

Molecular Cloning, Sequencing, and Identification of a Metalloprotease Gene from *Listeria monocytogenes* That Is Species Specific and Physically Linked to the Listeriolysin Gene

EUGEN DOMANN, MICHAELA LEIMEISTER-WÄCHTER, WERNER GOEBEL,
AND TRINAD CHAKRABORTY*

*Institut für Genetik und Mikrobiologie der Universität Würzburg,
8700 Würzburg, Federal Republic of Germany*

Received 11 July 1990/Accepted 16 October 1990

The entire nucleotide sequence of an open reading frame located immediately downstream of the listeriolysin gene from a virulent *Listeria monocytogenes* serotype 1/2a strain was determined. The product of the open reading frame was 510 amino acids with a predicted molecular weight of 57,400. The deduced amino acid sequence of this open reading frame is highly similar to that of a family of secreted metalloproteases produced by various members of the genus *Bacillus*, of which thermolysin is the prototype. Immunoblots performed with specific antisera raised against thermolysin from *Bacillus stearothermophilus* allowed the detection of a 60-kDa polypeptide, corresponding to the pro-form of the protease, in culture supernatants of *L. monocytogenes* strains. In maxicell experiments, *Escherichia coli* recombinants harboring this open reading frame also specifically directed production of a 60-kDa protein. Protease activity was low to undetectable in both *Listeria* strains and *E. coli* recombinants. This is due to lack of processing of the inactive pro-form of the protease to its mature active form in both species. We have designated this gene *mpl* for metalloprotease of *L. monocytogenes*. The gene was present only in pathogenic *L. monocytogenes* strains, in which it was physically linked to the listeriolysin gene.

Listeria monocytogenes is a motile gram-positive microorganism with a wide distribution in nature, having been isolated from soil, water, vegetation, and many animal species (30). It is an opportunistic intracellular pathogen responsible for a disease that in its severest form is a meningitis (17). Outbreaks of listeriosis have been traced to the consumption of food, mainly, dairy produce and vegetables (11). The fetus, the newborn, the old, and the immunocompromised individual are the main victims of this disease (17).

Relatively little is known of virulence factors that contribute to the pathogenesis of listerial infection. The secreted hemolysin, listeriolysin, is currently the only gene product that has been unequivocally assigned a role as a virulence factor (reviewed in reference 7). Yet, it is clear that, to cause disease, the intruding bacterium must have other factors that assist and promote evasion from immune defenses. These include factors for adhesion and invasion of the eucaryotic cell, survival and growth within the host cell, and cell-to-cell spread as well as factors that influence the regulation of the host immune response (5, 7).

In the course of our genetic studies on the listeriolysin gene of an *L. monocytogenes* serotype 1/2a strain, EGD, we found that sequences located downstream of the listeriolysin gene were unique to the pathogenic species *L. monocytogenes* (14). More recently, partial DNA sequencing of this region from an *L. monocytogenes* serotype 1/2c strain, LO28, indicated the presence of the start of an open reading frame (ORF), denoted ORF D, apparently coding for a secreted protein (19). Transcriptional mapping of the RNA start site of this ORF revealed the presence of sequence elements in its promoter region that were in common with

the promoter region of the *lisA* gene. This led to the suggestion that, in *L. monocytogenes*, the expression of this ORF is responsive to the same environmental stimuli and regulated in a manner similar to the listeriolysin gene (19).

In this study, we present the entire nucleotide sequence of an ORF located immediately downstream of the *lisA* gene from the virulent *L. monocytogenes* serotype 1/2a strain EGD. We report that the deduced amino acid sequence of this ORF is highly similar to that of members of a family of secreted metalloproteases produced by various members of the species *Bacillus*. We have designated this gene *mpl* for metalloprotease of *L. monocytogenes*.

MATERIALS AND METHODS

Bacterial strains and plasmids. *L. monocytogenes* serotype 1/2a strains EGD and NCTC 7973 and *L. innocua* strain NCTC 11288 have been described previously (14). *Escherichia coli* strain DH5 α (*endA1 hsdR17 supE44 thi-1 recA1 gyrA96 relA1*) was used for cloning and transformation (14). Maxicell analysis of plasmid-encoded polypeptides was performed with strain CSH26 Δ F6 [*ara thi* Δ (*lac-pro*) Δ (*recA-srlF6*) *rpsL*]. The plasmid cloning vector pUC18 has been described previously (14). Restriction analyses and plasmid constructions were done by standard techniques as outlined by Maniatis et al. (16).

Media and reagents. *Listeria* spp. were cultured in brain heart infusion broth (Difco Laboratories, Detroit, Mich.) at 37°C. *E. coli* strains were grown in LB at 37°C. LM plates consisted of L broth with 1% agar and 3% skim milk. Ampicillin (Sigma Chemical Co., St. Louis, Mo.) was used at a final concentration of 100 μ g/ml. Restriction endonucleases and ligase were purchased from Boehringer Mannheim GmbH, Mannheim, Federal Republic of Germany, and were used as suggested by the manufacturer. Avian myelo-

* Corresponding author.

blastosis virus reverse transcriptase and T7 DNA polymerase were from Pharmacia, Freiburg, Federal Republic of Germany. The peroxidase-conjugated anti-rabbit immunoglobulin (DAKO Laboratories, Hamburg, Federal Republic of Germany) was used at a 1:1,000 dilution. All other salts and ancillary agents were purchased from Merck, Darmstadt, Federal Republic of Germany.

DNA sequence analysis. The DNA sequence of the cloned *L. monocytogenes* DNA was determined from double-stranded plasmid templates by dideoxy-chain termination (22). Double-stranded templates were denatured, and the sequencing reactions were carried out with T7 DNA polymerase as suggested by the commercial supplier in the product literature. Sequencing reactions were primed from vector and custom-made oligonucleotide primers. [α - 32 P]dATP (800 Ci/mmol) was purchased from Amersham Büchler. The labeled reaction mixtures were separated by electrophoresis on urea-6% polyacrylamide gels, the gels were dried down, and the sequence was read from X-ray film autoradiograms. The sequence data were analyzed by using the University of Wisconsin Computer Group software (8).

Primer extension analysis. A synthetic oligonucleotide of 25 nucleotides from a region upstream of the *mpl* gene was 5' labeled and used as primer (see Fig. 1). Labeled primer (2 pmol) was annealed to 10 μ g of cellular RNA in 6 μ l of a solution containing 50 mM Tris hydrochloride (pH 8.3), 60 mM NaCl, 10 mM dithiothreitol, and 1 mM EDTA. Subsequently, 9 μ l of reaction mixture was added, resulting in a solution containing 0.5 μ l of avian myeloblastosis virus reverse transcriptase (Pharmacia), 50 mM Tris hydrochloride (pH 8.3), 60 mM NaCl, 10 mM dithiothreitol, 6 mM magnesium acetate, and 0.5 mM concentrations of each of the four deoxynucleoside triphosphates. The samples were incubated at 42°C for 90 min, vacuum dried, and resuspended in 10 μ l of formamide dye mixture. Aliquots, 3 μ l, were then heated at 95°C for 30 s and immediately loaded onto 6% polyacrylamide-urea sequencing gels for electrophoresis. Dideoxy sequencing reactions (22), using the same primer and an appropriate plasmid DNA template, were run in parallel to allow determination of the endpoints of the extension products.

PAGE. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed with a 12.5% separating gel, using 10 μ l of trichloroacetic acid-concentrated supernatants (1 ml) of *Listeria* strains in sample buffer. Polypeptides were visualized by staining the gel with Coomassie blue R-250.

Analysis of plasmid-coded translational products in a maxicell system has been described previously (6). Labeling of polypeptides was carried out in methionine assay medium (Difco) containing [35 S]methionine (Amersham-Büchler). Samples were analyzed as described above, dried, and used to expose Fuji RX X-ray film. For immunoblot reactions (27), proteins were transferred to nitrocellulose paper, reacted with rabbit anti-thermolysin antiserum (28) at a 1:1,000 dilution, and stained with horseradish peroxidase-conjugated second antibody, as described previously (14).

Southern hybridization. Bacterial chromosomal DNA was isolated as described in reference 14. Total genomic DNA was cleaved by *Sph*I endonuclease digestion and electrophoresed overnight at 30 V on 0.7% agarose. The DNA was transferred to nitrocellulose sheets by the method of Southern (24). DNA probes were isolated, purified, and labeled with [α - 32 P]dATP by the random priming technique of Feinberg and Vogelstein (10). The *lisA* probe consisted of a 651-bp *Hind*III fragment internal to the gene. A 673-bp

*Hind*III-*Bam*HI probe from within the *mpl* gene was used for detection of the presence of this gene. Hybridization conditions for the detection of the various listerial genes were as described in reference 14.

Nucleotide sequence accession number. The nucleotide sequence data reported in this paper have been submitted to EMBL and assigned the accession number X54619.

RESULTS

The DNA sequence of a region of *L. monocytogenes* chromosomal DNA from strain EGD, located between the *Hind*III restriction endonuclease site situated immediately downstream from the *lisA* gene to the *Hae*III restriction endonuclease site 2 kb away, was determined (Fig. 1). The nucleotide sequence revealed the presence of a large ORF of 1,533 nucleotides (nucleotides 2092 to 3624) located 329 nucleotides 3' to the UAA stop codon of the listerialysin gene. It was preceded by a putative ribosome binding site located seven nucleotides upstream of its initiation codon ATG. The reading frame encoded a protein of 510 amino acid residues with a calculated molecular weight of 57,400 and pI of 6.64. Sequences of a dyad symmetry that resemble a *rho*-independent terminator (21) were located 33 nucleotides downstream of the termination codon.

The first 72 nucleotides of the ORF encode a 24-amino-acid peptide that has all of the features of a procaryotic signal peptide sequence. It has an overall helical structure that is predicted to be transmembranous, positively charged amino acid immediately following the initiator methionine and a central hydrophobic core followed by more polar amino acids. This result suggested that the product of this ORF is, like listerialysin, a secreted polypeptide in *L. monocytogenes*.

Homology with bacterial neutral proteases. Analysis of the predicted polypeptide sequence of the ORF for homology to other sequences in the GenBank data base, using the FASTP algorithm of Lipman and Pearson (15), revealed strong similarities to secreted bacterial zinc metalloproteases from *Bacillus* spp. (13, 23, 25, 29). The similarities were substantial (Fig. 2), especially in the region corresponding to the mature forms of these proteases, and highly conserved with respect to residues forming the catalytic center of these enzymes (Fig. 8 of reference 12). The *Bacillus* proteases are all preproenzymes and possess a putative signal peptide at the N-terminal end of their full-length ORFs. Mature protease is produced by cleavage of an N-terminal pro-sequence and is approximately 60% of the total length predicted from their full-length ORFs. A comparison of the sequences available for the mature form of these enzymes suggests amino acid Glu-200 as the site at which the scissile bond is cleaved to give the mature form of the protease. In the case of the listerial protease, this would result in a mature protein of 35,000 Da. The ORF encoding this polypeptide has been designated *mpl*, denoting metalloprotease of *L. monocytogenes*.

Transcription start site of the *mpl* gene. The transcription start site of the *mpl* gene was mapped by primer extension, using avian myeloblast reverse transcriptase of RNAs extracted from *L. monocytogenes* EGD (Fig. 3). Two start sites for transcripts were detected, a minor transcript starting 149 nucleotides and a major transcript starting 150 nucleotides, respectively, upstream of the initiation ATG codon of the *mpl* gene. At the spatial regions corresponding to the -10 and -35 regions of either transcript, there is good fit at the -10 region but poor homology at the -35 region

TAACGACGATAAAGGGACAGCAGGACTAGAAATAAAGCTATAAAGCAAGCATATAATATG 60
CGTTTCATCTTTAGAACGGAAATTCGCCAATATTATAATATCAAAGAGAGGGGTGGCA 120
AACCGTATTTCGCATTATTAGGTTAAAAAATGTAGAAGGAGAGTGAACCCATG----- 174
HET-lisa-
---TAATTGTAAAGTAATAAAAAATTAAGAATAAAAAACCGCTTAACACACACGAAAAAAT 1815
---END
HindIII
AAGCTGTTTTGCACCTTCGTAATATTATTTGTGAAGAATGTAGAAACAGGCTTATTTT 1875
TTAATTTTTTTAGAGAATTAACAAATGTAAAAGAATATCTGACTGTTTATCCATAATAT 1935
-35 -10
ATAAGCATATCCCAAAGTTTAAAGCCACCTATAGTTTCTACTGCAAAAACGTATAATTTAGT 1995
+1 3'
TCCCACATATACTAAAAAGGTGCTTCTTACTCTCTGTGCAGATTAGTGTAGGTGGCT 2055
---GGGTGATATGATTTTTGACACAGG-5'
TAAACTTAGTTTTACGAATTAAGAGGAGCGGTGAATGAAAAGTAAACTTATTTGTATC 2115
RBS METLysSerLysLeuIleCysIle 8
ATCATGGTAATAGCTTTTCAGGCTCATTTCACTATGACGGTAAAAGCAGATTCTGTGCGG 2175
IleMetValIleAlaPheGlnAlaHisPheThrMetThrValLysAlaAspSerValGly 28
HindIII
GAAGAAAACCTCAAATAATACACAGCCAAAAAGCCCTGCTGATTAAAGCTTTG 2235
GluGluLysLeuGlnAsnAsnThrGlnAlaLysLysThrProAlaAspLeuLysAlaLeu 48
CCAGATTCCTCGCAAGCAAAAGATTTTTATAAAAAATTTAAATTTCTGATATGACAAA 2295
ProAspSerCysGluAlaLysAspPheLysAsnPheLysIleLeuAspMetGlyThrLys 68
GATAAGCTAGGCGTTACGCATTATACGCTCGCGCTAAGTCTGGTGTACTTTCAGCGAT 2355
AspLysLeuGlyValThrHisTyrThrLeuAlaLeuSerSerGlyGlyTyrLeuThrAsp 88
AATGATGAATCAAGGTCACCGCTCACACCGGATAAAAACTCACTTTCATAAATGGAGAT 2415
AsnAspGluIleLysValHisValThrProAspAsnLysIleThrPheIleAsnGlyAsp 108
TTGCAGCAAGGACAGCTTAGGATTACTAATCAATAAAAAATTACAGAAAAAATGCTATC 2475
LeuGlnGlnGlyGlnLeuArgIleThrAsnGlnIleLysIleThrGluLysAsnAlaIle 128
GAAAAGCAATTGAGCCATCGGTACAGCTGAAGCTCATGTGAAAAGTTATGTCGGAAAC 2535
GluLysAlaPheGluAlaIleGlnSerGluAlaHisValLysSerTyrValGluLys 148
CCAGTGAAGAAAAAGAAATCATCCTCAATTCAGACAACAAACGCTTAGTTTATAATATA 2595
ProValLysGluLysGluIleIleLeuAsnSerArgThrLysArgLeuValTyrAsnIle 168
AAATGATTTTTGCTGAGCCGGAAGTGGCAGTGGATGTTCAAGTGGACGTAGAAAACC 2655
LysLeuIlePheAlaGluProGluValAlaLysTrpIleValGlnValAspValGluLys 188
GGCGCAATTTAAAAAGCAAAATATGCTTTCCGAGGTAGAAGGGGTGATACCCACAAA 2715
GlyAlaIleLeuLysLysGlnAsnMetLeuSerGluValGluArgAlaAspThrHisLys 208
GATTTCAAGCGCTTGGTAAAGGACCAACCGATGCTCCAGAGGCCACTACATGTTATG 2775
AspPheGlnAlaLeuGlyLysGlyAlaAsnArgLeuLeuGlnArgProLeuHisValMet 228
AAAAATATGACCTATTACTTAGTGGATAGAATCAAAAGGACTCATAGAACCTTTT 2835
LysIleAsnAspLeuPheTyrLeuValAspArgThrHisLysGlyLeuIleArgThrPhe 248
GATTTAAAACATAATACCGACACTCGTTTGGGAAAGTGGTATCAATAAAAACGAATATG 2895
AspLeuLysHisAsnThrAspThrSerPheGlyLysValValSerAsnLysThrAsnMet 268
BamHI
TTTACGATCCAGAAATTTAGTTCCCGCAGTGGATGCTCATTTTTACGCAAGTGAAGTGTAC 2955
PheThrAspProGluPheSerSerAlaValAspAlaHisPheTyrAlaSerGluValTyr 288
GAATACTATAAAAATGCCCAACTAGAGAGTCTAGATGTTAAAGGTGGAGAAATGAT 3015
GluTyrLysAsnValIleGlnLeuLysLeuAspGlyLysGlyGlyGluIleAsp 308
Ball I
TCGTTTGCCATTATGGCTGAATGCAATAATGCCTTTTGGATGCCAAGAAATTCCT 3075
SerPheValHisTyrGlyLeuAsnCysAsnAsnAlaPheTrpAspGlyGlnGluIleLeu 328
TATGGAGATGGGCAAAAAGAAATTTCAAACCTTTTCATGCGCCAAAACATTTGTTGGT 3135
TyrGlyAspGlyAspLysLysAsnGluLysPheGlyLysLeuLysThrIleValGly 348
NruI
CATGAACTAACGCATCGGTTATCCAGTATTCGGCGGGATTGGAATACGAAGGGCAATCA 3195
HisGluLeuThrHisIleAlaValIleGlnTyrSerAlaGlyLeuGluTyrGluGlyGlnSer 368
GGTGCCTAAACGAGTCGTTGCGCGATGTTTTGGTTATTTTATGCGCCAAATCATTTG 3255
GlyAlaLeuAsnGluSerPheAlaAspValPheGlyTyrPheIleAlaProAsnHisTrp 388
TTGATTGGTGGAGATGCTGTGTGCGTGGGTCGCGAGATGGCGCAATAAGAAGCATTAAA 3315
LeuIleGlyGluAspValCysValArgGlySerArgAspGlyArgIleArgSerIleLys 408
GATCCTGACAAATATAATCAAGCGGCTCATATGAAAAGATTACGAATCGCTTCCACTCACA 3375
AspProAspLysTyrAsnGlnAlaAlaHisMetLysAspTyrGluSerLeuProLeuThr 428
GAGGAAGCGACTGGGCGGAGTTCATTATAATAGTGGTATCCCGAATAAAGCAGCCTAT 3435
GluGluGlyAspTrpGlyValHisTyrAsnSerGlyIleProAsnLysAlaAlaTyr 448
AATACTACTACTAACTGGAAAAGAAAAACAGACAGCTTTATTTTCGGCCCTTAAAG 3495
AsnThrIleThrLysLeuGlyLysGluLysPheGlyLysLeuLysThrIleValGly 468
TACTATTTAAAGAAAAATCCAGTTTACCGATGCGAAAAAGCGCTTCAACAAGCAGCG 3555
TyrTyrLeuThrLysLysSerGlnPheThrAspAlaLysLysAlaLeuGlnGlnAlaAla 488
HindIII
AAAGATTTATATGTTGAAGTCTCTAAAAAAGTTGCTGAAGCTGGGAAGCAGTTGGG 3615
LysAspLeuTyrGlyGluAspAlaSerLysLysValAlaGluAlaTrpGluAlaValGly 508
GTTAACTGATTAACAAATGTAGAGAAAAATTAATCTCCAAGTATATCTTAAATAA 3675
ValAsnEND 510
TCATGAAATTTTTTCTTATATTAGCTAATTAAGAAGATAACTAAGCTGCTAATCCAATT 3735
HaeIII
TTTAAACGAAACAAATTAGTGAATAAGAGCC 3767

when compared with consensus procaryotic promoter regions (18). Both the *lisA* and the *mpl* genes are transcribed by independent promoters from the same strand on the chromosome of *L. monocytogenes* (Fig. 4).

Detection and identification of the product of the *mpl* gene.
To identify the gene product encoded by the cloned *mpl* gene in *E. coli*, plasmid pLM47-300 containing only the reading frame corresponding to the *mpl* gene was transformed into the maxicell strain CSH26ΔF6. Polypeptides encoded by plasmid pLM47-300 and the vector plasmid pUC18 were detected by radioactive labeling with [³⁵S]methionine. A unique 60-kDa polypeptide was expressed in the strain carrying pLM47-300 (Fig. 5). Synthesis of all other identifiable proteins detected are directed by the vector plasmid. The detection of a 60-kDa form of the protease would be expected if no processing to the mature form were to occur. Indeed, we found that the recombinant plasmids pLM47, 47-210, and 47-300 harboring the *mpl* gene expressed very low proteolytic activities, showing small zones of clearing on LB-skim milk agar plates only after prolonged incubation (48 h) at 37°C. This activity was absent in the same *E. coli* strain harboring the disrupted version of the *mpl* gene on plasmid pLM47-200.

Detection and expression of protease activity in *Listeria* spp.
To date, no proteolytic activity has ever been reported for the species *L. monocytogenes*. When strain EGD was streaked on LB-skim milk agar plates, no proteolytic activity could be detected even after incubation at 37°C for several days. From several serotype 1/2a strains that were assayed for proteolytic activity, we found that only colonies of the type strain NCTC 7973 showed small zones of clearing after overnight incubation on skim milk agar plates at 37°C.

The reason for the absence of activity in *L. monocytogenes* EGD became apparent when trichloroacetic acid-precipitated culture supernatants were subject to SDS-PAGE, blotted onto nitrocellulose filters, and developed with antiserum raised against purified thermolysin of *Bacillus stearothermophilus* (Fig. 6). A polypeptide of 60-kDa size was specifically recognized by the antiserum in the supernatant fluids of *L. monocytogenes* serotype 1/2a strains EGD and NCTC 7973; no reaction with the antiserum was detected with supernatant fluids of *L. innocua*. It appears that, while the *mpl* gene is transcribed (Fig. 3) and expressed under standard laboratory conditions used for culturing *Listeria* spp., only the inactive proform of the protease is produced. The detection of small quantities of a 35-kDa polypeptide in the supernatants of strain NCTC 7973, absent in the supernatant fluids of strain EGD, would account for the proteolytic activity observed with this strain. These results also confirm that a 60-kDa polypeptide secreted by

FIG. 1. Nucleotide sequence and deduced amino acid sequence of the *L. monocytogenes* metalloprotease gene, including upstream and downstream flanking sequences. The coordinates of the nucleotide sequence used here are a continuation of the numbering system used for the DNA sequence of the listeriolysin gene of this strain (9). The initiation and the termination codons of the *lisA* gene are indicated, with dashed lines representing regions from the previously published sequence. The promoter region and the transcriptional start sites are indicated as -35, -10, and +1, respectively. The oligonucleotide sequence used for primer extension analysis is indicated. The ribosome binding site (RBS), the putative signal peptide, and transcriptional terminator are underlined and indicated in that order from the 5' start to the 3' end of the sequence presented.

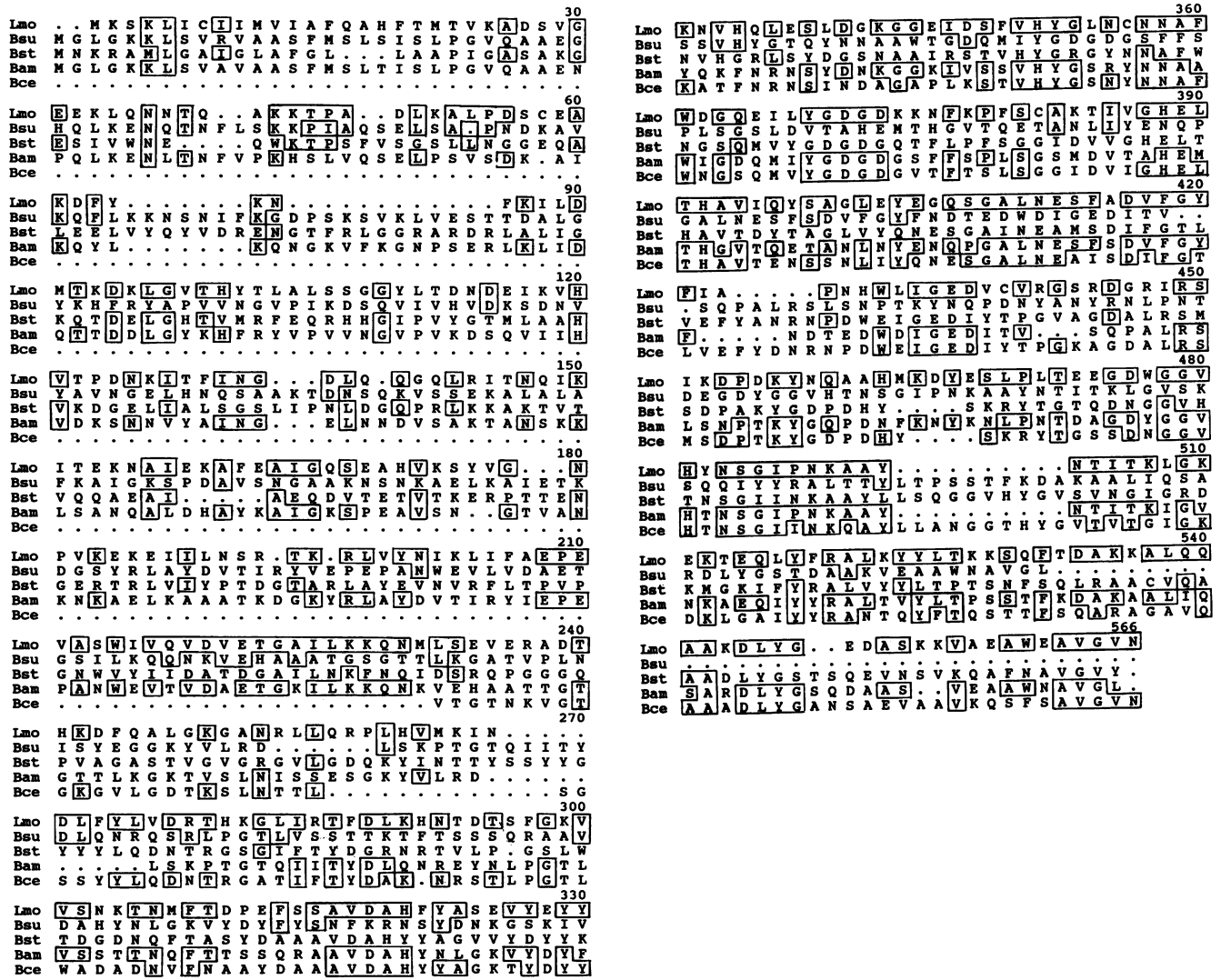


FIG. 2. Sequence alignments of metalloproteases from *Bacillus* spp. and the Mpl protease of *L. monocytogenes* (Lmo). Alignments were produced by using the Bestfit algorithm with a gap weight of 3.000 and length weight of 0.100 (8). Only amino acids identical to the Mpl protease of *L. monocytogenes* are boxed. Percent amino acid identity of Mpl protease to *B. subtilis* metalloprotease (Bsu) is 41.8%; *B. stearothermophilus* (Bst), 34.2%; *B. amyloliquefaciens* (Bam), 41.8%; and *B. cereus* (Bce), 38.3%.

strains of *L. monocytogenes* is in fact antigenically related to the metalloproteases produced by many *Bacillus* spp., of which thermolysin is a prototype (13).

The *mpl* gene is unique to the species *L. monocytogenes* and is physically linked to the listeriolysin gene in these strains. Genomic DNA of strains comprising several commonly occurring serotypes of *L. monocytogenes* and the species *L. ivanovii*, *L. innocua*, and *L. seeligeri* were isolated and digested with the restriction endonuclease *Sph*I. In the serotype 1/2a EGD strain, all of the *lisA* gene and part of the *mpl* gene are located on a common 3.86-kb *Sph*I fragment (14). Hence, if physical linkage of both genes also exists in strains of other serovars, DNA probes harboring sequences internal to each of these genes would be expected to hybridize to a common *Sph*I DNA fragment. The results obtained with two such DNA probes are depicted in Fig. 7A and B. In all *L. monocytogenes* strains tested, physical linkage between the *lisA* and *mpl* genes was detected. The results also demonstrated that, although the listeriolysin gene is present

in *L. ivanovii* and *L. seeligeri*, no sequences homologous to the *mpl* gene are present in these strains. Hence, the *mpl* gene is unique to the pathogenic species *L. monocytogenes*.

DISCUSSION

In this communication, we have reported on the molecular cloning, sequencing, and identification of a gene located immediately downstream of the listeriolysin gene in *L. monocytogenes* serotype 1/2a strain EGD. The product of the ORF was composed of 510 amino acids with a predicted molecular weight of 57,400. The amino acid sequence of the polypeptide showed strong similarities to those of other secreted bacterial metalloproteases, especially those belonging to the species *Bacillus* (13, 23, 25, 29). We have designated this gene *mpl* for metalloprotease of *L. monocytogenes*. No proteolytic activity has ever been described for the species *L. monocytogenes*, and our results document for the first time the presence of a metalloprotease gene in this

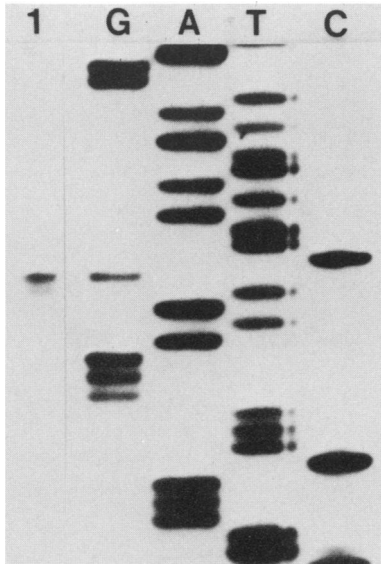


FIG. 3. Mapping and detection of *mpl* transcripts in *L. monocytogenes* EGD by primer extension analysis. An α -³²P-end-labeled, 25-oligonucleotide-long primer (Fig. 1) was used for reverse transcription of the transcript. The DNA products were separated on an 6% polyacrylamide gel simultaneously with a dideoxy sequencing reaction ladder, using the same DNA primer and plasmid pLM47 as template to allow determination of the extension product. Lane 1 denotes the primer extension reaction done with total RNA from strain EGD. Lanes denoted G, A, T, and C are tracts of the sequencing reaction.

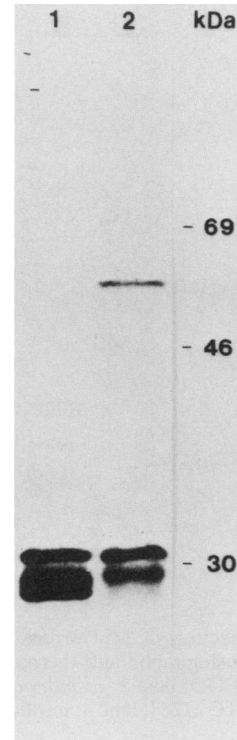


FIG. 5. Autoradiogram of ³⁵S-labeled polypeptides encoded by pUC18 (lane 1) and pLM47-300 (lane 2) in maxicells. The 60-kDa polypeptide in lane 2 is the product of the *mpl* gene. Molecular mass standards are indicated.

species. The gene is absent in all other species of the genus *Listeria*.

Primer extension analysis of RNA extracted from *L. monocytogenes* bacteria indicated the presence of two transcriptional start sites located 149 and 150 nucleotides upstream of the initiator ATG codon. The longer RNA transcript is the major form produced in this strain (Fig. 3). A previous analysis of the transcriptional start sites of the gene

from an *L. monocytogenes* serotype 1/2c strain revealed a single RNA start point for transcripts located 149 nucleotides upstream of the *mpl* gene of that strain (19). The reason for this discrepancy is not obvious, since the available DNA sequences from this region of the chromosome for both strains are identical.

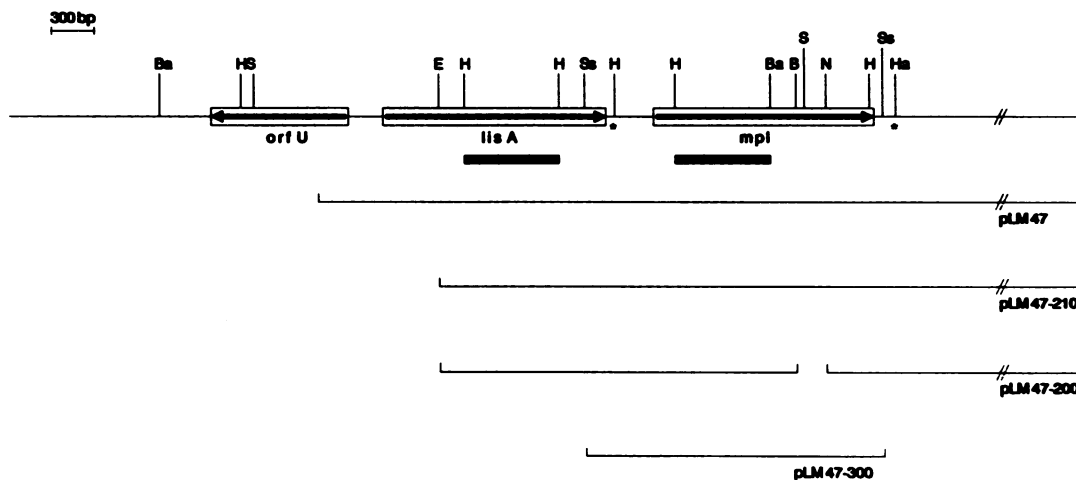


FIG. 4. Partial restriction map of the region around the *mpl* gene of *L. monocytogenes* EGD. Thin black lines represent various lengths of *L. monocytogenes* chromosomal DNA inserted into the plasmid pUC18 vector. Regions following the double slashed lines are not drawn to scale. Solid blocks represent the DNA probes used in the detection of the *lisA* and *mpl* genes. The location of ORF U (19), *lisA*, and *mpl* are shown by open boxes together with the direction of transcription of these genes. Restriction endonucleases are as follows: B, *BalI*; Ba, *Bam*HI; E, *Eco*RI; Ha, *Hae*III; H, *Hind*III; N, *Nru*I; S, *Sph*I; Ss, *Ssp*I.

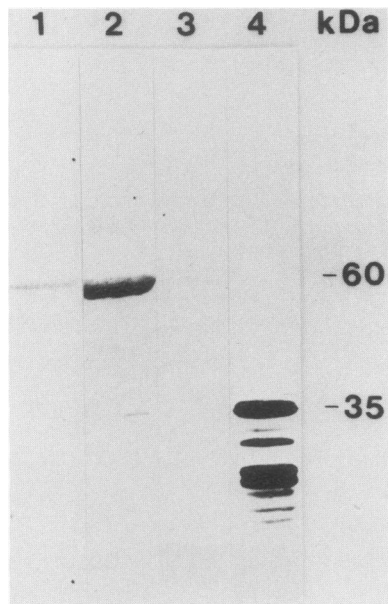


FIG. 6. Immunodetection of Mpl protease in culture supernatants of *Listeria* spp., using rabbit anti-thermolysin antiserum. Lane 1, *L. monocytogenes* EGD; lane 2, *L. monocytogenes* NCTC 7973; lane 3, *L. innocua* NCTC 11288; lane 4, purified thermolysin from *B. stearothermophilus*.

A palindromic sequence of 14 bp centered around the -40 regions of the promoters of genes detected within this region of the serotype 1/2c strain has invited speculation that it is a site at which DNA-binding proteins may act and given rise to the notion that these genes are commonly regulated (19). Such palindromic sequences are also present in the spatial region corresponding to the -40 regions of the *mpl* and *lisA* genes in strain EGD. We note, however, that a similar palindrome starting from nucleotides 3625 to 3638 also brackets the termination codon of the *mpl* gene. The role of the palindromic sequences in mediating coordinate regulation of these genes is not clear at this time; more studies will be required to address their significance.

The amino terminal of the predicted polypeptide contained a 24-amino-acid sequence with sequence features typical of a signal peptide. That this protein is in fact secreted was confirmed by the detection of a 60-kDa protein, cross-reactive with antiserum to purified thermolysin from *B. stearothermophilus*, in the supernatant fluids of *L. monocytogenes* strains. Little or no cross-reactive material was detected in total cell lysates of these bacteria (data not shown). The secreted bacterial metalloproteases are all preproenzymes and processed after secretion of the inactive pro-form to their respective active mature forms (13, 23, 25, 29). Amino acid sequence comparisons to the other metalloproteases suggest Glu-200 as the site of processing of the proform of the Mpl protease, to give a mature active protein of 35 kDa.

A secreted protease of 60,000 molecular weight would correspond to the inactive pro-form of the Mpl protease. This would explain why little or no protease activity was found in the supernatant fluids of various *L. monocytogenes* strains. It is apparent that, under the standard laboratory conditions used for culturing *Listeria* spp., production of the active form of the protease is not favored. Controlled production of active protease from its inactive precursor

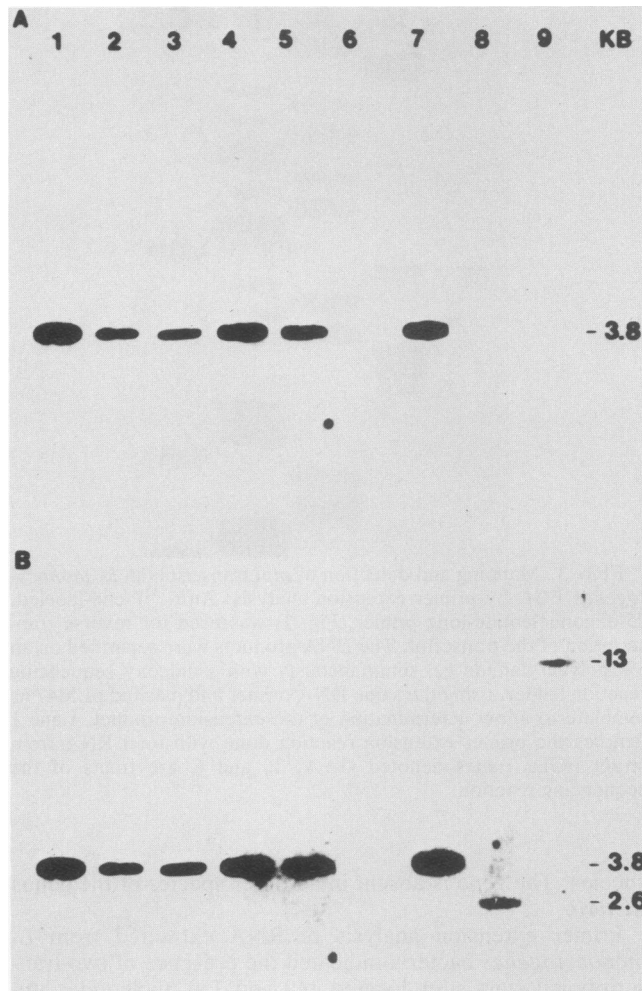


FIG. 7. Autoradiograph of a Southern blot of chromosomal DNA from different strains digested with restriction endonuclease *SphI* and hybridized to (A) a 673-bp *Bam*HI-*Hind*III fragment from the *mpl* gene and (B) a 651-bp *Hind*III fragment from the *lisA* gene. Lane 1, *L. monocytogenes* EGD serotype 1/2a; lane 2, *L. monocytogenes* SLCC 2755 serotype 1/2b; lane 3, *L. monocytogenes* NCTC 5348 serotype 1/2c; lane 4, *L. monocytogenes* SLCC 4013 serotype 4b; lane 5, *L. monocytogenes* L99 serotype 4a; lane 6, *L. innocua* NCTC 11288 serotype 6a; lane 7, *L. monocytogenes* SLCC 2487 serotype 7; lane 8, *L. seeligeri* SLCC 3954 serotype 1/2b; lane 9, *L. ivanovii* ATCC 19119 serotype 5.

within the environment of the host would be a means of effectively targeting it to its appropriate site of action. In the only strain in which activity was present, small but detectable amounts of a 35-kDa polypeptide were detected in its supernatant fluids (Fig. 6).

The low level of protease activity in *E. coli* recombinants harboring the *mpl* gene was likely the result of a combination of low expression of this gene in the recombinants used and the absence of any significant level of processing of the protease to its mature active form in *E. coli*. There is, in fact, precedent for the poor expression of protease activity in *E. coli* recombinants harboring cloned preproenzymes from *Bacillus* species (26).

Recently, a unique amino acid sequence motif representing the core of one of the two zinc-binding sites in ther-

<i>Bacillus subtilis</i>	V	T	A	H	E	M	T	H	G	V	T
<i>Bacillus amyloliquefaciens</i>	V	T	A	H	E	M	T	H	G	V	I
<i>Bacillus stearothermophilus</i>	V	V	G	H	E	L	T	H	A	V	I
<i>Bacillus thermoproteolyticus</i>	V	V	A	H	E	L	T	H	A	V	T
<i>Bacillus cereus</i>	V	I	G	H	E	L	T	H	A	V	T
<i>Listeria monocytogenes</i>	I	V	G	H	E	L	T	H	A	V	I
<i>Legionella pneumophila</i>	V	G	G	H	E	V	S	H	G	F	T
<i>Pseudomonas aeruginosa</i>	V	A	A	H	E	V	S	H	G	F	T
<i>Leishmania major</i>	V	V	T	H	E	M	A	H	A	L	G

FIG. 8. Comparison of the primary structures of zinc-dependent metalloproteases from various bacteria and protozoa, using the three residues (boxed) known to be required for hydrolytic activity. The alignment was generated as described in reference 12.

molysin, (uncharged)-(uncharged)-histidine-glutamine-(uncharged)-(uncharged)-histidine-(uncharged)-(hydrophobic), has been shown to be common to zinc-dependent metalloproteases (12). This motif is also preserved in the Mpl protease of *L. monocytogenes*. Using this signature, we found that similar regions in elastase of *Pseudomonas aeruginosa* (2) and the PEP 1 protease of *Legionella pneumophila* (3) are also conserved (Fig. 8). Both of these proteases are also metalloproteases and possess broad substrate specificity (3, 31). In both microorganisms, proteases have been implicated as virulence factors.

The role and significance of a metalloprotease in pathogenic *Listeria* spp. are unknown. Species other than *L. monocytogenes* are devoid of the gene, suggesting that it is required for neither growth nor persistence in nonanimal environments. Metalloproteases also form a part of the repertoire of secreted proteins of four other human and animal pathogens, namely, *P. aeruginosa* (31), *Legionella pneumophila* (1), *Serratia marcescens* (20), and *Leishmania major* (4), and it is possible that the adaptive significance is related to different stages of the infectious process within the animal host. Clearly, the scavenging function of protease could be an important prerequisite to intracellular survival of *L. monocytogenes* when considering its nutritional requirements. Alternatively, it could exert its effect by degrading host cell proteins and hence influence the response of the host in pathogenesis. Whatever the contribution of the protease to the ability of these bacteria to cause disease, the availability of the gene and its nucleotide sequence will enable us to address fundamental questions concerning its role and regulation during the pathogenesis of listerial infections in experimental animal models.

ACKNOWLEDGMENTS

We thank Brigitte Göpfert for excellent technical assistance and Bertus van den Burg (University of Groningen, Groningen, The Netherlands) for the kind gift of rabbit antiserum against purified

thermolysin. T.C. is indebted to Thomas Jarchau for discussions on many aspects of this work and the ideas expressed herein.

This research was financed by a grant from the Bundesministerium für Forschung und Technologie.

REFERENCES

- Baskerville, A., J. W. Conlan, I. A. Ashworth, and A. B. Dowsett. 1986. Pulmonary damage caused by a protease from *Legionella pneumophila*. Br. J. Exp. Pathol. 67:527-536.
- Bever, R. A., and B. H. Iglewski. 1988. Molecular characterization and nucleotide sequence of the *Pseudomonas aeruginosa* elastase structural gene. J. Bacteriol. 170:4309-4314.
- Black, W. J., F. D. Quinn, and L. S. Tompkins. 1990. *Legionella pneumophila* zinc metalloprotease is structurally and functionally homologous to *Pseudomonas aeruginosa* elastase. J. Bacteriol. 172:2608-2613.
- Button, L. L., and W. R. McMaster. 1988. Molecular cloning of the major surface antigen of *Leishmania*. J. Exp. Med. 167:724-729.
- Chakraborty, T., and W. Goebel. 1988. Recent developments in the study of virulence of *Listeria monocytogenes*. Curr. Top. Microbiol. Immunol. 138:41-58.
- Chakraborty, T., B. Huhle, H. Bergbauer, and W. Goebel. 1986. Cloning, expression, and mapping of the *Aeromonas hydrophila* aerolysin gene determinant in *Escherichia coli* K-12. J. Bacteriol. 167:368-374.
- Cossart, P., and J. Mengaud. 1989. *Listeria monocytogenes*: a model system for the molecular study of intracellular parasitism. Mol. Biol. Med. 6:463-474.
- Devereux, P., P. Haeblerli, and R. O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12:387-395.
- Domann, E., and T. Chakraborty. 1989. Nucleic acid sequence of the listeriolysin gene from a *Listeria monocytogenes* serotype 1/2a strain. Nucleic Acids Res. 17:6406.
- Feinberg, A. P., and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specificity. Anal. Biochem. 132:6-13.
- Fleming, D. W., S. L. Cochi, K. L. MacDonald, J. Brondum, P. S. Hayes, B. D. Plikaytis, M. B. Holmes, A. Audurier, C. V. Broome, and R. L. Reingold. 1985. Pasteurized milk as a vehicle of infection in an outbreak of listeriosis. N. Engl. J. Med. 312:404-407.
- Jongeneel, C. V., J. Bouvier, and A. Bairoch. 1989. A unique signature identifies a family of zinc-dependent metalloproteases. FEBS Lett. 242:211-214.
- Kubo, M., and T. Imanaka. 1988. Cloning and nucleotide sequence of a highly thermostable neutral protease gene from *Bacillus stearothermophilus*. J. Gen. Microbiol. 134:1883-1892.
- Leimeister-Wächter, M., and T. Chakraborty. 1989. Detection of listeriolysin, the thiol-dependent hemolysin in *Listeria monocytogenes*, *Listeria ivanovii*, and *Listeria seeligeri*. Infect. Immun. 57:2350-2357.
- Lipman, D. J., and W. R. Pearson. 1985. Rapid and sensitive protein similarity searches. Science 227:1435-1441.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Marget, W., and H. P. R. Seeliger. 1988. *Listeria monocytogenes* infections—therapeutic possibilities and problems. Infection 16(Suppl. 2):175-177.
- McClure, W. 1985. Mechanism and control of transcription in prokaryotes. Annu. Rev. Biochem. 54:171-204.
- Mengaud, J., M. F. Vicente, and P. Cossart. 1989. Transcriptional mapping and nucleotide sequence of the *Listeria monocytogenes hlyA* region reveal structural features that may be involved in regulation. Infect. Immun. 57:3695-3701.
- Nakahama, K., K. Yoshimura, R. Marumoto, M. Kikuchi, I. S. Lee, T. Hase, and H. Matsubara. 1986. Cloning and sequencing of the *Serratia* protease gene. Nucleic Acids Res. 25:5843-5855.
- Platt, T. 1986. Transcription termination and regulation of gene expression. Annu. Rev. Biochem. 55:339-372.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequenc-

- ing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.
23. Sidler, W., E. Niederer, F. Suter, and H. Zuber. 1986. The primary structure of *Bacillus cereus* metal proteinase and comparison with thermolysin and *Bacillus subtilis* neutral proteinase. Biol. Chem. Hoppe Seyler 367:643-657.
 24. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503-517.
 25. Takagi, M., T. Imanaka, and S. Aiba. 1985. Nucleotide sequence and promoter region for the neutral protease gene from *Bacillus stearothermophilus*. J. Bacteriol. 163:824-831.
 26. Takagi, H., Y. Morinaga, M. Tsuchiya, H. Ikemura, and M. Inouye. 1988. Control of folding of proteins secreted by a high secretion vector pIN III *ompA*: 16-fold increase in production of active subtilisin in *Escherichia coli*. Bio/Technology 6:948-950.
 27. Towbin, H., T. Staehlin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets. Proc. Natl. Acad. Sci. USA 76:4350-4354.
 28. Van Den Burg, B., V. G. H. Eijssink, B. K. Stulp, and G. Venema. 1989. One-step affinity purification of *Bacillus* neutral proteases using Bacitracin-silica. J. Biochem. Biophys. Methods 18:209-222.
 29. Vasantha, N., L. D. Thompson, C. Rhodes, C. Banner, J. Nagle, and D. Filpula. 1984. Genes for alkaline protease and neutral protease from *Bacillus amyloliquefaciens* contain a large open reading frame between the region coding for signal sequence and mature protein. J. Bacteriol. 159:811-819.
 30. Watkins, J., and K. P. Sleath. 1981. Isolation and enumeration of *Listeria monocytogenes* from sewage sludge and river water. J. Appl. Bacteriol. 50:1-9.
 31. Wretling, B., and O. R. Pavlovskis. 1983. *Pseudomonas aeruginosa* elastase and its role in pseudomonas infections. Rev. Infect. Dis. 5(Suppl):998-1004.