Identification of a Transcript and Its Promoter Region on the
Archaeobacterial 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 archaeobacterial 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 MgCl2–2 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). 32P-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 BclI-BsrEII fragment Ia and the 0.6-kilobase BclI-XhoI fragment Ib were 5' end labeled with [γ-32P]ATP and polynucleotide kinase at the BclI 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-32P]dGTP and Klenow fragment of DNA polymerase I at the BclI 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 XhoI site (Fig. 1).

The transcription start point was located 93 nucleotides upstream of the BclI site, close to the KpnI 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 DraII and HinfI 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.
VOL. 170, 1988

NOTES

749

HindIII-restricted, glyoxal-treated, (A) RNA probed followed translated fragments I (lane 2), 32P-labeled pME2001 examined, been tion is an increase in the 255-base pair (bp) in al We indicates that 3A (Fig. 1).
The DNA

FIG. 2. Northern blot analysis of total RNA from M. thermoautotrophicum. (A) RNA probed with undigested, nick-translated 32P-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, HindIII-restricted, 32P-end-labeled pBR322 DNA (panel A, lane 1) and glyoxal-treated, HindIII-restricted, 32P-end-labeled λ DNA (panel B, lane 1).

(Fig. 3A and 4). Transcription starts from a G, which is followed by CCGTG. 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 archaeabacterial 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) HindIII-DraII fragment at coordinate 3.7 (Fig. 1), which was 3' end labeled at the DraII site, are presented in Fig. 3B. They showed that the plasmid-encoded transcript ends closely downstream of the DraII site, at nucleotide 611 (Fig. 4). Unlike other archaeabacterial 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 archaeabacterial promoter 5'-GAANNTTCA. This sequence has been proposed to be involved in transcription initiation on the basis of a comparative analysis of sequences preceding archaeabacterial ORFs (7). In the 500-bp 5'-flanking region of the pME2001 transcript, there was no or only limited homology with other proposed archaeabacterial 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 AATGCTG-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 la and lb (Fig. 1) 5' end labeled with [γ-32P]ATP and polynucleotide kinase at the BcII site. Lane 1, 32P-end-labeled HpaII digest of pBR322 as size marker; lanes 2, 3, and 4, hybrids of fragment la after treatment with 1,000, 2,000, or 4,000 U of nuclease S1; lane 5, hybrid of fragment lb after digestion with 2,000 U nuclease S1. (B) Autoradiography of an 8% polyacrylamide sequencing gel with the 255-bp HindIII-DraII fragment which was 3' end labeled with [α-32P]dGTP and Klenow fragment of DNA polymerase I at the DraII site. Lane 2, hybrid after digestion with 2,000 U of nuclease S1; lane 1, HpaII-digested, 3' P-end labeled pBR322 as a size marker.
NOTES

MaeII
-564 GTGACACTTT ACTTCGGTAT ATCATTATAT TACGAAAGCTA AATCATGCTA CTGTGCGGAT GTTATGGGT TAAATGTTTA
-484 TACTGTGACA CACCATAGGT ATATGATARG ATCATAGGAG TGGATTTGAT GATGTGGGCG TGCCATTCAG
-404 TAACAGGTAA ACCAGGCAGG GAAACATTAA GTGGCTCCTG GACCTTGGAA ATGGTGGAG TGGTACCTCG
-324 TGGTTGCGGAA GATCAAAAA GTTGAAGCTG CATATGGACA CTCTCCCCTA TTGGCGAGGC TTTTAATTGCT
-244 TTACGAGTTT TGCAGAATT CCGATCTCCT GTAAAGATAA ATAAGAAAAA ATAATATAGT ATATAACTGT TTTTACAGGC
-164 CCGGCGTAAG TTTTTATAGC AGAACAACA ACAAAATAGTT ATGAAAAAGG GTGTTGGAG GGGGTTGAAA TACCCGCAAC

KpnI
-84 CAAAACCCA CCAACAGCT CACCTGCGC AAAAAATGAG GTTACACACG TTGCAAGAAGC GTACCCTGCA TAAGGTTTTG
-4 AAGACCTGG CTATTTCTCT GACCTGCGG GCAATAAAAA TCTTACCTCG ATCTTGCAGC TTGAATCTCG AGGGTGACAC

BstII
76 CAAAAACAAA AGGGGCTAGT ACCCTGAGA AAAAGGCTG AACAATGGAG ATAGGTATAG GAAACTGCA
156 GATTTAAGGG TTTGCAGCCT CAAGACTGGA AGTTTGAATG TGGCTGGGTT GTGCTGAGAG GATGCCCTTTA AACACTCCT
236 GACAGGAAGTT GACCTGAGAG AGAAAAATCC CCAGCGGAGG GAGGGAATAG GATAAATATTG AAGATATTTG
316 CACCCACCCA CCCCCAGAG CCACCTTAT GAAGGGGTAG TGCAATCCAC CCACAAAGGA GGTGAACAGC ATGAAAPAG
396 ATGTCAAAGA AGAAATATAA AATTTTGAG CGATCAATTT CCTCACACAT TAAACAGCCT GATTGAATAT GAGTAGGGCC
476 GAAATACAGG GGCAGAGACC TCTACGAGTA CTTGTCATCT AAGGGGTGTA AACAATCCTA CACAGGCTCA GCTAAAACCA
556 GCAACCTCTA CGAGATCAT TAAGAGGAGT AAGGCCTGGA TGAAGCCCTC AAAATCATAG GGGACACATT CACAGGAGT

DraII

XhoI
636 GATAAAAAAC TCCCTCTAG GGGAGGTTTC CATGATGATA ACAAAGACCC TGAAGGTTGC CTGGAGTTAA TCCAGCAGAT
716 CCTCCCTCAA GTTGAGGTGC TCCACCCAGG AGGGGATGTT ATGGAACCCC AATCTTCCAA GAGCTACCTG GAGGAGTAG

796 AAGCCCTAGC GACGATCTTCA ATCAATGATG ACAGCATGGA CATCTCCTGA TACCAAGCCTA GAGCATAGG GAGGTCCTAC

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 BstEII site at the end of fragment 1a and extends leftward towards the XhoI fragment delineating fragment 1b. 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(200,479),(914,830) 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 archaeabacterial consensus promoter for protein-encoding genes (7) has resulted from comparison of archaeabacterial gene sequences presently available. The nucleotide sequences used for this comparison were from distantly related archaeabacteria, 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
NOTES 481

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

LITERATURE CITED


