JOURNAL OF BACTERIOLOGY, Nov. 1994, p. 6885–6891
0021-9193/94/$04.00+0

Vol. 176, No. 22

# A Plasmid-Encoded Prepilin Peptidase Gene from Enteropathogenic *Escherichia coli*

HONG-ZHONG ZHANG,[1] STEPHEN LORY,[2] AND MICHAEL S. DONNENBERG[1,3]*

*Molecular and Cell Biology Program, Division of Infectious Disease, Department of Medicine, University of Maryland at Baltimore,[1] and Medical Service, Veterans Affairs Medical Center,[3] Baltimore, Maryland 21201, and Department of Microbiology, School of Medicine, University of Washington, Seattle, Washington 98195[2]*

Received 13 June 1994/Accepted 22 September 1994

**Enteropathogenic *Escherichia coli*, a leading agent of infantile diarrhea worldwide, adheres to tissue culture cells in a pattern called "localized adherence." Localized adherence is associated with bundle-forming pili encoded by the plasmid *bfpA* gene, the product of which is homologous with the major structural subunit proteins of type IV fimbriae in other bacteria. Several of these proteins have been shown to be processed from a precursor by a specific prepilin peptidase. We cloned restriction fragments downstream of the *bfpA* gene into an *E. coli-Pseudomonas aeruginosa* shuttle vector and mobilized them into a *P. aeruginosa* prepilin peptidase (*pilD*) mutant. A plasmid containing a 1.3-kb *Pst*I-*Bam*HI fragment was able to complement the *pilD* mutation, as demonstrated by restoration of sensitivity to the pilus-specific bacteriophage PO4. The DNA sequence of this fragment revealed an open reading frame, designated *bfpP*, the predicted product of which is homologous to other prepilin peptidases, including TcpJ of *Vibrio cholerae* (30% identical amino acids), PulO of *Klebsiella oxytoca* (29%), and PilD of *P. aeruginosa* (28%). A *bfpA*::Tn*phoA* mutant complemented with a *bfpA*-containing DNA fragment only partially processes the BfpA protein. When complemented with a larger fragment containing *bfpP* as well as *bfpA*, the mutant expresses the fully processed BfpA protein. *P. aeruginosa* PAK, but not a *pilD* mutant of PAK, expresses mature BfpA protein when the *bfpA* gene is mobilized into this strain. Thus, as in other type IV fimbria systems, enteropathogenic *E. coli* utilizes a specific prepilin peptidase to process the major subunit of the bundle-forming pilus. This prepilin peptidase contains sequence and reciprocal functional homologies with the PilD protein of *P. aeruginosa*.**

Enteropathogenic *Escherichia coli* (EPEC) infection is a major cause of infantile diarrhea around the world, especially in developing countries (5, 17, 26). When incubated with tissue culture cells, EPEC adheres and forms microcolonies rather than blanketing the cells (42). This specialized adherence pattern, called "localized adherence," is associated with the presence of a 50-to 70-MDa "EPEC adherence factor" plasmid, common to all classic EPEC strains (32, 33). Loss of this plasmid from EPEC is associated with loss of the capacity for localized adherence in vitro (1) and with dramatically reduced ability to cause diarrhea in volunteers (28).

Recently, a unique bundle-forming pilus (BFP) has been identified in EPEC and proposed to be the adhesin responsible for localized adherence (16). Donnenberg et al. (9) cloned the plasmid gene interrupted by Tn*phoA* in mutants of the O127:H6 EPEC E2348/69 deficient in localized adherence. The identified gene, *bfpA*, encodes a protein with a predicted N-terminal amino acid sequence nearly identical to that determined experimentally for the major structural protein (bundlin) of BFP. The same *bfpA* gene has also been cloned by another group from an O111:NM EPEC strain using a different approach and found to have an identical nucleotide sequence (43). The predicted amino acid sequence of bundlin is homologous with other members of the type IV pilus family.

Type IV fimbriae are found in a variety of gram-negative bacteria including *Pseudomonas aeruginosa*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Moraxella bovis*, *Moraxcella nonliquefaciens*, *Dichelobacter nodosus*, and *Vibrio cholerae* (for a review, see reference 19). These fimbriae are composed of repeating subunits of a single polypeptide pilin, which is synthesized as a precursor (prepilin). The type IV prepilin proteins typically show a high degree of homology with each other, especially within the first 25 to 35 amino acids. These polypeptides, which range in molecular mass from 15 to about 18 kDa, contain a short, positively charged leader sequence of six or seven amino acids, an unusual modified amino acid at the amino terminus of the mature protein (usually *N*-methylphenylalanine), and a conserved and hydrophobic amino-terminal domain (13, 38). Previous research has suggested that many components of the type IV fimbria biogenesis apparatus are likely to be functionally conserved across bacterial species (2, 21, 25, 31). In some species, several of the genes involved in biogenesis of type IV fimbriae are clustered together close to the pilin structural gene. Those genes identified often include the peptidase responsible for amino-terminal processing of the prepilin and a cytoplasmic protein with highly conserved nucleoside triphosphate binding sites (25, 30, 48). The cleavage site for the type IV prepilin is usually between glycine and phenylalanine residues (methionine in TcpA of *V. cholerae*) in the sequence Gly-Phe(Met)-Thr-Leu-Ile(Leu)-Glu, which is highly conserved among members of the type IV fimbria family (44). The prebundlin from EPEC has a 13-amino-acid leader sequence with a similar potential cleavage site (Gly-Leu-Ser-Leu-Ile-Glu) (9, 43).

Prepilin peptidase genes have been identified in several species including *P. aeruginosa* (PilD), *Klebsiella oxytoca* (PulO), *V. cholerae* (TcpJ), and *N. gonorrhoeae* (PilD) (11, 23, 25, 35, 39). One of these (*P. aeruginosa* PilD) has been shown to be bifunctional, capable of both cleavage and subsequent N

* Corresponding author. Mailing address: Division of Infectious Diseases, University of Maryland School of Medicine, 10 S. Pine St., MSTF 900, Baltimore, MD 21201. Phone: (410) 706-8330. Fax: (410) 706-8700.

TABLE 1. Bacterial strains, phage, and plasmids

| Strain, phage, or plasmid | Relevant characteristics | Source or reference(s) |
|---|---|---|
| *E. coli* | | |
| DH5α | *supE44 lacU169* (ϕ *lacZ* M15) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1* | 41 |
| HB101 | *supE44 hsdSD20* ($r_b^-$ $m_b^-$) *recA13 ara-14 proA2 lacY1 galK2 rpsL20 xyl-5 mtl-1* | 41 |
| BL21 (DE3) | $F^-$ *ompT* $r_b^-$ $m_b^-$, lambda D69 lysogen carrying the phage T7 gene 1 under the control of $P_{lacUV5}$ | 47 |
| E2348/69 | Prototypical virulent O127:H6 EPEC strain | 27 |
| 31-6-1(1) | E2348/69 with *bfpA4*::Tn*phoA* | 8, 9 |
| *P. aeruginosa* | | |
| PAK | *P. aeruginosa* wild-type strain | D. Bradley |
| PAK-DΩ | Mutant of PAK with the Ω interposon inserted in the *pilD* gene | 24 |
| Phage PO4 | *P. aeruginosa* pilus-specific lytic phage | D. Bradley |
| Plasmids | | |
| pRK2073 | $Str^r$ mobilizing plasmid | 4 |
| pUC19 | $Ap^r$ cloning vector | New England BioLabs |
| pBluescript KS or SK | $Ap^r$ phagemid cloning vector | Stratagene |
| pDN18 or -19 | $Tet^r$ broad-host-range cloning vector | 34 |
| pMMB67HE | $Ap^r$ broad-host-range cloning vector, $lacI^q$/*tac* promoter | 14 |
| pMSD207 | *bfpA* gene cloned in pUC19 | 15 |
| pBFP6 | 0.8-kb *Hin*dIII-*Eco*RI fragment containing *bfpA* cloned from pMSD207 in pMMB67HE | This study |
| pMSD201 | 5.2-kb EPEC *Bam*HI fragment cloned in pBR322 | 9 |
| pMSD217 | 9.3-kb EPEC *Bam*HI fragment cloned in pACYC184 | This study |
| pHZZ4B1 or -2 | 4.0-kb EPEC *Bam*HI fragment cloned in pACYC184 in both orientations | This study |
| pDN19PB | 1.3-kb EPEC *Pst*I-*Bam*HI fragment cloned in pDN19 | This study |
| pDN18BP | 1.3-kb EPEC *Pst*I-*Bam*HI fragment cloned in pDN18 | This study |
| pDN19PBΔ | Derivative of pDN19PB with the 93-bp *Bsr*FI fragment deleted from the *Pst*I-*Bam*HI insert | This study |
| pDNXB1 or -2 | 2.1-kb *Xba*I fragment, containing the upstream EPEC 2.0-kb *Bam*HI-*Xba*I fragment, cloned from pHZZ4B2 in pDN18 in both orientations | This study |
| pDNP1 or -2 | 2.3-kb EPEC *Pst*I fragment cloned in pDN18 in both orientations | This study |
| pDNBX1 or -2 | 1.7-kb *Xba*I fragment, containing the downstream EPEC 1.6-kb *Xba*I-*Bam*HI fragment, cloned from pHZZ4B1 in pDN18 in both orientations | This study |
| pMALc-2 | Expression vector for creation of MBP fusion, with the *malE* signal sequence deleted | New England BioLabs |
| pMAL-*bfpP* | PCR-amplified *bfpP* cloned in pMALc-2 between *Bam*HI and *Hin*dIII | This study |
| pMAL-*bfpP*Δ | PCR-amplified Δ*bfpP* cloned in pMALc-2 between *Bam*HI and *Hin*dIII | This study |

methylation of its substrate (46). These prepilin peptidases are homologous over the entire amino acid sequence (250 to 290 amino acids). In some but not all cases, these genes are interchangeable and can process heterologous substrates (6, 7, 11, 12).

In this article, we describe the identification and cloning of an EPEC prepilin peptidase gene (*bfpP*) by its ability to complement a *P. aeruginosa pilD* mutant. The predicted amino acid sequence of the cloned EPEC gene product is homologous with other members of the prepilin peptidase family. We show that a DNA fragment containing this gene is required for processing of prebundlin to bundlin. In parallel, we also demonstrate that PilD can process prebundlin expressed in a *P. aeruginosa* strain.

## MATERIALS AND METHODS

**Strains, phage, plasmids, and growth conditions.** The bacterial strains, phage, and plasmids used in this study are listed in Table 1. Bacteria were grown at 37°C with aeration in Luria-Bertani (LB) broth. The following antibiotic concentrations (in micrograms per milliliter) were used for *E. coli*: ampicillin, 200; chloramphenicol, 20; kanamycin, 25; tetracycline, 30. For *P. aeruginosa*, the concentrations (in micrograms per milliliter) were as follows: carbenicillin, 300; streptomycin, 1,000; tetracycline, 100; spectinomycin, 200.

**DNA and genetic techniques.** Restriction endonuclease digestion, ligation, transformation, and DNA electrophoresis were performed as described elsewhere (41). Electroporation was performed in 0.1-cm cuvettes using an *E. coli* Gene Pulser (Bio-Rad, Richmond, Calif.) set at 1.8 kV. Shuttle vectors pDN18 and pDN19 and all of their derivatives were mobilized from *E. coli* DH5α into *P. aeruginosa* PAK or PAK-DΩ by conjugation with the helper plasmid pRK2073 from strain HB101. The three parental strains were cross-streaked on LB plates and incubated overnight. Growth from the intersection of the streaks was transferred onto LB plates containing tetracycline and spectinomycin to select *Pseudomonas* exconjugants from parental strains. Individual clones were verified by plasmid minipreparation using Wizard miniprep kits (Promega, Madison, Wis.).

**DNA sequence analysis.** The 1.3-kb *Pst*I-*Bam*HI fragment from pMSD217 was cloned in pDN18 and pDN19, resulting in pDN18BP and pDN19PB. Double-stranded DNA sequencing was performed with the dideoxy chain termination method (3) using the sequenase 2.0 kit (United States Biochemical, Cleveland, Ohio) and a T7 promoter primer. *Acc*I and *Eag*I sites revealed by this sequence were used to subclone fragments into pBluescript KS or SK. The nucleotide sequences of both strands were completed by sequencing these constructs with T7 or reverse primers.

**Phage sensitivity assay.** *P. aeruginosa* strains to be tested

were grown overnight in the presence of antibiotics. Suspensions of the strains were streaked onto LB agar plates containing antibiotics. A 5-μl sample of a suspension of the pilus-specific phage PO4 (titer, approximately $10^{10}$ PFU/ml) was spotted onto the center of the streak. Sensitivity was defined as absence of growth in the area of phage inoculation.

**Western blot (immunoblot) analysis.** *P. aeruginosa* PAK(pBFP6), PAK(pMMB67HE), PAK-DΩ(pBFP6), and PAK-DΩ(pMMB67HE) were grown in LB broth to an optical density at 600 nm 0.5, induced with IPTG (isopropyl-β-D-thiogalactopyranoside) (0.5 mM), and incubated for 2 h. EPEC E2348/69 and derivatives 31-6-1(1), 31-6-1(1)(pMSD201), and 31-6-1(1)(pMSD217) were grown overnight in Eagle's minimal essential medium to induce *bfpA* expression (9). Bacteria were pelleted, resuspended in sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) loading buffer (60 mM Tris-HCl [pH 6.8], 2% SDS, 10% glycerol, 0.025% bromphenol blue, 5% β-mercaptoethanol), and boiled for 5 min. Samples (15 μl) were separated by discontinuous SDS-PAGE (15% polyacrylamide). Proteins were electrotransferred onto Immobilon-P polyvinyl difluoride membranes (Millipore Corp). Blots were treated with phosphate-buffered saline–Tween (0.5%), reacted with an anti-BFP monoclonal antibody, incubated with alkaline phosphatase-conjugated anti-mouse immunoglobulin G (Sigma Immuno Chemicals, St. Louis, Mo.), and developed with nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate (18).

**Creating an in-frame deletion in *bfpP*.** The 1.3-kb *Pst*I-*Bam*HI fragment was cloned into pUC19. The resulting plasmid, pUC19PB, was digested with *Bsr*FI, religated, and transformed into *E. coli* DH5α. Plasmids from transformants were screened for the loss of the *Spe*I site (which lies between the two *Bsr*FI sites in *bfpP*) in the insert. The deleted version of the *Pst*I-*Bam*HI fragment was then cloned back into the shuttle vector pDN19 to create pDN19PBΔ.

**Expression of an MBP-BfpP fusion protein.** The protein product encoded by the *bfpP* gene was identified by expressing it as a maltose-binding protein fusion protein. The *bfpP* open reading frame and its in-frame deletion derivative Δ*bfpP* were amplified with two synthetic primers containing engineered *Hin*dIII and *Bam*HI sites: Donne-62 (5′-GCGAAGCTTTTAATGATAAACTAAACATAT-3′) and Donne-63 (5′-CGCGGATCCATGCAAGAAAGTATATTTCTA-3′). The PCR products were digested with *Hin*dIII and *Bam*HI and cloned into the fusion protein expression vector pMALc-2 (New England BioLabs, Beverly, Mass.). Expression of the fusion proteins was detected by immunoblot analysis using anti-MBP antiserum after IPTG induction of cultures as described above.

**Nucleotide sequence accession number.** The nucleotide sequence of the *bfpP* gene has been deposited in the GenBank data library under accession number Z34464.

## RESULTS

**Functional complementation of a *P. aeruginosa pilD* mutant by DNA fragments from the EPEC adherence factor plasmid.** Wild-type *P. aeruginosa* PAK is sensitive to the pilus-specific bacteriophage PO4, while nonpiliated strains, including *pilD* mutants, are resistant. This phenotype can be rescued by reintroducing the intact *P. aeruginosa pilD* gene into the mutant and was used to test for complementation by EPEC DNA.

The *bfpA*-containing plasmid pMSD217 is composed of two contiguous *Bam*HI DNA fragments cloned into pACYC184 (Fig. 1). Several restriction fragments from this 9.3-kb region were subcloned into the *E. coli-P. aeruginosa* shuttle vector

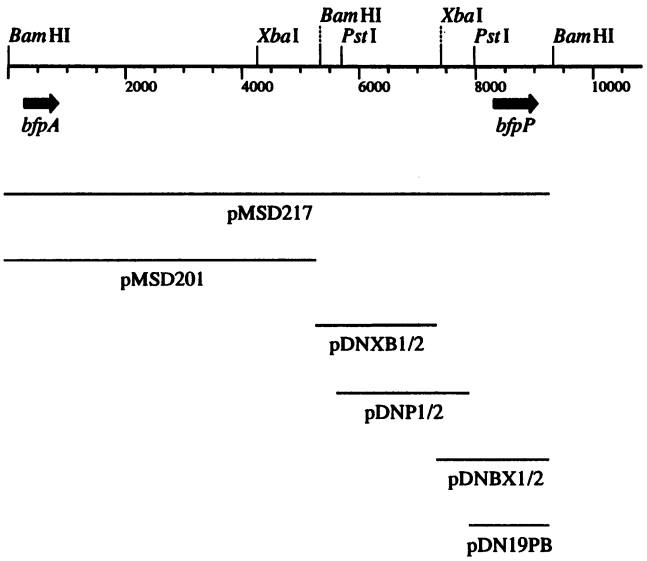


FIG. 1. Restriction map of the EPEC *bfpP* region and relevant subclones. The positions and orientations of the *bfpA* and *bfpP* loci are indicated (arrows).

pDN18 and mobilized into the *pilD* mutant by triparental mating. A 4-kb *Bam*HI DNA fragment was found to partially restore PO4 phage sensitivity to the mutant. Three subfragments covering different parts of the 4-kb *Bam*HI fragment were then cloned into pDN18 in both orientations and tested for their ability to complement the *pilD* mutant (pDNXB1/pDNXB2, pDNP1/pDNP2, and pDNBX1/pDNBX2) (Fig. 1). Of these, only pDNBX1 was able to restore PO4 phage sensitivity to the *pilD* mutant. Finally, a 1.3-kb *Pst*I-*Bam*HI subfragment in pDN19 (pDN19PB) was found to retain the ability to confer PO4 phage sensitivity to the *pilD* mutant. However, the same fragment cloned in the opposite orientation, relative to the *lac* promoter of the vector (pDN18BP), did not complement the *pilD* mutant. (Fig. 2).

**DNA sequence analysis of the *Pst*I-*Bam*HI fragment.** The nucleotide sequences of the ends of the *Pst*I-*Bam*HI fragment were obtained by using the T7 primer with pDN19PB and pDN18BP as templates. Restriction sites revealed by this sequence information were used to subclone restriction frag-

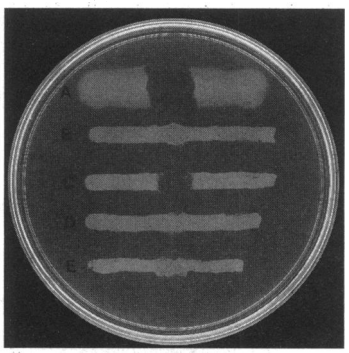

FIG. 2. PO4 phage sensitivity assay. Recombinant plasmids were tested for their ability to restore phage sensitivity to the *pilD* mutant. Sensitivity is defined as absence of bacterial growth at the site of phage inoculation. A, PAK(pDN18); B, PAK-DΩ(pDN18); C, PAK-DΩ(pDN19PB); D, PAK-DΩ(pDN18BP); E, PAK-DΩ(pDN19PBΔ).

```
gaaattgagcaacaacgaattatgttgagccatggaaaaaacctgattagctaaaatatt  60

ATGCAAGAAAGTATATTTCTATTATTGGTTTTTATATATGCGGCCACAATAACCAGTTTT  120
M  Q  E  S  I  F  L  L  L  V  F  I  Y  A  A  T  I  T  S  F
                                .AccI
ATATGGCTCGCTGTGGAACGTCTTCCCCATCAGTTAAAATGGGTAGACAATCCTGTTTCT  180
I  W  L  A  V  E  R  L  P  H  Q  L  K  W  V  D  N  P  V  S

GATATAACAATATGGTCTCCGGGAAGTAAATGTAATAATTGTGGTAAAAAAATTCGTTGG  240
D  I  T  I  W  S  P  G  S  K  C  N  N  C  G  K  K  I  R  W

TATTATCTTATTCCTGTTTTGGGGTATTTTTTGTGTCGGGGGGAATGTGGCTATTGTCAT  300
Y  Y  L  I  P  V  L  G  Y  F  L  C  R  G  E  C  G  Y  C  H

GCTAAAGTTCCCGTTCGGTATCCACTTACAGAGTTTATCTGTGGTGTTTGCAGTGTTATT  360
A  K  V  P  V  R  Y  P  L  T  E  F  I  C  G  V  C  S  V  I

ATATTTGTATTTTTGGGAGACAGACTGTATGATGCAGTAATAGTTTCGCTACTTTTTCTT  400
I  F  V  F  L  G  D  R  L  Y  D  A  V  I  V  S  L  L  F  L

TGCCTTGTTTTTCTTGCTCTGATTGATTTAAGAGAGAACTGGCTACCAGCTTGTGTTACG  460
C  L  V  F  L  A  L  I  D  L  R  E  N  W  L  P  A  C  V  T

TACCCTCTATTTTGGGCAGGTATGATTACTCCAGGGTTTGCATCCAGCGATGATAAGATT  520
Y  P  L  F  W  A  G  M  I  T  P  G  F  A  S  S  D  D  K  I
          .BsrFI                          .SpeI
TTCGGTGCATTTACCGGTTTTTTAATAATGTATATTTCAATGAAACTAGTTAGCGCCTTA  580
F  G  A  F  T  G  F  L  I  M  Y  I  S  M  K  L  V  S  A  L
                                      EagI. .BsrFI
AGAAAGGAGGATGTTTTTGCAGGCGGAGACATTGCTCTGGCTACAGCGGCCGGTGCATGG  640
R  K  E  D  V  F  A  G  G  D  I  A  L  A  T  A  A  G  A  W

TTAGGAATAGATAAAATGCCTTTTTTTCTTATTTTGTCTTCATTTATATTTATTCTTTAC  700
L  G  I  D  K  M  P  F  F  L  I  L  S  S  F  I  F  I  L  Y

TCTCTTCCAGCAAGGTTAAGAGGTCAGGTTTTTGTTCCTATGGGACCAGCTCTTTCCGCG  760
S  L  P  A  R  L  R  G  Q  V  F  V  P  M  G  P  A  L  S  A

TCATTTTTTATATGTTTAGTTTATCATTAAttacaggtgttgcagcatgtgttctcaatg  820
S  F  F  I  C  L  V  Y  H  Ter
```

FIG. 3. Nucleotide sequence of the EPEC *bfpP* coding region and flanking sequences. Amino acids deduced from the nucleotide sequences are specified by the standard one-letter abbreviation. A sequence similar to the proposed RBS of the *E. coli sdaA* gene is doubly underlined. Relevant restriction sites (*Acc*I and *Eag*I for subcloning for sequencing and *Bsr*FI for in-frame deletion construction) are indicated above the sequences.

ments into pBluescript KS or SK. The DNA sequences of both strands of the entire fragment were then determined. Figure 3 shows the sequence of the fragment, which reveals a single open reading frame (*bfpP*) of 747 bp. This reading frame has two potential in-frame initiation codons at positions 329 (ATG) and 401 (GTG) and a stop codon at position 1076. If translation starts from the first methionine, this open reading frame would encode a protein of 249 amino acids with a molecular mass of about 27.4 kDa. A search of the GenBank database with the TFASTA program (29) (Genetics Computer Group, University of Wisconsin) showed similarities with several members of the prepilin peptidase family, including PilD of *P. aeruginosa* (28% identity), PilD of *N. gonorrhoeae* (28%), TcpJ of *V. cholerae* (30%), and PulO of *K. oxytoca* (29%) (Fig. 4). Notably, BfpP contains near its center three CXXC motifs (common to members of the prepilin peptidase family) instead of two as in most other members of this family. However, only the first two are separated by the conserved distance of 21 amino acids, while the second two are separated by noncanonical 16 amino acids (Fig. 4).

**Identification of the polypeptide encoded by *bfpP*.** Attempts to detect the BfpP polypeptide either by an in vivo T7 expression system using *E. coli* BL21 (DE3) or by a commercially available in vitro transcription-translation kit were unsuccessful. Failure to visualize the BfpP protein could be due to the fact that *bfpP* lacks a consensus ribosome binding site (RBS) required for efficient translation. The closest match to a previously proposed RBS is the sequence TGATTAG at −10 with respect to the putative initiating methionine (Fig. 3) (40). PCR primers were used to amplify the *bfpP* gene and create an in-frame fusion to MBP in the commercially available expression vector pMALc-2. IPTG induction of synthesis of the fusion protein followed by immunoblot analysis using anti-MBP antiserum made it possible to visualize the MBP-BfpP polypeptide with the predicted $M_r$ of 69,000 (data not shown).

A 96-bp in-frame deletion in *bfpP* was created by excision of a *Bsr*FI restriction fragment. The *Pst*I-*Bam*HI fragment containing this deletion cloned in the pDN19 and mobilized into the *P. aeruginosa pilD* mutant failed to complement the mutant to restore PO4 phage sensitivity (Fig. 2). The *bfpP* gene containing this in-frame deletion was amplified by using primers Donne-62 and Donne-63, and the product was cloned into pMALc-2. As predicted, immunoblot revealed an MBP-ΔBfpP fusion with an $M_r$ of 66,000 (data not shown).

**A *bfpP*-containing DNA fragment can restore prebundlin processing to an EPEC *bfpA*::Tn*phoA* mutant.** EPEC 31-6-1(1) is a *bfpA*::Tn*phoA* mutant of E2348/69 that fails to express bundlin. When transformed with pMSD201 containing a 5.2-kb *Bam*HI fragment including *bfpA* and 4.5 kb of downstream DNA, proteins consistent with both prebundlin and bundlin are seen (9). As shown in Fig. 5, the same mutant transformed with pMSD217 expresses only fully processed bundlin (BfpA). Thus, the 4-kb *Bam*HI DNA, present in pMSD217 but not in pMSD201, contains *bfpP* and encodes the prepilin peptidase. This result is in agreement with that obtained from the in vivo complementation experiments with *Pseudomonas* spp.

**PilD is functionally homologous to BfpP.** The ability of *bfpP* to complement the *pilD* mutant provides evidence that PilD can be functionally substituted by BfpP. To demonstrate that this relationship is reciprocal, the *bfpA* gene was cloned into the shuttle expression vector pMMB67HE under the control of the *lac* promoter and mobilized into wild-type *P. aeruginosa* PAK and *pilD* mutant PAK-DΩ. Expression of bundlin was examined by immunoblot analysis (Fig. 6). In the wild-type strain PAK, bundlin was detected as the fully processed 18.5-kDa band. In the *pilD* mutant, however, the antibody recognized a larger protein consistent with failure of the mutant to process prebundlin. This indicates that the *P. aeruginosa* PilD protein is capable of using the EPEC prebundlin protein as a substrate.


FIG. 4. Comparison of the predicted products of *bfpP* and other members of the prepilin peptidase family in the region containing the conserved dithiol motif (shaded residues). Invariant amino acids are boxed.

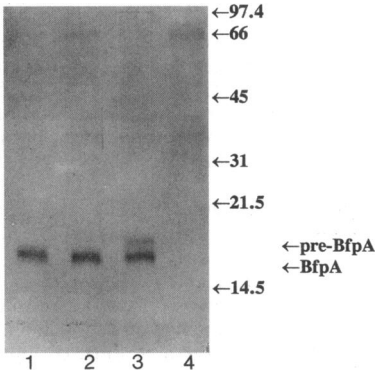


FIG. 5. Western analysis of expression of bundlin in the EPEC *bfpA*::Tn*phoA* mutant 31-6-1(1). Expression of bundlin (BfpA) was examined by Western blot following SDS-PAGE of whole-cell lysates from bacteria grown in Eagle's minimal essential medium with proper antibiotics. Lanes: 1, EPEC wild-type strain E2348/69; 2, 31-6-1(1)(pMSD217); 3, 31-6-1(1)(pMSD201); 4, 31-6-1(1). The sizes of molecular weight markers are indicated in kilodaltons.

## DISCUSSION

In this report we demonstrate the presence on the EPEC adherence plasmid of a prepilin peptidase gene with sequence and functional homologies to the *pilD* locus of *P. aeruginosa*. Our initial approach for cloning the EPEC prepilin peptidase gene was to use a probe derived from the *P. aeruginosa pilD* gene for hybridization to a DNA library of the EPEC adherence factor plasmid. A similar strategy was recently employed to identify the prepilin peptidase of *Pseudomonas putida* (7). However, the *pilD* probe failed to hybridize with the EPEC DNA, even under conditions of reduced stringency. Prepilin peptidase genes have been identified in other species by sequence analysis in the vicinity of pilin or pilin-like genes (11, 23, 25, 35). We therefore suspected that an EPEC prepilin peptidase gene might reside near the *bfpA* gene. We cloned the gene by functional complementation of a *P. aeruginosa* prepilin peptidase *pilD* mutant using DNA from the region downstream of *bfpA*. This is the first example of employing such a strategy for the cloning of a type IV fimbrial accessory gene, an approach that may prove fruitful for other components of the system and for other organisms.

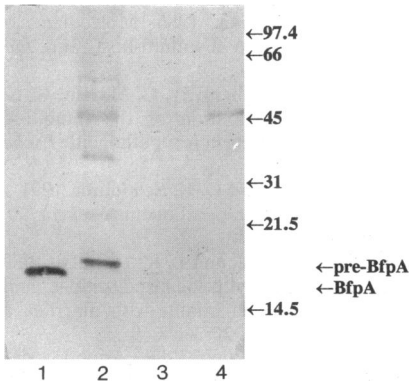


FIG. 6. Cleavage of EPEC prebundlin by PilD in vivo. The EPEC *bfpA* gene was cloned into the *E. coli-P. aeruginosa* shuttle vector pMMB67HE, resulting in pBFP6. Both pBFP6 and pMMB67HE were mobilized into the *P. aeruginosa* wild-type strain PAK and *pilD* mutant PAK-DΩ. Expression of bundlin was analyzed by Western blot. Lanes: 1, PAK(pBFP6); 2, PAK-DΩ(pBFP6); 3, PAK(pMMB67HE); 4, PAK-DΩ(pMMB67HE). The sizes of molecular weight markers are indicated in kilodaltons.

Since a partial prepilin peptidase-like sequence (*hopD*) had previously been identified in *E. coli* K-12 (19), it was not clear a priori that EPEC would require a specific prepilin peptidase gene to process prebundlin. Indeed, the partial processing of type IV prepilin and prepilin-like proteins, such as those from *N. gonorrhoeae*, *Aeromonas hydrophila*, and EPEC, observed upon expression in *E. coli* (9, 11, 20, 25) had been attributed to HopD (19). While in *P. aeruginosa* a single enzyme is responsible for processing the prepilin and components of a secretory apparatus (36), the existence of multiple prepilin peptidases in other organisms cannot be excluded.

In addition to cloning the *bfpP* gene in a *pilD* mutant of *P. aeruginosa*, we also showed that the bundlin protein of EPEC is fully processed when expressed in the wild-type but not *pilD* mutant *P. aeruginosa*, thereby demonstrating reciprocal functional homology between the prepilin peptidase enzymes of the two organisms. This result was also somewhat surprising since significant differences exist between the fimbriae of EPEC and *P. aeruginosa*. The type IV fimbria family has been classified into two groups on the basis of the amino acid sequence of pilin polypeptides and the first amino acid of the mature pilin subunit (44). The first group includes fimbriae from *P. aeruginosa*, *N. gonorrhoeae*, *N. meningitidis*, *M. bovis*, and *D. nodosus*. Prepilins in this group have very short (6 or 7 amino acids), basic, amino-terminal leader peptides. These peptides are cleaved by their respective prepilin peptidases between an invariant glycine and a phenylalanine. The amino acid sequences of the mature pilins are highly conserved in their amino-terminal domains. The second group currently has two members, the toxin-coregulated pilus of *V. cholerae* and BFP. While these two proteins are similar to each other and to members of the first group near their amino termini, they have longer leader peptides than those of the first group (TcpA [25 amino acids] and BfpA [13]). The first amino acid of the mature TcpA is methionine and that of BfpA is leucine. Only the former has been demonstrated to be N methylated. Previous reports have described the ability of prepilin peptidases from *K. oxytoca*, *B. subtilis*, and *P. aeruginosa* to process the prepilin of *N. gonorrhoeae* (11, 35). In addition, the gonococcal prepilin peptidase has been shown to process the PulG protein of *K. oxytoca* (10), and the enzyme from *P. aeruginosa* has been shown to process the ExeG protein of *A. hydrophila* (20). However, the TcpJ protein of *V. cholerae* is not able to process the prepilin of *P. aeruginosa* when both are cloned in *E. coli* (22). The functional conservation between PilD and BfpP demonstrated in this study provides evidence that the differences between the two groups of type IV fimbriae are not sufficient to preclude proper heterologous function. While complete assembly of fimbriae from pilin subunits of *N. gonorrhoeae*, *M. bovis*, and *D. nodosus* has been accomplished using the *P. aeruginosa* biogenesis functions, similar heterologous assembly of pilin subunits from group 2 (BFP or toxin-coregulated pilus) by the assembly machinery of group 1 has not been reported as yet. The results reported here suggest that such experiments are feasible.

The deduced amino acid sequence of BfpP indicates that it is the most distant relative in the prepilin peptidase family. Among the notable features of the EPEC enzyme is the presence of three, rather than the usual two, CXXC motifs. It would be interesting to know whether each of these cysteine pairs contributes to its activity. Another implication from the sequence is the fact that the ability to process heterologous substrates is not dependent on a high degree of sequence

conservation. Furthermore, our results show that the absence of signal with DNA probes does not exclude the existence of a fully functional homolog. Presuming that the extant prepilin peptidases evolved from a common ancestor, the regions of relatively high conservation among sequences of low similarity may provide clues as to the active site of the enzyme.

One of the more interesting properties of the *P. aeruginosa* PilD protein is its bifunctionality. The prepilin peptidase not only cleaves the leader peptide but also carries out a subsequent N methylation of the newly created N-terminal phenylalanine (46). Since many of the type IV fimbriae examined contain N-methylated amino termini and substitutions of the N-terminal phenylalanine by methionine, alanine, or tyrosine do not alter prepilin cleavage or N methylation (45), it is expected that the N-terminal leucine of bundlin is methylated as well. The fact that the amino acid sequence of bundlin indicates an N-terminal methionine (16) while the DNA sequence of the *bfpA* gene specifies leucine (9, 43) invites speculation that *N*-methyl leucine may not be resolved from methionine in amino acid sequence determinations.

We failed to detect the BfpP polypeptide either by T7 expression in BL21 (DE3) or by in vitro transcription-translation reaction. We suspect that failure to detect the protein may be due to inefficient translation of the *bfpA* mRNA because of a poor RBS. We have found only one similar sequence proposed to represent an RBS, that of the *sdaA* gene (40). A second possible initiation codon GTG (valine), present within the open reading frame, is also preceded by a poor potential RBS (TGGC). Other, less likely explanations for failure to detect the mature BfpP include rapid turnover of the BfpP protein and the short half-life of the mRNA. However, we were able to verify that the *bfpP* gene could encode a protein of the expected size by expressing an MBP-BfpP fusion protein. We obtained further evidence that the protein identified by immunoblotting with anti-MBP antiserum was the predicted fusion protein by identifying a similar fusion of slightly faster mobility from a clone containing an in-frame deletion within *bfpP*.

The *bfpP* gene was cloned from a 9.3-kb EPEC plasmid DNA fragment that also contains the *bfpA* gene. It is possible that these genes are part of an operon encoding the EPEC type IV fimbria biosynthetic machinery. One known type IV system, the Tcp gene cluster, contains more than 11 genes, with at least 9 of them belonging to the same transcriptional unit (37). The fact that the *Pst*I-*Bam*HI fragment can complement the *pilD* mutant in only one orientation suggests that expression of *bfpP* depends on a promoter from the vector (*lacZ* promoter in pDN19). This suggests that *bfpP* might be transcribed along with other upstream genes by the same promoter, possibly the one upstream of *bfpA*. Analysis of the DNA sequence upstream of *bfpP* reveals the existence of several open reading frames which are homologous with proteins involved in biogenesis of type IV fimbriae in other bacterial species (49). It remains to be determined whether additional loci of the EPEC *bfp* gene cluster encode proteins functionally homologous to those of other members of the type IV fimbria family.

## ACKNOWLEDGMENTS

We thank J. Albert for the mouse monoclonal antibody against BfpA and L. K. Carlson for performing the PilD probe hybridizations.

This study was supported by the Office of Research and Development, Medical Research Service, Department of Veterans' Affairs, by Public Health Service award AI32074 from the National Institutes of Health, and by a Designated Research Initiative Fund award from the University of Maryland. S. Lory was supported by NIH grant AI21451. Computer time was supported in full by the Maryland Biotechnology Institute.

## REFERENCES

1. **Baldini, M. M., J. B. Kaper, M. M. Levine, D. C. Candy, and H. W. Moon.** 1983. Plasmid-mediated adhesion in enteropathogenic *Escherichia coli*. J. Pediatr. Gastroenterol. Nutr. **2**:534–538.
2. **Beard, M. K., J. S. Mattick, L. J. Moore, M. R. Mott, C. F. Marrs, and J. R. Egerton.** 1990. Morphogenetic expression of *Moraxella bovis* fimbriae (pili) in *Pseudomonas aeruginosa*. J. Bacteriol. **172:** 2601–2607.
3. **Chen, E. Y., and P. H. Seeburg.** 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4**:165–170.
4. **Chikami, G. K., J. Fierer, and D. G. Guiney.** 1985. Plasmid-mediated virulence in *Salmonella dublin* demonstrated by use of a Tn*5-oriT* construct. Infect. Immun. **50**:420–424.
5. **Cravioto, A., A. Tello, A. Navarro, J. Ruiz, H. Villafán, F. Uribe, and C. Eslava.** 1991. Association of *Escherichia coli* HEp-2 adherence patterns with type and duration of diarrhoea. Lancet **337**:262–264.
6. **De Groot, A., A. Filloux, and J. Tommassen.** 1991. Conservation of *xcp* genes, involved in the two-step protein secretion process, in different *Pseudomonas* species and other gram-negative bacteria. Mol. Gen. Genet. **229**:278–284.
7. **De Groot, A., I. Heijnen, H. De Cock, A. Filloux, and J. Tommassen.** 1994. Characterization of type IV pilus genes in plant growth-promoting *Pseudomonas putida* WCS358. J. Bacteriol. **176:** 642–650.
8. **Donnenberg, M. S., S. B. Calderwood, A. Donohue-Rolfe, G. T. Keusch, and J. B. Kaper.** 1990. Construction and analysis of Tn*phoA* mutants of enteropathogenic *Escherichia coli* unable to invade HEp-2 cells. Infect. Immun. **58**:1565–1571.
9. **Donnenberg, M. S., J. A. Girón, J. P. Nataro, and J. B. Kaper.** 1992. A plasmid-encoded type IV fimbrial gene of enteropathogenic *Escherichia coli* associated with localized adherence. Mol. Microbiol. **6**:3427–3437.
10. **Dupuy, B., and A. P. Pugsley.** 1994. Type IV prepilin peptidase gene of *Neisseria gonorrhoeae* MS11: presence of a related gene in other piliated and nonpiliated *Neisseria* strains. J. Bacteriol. **176:** 1323–1331.
11. **Dupuy, B., M.-K. Taha, O. Possot, C. Marchal, and A. P. Pugsley.** 1992. PulO, a component of the pullulanase secretion pathway of *Klebsiella oxytoca*, correctly and efficiently processes gonococcal type IV prepilin in *Escherichia coli*. Mol. Microbiol. **6**:1887–1894.
12. **Dupuy, B., M. K. Taha, A. P. Pugsley, and C. Marchal.** 1991. *Neisseria gonorrhoeae* prepilin export studied in *Escherichia coli*. J. Bacteriol. **173**:7589–7598.
13. **Frost, L. S., M. Carpenter, and W. Paranchych.** 1978. N-methylphenylalanine at the N-terminus of pilin isolated from *Pseudomonas aeruginosa* K. Nature (London) **271**:87–89.
14. **Fürste, J. P., W. Pansegrau, R. Frank, H. Blocker, P. Scholz, M. Bagdasarian, and E. Lanka.** 1986. Molecular cloning of the plasmid RP4 primase region in a multi-host-range *tacP* expression vector. Gene **48**:119–131.
15. **Girón, J. A., M. S. Donnenberg, W. C. Martin, K. G. Jarvis, and J. B. Kaper.** 1993. Distribution of the bundle-forming pilus structural gene (*bfpA*) among enteropathogenic *Escherichia coli*. J. Infect. Dis. **168**:1037–1041.
16. **Girón, J. A., A. S. Y. Ho, and G. K. Schoolnik.** 1991. An inducible bundle-forming pilus of enteropathogenic *Escherichia coli*. Science **254**:710–713.
17. **Gomes, T. A. T., P. A. Blake, and L. R. Trabulsi.** 1989. Prevalence of *Escherichia coli* strains with localized, diffuse, and aggregative adherence to HeLa cells in infants with diarrhea and matched controls. J. Clin. Microbiol. **27**:266–269.
18. **Harlow, D. H., and D. Lane.** 1988. Antibodies, a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
19. **Hobbs, M., and J. S. Mattick.** 1993. Common components in the assembly of type 4 fimbriae, DNA transfer systems, filamentous phage and protein-secretion apparatus: a general system for the formation of surface-associated protein complexes. Mol. Microbiol. **10**:233–243.
20. **Howard, S. P., J. Critch, and A. Bedi.** 1993. Isolation and analysis of eight *exe* genes and their involvement in extracellular protein

secretion and outer membrane assembly in *Aeromonas hydrophila*. J. Bacteriol. **175**:6695–6703.
21. **Hoyne, P. A., R. Haas, T. F. Meyer, J. K. Davies, and T. C. Elleman.** 1992. Production of *Neisseria gonorrhoeae* pili (fimbriae) in *Pseudomonas aeruginosa*. J. Bacteriol. **174**:7321–7327.
22. **Kaufman, M. R.** 1993. Ph.D. thesis. University of Tennessee, Memphis.
23. **Kaufman, M. R., J. M. Seyer, and R. K. Taylor.** 1991. Processing of TCP pilin by TcpJ typifies a common step intrinsic to a newly recognized pathway of extracellular protein secretion by gram-negative bacteria. Genes Dev. **5**:1834–1846.
24. **Koga, T., K. Ishimoto, and S. Lory.** 1993. Genetic and functional characterization of the gene cluster specifying expression of *Pseudomonas aeruginosa* pili. Infect. Immun. **61**:1371–1377.
25. **Lauer, P., N. H. Albertson, and M. Koomey.** 1993. Conservation of genes encoding components of a type IV pilus assembly/two-step protein export pathway in *Neisseria gonorrhoeae*. Mol. Microbiol. **8**:357–368.
26. **Levine, M. M.** 1987. *Escherichia coli* that cause diarrhea: enterotoxigenic, enteropathogenic, enteroinvasive, enterohemorrhagic, and enteroadherent. J. Infect. Dis. **155**:377–389.
27. **Levine, M. M., E. J. Bergquist, D. R. Nalin, D. H. Waterman, R. B. Hornick, C. R. Young, S. Sotman, and B. Rowe.** 1978. *Escherichia coli* strains that cause diarrhoea but do not produce heat-labile or heat-stable enterotoxins and are non-invasive. Lancet **i**:1119–1122.
28. **Levine, M. M., J. P. Nataro, H. Karch, M. M. Baldini, J. B. Kaper, R. E. Black, M. L. Clements, and A. D. O'Brien.** 1985. The diarrheal response of humans to some classic serotypes of enteropathogenic *Escherichia coli* is dependent on a plasmid encoding an enteroadhesiveness factor. J. Infect. Dis. **152**:550–559.
29. **Lipman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227**:1435–1440.
30. **Lory, S.** 1992. Determinants of extracellular protein secretion in gram-negative bacteria. J. Bacteriol. **174**:3423–3428.
31. **Mattick, J. S., M. M. Bills, B. J. Anderson, B. Dalrymple, M. R. Mott, and J. R. Egerton.** 1987. Morphogenetic expression of *Bacteroides nodosus* fimbriae in *Pseudomonas aeruginosa*. J. Bacteriol. **169**:33–41.
32. **Nataro, J. P., K. O. Maher, P. Mackie, and J. B. Kaper.** 1987. Characterization of plasmids encoding the adherence factor of enteropathogenic *Escherichia coli*. Infect. Immun. **55**:2370–2377.
33. **Nataro, J. P., I. C. A. Scaletsky, J. B. Kaper, M. M. Levine, and L. R. Trabulsi.** 1985. Plasmid-mediated factors conferring diffuse and localized adherence of enteropathogenic *Escherichia coli*. Infect. Immun. **48**:378–383.
34. **Nunn, D., S. Bergman, and S. Lory.** 1990. Products of three accessory genes, pilB, pilC, and pilD, are required for biogenesis of *Pseudomonas aeruginosa* pili. J. Bacteriol. **172**:2911–2919.
35. **Nunn, D. N., and S. Lory.** 1991. Product of the *Pseudomonas aeruginosa* gene pilD is a prepilin leader peptidase. Proc. Natl. Acad. Sci. USA **88**:3281–3285.
36. **Nunn, D. N., and S. Lory.** 1992. Components of the protein-excretion apparatus of *Pseudomonas aeruginosa* are processed by the type IV prepilin peptidase. Proc. Natl. Acad. Sci. USA **89**:47–51.
37. **Ogierman, M. A., S. Zabihi, L. Mourtzios, and P. A. Manning.** 1993. Genetic organization and sequence of the promoter-distal region of the *tcp* gene cluster of *Vibrio cholerae*. Gene **126**:51–60.
38. **Paranchych, W., and L. S. Frost.** 1988. The physiology and biochemistry of pili. Adv. Microb. Physiol. **29**:53–114.
39. **Pugsley, A. P., and B. Dupuy.** 1992. An enzyme with type IV prepilin peptidase activity is required to process components of the general extracellular protein secretion pathway of *Klebsiella oxytoca*. Mol. Microbiol. **6**:751–760.
40. **Rudd, K. E., and T. D. Schneider.** 1992. Compilation of *E. coli* ribosome binding sites, p. 17.19–17.45. *In* J. H. Miller (ed.), A short course in bacterial genetics: a laboratory manual and handbook for *Escherichia coli* and related organisms. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
41. **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.
42. **Scaletsky, I. C. A., M. L. M. Silva, and L. R. Trabulsi.** 1984. Distinctive patterns of adherence of enteropathogenic *Escherichia coli* to HeLa cells. Infect. Immun. **45**:534–536.
43. **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.
44. **Strom, M. S., and S. Lory.** 1993. Structure-function and biogenesis of the type IV pili. Annu. Rev. Microbiol. **47**:565–596.
45. **Strom, M. S., D. Nunn, and S. Lory.** 1991. Multiple roles of the pilus biogenesis protein PilD: involvement of PilD in excretion of enzymes from *Pseudomonas aeruginosa*. J. Bacteriol. **173**:1175–1180.
46. **Strom, M. S., D. N. Nunn, and S. Lory.** 1993. A single bifunctional enzyme, PilD, catalyzes cleavage and N-methylation of proteins belonging to the type IV pilin family. Proc. Natl. Acad. Sci. USA **90**:2404–2408.
47. **Studier, F. W., and B. A. Moffatt.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189**:113–130.
48. **Whitchurch, C. B., M. Hobbs, S. P. Livingston, V. Krishnapillai, and J. S. Mattick.** 1991. Characterisation of a *Pseudomonas aeruginosa* twitching motility gene and evidence for a specialised protein export system widespread in eubacteria. Gene **101**:33–44.
49. **Zhang, H.-Z., and M. S. Donnenberg.** Unpublished data.