INFECTION AND IMMUNITY, Aug. 1996, p. 3038–3047
0019-9567/96/$04.00+0
Copyright © 1996, American Society for Microbiology
Vol. 64, No. 8

Novel Insights into the Genetics, Biochemistry, and Immunocytochemistry of the 30-Kilodalton Major Extracellular Protein of *Mycobacterium tuberculosis*

GÜNTER HARTH, BAI-YU LEE, JUN WANG, DANIEL L. CLEMENS, AND MARCUS A. HORWITZ*

Division of Infectious Diseases, Department of Medicine, University of California Los Angeles School of Medicine, Los Angeles, California 90095

Received 25 March 1996/Accepted 2 May 1996

The 30/32-kDa complex of major secretory proteins are among the most important and intensively studied proteins of *Mycobacterium tuberculosis*. The proteins have been demonstrated to be immunoprotective and to play a central role in the physiology of the mycobacterium. In this study, we present a series of novel insights into this key protein complex arising out of a combination of genetic, biochemical, and immunocytochemical analyses. Our genetic analyses (i) indicate that the genes are arranged as separate transcription units, (ii) demonstrate that the mature 30-kDa protein of *M. tuberculosis* differs from the corresponding 30-kDa proteins of two strains of *Mycobacterium bovis* BCG by only 1 and 5 amino acids, (iii) suggest that expression of the proteins is regulated at the transcriptional level, and (iv) map the transcriptional start site of the 30-kDa protein gene. Our biochemical analyses provide evidence that (i) the 30-kDa protein and the two 32-kDa proteins (i.e., 32A and 32B) are secreted at a ratio of ~3:2:1, respectively, (ii) the proteins exist as monomers, (iii) the proteins are not posttranslationally modified by the addition of carbohydrates and lipids, (iv) the 30-kDa and 32A proteins contain one disulfide bridge, and (v) high-level expression and leader peptide processing are achievable in *Escherichia coli*. Our immunocytochemical analyses demonstrate that the 30/32-kDa complex is expressed in human monocytes and that the proteins are localized to the phagosomal space and the mycobacterial cell wall. These analyses fill important gaps in our knowledge of this critical protein complex of *M. tuberculosis* and, at the same time, raise new and fundamental questions regarding regulatory mechanisms that control coordinate expression of the proteins at a fixed ratio.

Tuberculosis is a worldwide health problem and may be the most common cause of death from any single infectious agent. The annual number of deaths caused by this disease is anticipated to increase from its current level of ~3 million to ~3.5 million by 2005. The global emergence of multidrug-resistant strains of *Mycobacterium tuberculosis*, the primary causative agent of tuberculosis, has given new urgency to the need to develop better means to combat this disease.

Among the highest priorities of tuberculosis research are the identification of immunoprotective determinants for inclusion in a subunit vaccine and the identification of virulence determinants that may serve as new drug targets. In this regard, the 30-kDa major extracellular protein of *M. tuberculosis* is of great interest. With respect to its potential for immunoprotection, the 30-kDa protein has been demonstrated to be immunoprotective in the highly relevant guinea pig model of pulmonary tuberculosis (12). With respect to its potential as a drug target, the 30-kDa protein has been found to be among the most abundant proteins produced by *M. tuberculosis* growing either extracellularly in broth culture or intracellularly in human mononuclear phagocytes (17). This suggests that the protein plays a vital role in the physiology of the bacterium. A recent finding that the molecule has mycolic acid transferase activity indicates that it plays an important role in cell wall synthesis (2).

The 30-kDa protein is part of a complex of three highly related proteins. The two other members of the family have apparent molecular masses of approximately 32 kDa. The 30/32-kDa complex has often been referred to as the antigen 85 complex, the 30-kDa protein has been referred to as antigen 85B, and the two 32-kDa proteins have been referred to as antigens 85A and 85C. In this paper, we shall refer to the two 32-kDa proteins as the 32A (antigen 85A) and 32B (antigen 85C) proteins.

The 30/32-kDa protein complex has been the focus of intensive research over the past several years. For each of the three *M. tuberculosis* proteins, the nucleotide sequence of the gene encoding the protein (6, 8), the N-terminal amino acid sequence (12, 27), and other biochemical parameters such as mass and isoelectric point (27) have been reported. In addition, for the 32A protein, an analysis of the promoter structure (14), the mapping of T-cell epitopes (16), and the expression of a recombinant fusion molecule (3) have been described.

Much remains to be learned about this protein complex of *M. tuberculosis*. Seeking to fill some of the important gaps in our knowledge, we have undertaken a detailed genetic, biochemical, and immunocytochemical characterization of the 30-kDa protein of *M. tuberculosis* and extended some of these studies to the other two members of the 30/32-kDa complex to highlight similarities and differences among the three proteins. Our novel findings about this protein complex are organized into three parts. The first part presents our genetic analyses,

* Corresponding author. Mailing address: Division of Infectious Diseases, Department of Medicine, University of California Los Angeles School of Medicine, 10833 Le Conte Ave., Los Angeles, CA 90095. Phone: (310) 206-0074. Fax: (310) 794-7156.

including (i) sequence analyses of regions flanking the 30 and 32A protein genes, which together with mRNA analyses indicate that the genes are arranged as separate transcription units; (ii) homology studies showing only one or five amino acid differences between *M. tuberculosis* and two *Mycobacterium bovis* BCG strains; (iii) characterization of the transcription patterns of all three genes of the 30/32-kDa protein complex suggesting that expression is regulated at the transcriptional level; and (iv) the mapping of the transcriptional start site of the 30-kDa protein gene. The second part presents our biochemical characterization, including evidence that (i) the 30-kDa, 32A, and 32B proteins are secreted at a ratio of ~3:2:1, respectively; (ii) the three proteins exist as monomers; (iii) the three proteins are not posttranslationally modified; (iv) the 30-kDa and 32A proteins contain two cysteines linked by a disulfide bridge and one free cysteine; and (v) high-level expression and leader peptide processing of recombinant 30-kDa protein is achievable in *Escherichia coli*. The third part presents our immunocytochemical characterization of the 30/32-kDa protein complex. By using the cryosection immunogold technique, we demonstrate that the proteins of the 30/32-kDa complex are expressed in the *M. tuberculosis* phagosome in infected human monocytes and that most of the protein molecules are localized on the mycobacterial cell wall.

MATERIALS AND METHODS

***M. tuberculosis* cultures.** *M. tuberculosis* Erdman (ATCC 35801) was used for all procedures in this study and maintained in Middlebrook 7H9 medium (Difco) at pH 6.7 and 37°C in 5% CO_2–95% air as described previously (22).

***E. coli* strains.** *E. coli* DH5α [Gibco; F^- *endA1 hsdR17* (r_K^- m_K^+) *supE44 thi-1* lambda$^-$ *recA1 gyrA96 relA1 deoR* φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169] was used for general cloning purposes. *E. coli* BL21(DE3)pLysS [Novagen; DE3 lysogen, F^- *ompT* (r_B^- m_B^-)] was used to express recombinant-30 kDa protein in pET22b(+).

Reagents and enzymes. All chemicals and enzymes were purchased from Sigma unless indicated otherwise and were of the highest grade possible. Oligonucleotides were from either BioSynthesis or the UCLA School of Medicine Departmental Core Facility.

Genetic analyses of the 30/32-kDa complex. (i) Construction of *M. tuberculosis* Erdman genomic library. High-molecular-weight genomic DNA was isolated from a 3-week-old culture of *M. tuberculosis* Erdman by vortexing the cell pellet with 60-mesh crystalline alumina beads (Fisher) for 40 s at room temperature, extracting with phenol at 60°C, treating the extract with proteinase K and RNase A (Boehringer) for 2 h at 37°C at 100 μg/ml each, and precipitating the DNA twice with ethanol. The DNA was stored either as an ethanol precipitate or at 4°C in 10 mM Tris-HCl–1 mM EDTA (pH 7.5).

For the construction of the genomic library, high-molecular-weight mycobacterial genomic DNA was partially digested with *Hin*dIII and *Bam*HI and restriction fragments between 1.5 and 6 kb were ligated into pUC19 (28) that had been digested with the same restriction enzymes. Transformants of *E. coli* DH5α were amplified and stored at −70°C. The average insert size was ~3 kb, and the entire library contained ~25 genomic equivalents.

(ii) Isolation of the genes encoding the 30-kDa and 32A proteins. Two oligonucleotides were used for the amplification from genomic *M. tuberculosis* DNA of the full-length structural gene encoding the 30-kDa protein. The design of the two oligonucleotides was based on the DNA sequence of its *M. bovis* BCG homolog (6). The 5′ primer (5′-AAATGGATCCGGGGCACAGGTATGACA GACGTGAGCC-3′) contained a *Bam*HI restriction site and annealed to nucleotides −11 to +16 relative to the A residue of the initiator methionine codon ATG; the 3′ primer (5′-ACTAAAGCTTTAAGCAACCTTCGGTTGATCCCG TCAGC-3′), located downstream of the stop codon, contained a *Hin*dIII restriction site and annealed to nucleotides +1,002 to +974 relative to the initiator methionine codon ATG. Directional cloning into pUC19 and determination of the DNA sequence (23) across the plasmid-amplification product junctions with pUC19 forward and reverse sequencing primers (Gibco BRL) proved that the correct target sequence had been cloned.

The PCR product was radiolabelled with 5′-[α-^{32}P]dCTP (Amersham) and used as a probe in three successive colony hybridizations of the genomic DNA library at high stringency (9). Clones positives for the 30-kDa protein gene were further characterized by restriction enzyme mapping and DNA sequencing.

Cloning and sequencing of the gene encoding the 32A protein followed the procedures described for the 30-kDa protein. The DNA sequences of the amplification primers were as follows (6): the 5′ primer (5′-GGATCCATGCAGC TTGTTGACAGGGTTCGT-3′) carried a *Bam*HI restriction site and annealed to nucleotides +1 to +24 relative to the initiator methionine codon ATG; the 3′ end amplification primer (5′-GTTGTGTCTGTTCGGAGCTAGGCGCCTG-3′) annealed to nucleotides +1,034 to +1,006 relative to the initiator methionine codon ATG.

The nucleotide sequences for the genes encoding the 30-kDa and 32A proteins were determined independently for both DNA strands by using the double-stranded DNA sequencing strategy (4) and sets of primers for approximately every 250 nucleotides of gene sequence.

DNA sequence alignments were performed with the GCG sequence analysis software package version 7 from the University of Wisconsin, Madison.

(iii) Northern (RNA) blots and primer extensions. Total RNA was isolated from 3-week-old cultures by lysis of bacilli with sodium dodecyl sulfate and hot phenol extraction. RNA was digested with proteinase K and DNase I at 100 μg/ml for 30 min at 37°C, precipitated twice with ethanol, and stored at −20°C in 0.5% sodium dodecyl sulfate. The integrity of the RNA preparation was verified by the presence of two sharp rRNA bands, 1,550 and 3,150 nucleotides in length.

For Northern blots, RNA samples (25 μg) were run on 2.2 M formaldehyde-agarose gels and transferred in 20× SSC (1× SSC is 0.15 M NaCl and 0.015 M sodium citrate) to nitrocellulose filters (Schleicher & Schuell). Membranes were baked for 2 h at 60°C, and hybridization was performed in 500 mM sodium phosphate (pH 7.2)–7% sodium dodecyl sulfate at 55°C for 42 h with an anti-RNA sense oligonucleotide labelled with [γ-^{32}P]ATP to 10^8 cpm/μg. This oligonucleotide probe was specifically targeted to nucleotide positions +240 to +221 on the 30-kDa protein mRNA relative to the initiator methionine codon ATG (5′-GTCGAGCAGATAAACCGCAG-3′). The oligonucleotide specific for the 32A protein mRNA is located in the 3′ extension region at nucleotide positions +1013 to +991 relative to the initiator methionine codon ATG (5′-GCGCCC TGGGGCGCGGGCCCGGT-3′). The oligonucleotide specific for the 32B protein mRNA was targeted to the leader sequence of this gene at nucleotide positions +42 to +8 relative to the initiator methionine codon ATG (5′-CGCT GCGCTCCGCAACCTTCGCACCTGTTCGAAGA-3′). Membranes were washed with 50 mM sodium phosphate (pH 7.2)–0.1% sodium dodecyl sulfate for a total of 3 h at the hybridization temperature and exposed to X-ray film (Fuji RX) for 2 days.

Primer extension analyses were performed by annealing an anti-RNA sense oligonucleotide to *M. tuberculosis* RNA at 70°C and extending the primer for 15 min at 50°C with reverse transcriptase (Promega and Stratagene) in the presence of [α-^{32}P]dATP. The primer (5′-CGCGCCCGCGGTTGCCGCTCCGCCGGC AAG-3′) annealed to nucleotides +120 to +91 of the 30-kDa protein gene relative to the initiator methionine codon ATG. The primer extension products were electrophoresed on a 6% polyacrylamide–8 M urea gel alongside the sequencing products obtained with the same oligonucleotide primer. The gel was fixed in methanol-acetic acid and exposed to X-ray film for 2 days.

Biochemical analyses of the 30/32-kDa protein complex. (i) Purification and characterization of native proteins of the 30/32-kDa complex. Proteins of the 30/32-kDa complex were purified from culture filtrates of *M. tuberculosis* Erdman as described previously (12). Homogeneity of the proteins was ascertained by running aliquots on denaturing polyacrylamide gels stained with Coomassie brilliant blue R or silver nitrate.

The N-terminal amino acid sequences of the three homogeneous proteins were determined by running aliquots on denaturing polyacrylamide gels, transferring the proteins to polyvinylidene fluoride membranes in 10 mM 3-[cyclohexylamino]-1-propanesulfonic acid (CAPS), and subjecting the Coomassie brilliant blue R-stained protein bands to automated amino acid sequence determination at the UCLA Protein Microsequencing Facility.

Hydropathy patterns, structural features, and isoelectric points of the three proteins were determined and plotted as described previously (15, 18).

(ii) Carbohydrate and lipid analyses of the 30/32-kDa complex. The potential presence of carbohydrate modifications on proteins of the 30/32-kDa complex was assessed in two ways. Aliquots of homogeneously purified preparations of the three proteins were digested with endoglycosidases D, F, and H (Boehringer) under conditions of complete cleavage of carbohydrate residues from their covalently attached polypeptide chains as judged by digestion of test substrates. As a chemical alternative, the proteins were treated with 0.5 M NaOH–1 mM calcium chloride for 24 h at 40°C, which removes carbohydrates from the polypeptide chains irrespective of the nature of the carbohydrate residue linked to the polypeptide chain and the type and structure of this chemical bond. Aliquots of endoglycosidase- or NaOH-treated and untreated proteins of the 30/32-kDa complex were run on polyacrylamide gels and stained with Coomassie brilliant blue R.

The presence of modifications by lipid compounds was assessed by spotting 10 to 30 μg of each individual protein of the 30/32-kDa complex on thin-layer chromatography plates (Analtech), spraying the plates with 0.01% rhodamine 6G, and examining the plates for fluorescence at 365 nm. Control substrates included human plasma apolipoprotein A1, human plasma lipoprotein of the high-density type, Tween 80, *M. tuberculosis* glutamine synthetase (10), deoxycholic acid (Boehringer), digitoxin (Boehringer), sodium dodecyl sulfate, and bovine serum albumin. The sensitivity level of this system was determined by spotting defined amounts of human plasma lipoprotein of the high-density type and deoxycholic acid ranging from 50 to 0.05 μg on the thin-layer chromatography plates and analyzing the spots for fluorescence at 365 nm.

(iii) Analysis of intramolecular disulfide bonds. The 32B protein contains only one cysteine residue; the 30-kDa and 32A proteins, however, contain three cysteine residues which can potentially be in several configurations. Free thiol groups on the proteins were detected by coupling them to the fluorescent molecule *N*-(iodoacetylaminoethyl)-5-naphthylamine-1-sulfonic acid (I-AEDANS; Aldrich). Fluorescence of the I-AEDANS-labelled proteins was quantified with a fluorometer at an excitation wavelength of 300 nm (26). Fluorescence labelling of thiol groups involved in disulfide bonds was carried out by first modifying the free or potentially esterified thiol group with iodoacetamide and then incubating the proteins with I-AEDANS before and after reduction with dithiothreitol (DTT). Fluorescence labelling of free thiol groups on the proteins was carried out by incubating the proteins with I-AEDANS, followed by a second incubation step in the presence of 20 mM DTT and 50 mM iodoacetamide.

(iv) Cloning and expression of recombinant full-length and mature 30-kDa and 32A proteins. Recombinant plasmids carrying either the full-length structural gene for the 30-kDa or 32A protein, including its own leader sequence, or its truncated version coding for the mature protein but lacking the leader sequence were constructed from amplification products derived from the original genomic clone and vectors of the pET expression series (Novagen). The 5′ amplification primers contained the nucleotide sequences corresponding to amino acids −40 to −32 (the 30-kDa protein gene) or amino acids −43 to −35 (the 32A protein gene) for expression of the full-length proteins or to amino acids +1 to +8 for expression of the mature proteins. In addition, both primers contained restriction sites for the directional insertion of the amplification products into pET22b(+). The 3′ primers were the same oligonucleotides described above for the cloning of the 30-kDa and 32A protein genes from the genomic library. Two clones for each construct were sequenced to verify the presence of the correct amplification products.

For the expression of recombinant 30-kDa and 32A proteins in their full-length or truncated versions, constructs in pET22b(+) were expressed in *E. coli* BL21(DE3)pLysS upon induction with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Samples of the induced cultures were taken at hourly intervals for up to 8 h, and aliquots of the culture supernatants and cell pellets were run on 12.5% denaturing polyacrylamide gels and stained with Coomassie brilliant blue R. Recombinant protein was purified as described previously (12), with the exception that all chromatography steps included the addition of 8 M urea to the buffers. Purified recombinant protein was dialyzed against phosphate-buffered saline and remained soluble. The yield of both homogeneous recombinant proteins from 1 liter of *E. coli* culture was approximately 10 to 20 mg.

Immunocytochemical characterization of the 30/32-kDa complex. (i) Generation of antibodies against proteins of the 30/32-kDa complex. Antibodies specific for the proteins of the 30/32-kDa complex were raised in female New Zealand White rabbits by four subcutaneous immunizations at 3-week intervals with 4 × 100 μg of the 30-kDa protein in the presence of Syntex adjuvant formulation (1). At the first immunization, 100 μg of *N*-acetylmuramyl-L-alanine-D-isoglutamine was added to the protein sample. Two weeks after the fourth immunization, the rabbits were bled and the serum samples were stored at −20°C until use. Preimmune serum was obtained before the first immunization and stored at −20°C.

(ii) Specificity of antibodies against the 30/32-kDa complex. Antibodies raised in rabbits against homogeneously purified 30-kDa protein were assayed against both the purified protein and a concentrated culture filtrate of *M. tuberculosis* Erdman. The antibodies were highly specific for the 30/32-kDa complex. At a constant antigen concentration of 1 μg, the monospecific antibodies had a reciprocal titer of 20,000. At a 1:1,000 dilution, the antibodies recognized as little as 1 ng of antigen.

(iii) Western blot analyses of proteins of the 30/32-kDa complex. Examination of proteins of the 30/32-kDa complex by immunoblotting was carried out either with purified protein preparations or crude cell extracts. Typically, ~1 μg of each purified protein or 25 μg of *E. coli* cell extract containing recombinant protein was electrophoresed on denaturing polyacrylamide gels, transferred to nitrocellulose membranes, and probed with antibodies specific for the 30/32-kDa complex at a dilution of 1:2,500. Following incubation with peroxidase-labelled goat anti-rabbit antibodies, proteins of the 30/32-kDa complex were visualized by staining with 3,3′-diaminobenzidine.

(iv) Detection of in vivo-released proteins of the 30/32-kDa complex. Human mononuclear cells were isolated from fresh heparinized blood as described previously (5). After 5 days in monolayer culture, the monocytes were infected with *M. tuberculosis* at a multiplicity of infection of 1:10. Two days later, lysosomes were labelled by incubation with 20- to 25-nm-diameter mannan-gold particles (final optical density at 520 nm, 0.5) for 8 to 12 h at 37°C, washed four times with RPMI 1640 culture medium (Irvine Scientific), and incubated overnight at 37°C in RPMI 1640 medium containing 15% autologous serum. Monocytes were fixed in 0.1 M piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) (PIPES; pH 7.3), containing 4% paraformaldehyde and 0.01% glutaraldehyde, for 2 h at 4°C and washed with 0.1 M PIPES (pH 7.3). Monocytes were embedded in 10% gelatin and cryosectioned, and nonspecific antibody binding was blocked with 50% normal goat serum as described previously (5). The presence of proteins of the 30/32-kDa complex in phagosomes of infected human monocytes was examined by incubating cryosections of infected monocytes first with pre- or postimmune anti-30/32-kDa complex immunoglobulin G (IgG) at a 1:500 dilution and then with goat anti-rabbit IgG conjugated to 10-nm-diameter gold particles. In additional experiments, late endosomes and lysosomes were identified by incubation of sections with mouse monoclonal antibody to human CD63 (20 μg/ml; AMAC) followed by goat anti-mouse IgG conjugated to 10-nm-diameter gold particles. In these experiments, proteins of the 30/32-kDa complex were identified by incubation of cryosections with 1:500-diluted rabbit anti-30/32-kDa complex antibodies followed by goat anti-rabbit IgG conjugated to 15-nm-diameter gold particles. The stained sections were evaluated by transmission electron microscopy (5).

Nucleotide sequence accession numbers. The nucleotide sequences for the 30-kDa and 32A protein genes have been submitted to GenBank under the accession numbers MTB30-8 U38939 and MTB32-2 U47335, respectively.

RESULTS

Genetic analyses of the 30/32-kDa complex. (i) Identification of the genes encoding the 30-kDa and 32A proteins. The first step in our strategy to clone the gene for the 30-kDa protein involved amplification of the gene sequence from *M. tuberculosis* Erdman DNA by PCR. Oligodeoxyribonucleotide primers were therefore designed on the basis of the sequences of the 5′ and 3′ ends of the gene encoding the 30-kDa protein of *M. bovis* BCG (6), the only DNA sequence of a pathogenic mycobacterial species available at the time. A single amplification product of the expected size (~1 kb) was obtained and cloned into bacterial plasmids maintained in *E. coli* DH5α. Three clones were analyzed at the nucleotide level and shown by comparison with published data (6) to contain the DNA sequence coding for the 30-kDa protein. This PCR-derived sequence for the 30-kDa protein was then used as a probe to isolate the genomic clone encoding the 30-kDa protein from a genomic *M. tuberculosis* Erdman DNA library by three rounds of colony hybridizations (9). Restriction enzyme mapping and DNA sequence determination of several clones identified the 30-kDa protein gene as the longest continuous open reading frame on a 4.7-kb DNA fragment flanked by *Hin*dIII-*Bam*HI restriction sites. The strategy for the cloning and sequencing of the clone containing the 32A protein gene followed the steps outlined above by use of PCR primers derived from published DNA sequences (6).

(ii) Nucleotide sequence analysis of the genes of the 30/32-kDa complex. The homology among the proteins of the 30/32-kDa complex is very striking (12); in addition, this high degree of homology is also observed at the nucleotide level of the corresponding genes (6). On the basis of this observation, we established the DNA sequences not only for the structural genes of the 30-kDa and 32A proteins but also for large regions flanking the structural genes for several reasons. First, we wanted to see if the individual genes are arranged in tandem. This gene configuration is one of several ways microorganisms express proteins at high levels. This is obviously not the case, either for the 30-kDa or 32A protein gene. Second, we wanted to assess the possibility that these related genes are arranged in tight gene clusters, since these proteins are expressed at high levels throughout the entire 3-week growth phase of *M. tuberculosis* with no change in their relative ratios (30-kDa protein/32A protein/32B protein ratio, 3:2:1; see below). Both genes are flanked by unrelated DNA sequences, suggesting that this family of related genes is not organized into one transcription unit but rather broken up into single operons. For the 30-kDa protein gene, we have determined the DNA sequence of the aforementioned 4.7-kb clone; for the 32A protein gene, we have determined 5 kb of flanking DNA sequence. In both cases, there is no overlap of DNA sequences, indicating that these genes are separated by at least 5 kb. For an alignment of the DNA sequences of the 30-kDa and 32A protein genes, we focused on the structural genes and the adjacent 300 nucleotides at the 5′ and 3′ ends (Fig. 1). It is evident from the presented data that the DNA homology (~80% identity for the DNA sequence encoding the mature 30/32-kDa complex

FIG. 1. Alignment of DNA sequences of the 30-kDa (30K) and 32A (32K) protein genes. The sequences show the 5′ → 3′ DNA strand of the structural genes (boldface) of the 30-kDa protein gene (top strand) and 32A protein gene (bottom strand). The initiator methionine codons (*) of the two genes are aligned, necessitating the introduction of hyphens in the 30-kDa protein gene sequence since its leader is 3 amino acids shorter than that of the 32A protein gene. A solid black dot indicates the first codon of the mature proteins, TTC for the 30-kDa protein gene and TTT for the 32A protein gene, specifying phenylalanine in both cases. The structural genes are flanked by 300 nucleotides at both the 5′ and 3′ ends to highlight the homology of the structural genes and the sequence divergence in the flanking regions. Homologous DNA sequences between the two genes are indicated by dashes; divergent DNA sequences are shown by the actual single-letter base code.

proteins but only ~36% identity for the sequence encoding the leader peptides) of the genes of the 30/32-kDa protein family is confined to the open reading frames of the structural genes. In contrast, the DNA homology drops below significant levels in both the 5′ and 3′ flanking regions of the genes (~25 to 30% identity), even after optimizing the homology between these DNA sequences by a best-fit computer analysis. Furthermore, within the 30/32-kDa complex, the homology is significantly higher between the 30-kDa and 32A protein genes than between either of these two genes and the 32B protein gene (reference 6 and present study). The overall G+C content of the open reading frames is ~65%, and the codon usage of the 30-kDa and 32A protein genes is very similar to the one described for mycobacteriophage L5 (11). For the 30-kDa protein gene, a potential Shine-Dalgarno sequence (24) is present within a 15-bp stretch upstream of the first methionine codon of the structural gene. Assignment of promoter elements of mycobacterial genes is still very tentative at best; two sites at ~47 bp and ~71 bp upstream of the start codon of the structural gene were selected as potential candidates of promoter elements of the 30-kDa protein gene (Fig. 1). The arrangement of regulatory DNA sequences, including promoter elements and transcription start sites, has been described recently for the 32A protein gene (14).

(iii) Homology of the 30-kDa protein to its counterparts in other mycobacterial species. At the amino acid level, the highest degree of homology, nearly 100% identity, was observed between the *M. tuberculosis* Erdman 30-kDa protein and its *M.*

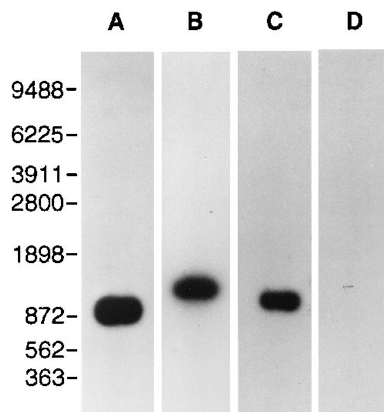

FIG. 2. Northern blot analyses of mRNAs encoding the 30-kDa, 32A, and 32B proteins. Total RNA from *M. tuberculosis* Erdman was probed with the following oligonucleotides: the 30-kDa protein mRNA probe (lane A); the 32A protein mRNA probe (lane B); the 32B protein mRNA probe (lane C); total RNA, RNase A digested and hybridized to the 30-kDa protein mRNA probe (lane D). The numbers on the left refer to RNA size standards in nucleotides.

bovis BCG counterpart (6, 20). The sequence differences involve changes of amino acid residues (Phe-100 → Leu, *M. tuberculosis* → *M. bovis* BCG 1137P2 and Tokyo; Asn-245/Ala-246 → Lys/Pro, *M. tuberculosis* → *M. bovis* BCG Tokyo) and the insertion of two *M. tuberculosis* amino acid residues, Pro and Ser, between amino acids 160 (Gly) and 161 (Leu) in the *M. bovis* BCG Tokyo sequence.

Alignment of the *M. tuberculosis* 30-kDa protein gene with the corresponding gene in other mycobacteria revealed homologies ranging from 90 to 82%: *Mycobacterium kansasii*, 90% (19); *Mycobacterium intracellulare*, 84% (13); *Mycobacterium avium*, 83% (21); and *Mycobacterium leprae*, 82% (7). No homologies to other DNA sequences deposited in GenBank and EMBL databases were detected.

(iv) Analysis of mRNA molecules encoding the 30/32-kDa proteins. Transcription of the genes coding for the 30/32-kDa proteins was analyzed by Northern blot analyses. These assays were designed to answer the following questions. (i) What are the sizes of the transcripts of the genes coding for the 30/32-kDa complex? (ii) Are precursors of these transcripts identifiable? (iii) Do the intensities of the hybridization signals reflect the relative amounts of each protein in the culture medium? To address these issues, we isolated total RNA from logarithmically growing cultures of *M. tuberculosis* Erdman and hybridized equal amounts of this RNA preparation to the following antisense oligonucleotides whose coordinates are given in their 5′ → 3′ direction relative to the A residue in the methionine initiator codon ATG: the 30-kDa protein gene, +240 → +221; the 32A protein gene, +1,013 → +991, in the 3′ extension region; the 32B protein gene, +42 → +8, in the gene's leader sequence. Single bands of ~1,000 to 1,050 nucleotides, ~1,100 to 1,150 nucleotides, and ~1,050 to 1,000 nucleotides were consistently observed for the 30-kDa, 32A, and 32B protein genes, respectively (Fig. 2). No larger transcripts were detected, although blots were exposed for longer periods of time than required to visualize the specific hybridizations. Since the hybridizations were carried out under comparable conditions with equivalent amounts of RNA per filter strip and probes labelled to the same intensity, the observed hybridization signals suggest that expression of the 30/32-kDa complex is regulated at the transcriptional level. In fact, densitometry performed on the X-ray films showed that the hybridization signals for the 30-kDa, 32A, and 32B protein mRNAs decreased from 1,200 to 660 to 400 arbitrary optical density units, respectively. This ratio is remarkably similar to the ratio with which these proteins are secreted into the culture medium (30-kDa protein/32A protein/32B protein ratio, 3:2:1). These results also strengthen the argument that each gene of the 30/32-kDa complex is arranged as a single transcription unit.

RNase A treatment of the RNA preparation abolished the hybridization signal, but DNase I treatment did not. The possibility of a chance hybridization to rRNA species was excluded by the hybridization pattern observed with rRNA-specific probes, i.e., one band of ~1,500 to 1,600 nucleotides for the small rRNA species and another band of ~3,000 to 3,200 nucleotides for the large rRNA molecule. These hybridizations yielded consistently strong signals of equal intensities, by use of 10 and 25 μg of total RNA, further indicating that the observed differences in the intensities of the hybridization signals reflect regulation of 30/32-kDa gene expression at the transcriptional level. The two rRNA species were also visible after ethidium bromide staining of the agarose gels, providing evidence that the RNA preparations used for the Northern analyses were of high integrity.

(v) Mapping of the transcriptional start site for the 30-kDa protein gene. The transcriptional start site for the 30-kDa protein gene was determined by primer extension studies using an antisense oligonucleotide primer whose 3′ end corresponded to nucleotide 91 relative to the A residue of the methionine initiator codon ATG. Annealing of this 30-mer oligonucleotide primer to total RNA followed by reverse transcription of the specifically recognized 30-kDa protein mRNA produced several extension products (Fig. 3). The largest product, a weak doublet, is consistent with a start at C or G, 6 or 7 nucleotides downstream of the motif 5′ TATGTTC 3′, a tentative promoter element. The strongest signal and possibly the

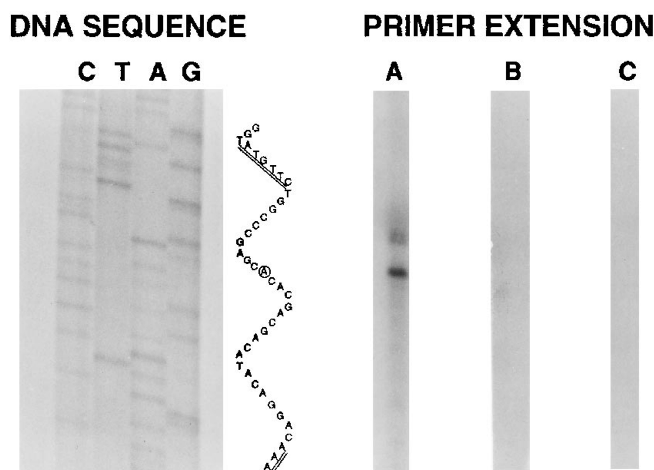

FIG. 3. Primer extension analysis of the 30-kDa protein mRNA. An anti-mRNA sense oligonucleotide, located at nucleotide positions +120 to +91 relative to the A residue of the initiator methionine codon ATG of the 30-kDa protein gene, was annealed to total *M. tuberculosis* RNA and extended with reverse transcriptase. (Right) Primer extension. Lanes: A, the 30-kDa protein mRNA specific extension product; B, primer extension products obtained with total RNA digested with RNase A; C, primer extension products obtained with no primer added. (Left) DNA sequence. The sequence ladder was obtained with the same oligonucleotide; nucleotide-specific lanes are marked CTAG. The tentatively assigned −10 promoter region, TATGTTC, and the start of the Shine-Dalgarno sequence, AAA, are underlined. The most likely initiator nucleotide of the mRNA, A, is circled.

most likely transcriptional start site maps to an A residue, just 4 nucleotides downstream of the weak doublet mentioned above. This site is 33 bp upstream of the ATG start codon of the structural gene. RNase treatment and omission of the primers abolished the observed extension products. No stable hairpin loop could be detected in this region; however, this result does not completely exclude the possibility that the strongest of the three signals represents a strong reverse transcriptase stop and not the actual 5′ end of the 30-kDa protein gene transcript.

Biochemical analyses of the 30/32-kDa protein complex. (i) Isolation of homogeneously purified 30-kDa, 32A, and 32B proteins and relative quantities in culture filtrates. The 30/32-kDa proteins were purified from concentrated culture filtrate of *M. tuberculosis* Erdman by column chromatography to apparent homogeneity, yielding a single band by denaturing polyacrylamide gel electrophoresis and a single N-terminal amino acid sequence extending to amino acid residue 40. The yield of purified 30-kDa protein from 1 liter of a culture grown for 2.5 weeks from an optical density (OD) at 540 nm of 0.05 to an OD at 540 nm of 0.5 was 600 μg or 60% of the amount originally present as extracellular protein. The yields of purified 32A and 32B protein from 1 liter of culture filtrate were 300 and 100 μg, respectively, approximately 50% of the material originally present. The proteins of the 30/32-kDa complex combined accounted for ~45% of the total amount of extracellular proteins, with the 30-kDa protein corresponding to 22%, the 32A protein corresponding to 15%, and the 32B protein corresponding to 8%.

(ii) Structural features of the 30-kDa protein. The 30-kDa protein is expressed as a preprotein which is processed as it is released from the mycobacterial cell into the culture filtrate (29). By gel filtration chromatography, the mature protein has an estimated molecular mass of ~30 kDa, suggesting that the 30-kDa protein does not contain any subunits and does not form higher-order structures combining several identical subunits (reference 12 and present study).

Interestingly, the *M. tuberculosis* 30-kDa protein appears to be rich in aromatic amino acid residues, which amount to ≥12% of the protein's amino acids. Another feature of the sequence is the presence at ≥20 positions of tandemly arranged, identical amino acid residues. A structural analysis revealed that 45% of the polypeptide chain is arranged as α-helices, 25% is arranged as β-sheets, and the rest is arranged as turns. The hydropathy pattern (15) displayed a fifty-fifty distribution of hydrophilic and hydrophobic domains. The isoelectric point of the 30-kDa protein was calculated to be 6.6.

The 32A and 32B proteins are also processed to their mature form as they are released from the mycobacterial cell into the culture filtrate. The 32A and 32B proteins share many of the biochemical features of the 30-kDa protein. They lack higher-order structure and have similar amino acid compositions, structural arrangements, hydropathy profiles, and isoelectric points (6.9 for the 32A protein and 6.5 for the 32B protein).

(iii) Evidence against posttranslational modifications of the 30-kDa protein. An important feature of proteins with regard to their immunogenicity is the presence of modifications such as carbohydrate and/or lipid moieties. We therefore undertook a detailed analysis of potential modifications using several approaches. The absence of any carbohydrate modification was ascertained by digesting the purified 30-kDa protein with endoglycosidases F, G, and H, or by treatment with 0.5 M NaOH and analyzing treated and untreated protein samples by gel electrophoresis. There was no indication of any modification by carbohydrate residues. The absence of lipids associated with the 30-kDa protein was demonstrated by subjecting the protein to thin-layer chromatography on plates and developing the protein spots with rhodamine 6G. In contrast to other known lipid-containing compounds, the 30-kDa protein did not show any UV fluorescence. Given a detection limit of 0.1 μg of lipid, a 25-μg sample of the 30-kDa protein should have produced a strong fluorescence signal, even if there was only one fatty acid chain of ~750 to 1,000 Da attached to the protein. Similarly, in the same assays, no posttranslational modifications of purified samples of the 32A and 32B proteins were detected.

(iv) Demonstration of a disulfide bridge in the 30-kDa and 32A proteins. The 30-kDa protein has three cysteines, two of which are separated from each other by only four amino acids. The presence of a disulfide bridge between two cysteines was demonstrated by labelling free −SH groups with I-AEDANS (26). In the first set of assays, the purified 30-kDa protein was derivatized with iodoacetamide to mask free or esterified thiol groups and then labelled with I-AEDANS either with or without DTT treatment. Without DTT treatment, no I-AEDANS molecules were coupled to the 30-kDa protein, whereas with DTT treatment, two fluorescence molecules were coupled to the protein. This result is consistent with the presence of one disulfide bridge. In a second set of assays, the 30-kDa protein was first labelled with I-AEDANS and then derivatized with iodoacetamide and DTT. In this case, one I-AEDANS molecule was coupled to the 30-kDa protein. This result is consistent with the presence on the protein of one free thiol group. A similar result was obtained for the 32A protein, which also contains three cysteine residues. Since there is only one cysteine in the 32B protein, no intramolecular disulfide bridge is possible.

(v) Expression and leader peptide processing of recombinant 30-kDa and 32A proteins. The proteins of the 30/32-kDa complex are synthesized with long leader peptides, namely, 40 amino acid residues in the case of the 30-kDa protein, 43 amino acids in the case of the 32A protein, and 46 amino acids in the case of the 32B protein (6, 29). Interestingly, sequences around the signal peptidase cleavage site follow the rule that the −1 and −3 positions in prokaryotic leader peptides are most likely occupied by alanine residues (25). Several bacterial expression systems were assessed for their capacity to express both the full-length and processed, mature 30-kDa and 32A proteins as recombinant molecules. Constructs in pET22b(+) were chosen for a detailed analysis of the expression of recombinant 30-kDa and 32A proteins, with both their own leader peptide and the *E. coli*-derived PelB leader as outlined for the 30-kDa protein gene constructs (Fig. 4).

For the expression of the full-length native proteins, the genes for the 30-kDa and 32A proteins were cloned into pET22b(+) such that the initiator methionine of each of the two proteins was the first methionine codon downstream of the expression vector's Shine-Dalgarno sequence. Strong expression of the recombinant protein was observed and verified by immunoblotting and N-terminal amino acid sequencing as described below (Fig. 5).

In a second set of recombinant vector constructs, the genes encoding the 30-kDa and 32A proteins were engineered such that the first residue of the mature proteins, phenylalanine, was fused to the carboxyl terminus of the PelB leader sequence, as shown for the 30-kDa protein gene constructs (Fig. 4). Since the genes for the mature proteins were cloned into the *Nco*I site at the 3′ end of the PelB leader, additional nucleotides specifying a glycine residue had to be inserted at this site to maintain the open reading frame extending from the PelB leader's methionine to the phenylalanine, the first amino acid residue of the mature proteins. Large quantities of the recombinant 30-kDa and 32A proteins were expressed in the cyto-

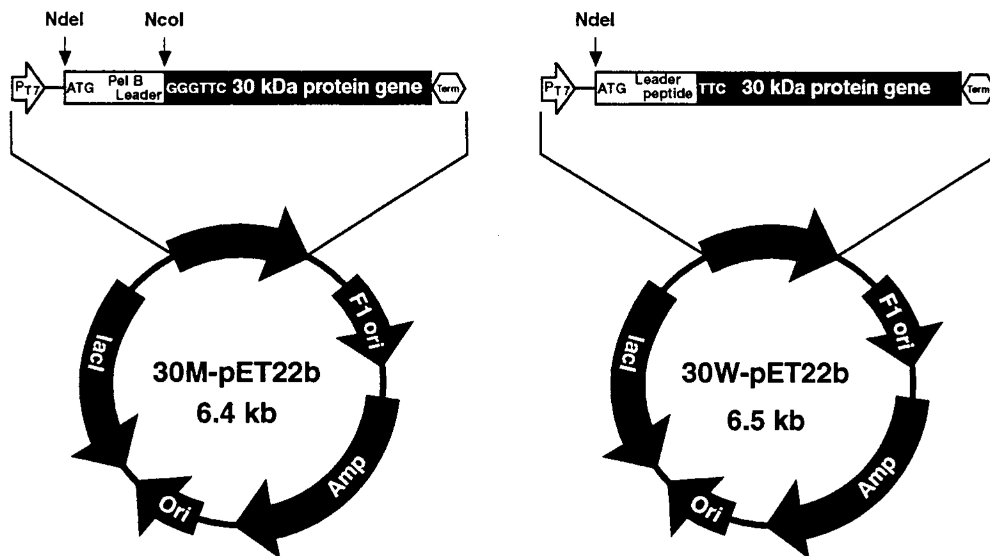

FIG. 4. Constructs used for the expression of recombinant 30-kDa and 32A proteins. The diagram depicts the pET22b(+) vector constructs used for the expression of the various forms of the recombinant 30-kDa protein. For the expression of the 32A protein, the 30-kDa gene cassettes were replaced with those encoding the 32A protein. The recombinant 30-kDa protein expressed with its own leader is labelled 30W-pET22b, while the form expressed with the plasmid-encoded PelB leader is labelled 30M-pET22b. Abbreviations: Ori, Col E1 type origin of replication; F1 ori, phage F1 origin of replication; Amp, ampicillin resistance gene; 30W and 30M, full-length (30W) or mature (30M) 30-kDa protein; lacI, lac repressor gene; P_{T7}, phage T7 RNA polymerase-specific promoter; *Nde*I and *Nco*I, restriction enzyme sites at vector-insert junctions.

plasm over a period of 4 h. Thereafter, expression of recombinant proteins reached a plateau. Expression of recombinant proteins continued for up to 8 h without exerting serious detrimental effects on the bacterial culture. A typical yield from 1 liter of *E. coli* culture was approximately 10 to 20 mg of recombinant proteins, amounting to ~5% of the total cell protein.

Whole-cell extracts from induced *E. coli* cultures expressing the various forms of the recombinant 30-kDa and 32A proteins were run on denaturing polyacrylamide gels, transferred to nitrocellulose membranes, and probed with rabbit anti-30/32-kDa complex antibodies. These specific antibodies very strongly and very specifically stained the different forms of the 30-kDa and 32A proteins in the crude cell extracts of induced bacteria but did not stain any protein in extracts from *E. coli* cultures carrying pET22b(+) alone. Preimmune antibodies did not react with any protein in extracts of induced bacteria. In a separate control blot with *M. tuberculosis* extracellular proteins, these antibodies were used to stain specifically the proteins of the 30/32-kDa complex.

As for the expression of the full-length 30-kDa protein, about 60% remained in the full-length form and 40% of the

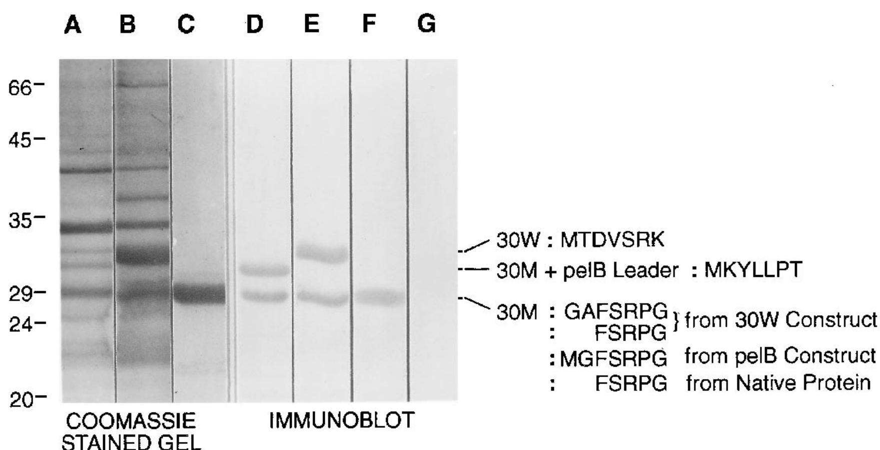

FIG. 5. Expression of full-length and mature 30-kDa protein in *E. coli* BL21(DE3)pLysS. The mature 30-kDa protein was expressed in the pET22b(+) vector with either its own or the plasmid-encoded PelB leader peptides. Lanes: A and B, Coomassie brilliant blue R-stained protein extracts upon IPTG induction of bacteria carrying the pET22b(+) vector with the mature 30-kDa protein gene fused to the PelB leader DNA sequence (A) and pET22b(+) containing the full-length 30-kDa protein gene (B); C, mature 30-kDa protein isolated from *M. tuberculosis* culture filtrates as a reference; D, E, and F, same protein extracts as those in lanes A, B, and C, respectively, probed with anti-30/32-kDa complex-specific antibodies; G, protein extract from *E. coli* cultures carrying pET22b(+) alone, probed with the anti-30/32-kDa complex-specific antibodies. Positions of full-length and mature 30-kDa proteins are marked 30W and 30M, respectively, and recombinant proteins are further identified by their first five or seven N-terminal amino acids. Recombinant forms of the 32A protein were analyzed in the same way (data not shown). The numbers on the left refer to molecular mass standards in kilodaltons.

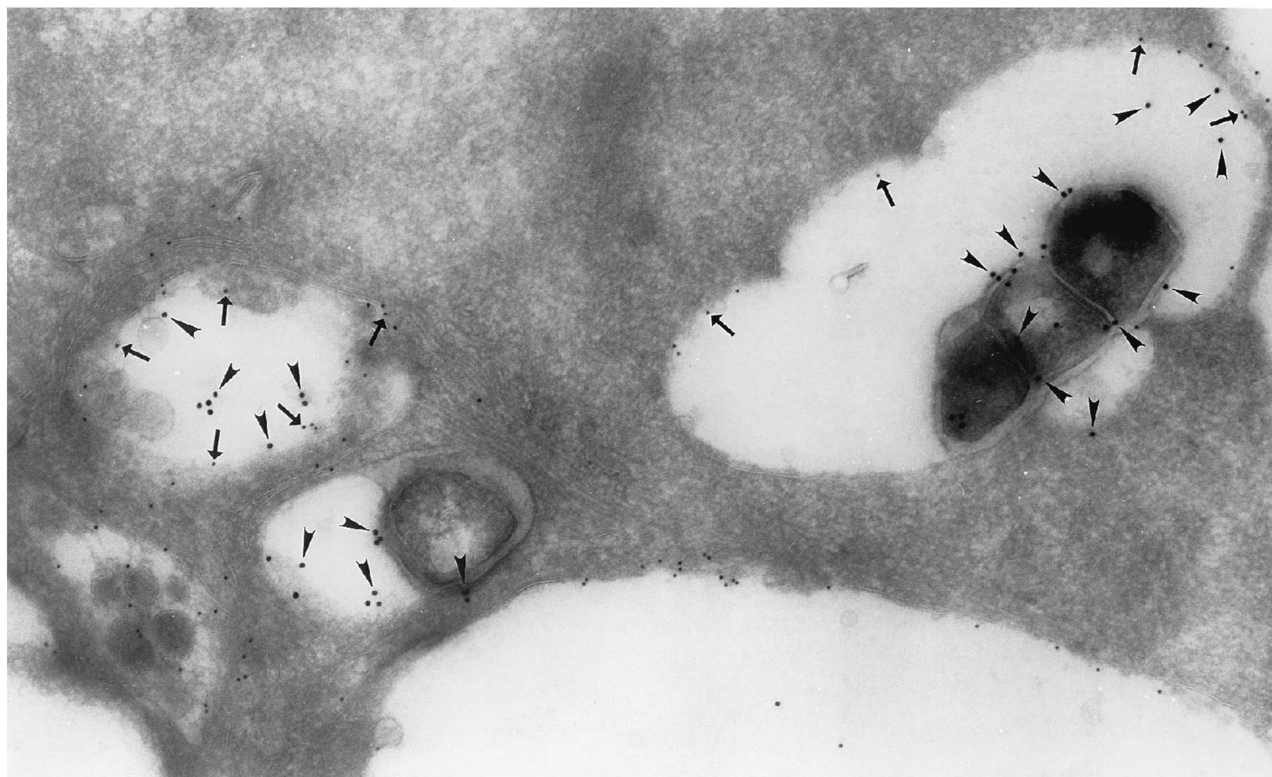

FIG. 6. Immunolocalization of *M. tuberculosis* 30/32-kDa complex. Ultracryomicroscopy of human monocytes reveals staining for proteins of the 30/32-kDa complex with 15-nm-diameter immunogold particles on the cell wall of the bacteria, in the phagosomal space, and in a cytoplasmic vesicle outside of the phagosome (arrowheads). The late endosomal-lysosomal marker CD63 has been stained with 10-nm-diameter immunogold particles (arrows). Magnification, ×56,000. Quantitation of gold particles is given in Table 1.

recombinant protein was processed (Fig. 5), while more than 90% of the full-length 32A protein was processed. The protein bands corresponding by migration to the proteins stained with the specific antibodies on the immunoblot were subjected to N-terminal amino acid sequence analysis. For both the 30-kDa and 32A proteins, one sequence corresponded to the N terminus of the respective full-length protein. Two other sequences were derived from the mature protein—one was identical to the sequence of the mature native *M. tuberculosis* 30-kDa or 32A protein and the other was 2 amino acid residues longer, suggesting that the *E. coli* processing enzymes did not uniformly cleave the correct peptide bond at the transition of the leader peptide and mature protein. This result indicates that secretory proteins are processed similarly in *E. coli* and *M. tuberculosis*. Processing of the *M. tuberculosis* protein in *E. coli* is evidently less efficient and less specific than it is in *M. tuberculosis*. An alanine residue is located at the −1 position of both cleavage sites, a fact which may have contributed to the somewhat sloppy processing of the recombinant preproteins in *E. coli* (25).

Expression of the mature 30-kDa and 32A proteins by constructs containing the PelB leader also yielded two forms of recombinant protein (Fig. 5), one form representing a fusion protein consisting of the PelB leader peptide and the mature protein and the second form representing the processed mature 30-kDa or 32A protein.

Immunocytochemical characterization of the 30/32-kDa protein complex: expression of the 30/32-kDa proteins in infected human monocytes and localization to the bacterial cell wall and phagosomal space. Although the 30-kDa protein is the most abundant extracellular protein of *M. tuberculosis*, no information has been available to date on its localization in *M. tuberculosis*-infected human cells. We therefore studied expression and release of the 30-kDa protein from the bacillus within infected human monocytes by immunogold electron microscopy. Cryosections of monocytes infected with *M. tuberculosis* at a multiplicity of infection of 0.1 were stained first with monospecific antibodies against the 30/32-kDa complex and then with gold-labelled secondary antibodies. Evaluation of approximately 200 *M. tuberculosis* phagosomes revealed that proteins of the 30/32-kDa complex are present on the bacterial cell wall and released into the phagosomal space (Fig. 6). The 30/32-kDa proteins were preferentially found within the phagosomal space and on the outer cell wall of the bacilli. For live bacilli, the combined number of gold particles found in the phagosomal space and associated with the mycobacterial cell wall exceeded the number of gold particles found inside the mycobacteria by a factor of 3 (Table 1). In these experiments, the number of gold particles identified as being released into the phagosomal space may be an underestimate because, in the case of *M. tuberculosis*, there is frequently little space between the bacterium and the phagosomal membrane. In addition to specific immunogold staining for the 30/32-kDa complex inside the phagosome, significant specific staining was also observed in cytoplasmic vacuoles outside the phagosome (Fig. 6 and Table 1). These vacuoles were often clearly identifiable as late endosomes or lysosomes by virtue of their labelling with the late endosomal-lysosomal marker CD63. A separate experiment revealed that heat-killed mycobacteria had significantly fewer gold particles (0.8 ± 0.2 gold particles per bacterium)

TABLE 1. Distribution of immunogold staining for the 30/32-kDa protein complex in infected human monocytes[a]

Serum	No. of immunogold particles/bacterium			No. of immunogold particles/μm^2	
	Phagosomal space[b]	Bacterial wall	Bacterial cytoplasm	Phagosome[c]	Host cell cytoplasm
Preimmune	0.01 ± 0.02	0.02 ± 0.03	0.09 ± 0.04	0.15 ± 0.09	0.06 ± 0.00
Immune	0.40 ± 0.05	4.70 ± 0.80	1.60 ± 0.40	11.80 ± 4.00	0.40 ± 0.10

[a] Human monocytes were infected with *M. tuberculosis* and stained for the presence of proteins of the 30/32-kDa complex by the cryosection immunogold technique as described in Materials and Methods. Data are the means ± standard deviations of ≥20 phagosomes on each of two electron microscopy grids for each serum condition.
[b] Area within phagosome exclusive of the bacterium.
[c] Entire area of phagosome including bacterium.

than live organisms (4.2 ± 2.1 gold particles per bacterium), and preimmune antibodies gave negligible background staining for the 30/32-kDa proteins.

DISCUSSION

The major secretory protein of *M. tuberculosis*, a protein with an apparent molecular weight of 30,000, is part of a complex of at least three proteins with a high degree of homology and nearly identical molecular weights. Although very closely related, the three *M. tuberculosis* proteins of the 30/32-kDa complex are separate entities encoded by separate genes and transcribed as distinct transcription units as demonstrated by our sequence and Northern blot analyses. These results confirm earlier reports which provided evidence for separate genes encoding the 30/32-kDa complex proteins on the basis of Southern blot analyses and pulsed-field gel electrophoresis (6). At both the amino acid and DNA levels, the 30/32-kDa complex proteins share a high degree of homology not only with each other but also with their counterparts in a number of pathogenic and nonpathogenic mycobacterial species. For example, there are only one and five amino acid differences between the *M. tuberculosis* and the two *M. bovis* BCG 30-kDa proteins. The proteins of the 30/32-kDa complex are among the most abundantly expressed proteins of *M. tuberculosis*. The regions upstream of the structural genes encoding the proteins do not share the high degree of homology observed for the coding region, yet expression of this protein complex is regulated at the transcriptional level. This raises two important questions. (i) What are the regulatory mechanisms that control the coordinate expression of these genes in such a way that the relative amount of each protein remains constant over a 2- to 3-week growth phase? In this context, it is important to point out that the promoter analyses for the 32A protein gene (14) did not reveal any consensus sequences with regions upstream of the 30-kDa protein structural gene. (ii) Are the promoter and other regulatory sequences outlined for the genes encoding the 30/32-kDa protein complex the correct RNA polymerase contact sites, especially in light of the fact that no relationship has been established to date between promoter elements in mycobacteria and prokaryotic consensus sequences? The primer extension studies suggest that the tentatively assigned promoter sequences for the 30-kDa protein gene might be correct but also allow for additional interpretation due to the presence of more than one apparent transcription start site.

The efficiency with which the proteins of this complex are released into the culture medium is truly remarkable, potentially making these proteins ideal target molecules for the elucidation of some of the key elements of the mycobacterial secretion machinery which remains poorly understood to date. Probably even more remarkable is the fact that the 30/32-kDa proteins are released into the culture medium at a constant ratio of 3:2:1 (30-kDa protein/32A protein/32B protein). All three proteins are very difficult to degrade by site-specific proteases (unpublished observation). This observation and the fact that highly active glutamine synthetase (10), which is inactivated by a single proteolytic cleavage event, is present in the culture medium argue against the ratio among the proteins of the 30/32-kDa complex reflecting different rates of protein degradation. The importance of the release of the 30/32-kDa complex at a constant ratio remains poorly understood, especially since very little information regarding the function of the protein complex or the relationship of each protein to the others is available. A recent report showed that the 30-kDa protein has mycolic acid transferase activity but did not extend this finding to the 32A and 32B proteins (2).

As purified from *M. tuberculosis* culture filtrates, the 30-kDa protein exists as a monomer and does not exert its function as a multimer combining homologous subunits or forming a complex with the 32A and 32B protein monomers. All three proteins of the 30/32-kDa complex lack posttranslational modifications, although several secretory *M. tuberculosis* proteins have been shown to be modified either by glycosylation or lipidation (29). Both the 30-kDa and 32A proteins contain a disulfide bridge, a potential source of problems for the expression of recombinant protein due to the formation of intermolecular disulfide bridges in the oxidative environment of the *E. coli* cytoplasm. Although not secreted into the culture medium, large fractions of expressed recombinant 30-kDa and 32A proteins were correctly processed in *E. coli*, either when expressed with their own leader or the vector-encoded PelB leader.

Our immunocytochemical analyses conclusively demonstrate that the 30/32-kDa proteins are abundantly expressed by intracellular bacteria within infected human monocytes. The release of proteins of the 30/32-kDa complex from the intraphagosomal bacillus and their localization to the mycobacterial cell wall and the intraphagosomal space are consistent with involvement of the proteins in the synthesis of the mycobacterial cell wall. With regard to *M. tuberculosis* pathogenesis, this finding has implications for the use of the proteins in a subunit vaccine (12), since their release makes them available for intracellular processing and the presentation of peptides on the surface of infected host cells in association with major histocompatibility complex molecules.

ACKNOWLEDGMENTS

We are grateful to Chalermchai Chaloyphian, Barbara Jane Dillon, Sasa Maslesa-Galic, Gita Surti, and Theresa Tuazon for expert technical assistance. We thank Audree V. Fowler of the UCLA Protein Microsequencing Facility for performing N-terminal amino acid sequencing and Anita Rambo of the Department of Microbiology and Immunology Oligonucleotide Synthesis Facility for the synthesis of oligonucleotides.

This work was supported by grants AI-33138 and AI-35275 from the National Institutes of Health. During the time this work was per-

formed, B.-Y. Lee was supported by National Institutes of Health training grant AI-07126 and Jun Wang was supported by a fellowship from the Will Rogers Memorial Fund.

REFERENCES

1. **Allison, A. C., and N. E. Byars.** 1986. An adjuvant formulation that selectively elicits the formation of antibodies of protective isotypes and of cell-mediated immunity. J. Immunol. Methods **95:**157–168.
2. **Belisle, J. T., T. Sievert, K. Takayama, and G. S. Besra.** 1995. Identification of a mycolyltransferase from *Mycobacterium tuberculosis* and the coincident definition of the physiological function of antigen 85B, p. 212–216. *In* Program of the 30th Joint Conference on Tuberculosis and Leprosy. U.S.-Japan Cooperative Medical Science Program, NIAID, NIH, Ft. Collins, Colo., 19–21 July 1995. Department of Health and Human Services, Washington, D.C.
3. **Borremans, M., L. De Wit, G. Volckaert, J. Ooms, J. De Bryun, K. Huygen, J.-P. Van Vooren, M. Stelandre, R. Verhofstadt, and J. Content.** 1989. Cloning, sequence determination, and expression of a 32-kilodalton-protein gene of *Mycobacterium tuberculosis*. Infect. Immun. **57:**3123–3130.
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. **Clemens, D. L., and M. A. Horwitz.** 1995. Characterization of the *Mycobacterium tuberculosis* phagosome and evidence that phagosomal maturation is inhibited. J. Exp. Med. **181:**257–270.
6. **Content, J., A. de la Cuvellerie, L. De Wit, V. Vincent-Levy-Frébault, J. Ooms, and J. De Bruyn.** 1991. The genes coding for the antigen 85 complexes of *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG are members of a gene family: cloning, sequence determination, and genomic organization of the gene coding for antigen 85-C of *Mycobacterium tuberculosis*. Infect. Immun. **59:**3205–3212.
7. **De Mendonça-Lima, L., J. Content, H. van Heuverswyn, and W. Degrave.** 1991. Nucleotide sequence of the gene coding for the 85-B antigen of *Mycobacterium leprae*. Nucleic Acids Res. **19:**5789.
8. **De Wit, L., M. Palou, and J. Content.** 1994. Nucleotide sequence of the 85B-protein gene of *Mycobacterium bovis* BCG and *Mycobacterium tuberculosis*. DNA Sequence **4:**267–270.
9. **Grunstein, M., and D. S. Hogness.** 1975. Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. USA **72:**3961–3966.
10. **Harth, G., D. L. Clemens, and M. A. Horwitz.** 1994. Glutamine synthetase of *Mycobacterium tuberculosis*: extracellular release and characterization of its enzymatic activity. Proc. Natl. Acad. Sci. USA **91:**9342–9346.
11. **Hatfull, G. F., and G. J. Sarkis.** 1993. DNA sequence, structure and gene expression of mycobacteriophage L5: a phage system for mycobacterial genetics. Mol. Microbiol. **7:**395–405.
12. **Horwitz, M. A., B.-W. E. Lee, B. J. Dillon, and G. Harth.** 1995. Protective immunity against tuberculosis induced by vaccination with major extracellular proteins of *Mycobacterium tuberculosis*. Proc. Natl. Acad. Sci. USA **92:** 1530–1534.
13. **Kitaura, H., N. Ohara, T. Matsuo, H. Tasaka, K. Kobayashi, and T. Yamada.** 1993. Cloning, sequencing and expression of the gene for alpha antigen from *Mycobacterium intracellulare* and use of PCR for the rapid identification of *Mycobacterium intracellulare*. Biochem. Biophys. Res. Commun. **196:**1466–1473.
14. **Kremer, L., A. Baulard, J. Estaquier, J. Content, A. Capron, and C. Locht.** 1995. Analysis of the *Mycobacterium tuberculosis* 85A antigen promoter region. J. Bacteriol. **177:**642–653.
15. **Kyte, J., and R. F. Doolittle.** 1982. Simple method for displaying the hydropathy character of a protein. J. Mol. Biol. **157:**105–132.
16. **Launois, P., R. DeLeys, M'. N'Diaye Niang, A. Drowart, M. Andrien, P. Dierckx, J.-L. Cartel, J.-L. Sarthou, J.-P. Van Vooren, and K. Huygen.** 1994. T-cell-epitope mapping of the major secreted mycobacterial antigen Ag85A in tuberculosis and leprosy. Infect. Immun. **62:**3679–3687.
17. **Lee, B.-Y., and M. A. Horwitz.** 1995. Identification of macrophage and stress-induced proteins of *Mycobacterium tuberculosis*. J. Clin. Invest. **96:**245–249.
18. **Lee, T. D., and S. Vemuri.** 1990. MacProMass: a computer program to correlate mass spectral data to peptide and protein structures. Biomed. Environ. Mass. Spectrom. **19:**639–645.
19. **Matsuo, K., R. Yamaguchi, A. Yamazaki, H. Tasaka, K. Terasaka, and T. Yamada.** 1990. Cloning and expression of the gene for the cross-reactive alpha antigen of *Mycobacterium kansasii*. Infect. Immun. **58:**550–556.
20. **Matsuo, K., R. Yamaguchi, A. Yamazaki, H. Tasaka, and T. Yamada.** 1988. Cloning and expression of the *Mycobacterium bovis* BCG gene for extracellular α-antigen. J. Bacteriol. **170:**3847–3854.
21. **Ohara, N., K. Matsuo, R. Yamaguchi, A. Yamazaki, H. Tasaka, and T. Yamada.** 1993. Cloning and sequencing of the gene for alpha antigen from *Mycobacterium avium* and mapping of B-cell epitopes. Infect. Immun. **61:** 1173–1179.
22. **Pal, P. G., and M. A. Horwitz.** 1992. Immunization with extracellular protein of *Mycobacterium tuberculosis* induces cell-mediated immune response and substantial protective immunity in a guinea pig model of pulmonary tuberculosis. Infect. Immun. **60:**4781–4792.
23. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
24. **Shine, J., and L. Dalgarno.** 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71:**1342–1346.
25. **Von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. **14:**4683–4690.
26. **Weimbs, T., and W. Stoffel.** 1992. Proteolipid protein (PLP) of CNS myelin: positions of free, disulfide-bonded, and fatty acid thioester-linked cysteine residues and implications for the membrane topology of PLP. Biochemistry **31:**12289–12296.
27. **Wiker, H. G., and M. Harboe.** 1992. The antigen 85 complex: a major secretion product of *Mycobacterium tuberculosis.* Microbiol. Rev. **56:**648–661.
28. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.
29. **Young, D. B., S. H. E. Kaufmann, P. W. M. Hermans, and J. E. R. Thole.** 1992. Mycobacterial protein antigens: a compilation. Mol. Microbiol. **6:**133–145.

Editor: S. H. E. Kaufmann