A conserved human T cell population targets mycobacterial antigens presented by CD1b

Ildiko Van Rhijn1,2, Anne Kasmar1, Annemieke de Jong1, Stephanie Gras3, Mugdha Bhati3, Marieke E Doorenspleet4, Niek de Vries4, Dale I Godfrey5, John D Altman6, Wilco de Jager7, Jamie Rossjohn3,8 & D Branch Moody1

Human T cell antigen receptors (TCRs) pair in millions of combinations to create complex and unique T cell repertoires for each person. Through the use of tetramers to analyze TCRs reactive to the antigen-presenting molecule CD1b, we detected T cells with highly stereotyped TCR α-chains present among genetically unrelated patients with tuberculosis. The germline-encoded, mycolyl lipid–reactive (GEM) TCRs had an α-chain bearing the variable (V) region TRAV1-2 rearranged to the joining (J) region TRAJ9 with few nontemplated (N)-region additions. Analysis of TCRs by high-throughput sequencing, binding and crystallography showed linkage of TCRα sequence motifs to high-affinity recognition of antigen. Thus, the CD1-reactive TCR repertoire is composed of at least two compartments: high-affinity GEM TCRs, and more-diverse TCRs with low affinity for CD1b-lipid complexes. We found high interdonor conservation of TCRs that probably resulted from selection by a nonpolymorphic antigen-presenting molecule and an immunodominant antigen.

Studies of T cells during vaccination, autoimmunity and infection rely on the model that highly diverse T cell antigen receptors (TCRs) recognize peptides bound to major histocompatibility complex (MHC) proteins. However, lipids presented by CD1 proteins expand the biochemical range of antigens for T cells1. T cells recognize foreign lipid antigens from major pathogens such as Mycobacterium tuberculosis, which indicates that the recognition of microbial lipids occurs during the natural history of infectious disease. Also, because CD1 and MHC proteins differ greatly in their rate of polymorphism in human populations, comparison of MHC and CD1-reactive TCRs provides an experimental system with which to study the role of host genetics in shaping human TCR repertoires. At present, understanding of the complexity of the overall T cell repertoire is dominated by knowledge of MHC-restricted T cells, which form highly complex patterns based on the pairing of rearranged TCR α- and β-chains that generate a unique TCR repertoire for each person. This diversity is generated through variable (V), diversity (D) and joining (J) segments, deletions and nontemplated (N)-region additions. MHC class I and II proteins show corresponding complexity, as they are highly polymorphic in human populations and any given protein sequence encoded by the HLA-A, HLA-B, HLA-C, HLA-DR, HLA-DP and HLA-DQ loci can bind many peptides.

In contrast, CD1 polymorphisms are rare and are usually functionally silent2. Thus, genetic polymorphism, which represents one major driver of the diversity of the MHC-reactive TCR repertoire, is largely absent from the human CD1 system. To some extent, the nonpolymorphic nature of CD1 antigen-presenting molecules results in an apparently simplified TCR repertoire. The main example of this is invariant natural killer T cells (iNKT cells), which express the TCR α-chain V region TRAV10 (Vα24) and α-chain J region TRAJ18 (Jα18) and are found in most human donors3,4. However, most CD1-reactive T cells do not express that TCR. Other CD1d-reactive T cells (diverse NKT cells)5,6 and all known T cells that recognize CD1a, CD1b or CD1c use different TCR V, D or J segments7–10. Accordingly, ‘consensus’ classification of the CD1 repertoire has categorized those that recognize three T cell types: iNKT cells, diverse NKT cells, and diverse T cells that recognize CD1a, CD1b and CD1c11. CD1a, CD1b and CD1c, known as the ‘group 1 CD1 proteins’, are broadly retained during mammalian evolution12, and they activate large numbers of T cells present in the blood and tissues of humans8,9. However, information about their biology and TCR diversity is extremely limited because no specific surface markers are known and efficient methods for capturing clones for TCR sequencing have not been developed. Taking advantage of CD1b tetramers as a new tool with which to study T cells in larger numbers, we identified highly stereotyped TCRs composed of nearly identical TCR α-chains. Those TCR amino acid sequences were derived from different nucleotide sequences, which indicated that conserved TCRs arose through distinct rearrangement events in individual donors. Furthermore, such TCRs were expressed on cell populations that were readily detectable.

1Division of Rheumatology, Immunology and Allergy, Brigham and Women’s Hospital, Harvard Medical School, Boston, Massachusetts, USA. 2Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands. 3Department of Clinical Immunology & Rheumatology, Academic Medical Center, University of Amsterdam, The Netherlands. 4Department of Microbiology & Immunology, University of Melbourne, Parkville, Australia. 5NIH Tetramer Core Facility, Emory Vaccine Center, Atlanta, Georgia, USA. 6Department of Pediatric Immunology, University Medical Centre Utrecht, Utrecht, The Netherlands. 7Institute of Infection and Immunity, Cardiff University School of Medicine, Cardiff, UK. Correspondence should be addressed to D.B.M. (bmoody@rics.bwh.harvard.edu) or I.V.R. (i.vanhijn@uu.nl).

Received 6 February; accepted 29 April; published online 2 June 2013; doi:10.1038/ni.2630
ex vivo among genetically unrelated donors. Our findings identify a previously unknown type of TCR-conserved αβ T cell type and point to interdonor conservation of CD1b-reactive T cell responses.

RESULTS

T cells with high and low avidity for CD1b

To study lipid-reactive T cells in a disease setting, we used lipid extracts of *M. tuberculosis* to activate T cells from patients with tuberculosis (Supplementary Fig. 1). By screening a panel of clones for responses to mycobacterial lipids and CD1b, we detected clone 18 (Fig. 1a). We identified the stimulating antigen as free mycolic acid (Fig. 1b), an essential long-chain mycobacterial lipid (Supplementary Fig. 2). Comparing the TCR α- and β-chains of clone 18 with five previously known TCRs that recognize mycolyl lipids, we did not observe conservation of TCR sequence or V- or J-segment use1,7,13–16 (Table 1). Thus, initial results based on conventional cloning methods supported the widespread view that TCRs in the group 1 CD1 system are diverse, similar to those of MHC-restricted T cells7,8,10,17 and different from NKT cells and mucosa-associated invariant T cells (MAIT cells), which express germline-encoded TCRs17,18.

However, existing data7,8,10 might inadequately represent the CD1b repertoire, given the small number of clones obtained so far and the need for extensive culture of such clones to expand them. To bypass such problems and analyze T cells in the ex vivo setting among patients immunized by natural infection, we generated CD1b tetramers19 and loaded them with glucose monomycolate (GMM; Supplementary Fig. 2). This mycobacterial glycolipid antigen is produced *in vivo* during infection20 and potently activates CD1b-reactive T cells in humans and cows15,19,21,22. We sorted tetramer-positive (Tet+) T cells from three patients with tuberculosis, cloned them at limiting dilution and screened them for GMM-dependent functional responses. Although clone 18 and other T cell clones of the group 1 CD1 system were collected over many years (Table 1), binding of tetramer to CD1b-reactive TCRs resulted in efficient enrichment for the relevant CD1b-reactive T cells and yielded multiple mycolyl lipid-reactive TCRs for all three patient samples (Fig. 2a). Thus, tetramer sorting was an efficient method with which to sample the natural CD1b-GMM-reactive T cell repertoire *in vivo* and demonstrated the existence of mycolyl lipid–reactive clones among genetically unrelated donors. In all cases, clones that produced interferon-γ (IFN-γ) after stimulation with the antigen GMM also stained with GMM-loaded CD1b tetramers (Fig. 2). On the basis of staining intensity, the analyzed clones segregated into two groups detected with intermediate staining (Tetmid) or high staining (Tethi). Tetmid clones (clones 2, 26, 34 and 71) had diverse patterns of expression of the monomorphic coreceptors CD4 and CD8, similar to that of clones generated by conventional methods (Table 1). In contrast, clone 18 and all Tethi clones (clones 1, 21 and 42) expressed CD4 but not CD8 (Figs. 2a and 3a). We isolated both Tetmid and Tethi cells from each patient, which suggested that the different avidity patterns were not patient specific but instead existed as distinct cell types present side by side in individual patients.

Antimicrobial effector functions of GMM-specific T cells

To gain insight into the helper or effector functions of CD1b-reactive T cells, we used a multiplex cytokine array to profile cytokines produced by clones that recognized CD1b (clones 1, 2, 26, 34 and 42) or control clones generated in the same way (clones 101 and 50; Fig. 2b). Our array design emphasized cytokines with defining roles in the helper T cell subsets T HEL1, T HEL17 and T HEL2, such as IFN-γ, interleukin 17 (IL-17) and IL-13. We also assessed whether the T cells produced two cytokines known to be necessary for effective human antitubercular responses *in vivo*: tumor-necrosis factor (TNF) and IFN-γ23,24. In all cases, T cells cultured for a short term produced TNF and IFN-γ, whereas we did not consistently detect helper cytokines such as IL-2 and IL-17; this suggested an antimicrobial function for GMM-specific T cells23,24. As observed before with clones cultured long term *in vitro*25,26, these *ex vivo* studies confirmed that IFN-γ and TNF were reliably produced.

Expression of invariant TCR α-chains

Similar to clones generated by conventional methods (Table 1), Tetmid clones expressed different TCR α-chains. In contrast, all Tethi clones expressed TCRs that were highly similar to each other and the TCR expressed by clone 18 (Fig. 3a). The TCR α-chains of the Tethi clones were identical in length and used the same variable (TRAV1-2) and joining (TRAJ9) segments, with few N-region additions, which yielded the complementarity-determining region 3 (CDR3) consensus sequence CAVRNTGGFKTIF (Fig. 3a). Because those similar rearrangements yielded largely germline-encoded TCR α-chains and mediated the recognition of mycolyl lipids, we designated TCRs with such α-chains ‘germline-encoded, mycolyl lipid–reactive’ (GEM) TCRs and the T cells that express GEM TCRs ‘GEM T cells’. Although they lacked strict sequence conservation, TCR β-chains were apparently biased in V-region (TRBV) use. TRBV6-2 and TRBV30 are expressed in 3.6% and 1% percent of human T cells27,28, respectively, yet they were present in all of the original set of clones expressing GEM TCRs (Figs. 2a and 3). In addition to results obtained with those clones from patients A22, C58, C52 and A14, single-cell sequencing of Tet+ cells from patients C40, C52, C58 yielded three TRBV6-2 sequences but no TRBV30 sequences (Supplementary Fig. 3).

Table 1 TCRs of mycolyl lipid–specific T cell clones

| Clone | TRAV | TRAJ | TRBV | TRBJ | Antigen |
|-------|------|------|------|------|---------|
| 18    | CD4+ | 1-2  | 9    | 6-2  | 2-2     | MA      |
| LDN5  | CD4−| 1-2  | 9    | 4-1  | 2-1     | GMM     |
| DN1   | CD4−| 1-2  | 9    | 5-1  | 2-1     | MA      |
| DN.ppt| CD4−| 3    | 31   | 7-3  | 2-1     | MA      |
| Z5B71 | CD4−| 35   | 52   | 3-1  | 2-5     | GMM     |

V- and J-segment use and coreceptor expression (Phenotype) of clone 18 and other, published CD1b-restricted, mycolyl lipid–specific T cell clones7,18. Glycol MM, glycercol monomycolate.
Thus, GEM TCRs were characterized by nearly invariant TCR α-chains and a bias in the TCR β-chain V region, a pattern similar to that of TCRs of iNKT cells and MAIT cells (Fig. 3b).

High-affinity binding of CD1b-GMM by GEM TCRs

Because we noted strict conservation of the TCR α-chain only among Tetα clones, we hypothesized that high-affinity binding of GEM TCRs to CD1b-GMM mediated the expansion of populations of T cells expressing GEM TCRs\(^\text{19}\). Binding of any TCR to CD1b has not been measured before, to our knowledge, so we developed a plasmon resonance assay to analyze the binding of disulfide-linked GEM TCRs with truncation of the transmembrane region\(^\text{29}\) to CD1b-GMM complexes. Each of the three GEM TCRs had a dissociation constant of nearly ~1 µM (Fig. 4). Their dissociation constants were at the high end of the range for TCRs that recognize natural peptide or lipid antigens and were similar to measurements of the binding of human NKT cell TCRs to the superagonist α-galactosylceramide in complex with CD1d\(^\text{30}\). Thus, the bright staining by CD1b tetramers for clones expressing GEM TCRs (Fig. 2a) was explained by high-affinity ternary interactions of αβ TCR heterodimers with CD1b-GMM complexes.

GEM TCRs mediate antigen recognition

Next we determined the role of GEM TCRs in the activation of T cells by transferring TCRs into the TCRβ-deficient human acute T cell leukemia cell line J.RT3-T3.5. Untransfected cells were not activated by antigens, whereas expression of native GMM-reactive TCRs that recognize natural peptide or lipid antigens and were similar to measurements of the binding of human NKT cell TCRs to the superagonist α-galactosylceramide in complex with CD1d\(^\text{30}\). Thus, the bright staining by CD1b tetramers for clones expressing GEM TCRs (Fig. 2a) was explained by high-affinity ternary interactions of αβ TCR heterodimers with CD1b-GMM complexes.

From the mycolic acid-specific TCR (clone 18) reconstituted recognition of mycolic acid but not of GMM in all cases (Fig. 5b). Thus, the TCR β-chain determined the fine specificity for two structurally related antigens.

Such functional dominance of the β-chain was unexpected, given that TCR α-chains were more conserved in structure among the GEM T cells. A converse set of TCR chain swaps, in which we paired the α-chain of clone 18 with the β-chain from clone 1, 21 or 42, yielded three TCR chimeras that did not respond to mycolic acid or GMM (Fig. 5c). Given the retained response of the transfectants to anti-CD3, this result was probably not due to failure of transfection or signaling but instead resulted from lack of antigen recognition by the chimeric receptors. Thus, in contrast to the preserved recognition of mycolic acid by chimeric TCRs (Fig. 5b), the failure to recognize GMM in all chimeras suggested that the recognition of GMM was a more stringent interaction that depended on both TCR chains. All four native TCR α-chains (from clones 1, 21, 42 and 18) were similar in sequence, but that of clone 18 differed from that of clone 42 only at position 107, with substitution of leucine for arginine (Fig. 5a).

Thus, position 107 in the CDR3α sequence was crucial for determining recognition of mycolic acid.
Insights from GEM TCR structures

To address the structural basis of TCRα and TCRβ pairing and gain general insights into the molecular basis of the antigen specificity of GEM TCRs, we expressed TCRs from clones 18 and 42 and solved their crystal structures at a resolution of 2.2Å (Fig. 6a and Supplementary Fig. 4). By superimposing those two GEM TCRs with the TCR of a human MAIT cell31, we were able to compare three TCRs that shared TRAV1-2 use but recognized CD1b and mycolic acid, CD1b and GMM, or the MHC class I–related molecule MRM and ribitol-lumazine antigens32 differently (Fig. 6d). Despite their shared use of TRAV1-2, the MAIT cell TCR and GEM TCRs differed in other aspects of their β-chain and CDR3α conformations (Fig. 6d), which presumably mediated their non–cross-reactive recognition of CD1b and MRM (refs. 31,33). However, this comparison provided specific insights into GEM TCR pairing that agreed with the results obtained by TCR swapping. Although similar in structure, the two GEM TCRs had notable differences at the CDR3α-CDR3β interface, which derived from perturbations in the germline-encoded CDR1α loop that arose mainly from the residue at position 107 in TCRβ (Fig. 6). Whereas Arg107 in the CDR3β of the mycolic acid-specific clone 18 packed against its CDR3α loop, the corresponding interaction was missing from the GMM-reactive clone, which resulted in greater mobility in its CDR3 loop (Fig. 6d). That structural feature of the β-chain, which probably determined the reactivity to mycolic acid versus GMM, was located near the TCRα residue 107 identified in the TCR chain-swapping experiments (Fig. 6d). Thus, these structural data indicated a specific mechanism for the influence of the β-chain on fine specificity whereby the β-chain pushed laterally on the invariant α-chain at a defined point on the CDRα-CDRβ interface, analogous to the fine-tuning of NKT cell TCR specificity34.

Population expansion of GEM T cells in vivo

In vitro studies of clones identified key features of GEM T cells, such as high-affinity TRAV1-2 TCRs and production of antimicrobial cytokines in response to mycobacterial lipids. To determine if such T cells were present in vivo and shared those properties, we tested polyclonal T cells freshly isolated from blood-donor patients or patients with tuberculosis in three types of experiments (Supplementary Fig. 1). We first used a monoclonal antibody that recognizes TCR α-chains with TRAV1-2 sequences35 to sort freshly isolated peripheral blood mononuclear cells (PBMCs) into four populations on the basis of their expression of TRAV1-2 and CD4. We expected CD1b- and GMM-reactive cells to be in the TRAV1-2+CD4+ gate and most MAIT cells to be in the TRAV1-2−CD4+ gate18,33,35. We also expected both TRAV1-2+ populations to contain lymphocytes that were CD1b independent. TRAV1-2−CD4+ cells from a blood-donor (BB2) had antigen-independent responses that led to IFN-γ production (Fig. 7a), which might have represented recognition of target cells with low expression of MHC class I by NK cells. The TRAV1-2−CD4+ control populations from all donors failed to recognize GMM. For cells from two blood-donor donors (BB12 and BB2) and one patient with tuberculosis (CS8), enzyme-linked immunospot (ELISPOT) assay of IFN-γ production (Fig. 7b) showed that only T cells from BB12 and BB2 expressed TRAV1-2+CD4+ in vitro. Thus, GEM T cells were present in vivo when we assayed them by surface plasmon resonance measurement (Fig. 6a).

Figure 3 Conserved TCRs and CD4 define GEM T cells. (a) Nucleotide sequences (lower-case) and amino-acid sequences (upper-case) of CDR3 regions of CD1b-restricted GEM T cell clones (GEM T cell) and a non–GEM T cell clone that recognizes CD1b (LDN5), presented by origin (left margin), as well as representative sequences from an NKT cell clone (J3N.5) and a MAIT cell clone (TRBV20). Light gray, germline-encoded V segment–derived nucleotides; dark gray, germline-encoded J segment–derived nucleotides; boxed areas, non–germline-encoded nucleotides in the α-chain. Transcripts containing sequence encoded by TRBV6-2 or TRBV6-3 cannot be distinguished because the coding sequence of these two genes is identical; called ‘TRBV6-2’ here. (b) Human T cells with conserved TCRs.
Figure 5 Antigen specificity of TCR transfectants. (a) ELISPOT assay of IL-2 in J.RT3-T3.5 cells transfected to express the α- and β-chains of the native GEM TCR of clone 18, 21, 1 or 42 (above plots), stimulated with no antigen, GMM or mycolic acid, presented by monocyte-derived dendritic cells. Right: CDR3α sequences; bold indicates position 107. (b,c) ELISPOT assay of IL-2 in J.RT3-T3.5 cells transfected to express chimeric TCRs generated by combination (above graphs) of the TCR α-chain of clone 18 with the TCR β-chain of clone 21, 42 or 1 (b) or the TCR α-chain of clone 21, 42, or 1 with the TCR β-chain of clone 18 (c), stimulated with no antigen, GMM or mycolic acid or activating OKT3 antibody to CD3 (α-CD3); for cells refractory to antigen). Data are representative of two independent experiments (error bars, s.d. of triplicate wells).

showed that TRAV1-2*CD4+ populations were activated by GMM or mycolic acid in a CD1b-dependent manner (Fig. 7a). Thus, the in vitro analysis of clones and ex vivo analysis of polyclonal T cells both showed that expression of TRAV1-2 and CD4 was a marker for a subset of cells in the repertoire that recognized CD1b and mycolyl lipids (Fig. 7a and Supplementary Fig. 1).

Contrasts among GEM T cells, NKT cells and MAIT cells

Next we evaluated whether certain aspects of the GEM T cell response initially observed in vitro might also occur with more physiological antigen-presenting cells and might apply to the ex vivo state. Activation of GEM T cells, as observed for transformed cells transfected to express CD1b (Figs. 2 and 7a), also occurred in response to autologous monocyte-derived dendritic cells, which more closely resemble the main cell population to express CD1b in the periphery (Supplementary Fig. 5a). Similar to results obtained with clones cultured in vitro (Fig. 2b), the response of polyclonal GEM T cells from patient C58 to GMM led to the secretion of IFN-γ and TNF when measured immediately after isolation of the cells (Supplementary Fig. 5b). Like the clones, freshly isolated TRAV1-2*CD4+ T cell populations showed detectable reactivity to both mycolic acid and GMM, but they produced more IFN-γ spots in response to GMM than in response to MA (Figs. 1 and 7a), which suggested that GMM was the immunodominant antigen.

Ex vivo analysis of GEM T cells from patients with tuberculosis showed the differences among those T cells, NKT cells and MAIT cells. NKT cells were discovered and named in part on the basis of their expression of CD161 and other NK cell markers 11, and showed the differences among those T cells, NKT cells and MAIT cells 18. Furthermore, we distinguished those T cells from most TRAV1-2*CD4+ MAIT cells 18. CD69, CD161 or PD-1 did not specifically mark GEM T cell populations, and CD161 expression, with cells from two donors expressing this marker at low frequency, similar to that of total CD3+ cells (Supplementary Fig. 6a). NKT cells circulate in a preactivated state and typically express the activation marker CD69 (ref. 36). Prolonged activation of T cells by antigens might additionally lead to expression of PD-1, a marker of T cell exhaustion 37. However, CD69 and PD-1 were not expressed above background expression by Tethi T cells from patients with tuberculosis (Supplementary Fig. 6b). Overall, CD4 was a defining marker of GEM T cells that further distinguished those T cells from most TRAV1-2*CD4+ MAIT cells 18. CD69, CD161 or PD-1 did not specifically mark GEM T cell populations ex vivo.

TCR sequence conservation is linked to high avidity

CD1b-reactive clones with diverse TRAV1-2 TCRs or conserved TRAV1-2 TCRs segregated into Tethi and Tein cells, respectively. Both clone types were present in each patient tested. This pattern suggested that GEM T cells and CD1b-restricted T cells of lower affinity might exist in the same repertoire of any patient. Polyclonal T cell analysis of patient C58 yielded a broad range of tetramer-staining intensity that extended from a low to a very high absolute signal (Fig. 7b). Furthermore, we detected that pattern of broadly ranging tetramer-staining intensity among cells from other patients with tuberculosis (patients C12, C32, C60 A14, A21 and A22; Supplementary Fig. 6c). Next we directly analyzed the relationship between tetramer-staining intensity and TRAV1-2 expression. We analyzed the frequency of TRAV1-2+ T cells.

Figure 6 GEM TCR structures. (a–c) Structures of the TCRs of clone 18 (a), clone 42 (b) and a human MAIT clone 31 (c), each with a surface representation of the antigen-binding face of the TCR: dashed lines, residues missing from Protein Data Bank coordinates; pink, TRAV1-2 α-chains; gray, TRBV6-2 β-chains; green, TRBV20 β-chains. (d) Overlay of the CDR1α–CDR3α and CDR1β–CDR3β loops of GEM TCRs of clone 18 (purple) and clone 42 (blue) with the MAIT TCR (green), presented over the Vα regions of the three TCRs.
among fresh CD3+ T cells in Tetneg, Tetint and Tethi populations. Whereas anti-TRA1-2 stained 1.8% of T cells and 3.6% of Tetint cells, it stained 71% of Tethi cells (Fig. 7b). Thus, this analysis of polyclonal T cells demonstrated linkage of TRA1-2 expression with high-avidity binding to CD1b-GMM and indicated that GEM TCRs can dominate the high-avidity repertoire.

Clonal expansion of GEM T cells in vivo

Next we sought to detect GEM T cells in a tetramer-independent manner by PCR. TRA1-2 was joined to TRAJ9 in all GEM T cells, but TRA1-2 was expressed in concert with other J regions in TCRs of M. tuberculosis (C34, C28, C58 and C52) and blood-bank donors (BB38 and BB36), presented relative to the abundance of the α-chain constant region. Data are representative of two (BB12 and BB2) or three (C58) experiments with similar results (a; error bars, s.d. of triplicate wells), two experiments with similar results (b), one experiment (c, left) or two experiments per donor (c, right) or two experiments (d).

GEM T cells in the naive repertoire

Mycobacterial lipids are foreign mycobacterial antigens that drive the population expansion of antigen-specific T cells during infection16,21,22,38,39. After detecting GEM T cells in patients with tuberculosis19, we sought to determine if such T cells also existed in the naive repertoire. Sorted TRA1-2+CD4+ T cells from two blood-bank donors (BB2 and BB12) yielded detectable GMM- and CD1b-dependent responses (Fig. 7a and Supplementary Fig. 1). Given that initial result, we quantitatively analyzed GEM-defining TCR sequences by real-time PCR and high-throughput sequencing. Real-time PCR analysis of TRA1-2+CD4+ T cells sorted from two blood-bank donors (BB36 and BB38) and four patients with tuberculosis (C28, C34, C52, C58) detected the TRA1-2–TRAJ9 junction in all samples, but the signals were much higher in cells from the patients with tuberculosis (Fig. 7d). Those findings were consistent with the presence of a small GEM T cell population in the naive repertoire and expansion after infection with M. tuberculosis. However, not every possible TRA1-2–TRAJ9 junction corresponds to a GEM TCR sequence. To determine if the candidate GEM TCR α-chains detected in cells from blood-bank donors actually met the length and sequence criteria for GEM TCRs, we did deep sequencing of the TCR α-chains of the sorted TRA1-2+CD4+ T cells from donors BB36 and BB38. This showed that among 7,282 (BB36) and 6,861 (BB38) productive TRA1-2 sequences analyzed, 0.48% and 0.90% of the TRA1-2+CD4+ T cells (or 0.0024% and 0.0023% of the total PBMC population, respectively) had the sequence CAVRXTGGKFTIF or CAVLXGKTGKFTIF (where X is any amino acid; Fig. 7d and Supplementary Fig. 7). Eight clones identified used limited N-region additions. Although the deep sequencing did not provide information on the TCR β-chain, analysis of the TRBV sequences proved that T cells bearing the defining GEM TCRα motif were present at low frequencies in blood from donors.

DISCUSSION

Our data have identified GEM T cells as one of three known cell types with conserved TCRs in humans. GEM T cells, iNKT cells and MAIT...
cells showed strict conservation of the TCRα sequence with biased selection of TCR β-chains. Furthermore, GEM T cells and MAIT cells were defined by the same TCR α-chain V region. TRAV1-2 is located at the most upstream position in the TRA-TRD locus, which suggests possibly shared mechanisms of TCR rearrangement or binding of target to CD1b and MR1. However, distinct differences in TCR structure and phenotype distinguished these three cell types. Such data indicated that GEM T cells differed from MAIT cells in their CD4 expression, J-segment use, antigen-binding surfaces and antigen specificity14,15. Relative to iNKT cells, GEM T cells had apparently lower baseline autoreactivity, lower precursor frequencies in the blood and low CD69 expression, which suggested that they would be more dependent on infection to drive their expansion as populations4. CD1b and CD1d differ in expression in the periphery, with CD1b expression restricted mainly to dendritic cells, in contrast to the broad expression of CD1d on gastrointestinal epithelia, macrophages and B cells. Whereas iNKT cells are preactivated by self antigens and have a patrolling function mediated by autoreactivity, the lower autoreactivity, lower precursor frequency and low CD69 expression of GEM T cells suggest a role for such cells in the host response to infection.

We propose a model of the CD1b repertoire with at least two compartments: high-affinity, conserved TCRs, and low-affinity, diverse TCRs. Because recognition of antigen with high affinity is linked to invariant TCR sequences, multiple TRAV1-2–TRAJ9 rearrangements create TCRs that probably drive the population expansion of GEM T cells through high-affinity interactions with CD1b-lipid. The amino-acid and nucleotide patterns suggested two mechanisms of such expansion. First, identical nucleotide sequences encoding TCR α- and β-chains were derived from different blood samples or culture wells from the same donors, which suggested that the GEM TCR-expressing T cells underwent clonal expansion in vivo before phlebotomy. Second, most sequences were distinct in ways that proved that GEM T cell populations in the blood were derived from many independent TCR-arrangement events. For example, three rearrangements with different underlying nucleotide sequences gave rise to identical CDR3α amino-acid sequences in donor BB36. Also, highly similar CDR3α amino-acid sequences in clones 1, 18, 21 and 42 were derived from different patterns of exonucleolytic trimming of the V and J regions. We did not observe autoreactivity to CD1b, and any self antigen that could support the thymic selection of GEM T cells is unknown. Therefore, the basis for the emergence of GEM T cells among donors without known mycobacterial infection remains unclear. However, it is possible that environmental or self antigens cause the population expansion of GEM T cells in a way similar to that of MHC-restricted, human immunodeficiency virus–specific T cells in unexposed donors.

The present views emphasizing the diversity of group 1 CD1-reactive TCRs are based on small panels of T cell clones2-10. We initially viewed clone 18 as further evidence of TCR diversity in the CD1b repertoire, but it was actually the prototype of a larger T cell population later identified in many unrelated donors. Given the larger number of clones and polyclonal cells analyze here and the nonpolyclonal nature of human CD1 molecules, we predict that interdonor TCR conservation in the CD1b repertoire is more common than is appreciated at present. Our hypothesis can be addressed through the analysis of many clones with newly generated tetramers of human CD1a, CD1b19 and CD1c42.

Interdonor CDR3 diversity is the dominant pattern for MHC-restricted TCRs, with the notable exception of ‘public’ TCRs—that is, TCRs shared by donors that share an MHC-encoding allele and specific antigenic exposure43. The GEM TCR could be considered a kind of ‘public’ TCR that recognizes CD1b. CD1b is expressed in an identical form in essentially all humans, which is not the case for even the most common MHC-encoding allele. In the MHC system, ‘public’ TCRs are proposed to arise via convergent recombination, a process whereby multiple recombination events with few N-nucleotide additions produce different nucleotide sequences that encode the same amino acid sequence44. GEM TCR α-chain sequences had such hallmarks of ‘public’ TCRs. Interdonor TCR conservation might be of practical use in tracking of CD1-restricted T cells or possibly TCR-based ‘immunodiagnosis’. In the MHC system, particular TCR sequences cannot normally be used to predict the antigens recognized by genetically unrelated donors. Even the conserved iNKT cell TCR does not correspond to any single antigen of known biological function. Here we found that the GEM TCR α-chain sequence was associated with the recognition of lipids produced by a defined genus of pathogenic bacteria. Thus, CD1b tetramers, PCR primers or monoclonal antibodies to the GEM TCR α-chain might now be evaluated for use as TCR-based tests for tuberculosis. GEM T cells are a plausible target for population expansion by vaccination with mycobacterial glycolipid antigens.

**METHODS**  
Methods and any associated references are available in the online version of the paper.

**Accession codes.** Protein Data Bank: clone 18 TCR, 4G8E; clone 42 TCR, 4G8F; GenBank: full-length TCR α- and β-chains of clones 1, 18, 21 and 42, JQ778257, JQ778258, JQ778259, JQ778260, JQ778261, JQ778262, JQ778263 and JQ778264.

Note: Supplementary information is available in the online version of the paper.

**ACKNOWLEDGMENTS**  
We thank the Tetramer Core Facility of the US National Institutes of Health for CD1b protein; M. Turner, C. Seshadri and the Shattuck Hospital for clinical collaboration; L. Tan for technical help; and the beamline staff at the Australian synchrotron for assistance with data collection. Supported by the US National Institute of Allergy and Infectious Diseases (AI04393, AR048632 to D.B.M., and K08 AI089858 to A.K.) the Burroughs Wellcome Fund, Nederlands Wetenschappelijk Onderzoek (Meerveld 836.08.001 to I.V.R.), the National Health and Medical Research Council (D.I.G. and J.R.) and the Australian Research Council (S.G.).

**AUTHOR CONTRIBUTIONS**  
I.V.R. designed and did experiments and prepared the manuscript; D.B.M. supervised the experiments and prepared the manuscript; A.K. developed tetramer methods and patient cohorts and analyzed PD-1, CD161 and CD69 by flow cytometry; A.D.I. designed experiments and provided technical advice; S.G., M.B., D.I.G. and J.R. designed experiments and contributed affinity and structural data for TCRs; D.I.G. J.R. and S.G. assisted in preparation of the manuscript; W.d.J. did Luminex experiments; M.E.D. and N.d.V. did deep-sequencing experiments; and W.d.J. provided CD1b monomers.

**COMPETING FINANCIAL INTERESTS**  
The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Beckman, E.M. et al. Recognition of a lipid antigen by CD1-restricted δβ T cells. *Nature* **372**, 691–694 (1994).
2. Han, M., Hännick, L.I., DiBrino, M. & Robinson, M.A. Polymorphism of human CD1 genes. *Tissue Antigens* **54**, 122–127 (1999).
3. Porcelli, S., Yockey, C.E., Brenner, M.B. & Balk, S.P. Analysis of T cell antigen receptor (TCR) expression by human peripheral blood CD4–δβ T cells demonstrates preferential use of several V beta genes and an invariant TCR α-chain. *J. Exp. Med.* **178**, 1–16 (1993).
4. Fowlkes, B.J. et al. A novel population of T-cell receptor δβ-bearing thymocytes which predominantly expresses a single V β gene family. Nature 329, 251–254 (1987).

5. Cardell, S. et al. CD1-restricted CD4+ T cells in major histocompatibility complex class II-deficient mice. J. Exp. Med. 182, 993–1004 (1995).

6. Van Rijn, I. et al. CD1d-restricted T cell activation by nonlipidic small molecules. Proc. Natl. Acad. Sci. USA 101, 13378–13383 (2004).

7. Grant, E.P. et al. Molecular recognition of lipid antigens by T cell receptors. J. Exp. Med. 189, 195–205 (1999).

8. de Jong, A. et al. CD1a-autoreactive T cells are a normal component of the human δβ T cell repertoire. Nat. Immunol. 11, 1102–1109 (2010).

9. de Lalla, C. et al. High-frequency and adaptive-like dynamics of human CD1 self-reactive T cells. Eur. J. Immunol. 41, 602–610 (2010).

10. Vincent, M.S., Xiong, X., Grant, E.P., Peng, W. & Brenner, M.B. CD1a-, b-, and c-restricted TCRs recognize both self and foreign antigens. J. Immunol. 175, 6344–6351 (2005).

11. Godfrey, D.I., MacDonald, H.R., Kronenberg, M., Smyth, M.J. & Van Kaer, L. NKT cells: what’s in a name? Nat. Rev. Immunol. 4, 231–237 (2004).

12. Kasmar, A., Van Rijn, I. & Moody, D.B. The evolved functions of CD1 during infection. Curr. Opin. Immunol. 21, 397–403 (2009).

13. Porcelli, S., Morita, C.T. & Brenner, M.B. CD1b restricts the response of human CD4+ T lymphocytes to a microbial antigen. Nature 360, 593–597 (1992).

14. Gilleron, M. et al. Diacylated sulfoglycolipids are novel mycobacterial antigens stimulating CD1-restricted T cells during infection with Mycobacterium tuberculosis. J. Exp. Med. 199, 649–659 (2004).

15. Moody, D.B. et al. Structural requirements for glycolipid antigen recognition by CD1b-restricted T cells. Science 278, 283–286 (1997).

16. Laye, E. et al. Mycolic acids constitute a scaffold for mycobacterial lipid antigens stimulating CD1-restricted T cells. Chem. Biol. 16, 82–92 (2009).

17. Cohen, N.R., Gang, S. & Brenner, M.B. Antigen Presentation by CD1 lipids, T cells, and NKT cells in microbial immunity. Adv. Immunol. 102, 1–94 (2009).

18. Tilloy, F. et al. An invariant T cell receptor alpha chain defines a novel TAP-independent major histocompatibility complex class Iβ-restricted alpha/beta T cell subpopulation in mammals. J. Exp. Med. 189, 1907–1921 (1999).

19. Kasmar, A.G. et al. CD1b tetramers bind δβ T cell receptors to identify a mycobacterial glycolipid-reactive T cell repertoire in humans. J. Exp. Med. 208, 1741–1747 (2011).

20. Moody, D.B. et al. CD1b-mediated T cell recognition of a glycolipid antigen generated from mycobacterial lipid and host carbohydrate during infection. J. Exp. Med. 192, 965–976 (2000).

21. Ulrichs, T., Moody, D.B., Grant, E., Kaufmann, S.H. & Porcelli, S.A. T cell responses to CD1-presented lipid antigens in humans with Mycobacterium tuberculosis infection. Infect. Immun. 71, 3076–3087 (2003).

22. Van Rijn, I. et al. Low cross-reactivity of T cell responses against lipids from Mycobacterium bovis and M. avium paratuberculosis during natural infection. Eur. J. Immunol. 39, 3031–3041 (2009).

23. Keane, J. et al. Tuberculosis associated with infliximab, a tumor necrosis factor α-neutralizing agent. N. Engl. J. Med. 345, 1098–1104 (2001).

24. Flynn, J.L. et al. An essential role for interferon γ in resistance to Mycobacterium tuberculosis infection. J. Exp. Med. 178, 2249–2254 (1993).

25. Stenger, S. et al. An antimicrobial activity of cytotoxic T cells mediated by granulysin. Science 282, 121–125 (1998).

26. Messi, M. et al. Memory and flexibility of cytokine gene expression as separable properties of human Tγ1 and Tγ2 lymphocytes. Nat. Immunol. 4, 78–86 (2003).

27. Mongkolratanap, J. et al. Antigen-specific expansion of cytotoxic T lymphocytes in acute measles virus infection. J. Virol. 73, 67–71 (1999).

28. Freeman, J.D., Warren, R.L., Webb, J.R., Nelson, B.H. & Holt, R.A. Profiling the T-cell receptor β-chain repertoire by massively parallel sequencing. Genome Res. 19, 1817–1824 (2009).

29. Gnas, S. et al. The shaping of T cell receptor recognition by self-tolerance. Immunity 30, 193–203 (2009).

30. Wun, K.S. et al. A minimal binding footprint on CD1d-glycolipid is a basis for selection of the unique human NKT TCR. J. Exp. Med. 205, 939–949 (2008).

31. Reantagoon, R. et al. Structural insight into MR1-mediated recognition of the mucosal associated invariant T cell receptor. J. Exp. Med. 209, 761–774 (2012).

32. Kjer-Nielsen, L. et al. MR1 presents microbial vitamin B metabolites to MAIT cells. Nature 491, 717–723 (2012).

33. Gold, M.C. et al. Human mucosal associated invariant T cells detect bacterially infected cells. PLoS Biol. 8, e1000407 (2010).

34. Pellicci, D.G. et al. Differential recognition of CD1d-α-galactosyl ceramide by the Vβ8.2 and Vβ7 semi-invariant NKT T cell receptors. Immunity 31, 47–59 (2009).

35. Martin, E. et al. Stepwise development of MAIT cells in mouse and human. PLoS Biol. 7, e45 (2009).

36. Cohen, N.R. et al. Shared and distinct transcriptional programs underlie the hybrid nature of INKT cells. Nat. Immunol. 14, 90–99 (2013).

37. Day, C.L. et al. PD-1 expression on HIV-specific T cells is associated with T-cell exhaustion and disease progression. Nature 443, 350–354 (2006).

38. Montamat-Sicotte, D.J. et al. Mycotic acid-specific CD1d-restricted T cell population contributes to acute and memory immune responses in human tuberculosis infection. J. Clin. Invest. 121, 2493–2503 (2011).

39. Felio, K. et al. CD1-restricted adaptive immune responses to Mycobacteria in human group 1 CD1 transgenic mice. J. Exp. Med. 206, 2497–2509 (2009).

40. Dougan, S.K., Kaser, A. & Blumberg, R.S. CD1 expression on antigen-presenting cells. Curr. Top. Microbiol. Immunol. 314, 113–141 (2007).

41. Su, L.F., Kidd, B.A., Han, A., Kotzin, J.L. & Davis, M.M. Virus-specific CD4+ memory-phenotype T cells are abundant in unexposed adults. Immunity 38, 373–383 (2013).

42. Ly, D. et al. CD1c tetramers detect ex vivo T cell responses to processed phosphorycoketide antigens. J. Exp. Med. 210, 729–741 (2013).

43. Turner, S.J., Doherty, P.C., McCluskey, J. & Rossjohn, J. Structural determinants of T-cell receptor bias in immunity. Nat. Rev. Immunol. 6, 883–894 (2006).

44. Venturi, V., Price, D.A., Douek, D.C. & Davenport, M.P. The molecular basis for public T-cell responses? Nat. Rev. Immunol. 8, 231–238 (2008).
Online Methods

Antigens. C32 GMM was purified from Rhodococcus equi and C60 GMM was purified from Mycobacterium phlei as described. Mycolic acid was from Sigma. For the preparation of lipid extract from M. tuberculosis, a pellet of M. tuberculosis strain H37Rv grown in 7H9 medium was extracted for 2 h at room temperature with chloroform-methanol (1:2, vol/vol), followed by another extraction with chloroform-methanol (2:1). Glycerol monomycolate was purified from a lipid extract from M. bovis BCG grown in Sauton’s medium, followed by loading on a silica column and elution with chloroform.

Flow cytometry. CD1b monomers (NIH Tetracer Core Facility) were loaded with R. equi–derived GMM and assembled into tetramers. Tetramers were incubated for 30 min at room temperature, followed by incubation for an additional 30 min on ice with monoclonal antibodies (Supplementary Table 1). Flow cytometry data were pregal for lymphocytes based on forward and side scatter. Cells were sorted on an 11-color FACSArray (Becton Dickinson).

T cell cloning and T cell assays. Blood was obtained from asymptomatic tuberculin-positive donors clinically assessed to have latent tuberculosis but with no clinical or radiographic evidence of active tuberculosis, from a patient with smear-positive pulmonary tuberculosis, and from blood-bank donors, after informed consent was obtained, as approved by the institutional review boards of the Lemuel Shattuck Hospital and Partners Healthcare. Sorted GMM-loaded Tet+ T cells were stored overnight in medium containing 0.2 ng/ml IL-15 and were plated the next day at a density of one cell per well in round-bottomed 96-well plates containing 2 × 10^4 irradiated allogeneic PBMC, 4 × 10^4 irradiated Epstein–Barr virus–transformed B cells and 30 ng/ml anti-CD3 (OKT3), with 1 ng/ml of IL-2 added on day 2 of the culture. Alternatively, PBMCs were stimulated with autologous monocyte-derived dendritic cells and 1 µg/ml total lipid extract of M. tuberculosis, and IFN-γ secreting cells were isolated (Miltenyi Biotech) and cloned. After 3 weeks, wells with visible growth were restimulated. Clones were tested for binding of GMM-loaded tetramer by flow cytometry and for antigen specificity by ELISPOT assay. For ELISPOT assays, coccults of antigen-presenting cells and T cell were incubated for 16 h in a Multiscreen-IP filter Plate (96 wells; Millipore) coated according to the manufacturer’s instructions (Mabtech). Antibodies are in Supplementary Table 1.

PCR and molecular cloning. From expanding T cell clones and sorted T cell populations, RNA was isolated with an RNeasy kit (Qiagen), and cDNA was synthesized with a QuantiTect reverse-transcription kit (Qiagen), including a genomic DNA–removal step. In vivo-sequence was determined by PCR with primer set IPS000029 and IPS000030 (described on the website of the International Immunogenetics Information System) in combination with TCRx constant-region reverse primer 5′-GTGGTACGCAGCCTCTGAGG-3′ and TCRβ constant-region reverse primer 5′-GTGGGAGCTCGGAGAGCGACCCCGAGG-3′. Taq polymerase was used in the supplied buffer (Denville) under the following cycling conditions: an initial denaturation of 5 min at 95 °C, followed by 35 cycles of 1 min at 94 °C, 1 min at 59 °C and 1 min at 72 °C, then a final elongation step of 7 min at 72 °C. For detection of TCRs with TRAJ, a V segment–specific forward primer was used in combination with the TRAJ9 reverse primer 5′-GGTGGCAGAGAGGCCCCTGTC-3′ and internal Cγ reverse primer 5′-CACTGAGAATACCTTACTCTTGGACACA TTGG-3′ by the Sybr green (BioRad) method. The TRAV1-2 PCR product of positive wells was isolated and sequenced and aliquots of the purified cDNA was amplified by individual PCR with the Vγ-specific forward primers and the Cγ reverse primer 5′-ACACGAGCCACTCGGAGG-3′. Positive reactions were detected by electrophoresis through agarose gels and the amplicons were sequenced.

Quantitative RT-PCR and deep sequencing. The TRAV1-2–TRAJ9 junction was quantified relative to the abundance of Cγ message by the TaqMan method (Applied Biosystems) with the following primers and probes (FAM, 5′-carboxyfluorescein; BHQ, black-hole quencher): TRAV1-2 forward primer, 5′-TCCTTATGGCTTAAAAGGTACAG-3′; TRAV1-2 reverse primer, 5′-TTGCTTTAAAATATGGTTCTTCTGCTTTCC-3′; TRAV1-2 probe, 5′-FAM-CCTCAAGTAAAGACTCTGCTTTACCTTGTG-BHQ-3′; Cγ forward primer, 5′-CTGACCTCTGGGTACCGCACCC-3′; Cγ reverse primer, 5′-CATGAGAATACCTTACTCCTTGTGACACATT-3′; Cγ probe, 5′-FAM-CGATGCAAGTGGTCTGACTTATGGAGTTTT-BHQ-3′. For deep sequencing of TRAV1-2 α-chains, V segment–specific linear amplification and next-generation sequencing was done as described.

TCR affinity measurement. Soluble TCR proteins were expressed and purified as described and were floated in increasing concentrations over GMM–loaded CD1b coupled to research-grade streptavidin-coated chips in a Biacore 3000. The final response was calculated by subtraction of the response from unloaded CD1b. AReal evaluation version 3.1 software (Biacore AB) was used to fit the data to the 1:1 Langmuir binding model and the equilibrium data were analyzed with the Prism program for biostatistics, curve fitting and scientific graphing (GraphPad).

TCR crystallization, data collection and structure determination. TCR crystals of clone 42 were grown by the vapor-diffusion method at 20 °C with a protein/reservoir drop ratio of 1:1 at a concentration of 5 mg/ml in 20% PEG 3350, 0.05 M Na-HEPES, pH 7, and 1% trypate (wt/vol). TCR crystals of clone 18 were grown by the same technique with 25% PEG15000 and 10% propionate cacockyoder–bis Tris propane (1.3-bis(tris(hydroxymethyl)methylamino)propane), pH 7. TCR crystals were soaked in a cryoprotectant solution containing crystallization solution with the PEG concentration increased to 25% and then were ‘flash-frozen’ in liquid nitrogen. Data were collected on the Micro-focus beamline (MX2) at the Australian Synchrotron with an ADSC-Q315r CCD detector at 100K. Data were processed as described, followed by maximum-likelihood refinement with the BUSTER crystallography refinement package. Structures were determined by molecular replacement, with the ELS4 TCR as model, by the Phaser program for phasing macro-

Single-cell sequencing. For single-cell TCR sequencing, cells were sorted in wells of a PCR plate containing 2.5 µl cDNA-synthesis reaction mix consisting of 0.1% Triton X-100 (Sigma–Aldrich) and 0.2 µl iScript reverse transcriptase in 1× iScript reaction buffer and were incubated for 5 min at 25 °C, for 30 min at 42 °C and for 5 min at 85 °C. Subsequently, preamplification targeting TRAV1-2 and all TRBV segments was done in GoTaq Flexi reaction mixture (Promega) containing 2.5 mM MgCl2, 0.2 mM each dNTP, 0.025 U/µl GoTaq polymerase, 0.14 µM each of 26 TRBV–specific forward primers, 0.25 µM β-chain constant region (Cγ) reverse primer 5′-GGTGGCAGAGAGGCCCCTGTC-3′ and internal Cγ reverse primer 5′-CACTGAGAATACCTTACTCTTGGACACA TTGG-3′ by the Sybr green (BioRad) method. The TRAV1-2 PCR product of positive wells was isolated and sequenced and aliquots of the purified cDNA was amplified by individual PCR with the Vγ-specific forward primers and the Cγ reverse primer 5′-ACACGAGCCACTCGGAGG-3′. Positive reactions were detected by electrophoresis through agarose gels and the amplicons were sequenced.

NATURE IMMUNOLOGY

doi:10.1038/nm.2630

© 2013 Nature America, Inc. All rights reserved.
45. Klarenbeek, P.L. et al. Human T-cell memory consists mainly of unexpanded clones. *Immunol. Lett.* 133, 42–48 (2010).

46. Klarenbeek, P.L. et al. Inflamed target tissue provides a specific niche for highly expanded T-cell clones in early human autoimmune disease. *Ann. Rheum. Dis.* 71, 1088–1093 (2012).

47. Smart, O.S. et al. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr. D Biol. Crystallogr.* 68, 368–380 (2012).

48. Tynan, F.E. et al. A T cell receptor flattens a bulged antigenic peptide presented by a major histocompatibility complex class I molecule. *Nat. Immunol.* 8, 268–276 (2007).