INFECTION AND IMMUNITY, Sept. 1990, p. 2895–2905
0019-9567/90/092895-11$02.00/0
Copyright © 1990, American Society for Microbiology
Vol. 58. No. 9

Cloning, Partial Sequence, Expression, and Antigenic Analysis of the Filamentous Hemagglutinin Gene of *Bordetella pertussis*

ANNE-MARIE DELISSE-GATHOYE,[1] CAMILLE LOCHT,[1]† FRANÇOISE JACOB,[1] METTE RAASCHOU-NIELSEN,[2] IVER HERON,[2] JEAN-LOUIS RUELLE,[1] MICHEL DE WILDE,[1] AND TERESA CABEZON[1]*

Department of Molecular and Cellular Biology, Smith Kline Biologicals, 89, Rue de l'Institut, B-1330 Rixensart, Belgium,[1] and Vaccine Department, Statens Seruminstitut, Copenhagen, Denmark[2]

Received 5 March 1990/Accepted 26 June 1990

The gene coding for the filamentous hemagglutinin (FHA), one of the main factors involved in mediating adherence of *Bordetella pertussis* to ciliated host cells, was cloned in *Escherichia coli*, and the 3,500-base-pair nucleotide sequence encoding the amino-terminal region was determined. Molecular cloning, together with the characterization of recombinant FHA-related proteins produced in *E. coli*, revealed that the primary translation product is a protein of about 370 kilodaltons (kDa). The mature 220-kDa FHA polypeptide secreted by *B. pertussis* is most probably generated by proteolytic processing that eliminates a carboxy-terminal portion of about 150 kDa. The 1,087 amino-terminal residues of the predicted FHA sequence showed a number of remarkable features. Extensive homology to the *Serratia marcescens* and *Proteus mirabilis* hemolysin proteins was found between amino acids 91 and 205 of the FHA sequence, suggesting involvement of this FHA domain in host cell binding or secretion of FHA from *B. pertussis*. In addition, two regions containing repetitive amino acid sequences were identified. One region, extending from residues 382 to 664, was formed by six repeats, and a second, extending from residues 701 to 912, contained three repeats. The reactivities of several recombinant FHA-derived proteins with a panel of monoclonal antibodies identified at least four epitopes composing an immunoreactive domain present in the carboxy-terminal moiety of the mature FHA.

Virulent *Bordetella pertussis*, the etiologic agent of whooping cough, is characterized by the coordinated synthesis of a variety of virulence factors (37). These include toxins and adherence factors such as pertussis toxin (PTX), extracellular adenylate cyclase, dermonecrotic toxin, filamentous hemagglutinin (FHA), and agglutinogens. Some of these factors are known to be important protective antigens in pertussis vaccines, and the major antigens included in the new acellular vaccines are PTX and FHA (12, 29). The amino acid sequence of PTX has been deduced from the cloning (16, 22) and sequencing (18, 22) of its structural gene. Individual subunits of PTX have been synthesized in heterologous hosts and analyzed for biochemical activity and immunological properties (1, 3, 17, 21, 26). Much less is known about the properties of FHA. Cloning and expression in *Escherichia coli* of the partial structural gene have been reported (2, 11, 20, 25). The important immunogenic regions on the FHA molecule have not yet been identified, despite the fact that immunization with FHA protects mice from aerosol challenge with virulent *B. pertussis* (23, 28). Furthermore, FHA has the ability to adhere to human ciliated respiratory epithelial cells (33) and probably plays an important role in the attachment of *B. pertussis* to the respiratory tract, one of the initial steps in the pathogenesis of whooping cough. Characterization of the epithelial cell-binding sites on FHA and identification of neutralizing epitopes are therefore important steps toward the definition of an FHA molecule useful for vaccine purposes. An understanding of the mechanisms intervening in the expression and posttranslational processing of the FHA could be useful for defining its function as a virulent factor.

We report here the cloning of the complete structural gene for FHA, the sequence of its 5′ third, and the analysis of its coding capacity by expression of gene fragments in *E. coli*. The expression products were characterized by polyclonal and monoclonal antibodies, and a 370-kilodalton (kDa) protein was identified as the primary translation product of the FHA gene. Finally, we identified by computer analysis extensive amino acid homology between the amino-terminal region of FHA and the hemolysins of *Serratia marcescens* and *Proteus mirabilis*, as well as the existence of two regions with repetitive amino acid sequences in the amino-terminal half of the FHA protein.

MATERIALS AND METHODS

Bacterial strains, plasmids, and transformation. The *B. pertussis* strains used were Tohama I, obtained from J. J. Munoz (Rocky Mountain Laboratories, Hamilton, Mont.), and BP353, a Tn*5*-induced FHA-defective mutant strain (38) provided by S. Falkow (Stanford University, Stanford, Calif.). *E. coli* MM294 (*endA hsr hsm*), obtained from M. Ptashne (Harvard University, Cambridge, Mass.), was used as recipient for recombinant plasmids. *E. coli* TG1 [Δ(*lac-pro*)*supE thi hsdD5*(F′ *traD36 proA*$^+$ *B*$^+$ *lacI*^{q} *lacZ*ΔM15)] was purchased from Amersham U.K. and used for FHA gene expression from pUC8- and pUC9-derived recombinant plasmids.

The vectors used for cloning and expression experiments were pBR327, pUC8, and pUC9 (35), purchased from Pharmacia, Inc. (Piscataway, N.J.), and pUC8-1, pUC8-2, pUC9-1, and pUC9-2 (9).

Transformation of *E. coli* was performed as described by Maniatis et al. (19, 27).

Growth media. *B. pertussis* was grown on Bordet-Gengou

* Corresponding author.
† Present address: Laboratory of Genetical and Molecular Microbiology, Institut Pasteur, 59019 Lille-Cédex, France.

agar plates (Difco Laboratories, Detroit, Mich.) or liquid Stainer-Scholte medium (32). *E. coli* cells were grown in Luria broth (LB) supplemented with ampicillin (200 μg/ml) when necessary (19, 27).

DNA manipulation. Restriction endonucleases were purchased from Amersham Corp. (Arlington Heights, Ill.), Bethesda Research Laboratories, Inc. (Gaithersburg, Md.), Boehringer Mannheim Biochemicals (Indianapolis, Ind.), New England BioLabs, Inc. (Beverly, Mass.), and P-L Biochemicals, Inc. (Milwaukee, Wis.) and used according to the specifications of the suppliers. Radiolabeling of DNA and DNA purification were performed as described by Maniatis et al. (19, 27).

DNA sequencing. Restriction fragments of recombinant plasmids containing the FHA structural gene were subcloned into different pUC plasmids. The DNA inserts were sequenced by the dideoxy-chain termination method adapted for plasmids (4), using Sequenase (United States Biochemical Corp., Cleveland, Ohio). To sequence through regions with DNA compression, dITP was used in all four sequencing reaction mixes (27).

Polyclonal antisera. The polyclonal antisera used were kindly provided by E. Simoen (Smith Kline Biologicals, Rixensart, Belgium). Anti-FHA polyclonal antiserum was obtained by immunization of a goat with FHA purified by standard techniques (11). Anti-*Bordetella* polyclonal antiserum was obtained by immunization of guinea pigs with heat-killed whole bacteria of *B. pertussis* Tohama I.

Generation of monoclonal anti-FHA antibodies. BALB/c mice were immunized intraperitoneally with one-third of a human dose of whole-cell pertussis vaccine and boosted intraperitoneally 3 weeks later with 500 μg of purified FHA adsorbed to $Al(OH)_3$. The fusion procedure was essentially as described by Köhler and Milstein (14), using X63 Ag8 myeloma cells as fusion partner. Screening of FHA-reactive hybridomas was performed in enzyme-linked immunosorbent assay and confirmed in Western blots (immunoblots) on purified FHA. Hybrids were cloned by limiting dilution until all wells derived from single cells were positive. Monoclonality was verified by isoelectric focusing. All antibodies used were of the immunoglobulin G1 subclass except antibody 13.1F5, which was of the immunoglobulin G2b subclass.

Immunoblot analysis of the recombinant proteins. *E. coli* cells carrying recombinant plasmids were grown at 37°C in LB medium containing ampicillin (200 μg/ml). At the end of the exponential growth phase, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 2 mM, and growth was continued for 4 h. The cells were then harvested by centrifugation, washed with phosphate-buffered saline, suspended in Laemmli buffer (15), and boiled. The proteins were then separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to nitrocellulose. FHA-related polypeptides were detected by using anti-FHA polyclonal antiserum or anti-FHA monoclonal antibodies.

Computer sequence analysis. Routine computer-aided analysis of nucleic and protein sequences was performed with the University of Wisconsin software package (5). The sequence data bases scanned for homology searches were NBRF (release 21) and GenBank (release 60).

Nucleotide sequence accession number. The sequence reported here was deposited in the EMBL/GenBank data base under accession number X53405 FHAB.

RESULTS

Molecular cloning of the FHA structural gene in *E. coli*. The *B. pertussis* gene library described by Jacob et al. (11) was screened with the following DNA probes. Probe A corresponds to *B. pertussis* DNA sequences flanking the Tn*5* insertion in the FHA-deficient mutant BP353 (38). To generate probe A, *B. pertussis* BP353 DNA was digested with *Sal*I restriction enzyme and inserted into pUC9 previously digested with *Sal*I. *E. coli* MM294 was transformed to ampicillin resistance with the recombinant plasmids and then screened by using the radiolabeled neomycin resistance gene carried by Tn*5*. Two clones both containing about 350 base pairs (bp) of *B. pertussis* DNA adjacent to the Tn*5* insertion were isolated. A 550-bp *Sal*I-*Hpa*I DNA fragment including the 350 bp of *B. pertussis* DNA was then purified, radiolabeled, and used as probe A. Probes B1 and B2 were synthetic oligonucleotides with the sequences CACATGC ACCTGGATGCGCCGCGC and GGACCAGCCTGGCCCG AGCGCTGC, corresponding to extremities I and II, respectively, of the 2.9-kilobase-pair (kb) recombinant *B. pertussis* DNA insert described by Mattei et al (20). Probe C was generated by reverse translation of a portion of the FHA protein sequence. This probe was a 32-residue-long synthetic oligonucleotide with the sequence CAGGACTTCTTC ACCCCGGGCAGCGTCGTCGT (11). The codon usage was chosen according to the sequence of the PTX gene (18). Several clones hybridizing to one or two of the probes described above were isolated from the *B. pertussis* gene library. Probe B2 hybridized to many clones in the gene library but did not hybridize to clones that simultaneously hybridized to any of the other probes. Therefore, probe B2 probably hybridized to DNA that is not part of the FHA gene, as discussed by Mattei et al. (20). Four of the isolated clones together cover a chromosomal region of about 20 kb. A restriction map of each cloned DNA insert was obtained, and the regions hybridizing to the individual DNA probes were identified (Fig. 1a to d). The *Bam*HI[a]-*Xho*I[a] (superscripts indicate restriction sites) fragment that hybridized to probe C was subcloned and sequenced. One of the three possible reading frames was in perfect correspondence to the determined protein sequence (11). However, neither an ATG initiation codon nor a sequence coding for a putative signal peptide (6) could be found in the region immediately upstream of the sequence corresponding to probe C. Instead, a potential cleavage site for trypsinlike proteases (Arg-Arg-Ala-Arg-Arg) was found at position 1074 of the sequence shown in Fig. 2. When the 800-bp *Bam*HI[a]-*Xho*I[a] (Fig. 1e, fragment 6) DNA fragment was subcloned into pUC8 such that the putative FHA-coding sequence was in frame with the first nine amino acids of the β-galactosidase open reading frame (ORF), a fusion protein was synthesized in *E. coli* that could be detected with anti-FHA polyclonal antiserum (11). This finding strongly indicated that the 20-kb DNA, covered by the four clones, encodes the structural gene of FHA.

Identification of the 5′ end of the FHA structural gene. Since the DNA sequence analysis discussed above suggested that the 800-bp *Bam*HI[a]-*Xho*I[a] fragment codes for an internal FHA fragment rather than the amino-terminal region, the upstream region was sequenced up to the *Eco*RI[a] site of pRIT13075 (Fig. 1d). The sequence revealed one ORF starting 253 bp downstream of this *Eco*RI site (Fig. 2) and extending through the sequenced region. Three in-frame ATG codons were found as possible candidates for the initiation of translation. An amino acid sequence with some features of a signal peptide was found between residues 33 to

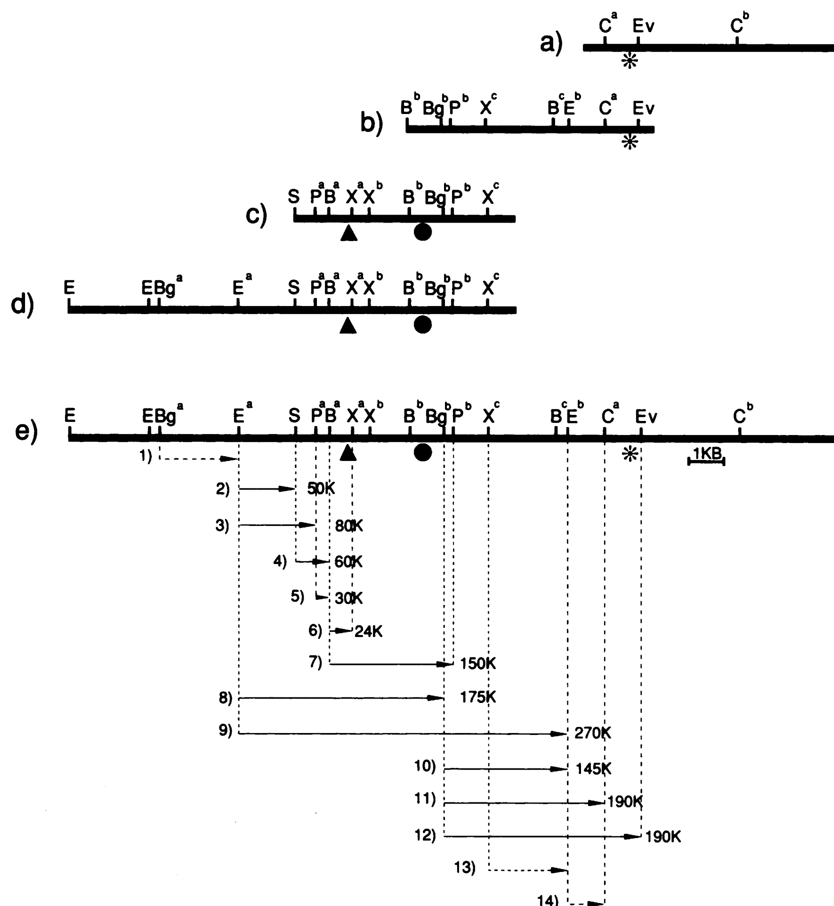

FIG. 1. Restriction map of the overlapping clones containing the FHA-coding sequence and schematic representation of 14 contiguous or overlapping *B. pertussis* DNA fragments expressed in *E. coli*: pRIT12990 (a), pRIT12988 (b), pRIT12987 (c), and pRIT13075 (d). In map e, the arrows indicate the length of the DNA fragments inserted into the pUC vectors and the direction of their transcription from the *lac* promoter of the vectors. The vertical dotted lines indicate the 5′ and 3′ ends of the inserted DNA fragments. The numbers following the arrows indicate the sizes (in kilodaltons [K]) of the proteins produced by the recombinant *E. coli* containing the respective DNA fragments. Note that the sizes of the recombinant proteins correspond well to the sizes of the DNA fragments except for fragment 12, as discussed in the text. Fragments 2 to 12, represented by the solid lines, are those recognized by polyclonal anti-FHA antibodies in immunoblot analysis. Fragments 1, 13, and 14, represented by the dotted arrows, are not recognized by anti-FHA polyclonal antibodies. Restriction sites: X, *Xho*I; B, *Bam*HI; E, *Eco*RI; C, *Cla*I; P, *Pst*I; Bg, *Bgl*II; S, *Sal*I; Ev, *Eco*RV. Sequences hybridizing to the three different probes are indicated by * (probe A), ● (probe B1), and ▲ (probe C).

71. A promoterlike structure was identified upstream of the first of the three ATG codons.

Because of the high C+G content of *B. pertussis* DNA, unambiguous sequence data are difficult to obtain and misinterpretation of the reading frame through sequence errors is difficult to avoid. An expression strategy was therefore developed to verify that the correct ORF had been identified. Several fragments of the 5′ region of the FHA gene were inserted in all three reading frames into the universal pUC vectors (9), and the expression products were analyzed with anti-FHA polyclonal antiserum. Six adjacent or overlapping fragments covering together almost 3,000 bp of the 5′ region were analyzed (fragments 1 to 6; Fig. 1e and Table 1) in this way. All fragments except fragment 1 (*Bgl*II[a]-*Eco*RI[a]) led to the synthesis of an anti-FHA-immunoreactive protein, and this in only one of the three phases of the vector series (Table 1). FHA-related fusion proteins were not detected in *E. coli* extracts harboring any of the three vectors containing no *B. pertussis* DNA insert. The migration of the different recombinant proteins is in good agreement with the lengths of the DNA inserts. From these results, it can be concluded that the FHA ORF start is located downstream and close to the *Eco*RI[a] site. Since several stop codons were found between the *Eco*RI[a] site and the first ATG (Fig. 2), it is surprising that expression of fragments 2 and 3 (Fig. 1e) was observed in only one reading frame. In addition, expression of these two fragments was independent of the presence of IPTG in the culture medium, suggesting constitutive expression. The mechanism by which polypeptides 2 and 3 are synthesized from these plasmids and the amino-terminal structures of these recombinant proteins remains to be investigated. The same feature was observed for fragments 9 and 8 (Fig. 3, lane b and c).

Identification of the 3′ end of the FHA structural gene. Insertion of fragment 7 (Fig. 1e) in pUC9, in the same reading frame that led to the synthesis of the 24-kDa polypeptide encoded by fragment 6, gave rise to a 150-kDa FHA-related protein, indicating that the ORF coding for the FHA extends downstream of the *Bgl*II[b] site (Fig. 1e) and that the total FHA ORF has a coding capacity going well beyond a 220-kDa protein, the largest FHA-related polypeptide observed to be secreted from *B. pertussis* (11). Frag-

```
     EcoRI
   1 GAATTCCTGCGCTGGCACCCGCGGCGGGCCGGGGAGCGGGTTGTCGGCGCACGCCTATAC   60

  61 GTGCCGGACAGGGTTTGATGGTTTGACTAAGAAATTTCCTACAAGTCTTGTATAAATATC  120

 121 CATTGATGGACGGGATCATTACTGACTGACGAAGTGCTGAGGTTTATCCAGACTATGGCA  180

 181 CTGGATTTCAAAACCTAAAACGAGCAGGCCGATAACGGATTCTGCCGATTACTTCACTTC  240

 241 GCTGGTCGGAATATGAACACGAACCTGTACAGGCTGGTCTTCAGCCATGTTCGCGGCATG  300
   1             M  N  T  N  L  Y  R  L  V  F  S  H  V  R  G  M    16

 301 CTTGTTCCCGTGAGCGAGCATTGCACCGTCGGAAACACCTTCTGTGGGCGCACGCGTGGT  360
  17 L  V  P  V  S  E  H  C  T  V  G  N  T  F  C  G  R  T  R  G    36

 361 CAAGCGCGAAGTGGGGCCCGCGCCACGAGCCTGTCCGTAGCGCCCAATGCGCTGGCCTGG  420
  37 Q  A  R  S  G  A  R  A  T  S  L  S  V  A  P  N  A  L  A  W    56
                                                 ↓
 421 GCCCTGATGTTGGCGTGTACGGGTCTTCCGTTAGTAACGCACGCCCAGGCTTGGTTCCT  480
  57 A  L  M  L  A  C  T  G  L  P  L  V  T  H  A  Q  G  L  V  P    76

 481 CAGGGGCAGACACAGGTGCTGCAGGGCGGGAACAAGGTTCCCGTTGTCAATATCGCCGAC  540
  77 Q  G  Q  T  Q  V  L  Q  G  G  N  K  V  P  V  V  N  I  A  D    96

I 541 CCAAATTCCGGCGGCGTCTCGCACAACAAGTTCCAGCAGTTCAACGTCGCCAACCCTGGC  600
  97 P  N  S  G  G  V  S  H  N  K  F  Q  Q  F  N  V  A  N  P  G   116

 601 GTGGTCTTCAACAACGGCCTGACCGACGGCGTGTCCAGGATCGGCGGGGCGCTGACCAAG  660
 117 V  V  F  N  N  G  L  T  D  G  V  S  R  I  G  G  A  L  T  K   136

 661 AACCCCAACCTGACTCGCCAGGCCTCGGCCATTCTTGCCGAAGTCACGGACACTTCGCCC  720
 137 N  P  N  L  T  R  Q  A  S  A  I  L  A  E  V  T  D  T  S  P   156

 721 AGTCGCCTGGCCGGTACGCTCGAAGTCTATGGCAAGGGCGCCGACCTCATCATCGCCAAC  780
 157 S  R  L  A  G  T  L  E  V  Y  G  K  G  A  D  L  I  I  A  N   176

 781 CCCAACGGCATCAGCGTCAACGGCCTGAGCACGCTCAACGCCAGCAACCTGACGCTCACG  840
 177 P  N  G  I  S  V  N  G  L  S  T  L  N  A  S  N  L  T  L  T   196

 841 ACGGGGCGTCCCAGCGTCAACGGCGGCCGCATCGGCCTTGATGTCCAACAGGGCACCGTC  900
 197 T  G  R  P  S  V  N  G  G  R  I  G  L  D  V  Q  Q  G  T  V   216

 901 ACGATCGAACGAGGCGGCGTCAATGCCACCGGCCTGGGCTATTTCGACGTGGTGGCGCGC  960
 217 T  I  E  R  G  G  V  N  A  T  G  L  G  Y  F  D  V  V  A  R   236

 961 CTGGTCAAGCTGCAGGGTGCCGTGTCGAGCAAGCAGGGCAAGCCCCTGGCCGACATCGCG 1020
 237 L  V  K  L  Q  G  A  V  S  S  K  Q  G  K  P  L  A  D  I  A   256

1021 GTGGTCGCCGGCGCCAACCGGTACGACCACGCAACCCGCCGCGCCACGCCGATCGCCGCA 1080
 257 V  V  A  G  A  N  R  Y  D  H  A  T  R  R  A  T  P  I  A  A   276

1081 GGCGCGCGGCGCGCCGCCGCGGCGGCCTACGCGATTGACGGCACGGCGGCGGGCGCCATG 1140
 277 G  A  R  R  A  A  A  A  A  Y  A  I  D  G  T  A  A  G  A  M   296

1141 TACGGCAAGCACATCACGCTGGTGTCCAGCGATTCAGGCCTGGGCGTGCGCCAGCTCGGC 1200
 297 Y  G  K  H  I  T  L  V  S  S  D  S  G  L  G  V  R  Q  L  G   316

1201 AGCCTGTCCTCGCCATCGGCCATCACCGTGTCGTCGCAGGGCGAAATCGCGCTGGGCGAC 1260
 317 S  L  S  S  P  S  A  I  T  V  S  S  Q  G  E  I  A  L  G  D   336

1261 GCCACGGTCCAGCGCGGCCCGCTCAGCCTCAAGGGCGCGGGGGTCGTGTCGGCCGGCAAA 1320
 337 A  T  V  Q  R  G  P  L  S  L  K  G  A  G  V  V  S  A  G  K   356
```

FIG. 2. DNA and deduced protein sequences of the 5′ 3,514 nucleotide residues of the FHA structural gene. The boxed regions correspond to the domain homologous to the Sh1A and HpmA genes (box I) or the two repetitive amino acid sequences (boxes II and III). The arrows indicate the putative cleavage sites in the FHA protein. The bold underlined protein sequence corresponds to the determined protein sequence used to generate probe C (11). The potential initiation ATGs discussed in the text are indicated in small boxes. The first *Eco*RI site (corresponding to E[a] in Fig. 1 and 5) and the putative promoter sequence are underlined.

ments 2, 3, 8, and 9 were expressed in the absence as well as in the presence of IPTG. In contrast, the expression of fragments 6 and 7, although observed in the absence of IPTG, was enhanced in the presence of IPTG, indicating that their expression is under the control of the *lac* promoter.

To more precisely locate the translational stop codon, expression of overlapping and contiguous *B. pertussis* sequences was carried out in *E. coli* (fragments 8 to 12; Fig. 1e). *E. coli* transformed with the recombinant plasmids pRIT13197, pRIT13202, pRIT13210, pRIT13296, and pRIT 13338 (Table 1) produced proteins of 175, 270, 145, 190, and 190 kDa, respectively, that were immunodetected with polyclonal anti-FHA antiserum (Fig. 3). Among the different recombinant proteins, the protein encoded by fragment 12 showed an M_r of 190,000, identical to that encoded by fragment 11. For all recombinant clones except the one containing fragment 12, the size of the anti-FHA-reactive polypeptide corresponded well to the size of the DNA insert. These results indicated that the translational stop codon for the FHA ORF is located in the vicinity of the *Cla*I[a] site. Taken together, the results of DNA sequencing and expression allowed us to conclude that the structural gene coding

```
     1321 CTGGCCTCCGGGGGGGGGGCGGTGAACGTCGCGGGCGGCGGGGCGGTGAAGATCGCGTCG 1380
      357 L  A  S  G  G  G  A  V  N  V  A  G  G  G  A  V  K  I  A  S  376

II   1381 GCCAGCAGCGTTGGAAACCTCGCGGTGCAAGGCGGCGGCAAGGTACAGGCCACGCTGTTG 1440
      377 A  S  S  V  G  N  L  A  V  Q  G  G  G  K  V  Q  A  T  L  L  396

     1441 AATGCCGGGGGGACGTTGCTGGTGTCGGGCCGCCAGGCCGTCCAGCTTGGCGCGGCGAGC 1500
      397 N  A  G  G  T  L  L  V  S  G  R  Q  A  V  Q  L  G  A  A  S  416

     1501 AGCCGTCAGGCGCTGTCCGTGAACGCGGGCGGCGCCCTCAAGGCGGACAAGCTGTCGGCG 1560
      417 S  R  Q  A  L  S  V  N  A  G  G  A  L  K  A  D  K  L  S  A  436

     1561 ACGCGACGGGTCGACGTGGATGGCAAGCAGGCCGTCGCGCTGGGGTCGGCCAGCAGCAAT 1620
      437 T  R  R  V  D  V  D  G  K  Q  A  V  A  L  G  S  A  S  S  N  456

     1621 GCGCTGTCGGTGCGTGCCGGCGGCGCCCTCAAGGCGGGCAAGCTGTCGGCGACGGGGCGA 1680
      457 A  L  S  V  R  A  G  G  A  L  K  A  G  K  L  S  A  T  G  R  476

     1681 CTGGACGTGGACGGCAAGCAGGCCGTCACGCTGGGTTCGGTTGCGAGCGACGGTGCGCTG 1740
      477 L  D  V  D  G  K  Q  A  V  T  L  G  S  V  A  S  D  G  A  L  496

     1741 TCGGTAAGCGCTGGCGGAAACCTGCGGGCGAACGAATTGGTCTCCAGTGCCCAACTTGTG 1800
      497 S  V  S  A  G  G  N  L  R  A  N  E  L  V  S  S  A  Q  L  V  516

     1801 GTGCGTGGGCAGCGGGAGGTCGCGCTGGATGACGCTTCGAGCGCACGCGGCATGACCGTG 1860
      517 V  R  G  Q  R  E  V  A  L  D  D  A  S  S  A  R  G  M  T  V  536

     1861 GTTGCCGCAGGAGCGCTGGCGGCCCGCAACCTGCAGTCCAAGGGCGCCATCGGCGTACAG 1920
      537 V  A  A  G  A  L  A  A  R  N  L  Q  S  K  G  A  I  G  V  Q  556

     1921 GGTGGAGAGGCGGTCAGCGTGGCCAACGCGAACAGCGACGCGGAATTGCGCGTGCGCGGG 1980
      557 G  G  E  A  V  S  V  A  N  A  N  S  D  A  E  L  R  V  R  G  576

     1981 CGCGGCCAGGTGGATCTGCACGACCTGAGCGCAGCGCGCGGCGCGGATATCTCCGGCGAG 2040
      577 R  G  Q  V  D  L  H  D  L  S  A  A  R  G  A  D  I  S  G  E  596

     2041 GGGCGCGTCAATATCGGCCGTGCGCGCAGCGATAGCGATGTGAAGGTCTCCGGGCACGGC 2100
      597 G  R  V  N  I  G  R  A  R  S  D  S  D  V  K  V  S  G  H  G  616

     2101 GCCTTGTCGATCGATAGCATGACGGCCCTCGGTGCGATCGGCGTCCAGGCAGGCGGCAGC 2160
      617 A  L  S  I  D  S  M  T  A  L  G  A  I  G  V  Q  A  G  G  S  636

     2161 GTGTCGGCCAAGGATATGCGCAGCCGTGGCGCCGTCACCGTCAGCGGCGGCGGCGCCGTC 2220
      637 V  S  A  K  D  M  R  S  R  G  A  V  T  V  S  G  G  G  A  V  656

     2221 AACCTGGGCGATGTCCAGTCGGATGGCAGGTCCGCGCCACCAGCGCGGGCGCCATGACG 2280
      657 N  L  G  D  V  Q  S  D  G  Q  V  R  A  T  S  A  G  A  M  T  676

     2281 GTGCGAGACGTCGCGGCTGCCGCCGACCTTGCGCTGCAGGCGGGCGACGCGTTGCAGGCC 2340
      677 V  R  D  V  A  A  A  A  D  L  A  L  Q  A  G  D  A  L  Q  A  696

III  2341 GGGTTCCTGAAATCGGCCGGTGCCATGACCGTGAACGGCCGCGATGCCGTGCGACTGGAT 2400
      697 G  F  L  K  S  A  G  A  M  T  V  N  G  R  D  A  V  R  L  D  716

     2401 GGCGCGCACGCGGGCGGGCAATTGCGGGTTTCCAGCGACGGGCAGGCTGCGTTGGGCAGT 2460
      717 G  A  H  A  G  G  Q  L  R  V  S  S  D  G  Q  A  A  L  G  S  736

     2461 CTCGCGGCCAAGGGCGAGCTGACGGTATCGGCCGCGCGCGCGGCGACCGTGGCCGAGTTG 2520
      737 L  A  A  K  G  E  L  T  V  S  A  A  R  A  A  T  V  A  E  L  756

     2521 AAGTCGCTGGACAACATCTCCGTGACGGGCGGCGAACGCGTGTCGGTTCAGAGCGTCAAC 2580
      757 K  S  L  D  N  I  S  V  T  G  G  E  R  V  S  V  Q  S  V  N  776

     2581 AGCGCGTCCAGGGTCGCCATTTCGGCGCACGGCGCGCTGGATGTAGGCAAGGTTTCCGCC 2640
      777 S  A  S  R  V  A  I  S  A  H  G  A  L  D  V  G  K  V  S  A  796
```

FIG. 2—*Continued.*

for the *B. pertussis* FHA extends from just downstream of the *Eco*RI[a] site to the *Cla*I[a] site shown in Fig. 1e. Indeed, expression of DNA fragments 8 and 11 in *E. coli* led to anti-FHA-reactive proteins of 175 and 190 kDa, respectively, which is in good correlation with the coding capacity of the DNA fragments. This result indicated that a 370-kDa anti-FHA-reactive polypeptide is encoded by the DNA fragment extending from the *Eco*RI[a] site to near the *Cla*I[a] site.

Antigenic delineation of the FHA gene. The FHA ORF extends over at least 12,100 bp, with an identified coding capacity for a protein of about 370 kDa (Fig. 1e). This calculated size is well above the estimated size of purified FHA from *B. pertussis* cultures. As indicated above, the amino-terminal and the middle portions of the FHA protein reacted with anti-FHA polyclonal antiserum. To investigate whether the carboxy-terminal region is also antigenic with respect to the same polyclonal antiserum, two adjacent DNA fragments (fragment 13 [*Xho*I[c]-*Eco*RI[b]] and fragment 14 [*Eco*RI[b]-*Cla*I[a]]; Fig. 1e) were subcloned in all three reading frames in the pUC vectors. However, Western blot analysis of the proteins encoded by these plasmids (pRIT13338, pRIT

```
2641  AAGAGCGGTATCGGGCTCGAAGGCTGGGGCGCGGTCGGAGCGGACTCCCTCGGTTCCGAC   2700
 797  K  S  G  I  G  L  E  G  W  G  A  V  G  A  D  S  L  G  S  D     816

2701  GGCGCGATCAGCGTGTCCGGGCGCGATGCGGTCAGGGTCGATCAAGCCCGCAGTCTTGCC   2760
 817  G  A  I  S  V  S  G  R  D  A  V  R  V  D  Q  A  R  S  L  A     836

2761  GACATTTCGCTGGGGGCGGAAGGCGGCGCCACGCTGGGCGCGGTGGAGGCCGCCGGTTCG   2820
 837  D  I  S  L  G  A  E  G  G  A  T  L  G  A  V  E  A  A  G  S     856

2821  ATCGACGTGCGCGGCGGATCCACGGTGGCGGCGAACTCGCTGCACGCCAATCGCGACGTT   2880
 857  I  D  V  R  G  G  S  T  V  A  A  N  S  L  H  A  N  R  D  V     876

2881  CGGGTCAGCGGCAAGGATGCGGTGCGCGTAACGGCCGCCACCAGCGGGGGCGGTCTGCAT   2940
 877  R  V  S  G  K  D  A  V  R  V  T  A  A  T  S  G  G  G  L  H     896

2941  GTGTCGAGCGGCCGCCAGCTCGATCTGGGCGCCGTGCAGGCGCGCGGCGCGCTGGCCCTG   3000
 897  V  S  S  G  R  Q  L  D  L  G  A  V  Q  A  R  G  A  L  A  L     916

3001  GACGGAGGCGCCGGCGTGGCGCTGCAATCGGCCAAGGCTAGCGGCACGCTGCATGTGCAG   3060
 917  D  G  G  A  G  V  A  L  Q  S  A  K  A  S  G  T  L  H  V  Q     936

3061  GGCGGCGAGCACCTGGACCTGGGCACGTTGGCCGCCGTAGGGGCGGTGGACGTCAATGGC   3120
 937  G  G  E  H  L  D  L  G  T  L  A  A  V  G  A  V  D  V  N  G     956

3121  ACGGGAGACGTGCGCGTTGCGAAGCTGGTGAGCGATGCAGGCGCCGATCTGCAAGCGGGG   3180
 957  T  G  D  V  R  V  A  K  L  V  S  D  A  G  A  D  L  Q  A  G     976

3181  CGCTCCATGACGCTGGGTATCGTCGACACGACCGGCGATCTGCAGGCGCGCGCGCAGCAG   3240
 977  R  S  M  T  L  G  I  V  D  T  T  G  D  L  Q  A  R  A  Q  Q     996

3241  AAGCTGGAGCTCGGGTCGGTTAAGAGCGATGGCGGCCTTCAGGCCGCCGCCGGCGGGGCC   3300
 997  K  L  E  L  G  S  V  K  S  D  G  G  L  Q  A  A  A  G  G  A    1016

3301  CTCAGCCTGGCGGCGGCGGAAGTCGCAGGGGCGCTGGAGCTCTCGGGCCAGGGCGTCACC   3360
1017  L  S  L  A  A  A  E  V  A  G  A  L  E  L  S  G  Q  G  V  T    1036

3361  GTGGACAGAGCCAGCGCTAGCCGGGCACGCATCGACAGCACCGGTTCGGTCGGCATCGGC   3420
1037  V  D  R  A  S  A  S  R  A  R  I  D  S  T  G  S  V  G  I  G    1056

3421  GCGCTGAAGGCAGGCGCTGTCGAGGCCGCCTCGCCACGGCGGGCGCGCCGCGCGCTGCGG   3480
1057  A  L  K  A  G  A  V  E  A  A  S  P  R  R  A  R  R  A  L  R    1076

3481  CAGGATTTCTTCACGCCCGGCAGCGTGGTGGTCC  3514
1077  Q  D  F  F  T  P  G  S  V  V  V    1087
```

FIG. 2—*Continued.*

13335, pRIT13336, pRIT13337, pRIT13272, pRIT13274, and pRIT13276; Table 1) revealed that they did not react with the anti-FHA polyclonal or monoclonal antibodies, nor did they react with polyclonal antiserum raised against whole *B. pertussis* organisms. This lack of reaction cannot be attributed to the lack of expression, since in one case (pRIT13276) one of the frames resulted in the production of a polypeptide clearly identifiable by Coomassie blue staining after SDS-PAGE. Synthesis of this polypeptide was inducible by IPTG (not shown). Moreover, the observed M_r of the produced polypeptide corresponded well with the M_r calculated from the size of the inserted DNA. These data suggest that DNA fragments 13 and 14 in Fig. 1e encode portions of the FHA protein that are not immunologically recognized or not present in the mature FHA.

To analyze the probable processing steps leading to mature FHA, various recombinant FHA-related proteins, as well as purified FHA, were analyzed by Western blot, using monoclonal anti-FHA antibodies. The five recombinant proteins used correspond to fragments 3, 6, 7, 10, and 11 (Fig. 1e). The results of this Western blot analysis except for fragment 6 are shown in Fig. 4. All of the monoclonal antibodies recognized fragment 7, suggesting that this fragment constitutes a highly immunogenic portion of the FHA molecule, at least in mice. By using the recombinant proteins as antigens, the monoclonal antibodies can be subdivided into two groups: those reacting with fragment 7 but not with fragments 3, 6, 10, and 11, and those reacting with fragments 7, 10, and 11 but not with fragments 3 and 6. Therefore, the epitopes for all of these monoclonal antibodies are located in the portion of the FHA molecule that is encoded by the *Xho*I^{a}-*Pst*I^{b} DNA fragment depicted in Fig. 1e.

FHA purified from the *B. pertussis* culture medium is usually composed of a major 220-kDa polypeptide and three minor (140-, 130-, and 90-kDa) polypeptides that can be resolved by SDS-PAGE (11). All of the monoclonal antibodies reacted strongly with the 220-kDa major polypeptide. Several monoclonal antibodies (12.5A9, 12.5D1, 13.1F5, 13.4A3, and 13.6E2) did not react with the 90-kDa protein. These antibodies reacted also with expression products of DNA fragment 7 but not of fragment 3, 6, 10, or 11. One antibody (12.1F9) reacted with both the 220- and the 90-kDa proteins as well as with the expression product of fragment 7, indicating that the components of the purified FHA are related and that they are also related to recombinant fragment 7. Two antibodies (12.1D3 and 12.2B11) reacted to the 220- and the 90-kDa polypeptides of purified FHA as well as to peptides coded by recombinant fragments 7, 10, and 11, whereas antibody 12.6F8 reacted with peptides coded by recombinant fragments 7, 10, and 11 and with the 220-kDa but not the 90-kDa polypeptide of purified FHA. This result indicated that the overlap of the 220- and 90-kDa polypeptides terminates within the region of FHA encoded by the *Blg*II^{b}-*Pst*I^{b} fragment. Taken together, these results allow us to subdivide the monoclonal antibodies into four groups (Table 2), to map the locations of their epitopes on the FHA

TABLE 1. *E. coli* plasmids encoding β-galactosidase-FHA fusion proteins[a]

Recombinant plasmid	DNA fragment[b]	Restriction site[c] 5′ end	Restriction site[c] 3′ end	Vector	Molecular size (kDa) of immunodetected protein
pRIT13200	1	Bg[a]	E[a]	pUC9	
pRIT13204	1	Bg[a]	E[a]	pUC9-1	
pRIT13206	1	Bg[a]	E[a]	pUC9-2	
pRIT13130	2	E[a]	S	pUC8	50
pRIT13122	3	E[a]	P[a]	pUC8	80
pRIT13124	3	E[a]	P[a]	pUC8-1	
pRIT13126	3	E[a]	P[a]	pUC8-2	
pRIT13120	4	S	B[a]	pUC9-2	60[d]
pRIT13118	4	S	B[a]	pUC9-1	
pRIT13115	4	S	B[a]	pUC9	
pRIT13106	5	P[a]	B[a]	pUC9-2	30[d]
pRIT13105	5	P[a]	B[a]	pUC9-1	
pRIT13059	5	P[a]	B[a]	pUC9	
pRIT13031	6	B[a]	X[a]	pUC8	24
pRIT13033	7	B[a]	P[b]	pUC8	150[d]
pRIT13197	8	E[a]	Bg[b]	pUC8	175
pRIT13202	9	E[a]	E[b]	pUC8	>270
pRIT13210	10	Bg[b]	E[b]	pUC9	145
pRIT13212	10	Bg[b]	E[b]	pUC9-1	
pRIT13214	10	Bg[b]	E[b]	pUC9-2	
pRIT13296	11	Bg[b]	C[a]	pUC8-1	190
pRIT13338	12	Bg[b]	Ev	pUC8-1	190
pRIT13335	13	X[c]	E[b]	pUC9	
pRIT13336	13	X[c]	E[b]	pUC9-1	
pRIT13337	13	X[c]	E[b]	pUC9-2	
pRIT13272	14	E[b]	C[a]	pUC8	
pRIT13274	14	E[b]	C[a]	pUC8-1	
pRIT13276	14	E[b]	C[a]	pUC8-2	

[a] Formed by the N-terminal region of *E. coli* β-galactosidase (from 4 to 14 N-terminal amino acids), followed by the FHA segment.
[b] See Fig. 4.
[c] Bg, *Bgl*II; B, *Bam*HI; E, *Eco*RI; S, *Sal*I; P, *Pst*I; X, *Xho*I; C, *Cla*I; Ev, *Eco*RV.
[d] Has the C-terminal extremity of the FHA fragment in frame with the α peptide of β-galactosidase; this increases the molecular size of the fusion by about 15 kDa.

protein, and to align the antigenic structure of FHA with its structural gene (Fig. 5).

The portion encoded by the *Xho*I[a]-*Pst*I[b] DNA fragment is highly immunogenic and contains at least four distinct epitopes immunodominant in mice. The portion of FHA encoded by the *Eco*RI[a]-*Xho*I[a] DNA fragment was not recognized by the monoclonal antibodies used here but was recognized by polyclonal anti-FHA antibodies. Finally, the carboxy-terminal portion of the FHA appears not to react with antibodies produced in rabbits, mice, or guinea pigs.

DISCUSSION

The *B. pertussis* FHA molecule is a likely component of the new generation vaccines against whooping cough (12, 29). To characterize this antigen at the molecular level, its structural gene was cloned in *E. coli*. By using three different DNA probes, four DNA fragments covering the entire structural gene were isolated. The 5′ portion of the gene was sequenced up to the region that hybridized to probe C. This sequence is identical to the one reported by Relman et al. (25) except at seven nucleotide positions (1089, 1090, 1103, 1104, 1799, 2093, and 3285), leading to four amino acids changes (Arg-280, Ala-284, Val-516, and Gly-614). A single long ORF was found that initiated with one of three possible in-frame ATG codons indicated in Fig. 2. Termination

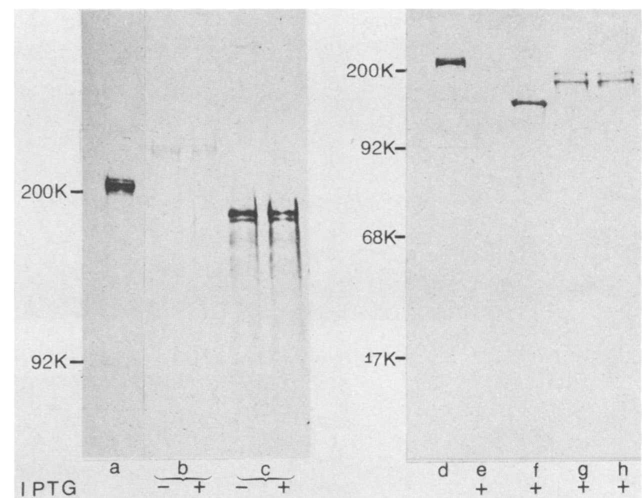

FIG. 3. Immunoblot analysis of the polypeptides encoded by the *B. pertussis* DNA fragments. Lanes: a and d, purified FHA; b, c, f, g, and h, FHA polypeptides encoded by DNA fragment 9, 8, 10, 11, and 12, respectively (Table 1 and Fig. 1); e, extract from *E. coli* with pUC vector. + and − indicate to the presence and absence, respectively, of IPTG in the culture medium. Numbers in the margins represent the M_r standards (in thousands [K]).

codons were found in all three reading frames upstream of the first ATG codon. None of the ATG codons were preceded by sequences homologous to the Shine-Dalgarno ribosome-binding sequences. Analysis by the statistical TESTCODE algorithm (7) to differentiate between true protein-coding sequences and fortuitous ORFs indicated that the ORF starting with the first of the three ATG codons has a high coding probability. This finding is in agreement with those of Relman et al. (25). A promoterlike structure was identified that contained the sequences TTGACTA and TATAAAT at positions 83 and 111, respectively, downstream of the first *Eco*RI site. The distance between these two sequences, which have homology to other bacterial promoter elements, is 21 nucleotides. This distance is significantly greater than that found in efficient *E. coli* promoters; however, the same distance was found in other *B. pertussis* promoters such as the PTX promoter (18). A sequence resembling somewhat a procaryotic signal peptide can be found in the near amino terminus of the proposed protein sequence, with a positively charged region from amino acids 33 to 43, a hydrophobic stretch from 49 to 68, and the (−3, −1) rule followed in the triplet at 69 to 71 (36). The hydrophobic region between residues 52 and 69 has previously been noticed by Relman et al. (25). The cleavage site of this unusually long signal peptide would then be located after Ala-71. A shorter signal peptide, but still longer than usual, could result from initiating the translation at the second Met residue. Unsuccessful attempts were made to sequence the amino terminus of the mature FHA protein, which seems to be blocked. It cannot be decided at this time whether the sequence features mentioned above characterize a true signal peptide or an uncleaved one, which would then act as an anchor of the antigen into the membrane. The amino-terminal portion of the FHA protein shows striking homology to the amino-terminal region of the hemolysins A of *S. marcescens* (24) and *P. mirabilis* (34). The homology to the *P. mirabilis* hemolysin is particularly interesting since, unlike *S. marcescens*, *P. mirabilis* is very different in C+G content (38%) from *B. pertussis* (67%). Therefore, the ho-

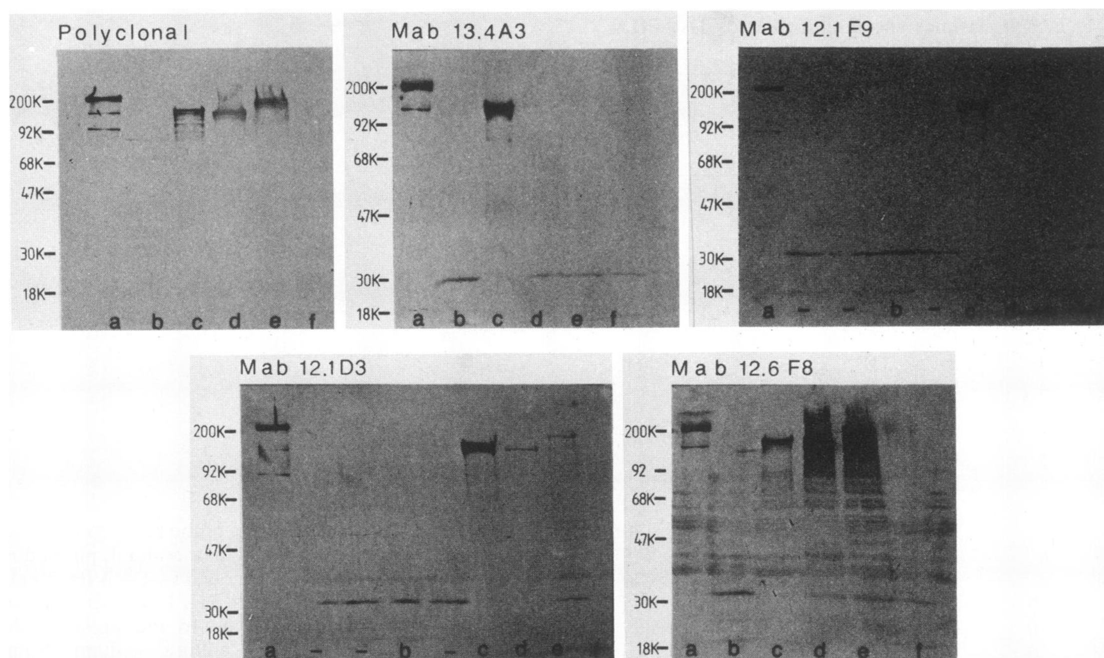

FIG. 4. Immunoblot analysis of recombinant FHA polypeptides, using monoclonal antibodies 13.4A3, 12.1F9, 12.1D3, 12.6F8, which represent one example of each group described in Table 2. Lanes: a, purified FHA; b to e, *E. coli* extracts with recombinant FHA polypeptides encoded by DNA fragments 3, 7, 10, and 11, respectively; f, extract from *E. coli* with the pUC vector. Numbers in the margins represent M_r standards (in thousands [K]).

mology at the amino acid sequence level can be considered highly significant. The sequence starting from amino acid residue 53 of the *S. marcescens* hemolysin and extending to residue 168 shows 48% identity with the *B. pertussis* FHA sequence, and most amino acid substitutions are conservative changes. The *P. mirabilis* hemolysin shares 47% identity with the same FHA region. The two hemolysins share 72% identity in that region (Fig. 6a; Fig. 2, box I).

Schiebel et al. (31) have recently shown that the secretion and activation mechanism of ShlA is dependent on the presence of the *shlB* gene product. In addition, ShlB-mediated secretion and activation of ShlA require the 261 amino-terminal residues of ShlA, indicating that the information for export to the cell surface and secretion into the culture medium is located in the amino-terminal domain of the protein. The high degree of homology of this region with the

TABLE 2. Immunoblot analysis with anti-FHA monoclonal antibodies

Group	Monoclonal antibody	Reaction with: Recombinant FHA protein 3	6	7	10	11	Purified FHA polypeptide 90 kDa	220 kDa
A	12.5A9							
	12.5D1							
	13.1F5	−	−	+	−	−	−	+
	13.4A3							
	13.6E2							
B	12.1F9	−	−	+	−	−	+	+
C	12.1D3 12.2B11	−	−	+	+	+	+	+
D	12.6F8	−	−	+	+	+	−	+

B. pertussis FHA amino-terminal portion suggests that the secretion of FHA also may be mediated by such a mechanism and therefore depend on other gene products. FHA$^-$ mutants with the intact FHA structural gene have indeed been identified. TnPhoA fusions made in *B. pertussis* that lead to a defective FHA phenotype and secretion of the *E. coli* phosphatase have been obtained (T. Finn and J. Mekalanos, unpublished data). However, these FHA$^-$ mutants are not affected in the FHA structural gene (T. Cabezon, unpublished results), and their involvement in secretion of the mature FHA awaits further investigation.

Furthermore, it was recently reported that the hemolysin activity is largely due to the carboxy-terminal domain of ShlA and that the erythrocyte-binding domain is located in the amino-terminal part of the polypeptide (30). It is therefore conceivable that the receptor-binding domain of the FHA protein is also located in the homologous amino-terminal portion. In that case, the receptors for the *B. pertussis* FHA and the hemolysins of *S. marcescens* and *P. mirabilis* may have similar features. Relman et al. (25) have proposed that the FHA region encoded by the DNA fragment between the two *Bam*HI sites (a and b) in Fig. 1 is required for binding of the FHA molecule. In addition, they found an RGD sequence and suggested that existence of an RGD-dependent fibronectinlike binding for FHA to integrin-related molecules. Using the recombinant protein encoded by DNA fragment 7 (Fig. 1e), we were unable to detect any hemagglutinating activity. Furthermore, the recombinant protein was also unable to inhibit FHA-mediated hemagglutination (data not shown). These findings suggest that although the RGD region of the FHA protein may be required for FHA activity on some cells (e.g., leukocytes), it is not sufficient. Alternatively, FHA may have developed more than one mechanism of binding, and different mechanisms may be used on different target cells. It is therefore tempting

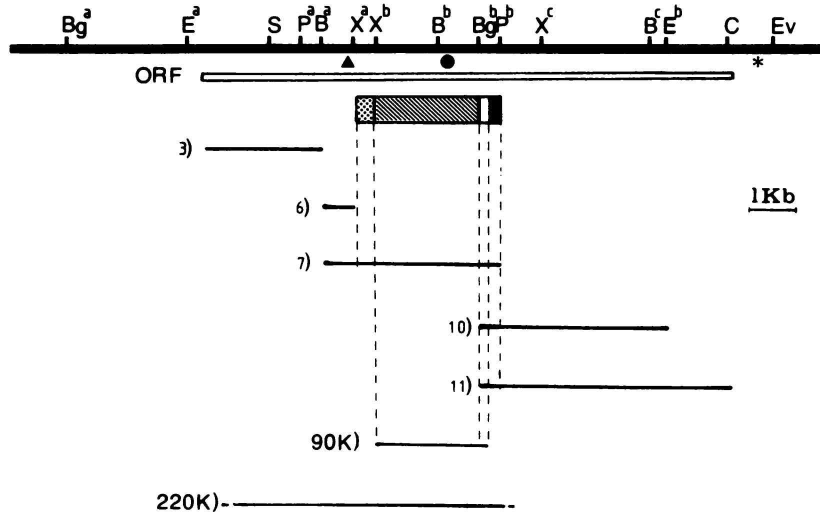

FIG. 5. Schematic representation of the FHA-related proteins recognized by the monoclonal antibodies shown in Table 2. ▩, ▨, □, and ■ represent groups A, B, C, and D respectively.

to speculate that the amino-terminal portion of the FHA molecule is also involved in the binding activity of FHA and may represent a second, independent carbohydrate recognition site.

Following the region of homology with the hemolysins, the FHA protein contains two regions of repeats. However, unlike the repeats in the adenylate cyclase or hemolysin protein of *B. pertussis* (8), the FHA repeats contain about 40

```
a) FHA    91  VVNIadPNsgGvSHNkFQqFNVanPGvVFNNgLtdGvSrigGaLtkNPNL  140
   Shla   53  VVdIVaPNgnGlSHNQYQDFNVnqPGAVlNNsrEAGlSQLAGqLgANPNL  102
   Hpma   52  ViNIVtPNneGiSHNQYQDFNVgkPGAVFNNaLEAGqSQLAGhLnANsNL  101

   Consensus  VvnIv-PN--G-SHNqYQdFNV--PGaVfNN-leaG-SqlaG-L-aNpNL

   FHA   141  .tRqASaILaEVtdtsPSrLaGtlEVyGkgADliiANPNGISvnGlstlN  189
   Shla  103  gGReASvILNEVigRNPSlLhGQQEiFGmAADYVLANPNGISCqsCGFIN  152
   Hpma  102  nGqaASlILNEVvsRNPSfLlGQQEVFGiAAeYVLsNPNGItCdGCGFIN  151

   Consensus  -gr-AS-ILnEV--rnPS-L-GqqEvfG-aAdyvlaNPNGIsc-gcgfiN

   FHA   190  aSnltLttGrPsVngG  205
   Shla  153  TShSSLVVGNPLVENG  168
   Hpma  152  TSrSSLVVGNPLfENG  167

   Consensus  tS-ssLvvGnPlvenG

b)       382  nLaVqqGGkvqAtlLnagGtLlVsGrQAVqLGaASSr  418
         420  ALSVnAGGALkAdkLsatrrvdVdGkQAVaLGsASSn  456
         457  ALSVrAGGALkAgkLsatGrLdVdGkQAVtLGsvaSD  493
         495  ALSVsAGGnLrAneLvssaqLvVrGqreVaLddASSa  531
         533  gmtVvAaGALaArnLqskGaigVqGgeAVsvanAnSD  569
         628  AigVqAGGsvsAkdmrsrGavtVsGggAVnLGdvqSD  664

   Consensus  alsV-agGal-A--l---g-l-V-G-qaV-lg-asSd

c)       701  SaGAmtVnGRDAVRlDgAhaGGqLrVSSdGQAaLGslaAkG  741
         815  SdGAisVSGRDAVRVDqArSladislgaeGgAtLGAVeAaG  855
         872  anrdvrVSGkDAVRVtaAtSGGgLhVSSgrQldLGAVqArG  912

   Consensus  s-ga--VsGrDAVRvd-A-sgg-l-vss-gqa-LGav-A-G
```

FIG. 6. Amino acid homologies in the N-terminal 1,087 residues of FHA. (a) Hemolysin homology; (b and c) FHA repeated regions. The amino acid sequence of the *S. marcescens* hemolysin (Shla) was reported by Pool et al. (24), and the sequence of the *P. mirabilis* hemolysin (Hpma) was reported by Uphoff and Welch (34).

amino acid residues per repeat unit. The first region of repeats extends from residues 382 to 664 and contains six repeat units (Fig. 6b; Fig. 2, box II). The second region extends from residues 701 to 912 and contains three repeat units (Fig. 6c; Fig. 2, box III). Although Relman et al. (25) also have identified repeats in the FHA structural gene, they are of a different kind than those identified in this study. It is not unusual for filamentous proteins to contain repetitive sequences; however, the role of these repeats is not known. Whether the repeats in the FHA protein are also important for its filamentous nature remains to be investigated.

Analysis of the FHA protein and the expression products of the different FHA DNA fragments in *E. coli* indicated that the carboxy-terminal third of the molecule does not react with polyclonal or monoclonal anti-FHA antibodies, nor does it react with anti-*B. pertussis* polyclonal antibodies. This absence of immunoreactivity may be interpreted as immunosilence of an important portion of the molecule with respect to antibodies. Alternatively, it is possible that the carboxy moiety of the FHA ORF is processed during the synthesis or assembly of mature FHA. This view is consistent with the fact that the isolated FHA protein from *B. pertussis* cultures has a molecular weight of 220,000, whereas the ORF has a coding capacity for a 370-kDa protein. The processing of the FHA molecule seems to be a complex multistep event.

Comparison of the DNA-derived protein sequence and the partial sequence published previously (11) indicated that trypsinlike proteolytic cleavage occurs after the RRARR sequence at amino acid position 1073 also noted by Relman et al. (25). Therefore, the smaller peptides in the purified FHA fraction are likely proteolytic cleavage products of the larger 220-kDa polypeptide. It has been previously reported that depending on storage conditions, the purified FHA appears to be composed by polypeptides with molecular masses ranging from 60 to 220 kDa. The different polypeptides species shared antigenic determinants when analyzed by specific anti-FHA monoclonal antibodies (10). Epitope mapping of the recombinant proteins and the polypeptides from purified *B. pertussis* FHA confirmed this finding and further suggested that the 90-kDa polypeptide represents the middle portion of the immature FHA molecule. Other FHA polypeptides (130 and 140 kDa) were also recognized by some of the monoclonal antibodies, indicating that they are also internal fragments of the FHA.

It is striking that all of the monoclonal antibodies reacted with the recombinant protein encoded by DNA fragment 7, the middle region of the FHA structural gene, suggesting that this region may be an important immunodominant domain of FHA. At least four different epitopes could be identified in this domain. Since the same region seems to be required for adherence to ciliated cells, the antibodies against it may be important in the prevention of binding of FHA to ciliated cells, and their putative role as antiadherence antibodies in the immune response against *B. pertussis* infection or by vaccination should be investigated. The monoclonal antibodies are now being used in mouse protective models as well as for cell-binding inhibition studies in vitro. We are also now investigating the potential of the recombinant proteins encoded by the different DNA fragments described in Fig. 1e in eliciting protective immunity to *B. pertussis* infection.

ACKNOWLEDGMENTS

We thank C. Van Roy and H. Monnig for typing, N. Harford, T. Rutgers, and J. Cohen for critical reading of the manuscript, E. Simoen for polyclonal antisera, B. Colau for oligonucleotides synthesis, and C. Capiau for providing purified FHA.

This work was supported by the Institut pour l'Encouragement de la Recherche Scientifique dans l'Industrie et l'Agriculture.

LITERATURE CITED

1. **Barbieri, J. T., R. Rappuoli, and R. J. Collier.** 1987. Expression of the S-1 catalytic subunit of pertussis toxin in *Escherichia coli*. Infect. Immun. **55**:1321–1323.
2. **Brown, D. R., and C. D. Parker.** 1987. Cloning of the filamentous hemagglutinin of *Bordetella pertussis* and its expression in *Escherichia coli*. Infect. Immun. **55**:154–161.
3. **Burnette, W. N., V. L. Mar, W. Cieplak, C. F. Morris, K. T. Kaljot, K. S. Marchitto, R. K. Sachdev, C. Locht, and J. M. Keith.** 1988. Direct expression of *Bordetella pertussis* toxin subunits to high levels in *Escherichia coli*. Bio/Technology **6**:699–706.
4. **Chen, E. Y., and P. H. Seeburg.** 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4**:165–170.
5. **Devereux, J., P. Haeberli, and D. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
6. **Duffaud, G., and M. Inouye.** 1988. Signal peptidases recognize a structural feature at the cleavage site of secretory proteins. J. Biol. Chem. **263**:10224–10228.
7. **Fickett, J. W.** 1982. Recognition of protein coding regions in DNA sequences. Nucleic Acids Res. **10**:5303–5318.
8. **Glaser, P., D. Ladant, O. Sezer, F. Pichot, A. Ullmann, and A. Danchin.** 1988. The calmodulin-sensitive adenylate cyclase of *Bordetella pertussis*: cloning and expression in *Escherichia coli*. Mol. Microbiol. **2**:19–30.
9. **Hanna, Z., C. Frageau, G. Prefontaine, and R. Brousseau.** 1984. Construction of a family of universal expression plasmid vectors. Gene **30**:247–250.
10. **Irons, L. I., L. A. E. Ashworth, and P. Wilton-Smith.** 1983. Heterogeneity of the filamentous haemagglutinin of *Bordetella pertussis* studied with monoclonal antibodies. J. Gen. Microbiol. **129**:2769–2778.
11. **Jacob, F., C. Capiau, A. M. Gathoye, C. Locht, and T. Cabezon.** 1988. Molecular cloning of the filamentous hemagglutinin structural gene from *Bordetella pertussis* in *Escherichia coli*, p. 62–68. *In* S. Mebel, H. Stompe, M. Drescher, and S. Rustenbach (ed.), FEMS Symposium on Pertussis. Berlin, German Democratic Republic.
12. **Kallings, L. O., P. Olin, J. Storsaeter, and ad hoc group for the study of pertussis vaccines.** 1988. Placebo-controlled trial of two acellular pertussis vaccines in Sweden—protective efficacy and adverse events. Lancet **i**:955–960.
13. **Knapp, S., and J. J. Mekalanos.** 1988. Two *trans*-acting regulatory genes (*vir* and *mod*) control antigenic modulation in *Bordetella pertussis*. J. Bacteriol. **170**:5059–5066.
14. **Köhler, G., and C. Milstein.** 1975. Continuous cultures of fused cells secreting antibody of predefined specificity. Nature (London) **256**:415–421.
15. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
16. **Locht, C., P. A. Barstad, J. E. Coligan, L. Mayer, J. J. Munoz, S. G. Smith, and J. M. Keith.** 1986. Molecular cloning of pertussis toxin genes. Nucleic Acids Res. **14**:3251–3261.
17. **Locht, C., W. Cieplak, K. S. Marchitto, H. Sato, and J. M. Keith.** 1987. Activities of complete and truncated forms of pertussis toxin subunits S1 and S2 synthesized in *Escherichia coli*. Infect. Immun. **55**:2546–2553.
18. **Locht, C., and J. M. Keith.** 1986. Pertussis toxin gene: nucleotide sequence and genetic organization. Science **232**:1258–1264.
19. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. **Mattei, D., F. Pichot, J. Bellalou, O. Mercereau-Puijalon, and A. Ullmann.** 1986. Molecular cloning of a coding sequence of *Bordetella pertussis* filamentous hemagglutinin gene. FEMS

Microbiol. Lett. **37**:73–77.

21. **Nicosia, A., A. Bartoloni, M. Perugini, and R. Rappuoli.** 1987. Expression and immunological properties of the five subunits of pertussis toxin. Infect. Immun. **55**:963–967.
22. **Nicosia, A., M. Perugini, C. Franzini, M. C. Casagli, M. G. Borri, G. Antoni, M. Almoni, P. Neri, G. Ratti, and R. Rappuoli.** 1986. Cloning and sequencing of the pertussis toxin genes: operon structure and genetic duplication. Proc. Natl. Acad. Sci. USA **83**:4631–4635.
23. **Oda, A., J. L. Cowell, D. G. Burstyn, and C. R. Manclark.** 1984. Protective activities of filamentous hemagglutinin and the lymphocytosis-promoting factor of *Bordetella pertussis* in mice. J. Infect. Dis. **150**:823–833.
24. **Poole, K., E. Schiebel, and V. Braun.** 1988. Molecular characterization of the hemolysin determinant of *Serratia marcescens*. J. Bacteriol. **170**:3177–3188.
25. **Relman, D. A., M. Domenighini, E. Tuomanen, R. Rappuoli, and S. Falkow.** 1989. Filamentous hemagglutinin of *Bordetella pertussis*: nucleotide sequence and crucial role in adherence. Proc. Natl. Acad. Sci. USA **86**:2637–2641.
26. **Runeberg-Nyman, K., O. Engström, S. Löfdahl, S. Ylöstalo, and M. Sarvas.** 1987. Expression and secretion of pertussis toxin subunit S1 in *Bacillus subtilis*. Microb. Pathog. **3**:461–468.
27. **Sambrook. J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning, a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
28. **Sato, H., and Y. Sato.** 1984. *Bordetella pertussis* infection in mice: correlation of specific antibodies against two antigens, pertussis toxin and filamentous hemagglutinin, with mouse protectivity in an intracerebral or aerosol challenge system. Infect. Immun. **46**:415–421.
29. **Sato, Y., H. Fukumi, and M. Kimura.** 1984. Development of a pertussis component vaccine in Japan. Lancet **i**:122–126.
30. **Schiebel, E., and V. Braun.** 1989. Integration of the *Serratia marcescens* heamolysin into human erythrocyte membranes. Mol. Microbiol. **3**:445–453.
31. **Schiebel, E., H. Schwarz, and V. Braun.** 1989. Subcellular location and unique secretion of the hemolysin of *Serratia marcescens*. J. Biol. Chem. **264**:16311–16320.
32. **Stainer, D. W.** 1988. Growth of *Bordetella pertussis*, p. 19–37 *in* A. C. Wardlaw and R. Parton (ed.), Pathogenesis and immunity in pertussis. J. Wiley & Sons, Chichester, United Kingdom.
33. **Tuomanen, E., and A. A. Weiss.** 1985. Characterization of two adhesins of *Bordetella pertussis* for human ciliated respiratory-epithelial cells. J. Infect. Dis. **152**:118–125.
34. **Uphoff, T. S., and R. A. Welch.** 1990. Nucleotide sequencing of the *Proteus mirabilis* calcium-independent hemolysin genes (*hpmA* and *hpmB*) reveals sequence similarity with the *Serratia marcescens* hemolysin genes (*shlA* and *shlB*). J. Bacteriol. **172**:1206–1216.
35. **Vieira, J., and J. Messing.** 1982. The pUC plasmids and M13 mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:259–268.
36. **von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. **14**:4683–4690.
37. **Weiss, A. A., and S. Falkow.** 1984. Genetic analysis of phase change in *Bordetella pertussis*. Infect. Immun. **43**:263–269.
38. **Weiss, A. A., E. L. Hewlett, G. A. Myers, and S. Falkow.** 1983. Tn*5*-induced mutations affecting virulence factors of *Bordetella pertussis*. Infect. Immun. **42**:33–41.