CLINICAL AND VACCINE IMMUNOLOGY, Jan. 2010, p. 168–175 Vol. 17, No. 1
1556-6811/10/$12.00 doi:10.1128/CVI.00208-09

Peptide Microarray-Based Identification of *Mycobacterium tuberculosis* Epitope Binding to HLA-DRB1*0101, DRB1*1501, and DRB1*0401▽†

Simani Gaseitsiwe,[1,2] Davide Valentini,[3] Shahnaz Mahdavifar,[1,2] Marie Reilly,[3] Anneka Ehrnst,[1] and Markus Maeurer[1,2]*

Department of Microbiology, Tumor and Cell Biology (MTC), Karolinska Institutet, Stockholm, Sweden[1]; Swedish Institute for Infectious Disease Control (SMI), Stockholm, Sweden[2]; and Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden[3]

Received 23 May 2009/Returned for modification 30 July 2009/Accepted 6 October 2009

A more effective vaccine against *Mycobacterium tuberculosis* is needed, and a number of *M. tuberculosis* vaccine candidates are currently in preclinical or clinical phase I and II studies. One of the strategies to select *M. tuberculosis* (protein) targets to elicit a $CD8^+$ or $CD4^+$ T-cell response is to gauge the binding of candidate peptides to major histocompatibility complex (MHC) class I or class II molecules, a prerequisite for successful peptide presentation and to expand antigen-specific T cells. We scanned 61 proteins from the *M. tuberculosis* proteome for potential MHC class II-presented epitopes that could serve as targets for $CD4^+$ T-cell responses. We constructed a peptide microarray consisting of 7,466 unique peptides derived from 61 *M. tuberculosis* proteins. The peptides were 15-mers overlapping by 12 amino acids. Soluble recombinant DRB1*0101 (DR1), DRB1*1501 (DR2), and DRB1*0401 (DR4) monomers were used to gauge binding to individual peptide species. Out of 7,466 peptides, 1,282, 674, and 1,854 peptides formed stable complexes with HLA-DR1, -DR2, and -DR4, respectively. Five hundred forty-four peptides bound to all three MHC class II molecules, 609 bound to only two, and 756 bound to only a single MHC class II molecule. This allowed us to rank *M. tuberculosis* proteins by epitope density. *M. tuberculosis* proteins contained "hot spots," i.e., regions with enriched MHC class II binding epitopes. Two hundred twenty-two peptides that formed MHC class II-peptide complexes had previously been described as exclusively recognized by IgG in sera from patients with active pulmonary tuberculosis, but not in sera from healthy individuals, suggesting that these peptides serve as B-cell and $CD4^+$ T-cell epitopes. This work helps to identify not only *M. tuberculosis* peptides with immunogenic potential, but also the most immunogenic proteins. This information is useful for vaccine design and the development of future tools to explore immune responses to *M. tuberculosis*.

$CD4^+$ T cells play a central role in *Mycobacterium tuberculosis*-directed cellular immune responses (2, 6, 7, 12). It is most likely that an effective tuberculosis (TB) vaccine would target the expansion of $CD8^+$ and $CD4^+$ T cells, which recognize *M. tuberculosis* peptides presented by major histocompatibility complex (MHC) class I and class II molecules.

The MHC locus is the most variable gene locus in the human genome, and the variability of MHC class II alleles in different populations is well documented (24). Certain MHC class II alleles have been shown to be associated with *M. tuberculosis* infection (1, 11, 15, 16, 23): DRB1*0803 and DQB1*0601 were found to be associated with TB disease progression, development of drug resistance, and disease severity in Koreans (16). In South Africa, DRB1*1302 and DQB1*0301 to -0304 were apparently associated with active TB compared to control individuals lacking these alleles (23). The prevalence of HLA-DRB1*0401 and HLA-DRB1*0801 was significantly decreased in Mexican patients with pulmonary TB compared to their prevalence in healthy controls (35).

The association of some MHC class II alleles with "better disease outcome" could be due to the fact that these alleles are "better" at binding and presenting a certain repertoire of peptide epitopes to $CD4^+$ T cells than other alleles. The identification of peptides binding to molecularly defined MHC class II alleles could therefore represent an important first step in identifying potential targets for TB vaccine design and the development of new diagnostic assays. More recently, De Groot and colleagues used a bioinformatics approach, followed by validation with functional assays to identify $CD4^+$ T-cell epitopes that were used to construct an epitope-based *M. tuberculosis* vaccine (5).

Only a few *M. tuberculosis* MHC class II binding peptides have been identified so far, and 7% of the *M. tuberculosis* open reading frames have been explored for both B-cell and T-cell epitopes (3). We described a peptide microarray assay that allowed us to visualize HIV peptide binding to molecularly defined MHC class II alleles (9). The assay has the major advantage that a high number of candidate peptides can be screened within a short time frame. In the current report, we describe *M. tuberculosis* peptide binding to the three most frequently encountered MHC class II alleles in different populations; DRB1*0101 (DR1), DRB1*1501 (DR2), and

* Corresponding author. Mailing address: Department of Microbiology, Tumor and Cell Biology (MTC), Karolinska Institutet, and the Swedish Institute for Infectious Disease Control (SMI), Nobels Väg 18, SE. 17182 Stockholm, Sweden. Phone: 46 84572650. Fax: 46 8337460. E-mail: markus.maeurer@ki.se.

† Supplemental material for this article may be found at http://cvi.asm.org/.

▽ Published ahead of print on 28 October 2009.

DRB1*0401 (DR4). DR1, DR2, and DR4 exhibit population frequencies of 15.4%, 32.9%, and 20.9% among Caucasians. In the Botswana population, HLA-DRB1*01, -DRB1*02, and -DRB1*04 show population frequencies of 21.7%, 21.3%, and 14.4%, respectively. The candidate test peptides are derived from 61 *M. tuberculosis* proteins that have been tested for IgG and IgA recognition in patients with active pulmonary TB. The data sets contribute to defining "immunogenicity" in *M. tuberculosis* candidate target proteins, visualize MHC class II epitope "hot spots," and allow us to link B-cell targets and potential MHC class II-presented *M. tuberculosis* epitopes.

MATERIALS AND METHODS

***Mycobacterium tuberculosis* peptides.** Sixty-one *M. tuberculosis* proteins were printed as overlapping peptide (15-amino-acid) stretches on microarray slides, as reported previously (10). Most of these proteins have not been mapped for MHC class II binding, except for antigen 85B, heat shock protein HSPX, and MPT63 (Rv1926c) (20). These data were therefore available for comparative analysis. The biological functions of the 61 proteins in the *M. tuberculosis* life cycle have been addressed in detail previously (10), and an overview is provided in Table 1.

Peptide microarray printing. The peptide microarray slides used in this experiment were produced by JPT, Germany. The peptides were synthesized as amino-oxy-acetylated peptides on cellulose membranes in a parallel manner using SPOT synthesis technology (8, 32). The printing process was carried out as reported previously (28), and the slides were stored at 4°C until they were ready for use.

Soluble HLA class II alleles. Three MHC class II alleles, HLA-DRB1*0101 (DR1), -DRB1*1501 (DR2), and -DRB1*0401 (DR4), were supplied by Beckman Coulter. The process for the production of these alleles has been described in detail elsewhere (29).

Sample processing. HLA-DR monomers were incubated with the peptide microarrays as described previously (9). Briefly, MHC class II monomers were diluted to a working concentration of 1 μg/ml using a binding buffer (36 mM phosphate, 14.4 mM citrate, 0.15% bovine serum albumin [BSA], 0.25% octyl beta-D-glucopyranoside, 0.02% NaN_3, pH 5.5). Three hundred microliters of the HLA-DR–buffer mixture was incubated with the peptide microarray slide for 48 h at 37°C in a humid chamber. The slides were then washed three times for 5 min each, two times with washing solution (phosphate-buffered saline [PBS] and 0.05% Tween 80), and once with PBS alone. Next, the slides were incubated for 1 h at room temperature with 300 μl of a Cy5-labeled monoclonal antibody (MAb) (clone L243, obtained from Beckman Coulter) diluted to 5 μg/ml in PBS to detect stable MHC class II-peptide complexes. The slides were dried by spinning them for 10 s using a slide spinner (Euro Tech, United Kingdom). Two slides were incubated with each monomer, and two slides were incubated with buffer and the detection antibody to identify peptides that were recognized by secondary antibody. These peptides were excluded from analysis.

Data acquisition. (i) Scanning and analysis. Each slide was scanned with the GenPix 4000B microarray scanner (Axon Instruments) at two wavelengths, 532 and 635 nm, and the images were saved in TIFF and JPG formats. Image analysis was performed utilizing the circular feature alignment of the GenePix Pro 5.1 software and the Genepix Array List (GAL) files supplied by JPT, Berlin, Germany. Spots with nonuniform foreground or background signals were flagged if they satisfied the following criteria: {[F635 mean] > (1.5 × [F635 median])} and ([F635 median] > 40) or {[B635 mean] > (1.5 × [B635 median])} and ([B635 median] > 40).

These and other flags assigned by GenePix resulted in four types of spots: "good" or "nonflagged" spots (labeled as 0′), "bad" or "flagged" spots (labeled as −100′), not-found spots (labeled as −50′), and empty spots (labeled as −75′). The image from each subarray was saved as a GenePix result (GPR) file, and the median foreground and background intensities for the 635-nm wavelength from individual peptide spots were used in the response analysis. All GPR files were saved in a common folder and imported into R/Bioconductor using the read.GenePix function from the marray R/bioconductor package.

We examined the distribution of the flags (listed above) to monitor the acquired data for quality control purposes. This quality control exercise was conducted for each of the four groups of slides ([i] slides incubated with buffer only, [ii] HLA-DRB1*0101 slides, [iii] HLA-DRB1*1501 slides, and [iv] HLA-DRB1*0401 slides), both overall and stratified by the type of feature (control or peptide spots). Visual inspection of the images from the individual subarrays was carried out using the Image function in Bioconductor in order to evaluate questionable responses that should be excluded from data analysis. For a measure of the strength of the response, we chose the ratio of the median foreground to background (on a log scale). This response index was computed for all spots with background greater than zero, and any spots with zero background were excluded. The data for each of the four groups of slides were arranged in a large matrix, with columns identifying slide, subarray, and block. All the analyses described below used these master data sets.

(ii) Data reduction. Using the distribution of the negative controls to define a cutoff for a "detectable" response, we removed the spots with no detectable response on any slide. The method used to define the cutoff has been described previously (23). Any peptide with a high response on slides incubated with buffer only and the Cy5-labeled MAb L243 was considered a false positive and discarded from analysis. After all valid (i.e., unflagged) peptide responses on the buffer slides were normalized using the same linear model as for the negative controls, the cutoff was determined for the definition of a false-positive event.

(iii) Analysis of peptide responses. For each group of slides incubated with soluble recombinant MHC class II molecules, we used the thresholds defined above to exclude from the analysis any peptide that (i) had no detectable response on any slide or (ii) had a false-positive response in at least 10% of replicates. The remaining peptide responses were normalized using a linear model to remove artifacts due to slide, subarray, and block. Since the systematic effects of slide, subarray, and block were removed, we refer to these as the "normalized responses." For any peptides that were replicated, the normalized values were averaged. Thus, the preprocessed data consist of a list of unique peptides with their normalized values for each slide.

RESULTS

Peptides binding to the three soluble HLA-DR alleles. Peptide microarray slides printed with 7,446 unique peptides derived from 61 *Mycobacterium tuberculosis* proteins were incubated with soluble MHC class II monomers, i.e., DRB1*0101, DRB1*1501, and DRB1*0401. The printed peptides were 15-mers overlapping by 12 amino acids. The peptide microarray slides also contained empty spots, which were used as negative control spots, and Cy3 spots, which were used for GAL file alignment.

The reported average index represents a function of both the binding affinity and the off rate of the MHC class II-peptide interaction. Each of these factors contributes to the signal intensity of the antibody that detects properly folded MHC class II-peptide complexes. Table 1 lists the *M. tuberculosis* proteins used to screen for MHC class II monomer interaction. The complete list of peptides binding to the MHC class II monomers is provided in Table S1 in the supplemental material. We observed binding of 1,282, 674, and 1,854 peptides to HLA-DRB1*0101, -DRB1*1501, and -DRB1*0401, respectively.

To evaluate the MHC class II-peptide binding pattern for the entire 61 *M. tuberculosis* proteins (and the three MHC class II alleles), we carried out a Pearson centered hierarchical clustering analysis (Fig. 1). The peptides are clustered into groups recognized by only one monomer, groups recognized by two monomers, and groups recognized by all three MHC class II monomers. There were more peptides binding to HLA-DRB1*0401 than to HLA-DRB1*1501 and -DRB1*0101.

MHC class II epitope densities on *M. tuberculosis* proteins. Next, we analyzed *M. tuberculosis* peptides that bind with a particular index value (i.e., the measure of MHC class II-peptide complex formation) in a reproducible fashion; we set an average index value cutoff of 0.00 in at least 2 of 3 repeats. We then calculated the epitope density of each individual *M. tuberculosis* protein, defined as the number of peptides binding to any MHC class II monomer per total number of peptides

TABLE 1. *M. tuberculosis* proteins tested for MHC class II peptide binding

Protein name	Gene name	Functional classification
10-kDa chaperonin GROES CAB01005	Rv3418c	Virulence, detoxification, adaptation
60-kDa chaperonin 2 P0A521	Rv0440	Virulence, detoxification, adaptation
Heat shock protein HSP CAA17343	Rv0251c	Virulence, detoxification, adaptation
Heat shock protein HSPX CAA17245	Rv2031c	Virulence, detoxification, adaptation
MCE family protein MCE1A YP_177701	Rv0169	Virulence, detoxification, adaptation
Possible hemolysin CAA16235	Rv3922c	Virulence, detoxification, adaptation
Possible hemolysin-like protein CAA17201	Rv1085c	Virulence, detoxification, adaptation
3-Oxoacyl-[acyl-carrier protein] synthase 2 KASB CAA94642	Rv2246	Lipid metabolism
Biotinylated protein TB7.3 CAB08316	Rv3221c	Lipid metabolism
M. Bovis acyl- coenzyme A synthase gene; U75685		Lipid metabolism
M. Bovis mycocerosic acid synthase gene M95808		Lipid metabolism
Putative cyclopropane-fatty-acyl-phospholipid synthase UFAA1 NP_854118	Rv0447c	Lipid metabolism
Secreted antigen 85-B FBPB (85B) CAB10044	Rv1886c	Lipid metabolism
Alternate RNA polymerase sigma factor SIGF CAB07069	Rv3286c	Information pathways
Probable 50S ribosomal protein L7/L12 RPLL CAB07109	Rv0652	Information pathways
RNA polymerase beta subunit AAA21416	Rv0667	Information pathways
RNA polymerase beta′ subunit AAA21417	Rv0667	Information pathways
10-kDa culture filtrate antigen ESXB	Rv3874	Cell wall and cell processes
Immunogenic protein MPT63 CAB06500	Rv1926c	Cell wall and cell processes
Cell surface lipoprotein MPT83 CAB08316	Rv2873	Cell wall and cell processes
Conserved hypothetical protein TB9.8 CAD93159	Rv0287	Cell wall and cell processes
ESAT-6-like protein ESXQ CAA16102	Rv3017c	Cell wall and cell processes
Immunogenic protein MPT64 CAA98382	Rv1980c	Cell wall and cell processes
Lipoprotein LPQH precursor P0A5J0	Rv3763	Cell wall and cell processes
Low-molecular-weight protein antigen 7 ESXH TB10.4 CAA17363	Rv0288	Cell wall and cell processes
Low-molecular-weight T-cell antigen TB8.4NP_215690	Rv1174c	Cell wall and cell processes
Major secreted immunogenic protein CAA98373	Rv2875	Cell wall and cell processes
Periplasmic phosphate-binding lipoprotein PSTS1 YP_177770	Rv0934	Cell wall and cell processes
Periplasmic phosphate-binding lipoprotein PSTS2 YP_177769	Rv0932c	Cell wall and cell processes
Periplasmic phosphate-binding lipoprotein PSTS3 YP_177768	Rv0928	Cell wall and cell processes
Probable cutinase precursor CFP21 NP_216500	Rv1984c	Cell wall and cell processes
Probable lipoprotein LPRJ CAB10947	Rv1690	Cell wall and cell processes
Putative ESAT-6-like protein ESXN YP_177838	Rv1793	Cell wall and cell processes
Secreted ESAT-6-like protein ESXR CAA16104	Rv3019c	Cell wall and cell processes
6-kDa early secretory antigenic target ESXA(ESAT-6) CAE55648	Rv3875	Cell wall and cell processes
PE family protein CAE55335	Rv0916c	PE/PPE
PPE family protein CAE55371	Rv1196	PE/PPE
PPE family protein YP 177963	Rv3347c	PE/PPE
PPE family protein CAE55334	Rv0915c	PE/PPE
PPE family protein CAE55489	Rv2430c	PE/PPE
PPE family protein CAE55504	Rv2608	PE/PPE
Secreted L-alanine dehydrogenase ALD CAA15575	Rv2780	Intermediary metabolism and respiration
Possible glycosyltransferase CAB05418	Rv2958c	Intermediary metabolism and respiration
Possible glycosyltransferase CAB05419	Rv2957	Intermediary metabolism and respiration
Probable isocitrate dehydrogenase CAA17111	Rv3339c	Intermediary metabolism and respiration
Probable isocitrate dehydrogenase CAA16247	Rv0066c	Intermediary metabolism and respiration
Probable molybdopterin-guanine dinucleotide biosynthesis protein CAA16030	Rv2453c	Intermediary metabolism and respiration
Probable serine protease PEPA CAB09453	Rv0125	Intermediary metabolism and respiration
Probable serine protease PEPD CAA17582	Rv0983	Intermediary metabolism and respiration
Possible glycosyltransferase CAB05415	Rv2962c	Intermediary metabolism and respiration
MTB48 AAK31576		Unknown
MTB81		Unknown
Transmembrane serine/threonine-protein kinasE D PKND NP_215446	Rv0931c	Regulatory proteins
Two-component transcriptional-regulatory protein DEVR NP_217649	Rv3133c	Regulatory proteins
Conserved hypothetical protein CFP17 CAB01474	Rv1827	Conserved hypotheticals
Conserved hypothetical protein TB16.3CAD97060	Rv2185c	Conserved hypotheticals
Conserved hypothetical protein TB18.5CAD93033	Rv0164	Conserved hypotheticals
Hypothetical protein NP_217139	Rv2623	Conserved hypotheticals
Hypothetical protein NP_217142	Rv2626c	Conserved hypotheticals
Iron-regulated conserved hypothetical protein CAB08889	Rv1636	Conserved hypotheticals
Conserved hypothetical protein CAB08634	Rv2629	Conserved hypotheticals

from the respective *M. tuberculosis* protein. Figure 2 shows the top 20 *M. tuberculosis* proteins with the highest epitope densities; the epitope densities of the entire set of 61 *M. tuberculosis* proteins are provided in Table 2. The "epitope density value" provides a good estimate of which proteins are likely to provide epitopes to DRB1*0101, DRB1*1501, and/or DRB1*0401 MHC class II molecules; they are also likely to provide more $CD4^+$ T-cell epitopes, which may lead to T-cell expansion if

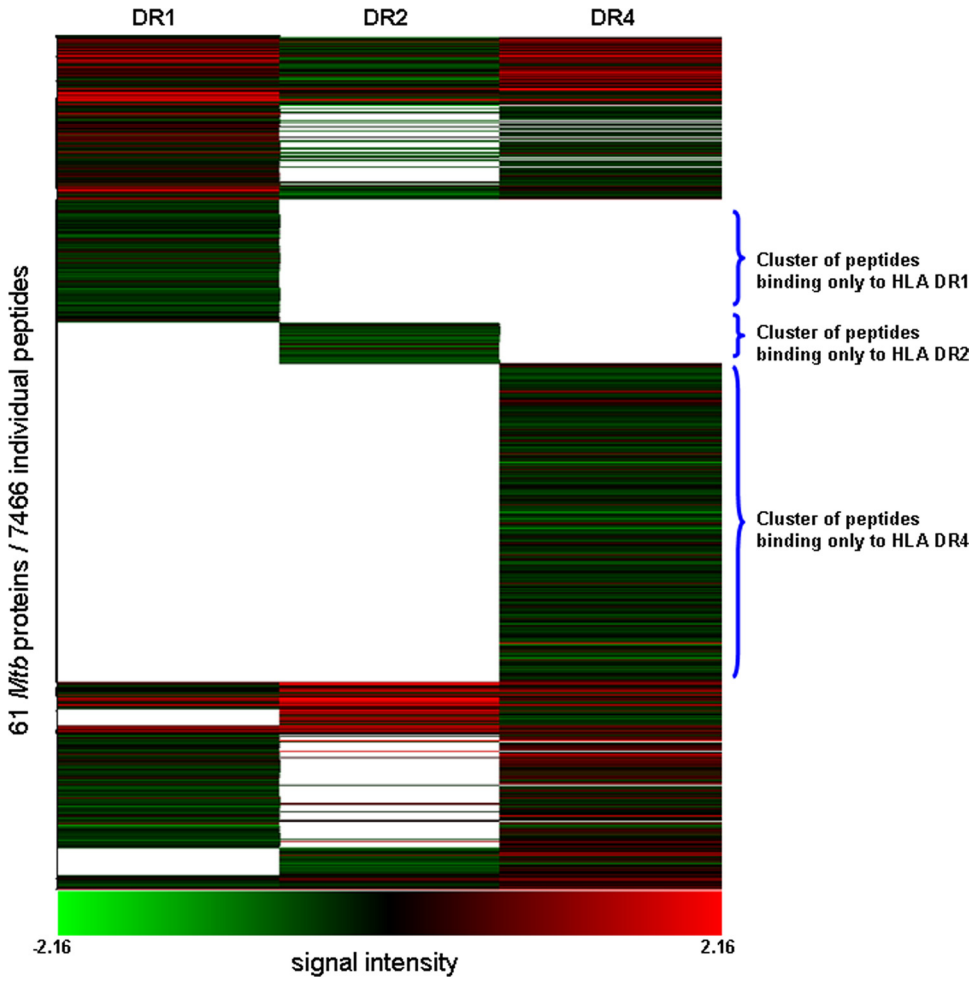

FIG. 1. Pearson centered hierarchical clustering analysis of *M. tuberculosis* peptides from 61 proteins binding to MHC class II monomers. Peptides binding to only one of the MHC class II alleles can be identified, as well as peptide groups that bind to only two or a to a single MHC class II molecule. *Mtb*, *M. tuberculosis*.

the appropriate T-cell receptors (TCRs) are present in the TCR repertoire at the time of vaccination or exposure to the nominal target antigen. We then compared the Ag85B and MPT63 peptides that tested positive for MHC class II monomer binding with previously published T-cell epitopes (Fig. 3). For MPT63, most of the peptides identified by use of the current approach have been described as $CD4^+$ T-cell epitopes.

MHC class II binding peptides represent commonly recognized Ig epitopes in sera from patients with pulmonary TB. We previously identified three patterns of IgG and IgA reactivity to *M. tuberculosis* target peptides: (i) epitopes that are exclusively recognized in individuals with pulmonary TB (and not in healthy individuals), (ii) epitopes that are recognized in healthy subjects and not in patients with pulmonary TB, and (iii) epitopes that are recognized in both TB patients and healthy controls, but in a differential manner, i.e., either strongly in one group and weakly in the other group or *vice versa*. Based on the observation that B- and T-cell epitopes can overlap, as defined by the SEREX approach in screening for tumor-specific B- and T-cell responses (20), we tested whether any MHC class II binding peptide identified in the current report would also serve as targets for an IgG response in sera from patients with acute pulmonary TB. Note that we screened only *M. tuberculosis* epitopes that were commonly recognized ($n = 35/35$ patients) in sera from patients with TB and not in any healthy individual ($n = 34$) for MHC class II binding. Two hundred twenty-two *M. tuberculosis* peptides that bound to any of the three MHC class II monomers were also defined as IgG epitopes in sera from patients with TB (14). Out of these 222 peptides, 33 bound to all three MHC class II monomers, 24 bound to only two monomers, and 165 bound to only a single MHC class II allele. Eighty peptides bound to HLA-DRB1*0101, 52 bound to HLA-DRB1*1501, and 185 peptides bound to HLA-DRB1*0401. These peptides are listed in Table S2 in the supplemental material.

DISCUSSION

Only 7% of the 4,000 open reading frames of *M. tuberculosis* have been explored for B-cell and T-cell epitopes. This is due to the size of the *M. tuberculosis* genome, i.e., the number of

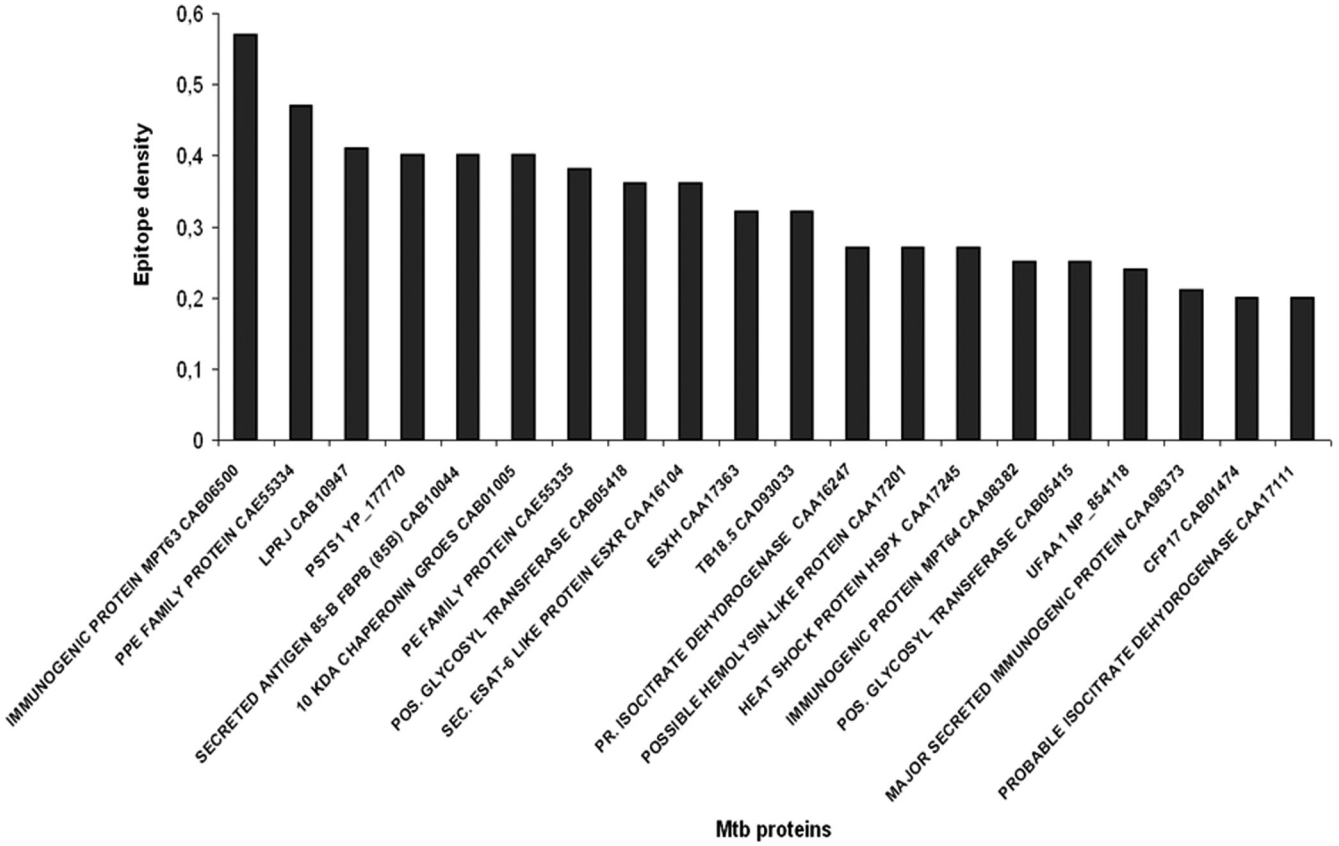

FIG. 2. Epitope densities of the top 20 *M. tuberculosis* proteins defined by the number of peptides binding to MHC class II molecules.

protein targets to be tested, and to the lack of appropriate technology to explore such a massive data set in an affordable manner. This report describes the detailed analysis of an *M. tuberculosis* peptide microarray using 7,446 overlapping peptides from 61 individual *M. tuberculosis* proteins to identify potential T-cell epitopes that could be presented by three common MHC class II alleles, HLA-DRB1*0101, -DRB1*1501, and -DRB1*0401. Most of the peptides that we identified bound to more than a single MHC class II molecule; only a few peptides bound to only one MHC class II allele. This is not surprising, since MHC class II peptide binding is quite promiscuous (17, 30), and all MHC class II binding *M. tuberculosis* peptides listed by Blythe and coworkers bound to three or more MHC class II alleles (3).

We identified more *M. tuberculosis* peptides that bound to HLA-DRB1*0401 than to HLA-DRB1*0101 or -DRB1*1501. HLA-DRB1*0401 was found to be associated with pulmonary TB in Italian patients (31). Thus, whether a broader *M. tuberculosis* peptide epitope presentation by HLA-DRB1*0401 is beneficial or detrimental to mounting a protective anti-*M. tuberculosis*-directed $CD4^+$ T-cell response has to be explored in future studies.

Ranking of *M. tuberculosis* proteins by epitope densities identified MPT63 Rv1926c and PPE CAE55334/Rv0915c as the 2 of 61 *M. tuberculosis* proteins with the highest epitope densities. Immunization of C57BL/6 mice with MTB41 Rv0915c DNA induced protection against *M. tuberculosis* infection comparable to the protection induced by *Mycobacterium bovis* BCG (34), and the cellular immune responses were dominated by $CD4^+$ T cells. Analysis of T-cell responses was carried out using the Rv0915c protein, and the immune responses were not determined on the peptide level.

Rv1926c, an *M. tuberculosis*-secreted protein, has recently been shown to be recognized (26) in healthy BCG-vaccinated subjects. Peptides binding to different MHC class II alleles were identified using a virtual matrix-based prediction program (ProPred). Nine (Rv1926c) peptides predicted to serve as promiscuous $CD4^+$ T-cell epitopes (24) show significant overlap with the peptides that we identified as binding to the three MHC class II alleles. This lends support to our approach that MHC class II binding peptide species, defined by the interaction of soluble MHC class II molecules and immobilized peptides on a microarray chip, serve as $CD4^+$ T-cell epitopes (9).

We were able to match MHC class II binding *M. tuberculosis* epitopes with peptide epitopes that were exclusively recognized by IgG from patients with TB (10). These candidate epitopes may represent clinically relevant targets for diagnostics. Conversely, MHC class II binding *M. tuberculosis* peptides recognized in individuals who have been exposed to *M. tuberculosis* but who are protected from development of disease may represent reasonable *M. tuberculosis* vaccine candidates. The identification of such *M. tuberculosis* targets (i.e., exclusive recognition in a clinically well-defined population) in association with "good immunogenicity," defined by MHC class II epitope density, is currently under way in our laboratory.

The current study is limited, since we did not test the iden-

TABLE 2. Number of peptides from each individual *M. tuberculosis* protein binding to three MHC class II alleles[a]

Protein	No. of peptides	No. of epitopes				No. of epitopes/peptide[b]
		DR1	DR2	DR4	Total	
Immunogenic protein MPT63 CAB06500	49	4	7	17	28	0.57
PPE family protein CAE55334	137	23	16	25	64	0.47
Probable lipoprotein LPRJ CAB10947	39	3	4	9	16	0.41
Periplasmic phosphate-binding lipoprotein PSTS1 YP_177770	121	17	12	20	49	0.40
Secreted antigen 85-B FBPB (85B) CAB10044	105	16	8	18	42	0.40
10-kDa chaperonin GROES CAB01005	30	6	2	4	12	0.40
PE family protein CAE55335	29	1	2	8	11	0.38
Possible glycosyltransferase CAB05418	139	12	10	28	50	0.36
Secreted ESAT-6-like protein ESXR CAA16104	28	5	0	5	10	0.36
Low-molecular-weight protein antigen 7 ESXH CAA17363	28	3	1	5	9	0.32
Conserved hypothetical protein TB18.5CAD93033	50	6	4	6	16	0.32
Probable isocitrate dehydrogenase CAA16247	245	30	9	28	67	0.27
Possible hemolysin-like protein CAA17201	77	9	3	9	21	0.27
Heat shock protein HSPX CAA17245	44	3	3	6	12	0.27
Immunogenic protein MPT64 CAA98382	72	8	2	8	18	0.25
Possible glycosyltransferase CAB05415	146	10	7	19	36	0.25
Putative cyclopropane-fatty-acyl-phospholipid synthase UFAA1 NP_854118	139	11	6	16	33	0.24
Major secreted immunogenic protein CAA98373	61	2	0	11	13	0.21
Conserved hypothetical protein CFP17 CAB01474	50	2	2	6	10	0.20
Probable isocitrate dehydrogenase CAA17111	133	12	5	9	26	0.20
MCE family protein MCE1AYP_177701	148	15	4	9	28	0.19
Heat shock protein HSP CAA17343	49	2	2	5	9	0.18
Conserved hypothetical protein TB16.3CAD97060	44	5	2	1	8	0.18
Possible hemolysin CAA16235	36	3	1	2	6	0.17
M. bovis mycocerosic acid synthase gene; M95808	700	35	16	61	112	0.16
Probable cutinase precursor CFP21 NP_216500	69	1	3	7	11	0.16
Two-component transcriptional regulatory protein DEVR NP_217649	69	2	5	4	11	0.16
Possible glycosyltransferase CAB05419	88	6	4	4	14	0.16
Probable molybdopterin-guanine dinucleotide biosynthesis protein CAA16030	63	1	3	6	10	0.16
Hypothetical protein RV2623 NP_217139	95	4	0	11	15	0.16
Transmembrane serine/threonine-protein kinase D PKND NP_215446	218	13	7	12	32	0.15
M. bovis acyl-coenzyme A synthase gene; U75685	190	7	6	14	27	0.14
ESAT-6-Like protein ESXQ CAA16102	36	0	1	4	5	0.14
Conserved hypothetical protein TB9.8 CAD93159	29	2	1	1	4	0.14
RNA polymerase beta subunit AAA21416	389	19	11	23	53	0.14
PPE family protein CAE55371	127	4	1	12	17	0.13
Conserved hypothetical protein CAB08634	121	2	2	11	15	0.12
Secreted L-alanine dehydrogenase ALD CAA15575	120	3	1	9	13	0.11
10-kDa culture filtrate antigen ESXB CAA17966	30	0	0	3	3	0.10
3-oxoacyl-[acyl-carrier protein] synthase 2 KASB CAA94642	142	4	3	7	14	0.10
PPE family protein YP_177963	1,037	33	24	44	101	0.10
MTB81	243	7	10	6	23	0.09
Low-molecular-weight T-cell antigen TB8.4 NP_215690	33	0	1	2	3	0.09
Iron-regulated conserved hypothetical protein CAB08889	45	0	0	4	4	0.09
MTB48 AAK31576	150	4	3	6	13	0.09
Probable serine protease PEPD CAA17582	151	4	3	6	13	0.09
Cell surface lipoprotein MPT83 CAA98350	70	2	0	4	6	0.09
Probable serine protease PEPA CAB09453	115	0	1	8	9	0.08
60-kDa chaperonin 2 P0A521	176	3	5	5	13	0.07
Conserved hypothetical protein CAA15739	39	0	0	2	2	0.05
Conserved hypothetical protein CAB06237	130	0	3	2	5	0.04
Probable 50S ribosomal protein L7/L12 RPLL CAB07109	40	0	0	1	1	0.03
Alternate RNA polymerase sigma factor SIGF CAB07069	83	0	0	2	2	0.02
Lipoprotein LPQH precursor P0A5J0	49	0	0	1	1	0.02

[a] DRB1*0101, DRB1*1501, and DRB1*0401.
[b] The "epitope density" of each protein.

tified HLA-DR binding peptides for $CD4^+$ T-cell recognition using peripheral blood mononuclear cells (PBMCs) from HLA-DR-matched patients with tuberculosis; corresponding $CD4^+$ T cells endowed with a clonotypic TCR may not be part of the TCR repertoire in individual patients. The disadvantages of testing peptide-specific T-cell recognition may include the facts that (i) a single cytokine, e.g., interferon, may not reflect the breadth of a $CD4^+$ T-cell response and (ii) individual peptide species may not be stable and may be quickly degraded in a standard assay gauging intracellular cytokine

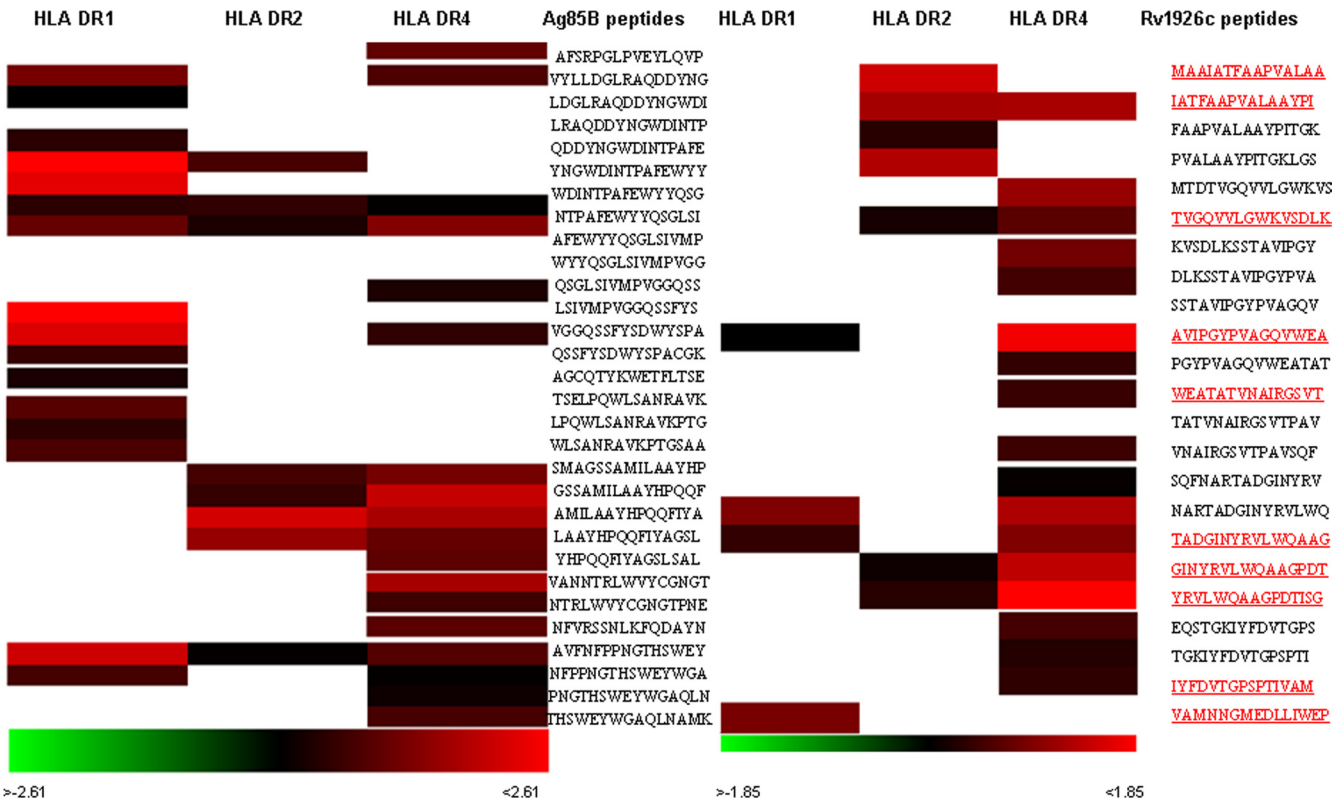

FIG. 3. Locations of peptide binding to MHC class II molecules for two selected epitope-rich proteins. The peptides are arranged from the N to the C terminus of the proteins Ag85B (left) and MPT63 (Rv1926c; right), and only positive binding results are shown. These two *M. tuberculosis* proteins have previously been explored for HLA class II binding peptides (3, 14, 26, 27). Peptides that have been previously described as $CD4^+$ T-cell epitopes, defined in either functional assays or MHC class II binding or by MHC class II tetramers, are underlined and shown in red.

production. Not mutually exclusive, anti-*M. tuberculosis* responses may also be anergic in individuals with active pulmonary TB (33, 36). Tetramer-guided analysis may represent a remedy to this problem. However, the fact that some of the identified peptides have been reported previously (3, 26, 27) using functional assays or tetramer-guided analysis (14) supports the validity of the approach reported here. It is also important to note that some of the candidate test peptides might not bind to the HLA-DR alleles in vivo due to differential protein processing and subsequent presentation: peptide processing is dependent on the three-dimensional structure of proteins (18, 21, 22, 25). Future tetramer-guided analysis of PBMCs from patients with TB will aid in determining which peptides are presented in vivo and lead to expansion of antigen-specific $CD4^+$ T cells.

Of note, MHC class II-presented peptides may also drive immunosuppressive immune responses associated with $CD4^+$ regulatory T cells (Tregs). This may be considered in rational vaccine design, since Tregs have been associated with the suppression of Th1-type immune responses in *M. tuberculosis* infection (13). In addition, instability of the transcription factor Foxp3 may lead to the generation of antigen-specific memory T cells with altered effector properties. This has recently been shown to be true for Tregs with an activated-memory T-cell phenotype, which gave rise to potentially autoreactive effector T cells (37).

We hypothesize that most of the *M. tuberculosis* peptides that formed stable complexes with HLA-DRB1*0101, -DRB1*1502, and -DRB1*0401 may serve as $CD4^+$ T-cell epitopes and that these peptides could be useful in designing a rational epitope-based vaccine against *M. tuberculosis*. We suggest that the integrated analysis of IgG-recognized targets from clinically very well-characterized patient cohorts will help make the best choice for *M. tuberculosis* vaccine targets. (Peptide) antigen-specific B cells may serve as professional antigen-presenting cells (19) for $CD4^+$ T cells. Conversely, $CD4^+$ T cells may provide help for B cells and $CD8^+$ T cells. Therefore, *M. tuberculosis* vaccine target identification should be accompanied by MHC class I peptide binding analysis, since $CD8^+$ T cells are instrumental in conferring long-term immune memory in TB (4), particularly in patients with HIV coinfection and decreased $CD4^+$ T-cell numbers.

ACKNOWLEDGMENTS

The work was supported in part by an EU Marie Curie Fellowship to S.G. and by grants from Vetenskapsrådet and SIDA, Sweden, to M.M.

REFERENCES

1. **Amirzargar, A. A., A. Yalda, M. Hajabolbaghi, F. Khosravi, H. Jabbari, N. Rezaei, M. H. Niknam, B. Ansari, B. Moradi, and B. Nikbin.** 2004. The association of HLA-DRB, DQA1, DQB1 alleles and haplotype frequency in Iranian patients with pulmonary tuberculosis. Int. J. Tuberc. Lung Dis. **8:**1017–1021.
2. **Beveridge, N. E., D. A. Price, J. P. Casazza, A. A. Pathan, C. R. Sander, T. E. Asher, D. R. Ambrozak, M. L. Precopio, P. Scheinberg, N. C. Alder, M.**

Roederer, R. A. Koup, D. C. Douek, A. V. Hill, and H. McShane. 2007. Immunisation with BCG and recombinant MVA85A induces long-lasting, polyfunctional Mycobacterium tuberculosis-specific CD4+ memory T lymphocyte populations. Eur. J. Immunol. **37:**3089–3100.

3. **Blythe, M. J., Q. Zhang, K. Vaughan, R. de Castro, Jr., N. Salimi, H. H. Bui, D. M. Lewinsohn, J. D. Ernst, B. Peters, and A. Sette.** 2007. An analysis of the epitope knowledge related to mycobacteria. Immunome Res. **3:**10.
4. **Day, C. L., N. Mkhwanazi, S. Reddy, Z. Mncube, M. van der Stok, P. Klenerman, and B. D. Walker.** 2008. Detection of polyfunctional Mycobacterium tuberculosis-specific T cells and association with viral load in HIV-1-infected persons. J. Infect. Dis. **197:**990–999.
5. **De Groot, A. S., J. McMurry, L. Marcon, J. Franco, D. Rivera, M. Kutzler, D. Weiner, and B. Martin.** 2005. Developing an epitope-driven tuberculosis (TB) vaccine. Vaccine **23:**2121–2131.
6. **D'Souza, S., M. Romano, J. Korf, X. M. Wang, P. Y. Adnet, and K. Huygen.** 2006. Partial reconstitution of the CD4+-T-cell compartment in CD4 gene knockout mice restores responses to tuberculosis DNA vaccines. Infect. Immun. **74:**2751–2759.
7. **Endsley, J. J., A. Hogg, L. J. Shell, M. McAulay, T. Coffey, C. Howard, C. F. Capinos Scherer, W. R. Waters, B. Nonnecke, D. M. Estes, and B. Villarreal-Ramos.** 2007. Mycobacterium bovis BCG vaccination induces memory CD4+ T cells characterized by effector biomarker expression and anti-mycobacterial activity. Vaccine **25:**8384–8394.
8. **Frank, R.** 2002. The SPOT-synthesis technique. Synthetic peptide arrays on membrane supports—principles and applications. J. Immunol. Methods **267:** 13–26.
9. **Gaseitsiwe, S., D. Valentini, R. Ahmed, S. Mahdavifar, I. Magalhaes, J. Zerweck, M. Schutkowski, E. Gautherot, F. Montero, A. Ehrnst, M. Reilly, and M. Maeurer.** 2009. Major histocompatibility complex class II molecule-human immunodeficiency virus peptide analysis using a microarray chip. Clin. Vaccine Immunol. **16:**567–573.
10. **Gaseitsiwe, S., D. Valentini, S. Mahdavifar, I. Magalhaes, D. F. Hoft, J. Zerweck, M. Schutkowski, J. Andersson, M. Reilly, and M. J. Maeurer.** 2008. Pattern recognition in pulmonary tuberculosis defined by high content peptide microarray chip analysis representing 61 proteins from M. tuberculosis. PLoS ONE **3:**e3840.
11. **Goldfeld, A. E., J. C. Delgado, S. Thim, M. V. Bozon, A. M. Uglialoro, D. Turbay, C. Cohen, and E. J. Yunis.** 1998. Association of an HLA-DQ allele with clinical tuberculosis. JAMA **279:**226–228.
12. **Goletti, D., O. Butera, F. Bizzoni, R. Casetti, E. Girardi, and F. Poccia.** 2006. Region of difference 1 antigen-specific CD4+ memory T cells correlate with a favorable outcome of tuberculosis. J. Infect. Dis. **194:**984–992.
13. **Guyot-Revol, V., J. A. Innes, S. Hackforth, T. Hinks, and A. Lalvani.** 2006. Regulatory T cells are expanded in blood and disease sites in patients with tuberculosis. Am. J. Respir. Crit. Care Med. **173:**803–810.
14. **Hohn, H., C. Kortsik, I. Zehbe, W. E. Hitzler, K. Kayser, K. Freitag, C. Neukirch, P. Andersen, T. M. Doherty, and M. Maeurer.** 2007. MHC class II tetramer guided detection of Mycobacterium tuberculosis-specific CD4+ T cells in peripheral blood from patients with pulmonary tuberculosis. Scand. J. Immunol. **65:**467–478.
15. **Kettaneh, A., L. Seng, K. P. Tiev, C. Toledano, B. Fabre, and J. Cabane.** 2006. Human leukocyte antigens and susceptibility to tuberculosis: a meta-analysis of case-control studies. Int. J. Tuberc. Lung Dis. **10:**717–725.
16. **Kim, H. S., M. H. Park, E. Y. Song, H. Park, S. Y. Kwon, S. K. Han, and Y. S. Shim.** 2005. Association of HLA-DR and HLA-DQ genes with susceptibility to pulmonary tuberculosis in Koreans: preliminary evidence of associations with drug resistance, disease severity, and disease recurrence. Hum. Immunol. **66:**1074–1081.
17. **Kobayashi, H., M. Wood, Y. Song, E. Appella, and E. Celis.** 2000. Defining promiscuous MHC class II helper T-cell epitopes for the HER2/neu tumor antigen. Cancer Res. **60:**5228–5236.
18. **Landry, S. J.** 2008. Three-dimensional structure determines the pattern of CD4+ T-cell epitope dominance in influenza virus hemagglutinin. J. Virol. **82:**1238–1248.
19. **Lanzavecchia, A.** 1996. Mechanisms of antigen uptake for presentation. Curr. Opin. Immunol. **8:**348–354.
20. **Lee, S. Y., and D. Jeoung.** 2007. The reverse proteomics for identification of tumor antigens. J. Microbiol. Biotechnol. **17:**879–890.
21. **Li, H., P. C. Chien, Jr., M. Tuen, M. L. Visciano, S. Cohen, S. Blais, C. F. Xu, H. T. Zhang, and C. E. Hioe.** 2008. Identification of an N-linked glycosylation in the C4 region of HIV-1 envelope gp120 that is critical for recognition of neighboring CD4 T cell epitopes. J. Immunol. **180:**4011–4021.
22. **Li, H., C. F. Xu, S. Blais, Q. Wan, H. T. Zhang, S. J. Landry, and C. E. Hioe.** 2009. Proximal glycans outside of the epitopes regulate the presentation of HIV-1 envelope gp120 helper epitopes. J. Immunol. **182:**6369–6378.
23. **Lombard, Z., D. L. Dalton, P. A. Venter, R. C. Williams, and L. Bornman.** 2006. Association of HLA-DR, -DQ, and vitamin D receptor alleles and haplotypes with tuberculosis in the Venda of South Africa. Hum. Immunol. **67:**643–654.
24. **Middleton, D., L. Menchaca, H. Rood, and R. Komerofsky.** 2003. New allele frequency database. Tissue Antigens **61:**403–407.
25. **Mirano-Bascos, D., M. Tary-Lehmann, and S. J. Landry.** 2008. Antigen structure influences helper T-cell epitope dominance in the human immune response to HIV envelope glycoprotein gp120. Eur. J. Immunol. **38:**1231–1237.
26. **Mustafa, A. S.** 2009. Th1 cell reactivity and HLA-DR binding prediction for promiscuous recognition of MPT63 (Rv1926c), a major secreted protein of Mycobacterium tuberculosis. Scand. J. Immunol. **69:**213–222.
27. **Mustafa, A. S., A. T. Abal, F. Shaban, A. M. El-Shamy, and H. A. Amoudy.** 2005. HLA-DR binding prediction and experimental evaluation of T-cell epitopes of mycolyl transferase 85B (Ag85B), a major secreted antigen of Mycobacterium tuberculosis. Med. Princ. Pract. **14:**140–146.
28. **Nahtman, T., A. Jernberg, S. Mahdavifar, J. Zerweck, M. Schutkowski, M. Maeurer, and M. Reilly.** 2007. Validation of peptide epitope microarray experiments and extraction of quality data. J. Immunol. Methods **328:**1–13.
29. **Novak, E. J., A. W. Liu, G. T. Nepom, and W. W. Kwok.** 1999. MHC class II tetramers identify peptide-specific human CD4(+) T cells proliferating in response to influenza A antigen. J. Clin. Invest. **104:**R63–R67.
30. **Panina-Bordignon, P., A. Tan, A. Termijtelen, S. Demotz, G. Corradin, and A. Lanzavecchia.** 1989. Universally immunogenic T cell epitopes: promiscuous binding to human MHC class II and promiscuous recognition by T cells. Eur. J. Immunol. **19:**2237–2242.
31. **Ruggiero, G., E. Cosentini, D. Zanzi, V. Sanna, G. Terrazzano, G. Matarese, A. Sanduzzi, F. Perna, and S. Zappacosta.** 2004. Allelic distribution of human leucocyte antigen in historical and recently diagnosed tuberculosis patients in Southern Italy. Immunology **111:**318–322.
32. **Scharn, D., H. Wenschuh, U. Reineke, J. Schneider-Mergener, and L. Germeroth.** 2000. Spatially addressed synthesis of amino- and amino-oxy-substituted 1,3,5-triazine arrays on polymeric membranes. J. Comb. Chem. **2:**361–369.
33. **Seitzer, U., K. Kayser, H. Hohn, P. Entzian, H. H. Wacker, S. Ploetz, H. D. Flad, J. Gerdes, and M. J. Maeurer.** 2001. Reduced T-cell receptor CD3zeta-chain protein and sustained CD3epsilon expression at the site of mycobacterial infection. Immunology **104:**269–277.
34. **Skeiky, Y. A., P. J. Ovendale, S. Jen, M. R. Alderson, D. C. Dillon, S. Smith, C. B. Wilson, I. M. Orme, S. G. Reed, and A. Campos-Neto.** 2000. T cell expression cloning of a Mycobacterium tuberculosis gene encoding a protective antigen associated with the early control of infection. J. Immunol. **165:**7140–7149.
35. **Teran-Escandon, D., L. Teran-Ortiz, A. Camarena-Olvera, G. Gonzalez-Avila, M. A. Vaca-Marin, J. Granados, and M. Selman.** 1999. Human leukocyte antigen-associated susceptibility to pulmonary tuberculosis: molecular analysis of class II alleles by DNA amplification and oligonucleotide hybridization in Mexican patients. Chest **115:**428–433.
36. **Weichold, F. F., S. Mueller, C. Kortsik, W. E. Hitzler, M. J. Wulf, D. M. Hone, J. C. Sadoff, and M. J. Maeurer.** 2007. Impact of MHC class I alleles on the M. tuberculosis antigen-specific CD8+ T-cell response in patients with pulmonary tuberculosis. Genes Immun. **8:**334–343.
37. **Zhou, X., S. L. Bailey-Bucktrout, L. T. Jeker, C. Penaranda, M. Martinez-Llordella, M. Ashby, M. Nakayama, W. Rosenthal, and J. A. Bluestone.** 2009. Instability of the transcription factor Foxp3 leads to the generation of pathogenic memory T cells in vivo. Nat. Immunol. **10:**1000–1007.