JOURNAL OF BACTERIOLOGY, Oct. 1990, p. 5586–5592
0021-9193/90/105586-07$02.00/0

Vol. 172, No. 10

Characterization of a Promoter and a Transcription Terminator of *Spiroplasma melliferum* Virus SpV4

C. STAMBURSKI, J. RENAUDIN,* AND J. M. BOVE

Laboratoire de Biologie Cellulaire et Moléculaire, Institut National de la Recherche Agronomique et Université de Bordeaux II, B.P. 81, 33883 Villenave d'Ornon Cedex, France

Received 29 December 1989/Accepted 7 July 1990

Spiroplasma virus 4 (SpV4) is an isometric virus with single-stranded, circular DNA infecting the helical mollicute *Spiroplasma melliferum*, a honeybee pathogen. Previous studies in our laboratory led to the determination of the base sequence of the SpV4 DNA. Nine open reading frames and three promoterlike sequences (P1, P2, and P3) were identified. An inverted repeat leading to the formation of a hairpin structure on the transcription product was also found and predicted to be a transcription terminator (T). We have now studied the in vivo transcription of the SpV4 genome by Northern (RNA) blot analysis of the total RNAs extracted from SpV4-infected spiroplasma cells. Transcripts of 7.8, 4.4, 3.4, and 2.7 kilobases (kb) were detected. The 3.4-kb RNA was the major transcript. The 5′ and 3′ ends of this transcript were determined by S1 mapping and primer extension. Characterization of the 3′ end by S1 mapping showed that the 3.4-kb transcript terminates within the stretch of uridine residues following the hairpin structure of terminator T. Characterization of the 5′ end by S1 mapping indicated that transcription proceeds from a newly recognized promoter, P0, located 36 nucleotides upstream of P1. Primer extension resulted in two cDNA signals. The short cDNA was probably a primer extension artifact due to the presence of a hairpin structure on the transcript. When reverse transcriptase stopped at this hairpin or read through, the short or the long cDNA, respectively, was obtained. The size of the long cDNA identified P0 as the transcription promoter. Promoter P0 was also shown to be functional in *Escherichia coli*. Indeed, when inserted upstream of the chloramphenicol acetyltransferase gene of a promoter selection vector, it promoted transcription of this gene. As in the case of *S. melliferum*, two cDNAs were obtained by primer extension, the longer cDNA identifying P0 as the promoter.

Mollicutes (mycoplasmas) are wall-less procaryotes that have arisen by degenerative evolution from ancestors of gram-positive eubacteria with DNA of low G+C content. As such, they are phylogenetically related to *Clostridium ramosum* and *Clostridium innocuum*. Among the mollicutes, spiroplasmas are the only organisms with helical morphology and motility. Ancestors of the spiroplasmas are believed to have played a central role in mollicute phylogeny. In particular, splitting of the spiroplasma branch has led to *Mycoplasma* spp. and *Ureaplasma* spp. (19). Most spiroplasmas are associated with insects or other arthropods. *Spiroplasma melliferum* (26 mol% G+C) is a honeybee pathogen. Spiroplasma virus 4 (SpV4), an isometric virus with single-stranded, circular DNA, was discovered in *S. melliferum* strain B63 (14). SpV4 is specific to *S. melliferum*. The 4.4-kilobase (kb) genome of SpV4 is one of the smallest genomes of procaryotic DNA viruses.

The replicative form (RF) of SpV4 was cloned in *Escherichia coli* and shown to be infectious by transfection (10). The SpV4 DNA was sequenced and found to contain 4,421 nucleotides (12). The capsid protein gene and eight other open reading frames have been located on the viral genome (12). Examination of the sequence also revealed three promoterlike sequences, P1, P2, and P3, each with −10 and −35 regions showing sequence homology with the consensus sequences TATAAT and TTGACA, respectively. These sequences are characteristic of eubacterial promoters recognized by RNA polymerases functioning with the general sigma factor σ^{70} for the *E. coli* enzyme and σ^{43} in *Bacillus subtilis* (13). On the SpV4 genome, P1 is located at nucleotide 544, P2 is at nucleotide 1292, and P3 is at nucleotide 3954 (nucleotide 1 is the cytidine residue of the unique *ClaI* site of the SpV4 RF). We have also found an inverted repeat sequence around nucleotide 3930 that could lead to the formation, on the transcription product, of a hairpin structure with three G · C pairs, followed by a sequence of seven uridine residues at the 3′-OH end. Such structures are considered characteristic of bacterial transcription terminators of the Rho-independent type. Similar transcription signals have been described for a tRNA gene cluster of *S. melliferum* (15) as well as for the spiralin gene (3) and the single ribosomal DNA operon (F. Laigret, O. Grau, and J. M. Bové, Zentralbl. Bakteriol. Parasitenkd. Infektionskr. Hyg. Abt. 1 Orig.) of *Spiroplasma citri*. However, evidence demonstrating that these regulatory sequences are functional is lacking. We have used Northern (RNA) blot analysis to study the in vivo transcription products of the SpV4 genome. In addition, we show by S1 mapping and primer extension experiments that the major transcript (3.4 kb) is initiated from a newly recognized promoter, P0, located 36 nucleotides upstream of P1, and is terminated within the stretch of uridine residues following the terminator hairpin at nucleotide 3930. Promoter P0 is functional in both *E. coli* and *S. melliferum*.

MATERIALS AND METHODS

Bacteria, spiroplasmas, and viruses. *E. coli* HB101 was used for amplification of the plasmid vector pKK232-8 (Pharmacia LKB Biotechnology Inc., Piscataway, N.J.) and its derivatives. *S. melliferum* G1 was used as a host for propagation of SpV4. Growth of *S. melliferum*, infection with SpV4, virus purification, and preparation of viral DNA have been previously described (10, 11).

* Corresponding author.

Enzymes. Restriction endonucleases and Moloney murine leukemia virus reverse transcriptase were purchased from Bethesda Research Laboratories, Inc. (Gaithersburg, Md.). T4 DNA ligase, *E. coli* DNA polymerase I (Klenow fragment), and S1 nuclease were from Boehringer GmbH (Mannheim, Federal Republic of Germany).

Construction of recombinant plasmids. Restriction fragments of the SpV4 RF were separated on and purified from low-melting-point agarose gels. They were eventually made blunt ended by using the Klenow enzyme and were ligated to the *Sma*I-linearized and dephosphorylated vector pKK232-8. The ligation mixture was used to transform *E. coli* HB101 competent cells according to standard procedures (7). Plasmid DNA was purified by the alkaline lysis procedure (7).

Southern hybridization. Recombinant plasmids were analyzed by agarose gel electrophoresis and by hybridization with SpV4-specific probes as described by Southern (17).

RNA isolation and Northern blot analysis. Total RNAs were isolated from SpV4-infected or noninfected *S. melliferum* cells in culture by using the guanidine thiocyanate method of Chirgwin et al. (4). Samples (5 μg) of RNA in 2.2 M formaldehyde–50% formamide–1× MOPS buffer [0.02 M 3-(N-morpholino)propanesulfonic acid (pH 7.0), 5 mM sodium acetate, 0.1 mM EDTA] were heated to 60°C for 10 min and then electrophoresed through 1% agarose gels containing 2.2 M formaldehyde and 1× MOPS buffer. After electrophoresis, the gel was soaked in distilled water for 4 h to remove the formaldehyde. The RNA was then transferred onto nitrocellulose paper with 20× SSC (SSC is 0.15 M NaCl plus 0.015 M sodium citrate). The RNA blots were prehybridized for 3 h at 45°C in a hybridization solution consisting of 50% formamide, 5× SSC, 5× Denhardt solution, 10% dextran sulfate, and 300 μg of yeast tRNA per ml in 100 mM Tris hydrochloride (pH 7.5). Hybridizations were carried out for 16 h at 45°C with 10^6 cpm of DNA probe per ml. After hybridization, the RNA blots were washed for 10 min at room temperature and four times for 10 min each time at 65°C in 0.1 SSC–0.1% sodium dodecyl sulfate.

S1 nuclease mapping. The probes were obtained by isolating SpV4 DNA fragments from a recombinant plasmid containing the entire SpV4 RF linearized at the unique *Cla*I site (10). The 5′ probe was prepared by isolating a 2,955-base-pair (bp) *Nde*I fragment that started at nucleotide 726 (728 on the minus strand) and extended left to outside of the *Cla*I site and labeling the 5′ ends with T4 polynucleotide kinase and [γ-^{32}P]ATP as described by Maniatis et al. (7). It was then recut with *Cla*I, and the 726-bp *Cla*I-*Nde*I end-labeled fragment (see Fig. 2) was isolated by gel purification. The 3′ probe was prepared by isolating a 2,656-bp *Sau*3AI fragment that started at nucleotide 3854 and extended right to the outside of the *Cla*I site and labeling the 3′ ends with [α-^{32}P]dGTP in a fill-in reaction as described by Maniatis et al. (7). The fragment was recut with *Cla*I, and the 568-bp *Sau*3A1-*Cla*I end-labeled fragment (see Fig. 2) was isolated. Thirty micrograms of total RNA from noninfected or SpV4 infected cells was analyzed by the S1 RNA mapping technique (1, 18). RNA was coprecipitated with the ^{32}P-end-labeled probe. After suspension in 30 μl of 0.4 M NaCl–40 mM piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) (PIPES; pH 6.4)–1 mM EDTA–80% formamide, the samples were incubated at 90°C for 5 min and allowed to cool to 45°C. Hybridization was performed for 3 h. For S1 digestion, the samples were mixed with 300 μl of a precooled mixture containing 0.3 M NaCl, 50 mM sodium acetate (pH 4.6), 4.5 mM $ZnSO_4$, 20 μg of heat-denaturated calf thymus DNA per ml, and 50 U of S1 nuclease (Boehringer). After 1 h of incubation at room temperature, the samples were put on ice and precipitated with ethanol. The pellets were washed twice with 70% ethanol, dried, and then suspended in 6 μl of 95% formamide–20 mM EDTA–0.1% bromophenol blue–0.1% xylene cyanol FF for analysis. The length of the DNA protected from single-strand digestion with S1 nuclease was determined by electrophoresis on a 6% denaturing polyacrylamide gel alongside sequencing reactions.

Primer extension analysis. Total RNA of *S. melliferum* or *E. coli* (60 μg) and 50 pmol of oligonucleotide were coprecipitated with ethanol and suspended in 30 μl of 0.4 M NaCl–40 mM PIPES (pH 6.4)–1 mM EDTA. The solution was incubated at 90°C for 3 min, cooled to 45°C, and incubated for 3 h. The nucleic acids were precipitated and washed with 70% ethanol and then suspended in 50 μl of 50 mM Tris hydrochloride (pH 8.3)–100 mM KCl–10 mM $MgCl_2$–10 mM dithiothreitol–50 μg of dactinomycin per ml–0.5 mM dATP–0.5 mM dGTP–0.5 mM dTTP–0.5 μM dCTP–20 μCi of [α-^{32}P]dCTP (3,000 Ci/mmol) with 4 U of reverse transcriptase and incubated for 1 h at 37°C. The elongation reaction was stopped by addition of 10 μl of 100 mM EDTA, and the DNA was precipitated with ethanol. The precipitate was washed and dried. The reaction products were analyzed on 6% denaturing polyacrylamide gels.

DNA sequencing. Sequencing ladders for sizing S1-protected fragments and cDNAs were produced by dideoxy sequencing reactions (16) of M13 recombinant clone DNA. Sequencing reactions were performed and analyzed by polyacrylamide gel electrophoresis as described by Renaudin et al. (12).

RESULTS

Analysis of in vivo SpV4 transcripts. The results of Northern blot analysis of total RNAs purified from SpV4-infected or noninfected spiroplasma cells and hybridized with a ^{32}P-labeled SpV4-DNA probe are shown in Fig. 1. In infected cells (lanes I-3, I-6, and I-9), four main transcripts of 2.7, 3.4, 4.4, and 7.8 kb were detected, regardless of the time (3, 6 or 9 h) after infection. The major transcript (3.4 kb) was detected very early after infection, even within the 20 min required for harvesting the cells by centrifugation (lane I-0). As expected, no hybridization signal was detected with RNAs extracted from noninfected cells (lane H).

Figure 2 indicates the location of the putative promoter sequences P1, P2, and P3 and that of the putative terminator sequence T on the SpV4 genome as determined from the SpV4 DNA sequence (12). The experimentally determined sizes of the transcription products are consistent with the locations of these regulatory sequences (Table 1). The major 3.4-kb transcript would be initiated at promoter P1. The 2.7- and 4.4-kb transcripts would be initiated at promoters P2 and P3, respectively. All four transcripts would be terminated at the single terminator T. The 7.8-kb transcript (Fig. 1, lanes I-3, I-6, and I-9) was much larger than the 4.4-kb SpV4 genome (Fig. 1, lane SpV4). This transcription product, like the 3.4-kb transcript, would be initiated at promoter P1 and would be terminated only when RNA polymerase encounters terminator T for the second time, after an extra full circle around the viral DNA (Fig. 2 and Table 1).

Transcription products were further characterized by hybridization experiments using cloned SpV4-DNA fragments A and B (Fig. 2) as specific probes. According to Fig. 2, probe B should hybridize with all four transcripts. In contrast, probe A should hybridize with the 4.4- and 7.8-kb

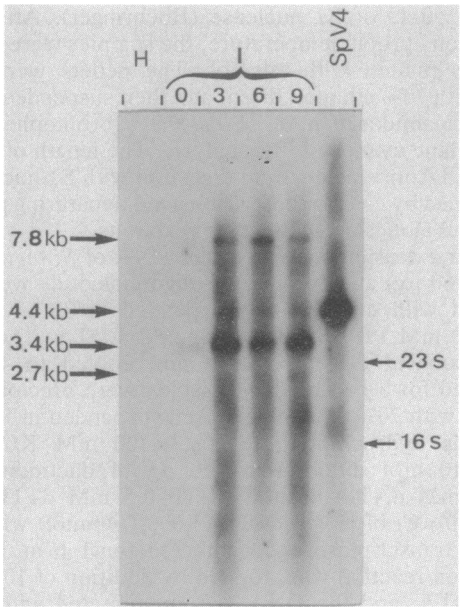

FIG. 1. Northern blot analysis of in vivo SpV4 RNA transcripts. Lanes: H, total RNAs extracted from healthy spiroplasma cells; I-0 to I-9, total RNAs extracted from SpV4-infected spiroplasma cells 0, 3, 6, and 9 h, respectively, after infection; SpV4, SpV4 single-stranded viral DNA. The probe was a ^{32}P-labeled recombinant plasmid containing the entire SpV4 RF. Positions of 16S and 23S *E. coli* rRNAs are indicated.

transcripts but not with the major 3.4-kb transcript. Indeed, despite the heavy background, both probes A and B gave a clear-cut hybridization signal at 7.8 kb (Fig. 3). In addition, probe B, but not probe A, hybridized with the 3.4-kb transcript. These results are consistent with the idea that the 3.4-kb transcript is initiated downstream of A close to P1, as shown in Fig. 2.

Determination of the 3′ end of the 3.4-kb transcript. S1 mapping was used to determine the termination site of the 3.4-kb transcript. When the 3′-end-labeled DNA probe encompassing the region of the putative terminator T was hybridized to total RNAs extracted from SpV4-infected cells and digested with S1 nuclease, DNA bands ranging from 87 to 94 bases were protected from degradation (Fig. 4a, lane I). These termination sites correspond to the poly(U) sequence following the hairpin structure of the terminator T (Fig. 4b). The two major bands correspond to nucleotides 3943 and 3944. The fact that the 3.4-kb transcript does not terminate at a single base position but ends heterogeneously within the stretch of uridine residues is characteristic of Rho-independent termination (6).

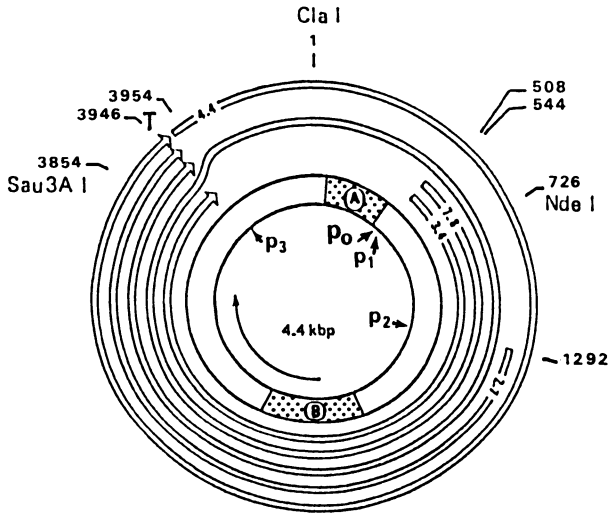

FIG. 2. Proposed organization of the SpV4 RNA transcripts. Positions of transcription promoter-like sequences P0, P1, P2, and P3 and of transcription terminator-like structure T are indicated. A and B represent the locations on the SpV4 genome of the cloned SpV4 RF DNA fragments used as specific probes in the experiment shown in Fig. 3.

TABLE 1. Comparison between the experimentally determined size of SpV4 transcripts and the theoretical size predicted from the nucleotide sequence

Zone transcribed	Transcript of size (bases) as determined by:	
	Base sequence of SpV4 genome	Northern blot analysis
P1–T	3,400	3,400
P2–T	2,652	2,700
P3–T	4,411	4,400
P1–(T)–T	7,821	7,800

Determination of the 5′ end of the 3.4-kb transcript. (i) S1 nuclease mapping. To map the transcription initiation point of the 3.4-kb transcript, the *Cla*I-*Nde*I fragment containing the P1 promoter region (Fig. 2) was labeled at the 5′ ends and used for hybridization with total RNAs extracted from SpV4-infected cells under conditions favoring RNA-DNA interactions. The hybridization products were submitted to S1 nuclease digestion. DNA fragments protected from the S1 digestion were analyzed on a denaturing polyacrylamide gel alongside sequencing reactions used as ladders. One major DNA band of 223 nucleotides and five additional, very faint DNA bands of 184, 193, 198, 202, and 220 nucleotides were detected (Fig. 5). The 184-nucleotide DNA band is hardly visible on the figure, even though it corresponds to adenine 544 of the chloramphenicol acetyltransferase (CAT) box of

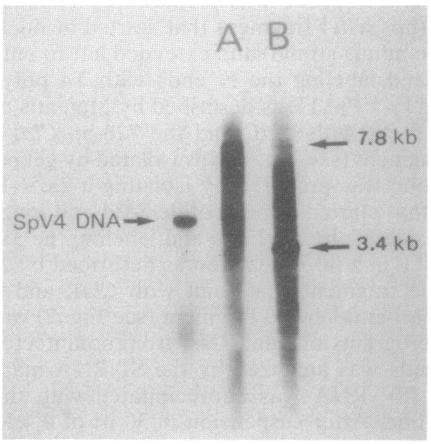

FIG. 3. Northern blot analysis of in vivo SpV4 RNA transcripts, using region-specific probes. Lanes: SpV4 DNA, SpV4 single-stranded viral DNA, hybridized with probe A (control); A and B, total RNAs extracted from SpV4-infected spiroplasma cells, hybridized with probe A (lane A) or probe B (lane B). Probes A and B are cloned SpV4 RF DNA fragments of, respectively, 391 and 598 bp. Their locations on the SpV4 genome are indicated in Fig. 2.

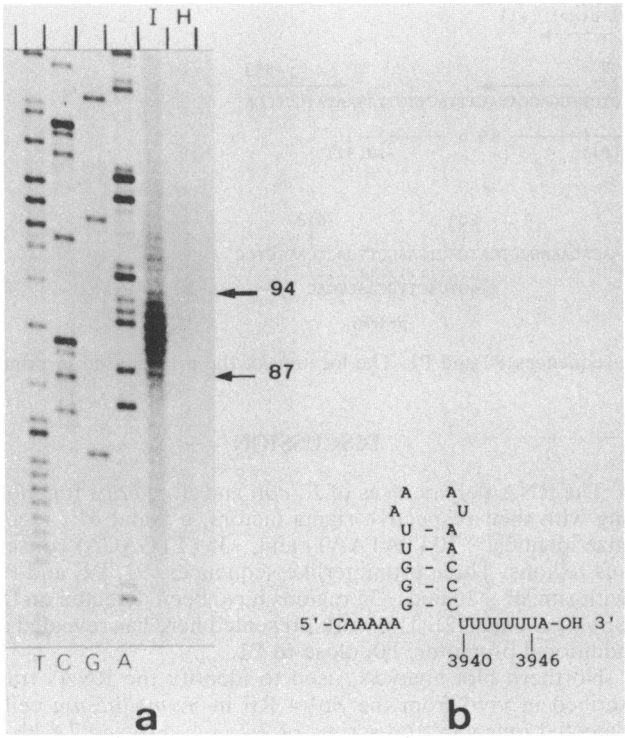

FIG. 4. S1 mapping of the 3′ terminus of the major SpV4 RNA transcript. (a) Analysis of S1-protected DNA fragments by denaturing polyacrylamide gel electrophoresis. Lanes: H and I, total RNAs extracted from healthy (H) or SpV4-infected (I) spiroplasma cells, hybridized with the 3′-end-labeled DNA probe (see Materials and Methods for design of the probe) and treated with S1 nuclease; T, C, G, and A, sequencing reactions used as ladders. (b) RNA structure of the SpV4 transcription terminator. The nucleotide corresponding to the major termination signal is indicated.

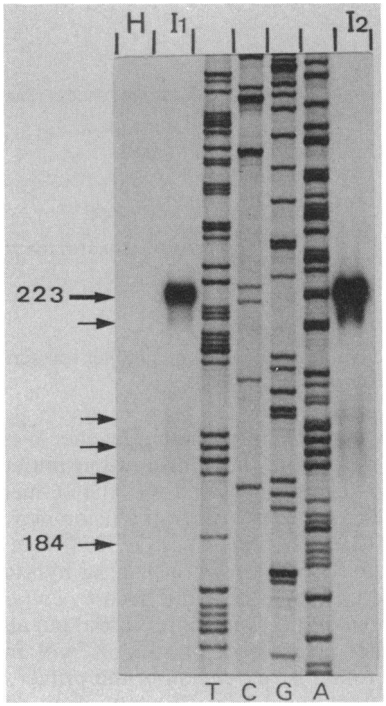

FIG. 5. S1 mapping of the 5′ terminus of the major SpV4 RNA transcript. Shown is analysis of S1-protected DNA fragments by denaturing polyacrylamide gel electrophoresis. Lanes: H and I, total RNAs extracted from healthy (H) or SpV4-infected (I) spiroplasma cells, hybridized with the 5′-end-labeled DNA probe (see Materials and Methods for design of the probe) and treated with S1 nuclease (2 and 4 μl of the reaction mixture were loaded in lanes I_1 and I_2, respectively); T, C, G, and A, sequencing reactions used as ladders.

the promoterlike sequence P1 (Fig. 6). However, unexpectedly, the major signal (223-nucleotide DNA band) corresponds to guanine 505 (Fig. 6). Examination of the nucleotide sequence revealed that position 505 was located three nucleotides downstream of the −10 region of a newly recognized promoterlike sequence, P0 (Fig. 6).

(ii) Primer extension analysis. To determine whether transcription starts from P0 or P1, the alternative technique of primer extension was used. A 16-mer oligonucleotide primer, complementary to nucleotides 603 to 618 (Fig. 6), was hybridized to the total RNAs from SpV4-infected cells and extended with reverse transcriptase. Two major extended cDNAs, 76 and 111 nucleotides long, were detected (Fig. 7), indicating two possible transcription initiation sites. One (the 76-nucleotide DNA band) corresponds to cytosine 543 of the CAT sequence of P1; the other (the 111-nucleotide DNA band) corresponds to guanine 508 of promoter P0 (Fig. 6).

In contrast to primer extension, S1 mapping yielded only one initiation site, that relative to promoter P0. Examination of the nucleotide sequence upstream of P1 revealed the presence of an inverted repeat that could lead upon transcript to the formation of a hairpin involving five adjacent G · C pairs (Fig. 8). Such a hairpin would probably be quite stable. As discussed below, this structure could explain our finding of two transcription start signals by the primer extension technique and only one by S1 nuclease mapping.

Promoters P0 and P1 are functional in *E. coli*. The −35 region of P0 (TTGACT) is only one nucleotide short of the eubacterial consensus sequence TTGACA; its −10 region (TAAAAA) is two nucleotides short of the consensus sequence TATAAT. In the case of P1, the −10 region is identical to the −10 consensus sequence and the −35 region (TTGTCT) is two nucleotides short. However, the spacing between the −10 and −35 regions of P1 with 19 nucleotides, instead of 17 nucleotides for P0, seems too long by *E. coli* standards to be a strong promoter. Therefore, the functioning of P0 and P1 in *E. coli* was investigated. Restriction fragment *Cla*I-*Nde*I (nucleotides 1 to 726) containing P0 and P1 was inserted in the promoter selection vector pKK232-8 (Fig. 9a) and used to transform *E. coli* cells.

In this vector, the CAT gene is free of its own promoter and can be expressed only when the DNA fragment inserted upstream of the gene contains a promoter sequence recognized by the bacterial RNA polymerase. Indeed, upon transformation of *E. coli* with the recombinant plasmid, 65 colonies resistant to chloramphenicol were obtained; no such colonies were produced in the case of the insert-free plasmid. One chloramphenicol-resistant clone was selected, and its recombinant plasmid pK1 was analyzed by Southern hybridization using SpV4 DNA as the probe. Resistance to chloramphenicol implies that the spiroplasmal *Cla*I-*Nde*I fragment containing P0 and P1 is inserted in plasmid pKK232.8 in the orientation shown in Fig. 9a. Treatment of the recombinant plasmid with *Taq*I should yield two fragments hybridizing with the probe, one of 260 nucleotides and the other of 630 nucleotides. This was found to be precisely the case (Fig. 9b). These results suggest that P0 and P1 carried by the recombinant plasmid pK1 are functional in *E.*

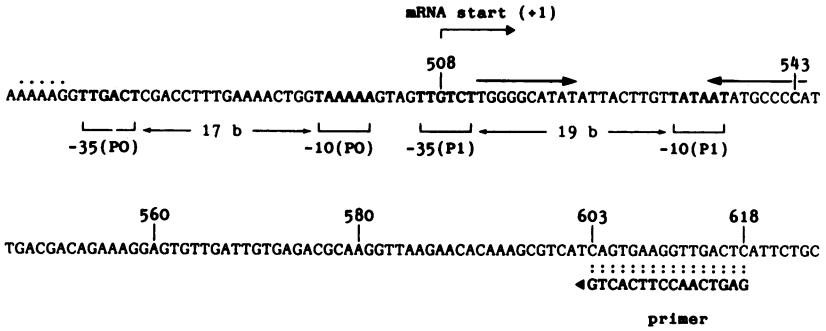

FIG. 6. Nucleotide sequence surrounding transcription promoter-like sequences P0 and P1. The location of the primer used in primer extension experiments is indicated.

coli and allow expression of the CAT gene. We used primer extension to determine the transcription initiation start on pK1 in *E. coli*. The primer used was the 16-mer oligonucleotide described above for primer extension in *S. melliferum*. Conditions of the experiment with *E. coli* were the same as those used with *S. melliferum*. RNAs for hybridization with the primer were isolated not only from *E. coli* cells containing the recombinant pK1 plasmid (Fig. 9a) but also from cells containing the insert-free plasmid. Two insert-specific cDNAs were found by extension of the primer, one with 79 and one with 111 nucleotides (Fig. 10). Thus, similar to the results obtained with the same primer in the case of *S. melliferum*, two cDNAs were obtained in *E. coli*. One of the cDNAs has the same size in *E. coli* and *S. melliferum*: the 111-nucleotide-long cDNA corresponding to the transcription start from promoter P0. The other cDNAs have very similar sizes, 79 nucleotides in *E. coli* and 76 nucleotides in *S. melliferum*.

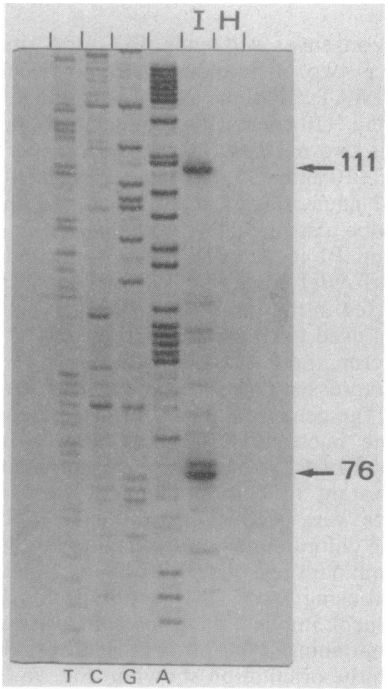

FIG. 7. Primer extension analysis of the 5′ terminus of the major SpV4 RNA transcript. Lanes: H and I, primer hybridized to total RNAs extracted from healthy (H) or SpV4-infected (I) spiroplasma cells and extended by reverse transcriptase; T, C, G, and A, sequencing reactions used as ladders.

DISCUSSION

The RNA polymerases of *E. coli* and *B. subtilis* functioning with their respective sigma factors, σ^{70} and σ^{43}, recognize identical −10 (TATAAT) and −35 (TTGACA) consensus regions. Three promoterlike sequences, P1, P2, and P3, with similar −10 and −35 regions have been detected on the SpV4 genome (12). The work presented here has revealed an additional promoter, P0, close to P1.

Northern blot analysis, used to identify the RNAs transcribed in vivo from the SpV4 RF in *S. melliferum* cells, detected four viral transcripts of 2.7, 3.4, 4.4, and 7.8 kb in infected but not uninfected cells. The sizes of these transcripts are in good agreement with the locations of the three promoterlike sequences P0 and/or P1, P2, and P3 and the position of the single terminatorlike sequence T. Indeed, transcripts initiated at P2, P0 and/or P1, and P3 and terminated at T would measure 2.7, 3.4, and 4.4 kb, respectively. The 7.8-kb transcript is probably also initiated from P0, P1, or both. Its transcription passes without stopping at terminator T and continues full circle around the RF before finally stopping at T. The reason for this behavior is not understood.

Since the 3.4-kb RNA is the earliest and the most abundant transcript, as shown by Northern blot analysis (Fig. 1), we have concentrated on determining the 5′ and 3′ ends of this transcript by S1 mapping and primer extension. Identification of the 5′ end should also indicate what promoter, P0, P1, or both, is functional.

From primer extension experiments, the 3.4-kb transcript appeared to be transcribed from two different initiation sites. One corresponds to cytosine 543 of the CAT sequence found five nucleotides downstream of the −10 region of P1 (Fig. 6). The second initiation site is at guanine 508 and corresponds

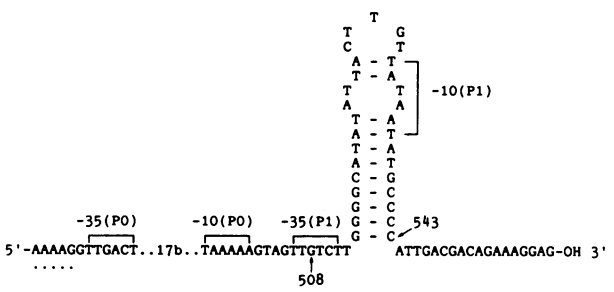

FIG. 8. Nucleotide sequence and putative hairpin structure of the SpV4 RNA transcript in the P1 promoterlike region. Arrows indicate positions of transcription start signals.

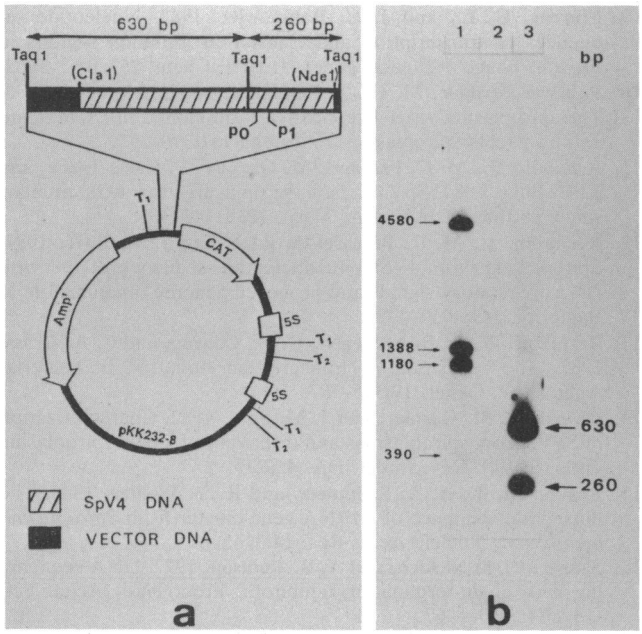

FIG. 9. Restriction map and Southern blot analysis of plasmid pK1. (a) Restriction map. Plasmid pK1 contains the 726-bp *Cla*I-*Nde*I restriction fragment of the SpV4 RF, inserted in the *Sma*I site of plasmid vector pKK232-8. (b) Hybridization of restricted DNAs with SpV4 DNA as the probe. Lanes: 1, *Hin*cII-restricted SpV4 RF; 2 and 3, plasmids pKK232-8 (lane 2) and pK1 (lane 3) restricted with *Taq*I.

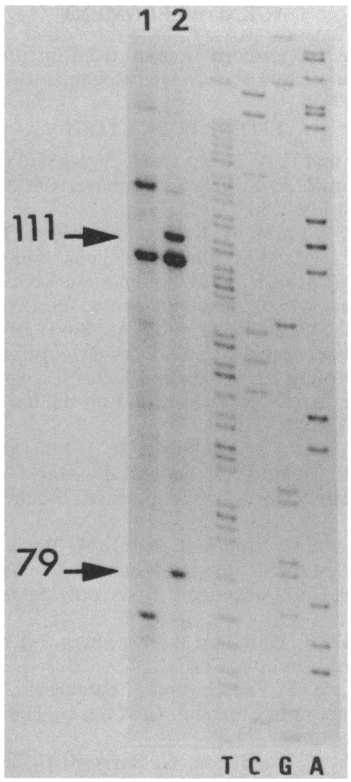

FIG. 10. Primer extension analysis of the 5′ terminus of the RNA transcribed in *E. coli* from SpV4 promoter P0. Lanes: 1 and 2, primer hybridized to total RNAs extracted from *E. coli* cells carrying plasmid pKK232-8 (in which the CAT gene is free of promoter [lane 1]) or plasmid pK1 (in which the CAT gene is under the control of a SpV4 promoter [lane 2]); T, C, G, and A, sequencing reaction used as ladders.

to P0. Thus, it would seem that transcription proceeds from two tandemly located promoters, P0 and P1.

S1 mapping of the 3.4-kb transcript gave unexpected results in that it revealed one strong transcription initiation signal at nucleotide 505 but a very weak signal, if any, at nucleotide 544, suggesting that P0, but not P1, is involved.

In primer extension, the signal apparently related to P1 could be an artifact resulting from the formation of a hairpin structure on the mRNA used as template by reverse transcriptase. The presence of this hairpin would prevent further elongation by the enzyme, yielding the short cDNA (76 nucleotides). Otherwise, the long cDNA (111 nucleotides) is obtained. Thus, P1 would not be used as a promoter. In agreement with this view, the sequence between the −10 and −35 regions is 19 bp long, too large by bacterial promoter standards.

The restriction fragment containing P0 and P1, when inserted into *E. coli* promoter selection vector pKK232.8, permits transcription of the CAT gene, since chloramphenicol-resistant bacterial transformants are obtained. Thus, this fragment possesses promoter activity in *E. coli*. We have determined by primer extension the transcription initiation point. Results very similar to those found in *S. melliferum* were obtained. Again, two cDNAs were detected, one 111 and one 79 nucleotides long. The first is identical in size to the large cDNA found with *S. melliferum*; the other is slightly longer, 79 nucleotides compared with 76. The 111-nucleotide cDNA designates guanine 508 as position +1. The location of promoter P0 fits well with this position. It therefore seems that P0 functions as a promoter in both *S. melliferum* and *E. coli* and that in both cases RNA transcription is initiated at nucleotide 508. This observation also demonstrates that the same promoter, P0, is recognized in both organisms. It is known from previous work that the core RNA polymerase of *S. melliferum* has a subunit structure of the eubacterial type: $\alpha_2\beta\beta'$ (5).

The −10 and −35 regions of P0 are TAAAAA and TTGACT, respectively. These sequences are close to the consensus sequences of bacterial promoters recognized by eubacterial RNA polymerases functioning with the general sigma factor. In addition, these regions are separated by 17 bp, a value within the range of eubacterial (*E. coli* and *B. subtilis*) promoter standards. Strong promoters of *B. subtilis* bacteriophages have an additional consensus sequence, AAAAG, at −40 (9), but in *E. coli* there is very little homology in this region. It is worth noting that P0 has also this additional −40 consensus sequence (Fig. 6). This finding agrees with the phylogenetic origin of the spiroplasmas as having arisen from ancestors of low-G+C gram-positive eubacteria such as *B. subtilis* and *Clostridium* spp. (19).

S1 mapping has shown that the 3′-OH end of the 3.4-kb transcript is within the stretch of uridine residues (nucleotides 3940 to 3946) following the sequence 5′-AACCC-3′, which can form a hairpin structure with the upstream sequence 5′-GGGUU-3′ (Fig. 4b). This result is indicative of transcription termination at a Rho-independent terminator.

Our experiments show that there is transcription initiation at nucleotide 508 and termination at nucleotides 3940 to 3946, leading to a transcript of 3,432 to 3,438 nucleotide residues. This value agrees well with the size (3.4 kb) of the major transcript detected by Northern blot analysis. Finally, the same initiation and termination sites could also be involved in transcription of the 7.8-kb transcript.

ACKNOWLEDGMENT

This work was supported by a grant from Institut National de la Recherche Agronomique and Conseil Régional d'Aquitaine.

LITERATURE CITED

1. **Berk, A. J., and P. A. Sharp.** 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell **12**:721–732.
2. **Bove, J. M., P. Carle, M. Garnier, F. Laigret, J. Renaudin, and C. Saillard.** 1988. Molecular and cellular biology of spiroplasmas, p. 244–364. *In* R. F. Whitcomb and J. G. Tully (ed.), The mycoplasmas, vol. 5. Academic Press, Inc., New York.
3. **Chevalier, C., C. Saillard, and J. M. Bové.** 1990. Organization and nucleotide sequences of the *Spiroplasma citri* genes for ribosomal protein S2, elongation factor Ts, spiralin, phosphofructokinase, pyruvate kinase, and an unidentified protein. J. Bacteriol. **172**:2693–2703.
4. **Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18**:5294–5299.
5. **Gadeau, A. P., C. Mouchès, and J. M. Bové.** 1986. Probable insensitivity of mollicutes to rifampin and characterization of spiroplasmal DNA dependent-RNA polymerase. J. Bacteriol. **166**:824–828.
6. **Homes, W. H., T. Platt, and M. Rosenberg.** 1983. Termination of transcription in *E. coli*. Cell **32**:1029–1032.
7. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
8. **Mouchès, C., T. Candresse, G. Barroso, C. Saillard, H. Wroblewski, and J. M. Bové.** 1985. Gene for spiralin, the major membrane protein of the helical molicute *Spiroplasma citri*: cloning and expression in *Escherichia coli*. J. Bacteriol. **164**: 1094–1099.
9. **Murray, C. L., and J. C. Rabinowitz.** 1982. Nucleotide sequences of transcription and translation initiation regions in *Bacillus* phage Φ29 early genes. J. Biol. Chem. **257**:1053–1062.
10. **Pascarel-Devilder, M. C., J. Renaudin, and J. M. Bové.** 1986. The spiroplasma virus 4 replicative form cloned in *Escherichia coli* transfects spiroplasmas. Virology **151**:390–393.
11. **Renaudin, J., M. C. Pascarel, M. Garnier, P. Carle-Junca, and J. M. Bové.** 1984. SpV4, a new spiroplasma virus with circular, single stranded DNA. Ann. Virol. **135E**:343–361.
12. **Renaudin, J., M. C. Pascarel-Devilder, and J. M. Bové.** 1987. Spiroplasma virus 4 (SpV4): nucleotide sequence of the viral DNA, regulatory signals and proposed genome organization. J. Bacteriol. **169**:4950–4961.
13. **Reznikoff, W. S., D. A. Siegele, D. W. Cowing, and C. A. Gross.** 1985. The regulation of transcription initiation in bacteria. Annu. Rev. Genet. **19**:355–387.
14. **Ricard, B., M. Garnier, and J. M. Bové.** 1982. Characterization of SPV3 from spiroplasmas and discovery of a new spiroplasma virus (SPV4). Rev. Infect. Dis. **4**:S275.
15. **Rogers, M. J., A. A. Steinmetz, and R. T. Wolker.** 1986. The nucleotide sequence of a tRNA gene cluster from *Spiroplasma melliferum*. Nucleic Acids Res. **14**:3145.
16. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
17. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
18. **Weaver, R. F., and C. Weissmann.** 1979. Mapping of RNA by a modification of the Berk-Sharp procedure: the 5′ termini of 15S beta-globin mRNA precursor and mature 10S beta-globin mRNA have identical map coordinates. Nucleic Acids Res. **7**:1175–1183.
19. **Woese, C. R.** 1987. Bacterial evolution. Microbiol. Rev. **51**:221–271.