INFECTION AND IMMUNITY, Apr. 1989, p. 1263–1270
0019-9567/89/041263-08$02.00/0

Vol. 57, No. 4

DNA Sequence of *mip*, a *Legionella pneumophila* Gene Associated with Macrophage Infectivity

N. CARY ENGLEBERG,[1,2,3]* CAROL CARTER,[4] DAVID R. WEBER,[2] NICHOLAS P. CIANCIOTTO,[3] AND BARRY I. EISENSTEIN[2,3]

Ann Arbor Veterans Administration Hospital,[1] and Departments of Internal Medicine[2] and Microbiology and Immunology,[3] University of Michigan Medical School, Ann Arbor, Michigan 48109, and Audie L. Murphy Memorial Veterans Hospital, San Antonio, Texas 78284[4]

Received 18 August 1988/Accepted 28 December 1988

In a previous study, a 24-kilodalton (kDa) protein surface antigen of *Legionella pneumophila* was cloned into *Escherichia coli* and found to be expressed on the host cell surface. Subsequently, a site-directed mutation in this gene (designated *mip*) in *L. pneumophila* was found to impair the capacity of this bacterium to initiate intracellular infection in human macrophages. The work presented here indicates that the antigenic gene product is distinct from the 24- to 29-kDa major outer membrane protein of *L. pneumophila*. In addition, the antigen was identified as a highly basic protein on two-dimensional nonequilibrium polyacrylamide gels and on two-dimensional monoclonal antibody immunoblots. When the DNA fragment encoding this protein was sequenced, a long open reading frame of 699 base pairs was identified within a region to which antigen expression was previously mapped. *mip* mRNA isolated from both *L. pneumophila* and transformed *E. coli* had the same 5′ end, as determined by primer extension analysis, indicating that the same promoter sequences are used in both species. A likely factor-independent transcriptional terminator was found 20 residues downstream of the stop codon, suggesting that *mip* is encoded on a monocistronic message. The inferred polypeptide began with a possible 20- to 24-residue signal sequence, and, as predicted by two-dimensional electrophoresis, had a molecular weight of 24,868 and was a potent polycation with an estimated pI of 9.8.

In previous reports we have described the identification and molecular cloning of a major, 24,000-dalton (Da) surface antigen of *Legionella pneumophila* (14, 35). This protein is expressed from a cloned fragment on a multicopy plasmid in *Escherichia coli* and is localized to the surface of the host *E. coli* strain (15). Recently, we exchanged a mutated copy of this gene for the wild-type gene in *L. pneumophila* to yield a mutant strain that produced no detectable 24-kDa antigen (8). Although this mutant replicated as well as its isogenic parent within macrophages, it was approximately 80-fold less infective for macrophages in vitro than its isogenic parent because of a defect in the initiation of intracellular infection. The gene responsible for this phenotype was designated *mip* (macrophage infectivity potentiator).

In this report we present an additional biochemical characterization of the 24-kDa antigen, including confirmation of its separate identity from the previously recognized major outer membrane protein (MOMP) with which it comigrates (5, 11, 16, 23, 35). We also present the DNA sequence of *mip* and the inferred amino acid sequence of the 24-kDa antigen. The analysis of the DNA sequence revealed transcriptional and secretory signals that can explain the expression and localization of the 24-kDa antigen in the *E. coli* host strain. The analysis also confirmed that the site-specific mutation described above is within the *mip* open reading frame. Analysis of the inferred amino acid sequence revealed several distinctive structural features of this protein that may be important clues to its mechanism of potentiating macrophage infection.

MATERIALS AND METHODS

Bacterial strains and plasmids. The gene encoding the 24-kDa antigen was cloned from *L. pneumophila* serogroup 1 (strain 130b), a clinical isolate from the Wadsworth Veterans Administration Hospital (West Los Angeles, Calif.) outbreak (20). *L. pneumophila* serogroup 5 strain (Dallas 1E) and strains representing several other serogroups were obtained from the Centers for Disease Control, Atlanta, Ga. *L. pneumophila* was grown either on buffered charcoal-yeast extract agar (10) or in yeast extract broth (36) as required for particular experiments.

pSMJ31.42 encodes the 24-kDa antigen on an 1,850-base-pair (bp) fragment of *L. pneumophila* DNA. It was derived from a larger plasmid in our original genomic library, pSMJ31 (14), by stepwise deletion of restriction fragments that did not affect 24-kDa antigen expression.

Antisera and antibodies. Antisera to *L. pneumophila* serogroup 5 were prepared in New Zealand rabbits by subcutaneous injection with 10^9 to 10^{10} CFU of live bacteria biweekly for 6 weeks. Monoclonal antibody (MAb) 12F4 was an immunoglobulin G1 that reacted with the 24-kDa antigen. To isolate this MAb, BALB/c mice were immunized with a crude outer membrane preparation from serogroup 1 and later boosted with killed, whole bacteria belonging to various other serogroups of *L. pneumophila*. This immunization schedule was used in order to minimize the immune response to the immunodominant, group-specific, nonprotein antigen of *L. pneumophila*.

Gel electrophoresis and immunoblotting. To prepare *L. pneumophila* lysates for electrophoresis, bacteria harvested from buffered charcoal-yeast extract agar were suspended in phosphate-buffered saline (PBS; pH 7.2) or sterile water at an optical density at 660 nm of 1.0.

One-dimensional sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) was used to demonstrate that the 24-kDa antigen was distinct from the comigrating MOMP by exploiting the tendency of the latter protein to remain insoluble in 2% SDS in the absence of a reducing

* Corresponding author.

agent. For this analysis, bacteria from 0.5 ml of the PBS suspension were pelleted by centrifugation; suspended in 2% SDS–67 mM Tris hydrochloride (pH 6.8), either with or without 5% 2-mercaptoethanol (2-ME); and boiled for 5 min. Boiled samples were centrifuged at 12,000 × *g* for 10 min, and the supernatants were designated as reduced and unreduced lysates, respectively. 2-ME was then added to the unreduced soluble protein sample to a final concentration of 5%, and the sample was reboiled before electrophoresis. The pellet from the unreduced sample was washed once in PBS and then boiled for 5 min in SDS buffer containing 5% 2-ME. After the final centrifugation, the supernatant was saved as the reduced pellet fraction. SDS-PAGE was conducted by using the discontinuous buffer system described by Laemmli (29).

Equilibrium and nonequilibrium two-dimensional gel electrophoresis (2DGE) was performed by the methods of O'Farrell and colleagues (33, 34). For both methods, the second dimension was a discontinuous SDS-PAGE system (4.5% stacking gel and 11.5% separating gel). Samples were prepared for 2DGE by pelleting cells from 1 ml of the PBS suspension and then boiling them for 4 min in 8 μl of lysing solution (2.6% SDS, 8.6% 2-ME, 4.3% glycerol, 50 mM Tris hydrochloride [pH 6.8]). The lysate was then diluted with 100 μl of buffer containing 9.5 M urea, 2% Nonidet P-40, 5% 2-ME, and 5% Ampholine (LKB Instruments, Inc., Rockville, Md.) mixture; and samples were either used immediately or quick-frozen in dry ice-ethanol and stored at −70°C. Gels were stained with silver by the method of Adams and Sammons (1).

Electrophoretic transfer of proteins from gels to nitrocellulose membranes was performed by the method of Towbin et al. (40). Antigens were visualized by using an enzyme immunoassay with a horseradish peroxidase-conjugated anti-immunoglobulin as the second antibody and 1-chloro-4-naphthol (Bio-Rad Laboratories, Richmond, Calif.) as a colorigenic substrate. In these immunoblots, all antibody incubations and washes were done in 50 mM Tris hydrochloride (pH 7.5) with either 500 mM NaCl (for assays with antisera) or 150 mM NaCl (for assays with MAb 12F4).

DNA sequencing. The dideoxy chain-termination method of sequencing was used with either *E. coli* DNA polymerase I (Klenow fragment; Bethesda Research Laboratories, Inc., Gaithersburg, Md.) or modified bacteriophage T7 DNA polymerase (Sequenase; U.S. Biochemical, Cleveland, Ohio). Initially, various restriction fragments of the pSMJ31.42 insert were cloned into bacteriophage M13 and sequenced from a universal primer. Additional M13 templates were generated by limited, unidirectional exonuclease III digestion of the inserts in the replicative form of certain, large M13 subclones (21). For selected sequences, double-stranded pSMJ31.42 was used as a template instead of M13, and sequencing was performed with synthetic oligonucleotide primers by the alkaline denaturation method (6).

Primer extension from total bacterial RNA. To map the 5′ end of *mip* mRNA, we used synthetic oligonucleotide primers to synthesize *mip* cDNA from purified bacterial RNA in the presence of a radiolabel. Total bacterial RNA was isolated from either *E. coli* or *L. pneumophila* by a modification of the method of Aiba et al. (2). Briefly, 100 ml of a mid- to late-log-phase broth culture was added to an equal volume of ice made with chloramphenicol (200 μg/ml). After centrifugation at 4°C, the bacteria were suspended in 3 ml of 20 mM sodium acetate (pH 5.5)–0.5% SDS–1 mM EDTA and then gently shaken at 60°C for 5 min with an equal volume of equilibrated phenol. The aqueous phase was separated and reextracted with phenol, and then the RNA was precipitated with 3 volumes of ethanol at −70°C for 30 min. The RNA pellet was recovered by centrifugation (17,500 × *g* for 15 min) and was washed three times in 70% ethanol–10 mM NaCl–1 mM Tris hydrochloride (pH 7.5). The final pellet was suspended in 1 ml of sterile water, and the RNA was quantified by UV spectroscopy.

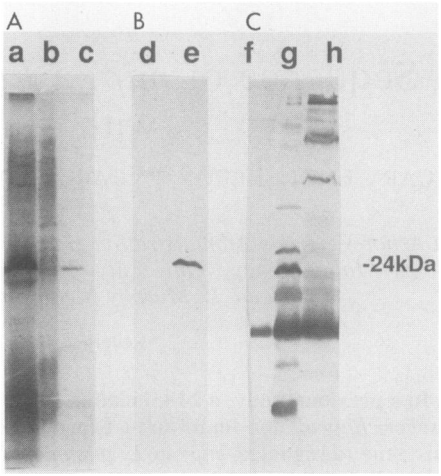

FIG. 1. Separation of the 24-kDa antigen and the MOMP by differential solubilization of *L. pneumophila* proteins and analysis by SDS-PAGE and immunoblotting. Gels were deliberately overloaded with samples prepared as described in the text. Lanes a and h, reduced lysate; lanes b, e, and g, unreduced lysate; lanes c, d, and f, reduced pellet fraction. (A) Lanes a, b, and c show proteins stained with Coomassie brilliant blue in a 15% polyacrylamide gel. (B and C) Lanes d and e and lanes f, g, and h are immunoblots with MAb 12F4 and serogroup 5 rabbit antisera, respectively.

For primer extension experiments, 30 to 40 μg of total RNA was hybridized with 25 ng of a specific oligonucleotide primer in 10 μl of 100 mM KCl–13 mM Tris hydrochloride (pH 8.3), first at 90°C for 1 min, then at 60°C for 2 min, and finally at 0°C for 30 min. cDNA was synthesized from one-half of this RNA-primer hybrid mixture with 1 U of reverse transcriptase (from avian myeloblastosis virus; Seikagaku America, St. Petersburg, Fla.) in a buffer of 50 mM Tris hydrochloride (pH 8.3)–40 mM KCl–6 mM magnesium chloride–2 mM dithiothreitol–200 μM of each deoxynucleotide (GTP, CTP, and TTP)–0.1 volume of [^{32}P]dATP (>400 Ci/mmol; Amersham Corp., Arlington Heights, Ill.)–0.4 U of RNasin (Promega Biotec, Madison, Wis.) per μl. After 10 min at 42°C, dATP was added to a final concentration of 200 mM and the incubation was continued for an additional 30 min. The reaction was stopped by the addition of 10 μl of deionized formamide, and the sample was boiled for 2 min prior to electrophoresis.

RESULTS

Separation of the 24-kDa antigen from the MOMP. The 24-kDa antigen could be separated from the MOMP by virtue of the solubility of the antigen in hot 2% SDS without reducing agents (Fig. 1, lanes e and g). Although some of the MOMP may also have been solubilized by this treatment, a considerable proportion of this protein remained bound to insoluble material after SDS extraction and was released only when the pellet was reboiled in SDS with 2-ME (Fig. 1, lane c). In contrast, none of the 24-kDa antigen was detected in this pellet fraction with either MAb 12F4 or heterologous

immune rabbit serum (Fig. 1, lanes d and f). The MOMP was not detected with rabbit antiserum, and its abundance in the reduced bacterial lysate apparently masked the reactivity of the 24-kDa antigen (Fig. 1, lane h). Results of these experiments confirm that the antigen and the MOMP that comigrated on SDS-polyacrylamide gels are, in fact, distinct proteins.

Nucleotide sequence of a DNA fragment that encodes the 24-kDa antigen. Both strands of the cloned insert on pSMJ31.42 were sequenced by the chain-termination method (Fig. 2). Only two potential open reading frames of greater than 200 bp in length were identified. One of these open reading frames was translated (Fig. 2); the other began at nucleotide (nt) 634 on the complementary strand, terminated in vector sequences 23 bp beyond the cloning junction, and encoded a polypeptide of no more than 21,803 daltons. In vitro transcription and translation of pSMJ31.42 showed that only one product, a single 24-kDa polypeptide, was produced in detectable amounts (data not shown).

Small deletions were introduced separately at the *Hin*dIII site within the first open reading frame (nt 1390) and at the *Ava*I site within the second open reading frame (nt 404) by digestion with the appropriate restriction endonuclease, limited digestion with S1 nuclease, and religation of the ends. Deletion of the *Hin*dIII site resulted in the loss of 24-kDa antigen expression from the plasmid, whereas deletion at the *Ava*I site had no effect on the expression of this antigen.

Characterization of a *mip* mutant by 2DGE. Expression of the 24-kDa antigen in *E. coli* was also abrogated when an insertion mutation was introduced at the unique *Hin*dIII site. When this mutated sequence was exchanged with the homologous sequences on the chromosome of wild-type *L. pneumophila*, the resulting recombinant strain failed to express the 24-kDa antigen and has been found to possess a defect in the initiation of intracellular infection in macrophages (8).

We compared lysates of the parent and isogenic mutant bacteria by 2DGE and found only one obvious difference in the pattern of silver-stained proteins. A large spot and a few smaller adjacent spots migrated with an M_r of ~24,000 at the basic end of the nonequilibrium gel profile of the parent strain proteins (Fig. 3B, arrow). This cluster of spots, which was produced by minor chemical modifications affecting the total charge of the protein, was absent from the profile of the mutant strain proteins (Fig. 3D, arrow). No other differences in the protein profiles of these two strains were detected by either equilibrium or nonequilibrium 2DGE that could not be accounted for by the expected gel-to-gel variation in migration and staining. The identity of the highly basic protein (and presumably one of its posttranslationally modified forms) was confirmed to be the 24-kDa antigen by immunoblot analysis of a nonequilibrium 2DGE of parent strain proteins run in parallel with the gel shown in Fig. 3B by using MAb 12F4 (Fig. 4). A comparable immunoblot of the companion gel of Fig. 3D yielded no detectable reactivity (data not shown). Based on these observations and an additional observation presented below, we concluded that the complete, long open reading frame translated in Fig. 2 is the coding sequence of the 24-kDa antigen and that an insertion mutation that interrupts this sequence is responsible for the *mip* mutation in *L. pneumophila*.

Analysis of *mip* transcriptional signals. To identify the *mip* promoter and to understand how *mip* is transcribed in *E. coli*, we identified the 5′ ends of *mip* mRNA isolated from both *L. pneumophila* and *E. coli* by synthesizing cDNA with synthetic oligonucleotide primers. Both primers were 18-mers that were complementary [(−)] to the sequence shown in Fig. 2; (−)nt 843 to 826 annealed to the putative initiation site on *mip* mRNA, and (−)nt 728 to 711 annealed 129 bp upstream of this site. Using the oligonucleotide (−)nt 843 to 826 as a primer, we observed identical cDNA bands synthesized from late-log-phase RNAs of *L. pneumophila* and *E. coli*(pSMJ31.42) (Fig. 5, lanes a and c). In contrast, no discrete band was seen in a control reaction done with *E. coli*(pBR322) (Fig. 5, lane b). By comparing these bands with a pSMJ31.42 sequencing reaction primed with the same oligonucleotide and loaded onto adjacent lanes on the gel, we mapped the 5′ end of the mRNA to the guanidine residue at nt 760. By contrast, identical experiments with primer (−)nt 728 to 711 yielded no detectable cDNA synthesis in any RNA sample (Fig. 5, lanes d, e, and f).

Accepting nt 760 as the transcriptional start site, the probable −10 and −35 promoter consensus sequences were identified and are indicated in boxes in Fig. 2. Although these sequences have only moderate homology with known *E. coli* promoters (49% homology by using the homology scoring system of Mulligan et al. [32]), the spacing of the two sequences is ideal with respect to one another and to the transcriptional start.

We identified a region of dyad symmetry (indicated by the broken underline in Fig. 2) 20 bp downstream from the *mip* termination codon that had features of a factor-independent transcriptional terminator (37). This sequence may form a stem-loop structure with a free energy of binding of −28.2 kcal (at 25°C), which was calculated by the method of Tinoco et al. (39). The theoretical stem-loop structure would have a GC-rich stem of 12 residues with no mismatches or loopouts, a hairpin loop of 5 residues (ATTTA), and a T-rich region immediately downstream of the stem with 6 T residues in a span of 9 bp. This region of dyad symmetry would also have been selected as a possible terminator by the dinucleotide frequency algorithm of Brendel and Trifonov (4).

Translational signals of *mip*. The longest possible open reading frame in the region to which *mip* mapped began with an ATG codon at nt 839 and ended with an ochre stop codon at nt 1540. We infer that translation begins at the first in-frame ATG (at nt 839), since this codon is associated with a putative ribosomal binding site and since it satisfies criteria for a translational start site in *E. coli* (38). Although there were four other in-frame ATG codons at the 5′ end of the open reading frame (and nt 845, nt 869, nt 881, and nt 893), none of these was situated in proximity to a possible ribosomal binding site. They are therefore presumed to encode methionine residues within the polypeptide sequence.

Codon usage within this gene was similar to that within *E. coli* genes (19). There was no extensive use of any particular tRNA species that is rare in *E. coli*.

Analysis of the polypeptide encoded by *mip*. The inferred gene product was a polypeptide of 233 amino acids and with a molecular weight of 24,868. There was a notable abundance of lysine residues in its amino acid composition (25 residues), and the pI predicted from the primary structure was 9.8. The molecular weight and pI inferred from the DNA sequence correlated well with the observed migration of this gene product by 2DGE (Fig. 3 and 4).

The first 20 to 24 residues of the inferred amino acid sequence suggested a possible secretory signal. The sequence began with a short series of charged amino acids, followed by a richly hydrophobic core and several possible cleavage sites, based on similarities with known sites in other secreted, bacterial proteins. Cleavage after the alanine

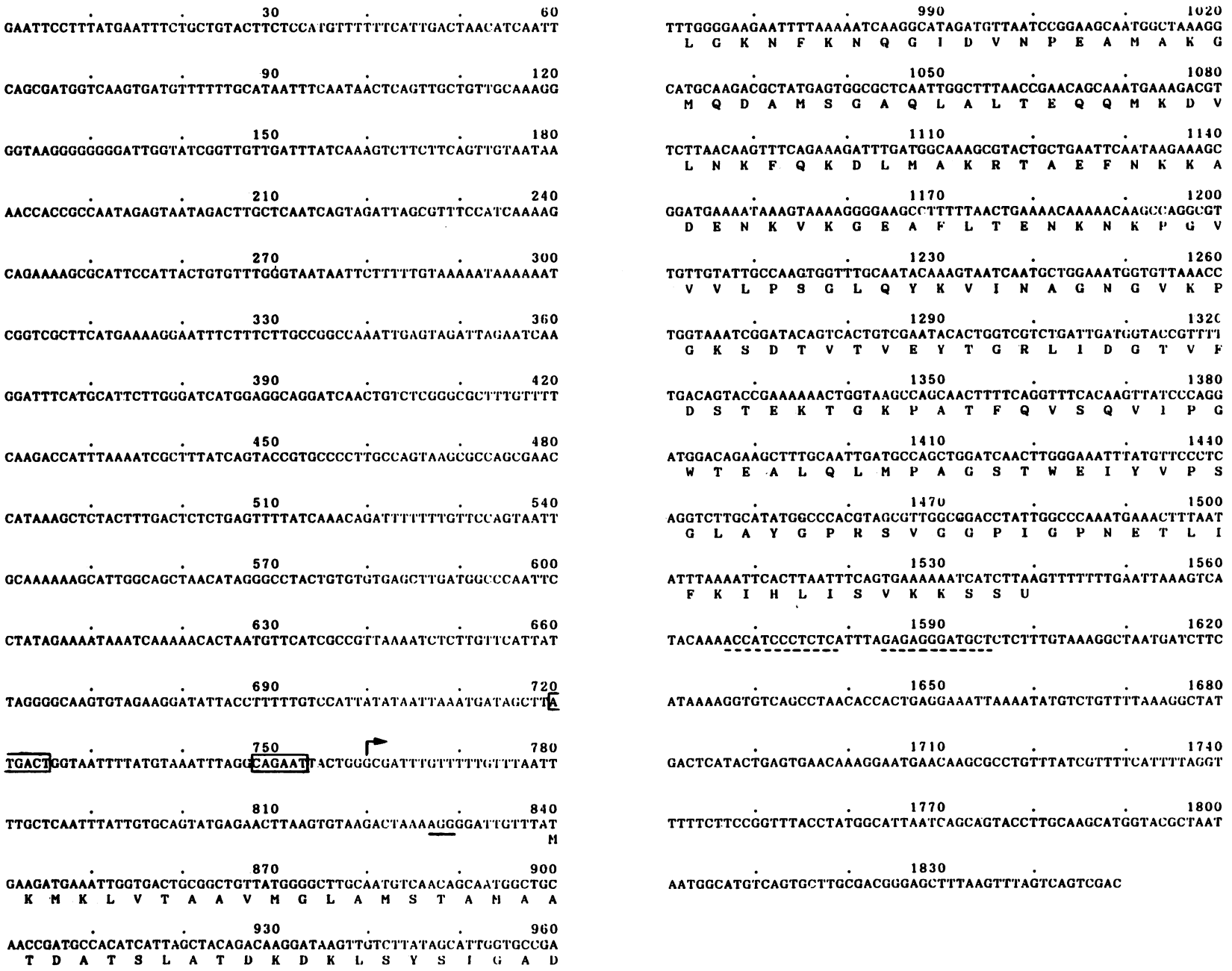

FIG. 2. DNA sequence of the 1,849-bp insert in pSMJ31.42. The sense strand sequence is shown, with the inferred translation product indicated in the single-letter amino acid code. The putative ribosomal binding site is underlined with a solid line. The transcriptional start site (determined by mRNA primer extension; see Fig. 5) is indicated by an arrow, and the −10 and −35 promoter sequences are indicated in boxes. A region of perfect dyad symmetry with features of a factor-independent transcriptional terminator is indicated by the broken underline immediately downstream of the *mip* open reading frame.

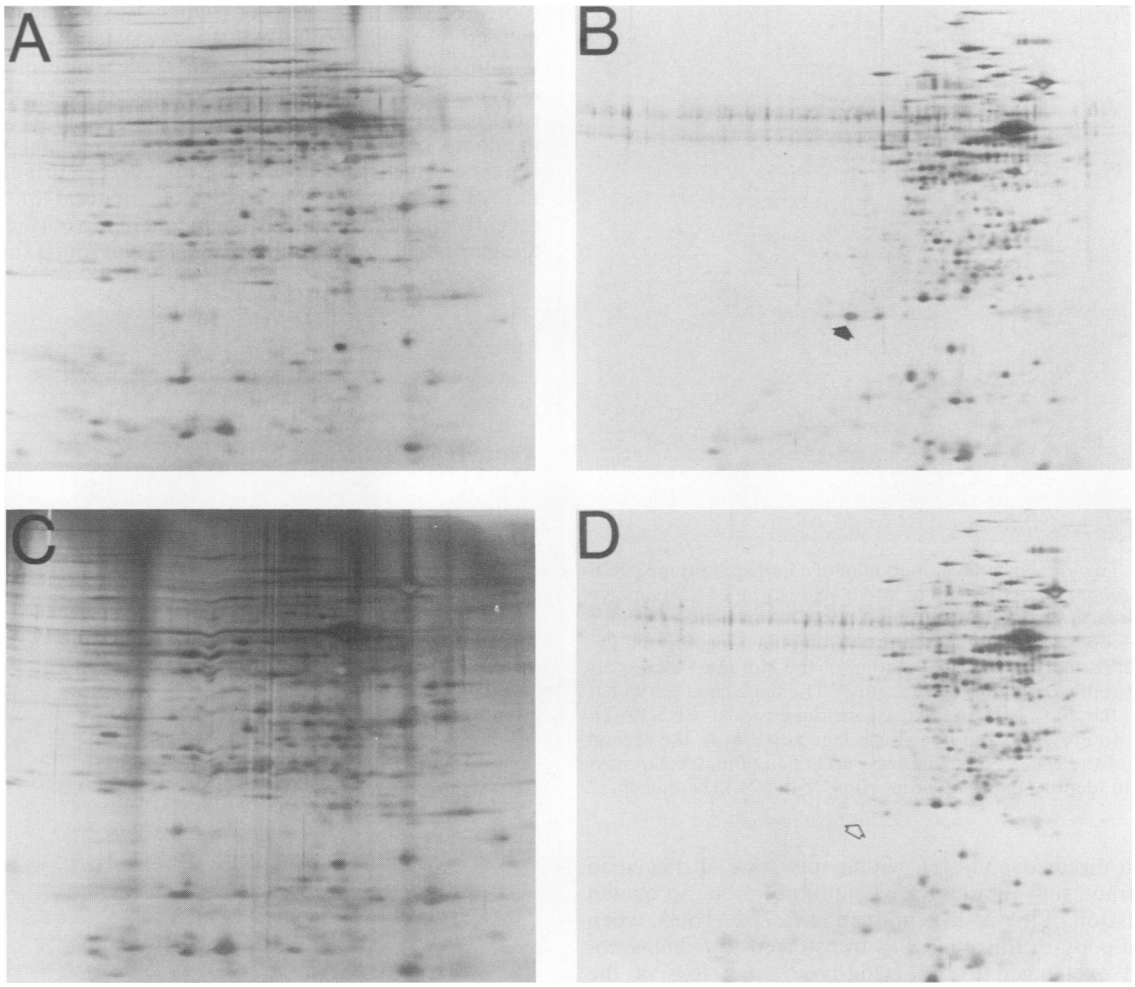

FIG. 3. 2DGE of *L. pneumophila* AA103 and the isogenic mutant strain AA105 that lacked the 24-kDa antigen. (A and C) Silver stains of bacterial proteins from strains AA103 (A) and AA105 (C) separated by equilibrium 2DGE. (B and D) Nonequilibrium 2DGE of strains AA103 (B) and AA105 (D). The arrows are explained in the text.

residues at positions 20 or 24 would most likely be favored (41).

An hydropathy plot of the polypeptide is shown in Fig. 6, along with a prediction of the secondary structure determined by the method of Garnier and co-workers (17). By these analyses, the initial 20 to 24 residues were predicted to be alpha-helical and strongly hydrophobic, consistent with the notion that they may represent a signal sequence. These analyses also suggest that the remainder of the polypeptide is divided into two large, distinct regions. Residues 21 through 114 are predicted to be rich in alpha helix and include an uninterrupted helical stretch of 60 residues (from residues 55 to 114). The remaining region (residues 115 through 233) is predicted to form beta sheets and turns. In addition, the hydropathy plot suggests that the first region is also more hydrophilic than the second region. Indeed, within the first region, 31% of the amino acids were charged and 24.5% were hydrophobic, whereas within the second region only 16.8% of the amino acids were charged and 32.8% were hydrophobic.

To determine whether the *mip* gene product might be a conserved analog of some previously characterized bacterial gene, we searched the Protein Sequence Database of the National Biomedical Research Foundation (June 1987) using the program FASTP (30). Three eucaryotic proteins and no bacterial proteins had homology scores outside of the normal distribution of random matching scores. We scrutinized the sequences of several bacterial proteins with homology scores at the upper end of the normal distribution and determined that none were similar to *mip*, based on the very limited extent and the position of the matched sequences and based on the comparative size and composition of the proteins relative to those of the *mip* gene product. We concluded that *mip* is not an analog of any previously sequenced bacterial protein.

DISCUSSION

In this study we sequenced a gene that is associated with a functional defect in the infectivity of *L. pneumophila* for human macrophages. Several lines of evidence support the conclusion that the open reading frame that we translated (Fig. 2) encodes *mip*. First, this gene was the only complete, long open reading frame in a cloned fragment of *L. pneumophila* DNA that expressed the 24-kDa antigen. The primer extension experiments (Fig. 5) proved that this gene is transcribed in vivo, and the sequence analysis suggested that it is most likely transcribed as a monocistronic message.

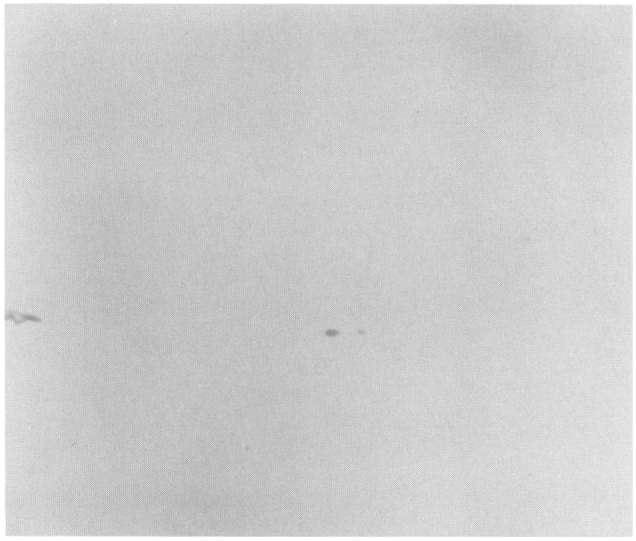

FIG. 4. Two-dimensional immunoblot of a nonequilibrium 2DGE of *L. pneumophila* AA103 with an MAb against the 24-kDa antigen (MAb 12F4). Note that the only two spots visualized are among those that were specifically lost in the *mip* mutant (Fig. 3B and D). The faint dark spots at the right margin of the blot are water spots and do not represent antibody reactivity. The dark band at the left margin of the blot is a standard, one-dimensional SDS-PAGE sample of strain AA103 run in a single lane adjacent to the second dimension of the 2DGE; this lane was run in lieu of molecular mass standards to identify the relative mobility of the 24-kDa antigen.

Second, a directed mutation within this gene, designed to create a frameshift mutation by deletion of 4 bp, abrogated the expression of the 24-kDa antigen in *E. coli*. Third, when a mutated copy of this gene was transferred to *L. pneumophila* and exchanged for the wild-type gene, loss of the 24-kDa protein antigen was demonstrated by 2DGE and immunoblotting. Fourth, the identified gene encoded a polypeptide of a size and total charge that was consistent with the migration of the antigen on two-dimensional gels.

We previously characterized the *L. pneumophila* 24-kDa antigen as a surface protein that induced strong antigenic reactivity in rabbits immunized with killed bacteria (35). Given the observation of others that the purified MOMP is also an important antigen (16), we inferred that these molecules are one and the same. With the isolation of an MAb to the 24-kDa protein, we were able to separate definitively these two proteins and prove that our initial inference was incorrect (Fig. 1). (A comparable situation was recently reported in an evaluation of MAbs that were originally thought to be directed at the gonococcal outer membrane protein III [3].) In the case of the 24-kDa protein, DNA sequencing added further confirmation to our conclusion since the absence of cysteine residues in the inferred sequence was inconsistent with the observed behavior of the MOMP under reducing conditions and with the published amino acid analysis of the purified MOMP (5).

When the gene encoding the 24-kDa antigen was cloned in *E. coli*, it was transcribed and translated and the gene product was translocated to the cell surface, as in the native state (13–15). The identification of a partial consensus promoter sequence 86 to 119 bp upstream of the initiation codon and a probable transcriptional stop signal 26 bp downstream of the termination codon suggests that the 24-kDa antigen is encoded on a monocistronic message. The 49% homology score for the *mip* promoter, which was calculated by the method of Mulligan et al. (32), is consistent with the notion that this promoter functions as a weak transcriptional initiator in *E. coli*; the mapping of the 5′ end of the mRNA isolated from an *E. coli* clone confirms this notion. Likewise, the compatibility of the putative secretory signal with signal sequences in *E. coli* probably explains the surface localization of this protein in our *E. coli* clones (42). We have recently performed amino acid sequencing on the Mip protein purified from *E. coli*(pSMJ31.42) and found that the first

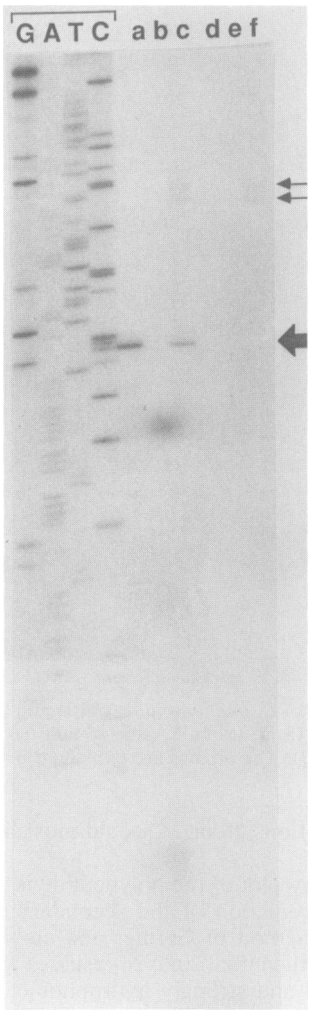

FIG. 5. Primer extension mapping of the 5′ end of the *mip* mRNA. Autoradiogram of a 7% polyacrylamide-urea sequencing gel is shown with a chain-termination sequencing reaction (lanes labeled G, A, T, and C) by using pSMJ31.42 as the template and oligonucleotide (−)nt 843 to 826 as the primer. cDNA was synthesized from total bacterial RNA in the presence of [^{32}P]dATP by using either primer (−)nt 843 to 826 (lanes a, b, and c) or primer (−)nt 728 to 711 (lanes d, e, and f). Template RNA was isolated from *E. coli* carrying pSMJ31.42 (lanes a and d) or pBR322 (lanes b and e) or from *L. pneumophila* AA103 (lanes c and f). The solid arrow indicates the 5′ end of the mRNA species isolated from 24-kDa protein-expressing bacteria corresponding to nt 760 on the DNA sense strand (Fig. 2). The smaller arrows indicate some minor incorporation of label into a naturally occurring *L. pneumophila* substrate for reverse transcriptase (lanes c and f); since the bands comigrated, this substrate is unrelated to the presence of the primers in the reaction but serve as an internal control confirming that transcriptase activity was present in both samples.

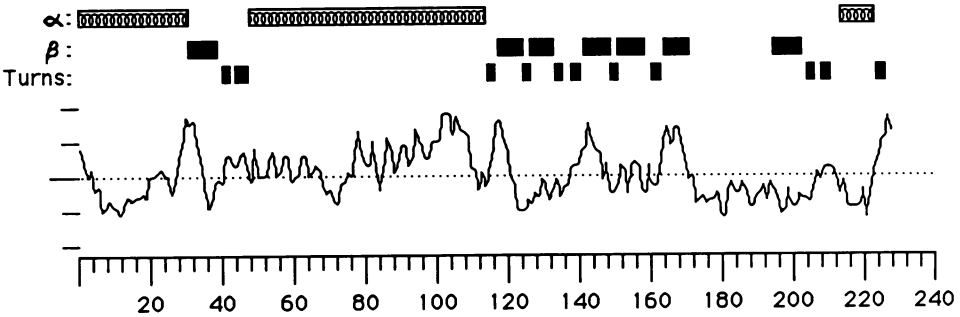

FIG. 6. Secondary structure and hydropathy prediction from the inferred 24-kDa polypeptide. The scale at the bottom corresponds to the polypeptide residue number. Above the scale is a hydropathy plot calculated by the method of Hopp and Woods (24) by using a window of six residues. The axis of the plot (on the left) indicates a range of hydrophilicity values from −2 to +2, and the dotted line through the plot indicates a value of 0. Above the hydropathy plot is a prediction of secondary structure that was made by the method of Garnier et al. (17). Similar results were obtained when the method of Chou and Fassman (7) was used (data not shown).

10 residues of its N-terminal sequence correspond with residues 21 to 30 of the inferred amino acid sequence shown in Fig. 2 (unpublished data). We are presently confirming this observation with the Mip protein purified from *L. pneumophila*.

Although we have not yet proven that the function impaired in the *mip* mutant is due to the direct action of the 24-kDa antigen, there are two circumstantial reasons why this is probably true. First, the phenomenology of intracellular infection with *L. pneumophila*, which has been described by Horwitz and colleagues (25–28), is different from that of other facultative, intracellular pathogens. Moreover, the molecular mechanisms of intracellular infection may differ markedly, even among closely related pathogens with similar intracellular life-styles (31). Therefore, it is reasonable to assume that the bacterial factors that mediate the novel phenomena seen during *L. pneumophila* infection may be unique. From results of DNA hybridization studies (12), immunoassays with monospecific antisera (13), and computer searches of large data banks, we found no similarity between the *mip* gene product and other known bacterial proteins, suggesting that it may be unique to the species *L. pneumophila* and associated with the production of a novel phenotype in that species alone. Second, the biochemical composition of the protein encoded by *mip* suggests a variety of possible functions relative to macrophage infection. For example, several proteins with extended alpha helices are known to fold into elongated structures (43). If the prediction of a 60-residue alpha helix in the N-terminal half of the 24-kDa protein is correct, then this protein may also form an elongated structure projecting from the bacterial surface. As a surface polycation, the protein could theoretically serve any of a number of important functions during infection. Polycations such as histones or polylysine are known to induce the phagocytosis of inert particles (9). Alternatively, following uptake of the bacterium by the phagocyte, this protein could be acting like known cationic lysosomotropic agents, such as ammonium chloride, to inhibit phagosome acidification and phagosome-lysosome fusion (18) or the protein could be interfering with lysosomal function in other ways (22, 44). Which, if any, of these theoretical functions is actually mediated by *mip* is not known, but the construction of an isogenic *mip* mutant (8) will make it possible to test each of these hypotheses in controlled experiments.

ACKNOWLEDGMENTS

We acknowledge Eric Pearlman and Doris Dixon for the isolation of the MAb used in this study. We also thank John Abraham, Brett Casey, Mark McClain, Jean Smith, and Vicki Vaughn for assistance in conducting this work.

N.C.E. is supported by Veterans' Administration Career Development and Merit Review Awards, and N.P.C. is supported by a National Research Service Award from the National Institutes of Health. We also acknowledge the support by Public Health Service grant RO1 AI24731 from the National Institutes of Health (to B.I.E.).

LITERATURE CITED

1. **Adams, L. D., and D. W. Sammons.** 1981. A unique silver staining procedure for color characterization of polypeptides, p. 155–165. *In* R. C. Allen and P. Arnaud (ed.), Electrophoresis '81. Walter deGruyter & Co., Berlin.
2. **Aiba, H., A. Sankar, and B. de Crombrugghe.** 1981. Evidence for two functional *gal* promoters in intact *Escherichia coli* cells. J. Biol. Chem. **256:**11905–11910.
3. **Barrera, O., and J. Swanson.** 1987. Antigens that are similar in apparent molecular weight to gonococcal outer membrane protein III. J. Clin. Microbiol. **25:**1155–1158.
4. **Brendel, V., and E. N. Trifonov.** 1984. A computer algorithm for testing potential prokaryotic terminators. Nucleic Acids Res. **12:**4411–4427.
5. **Butler, C. A., E. D. Street, T. P. Hatch, and P. S. Hoffman.** 1985. Disulfide-bonded outer membrane proteins in the genus *Legionella*. Infect. Immun. **48:**14–18.
6. **Chen, E. Y., and P. H. Seeburg.** 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4:**165–170.
7. **Chou, P. Y., and G. D. Fassman.** 1974. Prediction of protein conformation. Biochemistry **13:**222–245.
8. **Cianciotto, N. P., B. I. Eisenstein, C. H. Mody, G. B. Toews, and N. C. Engleberg.** 1989. A *Legionella pneumophila* gene encoding a species-specific surface protein potentiates initiation of intracellular infection. Infect. Immun. **57:**1255–1262.
9. **Deierkauf, F. A., H. Beukers, M. Deierkauf, and J. C. Riemersma.** 1977. Phagocytosis by rabbit polymorphonuclear leukocytes: the effect of albumin and polyamino acids on latex uptake. J. Cell. Physiol. **92:**169–176.
10. **Edelstein, P. H.** 1982. Comparative study of selective media for isolation of *Legionella pneumophila* from potable water. J. Clin. Microbiol. **16:**697–699.
11. **Ehret, W., and G. Ruckdeschel.** 1985. Membrane proteins of Legionellaceae. I. Membrane proteins of different strains and serogroups of *Legionella pneumophila*. Zentralbl. Bakteriol. Parasitenkd. Infektionskr. Hyg. Abt. 1 Orig. Reihe A **259:** 433–445.

12. **Eisenstein, B. I., and N. C. Engleberg.** 1986. Applied molecular genetics: new tools for the microbiologist and clinician. J. Infect. Dis. **153:**416–430.
13. **Engleberg, N. C., E. Pearlman, D. Dixon, and B. I. Eisenstein.** 1986. Antibodies isolated using cloned surface antigens recognize antigenically-related molecules in *Legionella pneumophila* and other *Legionella* species. J. Immunol. **136:**1415–1417.
14. **Engleberg, N. C., E. Pearlman, and B. I. Eisenstein.** 1984. Cloning and expression of *Legionella pneumophila* antigens in *Escherichia coli*. Infect. Immun. **44:**222–227.
15. **Engleberg, N. C., E. Pearlman, and B. I. Eisenstein.** 1984. *Legionella pneumophila* surface antigens cloned and expressed in *Escherichia coli* are translocated to the host cell surface and interface with specific anti-*Legionella* antibodies. J. Bacteriol. **160:**199–203.
16. **Gabay, J. E., and M. A. Horwitz.** 1985. Isolation and characterization of the cytoplasmic and outer membranes of the Legionnaires' disease bacterium (*Legionella pneumophila*). J. Exp. Med. **161:**409–422.
17. **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.
18. **Gordon, A. H., P. D'Arcy-Hart, and M. R. Young.** 1980. Ammonia inhibits phagosome-lysosome fusion in macrophages. Nature (London) **286:**79–80.
19. **Grantham, R., C. Gautier, M. Gouy, M. Jacobzone, and R. Mercier.** 1981. Codon catalog usage is a genome strategy modulated for gene expressivity. Nucleic Acids Res. **9:**r43–r74.
20. **Haley, C. E., M. L. Cohen, J. Halter, and R. D. Meyer.** 1979. Nosocomial Legionnaires' disease: a continuing common-source outbreak at Wadsworth Medical Center. Ann. Intern. Med. **90:**583–586.
21. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28:** 351–359.
22. **Hibbs, J. B., Jr., R. R. Taintor, and Z. Vavrin.** 1987. Macrophage cytotoxicity: role for L-arginine deiminase and imino nitrogen oxidation to nitrite. Science **235:**473–476.
23. **Hindahl, M. S., and B. H. Iglewski.** 1984. Isolation and characterization of the *Legionella pneumophila* outer membrane. J. Bacteriol. **159:**107–113.
24. **Hopp, T. P., and K. R. Woods.** 1981. Predictions of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA **73:**3824–3828.
25. **Horwitz, M. A.** 1983. Formation of novel phagosomes by the Legionnaires' disease bacterium in human monocytes. J. Exp. Med. **158:**1319–1331.
26. **Horwitz, M. A.** 1983. The Legionnaires' disease bacterium (*Legionella pneumophila*) inhibits phagosome-lysosome fusion in human monocytes. J. Exp. Med. **158:**2108–2126.
27. **Horwitz, M. A.** 1984. Phagocytosis of the Legionnaires' disease bacterium occurs by a novel mechanism: engulfment within a pseudopod coil. Cell **36:**27–33.
28. **Horwitz, M. A., and F. R. Maxfield.** 1984. *Legionella pneumophila* inhibits acidification of its phagosome in human monocytes. J. Cell Biol. **99:**1936–1943.
29. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
30. **Lipman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227:**1435–1441.
31. **Moulder, J. W.** 1985. Comparative biology of intracellular parasitism. Microbiol. Rev. **49:**298–337.
32. **Mulligan, M. E., D. K. Hawley, R. Entriken, and W. R. McClure.** 1984. *Escherichia coli* promoter sequences predict *in vitro* RNA polymerase selectivity. Nucleic Acids Res. **12:** 789–800.
33. **O'Farrell, P. H.** 1975. High resolution two-dimensional electrophoresis of proteins. J. Biol. Chem. **250:**4007–4021.
34. **O'Farrell, P. Z., H. M. Gooma, and P. H. O'Farrell.** 1977. High resolution two-dimensional electrophoresis of basic as well as acidic proteins. Cell **12:**1133–1142.
35. **Pearlman, E., N. C. Engleberg, and B. I. Eisenstein.** 1985. Identification of protein antigens of *Legionella pneumophila* serogroup 1. Infect. Immun. **47:**74–79.
36. **Ristroph, J. D., K. W. Hedlund, and R. G. Allen.** 1980. Liquid medium for growth of *Legionella pneumophila*. J. Clin. Microbiol. **11:**19–21.
37. **Rosenberg, M., and D. Court.** 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13:**319–353.
38. **Stormo, G. D., T. D. Schneider, and L. M. Gold.** 1982. Characterization of translational initiation sites in *Escherichia coli*. Nucleic Acids Res. **10:**2971–3011.
39. **Tinoco, I., P. N. Borer, B. Dengler, M. D. Levine, O. C. Uhlenbeck, D. M. Crothers, and J. Gralla.** 1973. Improved estimation of secondary structure of ribonucleic acids. Nature (London) **246:**40–41.
40. **Towbin, H., J. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76:**4350–4354.
41. **Von Henje, G.** 1984. How signal sequences maintain cleavage specificity. J. Mol. Biol. **173:**243–251.
42. **Von Henje, G.** 1985. Signal sequences; the limits of variation. J. Mol. Biol. **184:**99–105.
43. **Wilson, I. A., J. J. Skehel, and D. C. Wiley.** 1981. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3Å resolution. Nature (London) **289:**366–373.
44. **Zeya, H. I., and J. K. Spitznagel.** 1966. Antimicrobial specificity of leukocyte lysosomal cationic proteins. Science **154:**1049–1051.