Limited Diversity of the Protein D Gene (hpd) among Encapsulated and Nonencapsulated \textit{Haemophilus influenzae} Strains

HÅKAN JANSON, MAORONG RUAN, AND ARNE FORSGREN*

Department of Medical Microbiology, University of Lund, Malmö General Hospital, S-214 01 Malmö, Sweden

Received 3 June 1993/Returned for modification 5 July 1993/Accepted 5 August 1993

Protein D is a surface-exposed lipoprotein of the gram-negative bacterium \textit{Haemophilus influenzae} with affinity for human immunoglobulin D myeloma protein. The gene encoding protein D (hpd) in a serotype b strain of \textit{H. influenzae} was cloned. \textit{Escherichia coli} carrying the hpd gene bound human myeloma immunoglobulin D. Nucleotide sequence analysis identified an 1,092-bp open reading frame that was more than 99% identical to the hpd gene from a nontypeable \textit{H. influenzae} strain. In the deduced amino acid sequences for protein D, only 2 of 364 amino acid residues differed. The restriction fragment length polymorphism of the hpd region in different strains was analyzed by Southern blot analyses of \textit{PstI-} or \textit{EcoRI}-digested genomic DNA from 100 \textit{H. influenzae} strains. The analysis was performed by using isolated fragments of the cloned hpd gene, originating from the nontypeable \textit{H. influenzae} 772, as probes. All strains tested had DNA sequences with a high degree of homology to the hpd probes. The analysis also showed that restriction endonuclease sites within the gene were more conserved than sites adjacent to the hpd gene. An interesting difference between type b strains and unencapsulated strains was observed. The majority of type b strains seem to have a 1.4-kbp DNA fragment upstream of the hpd gene that is absent in nontypeable strains. On the basis of the high degree of conservation of the hpd gene among \textit{H. influenzae} strains, we conclude that protein D is a possible vaccine candidate.

\textit{Haemophilus influenzae} is a common colonizer of the mucosa of the upper respiratory tract and causes diseases by local spread or invasion. \textit{H. influenzae} serotype b (Hib) is a major cause of bacterial meningitis and other invasive infections among children under the age of 4 in Europe and the United States (23, 25). Although the bacterial capsule is recognized as an important virulence factor, the serotype b capsular polysaccharide alone has proven unsatisfactory as an immunogen in children under the age of 2 years (22). However, vaccines also effective in children have been developed by coupling type b carbohydrate capsule, polyribosyl-ribophosphate, to protein carriers such as tetanus or diphtheria toxoids or meningococcal outer membrane protein (1, 27–29). The Hib polysaccharide vaccines do not protect against infections caused by \textit{H. influenzae} strains belonging to serotypes other than type b (i.e., serotype a or c through f or nontypeable \textit{H. influenzae} [NTHi]). NTHi rarely causes invasive infections but is a frequent cause of otitis media and sinusitis in children (19). Patients with cystic fibrosis and chronic bronchitis often carry NTHi, which recently has been recognized also as an important cause of pneumonia (19).

To construct a vaccine that protects against all types of infections caused by \textit{H. influenzae}, surface-exposed \textit{H. influenzae} proteins such as pili (3) and outer membrane proteins (6) have been investigated. Several \textit{H. influenzae} outer membrane proteins have been shown to give rise to protective antibodies in immunized animals. These proteins include the P1 heat-modifiable outer membrane protein (14), the major P2 porin protein (18), the P4 lipoprotein (7), the P6 peptidoglycan-associated lipoprotein (8, 17), the PCP lipoprotein (5), and a surface protein with an apparent molecular weight of 98,000 (13). A protein chosen as a vaccine component must be highly conserved in order to give protection from infection by heterologous strains. Of the vaccine candidates mentioned above, at least P2 has been shown not to protect against challenge by heterologous strains (18).

Another possible candidate to be used as an vaccinogen is the 42-kDa, surface-exposed lipoprotein, named protein D (26), that was found in all of the 127 \textit{H. influenzae} strains tested for direct binding of human myeloma immunoglobulin D (IgD) 4490 to their surfaces (2). By using three different monoclonal antibodies directed against protein D, it was also shown that the IgD-binding molecule is antigenically conserved (2). In this study, we determined the DNA sequence of the structural gene (hpd) encoding protein D from Hib strain MinnA and compared it with the previously reported hpd sequence from NTHi strain 772 (11). We also showed that the gene is present in all \textit{H. influenzae} strains tested and that only a limited variation is found within the gene.

(This work was presented in part at the 92nd General Meeting of the American Society for Microbiology, New Orleans, La., 26–30 May 1992.)

\section*{MATERIALS AND METHODS}

\textbf{Bacterial strains.} A total of 100 \textit{H. influenzae} isolates belonging to different serotypes and biotypes of different geographical origins were examined. The study included 34 type b strains kindly provided by Mogens Kilian, Royal Dental College, Aarhus, Denmark (21, 24); 1 type b strain (MinnA) and 15 NTHi strains kindly obtained from Robert S. Munson, Jr., Washington University School of Medicine, St. Louis, Mo. (20, 30); 2 strains of each serotype a through f from the Culture Collection, University of Göteborg, Göteborg, Sweden; 1 type b strain, a kind gift of Eric J. Hansen,
University of Texas Southwestern Medical Center, Dallas; and 4 NTHi strains generously supplied by Timothy F. Murphy, State University of New York at Buffalo, Buffalo. The remaining 56 strains, including 2 Hib strains, were from our own collection and have been described previously (12). Each isolate was obtained from epidemiologically unrelated patients. All strains were serotyped by agglutination, using type-specific antisera, and were biotyped by the method of Kilian (12).

*Escherichia coli* JM83 (31) was used as the recipient for plasmids pUC18 and pUC19 (31) and derivatives thereof.

*H. influenzae* strains were grown in brain heart infusion broth (Difco Laboratories, Inc. Detroit, Mich.) supplemented with NAD and hemin (Sigma Chemical Co., St. Louis, Mo.), each at 10 μg/ml. *E. coli* JM83 was grown in LB broth (15) supplemented with ampicillin (Sigma) at 100 μg/ml when bacteria carrying plasmids were cultured.

**Cloning and sequencing of the hpd gene from influenzae MinnA.** A chromosomal DNA library from the type b strain *H. influenzae* MinnA was prepared as described previously (11). Transformants were screened for protein D expression by a colony immunobossay (11) with the human IgD myeloma 4499 (26), rabbit anti-human IgD (8 chains) (Dakopatts a/s, Glostrup, Denmark), and horseradish peroxidase-conjugated goat anti-rabbit immunoglobulins (Bio-Rad Laboratories, Richmond, Calif.). Plasmid DNA from positive clones was prepared by using a Quiagen plasmid DNA kit (Qiagen GmbH, Düsseldorf, Germany). Restriction endonuclease analysis, exonuclease III digestion, and DNA sequencing of plasmids containing the hpd gene were performed as described previously (11).

**Direct binding of radiolabelled IgD to intact bacteria.** Binding of 125I-labelled human IgD myeloma 4490 to intact bacteria was performed as described earlier (2).

**Preparation of genomic DNA for Southern blot analyses.** High-molecular-weight bacterial DNA was prepared by a modification (24) of the method described by Moxon et al. (16). Bacteria grown in 10 ml of supplemented brain heart infusion broth for 24 h were harvested by centrifugation followed by a washing step in Tris-EDTA (TE) buffer, pH 7.5 (15). The bacteria were suspended in 250 μl of 50 mM Tris-HCl (pH 8.0) containing 25% sucrose, and 50 μl of lysozyme (10 mg/ml) was added. The cells were left on ice for 5 min before EDTA was added to a final concentration of 0.1 M, and the incubation on ice was continued for 5 min. The proteins were degraded by adding 160 μl of 4 M NaCl, 50 μl of 10% sodium dodecyl sulfate (SDS), and 10 μl of proteinase K (10 mg/ml) and by incubating for 16 h at 55°C. After extraction three times with phenol-chloroform (1:1) and once with chloroform, the nucleic acids were precipitated with ethanol, suspended in 300 μl of TE buffer containing 0.1 mg of RNase A, incubated at 37°C for 3 h, and reprecipitated. The DNA (approximately 20 μg) was finally suspended in 200 μl of TE buffer.

**Genomic Southern blots.** Restriction endonucleases were obtained from Boehringer Mannheim GmbH, Mannheim, Germany. Approximately 1 μg of genomic DNA was digested with 10 U of the appropriate restriction enzyme in buffer provided with the enzyme in a total volume of 30 μl. After incubation at 37°C for 2 h, 5 U of fresh enzyme was added and incubation was continued for 2 h. The DNA fragments were separated by electrophoresis for 16 h at 1.5 V/cm on a horizontal 0.8% (wt/vol) agarose gel containing 0.5 μg of ethidium bromide per ml. The DNA was transferred onto Hybond N+ membranes (Amersham International, Amersham, United Kingdom) by the alkaline transfer method (4).

As probes for hybridization, we used a 316-bp *NheI-EcoRI* fragment and a 629-bp internal *EcoRI* fragment from the hpd gene of the nontypeable *H. influenzae* strain 772 cloned in pUC18 (11). DNA fragments were excised from agarose gels and extracted with Geneclean (Bio 101, Inc., La Jolla, Calif.). The excision-extraction procedure was repeated once to remove contaminating DNA before the fragments were labelled with [32P]dCTP, using a Megaprime random prime kit (Amersham). Prehybridization was carried out at 68°C for 1 h in a solution consisting of 1 M NaCl, 50 mM Tris-HCl (pH 7.5), and yeast RNA (200 μg/ml), and hybridization was carried out for 16 h in prehybridization solution supplemented with 10% dextran sulfate, 1% SDS, sheared denatured herring sperm DNA (200 μg/ml), and 10 6 cpn of 32P-labelled probe per ml. After hybridization, the filter was washed two times for 5 min each time at room temperature in 2× SSC (standard saline citrate; 1× SSC is 0.15 M NaCl, 0.015 M trisodium citrate [pH 7.0]) -0.1% SDS and two times for 20 min each time at 68°C in 0.1× SSC -0.1% SDS and autoradiographed.

**Nucleotide sequence accession number.** The protein D gene sequence of strain MinnA has been assigned GenBank accession number L12445.

**RESULTS**

**Cloning the gene encoding protein D.** Partially *Sau3A*-digested DNA isolated from Hib strain MinnA was ligated to *BamH1*-cut and phospatase-treated pUC18. *E. coli* JM83 cells, transformed with the ligation mixture by high-voltage electroporation, were selected for resistance to ampicillin. Individual colonies were transferred to nitrocellulose filters and screened with purified human IgD myeloma 4490. Among the 2,500 colonies tested, 9 were found to be positive. The positive colonies were picked, purified, and subjected to another two rounds of screening. Five clones remained positive after the purification.

Restriction enzyme analysis of plasmid DNA from the positive clones showed that three of the clones carried a 2.3-kbp insert, one clone carried a 2.9-kbp insert, and one clone had a 3.5-kbp fragment inserted into pUC18. One of the smaller recombinant plasmids, pHIM17, was chosen for further characterization. A partial restriction enzyme map was established for the insert of *H. influenzae* DNA in pHIM17. Previous work has shown that the hpd gene might be toxic to *E. coli* cells when its promoter is oriented in tandem with the lacZ promoter of the pUC vector (11). Therefore, to identify the region coding for protein D in pHIM17, restriction enzyme fragments were subcloned into pUC18 and pUC19. The resulting transformants were tested for expression of protein D by colony immunoblot analysis as described above. These experiments showed that expression of IgD-binding protein from plasmids carrying a 1.8-kbp *PstI*-XbaI fragment from pHIM17 was not dependent on the orientation of the insert since transformants bound IgD regardless of whether the fragment was ligated to *PstI*-XbaI-digested pUC18 or pUC19. Both recombinant plasmids were shown to express an IgD-binding protein of equivalent size to that expressed from the wild type *H. influenzae* strain (immunoblot) analysis (data not shown). The recombinant pUC19 plasmid carrying the 1.8-kbp *PstI*-XbaI fragment, called pHIM26 (Fig. 1), was kept for further experiments.

The nucleotide sequence was determined on both strands of the gene encoding protein D by direct plasmid sequencing of subclones and nested deletion constructs of pHIM26. The hpd gene from strain MinnA was compared with the previ-
FIG. 1. Partial restriction endonuclease map of pHIM26 encoding hpd from H. influenzae MinnA. Restriction endonucleases conserved in H. influenzae 772 ( ), direction of translation (→), and fragments used as probes (---) are indicated. Vector DNA is marked in boldface.

FIG. 2. Comparison between hpd gene sequence and deduced amino acid sequence of protein D from H. influenzae serotype b strain MinnA and nontypeable H. influenzae 772. Possible ribosomal binding sites are underlined. Dashes denote sequence identity. The H. influenzae 772 nucleotides are indicated where the sequences differ. Numbers in italics refer to deduced amino acid positions of protein D from Hib MinnA. Deduced amino acid substitutions are marked in boldface, with protein D from NTHi 772 shown below.

The nucleotide sequences were 99% identical and differed in only 7 of 1,092 bp. Nucleotide exchanges in positions 1009 and 1042 led to substitutions at the deduced amino acid level. The remaining nucleotide differences at positions 188, 329, 341, 452, and 1004 are silent substitutions (Fig. 2). Comparisons of sequences flanking hpd showed that the downstream region was conserved, while the sequence immediately upstream of the start codon showed no homology apart from the putative ribosomal binding site (Fig. 2). No putative promoter region could be found upstream of hpd in pHIM26 since the Sau3A site used for isolating the fragment is found only 29 nucleotides upstream of the ATG start codon (Fig. 2). The transcription of hpd from pHIM26 is therefore probably dependent on a promoter in pUC18.

Southern blot analysis of chromosomal DNA from various H. influenzae strains. To characterize the hpd locus at the nucleotide level further, chromosomal DNA from 100 strains of H. influenzae belonging to different serotypes and bio-types and from various geographical origins was digested with...
with PstI or EcoRI and subjected to restriction fragment length polymorphism (RFLP) analysis. Digested genomic DNA was separated on agarose gels and transferred to positively charged nylon membranes. Two probes were selected in order to determine the RFLP within and flanking the hpd gene and to cover as much as possible of the region encoding the mature gene product of the hpd gene, omitting the sequence encoding the signal sequence (10, 11) in order to minimize cross-hybridization to other signal sequences. Previous restriction endonuclease mapping experiments of the hpd gene from *H. influenzae* 772 cloned into pUC18 have shown that the hpd gene contains an internal 629-bp EcoRI fragment in the 3′ end of the protein D gene and a 316-bp NheI-EcoRI fragment upstream of this EcoRI fragment (11). These two restriction endonuclease fragments were also found when restriction endonuclease sites in pHI26 carrying the hpd gene from *H. influenzae* MinnA were mapped (Fig. 1). The probes correspond to 91% (945 of 1,035 bp) of the hpd gene encoding the mature protein D (10, 11). The membranes were hybridized with the 316-bp NheI-EcoRI probe, homologous to the 5′ end of hpd (Fig. 3). After washing and autoradiography, the membranes were stripped with alkali and rehybridized with the 629-bp EcoRI probe, homologous to the 3′ end of hpd (Fig. 4). DNA from each of the 100 strains contained sequences that hybridized with the two hpd probes under the stringent conditions used (Table 1).

FIG. 3. RFLP analysis of hpd gene sequences digested with PstI. Autoradiograms show the restriction fragments of chromosomal DNA hybridizing with a 316-bp NheI-EcoRI fragment of pHIC348 (A) and a 629-bp EcoRI fragment (B), which are specific for the 5′ and 3′ ends of the hpd gene, respectively. Molecular sizes are indicated in kilobase pairs. The sizes of fragments larger than 8 kbp are approximations as a result of the technique’s low resolution of large DNA fragments.

The number of hybridizing genomic DNA restriction fragments depends on the organization within the bacterial genome and the number of internal restriction endonuclease sites within the DNA that are homologous to the two hpd probes. If the hpd-specific chromosomal DNA sequences constitute a continuous stretch with no recognition sites for the restriction endonuclease used, then only one fragment will contain the hpd gene and subsequently both probes will hybridize to this fragment. The size of the fragment is determined by the location of restriction sites adjacent to the hpd gene. Restriction endonuclease PstI was chosen in order to determine the RFLP of sequences flanking hpd since neither *H. influenzae* 772 (11) nor *H. influenzae* MinnA (Fig. 1) contains any internal PstI sites within hpd. The *H. influenzae* strains analyzed were divided into groups (A to G) on the basis of their PstI RFLP (Table 1). DNA from 10 nontypeable strains (group G) was completely resistant to digestion with PstI (Table 1 and Fig. 3), suggesting a DNA modification system in these strains which includes the PstI recognition sequence, since EcoRI cleaved DNA from these strains readily. Resistance to PstI digestion has been reported previously by other authors (24). In 77 of 90 strains successfully digested with PstI, both probes recognized the same DNA fragment of a particular strain, whereas the remaining 13 strains hybridized with fragments of different sizes (Table 1 and Fig. 3). The patterns seen among these 13 strains belonging to groups C and D might indicate that there is an internal PstI site in the 5′ part of hpd in these strains, since the 3′-specific probe recognizes two fragments of different sizes while the 5′-specific probe only hybridizes to one of the fragments. Another possible explanation is that
the strains belonging to groups C and D carry a second gene with homology to the 3' end of hpd.

There are two internal restriction sites for EcoRI in the protein D gene in DNA from H. influenzae 772 (11) and MinnA (Fig. 1), both of which were used in this study to investigate whether any RFLP could be found within the hpd gene. The internal 629-bp EcoRI fragment which was used as a 3'-specific probe hybridized with a 0.63-kbp fragment after EcoRI digestion of genomic DNA in 93 of 100 strains (Table 1 and Fig. 4B), which suggests that the 3' end of hpd is highly conserved. The remaining seven strains had fragments that were recognized by both probes (Table 1 and Fig. 4), indicating that the EcoRI site separating the two probes was missing in these strains. The size of the fragment with which the 5'-specific, 316-bp NheI-EcoRI probe hybridizes after EcoRI digestion of genomic DNA is subsequently determined by the location of the EcoRI restriction site upstream of hpd. This region shows relatively high polymorphism.

![Image of autoradiograms showing RFLP analysis of hpd gene sequences digested with EcoRI.](http://iai.asm.org/)

**FIG. 4.** RFLP analysis of hpd gene sequences digested with EcoRI. Autoradiograms show the restriction fragments of chromosomal DNA hybridizing with a 316-bp NheI-EcoRI fragment of pHIC348 (A) and a 629-bp EcoRI fragment (B), which are specific for the 5' and 3' ends of the hpd gene, respectively. Molecular sizes are indicated in kilobase pairs. The sizes of fragments larger than 8 kbp are approximations as a result of the technique’s low resolution of large DNA fragments.

<table>
<thead>
<tr>
<th>Group</th>
<th>Sizes (kb) of DNA fragments hybridizing with:</th>
<th>No. of strains with given serotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PsI-digested DNA</td>
<td>EcoRI-digested DNA</td>
</tr>
<tr>
<td></td>
<td>5' probea</td>
<td>3' probea</td>
</tr>
<tr>
<td>A</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>B</td>
<td>3.3</td>
<td>3.3</td>
</tr>
<tr>
<td>C</td>
<td>2.7</td>
<td>2.7 + 20f</td>
</tr>
<tr>
<td>D</td>
<td>2.1</td>
<td>2.1 + 20f</td>
</tr>
<tr>
<td>E</td>
<td>2.1</td>
<td>2.1</td>
</tr>
<tr>
<td>F</td>
<td>1.9</td>
<td>1.9</td>
</tr>
<tr>
<td>G</td>
<td>Uncleaved</td>
<td>Uncleaved</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group</th>
<th>Sizes (kb) of DNA fragments hybridizing with:</th>
<th>No. of strains with given serotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PsI-digested DNA</td>
<td>EcoRI-digested DNA</td>
</tr>
<tr>
<td></td>
<td>5' probea</td>
<td>3' probea</td>
</tr>
<tr>
<td>A</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>B</td>
<td>3.3</td>
<td>3.3</td>
</tr>
<tr>
<td>C</td>
<td>2.7</td>
<td>2.7 + 20f</td>
</tr>
<tr>
<td>D</td>
<td>2.1</td>
<td>2.1 + 20f</td>
</tr>
<tr>
<td>E</td>
<td>2.1</td>
<td>2.1</td>
</tr>
<tr>
<td>F</td>
<td>1.9</td>
<td>1.9</td>
</tr>
<tr>
<td>G</td>
<td>Uncleaved</td>
<td>Uncleaved</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group</th>
<th>Sizes (kb) of DNA fragments hybridizing with:</th>
<th>No. of strains with given serotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PsI-digested DNA</td>
<td>EcoRI-digested DNA</td>
</tr>
<tr>
<td></td>
<td>5' probea</td>
<td>3' probea</td>
</tr>
<tr>
<td>A</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>B</td>
<td>3.3</td>
<td>3.3</td>
</tr>
<tr>
<td>C</td>
<td>2.7</td>
<td>2.7 + 20f</td>
</tr>
<tr>
<td>D</td>
<td>2.1</td>
<td>2.1 + 20f</td>
</tr>
<tr>
<td>E</td>
<td>2.1</td>
<td>2.1</td>
</tr>
<tr>
<td>F</td>
<td>1.9</td>
<td>1.9</td>
</tr>
<tr>
<td>G</td>
<td>Uncleaved</td>
<td>Uncleaved</td>
</tr>
</tbody>
</table>
(Table 1). The RFLP pattern after EcoRI digestion of DNA from NTHi strain 6-7572 is unique (Fig. 4). A plausible explanation is that this strain may contain two copies of hpd.

**Binding of IgD myeloma 4490 to intact cells.** Most *H. influenzae* strains used in this study have been shown previously to bind high levels of radiolabelled human myeloma IgD to their surfaces (2). Strains that were not tested previously for IgD binding were incubated with 125I-labelled human myeloma IgD 4490 for binding studies. *H. influenzae* 772 (NTHi) and MMa8 (Hib) and *E. coli* JM83 harboring plasmid pUC18 (negative control) were used as reference strains. All *H. influenzae* strains bound between 30 and 70% of the added radiolabelled IgD, in agreement with earlier results (2). *E. coli* carrying phIM26 bound 50% in comparison to *E. coli* JM83(pHIC348) and *E. coli* JM83 (pUC18), which bound 30% and 3%, respectively.

**DISCUSSION**

The present investigation describes the cloning and sequencing of the gene encoding protein D from the serotype b strain *H. influenzae* MMa8. The nucleotide and deduced amino acid sequences were found to be highly homologous to the protein D gene from the nontypeable *H. influenzae* strain 772. Only 2 of 364 amino acids differed in the deduced sequence when the two hpd genes were compared. Both substitutions were found in the C-terminal end of the deduced amino acid sequence. Results from previous investigations indicate that protein D is anchored to the membrane by fatty acids attached to the N-terminal cysteine residue (10). Preliminary data suggest that the IgD-binding region of protein D is located at the C-terminal part of protein D, since *E. coli* clones expressing protein D that is truncated in the C-terminal part of the protein do not bind radiolabelled human myeloma IgD 4490 to their surfaces (9). If the C-terminal end of protein D is exposed on the outside of the bacterium, there is a larger selective pressure on this part of protein D caused by the immune system of the host. Hence, this region could be more variable than the N-terminal part in order to avoid the immune system. The hpd genes of more *H. influenzae* strains must be sequenced in order to confirm this hypothesis.

Previous studies have demonstrated that all *H. influenzae* isolates tested express a 42-kDa IgD-binding protein (2). Our results show not only that protein D is antigenically and functionally conserved but that the organization of the hpd gene is homologous in *H. influenzae* strains belonging to different serotypes and biotypes. The 3' part of hpd shows a limited diversity since the 629-bp 3'-specific probe hybridizes to a 0.6-kbp fragment in 93 of the 100 strains after digestion with EcoRI. The remaining seven strains are apparently lacking the EcoRI site separating the two probes, while the site at the very 3' end of hpd is intact since the two probes hybridize to the same fragment in one given strain.

An interesting observation was found when the RFLPs of hpd of type b and nontypeable strains of *H. influenzae* were compared. The vast majority of Hib strains (34 of 39) belong to group B; i.e., they have a 3.3-kbp *PstI* fragment with which the two hpd-specific probes hybridize, whereas none of the 41 NTHi strains successfully digested with *PstI* show hybridization to a fragment of this size. The dominating size of *PstI* fragments from NTHi strains (24 of 41) with which hpd probes hybridize is a 1.9-kbp fragment (group F). The difference of 1.4 kbp probably lies upstream of hpd since the *PstI* site downstream of the protein D gene is conserved in *H. influenzae* MMa8 and 772 (Fig. 1), which belong to RFLP groups B and F, respectively (Fig. 3 and Table 1). The median sizes of EcoRI fragments that hybridize with the 5'-specific probe are 6.3 kbp for Hib strains belonging to group B and 5.0 kbp for NTHi strains belonging to group F. The EcoRI site separating the 5' and 3'-specific probes is conserved in most strains, and it is therefore the position of the EcoRI site upstream of hpd that determines the polymorphism. The difference of 1.3 kbp upstream of hpd between most Hib and NTHi strains correlates well with the difference of 1.4 kbp between Hib and NTHi fragments hybridizing to the hpd-specific probes after *PstI* digestion. The 1.4-kbp fragment upstream of hpd is probably not involved in capsule expression since the four Hib strains belonging to group F, lacking the additional 1.4-kbp fragment, are agglutinated by type b-specific antiserum. The region upstream of hpd is currently being analyzed in our laboratory in order to investigate the nature of the 1.4-kbp difference between nontypeable and serotype b strains.

The 10 nontypeable strains that were resistant to *PstI* digestion show EcoRI RFLPs similar to those of other NTHi strains, suggesting that the hpd region of these strains does not differ from that of the majority of nontypeable strains. The strains belonging to serotypes other than b and nontypeable strains were too few to allow us to draw any general conclusions about the organization upstream of the hpd gene among those strains. No correlation between RFLP patterns and country of isolation or biotype was observed.

On the basis of the homology between the nucleotide and deduced amino acid sequences of strains MMa8 and 772 together with the conservation of EcoRI sites within the protein D gene from 100 *H. influenzae* strains, it is concluded that there is only a limited diversity within the hpd gene. However, considerable variation in other regions within hpd could exist among the strains that were not sequenced. Therefore, the nucleotide sequence of hpd from more strains must be analyzed before any general conclusions can be made concerning the conservation of the amino acid sequence of protein D and its possibilities as a future vaccine against all types of *H. influenzae*.

**ACKNOWLEDGMENTS**

We thank Anders Grubb for sharing his limited amount of IgD with us, Gertrud Hansson for technical skillful assistance, and Lars-Olof Hedén for critically reviewing the manuscript. This work was supported by grants from Swedish Medical Research Council (projects 03163, and 05196), the Medical Faculty of Lund, the Alfred Osterlund Foundation, the Crafoordska Foundation, the G. and J. Kock Foundation, and Malmö Medical Service Foundation.

**REFERENCES**


