

Molecular characterization of the mycobacterial heparin-binding hemagglutinin, a mycobacterial adhesin

FRANCO D. MENOZZI*, RAINER BISCHOFF†‡, EMMANUELLE FORT*, MICHAEL J. BRENNAN§, AND CAMILLE LOCHT*¶

*Laboratoire de Microbiologie Génétique et Moléculaire, Institut National de la Santé et de la Recherche Médicale U447, Institut Pasteur de Lille, 1 rue Calmette, F-59019 Lille Cedex, France; †Transgène s.a., 11 rue de Molsheim, F-67082 Strasbourg Cedex, France; and §Laboratory of Mycobacteria, Center for Biologics Evaluation and Research, Food and Drug Administration, Rockville, MD 20852

Communicated by John B. Robbins, National Institute of Child Health and Human Development, Bethesda, MD, June 10, 1998 (received for review August 11, 1997)

ABSTRACT Although it generally is accepted that the interaction of *Mycobacterium tuberculosis* with alveolar macrophages is a key step in the pathogenesis of tuberculosis, interactions with other cell types, especially epithelial cells, also may be important. In this study we describe the molecular characterization of a mycobacterial heparin-binding hemagglutinin (HBHA), a protein that functions as an adhesin for epithelial cells. The structural gene was cloned from *M. tuberculosis* and bacillus Calmette–Guérin, and the sequence was found to be identical between the two species. The calculated M_r was smaller than the observed M_r when analyzed by SDS/PAGE. This difference can be attributed to the Lys/Pro-rich repeats that occur at the C-terminal end of the protein and to a putative carbohydrate moiety. Glycosylation of HBHA appears to protect the protein from proteolytic degradation, which results in the removal of the C-terminal Lys/Pro-rich region responsible for binding of HBHA to sulfated carbohydrates. Evidence suggests that glycosylation is also important for HBHA-mediated hemagglutination and for certain immunologic properties of the protein. Finally, the absence of a signal peptide in the coding region of HBHA raises the possibility that this protein is not secreted via the general secretion pathway.

Infection by the members of the *Mycobacterium tuberculosis* complex still remains an important cause of morbidity and mortality worldwide. Tuberculosis causes about 3 million annual deaths, and 8–10 million new cases are reported yearly (1). In addition to being a threat to human health, mycobacterial diseases also have a serious economical impact because of their importance in veterinary medicine. Other mycobacterial species, such as the members of the *M. avium/intracellulare* complex now are recognized as frequent opportunistic agents infecting immunocompromised individuals (2).

Specific interactions with pulmonary macrophages have been recognized as an important step in the pathogenesis of tuberculosis. Early interactions with other cells or with the extracellular matrix are less well documented, although binding to and invasion of HeLa cells by *M. tuberculosis* have been known since 1957 (3). More recently, it has been suggested that attachment of mycobacteria to epithelial cells and to extracellular matrix is related to virulence and to extrapulmonary dissemination of tubercle bacilli (4–7). Whereas binding of mycobacteria to macrophages involves mannose receptors and integrins such as the complement receptors (8–10), attachment to epithelial cells appears to occur through interactions with sulfated glycoconjugates (11). Similar to other pathogenic bacteria (12–16), viruses (17–19), or parasites (20–22), *M.*

tuberculosis and *M. bovis* express on their surface a heparin-binding hemagglutinin (HBHA), which is involved in attachment of the mycobacteria to epithelial cells (11). This 28-kDa protein also induces bacterial auto-aggregation, and patients with active tuberculosis produce antibodies against HBHA, indicating that the HBHA-encoding gene is expressed during the course of mycobacterial infections.

We describe here the cloning of the gene coding for HBHA from *M. bovis* bacillus Calmette–Guérin (BCG) and *M. tuberculosis* and the partial characterization of the gene product. We provide evidence that HBHA is a bacterial glycoprotein and that glycosylation may be crucial for the stability of the protein and for some of its immunologic properties. Although HBHA is a surface-exposed protein, the structural gene does not encode an amino-terminal signal sequence, suggesting a *sec*-independent mechanism of export. Finally, the domain responsible for the interactions of the protein with sulfated glycoconjugates was localized to the C-terminal region, which is composed of lysine-rich repeats.

MATERIALS AND METHODS

Bacterial Strains, Growth Conditions, and DNA Manipulations. Growth conditions for *M. bovis* BCG (strain 1173P2; World Health Organization, Stockholm, Sweden) and *M. tuberculosis* strains have been described elsewhere (11). *M. tuberculosis* H37Rv and H37Ra came from the culture collection of the Institut Pasteur. Briefly, the mycobacteria were cultured at 37°C in static conditions using 175-cm² Roux flasks that contained 150 ml of Sauton medium supplemented with Triton WR1339. All cloning steps were performed in *Escherichia coli* XL1-Blue (New England Biolabs) grown in Luria–Bertani broth (23). After transformation with pKK388–1 (CLONTECH), pUC18 (Boehringer Mannheim), or derivatives of these plasmids, antibiotic-resistant *E. coli* XL1-Blue was selected with ampicillin (150 µg/ml). Restriction enzymes, T4 DNA ligase, and other molecular biology reagents were purchased from Boehringer Mannheim, New England Biolabs, or Promega and were used as recommended by the suppliers. All DNA manipulations were carried out as described by Sambrook *et al.* (23). PCRs were performed in a Perkin–Elmer thermal cycler model 480 using 50 ng of mycobacterial chromosomal DNA and 1 µg of each primer (Table 1). DNA sequences were determined by using the dideoxy chain-termination method (24). Mycobacterial chromosomal DNA was prepared as described by Baulard *et al.* (25).

Abbreviations: HBHA, heparin-binding hemagglutinin; BCG, bacillus Calmette–Guérin; IPTG, isopropyl β-D-thiogalactoside; AEBSEF, 4-(2-aminoethyl)-benzenesulfonyl fluoride.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. AF07390).

‡Present address: Astra-Draco, AB, Department of Bioanalytical Chemistry, P.O. Box 34, S-221 00 Lund, Sweden.

¶To whom reprint requests should be addressed: e-mail: Camille.Locht@pasteur-lille.fr.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1998 by The National Academy of Sciences 0027-8424/98/9512625-6\$2.00/0 PNAS is available online at www.pnas.org.

Table 1. Oligonucleotide sequences

S1441: (5')AAG GC(G/C)GAG GG(G/C) TAC CT (3')
S1441 reverse:(5') AGG TA(G/C) CCC TC(G/C) GCC TT (3')
S1443: (5') GAC CAG GC(G/C) GT(G/C) GAG CT (3')
S1443 reverse: (5') AGC TC(G/C) AC(G/C) GCC TGG TC (3')
HBHA Seq1: (5') AGC CGG TAC AAC GAG CTG GTC (3')
HBHA Seq1 reverse: (5') GAC CAG CTC GTT GTA CCG GCT (3')
HBHA Seq2: (5') CAT CCA ACA CGT CGA CTC C (3')
HBHA Seq3: (5') TTG ATG TCA TCA ATG TTC G (3')
HBHA Seq4: (5') CGT GGA CCA GGC GGT GGA G (3')
HBHA Seq5: (5') GAC GAT CAG GAG GTT TCC CCG (3')
HBHA Seq6: (5') TGC CCC AAC GTC CAG ACC AAA GAT (3')
HBHA Seq7: (5') CAA GAC GGC GAC CAG CAA TAC CAG (3')
Universal reverse: (5') AGC GGA TAA CAA TTT CAC ACA GGA (3')

Production of Recombinant HBHA in *E. coli*. *E. coli* XL1-Blue transformed with pKK-HBHA, a pKK388-1 derivative containing the HBHA-encoding gene under the control of the *trc* promoter, was grown at 37°C in 500 ml of Luria-Bertani broth supplemented with 150 µg/ml of ampicillin. At an OD₆₀₀ of 0.5, isopropyl β-D-thiogalactoside (IPTG) was added at a final concentration of 1 mM, and incubation was continued for 5 hr. The cells then were harvested by centrifugation at 7,000 × *g* for 15 min at 4°C and stored at -20°C until further use.

Purification of Recombinant HBHA by Heparin-Sepharose Chromatography. Frozen IPTG-induced *E. coli* XL1-Blue(pKK-HBHA) cells from a 500-ml culture were thawed and then resuspended in 15 ml of PBS (pH 7.4) containing 1 mM 4-(2-aminoethyl)-benzenesulfonyl fluoride (AEBSF) (Pefabloc SC, Boehringer Mannheim). The bacteria were sonicated for 4 min at 4°C using a Branson Sonifier at an output of 5 delivered to a microtip, and the lysate was centrifuged at 10,000 × *g* for 15 min at 4°C. The supernatant was diluted to 200 ml with PBS containing 1 mM AEBSF and then applied onto a heparin CL-6B (Pharmacia) column (1 × 4 cm), previously equilibrated with 100 ml of PBS. The bound material was eluted with a 0–500 mM NaCl linear gradient in PBS containing 1 mM AEBSF (11). All chromatographic steps were carried out at 4°C using a flow rate of 1.5 ml/min.

SDS/PAGE Analysis. SDS/PAGE was performed as described by Laemmli (26) by using a 4% stacking gel and a 12.5% or 15% separating gel, as indicated. Before electrophoresis, samples were mixed with one-third (vol/vol) of solubilization buffer (6% SDS/15% β-mercaptoethanol/30% glycerol/0.005% bromophenol blue in 0.18 M Tris·HCl, pH 6.8) and heated at 95°C for 5 min. After electrophoresis, the gels were stained with Coomassie brilliant blue R-250 (ICN).

Protein Sequencing. The *M. bovis* BCG HBHA was purified by heparin-Sepharose chromatography as described (11). Twenty-five micrograms of the purified protein were subjected to SDS/PAGE on a 15% polyacrylamide gel. After electrophoresis, the protein was digested by trypsin within the gel, and the resulting peptides were separated by using reverse-phase

HPLC. Four major peptides were isolated and subjected to automated Edman degradation.

Determination of the Carbohydrate Content of HBHA. Determination of the carbohydrate content of HBHA was performed at the Carbohydrate Research Center at the University of Georgia. Purified *M. bovis* BCG HBHA was subjected first to preliminary aqueous hydrolysis (2 M trifluoroacetic acid for 2 hr at 105°C), and the liberated monosaccharides were converted to methyl glycosides with anhydrous methanolic HCl (1 M, 80°C for 16 hr), followed by *N*-acetylation by treatment with methanol, pyridine, and acetic anhydride for 6 hr at room temperature. The methyl glycosides then were treated with trimethylsilane (TMS) reagent (Tri-Sil, Pierce) to form TMS derivatives, as described (27, 28). GC/mass spectrophotometry analysis was performed on a 30-m DBI fused silica capillary column using a Hewlett Packard 5890 GC coupled to a 5970 mass spectrophotometer detector.

Other Techniques. Immunoblot analyses were performed by electrotransfer onto a poly(vinylidene difluoride) membrane (ProBlott; Applied Biosystems) using the method of Towbin *et al.* (29). Antigen-antibody complexes were detected by using the ProtoBlot Alkaline Phosphatase System (Promega). Mouse polyclonal anti-HBHA antibodies were produced as described for the production of rat anti-HBHA antibodies (11), and mAbs 4057D2 and 3921E4 were a kind gift from D. Rouse and S. Morris [Center for Biologics Evaluation and Research (CBER), Bethesda, MD] (30). Protein concentrations were determined by the method of Bradford (31), using BSA (Sigma) as a standard. Hemagglutination was carried out as described by Menozzi *et al.* (11).

RESULTS

Cloning and Sequencing of the *M. bovis* BCG HBHA-encoding Gene. A PCR approach was chosen to clone the HBHA-encoding gene of *M. bovis* BCG, based on peptide sequences of tryptic fragments of the purified protein. After digestion of purified HBHA with trypsin, four peptides were isolated by HPLC, and their amino-terminal sequences were determined (Table 2). The sequences of two peptides, S1441 and S1443, showed the least redundancy of the genetic code, and two pairs of oligonucleotides were designed and named S1441/S1441 reverse and S1443/S1443 reverse, respectively (see Table 1 for the oligonucleotide sequences). Because mycobacterial DNA has a high GC content (32), guanosine and cytosine were favored at the third codon position during the oligonucleotide synthesis. Two PCRs then were carried out by using 50 ng of chromosomal BCG DNA and oligonucleotides S1443 and S1441 reverse (1 µg each) or oligonucleotides S1443 reverse and S1441 (1 µg each). At an annealing temperature of 50°C and after 30 amplification cycles, only the PCR primed with oligonucleotides S1441 and S1443 reverse resulted in the amplification of a distinct DNA fragment, which had an approximate size of 150 bp (not shown). This fragment was amplified when the annealing temperature was increased to 57°C and therefore considered a specific amplification product. The PCR fragment was blunt-ended and inserted into pUC18 previously digested by *Hind*II. After transformation of

Table 2. Amino acid sequences of HBHA peptides

Peptide	Sequence	Reference
Amino terminus	N-Ala-Glu-Asn-Ser-Asn-Ile-Asp-Asp-Ile-Lys-Ala-Pro-Leu-Leu-Ala-Ala-C	11
Peptide S1441	N-Lys*-Ala-Glu-Gly-Tyr-Leu-Glu-Ala-Ala-Thr-C	This study
Peptide S1443	N-Xxx†-Glu-Gly-Tyr-Val-Asp-Gln-Ala-Val-Glu-Leu-Thr-Gln-Glu-Ala-Leu-Gly-C	This study
Peptide S1446	N-Xxx-Gln-Glu-Xxx-Leu-Pro-Glu-Xxx-Leu-C	This study
Peptide S1447	N-Phe-Thr-Ala-Glu-Glu-Leu-Arg-C	This study

*Lysine (4.6 pmol) and 11 pmol of alanine were found in the first sequencing cycle, probably because the tryptic peptide was partially cleaved at the Arg-Lys bond, giving rise to a mixture of two peptides, one starting with a lysine, one with an alanine.

†Xxx represents a position where no amino acid could be definitively assigned.

E. coli XL1-Blue, the recombinant plasmid, named pClone5, was purified, and the insert was sequenced. The DNA sequence corresponded to the amino acid sequences determined for peptides S1441 and S1443 (not shown), indicating that the PCR fragment contains part of the gene coding for HBHA. To clone the entire HBHA structural gene, the PCR product was isolated from pClone5 by digestion with *Bam*HI and *Hind*III, purified by electroelution from a polyacrylamide gel, labeled with digoxigenin, and used as a probe for Southern blot analyses of chromosomal BCG DNA cut with various restriction enzymes. Except for *Sma*I, each restriction enzyme yielded a single hybridizing band (Fig. 1), suggesting that the HBHA structural gene is present in a single copy in the BCG chromosome. As expected, two bands were obtained after digestion with *Sma*I, because the pClone5 insert contains a *Sma*I site. *Sph*I fragments from 2.3 to 2.7 kb, which were long enough to potentially contain the entire HBHA gene, were purified by electroelution and inserted into pUC18 cut with *Sph*I. After transformation of *E. coli* XL1-Blue, the recombinant colonies were analyzed by colony blot hybridization using the digoxigenin-labeled 150 bp probe of pClone5. Among the 300 colonies analyzed, one hybridized strongly with the probe. Restriction analysis of this clone revealed that its plasmid, named pHBHA, contained a 2.5-kb *Sph*I insert. Location of the HBHA gene within this fragment was confirmed by restriction and Southern blot analyses, as well as by DNA sequencing.

Analysis of the HBHA Structural Gene. The nucleotide sequence of the right side of the 2.5-kb *Sph*I insert of pHBHA is shown in Fig. 2. An ORF was found between nucleotides 331 and 927 with a coding potential of 199 amino acids. The deduced amino acid sequence of the first 16 codons following the first in-frame ATG codon was found to be identical to the determined amino-terminal sequence of purified HBHA (11). In addition, all internal peptide sequences of HBHA shown in Table 2 were found in the deduced amino acid sequence shown

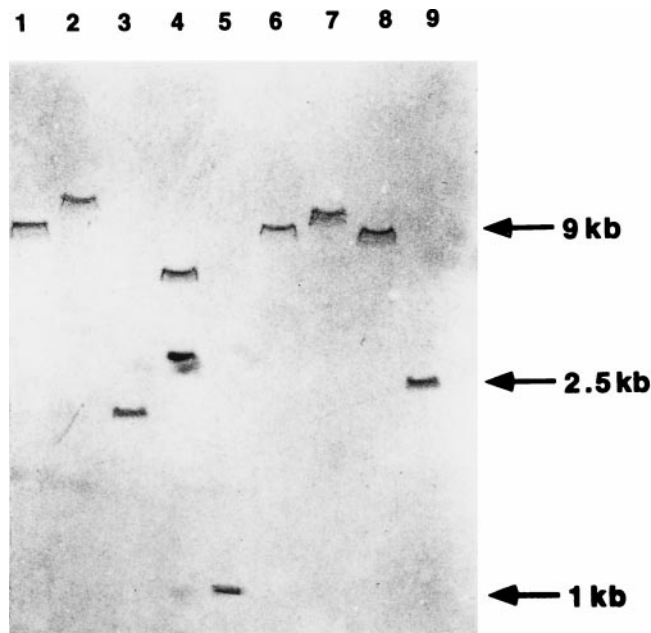


FIG. 1. Southern blot analysis of the HBHA-coding gene. Total DNA was extracted from *M. bovis* BCG, digested with *Bam*HI (lane 1), *Eco*RI (lane 2), *Pst*I (lane 3), *Sma*I (lane 4), *Acc*I (lane 5), *Nco*I (lane 6), *Not*I (lane 7), *Sac*I (lane 8), or *Sph*I (lane 9) and subjected to agarose-gel electrophoresis and transfer onto a nylon membrane. The membrane then was probed with the approximately 150-bp *Bam*HI-*Hind*III fragment of pClone5. The sizes of the markers are shown on the right.

```

1   cgg ctg gcg ggt aat caa acc tga agg aca gtc atc tgg gtg agg tgc acc gca ggc tga
61  tcc agc cga tcy gcc ggc gct ggc caa cag cga ctc cgt cga tga cgt gca gca aag gag
121 aca tgt agt gac cgg atc agc tgg gcc tga cat cta cga act cga ccg aca acc gac ccg
181 acg atc agg agg ttt ccc cgg caa gtc gcg tgc cat gtc aat ccg cgg gtc ttg act agt
241 cct ccc tgg agg agc cga cgc ttg ccc caa cgt cca gac caa aga tgt aag aac gcc gat
301 atc aga aaa tag tta atg aaa gga ata ccc atg gct gaa aac tgc aac att gat gac atc
361 aag gct ccg ttg ctt gcc gcg ctt gga gcg gcc gac ctg gcc ttg gcc act gtc aac gag
421 tgg atc agc aac ctg cgt gag cgt gcg gag gag act cgt acg gac acc cgc agc cgg gtc
481 gag gag agc cgt gct cgc ctg acc aag ctg cag gaa gat ctg ccc gag cag ctc acc gag
541 ctg cgt gag aag ttc acc gcc gag gag ctg cgt aag gcc gcc gag gcc tac ctc gag gcc
601 gcg act agc cgg tac aac gag ctg gtc gag cgc ggt gag gcc gct cta gag cgg ctg cgc
661 agc cag cag agc ttc gag gaa gtg tcy gcg cgc gcc gaa gcc tac gtg gac cag gcg gtg
721 gag ttg acc cag gag gcg tgg ggt acg gtc gca tcy cag acc cgc gcg gtc ggt gag cgt
781 gcc gcc aag ctg gtc gcc atc gag ctg cct aag aag gct gct ccg gcc aag aag gcc gct
841 ccg gcc aag aag gcc gct ccg gcc aag aag gcg gcg gcc aag aag ggc ccc gcc aag aag
901 gcg gcg gcc aag aag gtc acc cag aag tag tcy ggc tcc gaa tca cca tcy act ccg agt
961 cgc cca cgg ggc gac tcy gag tcy acg tgt tgg atg caa acc gca tag tct gaa tcy gtg
1021 agc cac ctc gty ggt acc gtc atg ctg gta ttg ctg gtc gcc gtc ttg gty aca gcg gty
1081 tac gcg ttt gty cat gc
      SphI
    
```

FIG. 2. Sequence of the HBHA-coding gene. The nucleotide sequence of the HBHA-coding gene is indicated by lowercase letters. The bold letters show the putative ribosomal-binding site, and the letters in italics indicate the *Sph*I site. The numbers refer to the numbers of the nucleotides. The capital letters represent the amino acid sequence in the single letter code. * denotes the stop of the amino acid sequence. The amino acids that were determined by peptide sequencing are underlined. The boxed sequences correspond to the Lys/Pro-rich repeats.

in Fig. 2, allowing us to conclude that this ORF contains the HBHA-coding sequence.

Codon 331-333 of the sequence is most likely the initiation codon, because no other in-frame ATG or GTG was found further upstream following the first in-frame stop codon at position 310-312. Moreover, a putative ribosome binding site (AGGAA) was found 6-10 nt upstream of this ATG. Because the first amino acid of purified HBHA is the first residue following the initiation methionine, the protein does not appear to contain a signal sequence, although it is found surface-associated and secreted in mycobacterial cultures (11). This observation suggests that HBHA is secreted in a *sec*-independent manner.

The predicted mature protein contains no methionine, no histidine, no cysteine, and no tryptophan. Interestingly, HBHA contains lysine-rich repeats in its carboxyl-terminal region. Two different strictly repeated motifs were identified (boxed in Fig. 2). The KKAAPA motif is directly repeated three times between residues 160 and 177, whereas the KKAAAKK motif is repeated two times between amino acids 178 and 194. An APA sequence separates the two repeats. The calculated molecular weight of HBHA is 21,331, substantially lower than its apparent molecular weight of approximately

28,000 as determined by SDS/PAGE. This discrepancy could be the result of an aberrant electrophoretic mobility of HBHA because of the K-A-P-rich carboxyl-terminal domain and/or to a posttranslational modification of the protein.

Cloning and Sequence of the *M. tuberculosis* HBHA Structural Gene. Because HBHA also is produced by *M. tuberculosis* (11), we used two synthetic oligonucleotides (oligonucleotides HBHA Seq6 and HBHA Seq7, Table 1) corresponding to the upstream and downstream regions of the BCG HBHA structural gene to amplify, by PCR, a 0.8-kb DNA fragment of the *M. tuberculosis* H37Rv chromosome. The amplified DNA fragment was ligated into pUC18 previously linearized by *Hind*II restriction. Sequence analysis of the cloned PCR fragment demonstrated that the *M. tuberculosis* HBHA gene sequence is the same as the sequence of the BCG HBHA gene.

Production of HBHA in *E. coli*. The expression vector pKK388-1 was used to produce HBHA in *E. coli*. The 705-bp *Nco*I-*Kpn*I fragment of pHBHA was inserted into pKK388-1 that had been previously restricted by *Nco*I and *Kpn*I. The recombinant plasmid, named pKK-HBHA, then was introduced into *E. coli* XL1-Blue. The synthesis of recombinant HBHA was induced by the addition of IPTG and detected by SDS/PAGE and immunoblot analysis of *E. coli* whole-cell lysates (Fig. 3). *E. coli* XL1-Blue (pKK-HBHA) produced a polypeptide of approximately 27 kDa (Fig. 3, Left, lanes B and D), which was not found in extracts of *E. coli* XL1-Blue(pKK388-1). Although the apparent molecular weight of this recombinant protein is still higher than the calculated molecular weight, it is lower than that of HBHA purified from BCG or *M. tuberculosis* H37Ra (compare lanes E to lanes D in Fig. 3). This observation suggests that the native mycobacterial HBHA is posttranslationally modified.

To confirm that the 27-kDa protein in *E. coli* XL1-Blue(pKK-HBHA) extracts corresponds to the recombinant HBHA, immunoblot analyses were carried out by using murine polyclonal anti-HBHA antibodies, as well as mABs 3921E4 and 4057D2 reactive with HBHA (30). The polyclonal antibodies and mAB 3921E4 recognized purified BCG HBHA (Fig. 3, lanes E) and the recombinant polypeptide produced by *E. coli* XL1-Blue(pKK-HBHA) (Fig. 3, lanes D). No immunoreactivity was detected in lysates of IPTG-induced *E. coli* XL1-Blue(pKK388-1), confirming that the 27-kDa protein corresponds to the recombinant HBHA. Surprisingly, mAB 4057D2 did not recognize the recombinant form of HBHA,

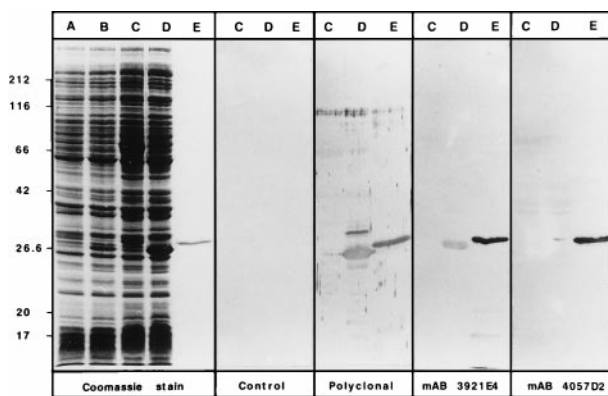


FIG. 3. Production of recombinant HBHA in *E. coli*. *E. coli* XL1-Blue containing pKK388-1 (lanes A and C) or containing pKK-HBHA (lanes B and D) were incubated in the presence of IPTG (lanes C and D). Cell extracts were prepared and analyzed by SDS/PAGE and Coomassie blue staining (Left) or by immunoblot analysis using only secondary conjugated antibodies (second panel), anti-HBHA polyclonal antibodies (third panel), or mABs 3921E4 and 4057D2 (fourth and fifth, panels, respectively). Lanes E contain authentic HBHA purified from BCG culture supernatants. The sizes of the M_r markers given in kDa are shown on the left.

indicating that its epitope is not present in the recombinant protein (Fig. 3, Right). Under the same electrophoresis and immunoblotting conditions mAB 4057D2 did recognize HBHA purified from mycobacteria, which suggests that its epitope may depend on the posttranslational modification of native HBHA.

Evidence for Glycosylation of HBHA. HBHA was purified from BCG culture supernatants and analyzed for potential monosaccharide composition by derivatization using trimethylsilane followed by combined GC/mass spectrophotometry analysis. The results of the glycosyl composition analysis indicate that glucose is the major component (41.6%), with lesser amounts of xylose (17%), mannose (10.2%), arabinose (8.8%), and galactose (6.2%). An unidentified component with a GC retention time of 29.02 min had a mass spectrum that was similar, but not identical with 2-keto-3-deoxy-D-manno-octulosonic acid.

Based on an estimate of the protein concentration of the purified HBHA sample, the total amount of carbohydrate represents 2.8% of the mass of HBHA, which is consistent with the observed difference in electrophoretic mobility between the authentic HBHA and the recombinant protein.

Susceptibility of Recombinant HBHA to Proteolytic Degradation. The IPTG-induced production of recombinant HBHA did not lead to the formation of inclusion bodies in *E. coli*, because most of the recombinant protein was recovered in the soluble fraction after sonication and subsequent clarification of the lysate by centrifugation (data not shown). However, soluble recombinant HBHA was rapidly degraded. During incubation of the lysate at room temperature, recombinant HBHA rapidly degraded into a 19-kDa polypeptide that was resistant to further degradation. This polypeptide was no longer recognized by mAB 3921E4, indicating that the epitope for mAB 3921E4 was lost after this cleavage (Fig. 4). After 2 hr of incubation at room temperature, the 27-kDa protein was completely cleaved into the 19-kDa form and was no longer reactive with mAB 3921E4. Because the carboxyl-terminal region of HBHA is rich in basic amino acids containing potential cleavage sites for serine endopeptidases, AEBSF, a strong, nontoxic inhibitor of serine proteases, was added to the *E. coli* lysate at a final concentration of 1 mM. Under these conditions, recombinant HBHA remained stable even after 18 hr of incubation at room temperature (Fig. 4, Right), suggesting that the recombinant HBHA was cleaved by the action of a serine protease. When the purified native HBHA was mixed with an equal volume of *E. coli* lysate, it did not undergo the

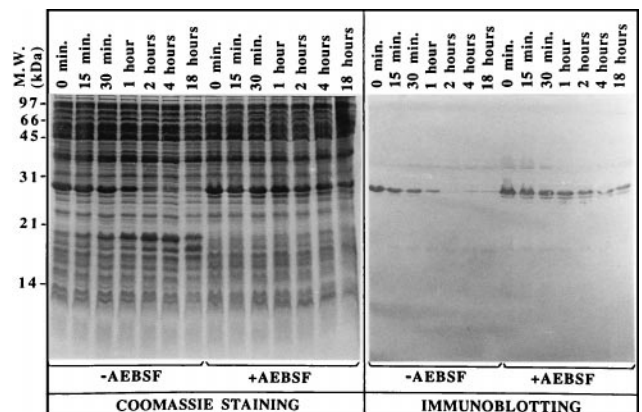


FIG. 4. Instability of recombinant HBHA. After induction with IPTG, *E. coli* (pKK-HBHA) cell extracts were incubated at room temperature either in the presence of AEBSF or in its absence. At the indicated time points the extracts then were analyzed by SDS/PAGE and Coomassie blue staining (Left) or by immunoblotting using mAB 3921E4 (Right). The sizes of the M_r markers given in kDa are shown on the left.

cleavage observed with recombinant HBHA (data not shown), suggesting that the glycosylation of native HBHA protects the protein against proteolytic cleavage.

Activities of Recombinant HBHA. To investigate whether recombinant HBHA is able to bind to sulfated carbohydrates, a lysate of IPTG-induced *E. coli* XL1-Blue(pKK-HBHA) was chromatographed on heparin-Sepharose at 4°C in the presence of 1 mM AEBSF. Recombinant HBHA eluted at approximately 350 mM NaCl (data not shown). This chromatographic behavior was identical to that of native HBHA (11), indicating that both HBHA forms interact similarly with sulfated polysaccharides.

When AEBSF was omitted during chromatography, HBHA was recovered as a 19-kDa protein, which no longer bound to heparin. This finding suggests that the cleaved region of HBHA is involved in binding of HBHA to heparin. To investigate which region of HBHA is involved in heparin-binding, heparin-Sepharose purified recombinant HBHA was incubated at 37°C for 2 hr. Three distinct peptides were detected upon SDS/PAGE and Coomassie blue staining, the 27-kDa protein, as well as two peptides with apparent molecular weights of approximately 26,000 and 25,000. Also, after 6 hr of incubation, the 19-kDa form became apparent (not shown). Partial amino acid sequencing of these peptides revealed that they all contained the amino-terminal sequence of HBHA, indicating that the degradation occurred at the carboxy-terminus of the protein. Moreover, immunoblot analysis showed that the 25-kDa and 19-kDa HBHA species were no longer detected by mAB 3921E4, indicating that its epitope is located in the carboxyl end of the protein. When these recombinant proteins were applied onto a heparin-Sepharose column, the 19-kDa form did not bind, whereas all other fragments of HBHA bound to heparin. After gradient elution, the smaller forms (25 and 26 kDa) eluted before the 27-kDa form (Fig. 5), indicating that heparin binding also requires the carboxyl-terminal domain of HBHA, and that progressive truncations from the C-terminal end diminish the affinity for heparin.

As shown previously (11), purified native HBHA agglutinates rabbit erythrocytes, whereas heparin-Sepharose-purified, full-length recombinant HBHA failed to induce

hemagglutination, suggesting that this activity depends on the glycosylation of HBHA.

DISCUSSION

In this study we describe the molecular characterization of a mycobacterial adhesin. This protein recently was identified as a surface-exposed HBHA involved in the adherence of *M. tuberculosis* to epithelial cells (11). The M_r deduced from the amino acid sequence was found to be smaller than the apparent M_r estimated by SDS/PAGE. This difference could be attributed to the Pro/Lys-rich C-terminal region of HBHA. However, in addition, SDS/PAGE analyses indicated that the apparent M_r of recombinant HBHA was approximately 1 kDa lower than that of native HBHA. This difference may be related to a posttranslational modification of the protein in mycobacteria, because GC/mass spectrophotometry analysis revealed that native HBHA contains *ca.* 2.8% carbohydrate. Combined GC/mass spectrophotometry analysis indicated that the carbohydrate moiety of HBHA was composed of glucose, some xylose, mannose, arabinose, galactose, and an unidentified component. It remains to be proven that an oligosaccharide is covalently linked to the protein core of HBHA. Interestingly, among the two anti-HBHA mABs one (4057D2) recognized native HBHA and not the recombinant form, suggesting that the glycosylation contributes to the epitope of this mAB. This finding may aid in the identification of functional sites on HBHA because the mABs have been shown to inhibit hemagglutination and mycobacterial attachment to eukaryotic cells (11).

Glycosylation is common among eukaryotic membrane and secretory proteins, but much less so in bacteria. Although glycoproteins have been known for some time to occur in archaeobacteria, their existence in eubacteria has only recently been established. Some of the eubacterial glycoproteins constitute virulence factors, especially adhesins, such as the fiber-forming pilin of *Neisseria gonorrhoeae* (33). In contrast to most bacteria, mycobacteria produce a number of different glycoproteins, and our evidence suggests that HBHA is an additional member of this group. In *M. tuberculosis*, several antigens have been identified as glycoproteins that range in size from 19 kDa (34) to 55 kDa (35). Some of these antigens also have been detected in *M. bovis* (36). So far, two of the glycoproteins also are known to contain lipid moieties at their amino terminus (34, 35). Although detailed chemical analyses have been reported for some of these glycoproteins (37), their function remains unknown, with the exception of the 38-kDa lipoprotein, which is likely involved in phosphate uptake (38). A recent study indicated that the glycosylation of the *M. tuberculosis* 19-kDa lipoprotein alters proteolytic cleavage of the molecule by conferring resistance to certain proteases (39).

Recombinant HBHA was found to be sensitive to proteolytic cleavage, whereas HBHA purified from BCG was not easily degraded, even when incubated in the presence of a crude *E. coli* extract. The proteolytic degradation of recombinant HBHA occurred primarily in the C-terminal region. Similar to the recent study on the role of glycosylation of the 19-kDa mycobacterial lipoprotein (39), glycosylation also appears to protect HBHA from proteolytic degradation. This finding suggests that the carbohydrate moiety of HBHA may be located within the C domain of the protein. Alternatively, glycosylation could occur elsewhere in the protein and induce a folding that confers resistance to proteolytic degradation. In eukaryotes, for example, there is evidence that glycosylation plays a role in protein folding (40).

In addition to its role in antigenicity and in protection against proteolysis, the glycosylation also appeared to be important for HBHA-mediated hemagglutination. Although the physiological significance of hemagglutination is not known, many bacterial adhesins that mediate attachment of

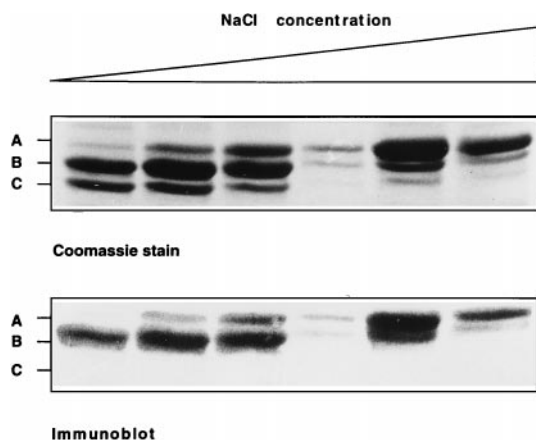


FIG. 5. Elution of recombinant HBHA proteins from heparin-Sepharose. Recombinant HBHA was applied onto a heparin-Sepharose column, and the bound material was eluted by a linear 0–500 mM NaCl gradient. The eluted fractions were analyzed by SDS/PAGE and Coomassie blue staining (*Upper*) and by immunoblotting using mAB 3921E4 (*Lower*). The three forms of recombinant HBHA (approximately 27 kDa, 26 kDa, and 25 kDa) are indicated by A, B, and C, respectively. The fractions analyzed go from low-salt (*Left*) to high-salt elution (*Right*). Note that the larger M_r species elute at higher salt concentrations and that the 25-kDa form is not recognized by mAB 3921E4.

the microorganisms to host cells express hemagglutination activity. It may be that the glycosylation of HBHA plays a role in promoting adherence of the mycobacteria, similar to what has been proposed for the role of glycosylation of the *N. meningitidis* pilin (41). However, a definitive answer to this question awaits careful comparisons of the binding activities of wild-type mycobacteria with those of isogenic mutant strains producing HBHA that is not glycosylated.

Because recombinant HBHA binds to heparin in a similar fashion as native HBHA, the loss of glycosylation appears not to influence the binding to sulfated carbohydrates. Amino acid sequence analysis revealed that the C-terminal region of HBHA contains Pro/Lys-rich repeats. One of them (KKAAAPA) is directly repeated three times, and a second one (KKAACK) is repeated two times separated by an APA sequence. Progressive removal of these repeats in the recombinant protein decreased the binding to heparin-Sepharose. The removal of all five repeats totally abolished heparin binding, indicating that they play an essential role in the interaction with sulfated carbohydrates. Lysines are known to play a key role in the binding activities of proteins that specifically interact with sulfated sugars (42). Because lysines are also prone to proteolytic attack by serine proteases, the protection against proteolysis by glycosylation of HBHA may be crucial to maintaining the biological activities of HBHA as an adhesin.

Although amino acid sequence analysis revealed the absence of a signal peptide, cell-fractionation experiments and immuno-electron microscopy analyses clearly established that HBHA is a surface-associated as well as a secreted protein in *M. tuberculosis* and *M. bovis* (11). Instead of a signal peptide sequence, the first amino acid of the mature form of the protein is immediately preceded by the initiation methionine. This observation indicates that HBHA is secreted in a signal-peptide independent manner. Because glycosylation has been shown to play a role in protein export in other systems (43), it is conceivable that the glycosylation of HBHA also is involved in the secretion of this protein. Production of a nonglycosylated mutant HBHA in mycobacteria may clarify this hypothesis. It is also possible that the secretion of HBHA requires the coordinated production of accessory proteins. In many cases, the genes encoding the accessory proteins are located in the vicinity of the structural gene (44). The cloning and sequence analysis of the sequences upstream and downstream of the HBHA structural gene hopefully will provide some evidence for the existence of possible genes.

We thank the Carbohydrate Research Center at the University of Georgia for the determination of the carbohydrate content of HBHA, and D. Rouse and S. Morris for the 4057D2 and 3921E4 hybridomas. The work was supported, in part, by Institut Pasteur de Lille, Institut National de la Santé et de la Recherche Médicale, Région Nord-Pas de Calais, Caisse Nationale d'Assurances Maladies des Travailleurs Salariés, and Ministère de l'Éducation Nationale, de l'Enseignement Supérieur, de la Recherche et de l'Insertion Professionnelle.

1. Raviglione, M. C., Snider, D. E. & Kochi, A. (1995) *J. Am. Med. Assoc.* **273**, 220–226.
2. Horsburgh, C. R. (1991) *N. Engl. J. Med.* **324**, 1332–1338.
3. Shepard, C. C. (1957) *J. Exp. Med.* **105**, 39–48.
4. Mapothar, M. E. & Sanger, J. G. (1984) *Infect. Immun.* **45**, 67–73.
5. McDonough, K. A. & Kress, Y. (1995) *Infect. Immun.* **63**, 4802–4811.
6. Bermudez, L. E. & Goodman, J. (1996) *Infect. Immun.* **64**, 1400–1406.

7. Ratliff, T. L., McCarthy, R., Telle, W. B. & Brown, E. J. (1993) *Infect. Immun.* **61**, 1889–1894.
8. Schlesinger, L. S., Bellinger-Kawahara, C. G., Payne, N. R. & Horwitz, M. A. (1990) *J. Immunol.* **144**, 2771–2780.
9. Schlesinger, L. S. & Horwitz, M. A. (1990) *J. Clin. Invest.* **85**, 1304–1314.
10. Hirsch, C. S., Ellner, J. J., Russell, D. G. & Rich, E. A. (1994) *J. Immunol.* **152**, 743–753.
11. Menozzi, F. D., Rouse, J. H., Alavi, M., Laude-Sharp, M., Muller, J., Bischoff, R., Brennan, M. J. & Loch, C. (1996) *J. Exp. Med.* **184**, 993–1001.
12. Menozzi, F. D., Mutombo, R., Renauld, G., Gantiez, C., Hannah, J. H., Leininger, E., Brennan, M. J. & Loch, C. (1994) *Infect. Immun.* **62**, 769–778.
13. Hannah, J. H., Menozzi, F. D., Renauld, G., Loch, C. & Brennan, M. J. (1994) *Infect. Immun.* **62**, 5010–5019.
14. Krivan, H. C., Olson, L. D., Barile, M. F., Ginsburg, V. & Roberts, D. (1989) *J. Biol. Chem.* **264**, 9283–9288.
15. Isaacs, R. D. (1994) *J. Clin. Invest.* **93**, 809–819.
16. Van Putten, J. P. M. & Paul, S. M. (1995) *EMBO J.* **14**, 2144–2154.
17. Nahmias, A. J. & Kibrick, S. (1964) *J. Bacteriol.* **87**, 1060–1066.
18. Compton, T., Nowlin, D. M. & Cooper, N. R. (1993) *Virology* **193**, 834–841.
19. Shieh, M.-T. & Spear, P. G. (1994) *J. Virol.* **68**, 1224–1228.
20. Ortega-Barria, E. & Pereira, E. A. (1991) *Cell* **67**, 411–421.
21. Frevert, U., Sinnis, P., Cerami, C., Shreffler, W., Takacs, B. & Nussenzweig, V. (1993) *J. Exp. Med.* **177**, 1287–1298.
22. Love, D. C., Esko, J. D. & Mosser, D. M. (1993) *J. Cell Biol.* **123**, 759–766.
23. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
24. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
25. Baulard, A., Kremer, L. & Loch, C. (1996) *J. Bacteriol.* **178**, 3091–3098.
26. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
27. York, W. S., Darvill, A. G., McNeil, M., Stevenson, T. T. & Albersheim, P. (1985) *Methods Enzymol.* **118**, 3–40.
28. Merkle, R. K. & Poppe, I. (1994) *Methods Enzymol.* **230**, 1–15.
29. Towbin, H., Staehlin, T. & Gordon, J. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4350–4354.
30. Rouse, D. A., Morris, S. L., Karpas, A. B., Mackall, J. C., Probst, P. G. & Chaparas, S. D. (1991) *Infect. Immun.* **59**, 2595–2600.
31. Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248–254.
32. Wayne, L. G. & Kubica, G. P. (1986) in *Bergey's Manual of Systematic Bacteriology*, eds. Sneath, P. H. A., Mair, N. S., Sharp, M. E. & Holt, J. G. (Williams & Wilkins, Baltimore), pp. 1436–1457.
33. Parge, H. E., Forest, K. T., Hickey, M. J., Christensen, D. A., Getzoff, E. D. & Tainer, J. A. (1995) *Nature (London)* **378**, 32–38.
34. Garbe, T., Harris, D., Vordermeier, M., Lathigra, R., Ivanyi, J. & Young, D. (1993) *Infect. Immun.* **61**, 260–267.
35. Espitia, C. & Mancilla, R. (1989) *Clin. Exp. Immunol.* **77**, 378–383.
36. Fifis, T., Costopoulos, C., Radford, A. J., Bacic, A. & Wood, P. R. (1991) *Infect. Immun.* **59**, 800–807.
37. Dobos, K. M., Swiderek, K., Khoo, K.-H., Brennan, P. J. & Belisle, J. T. (1995) *Infect. Immun.* **63**, 2846–2853.
38. Andersen, A. B., Ljungqvist, L. & Olsen, M. (1990) *J. Gen. Microbiol.* **136**, 477–480.
39. Herrmann, J. L., O'Gaora, P., Gallagher, A., Thola, J. E. R. & Young, D. B. (1996) *EMBO J.* **15**, 3547–3554.
40. Ruddon, R. W. & Bedows, E. (1997) *J. Biol. Chem.* **272**, 3125–3128.
41. Virji, M., Saunders, J. R., Sims, G., Makepeace, K., Maskell, D. & Ferguson, D. J. P. (1993) *Mol. Microbiol.* **10**, 1013–1028.
42. Cardin, A. D. & Weintraub, H. J. R. (1989) *Arteriosclerosis* **9**, 21–32.
43. Kubicek, C. P. (1982) *Adv. Biochem. Eng. Biotechnol.* **45**, 1–27.
44. Pugsley, A. P. (1993) *Microbiol. Rev.* **57**, 50–108.