The Gene Coding for Protein p60 of *Listeria monocytogenes* and Its Use as a Specific Probe for *Listeria monocytogenes*

STEPHAN KÖHLER,[1] MICHAELA LEIMEISTER-WÄCHTER,[1] TRINAD CHAKRABORTY,[1] FRIEDRICH LOTTSPEICH,[2] AND WERNER GOEBEL[1]*

Institut für Genetik und Mikrobiologie, Universität Würzburg, 8700 Würzburg,[1] and Max-Planck-Institut für Biochemie, 8033 Martinsried,[2] Federal Republic of Germany

The gene of *Listeria monocytogenes* that encodes a major extracellular protein (p60) was cloned in *Escherichia coli*. The gene was designated *iap*, as p60 was previously shown to represent an invasion-associated protein (M. Kuhn and W. Goebel, Infect. Immun. 57:55–61, 1989). The recombinant *E. coli* clone expressed p60, as shown by immunoblotting. The complete nucleotide sequence of *iap* was determined. The deduced amino acid sequence of p60 (484 amino acids) contains a putative N-terminal signal sequence of 27 amino acids and an extended repeat region consisting of 19 threonine-asparagine units. Hybridization with the entire *iap* gene revealed the presence of homologous sequences in most other *Listeria* species. In contrast, a 400-base-pair internal *iap* probe which contained the whole repeat region hybridized only with genomic DNA from *L. monocytogenes*. Four oligonucleotides previously described as specific probes for the detection of *L. monocytogenes* (A. R. Datta, B. A. Wentz, D. Shook, and M. W. Trucksess, Appl. Environ. Microbiol. 54:2933–2937, 1988) were shown to be part of the *iap* gene.

Listeria monocytogenes is a gram-positive, facultative intracellular bacterium that can cause severe infections in humans and many animal species. Special risk groups are pregnant women, newborns, and immunocompromised patients. The major clinical symptoms of *L. monocytogenes* infections are abortion, septicemia, and meningitis (37). Recent outbreaks of listeriosis show a correlation between infection and the ingestion of food, mainly dairy products, contaminated with *L. monocytogenes* (10, 36). This observation and electron microscopic studies of tissues of infected guinea pigs (32) provided evidence that the gastrointestinal tract is an important route of infection and that the epithelial cells of the intestine may be the primary site of entry of these bacteria.

Recent data indicate that listeriolysin, an SH-activated cytolysin produced and secreted by all virulent *L. monocytogenes* strains, is an essential virulence factor. Nonhemolytic mutants have significantly reduced rates of survival in the mouse infection model (12, 19) and in mouse peritoneal macrophages (24). Nevertheless, these mutants are still able to enter nonprofessional phagocytic cells, such as the human colon carcinoma cell line Caco-2 (11) or mouse fibroblast 3T6 cells (24), by the induction of their own phagocytosis. Listeriolysin therefore does not seem to be involved in invasion.

In the attempt to elucidate the molecular mechanisms leading to invasion of *L. monocytogenes* and to identify the factors involved, interest has been focused in this laboratory on an extracellular protein with a molecular mass of 60 kilodaltons (kDa). This protein, termed p60, is produced in relatively large amounts by all virulent *L. monocytogenes* strains (14). Spontaneously occurring mutants of *L. monocytogenes* which are impaired in the synthesis of p60 have lost the ability to invade 3T6 cells. These mutants form long cell chains which disaggregate after treatment with partially purified p60, and invasiveness of the treated mutant cells is restored (23). This finding led to the conclusion that p60 may

be involved in the invasion of nonprofessional phagocytic cells by *L. monocytogenes* (23).

Here we report on the cloning and sequence analysis of the gene *iap*, which encodes this invasion-associated protein (p60) of *L. monocytogenes*. Hybridization experiments showed that DNA sequences homologous to *iap* are found in most *Listeria* strains tested but that an internal *iap* fragment which contains the information for an extended repeat region of p60 reacts specifically with chromosomal DNA of *L. monocytogenes* strains belonging to different serogroups.

MATERIALS AND METHODS

Bacterial strains. *L. monocytogenes* Sv1/2a Mackaness (SLCC 5764), *L. monocytogenes* Sv3a (SLCC 5015), *L. monocytogenes* Sv3b (SLCC 5543), *L. monocytogenes* Sv4b (SLCC 4013), *Listeria innocua* Sv4f (SLCC 5025), *L. innocua* Sv6a (NCTC 11288), *L. innocua* Sv6b, *Listeria ivanovii* (ATCC 19119), *Listeria grayi*, *Listeria welshimeri*, *Listeria seeligeri*, and *Listeria murrayi* were obtained from the *Listeria* strain collection of the Institute of Hygiene and Microbiology at the University of Würzburg, Würzburg, Federal Republic of Germany. *L. monocytogenes* Sv1/2a EGD was obtained from S. H. E. Kaufmann, University of Ulm, Ulm, Federal Republic of Germany. The nonhemolytic transposon Tn916 mutant *L. monocytogenes* M20 used in this study has been described previously (19). The isolation of the spontaneous rough variants *L. monocytogenes* RI and *L. monocytogenes* RII and RIV, derived from *L. monocytogenes* Mackaness and EGD, respectively, has been reported earlier (18). *L. monocytogenes* RIII, another rough mutant derived from a smooth strain of serovar 1/2a, was obtained from J. Potel (Institute for Medical Microbiology, Medical Academy, Hannover, Federal Republic of Germany).

Escherichia coli JM109 [*recA endA1 gyrA96 thi hsdR17 supE44 relA1* Δ(*lac-proAB*) (F' *traD36 proAB lacI*q ZΔM15)] was used in the cloning experiments with the plasmid vector pUC18 (41) and as a host for the bacteriophages M13mp18 and M13mp19.

Media and reagents. *Listeria* strains were cultured in brain

---

* Corresponding author.

heart infusion broth (Difco Laboratories, Detroit, Mich.) at 37°C. *E. coli* strains were grown in Luria-Bertani (LB) broth or 2× yeast-tryptone (YT) broth at 37°C, and ampicillin was added at a concentration of 50 μg/ml when required. Restriction enzymes, T4 DNA ligase, and nucleotides were purchased from Boehringer GmbH, Mannheim, Federal Republic of Germany. The enzymes were used according to the suggestions of the manufacturer. Horseradish peroxidase-conjugated swine anti-rabbit antibodies were supplied by DAKO Laboratories, Hamburg, Federal Republic of Germany. ^{32}P-dATP (800 Ci/mmol and 3,000 Ci/mmol) was purchased from Amersham International plc.

DNA isolation and manipulations. The procedures for isolating plasmid DNA (3) and chromosomal DNA (8) were as previously described. DNA fragments were purified by electroelution with a biotrap apparatus (Schleicher & Schuell, Dassel, Federal Republic of Germany) following digestion with the appropriate restriction endonucleases and agarose gel electrophoresis carried out in Tris-borate buffer (90 mM Tris base, 90 mM H_3BO_3, 2 mM EDTA, pH 8.0). The ligation of DNA and transformation of competent *E. coli* cells by the calcium chloride technique were performed according to standard protocols (30).

Southern blot analysis. Chromosomal DNA was digested with *Hin*dIII, fractionated on a 1% Tris-borate agarose gel for approximately 3 h at 150 V, and transferred to nitrocellulose filters by the method of Southern (38). DNA probes were labeled by the random priming technique (7). Hybridization was carried out in a solution containing 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), 5× Denhardt reagent, 0.1% sodium dodecyl sulfate (SDS), 20 mM sodium phosphate buffer (pH 6.6), 2 mM EDTA (pH 8.0), and 100 μg of denatured salmon sperm DNA per ml in 50% deionized formamide at 40°C for 20 h. The blots were then washed twice in 2× SSC–0.1% SDS at room temperature; this was followed by two 30-min washes in 0.2× SSC–0.1% SDS at 60°C under stringent conditions or in 3× SSC–0.2% SDS at 55°C under nonstringent conditions. Nitrocellulose filters were exposed to X-ray films (Fuji RX-NIF) for 4 to 16 h.

DNA sequencing and sequence analysis. The cloned DNA fragment encoding the *iap* gene was digested with the restriction endonucleases *Pst*I and *Hin*dIII. The resulting fragments were ligated with linearized replicative forms of M13mp18 and M13mp19 and transfected into competent *E. coli* JM109 cells. Recombinant phages were detected by adding transfected cells to 3 ml of 2× YT–soft agar (0.7% agar) containing *E. coli* JM109 from an overnight culture, isopropylthio-β-galactoside, and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside and incubating this mixture on 2× YT plates overnight at 37°C. β-Galactosidase-negative plaques were picked, and after growth in *E. coli* JM109 for 8 h, single-stranded M13 DNA was isolated (34). For nucleotide sequence determination of both DNA strands, the dideoxy-chain termination method of Sanger et al. (35) was applied by using ^{32}P-dATP and either the Klenow fragment of DNA polymerase I (Boehringer GmbH) or T7 DNA polymerase (Pharmacia, Uppsala, Sweden) according to the instructions of the manufacturer. Oligonucleotide primers other than the standard M13 primer were prepared with a DNA synthesizer (Applied Biosystems, Inc., Foster City, Calif.) and purified by preparative polyacrylamide gel electrophoresis on 20% gels. Extension fragments of the sequencing reactions were separated by electrophoresis in 6% polyacrylamide–42% urea gels and visualized after drying and exposure of the gels to X-ray film for 16 to 20 h.

The DNA sequence and the derived protein sequence were analyzed with a microcomputer system using the PC/Gene (IntelliGenetics Inc.) software. Protein sequence comparison was done on a VAX computer system using the Genetics Computer Group sequence analysis software package, version 6.1 (6).

Polyacrylamide gel electrophoresis and immunoblotting. Proteins from overnight culture supernatants (0.7 ml) of *L. monocytogenes* were precipitated with 7% trichloroacetic acid (final concentration) on ice for 1 h, washed once with acetone, suspended in 20 μl of Laemmli sample buffer (26), and heated to 95°C for 3 min. Cell lysates of *E. coli* recombinants were made by harvesting cells from 100 μl of overnight culture and washing them once in phosphate-buffered saline. The cell pellet was lysed in 20 μl of Laemmli sample buffer and boiled for 3 min before it was loaded onto a polyacrylamide gel. Protein separation was achieved by SDS-polyacrylamide gel electrophoresis in 12.5% polyacrylamide gels as described by Laemmli (26). The transfer of proteins onto nitrocellulose sheets was performed by semidry electroblotting in a graphite chamber (25). For the antibody reaction, anti-p60 antiserum was diluted 1:1,000. Antibody binding was visualized by subsequent incubations of the nitrocellulose with horseradish peroxidase-conjugated swine anti-rabbit immunoglobulins and with 4-chloro-1-naphthol (0.5 mg/ml) and hydrogen peroxide (0.025%).

Preparation of anti-p60 antiserum. A rabbit antiserum was raised against p60 purified from polyacrylamide gels. Briefly, culture supernatant proteins of the nonhemolytic *L. monocytogenes* M20 were separated by SDS-polyacrylamide gel electrophoresis and stained in a 3.5 M sodium acetate solution (15) to visualize protein bands, and the p60 band was excised from the gel. The protein was then purified by electroelution (16). For the first injection, p60 was emulsified in Freund complete adjuvant, while subsequent booster injections were prepared in water. The rabbit was injected subcutaneously at multiple sites with approximately 300 μg of pure p60 at 2-week intervals and was bled 4 weeks after injection 3.

Amino acid sequencing. The N-terminal amino acid sequence of the gel-purified p60 was determined as described previously (22).

RESULTS

Identification of an *E. coli* clone expressing p60 and subcloning of the *iap* gene encoding p60. Anti-p60 antiserum raised against purified p60 reacted specifically in a Western blot with p60 when total protein of wild-type *L. monocytogenes* was separated by SDS-polyacrylamide gel electrophoresis, blotted onto nitrocellulose filters, and developed with this antibody preparation. An identical reaction occurred in the listeriolysin-negative mutant M20 (Fig. 1A), indicating that p60 is not related to listeriolysin (58 kDa). For the isolation of the *iap* gene encoding p60, we screened a preselected library of chromosomal DNA from *L. monocytogenes* inserted into pUC18 consisting of 49 recombinant *E. coli* clones which express soluble antigens of *L. monocytogenes* EGD (20). Because of the predominant role of p60 as a major secreted protein, we expected the *iap* gene to be present in this group of clones. Colony immunoblotting of these preselected clones with anti-p60 antiserum identified one recombinant, named pLM37, which produced a protein that cross-reacted with the antiserum. Western blot analysis of total proteins of clone pLM37 confirmed that this protein had an apparent molecular mass of 60 kDa (data not shown).
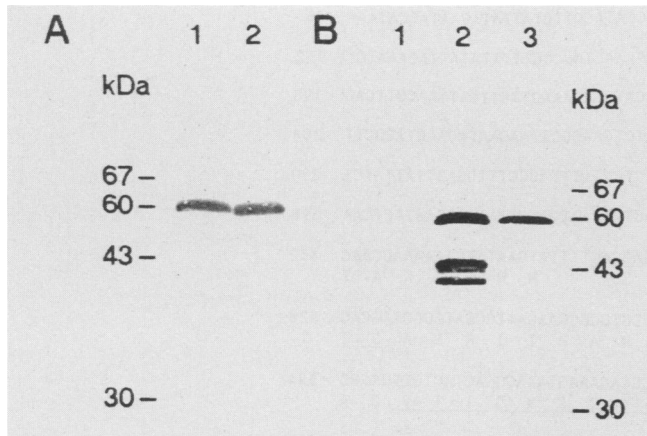
FIG. 1. Detection of p60 by immunoblot analysis with polyclonal anti-p60 antiserum. (A) Culture supernatant proteins of *L. monocytogenes*. Lanes: 1, strain Mackaness; 2, Hly⁻ M20. (B) Lanes: 1 and 2, cell lysate proteins of *E. coli* JM109 transformed with pUC18 and of the recombinant carrying plasmid pSK5, respectively; 3, culture supernatant proteins of *L. monocytogenes* EGD.

Restriction enzyme analysis of the pLM37 plasmid DNA indicated a listeria-specific DNA insert of 8.8 kilobases (kb). By subcloning, the localization and orientation of *iap* within the insert were determined. The subclones pSK2, pSK6, and pSK3 expressed truncated p60 derivatives. The molecular masses of these proteins, identified by immunoblotting with anti-p60 antiserum, increased with the size of the subcloned DNA fragment (Fig. 2A). The increase in the molecular masses of the truncated proteins expressed by pSK2, pSK6, and pSK3 (Fig. 2A) suggested that the 300-base-pair (bp)

*Pst*I-*Hind*III fragment and the adjacent 400-bp *Hind*III fragment are located within the *iap* gene. The recombinant plasmid pSK5 carried a 3.4-kb partial *Hind*III DNA fragment of pLM37 which seemed to encode the entire *iap* gene, since it expressed the full-length p60 in *E. coli* (Fig. 1B). Smaller cross-reacting proteins also visible in the immunoblot were degradation products of p60. Expression of the p60-derived proteins was not induced by isopropylthio-β-galactoside (data not shown) in any of the subclones, suggesting that *iap* was transcribed in *E. coli* from its own promoter and not as a *lacZ-iap* transcriptional fusion from the *lacZ* promoter of pUC18.

Nucleotide sequence of the *iap* gene. Based on the above-described localization of the *iap* gene, the determination of its nucleotide sequence was started at the *Hind*III site of the 300-bp *Pst*I-*Hind*III fragment. A stretch of 2,046 nucleotides extending 946 nucleotides to the left of the *Pst*I site and 459-bp to the right of the right *Hind*III site of the 400-bp *Hind*III-*Hind*III fragment (Fig. 2B) was determined. This DNA sequence revealed the presence of only one large open reading frame, extending from positions 443 to 1894 (Fig. 2B and 3). A purine-rich stretch upstream of the ATG initiation codon contained the sequence AGGAG (nucleotides 434 to 438). This sequence motif is complementary to the 3′ end of the 16S rRNA of *L. monocytogenes* (29) and may therefore represent the ribosome-binding site for the mRNA of *iap*. The open reading frame was terminated by a TAA stop codon. The inverted repeat followed by a short stretch of T downstream of the coding region (positions 1913 through 1945) may function as a *rho*-independent transcription terminator.

Analysis of the protein sequence and comparison with other proteins. The open reading frame coded for a protein con-
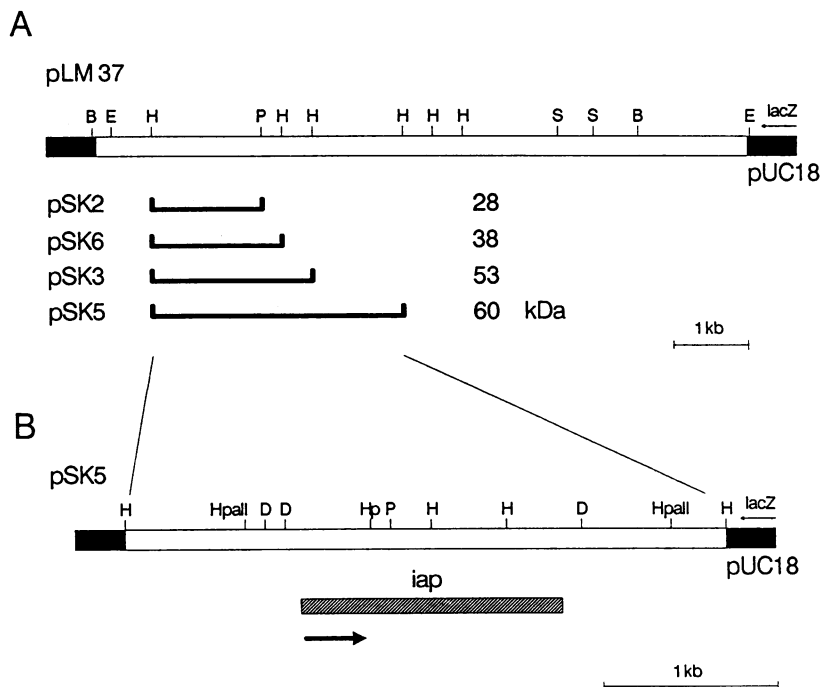

FIG. 2. Restriction map of plasmid pLM37 carrying the *iap* gene of *L. monocytogenes* EGD and subcloning of *iap* in plasmid pSK5 as a 3.4-kb partial *Hind*III fragment (A). In the *E. coli* subclones listed, proteins cross-reacting with anti-p60 antiserum ranged from 28 to 60 kDa in size, as determined by Western blot analysis (blot not shown). The exact location of *iap* in pSK5 (B) was deduced from sequence analysis data. The orientation of the *iap* gene is indicated by an arrow. Symbols: □, *Listeria* DNA; ■, vector DNA. Restriction sites indicated are *Bam*HI (B), *Dde*I (D), *Eco*RI (E), *Hind*III (H), *Hpa*I (Hp), *Hpa*II, *Pst*I (P), and *Sma*I (S).

```
   1  TCGATCATCATAATTCTGTCTCATTATATAACATCCTCCATACCTTCTATTATAGAATACCATAAA   66

  67  CTCATCTGGCAATTCATTTCGAGTCACGAAGAACGGAAAAACTGCCGGTTTTTATATTACAAATGT  132

 133  ATTAAGTTTTTCTATTAACAAAAAACAATAGGTTTCCCATAGCGAAAGTTGTTGATTAACGTTCAC  198

 199  ATCCCACTTACACTATAAAGGTTTACCCAGCAGTACATCTCAAGCCCTAAGAATACACGTTCGCTT  264

 265  TTCAACTGTTACAGAATTATTACAAATAGTTGGTATAGTCCTCTTTAGCCTTTGGAGTTATTATCT  330

 331  CATCATTTGTTTTTTAGGTGAAAACTGGGTAAACTTAGTATTAATCAATATAAAATTAATTCTCAA  396

 397  ATACTTAATTACGTACTGGGATTTTCTGAAAAAAGAGAGGAGTTTTATGAATATGAAAAAAGCAAC  462
                                        SD        M N M K K A T

 463  TATCGCGGCTACAGCTGGGATTGCGGTAACAGCATTTCGTGCGCCAACAATCCGATCCGCAAGCAC  528
       I A A T A G I A V T A F R A P T I R S A  S T

 529  TGTAGTAGTCGAAGCTGGTGATACTCTTTGGGGTATCGCACAAAGTAAAGGGACTACTGTTGACGC  594
        V V V E A G D T L W G I A Q S K G T T V  D A

 595  AATTAAAAAAGCAAACAATTTAACAACAGATAAAATCGTACCAGGTCAAAAATTACAAGTAAATAA  660
       I K K A N N L T T D K I V P G Q K L Q V N N

 661  TGAGGTTGCTGCTGCTGAAAAAACAGAGAAATCTGTTAGCGCAACTTGGTTAAACGTCCGTAGTGG  726
       E V A A A E K T E K S V S A T W L N V R S G

 727  CGCTGGTGTTGATAACAGTATTATTACGTCCATCAAAGGTGGAACAAAAGTAACTGTTGAAACAAC  792
       A G V D N S I I T S I K G G T K V T V E T T

 793  CGAATCTAACGGCTGGCACAAAATTACTTACAACGATGGAAAAACTGGTTTCGTTAACGGTAAATA  858
       E S N G W H K I T Y N D G K T G F V N G K Y

 859  CTTAACTGACAAAGCAGTAAGCACTCCAGTTGCACCAACACAAGAAGTGAAAAAAGAAACTACTAC  924
       L T D K A V S T P V A P T Q E V K K E T T T
                            Pst I                        ....................
 925  TCAACAAGCTGCACCTGCTGCAGAAACAAAAACTGAAGTAAAACAAACTACACAAGCAACTACACC  990
       Q Q A A P A A E T K T E V K Q T T Q A T T P
       ........                    ....................
 991  TGCGCCTAAAGTAGCAGAAACGAAAGAAACTCCAGTAGTAGATCAAAATGCTACTACACACGCTCT  1056
       A P K V A E T K E T P V V D Q N A T T H A V

1057  TAAAAGCGGTGACACTATTTGGGCTTTATCCGTAAAATACGGTGTTTCTGTTCAAGACATTATGTC  1122
       K S G D T I W A L S V K Y G V S V Q D I M S
                                 Hind III
1123  ATGGAATAATTTATCTTCTTCTTCTATTTATGTAGGTCAAAAGCTTGCTATTAAACAAACTGCTAA  1188
       W N N L S S S S I Y V G Q K L A I K Q T A N

1189  CACAGCTACTCCAAAAGCAGAAGTGAAAACGGAAGCTCCAGCAGCTGAAAAACAAGCAGCTCCAGT  1254
       T A T P K A E V K T E A P A A E K Q A A P V

1255  AGTTAAAGAAAATACTAACACAAATACTGCTACTACAGAGAAAAAAGAAACAGCAACGCAACAACA  1320
       V K E N T N T N T A T T E K K E T A T Q Q Q

1321  AACAGCACCTAAAGCACCAACAGAAGCTGCAAAACCAGCTCCTGCACCATCTACAAACACAAATGC  1386
       T A P K A P T E A A K P A P A P S T N T N A

1387  TAATAAAACAAATACAAATACAAATACAAATACAAATACAAACAATACTAATACAAATACACCATC  1452
       N K T N T N T N T N T N N T N T N T P S

1453  TAAAAAATACTAATACAAACTCAAATACTAATACGAATACAAACTCAAATACGAATGCTAATCAAGG  1518
       K N T N T N S N T N T N T N S N T N A N Q G

1519  TTCTTCCAACAATAACAGCAATTCAAGTGCAAGTGCTATTATTGCTGAAGCTCAAAAACACCTTGG  1584
       S S N N N S N S S A S A I I A E A Q K H L G
       Hind III
1585  AAAAGCTTATTCATGGGGTGGTAACGGACCAACTACATTTGATTGCTCTGGTTACACTAAATATGT  1650
       K A Y S W G G N G P T T F D C S G Y T K Y V

1651  ATTTGCTAAAGCGGGAATCTCCCTTCCACGTACTTCTGGCGCACAATACGCTAGCACTACAAGAAT  1716
       F A K A G I S L P R T S G A Q Y A S T T R I

1717  CTCTGAATCTCAAGCAAAACCTGGTGATTTAGTATTCTTTGACTATGGTAGCGGAATTTCTCACGT  1782
       S E S Q A K P G D L V F F D Y G S G I S H V

1783  TGGTATCTACGTTGGTAATGGTCAAATGATTAACGCGCAAGACAATGGCGTTAAATACGATAACAT  1848
       G I Y V G N G Q M I N A Q D N G V K Y D N I

1849  CCACGGCTCTGGCTGGGGTAAATATCTAGTTGGCTTCGGTCGCGTATAATTAATAACTTAAAGTAA  1914
       H G S G W G K Y L V G F G R V *          »»

1915  CCTGTGGAGCAAGCAGTTCGCTTCACAGGTTTTTTGTTGGAAATTTTATCTTAATGAAAGACGGTG  1980
       »»»»»»»»»»»         «««««««««««««

1981  TATGATGAAGAATCTAGTAAAAGTAAAAGTTTTCCTAAGTTCACAAAAGGCTATAAGGAAGGTTAT  2046
```

```
1373  ACAAAC ACAAAT GCTAAT AAA ACAAAT ACAAAT ACAAAT ACAAAT ACAAAT ACAAAC  1429
       T  N   T  N   A  N  K   T  N   T  N   T  N   T  N   T  N   T  N

1430  AAT ACTAAT ACAAAT ACA CCATCTAAA AAT ACTAAT ACAAAC TCAAAT ACTAAT      1483
      N   T  N   T  N   T   P  S  K   N   T  N   T  N   S  N   T  N

1484  ACGAAT ACAAAC TCAAAT ACGAAT                                          1507
       T  N   T  N   S  N   T  N
```

FIG. 4. Repeat region in the *iap* gene between sequence positions 1373 and 1507. Nucleotides forming the basic repeat unit ACAAAT were combined in groups of six and printed boldface. Corresponding repeat units on the protein level are underlined.

sisting of 484 amino acids (molecular mass, 50.7 kDa). Comparison of the amino acid sequence deduced from the DNA sequence with the experimentally determined N terminus of the extracellular p60 revealed that the N terminus of the mature protein is located at nucleotide 524. The 27 amino acids between the start codon and nucleotide 524 therefore function as a signal sequence directing the transport of p60 across the cell membrane. This putative signal sequence consists of a short, positively charged, hydrophilic N-terminal stretch followed by 22 mostly hydrophobic residues. Thus, the calculated molecular mass of the mature p60 is 47.5 kDa. This is considerably smaller than the molecular mass of p60 as determined by SDS-polyacrylamide gel electrophoresis (60 kDa). Analysis of the amino acid composition revealed an excess of positively charged residues (mainly lysine) equally distributed over the entire p60 sequence. This leads to a theoretical isoelectric point of 10.0, which is in good agreement with the experimentally determined isoelectric point of about 9.3 (23). The most striking region of p60 was a domain consisting of 19 Thr-Asn repeats from positions 1373 to 1507. This repeat domain was interrupted in three positions by short insertions encoding one or three amino acids. Serine replaced threonine in two positions (1472 and 1496) (Fig. 4). The hydropathy profile of protein p60 (Fig. 5A) indicates extended hydrophilic regions in the middle part of this protein, which included the repeat domain, whereas the N-terminal and particularly the C-terminal regions of p60 were more hydrophobic. Figure 5B indicates the prediction of secondary structural features of p60 (13). Alpha-helical regions were clustered in the middle part of p60, whereas the C-terminal region had a predominantly β-pleated sheet structure. The repeat domain of p60 had no obvious characteristic secondary structure, but it was flanked by two alpha-helical regions, which may allow this protein domain to become exposed on the protein surface.

The amino acid sequence of protein p60 deduced from the DNA sequence of the *iap* gene was compared with the protein sequences compiled in the NBRF protein database (release 21.0). No substantial sequence homology between p60 and other proteins was found. The extended repeat structure of p60 did not exhibit any similarity to the repeat regions of other proteins known to carry repeat domains within the protein sequence (21, 28, 39). Since p60 appears to be involved in the adhesion and uptake step of *L. monocytogenes* by nonprofessional phagocytic mammalian cells, we compared the p60 sequence with the known sequences of other proteins involved in invasion, i.e., Inv of *Yersinia pseudotuberculosis* (17) and IpaB and IpaC of *Shigella flexneri* (2). There was no substantial homology between the overall sequences or between a particular domain of p60 and these proteins.

Detection of sequences homologous to the *iap* gene in *Listeria* spp. In order to study the occurrence of the *iap* gene in *L. monocytogenes* strains and other *Listeria* species, we performed Southern hybridizations by using a 1.6-kb *Dde*I fragment which contained the entire *iap* gene of *L. monocytogenes* EGD (Fig. 2B). Strains of six different *Listeria* species were tested for homology to the cloned *iap* gene. Total DNA of these *Listeria* strains was isolated and digested with *Hin*dIII. The *iap* probe hybridized with DNA fragments of all strains except *L. grayi* (Fig. 6A). The hybridization pattern obtained with *L. monocytogenes* EGD, showing three hybridizing *Hin*dIII fragments of 1.8, 1.2, and 0.4 kb, is consistent with the restriction site data of the cloned gene (Fig. 2B). A different hybridization pattern was obtained with the *L. monocytogenes* strain belonging to serotype 4b. In that experiment, 1.2- and 2.3-kb fragments reacted with the probe. Two different hybridization patterns were also obtained with *L. innocua* strains. Strains of serotypes 6a and 6b hybridized to 7.5- and 2.0-kb DNA fragments, whereas *L. innocua* belonging to serotype 4f exhibited a hybridization pattern similar to the patterns of *L. monocytogenes* Sv4b and *L. welshimeri*. In *L. ivanovii*, 6.5- and 1.4-kb fragments showed homology with the *iap* probe. Two hybridizing fragments of 7.5 and 1.4 kb were observed with DNA from *L. seeligeri*.

Identification of an internal *iap* gene fragment specific for *L. monocytogenes*. Next, we used the internal 400-bp *Hin*dIII fragment, which included the Thr-Asn repeat region of *iap*, as a hybridization probe (positions 1164 to 1587 in Fig. 3). DNAs of *Listeria* strains, including those mentioned above, were again digested with *Hin*dIII. This probe proved to be highly specific for *L. monocytogenes* strains (Fig. 6B). No hybridization was observed with DNA of other *Listeria* species tested when the hybridization was carried out under stringent conditions. *Hin*dIII-cleaved DNA of all *L. monocytogenes* strains belonging to the serogroups 1/2a and 3a hybridized with the 0.4-kb fragment. The DNAs of serotypes 3b and 4b exhibited hybridization with the 2.3-kb *Hin*dIII fragment. Under less-stringent conditions, weak hybridization was also observed with the other *Listeria* species except *L. grayi* and *L. murrayi* (data not shown).

FIG. 3. Nucleotide sequence of the *iap* gene and its upstream and downstream regions. *Pst*I and *Hin*dIII sites present in the sequence are shown. The amino acid sequence deduced from *iap* is represented in the single-letter code. The putative ribosome-binding site SD is underlined, as well as the experimentally determined N terminus of the mature p60. The stop codon of *iap* is indicated by an asterisk, and the arrowheads downstream of the open reading frame mark the inverted repeats of a putative transcription terminator. Dotted lines above the nucleotide sequence indicate the locations of the four oligonucleotides described in reference 5.
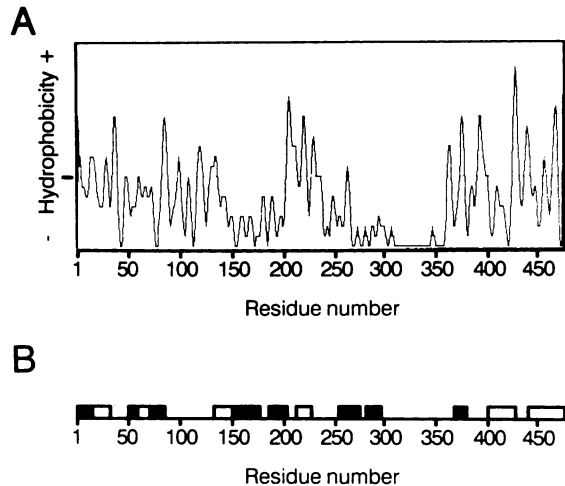
A

B

FIG. 5. (A) Hydropathy profile of the entire protein p60, by the method of Rose and Siddhartha (33). Values above the dotted line denote regions of relative hydrophobicity. (B) Secondary structure analysis of p60. Closed bars indicate alpha-helical structures, and open bars show regions with β-sheet properties. No characteristic secondary structure could be assigned to other parts of the protein.

In addition to the *Listeria* species, we tested a larger number of gram-positive and gram-negative bacteria, including *E. coli*, *Salmonella typhimurium*, *Staphylococcus aureus*, *Streptococcus pyogenes*, and *Bacillus subtilis*, for DNA homology to this *iap* gene probe. No hybridization was detected with any of these strains (data not shown).

The *iap* gene includes sequences previously described as probes for the detection of *L. monocytogenes*. A fragment of about 500 bp was previously described as a gene probe for the identification of *L. monocytogenes* by Datta et al. (4). From this fragment, four oligonucleotides were designed for the rapid detection of *L. monocytogenes* (5). We compared the reported sequences of these oligonucleotide probes with our *iap* sequence. All four oligonucleotides are part of the *iap* gene (Fig. 3, dotted lines). All four sequences are located upstream of the 400-bp *Hind*III fragment which we identified as a specific probe for *L. monocytogenes*.

DISCUSSION

The protein designated p60 was one of the major secreted proteins in all *L. monocytogenes* strains tested. Previous data indicated that this protein is involved in the invasion of nonprofessional phagocytic mammalian cells by *L. monocytogenes* (23). Members of most other *Listeria* species, all of which are nonpathogenic for humans, synthesize apparently similar proteins that cross-react with polyclonal antibodies raised against p60 of *L. monocytogenes*. As shown in this paper, the gene encoding p60, termed *iap* (gene for invasion-associated protein), also shares homology with the genomic DNAs of most other *Listeria* species, indicating a considerable conservation of this gene among members of the genus *Listeria*. Hybridization of the *iap* probe with DNAs of several other bacterial species was not detected.

The primary amino acid sequence of p60 deduced from the determined nucleotide sequence of the *iap* gene of *Listeria monocytogenes* shows some remarkable features. The hydrophilic protein is rich in positively charged lysine and arginine residues which are distributed over the entire polypeptide chain, which explains the experimentally determined high isoelectric point of p60 (23). The high positive charge of p60 may also explain the considerable discrepancy between the experimentally determined size (60 kDa by SDS-polyacrylamide gel electrophoresis) and the calculated molecular mass (50.7 kDa) (1). A posttranslational modification of the protein leading to this discrepancy cannot be ruled out rigorously, but there is no experimental evidence for such a possibility. The putative signal sequence for transport across the bacterial membrane consists of 27 amino acids. Typically, for such a sequence, two N-terminal amino acids are positively charged, whereas the 22 following amino acids are mostly hydrophobic. The cleavage site of the signal sequence as deduced from the N-terminal amino acid sequence of the mature p60 does not seem to follow the -3 -1 rule (40). We cannot rule out, however, the possibility that the first amino acid of the mature p60 has been cleaved off after removal of the signal sequence. In this case, the actual cleavage site of the signal peptide would be between alanine and serine, and serine and isoleucine would be in the -1 and -3 positions, respectively.

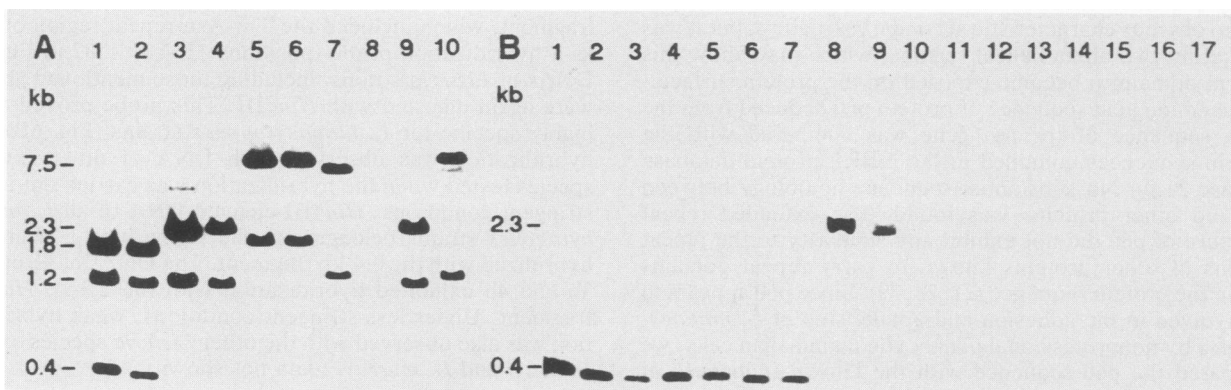The region between amino acids 311 and 355 consists of a

FIG. 6. Autoradiograph of a Southern blot of chromosomal DNA from different *Listeria* strains cleaved with *Hind*III and hybridized to the 1.6-kb *Dde*I fragment representing exactly the *iap* gene (A) and to the 400-bp *Hind*III fragment containing the repeat region of the gene (B). (A) Lanes: 1, *L. monocytogenes* Sv1/2a EGD: 2, *L. monocytogenes* Sv3a; 3, *L. monocytogenes* Sv4b; 4, *L. innocua* Sv4f; 5, *L. innocua* Sv6a; 6, *L. innocua* Sv6b; 7, *L. ivanovii*; 8, *L. grayi*; 9, *L. welshimeri*; 10, *L. seeligeri*. (B) Lanes: 1, *L. monocytogenes* Sv1/2a EGD; 2, *L. monocytogenes* Sv1/2a Mackaness; 3, *L. monocytogenes* RI; 4, *L. monocytogenes* RII; 5, *L. monocytogenes* RIII; 6, *L. monocytogenes* RIV; 7, *L. monocytogenes* Sv3a; 8, *L. monocytogenes* Sv3b; 9, *L. monocytogenes* Sv4b; 10, *L. innocua* Sv4f; 11, *L. innocua* Sv6a; 12, *L. innocua* Sv6b; 13, *L. ivanovii*; 14, *L. welshimeri*; 15, *L. grayi*; 16, *L. seeligeri*; 17, *L. murrayi*.

long repeat of the dipeptide Thr-Asn in which threonine is replaced by serine in two positions. This repeat occurs 19 times with three interruptions of one or three amino acids between repeats 2 and 3 and repeats 8 and 9 and within repeat 11. This repeat domain of p60 is apparently generated by tandem multiplication of the basic unit ACAAAT, which is interrupted in three positions by short insertions of 3 or 9 nucleotides. Interestingly, an internal *Hind*III fragment of the *iap* gene comprising the repeat region (r-*iap* probe) hybridized specifically with chromosomal DNA of *L. mono-cytogenes* but not with that of other *Listeria* species, whereas a *Dde*I fragment comprising the entire *iap* gene hybridized with all *Listeria* species tested except *L. grayi*. This indicates that part of the *iap* gene contains sequences (possibly the repeat region) which are specific for *L. mono-cytogenes*. This is in line with our previous observation that the p60-related proteins from *Listeria* species other than *L. monocytogenes* are unable to functionally replace p60 of *L. monocytogenes*; e.g., they cannot promote adhesion and penetration of these mutants into 3T6 fibroblast cells (M. Kuhn and W. Goebel, unpublished results).

A comparable repeat region is absent in the other sequenced proteins involved in cell invasion by intracellular bacteria, such as Inv of *Yersinia* spp. (17) or IpaB and IpaC of *Shigella* spp. (2). There are also no other substantial sequence similarities between these proteins and p60.

The cloning of a chromosomal fragment of *L. monocyto-genes* encoding a major secreted protein of *L. monocyto-genes* was recently described (9). A *Hind*III-*Hinc*II subfragment and four oligonucleotides derived from this sequence proved to be rather specific probes for *L. monocytogenes* (4, 5). We found that the reported sequences of the four oligonucleotide probes (5) are contained in our *iap* gene sequence, indicating that the fragment cloned by Flamm et al. (9) carries the *iap* gene. All four oligonucleotides are located 5' upstream of the r-*iap* fragment, which proved to be a specific probe for *L. monocytogenes*, in contrast to the entire *iap* gene. This finding suggests that regions of the *iap* gene outside of the r-*iap* fragment are also specific for *L. monocytogenes*.

The finding that the internal r-*iap* fragment is specific for *L. monocytogenes* adds to the number of DNA probes already available for the detection of this species. DNA probes encoding sequences from the listeriolysin gene (*hlyA*) are specific for *L. monocytogenes*, *L. ivanovii*, and *L. seeligeri* (27), while regions downstream of *hlyA* in *L. monocytogenes* are unique to this species (27). A probe from the *dth18* gene, which encodes a delayed-type hypersensitivity factor, detects all strains of *L. monocytogenes*, except those belonging to serovar 4a, and of *L. ivanovii* (31). More recently, we have found that sequences upstream of the *dth18* gene are unique to all *L. monocytogenes* strains, including those of serovar 4a (M. Leimeister-Wächter and T. Chakraborty, unpublished data). The availability of these DNA probes and the restriction fragment length polymorphism detected with the *iap* gene and the *hlyA* gene (27) may be useful in identifying and distinguishing virulent *L. monocytogenes* strains in epidemiological studies.

The nucleotide sequence data reported will appear in the EMBL, GenBank, and DDBJ nucleotide sequence data bases under the accession number X52268.

LITERATURE CITED

1. **Andrews, A. T.** 1986. Electrophoresis: theory, techniques and biochemical and clinical applications, p. 120–121. Oxford University Press, New York.
2. **Baudry, B., M. Kaczorek, and P. J. Sansonetti.** 1988. Nucleotide sequence of the invasion plasmid antigen B and C genes (ipaB and ipaC) of *Shigella flexneri*. Microb. Pathog. **4:**345–357.
3. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7:**1513–1523.
4. **Datta, A. R., B. A. Wentz, and W. E. Hill.** 1987. Detection of hemolytic *Listeria monocytogenes* by using DNA colony hybridization. Appl. Environ. Microbiol. **53:**2256–2259.
5. **Datta, A. R., B. A. Wentz, D. Shook, and M. W. Trucksess.** 1988. Synthetic oligodeoxyribonucleotide probes for detection of *Listeria monocytogenes*. Appl. Environ. Microbiol. **54:**2933–2937.
6. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
7. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specificity. Anal. Biochem. **132:**6–13.
8. **Flamm, R. K., D. J. Hinrichs, and M. F. Thomashow.** 1984. Introduction of pAMβ1 into *Listeria monocytogenes* by conjugation and homology between native *L. monocytogenes* plasmids. Infect. Immun. **44:**157–161.
9. **Flamm, R. K., D. J. Hinrichs, and M. F. Thomashow.** 1989. Cloning of a gene encoding a major secreted polypeptide of *Listeria monocytogenes* and its potential use as a species-specific probe. Appl. Environ. Microbiol. **55:**2251–2256.
10. **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 A. L. Reingold.** 1985. Pasteurized milk as a vehicle of infection in an outbreak of listeriosis. N. Engl. J. Med. **312:**404–407.
11. **Gaillard, J.-L., P. Berche, J. Mounier, S. Richard, and P. Sansonetti.** 1987. In vitro model of penetration and intracellular growth of *Listeria monocytogenes* in the human enterocyte-like cell line Caco-2. Infect. Immun. **55:**2822–2829.
12. **Gaillard, J. L., P. Berche, and P. Sansonetti.** 1986. Transposon mutagenesis as a tool to study the role of hemolysin in the virulence of *Listeria monocytogenes*. Infect. Immun. **52:**50–55.
13. **Garnier, J., D. J. Osguthorpe, and B. Robson.** 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. **120:**97–120.
14. **Goebel, W., S. Kathariou, M. Kuhn, Z. Sokolovic, J. Kreft, S. Köhler, D. Funke, T. Chakraborty, and M. Leimeister-Wächter.** 1988. Hemolysin from *Listeria*—biochemistry, genetics and function in pathogenesis. Infection **16**(Suppl. 2):149–156.
15. **Higgins, R. C., and M. Dahmus.** 1979. Rapid visualization of protein bands in preparative SDS-polyacrylamide gels. Anal. Biochem. **93:**257–260.
16. **Hunkapiller, M., E. Lujan, F. Ostrander, and L. Hood.** 1983. Isolation of microgram quantities of proteins from polyacrylamide gels for amino acid sequence analysis. Methods Enzymol. **91:**227–236.
17. **Isberg, R. R., D. L. Voorhis, and S. Falkow.** 1987. Identification of invasin: a protein that allows enteric bacteria to penetrate cultured mammalian cells. Cell **50:**769–778.
18. **Kathariou, S., J. Hacker, H. Hof, I. Then, W. Wagner, M. Kuhn, and W. Goebel.** 1987. Bacterial cytolysins—extracellular proteins and virulence factors, p. 141–150. *In* R. Rott and W. Goebel (ed.), Molecular basis of viral and microbial pathogenesis. Springer-Verlag KG, Berlin.
19. **Kathariou, S., P. Metz, H. Hof, and W. Goebel.** 1987. Tn*916*-induced mutations in the hemolysin determinant affecting viru-

lence of *Listeria monocytogenes*. J. Bacteriol. **169**:1291–1297.

20. **Kaufmann, S. H. E., M. M. Simon, and H. Hahn.** 1979. Specific Lyt 123 T cells are involved in protection against *Listeria monocytogenes* and in delayed-type hypersensitivity to listerial antigens. J. Exp. Med. **150**:1033–1038.

21. **Kochan, J., M. Perkins, and J. V. Ravetch.** 1986. A tandemly repeated sequence determines the binding domain for an erythrocyte receptor binding protein of *P. falciparum*. Cell **44**:689–696.

22. **Kreft, J., D. Funke, A. Haas, F. Lottspeich, and W. Goebel.** 1989. Production, purification and characterization of hemolysins from *Listeria ivanovii* and *Listeria monocytogenes* Sv 4b. FEMS Microbiol. Lett. **57**:197–202.

23. **Kuhn, M., and W. Goebel.** 1989. Identification of an extracellular protein of *Listeria monocytogenes* possibly involved in intracellular uptake by mammalian cells. Infect. Immun. **57**:55–61.

24. **Kuhn, M., S. Kathariou, and W. Goebel.** 1988. Hemolysin supports survival but not entry of the intracellular bacterium *Listeria monocytogenes*. Infect. Immun. **56**:79–82.

25. **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.

26. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.

27. **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.

28. **Ludwig, A., T. Jarchau, R. Benz, and W. Goebel.** 1988. The repeat domain of *Escherichia coli* haemolysin (HlyA) is responsible for its Ca^{2+}-dependent binding to erythrocytes. Mol. Gen. Genet. **214**:553–561.

29. **Ludwig, W., K. H. Schleifer, and E. Stackebrandt.** 1984. 16S rRNA analysis of *Listeria monocytogenes* and *Brochothrix thermosphacta*. FEMS Microbiol. Lett. **25**:199–204.

30. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

31. **Notermans, S., T. Chakraborty, M. Leimeister-Wächter, J. Dufrenne, K. J. Heuvelman, H. Maas, W. Jansen, K. Wernars, and P. Guinee.** 1989. Specific gene probe for detection of biotyped and serotyped *Listeria* strains. Appl. Environ. Microbiol. **55**:902–906.

32. **Racz, P., K. Tenner, and E. Mero.** 1972. Experimental *Listeria* enteritis. 1. An electron microscopic study of the epithelial phase in experimental *Listeria* infection. Lab. Invest. **26**:694–700.

33. **Rose, G. D., and R. Siddhartha.** 1980. Hydrophobic basis of packing in globular proteins. Proc. Natl. Acad. Sci. USA **77**:4643–4647.

34. **Sanger, F., A. R. Coulson, B. G. Bassell, A. J. H. Smith, and B. A. Roe.** 1980. Cloning in single stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. **143**:161–178.

35. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.

36. **Schlech, W. F., P. M. Lavigne, R. A. Bortolussi, A. C. Allen, E. V. Haldane, A. J. Wort, A. W. Hightower, S. E. Johnson, S. H. King, E. S. Nicholls, and C. V. Broome.** 1983. Epidemic listeriosis—evidence for transmission by food. N. Engl. J. Med. **308**:203–206.

37. **Seeliger, H. P. R.** 1961. Listeriosis. Hafner Press, New York.

38. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.

39. **Strathdee, C. A., and R. Y. C. Lo.** 1987. Extensive homology between the leukotoxin of *Pasteurella haemolytica* A1 and the alpha-hemolysin of *Escherichia coli*. Infect. Immun. **55**:3233–3236.

40. **Von Heijne, G.** 1984. How signal sequences maintain cleavage specificity. J. Mol. Biol. **173**:243–251.

41. **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.