INFECTION AND IMMUNITY, June 1991, p. 1991–1996
0019-9567/91/061991-06$02.00/0
Copyright © 1991, American Society for Microbiology
Vol. 59, No. 6

Comparison and Analysis of the Nucleotide Sequences of Pilin Genes from *Haemophilus influenzae* Type b Strains Eagan and M43

LARRY J. FORNEY,[1†*] CARL F. MARRS,[2] SUSAN L. BEKTESH,[1] AND JANET R. GILSDORF[3]

Synergen, Inc., Boulder, Colorado 80301,[1] and School of Public Health, University of Michigan,[2] and C. S. Mott Children's Hospital and University of Michigan Medical School,[3] Ann Arbor, Michigan 48109

Received 9 November 1990/Accepted 11 March 1991

Previous studies have demonstrated antigenic differences among the pili expressed by various strains of *Haemophilus influenzae* type b (Hib). In order to understand the molecular basis for these differences, the structural gene for pilin was cloned from Hib strain Eagan (p^+) and the nucleotide sequence was compared to those of strains M43 (p^+) and 770235 b^0f^+, which had been previously determined. The pilin gene of Hib strain Eagan (p^+) had a 648-bp open reading frame that encoded a 20-amino-acid leader sequence followed by the 196 amino acids found in mature pilin. The translated sequence was three amino acids larger than pilins of strains M43 (p^+) and 770235 b^0f^+ and was 78% identical and 95% homologous when conservative amino acid substitutions were considered. Differences between the amino acid sequences were not localized to any one region but rather were distributed throughout the proteins. Comparison of protein hydrophilicity profiles showed several hydrophilic regions with sequences that were conserved between strain Eagan (p^+) and pilins of other Hib strains, and these regions represent potentially conserved antigenic domains. Southern blot analyses using an intragenic probe from the pilin gene of strain Eagan (p^+) showed that the pilin gene was conserved among all type b and nontypeable strains of *H. influenzae* examined, and only a single copy was present in these strains. Homologous genes were not present in the phylogenetically related species *Pasteurella multocida*, *Pasteurella haemolytica*, and *Actinobacillus pleuropneumoniae*. These data indicate that the pilin gene was highly conserved among different strains of *H. influenzae* and that small differences in the pilin amino acid sequences account for the observed antigenic differences of assembled pili from these strains.

Strains of *Haemophilus influenzae* that elaborate a serotype b extracellular capsular polysaccharide cause bacteremia and serious focal infections, including meningitis, epiglottitis, and cellulitis (21, 22, 28), in young children. Colonization of the nasopharynx is apparently an important first step in pathogenesis by *H. influenzae* type b (Hib) since penetration of the nasopharyngeal epithelium appears to be the typical way by which the organism gains access to the bloodstream (22, 23). By analogy with other bacterial pathogens (3, 7), colonization by Hib is probably facilitated by specific adhesins, and previous investigators have postulated that pili expressed by the organism might have a role in this process (12, 24). This hypothesis is supported by the observation that Hib pili mediate adherence to buccal epithelial cells (12, 18, 27, 29), adenoid tissue (19), and respiratory epithelial cells (26).

Although pili from different strains of Hib apparently have a common function, they demonstrate immunological variability (5, 11). Gilsdorf et al. (11) found that antisera specific for native pili from Hib strains M43 (p^+) (also referred to as AO2) and Eagan (p^+) were not cross-reactive; i.e., antisera specific for the native pili of strain Eagan (p^+) did not react with M43 (p^+) and antisera specific for the pili of M43 (p^+) did not bind to Eagan (p^+) cells. These investigators also examined the ability of these antisera to recognize pili expressed by 21 other Hib strains and found that neither antiserum recognized all of the piliated strains tested but all strains were recognized by one of the antisera. Similar results were obtained by Brinton et al. (5), who showed that strains of Hib could be placed in several groups based on the reactivity of various strains with antisera raised against pili purified from Hib strains representative of each group. The pili of virtually all type b strains of *H. influenzae* (>90%) were placed in two serological groups designated LKP3 and LKP4. Hib strain Eagan (p^+) produces LKP3 pili, whereas strain M43 (p^+) is probably representative of the LKP4 group. Although antigenic differences exist between the pili expressed by various strains of Hib, the molecular basis for these differences is unknown and the immunodominant epitopes of the major subunit protein, pilin, have not been defined.

The nucleotide sequences of pilin genes from Hib strains M43 (p^+) (10, 16) and 770235 b^0f^+ (30) are identical, and thus the proteins they encode are likely to be representative of a single serological group. These pilin genes encode proteins that have a 20-amino-acid leader sequence followed by a mature protein comprising 193 amino acids. We have undertaken studies of pilin from Hib strain Eagan, which is immunologically distinct from pilin of strains M43 (p^+) and 770235 b^0f^+, in order to gain insight into the basis for the observed serological differences. In this study, we determined the DNA sequence of the pilin structural gene from strain Eagan (p^+) and confirmed the previously reported amino acid sequence (1). We compared the DNA and amino acid sequences of pilins from various Hib strains to define sequence differences and regions of sequence identity that were hydrophilic and thus constitute conserved domains that are potentially antigenic.

MATERIALS AND METHODS

Bacteria and restriction enzymes. Hib strain Eagan (p^-) was obtained from Terry Stull (Medical College of Pennsyl-

* Corresponding author.
† Present address: Center for Microbial Ecology, Michigan State University, East Lansing, MI 48824-1325.

vania), and Eagan (p^+) is a piliated variant obtained through enrichment by hemagglutination (8). Hib strains M43 (p^+), AA9, AA106, AAr122, and M9 and nontypeable (NT) *H. influenzae* M37 were clinical isolates provided by one of us (J. R. Gilsdorf), while NT *H. influenzae* 1128 was provided by Lauren Bakaletz (Ohio State University, Columbus). All Hib and NT *H. influenzae* strains were grown in brain heart infusion broth (Difco, Detroit, Mich.) supplemented with 10 μg of NAD per ml and 10 μg of hemin per ml at 37°C.

Escherichia coli DH5α [F^- f80d *lacZ*ΔM15 *endA1 recA1 hsdR17* (rK^- mK^+) *supE44 thi-1* d^- *gyrA96* Δ(*lacZYA-argF*)*U169*] and LE392 (*hsdR514 hsdM supE44 supF58 lacY1*) were used for all molecular cloning experiments.

Pasteurella multocida 055, *Pasteurella haemolytica* 018 and 046, and *Actinobacillus pleuropneumoniae* 026 were provided by Cathy Reese (Pfizer, Inc.), and Howard Rush (University of Michigan, Ann Arbor) provided *P. multocida* R1, 2929, 3761, and 3770.

All restriction enzymes were purchased from either Promega (Madison, Wis.) or New England BioLabs (Beverly, Mass.) and were used according to the manufacturers' instructions.

DNA library construction and screening. Chromosomal DNA isolated from Hib strain Eagen (p^+) (21) was partially digested with *Eco*RI and ligated to λGEM-11 (Promega) that had been previously digested with *Eco*RI and treated with calf intestinal alkaline phosphatase. The ligated DNA was packaged into λ phage particles by using Giga-pack Gold (Stratagene, LaJolla, Calif.). Aliquots of the library (10 μl) were mixed with 0.8 ml of *E. coli* LE392 that had been grown to mid-exponential phase in TB broth (20) with 10 mM $MgSO_4$ and 0.2% maltose, centrifuged, and resuspended in 10 mM $MgSO_4$ to an optical density (A_{600}) of 0.5. This suspension was incubated for 15 min at 37°C, added to molten (45°C) NZYCM medium (20) that contained 0.7% agarose, and plated on NZYCM agar in 150-mm plates. After the agarose had solidified, the plates were incubated for about 8 h at 37°C. A total of approximately 5,000 plaques were screened. Plaque lifts, made using BA-85 nitrocellulose membranes (New England Nuclear, Chicago, Ill.), were processed by methods described by Maniatis et al. (20). Duplicate plaque lifts were prepared from each plate.

The nitrocellulose membranes were probed with two degenerate oligonucleotides that were based on two widely separated regions of the amino acid sequence of pilin from Hib strain Eagan (1). Oligonucleotide probe 1 was a 20-mer that corresponded to amino acids 83 to 89 and had the following sequence (where N represents all four bases): TGGGA(A/G)TT(A/C)GCNGA(T/C)AA(A/G)GA. Oligonucleotide probe 2 was a 17-mer that corresponded to amino acids 17 to 23 and had the following sequence: GA(A/G) T(T/C)AACN(T/C)GTCA(A/G)GT. Each probe was end labeled by using polynucleotide kinase (Bethesda Research Laboratories, Gaithersburg, Md.) and [α-^{32}P]ATP (New England Nuclear). Unincorporated label was removed by using a Sep-pak cartridge (Whatman) according to the manufacturer's instructions.

One set of membranes (one from each plate) was prehybridized at 50°C in a hybridization solution that contained 6× SSC (0.9 M NaCl plus 0.09 M sodium citrate), 10× Denhardt's solution (20), 0.1% sodium dodecyl sulfate (SDS), 0.05% PP_i, and 100 μg of bakers' yeast tRNA per ml (Boehringer Mannheim). After overnight incubation, the solution was drained from the membranes and replaced with fresh hybridization solution that contained a ^{32}P-labeled oligonucleotide probe 1 (>1 μCi/ml). The membranes were hybridized overnight at 50°C and then washed twice at room temperature in 6× SSC–0.5% SDS and once at 50°C in the same solution. The duplicate set of membranes were treated in a similar manner except that the hybridization solution contained ^{32}P-labeled oligonucleotide probe 2 and all incubations and the final posthybridization wash were done at 42°C. The membranes were blotted dry, sandwiched in plastic wrap, and exposed overnight to XAR-5 film (Eastman Kodak Co., Rochester, N.Y.) at −76°C. After development of the film, the images of filters hybridized with oligonucleotide probes 1 and 2 were compared with two clones that hybridized to both probes were identified and plaque purified. One of these clones, λGEM-11-A1, was selected for further study.

Subcloning of the pilin gene. DNA from λGEM-11-A1 was prepared by previously described methods (21). The 16.5-kb insert DNA of λGEM-11-A1 was analyzed by restriction enzyme digestion and Southern blot analyses by standard methods (20). Oligonucleotide probes 1 and 2 both hybridized to a 1.1-kb *Xho*I-*Bgl*II fragment. This fragment was isolated by agarose gel electrophoresis and subcloned into pBSKM(−) (Stratagene) that had been digested with *Xho*I and *Bam*HI. *E. coli* DH5α was transformed with the ligated mixture, and the cells were plated on Luria agar that contained 5-bromo-4-chloro-3-indolyl-β-D-glucopyranoside (X-Gal; Sigma Chemical Co., St. Louis, Mo.) and isopropylthiogalactopyranoside (IPTG; Sigma). Clones that carried the pilin gene were identified by probing colony lifts (20) with ^{32}P-labeled oligonucleotide probes 1 and 2. One such clone, designated pPhb-1, was chosen for further study and found to form blue colonies on Luria agar that contained X-Gal. The blue coloration of colonies with this plasmid most likely resulted from the hydrolysis of X-Gal by a β-galactosidase fusion protein expressed by the cells. This was later substantiated by sequence analyses which showed that the insert DNA encoded the 5′ end of the Hib pilin gene cloned in frame to the *lacZ* gene of the vector.

The entire Hib pilin gene and flanking DNA of λGEM-11-A1 was cloned as a 16.5-kb *Bam*HI fragment into the *Bam*HI site of pXCosXV. pXCosXV was constructed as follows (5a). A kanamycin resistance cassette with *Bam*HI cohesive ends (PL Biochemicals, Piscataway, N.J.) was ligated to pRK290 (20) that had been digested with *Bgl*II. The resulting plasmid was digested with *Eco*RI and ligated to a 1.1.-kb *Eco*RI fragment that carried the *E. coli lacI^q* gene from pMMB22 (2) to yield pRK290 (Kan^r *LacI^q*). The *cos* sites of pXCosXV were obtained from pMMB33-tac. This vector was constructed through ligation of a synthetic *trp-lac* promoter (2) which had *Hinc*II cohesive ends and *Bam*HI and *Bgl*II sites downstream from the promoter into pMMB33 (9) that had been digested with *Hpa*I. Finally, pRK290 (Kan^r *lacI^q*) was digested with *Pst*I and ligated to similarly digested pMMB33-tac to give pXCosXV.

The Hib DNA was cloned into a unique *Bam*HI site located immediately downstream from the *tac* promoter sequence of pXCosXV. *E. coli* HB101 was infected with the recombinant cosmid that had been packaged in vitro (Packagene; Promega), and clones resistant to tetracycline (10 μg/ml) were selected on Luria agar. The cosmid DNA from one such clone was designated pXCosXV-12 and used in further subcloning experiments.

The entire Hib pilin gene and DNA downstream from the pilin gene was subcloned from pXCosXV-12 as a 10.5-kb *Xho*I-*Bam*HI fragment into pGEM7zf(+) (Promega) that had been digested with the same enzymes to yield pBX10.5 The resultant plasmid was used to transform *E. coli* DH5α.

Genomic Southern blots. Genomic DNAs isolated from type b and NT strains of *H. influenzae*, *P. multocida*, *P. haemolytica*, and *A. pleuropneumoniae* were digested with either *Taq*I or *Hae*III, and the DNA fragments were separated by agarose gel electrophoresis. The digested DNAs were transferred by capillary action to a Nytran membrane (Schleicher & Schuell, Keene, N.H.) according to the manufacturer's recommendations. The probe for Hib pilin was a 379-bp fragment that was purified by gel electrophoresis following digestion of pPhb-1 with *Dra*I and *Xba*I. The *Dra*I site is located at nucleotide 185 of the pilin gene, and the *Xba*I site is located immediately 3′ to the gene in the polylinker of the vector. (The *Xba*I site is 12 bp downstream from a *Bgl*II site within the pilin gene that was destroyed during subcloning.) The fragment was labeled with ^{32}P by priming with random hexanucleotides, using the protocol given by the manufacturer (Multiprime DNA Labelling Systems; Amersham, Chicago, Ill.), and added to a Hybridization solution that contained 6× SSC, 10× Denhardt's solution, (0.1% SDS, 50 mg of sodium PP_i and 100 μg of bakers' yeast tRNA (Sigma) per ml. The labeled DNA probe was hybridized to the digested genomic DNAs overnight at 65°C by procedures previously described (20). After hybridization, the membrane was washed twice at room temperature for 15 min with 6× SSC–0.1% SDS, twice at 37°C for 15 min in 1× SSC–0.1% SDS, and once for 60 min in 1× SSC–0.1% SDS. After being washed the membrane was blotted dry, exposed to XAR-5 film (Eastman Kodak) at −76°C, and developed.

DNA sequence determination and amino acid sequence analysis. The nucleotide sequence of the pilin gene was determined by dideoxy-chain termination (25), using T_7 polymerase (Sequenase; United States Biochemical Corp., Cleveland, Ohio). The majority of the gene sequence was determined by analysis of pPhb-1, which encodes the 5′ end of the pilin gene through nucleotide 552. Both strands of DNA were sequenced by using a primer complementary to the vector DNA of pPhb-1 and then a series of primers based on the determined sequence. The sequence of the extreme 3′ end of the gene was determined by sequencing portions of pBX10.5 with oligonucleotide primers based on the sequence of the insert DNA of pPhb-1.

Analysis of pilin hydrophilicity. The hydrophilicity of the amino acid sequences of pilins from Hib strains Eagan (p^+), M43 (p^+) (10, 16), and 770235 b^0f^+ (30) were analyzed by using the algorithm of Hopp and Woods (13), using a "window" of six amino acids.

Nucleotide sequence accession number. The nucleotide sequence reported here has been submitted to GenBank and assigned accession number M64334.

```
    TGC AAT TTT TGC AAA AAG GGC TTT AAT TAT ATA TAT ATA TAT ATA TAA TGG AAG GGT TTT

    TTA CTT TTT TGG GAA AAC AAA TCT TGC TGT TTA TTA AGG CTT TGG CAT TTT GAT AAA TGG

    TTT ATA TTC ATC TTT TGC CTT ATT TAT AAG GCG CAA ACC TCT TTT ATG GAG CAA TTT ATT

1   ATG AAA AAA ACA CTA CTT GGT AGC TTA ATT TTA TTG GCA TTT GCA GGG AAT GTG CAG GCG
    met lys lys thr leu leu gly ser leu ile leu leu ala phe ala gly asn val gln ala

61  GCA GCT AAT GCT GAT ACA AAA GGG ACT GTT ACT TTC TTT GGT AAG GTT GTT GAG AAT ACT
    ala ala asn ala asp thr lys gly thr val thr phe phe gly lys val val glu asn thr

121 TGT CAA GTA AAA ACA GAT CAT AAA AAT CGT AGT GTA GTG TTA AAT GAT GTC GGT AAA AAT
    cys gln val lys thr asp his lys asn leu ser val val leu asn asp val gly lys asn

181 AGT TTA AAA GAT AAA GGA AAT ACT GCT ATG CCA ACG CCA TTT ACT ATC ACA TTA CAA AAT
    ser leu lys asp lys gly asn thr ala met pro thr pro phe thr ile thr leu gln asn

241 TGT AAC CTA ACT GCG GCA AAT AGT TCT ACA AAT AAA GCT AAT AAA GTT GGG CTA TAT TTC
    cys asn leu thr ala ala asn ser ser thr asn lys ala asn lys val gly leu tyr phe

301 TAT TCT TGG GAG AAT GCG GAT AAA GAA AAT AAC TTT ACA TTG AAA AAT AAA ACA TCA ACT
    tyr ser trp glu asn ala asp lys glu asn asn phe thr leu lys asn lys thr ser thr

361 AGT AAT GAT TTT GCG ACC ATG GTT AAT ATT CAA CTT ATG GAA AGT GAT GGT ACA AAG GAA
    ser asn asp phe ala thr met val asn ile gln leu met glu ser asp gly thr lys glu

421 ATT AAA GTA GTA GGT AAA GAA ACA GAA GAT TTT GTT CAT AAG AAC GCT ACA GGA GCA GGA
    ile lys val val gly lys glu thr glu asp phe val his lys asn ala thr gly ala gly

481 GTA GCA TTA ACA CAA ACC CAC CCA GAT AAT GAT CAT ATT TCA GGA AGT ACA CAA TTA ACC
    val ala leu thr gln thr his pro asp asn asp his ile ser gly ser thr gln leu thr

                ▼
541 GGT GTT ACT GGA GAT CTT CCT CTC CAC TTT ATC GCC CAA TAT TAC TCA CTA GGT TCA ACA
    gly val thr gly asp leu pro leu his phe ile ala gln tyr tyr ser leu gly ser thr

601 ACC GCT GGT AAA GTA CAA TCC TCA GTT GAT TTC CAA ATT GCT TAC GAA TAA TTC CTA ATG
    thr ala gly lys val gln ser ser val asp phe gln ile ala tyr glu END

    TAA TGC AAG AGA AAC TAA GTT TAT GTG TTT GTT ATA TCT GCG TGG GCT
```

FIG. 1. DNA and translated amino acid sequences of the pilin gene from Hib strain Eagan (p^+). The inverted triangle indicates the *Bgl*II site used in the construction of pPhb-1.

RESULTS AND DISCUSSION

Cloning of the Hib pilin gene. The structural gene that encodes Hib pilin was cloned from a library of Hib chromosomal DNA in λGEM-11 by using two degenerate oligonucleotide probes. The probes used correspond to amino acids 83 to 89 (probe 1) and 17 to 23 (probe 2) of pilin from Hib strain Eagan (1). Among the ~5,000 plaques screened, two clones were found to hybridize to both probes, and one of these clones, λGEM-11-A1, was chosen for further study. A 1.1-kb *Bgl*II-*Xho*I fragment from λGEM-11-A1 was subcloned into pBKSM(−) that had been digested with *Xho*I and *Bam*HI to yield pPhb-1. Southern blot analysis of pPhb-1 and λGEM-11-A1 showed that both contained a 1.1-kb DNA fragment that hybridized to probes 1 and 2 (data not shown), suggesting that pPhb-1 contained at least a portion of the structural gene for Hib pilin

DNA sequence of the Hib pilin gene. A portion of the Hib DNA sequence of the insert carried by pPhb-1 was determined by using primers that were complementary to the vector DNA. Analysis of the sequence obtained showed that pPhb-1 encoded the 5′ end of the pilin gene through nucleotide 552 (a cleavage site for *Bgl*II) but not the entire pilin gene (Fig. 1). The entire pilin gene and flanking DNA that might be required for assembly and expression of Hib pili was obtained by subcloning a 10.5-kb *Xho*I-*Bam*HI fragment of Hib DNA from pXCosXV-12 into pGEM7 to yield plasmid pBX10.5 (See Materials and Methods). A series of synthetic oligonucleotide primers were used to progressively determine the sequence of overlapping regions of both strands of the insert DNA that encode the Hib pilin gene. The DNA and the translated amino acid sequences of the Hib pilin gene are shown in Fig. 1. The 648-bp open reading frame encodes a hydrophobic 20-amino-acid leader sequence followed by the 196 amino acids found in mature pilin. The calculated molecular mass of mature pilin was 21,152 Da, which was in reasonable agreement with that estimated by SDS-polyacrylamide gel electrophoresis of purified pilin (23.5 kDa [1]). The amino acid sequence of the translated gene completely agreed with the amino acid sequence of purified pilin that had been previously determined (1).

Homology of Hib pilins. The amino acid sequence of pilin from Hib strain Eagan was compared to those of pilins of Hib strains M43 (p^+) and 770235 b^0f^+, which have been previously reported (10, 16; Fig. 2). The mature forms of the proteins were similar in size. The pilin from strain Eagan (p^+) had 196 residues (M_r = 21,152), and strain M43 (p^+) had 193 residues (M_r = 21,101). Both proteins had only two cysteine residues; they were located at positions 21 and 61 of the mature proteins. The proteins were 78% identical but had 95% overall homology when conservative amino acid substitutions were considered and had identical 20-amino-acid

```
M43        MKKTLLGSLILLAFAGNVQADINTETSGKVTFFGKVVENT
CONSENSUS  MKKTLLGSLILLAFAGNVQA. N::T.G.VTFFGKVVENT
EAGAN      MKKTLLGSLILLAFAGNVQAAANADTKGTVTFFGKVVENT
           -20

M43        CKVKTEHKNLSVVLNDVGKNSLSTKVNTAMPTPFTITLQN
CONSENSUS  C:VKT:HKNLSVVLNDVGKNSL..K NTAMPTPFTITLQN
EAGAN      CQVKTDHKNLSVVLNDVGKNSLKDKGNTAMPTPFTITLQN
           +21

M43        CDPTTANGTANKANKVGLYFYSWKNVDKENNFTLKNEQTT
CONSENSUS  C: T:AN:::NKANKVGLYFYSW.N.DKENNFTLKN. :T
EAGAN      CNLTAANSSTNKANKVGLYFYSWENADKENNFTLKNKTST
           +61

M43        A-DYATNVNIQLMESNGTKAISVVGKETEDFMHTN--NNG
CONSENSUS  : D:AT VNIQLMES GTK.I.VVGKETEDF:H.N  ..G
EAGAN      SNDFATMVNIQLMESDGTKEIKVVGKETEDFVHKNATGAG
           +101

M43        VALNQTHPNNAHISGSTQLTTGTNELPLHFIAQYYATNKA
CONSENSUS  VAL.QTHP:N.HISGSTQLT. T.:LPLHFIAQYY::..:
EAGAN      VALTQTHPDNDHISGSTQLTGVTGDLPLHFIAQYYSLGST
           +141

M43        TAGKVQSSVDFQIAYE
CONSENSUS  TAGKVQSSVDFQIAYE
EAGAN      TAGKVQSSVDFQIAYE
           +181
```

FIG. 2. Amino acid sequence homology of the pilin genes of Hib strains Eagan (p^+) and M43 (p^+). Gaps within the consensus sequence indicate differences between the sequences, whereas dots and double dots represent conserved substitutions. Highly conserved hydrophilic regions of the sequences are underlined.

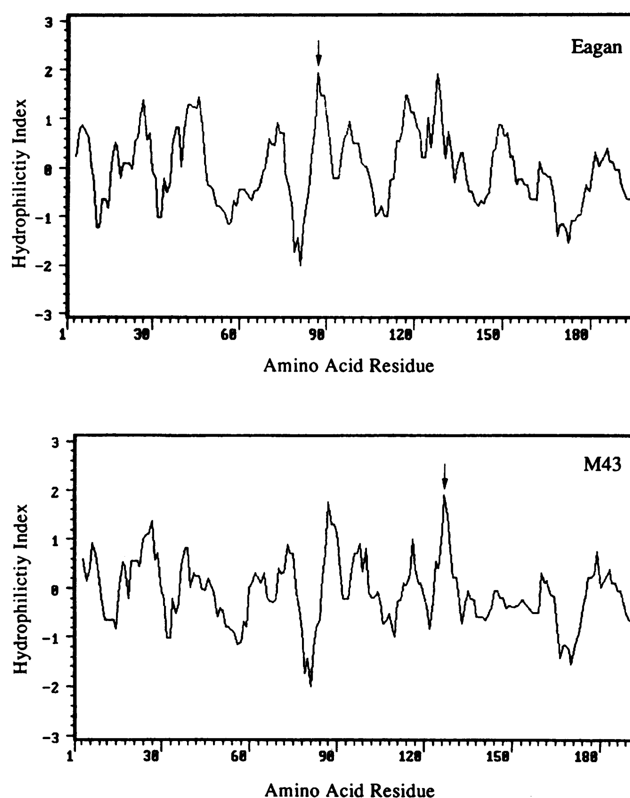

FIG. 3. Hydrophilicity plots of pilins from Hib strains Eagan (p^+) and M43 (p^+) determined by using the algorithm of Hopp and Woods (13). Arrows indicate regions of greatest hydrophilicity.

leader sequences. Differences in amino acid sequence were also reflected in the DNA sequences of the genes, which showed 84% homology. These data suggest that Hib pilins are highly conserved among different strains of Hib but that small numbers of differences in amino acid sequence account for the observed antigenic differences. Since the differences between the amino acid sequences were not localized to any particular region but rather were scattered throughout the proteins, the conformation-dependent epitopes that differ between the pili of Hib strain Eagan (p^+) and strains M43 (p^+) (11) and 770235 b^0f^+ may comprise amino acids from different regions of the proteins.

Analysis of pilin hydrophilicity. Differences in the amino acid sequences of pilins of Hib strains Eagan (p^+) and M43 (p^+) were expected since previously studies have shown that polyclonal antisera specific for the pili of these strains fail to bind to any significant extent to the native pili of the heterologous strain (11). This was consistent with the conclusion from previous studies that showed that the immunodominant epitopes of Hib pili were defined by the tertiary and quaternary structure of the organelle (11). Nonetheless, it may be possible to identify conserved immunogenic epitopes of Hib pili through analysis of the primary structures of the proteins. Several strategies for the prediction of antigenic epitopes within the primary sequences of proteins have been proposed (6, 13, 15). Using these methods, investigators have shown that peptides corresponding to hydrophilic regions of bacterial adhesins can elicit antibodies that cross-react with native proteins (4, 14, 17). Thus, it was of interest to identify conserved regions of pilins from Hib strains Eagan (p^+) and M43 (p^+) that were hydrophilic and therefore likely to be on the surface of the native protein. The algorithm described by Hopp and Woods (13) calculates local hydrophilicity in the primary sequence of proteins by averaging hydrophilicity values assigned to the different amino acids. The point of highest local hydrophilicity has been shown to be located in or adjacent to a known antigenic domain (13). As expected from the high degree of amino acid sequence conservation among Hib pilins, the hydrophilicity profiles were very similar (Fig. 3). Three hydrophilic domains common to the pilins of Hib strains Eagan (p^+) and M43 (p^+) were identified. These domains were between (i) residues 15 and 33, (ii) residues 79 and 96, and (iii) residues 120 and 131. Of these domains, residues 79 to 96 and 120 to 131 might constitute conserved Hib-specific epitopes. The other domain (residues 15 to 33) lies within a region that has extensive homology with *E. coli* pilins, including F17, type 1C, and members of the P-pilus family, including PapA(2), F7, $F7_2$, and F11, as well as with the major subunit proteins of type MR/K of *Klebsiella pneumoniae*, type US46 of *Serratia marcesens*, and serotype 2 of *Bordetella pertussis* (10) and thus is not specific to Hib. Studies are in progress to determine if antibodies raised against synthetic peptides corresponding to domains i and ii bind to native Hib pili from various strains.

Southern blot analysis of heterologous chromosomal DNA. Southern blot analyses of chromosomal DNA from heterologous strains of *H. influenzae* were done to determine (i) if the pilin gene was conserved among heterologous strains of *H. influenzae* and (ii) the number of gene copies in each

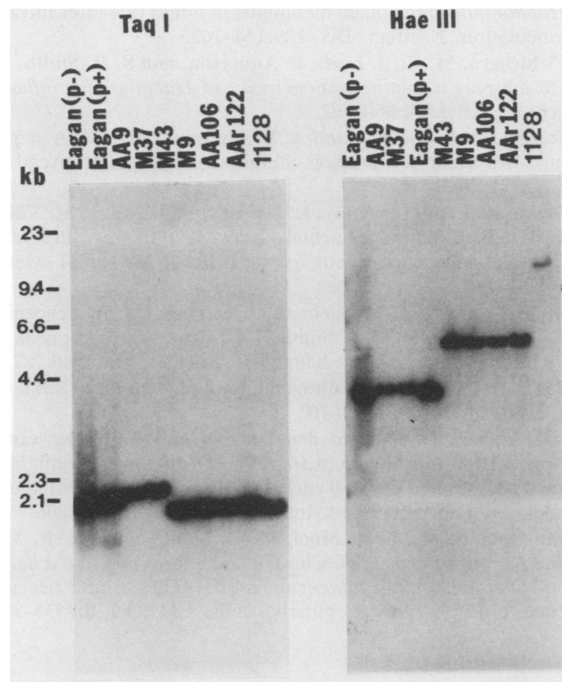

FIG. 4. Southern blot analyses of chromosomal DNA from Hib and NT strains digested with *Taq*I or *Hae*III, using an intragenic probe from the pilin gene of strain Eagan (p^+). Strains Eagan, AA9, M43, M9, AA106, AAr122 are serotype b, whereas M37 and 1128 are NT.

chromosome. Chromosomal DNAs from five type b and two NT strains of *H. influenzae* were digested with either *Taq*I or *Hae*III and analyzed by Southern blot by using a 379-bp *Dra*I-*Xba*I fragment from within the pilin gene as a probe. The chromosomal digests of all strains tested had a single *Taq*I or *Hae*III fragment that hybridized to the probe (Fig. 4). Restriction fragment length polymorphisms were observed among the strains tested. The pilin gene was found on 2.0-kb *Taq*I and 3.5-kb *Hae*III fragments from the chromosomes of Hib strains Eagan and AA9 and NT *H. influenzae* M37, whereas the gene was on 2.1-kb *Taq*I and 5.5-kb *Hae*III fragments from the chromosomes of Hib strains M43, M9, AA106, and AAr122. NT *H. influenzae* 1128 was the only strain in a third group in which the pilin gene was on a 2.0-kb *Taq*I fragment and a 12-kb *hae*III fragment. This finding suggests that pilins from various type b and NT strains appear to be highly conserved and are present as a single copy in the chromosome. Interestingly, *P. multocida* (five strains tested), *P. haemolytica* (two strains tested), and *A. pleuropneumoniae* (one strain tested), which are taxonomically related species that cause upper respiratory tract infections in other animals, did not have DNA that hybridized to the Hib pilin gene under the conditions used (data not shown). We also found that the failure of Eagan (p^-) to express pili was not due to loss of the pilin gene from the genome and was not associated with a substantial rearrangement of DNA (Fig. 4). These results confirm and extend those of Langermann and Wright (16), who showed that *H. influenzae* serotypes a, b, c, d, and f carried DNA sequences with homology to the pilin gene of Hib strain AO2 [M43 (p^+)] and that nonpiliated cells retain the pilin gene in the chromosome.

The results from the present study indicate that the signficant immunological differences between the pili of Eagan (p^+) and those of M43 (p^+) and 770235 b^0f^+ exist despite relatively few differences in the amino acid sequences. Furthermore, the pilin gene is largely conserved among different strains of *H. influenzae* but is not present in bacteria that cause respiratory tract infections in other animal species yet are phylogenetically related to *H. influenzae*. This suggests that the pili of all *H. influenzae* strains may be functionally similar and may have a role in determining the host specificity of the organism.

ACKNOWLEDGMENTS

We thank Jeffrey Bray, Meredith Hackel, and Donna Kersey for their technical assistance.

REFERENCES

1. **Armes, L. G., and L. J. Forney.** 1990. The complete primary structure of pilin from *Haemophilus influenzae* type b strain Eagan. J. Protein Chem. **9**:45–52.
2. **Bagdasarian, M. M., E. Amann, R. Lurz, B. Ruckert, and M. Bagdasarian.** 1983. Activity of the hybrid trp-lac (tac) promoter of *Escherichia coli* in *Pseudomonas putida*. Construction of broad-host-range, controlled-expression vectors. Gene **26**:273–282.
3. **Beachey, E. H.** 1981. Bacterial adherence: adhesion-receptor interactions mediating the attachment of bacteria to mucosal surfaces. J. Infect. Dis. **143**:325–345.
4. **Bessen, D., and V. A. Fischetti.** 1988. Influence of intranasal immunization with synthetic peptides corresponding to conserved epitopes of M protein on mucosal colonization by group A streptococci. Infect. Immun. **56**:2666–2672.
5. **Brinton, C. C., Jr., M. J. Carter, D. B. Derber, S. Kar, J. A. Kramarik, A. C.-C. To, S. C.-M. To, and S. W. Wood.** 1989. Design and development of pilus vaccines for *Haemophilus influenzae* diseases. Pediatr. Infect. Dis. J. **8**:554–561.

5a. **Capage, M.** Unpublished data.

6. **Chou, P. Y., and G. D. Fasman.** 1978. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. **47**:45–148.
7. **Christensen, G. D., W. A. Simpson, and E. H. Beachey.** 1985. Microbial adherence in infection, p. 6–23. *In* G. L. Mandell, R. Douglas, Jr., and J. E. Bennet (ed.), Principles and practice of infectious disease, 2nd ed. John Wiley & Sons, Inc., New York.
8. **Connor, E. M., and M. R. Loeb.** 1983. A hemadsorption method for detection of *Haemophilus influenzae* type b expressing fimbriae. J. Infect. Dis. **148**:855–860.
9. **Frey, J., M. Bagdasarian, D. Feiss, F. C. H. Franklin, and J. Deshusses.** 1983. Stable cosmid vectors that enable the introduction of cloned fragments into a wide range of Gram-negative bacteria. Gene **24**:299–308.
10. **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.
11. **Gilsdorf, J. R., K. W. McCrea, and L. J. Forney.** 1990. Conserved and nonconserved epitopes among *Haemophilus influenzae* type b pili. Infect. Immun. **58**:2252–2257.
12. **Guerina, N. G., S. Langermann, G. K. Schoolnik, T. W. Kesslor, and D. A. Goldmann.** 1985. Purification and characterization of *Haemophilus influenzae* pili, and their structural and serological relatedness to *Escherichia coli* P and mannose-sensitive pili. J. Exp. Med. **161**:145–159.
13. **Hopp, D., and K. R. Woods.** 1981. Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA **78**:3824–3828.
14. **Jacobs, A. A. C., J. Venema, R. Leeven, H. van Pelt-Heerschap, and F. K. de Graaf.** 1987. Inhibition of adhesive activity of K88 fibrillae by peptides derived from the K88 adhesin. J. Bacteriol. **169**:735–741.
15. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol.

157:105–109.

16. **Langermann, S., and A. Wright.** 1990. Molecular analysis of the *Haemophilus influenzae* type b pilin gene. Mol. Microbiol. **4:**221–230.
17. **Lee, K. K., W. Paranchych, and R. S. Hodges.** 1990. Cross-reactive and strain-specific antipeptide antibodies to *Pseudomonas aerugenosa* PAK and PAO pili. Infect. Immun. **58:**2727–2732.
18. **LiPuma, J. J., and J. R. Gilsdorf.** 1988. Strutural and serological relatedness of *Haemophilus influenzae* type b pili. Infect. Immun. **56:**1051–1056.
19. **Loeb, M. R., E. Connor, and D. Penney.** 1988. A comparison of the adherence of fimbriated and nonfimbriated *Haemophilus influenzae* type b to human adenoids in organ culture. Infect. Immun. **56:**484–489.
20. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. **Miller, H.** 1987. Practical aspects of preparing phage and plasmid DNA: growth, maintenance and storage of bacteria and bacteriophage. Methods Enzymol. **152:**145–170.
22. **Moxon, E. R.** 1985. *Haemophilus influenzae*, p. 1274–1279. *In* G. L. Mandell, R. Douglas, Jr., and J. E. Bennet (ed.), Principles and practice of infectious disease, 2nd ed. John Wiley & Sons, Inc., New York.
23. **Moxon, E. R., A. L. Smith, D. R. Averill, and D. H. Smith.** 1974. *Haemophilus influenzae* meningitis in infant rats after intranasal inoculation. J. Infect. Dis. **129:**154–162.
24. **Pichichero, M. E., J. Loeb, P. Anderson, and S. H. Smith.** 1982. Do pili play a role in pathogenicity of *Haemophilus influenzae* type b? Lancet **ii:**960–962.
25. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
26. **Smith, A. L., P. D. Penny, L. Langley, P. Verdugo, M. Villalon, J. Unadkat, and D. Luchtel.** 1989. Kinetic measurement of bacterial adherence to eukaryotic cells. J. Microbiol. Methods **10:**265–280.
27. **Stull, T. L., P. M. Mendelman, J. E. Haas, M. A. Schoenborn, K. D. Mack, and A. L. Smith.** 1984. Characterization of *Haemophilus influenzae* type b fimbriae. Infect. Immun. **46:**787–796.
28. **Turk, D. C.** 1984. The pathogenicity of *Haemophilus influenzae*. J. Med. Microbiol. **18:**1–16.
29. **van Alphen, L., N. van den Berghe, and L. Geelen-van den Broek.** 1988. Interaction of *Haemophilus influenzae* with human erythrocytes and oropharyngeal epithelial cells is mediated by a common fimbrial epitope. Infect. Immun. **56:**1800–1806.
30. **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.