Molecular Microbiology (1994) 13(2), 243–251

Cloning and molecular characterization of a *Legionella pneumophila* gene induced by intracellular infection and by various *in vitro* stress conditions

Yousef Abu Kwaik[1]* and N. Cary Engleberg[1,2]
[1]*Department of Microbiology and Immunology, and*
[2]*Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, Michigan 48109-0620, USA.*

Summary

The synthesis of a global stress protein (GspA) of *Legionella pneumophila* is induced in the intracellular environment of the phagocytic cell and by various *in vitro* stress stimuli. We used techniques of reverse genetics to isolate the *gspA* gene from a genomic library of *L. pneumophila*. Sequence analysis of approximately 1700 bp of a representative clone (pBSP1) showed the presence of two open reading frames (ORFs). ORF1 encoded for a polypeptide with an inferred molecular mass of 19 kDa and an isoelectric point of 6.1. These predictions correlated with the migration of the GspA protein on two-dimensional SDS–polyacrylamide gels. The predicted amino acid sequence of the GspA protein was identical to 22/23 residues of the *N*-terminal amino acid sequence derived by Edman degradation of the purified protein. The GspA protein was 41.3% and 36.5% identical to the 16 kDa IbpA and IbpB heat-shock proteins, respectively, of *Escherichia coli*. Primer extension from mRNA isolated from *L. pneumophila* showed that transcription of the *gspA* gene was controlled by two overlapping promoters. One of the promoters was a σ^{70} promoter, while the other was a heat-shock promoter and was regulated by the σ^{32} transcription factor in *E. coli*. Northern blot analysis showed that the level of *gspA* mRNA was elevated 3.4-, 5.0-, and 6.7-fold after exposure of *L. pneumophila* to heat shock, oxidative stress and osmotic shock, respectively. The *gspA* gene was conserved among 13 serogroups of *L. pneumophila*. Our data showed that the *gspA* gene of *L. pneumophila*, which is induced by intracellular infection and by various stress stimuli, is controlled transcriptionally by two overlapping and separately regulated promoters.

Received 22 February, 1994; revised 8 April, 1994; accepted 15 April, 1994. *For correspondence. Tel. (313) 936 0874; Fax (313) 764 3562.

Introduction

Legionella pneumophila is a ubiquitous organism in the aquatic environment and is a parasite of fresh water amoebae and other protozoa (Broome and Fraser, 1979; Dondero *et al.*, 1980; Morris *et al.*, 1979). Transmission of Legionnaires' pneumonia is thought to occur through inhalation of aerosolized organisms in contaminated water systems. The ability of *L. pneumophila* to cause pneumonia is dependent on its capacity to invade and to multiply within the host alveolar phagocytes (Payne and Horwitz, 1987; Horwitz, 1983a,b; 1984; Horwitz and Maxfield, 1984; Horwitz and Silverstein, 1980). The pathogenesis of the intracellular infection by *L. pneumophila* has been the subject of many recent reviews (Cianciotto *et al.*, 1989; Dowling *et al.*, 1992; Marra and Shuman, 1992).

Bacteria in general respond and adapt to environmental stimuli by co-ordinate regulation of gene expression (Miller *et al.*, 1989). Intracellular *Salmonella typhimurium* manifest a dramatic stress response to the intracellular environment (Abshire and Neidhardt, 1993; Buchmeier and Heffron, 1990). We previously showed that 35 bacterial proteins are induced and at least 32 proteins are repressed by *L. pneumophila* in response to the intracellular environment (Abu Kwaik *et al.*, 1993). Many of the macrophage-induced (MI) proteins of *L. pneumophila* are also induced in response to stress conditions *in vitro*. Despite the fact that intracellular *L. pneumophila* evade the microbicidal mechanisms of the phagocyte, our data clearly show a phenotypic stress response by intracellular bacteria. The role of stress-induced proteins in the intracellular infection is not known.

One of the MI proteins of *L. pneumophila*, with an apparent molecular mass of 19 kDa, is also induced by a variety of *in vitro* stress conditions (Abu Kwaik *et al.*, 1993). This protein was originally referred to as a 'universal stress protein' (UspA). In this paper, we describe the isolation and molecular characterization of the gene encoding this protein. We determined that this gene is unrelated to the *E. coli* gene (*uspA*) that was recently described as producing a 'universal stress protein' (Nyström and Neidhardt, 1992). Consequently, we have redesignated this *Legionella* gene as (*gspA*) or 'global stress protein' to avoid any confusion.

Results

Cloning of the gspA *gene*

A degenerate oligonucleotide corresponding to the *N*-terminal amino acid sequence of the GspA protein was used to probe a genomic library of *L. pneumophila*. Approximately 5000 clones were screened, and three clones hybridized to the probe. Restriction mapping and Southern hybridizations confirmed the origin of the *L. pneumophila* insert and showed that the clones were identical, each containing a 6.5 kb *Eco*RI insert (data not shown). A representative clone was designated pBSP1 and was chosen for further analysis (Fig. 1).

DNA sequencing and analysis

To sequence *gspA*, the degenerate oligonucleotide used in screening of the genomic library was the primer for the initial sequencing reaction using pBSP1 as a template. Subsequently, oligonucleotides were synthesized to sequence both strands. A DNA sequence of approximately 1700 bp was generated for both strands, and two open reading frames (ORFs) were identified (Fig. 2). The potential initiation codons of ORF1 and ORF2 are both preceded by a consensus Shine–Dalgarno sequence seven nucleotides upstream, strongly suggesting that these initiation codons represent translational initiation sites. We identified a region of perfect dyad symmetry (indicated by the broken lines in Fig. 2), located three nucleotides downstream of the termination codon of ORF1, that has features of a factor-independent transcriptional termination. The theoretical stem-loop structure would have a GC-rich stem of 13 nucleotides with no mismatches or loopouts and a hairpin loop of 12 AT-rich nucleotides. This potential stem-loop structure is immediately followed by an 11-nucleotide AT-rich region.

The first ORF encodes a polypeptide with a calculated size of 19 kDa and an isoelectric point (pI) of 6.1. The inferred size and pI of this putative gene product correspond to the observed migration of the GspA protein by two-dimensional SDS–PAGE (Abu Kwaik *et al.*, 1993). The amino acid sequence of the *N*-terminus of the gene product of ORF1 was identical to the amino acid sequence in 22/23 residues of the GspA protein derived by Edman degradation (Abu Kwaik *et al.*, 1993; data not shown). These findings confirm that ORF1 corresponds to the *gspA* gene product.

The amino acid sequence of the GspA protein does not contain a characteristic signal sequence, indicating that the protein is probably not secreted. A hydropathy profile of the GspA protein indicated that 64% of the protein is composed of polar and charged amino acid residues with no potential membrane-spanning domains, further indicating that the GspA protein is probably cytoplasmic (Kyte and Doolittle, 1982). A prediction of the secondary structure of the GspA protein by two algorithms showed that most of the *C*-terminus of the protein is composed of two α-helices (Kyte and Doolittle, 1982; Garnier *et al.*, 1978).

A homology search of GenBank using the FASTA program showed that the GspA protein was 41.3% and 36.5% identical to the two 16 kDa heat-shock proteins IbpA and IbpB, respectively, of *E. coli* (Fig. 3) (Allen *et al.*, 1992). Furthermore, the amino acid sequence of IbpA protein generates a similar hydropathy profile and similar secondary structure (i.e. two α-helices located in the *C*-terminus) (data not shown). These data indicated that the *gspA* gene is probably an analogue of the *E. coli ibpA*.

The predicted polypeptide encoded by *orf2* has a molecular mass of 23 kDa. There was no homology between the *orf2* product and other proteins in GenBank.

Mapping of gspA *transcriptional start sites*

To identify the *gspA* promoter(s) and to understand how the *gspA* gene is transcribed, we identified the 5′ ends of *gspA* mRNA by primer extension. In these experiments, two distinct cDNA species were synthesized from late-log-phase RNAs of *L. pneumophila* grown at 37°C, and no smaller transcripts were detected (Fig. 4, data not shown). By comparing these bands with a pBSP1 sequencing reaction primed with the same oligonucleotide primer, we mapped the 5′ ends of the two cDNAs to nucleotide 147 and nucleotide 159 (Figs 2 and 4). The integrity of the RNA was monitored by Northern blot hybridizations to the *gspA* and to the *mip* genes

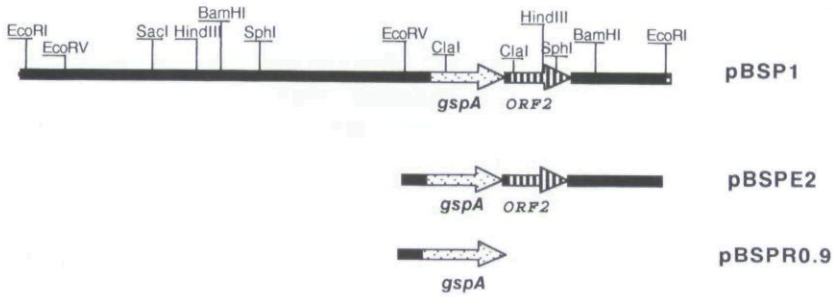

Fig. 1. Restriction map of the 6.5 kb *Eco*RI insert of pBSP1 and its subclones pBSPE2 and pBSPR0.9. The arrows indicate the location and direction of transcription of *gspA* and *orf2*.

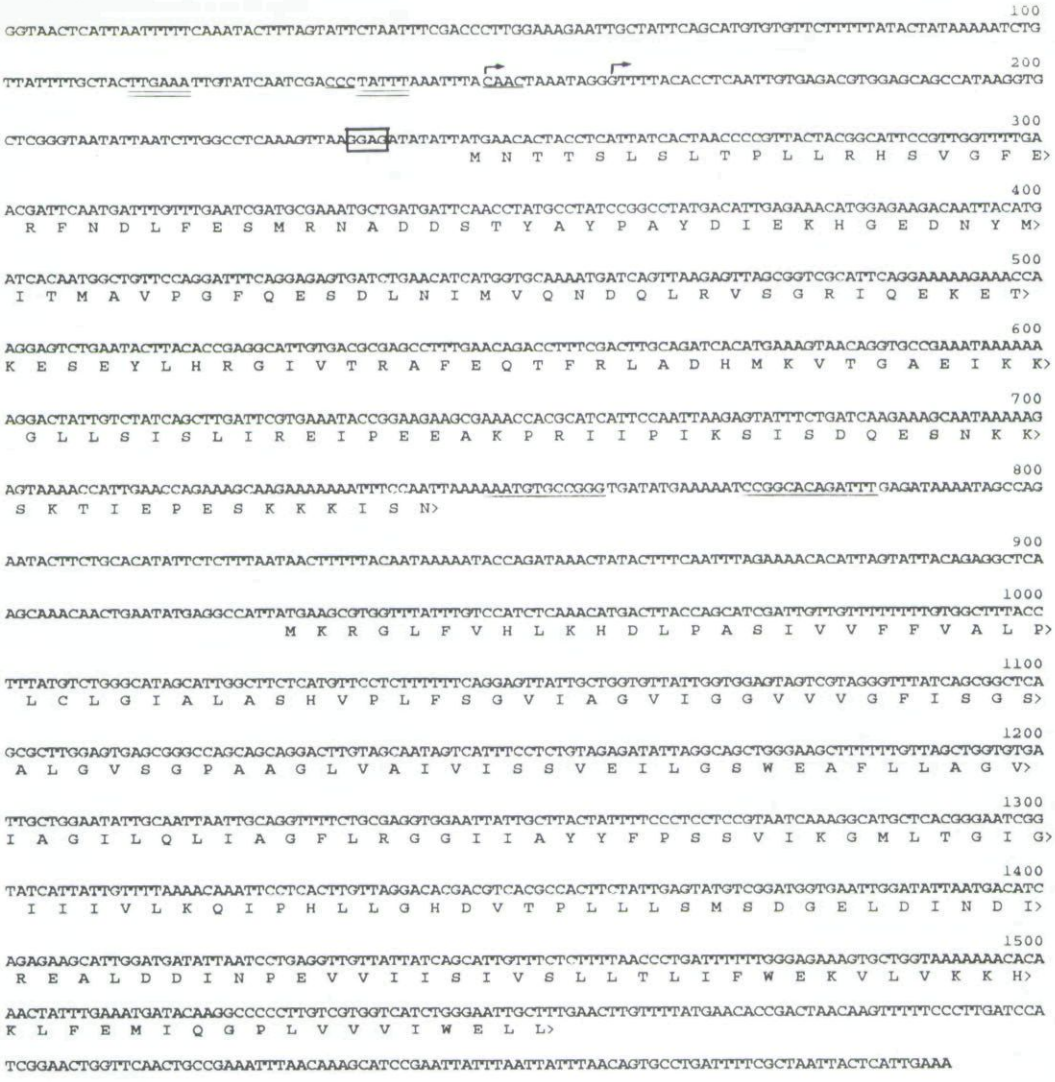

Fig. 2. Partial DNA sequence of pBSP1 insert. The sense-strand sequence is shown, with the inferred translation products indicated in the single-letter amino acid code. The putative ribosomal binding sites are indicated by boxes. The transcriptional start sites (determined by 5′-extension: see Figs 4 and 5) are indicated by arrows. The −10 and −35 regions of the σ^{32} and σ^{70} promoters are indicated by a single and double underline, respectively. A region of perfect dyad symmetry with features of a factor-independent transcription terminator is indicated by the broken underline immediately downstream of the *gspA* gene.

(Engleberg *et al.*, 1989), which showed that the RNA was intact (see Fig. 7 below).

Analysis of the DNA sequence upstream from the start sites suggested the presence of two overlapping promoters, P1 and P2. The probable −10 and −35 promoter consensus sequences are indicated in Fig. 2 by single and double underlining for P2 and P1, respectively. The sequences of the −10 and −35 regions of P2 are homologous to the consensus heat-shock promoter sequences of *E. coli*, suggesting that P2 is a heat-shock promoter. Upon heat shock of *L. pneumophila* (42°C), transcription of *gspA* was predominantly controlled by the P2 heat-shock promoter, while transcription through the P1 promoter was predominant at 28°C (Fig. 5). These data show that transcription of *gspA* is controlled by two overlapping promoters, which are differentially regulated.

Role of σ^{32} in transcription of gspA

To confirm that P2 is a heat-shock promoter and is regulated by the σ^{32} transcription factor, we repeated the primer-extension studies in an *E. coli* σ^{32} mutant and an isogenic wild-type strain harbouring pBSP1. Since the σ^{32} mutant is temperature sensitive, the RNAs were isolated from both strains grown at 30°C. A band representing a P2 transcript was clearly identical in the wild-type *E. coli*, but was absent in the σ^{32} mutant. Neither strain expressed a P1 transcript. However, both strains

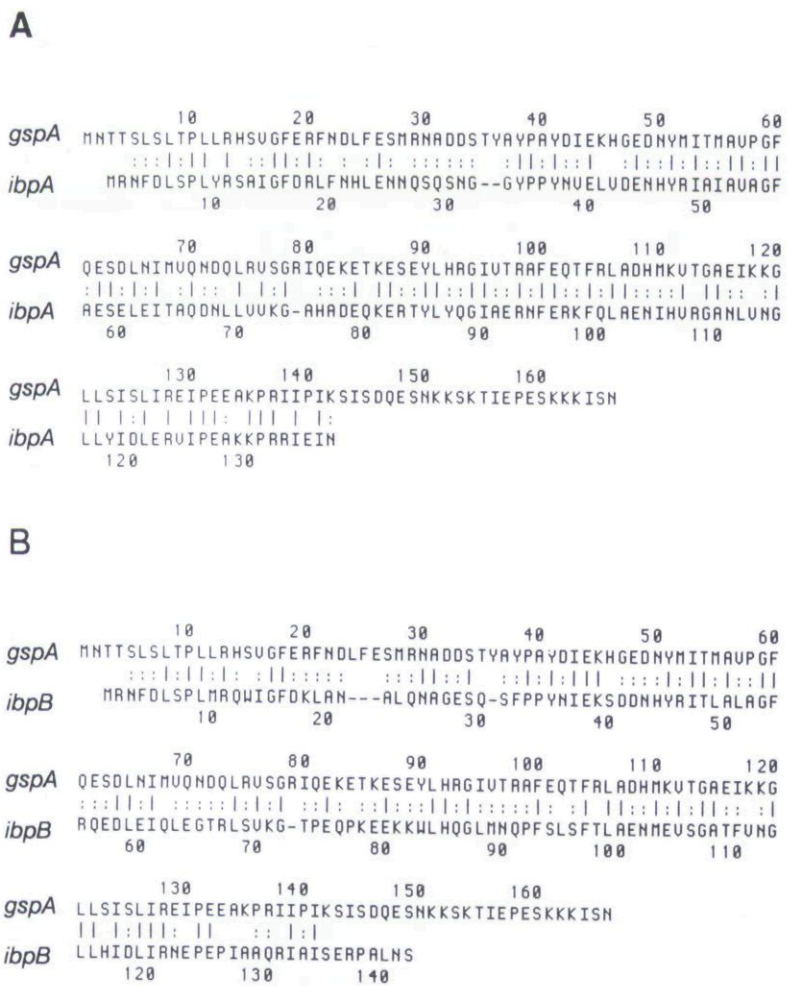

Fig. 3. Alignments of the inferred amino acid sequence of the *L. pneumophila* 19 kDa GspA protein with *E. coli* IbpA (A) and IbpB (B). I, indicates identity; :, indicates a conservative substitution.

expressed an abundant smaller transcript mapped to nucleotide 229, which does not correspond to a *L. pneumophila* transcript (Figs 2 and 6). The source of this band is unclear; it may represent a transcript from a cryptic promoter recognized by the *E. coli* transcription factors. These data confirmed that the σ^{32} heat-shock transcription factor was required for the expression from the P2 promoter of *gspA* in *E. coli*.

Induction of gspA *transcription in response to stress stimuli*

To measure the level of *gspA* mRNA in response to *in vitro* stress stimuli, Northern hybridizations were performed on total RNAs isolated from *L. pneumophila* after exposure to heat shock, osmotic shock and oxidative stress. The blot was sequentially probed with the *gspA* and *mip* genes, and the amount of *gspA* mRNA was expressed relative to *mip* mRNA as determined by densitometric analysis (see the *Experimental procedures*). There was no change in the amount of *mip* mRNA after exposure of the bacteria to stress stimuli (Fig. 7B). This is consistent with a previous observation that the expression of *mip* is not altered following exposure of *L. pneumophila* to various *in vitro* stress stimuli such as heat shock, osmotic shock and oxidative stress (N. Cianciotto, personal communication). Compared to levels of *mip* mRNA, there was a 3.4-, 5.0-, and 6.7-fold increase in the levels of *gspA* mRNA after exposure of *L. pneumophila* to heat shock, oxidative stress and osmotic shock, respectively (Fig. 7A). There was a minimal difference in the level of *gspA* mRNA isolated from *L. pneumophila* grown at 30°C or 37°C.

In vivo *and* in vitro *expression of* gspA

The protein profile of whole-cell lysates of *E. coli* harbouring pBluescript, pBSP1, pBSPE2 or pBSPR0.9 was examined by SDS–PAGE. A 19 kDa band was expressed by *E. coli* harbouring pBSP1, pBSPE2 and pBSPR0.9 but not the vector (Fig. 8A, data not shown). These data were further confirmed using an *in vitro* transcription–translation system (Fig. 8B). These data showed that the *gspA* gene contained in pBSPR0.9 was expressed in *E. coli*.

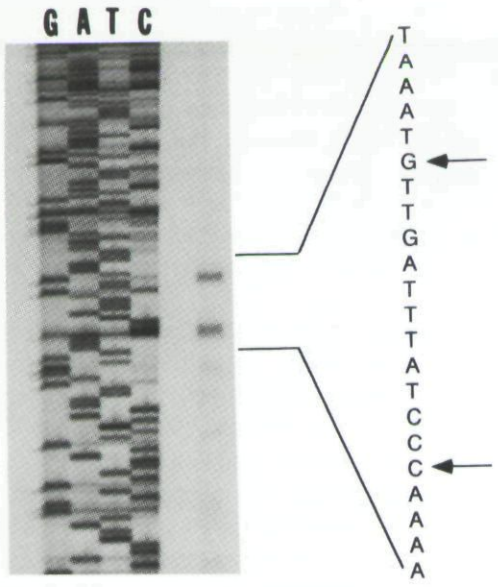

Fig. 4. 5′-Primer extension analysis of *gspA* transcripts. A radiolabelled oligonucleotide was annealed to the RNA isolated from *L. pneumophila* grown at 37°C, and reverse transcriptase was added to produce cDNA (see the *Experimental procedures*). The same oligonucleotide was used to prime the dideoxysequencing reaction of pBSP1 DNA. The letters above each lane indicate the dideoxy nucleotide used to terminate each reaction. The arrows indicate the 5′ end of the two mRNA species with the top arrow corresponding to nucleotide 147, while the bottom arrow corresponds to nucleotide 159 of the sequence shown in Fig. 2. No smaller transcripts were detected and, therefore, only the relevant portion of the gel is shown.

Conservation of the gspA *gene in* L. pneumophila

Although clinical isolates of *L. pneumophila* most often belong to serogroup 1 (SG1), strains of all the other 13 serogroups have been associated with cases of Legionnaires' disease. To determine whether the *gspA* gene is conserved in these other pathogenic strains, genomic DNAs were isolated from these strains, digested with *Eco*RI, and analysed by Southern hybridization using the pBSPR0.9 insert, which contains the *gspA* gene, as a probe. Under high-stringency hybridizations, the probe hybridized to single DNA fragments ranging in size from 6.5 to 9.5 kb from all of the 13 serogroups tested (data not shown). These data showed that the DNA encoding the *gspA* gene was conserved in all of the serogroups of pathogenic *L. pneumophila*, although restriction-fragment polymorphism was common in the DNA flanking *gspA*.

Discussion

We previously showed that intracellular *L. pneumophila* manifest a stress response during active growth within phagocytic cells with the induction in the synthesis of many proteins (Abu Kwaik *et al.*, 1993). The synthesis of a 19 kDa protein of *L. pneumophila*, designated GspA, was induced in response to the intracellular environment of the phagocytic cell and upon exposure of the bacteria to several stress stimuli *in vitro*. To define the role of the MI proteins in the intracellular infection, we used techniques of reverse genetics to isolate the *gspA* gene from a genomic library of *L. pneumophila*.

The *gspA* gene is regulated by two overlapping promoters, one of which is a heat-shock promoter. Although all the major heat-shock genes (i.e. *groEL*, *dnaK* and *grpE*) are regulated by a single heat-shock promoter, some heat-shock genes are regulated by two or more promoters (Missiakas *et al.*, 1993; Hemmingsen *et al.*, 1988; Saito and Uchida, 1978). The obvious advantage for *gspA* and other genes having multiple promoters is that each promoter can be regulated differently. This would allow the bacteria to control the expression through either promoter exclusively upon response and adaptation to certain environmental signals.

The 19 kDa GspA protein is highly homologous to the 16 kDa IbpA and IbpB heat-shock proteins of *E. coli* (Allen *et al.*, 1992). These two *E. coli* heat-shock proteins, of unknown function, are genetically and antigenically very similar and are probably the result of a gene duplication event. The synthesis of IbpA and IbpB proteins is induced by heat shock and in response to the

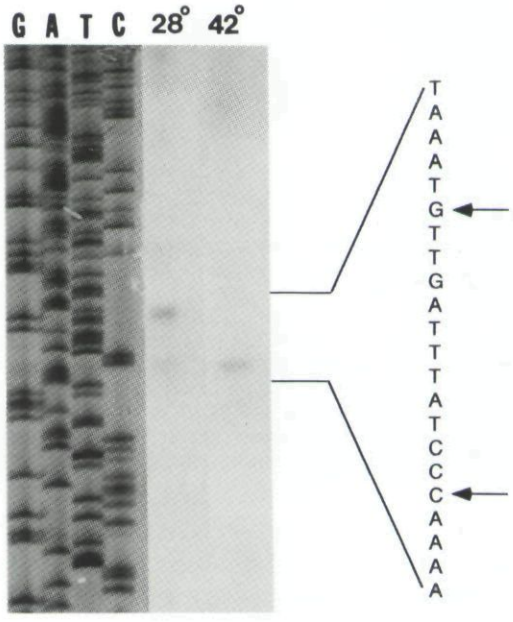

Fig. 5. 5′-Primer extension analysis of *gspA* transcripts following heat shock. A radiolabelled oligonucleotide was annealed to the RNA isolated from *L. pneumophila* grown at 28°C or after a heat shock at 42°C, and reverse transcriptase was added to produce cDNA. The oligonucleotide used for primer extension and sequencing of pBSP1 DNA is described in the legend to Fig. 4. The arrows indicate the 5′ end of the two mRNA species, as in the legend to Fig. 4. No smaller transcripts were detected and, therefore, only the relevant portion of the gel is shown.

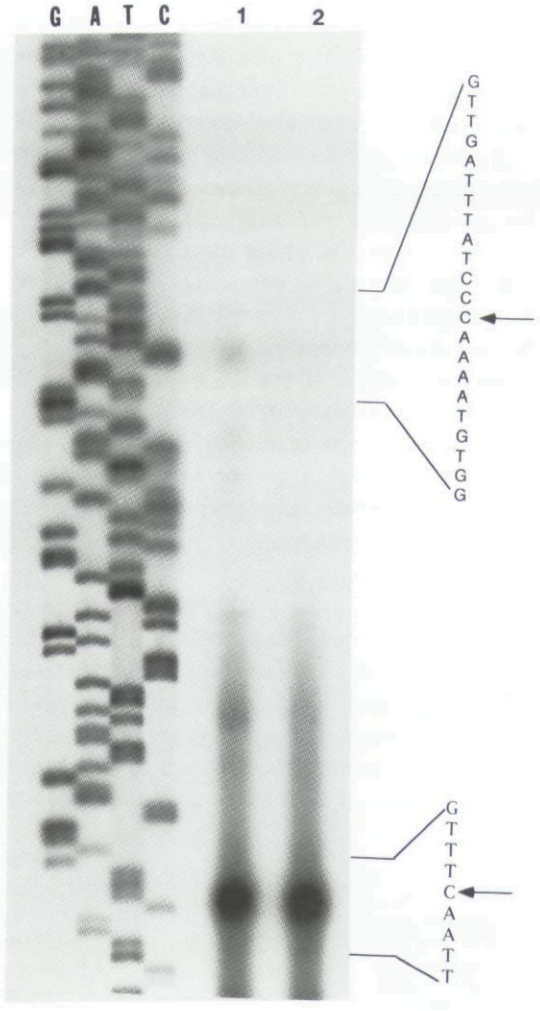

Fig. 6. 5′-Primer extension analysis of *gspA* transcripts expressed in *E. coli*. RNA was isolated from a σ^{32} *E. coli* mutant (strain CAG2046) and from its wild type (strain CAG2044), both harbouring pBSP1. The oligonucleotide used for primer extension and sequencing of pBSP1 DNA is described in the legend to Fig. 4. Lanes 1 and 2 contain the cDNA species synthesized from mRNAs of wild type and σ^{32} *E. coli* mutant, respectively. The arrows indicate the 5′ end of the two mRNA species with the top arrow corresponding to nucleotide 159, while the bottom arrow corresponds to nucleotide 229 of the sequence shown in Fig. 2. Although not depicted, there was no cDNA synthesized from RNA isolated from either strain harbouring the vector alone. No smaller transcripts were detected and only the relevant portion of the gel is shown.

formation of inclusion bodies upon overexpression of certain eukaryotic proteins in *E. coli* (Allen *et al.*, 1992). These two proteins may be involved in a chaperonin-like function in maintaining the biological functions of the cell by binding denatured proteins to renature them or to 'escort' them to degradation sites. We speculate that the exposure of *L. pneumophila* to intracellular stress stimuli may cause the denaturation of proteins, and the induction of GspA is needed to renature these proteins or to transport them to degradation sites.

There were significant increases in the levels of *gspA* mRNA after exposure of *L. pneumophila* to stress stimuli, consistent with our observation of an increased synthesis of the GspA protein in response to similar stimuli (Abu Kwaik *et al.*, 1993). However, the most potent stimulus for GspA biosynthesis is the intracellular environment of the phagocytic cell (Abu Kwaik *et al.*, 1993). These findings suggest that the intracellular environment may present multiple stress stimuli or a more potent stimulus than that tested *in vitro*.

Others have shown that stress proteins of *S. typhimurium* are induced in response to the intracellular environment of the phagocytic cell (Abshire and Neidhardt, 1993; Buchmeier and Heffron, 1990). However, in this model, the observed response is a summation of protein expression changes derived from a mixed population of intracellular *Salmonella*, since about 50% of the phagosomes containing bacteria are fused to lysosomes (Buchmeier and Heffron, 1991). Given the similarity of observed protein induction with intracellular *L. pneumophila*, the notion that this bacterium uniformly resides within a thoroughly protected phagosome may be an oversimplification. The induction of the GspA protein and other stress proteins by actively growing intracellular *L. pneumophila* is a clear indication that the population of organisms are exposed to certain unidentified stress stimuli. Characterization of the MI genes of *L. pneumophila* will allow us to study the regulation of expression of these genes in response to cellular uptake, and to correlate their expression with the intracellular fate of the bacteria.

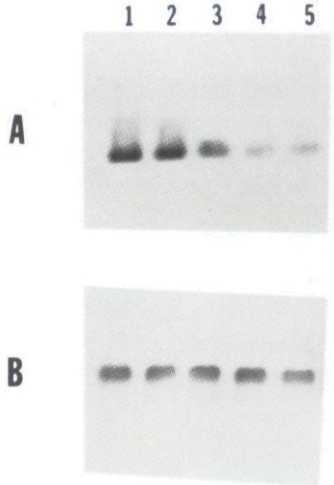

Fig. 7. A. Northern hybridization of *L. pneumophila* total RNA probed with the insert of pBSPR0.9 containing the *gspA* gene. B. The blot was stripped and rehybridized with a *mip* gene probe. RNA was isolated from *L. pneumophila* grown under various conditions: lane 1, osmotic shock; lane 2, oxidative stress; lane 3, heat shock; lane 4, grown at 28°C; lane 5, grown at 37°C.

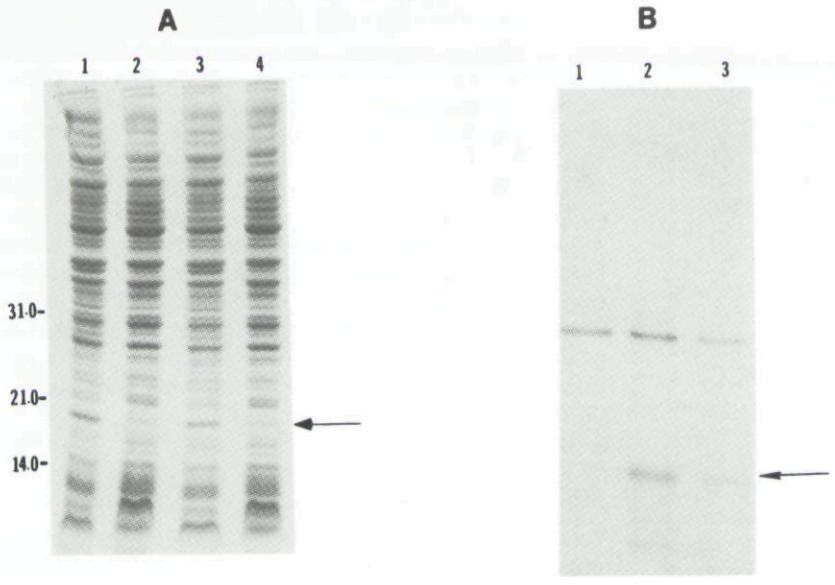

Fig. 8. A. SDS–PAGE stained with Coomassie brilliant blue of whole-cell lysates of *E. coli* harbouring pBSPR0.9 (lane 1); pBluescript (lane 2); pBSP1 (lane 3) and no plasmid (lane 4).
B. An autoradiograph of the polypeptides synthesized by a coupled *in vitro* transcription–translation system. Lane 1, pBluescript; lane 2, pBSPR0.9; and lane 3, pBSP1. The arrows indicate the 19 kDa protein.

The high degree of conservation of GspA and related genes through evolution suggests a vital function. As expected, strains representing 13 serogroups of *L. pneumophila* contained sequences highly homologous to *gspA*. The designation of the 16 kDa heat-shock proteins of *E. coli* as IbpA and IbpB (for inclusion-body-associated proteins) may be misleading, since these genes probably serve an endogenous function in the bacteria when not harbouring multicopy plasmids overexpressing foreign genes. Since the *Legionella* analogue is induced in response to the intracellular environment of the phagocytic cell and in response to *in vitro* stress stimuli, we chose global stress protein (GspA) as a better designation for the product of this gene.

Experimental procedures

Bacterial strains and vectors

The virulent AA100 strain of *L. pneumophila* has been described previously (Edelstein *et al.*, 1982; Abu Kwaik *et al.*, 1993). Strains belonging to all serogroups of *L. pneumophila* were kindly provided by Dr B. Fields at the Centers for Disease Control and are listed in Table 1. *E. coli* strain XL1-Blue was used as the host for the genomic library. The plasmid pBluescript was purchased from Stratagene. The *E. coli* strain CAG2046 is a σ^{32} mutant and this and its wild type, CAG2044, were kindly provided by K. Abshire and F. Neidhardt (University of Michigan).

DNA manipulations

Chromosomal DNA preparations, transfection, *in situ* colony hybridization, end radiolabelling of oligonucleotides, restriction enzyme digestion and DNA ligation were performed as described elsewhere (Sambrook *et al.*, 1989) unless specified. Restriction enzymes, T4 polynucleotide kinase and T4 DNA ligase were obtained from Bethesda Research Laboratories.

Plasmid DNA preparations were performed using the QIAGEN plasmid kit according to the manufacturer's recommendations (QIAGEN). Transformations were carried out by electroporation using a Bio-Rad Gene Pulser as recommended. DNA fragments from agarose gels for subcloning or for labelling for Southern hybridization were purified using the GeneClean Kit (Bio 101).

A Posiblot Pressure Blotter (Stratagene) was used to transfer the DNA from agarose gels on to nylon membranes. The DNA was cross-linked to the membrane using an Ultraviolet Stratalinker 1800 (Stratagene). For non-radioactive Southern hybridizations, the probe was labelled using an ECL random priming kit (Amersham). High-stringency hybridization and chemiluminescence detection were performed according to manufacturer's recommendations (Amersham). Autoradiography was performed using Kodak X-OMAT AR film (Kodak).

Table 1. *L. pneumophila* strains used in Southern hybridization.

Serogroup	Strain	Source
1[a]	AA100	Edelstein *et al.* (1982)
2	Togus 1	CDC[b]
3	Bloomington 2	CDC
4	Los Angeles 1	CDC
5	Dallas 1	CDC
6	Chicago 2	CDC
7	Chicago 8	CDC
8	Concord 3	CDC
9	IN 23	CDC
10	Leiden 1	CDC
11	797-PA-H	CDC
12	570-Co-H	CDC
13	82A3105	CDC

a. The serogroups of *L. pneumophila* are distinguished by their lipopolysaccharide antigen (Ciesielski *et al.*, 1986).
b. CDC, Centers for Disease Control, Atlanta, GA.

Cloning, screening, and subcloning

Chromosomal DNA of *L. pneumophila* strain AA100 was digested with *Eco*RI, and size-fractionated by sucrose gradient. A genomic DNA library was constructed in the lambda ZAPII vector (Stratagene) using a 4–10 kb fragment. Packaging was performed using Stratagene Gigapack II gold extracts, following the manufacturer's protocols (Stratagene). *In vivo* excision of the recombinant pBluescript phagemid library and transfection into *E. coli* XL1-Blue was performed according to the manufacturer's recommendations (Stratagene).

A degenerate oligonucleotide corresponding to the *N*-terminus amino acid sequence TPLLRHSVG of the GspA protein, derived by Edman degradation, was synthesized (University of Michigan Biomedical Research Core Facility). The sequence of the degenerate oligonucleotide was 5′-ACIC-CITTUTTUCGICAYTCIGTIGG-3′, where I = inosine; U = A and G; and Y = T and C. The oligonucleotide was radiolabelled and used as a probe for *in situ* colony hybridizations to screen the *L. pneumophila* genomic library transfected into *E. coli* (Sambrook *et al.*, 1989). Two steps were taken to minimize background and non-specific binding of the oligonucleotide to the debris of the colonies. First, the colonies were lifted on to nitrocellulose filters and the plasmid was amplified on Luria–Bertani (LB) agar plates supplemented with 200 $\mu g\,ml^{-1}$ chloramphenicol as described elsewhere (Sambrook *et al.*, 1989). Second, after baking the blots at 80°C, the blots were washed for 12–16 h at 65°C in 3× SSC, 0.1% SDS with multiple changes.

The subclone pBSPE2 contained the 2.7 kb *Eco*RV–*Eco*RI fragment of pBSP1 and was derived by a deletion of the left 3.8 kb *Eco*RV–*Eco*RI fragment of pBSP1 (Fig. 1). The DNA region spanning the sequence 1–920 in Fig. 2, which contained the *gspA* gene and its transcriptional initiation and termination signals, was amplified by the polymerase chain reaction (PCR) using the primers 5′-GGTAACTCATTAATTT-3′ and 5′-GGCCTCATATTCAGTTG-3′ for the sense and antisense strands, respectively. The PCR reactions were carried out in an automated DNA Thermal Cycler (Coy Laboratory Products) using the Vent™ DNA Polymerase (New England Biolabs). The 920 bp PCR product was subcloned into pBluescript and designated pBSPR0.9 (Fig. 1).

DNA sequencing

The dideoxy chain termination method of Sanger was employed using a Sequenase kit (United States Biochemical) (Sanger *et al.*, 1977). The degenerate oligonucleotide used for screening the genomic library was used for the initial sequencing reaction. Oligonucleotides were subsequently synthesized to sequence both strands (University of Michigan Biomedical Core Facility).

5′-primer extension

Total RNA was isolated using TRI reagent and used immediately, as recommended (MRC). For 5′-primer extensions, an oligonucleotide with the sequence 5′-TCATTGAATCGTT-CAAAACCAACGGAATGC-3′, which is complementary to the nucleotide sequence 283–312 of the sequence shown in Fig. 2, was synthesized. The 5′-primer extensions were performed as described elsewhere, with modifications (Freitag *et al.*, 1992). Briefly, 0.2 pmol of the 30-mer oligonucleotide was used to hybridize to 10 μg of RNA in 10 μl of 0.1 M KCl, 0.05 M Tris-HCl pH 8.3. The mixture was denatured at 95°C for 1 min prior to annealing for 2 min at 55°C, followed by incubation on ice for 15 min. cDNA was synthesized for 45 min at 48°C from one-half of this RNA–primer hybrid mixture with 1 U of Superscript Reverse Transcriptase in the recommended buffer supplied by the manufacturer (BRL). The reaction was terminated by the addition of 5 μl of Sequenase stop solution provided in the Sequenase kit (United States Biochemicals), and the sample was denatured at 95°C for 5 min prior to electrophoresis.

Northern hybridizations

Total RNA was isolated (as described above) from *L. pneumophila* grown at 28°C or 37°C, or after exposure of the bacteria for 30 min to heat shock, osmotic shock or oxidative stress as described previously (Abu Kwaik *et al.*, 1993). Equal amounts of total RNA (10 μg) was subjected to electrophoresis on a 1.2% agarose gel containing formaldehyde (Sambrook *et al.*, 1989). The RNA was transferred on to a nitrocellulose membrane using a Posiblot Pressure Blotter as recommended (Stratagene). The probes were labelled by a non-radioactive ECL random priming kit as describe above for Southern hybridization. The blot was first probed with the pBSPR0.9 insert and developed. The probe was released, and the blot was reprobed again with the *mip* gene, as recommended (Stratagene) (Engleberg *et al.*, 1989). The intensity of the bands were measured using a Gilford spectrophotometer equipped with densitometric analysis apparatus (Gilford Instrument Laboratories). The relative level of *gspA* mRNA from bacteria grown at 37°C to the level of *mip* mRNA was adjusted to 1, and the relative levels of the other mRNAs were calculated proportionally.

In vivo *and* in vitro *and expression of* gspA

To detect *in vivo* expression of *gspA* in *E. coli*, cell lysates of *E. coli* harbouring pBluescript vector pBSP1 or pBSPR0.9 were resolved by SDS–PAGE followed by Coomassie brilliant blue staining (Laemmli, 1970).

A coupled *in vitro* transcription–translation system was used to examine the polypeptides expressed by pBSP1 and pBSPR0.9 (Promega Biotech). [^{35}S]-methionine was used to radiolabel the polypeptides. The reaction products were resolved by SDS–PAGE and autoradiography.

Acknowledgements

We thank B. Fields (CDC, Atlanta, GA), K. Abshire and F. Neidhardt (University of Michigan) for providing bacterial strains. We thank D. Higgins and N. Freitag for their help with and comments on the 5′-primer extensions and M. Hurley for her help with the densitometric analysis. We thank N. C. Cianciotto for communication of his unpublished data about the regulation of *mip* expression. Y.A. is supported

by National Institute of Health training grant 5T32 AI07360. N.C.E. is supported by PHS grants RO1AI26232 and RO1AI24731.

References

Abshire, K.Z., and Neidhardt, F.C. (1993) Analysis of proteins synthesized by *Salmonella typhimurium* during growth within host macrophage. *J Bacteriol* **175:** 3734–3743.

Abu Kwaik, Y., Eisenstein, B.I., and Engleberg, N.C. (1993) Phenotypic modulation by *Legionella pneumophila* upon infection of macrophages. *Infect Immun* **61:** 1320–1329.

Allen, P.S., Polazzi, J.O., Gierse, J.K., and Easton, A.M. (1992) Two novel heat shock genes encoding proteins produced in response to heterologous protein expression in *Escherichia coli*. *J Bacteriol* **174:** 6938–6947.

Broome, C.V., and Fraser, D.W. (1979) Epidemiologic aspects of legionellosis. *Epidemiol Rev* **1:** 1–16.

Buchmeier, N.A., and Heffron, F. (1990) Induction of *Salmonella* stress proteins upon infection of macrophages. *Science* **248:** 730–732.

Buchmeier, N.A., and Heffron, F. (1991) Inhibition of macrophage phagosome–lysosome fusion by *Salmonella typhimurium*. *Infect Immun* **59:** 2232–2238.

Cianciotto, N., Eisenstein, B.I., Engleberg, N.C., and Shuman, H. (1989) Genetics and molecular pathogenesis of *Legionella pneumophila*, an intracellular parasite of macrophages. *Mol Biol Med* **6:** 409–424.

Ciesielski, C.A., Blaser, M.J., and Wang, W.L. (1986) Serogroup specificity of *Legionella pneumophila* is related to lipopolysaccharide characteristics. *Infect Immun* **51:** 397–404.

Dondero, T.J., Rendtorff, R.C., Mallison, G.F., Weeks, R.M., Levy, J.S., Wong, E.W., and Schaffner, W. (1980) An outbreak of Legionnaires' disease associated with a contaminated cooling tower. *N Engl J Med* **302:** 356–370.

Dowling, J.N., Saha, A.K., and Glew, R.H. (1992) Virulence factors of the family *Legionellaceae*. *Microbiol Rev* **56:** 32–60.

Edelstein, P.H., Brenner, D.J., Moss, C.W., Steigerwalt, A.G., Francis, E.M., and George, W.L. (1982) *Legionella wadsworthii* species nova: a cause of human pneumonia. *Ann Intern Med* **97:** 809–813.

Engleberg, N.C., Carter, C., Weber, D.R., Cianciotto, N.P., and Eisenstein, B.I. (1989) DNA sequence of *mip*, a *Legionella pneumophila* gene associated with macrophage infectivity. *Infect Immun* **57:** 1263–1270.

Freitag, N.E., Youngman, P., and Portnoy, D.A. (1992) Transcriptional activation of *Listeria monocytogenes* hemolysin gene in *Bacillus subtilis*. *J Bacteriol* **174:** 1293–1298.

Garnier, J., Osguthorpe, D.J., and Robson, B. (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J Mol Biol* **120:** 97–120.

Hemmingsen, S.M., Woolford, C., van der Vies, S.M., Tilly, K., Dennis, D.T., Georgopoulos, C., Hendrix, R.W., and Ellis, R.J. (1988) Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature* **333:** 330–334.

Horwitz, M.A. (1983a) The Legionnaires' disease bacterium (*Legionella pneumophila*) inhibits phagosome–lysosome fusion in human monocytes. *J Exp Med* **158:** 2108–2126.

Horwitz, M.A. (1983b) Formation of a novel phagosome by the Legionnaires' disease bacterium (*Legionella pneumophila*) in human monocytes. *J Exp Med* **158:** 1319–1331.

Horwitz, M.A. (1984) Phagocytosis of the Legionnaires' disease bacterium (*Legionella pneumophila*) occurs by a novel mechanism: engulfment within a pseudopod coil. *Cell* **36:** 27–33.

Horwitz, M.A., and Maxfield, F.R. (1984) *Legionella pneumophila* inhibits acidification of its phagosome in human monocytes. *J Cell Biol* **99:** 1936–1943.

Horwitz, M.A., and Silverstein, S.C. (1980) Legionnaires' disease bacterium (*Legionella pneumophila*) multiples intracellularly in human monocytes. *J Clin Invest* **66:** 441–450.

Kyte, J., and Doolittle, R.F. (1982) A simple method for displaying hydropathic character of a protein. *J Mol Biol* **157:** 105–132.

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

Marra, A., and Shuman, H. (1992) Genetics of *L. pneumophila* virulence. *Annu Rev Genet* **26:** 51–69.

Miller, J.F., Mekalanos, J.J., and Falkow, S. (1989) Coordinate regulation and sensory transduction in the control of bacterial virulence. *Science* **243:** 916–922.

Missiakas, D., Georgopoulos, C., and Raina, S. (1993). The *Escherichia coli* heat shock gene *htpy*: mutational analysis, cloning, sequencing, and transcriptional regulation. *J Bacteriol* **175:** 2613–2624.

Morris, G.K., Patton, C.M., Feeley, J.C., Johnson, S.E., Gorman, G., Martin, W.T., Skaliy, P., Mallison, G.F., Politi, B.D., and Mackel, D.C. (1979) Isolation of Legionnaires' disease bacterium from environmental samples. *Ann Intern Med* **90:** 664–666.

Nyström, T., and Neidhardt, F.C. (1992). Cloning , mapping and nucleotide sequence of a gene encoding a universal stress protein in *Escherichia coli*. *Mol Microbiol* **6:** 3187–3198.

Payne, N.R., and Horwitz, M.A. (1987) Phagocytosis of *Legionella pneumophila* is mediated by human monocyte complement receptors. *J Exp Med* **166:** 1377–1389.

Saito, H., and Uchida, H. (1978) Organization and expression of the *dnaJ* and *dnaK* genes of *Escherichia coli* K12. *Mol Gen Genet* **164:** 1–8.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

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