Genome-Based In Silico Identification of New Mycobacterium tuberculosis Antigens Activating Polyfunctional CD8⁺ T Cells in Human Tuberculosis


*J Immunol* 2011; 186:1068-1080; Prepublished online 17 December 2010; doi: 10.4049/jimmunol.1002212
http://www.jimmunol.org/content/186/2/1068

---

**Supplementary Material**
http://www.jimmunol.org/content/suppl/2010/12/17/jimmunol.1002212.DC1

**References**
This article cites 60 articles, 28 of which you can access for free at:
http://www.jimmunol.org/content/186/2/1068.full#ref-list-1

**Subscription**
Information about subscribing to *The Journal of Immunology* is online at:
http://jimmunol.org/subscription

**Permissions**
Submit copyright permission requests at:
http://www.aai.org/About/Publications/JI/copyright.html

**Email Alerts**
Receive free email-alerts when new articles cite this article. Sign up at:
http://jimmunol.org/alerts

---

*The Journal of Immunology* is published twice each month by
The American Association of Immunologists, Inc.,
1451 Rockville Pike, Suite 650, Rockville, MD 20852
All rights reserved.
Print ISSN: 0022-1767 Online ISSN: 1550-6606.
Genome-Based In Silico Identification of New Mycobacterium tuberculosis Antigens Activating Polyfunctional CD8+ T Cells in Human Tuberculosis


Although CD8+ T cells help control Mycobacterium tuberculosis infection, their M. tuberculosis Ag repertoire, in vivo frequency, and functionality in human tuberculosis (TB) remains largely undefined. We have performed genome-based bioinformatics searches to identify new M. tuberculosis epitopes presented by major HLA class I supertypes A2, A3, and B7 (covering 80% of the human population). A total of 432 M. tuberculosis peptides predicted to bind to HLA-A*0201, HLA-A*0301, and HLA-B*0702 (representing the above supertypes) were synthesized and HLA-binding affinities determined. Peptide-specific CD8+ T cell proliferation assays (CFSE dilution) in 41 M. tuberculosis-responsive donors identified 70 new M. tuberculosis epitopes. Using HLA/peptide tetramers for the 18 most prominently recognized HLA-A*0201-binding M. tuberculosis peptides, recognition by cured TB patients’ CD8+ T cells was validated for all 18 epitopes. Intracellular cytokine staining for IFN-γ, IL-2, and TNF-α revealed mono-, dual-, as well as triple-positive CD8+ T cells, indicating these M. tuberculosis peptide-specific CD8+ T cells were (poly)functional. Moreover, these T cells were primed during natural infection, because they were absent from M. tuberculosis-noninfected individuals. Control CMV peptide/HLA-A*0201 tetramers stained CD8+ T cells in M. tuberculosis-infected and noninfected individuals equally, whereas Ebola peptide/HLA-A*0201 tetramers were negative. In conclusion, the M. tuberculosis-epitope/Ag repertoire for human CD8+ T cells is much broader than hitherto suspected, and the newly identified M. tuberculosis Ags are recognized by (poly)functional CD8+ T cells during control of infection. These results impact on TB-vaccine design and biomarker identification. The Journal of Immunology, 2011, 186: 1068–1080.

One third of the world’s population is latently infected with Mycobacterium tuberculosis, and each year, 1.8 million people die of tuberculosis (TB) (1, 2). Bacillus Calmette-Guérin (BCG) vaccination has been used for almost a century, but its protective efficacy is highly variable and incomplete, especially against pulmonary TB, the main and contagious form of the disease. BCG’s effect wanes in adolescents, and revaccination does not confer additional protection. Currently, the TB problem is aggravated by HIV coinfection, and the emergence of multidrug-resistant and extensively drug-resistant M. tuberculosis strains, which threaten to set back TB control to the preantibiotic era (3, 4). New and better vaccines against TB are urgently needed, but to be able to develop these, a radically improved understanding of what constitutes protective host immunity is warranted. It is well established that CD4+ T cell responses play a major role in acquired immunity against M. tuberculosis (5–7). There is, however, mounting evidence that CD8+ T cells are also involved in the control of M. tuberculosis infection and in mediating optimal host defense in small animal models, nonhuman primates, and human infection (8–14).

The text represents the authors’ views and does not necessarily represent a position of the European Commission, which will not be liable for the use made of such information.

Address correspondence and reprint requests to Dr. Tom H. M. Ottenhoff, Prof. in Immunology, Head Group Immunology and Immunogenetics of Bacterial Infectious Diseases, Department of Infectious Diseases, Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, The Netherlands. E-mail address: T.H.M.Ottenhoff@lumc.nl

The online version of this article contains supplemental material.

Abbreviations used in this article: BCG, bacillus Calmette-Guérin; Bepl, proteins with B cell epitopes; BesPred, Best predicted epitopes; Cons, conserved candidate Ag; DOS/LAG, Mycobacterium tuberculosis DosR regulon encoded latency Ag; HC, healthy control; ICS, intracellular cytokine staining; IEDB, Immune Epitope Database; NCBI, National Center for Biotechnology Information; ORF, open reading frame; PPD, purified protein derivative; PredSecret, hypothetical proteins predicted to be secreted; Secret, secreted M. tuberculosis Ag; TB, tuberculosis; TB-CD8, M. tuberculosis CD8+ T cell epitope expressing Ag; TB-VAC, tuberculosis vaccine Ag.
Following recognition of mycobacterial Ags on infected cells, CD8+ T cells contribute to *M. tuberculosis* control through: 1) IFN-γ and TNF-α production (9, 15–17); 2) lysis of infected host cells (16–18); and 3) direct killing of mycobacteria (19–21). One study demonstrated clonal CD4+ and CD8+ T cell expansion in granulomas from subjects with latent TB infection (22), and similar changes in the TCR repertoire were reported in peripheral blood versus pleural fluid in TB patients (23). Furthermore, CD8+ T cells specific for a number of mycobacterial Ags have been isolated from human and mouse models, consistent with the hypothesis that CD8+ T cells are constantly being stimulated with Ag (18, 20). One of us reported that the frequency of *M. tuberculosis* Ag85A-specific CD8+ T cells correlated with therapy-induced curative responses in children: Ag85A epitope-specific CD8+ T cells during active TB produced low levels of IFN-γ and perforin, which normalized after therapy (24). In a later study, we reported similar findings for CD8+ T cells directed against six *M. tuberculosis* epitopes (two of which were newly identified). In that study, it was also found that *M. tuberculosis* peptide-specific IL-2/IFN-γ+ CD8+ T cell responses were associated with natural protection against developing TB disease (15). In parallel studies, Kaufmann and colleagues (25) found clonal expansion of effector memory CD8+ T cells in older children with TB, with potential impact on the course and severity of disease. Lewinsohn and colleagues (26) reported clonally expanded CD8+ T cells that recognized a number of *M. tuberculosis* epitopes in the context of HLA-A and -B alleles, and we ourselves reported Ag85A, B, and C epitopes activating human CD8+ T cells (13, 14, 27, 28). Despite these studies, little remains known about the size, quality, and specificity of *M. tuberculosis*-specific CD8+ MHC-Ia-restricted T cell responses in TB patients (20) and their relevance to control of infection (i.e., prevention of progression to TB disease).

The complete genome sequence of *M. tuberculosis* encodes 3985 coding open reading frames (ORFs) (29). Surprisingly, *M. tuberculosis*-specific epitopes have been identified in only 270 ORFs, representing only 7% of the *M. tuberculosis* genome. In fact, 30 ORFs account for 65% of all epitopes reported (30). This indicates that our current knowledge of the *M. tuberculosis* antigenome/epitome is limited and incomplete and that many more relevant *M. tuberculosis* Ags and epitopes likely remain to be discovered, particularly for CD8+ T cells. The search for Ags that stimulate human CD8+ T cells requires new search strategies, because classical protein expression technologies are not easily amenable to unravel CD8+ T cell responses. Recently developed new and powerful bioinformatics prediction tools may help to identify candidate epitopes and thus minimize the laborious screening of peptides for immunobiological characteristics of Ags capable of eliciting a protective immune response.

Using novel bioinformatics search tools in combination with functional immunological screening strategies, we have selected new *M. tuberculosis* proteins, which were likely to contain CD8+ T cell-stimulating HLA class Ia-restricted epitopes (reverse Ag discovery). We also used unbiased forward Ag discovery, in which expression libraries representing the whole *M. tuberculosis* genome were screened for proteins that are targets for B cell responses in TB patients. Within the selected proteins, potential peptide epitopes were identified that were restricted by any three of the major HLA class Ia supertypes (A2, A3, and B7), which cover >80% of individuals from different ethnic groups (31). These predicted epitopes were validated using specific tetramers and peptide stimulation combined with intracellular cytokine staining (ICS) to quantify polyfunctional CD8+ T cell responses in cured TB patients and immune donors.

**Materials and Methods**

**Study subjects**

Buffy coats from 41 in vitro purified protein derivative (PPD)-responsive (average IFN-γ responses to PPD: 1828 pg/ml, range 101 to >5000 pg/ml) HLA-A02*01+ healthy donors (Sanquin, Leiden, The Netherlands) were used for the study. Their HLA class I and their responses to PPD and ESAT6/CFP10 are shown in the Supplemental Table I. No clinical information is available for these donors other than that they were healthy and had no chronic viral infections or other contraindications for donating blood. BCG in The Netherlands is only administered to people at risk for TB exposure, and the TB incidence in The Netherlands is extremely low, such that the vast majority of our donors (>95%) is highly unlikely to have been vaccinated with BCG or to have had exposure to *M. tuberculosis*. All individuals gave written consent before blood donation. The study was approved by the Institutional Review Board of the Leiden University Medical Center.

In the second phase of the study, peripheral blood was obtained from 10 HLA-A02*01+–positive adults with recently cured TB disease (six males, four females; age range 46–67 y) from the Dipartimento di Medicina Clinica e delle Patologie Emergenti, University Hospital, Palermo, Italy, and 10 tuberculin (PPD)-negative healthy subjects (seven males and three females, age range, 32–51 y). The cured TB patients had had clinical and radiological findings consistent with active pulmonary TB. Diagnosis had been confirmed by bacteriological isolation of *M. tuberculosis* and by clinical and radiological features. All patients had been treated in accordance with Italian guidelines and had received multidrug therapy for 6 mo. TB treatment was successful in all participants as evidenced by the absence of any clinical or radiographic evidence of active disease, the confirmed completion of anti-TB chemotherapy, and sterile mycobacterial cultures. Blood samples were taken at time points well after cessation of treatment. None of the cured TB patients had been vaccinated during infancy with BCG, had evidence of HIV infection, or was being treated with steroid or other immunosuppressive or anti-TB drugs at the time of their sampling. Tuberculin (PPD) skin tests were considered positive when the induration diameter was larger than 5 mm at 72 h since injection of 1 U PPD (Statens Serum Institute, Copenhagen, Denmark). All Italian subjects were HLA typed serologically, whereas the HLA-type A02*01 was confirmed to be A*02*01 using sequence-specific oligonucleotide primers.

All individuals gave written consent before blood donation. The study was approved by the Ethical Committee of the Dipartimento di Medicina Clinica e delle Patologie Emergenti, University Hospital, Palermo, Italy. The 10 healthy controls had no symptoms or signs of active TB nor had they been infected with *M. tuberculosis*. All individuals gave written consent before blood donation. The study was approved by the Ethical Committee of the Dipartimento di Medicina Clinica e delle Patologie Emergenti, University Hospital, Palermo, Italy. The 10 healthy controls had no symptoms or signs of active TB nor had they been vaccinated with BCG. They were tuberculin skin test negative as well as in vitro QuantiFERON-TB Gold test (QFT-G, Cellestis, Victoria, Australia) negative.

**Bioinformatics strategy to identify CD8+ T cell epitopes**

Complete sequenced genome of *M. tuberculosis* strain H37Rv (National Center for Biotechnology Information [NCBI] Refseq Id: NC_000962) were retrieved from GenBank (http://www.ncbi.nlm.nih.gov/GenBank/). All 3918 candidate protein-encoding genomic sequences were submitted to the SubCell 1.0 (http://www.cbs.dtu.dk/services/SubCell/) and prediction performed using the option “gram positive bacteria” in the server. The SubCell server generated a list of 370 secreted proteins (the prediction program, SignalP, predicted 340 proteins with a signal peptide [the classical secretion pathway]; SecretomeP predicted 10 proteins for the nonclassical secretion pathway; and lastly, TatP predicted 20 proteins to have a twin-Arg signal peptide cleavage site characteristic for bacteria). NetCTL 1.0 server (http://www.cbs.dtu.dk/services/NetCTL-1.0/) was used for predicting 9-mer CD8+ T cell epitopes in *M. tuberculosis* protein sequences. Briefly, this method integrates prediction of MHC binding, proteasomal C-terminal cleavage and TAP transport efficiency (32). Prediction was performed for three main HLA supertypes: HLA-A2, -A3, and -B7. A threshold cutoff value of 0.42 in NetMHC corresponding to a predicted binding affinity of <500 nM was used.

**Epitope selection criteria**

Epitopes were selected using eight different selection criteria from proteins that were selected based on suspected immunogenicity. In bold below, the protein selections are shown; the numbers in parentheses represent the number of potential epitopes selected for further study within each protein group. A total of 498 peptides were selected initially but due to difficulties in peptide synthesis or purification, 66 peptides were either cancelled or discarded, leaving 432 peptides. A full list of proteins, which have been used...
for the peptide prediction for each selection, is provided in Supplemental Table II.

**TB-vaccine Ags (n = 27).** This group contains candidate epitopes from proteins being evaluated in TB-vaccine trials. These proteins were selected because they have been relatively well studied and are known to be immunogenic. The following proteins were used for the prediction: M. tuberculosis 7272 (Rv0125—NP_214639.1 [protein accession number according to NCBI]) and Rv1196—[YP_177951.1], TB10.4 (Rv0288—NP_214802.1), HBBHA (Rv0475—NP_214989.1), Ag85B (Rv1186c—NP_216402.1), HspX (Rv0251c—NP_216547.1), Ag85A (Rv3804c—NP_218321.1), and ESAT-6 (Rv3875—[YP_178023.1]).

M. tuberculosis CD8+ T cell epitope-expressing Ags (n = 60). This group contains candidate epitopes from proteins with already known CD8 T cell epitopes (33). This selection is included because proteins containing immunogenic CD8 T cell epitopes have been reported to be enriched in other CD8 T cell epitopes in HIV (33).

**Best predicted M. tuberculosis epitopes (n = 59).** This group contains candidate epitopes with the highest combined NetCTL score based on peptide-MHC binding, protosomal C-terminal cleavage, and TAP transport efficiency.

**Conserved candidate Ags (n = 65).** This group contains candidate epitopes that are conserved among related or closely related organisms. One rationale to include this selection is that heterologous immunity may exist to cross-reactive epitopes in other organisms or strains of the same organism (reviewed in Ref. 34). Furthermore, there is a general belief that vaccines may be more effective if they focus on Ags that are under functional or structural constraints, as variation in these regions may affect the fitness of the pathogen. The best predicted peptides are 100% conserved in the following bacterial strains (NCBI RefSeq ID in parentheses): M. tuberculosis CDC1551 (NC_002755), M. bovis AF212297 (NC_002945), M. leprae TN (NC_002677), M. avium subsp. paratuberculosis str. k10 (NC_002944), Co-nystebacterium glutamicum ATCC 13032 (NC_003008), M. efficiens YS-314 (NC_014889), C. glutamicum ATCC 13032 (NC_003040), C. diptheriae NCTC 13129 (NC_002935), Streptomyces coelicolor A3(2) (NC_003888), S. avermitilis MA-4680 (NC_003155), and Nocardia farcinica IFM 10152 (NC_006361).

M. tuberculosis DosR regulon-enabled latency Ags (n = 63). This group contains the best-predicted candidate epitopes from a list of recently discovered immunogenic proteins from the M. tuberculosis DosR regulon (28, 35). This selection was included because Ags of this dormancy regulon may contribute to the control of latent M. tuberculosis infection (28).

**Proteins with B cell epitopes (n = 52).** This group contains candidate epitopes predicted from novel proteins with B cell epitopes discovered by generation of shotgun expression library (U. Sahin and A. Zaigler, unpublished observations). Sera from TB patient blood samples were used for immunoscreening. This particular selection was used because proteins containing B cell epitopes are likely also to host T cell epitopes (e.g., see Ref. 36).

**Secreted M. tuberculosis Ags (n = 59).** This group contains candidate epitopes from potentially secreted M. tuberculosis Ags, because these are considered to be immunodominant and involved in protective immunity. Secreted proteins are readily available for Ag processing and presentation by MHC class I molecules on the infected host cell (37). Selection of peptides was based on a list of known secreted proteins obtained from the TuberculList database (http://www.genolist.pasteur.fr/TuberculList/).

**Hypothetical proteins predicted to be secreted (n = 47).** This group contains candidate epitopes from unknown hypothetical proteins predicted to be secreted by both including proteins using the classical (with an NHE terminal signal peptide (38)) and nonclassical secretion pathway (39). A prediction server, SubCell version 1.0, was used to predict various types of signal peptides and subcellular location in Gram-negative and Gram-positive bacteria (http://www.cbs.dtu.dk/services/SubCell; see above). Secreted proteins were then selected for M. tuberculosis.

**Peptides and biochemical peptide-HLA class I binding assay**

The 9-mer peptides were synthesized by standard 9-fluorenylmethyloxycarbonyl chemistry (Minotope; Schafer-N, Copenhagen, Denmark). Peptide sequences were synthesized by 90 µg/vial and stored lyophilized at −20°C as a lyophilized stock (for more details, see Ref. 32). Peptides were dissolved just before use. The biochemical peptide-HLA class I binding assay was performed as previously described (40).

**Isolation of PBMCs**

PBMC were isolated from buffy coats by density gradient centrifugation using Ficoll (Pharmacy, Leiden University Medical Center) and Leucosep (Aldrich, Zwijndrecht, The Netherlands). Cells were thawed, washed, and labeled with CFSE (Molecular Probes, Leiden, The Netherlands) at a final concentration of 5 µM for 10 min at 37°C. Washed, counted, and viable cells were seeded in six replicates in 96-well round-bottom plates at a concentration of 2 × 10^6 in the presence of control Ags (PPD, 5 µg/ml [Statens Serum Institute, Copenhagen, Denmark]), PHA 2 µg/ml (Remel; Oxoid, Haarlem, The Netherlands), or the test peptides (final concentration 10 µg/ml). Culture medium used was IMDM with glutamax supplemented with 100 U/ml penicillin, 100 µg/ml streptomycin (Invitrogen, Breda, The Netherlands), and 10% pooled human serum. After 7 d of incubation at 37°C in a 5% CO2 humidified incubator, cells were harvested and stained for further analysis on the flow cytometer.

**Flow cytometry for CFSE assay**

Cells were stained for 30 min at 4°C using anti-CD8-allophyocyanin (DakoCytomation, Everleee, Belgium), anti-CD3-PerCP, and anti-CD4-PE (BD Biosciences). Cells were then washed in PBS 0.1% BSA (Sigma-Aldrich, Zijnwardrecht, The Netherlands), fixed in 1% paraformaldehyde (Pharmacy, Leiden University Medical Center) and analyzed on an LSRII with an HITS plate loader (BD Biosciences). Analysis was done using FACSDiva software (BD Biosciences). Cells gated on live lymphocytes combined with gating on CD3+CD8+ T cells were analyzed for CFSE proliferation. The Δ geometric mean was used as a measure of proliferation and calculated as follows: Δ geometric mean = geometric mean (non-proliferated cells) − geometric mean (total cells) (see Supplemental Fig. 1 for explanation). The geometric mean was then used to calculate the relative proliferation that is the percentage of the maximal proliferation (PHA) corrected for spontaneous proliferation (medium): (Δ geometric mean sample − Δ geometric mean control medium)/[Δ geometric mean PHA − Δ geometric mean control medium] × 100% = percentage of maximal proliferation. The cutoff value for a positive proliferative response was arbitrarily set at 10% relative proliferation.

**Tetramer staining**

Tetramers were generated using a “one-pot, mix-and-read” strategy recently published by us (41). Tetramer staining was carried out as described in detail previously (15). PBMC (10^5/ml) were incubated in U-bottom 96-well plates, washed twice in PBS containing 1% FCS (Sigma-Aldrich), and stained for 30 min at 4°C with PE-labeled tetramers (35 µl each), washed, subsequently stained with FITC-labeled anti-CD8 mAb (clone HIT8a; BD Biosciences) and analyzed by flow cytometry on an FACS Calibur. Data were calculated with the use of FlowJo software (Tree Star, Ashland, OR). Viable lymphocytes were gated by forward and side scatter, and the analysis was performed on 100,000 acquired CD8 events for each sample. A cutoff of 0.01% was used as described previously (15); values below this were set to zero.

**Intracellular cytokine staining**

ICS was carried out as described in detail previously (15). PBMC (10^5/ml) were stimulated with peptides (1 µg/ml, final concentration) in the presence of monensin for 6 h at 37°C in 5% CO2. The cells were harvested, washed, and stained with PerCP Cy5.5-conjugated anti-CD8 mAb (BD Biosciences) in incubation buffer (PBS-1% FCS-0.1% Na azide) for 30 min at 4°C. The cells were washed twice in PBS-1% FCS and fixed with 4% paraformaldehyde overnight at 4°C. Fixation was followed by permeabilization with PBS-1% FACS-0.3% saponin-0.1% Na azide for 15 min at 4°C. Staining of intracellular cytokines was performed by incubation of fixed permeabilized cells with PE-labeled anti-IFN-γ (clone B27), FITC-labeled IL-2 Ab (clone MQ1-17H12), and allophycocyanin-labeled anti-TNF-α (MAb11) or an isotype-matched control mAb (all from BD Biosciences). Cells were acquired and analyzed by FACS as described above. Analysis was performed on a minimum of 100,000 acquired CD8 events for each sample.

Negative controls were background staining obtained with PBMC incubated with medium, in the absence of any stimulant or with control peptide, or similar experiments using PBMC of PPD nonresponsive donors. Cutoff values for a positive response were predetermined to be in excess of
The complete genome sequence of *M. tuberculosis* strain H37Rv was retrieved from GenBank for in silico epitope prediction using NetCTL. Proteins were included based on eight different selection criteria prior to peptide prediction. The numbers in each selection are given in parentheses. 1. TB-VAC: proteins used in TB vaccine trials; 2. TB-CD8: *M. tuberculosis* proteins with known CD8 T cell epitopes; 3. BestPred: proteins containing peptides with the best prediction values; 4. Cons: conserved *M. tuberculosis* proteins; 5. DOS/LAG: proteins encoded by the *M. tuberculosis* DosR regulon; 6. Bepi: 7. Secret: secreted proteins; and 8. PredSecret: *M. tuberculosis* proteins predicted to be secreted (see Materials and Methods for details). Epitope predictions were done for HLA-A2, A3, and B7. A total of 432 peptides were synthesized, and binding affinities were measured in an in vitro biochemical peptide-HLA class I binding assay (40). CD8* T cell proliferative responses were performed using a CFSE dilution assay on PBMC from PPD* donors in a flow cytometric analysis.

0.01% responsive cells. Results below this value were considered negative and set to zero.

**Statistics**

Nonparametric Mann–Whitney *U* test was used to determine statistical differences in the distribution of the results. The *p* values < 0.05 were considered significant. Data were analyzed using statistical software SYSTAT 11 (Systat Software). The significance of a difference between two proportions was calculated as described by Armitage et al. (42).

**Results**

Selection of *M. tuberculosis* Ags used for bioinformatics-based epitope identification

The complete genome sequence of *M. tuberculosis* strain H37Rv was retrieved from GenBank for in silico epitope prediction using NetCTL. Proteins were selected using eight different criteria (as described in detail in Materials and Methods) to evaluate how proteins could be selected for epitope discovery. Fig. 1 gives an overview of the strategy employed.

**Binding of predicted epitopes to HLA class Ia molecules**

A total of 432 peptides were synthesized and binding to the respective HLA molecule determined as summarized in Table I. A total of 373 (86%) bound to their respective HLA molecules with a *Kd* ≤ 500 nM, which is a normally accepted threshold for immunogenicity (43); 236 (55%) bound with high affinity with a *Kd* ≤ 50 nM. Only 7% of the peptides bound with an affinity < 5000 nM. The fraction of peptides binding to HLA-A2 and -B7 with high affinity (< 50 nM) was significantly higher than for HLA-A3 (*p* < 0.001).

**Immunogenicity of predicted new HLA class I-restricted *M. tuberculosis* epitopes**

PBMC from HLA-A2*, HLA-A3*, and/or HLA-B7*–positive PPD* responsive donors were then screened for possible CD8* T cell proliferative responses toward the novel peptides. Ten donors with the appropriate HLA specificity were tested for each peptide (see also Supplemental Table III).

The results are summarized in Table II (see also Fig. 1, Supplemental Table III). Peptide-induced CD8* T cell responses were scored positive when exceeding 10% relative proliferation in two or more PPD* donors (as we have described before; see Ref. 44 and Supplemental Fig. 1). In case there was a relatively high (ranging between 10 and 20%) response against a given peptide in only one single donor, the peptide-induced response was repeated and only included as hit when reproducible in that donor. According to these criteria, a total number of 70 out of the 432 tested peptides (16%) were found to elicit a CD8* T cell response (Table II).

When the results were compared, 34% (44 out of 130) of the predicted HLA-A2 binding peptides were found capable of eliciting a proliferative CD8* T cell response in HLA-A2* PPD* responsive donors compared with 6% (9 out of 140) for HLA-A3 and 10% (17 out of 163) for HLA-B7. That predicted HLA-A2 binding peptides were more frequently recognized was highly significant (*p* < 0.001 in both cases) and likely reflects the more accurately defined peptide-binding motif for this allele. The difference between the fractions of predicted HLA-A3 and -B7 binding peptides recognized was not significant.

**High-affinity HLA binding peptides are preferentially recognized compared with medium- or low-affinity binding peptides**

A total of 19% (46 out of 236) of the peptides binding to an HLA class Ia molecule with an affinity > 50 nM were recognized by the healthy immune donors (Table I). This fraction is significantly higher than that for the peptides with a lower binding affinity (> 50 nM; *p* < 0.02). The CD8* T cell response frequencies to these latter peptides were 9% (13 out of 137) for intermediate-affinity binders, 14% (4 out of 29) for low-affinity binders, and,

<table>
<thead>
<tr>
<th>HLA</th>
<th>High Binders (&lt;50 nM)</th>
<th>Intermediate Binders (50 nM &lt; <em>Kd</em> ≤ 500 nM)</th>
<th>Low Binders (500 nM &lt; <em>Kd</em> ≤ 5000 nM)</th>
<th>Low or Nonbinders (<em>Kd</em> &gt; 5000 nM)</th>
<th>Total/Supertype</th>
<th>Discarded or Cancelled</th>
</tr>
</thead>
<tbody>
<tr>
<td>A*0201</td>
<td>105 (33)</td>
<td>15 (5)</td>
<td>2 (1)</td>
<td>8 (5)</td>
<td>130</td>
<td>41</td>
</tr>
<tr>
<td>A*0301</td>
<td>18 (1)</td>
<td>82 (4)</td>
<td>20 (3)</td>
<td>19 (1)</td>
<td>139</td>
<td>16</td>
</tr>
<tr>
<td>B*0702</td>
<td>112 (12)</td>
<td>40 (4)</td>
<td>7 (0)</td>
<td>3 (1)</td>
<td>165</td>
<td>9</td>
</tr>
<tr>
<td>Total</td>
<td>236 (46)</td>
<td>137 (13)</td>
<td>29 (4)</td>
<td>30 (7)</td>
<td>432</td>
<td>66</td>
</tr>
</tbody>
</table>

The number of peptides that gave a positive CD8* T cell proliferative response are in parentheses.
Eight different selection criteria were used to select proteins for epitope prediction (Fig. 1). The results are shown in Table II. The highest overall proportional frequencies of CD8+ T cell responses were found in the sets of proteins used in the TB-vaccine Ags set (TB-VAC; 26%) as well as in the secreted M. tuberculosis Ag set (Secret; 25%). Two other broad selection criterions, best predicted epitopes (BestPred) and conserved candidate Ag (Cons), had the lowest frequencies of CD8+ T cell responses. BestPred and Cons together had a significantly lower response rate than the other more specific protein function-based selections combined (TB-VAC, M. tuberculosis CD8+ T cell epitope-expressing Ags [TB-CD8], M. tuberculosis DosR regulon-encoded latency Ags [DOS/LAG], proteins with B cell epitopes [Bepi], Secret, and hypothetical proteins predicted to be secreted [PredSecret]) (p < 0.01). When analyzing HLA-A2–associated epitopes only, we found that the selections DOS/LAG, Bepi, Secret, and PredSecret yielded a higher frequency of CD8+ T cell-stimulating epitopes compared with other selections, such as TB-VAC, TB-CD8, BestPred, and Cons (p < 0.001).

The most frequently recognized HLA-A*0201–restricted peptides were selected for testing in cured TB patients. The 18 most frequently recognized HLA-A*0201–restricted peptides were chosen for further studies. First, we selected 13 of the above-defined 44 HLA-A*0201–restricted peptides, namely those that were recognized in two or more donors (arbitrary cutoff was >10% relative CD8+ T cell proliferation). In addition, we also included five peptides (B118, C250, C255, B130, and B134) that were recognized in a single donor but with a very high CD8+ T cell response (>20% relative CD8+ T cell proliferation). The characteristics of all 18 selected HLA-A*0201–restricted epitopes are summarized in Table III. Sixteen are among the high-affinity HLA-A*0201 binders. The majority having a binding affinity ≤5 nM (n = 14).

Validation and quantitation of M. tuberculosis peptide-specific CD8+ T cell responses in cured TB patients using tetramers

Using HLA-A2 peptide tetramers for the 18 most frequently HLA-A*0201–restricted peptides, direct ex vivo recognition by cured TB patients’ CD8+ T cells was demonstrated for 16 of the 18 M. tuberculosis epitopes (Figs. 2, 3A). Of particular interest was that several epitopes were recognized by the majority of the cured TB patients: pmtb4 (A1) and pmtb15 (B182) were recognized together by 7 out of 10 cured TB patients, whereas several other peptides were recognized by a sizeable fraction of the patients as well. The epitopes studied in this paper thus constitute a significant expansion of the known antigenome for CD8+ T cells during M. tuberculosis infection. In all individuals tested, specificity of tetramer staining was confirmed by the negative data obtained using tetramers of an irrelevant specificity (the HLA-A*0201/Ebola peptide; Fig. 3B), a tetramer of an irrelevant positive control specificity (the HLA-A*0201/CMV peptide, positive in many cured TB patients and healthy controls (HC) PBMC; Fig. 3B), as well as absence of staining among PBMC from normal, uninfected HLA-A*0201–positive donors using the same M. tuberculosis tetramers (Supplemental Fig. 2).

Mono-, dual-, and triple-cytokine–producing peptide-specific CD8+ T cells at the single-cell level in TB

IFN-γ, IL-2, and TNF-α are relevant cytokines defining functional populations of Ag-specific CD4+ and CD8+ T cells (45, 46).
Table III. Characteristics of the 18 novel *M. tuberculosis* HLA-A2–restricted peptides

<table>
<thead>
<tr>
<th>Peptide No.</th>
<th>Selection</th>
<th>Peptide</th>
<th>Protein Description</th>
<th>Protein Id</th>
<th>Gene</th>
<th>Rv No.</th>
<th>Original Peptide No</th>
<th>$K_D$ (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pMtb1</td>
<td>TB-VAC</td>
<td>GLAGGAATA</td>
<td>Secreted Ag 85-B FBPB (85B) (Ag 85 complex B)</td>
<td>NP_216402.1</td>
<td>fbpB</td>
<td>Rv1886c</td>
<td>10851</td>
<td>308</td>
</tr>
<tr>
<td>pMtb2</td>
<td>BestPred</td>
<td>LLYDGSFAV</td>
<td>Hypothetical protein</td>
<td>NP_218356.1</td>
<td>NULL</td>
<td>Rv3839</td>
<td>11611</td>
<td>&lt;1</td>
</tr>
<tr>
<td>pMtb3</td>
<td>Cons</td>
<td>AYDTHMXYV</td>
<td>ATP-dependent Clp protease proteolytic subunit</td>
<td>NP_2146976.1</td>
<td>clpF2</td>
<td>Rv2460c</td>
<td>11679</td>
<td>&lt;1</td>
</tr>
<tr>
<td>pMtb4</td>
<td>DOS/LAG</td>
<td>AMAGSDDL</td>
<td>Probable trehalose-6-phosphate phosphatase OTSB1</td>
<td>YP_177855.1</td>
<td>ecb1</td>
<td>Rv2006</td>
<td>A-1</td>
<td>4</td>
</tr>
<tr>
<td>pMtb5</td>
<td>DOS/LAG</td>
<td>GMEANRWS</td>
<td>Probable metal cation transporter P-type ATPase</td>
<td>NP_216513.1</td>
<td>ctpF</td>
<td>Rv1997</td>
<td>A-4</td>
<td>27</td>
</tr>
<tr>
<td>pMtb6</td>
<td>Secret</td>
<td>YLPDPTGV</td>
<td>Hypothetical protein</td>
<td>NP_217329.1</td>
<td>NULL</td>
<td>Rv2813</td>
<td>B-118</td>
<td>1</td>
</tr>
<tr>
<td>pMtb7</td>
<td>Secret</td>
<td>YYDPGNPLV</td>
<td>MCE-family protein MCE3A</td>
<td>NP_177852.1</td>
<td>mce3A</td>
<td>Rv1966</td>
<td>B-119</td>
<td>2</td>
</tr>
<tr>
<td>pMtb8</td>
<td>Secret</td>
<td>ALLOGPRPV</td>
<td>MCE-family protein MCE4B</td>
<td>NP_218015.1</td>
<td>mce4B</td>
<td>Rv3498c</td>
<td>B-130</td>
<td>1</td>
</tr>
<tr>
<td>pMtb9</td>
<td>Secret</td>
<td>HLDDVGFLV</td>
<td>Possible esterase lipoprotein LPQC</td>
<td>NP_217815.1</td>
<td>lpqC</td>
<td>Rv3298c</td>
<td>B-131</td>
<td>&lt;1</td>
</tr>
<tr>
<td>pMtb10</td>
<td>Secret</td>
<td>SLIDLLIKI</td>
<td>MCE-family protein MCE4A</td>
<td>NP_177977.1</td>
<td>mce4A</td>
<td>Rv3498c</td>
<td>B-132</td>
<td>&lt;1</td>
</tr>
<tr>
<td>pMtb11</td>
<td>Secret</td>
<td>SLNRWIATL</td>
<td>Possible MCE-family lipoprotein LPRL (MCE-family lipoprotein MCE2E)</td>
<td>NP_215107.1</td>
<td>lpqL</td>
<td>Rv0593</td>
<td>B-134</td>
<td>66</td>
</tr>
<tr>
<td>pMtb12</td>
<td>PredSecret</td>
<td>FMYGGVTP</td>
<td>Probable ATP-dependent helicase LHR (larger helicase-related protein)</td>
<td>NP_217813.1</td>
<td>htr</td>
<td>Rv3296</td>
<td>B-178</td>
<td>&lt;1</td>
</tr>
<tr>
<td>pMtb13</td>
<td>PredSecret</td>
<td>ALDEGLPV</td>
<td>Possible conserved membrane protein</td>
<td>NP_218210.1</td>
<td>NULL</td>
<td>Rv3693</td>
<td>B-179</td>
<td>1</td>
</tr>
<tr>
<td>pMtb14</td>
<td>PredSecret</td>
<td>YLLATFIV</td>
<td>Probable membrane-associated phospholipase C2</td>
<td>NP_216866.1</td>
<td>plcB</td>
<td>Rv2350c</td>
<td>B-181</td>
<td>1</td>
</tr>
<tr>
<td>pMtb15</td>
<td>PredSecret</td>
<td>WLYPGAQNL</td>
<td>Probable amino acid decarboxylase</td>
<td>NP_177889.1</td>
<td>NULL</td>
<td>Rv2531c</td>
<td>B-182</td>
<td>&lt;1</td>
</tr>
<tr>
<td>pMtb16</td>
<td>TB-CD8</td>
<td>SLWKDGAPL</td>
<td>Glutamine synthetase GLNA1 (glutamine synthase)</td>
<td>NP_216736.1</td>
<td>glmA</td>
<td>Rv2220</td>
<td>C-249</td>
<td>5</td>
</tr>
<tr>
<td>pMtb17</td>
<td>TB-CD8</td>
<td>KLQERLKL</td>
<td>Chaperonin GroEL</td>
<td>NP_214954.1</td>
<td>groEL</td>
<td>Rv0440</td>
<td>C-250</td>
<td>18</td>
</tr>
<tr>
<td>pMtb18</td>
<td>TB-CD8</td>
<td>LLDSGTTSI</td>
<td>Secreted L-alanine dehydrogenase ALD (40 kDa Ag) (TB43)</td>
<td>NP_217296.1</td>
<td>ald</td>
<td>Rv2780</td>
<td>C-255</td>
<td>1</td>
</tr>
</tbody>
</table>

Peptide-induced T cell responses scored positive when they exceeded 10% CD8+ T cell relative proliferation in multiple donors (two or more donors) or, in the case of 10–20% proliferation in one donor, were confirmed twice in the same donor. In addition, five peptides were included that induced a very strong CD8+ T cell response (>20% relative proliferation) in a single donor.

*The type of protein the peptide is derived from (see Fig. 1).

*Peptide sequence.

*Protein accession number from GenBank (http://www.ncbi.nlm.nih.gov/).

*Rv names for *M. tuberculosis* (Tuberculist [http://genolist.pasteur.fr/TubercuList/] and Ref. 27).

*K_D* (nM) is the binding affinity value for the epitope in nanomolars.
Mono-, dual-, and triple-functional CD8+ T cells have been reported (45). The functionality of CD8+ T cells was assessed by peptide/tetramer analysis in cured TB patients and as control HC were used. The data are shown in Fig. 4, which is a composite figure combining tetramer with cytokine expression data.

With the exception of \textit{M. tuberculosis} peptide pMtb7 (B119) and pMtb6 (B118), single-, double-, and, in some cases, also triple-positive CD8+ T cells could be detected in the cured TB patients. CMV peptide stimulation revealed responses in \( \approx 60\% \) of the cured TB patients, whereas the Ebola peptide induced very low if any responses as expected (all \( \approx 0.2\% \)). Of interest, some peptides induced rather strong responses (\( \approx 0.3\% \)) in the majority of TB patients, including pmtb14 (B181) (80% of the patients), pmtb15 (B182) (80%), and pmtb4 (A1) (70%). These were also the peptides that typically gave the strongest tetramer responses, such that functional and tetramer results for the specific peptides are well in agreement, as visualized in Fig. 4.

Positive tetramer staining and cytokine expression is an almost exclusive property of the cured TB patient group, with, as expected, mostly negative results in the healthy \textit{M. tuberculosis} noninfected group (Figs. 4, 5). Control CMV-peptide/HLA-A2 tetramer CD8+ T cell staining was similarly positive in the \textit{M. tuberculosis}-infected and \textit{M. tuberculosis} noninfected groups, whereas control Ebola-peptide/HLA-A2 tetramer staining was invariably negative.

Taken together, these results indicate that the newly identified \textit{M. tuberculosis} epitopes and Ags are recognized by (poly)functional CD8+ T cells during (control of) infection in TB. The results also reveal a strong correlation between the presence of \textit{M. tuberculosis} peptide-tetramer–positive and polyfunctional CD8+ T cells in natural \textit{M. tuberculosis} infection in man.

**Discussion**

Whereas relatively much is known about the mycobacterial Ags recognized by CD4+ T cells (20), surprisingly little is known about the \textit{M. tuberculosis} Ags, which activate human CD8+ T cells in TB. We applied both forward and reverse Ag discovery approaches to identify new \textit{M. tuberculosis} epitopes for human CD8+ T cells. In this study, we report 70 \textit{M. tuberculosis} HLA-class Ia-restricted CD8+ T cell-stimulating epitopes (44 HLA-A*0201, 9 HLA*0301, and 17 HLA-B*0702), of which 58 are new. The results were validated by using a panel of 20 HLA-A*0201 tetramers and by ICS, which revealed mono-, double-, and triple-functional CD8+ T cell responses. There was a strong agreement between tetramer and cytokine positivity. Positive responses in the validation cohort were seen only in the cured TB patients, but absent from the
healthy non-*M. tuberculosis*-infected control individuals. In another recent study, we have reported the identification of a set of novel *M. tuberculosis* epitopes (>50) that are recognized by human CD8+ T cells in the context of the nonclassical HLA class Ib molecule HLA-E (44). Those results and the ones reported in this study very significantly increase our understanding of the human immune response to *M. tuberculosis* by identifying large sets of CD8+ T cell-recognized *M. tuberculosis* peptides.

We had expected newly discovered epitope frequencies to be highest for Ags from the TB vaccine Ag group and the *M. tuberculosis* CD8+ T cell epitope-expressing Ag group, because these contained highly immunogenic proteins (e.g., from *M.*
tuberculosis RD regions or culture filtrate proteins) (47–51) or proteins containing already known CD8+ T cell epitopes reported by others (13, 16, 22, 45, 52–58). The known CD8+ T cell epitopes were not deliberately deselected for, but our selection criteria were set to select those candidate peptide epitopes with the best predicted combined score and predicted binding affinity.

According to the Immune Epitope Database (IEDB), as of October 1, 2010, there were a total of 151 known MHC class I-
restricted epitopes of *M. tuberculosis* (http://www.immuneepitope.org/). Ninety-three of these known epitopes were found in humans and are restricted by 6 different HLA-A and 10 HLA-B alleles. The majority were HLA-A2–restricted epitopes. Twelve of our predicted and tested peptides can be found in the IEDB and were found to be MHC class I-restricted CD8+ T cell epitopes by other groups while this study was ongoing. These epitopes are highlighted in red in the arrayed tables. Four of these peptides (C228, 10853, 10861, and 10882) were confirmed as epitopes for CD8+ T cells in our study, whereas the remaining eight were either not recognized at all in the study reported in this paper or did not meet the pre-established criteria for positivity: they only scored positive in one donor and/or induced 10% CD8+ T cell proliferation. For example, the known epitope KLQERLAKL (#C250) was able to induce a very high response (55% CD8+ T cell proliferation) but only in 1 of the 10 donors in our experiments. Based on this exceptionally high response, the #C250 peptide was nevertheless included for tetramer construction.

We found that a higher fraction (34%) of the predicted HLA-A2 binding *M. tuberculosis* peptides elicited proliferative CD8+ T cell responses in HLA-A2+ PPD-responsive donors, when compared with the peptides predicted to bind to HLA-A3 (6%) and HLA-B7 (10%). There are several possible explanations for this. One is related to the prediction server used, as NetCTL might have a poorer performance for HLA-A3 and -B7. It is also possible that fewer *M. tuberculosis* peptides are able to bind to these super-types. However, more likely is that the lower frequency of HLA-A3–associated peptide hits is due to the larger diversity of

---

**FIGURE 5.** Representation of the combined tetramer staining and cytokine expression results for all cured TB patients and control individuals. In the table, results for the cured TB patients are shown in the *upper panel*, whereas those for the HC are shown in the *lower panel*. The left side shows the tetramer staining results of each individual donor with each of the 18 selected *M. tuberculosis* peptide/HLA-A2 tetramer, including the two control tetramers. White boxes, no detectable tetramer+ CD8+ T cells; light gray boxes, percentage of tetramer+ CD8+ T cells is between 0.05 and 0.10%; and dark gray boxes, percentage of tetramer+ CD8+ T cells is >0.10%. The right half of the figure shows the sum of the cytokine response. White boxes, no cytokine production could be measured; light gray boxes, the total percentage of cytokine+ CD8+ T cells ranges from 0.05–0.10%; dark gray boxes, the total percentage of cytokine+ CD8+ T cells ranges from 0.10–1.00%; and black boxes, the total percentage of cytokine+ CD8+ T cells is >1.00%.)
specificities among the alleles within the HLA-A3 supertype. Several studies have shown that peptides bind in larger numbers and at higher affinity to alleles of the B locus compared with alleles of the A and C loci (26, 59, 60). These studies therefore concluded that Ag presentation to CD8+ T cells is dominated by B locus alleles. The design of epitope-based vaccines should then be directed toward HLA-B–presented peptides. Our findings do not directly support this, given the higher number of responses to M. tuberculosis peptides predicted to bind to HLA-A2 supertype molecules. Interestingly, a similar distribution of epitopes restricted by different supertypes was also reported by Pasquetto et al. (61), who performed a large scale screening for pox-virus epitopes. In that study, the frequencies were 30, 6, and 21% for HLA-A2, A3, and B7, respectively.

Determination of the peptide-binding affinities for the respective supertypes revealed a significant correlation between high-affinity (<50 nM) peptide/MHC binding and the peptides’ ability to trigger CD8+ T cell responses (Table I). In poxvirus systems, similar associations were reported (62, 63). Unexpectedly, however, in our study several of the nonbinding peptides (23%; 7 out of 30) were able to induce positive CD8+ T cell responses (Table I). This may be because the cell-free peptide-MHC binding assay might have failed to capture key characteristics of (low-affinity) peptide/MHC binding (many other chaperones are involved next to MHC alone) or that these peptides actually bound to and were presented by HLA class I molecules other than A2/A3/B7.

We have used HLA-A2/pentamer peptide-specific single-cell multicytokine analyses to validate the above findings for the 18 most prominently recognized new M. tuberculosis HLA-A2–restricted epitopes. Importantly, direct ex vivo recognition by cured TB patients’ CD8+ T cells was found for 16 of the 18 M. tuberculosis epitopes (Figs. 2, 3A). The epitopes studied in this paper thus constitute a significant expansion of the known anti-M. tuberculosis CD8+ T cells is dominant of the cured TB patients as well as controls, whereas the Ebola peptide induced very low if any responses as expected. Of interest was that some peptides induced strong responses in the majority of the cured TB patients, including pmtb14 (B181) (80% of the patients), pmtb15 (B182) (80%), and pmtb4 (A1) (70%). These same peptides also showed the strongest tetramer responses. Thus, multifunctional cytokine production and tetramer results for the specific M. tuberculosis epitopes are well in agreement (Fig. 4).

Although more extensive phenotyping of M. tuberculosis–specific IFN-γ–, IL-2–, and TNF-α–secreting CD8+ T cells was beyond the scope of this study, previous studies have identified a relationship between the function and phenotype of memory CD4+ T cells and have proposed that the IL-2 only-secreting cells are typical of central memory T cells that persist after Ag clearance, whereas the IFN-γ/IL-2– and IFN-γ only-secreting T cells are typical of effector memory T cells (42).

In conclusion, we have identified 70 M. tuberculosis-specific CD8+ T cell epitopes in healthy PPD+ individuals. Fifty-eight of these epitopes were novel and have not previously been described by others. This is a considerable expansion of the existing list of known M. tuberculosis CD8+ T cell epitopes in the IEDB. Furthermore, our results indicate that the newly identified epitopes and Ags are recognized by (poly)functional CD8+ T cells during (control of) infection in TB. Finally, the results reveal a strong correlation between the presence of M. tuberculosis peptide-tetramer–positive and polyfunctional CD8+ T cells in natural M. tuberculosis infection in humans. These results provide a wealth of new M. tuberculosis Ags that may provide targets for TB vaccine development, particularly in the view of mounting evidence that CD8 T cells are important in controlling TB. Moreover, the epitopes we have identified may provide novel tools for monitoring the specific CD8+ T cell response in TB cohorts, providing potential novel TB biomarkers, analogous to what we have reported for a limited set of other CD8+ T cell epitopes in TB recently (15).

Disclosures
The authors have no financial conflicts of interest.

References


