Transcriptional Mapping and Nucleotide Sequence of the Listeria monocytogenes hlyA Region Reveal Structural Features That May Be Involved in Regulation

JERÔME MENGAUD,¹ MARIA F. VICENTE,² AND PASCALE COSSART^{1*}

Unité de Génie Microbiologique, Institut Pasteur, Paris, France,¹ and Servicio de Microbiologia, Hopital Ramon Y Cajal, Madrid, Spain²

Received 5 May 1989/Accepted 17 August 1989

DNA sequence analysis of the regions adjacent to the *hlyA* gene, which encodes listeriolysin O, an essential virulence factor of *Listeria monocytogenes*, revealed the presence of two open reading frames (ORFs): ORF D located 304 base pairs downstream from *hlyA*, and ORF U located 224 base pairs upstream from and in opposite direction to *hlyA*. Promoter mapping performed with RNAs extracted from cells growing exponentially in rich medium showed that the three ORFs are independently transcribed. *hlyA* is transcribed from two promoters separated by 10 base pairs (P1 *hlyA* and P2 *hlyA*). ORF U is transcribed in the opposite direction from an adjacent promoter. These two promoter regions are separated by a palindromic sequence T-T-A-A-C-A-A/T-A/T-T-G-T-T-A-A. This palindrome was also found upstream from the ORF D promoter, suggesting that all three genes are similarly regulated.

Listeria monocytogenes is a gram-positive bacterium responsible for severe infections (meningitis, septicemias) and abortion in humans and several animal species (34). It is one of the best models to study the virulence of facultative intracellular pathogens. L. monocytogenes survives and replicates in macrophages (23), and control of the infection requires the emergence of a T-cell-mediated response leading to macrophage activation. For this reason, Listeria infection in mice has been widely used to study induction of T-cell-mediated immunity (21).

Taxonomic studies established that the genus Listeria includes several species, L. monocytogenes, L. ivanovii, L. seeligeri, L. welshimeri, L. innocua, L. murrayi, and L. grayi (34), of which only the first two are pathogenic, L. ivanovii essentially affecting animal species. The low G+C content and 16S RNA sequence analysis (22) indicated that this genus is closely related to the genera Bacillus, Erysipelothrix, Lactobacillus, and Streptococcus (34).

Due to the good animal model provided by the experimental murine infection, genetic studies were initiated to identify the virulence factors of this intracellular pathogen (17, 20, 30). These studies took advantage of conjugative transposons initially isolated in streptococci, such as Tn1545 (10) and Tn916 (15), which were used to mutagenize genes onto the bacterial chromosome. Recently, we cloned the hlyA gene coding for listeriolysin O and showed its essential role in intracellular survival and therefore in virulence of L. monocytogenes by gene complementation of a nonhemolytic Tn917-induced mutant (9a, 25, 26, 41). The current hypothesis for the role of listeriolysin O is that this secreted thiol-activated protein would help the bacterium to escape from the phagosome compartment to the cytosol, where it can freely replicate. This hypothesis is strongly supported by the fact that, whereas both the wild type and nonhemolytic mutants are able to invade Caco2 cells, the latter are unable to escape from the phagosome (16) and to survive intracellularly.

Two observations indicate that listeriolysin O expression

must be regulated and not constitutively expressed: (i) it was reported that there is an inverse correlation between iron concentration in the medium and the hemolytic activity in the culture supernatants (11) and (ii) no increase in hemolytic titer was observed when a multicopy plasmid carrying only the hlyA gene was introduced in a wild-type strain (9a). To understand how hlyA is transcribed and regulated, we determined the nucleotide sequence of the 4.2-kilobase region containing the 1.5-kilobase hlyA gene. This sequence revealed the presence of two open reading frames (ORFs), one (ORF D) located downstream from hlyA and in the same orientation and another (ORF U) oriented in the opposite direction and located upstream from hlyA. To precisely describe the transcriptional organization of this region, we also mapped the promoters of hlyA and those of ORF U and ORF D. The three genes are transcribed independently. hlyA appears as a monocistronic unit. ORF U and the hlyA promoter regions are separated by a perfect 14-base-pair (bp) palindromic element. The fact that this same palindromic sequence is found upstream from ORF D suggests that expression of the three genes is similarly regulated.

MATERIALS AND METHODS

Strains, plasmids, and culture media. The L. monocytogenes strain used in this study was LO28, a wild-type isolate (26). Plasmid pUC18 (43) and phage M13mp18 (27) were used to clone DNA fragments in Escherichia coli MC1061 $[F^- araD139 \Delta(ara \ leu)7696 \Delta lacY74 \ galU \ galK \ hsr \ hsm$ strA] (9) or TG1 [K-12, Δ(lac-pro) supE thi hsdD5 F' traD36 proAB lacI^q lacZ Δ M15] (T. Gibson, Medical Research Council, Cambridge, England), respectively. E. coli strains were grown at 37°C in LB medium. For strains containing pUC derivatives, ampicillin was added at a final concentration of 25 µg/ml in liquid and 100 µg/ml in solid medium. Strain TG1 was routinely plated on minimal medium containing 0.2% (wt/vol) glucose and grown at 37°C. Isolated colonies were used to inoculate 2×YT liquid medium (tryptone [16 g/liter; Difco Laboratories, Detroit, Mich.], yeast extract [10 g/liter; Difco], sodium chloride [5 g/liter]) for overnight cultures, which were subsequently diluted for

^{*} Corresponding author.

infection with M13 derivatives or for transformation. L. monocytogenes strains were grown in brain heart infusion broth or agar (Difco) at 37° C.

Chemicals and enzymes. Restriction enzymes and ligase were purchased from Amersham Corp. (Les Ulis, France), Boehringer (Mannheim, Federal Republic of Germany), or Genofit SA (Geneva, Switzerland) and were used as recommended by the manufacturers. [³⁵S]dATP (800 Ci/mmol) and [³²P]ATP (3,000 Ci/mmol) were purchased from Amersham.

DNA techniques. Plasmid DNA was purified by ultracentrifugation in cesium chloride gradients (24). Rapid preparation of plasmid DNA was performed by the method of Birnboim and Doly (7). DNA fragments were purified (26) and recombinant DNA techniques were performed (24) as previously described. Hybridizations were performed on Hybond-N nylon membranes (Amersham). DNA probes were labeled by the multiprime labeling system (kit RPN.160 1Y; Amersham). Oligonucleotide primers, synthesized by using the phosphoramidite method (3) with a 380D DNA synthesizer (Applied Biosystems, Foster City, Calif.) were provided by the Unité de Chimie Organique, Institut Pasteur. They were 5' labeled with $[\gamma^{-32}P]ATP$ (24) and purified on NACS Prepac Cartridge 1526 NP (Gibco-BRL, Cergy-Pontoise, France). The DNA sequence was determined by the Sanger et al. technique (33) with the Sequenase kit (U.S. Biochemicals, Cleveland, Ohio) and oligonucleotide primers derived from the sequence. Plasmid DNA sequencing was performed as follows. Plasmid DNA (1 to 2 µg) was incubated for 5 min at room temperature in 20 µl of denaturation solution (0.2 M NaOH, 0.2 mM EDTA), ethanol precipitated, and directly sequenced as an M13 derivative.

RNA isolation. Total cellular RNAs were extracted from *Listeria* cultures in exponential growth ($A_{600} = 0.7$) as described previously (19), except that double-strength Kirby mixture was used. In the case of *L. monocytogenes*, yields were very low (50 µg/500 ml).

Promoter mapping. Both primer extension with reverse transcriptase mapping (6, 18) and S1 nuclease mapping (1) were used. For primer extension, total cellular RNAs (10 µg) were hybridized with 5'-labeled oligonucleotide primer (0.05)pmol) in 5μ l of H₂O by boiling the reaction mixture for 5 min, followed by slow cooling to 42°C. Then 1 µl of 0.5 mM dATP, dGTP, dCTP, and dTTP, 0.5 µl of reaction buffer (50 mM Tris hydrochloride [pH 8.3], 500 mM potassium chloride, 100 mM magnesium chloride, 10 mM dithiothreitol), and 9 U of avian myeloblastosis virus reverse transcriptase (Boehringer) were added to the hybridization mixture, which was then incubated for 30 min at 42°C and inactivated at 75°C for 10 min. In some controls, 0.5 µl of RNase solution (10 mg/ml) was added, and the sample was incubated for 1 h at 37°C. It was then twofold diluted in sequence loading buffer (24), denatured for 3 min at 90°C, and loaded on a 6% polyacrylamide-7.6 M urea sequencing gel.

For S1 mapping, single-stranded ³²P-labeled DNA probes were prepared by the modified prime-cut method (5, 13) with the same oligonucleotide primer as in the reverse transcription reaction. Before S1 protection studies, RNA samples (10 μ g) were coprecipitated with ³²P-labeled probes, suspended in 30 μ l of formamide containing 50% (vol/vol) 50 mM piperazine-*N*,*N'*-bis(2-ethanesulfonic acid) (Sigma, La Verpilliere, France)–480 mM sodium chloride–2.5 mM EDTA, denatured by heating at 75°C for 5 min, and incubated overnight at 42°C. Samples were divided in three parts, diluted 30-fold with S1 buffer (30 mM sodium acetate, 250 mM sodium chloride, 1 mM zinc sulfate [pH 4.5]). Unprotected probes were digested for 15 min at 37°C with 0, 2, and 100 U of S1 in each third of the sample, respectively. After ethanol precipitation, protected fragments were processed and analyzed on the same calibrated sequencing gel as for reverse transcriptase mapping samples. Sequence reactions obtained with the recombinant M13 single-stranded DNA and the oligonucleotide primer used in the probe preparation were loaded on the same gel and used to identify the transcription starting point. Samples were electrophoresed at 50 W for 3 h to determine the size of the primer extension and S1 digestion reaction products. Bands were examined by autoradiography.

Northern blotting. Formaldehyde agarose gel electrophoresis was performed as described previously (24), except that the final running buffer was 40 mM morpholinepropanesulfonic acid-10 mM sodium acetate-1 mM EDTA (pH 7.0) and that ethidium bromide was added to the gel at a final concentration of 0.6 µg/ml, RNA was blotted onto Hybond-N membranes in $20 \times$ SSC buffer (1× SSC is 0.15 M sodium chloride-0.015 M sodium citrate) as described previously (39). Prehybridization (3 h) and hybridization (16 h) with oligonucleotide primers were performed in $6 \times$ SSC-5× Denhardt solution-0.1% (wt/vol) sodium dodecyl sulfate buffer containing 100 µg of sonicated carrier DNA per ml in sealed plastic bags at $T_d - 5^{\circ}$ C, calculated according to the Wallace rule (38). Filters were then washed in 6× SSC-0.1% (wt/vol) sodium dodecyl sulfate-1× Denhardt buffer for 30 min at $T_d - 5^{\circ}$ C and exposed to film at -80° C with an intensifying screen.

RESULTS AND DISCUSSION

Cloning of overlapping fragments of the hlyA gene region in pUC18. A 3,454-bp BamHI chromosomal fragment containing hlyA had been cloned from L. monocytogenes LO28 in pUC18, giving rise to pLis3. Using as a probe an oligonucleotide primer (21-mer) internal to the hlyA gene (positions 1605 to 1625 in Fig. 1), we cloned a 1,886-bp EcoRI fragment containing part of hlyA and its upstream region, giving rise to plasmid pLis8 (see Fig. 3). Inserts of pLis3 and pLis8 spanned 4,223 bp of the hlyA gene region, i.e., 1,488 bp upstream and 1,145 bp downstream from hlyA.

Nucleotide sequence analysis, ORF assignments, and protein sequence analysis. The sequence of the 4,223-bp chromosomal region was determined on both strands of the DNA. The sequence of *hlyA* had been reported earlier (26). The presence of putative ORFs upstream and downstream from *hlyA* was screened for by use of a program based on codon usage (37), taking *hlyA* as the reference gene. Two major ORFs, ORF U and ORF D, having good probabilities of being coding sequences, were detected. ORF D, in the same orientation as *hlyA*, started 304 bp downstream from *hlyA* and was not interrupted by a stop codon in the insert of pLis3. ORF U, in an opposite orientation, began 224 bp upstream from the *hlyA* translation initiation codon and was 969 bp long (Fig. 1).

In the case of ORF U, a putative ribosome-binding site (AAAGGAGG) was detected 3 bp downstream from the beginning of the ORF (position 1265 in Fig. 1) and 9 bp upstream from a TTG codon. This codon is rarely used in *E. coli* as a translation initiation codon (14) and may not be used in *L. monocytogenes*. In addition, the region between this start codon and the putative Shine-Dalgarno sequence is rich in Gs and Cs, an unusual feature for translation initiation regions (14). The first ATG present in the ORF has no ribosome-binding site, whereas a second ATG, located 225 bp downstream from the beginning of the ORF, is preceded

nds to the region span ence. The deduced ar) amino acid sequence uctures in untranslate	FIG. 1. Nucleotide sequence of the <i>hlyA</i> region. The sequence presented here corresponds to the region span ORF D will be published elsewhere). Nucleotide coordinates are indicated below the sequence. The deduced an oordinates are given in the margin. Possible cleavage sites for signal peptidase in the ORF D amino acid sequence <i>i</i> , <i>Hin</i> dIII; B, <i>Bam</i> HI). The transcriptional start points (+1) are indicated. Palindromic structures in untranslate arminators are underlined (see the text).	FIG. 1.] DRF D will oordinates I, <i>Hin</i> dIII;
S V G E E TCTGTCGGGGAAGAA Agacagcccttctt 350 . 350	AMANTROCCATEGRAATERIGAGAITERICTATINICATAGAATERACERITICTAGIAGAATERIACERI TTITAACCEGTACCAITACAACTAAAAACGAATATATACTACAACAACGAATAACGAACG	50
M K S Ggtgaaatgaaaagta Ccactitactitica	ATTGAGTAATOGTITCTAATACACCTGAAAG TAACTCALTAAGCAAAGALTAAGTGGGACTITC Q T I T E L V G S L	77
TACTAAAAACGTGTC ATGATTTTTTGCACAC	TAMAMATAATCTITATAAAATGATAGATAGATGATAGATGATAAACTAACGATGATGATGATGATGATGATAAA ATTTITATAGAATATTACAACTAACCAACCTAAGCTATAGATAGCAACGAACAAAAGGAGCAGAAAATTT Y F Y D K Y I N I L F Q I R Y D F S D N S O E D K L 	104
ТАТССАТАТААТАТАТ АТАССТАТАТААТАТАТАТ	TICICIGAAAAGIAAIAATATITICCCCCGAACAACITITAATGAACGAITTAATGCCCGGTCAAGICAAGI	131
attitgtgaagaatg7 Taaaacacttcttac2	TANGTRAAATTAGTAGTGGGGGGGGGGGGGGGGGGGGGGG	158
gttaaaatttaga caatttttacatct1 tigtaaagtaataat aacatttcaattaatta	AGTTIATIGTCCGCTTIGGAAGCTIGAATAAGCAGTCTGGACAATCTTIGAATITIGTTICACACTCGGACCATTGTAG TCAAATAACAGCGAAAACCTACGAAACTATAGAGAAACTTAAAACAAAACTGTGAACCTGGTAACATC L K N D A K S A Q Y A T Q V I E K F K T K V S P G N Y 500 ~ H	185
CGTITCATCTITAGA GCAAAGTAGAAATCT	TGCTCTACTTIGTIGTTTAATGCTGCAGCATACTGAACGAGGTGTGAATGGTTAATGGAGGGCTAATATGGTTAAGAAAA Acgagangaacaacaacaatataacgacgtcgtaagacgctcacacgttacaattaacttcaccgcgattataccaattcttt Q E V K N N L A A A Y Q R P T F T L S T A S I H N L F	212
TGTTGAGGCATTAACI Acaactocgtaattgi 1300 .	TTAATTGTTTGTTTTTCGGGGAGGTCGATGATAGTAGATAGGAAGTAGGACTTGAACTTAGTTAG	239
ACTAATGTGCGTTGCJ TGATTACACGCAACGT	CCAAGGTIGCICGGAGATATACTAACCTITITIGTAGGAATAATAATAGTIAGTIGAAUTTIATIGTITITIATAGATGATT GGCICCAACGAGCCICTATATGATIGGGAAAAAAACAICCTTAITATATATACAACCAACGAGCCICTATATTAGATGGAAAAAAACAICCTTAITATATATACAACGATAATAACAAAAAAAAAA	248
G <u>ecttatt</u> ocagttat CCGAATAAGGTCAAT	ttgattogcamagaatagatartatataangttattattagatgttamattgamatagagtagtatggtt aactaagggtttrettingstatgatamattttggcetamatnaatctacaattataccaatgatastataccaa 200	
CCTGGTATAGAGAGCC GGACCATATCTCTCGC	gcastalcalatistiactostisaaasaansi'lacgastataistittitacatatassa tittataslaati Cstcatististiataaatcsacaactitcitticasiscsattiactacaaaaaatstatcctaaaataatattaa 100	
AATGACATOGTTTGTC TTACTGTAGCAAACAC 23 L S M T Q J	Ganitettetgettgagestichtstentessecantestitttintessectitttinaatasestittittaaleksittitetaaleksittitetaatas Cittagaagasessaatasessaatasessettaseaaatasessaaaaaatattittetaatate ^2	

S V G E E K L Q N N T G A K K T P A D L K A Tetgeogggamgamatittamatanggcamamagaccectattimangeti Agacagccectititittgagtittattatgigitogtititegggacgactamititega
M K S K L I C I I M V I A F Q A H F T M T V K A D Gotgamatgamagtamactiatifictaticategataticagetatificagetatiticactatigecagt Ccactitactiticatigamamaaangtagtagtagtagtaggagaggagagagagagagag
tactamamacstotocttmactotototototaattastto <u>tactssacttaakott</u> astittassaattamamagasc Atgattitttocaaassaattgasaasaacastotmatcmacatocaaatttgaatcamatgataattittoctos 3400
tatocatatatatatatatocca <u>agtittagccacta</u> tactitctactgcaaaacgtatattitagticcacata ataggtatatatatatocgataggg <u>tttcaartocgtggat</u> atciaagatgicgatattcaartataaatcaaggggat ^+1 3300
аттттотсилсилтеласыласыссттаттттталттттттасылсылтталсалатсталасылатстсастст Талалсастосттасалсттотоссбаагалалалагталалалалтсттотталтосттасатттосттатасыстсаса 3200
ттеглаластителькалалтельскала состателеское сталаститеся сторы состоятся состоят соста соста соста соста соста Ласатите соста сталитите соста со З 100 °E
K GTTAMAMTETAGAAGAGAGAGAGAGAAAAOOCATG/hlya/TM CMTTTTTACATCTTCCTCTACTTGGGTAC/////-///TT
CGTTTCATCTTAGAACCGANTATCGC <u>CATATTATAA</u> TATCAAAAGAGAGGGGGGGCAAACCGTATTIGCCATAATAA GCAAAGTAGAAATCTTCGCTTAAAGCGGTTATAATAATAATAATATCTCTCTC
tottgagcattaacattgttgttaacgacgataaagggacagcaggactagaataaagcaagc
actaatgegettgemataattettataemantggecceeteettgattagtatatteetatettaagtgeettta Tgattaclegemacgttattmgaatatgtttaccgggggggggmetmatcatatmggatagattteretgamaat ^+1
G <u>gettati</u> ccagit <u>atiaacgaata</u> tgettitccgecetaatgggaaagtaaaaagtataaaacgagtatiaa Ccgaataaggecaataattogeettatacgaaaaaggggattaceettteattteatattatatt
cctgetatagagacgactectagettistetististgencagetagacgegacatocattetittetastacagetect <u>tistt</u> ggaccatatectctcgcgacgatectaaccagetccatectctcgcctgtagetaacaaaacaa
TTACTOTAGCAAAACACAAACTGAATCACCAATTACAGGGAATAACGAGAATAACAGAGTAGCAATAAGAGGAATAATCGAGTAGCAATAAGAAGAACAATA 23 L S M T Q T Q A L P K T L T W T I D G N Y S M
Anteacancentrenettreagettreattreettreettangetcancentatectccettataectcatcetatcatetta

ORF D will be published elsewhere). Nuc coordinates are given in the margin. Possib H, *Hin*dIII; B, *Bam*HI). The transcription terminators are underlined (see the text). amining $\bigcup r$, nyA, and the 5' end of ORF D (the complete sequence of amino acid sequences of ORF U, hyA, and ORF D are presented. Their ce are after Ala-24 or Gly-28 (42). Restriction sites are indicated (E, *Eco*RI; ted leader sequences are indicated by arrows, and putative transcriptional

\$

25

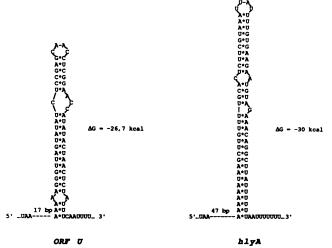


FIG. 2. Secondary structures of putative terminators for ORF U and ORF D. ΔGs were calculated as described previously (8, 40).

by a putative ribosome-binding site, AGG. If this second ATG is considered to be the translation initiation codon, the protein encoded by ORF U would be 248 amino acids long. The amino acid composition deduced from the sequence of

ORF U is not unusual, and the amino acid sequence does not present any characteristics of a secreted protein. Its hydropathicity pattern does not reveal either long hydrophobic or hydrophilic regions or alternance of such regions as in transmembrane proteins. However, interestingly, the most hydrophobic region contains a stretch of 18 amino acids, from amino acid position 49 to amino acid position 68, which has homology with the helix-turn-helix motif of one class of DNA-binding protein (28): Ile - Phe-Leu - Asn - Ala - Ser - Leu-Ser-Gly-Val-Leu-Glu-Thr-Ile-Thr-Gln-Phe-Leu-Lys-Lys (Fig. 1). This indicates that the protein could be a DNAbinding protein.

In the case of ORF D, three ATG codons are located 27, 54 and 84 bp downstream from the beginning of the ORF (position 3382 in Fig. 1). Only the first is preceded, 8 bp upstream, by an heptanucleotide (AAAGGAG) that is complementary to the 3' end of 16S rRNA (22) and is considered to be a ribosome-binding site. This ATG is probably the initiation codon of ORF D. In addition, the beginning of the protein sequence deduced from the ORF D sequence has all the characteristics of signal sequences of gram-positive bacteria (44), indicating that ORF D probably encodes a secreted protein. This may be relevant for virulence, since ORF D is unique to *L. monocytogenes*: an *L. monocytogenes* ORF D probe did not hybridize with *L. ivanovii* and *L. seeligeri*. This was not the case for *hlyA*, which can be detected in *L. ivanovii* and *L. seeligeri*, or for ORF U, which

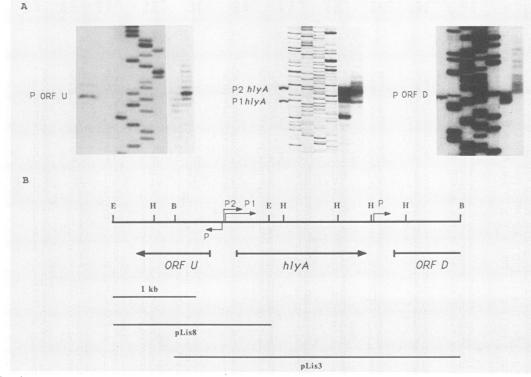


FIG. 3. Mapping of hlyA, ORF U, and ORF D transcripts by primer extension and S1 mapping analysis. (A) ³²P end-labeled oligonucleotide primers corresponding to positions 1226 to 1245, 1605 to 1625, and 3435 to 3454 in Fig. 1, respectively, were used for reverse transcription and S1 mapping analysis of ORF U, hlyA, and ORF D transcripts as described in Materials and Methods. The DNA products were separated by electrophoresis on an 6% polyacrylamide gel simultaneously with the Sanger dideoxy-chain termination ladder produced with the identical primer and the sense or antisense strand of the pLis3 insert cloned into M13mp19 as the template. Lanes (from left to right): reverse transcription reactions, G, A, T, and C tracks of sequencing reactions, S1 mapping reactions obtained with 100 and 2 U of S1 nuclease. (B) Schematic representation of the transcriptional organization of the hlyA region. Under the line representing the region, the three ORFs identified by the sequence have been indicated as well as the inserts of pLis8 and pLis3. Restriction sites are indicated (E, *Eco*RI; H, *Hin*dIII; B, *Bam*HI).

is also detected in *L. ivanovii* at low stringency (Gormley et al., manuscript in preparation).

Palindromic sequences of 69 and 54 bp were found downstream from hlyA and ORF U stop codons, respectively (Fig. 2). These structures have calculated ΔGs of -30 and -26.7kcal (ca. 125.5 and 111.7 kJ), respectively. They are indicative of putative transcription termination signals (29).

Promoter mapping and transcriptional organization of the region. Promoters of hlyA, ORF U, and ORF D were mapped by both S1 mapping and primer extension analysis with RNAs extracted from cells growing exponentially in rich medium. Both techniques gave identical results; the reverse transcription allowed a more precise determination of transcriptional start sites (Fig. 3). Two transcriptional start sites for hlyA were reproducibly identified at positions 1367 (P1 transcript) and 1356 or 1357 (P2 transcript) (Fig. 1). These transcriptional start sites are separated by 10 bp and located 122 and 133 bp, respectively, from the translation initiation codon of hlyA. The P1 transcript starts with an A, and the P2 transcript starts with an A or a T. Both transcripts detected were of similar intensity; the longer one (originating at P2) was slightly more abundant than the shorter one (Fig. 3).

Transcriptional start sites for the two putative ORFs were identified, implying that not only hlyA but also ORF U and ORF D were transcribed independently in L. monocytogenes. The transcriptional start site of ORF U was localized at A or T in position 1272 or 1273 (Fig. 1). This position is 232 bp from the presumed translation initiation codon of ORF U. Thus the two divergent ORFs, hlyA and ORF U, have their transcriptional start sites separated by 84 bp when considering P2 of hlyA or 95 bp in the case of P1 of hlyA. The -35position of the ORF U promoter is thus 14 bp from the -35position of P2 hlyA. The characteristic features of this intergenic region will be described below. The transcriptional start site of ORF D was localized at an A in position 3259, 181 bp from the translation termination codon of hlyA and 65 bp downstream from the putative terminator of hlyA. The beginning of this transcript is located 150 bp upstream from the putative translation initiation codon of ORF D. No longer transcript was detected for ORF D, which might have started at either one of the two promoters of hlyA or elsewhere.

Northern blot analysis allowed sizing of the hlyA transcripts. When total RNAs extracted in the same conditions as those used for the promoter mapping experiments were separated by electrophoresis, transferred onto nylon membranes, and hybridized with the oligonucleotide primer (positions 1605 to 1625 in Fig. 1) used for the primer extension analysis of the hlyA transcripts, one single band was detected that corresponded to transcripts of about 1,800 nucleotides (Fig. 4). This is exactly the size calculated for RNAs originating at the identified hlyA promoters and ending at the putative terminators. Thus, hlyA appears to be a monocistronic unit, in agreement with our previous complementation studies: when a plasmid carrying only hlyA was introduced in a transposon-induced nonhemolytic avirulent mutant, both the hemolytic phenotype and the virulence were restored, demonstrating that the transposon insertion had no polar effect on adjacent genes necessary for virulence (9a). As a corollary, it appears that terminators in L. monocytogenes would have the same structural features as E. coli rho-independent terminators (29). A schematic drawing of the transcriptional organization of the region is presented in Fig. 3B and 5B.

In regions of divergent transcription with back-to-back

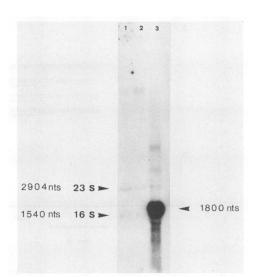


FIG. 4. Northern blot analysis of *hlyA* transcripts. Lanes: 1 and 2, *E. coli* total RNAs; 3, *L. monocytogenes* LO28 total RNAs. nts, Nucleotides.

promoters in procaryotes and their viruses (reviewed in reference 4), the distances between the transcription initiation sites range from 75 bp (X-rpsU of E. coli) to 513 bp (malT-malP of E. coli), and the distances between the translational start sites range from 103 bp (cI-cro of bacteriophage lambda) to 611 bp (malT-malP of E. coli). In the case described here, these two distances are 84 and 449 bp, respectively, and therefore are not exceptional. In contrast, the untranslated leader regions were rather long, 122 and 133 bp in the case of hlyA and 232 bp in the case of ORF U. This was also the case for the ORF D leader, which is 150 bp long. The potential secondary structures detected in these leaders (Fig. 1) are an indication of putative regulations in those regions. Since divergent transcription units often code for a regulatory protein (R) and a nonregulatory polypeptide (S) (4), the homology of the protein encoded by ORF U with some DNA-binding proteins suggests that the ORF U-hlyA region could be of the type R-S. ORF U would play a regulatory role, controlling either the expression of hlyA and of ORF U itself or that of other loci.

Comparison of the promoter sequences for hlyA, ORF U, and ORF D. Four promoter sequences were identified: two for hlyA, one for ORF U, and one for ORF D. The four promoters have -10 regions related to the *E*. *coli* consensus sequence, but only one of them (P1 hlyA) has a -35 region similar to the E. coli consensus sequence (32). The three others lack a -35 region, a feature characteristic of positively regulated promoters (31). This would suggest that hlyA has one constitutive promoter and a second promoter that requires an activator. By comparing the sequences of P1 hlyA, P ORF U, and P ORF D, we identified a palindromic sequence of 14 bp that is conserved in the three promoters and located just upstream from the -35 region (Fig. 5). Moreover, in the intergenic sequence ORF U-hlyA, the palindrome T-T-A-A-C-A-A/T-A/T-T-G-T-T-A-A is shared by the two back-to-back promoters. This palindromic structure might be recognized by regulatory DNA-binding proteins. Since P2 hlyA and P ORF U lack an E. coli consensus -35 region, an activator is possibly involved (31). Moreover, the homologies between the -40 regions of P2 hlyA and P ORF U promoters and that of ORF D suggest that hlyA, ORF U, and ORF D could have the same activator. How-

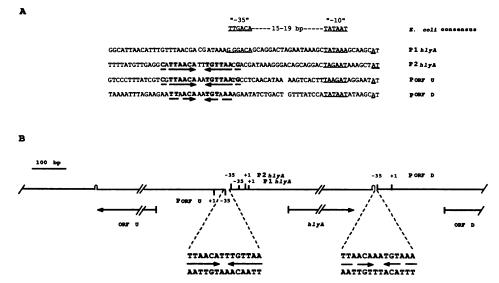


FIG. 5. (A) Promoter comparison. (B) Schematic drawing of the hlyA region.

ever, since no other promoter has been mapped and sequenced in *L. monocytogenes*, the possibility cannot be excluded that the palindrome is a characteristic of all *Listeria* promoter regions.

The hypothesis of a requirement for a positive regulator for hlyA expression would explain the absence of an increase in the hemolytic titer when a multicopy plasmid carrying hlyA was introduced in a wild-type L. monocytogenes. The simplest explanation for such an observation is that the activator concentration would be a limiting factor. Nevertheless, the regulation of hlyA is probably more complex than a simple activation. It has been reported that iron down-regulates the expression of listeriolysin O (11). Such a regulation could be mediated by a repressor as in E. coli, where iron-regulated promoters are repressed by the fur gene product in the presence of iron (2). Fur has also been shown to regulate expression of the diphtheria toxin gene tox cloned in E. coli and a Fur-binding site was identified in the tox promoter region (35). In the case presented here, no site homologous to the Fur consensus binding site was detected in the promoter regions.

It has been recently reported that listeriolysin O expression is under heat shock control (36). The promoters identified here do not share homologies with E. *coli* heat shock promoters (12), and there is no evidence that the 14-bp palindromic sequence identified in this study has any relationship to a heat shock response.

In conclusion, to understand the regulation of the expression of an essential virulence factor in L. monocytogenes, we have described the structural organization of the region encoding listeriolysin O. hlyA appears as a monocistronic unit that can be transcribed from two promoters spaced by 10 bp. hlyA is adjacent to another gene, ORF U, transcribed in opposite direction. These two transcription units are separated by the perfect palindrome T-T-A-A-C-A-A/T-A/ T-T-G-T-T-A-A. Since this palindrome is also found in the promoter region of ORF D, a species specific gene located downstream from hlyA, it can be suggested that the three genes are under a similar regulation.

ACKNOWLEDGMENTS

We thank Fernando Baquero and Julian Davies for constant support in this work. We are grateful to S. Cole, T. Garnier, A. Raibaud, and C. J. Thompson for advice in the experiments. We thank E. Gormley, K. Reich, and C. Thompson for helpful criticisms in the manuscript preparation and I. Sauvaget for computer assistance.

This work was supported by a contrat de jumelage from the CEE (STJ2-0319-C), an action intégrée franco espagnole (programme 1988) from the Institut National de la Santé et de la Recherche Médicale and the Institut Pasteur. J.M. was the recipient of a short-term Académie d'Agriculture de France fellowship.

LITERATURE CITED

- Aiba, H., S. Adhya, and B. De Crombrugghe. 1981. Evidence for two functional gal promoters in intact Escherichia coli cells. J. Biol. Chem. 256:11905–11910.
- Bagg, A., and J. B. Neilands. 1987. Ferric uptake regulation protein acts as a repressor, employing iron (II) as a cofactor to bind to the operator of an iron transport operon in *Escherichia coli*. Biochemistry 26:5471-5477.
- 3. Beaucage, S. L., and M. H. Caruthers. 1981. Deoxynucleoside phosphoramidites—a new class of key intermediates for deoxy-polynucleotides synthesis. Tetrahedron Lett. 22:1859–1862.
- Beck, C. F., and R. A. J. Warren. 1988. Divergent promoters, a common form of gene organization. Microbiol. Rev. 52:318– 326.
- Biggin, M. D., P. J. Farrell, and B. G. Barrell. 1984. Transcription and DNA sequence of the *Bam*HI L fragment of P95-8 Epstein-Barr virus. EMBO J. 3:1083–1090.
- Bina-Stein, N., M. Thoren, N. Salzman, and J. Thomson. 1979. Rapid sequence determination of late SV40 16S mRNA leader by using inhibitors of reverse transcriptase. Proc. Natl. Acad. Sci. USA 76:731-735.
- Birnboim, H. C., and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7:1513–1523.
- Borer, P. N., B. Dengler, I. Tinoco, Jr., and O. C. Uhlenbeck. 1974. Stability of ribonucleic acid double-stranded helices. J. Mol. Biol. 86:843–853.
- Casadaban, M. J., and S. N. Cohen. 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J. Mol. Biol. 138:179–207.
- 9a.Cossart, P., M. F. Vicente, J. Mengaud, F. Baquero, J. C. Perez-Diaz, and P. Berche. 1989. Listeriolysin O is essential for virulence of *Listeria monocytogenes*: direct evidence obtained by gene complementation. Infect. Immun. 57:3629–3636.
- Courvalin, P., and C. Carlier. 1986. Transposable multiple antibiotic resistance in *Streptococcus pneumoniae*. Mol. Gen. Genet. 205:291-297.
- 11. Cowart, R. E. 1987. Iron regulation of growth and haemolysin

production by Listeria monocytogenes. Ann. Inst. Pasteur Microbiol. 138:246–249.

- Cowing, D. W., J. C. A. Bardwell, E. A. Craig, C. Woolford, R. W. Hendrix, and C. A. Gross. 1985. Consensus sequence for *Escherichia coli* heat shock gene promoters. Proc. Natl. Acad. Sci. USA 82:2679-2683.
- Danos, O. 1985. Fine structure of cotton tail rabbit papilloma virus mRNAs expressed in the transplantable VX2 carcinoma. J. Virol. 53:735-741.
- Dreyfus, M. 1988. What constitutes the signal for the initiation of protein synthesis on *E. coli* mRNAs? J. Mol. Biol. 204:79–94.
- Franke, A., and D. Clewell. 1980. Evidence for conjugal transfer of a *Streptococcus faecalis* transposon Tn916 from a chromosomal site in the absence of plasmid DNA. Cold Spring Harbor Symp. Quant. Biol. 45:77-80.
- Gaillard, J. L., P. Berche, J. Mounier, S. Richard, and P. Sansonnetti. 1987. In vitro model of penetration and intracellular growth of *Listeria monocytogenes* in the human enterocyte-like cell line Caco2. Infect. Immun. 55:2822-2829.
- 17. Gaillard, J. L., P. Berche, and P. J. Sansonetti. 1986. Transposon as a tool to study the role of hemolysin in the virulence of *Listeria monocytogenes*. Infect. Immun. 52:50-55.
- Garnier, T., and S. T. Cole. 1988. Studies of UV-inducible promoters from *Clostridium perfringens in vivo* and *in vitro*. Mol. Microbiol. 2:607-614.
- Hopwood, D. A., M. J. Bibb, K. F. Chater, T. Kieser, C. J. Bruton, H. M. Kieser, D. J. Lydiate, C. P. Smith, J. M. Ward, and H. Schrempf. 1985. Genetic manipulation of *Streptomyces*, p. 213–220. The John Innes Foundation, Norwich, England.
- Kathariou, P. Metz, H. Hof, and W. Goebel. 1987. Tn916induced mutations in the hemolysin determinant affecting virulence of *Listeria monocytogenes*. J. Bacteriol. 169:129–137.
- Kaufmann, S. H. E. 1988. Immunity against intracellular bacteria: biological effector functions and antigen specificity of T lymphocytes. Curr. Top. Microbiol. Immunol. 138:141–176.
- 22. Ludwig, W., K. H. Schleifer, and E. Stackebrandt. 1984. 16S rRNA analysis of *Listeria monocytogenes* and *Brochothrix thermosphacta*. FEMS Microbiol. Lett. 25:199-204.
- Mackaness, G. B. 1962. Cellular resistance to infection. J. Exp. Med. 116:381-406.
- 24. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Mengaud, J., J. Chenevert, C. Geoffroy, J. L. Gaillard, and P. Cossart. 1987. Identification of the structural gene encoding the SH-activated hemolysin of *Listeria monocytogenes*: listeriolysin O is homologous to streptolysin O and pneumolysin. Infect. Immun. 55:3225–3227.
- Mengaud, J., M. F. Vicente, J. Chenevert, J. Moniz Pereira, C. Geoffroy, B. Gicquel-Sanzey, F. Baquero, J. C. Perez-Diaz, and P. Cossart. 1988. Expression in *Escherichia coli* and sequence analysis of the listeriolysin O determinant of *Listeria monocytogenes*. Infect. Immun. 56:766-772.

- 27. Messing, J. 1983. New M13 vectors for cloning. Methods Enzymol. 101:20-78.
- Pabo, C. O., and R. T. Sauer. 1984. Protein-DNA recognition. Annu. Rev. Biochem. 53:293–321.
- 29. Platt, T. 1986. Transcription termination and the regulation of gene expression. Annu. Rev. Biochem. 55:339–372.
- Portnoy, D., P. S. Jacks, and D. Hinrichs. 1988. Role of hemolysin for the intracellular growth of *L. monocytogenes*. J. Exp. Med. 167:1459-1471.
- Raibaud, O., and M. Schwartz. 1984. Positive control of transcription initiation in bacteria. Annu. Rev. Genet. 18:173-206.
- 32. 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.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.
- 34. Seeliger, H. P. R., and D. Jones. 1986. Genus Listeria Pirie 1940, p. 1235–1245. In P. H. A. Sneath, N. S. Mair, N. E. Sharpe, and J. G. Holt (ed.), Bergey's manual of systematic bacteriology, vol. 2. The Williams & Wilkins Co., Baltimore.
- Shih-Peng, S., and R. K. Holmes. 1988. Iron regulation of the cloned diphtheria toxin promoter in *Escherichia coli*. Infect. Immun. 56:2430-2436.
- Sokolovic, Z., and W. Goebel. 1989. Synthesis of listeriolysin in Listeria monocytogenes under heat shock conditions. Infect. Immun. 57:295-298.
- 37. Staden, R., and A. D. McLachlan. 1979. Codon preference and its use in identifying protein coding regions in long DNA sequences, p. 141–153. *In* D. Soll and R. J. Roberts (ed.), The applications of computers to research on nucleic acids. IRL Press, Oxford.
- 38. Suggs, S. V., T. Hirose, T. Miyake, E. H. Kawashima, M. J. Johnson, K. Itakura, and R. B. Wallace. 1981. Use of synthetic oligonucleotides for the isolation of specific cloned DNA sequences, p. 683–697. *In* D. Brown (ed.), Developmental biology using purified genes. Academic Press, Inc., New York.
- Thomas, P. S. 1980. Hybridization of denatured RNA, and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA 77:5201-5205.
- Tinoco, I., Jr., P. N. Borer, B. Dengler, M. D. Levine, O. C. Uhlenbeck, D. M. Crothers, and J. Gralla. 1973. Improved estimation of secondary structure in ribonucleic acids. Nature (London) New Biol. 246:40-41.
- Vicente, M. F., F. Baquero, and J. C. Perez-Diaz. 1985. Cloning and expression of the *Listeria monocytogenes* haemolysin in *E. coli*. FEMS Microbiol. Lett. 30:77–79.
- 42. Von Heijne, G. 1983. Patterns of amino acids near signalsequence cleavage sites. Eur. J. Biochem. 133:17-21.
- 43. Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103–119.
- Watson, M. E. E. 1984. Compilation of published signal sequences. Nucleic Acids Res. 12:5145-5163.