JOURNAL OF BACTERIOLOGY, Nov. 1995, p. 6601–6609 Vol. 177, No. 22
0021-9193/95/$04.00+0

Organization and Transcriptional Analysis of the *Listeria* Phage A511 Late Gene Region Comprising the Major Capsid and Tail Sheath Protein Genes *cps* and *tsh*

MARTIN J. LOESSNER* AND SIEGFRIED SCHERER

Institut für Mikrobiologie, Forschungszentrum für Milch und Lebensmittel Weihenstephan, Technische Universität München, D-85350 Freising, Germany

Received 21 June 1995/Accepted 10 September 1995

A511 is a broad-host-range, virulent myovirus for *Listeria monocytogenes*. The genes encoding major structural proteins of the capsid (*cps*) and tail sheath (*tsh*) were mapped to a 10.15-kb late gene fragment. We have determined the complete nucleotide sequence of this region and confirmed the identities of Cps (48.7 kDa) and Tsh (61.3 kDa) by N-terminal amino acid sequencing of both proteins. In addition, nine other open reading frames were identified. On the basis of amino acid sequence homologies to known phage-encoded proteins, some putative functions and locations could be assigned to some of the deduced gene products. We present evidence that the *cps* product is proteolytically cleaved between Lys-23 and Ser-24 to yield the 444-residue polypeptide found in the mature viral capsid. We also found that the N-terminal methionine is absent from the mature tail sheath protein. *cps* and *tsh* are late genes; mRNAs first appear 15 to 20 min after infection of *L. monocytogenes*. Northern (RNA) hybridizations of total late mRNA with specific oligonucleotide probes were used to determine the sizes of respective transcripts. Primer extension analyses enabled the positive identification of six late promoters, which were found to differ from those identified in the chromosome of *Listeria* spp. The bulk of transcripts from *cps* and *tsh* arise from two phage promoters with identical 13-nucleotide sequences (TGCTAGATTATAG [core region underlined]) in the −10 region which we speculate determines specific and timed expression of these genes. A 123-nucleotide leader sequence at the 5′ end of the *cps* transcript was predicted to form a strong secondary structure ($\Delta G = -40.7$ kcal [−170.3 kJ]/mol). Our results show that the strongly expressed A511 *cps* and *tsh* genes are included in two separate gene clusters and are independently regulated at the transcriptional level.

Listeria monocytogenes is an opportunistic gram-positive pathogen responsible for a variety of severe infections in both animals and humans (reviewed in reference 17). Much research is presently directed towards understanding the molecular mechanisms of its pathogenicity, as well as developing tools for molecular research and rapid detection methods.

Listeria bacteriophage A511 is the only known virulent phage for this genus and is unrelated to all other *Listeria* phages (31, 43). A511 has an extremely broad host range, being capable of lysing approximately 95% of all *L. monocytogenes* strains of serovars 1/2 and 4b (30, 33). This makes it useful in bacteriophage typing of listeriae (29, 30) and renders it a promising candidate for development of a specific reporter vehicle for rapid detection of *Listeria* spp. in foods and environmental samples.

The A511 virion is a myovirus with an isometric capsid 88 nm in diameter and a straight tail 200 nm long with a contractile sheath around the inner tail tube (43). Upon recognition of its cell surface receptor (42), the sheath is contracted to two-thirds of its original length and the linear double-stranded DNA molecule of approximately 116 kb (31) is injected into the host cell. After a latent period of 55 to 60 min, the multiplication cycle is terminated by cell lysis through the *ply* gene product, a *Listeria* cell wall-specific *N*-acetylmuramoyl-L-alanine amidase (35). Burst size is between 40 and 50 new particles per infected cell (33).

The strongest promoters in phages, leading to the highest levels of protein synthesis, are often those driving expression of the late genes encoding major structural proteins of the virion. Identification of their locations, sequences, and regulation is useful not only for understanding basic biological processes of phage-host interaction but also for practical applications, such as efficient and tightly regulated expression of heterologous genes in the host organism (40).

As a step towards this, we have cloned and analyzed a large DNA region containing the genes for the most abundant proteins in the A511 particle: the major capsid protein (Cps) and the tail sheath element (Tsh). Nine other open reading frames (ORFs) were also identified. We found evidence that the primary *cps* gene product is posttranslationally processed to yield the mature protein found in the viral capsid. We also present evidence that A511 late transcription is initiated at promoter sequences which differ significantly from *Listeria* promoter sequences. A hypothetical model regarding enhanced stability of the highly expressed *cps* transcript through 5′-end secondary structure is proposed.

MATERIALS AND METHODS

Organisms, plasmids, and culture conditions. *L. monocytogenes* WSLC 1001 (ATCC 19112) was used for propagation of *Listeria* phage A511 (30). *Escherichia coli* DH5αMCR (Life Technologies) was used for subcloning A511 DNA fragments inserted into the vector pBluescript II SK− (Stratagene).

L. monocytogenes was grown in tryptose broth at 30°C, phage A511 was propagated and purified on preformed CsCl gradients as described earlier (43),

* Corresponding author. Mailing address: Institut für Mikrobiologie, Forschungszentrum für Milch und Lebensmittel Weihenstephan, Technische Universität München, Vöttinger Str. 45, D-85350 Freising, Germany. Phone: (8161) 71-3859. Fax: (8161) 71-4492.

and *E. coli* was grown in standard LB medium (38) at 37°C. For selection of plasmid-bearing cells, ampicillin was added at 100 μg/ml.

Western blotting (immunoblotting) and amino-terminal sequencing. Phage structural proteins were separated by horizontal sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to polyvinylidene difluoride membranes as outlined elsewhere (24, 31). Coomassie blue-stained protein bands (Cps and Tsh) were excised from the membranes and used for determination of the N-terminal amino acids with an ABI 477A protein sequencer.

Restriction enzyme analysis and Southern hybridization. A511 DNA was extracted and purified by described methods (31, 38). Enzymatic digestions with various endonucleases (*Bsp*HI, *Hin*dIII, *Xba*I, *Mun*I, *Spe*I, *Nde*I, *Eco*RV, *Sna*BI, *Psp*1406I, and *Bsp*1407I) were carried out according to the suppliers' recommendations (United States Biochemical, Stratagene, and New England Biolabs). Conditions for agarose gel electrophoresis and vacuum blotting of the DNA fragments to nylon membranes were as reported earlier (31). Degenerated oligonucleotide probes were designed according to the N-terminal sequences of both Cps and Tsh and 3′ end labeled with digoxigenin by using terminal transferase (Boehringer). They were then hybridized to the membrane-bound A511 DNA fragments in standard 5× SSC hybridization buffer (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) at 53°C for 16 h. Hybrids were detected by chemiluminescence reaction as described in the manufacturer's manual (Boehringer).

Cloning procedures. Selected DNA fragments of a convenient size (2 to 4 kb) which had hybridized to the respective probes were isolated from low-melting-point agarose gels as previously described (37). Fragments were made blunt ended by filling in with Klenow polymerase (United States Biochemical) and ligated into the cloning vector, which had been linearized with *Eco*RV and treated with shrimp alkaline phosphatase (United States Biochemical). Reaction mixtures were electroporated into the *E. coli* host, and colonies harboring recombinant plasmids were selected by blue-white screening (38). The identity of the cloned fragments was confirmed by colony hybridizations and restriction endonuclease analysis of plasmid minipreps.

DNA sequencing and computer analysis. Large-scale plasmid purification was done by using ion-exchange columns (Qiagen). Nucleotide sequences of the cloned phage DNA fragments were determined from both strands, by primer walking with synthetic oligonucleotides as sequences became available. Sequenase version 2.0 (United States Biochemical) and α-^{35}S-dATP (Amersham) were used in all sequencing reactions. DNA sequences and proteins translated from open reading frames were analyzed by using the DNASIS and PROSIS programs from Hitachi.

Isolation of total RNA. Isolation of total RNA was carried out essentially as reported earlier (34, 35) by using the recombinant *Listeria* bacteriophage A118 endolysin PLY118 for instant lysis of cells. Liberated RNA was purified on spin columns (Qiagen), digested with RNase-free DNase (Promega), resuspended in Tris-EDTA buffer, and stored at −30°C. Concentrations of the various preparations were measured photometrically, and agarose gel electrophoresis demonstrated the absence of contaminating DNA.

Northern (RNA) hybridizations. In order to determine the time course of *cps* and *tsh* transcription, dot blots were made from RNA preparations isolated at various time points after infection of WSLC 1001 with A511. Membrane-bound mRNAs were then hybridized to digoxigenin-labeled oligonucleotide probes complementary to regions in the coding sequences of *cps* (sequence positions 3013 to 3039) and *tsh* (positions 8716 to 8740). Hybridization and detection of hybrids were carried out as described above.

For determination of the approximate transcript sizes, denaturing agarose gel electrophoresis (7% formaldehyde [38]) was performed. Three micrograms of total RNA preparation isolated 60 min postinfection was electrophoretically separated, and then the fragments were vacuum transferred to nylon membranes. Oligonucleotides complementary to regions of the presumptive transcripts generated from the 10.15-kb fragment of A511 DNA (see legend to Fig. 6) were selected for RNA-DNA hybridization, which was carried out as described above.

Primer extensions. According to the approximate transcription initiation sites suggested from the Northern analysis, we designed oligonucleotide primers complementary to the near 5′ ends of the proposed *cps* and *tsh* transcripts: CPS-PX (positions 1794 to 1820) and TSH-PX (position 6074 to 6100). Additional primers were designed to hybridize approximately 50 to 100 nucleotides (nt) downstream (ORF3-PX, positions 3664 to 3690; ORF9-PX, positions 9516 to 9540) from putative promoters, which we assumed should be located downstream of the predicted RNA polymerase termination signals.

The reactions were carried out to allow continuous incorporation of α-^{35}S-dATP. Total RNA (6 μg, isolated at 60 min postinfection) was annealed with 5 pmol of primer in a volume of 8 μl at 75°C for 5 min and incubated at 52°C for 10 min. To each reaction mixture was then added 4.0 μl of 5× reverse transcriptase buffer (Life Technologies), 2.0 μl of 0.1 M dithiothreitol, 0.5 μl of RNAsin (Promega), 3.0 μl of labeling mix (10 μM [each] dCTP, dGTP, and dTTP), 1.0 μl (10 μCi) of α-^{35}S-dATP, and 2.5 μl (500 U) of SuperScript II reverse transcriptase (Life Technologies). After 10 min at 42°C, 1 μl of 10 mM deoxynucleoside triphosphate solution was added and the incubation was continued for another 30 min. Reactions were terminated by the addition of 9 ml (0.4 volumes) of Sequenase Stop-Solution (United States Biochemical). Two to 3 μl of each reaction mixture was then loaded onto a standard sequencing gel alongside a sequence ladder derived from the corresponding segment of DNA with the same primer.

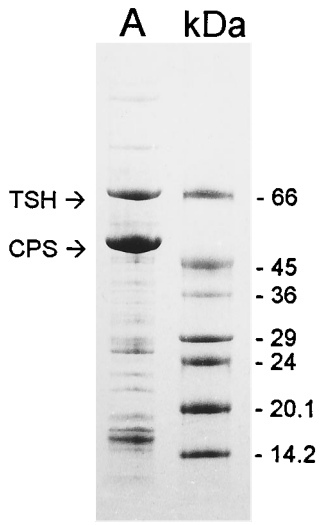

FIG. 1. Structural proteins of purified A511 separated by SDS-PAGE (Lane A). The molecular size marker SDS-7L (Sigma) was used for calibration (lane kDa); protein sizes are indicated on the right in kilodaltons.

Nucleotide sequence accession number. The DNA sequence reported in this paper has been assigned accession number X91069 in the EMBL, GenBank, and DDBJ databases.

RESULTS

Identification and cloning of *cps* and *tsh* gene clusters. The most abundant structural proteins of A511 are the major capsid protein (Cps) and the tail sheath protein (Tsh), whose identities have been shown earlier (31, 33). Laser-densitometrical scanning of the protein profile shown in Fig. 1 revealed that the percentages of the total protein content of the viral particle of Cps (49 kDa) and Tsh (65 kDa) are approximately 34 and 14%, respectively. The N-terminal sequences are S-F-T-T-G-Y-G-I-T-P-D-T-Q-T-D-A-G-A-L-R-R for Cps and A-I-E-I-Y-P-R for Tsh. The sequences allowed the design of degenerated oligonucleotide probes corresponding to the 5′ region of the genes: A511CPS (30-mer), TA(T/C)GG(T/A)AT (T/C)AC(T/A)CC(T/A)GATAC(T/A)CAAAC(A/T) GAT, and A511TSH (18-mer), GC(T/A)AT(T/C)GAAAT(T/C)TA (T/C)CC(T/A).

Southern hybridization of digested A511 DNA to the labeled probes led to the identification of the 5′ end of *cps* on a 2.3-kb *Xba*I-*Hin*dIII fragment and the 5′ end of *tsh* on a 2.9-kb *Nde*I fragment. Subcloning in *E. coli* yielded plasmids pCA511-A and pTA511-A, respectively (Fig. 2). Following nucleotide sequencing of the A511-derived inserts it became clear that the cloned fragments did not include the entire coding regions for both genes (Fig. 2). This necessitated subcloning and sequencing of two additional fragments which had hybridized to the probes for *cps* (4.2-kb *Spe*I fragment in pCA511-B) and *tsh* (3.5-kb *Bsp*HI fragment in pTA511-B). Alignment of the four sequences yielded a contiguous stretch of 10,152 bp.

Sequence analysis. In addition to *cps* and *tsh*, computer analysis revealed the presence of nine other ORFs (Fig. 2D). The nucleotide sequence and inferred amino acid sequence for *cps*, including upstream and downstream regulatory regions, is shown in Fig. 3. The major structural protein genes are preceded by ribosome binding sites (RBS) of perfect (*cps*) or

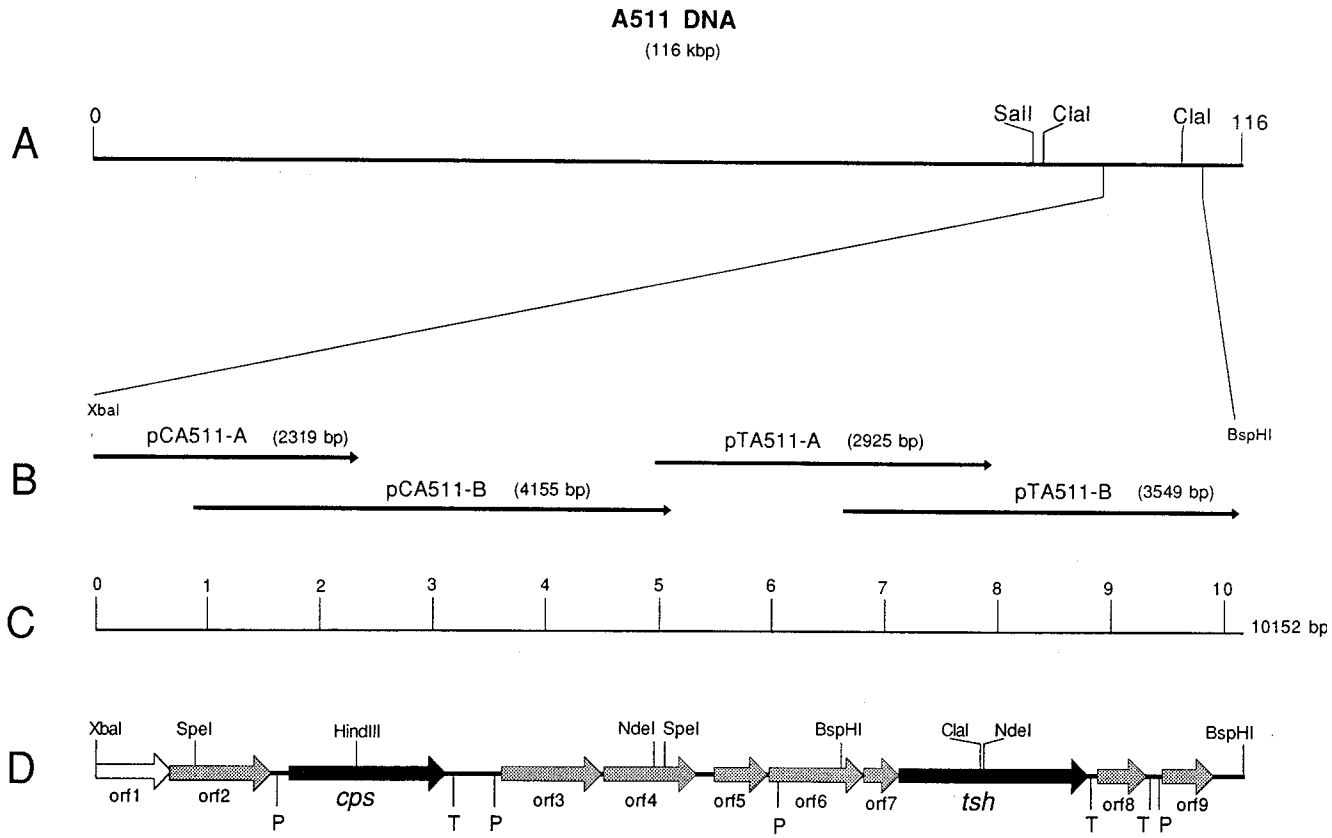

FIG. 2. Locations of the cloned DNA fragments on the physical map of A511 (A). Plasmid names and individual sizes of their inserts are indicated (B) and drawn to scale. Overlapping sequences were assembled into a contiguous stretch of 10,152 bp (C). In addition to the genes (black arrows) for the major capsid (*cps*) and tail sheath (*tsh*) proteins, further ORFs (ORF2 to ORF9; shaded arrows) could be identified (D). ORF1 is incomplete (white arrow). Transcription proceeds rightward; identified promoter sequences are indicated by P, and the positions of possible transcription terminators are indicated by T.

near-perfect (*tsh*) complementarity to the 3′ end of *L. monocytogenes* 16S rRNA (16), with a 9- to 10-nt spacer (A and T only) to the ATG start codon. Furthermore, both genes are followed by inverted repeats which are presumed to function as transcription terminators. A third inverted repeat was found downstream of ORF8. Free energies of the predicted secondary structures were calculated to be −19.6 kcal [−82.0 kJ]/mol (T_{cps}; nt 3159 to 3187), −16.5 kcal [−69.0 kJ]/mol (T_{tsh}; nt 8821 to 8850), and −22.0 kcal [−92.0 kJ]/mol (T_{ORF8}; nt 9303 to 9353). The stem-loop downstream of *tsh* is less than perfect; therefore, its significance is questionable. Another interesting sequence motif, which is located in the region between *cps* and the downstream promoter for ORF3 (Fig. 3), contains a 15-bp sequence repeated perfectly six times. However, the function of this repeat is unclear. Table 1 summarizes the basic characteristics of the identified genes and ORFs and their possible products.

Evidence for posttranslational processing of the *cps* product. Comparison of the experimentally determined N terminus of Cps with the protein sequence translated from the gene reveals that the first 23 amino acids (aa) (Fig. 3) of the protein are removed when the phage capsid is matured. Nucleotide sequencing of the same region present in three independent clones ruled out the possibility of an error in the gene sequence.

Evidence that the position between Lys-23 and Ser-24 is a potential cleavage site in Cps was also derived from local amino acid sequence homologies, presented in Fig. 4. The 23-aa N-terminal peptide of the primary translation product (Pro-Cps) shows significant homology to the automodification domain of bovine ADP-ribosyltransferase (37). Moreover, the sequence around and directly downstream from the proposed processing site shows a high degree of sequence identity to the C-terminal domain of the putative ORF1 product. Sequence homology of Cps to the N-terminal domain of *E. coli* methionine aminopeptidase was also found (2) (23.5% identity and 68% similarity in a 95-aa overlap). The predicted M_r of the processed *cps* product (48,711) agrees well with the molecular mass observed on SDS-PAGE (49.5 kDa).

Regarding Tsh, the experimentally determined amino acid sequence was identical to that derived from the DNA sequence, except that the N-terminal methionine was absent from the mature protein. The observed molecular mass (Fig. 1; 65 kDa) was in reasonable agreement with the M_r predicted (61,300).

***cps* and *tsh* are late genes.** The latent phase for A511 (at 30°C) is about 55 to 60 min, and cell lysis is usually complete after 90 min (33, 35). In order to determine whether *cps* and *tsh* are true late genes regulated at the transcriptional level, we performed dot blot hybridizations of total RNA isolated at different time intervals after A511 infection, with labeled oligonucleotide probes complementary to the respective transcripts (Fig. 5). It is evident that transcription starts at approximately 20 min postinfection and that the number of both transcripts increases until lysis occurs. This clearly demonstrates the late nature of these genes.

Transcriptional analysis. Northern hybridizations were carried out to determine the approximate sizes of the major late

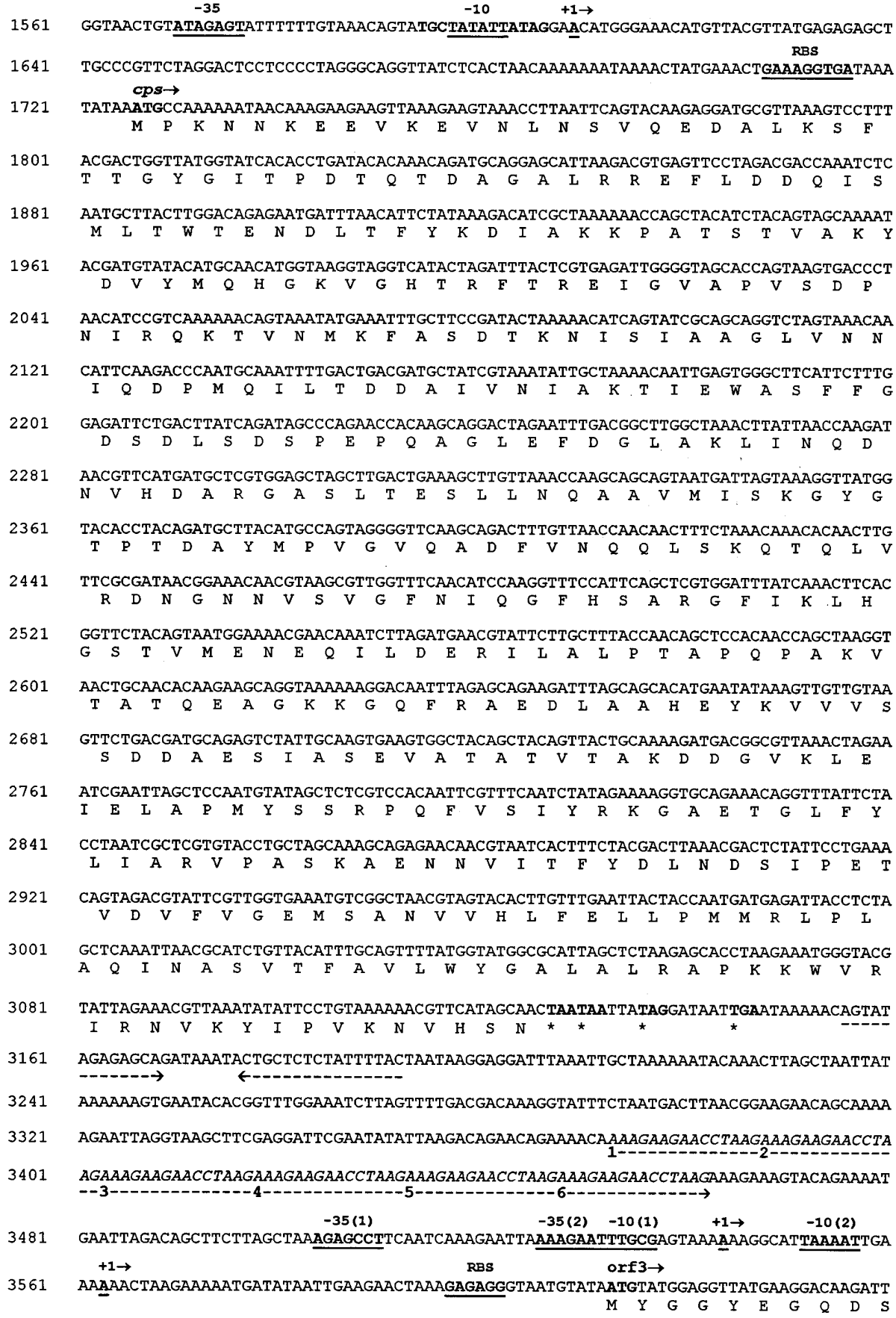
-35
-10
+1→
RBS
cps→
-35(1)
-35(2)
-10(1)
+1→
-10(2)
RBS
orf3→

FIG. 3. Nucleotide sequence of the A511 DNA region including *cps* and upstream and downstream regulatory regions. The predicted amino acid sequence of Cps is given below the DNA sequence. Promoter sequences and presumed RBS are underlined and marked. Stop codons are marked with asterisks, the inverted repeat downstream of *cps* (presumed transcription terminator T_{cps}) is indicated by arrows, and the six-times-repeated 15-bp perfect direct repeat upstream of the tandem ORF3 promoter is shown in italics and indicated by broken underlining. Numbering of the sequence is according to Fig. 2.

transcripts encoding *cps* and *tsh*. Similar analyses were performed for the other identified ORFs in this region. The results obtained by hybridization of late RNA with six different oligonucleotide probes are shown in Fig. 6. The transcript covering ORF1 and ORF2 is approximately 3.7 kb long. Since the probe specific for *cps* mRNA detects both a high-abundance *cps* transcript (1.6 kb) and the 3.7-kb mRNA for the upstream reading frames, both transcripts must be terminated at the same point, i.e., T_{cps}. This also indicates a promoter about 0.5 kb upstream from the *Xba*I site at sequence position 1. The probe specific for ORF3 yielded a relatively weak signal in the range of 6 kb. This corresponds to the presumed termination signal downstream of ORF8. An intense band of approximately 2.7 kb was seen after hybridization with the *tsh*-specific probe, which indicated transcription initiation from a promoter within ORF6, regardless of whether T_{tsh} or T_{ORF8} was used for termination. The ORF8-specific oligonucleotide revealed different mRNA sizes of approximately 3 and 0.5 kb. This may be explained by mRNA processing, possibly in the region of the stem-loop structure downstream of *tsh*, which would result in a 500-nt fragment of mRNA encoding the ORF8 product. The rightmost coding region (ORF9) is transcribed on a 1-kb message, the 3′ end of which extends approximately 300 nt beyond the 3′ end of the cloned *Bsp*HI fragment in pTA511-B.

Identification of promoter regions by primer extension analysis. The final step in our effort to identify specific promoter sequences responsible for A511 late gene transcription was defining the 5′ ends of the late transcripts. Primer extension reactions were performed with total RNA isolated from *L. monocytogenes* cells 60 min after infection with A511, and our results are presented in Fig. 7. A very intense signal was obtained by extension of the *cps* transcript, such that electrophoresis had to be repeated with a fivefold-diluted sample to obtain a band which could then be correlated to nt 1610 (Fig. 7A). The start of the transcript covering ORF7 and *tsh* (and probably ORF8) lies within ORF6 at nt 6050 (Fig. 7B). Apparently, the initiations upstream of both ORF3 and ORF9 occur from tandem promoters, as evidenced by the extension reactions shown in Fig. 7C and D.

These results enabled the identification of −10 and −35 promoter sequence motifs. Table 2 is a compilation of identified *Listeria* promoter sequences and compares them with the positively identified A511 late promoters. The two promoters specifying *cps* and *tsh* mRNAs show identical 13-nt sequences with a hexanucleotide core (TATATT) in the −10 region. Except for $P_{ORF9\text{-}A}$, the other promoters show more or less variation within this motif. Regarding the −35 region, we observed some conserved bases: positions 1, 3, and 4 of the 7-mer are identical in all six promoters. While the −35 motifs are significantly different from those so far identified to occur in *Listeria* spp., the −10 regions of both promoters seem to be more conserved. An interesting aspect is the predicted strong secondary structure (ΔG, −40.7 kcal [−170.3 kJ]/mol) in the 5′ end of the *cps* transcript starting at nt 1610 (Fig. 8). Folding of the mRNA in this way would also result in an exposed RBS.

DISCUSSION

In this study, we identified, cloned, and sequenced two A511 late genes which encode the major capsid protein (gene *cps*) and the tail sheath protein (gene *tsh*). The genes are located on two different transcriptional units (operons) and are separated by approximately 4 kb. Nine additional ORFs were found on

TABLE 1. Characteristics of *cps* and *tsh* and other ORFs (and corresponding gene products) present on the cloned 10.15-kb late gene region of A511

Genes or ORFs		Gene products (proteins)			
Name	Length (nt)	Length (aa)	Size (kDa)	Predicted pI	Location, function, and comments
ORF1	>670	>222	>25.0		Incomplete ORF; probably capsid; partial homology to N terminus of *cps* product (Fig. 4); may be proteolytically cleaved
ORF2	894	297	33.5	4.3	Probably capsid; highly charged; rich in glu; partial homologies to various DNA-binding proteins; partial homology to prehead core component PIP (gp67) of phage T4 (41) (24% identity and 86% similarity in a 58-aa overlap)
cps	1,404	467→23 + 444	51.4→2.7 + 48.7	5.3	Major capsid protein Cps; posttranslationally processed; homologies to gpORF1 and other proteins (Fig. 4)
ORF3	882	293	33.2	4.9	Probably tail; partial homology to tail tubular protein A (gp11) of bacteriophage T7 (15) (47% identity and 76% similarity in a 17-aa overlap)
ORF4	819	272	31.4	8.8	Probably tail; partial homologies to rod-shape-determining protein MREB from *E. coli* (10) (26% identity and 63% similarity in a 156-aa overlap) and to the upper collar protein (gp10) of bacteriophage PZA (36) (27.5% identity and 76% similarity in a 33-aa overlap)
ORF5	477	158	18.7	10.5	Possibly tail?
ORF6	840	279	31.5	4.5	Possibly tail?
ORF7	321	106	12.3	9.2	Possibly tail? Highly charged; rich in Lys and Glu; partial homology to gpORF2 (33.5% identity and 70% similarity in a 66-aa overlap)
tsh	1,689	562→1 + 561	61.3	4.7	Major tail sheath protein Tsh
ORF8	423	140	15.5	6.2	
ORF9-A	444	147	17.3	4.7	May be expressed in two different forms (see text)
ORF9-B	390	129	15.3	5.2	

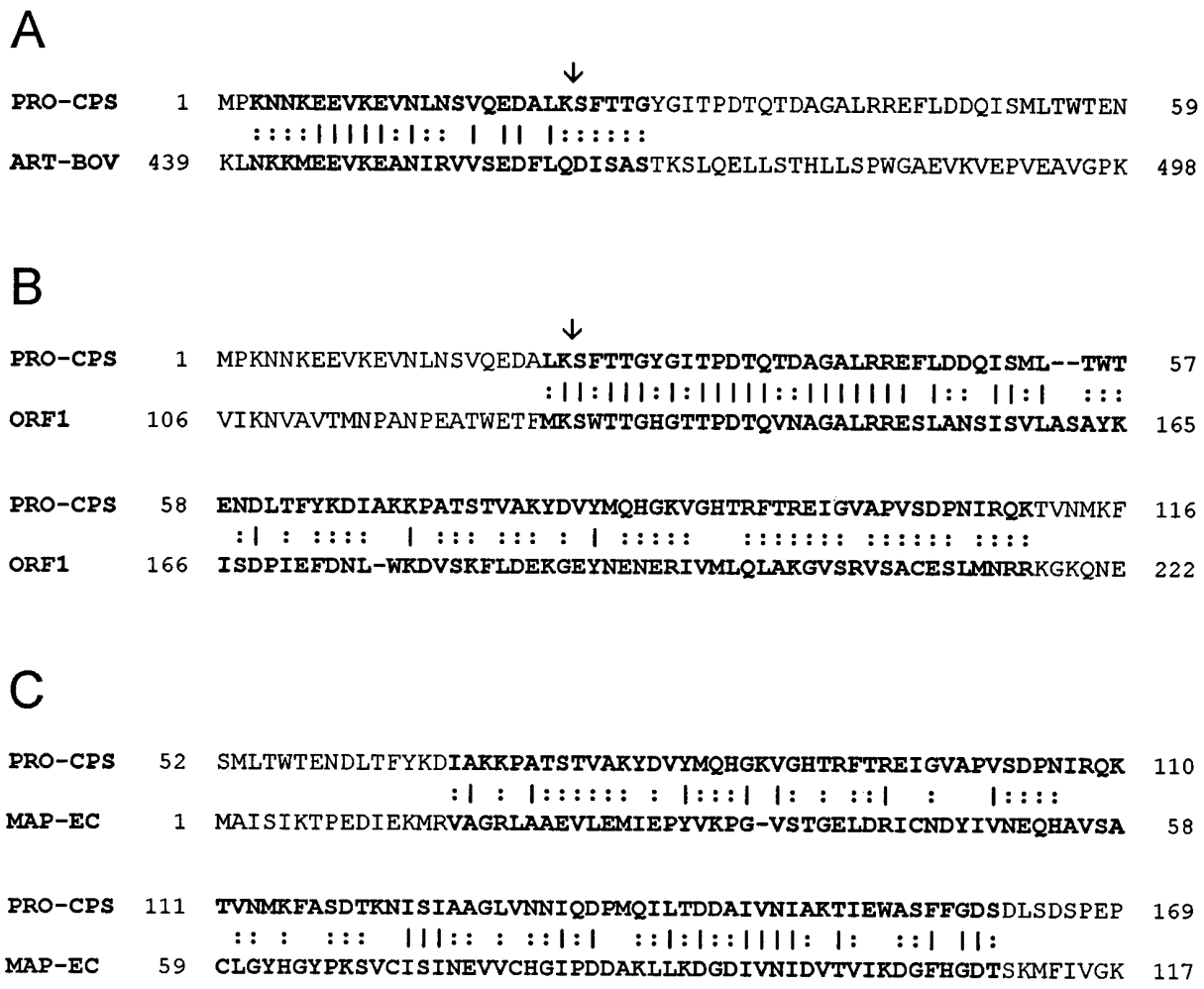

FIG. 4. Alignment of the primary amino acid sequence of the *cps* gene product (Pro-Cps) with (A) the automodification domain of bovine ADP-ribosyltransferase (ART-BOV) (35), (B) the C terminus of the presumed protein encoded by ORF1 and (C) the N-terminal domain of *E. coli* methionyl aminopeptidase (MAP) (2). Homologous domains are shown in bold letters. The cleavage site for processing of Pro-Cps to the form found in the viral capsid is marked by an arrow. Identical residues (|) and conservative replacements (:) are indicated.

the sequenced 10.15-kb fragment. A genome composed of a specifically ordered set of genetic modules is common to the tailed phages (see references 4, 6, and 7). The genes for related structural components are often clustered together in the viral genome. We conclude that this is the case for A511 as well. The locations of Cps and Tsh are clear. However, only tentative functions can be assigned to the gene products of the remaining ORFs (Table 1). Local amino acid sequence homologies suggest possible roles for some of the putative proteins: the products of ORF1 and ORF2 may be involved in head morphogenesis, while the polycistronic message transcribed from the tandem P_{ORF3} promoter probably specifies tail-associated proteins. Another tandem promoter, P_{ORF9}, was positively identified upstream of ORF9, where the shorter transcript excludes the proximal RBS and start codon present on the longer mRNA. The roles or locations of gpORF8 and the slightly different proteins made from ORF9-A and ORF9-B are unclear, since we found no similarity to other known proteins in the databases.

Amino-terminal sequencing of Cps and Tsh showed post-translational modifications in both proteins. Evidence presented here suggests that the primary product of *cps* is cleaved

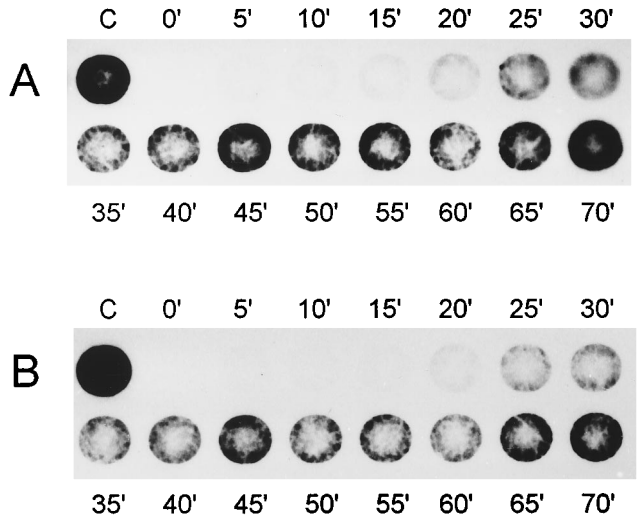

FIG. 5. Detection of specific transcripts of *cps* (A) and *tsh* (B) following gene expression of A511. Total RNA was isolated from infected cells at the time points indicated, and specific mRNAs were detected by dot blot hybridization to oligonucleotide probes complementary to the respective genes. C, Positive control (A511 DNA); 0′, immediately prior to phage infection; 5′ to 70′, times in minutes after infection.

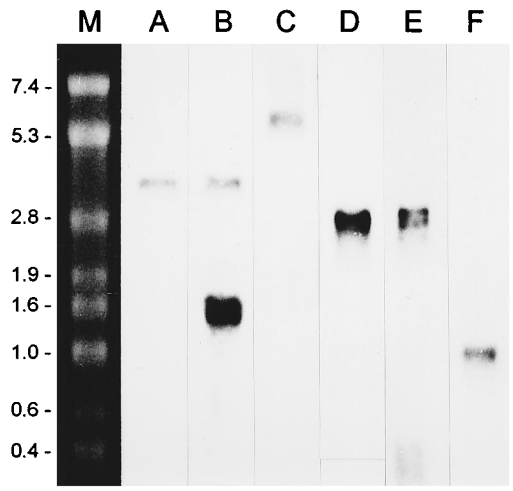

FIG. 6. Northern hybridizations showing the approximate sizes of mRNAs transcribed during A511 late gene expression. Total late RNA was separated by denaturing agarose gel electrophoresis, transferred to nylon membranes, and hybridized to oligonucleotide probes complementary to the mRNA transcripts. Lane M shows an ethidium bromide-stained RNA size marker (in kilobases, indicated on the left). Other lanes show signals obtained after hybridization of probes complementary to the 3′ end of ORF2 (sequence positions 1425 to 1444) (A), *cps* (positions 3013 to 3039) (B), ORF3 (positions 4161 to 4183) (C), *tsh* (positions 8716 to 8740) (D), ORF8 (positions 9118 to 9140) (E), and the region immediately downstream of ORF9 (positions 9947 to 9966) (F).

proximal to Ser-26, yielding the form found in the mature phage capsid. We speculate that this process is analogous to the situation in a number of other phages (reviewed in reference 7). Processing of gp23 in phage T4 (3, 25) results in removal of an N-terminal peptide and is crucial for maturation of the head. In T4, all but one prohead component are proteolytically cleaved. The ORF1 product from A511 shows convincing amino acid sequence homology to the N terminus of mature cps, including the processing site (Fig. 4). Therefore, it seems likely that gpORF1 is also processed between the corresponding Lys and Ser residues, which would result in a polypeptide of 94 aa with an M_r of 10,500.

Whether the observed sequence homology of the Cps 23-aa N-terminal peptide to the automodification domain of bovine ADP-ribosyltransferase is actually meaningful for function or in processing of this peptide remains to be investigated. Interestingly, enzymes of this type have been found in other phages (e.g., T4 and T7; see reference 18) and were determined to be involved in covalent modification of host RNA polymerases. The possible proteolytic activity of a specific Cps domain, implied by local sequence homology to the methionine-removing methionyl aminopeptidase from *E. coli* (2), also requires further study. This activity, however, would correspond to our observation that the N-terminal methionine is absent from the mature *tsh* product and seems to be generally removed from capsid and tail proteins in other *Listeria* phages (31).

The onset of *cps* and *tsh* transcription is approximately 15 to 20 min postinfection. This corresponds well to the timing of the appearance of the *ply* transcript, which encodes the A511 endolysin. We have recently constructed a derivative of A511 (A511::*luxAB*), which has a *Vibrio harveyi luxA-luxB* gene fusion (21) inserted directly downstream from the *cps* 3′ end (32). Luciferase expression can be first detected 15 min after infection, with a steady increase in light output until lysis (data not shown). This corresponds very well to the data presented herein (Fig. 5).

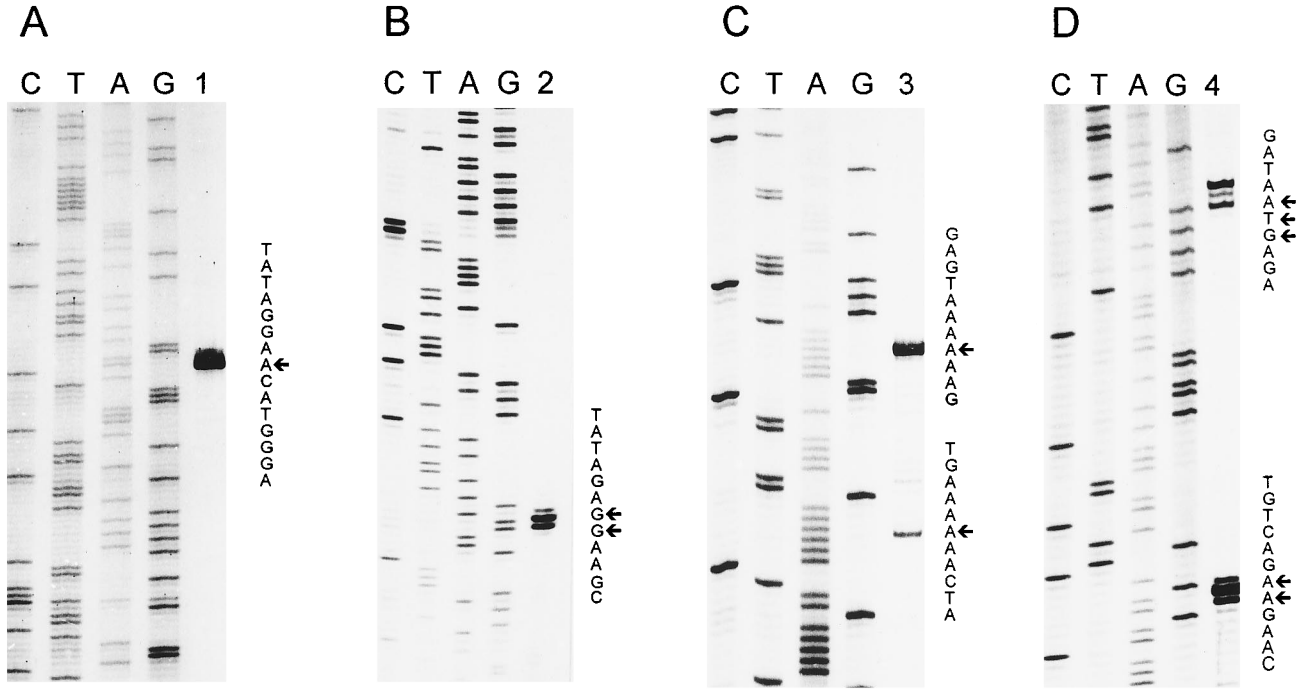

FIG. 7. Mapping of 5′ ends of four A511 late transcripts by primer extension analysis. Nucleotide sequencing ladders shown in lanes C, T, A, and G (complementary) were generated with the same primers as used for the primer extension reaction mixtures in lanes 1, 2, 3, and 4, respectively. (A) mRNA start site upstream of *cps* (sequence position 1610 [Fig. 2]); (B) transcription start site within the coding region of ORF6 (position 6050); (C) dual start sites upstream of ORF3 (positions 3544 and 3563); (D) dual start sites upstream of ORF9 (positions 9424 and 9464). The mRNA 5′ ends are indicated by arrows in the sequences shown alongside the extension lanes.

TABLE 2. Known promoter core sequences from *Listeria* spp. and comparison with A511 late gene promoters[a]

Organism	Gene	−35 Region	Spacer (nt)	−10 Region	Source or reference
L. monocytogenes	*cadA*	**TTGA**CT	17	**TATAAT**	27
L. monocytogenes	*tet*	**TTTA**CA	17	**TAGAAT**	8
L. monocytogenes	*sod*	**TTGA**AA	22	**TATACG**	5
L. monocytogenes	*hly*	**TTTA**GA	18	**TATAAT**	11
L. monocytogenes	*rep*	**TT**GCTT	17	**TAAAAT**	20
L. monocytogenes	*cat*	**TT**GTTT	20	**TATAAT**	20
L. monocytogenes	*fla*	**TTGA**TT	17	**TATAAT**	14
L. monocytogenes	*mlp*	AAA**A**GA	19	**TATAAT**	12
L. monocytogenes	*pic*	**TTAA**TG	20	**TAAGAT**	28
L. monocytogenes	*actA*	GTTAGA	22	**TA**TTCT	13
L. ivanovii	*liactA*	GTTAAA	22	**TA**TTCT	22
L. ivanovii	*plcA*	**TAAA**GA	19	**TAAGAT**	26
L. ivanovii	*prfA*	**TT**GCGA	17	**TAAAAT**	26
L. seeligeri	*kat*	**TTGA**AG	17	**TAAAAT**	19
A511	*cps*	**ATAGAGT**	20	TGC**TATATT**ATAG	This study
A511	*tsh*	**AGAGCCT**	22	TGC**TATATT**ATAG	This study
A511	ORF3-A	**AGAGCCT**	21	**TTTGCG**	This study
A511	ORF3-B	**AAAGAAT**	19	**TAAAAT**	This study
A511	ORF9-A	**AGAG**AAC	21	**TATATT**	This study
A511	ORF9-B	**AGAGAGT**	19	**TA**ACAT	This study

[a] Apparently conserved nucleotides (for each of the two groups) are printed in bold letters.

The most abundant structural proteins in A511 particles are Cps and Tsh. It is evident that transcription and/or translation of their genes is favored over that of the other genes in this region. Both transcripts are initiated from promoters with identical 13-nt sequences in the −10 region (Table 2), which differ from characterized host promoter sequences. The −10 sequences of the other identified A511 promoters are more or less different, which may help to explain the relatively lower level of expression of the downstream genes. The 13-nt sequence may be specifically recognized by an A511-encoded RNA polymerase, as in T7 (38). More likely, however, is the modification of host RNA polymerase by phage-encoded sigma-like factors, which regulate sequential gene expression by altering the recognition properties of the core enzyme. This topic is well studied for the virulent *Bacillus subtilis* phage SPO1 (39), which is morphologically similar to A511.

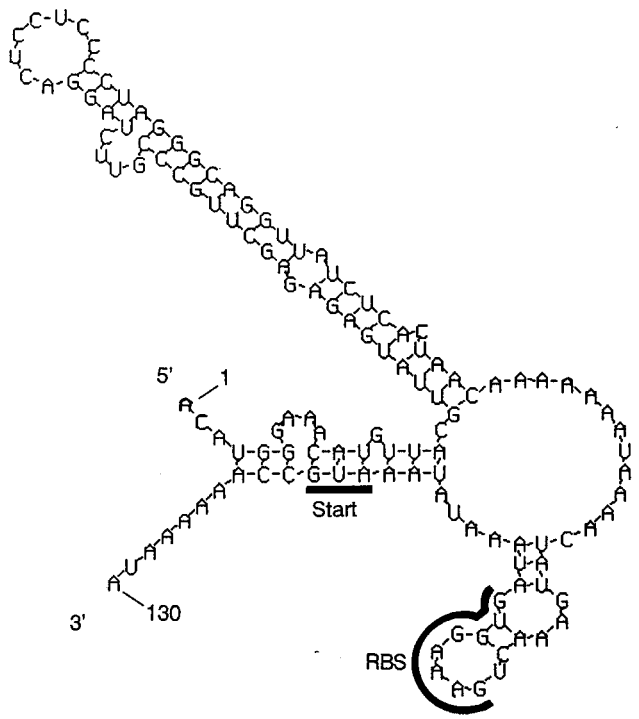

FIG. 8. Predicted secondary structure of the 5′ end (first 130 nt) of the *cps* mRNA transcript. Calculated free energy for this structure is −40.7 kcal [−170.3 kJ]/mol. The exposed RBS and the start codon are indicated.

Another important aspect of gene expression is stability of the transcripts. It is tempting to speculate that the predicted strong secondary structure of the major (1.6-kb) *cps* mRNA 5′ end allows enhanced mRNA stability by increased resistance to degradation (1, 23). Moreover, the predicted folding would expose the RBS, which may promote efficient recognition by the ribosomal subunits. It has been argued that loops in translation initiation regions can decrease efficiency of translation (9). However, the *cps* gene RBS seems to be exposed rather than sequestered within a structure.

No secondary structure was predicted for the major (2.7-kb) *tsh* transcript whose promoter is localized inside a putative coding region (ORF6), 1.1 kb upstream of the *tsh* gene. Approximately 90% of the ORF6 coding region and the entire ORF7 are included on this mRNA. Whereas there can be no expression of truncated ORF6 products (because of a lack of RBS upstream of possible starts), there should be full expression of gpORF7, a small and highly charged protein of unknown function. In terms of transcript stability, the long leader may serve to increase the half-life of the promoter-distal *tsh* gene.

Aside from being regulated by the above-mentioned factors, the relative rates of gene expression can also be influenced by individual efficiencies of RBS. The (complementary) 3′ end of *L. monocytogenes* 16S rRNA is 3′-TAGAAAGGAGGTGAT CCA-5′ (16). The *cps* gene RBS appears optimized for binding, since there are two possible (overlapping) perfect matches of 6 and 7 nt (GAAAGG and AGGTGAT). In contrast, the *tsh* gene RBS is less perfect (AAGGGAAGT). However, the stop codon of the upstream ORF7 is only one codon apart (same reading frame) from the *tsh* initiation ATG and overlaps the *tsh* RBS. This suggests translational coupling between these two genes, where translation is initiated via the ORF7 perfect

RBS (AGGAG), and therefore could obviate the need of *tsh* for a strong RBS.

In conclusion, there is good evidence for regulation of the strongly expressed A511 *cps* and *tsh* genes at the transcriptional level. Strong promoters and efficient translation are the mechanisms that provide the large relative amounts of these proteins necessary to build virions.

ACKNOWLEDGMENTS

We thank Patrick Schiwek for excellent technical assistance in DNA sequencing work, Anette Schneider for purifying PLY118 endolysin used in RNA work, and Ralf Einspanier and Ingo Krause for help with some of the computer analyses. We are grateful to Cath Rees (Nottingham, United Kingdom) for helpful discussions and critical reading of the manuscript.

REFERENCES

1. **Belasco, J. G., and C. T. Higgins.** 1988. Mechanisms of mRNA decay in bacteria: a perspective. Gene **72:**15–23.
2. **Ben-Bassat, A., K. Bauer, S. Y. Chang, K. Myambo, A. Boosman, and S. Chang.** 1987. Processing of the initiation methionine from proteins: properties of the *Escherichia coli* methionine aminopeptidase and its gene structure. J. Bacteriol. **169:**751–757.
3. **Black, L. W., M. K. Showe, and A. C. Steven.** 1994. Morphogenesis of the T4 head, p. 218–258. *In* J. D. Karam et al. (ed.), Molecular biology of bacteriophage T4. American Society for Microbiology, Washington, D.C.
4. **Botstein, D.** 1980. A theory of modular evolution for bacteriophages. Ann. N.Y. Acad. Sci. **354:**484–491.
5. **Brehm, K., A. Haas, W. Goebel, and J. Kreft.** 1992. A gene encoding a superoxide dismutase of the facultative intracellular bacterium *Listeria monocytogenes*. Gene **118:**121–125.
6. **Campbell, A.** 1988. Phage evolution and speciation, p. 1–14. *In* R. Calendar (ed.), The bacteriophages, vol. 1. Plenum Press, New York.
7. **Casjens, S., and R. Hendrix.** 1988. Control mechanisms in dsDNA bacteriophage assembly, p. 15–91. *In* R. Calendar (ed.), The bacteriophages, vol. 1. Plenum Press, New York.
8. **Charpentier, E., G. Gerbaud, and P. Courvalin.** 1993. Characterization of a new class of tetracycline-resistance gene *tet*(s) in *Listeria monocytogenes* BM4210. Gene **131:**27–34.
9. **de Smith, M. H., and J. van Duin.** 1990. Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. Proc. Natl. Acad. Sci. USA **87:**7668–7672.
10. **Doi, M., M. Wachi, F. Ishino, S. Tomioka, M. Ito, Y. Sakagami, A. Suzuki, and M. Matsuhashi.** 1988. Determinations of the DNA sequence of the *mreB* gene and of the gene products of the *mre* region that function in formation of the rod shape of *Escherichia coli* cells. J. Bacteriol. **170:**4619–4624.
11. **Domann, E., and T. Chakraborty.** 1989. Nucleotide sequence of the listeriolysin gene from a *Listeria monocytogenes* serotype 1/2a strain. Nucleic Acids Res. **17:**6406.
12. **Domann, E., M. Leimeister-Waechter, W. Goebel, and T. Chakraborty.** 1991. Molecular cloning, sequencing, and identification of a metalloprotease gene from *Listeria monocytogenes* that is species specific and physically linked to the listeriolysin gene. Infect. Immun. **59:**65–72.
13. **Domann, E., J. Wehland, M. Rohde, S. Pistor, M. Hartl, W. Goebel, M. Leimeister-Waechter, M. Wuenscher, and T. Chakraborty.** 1992. A novel bacterial virulence gene in *Listeria monocytogenes* required for host cell microfilament interaction with homology to the proline-rich region of vinculin. EMBO J. **11:**1981–1990.
14. **Dons, L., O. F. Rasmussen, and J. E. Olsen.** 1992. Cloning and characterization of a gene encoding flagellin of *Listeria monocytogenes*. Mol. Microbiol. **6:**2919–2929.
15. **Dunn, J. J., and F. W. Studier.** 1983. Complete nucleotide sequence of bacteriophage T7 DNA and the location of T7 genetic elements. J. Mol. Biol. **166:**477–535.
16. **Emond, E., I. Fliss, and S. Pandian.** 1993. A ribosomal DNA fragment of *Listeria monocytogenes* and its use as a genus-specific probe in an aqueous-phase hybridization assay. Appl. Environ. Microbiol. **59:**2690–2697.
17. **Farber, J. M., and P. I. Peterkin.** 1991. *Listeria monocytogenes*, a food-borne pathogen. Microbiol. Rev. **55:**476–511.
18. **Geiduschek, E. P., and G. A. Kassavetis.** 1988. Changes in RNA polymerase, p. 93–116. *In* R. Calendar (ed.), The bacteriophages, vol. 1. Plenum Press, New York.
19. **Haas, A., K. Brehm, J. Kreft, and W. Goebel.** 1991. Cloning, characterization, and expression in *Escherichia coli* of a gene encoding *Listeria seeligeri* catalase, a bacterial enzyme highly homologous to mammalian catalases. J. Bacteriol. **173:**5159–5167.
20. **Hadorn, K.** 1992. *L. monocytogenes* plasmid delta pWDB100 *cat* and *rep* genes for chloramphenicol acetyltransferase and replication initiator protein. EMBL database accession number X68412, release 40.
21. **Hill, P. J., S. Swift, and G. S. A. B. Stewart.** 1991. PCR based engineering of the *Vibrio harveyi lux* operon and the *Escherichia coli trp* operon provides for biochemically functional native and fused gene products. Mol. Gen. Genet. **226:**41–48.
22. **Kreft, J., M. Dumbsky, and S. Theiss.** 1995. The actin polymerization protein from *Listeria ivanovii* is a large repeat protein which shows only limited amino acid sequence homology to ActA from *Listeria monocytogenes*. FEMS Microbiol. Lett. **126:**113–122.
23. **Kubo, M., Y. Higo, and T. Imanaka.** 1990. Biological threshold values of procaryotic gene expression which is controlled by the DNA inverted repeat sequence and the mRNA secondary structure. J. Ferment. Bioeng. **69:**305–307.
24. **Kyhse-Andersen, J.** 1984. Electroblotting of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide to nitrocellulose. J. Biochem. Biophys. Methods **10:**203–209.
25. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
26. **Lampidis, R., R. Gross, W. Goebel, and J. Kreft.** 1994. The virulence regulator protein of *Listeria ivanovii* is highly homologous to PrfA from *Listeria monocytogenes* and both belong to the Crp-Fnr family of transcription regulators. Mol. Microbiol. **13:**141–151.
27. **Lebrun, M., A. Audurier, and P. Cossart.** 1994. Plasmid-borne cadmium resistance genes in *Listeria monocytogenes* are similar to *cadA* and *cadC* of *Staphylococcus aureus* and are induced by cadmium. J. Bacteriol. **176:**3040–3048.
28. **Leimeister-Waechter, M., E. Domann, and T. Chakraborty.** 1991. Detection of a gene encoding a phosphatidylinositol specific phospholipase C that is coordinately expressed with listeriolysin in *Listeria monocytogenes*. Mol. Microbiol. **5:**361–366.
29. **Loessner, M. J.** 1991. Improved procedure for bacteriophage typing of *Listeria* strains and evaluation of new phages. Appl. Environ. Microbiol. **57:**882–884.
30. **Loessner, M. J., and M. Busse.** 1990. Bacteriophage typing of *Listeria* species. Appl. Environ. Microbiol. **56:**1912–1918.
31. **Loessner, M. J., I. B. Krause, T. Henle, and S. Scherer.** 1994. Structural proteins and DNA characteristics of 14 *Listeria* typing bacteriophages. J. Gen. Virol. **75:**701–710.
32. **Loessner, M. J., C. E. D. Rees, G. S. A. B. Stewart, and S. Scherer.** Construction of luciferase reporter phage A511::*luxAB* for rapid and sensitive detection of viable *Listeria* cells. Submitted for publication.
33. **Loessner, M. J., and S. Scherer.** Unpublished data.
34. **Loessner, M. J., A. Schneider, and S. Scherer.** 1995. A new procedure for efficient recovery of DNA, RNA, and proteins from *Listeria* cells by rapid lysis with a recombinant bacteriophage endolysin. Appl. Environ. Microbiol. **61:**1150–1152.
35. **Loessner, M. J., G. Wendlinger, and S. Scherer.** 1995. Heterogeneous endolysins in *Listeria monocytogenes* bacteriophages: a new class of enzymes and evidence for conserved holin genes within the siphoviral lysis cassettes. Mol. Microbiol. **16:**1231–1241.
36. **Paces, V., C. Vlcek, and P. Urbanek.** 1986. Nucleotide sequence of the late region of *Bacillus subtilis* phage PZA, a close relative of phi-29. Gene **44:** 107–114.
37. **Saito, I., K. Hatakeyama, T. Kido, H. Ohkubo, S. Nakanishi, and K. Ueda.** 1990. Cloning of a full-length cDNA encoding bovine thymus poly(ADP-ribose) synthetase: evolutionary conserved segments and their potential functions. Gene **90:**249–254.
38. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
39. **Stewart, C.** 1988. Bacteriophage SPO1, p. 477–516. *In* R. Calendar (ed.), The bacteriophages, vol. 1. Plenum Press, New York.
40. **Tabor, S., and C. C. Richardson.** 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA **82:**1074–1078.
41. **Volker, T. A., J. Gafner, T. A. Bickle, and M. K. Showe.** 1982. Gene 67, a new essential bacteriophage T4 head gene, codes for a prehead core component, PIP. I. Genetic mapping and DNA sequence. J. Mol. Biol. **161:**479–489.
42. **Wendlinger, G., M. J. Loessner, and S. Scherer.** Bacteriophage receptors on *Listeria monocytogenes* cells are the N-acetyl-glucosamine and rhamnose substituents of teichoic acids or the peptidoglycan. Microbiology, in press.
43. **Zink, R., and M. J. Loessner.** 1992. Classification of virulent and temperate bacteriophages of *Listeria* spp. on the basis of morphology and protein analysis. Appl. Environ. Microbiol. **58:**296–302.