INFECTION AND IMMUNITY, July 2001, p. 4447–4457 Vol. 69, No. 7
0019-9567/01/$04.00+0 DOI: 10.1128/IAI.69.7.4447–4457.2001
Copyright © 2001, American Society for Microbiology. All Rights Reserved.

Novel Type of Fimbriae Encoded by the Large Plasmid of Sorbitol-Fermenting Enterohemorrhagic *Escherichia coli* O157:H$^-$

WERNER BRUNDER,[1]* A. SALAM KHAN,[2] JÖRG HACKER,[2] AND HELGE KARCH[1]

Institut für Hygiene und Mikrobiologie der Universität Würzburg, D-97080 Würzburg,[1] and Institut für Molekulare Infektionsbiologie der Universität Würzburg, D-97070 Würzburg,[2] Germany

Received 22 December 2000/Returned for modification 7 March 2001/Accepted 9 April 2001

Sorbitol-fermenting (SF) enterohemorrhagic *Escherichia coli* (EHEC) O157:H$^-$ have emerged as important causes of diarrheal diseases and the hemolytic-uremic syndrome in Germany. In this study, we characterized a 32-kb fragment of the plasmid of SF EHEC O157:H$^-$, pSFO157, which differs markedly from plasmid pO157 of classical non-sorbitol-fermenting EHEC O157:H7. We found a cluster of six genes, termed *sfpA*, *sfpH*, *sfpC*, *sfpD*, *sfpJ*, and *sfpG*, which mediate mannose-resistant hemagglutination and the expression of fimbriae. *sfp* genes are similar to the *pap* genes, encoding P-fimbriae of uropathogenic *E. coli*, but the *sfp* cluster lacks homologues of genes encoding subunits of a tip fibrillum as well as regulatory genes. The major pilin, SfpA, despite its similarity to PapA, does not cluster together with known PapA alleles in a phylogenetic tree but is structurally related to the PmpA pilin of *Proteus mirabilis*. The putative adhesin gene *sfpG*, responsible for the hemagglutination phenotype, shows significant homology neither to *papG* nor to other known sequences. Sfp fimbriae are 3 to 5 nm in diameter, in contrast to P-fimbriae, which are 7 nm in diameter. PCR analyses showed that the *sfp* gene cluster is a characteristic of SF EHEC O157:H$^-$ strains and is not present in other EHEC isolates, diarrheagenic *E. coli*, or other *Enterobacteriaceae*. The *sfp* gene cluster is flanked by two blocks of insertion sequences and an origin of plasmid replication, indicating that horizontal gene transfer may have contributed to the presence of Sfp fimbriae in SF EHEC O157:H$^-$.

Adherence is one of the prerequisites for the successful colonization of a eukaryotic host by a bacterium. One of the best understood mechanisms of adherence is that mediated by rod-shaped proteinaceous appendages of the bacterial surface called fimbriae or pili. Well-studied adhesion systems of pathogenic bacteria are the S-fimbria superfamily (30, 31) and the P (pyelonephritis-associated)-fimbriae of uropathogenic *Escherichia coli*. The latter are composed of a thin tip fibrillum (2 nm in diameter) carrying the adhesin at its distal end and joined at its proximal end to a more rigid, 7-nm-diameter pilus rod (10). The *pap* gene cluster consists of 11 genes encoding the main component of the pilus rod (PapA), several minor fimbrial subunits (PapHKEF), the adhesin (PapG), the assembly machinery (PapCDJ), and two regulatory proteins (PapIB) (for a review see reference 17). P-fimbriae are part of a family of adhesive organelles that are characterized by an assembly machinery consisting of a periplasmic chaperone (PapD) and a pore-forming outer membrane usher (PapC) (22).

Enterohemorrhagic *E. coli* (EHEC) O157:H7 are well known as causative agents of diarrhea, hemorrhagic colitis, and the hemolytic-uremic syndrome (HUS) (29, 45). Whereas the Shiga toxins, the most well-established virulence factor of EHEC, have been extensively studied, the mechanisms underlying the adherence of the bacteria to epithelial cells are only partly understood. One adherence system shared by *E. coli* O157:H7 and enteropathogenic *E. coli* (EPEC) is the attaching and effacing mechanism encoded by a pathogenicity island called the locus of enterocyte effacement (42). Recently, Tarr et al. (53) described an adherence-conferring protein of EHEC O157:H7 termed Iha, which is similar to the product of an iron-regulated gene of *Vibrio cholerae*. The presence of fimbriae on the surface of *E. coli* O157:H7 has been reported by several investigators (15, 18, 27, 58), but their genetic representation and their molecular structure are still unknown.

A biochemical characteristic of *E. coli* O157:H7 is that they do not ferment sorbitol, and such non-sorbitol-fermenting (NSF) strains have been isolated throughout the world. In Germany, however, nonmotile EHEC O157, which are able to ferment sorbitol (SF EHEC O157:H$^-$), have also been implicated in outbreaks of HUS (1, 28). SF EHEC O157:H$^-$ strains are thought to represent a clonal lineage which has branched off at an early stage during the evolution from an EPEC-like *E. coli* O55:H7 ancestor to EHEC O157:H7 (16). Therefore, detailed comparison of the genomes of SF O157:H$^-$ and NSF O157:H7 strains could give further insight into the emergence of highly pathogenic EHEC. In a recent study striking differences were found in the gene compositions of plasmid pO157 of NSF EHEC O157:H7 and the plasmid of SF EHEC O157:H$^-$, henceforth called pSFO157 (7). During our investigations to further characterize these differences, we discovered a gene cluster present only on pSFO157 and mediating mannose-resistant hemagglutination and the expression of a novel type of fimbriae.

* Corresponding author. Mailing address: Institut für Hygiene und Mikrobiologie der Universität Würzburg, Josef-Schneider-Str. 2, D-97080 Würzburg, Germany. Phone: 49-931-2013981. Fax: 49-931-2013445. E-mail: wbrunder@hygiene.uni-wuerzburg.de.

MATERIALS AND METHODS

Bacterial strains and culture conditions. *E. coli* K-12 strains DH5α and HB101 were used as hosts for recombinant plasmids. The SF EHEC O157:H$^-$

TABLE 1. PCR primers and amplification conditions used in this study

Primer	Sequence	Amplification conditions (temp [°C], time [s])			Product length (bp)
		Denaturing	Annealing	Extension	
sfpA-U	5′-AGCCAAGGCCAAGGGATTATTA-3′	94, 30	59, 60	72, 60	440
sfpA-L	5′-TTAGCAACAGCAGTGAAGTCTC-3′				
sfpDG-U	5′-GATGCTGGCACCGAAGTCAA-3′	94, 30	57, 60	72, 60	1,682
sfpDG-L	5′-TCCTGCAGCATCCAAGTCAC-3′				
wprom-3	5′-GCTGGACGCCGGTGCTTATT-3′	94, 30	61, 60	72, 150	1,900
wprom-4	5′-CCAGATTGCCGCTCGCTTCAG-3′				
sfpG-UApa	5′-CCCGGGCCCGGCTGCGCTGCTGGTGAG-3′	94, 15	55, 30	68, 180[a]	3,162
sfpG-LApa	5′-CCCGGGCCCTCATCTGCCGGTCCCTTAC-3′				

[a] This PCR was performed using the Expand High Fidelity PCR system. From cycle 11 to 30 the extension time was prolonged 5 s with each cycle. The PCR was completed by a final elongation step of 7 min at 72°C.

strain 3072/96 was isolated from a patient suffering from HUS in Germany in 1996. This strain harbors the *stx*$_2$ gene as well as the *eae* gene. The clinical *E. coli* isolates and the other *Enterobacteriaceae* used in this study were from our strain collection. Characteristics of the diarrheagenic *E. coli* strains were described earlier (3, 4, 7).

Bacteria were grown in Luria broth (1% [wt/vol] tryptone, 0.5% [wt/vol] yeast extract, 1% [wt/vol] sodium chloride, pH 7.5) or in CDMT [13 mM K_2HPO_4, 6 mM KH_2PO_4, 8 mM $(NH_4)_2SO_4$, 2 mM sodium citrate, 0.4 mM $MgSO_4$, 0.2% Casamino Acids, 0.2% glucose, 5 μM $CaCl_2$, 0.01% tryptone, pH 7.4]. Solid media were prepared by the addition of 1.6% (wt/vol) Bacto agar.

General recombinant DNA techniques. Plasmid DNA was purified with Qiagen tip-100 cartridges according to the instructions of the supplier. Purification of DNA from agarose gels was performed using a Prep-A-Gene kit (Bio-Rad). Restriction enzymes and T4 DNA ligase were purchased from Gibco-BRL and New England Biolabs. Plasmids pK18 (43) and pBluescript II KS(+) (Stratagene) were used as cloning vectors. Restriction enzyme analysis, ligation, and transformation were conducted according to standard procedures (49).

Sequence analysis. Nucleotide sequencing was performed with an ABI Prism Big Dye Terminator Cycle Sequencing kit (Perkin-Elmer Applied Biosystems) according to the instructions of the supplier. Sequencing reactions were run on an ABI Prism 377 automatic DNA sequencer (Perkin-Elmer). Both strands of the DNA were sequenced stepwise using customized oligonucleotide primers (ARK Scientific), and each base was determined three times on average.

Sequence analysis was conducted with the Wisconsin Package, version 10.0 (Genetics Computer Group, Madison, Wis.) as well as with the Dnasis program (Hitachi Software). Promoter prediction was performed using the NNPP2.1 algorithm (promoter prediction by neural network; http://www.fruitfly.org/seq_tools/promoter.html). To illustrate phylogenetic relationships between SfpA and other major fimbrial subunits, an alignment of the predicted mature peptides was made using the CLUSTAL W program (21). This alignment was then used to calculate a distance matrix from which an unrooted phylogenetic tree was inferred according to the neighbor-joining method (48) using the Phylip 3.57c program package (http://evolution.genetics.washington.edu/phylip.html).

PCR. PCR was performed using the GeneAmp 9600 PCR system (Perkin-Elmer). Table 1 shows the primers and conditions used for amplification of the *sfpA* gene (primers sfpA-U and sfpA-L), the *sfpDG* region (primers sfpDG-U and sfpDG-L), and the region connecting the cloned fragments pSFO157-E14 and pSFO157-E11 (primers wprom-3 and wprom-4). Amplification was carried out in a total volume of 50 μl containing 5 μl of bacterial cell suspension (about 10^3 bacterial cells suspended in 0.85% [wt/vol] sodium chloride), 30 pmol of each primer, 200 μM concentrations of each deoxynucleoside triphosphate, 5 μl of 10× GeneAmp PCR buffer II, 3 μl of 25 mM $MgCl_2$, and 2 U of Ampli*Taq* DNA polymerase (Perkin-Elmer). After an initial denaturation step of 5 min at 94°C, the samples were subjected to 30 cycles of denaturing, annealing, and extension (see Table 1). The reaction was completed with a final extension step of 5 min at 72°C. Each reaction was conducted at least twice in independent reactions.

Amplification of the *sfpG* gene (primers sfpG-UApa and sfpG-LApa) for complementation purposes was performed using the Expand High Fidelity PCR system (Roche Molecular Biochemicals) according to the recommendations of the supplier (Table 1).

Colony blot hybridization. Probes specific for *sfpA* and *sfpDG* for colony blot hybridization were prepared by random labeling of amplicons derived from *sfpA* and *sfpDG* PCR using DNA of strain 3072/96 as the template. The PCR products were purified from agarose gels and labeled with a Digoxigenin Labeling and Detection kit (Roche Molecular Biochemicals) according to the instructions of the supplier. Colony blot hybridization was performed as described previously (7).

Preparation of fimbriae. Strains were grown in CDMT for 20 h at 37°C with shaking. Bacteria were harvested by centrifugation (20 min, 1,700 × *g*, 4°C) and resuspended in 1/10 volume of a solution of 75 mM NaCl and 0.5 mM Tris-HCl, pH 7.4 (32). The cell suspension was incubated for 90 min at 60°C and the bacteria were removed by centrifugation. The thermoeluted fimbrial proteins were concentrated from the supernatant by precipitation with trichloroacetic acid and dissolved in Laemmli sample buffer.

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Electrophoresis was performed on SDS–15% polyacrylamide gels in a Mini-PROTEAN II Dual Slab Cell (Bio-Rad) according to the method of Laemmli (34). Gels were either stained with Coomassie blue G or the proteins were blotted onto nitrocellulose membranes for immunologic detection.

Amino acid sequencing. N-terminal sequencing of proteins was performed by automated Edman degradation. Heat-extracted fimbrial proteins were separated by SDS-PAGE and blotted onto a ProBlott membrane (Applied Biosystems). After Coomassie blue staining of the proteins, the 17-kDa band was excised and subjected to protein sequencing employing an Applied Biosystems 476A sequencer.

Preparation of anti-SfpA antiserum and immunoblot analysis. For preparation of a polyclonal serum to SfpA, we purified SfpA from thermoeluted fimbrial extracts using SDS-PAGE followed by electroelution of the 17-kDa protein from gel slices as described previously (9). The protein eluate was emulsified with ABM-S (Linaris) as the adjuvant and administered subcutaneously to a New Zealand White rabbit. One booster injection was given 4 weeks later. A further 4 weeks later the rabbit was bled, and the resulting serum was investigated for reactivity against SfpA.

Immunoblot analysis was performed as described previously (9). In order to minimize nonspecific reactions, the sera were adsorbed against a vector control strain as follows. A crude extract of *E. coli* HB101/pK18 was diluted to a final concentration of 10 mg/ml in phosphate-buffered saline and mixed with 1 volume of the serum diluted 1:10. The mixture was incubated for 1 h at 4°C and centrifuged (15 min, 15,000 × *g*, 4°C), and the supernatant was used in immunoblot analysis. Rabbit normal serum was used at a final dilution of 1:100, and rabbit anti-SfpA-serum was used at a dilution of 1:20,000.

Hemagglutination assays. Quantitative hemagglutination assays were performed in 96-well, round-bottom microtiter plates. Bacteria were grown in CDMT broth for 20 h at 37°C with shaking, harvested by centrifugation (15 min, 1,700 × *g*, 4°C), and adjusted in saline to 10^{10} cells ml^{-1}. Twofold serial dilutions of the bacterial suspensions were prepared in saline. Human erythrocytes were diluted in saline to a 0.5% (vol/vol) suspension, and 100 μl of each bacterial dilution and 100 μl of the erythrocyte suspension were mixed in microtiter plates. After being shaken, the plates were incubated for 4 h at room temperature and then examined for hemagglutination. Wells containing a small pellet of erythrocytes at the bottom were considered hemagglutination negative, and those containing an even sheet of erythrocytes across the well were considered hemagglutination positive. The hemagglutination titer was expressed as the reciprocal of the greatest dilution of bacterial cells that resulted in positive hemagglutination.

For fast qualitative tests of hemagglutination activity, 15 μl of a bacterial culture or a colony suspended in saline was mixed with 15 μl of human erythrocytes (diluted 1:50 in saline) on glass slides. Agglutination of the erythrocytes was examined by eye after 1 min of rocking at room temperature. Hemagglutination assays were performed both in the presence and in the absence of 0.5% methyl α-D-mannopyranoside, a nonmetabolizable mannose analog.

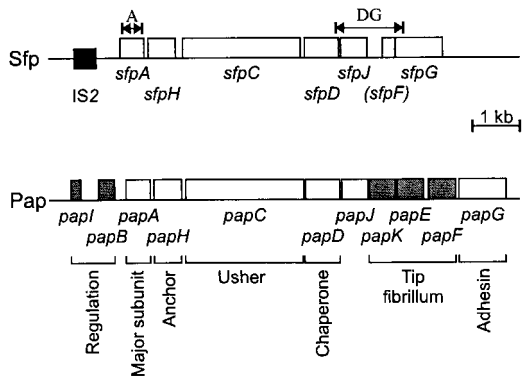

FIG. 1. Genetic organization of the *sfp* fimbrial gene cluster and comparison with the *pap* cluster. Functions of Pap proteins (22) are shown below the map. Shaded boxes indicate components only present in the *pap* cluster. The black box depicts a DNA region with high homology to IS*2*. The double arrows indicate regions amplified by *sfpA* PCR (A) and *sfpDG* PCR (DG).

Electron microscopy. Bacteria grown on Luria broth agar plates were suspended in phosphate-buffered saline and allowed to adhere to a Pioloform-coated grid for 2 min. The grid was then negatively stained with 0.25% uranyl acetate for 60 min and examined with a Zeiss EM10 transmission electron microscope operated at 80 kV. Electron microscopy was performed at the Central Division of Electron Microscopy of the University of Würzburg.

Nucleotide sequence accession number. The nucleotide sequence data reported in this paper have been submitted to the DDBJ, EMBL, and GenBank databases under accession number AF228759.

RESULTS

Cloning of a novel fimbrial gene cluster. In order to compare the properties of plasmid pSFO157 of SF EHEC O157:H⁻ with those of pO157, we performed shotgun cloning of pSFO157 fragments derived from SF EHEC O157:H⁻ strain 3072/96 in the vector pBluescript II KS. The ends of the cloned fragments were sequenced using M13/pUC forward and reverse primers. The sequences were compared to that of pO157 of NSF EHEC O157:H7 reference strain EDL933 (11). One 10.9-kb *Eco*RI fragment, termed pSFO157-E11, showed no homology to pO157. After subcloning, we obtained a sequence homologous to that of the *papA* gene, which encodes the major subunit of P-fimbriae of uropathogenic *E. coli* (UPEC). In addition, we found that a recombinant *E. coli* K-12 strain harboring pSFO157-E11, but not a vector control strain, was able to agglutinate human red blood cells. From this we hypothesized the presence of *pap*-like adhesin genes present on pSFO157-E11 and determined the complete sequence of this fragment.

Sequence analysis revealed nine open reading frames (ORFs), eight of which were transcribed in the same direction. The amino acid sequences deduced from six of these were markedly homologous to those of the F13 P-fimbrial proteins of UPEC strain J96, PapA, PapH, PapC, PapD, PapJ, and PapF (36). We therefore termed the ORFs in analogy to the *pap* fimbrial genes *sfpA*, *sfpH*, *sfpC*, *sfpD*, *sfpJ,* and *sfpF*, standing for "sorbitol-fermenting EHEC O157 fimbriae, plasmid-encoded" (Fig. 1). A further ORF 21 bp downstream of *sfpF* and without homology to known proteins was termed *sfpG*, because it shared some properties with the *papG* gene encoding the adhesin molecule of P-fimbriae (see below).

The properties of the proteins deduced from *sfp* genes are summarized in Table 2 and are compared to those of corresponding Pap proteins. As shown by the method of von Heijne (60), all Sfp proteins, with the exception of SfpF, possess putative N-terminal signal peptides like the Pap pilus subunits and assembly proteins (Table 2).

Major pilus subunit SfpA. The predicted mature SfpA protein had 62% amino acid similarity to the major pilin of F13 P-fimbriae, $PapA_{F13}$ (36), 41% similarity to the *E. coli* type I pilin, FimA (41), and 38% similarity to the major pilin of *E. coli* S-fimbriae, SfaA (51). SfpA showed structural properties of the so-called class I pilins (12): a pair of cysteine residues (positions 19 and 58 of the predicted mature protein) known to establish a disulfide bridge in pilins of type I, P-fimbriae, and S-fimbriae; a variable inner region; and a conserved C-terminal β-zipper-forming domain (residues 161 to 174) responsible for the interaction of pilins with the periplasmic chaperone during the assembly of the pilus (Fig. 2). This domain contains invariant glycine and aromatic residues separated by a series of alternating hydrophobic amino acids (33). Another site of interaction between the chaperone and the pilin, the so-called second site near the amino terminus (17), was also conserved in SfpA (Fig. 2).

Recently, Girardeau et al. (19) classified class I pilins into seven subfamilies and described intrasubfamily signature motifs. Based on these characteristics, the SfpA pilin is a member of the Ic subfamily, together with *E. coli* PapA, *Serratia marcescens* SmfA, and *Proteus mirabilis* MrpA and PmpA. However, the SfpA characteristics differed from the subfamily's signature: the S1 motif, **G**x**G**[KT]**V**[TS]**F**x**G**[TS]**V**[VI]**DAP** (strongly

TABLE 2. Properties of putative *sfp* gene products and comparison to the corresponding Pap proteins[a]

Protein	Molecular mass overall (kDa)	Signal peptide (amino acids)[c]	Molecular mass of mature protein (kDa)	pI of mature protein	G+C content of respective gene (%)
SfpA	17.9 [18.7]	20	15.9 [16.6]	4.93 [4.72]	41.9 [44.1]
SfpH	21.3 [21.8]	19	19.1 [19.3]	7.17 [5.36]	53.2 [48.8]
SfpC	91.9 [91.1]	27	89.1 [88.3]	6.08 [5.92]	56.1 [52.5]
SfpD	26.9 [26.8]	26	24.2 [24.6]	9.72 [9.01]	47.7 [45.6]
SfpJ	21.1 [20.7]	20	19.1 [17.8]	4.80 [7.11]	40.2 [55.5]
SfpF[b]	9.3 [17.6]	None			34.5 [42.9]
SfpG	36.2 [38.3]	22	33.5 [35.9]	5.52 [5.59]	36.0 [38.3]

[a] Numbers in brackets indicate values for the corresponding *pap* protein.
[b] Translated starting from the first ATG codon of the putative nonfunctional *sfpF* gene; see the text for further details.
[c] Predicted using the method of von Heijne (60).

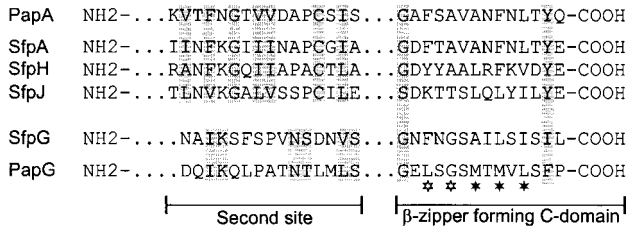

FIG. 2. Conserved structural motifs within Sfp subunits. Putative amino acid sequences of the C-terminal β-zipper-forming motif as well as the second site of interaction between pilins and the periplasmic chaperone are compared to the respective motifs of PapA and PapG (17). The most conserved residues are indicated by shaded boxes, and the asterisks mark the alternating hydrophobic amino acids in the C domain (open asterisks stand for residues conserved in all pilins with the exception of SfpJ).

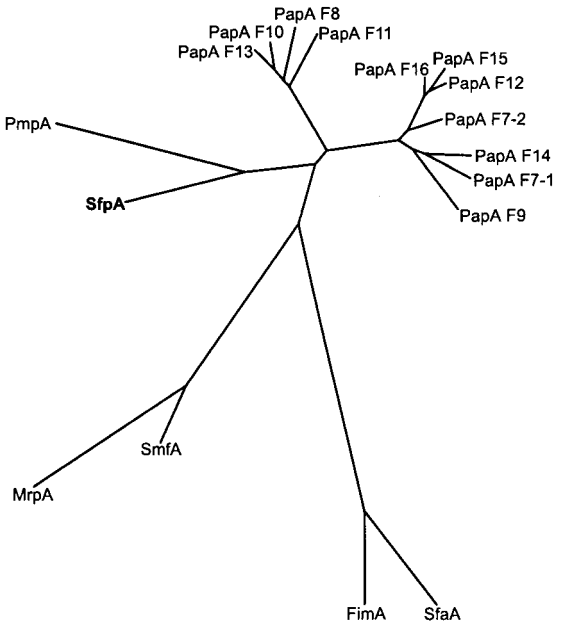

FIG. 3. Phylogenetic relationship between SfpA and different PapA alleles. Included are 11 PapA variants of *E. coli* (F7–1, F7–2, F8, F9, F10, F11, F12, F13, F14, F15, and F16), the Pap-related SmfA, MrpA, and PmpA pilins of *S. marcescens* and *P. mirabilis*, and the *E. coli* type I (FimA) and S-fimbriae (SfaA) major subunits. The unrooted tree was based on (predicted) mature peptides and was constructed using the neighbor-joining method. The presumed evolutionary distance between any two members of the tree equals the sum of the lengths of the branches connecting them.

conserved residues are shown in boldface), is present as **GQ GIINFKGIIINAP** in SfpA (differences from the conserved residues are underlined). In contrast to all other members of the Ic subfamily, SfpA carries asparagine in place of aspartate at position 13 of the motif. The valine residues at positions 5 and 11 are replaced by isoleucine in contrast to PapA but similar to the *P. mirabilis* pilins PmpA and MrpA as well as SmfA of *S. marcescens*. Furthermore, SfpA did not possess an insertion of 8 to 11 residues in loop L4, which is a further characteristic of the Ic subfamily and especially of the Pap pilins. Again, the *S. marcescens* and *P. mirabilis* pilins SmfA and MrpA are similar to SfpA in lacking this insertion. Obviously, SfpA is not a close member of the Pap pilin family, despite having the highest sequence similarity to $PapA_{F13}$. Phylogenetic analyses using CLUSTAL and Phylip programs supported this conclusion. As shown in Fig. 3, SfpA did not cluster together with the most homologous $PapA_{F13}$ or other PapA alleles (24) but had a common branch with PmpA from canine uropathogenic *P. mirabilis* (5), a protein showing 58% amino acid similarity to the mature SfpA protein.

Minor subunits and assembly genes. A second ORF of the gene cluster, termed *sfpH*, is located 79 bp downstream of *sfpA* (Fig. 1). The predicted mature SfpH protein showed 71% amino acid sequence similarity to $PapH_{F13}$ (36). However, the pI of the mature SfpH protein is markedly different from that of PapH (Table 2). Like SfpA, SfpH possessed a C-terminal chaperone binding motif as well as a conserved second site (Fig. 2). In view of the homology to PapH, we suggested that SfpH is a minor component involved in the anchoring of the pilus to the outer membrane and modulating its length (2).

We detected two genes downstream of *sfpH*, *sfpC,* and *sfpD*, highly homologous to genes encoding components involved in the assembly of P-fimbriae. The predicted mature SfpC protein showed 76% similarity to the PapC outer membrane usher protein and also had a similar molecular weight and pI (Table 2). PapC has been shown to form oligomeric channels in the outer membrane which facilitate the transport of the nascent pilus out of the cell (56).

The mature SfpD protein showed 83% sequence similarity to the periplasmic chaperone PapD. This protein forms stable complexes with pilus subunits in the periplasm and presents these subunits to the outer membrane usher (17). PapD is a member of a large group of periplasmic chaperones of fimbrial systems of gram-negative bacteria. All these chaperones are characterized by invariant amino acid residues Arg-8 and Lys-112 necessary for the binding of subunits during pilus assembly (33, 52). In SfpD, Arg-8 and Lys-112 are conserved, indicating that this protein is the chaperone necessary for the assembly of Sfp fimbriae.

The genes in the 3′ part of the *sfp* cluster differed more from their corresponding *pap* genes than did *sfpA*, *sfpH*, *sfpC,* and *sfpD*. The sequence of SfpJ showed only 38% overall sequence similarity to PapJ, and the mature SfpJ had a pI of 4.8, whereas that of PapJ was 7.11. However, the molecular masses of the two proteins, about 21 kDa, are similar. Interestingly, the sequence homology was not equally distributed over the whole protein. Whereas the N-terminal 76 residues of the mature protein (amino acids 21 to 96) showed 53% sequence similarity, the C-terminal part (residues 97 to 187) had only 25% similarity. SfpJ possessed structural motifs for the interaction with the periplasmic chaperone at the N and C termini (Fig. 2). The C-terminal motif, however, seemed to be conserved to a lesser extent than that of SfpA, SfpH, and SfpG. Differences similar to these are described in a comparison of PapJ to Pap pilus subunits (55). On the other hand, some characteristics which distinguish PapJ from typical pilins were not detected in SfpJ. The mature SfpJ proteins possess only two cysteine residues, whereas PapJ has five. Furthermore, a hydrophobic domain and a sequence suggested to act as an ATP binding site

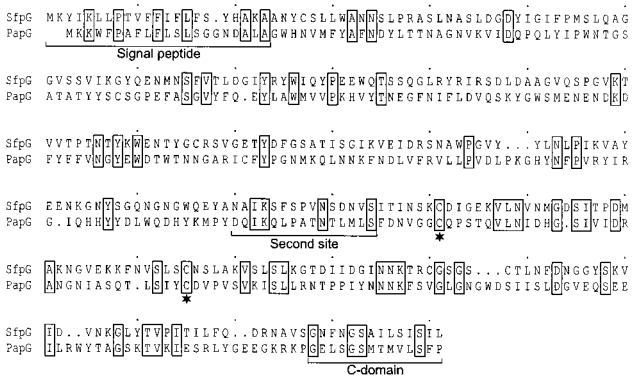

FIG. 4. Sequence comparison between SfpG and $PapG_{F13}$ (36). The alignment was produced using CLUSTAL W. Identical amino acid residues are depicted by boxes. The predicted signal peptide as well as the sites responsible for the interaction between PapG and the periplasmic chaperone (second site and C domain) are indicated by the brackets. Conserved pairs of cysteine residues in the C-terminal protein domain are marked by asterisks.

in PapJ were not detected in SfpJ. We assume, therefore, that SfpJ is a more typical pilin than PapJ and might be a component of the pilus structure. PapJ, however, has been shown not to be part of the Pap pilus but to be necessary for its integrity and is thought to work as a cochaperone (55).

The region downstream of *sfpJ* showed 53% nucleotide sequence identity over 516 bp with the sequence of the P-pilus minor subunit gene *papF*. However, the ATG start codon seemed to be changed to a TTG codon, and the protein deduced from the ORF, starting at the first ATG codon, represents only the C-terminal half of a protein homologous to PapF. In addition, the deduced amino acid sequence did not show the properties of a signal peptide necessary for transport through the cytoplasmic membrane. We suggested, therefore, that *sfpF* is a nonfunctional ORF.

The gene located 21 bp downstream of *sfpF* and termed *sfpG* showed significant homology to neither a *pap* gene nor any other known sequence when its product was compared to amino acid or nucleotide sequence databases using Fasta or Blast algorithms. However, in addition to its position immediately downstream of *sfpF* (Fig. 1), it shared properties with the *papG* gene, encoding the adhesin of P-fimbriae. The protein deduced from *sfpG* showed a molecular mass and pI similar to those of PapG (Table 2). Pairwise comparison of SfpG and $PapG_{F13}$ sequences revealed local similarities, especially in the C-terminal half (Fig. 4). This C-terminal domain is known as the assembly domain of PapG and is necessary for the interaction with other pilus subunits and the assembly machinery, whereas the N-terminal half of PapG is responsible for adhesin specificity (23). As for SfpA, SfpH, and SfpJ, a C-terminal chaperone recognition motif was conserved in SfpG with the exception of the penultimate aromatic residue, which is changed to an isoleucine residue (Fig. 2). Nevertheless, the P-fimbrial adhesin allele PrsG similarly possesses a penultimate leucine in place of an aromatic residue (36). Some residues of the second site of interaction between PapG and the periplasmic chaperone (61) were also conserved in SfpG (Fig. 2). Furthermore, a pair of cysteine residues (positions 195 and 229 of the mature protein) with a 33-residue spacing in the C-terminal half of the protein is similar to the cysteine pair with a 31-residue spacing establishing a conserved disulfide bridge in PapG (22) (Fig. 4).

Structure of the *sfp* gene cluster and comparison with the *pap* cluster. Sequence comparison of the whole *sfp* gene cluster with the *pap* cluster showed that several *pap* genes had no equivalent in the *sfp* system (Fig. 1). DNA regions, homologous to *papK* and *papE,* which encode subunits of the tip fibrillum of P-fimbriae, could not be detected. In addition, the homologue of *papF*, the third gene encoding a component of the tip fibrillum, appeared to be nonfunctional. From these data we suggest that the structure of the Sfp pilus is different from that of the Pap pilus, especially with regard to the tip fibrillum. We also found no genes corresponding to the regulatory genes *papI* and *papB* and also no homology to an ORF encoding a 17-kDa protein which is typically present downstream of the *papG* gene (36). Moreover, the respective intergenic regions between the *sfp* and *pap* genes were much more different from one another, both in length and sequence, than the genes themselves. The region between *sfpH* and *sfpC* in particular was three times longer than that between *papH* and *papC*. A search using the program NNPP revealed putative promoter sequences starting 111 bp upstream of *sfpA*, 91 bp upstream of *sfpC,* and 117 bp upstream of *sfpG*.

The G+C content of the *sfp* genes varies over the gene cluster (Table 2). Whereas the *sfpH*, *sfpC,* and *sfpD* genes have a G+C content close to the typical G+C content of *E. coli* genes, the *sfpA*, *sfpJ,* and *sfpG* genes contain only 36 to 42% G+C. Such variations in the base composition have also been reported for the *pap* genes (36). The *sfpJ* gene, however, showed a G+C content of 40.2%, whereas *papJ* has 55.5%. As shown for the homology between *sfpJ* and *papJ* (see above), the G+C content also was not equally distributed over the *sfpJ* gene: the *papJ* homologous N-terminal part had a G+C content of 48% compared to only 32% G+C in the C-terminal half of the gene.

DNA sequences adjacent to the *sfp* gene cluster. In order to investigate the DNA regions neighboring the *sfp* gene cluster and to look for possible regulatory genes of the system, we cloned and sequenced restriction fragments of pSFO157 contiguous to pSFO157-E11 (Fig. 5A). A 7.7-kb *Bam*HI fragment, termed pSFO157-B8 and located downstream of the *sfp* gene cluster, was identified by Southern blot hybridization of restricted pSFO157 DNA using pSFO157-E11 as the probe. Based on continuous sequence homologies to Tn*2501* (see below), the 13.9-kb *Eco*RI fragment pSFO157-E14 was hypothesized to be located upstream of *sfpA*. In order to examine this hypothesis and to sequence the connection between the two fragments, we generated PCR primers wprom-3 and wprom-4 (Table 1 and Fig. 5A). A PCR using this primer pair and DNA of strain 3072/96 as the template revealed the expected product as having a length of 1,900 bp. The nucleotide sequence of the PCR product was identical to the ends of pSFO157-E14 and pSFO157-E11, and there was no additional DNA present between the two fragments.

Nucleotide sequencing of pSFO157-E14 and pSFO157-B8 revealed no further sequences similar to that of fimbrial subunit or regulatory genes. We found a DNA region of 2,054 bp which was 97% identical to a RepFIB origin of replication 686

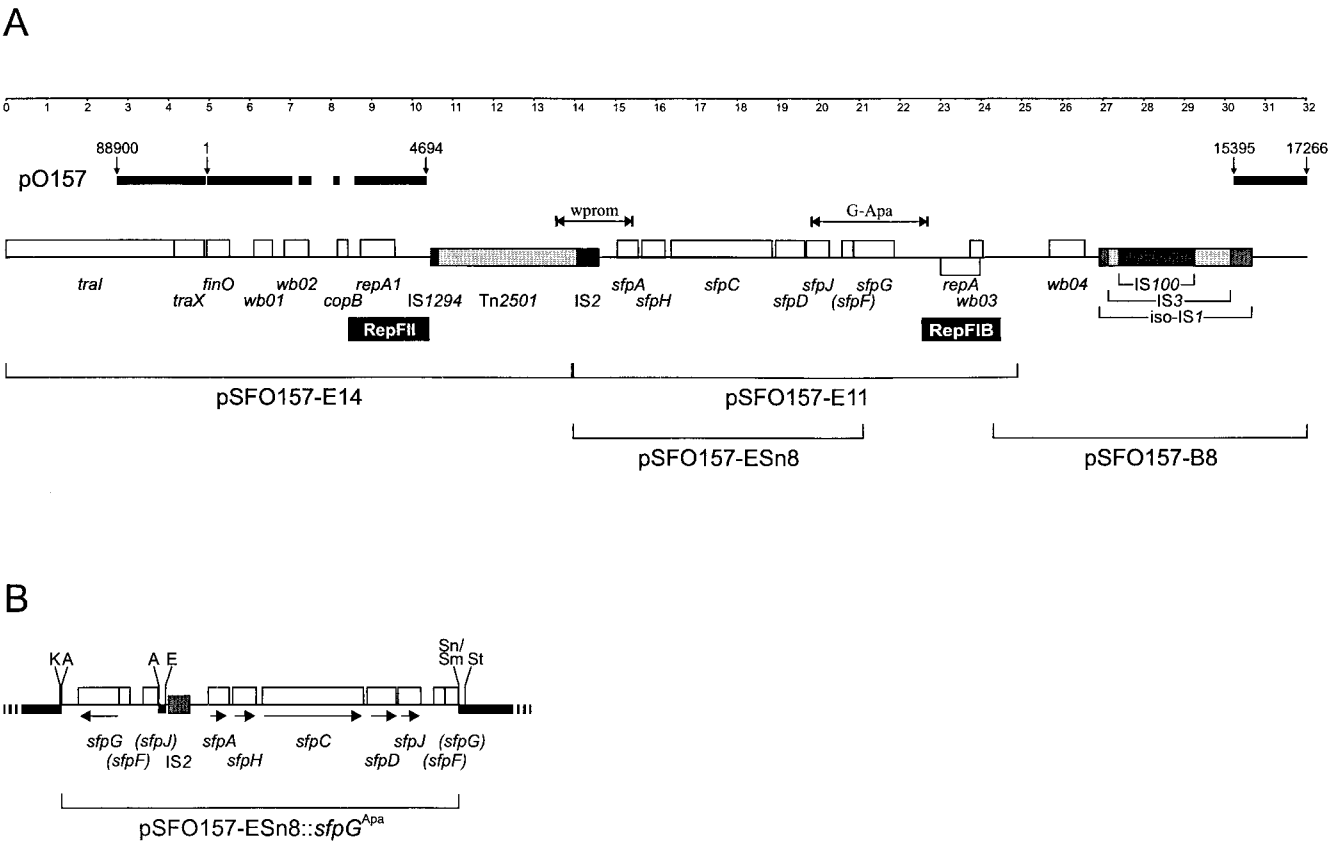

FIG. 5. (A) Genetic organization of the *sfp* gene cluster within plasmid pSFO157. White boxes indicate open reading frames, and shaded boxes depict regions with homology to IS elements. Putative origins of replication (RepFII and RepFIB) are indicated below the map. The brackets in the lower row show the localization of plasmid fragments cloned in this study. The double arrow labeled "wprom" indicates the PCR amplicon used for the sequencing of the connection between fragment pSFO157-E14 and pSFO157-E11. The arrow labeled "G-Apa" depicts the location of the PCR product used for the cloning of *sfpG*. Above the gene map, nearly identical sequences of plasmid pO157 of NSF EHEC O157:H7 strain EDL933 are depicted by the bold line together with their sequence positions (GenBank AF074613). The ruler depicts map positions of pSFO157 in kilobase pairs. (B) Map of the recombinant plasmid pSFO157-ESn8::*sfpG*^{Apa}. The G-Apa PCR amplicon was inserted into the *Apa*I restriction site of the *sfpG* mutant plasmid pSFO157-ESn8. Plasmid parts derived from the vector pBluescript II KS are depicted as a bold black line together with relevant restriction sites (A, *Apa*I; E, *Eco*RI; K, *Kpn*I; Sm, *Sma*I; Sn, *Sna*BI; St, *Sst*I). Arrows indicate the transcription direction of the ORFs. Nonfunctional ORFs are given in parentheses.

bp downstream of the *sfp* cluster. This sequence enclosed the complete minimal replication region of IncFI plasmids, including an ORF for the replication protein RepA (50) (Fig. 5A). The *sfp* gene cluster, together with this RepFIB origin, is flanked by two blocks composed of different insertion sequences (IS). Upstream of *sfpA*, a region of 3,415 bp has 76% similarity to the transposase and resolvase genes of Tn*2501* (38, 39). Adjacent to the Tn*2501*-like sequences we detected homologies to IS*2* (46) and IS*1294* (54) (Fig. 5A). We found another IS-related region downstream of the RepFIB origin composed of an IS*100* (99.6% identity but with an internal deletion of 95 bp) (37) inserted into an IS*3* element (80% similarity) (57), which is itself inserted into an iso-IS*1*-like sequence (68% similarity) (40). None of the IS elements flanking the *sfp* gene cluster seemed to be functional, because in all cases the transposase gene was not fully conserved due to partial deletion, insertion, or frameshift mutations. Nevertheless, the terminal inverted repeats of Tn*2501*, iso-IS*1*, IS*3,* and IS*100* are conserved. In addition, the IS*100* and IS*3* elements are flanked by direct repeats, indicating that these IS elements have been inserted into the plasmid by regular transposition events.

Upstream of the IS*1294*-like sequence we found a region of 2,017 bp homologous to the origin of replication of IncFII plasmid R100 (47), including an ORF encoding the replication protein RepA1 (Fig. 5A). At the distal end of pSFO157-E14 we detected further homologies to parts of plasmid R100, comprising *traI*, *traX,* and *finO* (Fig. 5A).

A DNA region similar to that of plasmid R100 is also present in plasmid pO157 of NSF EHEC O157:H7. We therefore compared the sequences adjacent to the *sfp* cluster to the complete sequence of pO157 (11) and found a large region upstream of the *sfp* cluster which was nearly identical in the two sequences and included *traI*, *traX*, *finO*, and the RepFII origin of replication (Fig. 5A). These identical sequences, however, are interrupted by segments which show differences. Downstream of the *sfp* gene cluster we also detected DNA identical to the sequence of pO157, beginning from the right part of iso-IS*1* (Fig. 5A). In pO157, the region between the RepFII origin and the iso-IS*1*-like sequences is covered by the

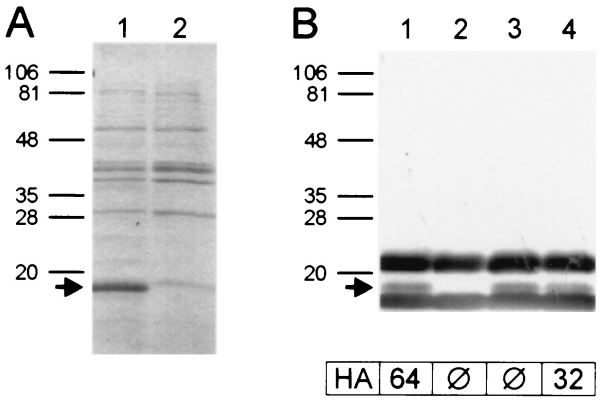

FIG. 6. Coomassie blue-stained SDS-PAGE (A) and immunoblot analysis using anti-SfpA antiserum (B) of thermoeluted fimbrial extracts taken from recombinant *E. coli* strains HB101/pSFO157-E11 (lane 1), HB101/pBluescript II KS (vector control) (lane 2), HB101/pSFO157-ESn8 (*sfpG* mutant) (lane 3), and pSFO157-ESn8::*sfpG*^{Apa} (lane 4). Positions and sizes of marker proteins (in kilodaltons) are given on the left. The arrows indicate the SfpA protein as identified by amino acid sequencing. Titers of mannose-resistant hemagglutination (HA) of the strains are given below the immunoblot. The bands strongly reacting with anti-SfpA antiserum above and below the SfpA protein and not visible in Coomassie blue-stained SDS-PAGE are suspected to represent nonproteinaceous components of the bacterial cell envelope present in the SfpA preparation used for immunization. These bands were not detected using rabbit normal serum, nor were they visible in silver-stained gels.

katP (8) and *espP* (9) genes together with the incomplete insertion elements IS*91*, IS*600,* and IS*1203*. We suggested, therefore, that the *sfp* cluster, together with an RepFIB origin of replication and several IS elements, is inserted into pSFO157 in the region where *espP* and *katP* reside in pO157.

Expression of Sfp fimbriae by recombinant *E. coli* strains. In order to investigate the expression of Sfp fimbriae, we transformed pSFO157-E11 into *E. coli* K-12 strain HB101. This strain was used because it is known to lack type I fimbriae (6). The Fim⁻ phenotype of our laboratory stock of *E. coli* HB101 was confirmed by its inability to agglutinate yeast cells.

Analysis of thermoeluted extracts of strain HB101/pSFO157-E11 by SDS-PAGE revealed a protein with an apparent molecular mass of about 17 kDa which was not expressed by the vector control strain HB101/pBluescript II KS (Fig. 6A). N-terminal amino acid sequencing of this protein revealed the sequence ASQGQGIINF, corresponding to residues 21 to 30 of the deduced amino acid sequence of SfpA. The starting point corresponds to the predicted cleavage site of the signal peptide (Table 2).

Recombinant *E. coli* HB101/pSFO157-E11 and the vector control strain HB101/pBluescript II KS were also analyzed by transmission electron microscopy for the presence of fimbriae. Whereas the vector control strain was not fimbriated (data not shown), we detected fimbriae on cells of strain HB101/pSFO157-E11 (Fig. 7A). The fimbriae were up to 0.4 μm in length and had a diameter of 3 to 5 nm.

Taken together, our data showed that the *sfp* gene cluster in plasmid pSFO157-E11 is sufficient to mediate the expression of fimbriae in *E. coli* K-12 and that SfpA is the major subunit of Sfp fimbriae.

Construction and analysis of an *sfpG* deletion mutant. In order to test our hypothesis that *sfpG* is part of the *sfp* cluster, even though it lacks homology to known fimbrial proteins, and that it encodes the adhesin subunit of the *sfp* fimbriae, we constructed a mutant lacking a functional *sfpG* gene. For this purpose we restricted plasmid pSFO157-E11 with the enzymes *Eco*RI and *Sna*BI and ligated the resulting 7,259-bp fragment into the vector pBluescript II KS, which had been restricted with *Eco*RI and *Sma*I. The resulting plasmid, pSFO157-ESn8, was transformed into *E. coli* HB101, and the deletion of *sfpG* was confirmed by nucleotide sequencing. Compared to plasmid pSFO157-E11, pSFO157-ESn8 lacked the RepFIB origin of replication and 683 bp of the 3′ part of *sfpG* but still retained *sfpA*, *sfpH*, *sfpC*, *sfpD,* and *sfpJ* (Fig. 5A).

The mutant strain HB101/pSFO157-ESn8 as well as strain HB101/pSFO157-E11 and the vector control strain HB101/pBluescript II KS were investigated for hemagglutination by a quantitative assay. Whereas the strain harboring the whole *sfp* cluster showed a mannose-resistant hemagglutination titer of 64, the mutant as well as the vector control strain was unable to agglutinate human red blood cells. On the other hand, immunoblot analysis of thermoeluted fimbrial proteins of the strains revealed that the mutant strain still possessed SfpA on its surface (Fig. 6B). Furthermore, we could detect fimbriae on cells of the *sfpG* mutant strain by electron microscopy (no more than a few per cell were detected, however) (Fig. 7B). The morphology of these fimbriae showed no obvious differences to those observed on strain HB101/pSFO157-E11 (Fig. 7A).

Attempts to clone *sfpG* in order to *trans*-complement the gene on a second plasmid failed, possibly due to a toxic side effect of SfpG when expressed without the chaperone and usher system (25). We therefore performed *cis*-complementation of *sfpG*. For this purpose we amplified *sfpG* by high-fidelity PCR using primers SfpG-UApa and SfpG-LApa, which contain *Apa*I restriction sites (Table 1 and Fig. 5A). The 3,162-bp amplicon was restricted with *Apa*I, ligated into *Apa*I-restricted plasmid pSFO157-ESn8, and transformed into *E. coli* HB101. The resulting plasmid, pSFO157-ESn8::*sfpG*^{Apa}, was verified using restriction enzyme analysis and nucleotide sequencing. In this construct the *sfpG* gene is inserted upstream and in the opposite direction of the other *sfp* genes (Fig. 5B).

As shown by immunoblot analysis, strain HB101/pSFO157-ESn8::*sfpG*^{Apa} produced the SfpA protein (Fig. 6B). The strain also exhibited a hemagglutination titer of 32, one step lower than that of strain HB101/pSFO157-E11. Electron microscopy revealed fimbriated cells; however, the number of fimbriae seemed to be reduced compared to that of strain HB101/pSFO157-E11 (Fig. 7C). It is possible that the coordinated expression of all fimbrial subunits is impaired in the artificial complementation construct, leading to problems in the assembly of the fimbriae.

Prevalence of the *sfp* gene cluster in *E. coli* and other *Enterobacteriaceae*. In order to determine the prevalence of the novel fimbrial gene cluster in pathogenic *E. coli*, we constructed PCR primer pairs specific for two different regions of the *sfp* cluster, namely, the *sfpA* gene and the region from *sfpD* to *sfpG* (Fig. 1 and Table 1). Using these primer pairs we examined 107 clinical isolates of *E. coli,* including 74 EHEC

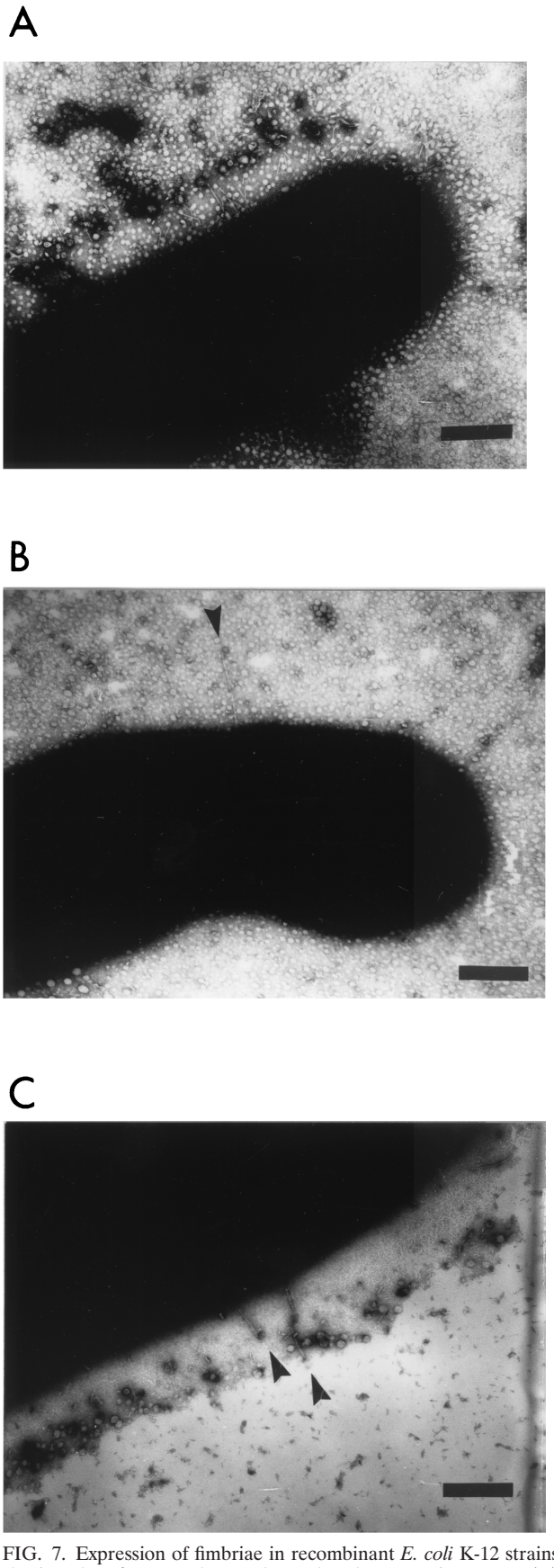

FIG. 7. Expression of fimbriae in recombinant *E. coli* K-12 strains harboring the *sfp* gene cluster shown by electron microscopy. (A) Strain HB101/pSFO157-E11; (B) *sfpG* mutant strain HB101/pSFO157-ESn8; (C) complemented mutant HB101/pSFO157-ESn8::*sfpG*^{Apa}. The bar represents 200 nm.

TABLE 3. Prevalence of the *sfp* gene cluster in human pathogenic *E. coli* and other *Enterobacteriaceae* examined by PCR and colony blot hybridization

Pathogen[a] or serotype	Total no.	No. *sfpA* positive	No. *sfpDG* positive
SF EHEC O157:H$^-$	14	14	14
NSF EHEC O157:H7/H$^-$	7	0	0
EHEC O26:H11/H$^-$	16	0	0
EHEC O111:H2/H$^-$	7	0	0
EHEC O103:H2	6	0	0
Other EHECs	24	0	0
EPEC	9	0	0
EAEC	6	0	0
ETEC	4	0	0
EIEC	4	0	0
UPEC	10	0	0
Other *Enterobacteriaceae*	15	0	0

[a] EAEC, enteroaggregative *E. coli*; ETEC, enterotoxigenic *E. coli*; EIEC, enteroinvasive *E. coli*.

strains of different serotypes, 23 strains of other diarrheagenic *E. coli* (EPEC, enteroaggregative *E. coli*, enterotoxigenic *E. coli*, and enteroinvasive *E. coli*), and 10 UPEC strains. In addition, 15 isolates of other enterobacterial species, such as *Enterobacter* spp., *Klebsiella pneumoniae*, *P. mirabilis*, *Salmonella enterica*, *S. marcescens*, *Yersinia enterocolitica,* and *Yersinia pseudotuberculosis* were included in the study. Whereas all 14 isolates of SF EHEC O157:H$^-$, including one strain isolated from a cow in the Czech Republic (4), possessed both the *sfpA* and *sfpDG* region, all the other 108 strains tested were negative in both PCRs (Table 3). The *E. coli* K-12 strains DH5α and HB101 used as hosts for cloning and expression of the *sfp* cluster were also negative by PCR. In order to exclude the possibility that the PCR had failed to detect the *sfp* genes due to minor sequence variations within the primer binding sites, we performed colony blot hybridization using probes specific for *sfpA* and *sfpDG*. The results of these hybridization experiments confirmed the PCR results in all cases. Therefore, we conclude that the *sfp* gene cluster is specific for the SF EHEC O157:H$^-$ group among human pathogenic *E. coli*.

DISCUSSION

In this paper we describe a novel fimbrial gene cluster harbored by the large plasmid of SF EHEC O157:H$^-$ strains. Sfp fimbriae are highly homologous to P-fimbriae of UPEC regarding their assembly machinery. However, in phylogenetic analyses the major pilus subunit, SfpA, does not cluster together with typical PapA serotypes and also shares structural similarity with the major subunit of *P. mirabilis* Pmp fimbriae. Sfp fimbriae are also distinct from P-fimbriae with regard to the organization of their gene clusters. Compared to *pap*, the *sfp* cluster lacks genes homologous to the regulatory elements *papI* and *papB* and in particular lacks all genes coding for parts of the tip fibrillum (*papK*, *papE*, and *papF*). The absence of functional homologues of *papKEF* is not unique to strain 3072/

96, from which the *sfp* cluster has been cloned. As we have shown by PCR, all SF O157:H$^-$ EHEC isolates possessing *sfpA* also gave an *sfpDG* PCR product having the same length as that obtained with strain 3072/96. Furthermore, nucleotide sequencing of the *sfpDG* PCR product of strain 493/89 isolated from a patient suffering from HUS in 1989 (26) revealed no differences compared to that of strain 3072/96, and this included the lack of an ATG start codon in the *papF* homologous region. The absence of tip subunit genes led us suppose either that Sfp fimbriae lack a tip fibrillum completely or that a similar structure is composed of other subunits, possibly SfpJ and/or SfpG.

The *sfp* gene differing most from *pap* is the ORF we have named *sfpG*. Using a mutated *sfpG* we showed that this gene actually contributes to the Sfp fimbriae in recombinant *E. coli* K-12 strains. Although strains lacking *sfpG* are able to express *sfpA* on their surface and to assemble fimbriae, the cells do not show mannose-resistant hemagglutination. Therefore, it seems likely that SfpG is the adhesin necessary for binding to erythrocytes and possibly to other eukaryotic cells, and this might be a prerequisite for the colonization of a metazoic host by SF EHEC O157:H$^-$. Nevertheless, the observation that *sfpG* mutant bacteria express a reduced number of fimbriae even though the major pilus subunit, SfpA, is present on the surface led us to suppose that SfpG also has structural functions as a starter for the correct assembly of the fimbriae, as in the case of PapG and P-fimbriae (13).

The structure of the *sfp* cluster, which is not dissimilar to that of a reduced variant of a *pap* system, resembles the F17 fimbrial gene cluster. This cluster includes only four genes, *f17-D* and *f17-C*, encoding the chaperone and usher proteins, *f17-A* for the major pilus subunit, and the *f17-G* adhesin gene. A cluster of five genes is required for the expression of the F17-related fimbriae of *Haemophilus influenzae*, Hif and Haf. In addition to four proteins with functions similar to those of F17, the *hif* cluster encodes a minor subunit, HifD, probably serving as a terminator of pilus polymerization (35, 59). Another fimbrial system with only four structural genes (major subunit, adhesin, chaperon, and usher) and an additional regulator is involved in the expression of the SEF14 fimbriae of *Salmonella enteritidis* (14). In addition to the genetic structure, the morphology of Sfp fimbriae resembled that of F17 fimbriae and differed from that of P-fimbriae. Whereas the diameters of Sfp fimbriae and F17 fimbriae are 3 to 5 nm and 3 to 4 nm, respectively, P-pili have a diameter of 7 nm (10, 35). On the other hand, the sequence similarity of F17 and Sfp pilus subunits does not exceed 35%.

The role of plasmid-encoded Sfp fimbriae in the biology of SF EHEC O157:H$^-$ strains needs further investigation. It was not possible to detect Sfp fimbriae on wild-type EHEC O157:H$^-$ isolate 3072/96 under standard laboratory conditions, and variations in growth parameters such as type of media, temperature, aeration, osmolarity, and pH had no effect on their expression. The strain showed no mannose-resistant hemagglutination, and we were unable to detect the SfpA protein by immunoblot analysis. In order to exclude the possibility that the lack of fimbrial expression is a characteristic only of strain 3072/96, from which the *sfp* cluster had been cloned, we investigated 13 strains of SF EHEC O157:H$^-$ previously shown to harbor the fimbrial gene cluster. As in the case of strain 3072/96, we were also unable to detect Sfp fimbriae on these strains. We therefore concluded that, in contrast to the *E. coli* K-12 background of recombinant strains, the expression of Sfp fimbriae is strictly repressed in wild-type EHEC O157:H$^-$ strains. These findings resemble those of Elliott et al. (15), who examined the expression of fimbriae in EHEC O157:H7 and EPEC strains. The authors mutated a global regulator, termed Ler (locus of enterocyte effacement-encoded regulator), and found enhanced fimbrial expression and expression of novel types of fimbriae in the mutant strains. It is therefore possible that a similar repression of *sfp* genes by a strong repressor also takes place in SF EHEC O157:H$^-$ strains. If this is the case, then Sfp fimbriae may be expressed only under environmental conditions not easily mimicked in the laboratory and which occur only during an infection stage in humans, in an animal reservoir, or during a free phase in the environment.

Strains of SF EHEC O157:H$^-$ have been implicated in several outbreaks, as well as in sporadic cases, of HUS and diarrhea in Germany and the Czech Republic (1, 3, 20, 28). Although SF EHEC O157:H$^-$ strains represent a distinct clone within *E. coli* O157 (26), multilocus enzyme electrophoresis, sequence data of the β-glucuronidase gene, and multilocus sequence typing indicate that they are closely related to NSF EHEC O157:H7 strains (16, 44). It was hypothesized that both groups have evolved from a common EPEC-like O55:H7 ancestor and that the as-yet-unknown most recent common ancestor of SF EHEC O157:H$^-$ and NSF EHEC O157:H7 strains already possessed the large EHEC plasmid pO157. However, the sequence data provided in this study represent a segment of plasmid pSFO157 of SF EHEC O157:H$^-$ which differs markedly from pO157. The *sfp* fimbrial gene cluster, as we have shown, is unique to pSFO157 and seems to be inserted into a region corresponding to that where *katP* and *espP* reside in pO157. This is in accord with earlier observations showing that plasmids of SF EHEC O157:H$^-$ strains are distinct from those of NSF EHEC O157:H7/H$^-$ strains with regard to the *katP* and *espP* genes, which are present in NSF but not in SF strains (7). The remnants of different IS flanking the *sfp* cluster in pSFO157, as well as the *katP* and *espP* genes in pO157, lead us suppose that transposition processes took part in the creation of these differences. On the other hand, the existence of the RepFIB-like origin of replication downstream of the *sfp* cluster indicates that a second plasmid is the source of the fimbrial gene cluster and that this plasmid underwent replicon fusion with a pO157-like precursor of pSFO157. However, the gene pool from which SF EHEC O157:H$^-$ acquired the *sfp* fimbrial genes is not known, and the question of why this gene cluster, despite its putative mobility, is so unique among our collection of *Enterobacteriaceae* cannot be answered at this time.

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (grant Ka 717/2–4).

We thank Stefanie Amersbach for excellent technical assistance.

REFERENCES

1. **Ammon, A., L. R. Petersen, and H. Karch.** 1999. A large outbreak of hemolytic uremic syndrome caused by an unusual sorbitol-fermenting strain of *Escherichia coli* O157:H$^-$. J. Infect. Dis. **179:**1274–1277.

2. **Baga, M., M. Norgren, and S. Normark.** 1987. Biogenesis of *E. coli* Pap pili: PapH, a minor pilin subunit involved in cell anchoring and length modulation. Cell **49:**241–251.
3. **Bielaszewska, M., H. Schmidt, M. A. Karmali, R. Khakhria, J. Janda, K. Bláhová, and H. Karch.** 1998. Isolation and characterization of sorbitol-fermenting Shiga toxin (Verocytotoxin)-producing *Escherichia coli* O157:H$^-$ strains in the Czech Republic. J. Clin. Microbiol. **36:**2135–2137.
4. **Bielaszewska, M., H. Schmidt, A. Liesegang, R. Prager, W. Rabsch, H. Tschäpe, A. Cízek, J. Janda, K. Bláhová, and H. Karch.** 2000. Cattle can be a reservoir of sorbitol-fermenting Shiga toxin-producing *Escherichia coli* O157:H$^-$ strains and a source of human diseases. J. Clin. Microbiol. **38:** 3470–3473.
5. **Bijlsma, I. G., L. Van Dijk, J. G. Kusters, and W. Gaastra.** 1995. Nucleotide sequences of two fimbrial major subunit genes, *pmpA* and *ucaA*, from canine-uropathogenic *Proteus mirabilis* strains. Microbiology **141:**1349–1357.
6. **Blomfield, I. C., M. S. McClain, and B. I. Eisenstein.** 1991. Type 1 fimbriae mutants of *Escherichia coli* K12: characterization of recognized afimbriate strains and construction of new *fim* deletion mutants. Mol. Microbiol. **5:**1439–1445.
7. **Brunder, W., H. Schmidt, M. Frosch, and H. Karch.** 1999. The large plasmids of Shiga-toxin-producing *Escherichia coli* (STEC) are highly variable genetic elements. Microbiology **145:**1005–1014.
8. **Brunder, W., H. Schmidt, and H. Karch.** 1996. KatP, a novel catalase-peroxidase encoded by the large plasmid of enterohaemorrhagic *Escherichia coli* O157:H7. Microbiology **142:**3305–3315.
9. **Brunder, W., H. Schmidt, and H. Karch.** 1997. EspP, a novel extracellular serine protease of enterohaemorrhagic *Escherichia coli* O157:H7 cleaves human coagulation factor V. Mol. Microbiol. **24:**767–778.
10. **Bullitt, E., and L. Makowski.** 1995. Structural polymorphism of bacterial adhesion pili. Nature **373:**164–167.
11. **Burland, V., Y. Shao, N. T. Perna, G. Plunkett, H. J. Sofia, and F. R. Blattner.** 1998. The complete DNA sequence and analysis of the large virulence plasmid of *Escherichia coli* O157:H7. Nucleic Acids Res. **26:**4196–4204.
12. **Denich, K., L. B. Blyn, A. Craiu, B. A. Braaten, J. Hardy, D. A. Low, and P. D. O'Hanley.** 1991. DNA sequences of three *papA* genes from uropathogenic *Escherichia coli* strains: evidence of structural and serological conservation. Infect. Immun. **59:**3849–3858.
13. **Dodson, K. W., F. Jacob-Dubuisson, R. T. Striker, and S. J. Hultgren.** 1993. Outer-membrane PapC molecular usher discriminately recognizes periplasmic chaperone-pilus subunit complexes. Proc. Natl. Acad. Sci. USA **90:**3670–3674.
14. **Edwards, R. A., D. M. Schifferli, and S. R. Maloy.** 2000. A role for *Salmonella* fimbriae in intraperitoneal infections. Proc. Natl. Acad. Sci. USA **97:**1258–1262.
15. **Elliott, S. J., V. Sperandio, J. A. Girón, S. Shin, J. L. Mellies, L. Wainwright, S. W. Hutcheson, T. K. McDaniel, and J. B. Kaper.** 2000. The locus of enterocyte effacement (LEE)-encoded regulator controls expression of both LEE- and non-LEE-encoded virulence factors in enteropathogenic and enterohemorrhagic *Escherichia coli*. Infect. Immun. **68:**6115–6126.
16. **Feng, P., K. A. Lampel, H. Karch, and T. S. Whittam.** 1998. Genotypic and phenotypic changes in the emergence of *Escherichia coli* O157:H7. J. Infect. Dis. **177:**1750–1753.
17. **Fernández, L. A., and J. Berenguer.** 2000. Secretion and assembly of regular surface structures in gram-negative bacteria. FEMS Microbiol. Rev. **24:**21–44.
18. **Fratamico, P. M., S. Bhaduri, and R. L. Buchanan.** 1993. Studies on *Escherichia coli* serotype O157:H7 strains containing a 60-MDa plasmid and on 60-MDa plasmid-cured derivatives. J. Med. Microbiol. **39:**371–381.
19. **Girardeau, J. P., Y. Bertin, and I. Callebaut.** 2000. Conserved structural features in class I major fimbrial subunits (Pilin) in gram-negative bacteria. Molecular basis of classification in seven subfamilies and identification of intrasubfamily sequence signature motifs which might be implicated in quaternary structure. J. Mol. Evol. **50:**424–442.
20. **Gunzer, F., H. Böhm, H. Rüssmann, M. Bitzan, S. Aleksic, and H. Karch.** 1992. Molecular detection of sorbitol-fermenting *Escherichia coli* O157 in patients with hemolytic-uremic syndrome. J. Clin. Microbiol. **30:**1807–1810.
21. **Higgins, D. G., J. D. Thompson, and T. J. Gibson.** 1996. Using CLUSTAL for multiple sequence alignments. Methods Enzymol. **266:**383–402.
22. **Hultgren, S. J., C. H. Jones, and S. Normark.** 1996. Bacterial adhesins and their assembly, p. 2730–2756. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.
23. **Hultgren, S. J., F. Lindberg, G. Magnusson, J. Kihlberg, J. M. Tennent, and S. Normark.** 1989. The PapG adhesin of uropathogenic *Escherichia coli* contains separate regions for receptor binding and for the incorporation into the pilus. Proc. Natl. Acad. Sci. USA **86:**4357–4361.
24. **Johnson, J. R., A. L. Stell, F. Scheutz, T. T. O'Bryan, T. A. Russo, U. B. Carlino, C. Fasching, J. Kavle, L. Van Dijk, and W. Gaastra.** 2000. Analysis of the F antigen-specific *papA* alleles of extraintestinal pathogenic *Escherichia coli* using a novel multiplex PCR-based assay. Infect. Immun. **68:**1587–1599.
25. **Jones, C. H., P. N. Danese, J. S. Pinkner, T. J. Silhavy, and S. J. Hultgren.** 1997. The chaperone-assisted membrane release and folding pathway is sensed by two signal transduction systems. EMBO J. **16:**6394–6406.
26. **Karch, H., H. Böhm, H. Schmidt, F. Gunzer, S. Aleksic, and J. Heesemann.** 1993. Clonal structure and pathogenicity of Shiga-like toxin-producing, sorbitol-fermenting *Escherichia coli* O157:H$^-$. J. Clin. Microbiol. **31:**1200–1205.
27. **Karch, H., J. Heesemann, R. Laufs, A. D. O'Brien, C. O. Tacket, and M. M. Levine.** 1987. A plasmid of enterohemorrhagic *Escherichia coli* O157:H7 is required for expression of a new fimbrial antigen and for adhesion to epithelial cells. Infect. Immun. **55:**455–461.
28. **Karch, H., R. Wiβ, H. Gloning, P. Emmrich, S. Aleksic, and J. Bockemühl.** 1990. Hämolytisch-urämisches Syndrom bei Kleinkindern durch Verotoxin-produzierende *Escherichia coli*. Dtsch. Med. Wochenschr. **115:**489–495.
29. **Karmali, M. A., B. T. Steele, M. Petric, and C. Lim.** 1983. Sporadic cases of haemolytic-uraemic syndrome associated with faecal cytotoxin and cytotoxin-producing *Escherichia coli* in stools. Lancet **i:**619–620.
30. **Khan, A. S., B. Kniep, T. A. Oelschlaeger, I. van Die, T. Korhonen, and J. Hacker.** 2000. Receptor structure for F1C fimbriae of uropathogenic *Escherichia coli*. Infect. Immun. **68:**3541–3547.
31. **Khan, A. S., I. Mühldorfer, V. Demuth, U. Wallner, T. K. Korhonen, and J. Hacker.** 2000. Functional analysis of the minor subunits of S fimbrial adhesion (SfaI) in pathogenic *Escherichia coli*. Mol. Gen. Genet. **263:**96–105.
32. **Khan, A. S., and D. M. Schifferli.** 1994. A minor 987P protein different from the structural fimbrial subunit is the adhesin. Infect. Immun. **62:**4233–4243.
33. **Kuehn, M. J., D. J. Ogg, J. Kihlberg, L. N. Slonim, K. Flemmer, T. Bergfors, and S. J. Hultgren.** 1993. Structural basis of pilus subunit recognition by the PapD chaperone. Science **262:**1234–1241.
34. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680–685.
35. **Le Bouguenec, C., and Y. Bertin.** 1999. AFA and F17 adhesins produced by pathogenic *Escherichia coli* strains in domestic animals. Vet. Res. **30:**317–342.
36. **Marklund, B. I., J. M. Tennent, E. Garcia, A. Hamers, M. Baga, F. Lindberg, W. Gaastra, and S. Normark.** 1992. Horizontal gene transfer of the *Escherichia coli pap* and *prs* pili operons as a mechanism for the development of tissue-specific adhesive properties. Mol. Microbiol. **6:**2225–2242.
37. **McDonough, K. A., and J. M. Hare.** 1997. Homology with a repeated *Yersinia pestis* DNA sequence IS*100* correlates with pesticin sensitivity in *Yersinia pseudotuberculosis*. J. Bacteriol. **179:**2081–2085.
38. **Michiels, T., and G. Cornelis.** 1984. Detection and characterization of Tn*2501*, a transposon included within the lactose transposon Tn*951*. J. Bacteriol. **158:**866–871.
39. **Michiels, T., G. Cornelis, K. Ellis, and J. Grinsted.** 1987. Tn*2501*, a component of the lactose transposon Tn*951*, is an example of a new category of class II transposable elements. J. Bacteriol. **169:**624–631.
40. **Ohtsubo, H., K. Nyman, W. Doroszkiewicz, and E. Ohtsubo.** 1981. Multiple copies of iso-insertion sequences of IS*1* in *Shigella dysenteriae* chromosome. Nature **292:**640–643.
41. **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.
42. **Perna, N. T., G. F. Mayhew, G. Pósfai, S. Elliott, M. S. Donnenberg, J. B. Kaper, and F. R. Blattner.** 1998. Molecular evolution of a pathogenicity island from enterohemorrhagic *Escherichia coli* O157:H7. Infect. Immun. **66:**3810–3817.
43. **Pridmore, R. D.** 1987. New and versatile cloning vectors with kanamycin-resistance marker. Gene **56:**309–312.
44. **Reid, S. D., C. J. Herbelin, A. C. Bumbaugh, R. K. Selander, and T. S. Whittam.** 2000. Parallel evolution of virulence in pathogenic *Escherichia coli*. Nature **406:**64–67.
45. **Riley, L. W., R. S. Remis, S. D. Helgerson, H. B. McGee, J. G. Wells, B. R. Davis, R. J. Hebert, E. S. Olcott, L. M. Johnson, N. T. Hargrett, P. A. Blake, and M. L. Cohen.** 1983. Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. N. Engl. J. Med. **308:**681–685.
46. **Ronecker, H. J., and B. Rak.** 1987. Genetic organization of insertion element IS*2* based on a revised nucleotide sequence. Gene **59:**291–296.
47. **Rosen, J., T. Ryder, H. Inokuchi, H. Ohtsubo, and E. Ohtsubo.** 1980. Genes and sites involved in replication and incompatibility of an R100 plasmid derivative based on nucleotide sequence analysis. Mol. Gen. Genet. **179:**527–537.
48. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:**406–425.
49. **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.
50. **Saul, D., A. J. Spiers, J. McAnulty, M. G. Gibbs, P. L. Bergquist, and D. F. Hill.** 1989. Nucleotide sequence and replication characteristics of RepFIB, a basic replicon of IncF plasmids. J. Bacteriol. **171:**2697–2707.
51. **Schmoll, T., J. Hacker, and W. Goebel.** 1987. Nucleotide sequence of the

sfaA gene coding for the S-fimbrial protein subunit of *Escherichia coli*. FEMS Microbiol. Lett. **41:**229–235.
52. **Slonim, L. N., J. S. Pinkner, C. I. Brändén, and S. J. Hultgren.** 1992. Interactive surface in the PapD chaperone cleft is conserved in pilus chaperone superfamily and essential in subunit recognition and assembly. EMBO J. **11:**4747–4756.
53. **Tarr, P. I., S. S. Bilge, J. Vary, S. Jelacic, R. L. Habeeb, T. R. Ward, M. R. Baylor, and T. E. Besser.** 2000. Iha: a novel *Escherichia coli* O157:H7 adherence-conferring molecule encoded on a recently acquired chromosomal island of conserved structure. Infect. Immun. **68:**1400–1407.
54. **Tavakoli, N., A. Comanducci, H. M. Dodd, M. C. Lett, B. Albiger, and P. Bennett.** 2000. IS*1294*, a DNA element that transposes by RC transposition. Plasmid **44:**66–84.
55. **Tennent, J. M., F. Lindberg, and S. Normark.** 1990. Integrity of *Escherichia coli* P pili during biogenesis: properties and role of PapJ. Mol. Microbiol. **4:**747–758.
56. **Thanassi, D. G., E. T. Saulino, M. J. Lombardo, R. Roth, J. Heuser, and S. J. Hultgren.** 1998. The PapC usher forms an oligomeric channel: implications for pilus biogenesis across the outer membrane. Proc. Natl. Acad. Sci. USA **95:**3146–3151.
57. **Timmerman, K. P., and C. P. Tu.** 1985. Complete sequence of IS*3*. Nucleic Acids Res. **13:**2127–2139.
58. **Toth, I., M. L. Cohen, H. S. Rumschlag, L. W. Riley, E. H. White, J. H. Carr, W. W. Bond, and I. K. Wachsmuth.** 1990. Influence of the 60-megadalton plasmid on adherence of *Escherichia coli* O157:H7 and genetic derivatives. Infect. Immun. **58:**1223–1231.
59. **van Ham, S. M., L. van Alphen, F. R. Mooi, and J. P. M. van Putten.** 1994. The fimbrial gene cluster of *Haemophilus influenzae* type b. Mol. Microbiol. **13:**673–684.
60. **von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. **14:**4683–4690.
61. **Xu, Z., C. H. Jones, D. Haslam, J. S. Pinkner, K. Dodson, J. Kihlberg, and S. J. Hultgren.** 1995. Molecular dissection of PapD interaction with PapG reveals two chaperone-binding sites. Mol. Microbiol. **16:**1011–1020.

Editor: V. J. DiRita