Molecular Cloning, Expression, and Immunogenicity of MTB12, a Novel Low-Molecular-Weight Antigen Secreted by Mycobacterium tuberculosis

JOHN R. WEBB, † THOMAS S. VEDVICK, MARK R. ALDERSON, JEFFREY A. GUDERIAN, SHYLAN S. JEN, PAMELA J. OVENDALE, STEPHEN M. JOHNSON, STEVEN G. REED, and YASIR A. W. SKEIKY

Infectious Disease Research Institute and Corixa Corporation, Seattle, Washington 98104, and Department of Pathobiology, University of Washington, Seattle, Washington 98195

Received 4 March 1998/Returned for modification 20 April 1998/Accepted 10 June 1998

Proteins secreted into the culture medium by Mycobacterium tuberculosis are thought to play an important role in the development of protective immune responses. In this report, we describe the molecular cloning of a novel, low-molecular-weight antigen (MTB12) secreted by M. tuberculosis. Sequence analysis of the MTB12 gene indicates that the protein is initially synthesized as a 16.6-kDa precursor protein containing a 48-amino-acid hydrophobic leader sequence. The mature, fully processed form of MTB12 protein found in culture filtrates has a molecular mass of 12.5 kDa. MTB12 protein constitutes a major component of the M. tuberculosis culture supernatant and appears to be at least as abundant as several other well-characterized culture filtrate proteins, including members of the 85B complex. MTB12 is encoded by a single-copy gene which is present in both virulent and avirulent strains of the M. tuberculosis complex, the BCG strain of M. bovis, and M. leprae. Recombinant MTB12 containing an N-terminal six-histidine tag was expressed in Escherichia coli and purified by affinity chromatography. Recombinant MTB12 protein elicits in vitro proliferative responses from the peripheral blood mononuclear cells of a number of purified protein derivative-positive (PPD+) human donors but not from PPD− donors.

Tuberculosis remains one of the world's most serious health threats, with approximately 2 billion people infected worldwide and an estimated 2.9 million deaths due to tuberculosis annually (20). The recent increase in the incidence of tuberculosis, particularly antibiotic-resistant tuberculosis, underscores the need for an effective vaccine against this important disease (19). The only vaccine currently in use is the live, attenuated strain BCG and its extensively characterized, a recent analysis of CFP by two-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) revealed that CFP is comprised of up to 205 distinct proteins (32). We are currently characterizing culture supernatants of M. tuberculosis with the aim of identifying novel antigenic proteins. In this report, we describe the identification, molecular cloning, and expression of a novel, low-molecular-weight antigen (MTB12) from M. tuberculosis culture supernatants. MTB12 is a highly abundant component of M. tuberculosis culture supernatant that is readily detectable in Coomassie blue-stained gels of CFP. The MTB12 gene is present in virulent and avirulent strains of the M. tuberculosis complex, M. leprae, and M. bovis (BCG), and recombinant MTB12 protein elicits in vitro responses from the peripheral blood mononuclear cells (PBMC) of PPD+ human donors.

MATERIALS AND METHODS

Strains. M. tuberculosis H37Rv, H37Ra, and Erdman were provided by Sean Skerritt (Seattle VA Hospital). M. tuberculosis strain C is a clinical isolate provided by Lee Riley (University of California, Berkeley). Pelleted samples of M. bovis BCG and M. leprae were kindly provided by Paul Tan (Genesis Corp.). Mycobacterial genomic DNA was prepared as previously described (18). Genomic DNA from M. tuberculosis H37Ra and H37Rv was fragmented for library generation by using either partial digestion with SmaI (H37Rv library) or sonication (H37Ra library). In both cases, DNA fragments in a size range of 300 to 4,000 bp were blunt-ended with Klenow polymerase, ligated to EcoRI adapters, and subcloned into EcoRI-predigested λZAP bacteriophage arms as specified by...
the manufacturer (Stratagene). Phage were packaged by using Gigapack II packaging extracts (Stratagene) as recommended by the manufacturer.

**HPLC purification of CFP.** Concentrated CFP of *M. tuberculosis* Erdman was provided by John Belisle (Colorado State University) and purified by a two-step procedure. CFP was fractionated by preparative liquid chromatography (HPLC) on a 4.6- by 25-cm Aquapore C8 column (Brownlee) at a flow rate of 1 mL/min with a 0 to 60% acetonitrile gradient in 30 min. One of the major peaks resolved by this method was shown by protein sequence analysis to be a mixture of proteins and was therefore subjected to further purification using microbore HPLC. The sample was resolved on a 1.1- by 100-mm Aquapore C8 column (Brownlee) at a flow rate of 80 μL/min with a 20 to 70% acetonitrile gradient in 70 min. Peak fractions from the microbore HPLC were loaded onto biotin – streptavidin fiber filters (Perkin-Elmer/Applied Biosystems). The purified recombinant proteins were then placed in a Precise 494 protein sequencer (Perkin-Elmer/Applied Biosystems) and sequenced from the amino terminus, using traditional Edman chemistry.

**Cloning of the *M. tuberculosis* MTB12 gene.** The *M. leprae* homolog of the MTB12 gene was amplified from *M. leprae* genomic DNA by PCR. PCR primers (5′-ATGAAATGCACCATACTATGACA-3′ and 5′-TCAAGCCTCCGGCGCTGCAACACGATC-3′) were designed based on sequence obtained from GenBank accession no. U00016_13. The PCR program consisted of 30 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 1 min. A single amplification product of the expected size (495 bp) was subcloned into the pCR vector (Invitrogen), and the insert identity was confirmed by DNA sequence analysis. The cloned amplification product was then digested with XbaI and agarose gel purified and was labeled to high specific activity (~10⁶ cpm/μg) with [α-³²P]dCTP, using the random primer method (14). This probe was used to screen an *M. tuberculosis* H37Rv genomic library prepared in the ZAPφ vector (Stratagene). Approximately 1 x 10⁶ PFU were screened by plaque hybridization. Filters were probed to a final stringency of 0.2× SSC (1× SSC = 0.15 M NaCl plus 0.015 M sodium citrate) at 65°C. Hybridizing plaques were purified to homogeneity by two subsequent rounds of low-density plaque screening, and Bluescript phagemids were excised from positive clones as specified by the manufacturer (Stratagene). Sequence analysis revealed that one of the clones contained the complete *M. tuberculosis* MTB12 open reading frame (ORF) plus 1.2 kb of 5′ untranslated sequence and 2 kb of 3′ untranslated sequence.

For serological screening, a polyclonal rabbit antiserum was raised against the concentrated culture filtrates of *M. tuberculosis* MTB12 and mature protein (lacking signal sequence) were prepared by using the pET17b expression system (Novagen). MTB12 DNA was amplified by PCR using the primers 5′-AATTACATATGACCGGTAGTTTGAACCTACCATCACAC-3′ and 5′-CATGGAATTCTCAGTTCCCTGCGGCCTGCAGCAA-3′ and was labeled to high specific activity ([¹⁴C]dCTP, Amersham) at a flow rate of 80 μL/min with a 20 to 70% acetonitrile gradient in 70 min. Peak fractions from the microbore HPLC were loaded onto biotin – streptavidin fiber filters (Perkin-Elmer/Applied Biosystems). The loaded filters were treated with CFP.

**HPLC purification of CFP.** CFP was purified by preparative liquid chromatography (HPLC) on a 4.6- by 25-cm Aquapore C8 column (Brownlee) at a flow rate of 80 μL/min with a 20 to 70% acetonitrile gradient in 70 min. Peak fractions from the microbore HPLC were loaded onto biotin – streptavidin fiber filters (Perkin-Elmer/Applied Biosystems). For amino acid sequence determination, CFP was separated by SDS-PAGE on a 12% gel and transferred to a polyvinylidene difluoride (PVDF) membrane. Membrane strips were stained with Coomassie blue to detect protein bands. Specific bands were excised from the stained strip and subjected to N-terminal sequence analysis using a Procise 494 protein sequencer (Perkin-Elmer/Applied Biosystems). A second and a third strip were completely destained in methanol and were analyzed by Western blotting using an anti-MTB12 polyclonal rabbit antiserum or corresponding preimmune serum as described above.

**Immunological reactivity of MTB12.** Proteins isolated from CFP by microbore HPLC and recombinant MTB12 proteins were assayed for the ability to elicit in vitro proliferative responses from whole PBMC of healthy PPD− and PPD+ donors. PBMC were obtained from heparinized blood by ficoll gradient centrifugation or by leukapheresis. PBMC (2 x 10⁶) were incubated in 96-well round-bottom plates (Costar) in medium only (RPMI 1640 with 10% pooled human serum) or in medium containing specific antigens at the indicated concentrations. Plates were cultured for 5 days at 37°C in 5% CO₂ and were pulsed with 1 μCi of [³¹P]orthophosphate (Amersham) for the final 18 h. Cells were harvested onto filter mats and counted in a Matrix 9600 direct beta counter (Packard).

**Nucleotide sequence accession number.** The nucleotide and the deduced amino acid sequences of MTB12 clones have been entered in the GenBank database under accession no. AF602606.

**RESULTS**

**Purification of MTB12 from *M. tuberculosis* CFP.** The CFP of *M. tuberculosis* has been repeatedly shown to be rich in immunologically reactive protein antigens (25, 35). To identify novel antigenic components of this material, we fractionated CFP by C₈ reverse-phase HPLC followed by microbore HPLC and characterized protein peaks eluting from the second fractionation step by N-terminal amino acid acid sequencing. Peak fractions were also assayed for the ability to elicit in vitro proliferative responses from the PBMC of PPD− human donors. We focused our efforts on a single peak (Fig. 1, fraction 7) that eluted just prior to the 45/47-kDa secreted antigen MPT32 (21). Fraction 7 contained a single protein with an N-terminal sequence (DPASAPDVPPTAQLTSLNSLADPVNSF) that was not present in any previously reported secreted protein of *M. tuberculosis*. In addition, this fraction elicited in vitro proliferative responses in the PBMC from two of two PPD− donors but not from PPD+ donors (Table 1).

**Molecular cloning of the gene encoding MTB12.** A search of the gene data banks with the N-terminal sequence obtained from fraction 7 indicated that this sequence was derived from a novel protein of *M. tuberculosis*. However, the search did reveal homology with the product of a putative ORF from the *M. leprae* genome sequencing project (accession no. U00016_13). The predicted amino acid sequence of the *M. leprae* ORF contained a region that was identical to the fraction 7 N-terminal sequence at 21 of 29 residues. To facilitate the isolation of the *M. tuberculosis* fraction 7 gene, the *M. leprae* homolog was amplified from *M. leprae* genomic DNA by PCR and was used as a probe to screen an *M. tuberculosis* H37Rv genomic library by plaque hybridization. Five positive plaque clones were obtained. Sequence analysis of specific antigens captured from these clones contained a 507-bp ORF encoding a 168-amino acid protein with a predicted molecular mass of 16.6 kDa (Fig. 2A). Residues 94 to 78 of the predicted protein sequence had 100% identity with the N-terminal sequence obtained from the microbore HPLC fraction 7 residues. These sequences were preceded by...
a stretch of 48 highly hydrophobic amino acids that are presumed to constitute an N-terminal signal sequence, directing the transport of the protein to the extracellular space. Thus, the N-terminal sequence of the protein identified as HPLC fraction 7 likely represents the amino terminus of the mature fully processed protein found in the CFP. The mature, processed protein had a predicted molecular mass of 12.5 kDa and a pI of 5.03. Based on the size of the mature protein, we have designated this protein MTB12. Similarity between the M. tuberculosis MTB12 protein and the M. leprae homolog was significant: 107 of 168 (63.7%) amino acid identity and 22 conservative substitutions (Fig. 2B). The most significant difference between the two species was a four-residue deletion within the putative signal sequence region of the M. leprae homolog. Other conservative and nonconservative substitutions occurred throughout the protein and were not restricted to any particular region. Further database searching did not reveal the presence of any structural domains that might provide evidence of the biological function of MTB12.

The MTB12 gene was also isolated by using a serological screening approach that was being performed concurrently to the biochemical characterization of M. tuberculosis CFP. In this second approach, a polyclonal rabbit antiserum generated against total M. tuberculosis CFP was used to screen an M. tuberculosis H37Ra genomic expression library. Of the 27 clones thus isolated, two contained inserts corresponding to MTB12. One of these clones (designated Ra-1) contained the complete MTB12 ORF (including the signal sequence) plus 63 bp of 5' untranslated sequence in frame with the vector-encoded β-galactosidase. This clone expressed a 21-kDa fusion protein that.

### TABLE 1. In vitro proliferative responses of PBMC obtained from PPD⁺ or PPD⁻ donors in the presence of M. tuberculosis CFP fraction 7

<table>
<thead>
<tr>
<th>Fraction 7 (µg/ml)</th>
<th>Mean[^H]thymidine incorporation (cpm) of triplicate wells</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PPD⁺</td>
</tr>
<tr>
<td></td>
<td>D7</td>
</tr>
<tr>
<td>5</td>
<td>2260 (11.3)</td>
</tr>
<tr>
<td>1</td>
<td>834 (4.2)</td>
</tr>
<tr>
<td>0.2</td>
<td>464 (2.3)</td>
</tr>
<tr>
<td>0</td>
<td>200</td>
</tr>
</tbody>
</table>

[^H]: Stimulation index (mean cpm incorporated in the presence of antigen divided by mean cpm incorporated in the absence of antigen).
exhibited strong immunoreactivity with the anti-CFP antiserum (data not shown). A number of other clones corresponding to previously characterized secreted proteins, including 85B (9), 45/47-kDa secreted antigen (21), and MPT64 (24), were also isolated by using this antiserum.

The organization of the MTB12 gene in various members of the *M. tuberculosis* complex was analyzed by Southern blotting. A probe comprising the insert of clone Ra-1 (containing the complete MTB12 coding region plus a small amount of 5′ untranslated sequence) was hybridized to a blot containing *Pst* I-digested genomic DNAs from *M. tuberculosis* H37Ra, H37Rv, and Erdman; a recent *M. tuberculosis* clinical isolate referred to as strain C; and the *M. bovis* BCG. Hybridizing bands of approximately 1.0 and 1.6 kbp were observed in all strains (data not shown). Sequence analysis of the 3.6-kbp genomic clone from which the *M. tuberculosis* MTB12 gene was isolated revealed the presence of three *Pst*I restriction sites in the vicinity of the MTB12 ORF. The first site was located near the 3′ end of the MTB12 ORF, and the remaining two sites were located 988 bp upstream and 1,579 bp downstream of the first *Pst*I site. Thus, the sizes of the predicted *Pst*I restriction fragments were in complete agreement with the hybridization pattern observed during Southern blot analysis of genomic DNAs. Southern blot hybridizations to genomic DNAs from a panel of other mycobacterial species (using the same stringency conditions) were negative (data not shown), suggesting that these species either lack an MTB12 homolog or contain a homolog with sufficient sequence divergence to avoid detection at the stringencies used (0.1× SSC at 65°C).

**Recognition of rMTB12 protein by human PBMC.** rMTB12 protein containing an N-terminal six-histidine affinity tag was expressed in the pET17b *E. coli* high-level expression system (Novagen). Two recombinant proteins were prepared, one corresponding to the mature protein (lacking a signal sequence) and the other corresponding to the full-length MTB12 protein expressed by clone Ra-1 (containing the leader sequence plus an additional 21 residues derived from the 5′ untranslated region). Recombinant proteins were purified to homogeneity by Ni-NTA affinity chromatography and were shown by SDS-PAGE to be free of contaminating *E. coli* protein (Fig. 3). Recombinant MTB12 protein was assayed for the ability to elicit in vitro proliferative responses from PPD+ and PPD− human donor PBMC. Although proliferative responses to recombinant protein were weaker than the responses elicited by total CFP, recombinant MTB12 consistently elicited in vitro proliferation from a number of PPD− donors (Table 2). Of the 10 PPD+ donors tested, 4 responded to recombinant MTB12 protein with a stimulation index of between 4 and 12. Recombinant MTB12 protein did not elicit substantial proliferation from any of the six PPD+ individuals tested. Viability of all PBMC was confirmed by proliferation in the presence of CFP (PPD+ individuals), tetanus toxoid (PPD− individuals), and phytohemagglutinin (data not shown). Interestingly, full-length recombinant MTB12 protein elicited a slightly stronger re-
FIG. 4. Immunoblot analyses of MTB12. Western blots containing M. tuberculosis lysate (2.5 μg), CFP (2.5 μg), recombinant mature MTB12 (50 ng), full-length MTB12 (Fl-MTB12; 50 ng), and 85B (50 ng) were incubated with anti-CFP (A), anti-MTB12 (B), or anti-85B (C) polyclonal rabbit antisera. Immunoreactive proteins were detected by using [125I]protein A followed by autoradiography. Positions of size markers are indicated in kilodaltons.

FIG. 5. Relative abundance of MTB12 in M. tuberculosis culture filtrates. M. tuberculosis CFP was separated by SDS-PAGE, transferred to a PVDF membrane by Western blotting, and stained with Coomassie blue. A triplet of proteins with the same approximate molecular weight as mature MTB12 was readily visible after staining (Fig. 5, lane 1). The middle band of the triplet was determined to be MTB12 according to two different criteria. First, this band (on a duplicate blot) was reactive with the polyclonal anti-MTB12 antiserum by Western blotting (Fig. 5, lane 2) but not with the corresponding preimmune serum (Fig. 5, lane 3). Second, the identity of a number of predominant CFP proteins was determined by N-terminal sequencing of bands excised from the Coomassie blue-stained PVDF membrane. The middle band of the triplet contained the N-terminal sequence of mature MTB12, whereas the lower band of the triplet corresponded to GroES. The uppermost band of the triplet could not be identified, as it was blocked at the N terminus. The locations of three other abundant CFP proteins (85B complex, MPT64, and GroEL) that were identified by N-terminal sequencing are indicated at the right in Fig. 5. Together, these data suggest that MTB12 is a highly abundant component of M. tuberculosis CFP and that MTB12 is at least as abundant as several other well-characterized CFP proteins.

DISCUSSION

In this study, a novel protein present in M. tuberculosis CFP was identified, cloned, and characterized in terms of abundance and immunological reactivity. The crude CFP from M. tuberculosis has been extensively characterized as a rich source of antigens that elicit protective responses in various models of tuberculosis (2, 26). Furthermore, it has been speculated that active synthesis and secretion of CFP components are responsible for the greater efficacy of vaccination using live attenuated mycobacteria than of vaccination using killed organisms. Consequently, the CFP is currently being characterized by a number of labs in an effort to define specific proteins that may be useful as subunit vaccine reagents. The MTB12 protein
reported herein is a highly abundant component of \textit{M. tuberculosis} CFP that appears to be actively secreted. The presence of a consensus signal sequence at the N terminus and the removal of this sequence from the mature, extracellular form of the protein confirm that MTB12 is indeed targeted to the extracellular space by \textit{M. tuberculosis}. However, it is not clear from this analysis whether the MTB12 protein is released directly into the culture medium or whether it represents a surface protein that is fortuitously shed into the culture medium by membrane turnover. The lack of a recognizable lipid attachment site and the inability to detect MTB12 protein in \textit{M. tuberculosis} lysate preparations suggest that the membrane turnover hypothesis is less likely. Also, the MTB12 protein contains two consensus sites for N-linked glycosylation. Although there is limited evidence to suggest that glycosylation does occur in mycobacteria (11–13, 15), only O-linked glycosylation has been reported to date.

It is intriguing that MTB12 has not been previously identified as a component of CFP considering the apparent abundance of the protein. Various preparations of CFP are known to vary in composition dependent on the culture media used (3, 10, 23, 32) and the growth phase of the culture (i.e., early log versus late log phase) (3). Therefore, it is possible that secretion of MTB12 is dependent on one or both of these parameters. The CFP analyzed in this study corresponded to filter proteins of a late-log-phase culture grown in GAS medium (32). Also, mature MTB12 is found in close proximity to the GroES protein after one-dimensional SDS-PAGE and has a pI (5.03) which is similar to that determined for GroES (32). It is therefore plausible that the MTB12 protein has been previously overlooked due to its proximity to the well-characterized GroES protein.

During preparation of this report, a search of a recently released database containing portions of the \textit{M. tuberculosis} genome revealed the presence of a cosmid (MTCy27) that harbored the complete MTB12 ORF (sequence accession no. MTCy27.04). Analysis of the cosmid sequence in the region of the MTB12 ORF confirmed that digestion with \textit{PstI} would produce the 1- and 1.6-kbp hybridizing bands that we observed by Southern blotting. Although the complete \textit{M. tuberculosis} genome is not yet available, this result also suggests that MTB12 is likely to be encoded by a single-copy gene. Furthermore, our Southern blot results indicate that these same two \textit{PstI} sites are conserved in other strains of \textit{M. tuberculosis} (11–13, 15), only O-linked glycosylation has been reported to date.

Comparison of the antigenicity of recombinant mature MTB12 with that of the equivalent full-length protein indicated that for some human donors the full-length MTB12 protein elicited a slightly stronger in vitro response than did mature protein (data not shown). This finding suggests that the hydrophobic leader sequence, which is normally cleaved during processing, may contribute to the antigenicity of the full-length molecule. We have also observed a similar phenomenon in mouse immunization experiments (unpublished data). Although leader sequence peptides derived from endogenously synthesized proteins are known to be bound and presented by classical (16, 34) and nonclassical (1, 5) class I MHC molecules, presentation of exogenous leader sequence peptides by class I or class II MHC molecules has not been well described. Thus, it would be interesting to determine the relative contributions of CD4+ and CD8+ T cells during the augmented proliferative response observed in the presence of full-length MTB12 protein.

The abundance of MTB12 in culture supernatants together with the immunological data presented herein suggests that MTB12 may have potential value as a subunit vaccine component to protect against infection by \textit{M. tuberculosis}. We are currently assessing the protective capability of MTB12 vaccination in murine models of tuberculosis, using both recombinant protein in conjunction with specific adjuvants and DNA vaccine approaches.

ACKNOWLEDGMENTS

We thank John Belisle for providing \textit{M. tuberculosis} CFP and Teresa Bement for performing human PBMC proliferation assays. S.M.J. was supported in part by NIH training grant 5T32HD07223. J.R.W. and Y.A.W.S. were supported in part by fellowships from the Medical Research Council of Canada.

REFERENCES


Editor: S. H. E. Kaufmann