DNA Sequence Analysis and Restriction Fragment Length Polymorphisms of the P1 Gene of *Haemophilus influenzae* Biogroup aegyptius Associated with Brazilian Purpuric Fever

ROBYN B. REED,¹ JEFF B. FROST,² KRISTY KORT,² SHEILA D. MYERS,² AND ALAN J. LESSE^{1,2*}

*Departments of Medicine and Pharmacology and Toxicology, State University of New York at Buffalo,¹ and
Department of Veterans Affairs Medical Center,² Buffalo, New York*

Received 11 March 1996/Returned for modification 18 April 1996/Accepted 13 June 1996

Brazilian purpuric fever (BPF) is a fulminant pediatric disease caused by specific strains of *Haemophilus influenzae* biogroup aegyptius. A conserved epitope on the P1 protein of strains of *H. influenzae* biogroup aegyptius is seen on most virulent isolates. The P1 protein from a Brazilian case-clone strain of *H. influenzae* biogroup aegyptius was analyzed by cloning and sequencing the gene. Three major variable regions are present within the P1 gene of the BPF clone in an architecture similar to that of the previously sequenced P1 genes from *H. influenzae*. The DNA sequence data of the P1 gene provided information for restriction fragment length polymorphism analyses among strains of *H. influenzae* biogroup aegyptius. Using PCR for amplification of the P1 gene, we found that *AlwI* restriction of this gene allowed for a highly accurate segregation of virulent strains of *H. influenzae* biogroup aegyptius associated with BPF. The strong association of virulent phenotypes with specific *AlwI* restriction patterns of the P1 gene provides a basis for the convenient and accurate identification of strains of *H. influenzae* biogroup aegyptius which cause BPF.

Brazilian purpuric fever (BPF) is a fatal childhood illness which was first recognized in a southern state of São Paulo, Brazil, in 1984 (4, 5, 10). The initial stage of this illness is characterized by conjunctivitis, which is followed by high fever, hemorrhagic skin lesions, vascular collapse, and death. *Haemophilus influenzae* biogroup aegyptius (*Haemophilus aegyptius*), the etiologic agent of BPF, is an unencapsulated, nontypeable *H. influenzae* biogroup. The initial outbreak in Brazil was caused by a clonal strain of *H. influenzae* biogroup aegyptius with several unique genotypic and phenotypic markers, including a unique 24-MDa plasmid and a specific electrophoretic typing (ET) pattern. These strains are referred to as the Brazilian case-clone strains (7, 9). Subsequent outbreaks from Mato Grosso were also caused by this clone (6, 11). Recent outbreaks in Valparaíso have been caused by isolates with a similar ET pattern but which lack the plasmid (34).

Previous work in this laboratory has shown that a specific epitope, recognized by the monoclonal antibody (MAb) 8G3, is present in over 95% of Brazilian case-clone strains (21). This antibody recognizes an epitope on the 48-kDa, heat-modifiable P1 protein located on the outer membrane of *H. influenzae* (27). Given the strong association of the 8G3 epitope with virulent strains, we chose to analyze further the P1 genes of these strains.

To accomplish this, the P1 gene from *H. influenzae* biogroup aegyptius strain 16 was cloned and sequenced. In addition to database comparisons and comparisons among the previously sequenced P1 genes, restriction fragment length polymorphism (RFLP) analyses using PCR and restriction enzyme digestions of strains of *H. influenzae* biogroup aegyptius were performed in an attempt to correlate strains on the basis of virulence and RFLP patterns.

(This work was presented in part at the 94th General Meeting of the American Society for Microbiology, 1994.)

MATERIALS AND METHODS

Bacterial strains. The Epidemic Investigations Laboratory, Centers for Disease Control and Prevention, provided strains of *H. influenzae* biogroup aegyptius. A complete list of the strains tested is given in Table 1. Bacteria were grown on chocolate agar at 37°C in 5% CO₂. The strains were stored in skim milk at –85°C.

Description of strains. The Brazilian case-clone strains and other *H. influenzae* biogroup aegyptius strains were described by Brenner et al. (7). The Australian strains were described by Mayer et al. (24), Wild et al. (39), and McIntyre et al. (25). The Mato Grosso isolates were characterized by the BPF Study Group (6, 11), and the Pradópolis strains were described by Mayer et al. (24). The Valparaíso strains were described by Tondella et al. (34).

Classification of virulent and avirulent strains. Table 1 highlights some shared features of these strains. (See below for a description of *AlwI* digests.) The Brazilian case-clone strains (including Mato Grosso strains) are virulent by virtue of their recovery from sterile sites such as blood. They possess the unique 24-MDa plasmid of plasmid restriction endonuclease type 3031, have ET-2 and sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) protein pattern 3031, and react with a MAb to a 25-kDa pilin protein (7, 37).

The isolates from Australia are virulent by virtue of being bloodstream isolates. These isolates have ET-30 and SDS-PAGE type 1947, generate unique ribosomal DNA (rDNA) patterns, do not have a plasmid, and fail to react with either the MAb to the pilin protein or MAb 8G3 (24).

The Valparaíso isolates are clearly virulent as documented by recovery from the bloodstream. These isolates have an ET pattern similar to that of Brazilian case-clone isolates but lack the 24-MDa plasmid, and they have not incorporated the plasmid into the chromosomal DNA (34).

The Pradópolis isolates are questionably virulent. They have been shown to produce a sustained bacteremia in an animal model (30), but no confirmed cases of BPF caused by these strains have been reported (24, 32). Pradópolis strains lack the plasmid seen in the Brazilian case-clone strains but have ET-2, SDS-PAGE pattern 2, and rDNA pattern 4 as seen in the Brazilian case-clone strains (24).

A group of eight isolates from nonsterile sites, predominantly the conjunctiva, are classified as in vitro false positives, since these strains invade and cause a cytopathic effect in the human microvascular endothelial cell (HMEC) tissue culture model (29, 38).

Avirulent *H. influenzae* biogroup aegyptius strains were never recovered from sterile body sites and differ from case-clone strains in having various ETs, rDNA restriction patterns, and SDS-PAGE protein patterns (7).

Strains 1479, 2019, 3198, 5657, and 7502 are representative nontypeable *H. influenzae* strains, are not sterile-site isolates (avirulent), and are not *H. influenzae* biogroup aegyptius. Of note, strain 2019 is, however, the only non-BPF strain

* Corresponding author. Mailing address: Division of Infectious Diseases, Buffalo VAMC, 151, 3495 Bailey Ave., Buffalo, NY 14215. Phone: (716) 862-3304. Fax: (716) 862-3419. Electronic mail address: camlesse@acsu.buffalo.edu.

TABLE 1. Strain classification and phenotypic and genotypic strain descriptions

Strain classification	Phenotype	Sustained bacteremia in infant rats ^a	Cytopathic in tissue culture ^b	No. with 8G3 reactivity/no. tested ^c	No. restricted with <i>AlwI</i> /no. tested ^{d,e}	No. with <i>AlwI</i> restriction pattern ^f :			
						AA	AB	AC	AD
Brazilian case-clone strains ^f	ET-2, 24-MDa plasmid	Yes	Yes	18/21	20/21	19	0	1 ^g	1 ^h
Australian strains ⁱ	ET-30, no plasmid	Yes	Yes	0/2	2/2	0	0	2	0
Valparaiso strains ^j	ET-2, no plasmid	Yes	Yes	0/5	5/5	0	0	5	0
Pradopolis strains ^k	ET-2, no plasmid	Yes	No	0/7	0/7	0	0	0	7
2019	MAB 8G3 ⁺ , biotype II	Yes	NT ^l	1/1	1/1	0	1	0	0
Strains false positive in tissue culture model ^m	Varied	NT	Yes	0/8	6/8	0	0	6	2
Conjunctival isolates ⁿ	Varied	No	No	0/13	0/13	0	0	0	13

^a Data are from reference 30.^b Data are from reference 38.^c Data are from reference 21.^d Data are from this study.^e The PCR-amplified P1 gene product has at least one *AlwI* site.^f Strains 7, 10, 14, 15, 16, 17, 18, 24, 34, 35, 71, 76, 77, 78, 126, 130, 148, 156, 413, 415, and 416.^g Strain 15.^h Strain 14.ⁱ Strains 171 and 349.^j Strains 259, 260, 261, 263, and 264.^k Strains 194, 195, 196, 197, 200, 201, and 203.^l NT, not tested.^m Strains 54, 57, 84, 88, 102, 103, 104, and 107.ⁿ Strains 3, 4, 5, 6, 8, 30, 37, 38, 56, 65, 66, 83, and 94.

to react with the MAB 8G3 and the only non-BPF strain to produce a sustained bacteremia in the infant rat model (21). Serotype a to f encapsulated strains were also tested.

PCR-based cloning of the P1 gene. Genomic DNA from *H. influenzae* biogroup aegyptius strain 16 was purified by the cetyltrimethylammonium bromide method (40). The P1 gene was amplified from the genomic DNA by using primers based on the published sequence information for the three P1 genes of *H. influenzae* type b (28). The N-terminal primer was 5'-CTCCGGATCCCCA TGGCAGCGTTTCAATTGGCG-3', and the C-terminal primer was 5'-ATTC GAATTCATTAGAACTATAATT-3'. *Bam*HI and *Eco*RI sites were added to the primers (underlined bases) to facilitate the subsequent ligation into the plasmid. The initial PCR cycles were less stringent to allow for base pair mismatching from the published sequences. The PCR conditions were as follows: cycles 1 to 5, 95°C for 1.5 min, 37°C for 2 min, and 50°C for 3 min; cycles 6 to 30, 95°C for 1.0 min, 50°C for 1 min, and 72°C for 2 min; and cycle 31, 95°C for 1.5 min, 50°C for 1 min, and 72°C for 10 min. *Taq* polymerase (Promega, Madison, Wis.) was used for amplification. The PCR product was electrophoresed on a 1.5% agarose gel, and the product was purified from the agarose by using a commercial kit (GeneClean; Bio 101, La Jolla, Calif.).

Restriction enzymes, ligases, and polymerases were used according to the manufacturer's recommendations.

Sequencing. Sequencing was done manually by using the nonisotopic Sequenase sequencing system (United States Biochemicals, Cleveland, Ohio), and automated sequence analysis was performed with the Applied Biosystems model 373 apparatus. Gel compressions were resolved by using inosine manual sequencing according to the manufacturer's recommendations. Sequence "walking" was performed with primers constructed from the sequence data. The automated sequencing and primer synthesis were performed at the DNA Sequencing and Oligonucleotide Synthesis Facility at the State University of New York at Buffalo.

Since the clones sequenced contained PCR-derived inserts, one clone, pJBF3B, was sequenced on both strands manually and the sequence was confirmed by automated sequencing of one strand of another clone, pJBF4C. When sequence discrepancies were detected, manual sequencing gels were run with plasmids pJBF3B, pJBF4C, and pJBF2B and the appropriate primer. The majority base at each of the five discrepancies was recorded and identified in the database with the submission. It should be noted that the sequences of the first 18 nucleotides and the last 14 nucleotides represent primer sequence because of the primer excess in the initial PCR.

P1 PCR-RFLP. (i) Primer selection. The P1 gene was amplified directly from bacterial colonies by PCR. Primers were synthesized after a careful inspection of the sequence data from the *H. influenzae* biogroup aegyptius strain 16 P1 gene and the P1 genes of *H. influenzae* Minn A, 1613, and 8358 (28). Regions which were conserved in all three type b P1 genes and the strain 16 P1 gene were considered. P1N RR was constructed from the sense strand at the initiation of the first amino acid of the mature peptide from bases 121 to 141 of the Minn A P1 gene as follows: 5'-GCAGCGTTTCAATTGGCGGA-3'. P1C RR was constructed 22 bases in from the carboxy terminus of the Minn A P1 gene. The P1C primer was bases 1410 to 1390 of the antisense strand: 5'-CCGTAAAGATT

CGGTGTGC-3'. The C-terminal primer was chosen to avoid single base pair substitutions as seen in isolate 8358 (28). The primers used were synthesized at the DNA Sequencing and Oligonucleotide Synthesis Facility at the State University of New York at Buffalo.

(ii) P1 RFLP-PCR amplification. Conditions were optimized by varying the pH and magnesium concentration by using the PCR Optimizer kit (Invitrogen, San Diego, Calif.). Optimization with *H. influenzae* biogroup aegyptius strain 16 found that a reaction buffer at pH 9.5 and a magnesium concentration of 1.5 mM were the ideal conditions for the amplification of the P1 gene (data not shown).

The DNA template was prepared by boiling three to five colonies of a bacterial strain in 100 µl of sterile distilled water for 10 min and then centrifuging for 10 min at 14,000 × g. Five microliters of the supernatant provided the template for a 50-µl PCR mixture. Deoxynucleoside triphosphates (Boehringer Mannheim Corporation, Indianapolis, Ind.) were used at a concentration of 2.5 mM. *Taq* DNA polymerase (Boehringer Mannheim and Promega) was used at a concentration of 1.25 U/50-µl reaction mixture. The primers were added at a concentration of 0.25 µg/µl. Each of the reaction mixtures was covered with 30 µl of Chill-out 7 liquid wax (MJ Research, Watertown, Mass.). The initial PCR denaturation step was 2 min at 94°C. The cycles consisted of 30 cycles of 94°C for 1 min, 55°C for 2 min, and 72°C for 3 min. The final extension was for 7 min at 72°C. The thermal reactors used were a Hybaid thermal reactor (Hybaid, Rochester, N.Y.) and a MiniCycler (MJ Research). The PCR products were precipitated and resuspended in sterile Tris-EDTA (pH 8.0).

(iii) Restriction enzyme digestions. Prior to restriction enzyme digestions, the size, purity, and approximate concentration of the PCR product were confirmed on a 0.7% agarose gel. The DNA was digested overnight to completion, and the products were electrophoresed on an HE 100 Supersub Horizontal Unit (Hoefler, San Francisco, Calif.) in a 2% MetaPhor agarose gel (FMC, Rockland, Maine) with 1× Tris-borate-EDTA as the running buffer. The DNA molecular size markers used were a 123-bp ladder (Gibco BRL, Grand Island, N.Y.). The gel was stained with ethidium bromide, visualized with UV light (300 nm), and photographed.

Images from the negatives were digitized with the Molecular Dynamics Personal Densitometer (Molecular Dynamics, Sunnyvale, Calif.). The resulting TIFF images were adjusted with NIH Image (WWW: <http://rsb.info.nih.gov/nihi-image/>) and labelled in Freehand (Macromedia, San Francisco, Calif.) on a Power Macintosh 7100/66 (Apple Computer, Cupertino, Calif.).

Computer analysis of DNA sequences. Sequence analysis was performed with the MacVector Program (Kodak/IBI) for sequence assembly and alignment. BLASTN, BLASTP, and TBLASTN searches (1) were performed via the internet at the National Center for Biotechnology Information. BLASTN and BLASTP searches compare nucleotide or protein sequences, respectively, with the known database, while TBLASTN searches compare a protein sequence with the translation of the nucleotide database in all six frames to search for homology to proteins not entered in the protein database. Multiple sequence alignments were constructed with the PILEUP program, and analyses for motifs were run with the MOTIF program from the Genetics Computer Group (14).

Nucleotide sequence accession number. The sequence data reported here have been submitted to GenBank (accession number HIU28808).

RESULTS

Cloning and expression of the P1 gene. The P1 gene from *H. influenzae* biogroup aegyptius strain 16 was amplified by using PCR with primers based on the known sequence of three P1 genes from *H. influenzae* type b (27, 28). The N-terminal primer was also constructed to insert the gene product out of frame for the LacZ fusion protein to allow for the possible cloning of a toxic gene. PCR amplification of the P1 gene resulted in an approximately 1,300-bp fragment. The purified product was then restricted with *Eco*RI and *Bam*HI and ligated into *Eco*RI-*Bam*HI vector pGEM7Zf⁺ after treatment with calf alkaline phosphatase. The resulting plasmid was introduced into competent *Escherichia coli* DH5 α and screened for white colonies after growth on LB-ampicillin with IPTG (isopropyl- β -D-thiogalactopyranoside) and X-Gal (5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside). Four independent clones were selected for further analysis and designated pJBF1, pJBF2, pJBF3, and pJBF4.

These plasmids were not reactive with MAb 8G3 on colony blotting, as the cloned P1 gene resulted in an out-of-frame LacZ fusion protein. Restriction of the plasmid with *Bam*HI, Klenow end filling, religation, and transformation into DH5 α placed the fragment in frame for a fusion protein with LacZ and resulted in plasmids that were reactive on colony lifts with MAb 8G3. This confirmed that the cloned fragment was responsible for 8G3 reactivity. Three individual clones from each of the plasmids were saved and designated pJBF1A, pJBF1B, and pJBF1C, etc.

Since the original DNA was of PCR origin, one clone, pJBF3B, was sequenced in both directions, and the sequence was confirmed on a single strand of pJBF4C. Conflicts were resolved by simultaneous manual sequencing of the ambiguous nucleotide sequences of the two plasmids along with another independent clone, pJBF2B. Five discordant bases, at positions 181, 411, 762, 793, and 1153, were resolved by majority analysis, and the frequency of 5 bases per 1,284 nucleotides is within the published error rates for *Taq* polymerase (8, 23, 33).

Sequence analysis. Figure 1 details the DNA sequence of the P1 gene of *H. influenzae* biogroup aegyptius strain 16 compared with those of the previously sequenced P1 genes. Three P1 genes from *H. influenzae* type b have been sequenced (27, 28), and a fourth P1 gene, from the nontypeable strain Rd, was sequenced in the *Haemophilus* genome project (13). Previous work by Munson et al. described three variable regions within the P1 genes of *H. influenzae* type b strains Minn A, 8358, and 1613, which represent the 1H, 3L, and 6U phenotypes, respectively (28). Three distinct variable portions were also noted in the sequence of *H. influenzae* biogroup aegyptius strain 16. These areas, designated V1, V2, and V3, are bp 187 to 228, 514 to 609, and 1189 to 1213, respectively. Specific restriction enzyme recognition sequences were found within these variable regions. The sequence data revealed a *Bsr*GI site in V1, an *Alw*I site in V2, and an *Hae*III site in V3. (Additional *Alw*I sites were found at bp 237 and 1043 in strain 16, as shown in Fig. 1.) Nonconserved nucleotide substitutions outside the variable regions that result in altered amino acid are shown in black in Fig. 1.

A comparison of the protein sequence with known protein sequences in the National Center for Biotechnology Information database showed significant homology between this protein and the previously sequenced P1 proteins of other *Haemophilus* strains. Homology was also seen with the FadL protein of *E. coli* (3) and the TodX and XylN proteins of *Pseudomonas putida* (14a, 15, 36). Lesser homology was seen with the OmpE protein of *Moraxella catarrhalis* (2). A

TBLASTN search of the nucleotide database did not reveal any additional protein homologies with previously untranslated DNA sequences.

No significant sequence homology could be found by TBLASTN analysis of the variable regions V1 and V3 of this protein, including the corresponding regions of the previously published P1 sequences. This suggests that these variable regions are significantly different from the previously sequenced P1 proteins and other proteins in the database. The V2 region of *H. influenzae* biogroup aegyptius strain 16 has significant homology to the previously published P1 sequences.

All of the sequenced P1 proteins contain a conserved ATP/GTP-binding site motif referred to as the P-loop (19, 31, 35). The motif, identified by the signature (A,G)X₄GK(S,T), is represented in these proteins as AG(I,V)FGGKT. The P-loop is not present in either FadL, TodX, or XylN.

PCR-RFLP. PCR amplification of P1 genes from all typeable and nontypeable strains of *H. influenzae*, including all *H. influenzae* biogroup aegyptius strains, demonstrated conservation of the P1 gene at both the carboxy terminus and the amino terminus. Amplification of the P1 gene of *Haemophilus paraphrohaemolyticus* 29237 was also successful; however, no detectable PCR product was generated when *Haemophilus parainfluenzae* 1365 and *Haemophilus segnis* 10977 were amplified (data not shown).

The *Alw*I enzyme digestions of several *H. influenzae* biogroup aegyptius strains (both disease associated and non-disease associated) and the encapsulated *H. influenzae* strains produced three separate patterns, as shown in Fig. 2A. The P1 genes from the Brazilian case-clone strains (except strains 14 and 15) produced a four-band pattern. The Valparaíso and Australian strains and Brazilian case-clone strain 15 generated a two-band pattern when restricted with *Alw*I. (Slight variations in the migration of the two-band pattern were seen among strain 15 and the Valparaíso and Australian isolates, presumably because of the differences in the molecular weights of the P1 products.) No *Alw*I restriction sites were seen in Pradopolis or non-case-clone *H. influenzae* biogroup aegyptius strains.

Figure 2B shows a gel summarizing the above-described three patterns designated AA, AC, and AD, respectively, and the remaining fourth pattern from strain 2019, a nontypeable, non-biogroup aegyptius strain, designated AB. These four patterns represented all of the *Alw*I restriction patterns seen with our analysis.

Since our original collection of strains did not contain the strains described by Weyant et al. as causing a false-positive reaction in the HMEC tissue culture model (38), we obtained these strains and analyzed them by PCR-RFLP. As shown in Table 2, six of the eight strains that were in vitro false-positive strains were restricted with *Alw*I despite their marked genotypic, phenotypic, and geographic differences.

Only one clearly pathogenic P1 gene product of a strain of *H. influenzae* biogroup aegyptius associated with BPF, that of strain 14, failed to be restricted with *Alw*I. All other virulent BPF isolates had at least one *Alw*I restriction site. All non-BPF strains which had at least one *Alw*I site behaved similarly to virulent isolates in vitro in either the animal model (strain 2019) or the HMEC tissue culture model (six of the eight false-positive isolates).

DISCUSSION

The molecular analysis of the P1 gene from a strain of *H. influenzae* biogroup aegyptius associated with BPF revealed a conserved architecture similar to that of previously sequenced

The P1 protein has been shown to be highly immunogenic following natural infection with type b strains in humans and is

the target for C3 binding in both type b and nontypeable *H. influenzae* (17, 18). Anti-P1 antibody has also been shown to be protective in animal models of *H. influenzae* type b (12, 22). Chong et al. have shown that several areas of the P1 protein are highly immunogenic (12). Conserved linear antigenic domains exist on the P1 protein in both type b and nontypeable *H. influenzae*; however, antibodies to these linear epitopes gen-

FIG. 2. (A) Composite of three gels showing the *AlwI* digests of various strains of *H. influenzae* biogroup *aegyptius* and encapsulated strains. Strain classifications (see text) are indicated. A 123-bp ladder is shown in the left lane of each gel. Brazilian case-clone strains from left to right on the first gel are strains 14, 15, 7, 10, 16, 17, 18, 24, 34, 35, 71, 76, 77, 78, 126, 130, 148, and 156. The second gel contains Brazilian case-clone strains 16, 413, 415, and 416. The Valparaíso strains are 259, 260, 261, 263, and 264. The Australian strains are 171 and 349. The Pradópolis strains are 194, 195, 196, 197, 200, 201, and 203. The third gel has the non-BPF strains 3, 4, 5, 6, 8, 30, 37, 38, 56, 65, 66, 83, and 94. The encapsulated strains are types a, b, c, d, e, and f. (B) Summary of the four P1 PCR-RFLP patterns resulting from *AlwI* digestion of PCR-amplified P1 genes. Patterns AA, AB, AC, and AD are shown along with the 123-bp ladder.

erated by immunization with peptides were unable to generate bactericidal activity and were unable to absorb bactericidal activity from a polyclonal antibody generated against purified P1 protein, suggesting that the protective antibody to P1 may be conformational in nature (12).

TABLE 2. Phenotypic characterization of the eight false-positive isolates in the HMEC tissue culture model^a

Strain	<i>AlwI</i> restriction pattern	ET ^b	Place of isolation ^b	rDNA ^{a,c}	Plasmid size (MDa) ^b
54	AC	12	Lins, Brazil	13	None
57	AD	15	Bauru, Brazil	9	None
84	AC	6	Presidente, Brazil	10	2
88	AD	18	Egypt	2	24 ^d
102	AC	23	Ribeirao, Brazil	10	2
103	AC	23	Ribeirao, Brazil	10	2
104	AC	23	Ribeirao, Brazil	10	2
107	AC	23	Ribeirao, Brazil	10	2

^a Data are from reference 38.

^b Data are from reference 7.

^c Ribosomal DNA (rDNA) restriction pattern.

^d This plasmid is not plasmid pattern 3031, which is specific to Brazilian case-clone isolates.

The specific epitope recognized by MAb 8G3 was not identified from these studies. We are currently evaluating whether the epitope recognized by MAb 8G3 is linear or conformational, but on the basis of the inability of the MAb to induce bactericidal activity or protect in the infant rat model (21), it is likely that the epitope is linear.

The function of the P1 protein has not been determined. Significant homology to the fatty acid transport protein FadL of *E. coli* on the protein sequence level has been noted, but a plasmid containing the P1 gene was not able to complement a FadL-deficient strain of *E. coli* (26). The homology to the TodX and XylN proteins of *P. putida* is interesting in that these proteins are thought to function in transporting the hydrophobic molecule toluene across the cell membrane. Since the homologs (intraspecies homologous genes) FadL, TodX, and XylN function in transport of hydrophobic molecules across the membrane, it is interesting to postulate an as-yet-undefined porin or transport function for P1.

The significance of the ATP/GTP-binding site motif, or P-loop, is questionable. Many bacterial proteins have this motif, and a sequence similarity analysis of *E. coli* proteins in the published databases showed this motif to be the second largest paralogous supercluster (20). However, this motif is neither 100% sensitive nor 100% specific for nucleoside triphosphate-

binding proteins, and a number of proteins which have this motif do not bind ATP or GTP (14). Additionally, a functional P-loop has not been described with bacterial outer membrane proteins, and the corresponding homologs of the P1 protein do not have this functional motif, making it unlikely that the identified motif has any functional significance.

Whatever the function of the P1 protein, it is interesting that the *AlwI* restriction pattern of the P1 gene so closely correlates with an *in vitro* measure of virulence, namely, the ability of these strains to invade the HMEC line (29, 38). This correlation is present despite the fact that not all strains causing a purpuric disease are clonal, as evidenced by the different ET patterns of the Brazilian case-clone and Australian strains and the false positives in the tissue culture assay. It may be that certain nonvirulent isolates possess the machinery to invade cells without the additional pathogenic machinery necessary to survive in the host and cause disseminated disease. Acquisition of these additional factors by the false-positive tissue culture isolates could then result in bacteria which could cause systemic disease.

What role the P1 protein might play in this hypothesis is currently speculative. Alternatively, this association may be the result of a genetic linkage of the P1 protein with another protein that is responsible for the virulent phenotype. If this was true, then the association of specific P1 PCR-RFLP patterns with virulent strains would be more of an epiphenomenon and not directly linked to the ability to cause disease.

H. influenzae biogroup *aegyptius* strain 14, *H. influenzae* 2019, and the six *in vitro* false positives in the HMEC tissue culture assay were exceptions to the general results of the *AlwI* analysis. Strain 14 is a virulent, bloodstream, Brazilian case-clone strain which had the only "virulent" P1 gene which was not restricted with *AlwI*. To confirm that a storage or labelling error did not account for the discrepant results, a second sample of strain 14 was obtained from the Centers for Disease Control and Prevention and behaved in a manner identical to that of the strain in our collection. Strain 14 is also one of three MAb 8G3-negative Brazilian case-clone strains.

H. influenzae 2019 produced a unique RFLP pattern (AB) when its P1 gene was restricted with *AlwI*. *H. influenzae* 2019 is the only MAb 8G3-positive, non-biogroup *aegyptius* *H. influenzae* strain tested and the only non-biogroup *aegyptius*, non-BPF strain to produce bacteremia in the infant rat model (21). While this strain was clearly virulent in the infant rat model, it has not been tested in the HMEC tissue culture model. Since the strain is biotype II, it is not a member of the *aegyptius* biogroup. The unique ability to produce bacteremia in the infant rat model, the reactivity with MAb 8G3, and the unique restriction pattern of the P1 gene in a non-biogroup *aegyptius*, non-BPF strain of *H. influenzae* are intriguing with regard to a potential role of P1 in virulence. Further work in identifying the 8G3 epitope and a P1-deficient mutant will be needed to address this question.

Interestingly, the only other non-BPF isolates which showed *AlwI* restriction of their P1 genes were six of eight isolates that were classified as biologically nonvirulent on the basis of the absence of invasive disease in humans. However, these isolates were able to attach to and invade human endothelial cells and behave in this biologic model as virulent (38). Despite the fact that all eight isolates were 8G3 negative, the single P1 gene restriction pattern of six of these strains paralleled the RFLP pattern of biologically virulent isolates. It is unclear why such genotypically diverse strains would share a common restriction pattern of a single gene and behave in a similar manner in a biologic *in vitro* assay if that gene played no role in the observed phenotype. While these data cannot be used to associ-

ate the P1 gene with this invasive phenotype, it is interesting to note the strong association of P1 PCR-RFLP patterns with the behavior in the model system. Further work with isogenic, P1-deficient strains will be necessary to answer these questions.

The P1 genes of the Pradopolis strains were not restricted in an *AlwI* digestion. These problematic isolates were obtained from the town of Pradopolis in Brazil in 1987 and 1988, when no BPF cases were confirmed (24, 32). One child exhibited fever, vomiting, and questionable perioral petechiae 15 days after conjunctivitis; however, the condition never progressed to clinical BPF (16). Pradopolis strains differ from BPF case-clone strains in that they lack a plasmid. While Pradopolis strains have demonstrated virulence in an infant rat model (30), Weyant et al. showed that cytotoxicity was not induced by these strains in the HMEC model (38). Thus, the PCR-RFLP data correlate strongly with the *in vitro* tissue culture virulence assay and group the Pradopolis strains with avirulent isolates.

Our current study does not define whether P1 is a virulence factor but rather provides a means of classifying the groups of strains on the basis of RFLP patterns of a PCR-amplified P1 gene. The results of this investigation can serve as a diagnostic tool for the rapid identification of geographic classes of strains of *H. influenzae* biogroup *aegyptius*. Brazilian case-clone strain 14 is the single exception and could lead to misclassification of a potentially virulent strain. Current screenings for this bacterium in Brazil include agglutination by polyclonal antisera and confirmation by labor-intensive techniques, including rDNA and electrophoretic typing. The PCR technique described in this study is a fast and accurate means of initial screening. It remains to be ascertained why a single outer membrane protein gene would have conservation of sequence that allows for virulence categorization.

ACKNOWLEDGMENTS

This work was supported by the Office of Research and Development, Department of Veterans Affairs, by National Institutes of Health training grant GM 07145-20 to R.B.R., and by a Veterans Administration Merit Review Grant to A.J.L.

We thank Timothy Murphy and George Carlone for their comments regarding the manuscript. We also thank the Centers for Disease Control and Prevention for providing the strain collection for analysis, and we acknowledge Eugene Koonin of the National Center for Biotechnology Information for discussions concerning the relevance of the P-loop motif.

REFERENCES

1. Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403-410.
2. Bhushan, R., R. Craigie, and T. F. Murphy. 1994. Molecular cloning and characterization of outer membrane protein E of *Moraxella (Branhamella) catarrhalis*. *J. Bacteriol.* **176**:6636-6643.
3. Black, P. N. 1991. Primary sequence of the *Escherichia coli* *fadL* gene encoding an outer membrane protein required for long-chain fatty acid transport. *J. Bacteriol.* **173**:435-442.
4. Brazilian Purpuric Fever Study Group. 1987. *Haemophilus aegyptius* bacteraemia in Brazilian purpuric fever. *Lancet* **ii**:761-763.
5. Brazilian Purpuric Fever Study Group. 1987. Brazilian purpuric fever: epidemic purpura fulminans associated with antecedent purulent conjunctivitis. *Lancet* **ii**:757-761.
6. Brazilian Purpuric Fever Study Group. 1992. Brazilian purpuric fever identified in a new region of Brazil. *J. Infect. Dis.* **165**(Suppl. 1):S16-S19.
7. Brenner, D. J., L. W. Mayer, G. M. Carlone, L. H. Harrison, W. F. Bibb, M. C. 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. Ghesling, R. C. Cooksey, R. J. Arko, C. V. Broome, and 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.

8. Cariello, N. F., J. A. Swenberg, and T. R. Skopek. 1991. Fidelity of *Thermococcus litoralis* DNA polymerase (Vent) in PCR determined by denaturing gradient gel electrophoresis. *Nucleic Acids Res.* **19**:4193–4198.
9. Carlone, G. M., L. Gorelkin, L. L. Gheesling, A. L. Erwin, S. K. Hoiseth, M. H. Mulks, S. P. O'Connor, R. S. Weyant, J. Myrick, L. Rubin, R. S. Munford III, E. H. White, R. J. Arko, B. Swaminathan, L. M. Graves, L. W. Mayer, M. K. Robinson, S. P. Caudill, and Brazilian Purpuric Fever Study Group. 1989. Potential virulence-associated factors in Brazilian purpuric fever. *J. Clin. Microbiol.* **27**:609–614.
10. Centers for Disease Control. 1986. Brazilian purpuric fever: *Haemophilus aegyptius* bacteremia complicating purulent conjunctivitis. *Morbidity Mortal. Weekly Rep.* **35**:553–554.
11. Centers for Disease Control. 1990. Brazilian purpuric fever—Mato Grosso, Brazil. *Morbidity Mortal. Weekly Rep.* **39**:903–905.
12. Chong, P., Y. P. Yang, D. Persaud, M. Haer, B. Tripet, E. Tam, C. Sia, and M. Klein. 1995. Immunogenicity of synthetic peptides of *Haemophilus influenzae* type b outer membrane protein P1. *Infect. Immun.* **63**:3751–3758.
13. 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. McKenney, 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. Furhman, 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.
14. Genetics Computer Group, Inc. 1995. Wisconsin sequence analysis package program manual, 8th ed. Genetics Computer Group, Madison, Wis.
- 14a. Harayama, S. Unpublished results (NCBI locus PSEXYLU; NCBI gi:939839).
15. Harayama, S., M. Rekik, M. Wubolts, K. Rose, R. A. Leppik, and K. N. Timmis. 1989. Characterization of five genes in the upper-pathway operon of TOL plasmid pWWO from *Pseudomonas putida* and identification of the gene products. *J. Bacteriol.* **171**:5048–5055.
16. Harrison, L. H., G. A. da Silva, M. Pittman, D. W. Fleming, A. Vranjac, C. V. Broome, and Brazilian Purpuric Fever Study Group. 1989. Epidemiology and clinical spectrum of Brazilian purpuric fever. *J. Clin. Microbiol.* **27**:599–604.
17. Hetherington, S. V., and C. C. Patrick. 1992. Complement component 3 binding to *Haemophilus influenzae* type b in the presence of anticapsular and anti-outer membrane antibodies. *Infect. Immun.* **60**:19–24.
18. Hetherington, S. V., C. C. Patrick, and E. J. Hansen. 1993. Outer membrane protein binding sites of complement component 3 during opsonization of *Haemophilus influenzae*. *Infect. Immun.* **61**:5157–5163.
19. Koonin, E. V. 1993. A superfamily of ATPases with diverse functions containing either classical or deviant ATP-binding motif. *J. Mol. Biol.* **229**:1165–1174.
20. Koonin, E. V., R. L. Tatusov, and K. E. Rudd. 1995. Sequence similarity analysis of *Escherichia coli* proteins: functional and evolutionary implications. *Proc. Natl. Acad. Sci. USA* **92**:11921–11925.
21. Lesse, A. J., L. L. Gheesling, W. E. Bittner, S. D. Myers, and G. M. Carlone. 1992. Stable, conserved outer membrane epitope of strains of *Haemophilus influenzae* biogroup aegyptius associated with Brazilian purpuric fever. *Infect. Immun.* **60**:1351–1357.
22. Loeb, M. R. 1987. Protection of infant rats from *Haemophilus influenzae* type b infection by antiserum to purified outer membrane protein a. *Infect. Immun.* **55**:2612–2618.
23. Loewen, P. C., and J. Switala. 1995. Template secondary structure can increase the error frequency of the DNA polymerase from *Thermus aquaticus*. *Gene* **164**:59–63.
24. Mayer, L. W., W. F. Bibb, K. A. Birkness, K. Irino, R. S. Weyant, M. W. Reeves, J. M. Swenson, and Brazilian Purpuric Fever Study Group. 1989. Distinguishing clonal characteristics of the Brazilian purpuric fever-producing strain. *Pediatr. Infect. Dis. J.* **8**:241–243.
25. McIntyre, P., G. Wheaton, J. Erlich, and D. Hansman. 1987. Brazilian purpuric fever in central Australia. *Lancet* **ii**:112. (Letter.)
26. Munson, R., Jr., B. Brodeur, P. Chong, S. Grass, D. Martin, and C. Proulx. 1992. Outer membrane proteins P1 and P2 of *Haemophilus influenzae* type b: structure and identification of surface-exposed epitopes. *J. Infect. Dis.* **165**(Suppl. 1):S86–S89.
27. Munson, R., Jr., and S. Grass. 1988. Purification, cloning, and sequence of outer membrane protein P1 of *Haemophilus influenzae* type b. *Infect. Immun.* **56**:2235–2242.
28. Munson, R., Jr., S. Grass, M. Einhorn, C. Bailey, and C. Newell. 1989. Comparative analysis of the structures of the outer membrane protein P1 genes from major clones of *Haemophilus influenzae* type b. *Infect. Immun.* **57**:3300–3305.
29. 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.
30. Rubin, L. G., E. S. Gloster, G. M. Carlone, and Brazilian Purpuric Fever Study Group. 1989. An infant rat model of bacteremia with Brazilian purpuric fever isolates of *Haemophilus influenzae* biogroup aegyptius. *J. Infect. Dis.* **160**:476–482.
31. Saraste, M., P. R. Sibbald, and A. Wittinghofer. 1990. The P-loop—a common motif in ATP- and GTP-binding proteins. *Trends Biochem. Sci.* **15**:430–434.
32. Swaminathan, B., L. W. Mayer, W. F. Bibb, G. W. Ajello, K. Irino, K. A. Birkness, C. F. Garon, M. W. Reeves, M. C. de Cunto Brandileone, F. O. Sottnek, et al. and Brazilian Purpuric Fever Study Group. 1989. Microbiology of Brazilian purpuric fever and diagnostic tests. *J. Clin. Microbiol.* **27**:605–608.
33. Tindall, K. R., and T. A. Kunkel. 1988. Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry* **27**:6008–6013.
34. Tondella, M. L., F. D. Quinn, and B. A. Perkins. 1995. Brazilian purpuric fever caused by *Haemophilus influenzae* biogroup aegyptius strains lacking the 3031 plasmid. *J. Infect. Dis.* **171**:209–212.
35. Walker, J. E., M. Saraste, M. J. Runswick, and N. J. Gay. 1982. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**:945–951.
36. Wang, Y., M. Rawlings, D. T. Gibson, D. Labbé, H. Bergeron, R. Brousseau, and P. C. K. Lau. 1995. Identification of a membrane protein and a truncated LysR-type regulator associated with the toluene degradation pathway in *Pseudomonas putida* F1. *Mol. Gen. Genet.* **246**:570–579.
37. 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.
38. 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.
39. Wild, B. E., J. W. Pearman, P. B. Campbell, P. K. Swan, and D. L. Garry. 1989. Brazilian purpuric fever in Western Australia. *Med. J. Aust.* **150**:344, 346. (Letter.)
40. Wilson, A. 1989. Preparation of genomic DNA from bacteria, p. 2.4.1–2.4.2. In F. M. Ausubel, R. Brent, R. E. Kingston, et al. (ed.), *Current protocols in molecular biology*. John Wiley & Sons, New York.