucleotide sequence positioned immediately in front of ORF 4 (Fig. 4).

The 500-bp region upstream of the transcription initiation site did not contain sequences with homology to the putative archaebacterial promoter 5′-GAANTTTCA. This sequence has been proposed to be involved in transcription initiation on the basis of a comparative analysis of sequences preceding archaebacterial ORFs (7). In the 500-bp 5′-flanking region of the pME2001 transcript, there was no or only limited homology with other proposed archaebacterial promoter motifs detected in front of *Halobacterium* rRNA genes (8) or preceding tRNA genes of methanococci (16). However, nucleotides −28 to −8 were very similar to the sequence AATGGT-CCCTGC preceding the *purE* gene of *M. thermoautotrophicum* ΔH (6). Since the DNA-dependent

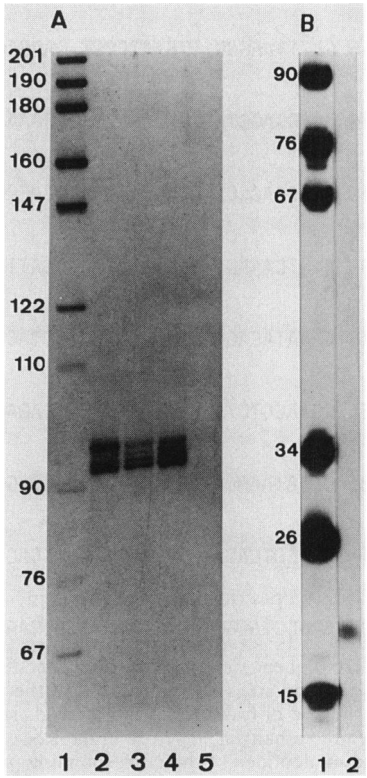

FIG. 3. Nuclease S1 mapping of the 5′ and 3′ termini of the pME2001-encoded transcript. (A) Autoradiography of a 4% polyacrylamide sequencing gel with DNA fragments Ia and Ib (Fig. 1) 5′ end labeled with [γ-^{32}P]ATP and polynucleotide kinase at the *Bcl*I site. Lane 1, 3′-^{32}P-end-labeled *Hpa*II digest of pBR322 as size marker; lanes 2, 3, and 4, hybrids of fragment Ia after treatment with 1,000, 2,000, or 4,000 U of nuclease S1; lane 5, hybrid of fragment Ib after digestion with 2,000 U nuclease S1. (B) Autoradiography of an 8% polyacrylamide sequencing gel with the 255-bp *Hin*fI-*Dra*II fragment which was 3′ end labeled with [α-^{32}P]dGTP and Klenow fragment of DNA polymerase I at the *Dra*II site. Lane 2, hybrid after digestion with 2,000 U of nuclease S1; lane 1, *Hpa*II-digested, 3′-^{32}P-end labeled pBR322 as a size marker.

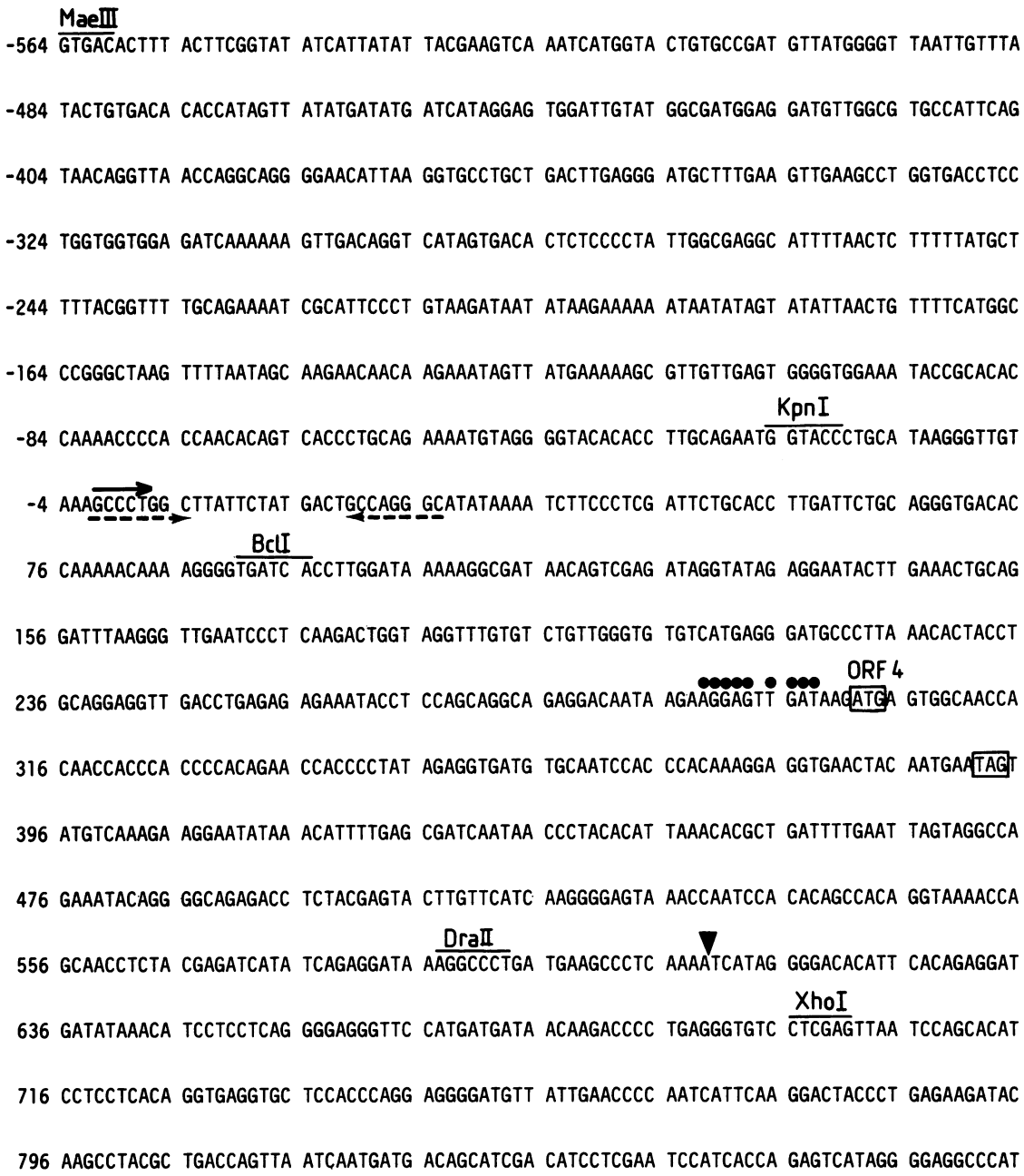

FIG. 4. DNA sequence of the pME2001-encoded transcript and its flanking regions. Nucleotides are numbered consecutively from the transcription initiation site (→). The 3′ end of the transcript is marked (▼). With respect to Fig. 1, the sequence starts downstream of the *Bst*EII site at the end of fragment Ia and extends leftward towards the *Xho*I fragment delineating fragment Ib. An inverted repeat sequence which could lead to a hairpin structure in the RNA is shown (- - - -› and ‹- - - -). The initiation and termination codons of a potential ORF are boxed, and the nucleotides of the corresponding putative ribosome binding site are indicated (●●●).

RNA polymerase of *M. thermoautotrophicum* binds to the 5′-flanking region of the *purE* gene, which contains this motif (J. W. Brown and J. N. Reeve, personal communication), the common sequence may be specifically involved in transcription initiation.

The putative archaebacterial consensus promoter for protein-encoding genes (7) has resulted from comparison of archaebacterial gene sequences presently available. The nucleotide sequences used for this comparison were from distantly related archaebacteria, and transcription start points were not determined in the genes that were analyzed. More sequence data, experiments to map 5′ ends of mRNA made in vivo, and in vitro transcription experiments are thus necessary for the definite identification of methanogen promoters. Our characterization of a plasmid-encoded transcript from *M. thermoautotrophicum* contributes to these efforts in that it has revealed a promoter region with homology to a sequence preceding the *purE* gene of a closely related organism.

We thank D. Haas for discussions.

This work was supported by project 3.193-85 from the Swiss

National Foundation for Scientific Research and by the Roche Research Foundation.

LITERATURE CITED

1. **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. **43**:260–296.
2. **Betlach, M., J. Friedman, H. W. Boyer, and F. Pfeifer.** 1984. Characterization of a halobacterial gene affecting bacterio-opsin gene expression. Nucleic Acids Res. **12**:7949–7959.
3. **Bollschweiler, C. R., C. Kühn, and A. Klein.** 1985. Non-repetitive AT-rich sequences are found in intergenic regions of *Methanococcus voltae* DNA. EMBO J. **4**:805–809.
4. **DasSarma, S., U. L. RajBhandary, and H. G. Khorana.** 1984. Bacterio-opsin mRNA in wild-type and bacterio-opsin-deficient *Halobacterium halobium* strains. Proc. Natl. Acad. Sci. USA **81**:125–129.
5. **Favaloro, J., R. Treisman, and R. Kamen.** 1980. Transcription maps of polyoma virus-specific RNA analysis by two dimensional nuclease S1 gel mapping. Methods Enzymol. **65**:718–749.
6. **Hamilton, P. T., and J. N. Reeve.** 1985. Sequence divergence of an archaebacterial gene cloned from a mesophilic and a thermophilic methanogen. J. Mol. Evol. **22**:351–360.
7. **Hamilton, P. T., and J. N. Reeve.** 1986. Molecular biology of archaebacteria, p. 338–348. *In* L. Leive (ed.), Microbiology—1986. American Society for Microbiology, Washington, D.C.
8. **Hui, L., and P. P. Dennis.** 1985. Characterization of the ribosomal RNA gene clusters in *Halobacterium cutirubrum*. J. Biol. Chem. **260**:899–906.
9. **Kiener, A., H. König, J. Winter, and T. Leisinger.** 1987. Purification and use of *Methanobacterium wolfei* pseudomurein endopeptidase for lysis of *Methanobacterium thermoautotrophicum*. J. Bacteriol. **169**:1010–1016.
10. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
11. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
12. **Meile, L., A. Kiener, and T. Leisinger.** 1983. A plasmid in the archaebacterium *Methanobacterium thermoautotrophicum*. Mol. Gen. Genet. **191**:480–484.
13. **Meile, L., and J. N. Reeve.** 1985. Potential shuttle vectors based on the methanogen plasmid pME2001. Bio/Technology **3**:69–72.
14. **Müller, B., R. Allmansberger, and A. Klein.** 1986. Termination of a transcription unit comprising highly expressed genes in the archaebacterium *Methanococcus voltae*. Nucleic Acids Res. **13**:6439–6445.
15. **Sharp, P. A., A. J. Berk, and S. M. Berget.** 1980. Transcription maps of adenovirus. Methods Enzymol. **65**:750–768.
16. **Wich, G., L. Sibold, and A. Böck.** 1986. Genes for tRNA and their putative expression signals in *Methanococcus*. Syst. Appl. Microbiol. **7**:18–25.