

Identification and Sequence Analysis of *lpfABCDE*, a Putative Fimbrial Operon of *Salmonella typhimurium*

ANDREAS J. BÄUMLER* AND FRED HEFFRON

Department of Molecular Microbiology and Immunology, Oregon Health Sciences University,
Portland, Oregon 97201-3098

Received 20 September 1994/Accepted 8 February 1995

A chromosomal region present in *Salmonella typhimurium* but absent from related species was identified by hybridization. A DNA probe originating from 78 min on the *S. typhimurium* chromosome hybridized with DNA from *Salmonella enteritidis*, *Salmonella heidelberg*, and *Salmonella dublin* but not with DNA from *Salmonella typhi*, *Salmonella arizonae*, *Escherichia coli*, and *Shigella* serotypes. Cloning and sequence analysis revealed that the corresponding region of the *S. typhimurium* chromosome encodes a fimbrial operon. Long fimbriae inserted at the poles of the bacterium were observed by electron microscopy when this fimbrial operon was introduced into a nonpiliated *E. coli* strain. The genes encoding these fimbriae were therefore termed *lpfABCDE*, for long polar fimbriae. Genetically, the *lpf* operon was found to be most closely related to the *fim* operon of *S. typhimurium*, both in gene order and in conservation of the deduced amino acid sequences.

All members of the family *Enterobacteriaceae* express fimbriae, or pili, which allow them to attach to host surfaces. This attachment is the first step in the colonization of their preferred host niche. The various fimbrial adhesins were originally distinguished by their receptor specificity. Hemagglutination of erythrocytes was used as a criterion to define three fimbrial classes (for a recent review, see reference 14). Fimbriae whose agglutination is prevented by coinubation with D-mannose were designated type 1 fimbriae (41). Type 2 fimbriae, a non-hemagglutinating fimbrial class, were later found to resemble nonadhesive variants of type 1 fimbriae (20, 36, 43). Finally, pili able to agglutinate animal erythrocytes only after the erythrocytes have been treated with tannic acid were designated type 3 fimbriae (22, 23). This early classification was supported by data on the morphological characteristics of these fimbriae. Type 1 and type 2 fimbriae have a width of 7 to 8 nm and display a channeled rodlike appearance due to the arrangement of subunits around a hollow core (43). Type 3 fimbriae are 4 to 5 nm in diameter and lack an axial hole on the basis of electron microscopic examination (42). Recently, thin fimbriae (3 to 4 nm) with a unique curly morphology in members of the *Enterobacteriaceae* have been described (4, 44). A comparison of sequence data obtained from several fimbrial systems with similar morphologies revealed that the major fimbrial subunits of curly fimbriae show a high degree of sequence conservation. These fimbriae are therefore now considered to form a distinct fimbrial class. The designation GVVPO has been proposed, in reference to a short amino acid sequence conserved at the N terminus of their major fimbrial subunit (19). Type 1, type 2, type 3, and GVVPO fimbriae have the common characteristic that their fimbrial subunits contain signal sequences which are cleaved by signal peptidase I. In contrast, a fifth enterobacterial class of pili, termed type 4 fimbriae, has been shown to use a different export system for translocation of its fimbrial subunits across the cytoplasmic membrane. Fimbrial subunits of type 4 fimbriae contain signal sequences which are recognized and cleaved by type IV prepilin signal peptidase (for a recent

review, see reference 54). As with that for the other fimbrial types, this classification has been supported by several morphological characteristics which distinguish enterobacterial type 4 fimbriae from other fimbrial classes. While type 1, type 2, type 3, and GVVPO fimbriae from enterobacteria have a length of 0.5 to 2 μm and are arranged peritrichously around the bacterial cell, enterobacterial type 4 fimbriae are polar and have a length of 10 to 20 μm (28).

Here we describe a fimbrial operon of *Salmonella typhimurium* termed *lpf*. Introduction of the *lpf* operon into *Escherichia coli* results in the expression of fimbriae with a morphology similar to that of type 4 fimbriae. However, sequence analysis indicated that the *lpf* operon is not closely related to that expressing type 4 fimbriae but rather shows similarity to that expressing *S. typhimurium* type 1 fimbriae.

MATERIALS AND METHODS

Bacterial strains, media, and growth conditions. *E. coli* ORN172 has been described previously (59). *E. coli* DH5 α and LE392 were purchased from Gibco BRL and New England Biolabs, respectively. *S. typhimurium* 14028, 6994, 7823, and LT2 were obtained from the American Type Culture Collection (ATCC). *E. coli* W3110, *Salmonella dublin* LANE, and *S. typhimurium* SL1344 and SR-11 were from laboratory stocks. Clinical isolates of *Shigella sonnei*, *Shigella boydii*, *Shigella dysenteriae*, enteroinvasive *E. coli*, enterotoxigenic *E. coli*, enteropathogenic *E. coli*, *Salmonella arizonae*, *Salmonella typhi*, *Salmonella heidelberg*, and *Salmonella enteritidis* were obtained from the State of California Health Laboratory or the County of San Diego Health Laboratory and have been described previously (34). *Shigella flexneri* BS176 and M90T have been described elsewhere (61).

All bacteria were cultured aerobically in Luria-Bertani broth (5 g of yeast extract per liter, 8 g of tryptone per liter, and 5 g of NaCl per liter) or on plates (Luria-Bertani broth containing 15 g of agar per liter) at 37°C. Antibiotics, when required, were included in the culture medium or plates at the following concentrations: carbenicillin, 100 mg/liter; kanamycin, 60 mg/liter; tetracycline, 15 mg/liter; and streptomycin, 50 mg/liter. Analytical grade chemicals were purchased from Sigma. All enzymes were purchased from Boehringer Mannheim.

Pulsed-field gel electrophoresis. The preparation of agarose-embedded chromosomal DNA for pulsed-field gel electrophoresis was based on a protocol from Liu and Sanderson (35). In brief, cells were harvested from 5 ml of an early-log-phase bacterial culture, resuspended in 0.5 ml of prewarmed (37°C) cell suspension buffer (10 mM Tris-HCl [pH 7.2], 20 mM NaCl, 100 mM EDTA), and mixed with 0.5 ml of 2% InCert agarose (InCert) which was dissolved in phosphate-buffered saline and precooled to 37°C. The mixture was poured into molds (Bio-Rad) and allowed to solidify. The molds were digested with lysozyme (1 mg/ml; New England Biolabs) in a volume of 7.5 ml for 1 h at room temperature in buffers provided by the manufacturer. The lysozyme solution was replaced by 7.5 ml of proteinase K (1 mg/ml; New England Biolabs) in buffer provided by the

* Corresponding author. Mailing address: Department of Molecular Microbiology and Immunology, Oregon Health Sciences University, 3181 S.W. Sam Jackson Park Rd., L220, Portland, OR 97201-3098. Phone: (503) 494-7768. Fax: (503) 494-6862.

manufacturer and incubated at 55°C for 48 h with shaking. Agarose-embedded DNA prepared in this manner can be stored in 7.5 ml of TE buffer (10 mM Tris-HCl [pH 7.2], 5 mM EDTA) at 4°C for several months without degradation.

For restriction digests, an agarose block equivalent to about 0.05 to 0.1 ml was incubated for 15 min at room temperature in 0.2 ml of 1× KGB buffer (38) in which sodium acetate was replaced by potassium acetate. The buffer was replaced with 1× KGB buffer containing 20 U of *Xba*I, and the samples were incubated for 4 h at 37°C. The agarose blocks were then loaded on an agarose gel for pulsed-field gel electrophoresis.

Pulsed-field gel electrophoresis of a 1% agarose gel was performed in 0.5% Tris-borate-EDTA at 200 V with pulse lengths of 6 to 150 s for 6 h, 6 to 12 s for 6 h, 26 to 36 s for 6 h, and 120 to 180 s for 6 h with a CHEF-DR11 apparatus from Bio-Rad.

Recombinant DNA techniques. Plasmid DNA was isolated with ion exchange columns from Qiagen. Standard methods were used for restriction endonuclease analyses, ligation and transformation of plasmid DNA, and isolation of chromosomal DNA from bacteria (39, 48). The construction of the gene bank from *S. typhimurium* ATCC 14028 in the cosmid vector pLAFRII has been described elsewhere (34). Subclones of cosmid inserts were constructed with the vector pBluescript SK+ (52). Sequencing was performed by the dideoxy chain termination method (49) with an AutoRead Sequencing Kit (Pharmacia) and an ALF automatic sequencer or according to a protocol of Kraft et al. (33) with α -³⁵S-dATP (Amersham) being used for labeling.

Southern hybridization. Southern transfer of DNA onto a nylon membrane was performed as previously described (48). The labeling of DNA probes, hybridization, and immunological detection were performed with the DNA labeling and detection kit (nonradioactive) from Boehringer Mannheim. The DNA was labeled by random primed incorporation of digoxigenin-labeled dUTP. Hybridization was performed at 65°C in solutions without formamide. For Southern hybridization with chromosomal DNA of different bacterial species, two 15-min washes were performed under nonstringent conditions at room temperature in 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.1% sodium dodecyl sulfate (SDS). For Southern hybridization with cosmids of a gene bank or with DNA separated on a pulsed-field gel, a nonstringent wash (10 min at room temperature in 2× SSC–0.1% SDS) and a stringent wash (30 min at 65°C in 0.2× SSC–0.1% SDS) were performed. Hybrids were detected by an enzyme-linked immunosorbent assay with an antibody conjugate (anti-digoxigenin alkaline phosphatase conjugate) and the substrate AMPPD [3-(2'-spiroadamantane)-4-methoxy-4-(3'-phosphoryloxy)phenyl-1,2-dioxethane; Boehringer Mannheim]. The light emitted by the dephosphorylated AMPPD was detected by X-ray film.

Preparation of fimbrial crude extracts. The protocol used for the preparation of fimbrial crude extracts was based on a protocol by van Zijderfeld et al. (58). In brief, bacteria were grown in 100 ml of Luria-Bertani broth to mid-log phase and were harvested by centrifugation (3,000 × g for 10 min). After the bacteria were resuspended in 10 ml of 75 mM NaCl–0.5 mM Tris-HCl (pH 7.0), they were incubated at 60°C for 30 min and then pelleted by centrifugation (3,000 × g for 10 min). The supernatant was transferred to a fresh tube, and solid ammonium sulfate was added to reach a final saturation of 60%. The mixture was incubated for 1 h at 4°C with stirring. Proteins were harvested by centrifugation (40,000 × g for 30 min), and the supernatant was discarded. The fimbrial crude extracts were resuspended in 0.1 ml of 75 mM NaCl–0.5 mM Tris-HCl (pH 7.0) and dialyzed overnight at 4°C against 0.1 M Tris-HCl, pH 7.0. Fimbrial crude extracts were boiled for 10 min in sample buffer containing 4% SDS and separated on a 15% Tricine-SDS-polyacrylamide gel (51). After electrophoresis, the proteins were visualized by silver staining.

Electron microscopy. Bacteria were grown overnight in a static culture and were allowed to adhere to a Formvar-coated grid for 2 min. The bacteria were fixed with 0.1% glutaraldehyde in sodium cacodylate buffer (100 mM, pH 7.4) for 1 min. The grid was rinsed with water, and the fimbriae were negatively stained with 0.5% (wt/vol) aqueous uranyl acetate, pH 4.6, for 30 s. The grids were allowed to dry before they were analyzed by electron microscopy.

Computer analysis. The nucleotide sequences were compared with nonredundant updates of SWISS-PROT, PIR(R), and GenPept with the program blastX and were compared with nonredundant updates of GenBank and EMBL with the program blastN (3). The nucleotide sequences were further analyzed with the PC/GENE software package.

Nucleotide sequence accession number. The nucleotide sequence of the 7,065-bp DNA fragment reported on in this paper has been assigned the GenBank accession number U18559. The nucleotide sequence of a region of the *E. coli* K-12 chromosome which corresponds to 78 min on the *S. typhimurium* genome has been determined previously. It has been assigned GenBank accession number U00039.

RESULTS

Identification and cloning of a *Salmonella*-specific chromosomal region from *S. typhimurium*. Several short chromosomal DNA fragments from *S. typhimurium* were obtained by inverse PCR from chromosomal DNA of strains carrying transposon Tn10 insertions at different map locations (8). In order to

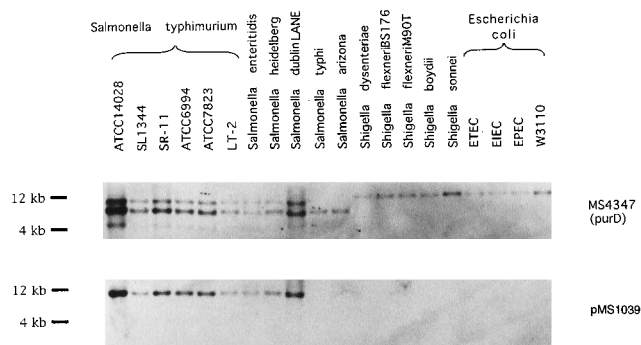


FIG. 1. Southern blot of chromosomal DNA from enterobacteria digested with *Eco*RI, separated on an agarose gel, and hybridized with part of the *S. typhimurium purD* gene (MS4347) and with part of the *lpf* operon (pMS1039). The designations of the bacterial strains used for the preparation of the chromosomal DNA are indicated above the blots. The positions of the DNA fragments used as size standards are given on the left. ETEC, enterotoxigenic *E. coli*; EIEC, enteroinvasive *E. coli*; EPEC, enteropathogenic *E. coli*.

identify regions on the *S. typhimurium* chromosome which were not present in related members of the *Enterobacteriaceae*, we used these DNA fragments as probes in a Southern hybridization with DNA from several related species. Under the conditions used, a hybridization signal should be obtained if a DNA fragment showed a more than 65% nucleotide sequence identity with a probe. A 210-bp DNA probe (the insert of plasmid pMS10) flanking the Tn10 insertion of strain MS1521 hybridized with DNA from *Salmonella choleraesuis* subsp. *choleraesuis*, including the serotypes *S. enteritidis*, *S. heidelberg*, *S. dublin*, and several frequently used strains of serotype *S. typhimurium* (for current nomenclature on the *Salmonella* genus, see reference 31). No hybridization signal was detected with DNA from the *S. choleraesuis* subsp. *choleraesuis* serotype *S. typhi*, *S. choleraesuis* subsp. *arizonae*, several *E. coli* isolates, and *Shigella* serovars (Fig. 1). A 317-bp DNA fragment containing an internal part of the *S. typhimurium purD* gene (8) was used to probe a second blot containing chromosomal DNA from the same enterobacterial species. The *stn* gene, which is located only 1.4 kb downstream of *purD* (13, 45), has been shown to be duplicated in *S. typhimurium* (12). The two bands hybridizing with the *purD* probe in DNA from *Salmonella* serotypes indicate that this duplication includes part of the *purD* gene. Unlike the pMS10 probe, the *purD* probe gave a hybridization signal with all strains tested (Fig. 1). These data indicated that the insert of plasmid pMS10 originates from an *S. typhimurium* region which is present in serotypes of *S. choleraesuis* subsp. *choleraesuis* except serovar *S. typhi*.

The corresponding region of the *S. typhimurium* chromosome was cloned from a previously described cosmid bank constructed in pLAFRII and propagated in *E. coli* LE392 (34). The cosmid bank was spread on plates, and 300 colonies were picked and grown up individually overnight. Cosmid DNA was prepared from 30 pools, each containing 10 overnight cultures, and digested with *Eco*RI, separated on an agarose gel, and hybridized with probe pMS10. The DNA of two pools hybridized with probe pMS10. The 20 cosmids included in these two pools were then isolated individually and digested with *Eco*RI and *Pst*I, and the hybridization was repeated with probe pMS10. Each pool contained one cosmid, termed pMS1000 and pMS1001, respectively, which gave a hybridization signal corresponding to a 2-kb *Eco*RI-*Pst*I DNA fragment (data not shown). The 2-kb *Eco*RI-*Pst*I fragment was cloned into the vector pBluescript SK+ (52) to give rise to plasmid pMS1039

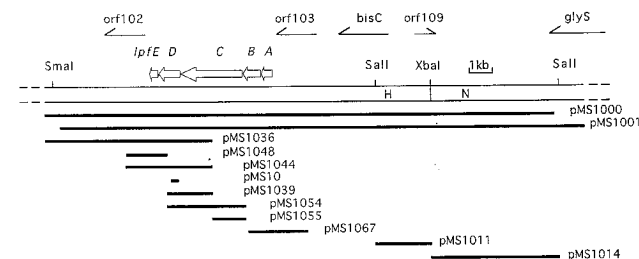


FIG. 2. Restriction map of an *S. typhimurium* chromosomal region located around 78 min on the genetic map. The region shown contains the ends of *Xba*I fragment H (H) and *Xba*I fragment N (N) from *S. typhimurium* (34). The inserts of cosmids pMS1000 and pMS1001 are shown below. The positions of the inserts contained in several subclones (pMS1036, pMS1048, pMS1044, pMS10, pMS1039, pMS1054, pMS1055, pMS1067, pMS1011, and pMS1014) are given at the bottom of the figure. The positions and orientations of the genes *lpfABCDE* are indicated by outlined arrows above the restriction map. The approximate positions of genes showing short stretches of homology to those of *E. coli* are indicated at the top of the figure by thin arrows and the designation of the *E. coli* homolog.

(Fig. 2). In order to confirm the hybridization data obtained with probe pMS10, the 2-kb insert of pMS1039 was used as a probe in a Southern hybridization with DNA from several enterobacterial species (Fig. 1). The hybridization pattern obtained with probe pMS1039 was found to be identical to the one obtained with probe pMS10.

Restriction map and physical map location of the cloned *Salmonella*-specific region. In order to construct a restriction map, cosmids pMS1000 and 1001 were digested with the restriction enzymes *Xba*I, *Sal*I, *Sma*I and with combinations of these. After separation on an agarose gel, DNA fragments were hybridized with probe pMS10 (data not shown). In addition, a 5.7-kb *Sal*I-*Xba*I fragment from pMS1000 was cloned into pBluescript SK+ (pMS1014) and used as a probe for hybridization with the same blot (data not shown). The information obtained from these experiments allowed the construction of a restriction map of cosmids pMS1000 and 1001 (Fig. 2).

Cosmids pMS1000 and pMS1001 contained an *Xba*I site located a distance of about 10 kb from the insert of pMS10 (Fig. 2). The restriction endonuclease *Xba*I cuts the *S. typhimurium* chromosome only 23 times, creating fragments of between 18 and 800 kb which have been designated alphabetically according to their size. The map positions of all *Xba*I sites on the *S. typhimurium* chromosome have been determined so that a macrorestriction map of the *S. typhimurium* genome can be constructed (35). The *Tn10* insertion in strain MS1521 has previously been mapped to 78 min on the *S. typhimurium* chromosome, close to an *Xba*I site located between *Xba*I fragments N and H (8). In order to determine whether the *Salmonella*-specific region defined by pMS10 and pMS1039 is located on *Xba*I fragment N or H, chromosomal DNA of *S. typhimurium* was digested with *Xba*I, separated by pulsed-field electrophoresis, and hybridized with probe pMS1039 and probe pMS1014. A hybridization signal corresponding to *Xba*I fragment G, H, I, or J was obtained with probe pMS1039, while probe pMS1014 gave a hybridization signal corresponding to *Xba*I fragment N, M, O, or P (Fig. 3). With the map location determined previously for the *Tn10* insertion in strain MS1521 (8), these data show that the inserts of pMS10 and pMS1039 originate from *Xba*I fragment H at 78 min on the *S. typhimurium* genome (Fig. 2).

Sequence analysis of the *Salmonella*-specific DNA region. The nucleotide sequence of a 7,065-bp DNA region was determined (Fig. 4). The region was found to encode five open

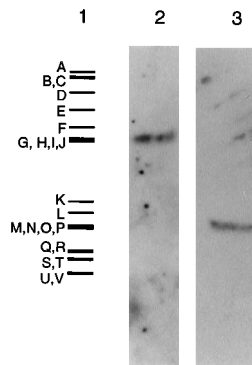


FIG. 3. Southern blot of chromosomal DNA from *S. typhimurium* 14028 digested with *Xba*I, separated by pulsed-field agarose gel electrophoresis, and hybridized with probe pMS1039 (lane 2) or probe pMS1014 (lane 3). The positions and designations (according to reference 34) of the *Xba*I fragments of *S. typhimurium* 14028 are given in lane 1.

reading frames, designated *lpfABCDE*, which are transcribed counterclockwise from the *S. typhimurium* chromosome (Fig. 2). In addition, the sequence contained the beginning and the end of two further open reading frames, designated *orf1* and *orf2*, respectively (Fig. 4).

The positions of open reading frames *lpfABCD* and *-E* on the nucleotide sequence and the sizes of the predicted polypeptide are given in Table 1 and Fig. 4. All predicted polypeptides were found to contain typical signal sequences (Table 1).

Sequence comparison with *E. coli* defines the endpoints of the *Salmonella*-specific DNA region. The nucleotide sequence of a region of the *E. coli* K-12 chromosome which corresponds to 78 min on the *S. typhimurium* genome was compared with the nucleotide sequence determined from pMS1000 in order to identify regions of DNA which are conserved between *E. coli* and *S. typhimurium* and those which are present in *S. typhimurium* only.

The deduced amino acid sequences of *orf1* and *orf2* shared 64 and 91% identities with the deduced amino acid sequences of the *E. coli* genes *orf103* and *orf102*, respectively. In *E. coli*, *orf103* and *orf102* are separated by 327 bp of noncoding DNA. In contrast, *orf1* and *orf2* are separated by a 6,053-bp DNA region which contains the open reading frames *lpfABCDE* (Fig. 5A). The GC content of open reading frames *lpfABCDE* is close to the GC content of 51 to 53% described for the entire *Salmonella* genome (Table 1) (24). However, this DNA region showed no homology to DNA from this region of the *E. coli* chromosome. The DNA-DNA homology of *E. coli* and *S. typhimurium* ends on one side close to the beginning of *orf2* (Fig. 5B) and on the other side close to the end of *orf1* (Fig. 5C). A short imperfect inverted repeat (13 bp with 11 matches) was found at the boundaries of the sequence homology in the *S. typhimurium* sequence (Fig. 5B and C). The above-mentioned results show that *orf1* and *orf2* of *S. typhimurium* are separated by a 6,053-bp region of DNA which is not present in *E. coli* and therefore confirm data obtained by hybridization with probes pMS10 and pMS1039 (Fig. 1).

In order to determine whether any other major differences in gene order between DNA present in pMS1000 or pMS1001 and the corresponding region of *E. coli* can be detected, the nucleotide sequences of short stretches of DNA were determined from the ends of plasmids pMS1011 and pMS1014 (Fig. 2). By comparison of the deduced amino acid sequences, *Salmonella* homologs of *bisC*, *orf109*, and *glyS* were found at

1	AAA AGC TCT TGT GGC GGA AGT AGG GTG CTA ACC TTC TGC GTT ATC TTT Lys Ser Ser Cys Gly Gly Ser Arg Val Leu Thr Phe Cys Val Ile Phe	1873	GCG CTT GCG CTG ATG ATA GCG GGT ACG CTC CCC GCG TAT GCG GGA ACA Ala Leu Ala GCG Leu Met Ile Ala Gly Thr Ala Tyr Ala Gly Thr
49	GCG CTG TTA ATT ATC TCG CTG GCA CTT TCC ACC ACT ATT CGC CAG CCG Ala Leu Leu Ile Ile Ser Leu Ala Leu Ser Thr Thr Ile Arg Gln Pro	1921	TTT AAC CCG CGC TTT CTG GAG GAT GTG CCG GGT ATT GAT CAG CAC GTT Phe Asn Pro Arg Phe Leu Glu Asp Val Pro Gly Ile Asp Gln His Val
97	CAG CGG GAG GGG GTT AGC GAA GCG TCT GCG TAA CAC ATG CAA CAC CAG Gln Arg Glu Gly Val Ser Glu Ala Ser Ala	1969	GAC CTT TCA ATG TAT GAA TCC AAT AAA GCT GAA CAC CTG CCA GGT AAA Asp Leu Ser Met Tyr Glu Ser Asn Lys Ala Glu His Leu Ser Gly Lys
145	GCG TAC AGC GCA CCC ATT TGG TGC CCT TTT TTT TAT TTA GCA CAA ATA 193 CCT AAT CAA TTG TAG TTA AAA AAA CGT CTA ATA AAT AAG GAA GAC ATT 241 TAA CTT ATT TAT GAA TAG GAA GAA ATA ATA TAT TAA TTA TAT TAA TTT 289 ATT CTT AAT AAA AAA TTA CAT TTA TGT ACA TTC CAT TTG TAA TAT ATT 337 GAT TTC TAT TCT TTT TAA GAT CTA CTA ACA ATT ATT TTT ATA TAT ACT 385 AAT TAT AGT ATT CAA TAC CCA TCT CTA TAC ATT CCT CAC AGA	2017	TAC CGC GTC TCG GTG GTG GTC AAC GAA AAA AAA ATG GAG TCT CGC ACC Tyr Arg Val Ser Val Val Val Val Val Val Val Val Val Val Val Val Val
433	ATG CAG ATA ATC CTA AGG ATG CGT TCT GTT ATC TAC CGT CAT AAA TGG Met	2065	CTG GAG TTT AAG GCA GCG ACA GAG GCG CAG GCG GCA AAA ATG GGT GAA Leu Glu Phe Lys Ala Ala Thr Glu Ala Gln Arg Ala Lys Met Gly Glu
481	AGT TTT TAA TGA AAA AGG TTG TTT TTG CTC TGT CTG CTC TCG CTG TAG Glu Phe Leu Met Lys Lys Val Val Phe Ala Leu Ser Ala Leu Ala Val	2113	TCC CTG GTG CCG TGC TTA AGT CGC GTG CAG CTT GAA GAT ATG GGC GTG Ser Leu Val Pro Cys Leu Ser Arg Val Gln Leu Glu Asp Met Gly Val
529	TTT CCA CTT CTG CTT TCG CTG CTG AAT CTG GTG ACG GCA CCA TTA AAT Val Ser Thr Ser Ala Phe Ala Ala Glu Ser Gly Asp Gly Thr Ile Lys	2161	CGT ATT GAT AGC TTC CCG GCG CTG AAA ATG GCC CCG CCT GAA GCC TGT Arg Ile Asp Ser Phe Pro Ala Leu Ser Phe Ala Leu Ser Ala Met Lys Cys
577	TCA CCG GTG AAA TCG TTG ACG GCG CAT GCG TTT CTA CTG ACT CCC Phe Thr Gly Glu Ile Val Asp Ala Pro Cys Val Val Ser Thr Asp Ser	2209	GTT GCT TTT GAC GAC ATT ATT CCC CAG GCC GCC AGC CAT TTC GAC TTT Val Ala Phe Asp Asp Ile Ile Pro Gln Ala Ala Ser His Phe Asp Phe
625	AGA ACC AGG AAG TTG TGC TGG GTC AGG TTA AGA AAA ATA TCT TCA AAG Gln Asn Gln Glu Val Val Leu Glu Val Lys Lys Asn Ile Phe Lys	2257	GCA GAC CAG ACC CTG ATC ATG AGC TTC CCG CAG GCT GCG ATG AAG CAG Ala Asp Gln Thr Leu Ile Met Ser Phe Pro Gln Ala Ala Met Lys Gln
673	CCA TTG GCG ACA AGT CTT CTT CTA AGC CTT TCC AGA TCA AAC TGG AAG Ala Ile Gly Asp Lys Ser Ser Ser Lys Pro Phe Gln Ile Lys Leu Glu	2305	ACA GCG CGC GGT ACG GTG CCA GAA TCG CAG TGG CAG GAA GGG GTG AAT Thr Ala Arg Gly Thr Val Pro Glu Ser Gln Trp Asp Glu Gly Val Asn
721	ACT GTG ACA TCA CCT CTA ATA CCA AAG TTA ACG TAA GCT TCA ATG GCG Asp Cys Asp Ile Thr Ser Asn Thr Lys Val Asn Val Ser Phe Asn Gly	2353	GCC CTG CTG GTG GAT TAT AAC TTT TCC GGC AGC AAC GCC AGC TAT GAC Ala Leu Leu Val Asp Tyr Asn Phe Ser Gly Ser Asn Ala Ser Tyr Asp
769	TTG GTG ATA CAG ACG ATG CGA CAC TGG TTT CTG TTA ACA CTG AAG CAG Val Gly Asp Thr Asp Asp Ala Thr Leu Val Ser Val Asn Thr Glu Ala	2401	GCA CAC GAC AGT GAA ACC AGC TAC AAC AGC CAG ACC TAT TCT ATG AAT Ala His Asp Ser Glu Thr Ser Tyr Asn Ser Asp Ser Tyr Tyr Leu Asn
817	GTG CCG CAA CTG GCG TGG GCA TCG GTA TCT ACG ACA AGC CTA ACA ACC Gly Ala Ala Thr Gly Val Gly Ile Gly Ile Tyr Asp Asn Ala Asn Lys	2449	CTG CGC AGC GGT ATG AAC CTG GGG GCA TGG CCG TTA CGT AAC TAT AGC Leu Arg Ser Gly Met Asn Leu Glu Ala Trp Arg Ala Trp Arg Ser Tyr Ser
865	TTG TTG AAA TGA ACA CCG GTA AAT CCA CCA CTA CGT TGG CTG CTG GTC Leu Val Glu Met Asn Thr Gly Lys Ser Thr Thr Thr Leu Ala Ala Gly	2497	ACC TGG ACG CGA AAC GAC GGT AAC AAC ACA TGG GAT AAC ATT GGC ACA Thr Trp Thr Arg Asn Asp Gly Asn Asn Thr Trp Asp Asn Ile Gly Thr
913	AGA CCG TGC TGT ACT ACA CCG CTA ACT ACG TTG CAA CAA AAG ATA CTG Gln Thr Val Leu Tyr Tyr Thr Ala Asn Tyr Val Ala Thr Lys Asp Thr	2545	TCC TTA AGC CGT GCC ATT GTA CCG GTG AAA TCA CAG CTG ACG TTG GGG Ser Leu Ser Arg Ala Ile Val Pro Leu Lys Ser Gln Leu Thr Leu Gly
961	TAA CCA CTG GTT ACG GTA ACG CAG AAG TGG ACT TCA ACC TGT CCT ACG Val Thr Thr Gly Tyr Gly Lys Ala Glu Val Asp Phe Asn Leu Ser Tyr	2593	GAT ACC TCC ACT GCC GGT GAT ATT TTT GAC AGC GTT CAG ATG CCG GGT Asp Thr Ser Thr Ala Gly Asp Ile Phe Asp Ser Val Gln Met Arg Gly
1009	AAT AAT CGA ATT TTC GTT AAT ACA GAC AAT CAT AAT GGC AAC GGA AAT	2641	GTG CAG TTA ACT TCC GAC GAA GAG ATG CTG CCT GAC AGC CAG CGC GGG Val Gln Leu Thr Ser Asp Glu Glu Met Leu Thr Pro Asp Ser Gln Arg Thr
1057	CCC GTT GCC ATT TTT TCC AGC GGA GGC TCA GGA AGA GAA TCA TGA ACC Met Asn	2689	TTT GCG CCC GTC ATC CCG GGT ATT GCC AAA AGT AAC GCC GAA GTT ACC Phe Ala Pro Val Ile Arg Gly Ile Ala Lys Ser Asn Ala Glu Val Thr
1105	GCT CAC GTT TGA TAT CTT GCA CAG CAC TGG TGC TGG CGT TGA TTG CTC Arg Ser Arg Leu Ile Ser Cys Thr Val Leu Ala Leu Ile Ala	2737	GTT GAG CAG AAC AAC TAC GTT ATT TAC CGT ACG TTT GTT CAG CCG GGT Ala Phe Gly Ile Asn Asp Leu Tyr Val Ile Tyr Arg Thr Val Gln Pro Gly
1153	AAA ACA GTT TTG CCG GAG GCG TGG CAT TAA GCA GCA CGC GTG TTA TTT Gln Asn Ser Phe Ala Gly Gly Val Ala Leu Ser Ser Thr Arg Val Ile	2785	GCG TTT GAA ATT AAC GAC CTG TAT CCA ACC TCA AAC AGC GGC GAC CTG Ala Phe Gly Ile Asn Asp Leu Tyr Pro Thr Ser Asn Ser Gly Asp Leu
1201	ATG ACG GTA GTA GAA AGG AAG CTT CTC TTA CCG TAA ATA ATA AAA GCA Tyr Asp Gly Ser Arg Lys Glu Ala Ser Leu Thr Val Asn Asn Lys Ser	2833	ACG GTC ACC ATT AAA GAA TCG GAC GGC AGT GAG CAG AAG TTT CTT CAG Thr Val Thr Ile Lys Glu Ser Asp Gly Ser Glu Gln Lys Phe Val Gln
1249	CCA CCG ATG AAT TTC TCA TFC AGT CAT GGA TTG ATG ATG CTA ACG GTA Thr Thr Asp Glu Thr Phe Leu Ile Gln Ser Trp Ile Asp Asp Ala Asn Gly	2881	CCG TTC TCC TCG GTG GCG CTC CTC CAG CGT GAA GGC CAT CTC AAA TAC Pro Phe Ser Ser Val Ala Leu Leu Gln Arg His Lys Tyr Lys Tyr
1297	ATA AAA AGA GCG CCT TTA TCA TCA CTC CAC CGT TAT TTA AAT TAA GCC Asn Lys Lys Thr Pro Phe Ile Ile Thr Pro Pro Leu Phe Lys Leu Ser	2929	AGC CTT TCC GCC GGG GAA TAC CGT GCC GGG AAC TAT AAC AGC GCC GAG Ser Leu Ser Ala Gly Glu Tyr Arg Ala Gly Asn Ala Ser Asn Ala Ser
1345	GCA CTA AAA ATA ACG TTT TAC GTA TTG TTA ATA CGA CGA ACA CGT TAC Pro Thr Lys Asn Asn Val Leu Arg Ile Val Asn Thr Thr Asn Thr Leu	2977	CCG AAA TTC GGG CAG CTT GAT GCC ATG TAC GGC CTG CCG TAT GGC TTT Pro Lys Phe Gly Gln Leu Asp Ala Met Tyr Gly Leu Pro Tyr Gly Phe
1393	CGC AGG ATC GCG AGT CCG TTT ATT GGA TTA ACG TAA AAG CTA TTC CTG Pro Gln Asp Arg Glu Ser Val Tyr Trp Ile Asn Val Lys Ala Ile Pro	3025	ACC GTT TAC GGT GGT GCG ATC TTC TCT GAC GAC TAT TAC TCG CTG CCG Thr Val Tyr Gly Gly Ala Ile Phe Ser Asp Asp Tyr Tyr Ser Leu Ala
1441	CCA AAA GTG AAG ACG CCG AAG CTA AAA ACG TAC TGC AGA TCG CCG TAC Ala Lys Ser Glu Asp Ala Glu Ala Lys Asn Val Leu Gln Ile Ala Val	3073	GGA GGA TTA GGT AAA AAC TTC GGT TAT ATC GGC GCG ATC TCC ATC GAT Gly Gly Leu Gly Lys Lys Asn Phe Gly Tyr Ile Gly Ala Ile Ser Ile Asp
1481	GTA CCC GCT TAA AAC TGT TCT ATC GCC CCG GCG TGA AAG CGA ATA Arg Thr Arg Leu Lys Leu Phe Tyr Arg Pro Ala Ala Leu Lys Ala Asn	3121	GTA ACC CAG GCA AAA AGC AAG CTG GCA AAT GAG GAG AAT TCG GAA GGT Val Thr Gln Ala Lys Ser Lys Leu Ala Asn Glu Glu Asn Ser Glu Gly
1537	GCA TGG ACG GCT GGA ACA AAC TGC AGT TCA CCA GCG CAG GGG CTA ACC Ser Met Asp Gly Trp Asn Lys Leu Gln Phe Thr Ser Ala Gly Ala Asn	3169	CAG TCT TAT CGT TTC CTC TAC TCC AAG AGC TTT AAC AGC GGT ACA GAT Gln Ser Tyr Arg Phe Leu Tyr Ser Lys Ser Phe Asn Ser Gly Thr Asp
1585	AGA TCA AAG TGG AAA ACC CAT CTG CCT TTA ACC TGA CGT TTA ATA AAT Gln Ile Lys Val Glu Asn Pro Ser Ala Phe Asn Leu Thr Phe Asn Lys	3217	TTC CGT CTG CTG GGT TAC AAG TAT TCG ACC AGC GCG TAT TAC ACC TTC Phe Arg Leu Leu Gly Tyr Lys Tyr Ser Thr Ser Gly Tyr Tyr Thr Phe
1633	TTT ATG CCA ACG GCC GTG ATA TTG AAA AAA CCG GAA TGG TTC CCG CAA Phe Tyr Ala Asn Gly Arg Asp Ile Glu Lys Thr Gly Met Val Pro Ala	3265	CAG GAA GCG ACG GAT GTG CCG ACC GAT GCG GAG TCT TAT AGC CAG Gln Glu Ala Thr Asp Val Arg Thr Asp Ala Asp Ser Ser Tyr Ser Gln
1681	AAG GCT CAT TGA ATA TTG AAC TGC CAG CCG GCA CCG GCA AGG TAA GCG Lys Gly Ser Leu Asn Ile Glu Leu Pro Ala Gly Thr Gly Lys Val Ser	3313	TAC CAC AAA CGT AGT CAG ATT CAG GGC AAC GTG ACG CAG CAA CTG GGC Tyr His Lys Arg Ser Gln Ile Gln Gly Asn Val Thr Gln Gln Leu Gly
1729	AAG TTA AAT ACA ACA TTA TTA ATG ACT TTG GCA CTG CTG GCG ACA TGT Glu Val Lys Tyr Asn Ile Ile Asn Asp Phe Gly Thr Ala Gly Asp Met	3361	GCC TGG GCG TCG GTC TAT TTT AAC GTC ACG CAG GAC TAC TGG AAC Ala Trp Gly Ser Val Tyr Phe Asn Val Thr Gln Gln Asp Tyr Trp Asn
1777	TGA CAC AGC GCG TTA ACT AAC ACG TTT TAA AGG ATT ATT ACT ATG ACA Leu Thr Gln Arg Val Asn	3409	GAT GAA GGT AAA CAG CGT TCG ATG AAT GCC GGT TAT AAC GGC CGT ATT Asp Glu Gly Lys Gln Arg Ser Leu Asn Ala Gly Tyr Asn Gly Arg Ile
1825	TGG ACG CAT CTT CCT CTG GGC AAT AAG ACC TCG CGT TTC ACG CAG TCT Trp Thr His Leu Pro Leu Gly Asn Lys Thr Ser Arg Phe Thr Gln Ser	3457	GCC CGC GTG AAC TAC AGC GTT GCT TAC ACC TGG ACG AAA AGC CCG GAG Gly Arg Val Asn Tyr Ser Val Ala Tyr Thr Thr Thr Thr Thr Thr Thr Thr
		3505	TGG GAT GAG AGC GAT CGT TTA CTG TCA TTC TCC ATG TCG ATT CCA CTG Trp Asp Gly Ser Asp Arg Leu Leu Ser Phe Ser Met Ser Ile Pro Leu
		3553	GGA CCG GTG TGG AGT AAC TAC CAC CTC ACG ACC GAT CAG CAT GGC CGA Gly Arg Val Trp Ser Asn Tyr His Leu Thr Thr Asp Gln His Gly Arg

FIG. 4. Nucleotide sequence of a 7,065-bp DNA fragment encoding LP fimbriae. The positions and deduced amino acid sequences of *orf1*, *lpfA*, *lpfB*, *lpfC*, *lpfD*, *lpfE*, and *orf2* are indicated. The numbers on the left indicate the position of the first nucleotide in each line.

positions indicated in Fig. 2 (data not shown). These genes were found to be located at the same positions as the corresponding *E. coli* genes, indicating that no other large insertions are present in either species.

Homology of LpfABCDE to other fimbrial proteins. The deduced amino acid sequences of all five *lpf* genes showed significant homology to those of proteins from various bacterial fimbrial systems (Fig. 6). Furthermore, the gene order

3601 ACC AAC CAG CAG TTA GGG GTG AGC GGC ACC GCG CTG GAA GAC CAC AAC Thr Asn Gln Gln Leu Gly Val Ser Gly Thr Ala Leu Gly Asp His Asn	5329 CCG CCT CAA CGA CCA TTT CGG CTT ACC CTA TCA GTA CCA CCG GCA AAT Thr Ala Ser Thr Thr Ile Ser Ala Tyr Pro Ile Ser Thr Thr Gly Lys
3649 CTG AAC TAT AGT GTG CAG GAA GGC TAC GGC AGC AAC GGC GTG GGT AAC Leu Asn Tyr Ser Val Gln Gly Tyr Gly Ser Asn Gly Val Gly Asn	5377 TGC CGG CCG CCG GGG ATT TCG AGG GAA TTG CCA CCA TGC GTA TTG ATG Leu Pro Ala Ala Gly Asp Phe Gly Ile Ala Thr Met Arg Ile Asp
3697 AGC GGC AGC GTG AAC CTG GAT TAC CAG GGC GGC GTG GGT AGC GCC AGC Ser Gly Ser Val Asn Leu Asp Tyr Gln Gly Gly Val Gly Ser Ala Ser	5425 TGG AGT AAG CAG GAT GAA AAA CCT TCA TGC TTT GAT GCC AGC GTG TTTA Val Glu Met Lys Asn Leu His Ala Leu Met Pro Ala Cys Leu
3745 CTG GGT TAC AAC CAC AAC CGT GAC GGC CAG CAG GTA GAA CTA CCG TTT Leu Gly Tyr Asn His Asn Arg Asp Gly Gln Gln Val Glu Leu Arg Phe	5474 CTG CTT ACC GCT TCC GCG ATG GCG GCA CCG TCG AAT ATC GGT TCT GCT Gly Gly Val Ile Ala His Ser Glu Gly Ile Thr Leu Ser Gln Pro
3793 GGC GGC GGT GTA ATA GCC CAT AGC GAA GGT ATC ACT CTT TCT CAA CCG Gly Gly Val Ile Ala His Ser Glu Gly Ile Thr Leu Ser Gln Pro	5522 GGT GAT ATC CAC TTT ACC ATT ACT ATT AAG GCG GCT ACC TGT GAA CTG Leu Asp Ile His Phe Thr Thr Ile Thr Ile Lys Ala Ala Thr Cys Glu Leu
3841 CTG GGT GAA TCC ARG GCC ATT ATC TCC GCG CCG GGC GCG CCG CCG CCG Leu Gly Glu Ser Met Ala Ile Ile Ser Ala Pro Gly Ala Arg Phe Ser Ala Arg	5570 GAA AAC GAC AGT ATC GAC GTC AAT GAT GAG ACC GTG GTG CTT CAG CCG Glu Asn Asp Ser Ile Asp Val Asn Met Glu Thr Val Val Leu Gln Arg
3889 CAC GTG ATC AAC AAC GGT GGT GTG GAA GTG GAC TGG ATG GGT AAT GCG His Val Ile Asn Asn Gly Gly Val Glu Val Asp Trp Met Gly Asn Ala	5618 CCG GTA AAA GTG GGT AAA GAG CTG AAC CAG AAA AAC TTT AGC ATC GGC Pro Val Lys Val Gly Lys Glu Leu Lys Ala Ser Lys Thr Ser Phe Ile Gly
3937 GTC GTA CCT TAC CTT ACT CCG TAC CGT GAA ACG GAA GTC TCA CTG CGA Val Val Pro Tyr Leu Thr Pro Tyr Arg Glu Thr Glu Val Ser Leu Arg	5666 TTA AAA GAT TGC GCG TAT GCC ACA AAG GCC AGC GTT ACG ATG GAC GGT Leu Lys Asp Cys Ala Tyr Ala Thr Lys Ala Ser Val Thr Met Asp Gly
3985 AGC GAC AGC CTG AAC AAC CAG GTT GAC CTG GAT ACC GCC TCC GTG AAC Ser Asp Ser Leu Asn Asn Gln Val Asp Leu Asp Thr Ala Ser Val Asn	5714 TCT CCG GAC CCG ACT GAC CCC TCG CTT TTT GCC CTG AAT GAC GCG GGC Ser Pro Asp Pro Thr Asp Pro Ser Leu Phe Ala Leu Asp Ser Gly Gly
4033 GTA GTG CCG ACA CCG GCG GCG ATT GTT CGT GCC CCG TTC GAT ACC CGA Val Val Pro Thr Arg Ala Ile Val Arg Ala Thr Arg Ala Thr Ser Arg	5762 GCG ACG GGC GTG GCG TTA AAA ATT AAA ACA TCT GGT GGG GAG CAA CAA Ala Thr Gly Val Ala Leu Lys Ile Lys Thr Ser Gly Gly Glu Gln Gln
4081 GTG GGC TAT CGT GTG CTG ATG AAT CTG ACG CAG GCC AAT GGC AAA GCG Val Gly Tyr Arg Val Leu Met Asn Leu Thr Gln Ala Asn Gly Lys Ala	5810 TAC CCC TCC AGT ACC GAC TCT ACG CCT GTC GAA CAC ACT GTC TGG TTT Tyr Pro Ser Ser Thr Thr Asp Ser Thr Pro Val Glu His Thr Val Trp Phe
4129 GTG CCG TTT GGT GCT ACC GCG ACG CTG CTG GAT ACC ACA AAA GAG TCC Val Pro Phe Gly Ala Thr Ala Thr Leu Leu Asp Thr Thr Lys Glu Ser	5858 GAT GGT ACG AAC AAG CTG AAC TAT ATC GCC AGC TAT GTG CCT GTT AAG Asp Gly Thr Asn Lys Leu Asn Tyr Ile Ala Ser Tyr Val Pro Val Lys
4177 AGC AGC ATT GTG GGT GAA GAC GGT CAG CTT TAT ATC AGC GGG ATG CCG Ser Ser Ile Val Gly Glu Asp Ser Gly Met Arg	5906 CCG GAT GCC ACC GTT GGC ACA GCG AAT GCG ACG GTG AAT TTT AGC GTC Pro Asp Ala Thr Val Gly Thr Ala Asn Ala Thr Val Asn Phe Ser Val
4225 CAG AAA GGT GCC CTT CAG GTG AAC TGG GGT AAA GAC CAG GCA CAG CAA Gln Lys Gly Ala Leu Gln Val Asn Trp Gly Lys Asp Gln Ala Gln Gln	5954 ACA TAC GAA TAA TCA CTG AGG GCC AGT TCG CTG GCC TTT TCC ATT TT
4273 TGC CCG GTG CCG TTT CAG CTG CCG GAA CAA CAG GAT AAT ACC GCG GTG Cys Arg Val Ala Phe Thr Leu Pro Glu Gln Gln Asp Asn Thr Gly Val	6001 AGT GAT TTT TTG TAA AAA CTT CTC CGA TCA CAC TCT CCG TTG CCA CTT 6049 TCC CTC TGC TGG TGG TCT ACT TAA CCG TGC TTG TAG ACT TTC TAA CAT 6097 TCT CCG TTA GCC CCA TAT TTA CCT GTT TGA TTA AAG AAT CGT TAC CAT
4321 GTG ATG GCG AAT GCC GTC TGC CCG TAA CAG GGA AGG AAA CGA TTA TGT Val Met Ala Asn Ala Val Cys Arg Met	6145 GGT CGT TTT TAT CCC GAT TCC CCA GGG TTT GTT TGC ATG AGA TAT ATT Met Arg Tyr Ile
4369 TGA AAA AGT TGA TAA TGT TTA CCG GCC TGT TGG CCG GGT CCG TGC TGT Leu Lys Lys Leu Ile Met Phe Thr Gly Leu Leu Gly Gly Ser Val Leu	6193 AAA TCG ATG ACG CAA CAG AAA CTT AGT TTC TTG CTT GCG CTC TAT ATC Lys Ser Ser Met Thr Gln Gln Lys Leu Ser Phe Thr Ser Ala Thr Tyr Ile
4417 TTT CCG GGC AGG CCG TGG CAG CCG CAG ATT TTG GAC CAT GTA CTC CTG Phe Ser Gly Gln Ala Leu Ala Asp Phe Gly Pro Cys Thr Pro	6241 GGT CTG TTT ATG AAT TGC GCC GTG TTT TAC CGC CGT TTC GGC AGT TAT Gly Gly Thr His Ile Phe Ser Ala Thr Ile Asn Lys Thr Val Ser
4465 AAG GTG GAA CAC ATA TCT TCA GTG CCA CCA TAA ATA AAA CAG TTT CAG Gly Gly Thr His Ile Phe Ser Ala Thr Ile Asn Lys Thr Val Ser	6289 GCA CAA GAA TTT ACC ATT TGG AAA GGC CTC TCC GCA GTT GTC GAA CTG Ala Gln Glu Phe Thr Ile Trp Lys Gly Leu Ser Ala Val Val Glu Leu
4513 ATA CGT CAA AGA ACA CCG GTG CGA CCT TCG TAG ATT TCG ATA GCT Asp Thr Ser Lys Asn Thr Thr Gly Ala Thr Phe Val Asp Phe Asp Ser	6337 GGC GCC ACG GTG CTG GTC ACT TTC TTA CTT CGT CTT CTT TCA CTG Thr Asn Leu Gly Thr Tyr Ala Leu Met Ser Cys Glu Cys Pro Asp ATA
4561 GGA ATT TAG GTG GAA CCT ATG CGA TGT CCT GTG AAT GCC CTG ATG ATA Thr Asn Leu Gly Thr Tyr Ala Leu Met Ser Cys Glu Cys Pro Asp ATA	6385 TTT GGC CGA CCG GTC TGG CGT GTG CTG GCC ACG CTG GTG GTG CTG TTT Phe Gly Arg Arg Val Trp Arg Val Leu Ala Thr Leu Val Val Leu Phe
4609 CCT CTC TTA TAA ATG ACA CCT TAT TTA AGG CTG TGG TTC CTC TGG CCT Thr Ser Leu Ile Asn Asp Thr Leu Phe Lys Ala Val Val Pro Leu Ala	6433 TCC GCT GGC GCC AGT TAT TAC ATG ACC TTC CTG AAC GTG ATG ATT GGC Ser Ala Gly Ala Ser Tyr Tyr Met Thr Phe Leu Asn Val Val Ile Gly
4657 TTG TTA CGA ATA TAG AGA GTC GCT CCT ATT ACC AGA TCA ATA ATA ATA Phe Val Thr Asn Ile Glu Ser Arg Ser Tyr Tyr Gln Ile Asn Asn Asn	6481 TAC GGC ATT ATT GCG TCT GTT ATG ACC ACC GAT ATC GAT CTC TCG AAA Thr Gly Ile Ala Ser Val Met Thr Thr Asp Ile Asp Leu Ser Lys
4705 TTG CCA TTG CGA CCG ATG TAC TGA TTT CCG GGG ACG AGG AGA ATA CGT Ile Ala Ile Ala Thr Asp Val Leu Ile Ser Gly Asp Glu Glu Ser Phe	6529 GAG GTG GTG GGG CTG CAC TTT GFA TTG TGG CTG ATT GCC GTG AGC GTG Glu Val Val Gly Leu His Phe Val Leu Trp Leu Ile Ala Val Ser Val
4753 TAA CAC ACC GTA GTA ACC TGA CAA ACA ACC GCT CTC AGT GTT CCG AAA Leu Thr His Arg Ser Asn Leu Thr Asn Asn Arg Ser Gln Cys Ser Gln	6577 CTT CCG CTC ATC TTT ATC TGG AGT AAC CAC TGT CCG TAC ACG TTG TTG Leu Pro Leu Ile Phe Ile Trp Ser Asn His Cys Arg Tyr Thr Leu Leu
4801 ATG CAA GTA AAG ATG CAA TAT GGA CAT CCG GTG GCA AAG GTC ACT Asn Ala Ser Ser Lys Asp Ala Ile Trp Thr Ser Gly Gly Lys Gly His	6625 CCG CAG CTA CGT ACG CCG GGG CAG CGT TTT CCG AGC GCC GCT GTA GTG Leu Ser Leu Tyr Ile Leu His Pro Phe Val Gly Glu Ser Ile Ile Pro
4849 TAT CCG TCT ATA TTC TCC ATC CGT TTG TGG GTG AAA GTA TTA TAC CTA Leu Ser Leu Tyr Ile Leu His Pro Phe Val Gly Glu Ser Ile Ile Pro	6673 GTA CTC GCA GGC GTA ATG GTG TGG CCG CCT ATT CCG CTG CTG GAT ATA Val Leu Ala Gly Val Met Val Trp Ala Pro Ile Arg Thr Leu Asp Ile
4897 GCA CCA AAA TAA TGG ACC TTT TTG TGA CAA AGA AAC CCA GTG TAT ATG Ser Thr Lys Ile Met Asp Leu Phe Val Thr Lys Lys Pro Ser Val Tyr	6721 CAG CAA AAA AAG GTT GAA CCG GCG ACA GGC ATC GAC TTA CCC AGT TAT Gln Gln Lys Lys Val Glu Arg Ala Thr Gly Ile Asp Leu Pro Ser Tyr
4945 GCA GTA TAC CTG CGT COT CTG TAT ATA TCA GTG GTT CAA TTA CCG TAC Gly Ser Ile Pro Ala Ser Ser Val Tyr Ile Ser Gly Ser Ile Thr Val	6769 GGC GGC GTG GTG CCG AAC TCC TAC CTG CCC TCA AAC TGG TTA TCT GCG Gly Gly Val Val Ala Asn Ser Tyr Leu Pro Ser Asn Trp Leu Ser Ala
4993 CTC AGG GCT GTG AAC TCT CCA GCG GCA GCA CCG TGG AAA TTC CGT TTG Pro Gln Gly Cys Glu Leu Ser Ser Thr Leu Glu Ile Pro Phe	6817 TTA GGG CTG TAT GCC TGG CCG CAG GTA GAT GAG TCG TCG GAC AAT AAT Leu Gly Leu Tyr Ala Trp Ala Gln Val Asp Glu Ser Ser Asp Asn Asn
5041 GGG AAT TTA AGG CCA CTG ATT TTA ARG ATC GCA AAG GAC AAG TTG CAA Gly Glu Phe Lys Ala Thr Asp Phe Lys Asp Arg Lys Gly Gln Val Ala	6865 TCG TTA ATA AAC CCG GCC AGG AAA TTT ACC TAT GTT GCG CCG AAA GAT Ser Leu Ile Asn Pro Ala Arg Lys Phe Thr Tyr Val Ala Pro Lys Asp
5089 AGA ACG CCA CGA AAT TCA CCA AAG AGC TGC AGT TTA AAT GCA CCA ATA Lys Asn Ala Thr Lys Phe Thr Lys Glu Leu Gln Phe Lys Cys Thr Asn	6913 GGG GAT GAC ACC TAC GTC GTT TTC ATT ATC GGT GAG ACG ACC CGT TGG Ile Ser Asp Gly Val Lys Ile Phe Leu Met Ser Tyr Val Val Phe Ile Ile Leu Gly Thr Arg Trp
5137 TTT CCG ATG GCG TAA AGA TCT TCC TGC GTA TTG AGG GAA TGC CAA ACG Ala Asn Asp Ser Ser Asn Ala Ile Asp Met Gly Asn Pro Asp Ile Gly Ala	6961 GAT CAC ATG GGG ATT TTC GGC TAC GAG COT AAT ACC ACC CCG AAG CTG Gly His Met Gly Ile Phe Gly Tyr Glu Arg Asn Thr Thr Pro Lys Leu
5185 CTA ATG ATT CGA ATG CCA TCG ACA TGG GCA ACC CCG ATA TCG GTG CCG Ala Asn Asp Ser Ser Asn Ala Ile Asp Met Gly Asn Pro Asp Ile Gly Ala	7009 GCG CAG GAA AAA AAT CTG GCG CCG TTC CCG GCG TAT TCC TGC GAT ACC Ala Gln Glu Lys Asn Leu Ala Ala Phe Arg Gly Tyr Ser Cys Asp Thr
5233 TCA TTG AGG GCG CTA ACG GTA AAA TTT TGG TGC CAA ATG ACG CCA GTG Val Ile Glu Gly Ala Asn Gly Lys Ile Leu Val Pro Asn Asp Ala Ser	7057 GCG ACG AAG Ala Thr Lys
5281 TTA ATC AGG AGC TGA CCG TAT CCG GTC TTG TTG ACG ACA CCG ACC GTA Val Asn Gln Glu Leu Ser Val Ser Gly Leu Val Asp Asp Thr His Arg	

FIG. 4—Continued.

between *lpfABCDE* and the *fim* operon of *S. typhimurium* was found to be conserved (Fig. 6). These data strongly suggest that *lpfABCDE* constitutes a fimbrial operon. Start and stop codons of the genes *lpfBCDE* are separated by only a few base pairs;

thus, this organization also substantiates the proposed operon structure (Fig. 4). A stem-loop structure marking a potential end of the operon was found downstream of *lpfE* (5' stem bp 5974 to 5979; loop bp 5980 to 5983; 3' stem bp 5984 to 5989;

TABLE 1. Sizes and positions of genes and gene products of the *lpf* operon

Open reading frame	Position (bp) ^a	GC content (%)	Length of polypeptide (amino acids)	Predicted signal sequence (amino acids)	Calculated mass of mature protein (Da)
<i>lpfA</i>	477-1010	47	179	1-24	16,138
<i>lpfB</i>	1098-1793	45	233	1-26	22,898
<i>lpfC</i>	1819-4344	55	843	1-21	90,158
<i>lpfD</i>	4365-5429	47	356	1-24	35,234
<i>lpfE</i>	5438-5962	51	175	1-20	16,335

^a Position in the nucleotide sequence shown in Fig. 4.

$\Delta G(25^\circ\text{C}) = -12.6$ kJ). For several fimbrial operons, a second stem-loop structure between the gene encoding the major fimbrial subunit and downstream genes has been found (7, 18, 26). Similarly, a stem-loop structure between *lpfA* and *lpfB* was found (5' stem bp 1042 to 1052; loop bp 1053 to 1057; 3' stem bp 1058 to 1068; $\Delta G(25^\circ\text{C}) = -22.2$ kJ). By analogy, this suggests that *lpfA* may encode the major fimbrial subunit. This assumption is further supported by the sequence homology

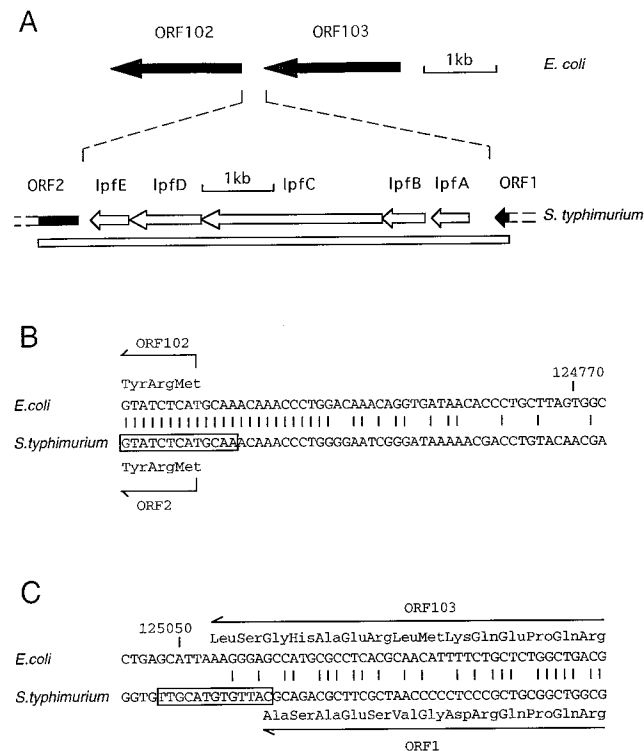


FIG. 5. (A) Comparison of the region of the *S. typhimurium* chromosome containing the *lpf* operon with the corresponding *E. coli* region. The positions of the genes are indicated by arrows. Solid arrows represent genes which are present in *E. coli* and *S. typhimurium*, while open arrows represent genes which are present only in *S. typhimurium*. The bar indicates the position of the nucleotide sequence shown in Fig. 4. The boundaries of sequence homology between DNA from *S. typhimurium* and *E. coli* are indicated by dashed lines. (B) Comparison of the nucleotide sequences at the left boundaries of sequence homology of *S. typhimurium* and *E. coli*. The start points of *orf2* and *orf102* are indicated by arrows. The numbers above the *E. coli* sequence refer to the position in the database entry U00039. The position of an inverted repeat found at the boundary of the sequence homology is boxed. (C) Comparison of the nucleotide sequences at the right boundaries of sequence homology of *S. typhimurium* and *E. coli*. The ends of *orf1* and *orf103* are indicated by arrows. The numbers above the *E. coli* sequence refer to the position in the database entry U00039. The position of an inverted repeat found at the boundary of the sequence homology is boxed.

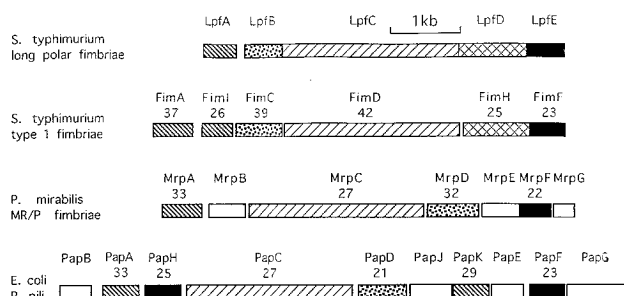


FIG. 6. Comparison of the deduced amino acid sequences and the gene orders of fimbrial operons from *E. coli* (*pap* operon), *Proteus mirabilis* (*mrp* operon), and *S. typhimurium* (*fim* operon) with those of the *lpf* operon. The position and length of a gene in an operon are indicated by bars. Open bars indicate genes which have no homologous partner in the *lpf* operon. Homologs to *lpf* genes in other fimbrial operons are indicated by identical bars. The numbers indicate the percent sequence identities between the deduced amino acid sequences of the *lpf* genes and homologs in other fimbrial operons.

found between *LpfA* and major subunits from other fimbrial systems (Fig. 7A and B). The sequence homology to other fimbrial systems suggests that *LpfB*, *LpfC*, and *LpfE* function as chaperonin, the usher, and the minor fimbrial subunit, respectively (Fig. 7C, D, and F). *S. typhimurium* *FimH* was the only homolog of *LpfD* found in the current protein database (Fig. 7E). So far, no function has been proposed for this protein.

Analysis of the *lpfA* promoter region. In order to identify potential regulatory sequences, the *lpfA* promoter region was assayed for the presence of GATC sites and direct or inverted repeats. Recently, phase variation of fimbrial operons mediated by differential methylation of GATC sites has been described (9, 57). However, no GATC sites were found in the promoter region of *lpfA*, making a regulation by this mechanism unlikely.

A 9-bp DNA sequence was found to form four direct repeats and one inverted repeat located between bp 263 and 388 in the *lpfA* promoter region (Fig. 8A). This sequence was also found to form an inverted repeat in the *fimA* promoter region of *S. typhimurium* and *S. typhi* (46, 47). The positions of these short repeats overlap a putative $\sigma 70$ promoter found upstream of *lpfA* (-35 sequence, bp 293 to 298; -10 sequence, bp 315 to 320). However, further experiments are necessary to establish a regulatory role for these repeats.

An inverted repeat was found to flank the open reading frame *lpfA* (Fig. 8B). The nucleotide sequence of this inverted repeat was similar to that of the inverted repeat which is thought to be involved in an inversion event controlling fimbrial phase variation in the *fim* operon of *S. typhimurium* and *S. typhi* (46, 47).

Expression of the *lpf* operon in *E. coli* and identification of long polar (LP) fimbriae. In order to determine whether pMS1000 encodes *S. typhimurium* fimbriae, this cosmid was introduced into nonfimbriated *E. coli* ORN172 (59). This strain carries a deletion in the *fim* operon and has been shown by electron microscopy not to express fimbriae. Fimbrial crude extracts from strains ORN172 and ORN172(pMS1000) were prepared, and the protein profiles were analyzed (Fig. 9). The fimbrial crude extract of strain ORN172(pMS1000) contained a band with an apparent molecular mass of 22 kDa which was not present in the crude extract of strain ORN172 (Fig. 9). The size of this protein was slightly larger than the calculated size for the major form of *LpfA* and *LpfE* (Table 1). The presence of proteins in a fimbrial crude extract is generally considered evidence for the expression of fimbriae. Therefore, these data

A		B	
FimA LpfA	MRHKLMSTIASLMPVAGA...ADPTVSVSGGTIFHFGKLVNAACAVST MEF-LMKKVVFAL--SALAVVSTSAFAAESGDGTIKFTGETVDPACVVSQ * * * * *	FimI LpfA	MSPVIAHAVM-----ESGRVHLRGLQVNVGGCAVATESQ MEFLMKKVVFALSALAVVSTSAFAAESGDGTIKFTGETVDPACVVSQ * * * * *
FimA LpfA	KSADQVTVLGGQYRTASFTAIGNTTAQVFFSIVLNDCKPKVAANAFAVSG DSQNEVVLGGQYKKNIFKAIKDGKSSKPFQIKLEDCDITSNKVNVSFNG * * * * *	FimI LpfA	NLRVLMGQYRTAFTGPGSFAPVSVFSLRLISCSADEVWRHVGIFAGVGT NQEVVLGGQYKKNIFKAIKDGKSSKPFQIKLEDCDITSNKVNVSFNG * * * * *
FimA LpfA	QADNTNPNLLAVSSADNSTTATGVGIEILDNNTSPLKPDGATFSAKQSLV VGDTDATLVSVNT--EAGAAATGVGIGIYDNANKLVEMNTG--KSTTTLA * * * * *	FimI LpfA	PAEDPHVFLASGEGIGNAGLGLFDDQQRQIIPNTLPLHPTILTSEM DTDDATLVSVNTTGAAGATGVGIGIYDNANKLVEMNT--GKSTTTLAAGQTV * * * * *
FimA LpfA	EGTNTLRFTARYKATAAATTPQANADATFIMKYE AGQTVLYYTANYVATKDTVTVTGYGNAEVDNFLSYE * * * * *	FimI LpfA	LHFTARYRAISENMTPGRIHSEVWFVLVYP LYYTANYVATKDTVTVTGYGNAEVDNFLSYE * * * * *
identity:	37%	identity:	26%
conservative amino acid changes:	24%	conservative amino acid changes:	26%
C		D	
FimC LpfB	MLNSIKVGFIVL-LTLFTSLNVQAAGGIALGATRVIPYSAAKQTSLAISN MNRSLISCTALVLAIAQNSP--AGGVALSSTRVYDGSRRKASLTVNN * * * * *	fimD LpFC	MKKTWTFAGRFPYVGSPLSSVALSVALAALPLTSRGESYFNFPAFLS--AD M--TWTHLPLGNKTSRFTQSSALALMIA-GTLPAYAGT-FNPRFLBDVPG * * * * *
FimC LpfB	SDTQERYLVNSWIENAGQKERTFIVTTPPLFVSEPKSENTLRIIYAGQPL KSTTDEFLIQSWIDDANGNKTPFIIITPPLKLSPTKNNVLRIVNTTNTL * * * * *	fimD LpFC	TASVADLSRFEGKGNHP-PGIRYVDLWRNDEFVATQDIRFEAGAVCTGDK IDQVLDLSMYSNKAEHLPGKYRVSVVNEKMEKSRLEFKAATPEAQRK * * * * *
FimC LpfB	PGDRESLFWNNKAIKPSVDKSHIEGKLVQLAAILSRKLFVRPANLPQTP PQDRSVYWINVAIPA-KSEDAEAKNVLIQAVRRLKLFYRPAALKANS * * * * *	fimD LpFC	SG-GLMPCPTFEWIKRIGVNTAAFPVSDKQVDTTCHLPEKIPGAEVAFD MGSVLVCLSRVQLQEDMGRVDEFPALMAPPEACVAFDDIIIPQAASHFD * * * * *
FimC LpfB	EDAPTLLKFSRVG-NHLKITNPSAYYLTAVNIVSAGAKKIDNV-MIAPKSD MDGNLQPTSGAGNQIKVENSPAFNLTFNKFYANGRDIERTGMVPAKGS * * * * *	fimD LpFC	FASMRNLISLPOASLNSARGYIPPEWDEGIPALINVSFTGSRGT--- FADQTLIMSFPQAKMKTQARTGTVPESQWDEGVNALLVDYNSFGSNAYDA * * * * *
FimC LpfB	MQIPLPTGAQ--GNVTFQSVNDYGALTSATTASLG LNIELPAGTGVSEVKYNIINDPFGTAGDMLTQRVN * * * * *	fimD LpFC	-----DSISYFLSLLSGLNYGFWLRNNGANWYKSGDGHYSQRWNIG HDSETSNSDSYLNLRSGMNLGAWRLRNYSTW--TRNDG--NNTWDNIG * * * * *
identity:	39%	fimD LpFC	TWQRAIIPKSELVMDGNTGNDVDFSGFRGARLYSSDNMYPDSLQGY TSLRAIVPLKSLTLGDTSTAGDIFDSVQMRGQLTSDRMLPDSQRGF * * * * *
conservative amino acid changes:	30%	fimD LpFC	APTVRGIARTAAKLTIRQNGYVIVYQSVSPGAFITDLNPTSSGGDLEVT APVIRGLIASKNAEIVTEQNMYVYRFTVQPGAFIENDLYPTNSGDLVVT * * * * *
E		fimD LpFC	VDEKQGSQRYTVPYSTVPLLRQREARVYDLVAGDFRSGNSQQSSPFFQ IKESDGEQKFPVQPPSSVALLQREGLHKYLSLAGEYRAGNYSABPKFGQ * * * * *
FimH LpFD	M---KIYALLAGTALFFTHPALATV---CRNSNGTADTIDYFLSDV MLKKLIMPTGLL--GSSVLFSGQALAAADFGPCTPEGG--THIFSATINK * * * * *	fimD LpFC	GTVIAGLPAGLTAYGGIQLADRYRAVYVAGARNLGDWGAVSVVDTHARQ LDAMIYGLPYGFTVGGAIPTDDYISLAGLGNKNGYICAIISIDVTQAKSK * * * * *
FimH LpFD	FTSGNPPQGVVTLPEKSAW----GVNATCPAGTIV--NYTYSYVSE TVSDTSKNTGATFVDPDSNLGCTYAMSCBDDTSLINDTLKPAVVL * * * * *	fimD LpFC	LADDSTHQQSLRPLYAKSLANNYGTNFQLGYRSTRGTYTLDVAYRSM LANEENSEGGYRFLYSKFSN-GTDFRLGLKYSTSGYTFQEA--- * * * * *
FimH LpFD	LPVQSTEGNFYKLNDYLLGAMSITDSVAGVSYPPRNYILMGVDY---N AFVTNIESRSYYQINNNIATADVLISGDEENTLTHRSNLTNRSQCSON * * * * *	fimD LpFC	EGYDYEDSDGRRHKVPAQSYHNLRYSKGRFQVNISQNLGDYGSGLYLS ---DVRTDADS-----SYSQ--YHKRSQIQGNVTQQLGAWGVSYFN * * * * *
FimH LpFD	VSQKFPFGVQDSKLVFLKLVIRPFNMVITIPRQTFVYVYTTSGDALST ASSKDAIWTSGGKHLISLYILHFPVGESIIPSTKIMDLVFTKPPSVYGS * * * * *	fimD LpFC	GSQNYWNTADTNTWYQLGYASGWQGISYLSWSWSSESVSGSGADRILAF VTQQDYWNBEGQRSLNAGYNGRIGRVNYSVAYTWKSPWEDSDRLLSF * * * * *
FimH LpFD	PVYTISSYSGKVEVPPQNCVENAGQVVEFDGIGASLFSQAGAGNRQGV PASSVYISGSIIVPQGCLESSTLELIPFGEFKATDPKDR-KGQVAKNAT * * * * *	fimD LpFC	NMSVFPVSLTGRRYARDTILDRTYATFNANRNRDGNSWQTGVGGTLEEG SMSIP-----LGRVMSNYHLTDQHGRTNQQLGWSGTALED * * * * *
FimH LpFD	PQTKTIAIKTNVAQAQYLSMRLEA--EKASGQAMVSDNPLDGLFVVANS KFTKELQPKCTNISDGVKIFLRIGMPPNANDSNAIDMGNPDIGAVLEGAN * * * * *	fimD LpFC	RNLSYSVTQGRSSN-GYSGSASASQATYGTGLGVYNYDRDQHDYMQ HNLNYSVQEGYSGNVSNGSGLVLDYQGVGASLGYNHNDRDQGVLELRF * * * * *
FimH LpFD	GTPLTPNNLSSKIPHL-----DNAAARVGRAPVSVTGKIPAGPFT GKILVNDASVQELSVSGLVDDTHRTASTTISAPYISTTGLPAAGDFE * * * * *	fimD LpFC	SGGVGHADGITFSQPLGDTNVLIKAPGARGVRIENQTVKTDWRGYAVM GGGVIAHSEGITLSQPLGESMAIIISAPGARARHVINNGVEVDWMGNVAV * * * * *
FimH LpFD	ARGYLVDYD GIATMRIDVE * * * * *	fimD LpFC	DYATVYVYRNRVALDNTNMDNHDVEMNVSVVVTEGALVRAAFDTRIGVR PYLTPYRETEVSLRSDSLANQVLDLTSVNVVPTGAIIVRARFDRVGYR * * * * *
identity:	25%	fimD LpFC	AIIT-ARLGGRLPFGAIVR--EFSAGTSMVGGDQGLVLSGLPLKGLF VLMNLTQANGKAVPFGATATLDTTKESSIVGEGDQLVYSGMQRKALQ * * * * *
conservative amino acid changes:	27%	fimD LpFC	IQWEGKNARCIPYALAEDSLQKAITIASATCIRPAS VNWGDQAOQCRVAFPLPEQDNTGVVMAVAV---- * * * * *
		identity:	42%
		conservative amino acid changes:	29%
		F	
		FimF LpFE	MILRRVFAIGCVLF--SPLSQANSSLEGVNIELRGNVWVPTCAVVAGDS MKNLHALMP-ACLLTASAMAAPSNIGSAGDIHFTITIKAACTEL--END * * * * *
		FimF LpFE	NKSVNLGTWPTQLHAAGDATQPVAFSLKLEGCP-PGSASITFSGTPAPG SIDVNMETVVLQRVVKVGLNQLNPNFISGLKDCAYTKASVMTGSPDPT * * * * *
		FimF LpFE	TALLALADTAMAQKLAIEIRD-GDQRRLEQAKAVDIDN--NGNATLK DPSLFDLSDGGATGVALKIKTSGGQQYPSSTDPVHEVHTVWDFDCTNKL * * * * *
		FimF LpFE	FYANYIALADGVQGLANADATFLINYN YIASYVPVKPDATVGTANATVNFSVIYE * * * * *
		identity:	23%
		conservative amino acid changes:	30%

FIG. 7. Comparison of the primary structures of LpfA and FimA (A), LpfA and FimI (B), LpfB and FimC (C), LpfC and FimD (D), LpfD and FimH (E), and LpfE and FimF (F) by means of the program CLUSTAL. Identical amino acids are indicated by stars, and conservative amino acid changes are indicated by dots. Gaps introduced by the program to optimize the alignment are indicated by dashes.

supported the idea that *lpfABCDE* encode proteins involved in the biosynthesis of fimbriae, as suggested by the sequence homology of their gene products to those of the *fim* genes of *S. typhimurium*.

In order to detect expression of fimbriae directly, strains ORN172 and ORN172(pMS1000) were analyzed by electron microscopy. Fimbriae which were about 7 to 8 nm in width and which displayed a rodlike appearance were detected in strain ORN172(pMS1000) (Fig. 10) but not in strain ORN172 (data not shown). These morphological features are typical for enterobacterial type 1 and type 4 fimbriae. However, unlike type 1 fimbriae, the fimbriae detected on strain ORN172(pMS1000)

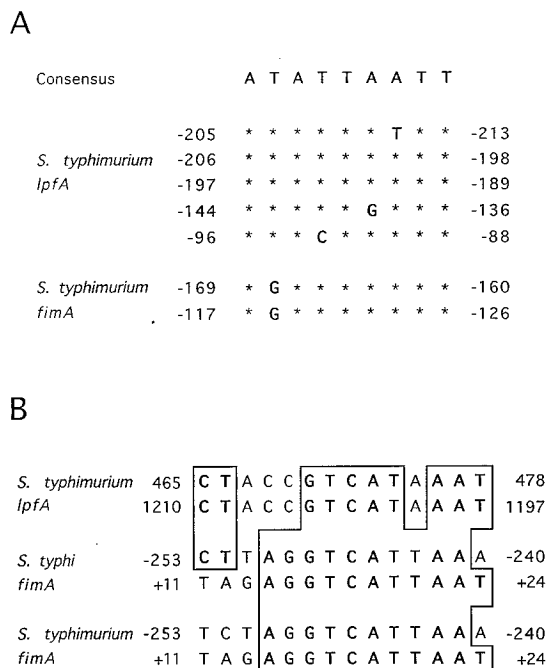


FIG. 8. (A) Comparison of repeats found in the promoter regions of *S. typhimurium* *lpfA* and *fimA*. The numbers indicate the position of the repeat relative to the start codon (+1) of *lpfA* and *fimA*. A consensus sequence of the repeat is given at the top. Bases in the repeats which correspond to those of the consensus sequence are indicated by stars. (B) Comparison of inverted repeats found in front of *fimA* in *S. typhimurium* and in *S. typhi* and flanking *lpfA* in *S. typhimurium*. The numbers for *lpfA* indicate the position of the repeat relative to the sequence shown in Fig. 4. The numbers for *fimA* indicate the position of the repeat relative to the start codon (+1). Conserved nucleotides are boxed.

were polar (Fig. 10) and between 2 and 10 μm long. Sometimes, fimbriae were detected at only one pole of the cell. Occasionally, more than one filament was inserted at the pole of a bacterial cell. About two-thirds of the bacteria examined displayed no fimbriae on their surfaces. Because of their morphology, these fimbriae were termed LP fimbriae.

DISCUSSION

We have identified a new fimbrial operon consisting of five genes, *lpfABCDE*, located at 78 min on the *S. typhimurium* chromosome. Expression of LP fimbriae could be detected by electron microscopy when the *S. typhimurium* operon was introduced into nonfimbriated *E. coli* ORN172. Six distinct fimbriae in *S. typhimurium* and in *S. enteritidis* have so far been described, and one in *S. typhi* has been described (Table 2). However, nucleotide sequence analysis showed that the *lpf*

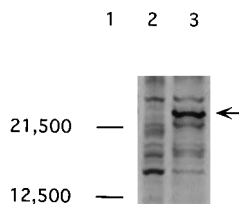


FIG. 9. Tricine-SDS-polyacrylamide gel electrophoresis of fimbrial crude extracts from *E. coli* ORN172 (lane 2) and ORN172(pMS1000) (lane 3). The positions of molecular weight standard proteins are given in lane 1. The position of a fimbrial protein in the crude extract of ORN172(pMS1000) is indicated by the arrow.

operon is distinct from other *Salmonella* operons for which sequence information is available, including *pef*, *fim*, *sef*, and *agf*. In addition, their polar distribution suggested that LP fimbriae are distinct from previously described *Salmonella* type 1, type 3, and GVVPO fimbriae, all of which originate peritrichously from the bacterial surface (Table 2). In fact, LP fimbriae morphologically resemble type 4 fimbriae from other members of the *Enterobacteriaceae* in length, width, and polar distribution (28). However, there is so far no direct evidence that LP fimbriae are coded for by the *lpf* operon. Although unlikely, it is possible that the introduction of the *lpf* operon induces a cryptic fimbrial operon of *E. coli*. A comparison of the primary structures of Lpf proteins indicated that the *lpf* operon is more closely related to the *fim* operon encoding type 1 fimbriae of *S. typhimurium* than to operons encoding type 4 fimbriae. In fact, the *fim* and the *lpf* operons share the same gene order (55, 56).

An unusual feature of the *lpf* gene cluster is the lack of genes encoding regulatory proteins. In the *fim*, *pap*, and *mrp* gene clusters, one or more regulatory genes transcribed in opposite orientations are located upstream or downstream of the fimbrial operon (6, 7, 55). In the *fim* operon, these regulatory genes are thought to mediate the inversion of the *fimA* promoter, which is flanked by a 10-bp inverted repeat. The sequence of this inverted repeat is conserved in *S. typhimurium* and *S. typhi* (46, 47). A similar inverted repeat was found in the *lpf* operon, but instead of flanking the promoter of the *lpf* operon, the repeats were found up- and downstream of *lpfA*, the first gene in the operon (Fig. 8B). Thus, the positions of these repeats suggest that phase variation of the LP fimbriae may be mediated by inversion of the *lpfA* gene. The absence of regulatory genes implies that if inversion of *lpfA* occurs, it must be mediated by proteins coded for at a different locus. The similarity between the inverted repeats found in the *fim* operon and those of the *lpf* operon suggests that the latter may utilize the inversion machinery of type 1 fimbriae. In order to substantiate regulation of the *lpf* operon by phase variation, however, further experiments are necessary.

The presence of *lpf*-related sequences is restricted to *S. choleraesuis* serotypes belonging to *S. choleraesuis* subsp. *choleraesuis* (for current nomenclature on *Salmonella* spp., see reference 31). However, since only a limited number of strains was tested, we cannot exclude the possibility that *lpf*-related sequences are distributed more widely among enterobacteria. Two other fimbrial systems seem also to be restricted to *Salmonella* spp. Plasmid-encoded fimbriae are encoded by the *S. typhimurium* virulence plasmid (26). Related virulence plasmids are present in 11 other serotypes of *S. choleraesuis* subsp. *choleraesuis* but not in *S. typhi*, indicating that *pef* and *lpf* operons show similar distributions (10, 60). Type 1 fimbriae are coded for by a third fimbrial operon conserved only among *Salmonella* serotypes, as has been shown by hybridization (11, 15, 16, 56). Although the mannose-binding minor subunits of enterobacterial type 1 fimbriae are serologically conserved (1), the deduced amino acid sequences of the proteins forming this *Salmonella* adhesin have been shown to be distinct from those of type 1 fimbriae of other members of the *Enterobacteriaceae* (40, 46, 55). In addition, type 1 fimbriae are coded for by genes located at 13 and 98 min on the chromosomes of *E. coli* and *S. typhimurium*, respectively, indicating that *Salmonella* type 1 adhesins were acquired and evolved independently from those present in other genera of the *Enterobacteriaceae*. In contrast to these *Salmonella*-specific adhesin systems, several other *Salmonella* fimbrial systems were found to be distributed more widely among members of the *Enterobacteriaceae*. DNA probes from the *agf* operon of *S. enteritidis* have been used to find

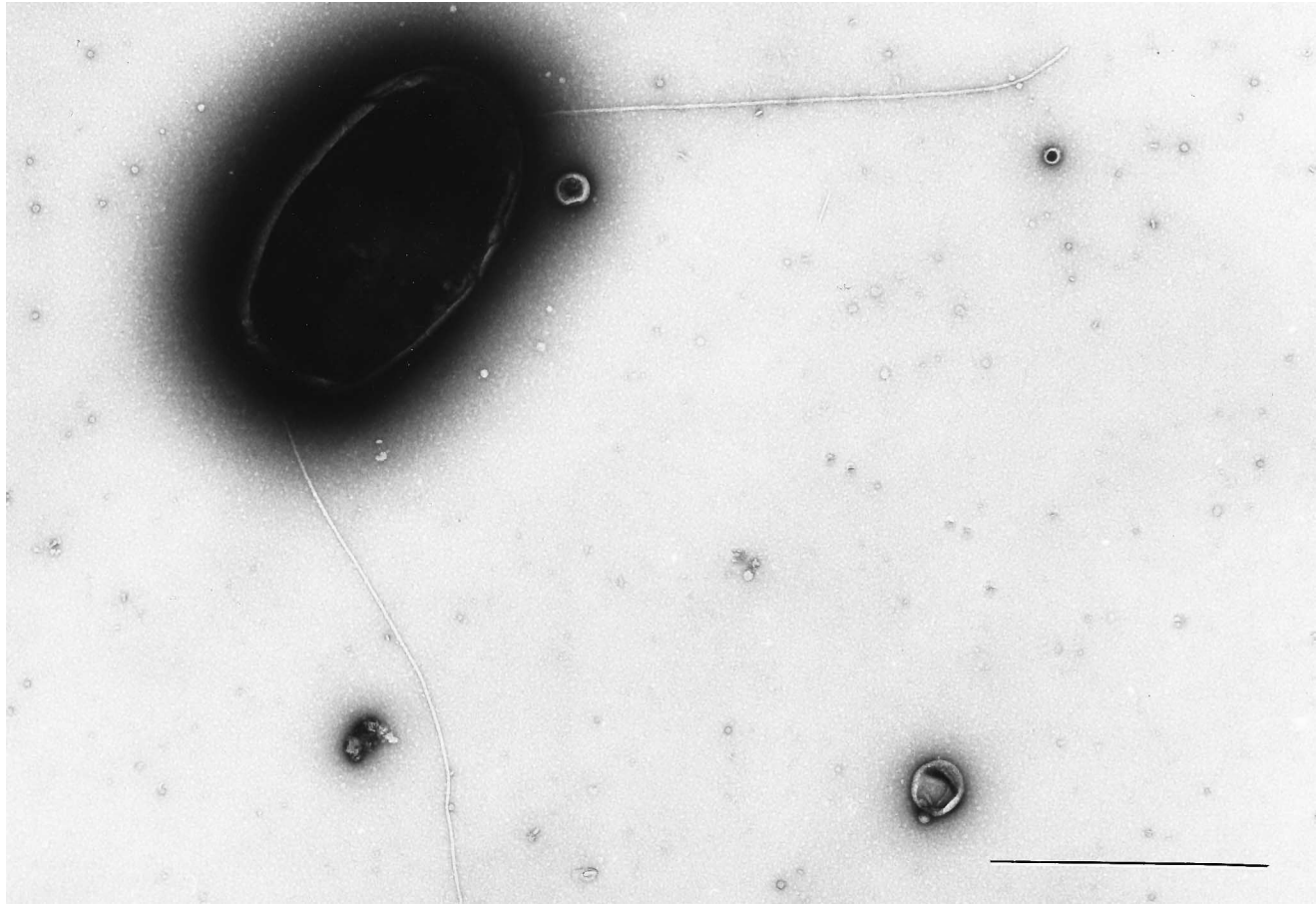


FIG. 10. Electron micrograph of *E. coli* ORN172(pMS1000). Bar, 1 μ m. Final magnification, $\times 36,700$.

related sequences not only in most *Salmonella* serotypes but also in most *E. coli* strains and in some *Shigella*, *Enterobacter*, and *Citrobacter* isolates (21). Similarly, sequences related to genes encoding bundle-forming pili of enteropathogenic *E. coli*

were found in most *Salmonella* serotypes (53). Sequences related to the *sefD* gene of *S. enteritidis* were even more widely distributed among members of the *Enterobacteriaceae*, with homologs being detected in *Salmonella*, *Citrobacter*, *Enterobacter*, and *Shigella* (53).

TABLE 2. Fimbrial systems of *S. typhimurium*, *S. enteritidis*, and *S. typhi*

<i>Salmonella</i> serovar	Fimbrial form	Fimbrial class	Distribution	Width (nm)	Genetic location (min)	Fimbrin gene	Mol wt of fimbrin	Reference(s)
<i>S. typhimurium</i>	LP fimbriae ^a	ND ^b	Polar ^a	7–8 ^a	Chromosome (78)	ND	ND	This study
<i>S. typhimurium</i>	E4	Type 1	ND	ND	ND	ND	17	27
<i>S. typhimurium</i>	F1	Type 1	Peritrichous	7–8	Chromosome (98)	<i>fimA</i>	21	13, 46
<i>S. typhimurium</i>	Type 3 fimbriae	Type 3	Peritrichous	4–5	ND	ND	22	2, 23
<i>S. typhimurium</i>	Bundle-forming pili	Type 4	Polar	7	ND	ND	ND	53
<i>S. typhimurium</i>	Thin aggregative fimbriae	GVVPQ	Peritrichous	3	ND	ND	17	21, 29
<i>S. typhimurium</i>	Plasmid-encoded fimbriae	ND	Peritrichous	ND	Plasmid	<i>pefA</i>	15	26, 32
<i>S. enteritidis</i>	SEF 21	Type 1	Peritrichous	7–8	ND	<i>fimA</i>	21	40
<i>S. enteritidis</i>	Type 3 fimbriae	Type 3	Peritrichous	4–5	ND	ND	22	5, 23
<i>S. enteritidis</i>	SEF 14	Type 3-like	Peritrichous	ND	Chromosome	<i>sefA</i>	14	18, 25
<i>S. enteritidis</i>	SEF 18	Type 3-like	Peritrichous	Thin	Chromosome	<i>sefD</i>	18	17
<i>S. enteritidis</i>	Bundle-forming pili	Type 4	Polar	7	ND	ND	ND	53
<i>S. enteritidis</i>	SEF 17	GVVPQ	Peritrichous	3–4	ND	<i>agfA</i>	17	19
<i>S. typhi</i>	Type 1 fimbriae	Type 1	Peritrichous	7–8	Chromosome	<i>fimA</i>	16	47

^a When expressed in *E. coli*.

^b ND, no data.

bacter, *Erwinia*, *Escherichia*, *Shigella*, *Hafnia*, *Klebsiella*, *Proteus*, and *Providencia* spp. (17).

So far, only the role of type 1 fimbriae in the colonization of host niches has been studied for *Salmonella* spp. The lack of these adhesins was found to have no effect on the colonization of *S. typhimurium* in mice (37). Interestingly, type 1 fimbriae appear to be the only fimbrial system present in *S. typhi* (23, 30). This is also consistent with our finding that *lpf*-related sequences are absent from *S. typhi*. The lack of colonization factors other than type 1 fimbriae in *S. typhi* may reflect the adaptation of this pathogen to humans. The adaptation to a single mammalian host may have abolished the need for multiple fimbrial systems. In fact, type 1 fimbriae of *S. typhi* have been shown to mediate binding to human epithelial cells in vitro (50), supporting a role for these adhesins in the colonization of human host cell surfaces. In contrast to *S. typhi*, most other *S. choleraesuis* serotypes have a wide host range, including various species of livestock and wild animals. The large number of fimbrial systems described for these serotypes may thus reflect a genetic flexibility which is a prerequisite for the ability to colonize such a variety of hosts (Table 2).

The distribution of *lpf*-related sequences among *Salmonella* and related species can be explained by two possible evolutionary scenarios. One possibility is that the *lpf* operon is a recent acquisition of *S. choleraesuis* subsp. *choleraesuis* which occurred after the separation of the lineage into mouse pathogenic serotypes and *S. typhi*. Alternatively, a common ancestor of all *S. choleraesuis* subsp. *choleraesuis* serotypes may have possessed the *lpf* operon, but the operon was subsequently lost from *S. typhi* during its adaptation to its human host. Short inverted repeats flanking the *lpf* operon may be remnants of the ancient recombination event which led to the acquisition of this operon.

ACKNOWLEDGMENTS

We thank P. Orndorff, S. Libby, and J. Kusters for providing bacterial strains; S. Anic for performing the sequencing with the ALF sequenator; P. Stenberg and R. Jones for performing the electron microscopy; R. Tsolis and I. Stojiljkovic for critical reading of the manuscript; and D. Schiefferl for experimental advice and helpful suggestions during this work.

This work was supported by NIH grant ROI AI 22933 to F. Heffron. A.J.B. was supported by fellowship Ba1337/1-2 from the Deutsche Forschungsgemeinschaft.

REFERENCES

- Abraham, S. N., D. Sun, J. B. Dale, and E. H. Beachey. 1988. Conservation of the D-mannose adhesin protein among type 1 fimbriated members of the family *Enterobacteriaceae*. *Nature (London)* **336**:682-684.
- Adegbola, R. A., D. C. Old, and S. Aleksic. 1983. Rare MR/K-like haemagglutinins (and type 3-like fimbriae) of *Salmonella* strains. *FEMS Microbiol. Lett.* **19**:233-238.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403-410.
- Arnqvist, A., A. Olsén, J. Pfeifer, D. G. Russel, and S. Normark. 1992. The Crl protein activates cryptic genes for curly formation and fibronectin binding in *Escherichia coli*. *Mol. Microbiol.* **6**:2443-2452.
- Aslanzadeh, J., and L. J. Paulissen. 1992. Role of type 1 and type 3 fimbriae on the adherence and pathogenesis of *Salmonella enteritidis* in mice. *Microbiol. Immunol.* **36**:351-359.
- Båga, M., M. Göransson, S. Normark, and B. E. Uhlin. 1985. Transcriptional activation of a Pap pilus virulence operon from uropathogenic *Escherichia coli*. *EMBO J.* **4**:3887-3893.
- Bahrani, F. K., and H. L. T. Mobley. 1994. *Proteus mirabilis* MR/P fimbrial operon: genetic organization, nucleotide sequence, and conditions for expression. *J. Bacteriol.* **176**:3412-3419.
- Bäumler, A. J., J. G. Kusters, I. Stojiljkovic, and F. Heffron. 1994. *Salmonella typhimurium* loci involved in survival within macrophages. *Infect. Immun.* **62**:1623-1630.
- Braaten, B. A., X. Nou, L. S. Kaltenbach, and D. A. Low. 1994. Methylation patterns in *pap* regulatory DNA control pyelonephritis-associated pili phase variation in *E. coli*. *Cell* **76**:577-588.
- Brown, D. J., D. S. Munro, and D. J. Platt. 1986. Recognition of the cryptic plasmid, pSLD, by restriction fingerprinting and a study of its incidence in Scottish salmonella isolates. *J. Hyg. Camb.* **97**:193-197.
- Buchanan, K., S. Falkow, R. A. Hull, and S. I. Hull. 1985. Frequency among *Enterobacteriaceae* of the DNA sequences encoding type 1 pili. *J. Bacteriol.* **162**:799-803.
- Chopra, A. K., C. W. Houston, J. W. Peterson, R. Prasad, and J. J. Mekalanos. 1987. Cloning and expression of the *Salmonella* enterotoxin gene. *J. Bacteriol.* **169**:5095-5100.
- Chopra, A. K., J. W. Peterson, P. Chary, and R. Prasad. 1994. Molecular characterization of an enterotoxin from *Salmonella typhimurium*. *Microb. Pathog.* **16**:85-98.
- Clegg, S., and G. F. Gerlach. 1987. Enterobacterial fimbriae. *J. Bacteriol.* **169**:934-938.
- Clegg, S., S. Hull, R. Hull, and J. Pruckler. 1985. Construction and comparison of recombinant plasmids encoding type 1 fimbriae of members of the family *Enterobacteriaceae*. *Infect. Immun.* **48**:275-279.
- Clegg, S., B. K. Purcell, and J. Pruckler. 1987. Characterization of genes encoding type 1 fimbriae of *Klebsiella pneumoniae*, *Salmonella typhimurium*, and *Serratia marcescens*. *Infect. Immun.* **55**:281-287.
- Clouthier, S. C., S. K. Collinson, and W. W. Kay. 1994. Unique fimbriae-like structures encoded by *sefD* of the SEF14 fimbrial gene cluster of *Salmonella enteritidis*. *Mol. Microbiol.* **12**:893-903.
- Clouthier, S. C., K.-H. Müller, J. L. Doran, S. K. Collinson, and W. W. Kay. 1993. Characterization of three fimbrial genes, *sefABC*, of *Salmonella enteritidis*. *J. Bacteriol.* **175**:2523-2533.
- Collinson, S. K., S. K. Emödy, T. J. Trust, and W. W. Kay. 1992. Thin, aggregative fimbriae from diarrheagenic *Escherichia coli*. *J. Bacteriol.* **174**:4490-4495.
- Crichton, P. B., D. E. Yakubu, D. C. Old, and S. Clegg. 1989. Immunological and genetical relatedness of type 1 and type 2 fimbriae in salmonellas of serotype gallinarum, pullorum and typhimurium. *J. Appl. Bacteriol.* **67**:283-291.
- Doran, J. L., S. K. Collinson, J. Burian, G. Sarlós, E. C. D. Todd, C. K. Munro, C. M. Kay, P. A. Banser, P. I. Peterkin, and W. W. Kay. 1993. DNA-based diagnostic tests for *Salmonella* species targeting *agfA*, the structural gene for thin, aggregative fimbriae. *J. Clin. Microbiol.* **31**:2263-2273.
- Duguid, J. P. 1959. Fimbriae and adhesive properties in *Klebsiella* strains. *J. Gen. Microbiol.* **21**:271-286.
- Duguid, J. P., E. S. Anderson, and I. Campbell. 1966. Fimbriae and adhesive properties in *Salmonellae*. *J. Pathol. Bacteriol.* **92**:107-137.
- Fasman, G. D. 1976. CRC handbook of biochemistry and molecular biology. CRC Press, Cleveland.
- Feutrier, J., W. W. Kay, and T. J. Trust. 1986. Purification and characterization of fimbriae from *Salmonella enteritidis*. *J. Bacteriol.* **168**:221-227.
- Friedrich, M. J., N. E. Kinsey, J. Vila, and R. J. Kadner. 1993. Nucleotide sequence of a 13.9 kb segment of the 90 kb virulence plasmid of *Salmonella typhimurium*: the presence of fimbrial biosynthetic genes. *Mol. Microbiol.* **8**:543-558.
- Ghosh, S., A. Mittal, and N. K. Ganguly. 1994. Purification and characterization of distinct type of mannose-sensitive fimbriae from *Salmonella typhimurium*. *FEMS Microbiol. Lett.* **115**:229-234.
- Girón, J. A., M. M. Levine, and J. B. Kaper. 1994. Longus: a long pilus ultrastructure produced by human enterotoxigenic *Escherichia coli*. *Mol. Microbiol.* **12**:71-82.
- Grund, S., and A. Seiler. 1993. Elektronenmikroskopische Untersuchungen zu Fimbrien und Lektinphagozytose bei *Salmonella typhimurium* varieto copenhagen (STMVC). *J. Vet. Med. B* **40**:105-112.
- Haluala, M. C., and B. A. D. Stocker. 1987. Distribution and properties of the mannose-resistant hemagglutinin produced by *Salmonella* species. *Microb. Pathog.* **3**:455-459.
- Holt, J. G., N. R. Krieg, P. H. A. Sneath, J. T. Staley, and S. T. Williams. 1994. Bergey's manual of determinative bacteriology, 9th ed., p. 186-187. Williams & Wilkins, London.
- Jones, G. W., D. K. Svinarich, and H. J. Whitfield. 1982. Association of adhesive, invasive, and virulent phenotypes of *Salmonella typhimurium* with autonomous 60-megadalton plasmids. *Infect. Immun.* **38**:476-486.
- Kraft, R., J. Tardiff, K. S. Krauter, and L. A. Leinwand. 1988. Using miniprep plasmid DNA for sequencing double stranded templates with sequenase. *BioTechniques* **6**:544-546.
- Libby, S. J., W. Goebel, A. Ludwig, N. Buchmeier, F. Bowe, F. C. Fang, D. G. Guiney, J. G. Songer, and F. Heffron. 1994. A cytolysin encoded by *Salmonella* is required for survival within macrophages. *Proc. Natl. Acad. Sci. USA* **91**:489-493.
- Liu, S.-L., and K. E. Sanderson. 1992. A physical map of the *Salmonella typhimurium* LT2 genome made by using *XbaI* analysis. *J. Bacteriol.* **174**:1662-1672.
- Lockman, H. A., and R. Curtiss III. 1992. Isolation and characterization of conditional adherent and non-type 1 fimbriated *Salmonella typhimurium* mutants. *Mol. Microbiol.* **6**:933-945.
- Lockman, H. A., and R. Curtiss III. 1992. Virulence of non-type 1-fimbriated

- and nonfimbriated nonflagellated *Salmonella typhimurium* mutants in murine typhoid fever. *Infect. Immun.* **60**:491–496.
38. McClelland, M., J. Hanish, M. Nelson, and Y. Patel. 1988. KGB: a single buffer for all restriction endonucleases. *Nucleic Acids Res.* **16**:364.
 39. Miller, J. H. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
 40. Müller, K.-H., K. Collinson, T. J. Trust, and W. W. Kay. 1991. Type 1 fimbriae of *Salmonella enteritidis*. *J. Bacteriol.* **173**:4765–4772.
 41. Old, D. C. 1972. Inhibition of the interaction between fimbrial hemagglutinin and erythrocytes by D-mannose and other carbohydrates. *J. Gen. Microbiol.* **71**:149–157.
 42. Old, D. C., and R. A. Adegbola. 1983. A new mannose-resistant haemagglutinin in *Klebsiella*. *J. Appl. Bacteriol.* **55**:165–172.
 43. Old, D. C., and S. B. Payne. 1971. Antigens of the type 2 fimbriae of *Salmonellae*: “cross reacting material” (CRM) of type 1 fimbriae. *J. Med. Microbiol.* **4**:215–225.
 44. Olsén, A., A. Jonsson, and S. Normark. 1989. Fibronectin binding mediated by a novel class of surface organelles on *Escherichia coli*. *Nature (London)* **338**:652–655.
 45. Peterson, J. W., A. K. Chopra, and R. Prasad. 1991. Fine mapping of the *rmE*, *purHD*, and *hydGH* operons on the *Escherichia coli* chromosome. *J. Bacteriol.* **173**:3274–3275.
 46. Purcell, B. K., J. Pruckler, and S. Clegg. 1987. Nucleotide sequences of the genes encoding type 1 fimbrial subunits of *Klebsiella pneumoniae* and *Salmonella typhimurium*. *J. Bacteriol.* **169**:5831–5834.
 47. Rossolini, G. M., P. Muscas, A. Chiesurin, and G. Satta. 1993. Analysis of the *Salmonella fim* gene cluster: identification of a new gene (*fimI*) encoding a fibrin-like protein and located downstream from the *fimA* gene. *FEMS Microbiol. Lett.* **114**:259–266.
 48. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
 49. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
 50. Satta, G., I. Ingianni, P. Muscas, G. M. Rossolini, and R. Pompei. 1993. The pathogenicity determinants of *Salmonella typhi*: potential role of fimbrial structures, p. 83–90. *In* C. F. Cabello Hormaeche, P. Mastroeni, and L. Bonina (ed.), *Biology of Salmonella*. Plenum Press, New York.
 51. Schägger, H., and G. von Jagow. 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**:368–379.
 52. Short, J. M., J. M. Fernandez, J. A. Sorge, and W. D. Huse. 1988. λZAP: a bacteriophage expression vector with in vivo excision properties. *Nucleic Acids Res.* **16**:7583–7600.
 53. Sohel, I., J. L. Puente, W. J. Murray, J. Vuopio-Varkila, and G. K. Schoolnik. 1993. Cloning and characterization of the bundle-forming pilin gene of enteropathogenic *Escherichia coli* and its distribution in *Salmonella* serotypes. *Mol. Microbiol.* **7**:563–575.
 54. Strom, M. S., and S. Lory. 1993. Structure-function and biogenesis of the type IV pili. *Annu. Rev. Microbiol.* **47**:565–596.
 55. Swenson, D. L., and S. Clegg. 1992. Identification of ancillary *fim* genes affecting *fimA* expression in *Salmonella typhimurium*. *J. Bacteriol.* **174**:7697–7704.
 56. Swenson, D. L., S. Clegg, and D. C. Old. 1991. The frequency of *fim* genes among *Salmonella* serovars. *Microb. Pathog.* **10**:487–492.
 57. van der Woude, M. W., and D. A. Low. 1994. Leucine-responsive regulatory protein and deoxyadenosine methylase control the phase variation and expression of the *sfa* and *daa* pili operons in *Escherichia coli*. *Mol. Microbiol.* **11**:605–618.
 58. van Zijderveld, F. G., J. Anakotta, R. A. M. Brouwers, A. M. van Zijderveld, D. Bakker, and F. K. de Graaf. 1990. Epitope analysis of the F4 (K88) fimbrial antigen complex of enterotoxigenic *Escherichia coli* by using monoclonal antibodies. *Infect. Immun.* **58**:1870–1878.
 59. Woodall, L. A., P. W. Russell, S. L. Harris, and P. E. Orndorff. 1993. Rapid, synchronous, and stable induction of type 1 piliation in *Escherichia coli* by using a chromosomal *lacUV5* promoter. *J. Bacteriol.* **175**:2770–2778.
 60. Woodward, M. J., I. McLaren, and C. Wray. 1989. Distribution of virulence plasmids within salmonellae. *J. Gen. Microbiol.* **135**:503–511.
 61. Zychlinski, A., M. C. Prevost, and P. J. Sansonetti. 1992. *Shigella flexneri* induces apoptosis in infected macrophages. *Nature (London)* **358**:167–168.