JOURNAL OF BACTERIOLOGY, Nov. 1996, p. 6564–6570 Vol. 178, No. 22
0021-9193/96/$04.00+0
Copyright © 1996, American Society for Microbiology

Duplication of Pilus Gene Complexes of *Haemophilus influenzae* Biogroup aegyptius

TIMOTHY D. READ, MELANIE DOWDELL,† SARAH W. SATOLA, AND MONICA M. FARLEY*

Veterans Affairs Medical Center and Department of Medicine, Emory University School of Medicine, Decatur, Georgia 30033

Received 11 March 1996/Accepted 13 September 1996

Brazilian purpuric fever (BPF) is a recently described pediatric septicemia caused by a strain of *Haemophilus influenzae* biogroup aegyptius. The pilus specified by this bacterium may be important in BPF pathogenesis, enhancing attachment to host tissue. Here, we report the cloning of two *haf* (for *H. influenzae* biogroup aegyptius fimbriae) gene clusters from a cosmid library of strain F3031. We sequenced a 6.8-kb segment of the *haf1* cluster and identified five genes (*hafA* to *hafE*). The predicted protein products, HafA to HafD, are 72, 95, 98, and 90% similar, respectively, to HifA to HifD of the closely related *H. influenzae* type b pilus. Strikingly, the putative pilus adhesion, HifE, shares only 44% identity with HafE, suggesting that the proteins may differ in receptor specificity. Insertion of a mini-γδ transposon in the *hafE* gene eliminated hemadsorption. The nucleotide sequences of the *haf1* and *haf2* clusters are more than 99% identical. Using the recently published sequence of the *H. influenzae* Rd genome, we determined that the *haf1* complex lies at a unique position in the chromosome between the *pmbA* gene and a hypothetical open reading frame, HI1153. The location of the *haf2* cluster, inserted between the *purE* and *pepN* genes, is analogous to the *hif* genes on *H. influenzae* type b. BPF fimbrial phase switching appears to involve slip-strand mispairing of repeated dinucleotides in the pilus promoter. The BPF-associated *H. influenzae* biogroup aegyptius pilus system generally resembles other *H. influenzae*, but the possession of a second fimbrial gene cluster, which appears to have arisen by a recent duplication event, and the novel sequence of the HafE adhesin may be significant in the unusual pathogenesis of BPF.

Haemophilis influenzae biogroup aegyptius (Hae, formerly *H. aegyptius*) is a nontypeable (nonencapsulated) gram-negative bacterium that is an important cause of contagious purulent conjunctivitis. In 1984, Hae was found to be the etiologic agent of Brazilian purpuric fever (BPF), a septicemic illness occurring in young children (3). The disease usually follows recovery from purulent conjunctivitis and is associated with acute fever, the development of petechiae, purpura fulminans, and vascular collapse. Molecular epidemiological studies suggested that a single clone of Hae was responsible for nearly all BPF cases identified in Brazil (3). This BPF-associated strain has been distinguished from other Hae strains by its resistance to killing by human serum (26), toxicity for human endothelial cells (27, 44), and ability to cause bacteremia in an infant rat model (30). The bacterium is also capable of invading cells of a primary human adenoid tissue culture (7). Weyant et al. (43) reported that BPF-associated clone strains can express an antigenically distinct pilus (fimbrium) that mediates hemagglutination and is expressed in a phase-variable manner. In contrast to the pili of closely related *H. influenzae* type b (Hib), pili of BPF-associated *H. influenzae* strains are generally expressed at the time of isolation from systemic sites (23), suggesting that the pilus may play a novel role in BPF pathogenesis.

Pilus-mediated adhesion is a well-described mechanism for bacterial attachment to epithelial surfaces. Pili are surface appendages of gram-negative bacteria that are usually composed of repeated major pilin subunits and one or more minor subunits. The best-studied example is Pap (pyelonephritis-associated pilus) of uropathogenic *Escherichia coli* (16). Major pilus gene sequences of a number of *H. influenzae* strains have also been reported (10, 12, 17, 19, 37). The *H. influenzae* genes are each present in a single copy, are very similar to each other (78 to 100% identity at the amino acid level), and show some resemblance to *E. coli* type 1 and Pap pili. A pilus gene cluster comprising five genes has now been characterized in Hib (24, 38, 42). The cluster encodes HifA, a major pilin protein; HifB, a periplasmic chaperone; HifC, an outer membrane usher; and two minor subunits, HifD and HifE. Phase variation of pilus gene expression is mediated by slip-strand mispairing altering the spacing between the overlapping, divergently oriented −10 and −35 promoters of *hifA* and *hifB* (39). The major pilus gene of BPF-associated *H. influenzae* has also been characterized and shown to encode a pilin with distinctive antigenic features (34, 45).

In this work, we report on the *haf* (for *H. influenzae* biogroup aegyptius fimbrium) genes of the prototypic BPF-associated strain F3031. The BPF system differs significantly from that of Hib because there are two identical, independently phase-variable copies of the pilus gene cluster and also because the final gene of the complex encodes a novel version of the putative pilus adhesin protein.

* Corresponding author. Mailing address: Atlanta VA Medical Center, Medical Research Service 151, 1670 Clairmont Rd., Decatur, GA 30033. Phone: (404) 728-7688. Fax: (404) 329-2210. Electronic mail address: mfarley@emory.edu.

† Present address: PERL Laboratory, Decatur, GA 30030.

MATERIALS AND METHODS

Bacteria. BPF case strains from Brazil, F1946, F2052, F3031, F3034, F3049, F3114, F3224 (3) and G1097, and strain 34b, a piliated phase variant of F3049, were provided kindly by the Meningitis and Special Pathogens Branch, Centers for Disease Control and Prevention, Atlanta, Ga. Hib strain 1007 has been described previously (6). Hae and Hib strains were grown at 37°C on chocolate agar supplemented with IsoVitaleX and hemoglobin in a 4% CO_2 atmosphere or in brain heart infusion broth supplemented with Fildes enrichment (5%, vol/vol) (Becton Dickinson, Cockeysville, Md.).

E. coli HB101 (2) and XL-1 Blue MRA (Stratagene Cloning Systems, La Jolla, Calif.) were grown in Luria broth or on Luria agar plates (Becton Dickinson). Ampicillin was added to a final concentration of 50 μg/ml when necessary for selection.

Cloning and sequencing the Haf pilus gene clusters. A segment of the *haf1*

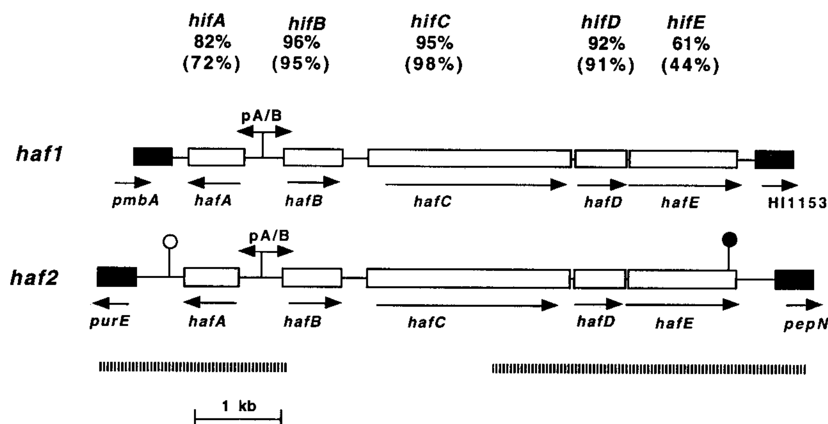

FIG. 1. Map of the *haf1* and *haf2* genes of strain F3031. Genes flanking the clusters are shown as black rectangles. pA/B is the predicted bidirectional promoter. The complete nucleotide sequence of the *hafB* and *hafC* genes on *haf2* has not been determined. The percent identity of each gene with its counterpart in the *hif* cluster (38) is shown at the top; the figure in parentheses following each identity percentage is the similarity of each predicted amino acid sequence. The closed circle represents mini-γδ transposon insertion D6 in the *hafE* gene of the *haf2* cluster; the open circle represents insertion F5. The stippled lines at the bottom of the figure depict the extent of the *haf2* cluster sequenced.

gene cluster was originally cloned in plasmid pAW30 (45). We determined that BPF-associated strain F3031 carried two *haf* pilus gene clusters when *Bgl*II chromosomal digests probed with a fragment of the major pilus gene demonstrated two hybridizing bands of approximately 12 and 25 kb. In order to isolate the individual *haf* clusters, a cosmid library of F3031 was then prepared. *Bam*HI-digested pHC74 cosmid DNA (4) was ligated with a partial *Sau*3A1 digest of F3031 chromosomal DNA, packaged with the GigaPak Gold II extract (Stratagene), and used to transfect XL-1 Blue MCR cells. pMF1 and pMF6, two recombinant plasmids that hybridized with a major pilus gene probe, were shown by restriction enzyme mapping to contain different regions of the *H. influenzae* chromosome. pMF1 carries the *haf1* genes, and pMF6 carries *haf2*.

Sequencing of pAW30 and the pMF1 and pMF6 cosmids was accomplished by a combination of the dideoxy chain termination method of Sanger et al. (32) and double-stranded DNA Cycle sequencing (Bethesda Research Laboratories), using the Sequenase 2.0 enzyme (U.S. Biochemicals, Cleveland, Ohio). A set of primers was synthesized to allow both strands of the DNA to be sequenced. DNA sequence analysis was performed with Geneworks software (IntelliGenetics, Mountain View, Calif.). Nucleotide and amino acid alignments were computed by a method based on that of Smith and Waterman (33).

Hemadsorption assays. Strains were evaluated for the ability to adsorb type O-positive human erythrocytes in a 5% suspension by using a nitrocellulose overlay technique described previously (5, 6).

Determination of hemagglutination titer. The hemagglutination proficiency of *H. influenzae* strains was determined by a procedure based on that of St. Geme et al. (35). Bacteria were picked as single colonies from stock plates and spread on chocolate agar plates. After overnight incubation, cells were resuspended in phosphate-buffered saline (PBS) (31) and standardized to an A_{600} of 1.0. Serial twofold dilutions of the suspension with PBS were made in a Costar 96-well V-bottom microtiter plate. An equal volume (50 μl) of heparinized human O-positive erythrocytes, resuspended to 1.0% (vol/vol) in PBS, was added to each well. After incubating the plate for 60 min at room temperature, the highest dilutions of bacteria giving visible hemagglutination were scored. Each experiment was repeated at least three times.

DNA extractions. *H. influenzae* chromosomal DNA was prepared by the method of Moxon et al. (25). Cosmid DNA was isolated using protocols and reagents of Qiagen (Chatsworth, Calif.). Restriction enzymes were purchased from New England Biolabs, Inc., Beverly, Mass., and Bethesda Research Laboratories, Inc., Gaithersburg, Md.

PCR. Oligonucleotide DNA primers were synthesized by the Emory University Microchemical Facility, Atlanta, Ga. Reagents were purchased from Perkin-Elmer Cetus Corporation (Norwalk, Conn.). PCR reactions were performed for 30 cycles consisting of a 1-min denaturation step at 94°C followed by 1 min of annealing at 55°C and a 1-min extension at 72°C.

Mini-γδ transposon mutagenesis of the *haf2* gene cluster. An 11-kb *Bgl*II fragment of pMF6 containing the *haf2* gene cluster was inserted into the *Bgl*II polylinker site of the kanamycin resistance vector pOK12 (40), producing plasmid pMF20. The 11-kb *haf2* fragment was excised from the polylinker of pMF20 by *Xho*I and *Bam*HI digestion and ligated to *Bam*HI and *Sal*I-digested ampicillin resistance vector pNEB193 (New England Biolabs), creating pMF21. Random mini-γδ mutagenesis of pMF21 was performed by the method of Berg et al. (1). Kanamycin-resistant hemadsorption-deficient transposon mutants of pMF21 were identified by the protocol outlined earlier (see Materials and Methods). The positions of the mini-γδ insertions were determined by restriction digestion of the plasmids and sequencing with the mγδ-1 and mγδ-2 primers (1).

Nucleotide sequence accession numbers. The GenBank accession number of the sequence of the *haf1* genes is U54780. The *hafE2* gene has also been deposited as U58657.

RESULTS

Nucleotide sequence of the Haf pilus gene clusters. A cosmid library from BPF-associated strain F3031 was screened for sequences that hybridized to a probe from the pilus major subunit gene cloned in plasmid pAW30 (see Materials and Methods) (45). We sequenced both strands of a 6,457-bp segment from one cosmid, pMF1, containing a pilus gene cluster (designated *haf1*) homologous to that encoding the LKP pilus of Hib (38). Five putative genes, *hafA* to *hafE*, were identified by their similarity to the *hif* genes (Fig. 1). We found that another cosmid, pMF6, carried a second set of pilus genes (*haf2*), based on differences in the restriction pattern of the chromosomal region flanking the cluster. Three *haf2* open reading frames (ORFs) were sequenced (*hafA*, *hafD*, and *hafE*) and found to be more than 99% similar to the corresponding *haf1* DNA. There were only two disparities between the *haf* clusters. First, there were 10 TA dinucleotide repeats between the overlapping −10 and −35 consensus sequences of the *haf2* promoter but only 9 such repeats between the consensus sequences of the *haf1* promoters (Fig. 2). Also, nucleotides 1225 and 1226 of *hafE* are T and C on *haf2* rather than C and A as in *haf1*, with the consequence that residue 409 of HafE of *haf1* was a serine rather than a glutamine.

The general organization of the *haf* complex is similar to that of the *hif* genes (Fig. 1). The first gene, *hafA*, encodes a protein with a molecular mass of 20.5 kDa that has been shown to be the major subunit of the pilus produced by BPF-associated strains (34, 45). Within the 463-bp intergenic region between the *hafA* and *hafB* ORFs are two sets of overlapping, divergently aligned −10 and −35 consensus promoter sequences separated by a string of TA dinucleotide repeats (Fig. 2). van Ham et al. (39) showed in the Hif system that similarly arranged promoters are responsible for transcription of *hifA* and the *hifBCDE* operon. Upstream of the translational start of *hafB* there are two imperfect direct repeats of a 44-bp sequence (marked REP in Fig. 2) composed partly of a 23-bp sequence element (5′ GTA GGG TGG GCG TYY GCC CAC MT) found several times on both strands of *H. influenzae* pilus clusters and also found at intergenic positions within the

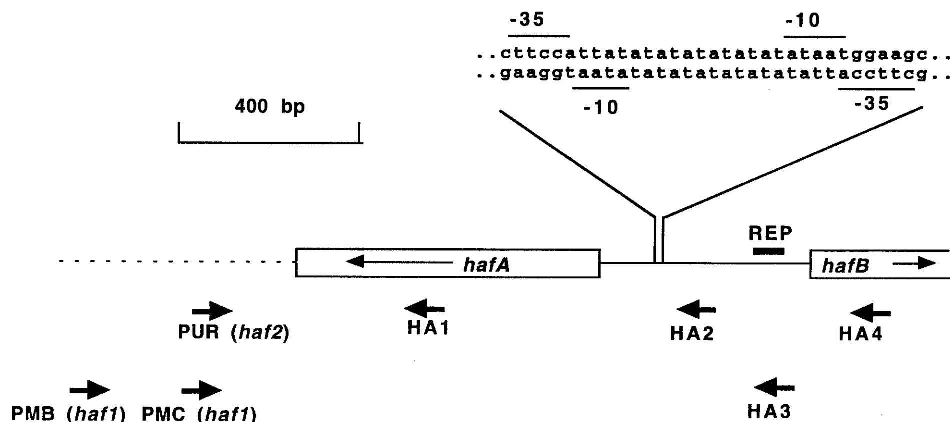

FIG. 2. PCR primer map of the *hafA-hafB* region. The PCR oligonucleotides used in this study are depicted as thick arrows underneath their approximate positions of binding in the *hafA-hafB* region. The arrows indicate the direction of DNA elongation from each primer. The primer PUR is specific for sequences downstream of *haf2*; PMB and PMC are specific for DNA flanking the *haf1* cluster. The nucleotide sequences of the primers are as follows: HA1, 5′ GGT AAG GTT GTT GAG AAT ACT TG; HA2, 5′ GCT AAA GCC AAC CTA CGG; HA3, 5′ GCT TTA GCC TCC ACG TGG TTA C; HA4, 5′ CAC TCT GGT GCC AGT GAT AAT C; PMB, 5′ GGG CGA ATT GCA AGA TAT GTT G; PMC, 5′ GCG CAA GAA GAC AAA AAT CAC C; and PUR, 5′ CGT GGC GTA GAA TAG CAT AG. The sequence and −10 and −35 consensus sites of the bidirectional *hafA-hafB* promoter of *haf1* are also shown. There are 9 iterated TA repeats in the *haf1* cluster of F3031 and 10 on the *haf2* cluster. The position of the two imperfect 44-bp direct repeats is depicted by the short bar (REP).

H. influenzae Rd chromosome (8, 27a). In comparison, at the same position in the *hif* cluster, the 44-bp sequence is repeated directly 10 times (38).

The next four ORFs are on the DNA strand opposite *hafA* (Fig. 1). The putative *hafB* gene encodes a protein of 232 amino acids that is homologous to the HifB periplasmic chaperone. The putative HafB protein contains all 13 invariant and 23 highly conserved residues of the consensus sequence for the PapD family of chaperones (48). There is only one possible translational start site for the HafB protein, whereas there are two for HifB, separated by three codons. However, the predicted signal peptidase cleavage sites, calculated by the algorithm of von Heijne (41), are identical for both proteins, giving HafB a 19-amino-acid signal sequence. After a 177-bp intergenic region, the next ORF is *hafC*, the first codon of which is for valine (GTG). The putative HafC protein has 98% identity with the *H. influenzae* pilus outer membrane usher, HifC (42). The next predicted protein, HafD (23.3 kDa), like the corresponding minor pilus subunit protein HifD, has a TASC motif for lipoprotein-specific protease between residues 17 and 20 (38). Although more than 90% identical, the putative HafD has an insertion of the amino acid sequence SNYSS between residues 22 and 23 of HifD.

Surprisingly, given the similarity of the first four genes to their counterparts in the Hib pilus cluster, *hafE* is markedly different from *hifE*. There is only 44% amino acid identity when the predicted proteins are aligned (Fig. 3). Much of the similarity is at the C termini of the proteins, a region where PapD-like chaperones are proposed to bind to pilus structural components (18, 47). Although the N terminus of HafE (48.1 kDa) is particularly divergent from HifE, computation of the preferred signal peptidase cleavage site by the von Heijne algorithm suggests that both proteins have signal peptides of about 30 amino acids (Fig. 3). Neither the *hafE* gene nor the predicted protein encoded by it have strong sequence similarity to sequences in the GenBank database (excluding *H. influenzae* pilus clusters).

Chromosomal location of the pilus gene clusters. The *hif* gene cluster is inserted in the intergenic region between *H. influenzae* chromosomal genes homologous to *purE* (36) and *pepN* (9) of *E. coli* (38, 42). In order to determine the location of the *haf* genes, we compared the sequence of about 500 bp of DNA flanking each cluster to the recently published nucleotide sequence of the *H. influenzae* Rd genome (8) by using the GRASTA search program, which has been made available on-line by The Institute for Genomic Research. The search showed that the *haf1* genes have been inserted at a novel site in the *H. influenzae* genome, between a gene 78% similar to *pmbA* of *E. coli* (29), which encodes a product believed to be involved in maturation of microcin B17, and a hypothetical ORF, HI1153 (Fig. 1). There are 32-bp direct repeats of genomic coordinates 1220363 to 1220394 flanking the *haf1* cluster. The *haf2* gene complex is flanked by DNA with greater than 95% similarity to the intergenic region between *purE* and *pepN* on the chromosome, with direct 59-bp repeats, corresponding to coordinates 1682304 to 1682362, downstream of *hafA* and *hafE*. Because of its location and association with identical direct repeats, the *haf2* cluster of BPF-associated *H. influenzae* appears to be the homolog of the *hif* complex in Hib.

Phase variation of pilus gene expression. We noted earlier in Results that the overlapping *hafA* and *hafBCDE* promoters are separated by a string of TA dinucleotide repeats (Fig. 2). A similar arrangement of the promoters is responsible for reversible phase variation in expression of the Hib LKP pilus (39). Alterations in the number of TA units, mediated by slip-strand mispairing, change the spacing between the −10 and −35 sequences and affect the efficiency of bidirectional-transcription initiation. With 9 TA units, little or no pilus production results, but a spacing of 10 units allows fimbrial expression. Expression of the cloned *haf* clusters in *E. coli*, with hemadsorption being used as a marker for pilus production, appears to follow the same rule. We evaluated pilus expression in cosmids pMF1 and pMF6 in an *E. coli* HB101 background by using colony hemadsorption. In a control experiment, colonies of HB101 carrying the vector pHC74 did not adsorb human O-positive erythrocytes (HA^- phenotype). pMF1, carrying the *haf1* cluster, has 9 TA units between the divergent −10 and −35 sequences and conferred an HA^- phenotype, whereas pMF6, carrying the *haf2* genes, has 10 TA units and conferred hemadsorption proficiency.

In an investigation of phase-variable Haf pilus expression in vivo, we determined the *haf1* and *haf2* promoter sequences of nine BPF-associated isolates (see Materials and Methods). These nine strains are indistinguishable by serological, multilo-

```
                                                   ↓
HifE       mktlttyaky ftpiskiafl fcflmgniae aTIKRAKFTN GFSGINRIIT YTFEGSSTMI  60
HafE       mktlkssplh lsllpylmgl l-flfsaypa maGPKQASSG QVQGLRKVFT FN-GDNSRLI  58
Consensus  mktl...... .........l ..fl...... .......... ...G.....T ......S..I  60
                                            ↑

HifE       ASATTPEQIL FSKARDNTVI DPSYSNNVQQ WSV-FNNWID TTVSGDTGYS FAGFSCVSNP  119
HafE       ALSTKPEIVL FSKALNNPKY DIFDGNIATT LPTPFTDWIP TNVTGDSGYS FRDFTCGG--  116
Consensus  A..T.PE..L FSKA..N... D....N.... ....F..WI. T.V.GD.GYS F..F.C....  120

HifE       CAQMQLPLRF YLDSAILEAT SMRSAD-NQV IFKIRQHPEL GVSFQLGMK- KGIED-VKWL  176
HafE       CSPLTLPLNF HLNTAKLENT NMKDPNTGDS IFKIKDHPEL GVSFQLGTRY QGRDDYLAPL  176
Consensus  C....LPL.F .L..A.LE.T .M........ IFKI..HPEL GVSFQLG... .G..D....L  180

HifE       SNLQQEDFLL TTLQIYFGDA ADISFKVRAK LHLL--KLPT ENTELPTMKL NLGQIKLQSW  234
HafE       NFGINNTARI GSVQLYFGNT VSIFFLLRAK LHLIDNSIPN GMKRIPTLDL DLGYIKMEP-  235
Consensus  .......... ...Q.YFG.. ..I.F..RAK LHL.....P. .....PT..L .LG.IK....  240

HifE        GINNWGRTKV SYRVQDVGSL NVQLKTPKIY FIQQQRQCIL NSTYKKIPVT LKSVKKREFE 294
HafE        -V-SFKPKVI TFAIRDNGRI NVNLKTPQVM LLEKQRQCFL VSREKNKTVP LKEVKKSVFD 293
Consensus   .......... .....D.G.. NV.LKTP... ....QRQC.L .S..K...V. LK.VKK..F. 300

HifE       TNTEIEGGQF KLRVNCEDTT YNKFNGKWLF PVVKVTFRGE DGTMNDGTNE LLRTQTGTGQ  354
HafE       NIEEIEGGEF RLSVSCDKTE -NQINRQWLF PRVMITFKGE NGTNNNGLSD LLKTEKGNDQ  352
Consensus  ...EIEGG.F .L.V.C..T. .N..N..WLF P.V..TF.GE .GT.N.G... LL.T..G..Q  360

HifE       ATGVSLKIKR DSGNGDSVKY GLDSANMNNH GQFELKKQPS PAGGDQSAEE TFKVYYVKDT  414
HafE       AKGVSLRIKR QNG-TDTVKY GLDSPQMNNP GQFQLQKQPS --EADKNAQE TFKVYYVKDQ  409
Consensus  A.GVSL.IKR ..G..D.VKY GLDS..MNN. GQF.L.KQPS ....D..A.E TFKVYYVKD.  420

HifE       TRGALTEGKV KAAATFTMSY Q  435
HafE       TRGTLTEGKV NAAATFTMSY Q  430
Consensus  TRG.LTEGKV .AAATFTMSY Q  441
```

FIG. 3. Comparison of the HifE and HafE (*haf1*-encoded) amino acid sequences. The predicted signal peptidase site for each protein is indicated with an arrow, and the resulting signal peptides are shown in lowercase letters. The four conserved cysteine residues are shown in boldface type. Dots in the consensus sequence indicate nonidentity; dashes are gaps inserted by the alignment program.

cus enzyme mobility, and sodium dodecyl sulfate-polyacrylamide gel electrophoresis typing (3). DNA fragments specific for either the *haf1* or *haf2* promoter region were amplified from each strain by selecting primer sets with one oligonucleotide complementary to chromosomal sequences downstream of *hafA* specific for each gene cluster and one oligonucleotide on the *hafB*-proximal side of the promoter. Primers PMC and HA4 were used to amplify *haf1*; PUR and HA3 were used for *haf2* (Fig. 2). The DNA sequence of the promoter on the PCR fragments was then determined by cycle sequencing with the primer HA2. The number of TA repeats on each cluster is shown in Table 1. Our results support the model of van Ham for *H. influenzae* pilus phase variation; all of the HA^+ strains tested have at least one *haf* promoter with 10 TA repeats, while F3049 and F1946, which are HA^- have 9 TA units in both clusters. Moreover, our data indicate that either *haf* cluster in the phase "on" mode can support pilus synthesis. In the case of strains 34b, F2052, and F3224, the *haf1* promoter has 10 TA repeats and is therefore predicted to express pili, whereas *haf2* has 9 TA repeats. For F3031 and F3114, only the *haf2* genes would express detectable pili, supporting the results of the earlier hemadsorption experiment with the pMF1 and pMF6 clusters.

Transposon insertion in *hafE* eliminates hemadsorption. To demonstrate that the predicted *hafE* gene sequence specified a protein necessary for pilus function, we constructed a *hafE*::mini-$\gamma\delta$ insertion in the *E. coli* high-copy-number plasmid pNEB193 containing the *haf2* cluster and found it to confer an HA^- phenotype in HB101. In order to verify that the HA^- phenotype was a result of the *hafE* transposon insertion and not due to a switch to phase "off" pilus expression, we compared phase shifting of two HA^- mini-$\gamma\delta$ mutants in HB101. Insertion D6 is located in the *hafE* gene at position 1247 (out of 1289 nucleotides), while insertion F5 is situated outside of the pilus cluster, downstream of the *hafA* coding

TABLE 1. Phase variation of pilus gene expression in BPF strains

Strain	Data for pilus gene phase variation[a]:				Hemagglutination[b]
	haf1		*haf2*		
	TA units	Expression	TA units	Expression	
34b	10	On	9	Off	+
F1946	9	Off	9	Off	−
F2052	10	On	9	Off	+
F3031	9	Off	10	On	+
F3034	10	On	10	On	+
F3049	9	Off	9	Off	−
F3114	9	Off	10	On	+
F3224	10	On	9	Off	+
G1097	11	On	10	On	+

[a] Shown are the number of TA repeats between the overlapping promoters of each gene cluster in each strain and the pilus expression predicted by the phase variation model of van Ham et al. (39).

[b] Hemagglutination-positive strains, standardized to an A_{600} of 1.0 after overnight growth, produced visible hemagglutination of human O-positive erythrocytes in a microplate assay at dilutions of 1:16 or greater. The end-point dilutions for each strain were as follows: 34b, 1:32; F1946, <1:1; F2052, 1:64; F3031, 1:16; F3034, 1:64; F3049, <1:1; F3114, 1:32; F3224, 1:32; and G1097, 1:128. Results represent the average of data from at least two experiments. +, positive for hemagglutination; −, hemagglutination negative.

sequence (Fig. 1). In two independent experiments, single HA^- colonies of each mutant were diluted and grown on Luria agar plates with selection for ampicillin resistance. The number of HA^+ colonies was determined by hemadsorption. Of the F5 colonies, 650 out of 17,200 were HA^+ (frequency, 0.038), but there were no HA^+ revertants in 18,000 colonies with the D6 insertion. These data indicate that the *hafE* mutant is defective in pilus-mediated hemagglutination.

DISCUSSION

Because of the potential importance of pili in the pathogenesis of BPF, we have cloned and sequenced two fimbrial gene clusters from BPF-associated Hae strain F3031. We have identified five putative ORFs (*hafA* to *hafE*) on the nearly identical *haf1* and *haf2* gene clusters. The presence of a second fimbrial gene cluster and the novel sequence of the putative adhesin, HafE, distinguish the BPF-associated pilus system from those previously described for Hib. Examination of nine wild-type BPF-associated strains suggested that both Hae fimbrial gene clusters undergo independent phase variation in vivo by a mechanism similar to that of the *H. influenzae* LKP pilus (39). Elimination of hemadsorption by transposon insertion within the *haf* cluster suggested that the putative ORFs encode functional pilus determinants and that HafE is essential to erythrocyte adherence.

The pilus genes of Hae appear to share a close evolutionary relationship to the genes specifying pili of Hib (Fig. 1) (24, 38, 42). Presumably, biogenesis of the BPF-associated pilus is similar to the LKP pilus model proposed by van Ham et al. (38). The putative HafB protein, which is 95% identical to HifB, also has 34 and 45% similarity to chaperones from the *E. coli* Pap and *Bordetella pertussis* pilus clusters, respectively (16, 38, 46). The HafC usher is homologous to PapC and FimC (26 and 24% similarity). HafA, HafD, and HafE have conserved C termini with a glycine and a tyrosine 1 and 13 residues, respectively, from the final amino acid, typical of structural subunits of type 1 pili (10, 13, 21, 24). HafE is approximately twice the size of the other pilin proteins, a characteristic of pilus adhesins, such as PapG, F-17G, and MrkD, of gram-negative organisms. Another shared characteristic of adhesins is the presence of two pairs of conserved cysteine residues (Fig. 3), which give the pilin a two-domain structure (11, 20, 21). Insertion of the mini-$\gamma\delta$ element in the $3'$ end of the *hafE* gene (Fig. 1) encoding the conserved C terminus might have interfered with chaperone binding, with resultant inactivation of HafE. The fact that pilus-mediated hemagglutination is eliminated by this mutation indicates that HafE is required either for pilus assembly or for adhesion, or for both functions.

The most conspicuous difference between the *H. influenzae* pilus complexes is in the primary sequences of HafE and HifE. The proteins show sequence divergence predominantly over their N-terminal two-thirds. The same region of the PapG adhesin is involved in binding mammalian cell receptors (14, 15). Therefore, the Hib and BPF-associated Hae pili, although similar in structure and assembly, may have different tissue specificities for attachment, owing to their distinctive putative adhesins. For example, fimbriae might be important for BPF-associated clones to adhere to and damage endothelial cells in the later stages of disease, following the initial epithelial cell attachment and invasion. Whether HafE receptor specificity is important in the pathogenesis of BPF-associated Hae is a subject for further investigation.

In comparing the pilus proteins, it is notable that the amino acid sequences of the Hae pilus structural proteins are significantly more diverged from their Hib counterparts than the chaperone and usher proteins dedicated to assembly (Fig. 1). The difference in the rate of evolutionary change between the primary sequences may reflect their function. The mechanism of pilus assembly involves specific interactions between pilus proteins and the inner and outer membranes; therefore, alterations in HafB and HafC may prove deleterious. In contrast, mutations that change surface-exposed components of the pilus structure can be selectively neutral or even result in beneficial change in adherence specificity. Further, the novel pilus structure might help *Haemophilus* cells to evade the host immune response, enhancing survival of the organism.

The carriage of two pilus gene clusters by the BPF-associated clone strain is a significant deviation from Hib. One reason for a pathogenic bacterium to maintain two or more pilus complexes is that the fimbriae could conceivably interact with different mammalian cell receptors, thereby broadening the range of attachment of the microorganism. Examples of this strategy are found in strains of uropathogenic *E. coli*, such as J96, which carry the *pap* and *prs* genes (22). These pilus clusters, apparently arising from an ancestral duplication, encode structurally similar fimbriae that bind to distinct classes of carbohydrate receptors. The carriage of two or more different phase-switching pilus complexes could also permit antigenic variation that circumvents the host immune response. In *Neisseria gonorrhoeae*, homologous recombination between promoterless silent copies of pilin genes and the pilus expression locus may result in phase variation as well as antigenic variation (28). However, the finding that the sequences of the genes encoding the pilin subunits of *haf1* and *haf2* were more than 99% conserved and hence produce structurally identical pili with identical adhesins discounts these explanations. An alternative explanation is that *H. influenzae* strains containing two independently phase-variable clusters will more likely express pili than strains with only one cluster. This is because only one cluster in phase "on" is required for pilus expression whereas both sets of genes have to be in phase "off" to give a nonpiliated phenotype. Perhaps the finding of two phase-variable pilus complexes in strain F3031 explains why BPF-associated Hae isolates from blood are mostly piliated while Hib strains are generally nonpiliated (23). Additionally, strains with two functional sets of pilus genes may produce significantly greater amounts of fimbriae on their cell surface. In the pathogenesis of BPF, abundant piliation could improve the efficiency of mucosal colonization or perhaps contribute to pathogenesis after systemic invasion.

Making use of the recently published sequence of the entire *H. influenzae* Rd genome (8), we showed that the *haf1* pilus cluster lies between a gene homologous to *pmbA* of *E. coli* and a hypothetical ORF (HI1153). This is the first time pilus genes of *H. influenzae* have been identified at this location, and it illustrates the value of whole-genome sequences of pathogenic bacteria. The extremely close nucleotide similarity between *haf1* and *haf2* ORFs suggests that duplication of the clusters occurred by a recent intragenic recombination event rather than acquisition of DNA from another strain through natural transformation.

The pilus gene clusters of BPF-associated Hae represent a variation on the Hib pilus complexes characterized previously. The presence of two nearly identical copies of pilus genes and differences in the primary sequences of the putative adhesins may all contribute to the virulence of Hae. Future studies are needed to determine any relationship between Haf pilus expression and features of BPF pathogenesis, such as endothelial cytotoxicity, resistance to serum bactericidal activity, and human epithelial cell invasion.

ACKNOWLEDGMENTS

We thank Leonard Mayer, Frederick Quinn, and Claire Berg for providing strains and William Shafer for review of the manuscript. Kavitha Kosuri, Phalasy Juieng, Maria-Elena Grogan, and Jason Opdyke contributed technical support.

This work was funded by a Veterans Administration Merit Award to M.M.F.

REFERENCES

1. **Berg, C. M., G. Wang, L. D. Strausbaugh, and D. E. Berg.** 1993. Transposon-facilitated sequencing of DNAs cloned in plasmids. Methods Enzymol. **218:** 279–306.
2. **Boyer, H. W., and D. Roulland-Dussoix.** 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. **41:**459.
3. **Brenner, D. J., L. W. Mayer, G. M. Carlone, L. H. Harrison, W. F. Bibb, M. C. de Cuento Brandileone, F. O. Sottnek, K. Irino, M. W. Reeves, J. M. Swenson, K. A. Birkness, R. S. Weyant, S. F. Berkley, T. C. Woods, A. G. Steigerwalt, P. A. D. Grimont, R. M. McKinney, D. W. Fleming, L. L. Gheesling, R. C. Cooksey, R. J. Arko, C. V. Broome, and The Brazilian Purpuric Fever Study Group.** 1988. Biochemical, genetic, and epidemiologic characterization of *Haemophilus influenzae* biogroup aegyptius (*Haemophilus aegyptius*) strains associated with Brazilian purpuric fever. J. Clin. Microbiol. **26:**1524–1534.
4. **Collins, J., and H. J. Bruning.** 1978. Plasmids useable as gene-cloning vectors in an in vitro packaging by coliphage lambda: "cosmids". Gene **4:**85–107.
5. **Connor, E. M., and M. R. Loeb.** 1983. A hemadsorption method for detection of colonies of *H. influenzae* type b expressing fimbriae. J. Infect. Dis. **148:** 855–860.
6. **Farley, M. M., D. S. Stephens, S. L. Kaplan, and E. O. Mason, Jr.** 1990. Pilus- and non-pilus-mediated interactions of *Haemophilus influenzae* type b with human erythrocytes and human nasopharyngeal mucosa. J. Infect. Dis. **161:**274–280.
7. **Farley, M. M., A. M. Whitney, P. Spellman, F. D. Quinn, R. S. Weyant, L. Mayer, and D. S. Stephens.** 1992. Analysis of attachment and invasion of human epithelial cells by *Haemophilus influenzae* biogroup *aegyptius*. J. Infect. Dis. **165:**S111–S114.
8. **Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J.-F. Tomb, B. A. Dougherty, J. M. Merrick, K. McKenny, G. Sutton, W. FitzHugh, C. Fields, J. D. Gocayne, J. Scott, R. Shirley, L.-I. Liu, A. Glodek, J. M. Kelley, J. F. Weidman, C. A. Phillips, T. Spriggs, E. Hedblom, M. D. Cotton, T. R. Utterback, M. C. Hanna, D. T. Nguyen, D. M. Saudek, R. C. Brandon, L. D. Fine, J. L. Fritchman, J. L. Fuhrman, N. S. M. Geoghagen, C. L. Gnehm, L. A. McDonald, K. V. Small, C. M. Fraser, H. O. Smith, and J. C. Venter.** 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science **269:**496–512.
9. **Foglino, M., S. Gharbi, and A. Lazdunski.** 1986. Nucleotide sequence of the *pepN* gene encoding aminopeptidase N of *Escherichia coli*. Gene **49:**303–309.
10. **Forney, L. J., C. F. Marrs, S. L. Bektesh, and J. R. Gilsdorf.** 1991. Comparison and analysis of the nucleotide sequences of pilin genes from *H. influenzae* type b strains Eagan and M43. Infect. Immun. **59:**1991–1996.
11. **Gerlach, G. F., S. Clegg, and B. L. Allen.** 1989. Identification and characterization of the genes encoding the type 3 and type 1 fimbrial adhesins of *Klebsiella pneumoniae*. J. Bacteriol. **171:**1262–1270.
12. **Gilsdorf, J. R., C. F. Marrs, K. W. McCrea, and L. J. Forney.** 1990. Cloning, expression, and sequence analysis of the *Haemophilus influenzae* type b strain M43p^{+} pilin gene. Infect. Immun. **58:**1065–1072.
13. **Hacker, J.** 1990. Genetic determinants coding for fimbriae and adhesins of extraintestinal *Escherichia coli*. Curr. Top. Microbiol. Immunol. **151:**1–27.
14. **Haslam, D. B., T. Boren, P. Falk, D. Ilver, A. Chou, Z. Xu, and S. Normark.** 1994. The amino-terminal domain of the P-pilus adhesin determines receptor specificity. Mol. Microbiol. **14:**399–409.
15. **Hultgren, S. J., F. Lindberg, G. Magnusson, J. Kihlberg, J. M. Tennant, 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.
16. **Hultgren, S. J., and S. Normark.** 1991. Biogenesis of the bacterial pilus. Curr. Opin. Genet. Dev. **1:**313–318.
17. **Kar, S., S. C.-M. To, and C. C. Brinton, Jr.** 1990. Cloning and expression in *Escherichia coli* of LKP pilus genes from a nontypeable *Haemophilus influenzae* strain. Infect. Immun. **58:**903–908.
18. **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.
19. **Langermann, S., and A. Wright.** 1990. Molecular analysis of the *Haemophilus influenzae* type b pilin gene. Mol. Microbiol. **4:**221–230.
20. **Lindberg, F. P., B. Lund, and S. Normark.** 1984. Genes of pyelonephritogenic *E. coli* required for digalactoside-specific agglutination of human cells. EMBO J. **3:**1167–1173.
21. **Lintermans, P. F., A. Bertels, C. Schlicker, F. Deboeck, G. Charlier, P. Pohl, M. Norgren, S. Normark, M. Van Montagu, and H. De Greve.** 1991. Identification, characterization, and nucleotide sequence of the F17-G gene, which determines receptor binding of *Escherichia coli* F17 fimbriae. J. Bacteriol. **173:**3366–3373.
22. **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.
23. **Mayer, L. W., W. F. Bibb, K. A. Birkness, K. Irino, R. S. Weyant, M. W. Reeves, and J. M. Swenson.** 1989. Distinguishing clonal characteristics of the Brazilian purpuric fever-producing strain. The Brazilian Purpuric Fever Study Group. Pediatr. Infect. Dis. J. **8:**241–243.
24. **McCrea, K. W., W. J. Watson, J. R. Gilsdorf, and C. F. Marrs.** 1994. Identification of *hifD* and *hifE* in the pilus gene cluster of *Haemophilus influenzae* type b strain Eagan. Infect. Immun. **62:**4922–4928.
25. **Moxon, E. R., R. A. Diech, and C. Connelly.** 1984. Cloning of chromosomal DNA from *Haemophilus influenzae*—its use for studying the expression of type b capsule and virulence. J. Clin. Invest. **73:**298–306.
26. **Porto, M. H. O., G. J. Noel, P. J. Edelson, and The Brazilian Purpuric Fever Study Group.** 1989. Resistance to serum bactericidal activity distinguishes Brazilian purpuric fever (BPF) case strains of *Haemophilus influenzae* biogroup aegyptius (*H. aegyptius*) from non-BPF strains. J. Clin. Microbiol. **27:**792–794.
27. **Quinn, F. D., R. S. Weyant, M. J. Worley, E. H. White, E. A. Utt, and E. A. Ades.** 1995. Human microvascular endothelial tissue culture cell model for studying pathogenesis of Brazilian purpuric fever. Infect. Immun. **63:**2317–2322.

27a. **Read, T. D., and M. M. Farley.** Mol. Microbiol., in press.

28. **Robertson, B. D., and T. F. Meyer.** 1992. Genetic variation in pathogenic bacteria. Trends Genet. **8:**422–428.
29. **Rodriguez-Sainz, M. C., C. Hernandez-Chico, and F. Moreno.** 1990. Molecular characterization of *pmbA*, an *Escherichia coli* chromosomal gene required for the production of the antibiotic peptide MccB17. Mol. Microbiol. **4:**1921–1932.
30. **Rubin, L. G., E. S. Gloster, G. M. Carlone, and The Brazilian Purpuric Fever Study Group.** 1989. An infant rat model of bacteremia with Brazilian purpuric fever isolates of *Haemophilus aegyptius* biogroup *aegyptius*. J. Infect. Dis. **160:**476–482.
31. **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.
32. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
33. **Smith, T. F., and M. S. Waterman.** 1981. Identification of common molecular subsequences. J. Mol. Biol. **147:**195–197.
34. **St. Geme, J. W., III., and S. Falkow.** 1993. Isolation, expression, and nucleotide sequencing of the pilin structural gene of the Brazilian purpuric fever clone of *Haemophilus influenzae* biogroup aegyptius. Infect. Immun. **61:** 2233–2237.
35. **St. Geme, J. W., III., J. R. Gilsdorf, and S. Falkow.** 1991. Surface structures and adherence properties of diverse strains of *Haemophilus influenzae* biogroup aegyptius. Infect. Immun. **59:**3366–3371.
36. **Tiedeman, A. A., J. Keyhani, J. Kamholz, H. A. Daum III., J. S. Gots, and J. M. Smith.** 1989. Nucleotide sequence of the *purEK* operon encoding 5′-phosphoribosyl-5-aminoimidazole carboxylase of *Escherichia coli* K-12. J. Bacteriol. **171:**205–212.
37. **van Ham, S. M., F. R. Mooi, M. G. Sindhunata, W. R. Maris, and L. van Alphen.** 1989. Cloning and expression in *Escherichia coli* of *Haemophilus influenzae* fimbrial genes establishes adherence to oropharyngeal epithelial cells. EMBO J. **8:**3535–3540.
38. **van Ham, S. M., L. van Alphen, F. R. Mooi, and J. P. van Putten.** 1994. The fimbrial gene cluster of *Haemophilus influenzae* type b. Mol. Microbiol. **13:**673–684.
39. **van Ham, S. M., L. van Alphen, F. R. Mooi, and J. P. M. van Putten.** 1993. Phase variation of *H. influenzae* fimbriae: transcriptional control of two divergent genes through a variable combined promoter region. Cell **73:**1187–1196.
40. **Viera, J., and J. Messing.** 1991. New pUC-derived vectors with different selectable markers and DNA replication origins. Gene **100:**189–194.
41. **von Heijne, G.** 1986. Patterns of amino acids near signal-sequence cleavage sites. Eur. J. Biochem. **133:**17–21.
42. **Watson, W. J., J. R. Gilsdorf, M. A. Tucci, K. W. McCrea, L. J. Forney, and C. F. Marrs.** 1994. Identification of a gene essential for piliation in *Haemophilus influenzae* type b with homology to the pilus assembly platform genes of gram-negative bacteria. Infect. Immun. **62:**468–475.
43. **Weyant, R. S., W. F. Bibb, D. S. Stephens, B. P. Holloway, W. F. Moo-Penn, K. A. Birkness, L. O. Helsel, and L. W. Mayer.** 1990. Purification and characterization of a pilin specific for Brazilian purpuric fever-associated *Haemophilus influenzae* biogroup aegyptius (*H. aegyptius*) strains. J. Clin. Microbiol. **28:**756–763.
44. **Weyant, R. S., F. D. Quinn, E. A. Utt, M. Worley, V. G. George, F. J. Candal, and E. W. Ades.** 1994. Human microvascular endothelial cell toxicity caused

by Brazilian purpuric fever-associated strains of *Haemophilus influenzae* biogroup *aegyptius*. J. Infect. Dis. **169:**430–433.
45. **Whitney, A. M., and M. M. Farley.** 1993. Cloning and sequence analysis of the structural pilin gene of Brazilian purpuric fever-associated *Haemophilus influenzae* biogroup aegyptius. Infect. Immun. **61:**1559–1562.
46. **Willems, R. J., H. G. van der Heide, and F. R. Mooi.** 1992. Characterization of a *Bordetella pertussis* fimbrial gene cluster which is located directly downstream of the filamentous haemagglutinin gene. Mol. Microbiol. **6:**2661–2671.
47. **Xu, Z., C. H. Jones, D. Haslam, J. S. Pinker, 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.
48. **Zav'yalov, V. P., G. A. Zav'yalova, A. I. Denesyuk, and T. Korpela.** 1995. Modelling of steric structure of a periplasmic molecular chaperone Caf1M of *Yersinia pestis*, a prototype member of a subfamily with characteristic structural and functional features. FEMS Immunol. Med. Microbiol. **11:**19–24.