JOURNAL OF BACTERIOLOGY, Dec. 1987, p. 5831–5834
0021-9193/87/125831-04$02.00/0
Copyright © 1987, American Society for Microbiology
Vol. 169, No. 12

Nucleotide Sequences of the Genes Encoding Type 1 Fimbrial Subunits of *Klebsiella pneumoniae* and *Salmonella typhimurium*

BRET K. PURCELL, JANET PRUCKLER, AND STEVEN CLEGG*

Department of Microbiology, University of Iowa, Iowa City, Iowa 52242

Received 20 July 1987/Accepted 8 September 1987

The nucleotide sequences of the genes encoding the subunits of *Klebsiella pneumoniae* and *Salmonella typhimurium* type 1 fimbriae were determined. Comparison of the predicted amino acid sequences of the two subunits revealed domains in which the sequences were highly conserved. Both gene products possessed signal peptides, a fact consistent with the transport of the fimbrial subunit across the membrane, but these regions showed no amino acid homology between the two proteins. The predicted N-terminal amino acid sequences of the processed fimbrial subunits were in good agreement with those obtained by purification of the fimbrial subunits.

Type 1 fimbriae of many genera of the family *Enterobacteriaceae* appear morphologically identical, yet antigenic diversity exists among these fimbriae (5, 7, 12). This may result from evolutionary selection pressure to preserve the fimbrial subunits of *K. pneumoniae* and *S. typhimurium*. We also compare these sequences with the previously characterized fimbrial subunit of *Escherichia coli* (13, 14).

The physical maps of the chimeric plasmids pBP7 and

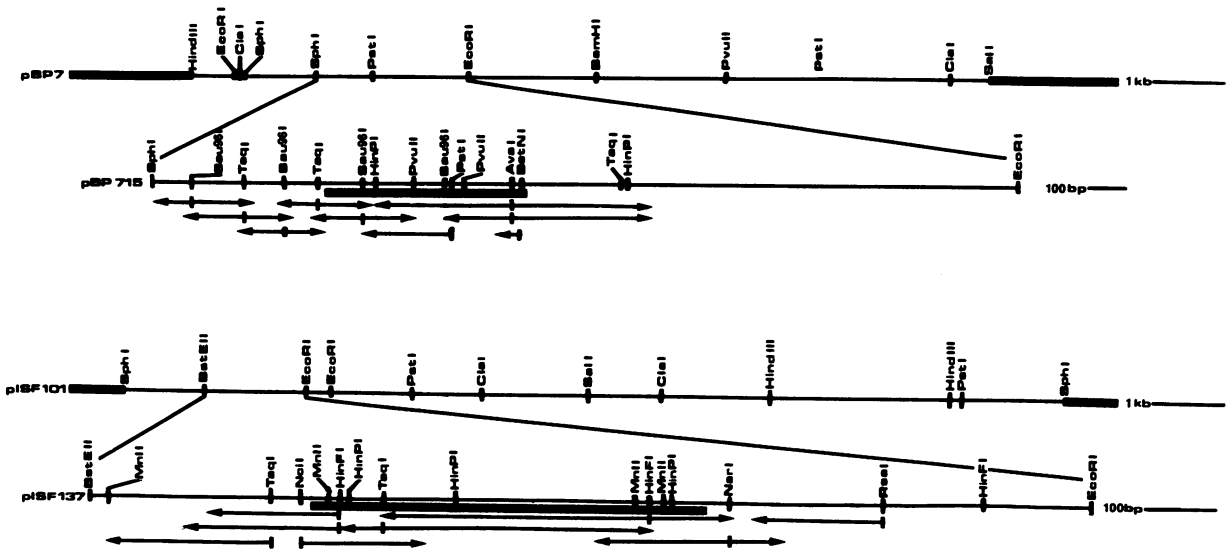

FIG. 1. Physical maps and DNA sequencing strategies. The physical maps of plasmids pBP7 and pISF101 are shown, and the restriction maps of their deletion derivatives, pBP715 and pISF137, are represented directly beneath each parent plasmid. The dark lines below plasmid pBP715 and pISF137 depict the relative locations of the genes encoding the respective fimbrial subunits. DNA sequencing strategies for each gene are outlined beneath plasmids pBP715 and pISF137. Both strands of the DNA for each gene have been sequenced. kb, Kilobase; bp, base pair.

structure and function of these organelles while concomitantly allowing for antigenic variation. Thus, the type 1 fimbriae of *Klebsiella pneumoniae* and *Salmonella typhimurium* are serologically distinct but maintain a similar morphology. The fimbrial gene clusters of *K. pneumoniae* and *S. typhimurium* have been cloned and described elsewhere (4, 15). Each cluster is composed of several genes, the expression of which is required for phenotypic production of type 1 fimbriae. In this report, we describe the nucleotide and amino acid sequences of the genes encoding the type 1 pISF101, encoding phenotypic expression of *K. pneumoniae* and *S. typhimurium* fimbriae, respectively, are shown in Fig. 1. The deletion derivatives pBP715 and pISF137 were derived from the parental plasmids and contain the appropriate fimbrial subunit genes. The complete nucleotide sequences of both fimbrial genes are shown in Fig. 2, and a comparison of the predicted amino acid sequences for the *K. pneumoniae*, *S. typhimurium*, and previously published *E. coli* type 1 fimbriae (13) is shown in Fig. 3. Although the DNA relatedness of *E. coli* K-12 and *S. typhimurium* LT2 has been reported to be greater than that between *E. coli* K-12 and *K. pneumoniae* II (2), the fimbrial subunit genes of *K.*

* Corresponding author.

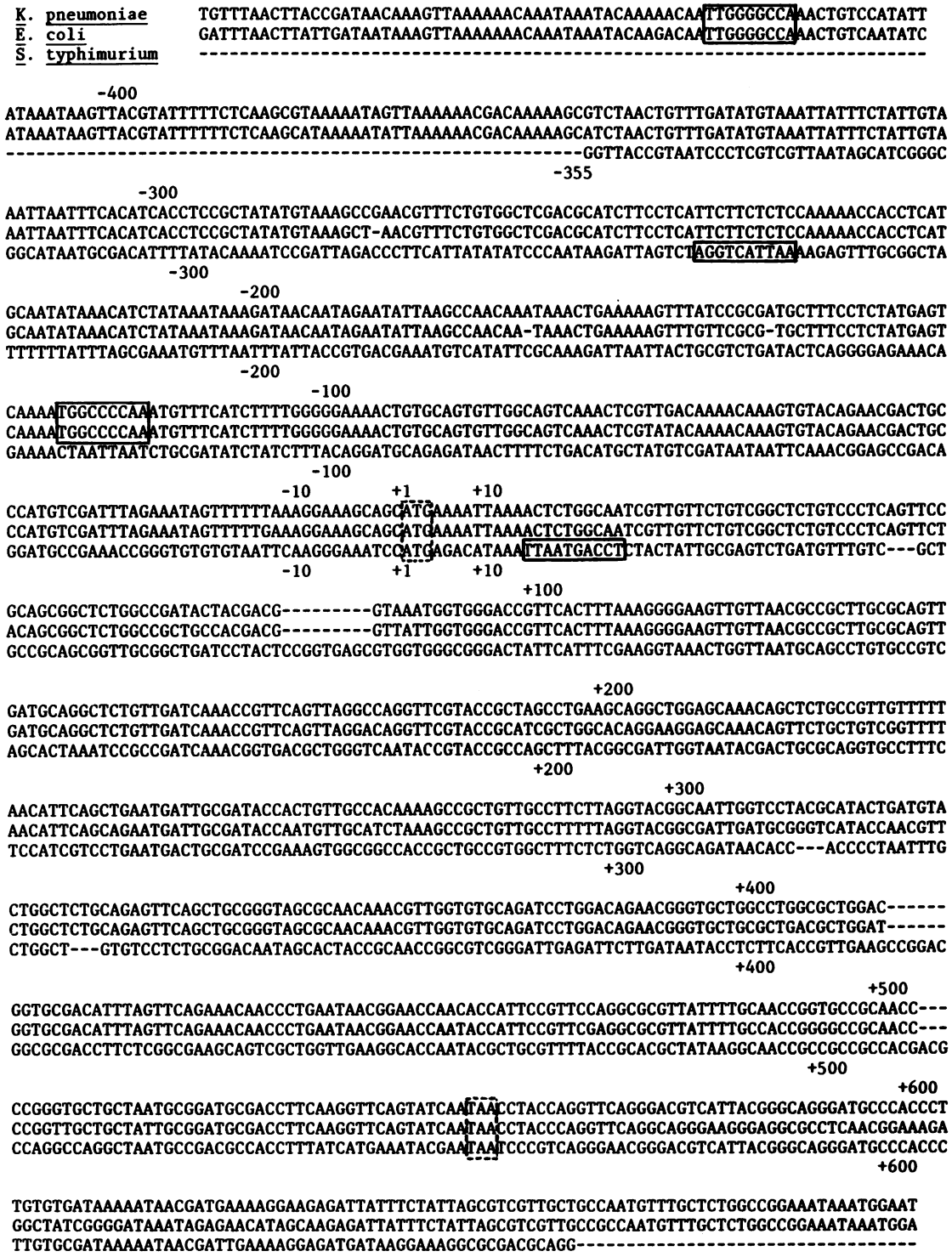

FIG. 2. Comparison of the nucleotide sequences for type 1 fimbrial subunit genes. The nucleotide sequences are given for the type 1 fimbrial subunit genes of *K. pneumoniae*, *E. coli* (13), and *S. typhimurium*. The inversion repeats are indicated as solid-line boxes, and the dashed-line boxes represent the start and stop codons for all three genes. The nucleotides of the *K. pneumoniae* and *E. coli* genes are numbered above the sequences, whereas the nucleotides of the *S. typhimurium* gene are numbered below. The nucleotide sequences are aligned for the best possible comparison of all three genes. The dashed lines represent sequences not present in the appropriate gene.

pneumoniae and *E. coli* demonstrate more nucleotide sequence homology (91%) than those of *S. typhimurium* and *E. coli* (49%). Presumably, the *S. typhimurium* gene has undergone significant divergence from both the *K. pneumoniae* and the *E. coli* genes.

The frequency of codon usage in *K. pneumoniae* and *E. coli* is similar for both fimbrial genes, whereas several codons in *S. typhimurium* demonstrate significant variation from usage in either *E. coli* or *K. pneumoniae*. For example, the GTT codon, determining valine utilization, is preferred in *K. pneumoniae* and *E. coli*, whereas GTC and GTG are preferentially used in *S. typhimurium*. Other amino acids which vary in their codon usage for these fimbrial genes are Leu, Gln, Asn, Thr, Ala, and Gly.

Inverted repeat sequences of 9 and 10 base pairs were found upstream from or within the *K. pneumoniae* and *S. typhimurium* fimbrial genes, respectively (Fig. 2). The inversion of a DNA segment has been found to play a role in the control of type 1 fimbrial phase variation in *E. coli* (1, 16). Since fimbrial phase variation also occurs in both *K. pneumoniae* and *S. typhimurium*, a similar mechanism may control fimbrial expression. The predicted inversion sequence for *K. pneumoniae* is similar to the invertible fragment documented for the *E. coli* type 1 fimbrial gene (1),

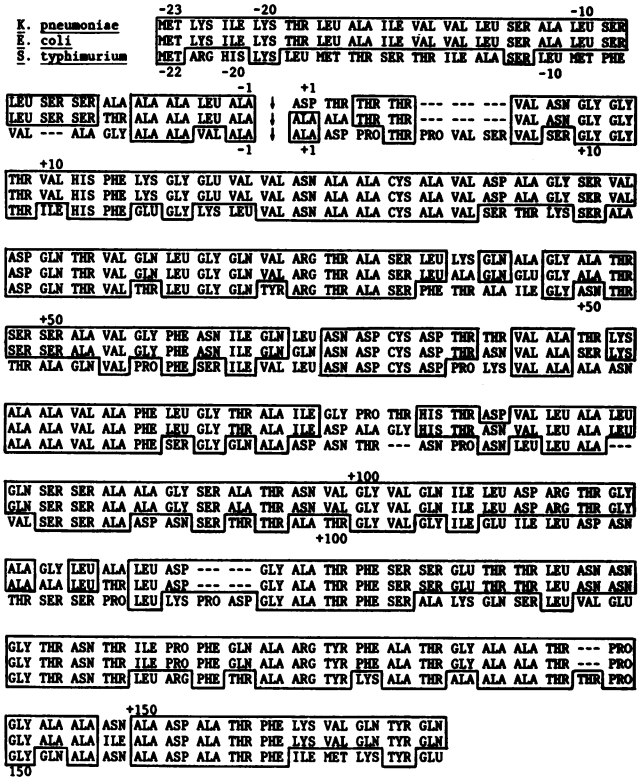

FIG. 3. Comparison of the predicted amino acid sequences of the type 1 fimbrial subunit. The amino acid sequences for the *K. pneumoniae*, *E. coli*, and *S. typhimurium* fimbrial subunits were derived from the nucleotide sequences of their respective genes. The boxed areas represent common amino acids, and the dashed lines are amino acids absent from the predicted sequences. The amino acid sequences for *K. pneumoniae* and *E. coli* are numbered above the sequences, whereas the sequences for *S. typhimurium* are numbered below. The signal peptide cleavage point for each polypeptide is indicated as an arrow between −1 and +1.

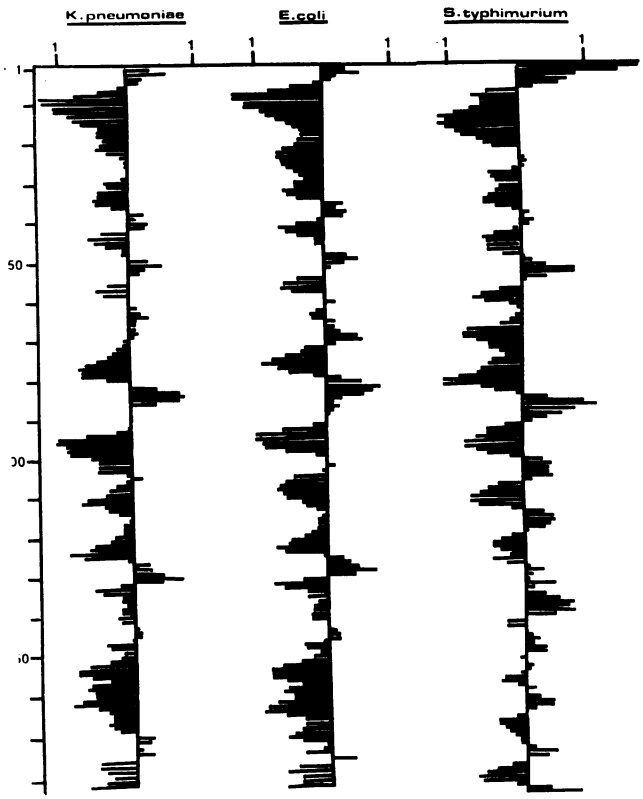

FIG. 4. Predicted hydropathicity profiles of the three fimbrial subunits. The method of Chou and Fasman (3) was used to calculate the hydropathicity profiles for the fimbrial subunits of *K. pneumoniae*, *E. coli*, and *S. typhimurium*. The amino acid sequence is numbered on the *x* axis, and the hydropathicity values are indicated on the *y* axis. The bars below the midpoint of each hydropathicity profile represent hydrophobic regions; the bars above the midpoint represent hydrophilic regions.

except that 16 base pairs vary within the 314-base-pair invertible region. The possible inverted repeats of the *S. typhimurium* gene are different from those of *K. pneumoniae* and *E. coli* in both size and location. The possible role of this DNA region in *S. typhimurium* fimbrial phase variation is currently being investigated.

The work of Waalen and co-workers (17) established the first 27 amino acid residues of the mature fimbrial subunit prepared from *S. typhimurium* LT2. The predicted amino acid sequence derived from the DNA sequence was in good agreement with our results with two exceptions. Our data predicted that the residue 5 would be a valine rather than a threonine and that the residue 18 would be a lysine, although the N-terminal sequence data do not indicate that this amino acid is present. In addition, the unknown residue at position 24 reported by the European workers was shown by our data to be a cysteine. The predicted molecular mass of the *S. typhimurium* fimbrial subunit was 20,140, which compares favorably with the molecular mass of 21,500 calculated from the electrophoretic mobility of the purified fimbriae. The sequence predicts that the fimbrial subunit would be synthesized as a precursor with 22 amino acids in the signal peptide region (Fig. 3). This is consistent with transport of the subunit across the bacterial membrane during fimbrial assembly. Like other previously reported signal peptides, the

S. typhimurium fimbrial polypeptide possesses positively charged residues immediately adjacent to the translation initiation codon, a relatively high number of hydrophobic residues (10 of the last 12 amino acids of the signal peptide are hydrophobic), and a point of cleavage between two amino acids (alanines) with short carbon side chains (10).

The predicted amino acid sequence of the N-terminal region of the *K. pneumoniae* gene was also in good agreement with the sequence derived from the analysis of the biochemically purified protein (6). Thus, only three discrepancies within the first 25 amino acids of the N-terminal region were found when all data were compared. In the *K. pneumoniae* gene, the putative signal peptide consists of 23 amino acids, 16 of which are hydrophobic. Also, the residue following the initiation codon is positively charged, and an alanine immediately precedes the site of cleavage.

The hydropathicity profiles for *K. pneumoniae* and *S. typhimurium* fimbrial polypeptides are shown in Fig. 4 (a profile for *E. coli* is also included for the sake of comparison). As mentioned above, the signal peptides of the fimbrial subunits were primarily composed of hydrophobic residues. Each mature fimbrial subunit was composed of more that 50% hydrophobic residues, and each subunit contained two cysteine residues. The first of these, for both *K. pneumoniae* and *E. coli* fimbriae, was located at position 44 within a predicted hydrophobic region with secondary α-helical structure. The first cysteine residue for *S. typhimurium* fimbriae was at position 46 within a hydrophobic region, but positioned in a secondary β-sheet structure. All three fimbrial subunits had the second cysteine residue within hydrophilic regions, located in a turn of their predicted secondary structures.

An analysis of the predicted amino acid sequences indicated that strings of amino acid residues were common throughout the *K. pneumoniae* and *S. typhimurium* polypeptides. For example, in *K. pneumoniae*, residues 7 to 12 were -Gly-Gly-Thr-Val-His-Phe-, whereas in the *S. typhimurium* gene, the string of amino acids -Gly-Gly-Thr-Ile-His-Phe- appeared at positions 10 to 15 (Fig. 3). As can be seen, these two stretches of amino acids differed only in a conservative valine-to-isoleucine change at amino acid 4. Similarly, at amino acid positions 17 to 23, a string of seven amino acids (-Val-Asn-Ala-Ala-Cys-Ala-Val-) was found in the *K. pneumoniae* system, and an identical sequence appeared at positions 20 to 26 in the *S. typhimurium* gene. The roles of these conserved amino acid sequences are unknown. However, it is now known that, in the *E. coli* system, more than one component is necessary for the formation of wild-type, receptor-binding fimbriae (8, 9, 11). In addition, we have recently constructed Fim^+HA^- and Fim^-HA^+ phenotypes with the *K. pneumoniae* system (unpublished results). Therefore, if the type 1 fimbriae of enteric bacteria are in fact composed of a number of distinct polypeptides, it is possible that such conserved sequences in the fimbrial subunit are necessary for maintenance of the correct configurations facilitating protein-protein interactions.

LITERATURE CITED

1. **Abraham, J. M., C. S. Freitag, J. R. Clements, and B. I. Eisenstein.** 1985. An invertible element of DNA controls phase variation of type 1 fimbriae of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **82**:5724–5727.
2. **Brenner, D. J.** 1978. Characterization and clinical investigation of *Enterobacteriaceae* by DNA hybridization. Prog. Clin. Pathol. **7**:71–117.
3. **Chou, P. Y., and G. D. Fasman.** 1978. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. Relat. Areas Mol. Biol. **47**:45–148.
4. **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.
5. **Duguid, J. P., and I. Campbell.** 1969. Antigens of the type 1 fimbriae of *Salmonella* and other enterobacteria. J. Med. Microbiol. **2**:535–553.
6. **Fader, R. C., L. K. Duffy, C. P. Davis, and A. Kurosky.** 1982. Purification and characterization of type 1 pili isolated from *Klebsiella pneumoniae*. J. Biol. Chem. **257**:3301–3305.
7. **Gilles, R. R., and J. P. Duguid.** 1958. Fimbrial antigens of *Shigella*. J. Hyg. **56**:303–318.
8. **Maurer, L., and P. E. Orndorff.** 1985. A new locus, *pil*E, required for the binding of type 1 piliated *Escherichia coli* to erythrocytes. FEMS Microbiol. Lett. **30**:59–66.
9. **Maurer, L., and P. E. Orndorff.** 1987. Identification and characterization of genes determining receptor binding and pilus length of *Escherichia coli* type 1 pili. J. Bacteriol. **169**:640–645.
10. **Michaelis, S., and J. Beckwith.** 1982. Mechanism of incorporation of cell envelope proteins in *Escherichia coli*. Annu. Rev. Microbiol. **36**:435–465.
11. **Minion, F. C., S. N. Abraham, E. H. Beachey, and J. D. Gaugen.** 1986. The genetic determinant of adhesive function in type 1 fimbriae of *Escherichia coli* is distinct from the gene encoding the fimbrial subunit. J. Bacteriol. **165**:1033–1036.
12. **Nowotarska, M., and M. Mulczyk.** 1977. Serologic relationship of fimbriae among the *Enterobacteriaceae*. Arch. Immunol. Ther. Exp. **25**:7–16.
13. **Orndorff, P. E., and S. Falkow.** 1985. Nucleotide sequence of *pilA*, the gene encoding the structural component of type 1 pili in *Escherichia coli*. J. Bacteriol. **162**:454–457.
14. **per Klemm, P.** 1984. The fimA gene encoding the type 1 fimbrial subunit of *Escherichia coli*. Eur. J. Biochem. **143**:393–399.
15. **Purcell, B. K., and S. Clegg.** 1983. Construction and expression of recombinant plasmids encoding type 1 fimbriae of a urinary *Klebsiella pneumoniae* isolate. Infect. Immun. **39**:1122–1127.
16. **Spears, P. A., D. Schauer, and P. E. Orndorff.** 1986. Metastable regulation of type 1 piliation in *Escherichia coli* and isolation and characterization of a phenotypically stable mutant. J. Bacteriol. **168**:179–185.
17. **Waalen, K., K. Sletton, L. Frohalm, V. Vaisamen, and T. K. Korhonen.** 1983. The N-terminal amino acid sequence of type 1 fimbriae (pili) of *Salmonella typhimurium* LT2. FEMS Microbiol. Lett. **16**:149–151.