Identification of $CD4^+$ effector memory T cells that recognize a mycobacterial adjuvant

Yuki Sakai

Expected graduation: Sep 2024

Abstract

Mycobacteria causes serious diseases such as tuberculosis (TB), leprosy, Buruli ulcer and nontuberculous mycobacterial (NTM) lung disease. These bacteria are protected by a thick lipid cell envelope and hosts have developed immune systems targeting their unique lipids. Screening for responses of human cells to complex *Mycobacterium tuberculosis* lipids mixture revealed specific recognition of trehalose monomycolate (TMM). Whereas TMM has known adjuvanticity through the innate immune receptor, Mincle, I determined that responses were mediated by human T cells in a CD1b-restricted manner. Single cell profiling showed that TMM-specific T cells were $CD4^+$ effector memory subtype, exhibiting cytotoxic and proinflammatory signatures immediately after TMM stimulation. These T cells were detected by CD1b-TMM tetramers at higher frequencies in TB patients than in controls. TMM-specific TCRs were conserved among donors and displayed cationic CDR3α and long CDR3β regions, pointing to unique antigen recognition mechanisms that were supported by mutagenesis analyses, TCR crystallization and a cryo-EM ternary structure. These data demonstrate a pre-formed T cell subset against mycobacteria that is shared across humans. My findings expand the known bioactivity of a well-studied adjuvant molecule and justify reconsideration of the mechanism of action of adjuvants.

Contents

General Introduction

Diseases caused by mycobacteria, including tuberculosis, leprosy, Buruli ulcer and nontuberculous mycobacterial (NTM) lung disease, rank among the top causes of death and disability worldwide. Mycobacteria are distinguished from other bacteria by a thick cell envelope comprised of unique outer membrane of neutral lipids and glycolipids, which forms the primary barrier with the host and contributes to resistance to host immunity (Dulberger et al., 2020; Stenger et al., 1998). The mycobacterial cell envelope has been broadly applied *in vivo* in animals as 'complete Freund's adjuvant' (CFA) that can promote a strong vaccine response (Freund et al., 1937). Yet, limited understanding of the defined immunogenic components of CFA has to date prevented its therapeutic use as a vaccine adjuvant in humans (Dube et al., 2020). Recently, the receptors for some mycobacterial immunogens have been defined, which involve lipopeptide–NOD2 and trehalose dimycolate (TDM)/TMM–Mincle axes (Brown et al., 2018; Ishikawa et al., 2017). Some other lipids are presented by CD1 to innate-type T cells (De Libero et al., 2014 ; Van Rhijn and Moody, 2015). It is therefore possible that a lipid recognized by both innate and acquired immune receptors may possess potent and durable immunostimulatory activity.

In this study, I sought to identify mycobacterial cell wall components leading to human T cell activation *ex vivo*. After demonstrating strong T cell activation, I purified the stimulatory component and established its structure as TMM. T cell activation was not seen by another Mincle agonist TDM, which led to the identification of TMM as a CD1b presented antigen. Using TMM-loaded CD1b tetramers, I identified a naturally

occurring CD1b-restricted effector memory T cell population, which exists in naïve donors and is suggested to expand in humans during Mycobacteria infection. Single cell TCR sequencing (sc-TCR-seq) identified a cationic TCR motif for CD1b and TMM recognition and the action mechanism was validated by mutagenesis and structural analysis. Finally, single cell measurements provided a broad and unbiased insight into natural effector functions of CD1b-reactive T cells *ex vivo*, demonstrating that these CD4 T cells produce typical mycobactericidal effectors including granulysin (Stenger et al., 1998), IFNγ (Flynn et al., 1993) and TNF (Flynn et al., 1995; Keane et al., 2001).

Results

Identification of mycobacterial lipid-reactive T cells in human PBMCs

As a relatively unbiased and comprehensive way to search for T cells reactive to the mycobacterial cell envelope, PBMCs from healthy donors were stimulated with total *M. tuberculosis* compounds extracted into chloroform/methanol (2:1, vol/vol) and coated onto tissue culture plates (Ishikawa et al., 2009). Based on cell trace violet (CTV) dilution, I subjected lipid-responsive T cells to single cell-TCR-RNA sequencing (sc-TCR-RNA-seq) to identify potentially diverse clonotypes along with their effector function and TCR profiles (Fig. 1A)(GSE260931). Among 26,502 detected clonotypes, I selected 52 clonotypes that are expanded by *M. tuberculosis* lipids and reconstituted their TCRαβ pairs in NFAT-GFP reporter cells (Fig. 1A and Supplementary Table 1).

After surface expression of the TCR complex was confirmed for 44 clonotypes (Fig. 1B), I tested these cells for responses to plate-coated *M. tuberculosis* lipids in the presence of cytokine-differentiated human monocytes as antigen-presenting cells (APCs). One clonotype derived from $CD4^+$ T cells, Y-50, responded strongly to mycobacterial lipids based on GFP and CD69 upregulation (Fig. 1C). Analysis of sc-TCR-RNA-seq data revealed that the Y-50 clonotype was expressed by fourteen individual cells within the T helper 1 cell (Th1) cluster, located close to $CD4^+$ cytotoxic T lymphocyte (CTL)-like clusters (Fig. 1D). These cells were characterized by the expression of granzyme B (*GZMB*), perforin-1 (*PRF1*), granulysin (*GNLY*), TNF (*TNF*) and IFNγ (*IFNG*) (Fig. 1E, Extended Data Fig. 1). Y-50 also expressed innate-like T cell markers, like CD161/killer cell lectin like receptor B1 (*KLRB1*) and CCAAT enhancer binding protein delta

(*CEBPD*) (Fig. 1E). Some Y-50 cells were also detected in a $Ki67^{hi}$ $CD4^{+}$ CTL cluster (Fig. 1D), in agreement with their CTV^{lo} status used for sorting (Fig. 1A). These results suggest that mycobacterial lipid-reactive Y-50-expressing $CD4^{+}$ T cells exhibit an innate- and cytotoxic-like signature after *ex vivo* lipid stimulation.

TMM activates the Y-50 clonotype T cells

To identify the lipid stimulus, I separated crude lipids by thin layer chromatography (TLC) and measured responses to each fraction. I collected 16 fractions and found potent antigen activity in fraction 11 (Fig 2A). The active peak shifted to lower Rf values (fraction 2) under more hydrophobic solvent conditions (Fig. 2B), suggesting that the antigenic component was likely a moderately polar lipid. Thus, the fraction were analyzed by matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS) (Fig. 2C), finding ions that matched and largely overlapped in chain length and saturation patterns with purified trehalose monomycolate (TMM) from *M. tb* H37 Rv (Harris et al., 2010)(Fig. 2D and Extended Data Fig. 2). In addition to purified TMM, Y-50 TCR-expressing reporter cells were also activated by APCs cocultured with *M. tuberculosis* H37 Rv and *M. bovis* BCG (Fig. 2E), demonstrating the origin of the stimulus from intact bacteria.

CD1b restricts TMM recognition by Y-50 TCR

TMM is most well known as a major cell surface glycolipid and macrophage adjuvant that is a biosynthetic intermediate to TDM, also known as cord factor. As TDM

and TMM were both known to potently activate myeloid cells through the innate receptor, Mincle (Decout et al., 2017; Ishikawa et al., 2009), TMM might activate T cells via Mincle on APCs. However, the Y-50 T cells selectively recognized TMM but not TDM (Fig. 2F), suggesting that the response was specific to TMM structure and not Mincle mediated. The major alternative hypothesis was CD1 presentation of TMM to TCRs, as prior studies reported that structurally similar glucose monomycolate (GMM), which was not a Y-50 antigen (Fig. 2F), is presented by CD1b (Moody et al., 1997). It is possible that TMM might be a novel T cell antigen presented by CD1b.

As all four isoforms of human CD1 antigen presenting molecules can present many types of lipids (Huang et al., 2023), I examined the effect of blocking antibodies against human CD1a, CD1b, CD1c and CD1d and found that only anti-CD1b selectively suppressed TMM-induced activation of Y-50 reporter cells in the presence of APCs (Fig. 3A). Conversely, ectopic expression of CD1b on HEK293T cells conferred Y-50 TCR reactivity to TMM (Fig. 3B). Thus, only CD1b is necessary and sufficient for the presentation of TMM to the Y-50 TCR.

Clone Y-50 broadly recognize TMM from various mycobacterial species

TMM is produced broadly among mycobacterial species (Fujita et al., 2005). To further characterize the selectivity of antigenic lipids recognized by the Y-50 TCR, I examined the purified TMMs possessing different lipid lengths from *M. intracellulare* and *M. smegmatis* (C60-C88) and *Rhodococcus* species with shorter mycolate moieties (C28-C36), finding that all showed antigenic activity (Fig. 3C). However, previously

identified glycolipid antigens that varied in the head group moiety, GMM, mannose monomycolate (MMM), glycerol monomycolate (GroMM) and free mycolic acid (MA)(Beckman et al., 1994; Layre et al., 2009), lacked antigenicity (Fig. 3C). Thus, the T cell reactivity identified here was new, and these data indicate that the head group comprised of the trehalose disaccharide is required for Y-50 TCR recognition.

To exclude the possibility of contaminants or mitogens in natural TMM preparations, I examined the reactivity of synthetic TMM (*R, R*) (Fig. 3D). Synthetic TMM also induced T cell activation, formally ruling in this structure as an antigen (Fig. 3E). Further, other synthetic TMM analogues lacking an α-branched alkyl chain or a β-hydroxy group did not activate, defining components of the lipid tail that are necessary for recognition (Fig. 3E). Thus, while Y-50 broadly recognizes TMM with various lipid-chains present across mycobacterial species, it discriminates the natural configuration of mycobacterial TMM with regard to its lipid anchor and the carbohydrate head group.

Cationic residues in Y-50 TCR are critical for TMM recognition

A distinguishing characteristic of Y-50 TCRα is the presence of four positively-charged arginine residues (R107, R108, R113 and R114) within the CDR3α region (Fig. 4A). To investigate the contribution of these residues to TMM recognition, I introduced alanine mutations and evaluated their effects using TCR-reconstituted reporter cells. TCRα containing alanine substitutions at R107, R113 or R114 showed impaired reporter activity, whereas R108A had no impact (Fig. 4A).

Compared to the average CDR3β (14.4 residues), the Y-50 CDR3β was much longer, encoded by 20 residues (Fig. 4B). To assess the contribution of amino acid insertion during VDJ recombination to antigen recognition, I removed amino acids from the junctional region to create shorter Y-50 TCRβ chains. None of these TCRβ mutants recognized TMM (Fig. 4B), suggesting that certain aspects of this long CDR3β are required for the recognition of TMM by the Y-50 TCR.

Structural characterization of Y-50 TCR

The cationic and long loop motifs suggested a TCR binding mechanism dominated by electrostatic interactions and a flexible TCR surface. To gain structural insight into the Y-50 TCR, a soluble TCRαβ was constructed and a crystal structure of the TCRαβ complex (PDB: 8XUB) that diffracted to a resolution of 2.5 Å were obtained (Fig. 4C and Supplementary Table 2). Three CDR3α arginine residues that were found to be critical for TMM recognition were positioned toward TCRβ, whereas R108, which was dispensable, was oriented away from the TCRα-β interface (Extended Data Fig. 3A). The electron density of the TCRαβ was clear except for the CDR3β loop region, implying that the extra-long CDR3β loop may be highly flexible (Fig. 4D, E).

Determination of ternary complex structure of Y-50 TCR-TMM-CD1b

However, the definitive recognition mechanism requires a ternary structure, so cryo-electron microscopy (cryo-EM) analysis were conducted. Recombinant CD1b was refolded with synthetic TMM (Extended Data Fig. 3B) and incubated with soluble

TCRαβ. The cryo-EM map of the ternary complex was successfully reconstructed to a resolution of 3.18 Å (PDB: 8ZOX)(Fig. 5A, Extended Data Fig. 4 and Supplementary Table 3). In the area between TCRαβ and CD1b, the clear density map that exactly overlapped with the chemical structure of TMM were observed (Extended Data Fig. 3C). TMM lipid chains were buried within CD1b pockets like GMM (Gras et al., 2016) with its sugar head exposed toward TCR (Extended Data Fig. 3D). CDR3αβ regions appeared to accommodate its protruded bulky trehalose moiety (Extended Data Fig. 3D).

This observation was supported by the superimposition of the Y-50 TCR structure within the ternary complex (Fig. 5A, PDB 8ZOX) with Y-50 TCR alone (Fig. 4C, PDB 8XUB). The position of backbone Cα atoms are largely overlapped; however, the location of long CDR3 loop, particularly the CDR3β region, was noticeably shifted (Fig. 5B). Whereas the CDR3β loop hung 'downwards' in the structure without antigen, it "rolled up" to allow the recognition of TMM presented by CD1b (Fig. 5B), supporting the importance of a long CDR3β which creates the flexible interface with CD1b-TMM complex by avoiding steric hindrance.

I next determined the mode of specific TMM recognition by Y-50 TCR. The experimentally observed functional importance of cationic CDR3α residues (Fig. 4A) was explained by the ternary structure. R114 formed a hydrogen bond with β-hydroxy residue of TMM (Fig. 5C), in agreement with both the strong effects of alanine mutation (Fig. 4A) and altered recognition TMM lacking the β-hydroxy group (Fig. 3E). To further test the significance of this interaction, TMM stereoisomers that differs in the stereochemistry of the acyl group (Extended Data Fig. 5A, 5B) were synthesized. Y-50

TCR recognized natural TMM (*R,R*) but not non-natural isomers (Fig. 5D). Thus, the Y-50 TCR is highly specific for the natural stereo-configuration of the TMM lipid moiety through the cationic residue R114.

Trehalose contains two glucose units. Another critical cationic residue, TCRα R107, formed a hydrogen bond with hydroxy residue at C-2′ atom of the distal glucose in trehalose, while the proximal glucose interacted with TCRβ D114 through hydroxy residue at C-4 atom (Fig. 5E). Furthermore, α^{R107A} or β^{D114A} formed a salt bridge (Fig. 5E). Thus, Y-50 TCRα and TCRβ may cooperate by interacting each other and recognizing distinct epitopes of TMM. In line with this interpretation, Y-50 TCR recognition of TMM was impaired in α^{R107A} or β^{D114A} single mutation and more severely in double mutation (α^{R107A}-β^{D114A})(Fig. 5F).

These structural analyses revealed the molecular basis in which the characteristic features of Y-50 TCR, including the cationic CDR3α and long CDR3β, mediate recognition of TMM-CD1b complex. Alanine scanning confirmed this mode of antigen recognition and found additional critical residues (Extended Data Fig. 6). TCRβ G110 was also important due to an interaction with E80 of CD1b (Extended Data Fig. 3E). In addition to CDR3β, other CDRβ regions derived from TRBV4-1 also contributed to the binding to CD1b independent of antigen. R37 in CDR1β formed salt bridge with D83 of CD1b. CDR2β also interacted with CD1b α1 helix through salt bridge (E63–R79) and hydrogen bond (Y58–E80) (Extended Data Fig. 3E). Thus, multiple bonds between CD1b ^{79}RExxD^{83} sequence (Extended Data Fig. 3F) and TCRVβ loops may stabilize preferential interaction of CD1b to TRBV4-1 (Reinink et al., 2019; Van Rhijn et al., 2014).

Characterization of TMM-specific T cells in the periphery using tetramers

Little is known about the profile of CD1b reactive T cells in the periphery without *in vitro* expansion, so, I sought to generate TMM-loaded CD1b tetramers and combine them with single cell analysis. These TMM-CD1b and unloaded control tetramers were validated by binding to cell lines expressing the Y-50 TCR and human TMM-expanded polyclonal T cells (Extended Data Fig. 7). I thus investigated gene expression profiles of freshly-isolated and TMM-stimulated T cells bearing Y-50 TCR using sc-TCR-RNA-seq. Y-50 T cells were separated into different clusters before and after stimulation in UMAP plot (Fig. 6A), implying that gene expression signatures were altered by antigen stimulation. Before stimulation, Y-50 T cells expressed typical effector memory markers, such as CD44, IL-7 receptor α chain and integrin β_1 (Fig. 6B and Extended Data Fig. 8A, B), suggesting that TMM-specific T cells are naturally occurring memory T cells without antigen exposure, which is the characteristic feature of NKT and other innate T cells (Bendelac et al., 2007). Upon TMM stimulation, Y-50 T cells moved to the cluster characterized by the expression of cytotoxic effector genes (*GZMB, PRF1, GNLY*) and pro-inflammatory cytokines and chemokines (*IFNG, TNF* and *CCL5*), whereas the expression levels of stemness-related molecules, such as *IL7R*, *TCF7* and *CXCR4*, were downregulated (Fig. 6C-E). These results suggest that Y-50 T cells are innate-like T cells exhibiting cytotoxic potential in the periphery, whose effector signature is markedly enhanced upon antigen stimulation.

TMM-specific T cells are shared across humans

To examine whether TMM-specific T cells are shared across genetically unrelated individuals, I sorted TMM-CD1b-tet$^+$ cells from fresh PBMCs of additional donors and examined their characteristics by sc-TCR-RNA-seq. Frequent TMM-CD1b-tet$^+$ clonotypes from unrelated donors were mainly localized within the CD4$^+$ effector memory clusters (Fig. 7A, B), which expressed CD44 (*CD44*), IL-7 receptor (*IL7R*) or integrin β_1 (*ITGB1*) but not homing receptor CCR7 (*CCR*7)(Extended Data Fig. 8C), similar to unstimulated Y-50 T cells (Extended Data Fig. 8B). TCRs expressed by these clonotypes were reconstituted in reporter cells and confirmed to react to TMM (Fig. 7C). Importantly, sequences of these TCRs showed them not be identical to Y-50 but they possessed similar characteristics, including positively-charged CDR3α, biased TCRβ that is encoded by TRBV4-1 and TRBV6-2, and long CDR3β sequences (Fig. 7C). Further, these patterns could be seen in TCR analysis that were either unsorted or sorted with TMM-CD1b tetramers (Fig. 7D, E). Thus, TMM and CD1b-reactive T cells show clear evidence for conserved features across numerous clonotypes from different donors, constituting a new donor unrestricted T cell subset in humans (Extended Data Fig. 9B).

Quantification of TMM-specific T cells during Mtb infection

Finally, I used tetramers to examine the frequency of TMM-specific T cells in PBMCs from uninfected and TB infected donors (Table S4). TMM-CD1b-tet$^+$ T cells were detected in most uninfected donors, consistent with prior experiments of single cell or tetramer-based outcomes (Fig. 6 and 7), further suggesting that these cells being a

preformed effector memory T cell type. The frequency was significantly increased in active TB patients (Fig. 7F), suggesting that these T cells may react to TMM during mycobacterial infection in the host. In addition to Japanese healthy donors who had received BCG vaccination, I also examined PBMCs and cord blood cells of healthy donors from North America where BCG is no longer widely used. The frequency of tetramer^{+} T cells was comparable between these three groups, indicating that exposure to BCG or environmental antigens is unlikely to be a trigger for durable expansion of TMM-specific T cells (Fig. 7F).

Discussion

A basic paradigm for adaptive MHC-restricted T cells is that naïve cells are primed by antigen to differentiate to memory T cells, including effector memory T cells, that persist in the periphery. In contrast, CD1d-restricted iNKT cells express memory markers in the absence of defined antigenic stimulation, and functionally circulate in larger numbers and respond rapidly as a cohort to antigen challenge (Bendelac et al., 2007; Engel and Kronenberg, 2012). However, the extent to which human group 1 CD1 (CD1a, CD1b and CD1c)-restricted T cells behave as adaptive or innate T cells, which are not present in mice, remains unknown owing mainly to the lack of defined infection model and the limited ability to track human T cells directly and quantitatively *ex vivo*. Through the discovery of TMM antigen, generation of CD1b-TMM tetramers applied across unrelated donors, discovery of a new TCR motif and structural determination of ternary complex, this study begins to tease out the natural history of a human pathogen-specific T cell responses in the CD1b system.

Key findings are the identification of pathogen-specific TCRs without infection, and the rapid upregulation of anti-mycobacterial effector functions by stimulation *ex vivo*. IFNγ and TNF are canonical anti-mycobacterial protective cytokines produced from $CD4^+$ Th1 cells (Flynn et al., 1993; Flynn et al., 1995; Keane et al., 2001). Recently, granulysin, granzyme B and perforin secreted from $CD8^+$ T cells are also recognized to be important for protection against mycobacterial infection (Nunes-Alves et al., 2014; Stenger et al., 1998; Vesosky et al., 2010). TMM-specific T cells are unique cell subset that rapidly upregulated all of these effector genes simultaneously in response to *M.*

tuberculosis antigen. While host protection is difficult to demonstrate in human systems, this evidence supports that TMM-reactive T cells express a host protective effector function.

Although TMM-reactive T cells expressed CD4, this coreceptor seems dispensable for the recognition of CD1b-TMM complex, as the reporter cells do not express human CD4. However, I cannot fully exclude the possibility that, like conventional T cells, CD4 is required for those T cells to be selected by MHC class II. Clinically, the well-known susceptibility of human immunodeficiency virus (HIV)-infected patients to tuberculosis resulting from the reduction of $CD4^+$ T cells (Geldmacher et al., 2010) might also be partly due to the loss of $CD4^+$ TMM-specific T cells.

Currently, the advantage of targeting TMM by both innate and acquired immunity for the host is unclear. TMM is expressed by most mycobacterial species and is used for further biosynthesis of other cell wall components, including arabinogalactan and TDM (Su et al., 2019). Unlike TDM, which is downregulated in mycobacteria upon infection of the host, the level of TMM is relatively constant (Matsunaga et al., 2008). Compared with GMM, TMM may be resistant to stresses such as oxidation because the reducing ends of both glucoses are occupied. Thus, given the importance of TMM for multiple stages of the mycobacterial life cycle, the presence of T cells that recognize TMM with various lipid chain lengths plausibly could allow effective induction of responses to a broad spectrum of mycobacteria.

Taking advantage of the direct detection by tetramers and single cell analysis, I provide several lines of evidence that TMM-specific T cells exist before the host is

exposed to mycobacteria, as they were detected in random blood donors, as well as non-TB or non-BCG vaccinated donors and even in cord blood cells. However, as *PLZF* was not highly expressed in TMM-specific T cells, it is unclear whether they were selected by DP thymocytes like other innate-like T cells (Savage et al., 2008; Seach et al., 2013). Even assuming the involvement of CD1b for selection, the selecting ligand(s) is unclear. It is also possible that the intrinsic affinity of TRBV4-1 or TRBV6-2 to CD1 molecules (Reinink et al., 2019) might allow less ligand-dependent selection.

Comparison of the ternary structure of TMM-specific TCR with the previously reported structure of GEM TCR-GMM-CD1b (Gras et al., 2016) provides similarities and differences in the mode of glycolipid recognition (Extended Data Fig. 9A, B). Arginine residues within CDR3α region of both TCRs interact with β-hydroxy residue of GMM and TMM, which allow them to discriminate natural configuration of mycolyl lipids. Compared to the monosaccharide in GMM, two sugar moieties of TMM interacted more extensively with TCR residues, possibly determining its antigen specificity. Furthermore, long CDR3β region uniquely found in TMM-specific TCRβ and the demonstrated compression of the TCRβ CDR3 loop in the two Y-50 structures shows how the TCR β chain moves to accommodates the bulky TMM presented by CD1b.

Biased usage of TRBV4-1 has been reported for CD1b-restricted T cells (Reinink et al., 2019; Van Rhijn et al., 2014). Shao et al. also reported the increase of TRBV4-1^{+} T cells in tuberculous pleural effusions (Shao et al., 2022). Our first complex structure of CD1b and TRBV4-1 TCR may provide mechanistic insight, including the presence of bonds between CD1b RExxD sequence and TRBV4-1-derived amino acids.

Conservation of this motif among CD1b and even CD1c, but not in other human CD1 molecules (Extended Data Fig. 3F), may support the high frequency of TRBV4-1 in CD1b- and CD1c-restricted T cells (Guo et al., 2018).

TMM is recognized by pattern recognition receptor Mincle (Ishikawa et al., 2009). Thus, this mycobacterial lipid represents a 'dual ligand' that can activate both PRR and TCR, so our data propose that TMM can act simultaneously as a PAMP and an T cell antigen, respectively (Extended Data Fig. 9C). TMM analogues could therefore be developed as lipid immunogens with a novel mode of action, as both adjuvant and antigen, to prevent various diseases caused by mycobacterial species. In addition to tuberculosis, NTM lung disease is one of the most urgent targets as cases are dramatically increasing and current drug treatments are ineffective (Daley et al., 2020; Wu et al., 2018). TMM from NTM species are also – and more strongly – recognized by TMM-specific T cells. In addition, the potent activity and higher hydrophilicity of short-chain TMMs is advantageous in terms of efficacy and formulation/administration modalities. More detailed analysis on the structure-activity relationships and the protective role of TMM-specific T cells against infection are important future issues to be addressed.

Materials and methods

Human subjects

This project was approved by Osaka University Institutional Review Board (IRB) (approval number 898-4). Informed consent was obtained from all participants before the first blood sampling. Peripheral and cord blood cells from healthy donors were also obtained from VERITAS Corporation (Tokyo, Japan; Batch 210570303C, 220771201C, 220771404C, 220772503C, 220781001C, 2208409001, 2208411000, 220873101C, 220880801C, 220881703C (PBMC); Batch 2211410002, 2211416002, 2211422002, 2211422002, 2211422003, 2211423000, 2211423001, 2212406005, 2212414001, 2212414003, 2212420000 (CBMC)). Active tuberculosis cases (13 cases) are those that were admitted to the hospital as sputum smear and culture positive pulmonary tuberculosis cases that consecutively included in the study from April 2023 to July 2023. All participants were enrolled after giving written informed consent. Blood was taken for active tuberculosis before starting treatment. This study was reviewed and approved by the medical research ethics committee of the National Institute of Infectious Diseases for inclusion of human subjects (#1343 and #1491) and of Fukujuji Hospital (#22034).

Bacteria

M. tuberculosis strain H37Rv was kindly provided by Dr. Ikuya Yano. For inactivation, the bacterium was heated at 65°C for 1 h, followed by incubation at 60°C for overnight. *M. bovis* BCG was purchased from Japan BCG Laboratory.

Lipid extraction and purification for stimulation

Mycobacterium tuberculosis strain H37Rv lipids were extracted as previously described (Nishimura et al., 2023). Briefly, 10 ml of chloroform/methanol (2:1, vol/vol) or acetone were added to 100 mg of bacteria and sonicated at 40°C for 10 min. The organic phase was collected and dried and dissolved in chloroform/methanol (2:1, vol/vol) for storage and aliquoting into various assays as a crude lipid. For lipid fractionation, crude lipids were separated by high performance thin layer chromatography (Merck) followed by charring with copper(II) acetate-phosphoric acid. TMM was purified from *M. tuberculosis* H37Rv, *M. tuberculosis* CDC1511, *M. bovis* BCG, *M. intracellulare*, *M. smegmatis*, *R. equi* and *R. sp* 4306, TDM was purified from *M. tuberculosis* CDC1511, GMM and MMM was purified from *R. ruber* and GroMM was purified from *M. bovis* BCG as previously described (Fujiwara et al., 2012; Fujiwara et al., 1999; Harris et al., 2010; Matsunaga et al., 1996; Naka et al., 2011; Ueda et al., 2001). Briefly, the heat-killed bacteria were sonicated in chloroform/methanol (2:1, vol/vol) for 15 min on ice, and water was added (1/20 total volume). The organic layer was collected and evaporated completely. The crude lipids were separated by thin-layer chromatography (Merck) and fractions were extracted. Mycolic acid was purified from *M. tuberculosis* H37Rv as described (Nishimura et al., 2023). Synthetic GMM was provided by Adriaan Minnaard (Groningen University) (Tahiri et al., 2020). For stimulation of cells, lipids dissolved in chloroform/methanol (2:1, vol/vol) were diluted in isopropanol, applied to 96-well plates at 20 μL/well and air-dried prior to adding media.

MALDI-TOF MS analysis

TMM was detected by matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) with an UltrafleXtreme (Bruker Daltonics, Billerica, MA). In brief, purified lipid fractions and TMM standards were dissolved in chloroform/methanol (3:1, vol/vol) at a concentration of 1 mg/mL, and 1 μL of sample was applied directly to the sample plate, followed by addition of 1 μL 2,5-dihydroxybenzoic acid (10 mg/mL in chloroform/methanol, 1:1, vol/vol) as a matrix. The samples were analyzed in the reflectron mode with an accelerating voltage operating in positive mode of 20 kV (Bhatt et al., 2007; Harris et al., 2010).

Chemical synthesis

Reactions were carried out under a nitrogen atmosphere unless noted otherwise and monitored by thin-layer chromatography using Merck Silica Gel 60 F254 plates. Flash chromatography was performed using flash silica gel 60N (spherical neutral, particle size 40–50 μm,Kanto Chemical Co. Ltd). NMR spectra were recorded using a Bruker Avance III (500 MHz) device with a Prodigy (nitrogen-based) cryoprobe or a JNM-ECZL600R (600 MHz) device with a ROYAL HFX probe. Chemical shifts were reported in the scale relative to $CHCl_3$ (δ 7.26 ppm for ^{1}H NMR, 77.16 ppm for ^{13}C NMR) or pyridine (δ 7.58 ppm for ^{1}H NMR, 135.91 ppm for ^{13}C NMR) as an internal reference. Splitting patterns are designated as s: singlet, d: doublet, t: triplet, q: quartet, br: broadening, and m: multiplet. High-resolution mass spectrometry (HRMS) was obtained with a Bruker MicrOTOF II detector or a Bruker MALDI-TOF MS Autoflex Speed device. Gel

permeation chromatography (GPC) was executed using LaboACE LC–5060 equipped with JAIGEL-1HR and JAIGEL-2HR ($CHCl_3$).

TMM (C32, *RR*): $[\alpha]_D^{27}$ +110.01 (c = 0.28, $CHCl_3$/MeOH 4/1); ^{1}H NMR (500 MHz, $CDCl_3$/CD_3OD 4/1) δ 4.97 (d, *J* = 3.7 Hz, 1H), 4.93 (d, *J* = 3.7 Hz, 1H), 4.48 (dd, *J* = 11.9, 1.8 Hz, 1H), 4.05 (ddd, *J* = 10.0, 7.0, 1.8 Hz, 1H), 3.94 (dd, *J* = 11.9, 7.0 Hz, 1H), 3.75 (m, 1H), 3.72 (dd, *J* = 11.0, 1.8 Hz, 1H), 3.69 (dd, *J* = 9.8, 9.2 Hz, 1H), 3.67 (dd, *J* = 9.8, 9.2 Hz, 1H), 3.56–3.50 (m, 2H), 3.42 (dd, *J* = 9.8, 3.7 Hz, 1H), 3.37 (dd, *J* = 9.8, 3.7 Hz, 1H), 3.18 (dd, *J* = 9.6, 9.2 Hz, 1H), 3.14 (dd, *J* = 10.0, 9.2 Hz, 1H), 2.29 (ddd, *J* = 10.1, 7.8, 4.6 Hz, 1H), 1.50–1.39 (m, 2H), 1.38–1.28 (m, 2H), 1.28–1.05 (m, 50H), 0.75 (t, *J* = 7.0 Hz, 6H); ^{13}C NMR (126 MHz, $CDCl_3$/CD_3OD 4/1) δ 175.5, 94.43, 94.37, 72.7, 72.62, 72.58, 72.3, 71.6, 71.4, 71.0, 70.8, 70.0, 64.0, 61.9, 52.5, 34.7, 31.8 (2C), 29.64–29.48 (15C), 29.4, 29.34, 29.26 (2C), 29.2, 27.2, 25.2, 22.6 (2C), 13.9 (2C); HRMS-MALDI (*m/z*): $[M+Na]^+$ calcd for $C_{44}H_{84}NaO_{13}$, 843.5810; found 843.58.

TMM (C32, *SS*): $[\alpha]_D^{27}$ +8.03 (c = 0.22, $CHCl_3$/MeOH 4/1); ^{1}H NMR (500 MHz, $CDCl_3$/CD_3OD 4/1) δ 4.98 (d, *J* = 3.7 Hz, 1H), 4.95 (d, *J* = 3.7 Hz, 1H), 4.29 (dd, *J* = 11.7, 2.0 Hz, 1H), 4.19 (dd, *J* = 12.1, 5.0 Hz, 1H), 3.92–3.87 (m, 1H), 3.72–3.66 (m, 4H), 3.57–3.50 (m, 2H), 3.41 (dd, *J* = 9.8, 3.7 Hz, 1H), 3.37 (dd, *J* = 9.8, 3.7 Hz, 1H), 3.24 (dd, *J* = 9.3, 9.2 Hz, 1H), 3.19 (dd, *J* = 9.3, 9.2 Hz, 1H), 2.29 (ddd, *J* = 10.2, 7.8, 4.6 Hz, 1H), 1.49–1.39 (m, 2H), 1.38–1.28 (m, 2H), 1.28–1.03 (m, 50H), 0.74 (t, *J* = 7.0 Hz, 6H); ^{13}C NMR (126 MHz, $CDCl_3$/CD_3OD 4/1) δ 175.5, 94.0, 93.9, 72.8, 72.7, 72.5, 72.2, 71.6 (2C), 70.7, 70.3, 70.2, 63.1, 61.9, 52.9, 34.7, 31.8 (2C), 29.6–29.5 (15C), 29.4, 29.3, 29.2

(2C), 29.1, 27.2, 25.1, 22.6 (2C), 13.9 (2C); HRMS-MALDI (*m/z*) $[M+Na]^+$ calcd for $C_{44}H_{84}NaO_{13}$, 843.5810; found 843.58.

TMM (C32, *RS* +*SR*): $[\alpha]_D^{27}$ +9.02 (c = 0.23, $CHCl_3$/MeOH 4/1); ^1H NMR (500 MHz, $CDCl_3/CD_3OD$ 4/1) δ 4.97 (d, *J* = 3.7 Hz, 1H), 4.96 (d, *J* = 3.4 Hz, 1H), 4.94 (d, *J* = 3.7 Hz, 1H), 4.92 (d, *J* = 3.7 Hz, 1H), 4.51 (dd, *J* = 11.8, 1.8 Hz, 1H), 4.23–4.17 (m, 2H), 4.00–3.88 (m, 3H), 3.74–3.63 (m, 10H), 3.540 (dd, *J* = 12.7, 6.7 Hz, 1H), 3.538 (dd, *J* = 11.9, 6.1 Hz, 1H), 3.404 (dd, *J* = 9.8, 3.7 Hz, 1H), 3.399 (dd, *J* = 9.8, 3.7 Hz, 1H), 3.37 (dd, *J* = 9.6, 3.7 Hz, 1H), 3.35 (dd, *J* = 9.8, 3.7 Hz, 1H), 3.25–3.12 (m, 4H), 2.40–2.32 (m, 2H), 1.60–1.50 (m, 2H), 1.38–1.28 (m, 6H), 1.28–1.03 (m, 100H), 0.74 (t, *J* = 6.9 Hz, 12H); ^{13}C NMR (126 MHz, $CDCl_3/CD_3OD$ 4/1) δ 175.04, 174.98, 94.3, 94.1 (2C), 94.0, 73.0, 72.8 (2C), 72.7, 72.4, 72.3, 72.2 (2C), 71.6 (3C), 71.5, 70.9, 70.8, 70.7, 70.3, 70.0, 69.9, 63.8, 63.3, 61.9 (2C), 51.6, 51.3, 33.8, 33.5, 31.8 (4C), 29.65–29.48 (30C), 29.46 (2C), 29.4 (2C), 29.2 (4C), 27.9, 27.8, 26.6, 26.1 (2C), 25.8, 22.6 (4C), 13.9 (4C); HRMS-MALDI (*m/z*) $[M+Na]^+$ calcd for $C_{44}H_{84}NaO_{13}$, 843.5810; found 843.58.

Other synthetic precursors and TMM analogues were synthesized as described in the supplementary information.

The stereoselective synthesis of **TMM (C32, *RR*)** was carried out by a modified method of Nishizawa et al. (Nishizawa et al., 2007): $[\alpha]_D$ +28.8 (c 0.49, $CHCl_3$/MeOH = 1:1); ^1H NMR (500 MHz, $CDCl_3/CD_3OD$ = 1:1) δ 5.10 (d, *J* = 4.0 Hz, 2H), 4.49 (dd, *J* = 12.0, 2.3 Hz, 1H), 4.19 (dd, *J* = 12.0, 5.7 Hz, 1H), 4.11–4.04 (ddd, *J* = 9.7, 5.7, 2.3 Hz, 1H), 3.85–3.76 (m, 4H), 3.72–3.66 (m, 2H), 3.55–3.46 (m, 2H), 3.38–3.30 (m, 2H), 2.48–2.40 (m, 1H), 2.21–1.94 (m, 2H), 1.68–1.15 (m, 52H), 0.89 (t, *J* = 6.9 Hz, 6H); $^{13}C\{^1H\}$

NMR (150 MHz, $CDCl_3/CD_3OD$ = 1:1) δ 174.7, 93.4, 93.3, 72.6, 72.5, 71.8, 71.22, 71.17, 70.2, 70.1, 69.5, 62.8, 61.0, 52.2, 33.9, 31.2, 29.0, 28.94, 28.89, 28.86, 28.8, 28.7, 28.6, 28.2, 26.7, 24.7, 21.9, 21.6, 13.0; HRMS (ESI-TOF) *m/z*: $[M + Na]^+$ calcd for $C_{44}H_{84}NaO_{13}$, 843.5804; found, 843.5778.

Antibodies

Anti-human CD3 (HIT3a), anti-human CD19 (SJ25C1), TotalSeq-C Hashtags (LNH-94; 2M2), anti-mouse CD3 (2C11, 17A2), anti-mouse CD69 (H1.2F3), and anti-rat CD2 (OX-34), anti-human CD1a (HI149), anti-human CD1b (SN13), anti-human CD1c (L161), anti-human CD1d (51.1), mouse IgG1 κ isotype control (MG1-45) and mouse IgG2b κ isotype control (MPC-11) antibodies were purchased from BioLegend.

***In vitro* stimulation of PBMCs**

Cryopreserved human PBMCs were thawed and labeled by CellTrace Violet (CTV, Thermo Fisher Scientific), then quenched and washed by RPMI 1640 medium (Sigma) supplemented with 5% human AB serum (Gemini Bio), penicillin (Sigma), streptomycin (MP Biomedicals), and 2-mercaptoethanol (Nacalai Tesque). 10^6 CTV labeled PBMCs were stimulated in the same medium with plate-coated 3 μg of Mtb-crude lipids, 3 μg of synthetic GMM or 3 μg of heat-killed Mtb H37Rv for 10 days. Recombinant human IL-2 (1 ng/m L, PeproTech), human IL-7 (5 ng/m L, PeproTech) human IL-15 (5 ng/m L, PeproTech) were added at day 2, day 5 and day 8. After staining with anti-human CD3

antibody, $CTV^{low}CD3^{+}$ cells were sorted by SH800 Cell Sorter (Sony Biotechnology) and used for single-cell TCR- and RNA-seq analyses.

Single-cell–based transcriptome and TCR repertoire analysis

Single-cell transcriptome and TCR repertoire analysis were performed using Chromium Controller (10x Genomics) according to the manufacturer's instructions as previously described (Lu et al., 2021). Libraries were sequenced on Illumina NovaSeq 6000 in the paired-end mode. The raw reads were processed by Cell Ranger v6.0.0-7.1.0 (10x Genomics). TCR repertoire analysis was conducted using Scirpy (v0.11.1) and gene expression-based clustering was performed using Scanpy (1.9.1). UMAP plot, heatmap, volcano plot and differential expression analysis were performed by Seurat R package (v5.0.1).

Bulk TCR-sequencing

3×10^5 PBMCs were lysed in QIAzol (QIAGEN). Full-length cDNA was then synthesized using SMARTer technology (Takara Bio), and the variable regions of TCR α and TCRβ genes were amplified using TRAC/TRBC-specific primers. After sequencing of the variable region amplicons, each pair of reads was assigned a clonotype (defined as TR(A/B)V and TR(A/B)J genes and CDR3) using MiXCR software (Bolotin et al., 2015).

APCs

For the preparation of cytokine-differentiated human monocytes, $CD14^+$ monocytes were sorted from freshly isolated human PBMCs using MACS cell separation column (Miltenyi Biotec), followed by cultured in RPMI 1640 supplemented with 10% FBS, non-essential amino acids, 10 ng/mL human GM-CSF, and 10 ng/mL human IL-4 for 7 days. Human CD1b was cloned into the retroviral vector pMX-IRES-hCD8 (Yamasaki et al., 2006) using Phoenix packaging cells and PEI MAX (Polysciences). Supernatant containing retroviruses was used for infection into the mouse dendritic cell line DC2.4 (ATTC).

TCR reconstitution and stimulation

TCR α and β chain cDNA sequences were synthesized with eblock (IDT) and cloned into retroviral vectors pMX-IRES-rat CD2. TCRα mutants were constructed by site-directed mutagenesis. The vectors were transduced into mouse T cell hybridoma with an NFAT-GFP reporter gene (Matsumoto et al., 2021) (Lu et al., 2021) using retroviruses described above to reconstitute TCRαβ pairs. For antigen stimulation, TCR-reconstituted cells were cocultured with stimulants in the presence of APCs unless indicated otherwise. After 20 h, T cell activation was assessed by GFP and CD69 expression.

NGS-based mutagenesis scanning

For the mutant libraries, synthesized mutant TCRα or β cDNA sequences were pooled, and reconstituted into reporter cell lines with WT TCRβ or TCRα, respectively. Library cells were left unstimulated or stimulated with TMM for 20 h and then sorted by GFP

negative/positive populations. Each sorted population was analyzed by bulk TCR-sequencing (GSE261269). The proportion of the read counts of each mutant within GFP+ or GFP$^-$ population were shown as percentage.

CD1b tetramers

Unloaded human CD1b monomers (biotinylated) were obtained from the NIH tetramer facility. For TMM loading, 16 μg of Mtb TMM was sonicated at 45°C for 1 h in 45 μL of 0.5% CHAPS 50 mM sodium citrate buffer (pH 4.5), added to 5 μL of CD1b monomers (2 mg/mL) and incubated overnight at 37°C. For the preparation of control CD1b tetramers (endo-CD1b tet), CD1b monomers were treated as described above without TMM loading. Monomers were then neutralized by 5 μL 1 M Tris (pH 8) and tetramerized using streptavidin-PE (Biolegend) or streptavidin-APC (eBioscience).

Tetramer staining and isolation of CD1b tetramer$^+$ T cells

10^7 human PBMCs were incubated with 20 μg/mL of PE-conjugated TMM-CD1btetramers and 20 μg/mL of APC-conjugated endo-CD1b tetramers in 40 μL 1%BSA/PBS at room temperature in the dark for 15 min. Without washing, 2 μL of human Fc block (50 μg/mL), 60 μL of anti-human CD3-FITC (2 μg/mL) and 2 μL of TotalSeq-C Hashtags were added and incubated on ice for 20 min. Before sorting, cells were washed and filtered by nylon mesh, and incubated with propidium iodide. TMM-tetramer$^+$ cells within the CD3$^+$ gated population were sorted by SH800 Cell Sorter (Sony Biotechnology) and subjected to sc-TCR-RNA-seq. Tetramer$^+$ clonotypes were further

validated by its enrichment score derived from clonotype frequencies in tetramer$^+$ T cells/clonotype frequencies in the original T cell population.

Crystal structural analysis

Complementary DNA encoding the ectodomains of Y-50 TCR α (from Ala-1 to S-205) and β (from Asp-1 to Asp-249) with a nidogen signal sequence, a 6 × His-tag and a tobacco etch virus protease cleavage site at the N-terminal were synthesized (Thermo GeneArt) and cloned into pcDNA3.1(+) vector. To improve the efficiency of protein expression, artificial disulfide bond and stabilizing mutations were introduced as described previously (Boulter et al., 2003; Froning et al., 2020). The plasmids were transformed into 293expi cells in the presence of the mannosidase inhibitor kifunensine. The cells were cultured with shaking at 120 rpm 37°C 8% CO_2 for 4 days. After being filtered through a 0.22 μm, the supernatant was applied onto 5 mL nickel-nitrilotriacetic acid agarose (FUJIFILM Wako), and His-tagged TCRαβ were eluted with elution buffer (50 mM Tris-HCl [pH 8.0], 300 mM NaCl, and 250 mM imidazole). After removal of His-tag by tobacco etch virus protease, the eluted protein was concentrated and further applied to Superdex 75 (Cytiva) equilibrated with 20 mM Tris-HCl (pH 8.0) buffer containing 100 mM NaCl. The crystals were formed by the sitting-drop vapor-diffusion method. 0.4 μL of protein solution (5 mg/mL in 100 mM NaCl, 20 mM Tis-HCl (pH8.0)) was mixed with 0.4 μL of mother liquid containing 0.2 M potassium sulfate 0.1M Bis-Tris pH 5.5, 25% PEG3350 and incubated at 20°C. The diffraction data were collected in a cold nitrogen gas stream on an EIGER X 9M detector (DECTRIS) with a wavelength

of 1.0 Å. The resulting datasets were processed, merged, and scaled using XDS (Kabsch, 2010). Structures were clarified by molecular replacement with the TCR complex (PDB 8ZO4) as a search model, by MOLREP) as implemented in CCP4i software (Vagin and Teplyakov, 2010). The models were refined using REFMAC5 and PHENIX1.20 software (Vagin et al., 2004) , (Liebschner et al., 2019). The structures were rebuilt using COOT 0.9.8.92 and further modified based on σ-weighted ($2|F_{obs}|—|F_{calc}|$) and ($|F_{obs}|—|F_{calc}|$) electron density maps. Crystallographic images were created using PyMOL software (Schrödinger, LLC.).

Cryo-EM structural analysis

Trehalose monomycolate (TMM) loaded CD1b ectodomain was refolded as follows: Denatured proteins of CD1b (24.8 mg) and β2m (9.6 mg) were mixed with 3.28 mg TMM and refolded in the buffer containing 1 M L-arginine (pH 8.0), 5 M urea, 5 mM reduced glutathione, and 0.5 mM oxidized glutathione. The refolded proteins were then dialyzed four times against 0.01 M Tris-HCl (pH 8.0) and applied onto HiTrapTM Q HP 5 mL column (Cytiva). Purified TMM-loaded CD1b was mixed with Y-50 TCR at the ratio of 1:1. 2.2 μl of sample (1.0 mg/ml) was applied onto the glow-discharged Quantifoil Au 0.6/1.0 200 mesh grid (Quantifoil Micro Tools GmbH, Germany) and frozen in liquid ethane using a Vitrobot IV (FEI, 4°C and 95% humidity). Cryo-EM data collection was performed on a Titan Krios cryo-TEM equipped with a Cs corrector (Thermo Fisher Scientific, USA) operating at 300 keV in EFTEM nanoprobe mode at Institute for Protein Research, Osaka University. Images were acquired as movies using Gatan BioQuantum

energy filter (slit width of 20 eV) and K3 direct electron detector camera (Gatan, Inc., USA) in electron counting mode. A total of 8,533 movies were collected at a dose rate of 8.532 e^-/pixel/s, a pixel size of 0.675 $Å^2$, and a total dose of 60 $e^-/Å^2$. SerialEM software (Schorb et al., 2019) was used for automated data collection using a 3 × 3-hole pattern beam-image shift scheme with a nominal defocus range of −0.6 to −1.8 μm. All of image processes were carried out on cryoSPARC v4.4.1 software (Punjani et al., 2017). After motion correction of movies and CTF parameter estimation, an initial round of particle picking was performed using the blob picker tool (diameter 100 - 140 Å). After four iterations of 2D class and manual selection, 12,424 particles were selected. Classification into three classes using Ab Initio Reconstruction and manual selection was repeated twice, and the resulting 6,961 particles were used as training data for Topaz picks. The 2,266,363 particles were automatically picked using Topaz picking algorithm. After two rounds of 2D classification and 3D classification using Ab Initio Reconstruction and Heterogeneous refinement, 599,402 particles were selected. A subsequent round of 2D classification further narrowed the selection to 232,909 particles. Ab Initio Reconstruction with C1 symmetry. As the result, three distinguishable classes were obtained. One of these classified particles (114,111particles) were applied to Non-uniform Refinement. The density maps from the refinement were obtained at 3.31 Å resolution. Each particle was subjected to Reference Based Motion Correction. As the results of Non-uniform Refinement, a map of the complex at 3.18 Å resolution was obtained. Local resolution of the obtained map was estimated by Local resolution estimation job on cryoSPARC. 3D structures of Y-50 (PDB: 8XUB) and CD1b

(PDB: 5L2K) were automatically fitted into map with program phenix.dock_in_map in PHENIX program suite (Liebschner et al., 2019). Chemical structure of TMM was idealized by phenix.elbow. The atomic model of ternary complex was manually modified using COOT (Emsley et al., 2010) and refined with phenix.real_space_refine of PHENIX suite. Stereochemistry of the refined structure was evaluated with MolProbity (Williams et al., 2018). Validation of the final model is summarized in Supplementary Table 3.

References

Beckman, E.M., S.A. Porcelli, C.T. Morita, S.M. Behar, S.T. Furlong, and M.B. Brenner. 1994. Recognition of a lipid antigen by CD1-restricted alpha beta+ T cells. *Nature* 372:691-694.

Bendelac, A., P.B. Savage, and L. Teyton. 2007. The biology of NKT cells. *Annu Rev Immunol* 25:297-336.

Bhatt, A., N. Fujiwara, K. Bhatt, S.S. Gurcha, L. Kremer, B. Chen, J. Chan, S.A. Porcelli, K. Kobayashi, G.S. Besra, and W.R. Jacobs, Jr. 2007. Deletion of kasB in Mycobacterium tuberculosis causes loss of acid-fastness and subclinical latent tuberculosis in immunocompetent mice. *Proc Natl Acad Sci U S A* 104:5157-5162.

Boulter, J.M., M. Glick, P.T. Todorov, E. Baston, M. Sami, P. Rizkallah, and B.K. Jakobsen. 2003. Stable, soluble T-cell receptor molecules for crystallization and therapeutics. *Protein Eng* 16:707-711.

Brown, G.D., J.A. Willment, and L. Whitehead. 2018. C-type lectins in immunity and homeostasis. *Nat Rev Immunol* 18:374-389.

Daley, C.L., J.M. Iaccarino, C. Lange, E. Cambau, R.J. Wallace, Jr., C. Andrejak, E.C. Bottger, J. Brozek, D.E. Griffith, L. Guglielmetti, G.A. Huitt, S.L. Knight, P. Leitman, T.K. Marras, K.N. Olivier, M. Santin, J.E. Stout, E. Tortoli, J. van Ingen, D. Wagner, and K.L. Winthrop. 2020. Treatment of nontuberculous mycobacterial pulmonary disease: an official ATS/ERS/ESCMID/IDSA clinical practice guideline. *Eur Respir J* 56:

De Libero, G., A. Singhal, M. Lepore, and L. Mori. 2014. Nonclassical T cells and their antigens in tuberculosis. *Cold Spring Harb Perspect Med* 4:a018473.

Decout, A., S. Silva-Gomes, D. Drocourt, S. Barbe, I. Andre, F.J. Cueto, T. Lioux, D. Sancho, E. Perouzel, A. Vercellone, J. Prandi, M. Gilleron, G. Tiraby, and J. Nigou. 2017. Rational design of adjuvants targeting the C-type lectin Mincle. *Proc Natl Acad Sci U S A* 114:2675-2680.

Dube, J.Y., F. McIntosh, J.G. Zarruk, S. David, J. Nigou, and M.A. Behr. 2020. Synthetic mycobacterial molecular patterns partially complete Freund's adjuvant. *Sci Rep* 10:5874.

Dulberger, C.L., E.J. Rubin, and C.C. Boutte. 2020. The mycobacterial cell envelope - a

moving target. *Nat Rev Microbiol* 18:47-59.

Emsley, P., B. Lohkamp, W.G. Scott, and K. Cowtan. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486-501.

Engel, I., and M. Kronenberg. 2012. Making memory at birth: understanding the differentiation of natural killer T cells. *Curr Opin Immunol* 24:184-190.

Flynn, J.L., J. Chan, K.J. Triebold, D.K. Dalton, T.A. Stewart, and B.R. Bloom. 1993. An essential role for interferon gamma in resistance to Mycobacterium tuberculosis infection. *J Exp Med* 178:2249-2254.

Flynn, J.L., M.M. Goldstein, J. Chan, K.J. Triebold, K. Pfeffer, C.J. Lowenstein, R. Schreiber, T.W. Mak, and B.R. Bloom. 1995. Tumor necrosis factor-alpha is required in the protective immune response against Mycobacterium tuberculosis in mice. *Immunity* 2:561-572.

Freund, J., J. Casals, and E.P. Hosmer. 1937. Sensitization and antibody formation after injection of tubercle bacilli and paraffin oil. *Proceedings of the Society for Experimental Biology and Medicine* 37:509-513.

Froning, K., J. Maguire, A. Sereno, F. Huang, S. Chang, K. Weichert, A.J. Frommelt, J. Dong, X. Wu, H. Austin, E.M. Conner, J.R. Fitchett, A.R. Heng, D. Balasubramaniam, M.T. Hilgers, B. Kuhlman, and S.J. Demarest. 2020. Computational stabilization of T cell receptors allows pairing with antibodies to form bispecifics. *Nat Commun* 11:2330.

Fujita, Y., T. Naka, T. Doi, and I. Yano. 2005. Direct molecular mass determination of trehalose monomycolate from 11 species of mycobacteria by MALDI-TOF mass spectrometry. *Microbiology (Reading)* 151:1443-1452.

Fujiwara, N., T. Naka, M. Ogawa, R. Yamamoto, H. Ogura, and H. Taniguchi. 2012. Characteristics of Mycobacterium smegmatis J15cs strain lipids. *Tuberculosis (Edinb)* 92:187-192.

Fujiwara, N., S. Oka, M. Ide, K. Kashima, T. Honda, and I. Yano. 1999. Production and partial characterization of antibody to cord factor (trehalose 6,6'-dimycolate) in mice. *Microbiol Immunol* 43:785-793.

Geldmacher, C., N. Ngwenyama, A. Schuetz, C. Petrovas, K. Reither, E.J. Heeregrave, J.P. Casazza, D.R. Ambrozak, M. Louder, W. Ampofo, G. Pollakis, B. Hill, E. Sanga, E. Saathoff, L. Maboko, M. Roederer, W.A. Paxton, M. Hoelscher, and R.A. Koup. 2010. Preferential infection and depletion of Mycobacterium

tuberculosis-specific CD4 T cells after HIV-1 infection. *J Exp Med* 207:2869-2881.

Gherardin, N.A., S.J. Redmond, H.E.G. McWilliam, C.F. Almeida, K.H.A. Gourley, R. Seneviratna, S. Li, R. De Rose, F.J. Ross, C.V. Nguyen-Robertson, S. Su, M.E. Ritchie, J.A. Villadangos, D.B. Moody, D.G. Pellicci, A.P. Uldrich, and D.I. Godfrey. 2021. CD36 family members are TCR-independent ligands for CD1 antigen-presenting molecules. *Sci Immunol* 6:

Gras, S., I. Van Rhijn, A. Shahine, T.Y. Cheng, M. Bhati, L.L. Tan, H. Halim, K.D. Tuttle, L. Gapin, J. Le Nours, D.B. Moody, and J. Rossjohn. 2016. T cell receptor recognition of CD1b presenting a mycobacterial glycolipid. *Nat Commun* 7:13257.

Guo, T., M.Y. Koo, Y. Kagoya, M. Anczurowski, C.H. Wang, K. Saso, M.O. Butler, and N. Hirano. 2018. A Subset of Human Autoreactive CD1c-Restricted T Cells Preferentially Expresses TRBV4-1(+) TCRs. *J Immunol* 200:500-511.

Harris, S.P., N. Fujiwara, R.H. Mealey, D.C. Alperin, T. Naka, R. Goda, and S.A. Hines. 2010. Identification of Rhodococcus equi lipids recognized by host cytotoxic T lymphocytes. *Microbiology (Reading)* 156:1836-1847.

Huang, S., A. Shahine, T.Y. Cheng, Y.L. Chen, S.W. Ng, G.R. Balaji, R. Farquhar, S. Gras, C.S. Hardman, J.D. Altman, N. Tahiri, A.J. Minnaard, G.S. Ogg, J.A. Mayfield, J. Rossjohn, and D.B. Moody. 2023. CD1 lipidomes reveal lipid-binding motifs and size-based antigen-display mechanisms. *Cell* 186:4583-4596 e4513.

Ishikawa, E., T. Ishikawa, Y.S. Morita, K. Toyonaga, H. Yamada, O. Takeuchi, T. Kinoshita, S. Akira, Y. Yoshikai, and S. Yamasaki. 2009. Direct recognition of the mycobacterial glycolipid, trehalose dimycolate, by C-type lectin Mincle. *J Exp Med* 206:2879-2888.

Ishikawa, E., D. Mori, and S. Yamasaki. 2017. Recognition of Mycobacterial Lipids by Immune Receptors. *Trends Immunol* 38:66-76.

Kabsch, W. 2010. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr* 66:133-144.

Keane, J., S. Gershon, R.P. Wise, E. Mirabile-Levens, J. Kasznica, W.D. Schwieterman, J.N. Siegel, and M.M. Braun. 2001. Tuberculosis associated with infliximab, a tumor necrosis factor alpha-neutralizing agent. *N Engl J Med* 345:1098-1104.

Kiyotake, R., M. Oh-Hora, E. Ishikawa, T. Miyamoto, T. Ishibashi, and S. Yamasaki. 2015. Human Mincle Binds to Cholesterol Crystals and Triggers Innate Immune

Responses. *J Biol Chem* 290:25322-25332.

Layre, E., A. Collmann, M. Bastian, S. Mariotti, J. Czaplicki, J. Prandi, L. Mori, S. Stenger, G. De Libero, G. Puzo, and M. Gilleron. 2009. Mycolic acids constitute a scaffold for mycobacterial lipid antigens stimulating CD1-restricted T cells. *Chem Biol* 16:82-92.

Liebschner, D., P.V. Afonine, M.L. Baker, G. Bunkoczi, V.B. Chen, T.I. Croll, B. Hintze, L.W. Hung, S. Jain, A.J. McCoy, N.W. Moriarty, R.D. Oeffner, B.K. Poon, M.G. Prisant, R.J. Read, J.S. Richardson, D.C. Richardson, M.D. Sammito, O.V. Sobolev, D.H. Stockwell, T.C. Terwilliger, A.G. Urzhumtsev, L.L. Videau, C.J. Williams, and P.D. Adams. 2019. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* 75:861-877.

Lu, X., Y. Hosono, M. Nagae, S. Ishizuka, E. Ishikawa, D. Motooka, Y. Ozaki, N. Sax, Y. Maeda, Y. Kato, T. Morita, R. Shinnakasu, T. Inoue, T. Onodera, T. Matsumura, M. Shinkai, T. Sato, S. Nakamura, S. Mori, T. Kanda, E.E. Nakayama, T. Shioda, T. Kurosaki, K. Takeda, A. Kumanogoh, H. Arase, H. Nakagami, K. Yamashita, Y. Takahashi, and S. Yamasaki. 2021. Identification of conserved SARS-CoV-2 spike epitopes that expand public cTfh clonotypes in mild COVID-19 patients. *J Exp Med* 218:

Matsumoto, Y., K. Kishida, M. Matsumoto, S. Matsuoka, M. Kohyama, T. Suenaga, and H. Arase. 2021. A TCR-like antibody against a proinsulin-containing fusion peptide ameliorates type 1 diabetes in NOD mice. *Biochem Biophys Res Commun* 534:680-686.

Matsunaga, I., T. Naka, R.S. Talekar, M.J. McConnell, K. Katoh, H. Nakao, A. Otsuka, S.M. Behar, I. Yano, D.B. Moody, and M. Sugita. 2008. Mycolyltransferase-mediated glycolipid exchange in Mycobacteria. *J Biol Chem* 283:28835-28841.

Matsunaga, I., S. Oka, N. Fujiwara, and I. Yano. 1996. Relationship between induction of macrophage chemotactic factors and formation of granulomas caused by mycoloyl glycolipids from Rhodococcus ruber (Nocardia rubra). *J Biochem* 120:663-670.

Moody, D.B., B.B. Reinhold, M.R. Guy, E.M. Beckman, D.E. Frederique, S.T. Furlong, S. Ye, V.N. Reinhold, P.A. Sieling, R.L. Modlin, G.S. Besra, and S.A. Porcelli. 1997. Structural requirements for glycolipid antigen recognition by CD1b-

restricted T cells. *Science* 278:283-286.

Nagata, M., Y. Izumi, E. Ishikawa, R. Kiyotake, R. Doi, S. Iwai, Z. Omahdi, T. Yamaji, T. Miyamoto, T. Bamba, and S. Yamasaki. 2017. Intracellular metabolite beta-glucosylceramide is an endogenous Mincle ligand possessing immunostimulatory activity. *Proc Natl Acad Sci U S A* 114:E3285-E3294.

Naka, T., S. Maeda, M. Niki, N. Ohara, S. Yamamoto, I. Yano, J.I. Maeyama, H. Ogura, K. Kobayashi, and N. Fujiwara. 2011. Lipid phenotype of two distinct subpopulations of Mycobacterium bovis Bacillus Calmette-Guerin Tokyo 172 substrain. *J Biol Chem* 286:44153-44161.

Nishimura, N., N. Tomiyasu, S. Torigoe, S. Mizuno, H. Fukano, E. Ishikawa, H. Katano, Y. Hoshino, K. Matsuo, M. Takahashi, Y. Izumi, T. Bamba, K. Akashi, and S. Yamasaki. 2023. Mycobacterial mycolic acids trigger inhibitory receptor Clec12A to suppress host immune responses. *Tuberculosis (Edinb)* 138:102294.

Nishizawa, M., H. Yamamoto, H. Imagawa, V. Barbier-Chassefiere, E. Petit, I. Azuma, and D. Papy-Garcia. 2007. Efficient syntheses of a series of trehalose dimycolate (TDM)/trehalose dicorynomycolate (TDCM) analogues and their interleukin-6 level enhancement activity in mice sera. *J Org Chem* 72:1627-1633.

Nunes-Alves, C., M.G. Booty, S.M. Carpenter, P. Jayaraman, A.C. Rothchild, and S.M. Behar. 2014. In search of a new paradigm for protective immunity to TB. *Nat Rev Microbiol* 12:289-299.

Punjani, A., J.L. Rubinstein, D.J. Fleet, and M.A. Brubaker. 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* 14:290-296.

Reinink, P., A. Shahine, S. Gras, T.Y. Cheng, R. Farquhar, K. Lopez, S.A. Suliman, J.F. Reijneveld, J. Le Nours, L.L. Tan, S.R. Leon, J. Jimenez, R. Calderon, L. Lecca, M.B. Murray, J. Rossjohn, D.B. Moody, and I. Van Rhijn. 2019. A TCR beta-Chain Motif Biases toward Recognition of Human CD1 Proteins. *J Immunol* 203:3395-3406.

Savage, A.K., M.G. Constantinides, J. Han, D. Picard, E. Martin, B. Li, O. Lantz, and A. Bendelac. 2008. The transcription factor PLZF directs the effector program of the NKT cell lineage. *Immunity* 29:391-403.

Schorb, M., I. Haberbosch, W.J.H. Hagen, Y. Schwab, and D.N. Mastronarde. 2019. Software tools for automated transmission electron microscopy. *Nat Methods*

16:471-477.

Seach, N., L. Guerri, L. Le Bourhis, Y. Mburu, Y. Cui, S. Bessoles, C. Soudais, and O. Lantz. 2013. Double-positive thymocytes select mucosal-associated invariant T cells. *J Immunol* 191:6002-6009.

Shao, M.M., F.S. Yi, Z.Y. Huang, P. Peng, F.Y. Wu, H.Z. Shi, and K. Zhai. 2022. T Cell Receptor Repertoire Analysis Reveals Signatures of T Cell Responses to Human Mycobacterium tuberculosis. *Front Microbiol* 13:829694.

Stenger, S., D.A. Hanson, R. Teitelbaum, P. Dewan, K.R. Niazi, C.J. Froelich, T. Ganz, S. Thoma-Uszynski, A. Melian, C. Bogdan, S.A. Porcelli, B.R. Bloom, A.M. Krensky, and R.L. Modlin. 1998. An antimicrobial activity of cytolytic T cells mediated by granulysin. *Science* 282:121-125.

Su, C.C., P.A. Klenotic, J.R. Bolla, G.E. Purdy, C.V. Robinson, and E.W. Yu. 2019. MmpL3 is a lipid transporter that binds trehalose monomycolate and phosphatidylethanolamine. *Proc Natl Acad Sci U S A* 116:11241-11246.

Tahiri, N., P. Fodran, D. Jayaraman, J. Buter, M.D. Witte, T.A. Ocampo, D.B. Moody, I. Van Rhijn, and A.J. Minnaard. 2020. Total Synthesis of a Mycolic Acid from Mycobacterium tuberculosis. *Angew Chem Int Ed Engl* 59:7555-7560.

Ueda, S., N. Fujiwara, T. Naka, I. Sakaguchi, Y. Ozeki, I. Yano, T. Kasama, and K. Kobayashi. 2001. Structure-activity relationship of mycoloyl glycolipids derived from Rhodococcus sp. 4306. *Microb Pathog* 30:91-99.

Vagin, A., and A. Teplyakov. 2010. Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* 66:22-25.

Vagin, A.A., R.A. Steiner, A.A. Lebedev, L. Potterton, S. McNicholas, F. Long, and G.N. Murshudov. 2004. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr D Biol Crystallogr* 60:2184-2195.

Van Rhijn, I., N.A. Gherardin, A. Kasmar, W. de Jager, D.G. Pellicci, L. Kostenko, L.L. Tan, M. Bhati, S. Gras, D.I. Godfrey, J. Rossjohn, and D.B. Moody. 2014. TCR bias and affinity define two compartments of the CD1b-glycolipid-specific T Cell repertoire. *J Immunol* 192:4054-4060.

Van Rhijn, I., and D.B. Moody. 2015. CD1 and mycobacterial lipids activate human T cells. *Immunol Rev* 264:138-153.

Vesosky, B., E.K. Rottinghaus, P. Stromberg, J. Turner, and G. Beamer. 2010. CCL5

participates in early protection against Mycobacterium tuberculosis. *J Leukoc Biol* 87:1153-1165.

Williams, C.J., J.J. Headd, N.W. Moriarty, M.G. Prisant, L.L. Videau, L.N. Deis, V. Verma, D.A. Keedy, B.J. Hintze, V.B. Chen, S. Jain, S.M. Lewis, W.B. Arendall, 3rd, J. Snoeyink, P.D. Adams, S.C. Lovell, J.S. Richardson, and D.C. Richardson. 2018. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci* 27:293-315.

Wu, M.L., D.B. Aziz, V. Dartois, and T. Dick. 2018. NTM drug discovery: status, gaps and the way forward. *Drug Discov Today* 23:1502-1519.

Yamasaki, S., E. Ishikawa, M. Sakuma, K. Ogata, K. Sakata-Sogawa, M. Hiroshima, D.L. Wiest, M. Tokunaga, and T. Saito. 2006. Mechanistic basis of pre-T cell receptor-mediated autonomous signaling critical for thymocyte development. *Nat Immunol* 7:67-75.

Fig. 1

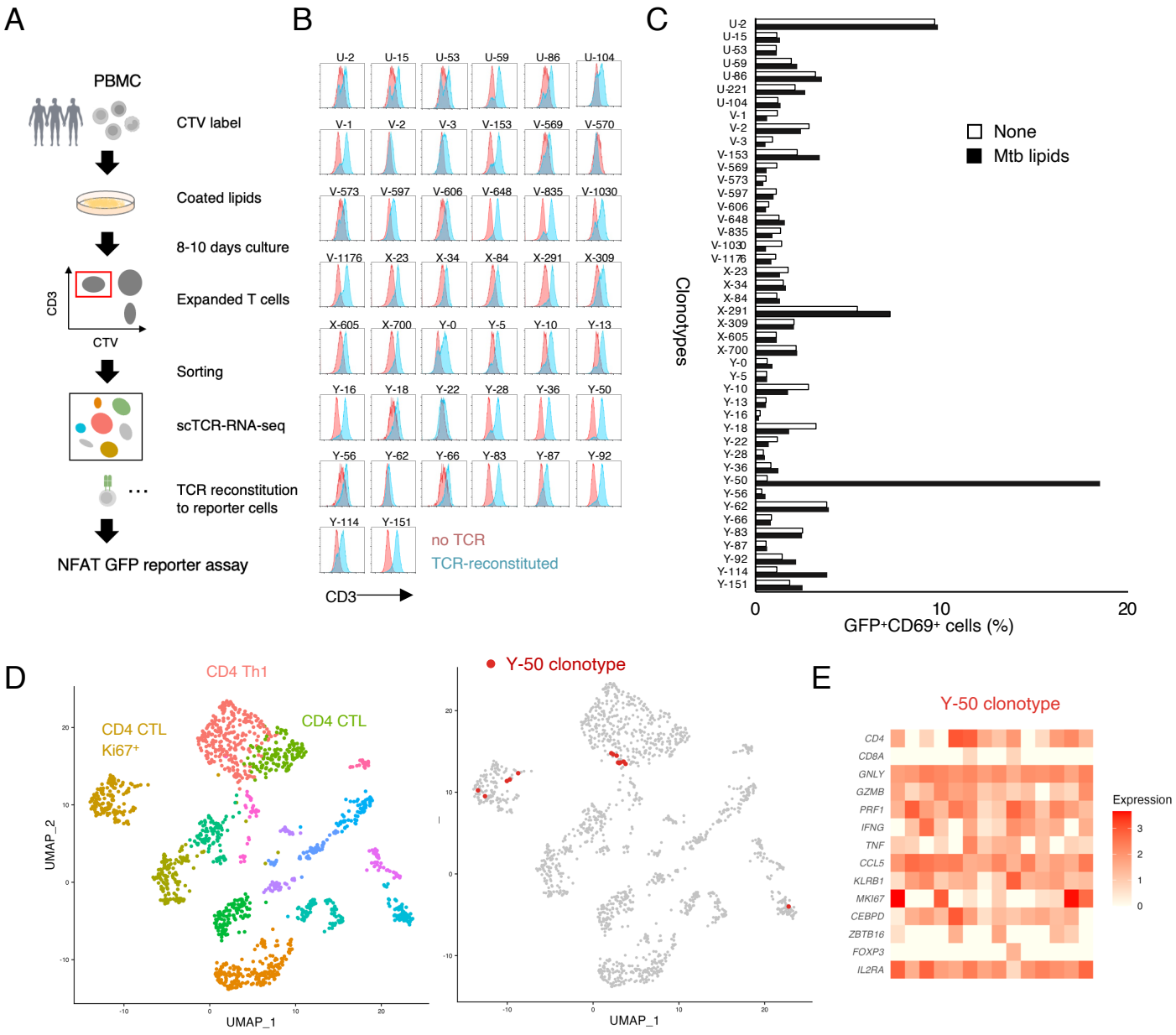


Figure 1. Identification of mycobacterial lipid-reactive T cells.

A) Schematic representation of experimental procedure. Human PBMCs were cultured with plate-coated crude lipids extracted from *M. tb*. The expanded T cells were sorted and analyzed by single-sell RNA-TCR-sequencing. Highly expanded CTV^{lo} TCR clonotypes were reconstituted into NFAT-GFP reporter cells to examine the reactivity to *M. tb* lipids.

B) Fifty-two TCR clonotypes were reconstituted into reporter cells and analyzed for their surface expression using anti-CD3 antibody.

C) NFAT-GFP reporter cells (44 clonotypes) expressing each different TCR were stimulated with *M. tb* crude lipids in the presence of PMBCs or cytokine-differentiated monocytes as APCs, and after 20 h incubation, analyzed for GFP and CD69 expression. Representative results from two independent experiments are shown.

D) UMAP plot of T cells expanded in response to *M. tb* lipids (upper panel). T cell clones expressing Y-50 clonotype are highlighted in red dots (lower panel).

E) Gene expression signature of Y-50 cells. The expression of characteristic genes in each cell expressing Y-50 TCR clonotype are shown. The average expression in cells within CD4 Th1 and CD4 CTL clusters are shown as references.

Fig. 2

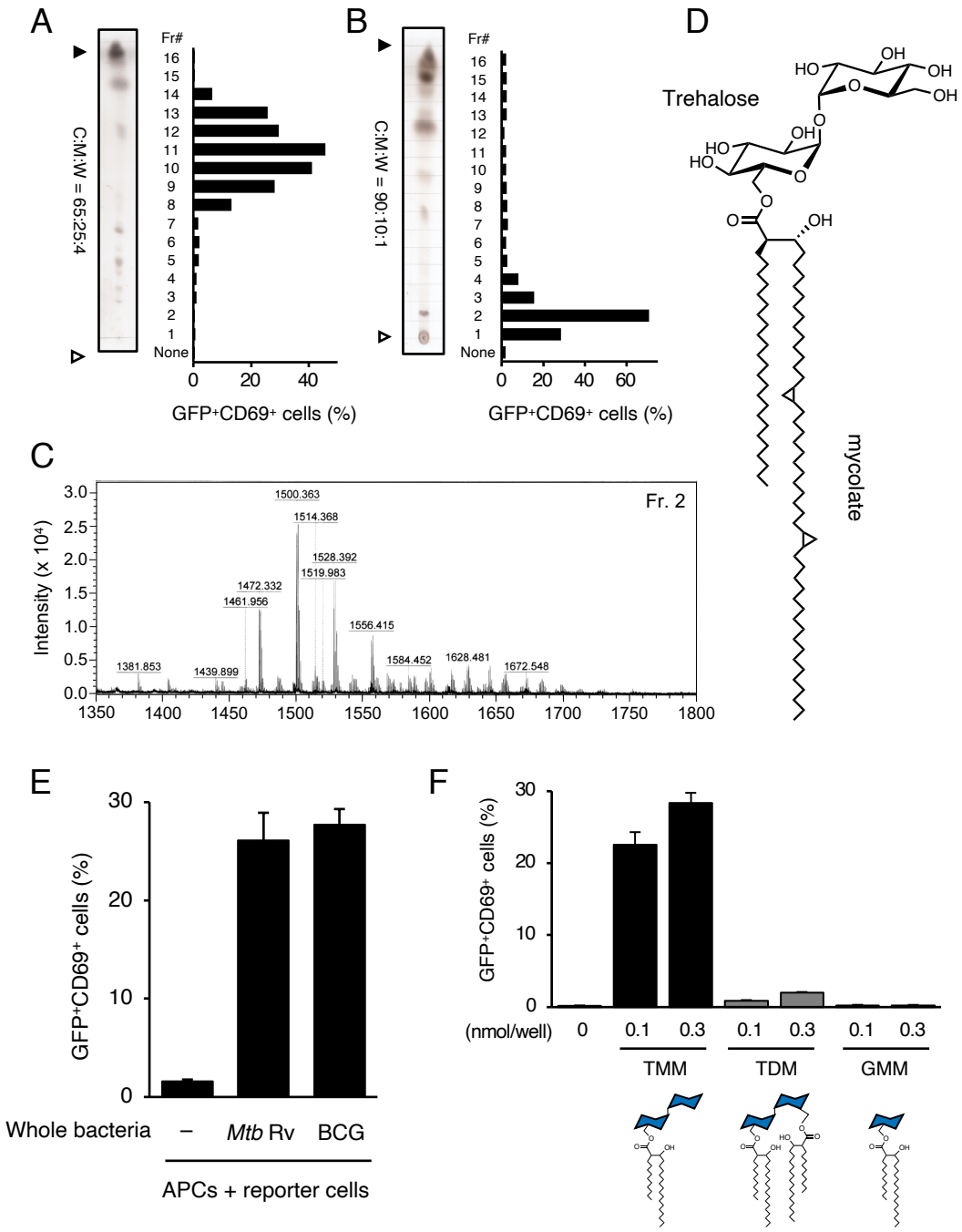


Figure 2. Identification of TMM as a T cell antigen.

A-B) *M. tuberculosis* H37Rv crude lipids were fractionated by HPTLC using chloroform/methanol/water (65:25:4; vol/vol/vol (A) and 90:10:1; vol/vol/vol (B)) and stained with copper(II) acetate-phosphoric acid. Y-50 reporter cells were stimulated with each fraction in the presence of APCs and analyzed for GFP and CD69 expression. Open and closed arrowheads denote the origin and the solvent front, respectively.

C) MALDI-TOF MS spectrum of lipid fraction 2 (Fr2).

D) Chemical structure of TMM of α-mycolate is shown, and methoxy-mycolate, and keto-mycolate are the other major subclasses of mycolate found in *M. tuberculosis* TMM.

E) Y-50 reporter cells were co-cultured with cytokine-differentiated human monocytes pre-incubated with whole bacteria (heat killed *M. tb* H37Rv or living BCG) and analyzed for GFP and CD69 expression.

F) Y-50 reporter cells were stimulated with the indicated concentration of TMM, TDM or GMM. GFP and CD69 expression are shown in dot plots (upper) and bar graphs (lower). Schematic ligand structures are shown below.

Data are shown as the mean ± SD for triplicate assays (E, F) and representative results from two independent experiments are shown (A, B, E and F).

Fig. 3

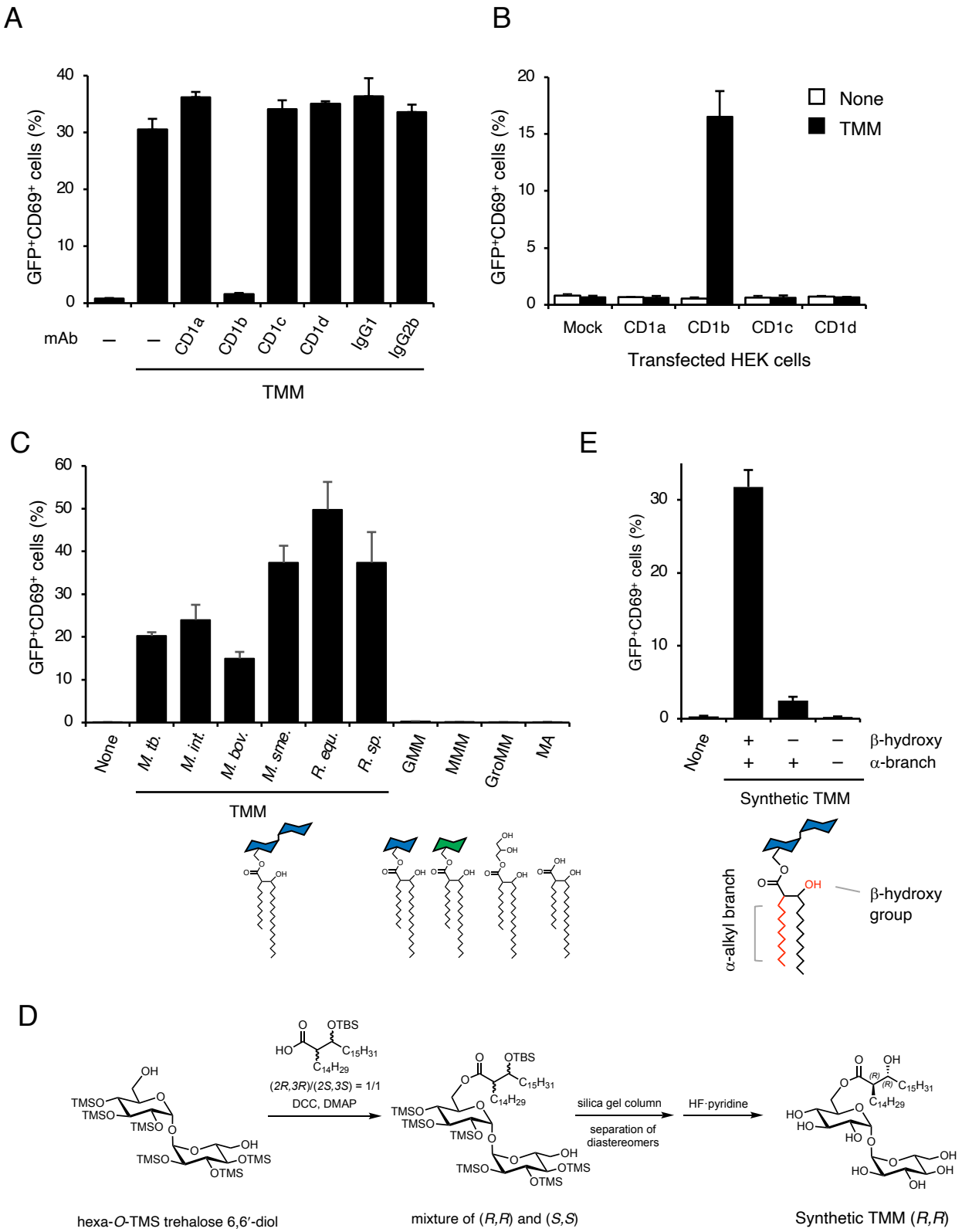


Figure 3. CD1b restricts TMM recognition by Y-50 T cells.

A) Y-50 reporter cells were co-cultured with cytokine-differentiated human monocytes and TMM (0.3 nmol /well) in the presence of 5 μg/mL of anti-CD1a, CD1b, CD1c, CD1d or isotype control antibodies (IgG1 and IgG2b) and analyzed for GFP and CD69 expression.

B) The reporter cells expressing Y-50 TCR were stimulated with TMM (1 nmol/well) in the presence of HEK293T cells transfected with human CD1a, CD1b, CD1c or CD1d.

C) Y-50 reporter cells were stimulated with TMM (1 nmol/well) purified from *M. tuberculosis* CDC1551, *M. bovis* BCG, *M. intracellulare, M. smegmatis*, *Rhodococcus equi* and *R. sp* 4306. Also, GMM, mannose-monomycolate (MMM), glycerol monomycolate (GroMM) and mycolic acid (MA) were tested in the presence of human CD1b-expressing DC2.4 cells (CD1b-DC2.4).

D) Synthetic scheme for TMM.

E) Y-50 reporter cells were stimulated with synthetic TMM harboring a β-hydroxy group (hydroxy) and α-branched alkyl chains (branch) or synthetic analogues lacking hydroxy (–, +) or both moieties (–, –) in the presence of CD1b-DC2.4 cells as APCs.

Data are means ± SD for triplicate assays and representative results from two independent experiments are shown (A, B, C and E).

Fig. 4

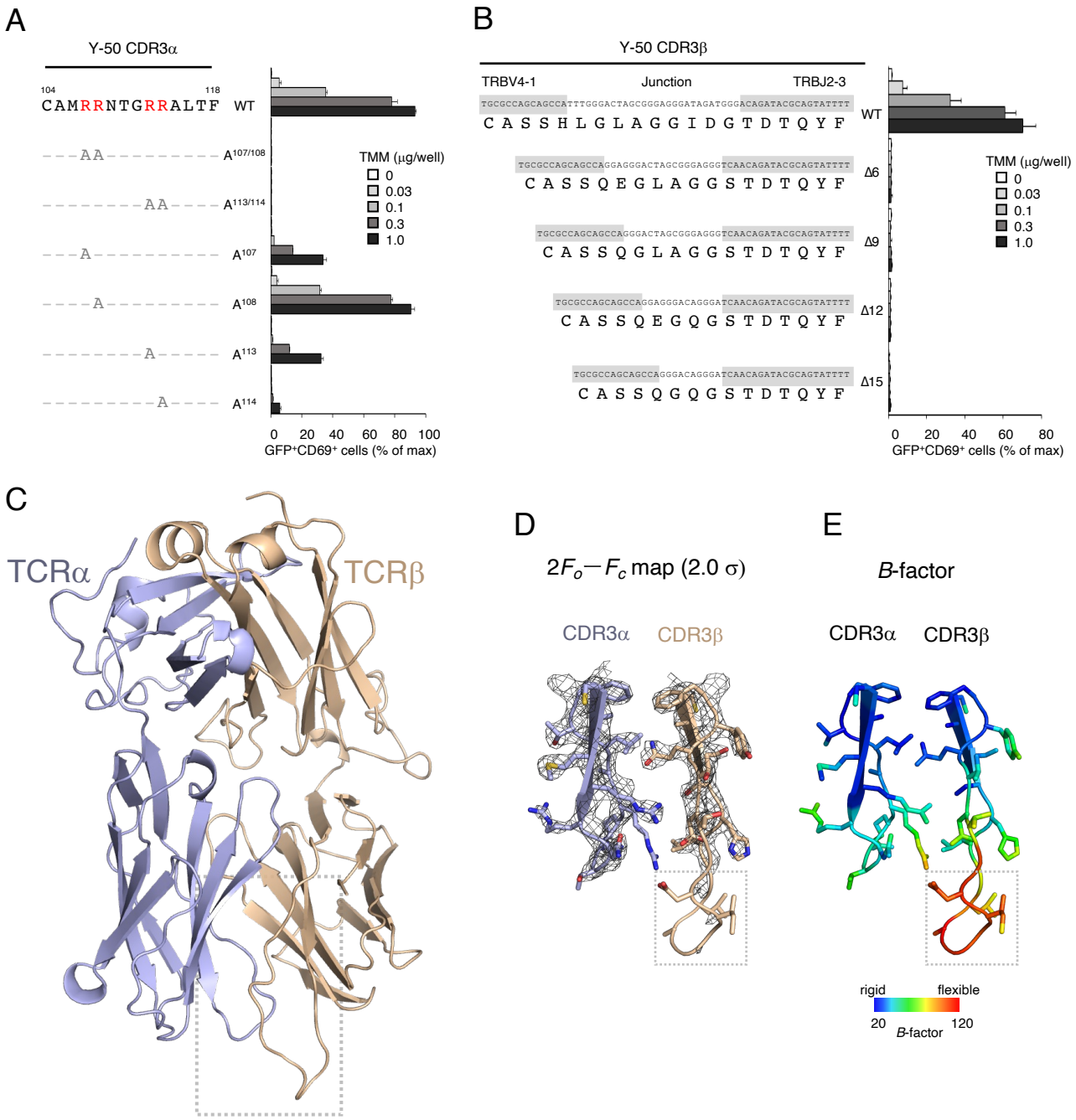


Figure 4. Mutagenesis and structural analysis of TMM-reactive TCR.

A) The amino acid sequences of Y-50 CDR3α within the arginine mutants shown in red. TMM reactivities of each mutant were shown as percentage of the maximum response induced by plate coated anti-CD3 Ab. The number of amino acids were shown in accordance with the ImMunoGeneTics (IMGT) definition (https://imgt.org/IMGTScientificChart/).

B) Nucleotide and amino acid sequences of the Y-50 TCR CDR3β region and its junction-deletion mutants (Δ). D region and N or P nucleotide sequences that constitute junctional sequences were unshaded. Cells were stimulated as indicated in (A).

C) Crystal structure of the Y-50 TCRαβ heterodimer (PDB 8XUB). The main chains of TCRα and β were drawn in violet and brown, respectively.

D) 2F*o*-F*c* map of CDR3αβ (gray mesh) contoured at 2.0 σ.

E) *B*-factor diagram of CDR3αβ. The *B*-factor values are illustrated by color, ranging from blue (rigid) to red (flexible).

Data are shown as mean ± SD of triplicates (A and B) and representative results from two independent experiments are shown.

Fig. 5

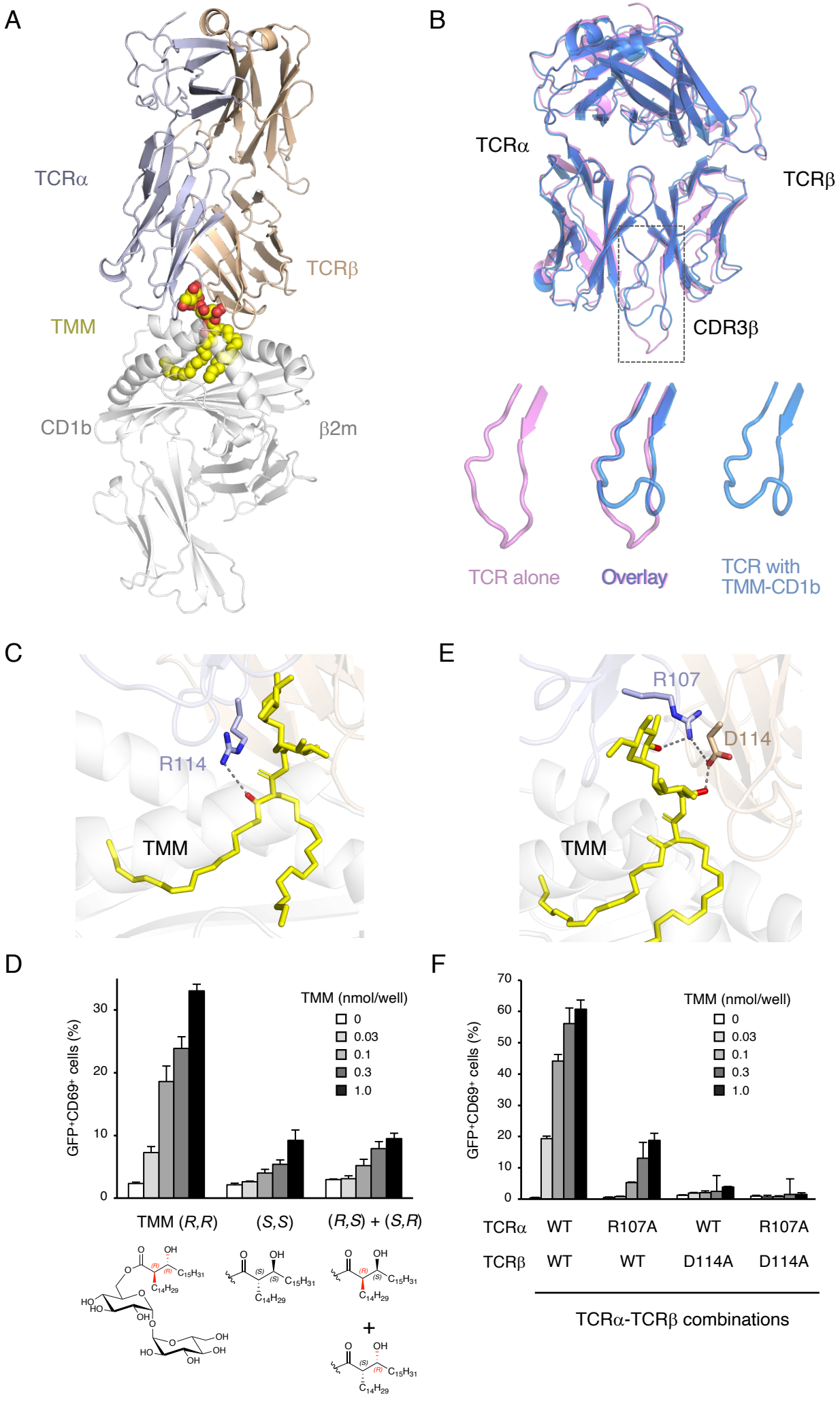

A
TCRα
TCRβ
TMM
CD1b
β2m
B
TCRα
TCRβ
CDR3β
TCR alone
Overlay
TCR with TMM-CD1b
C
R114
TMM
E
R107
D114
TMM
D
GFP+CD69+ cells (%)
TMM (nmol/well)
0
0.03
0.1
0.3
1.0
TMM (R,R)
(S,S)
(R,S) + (S,R)
F
GFP+CD69+ cells (%)
TMM (nmol/well)
0
0.03
0.1
0.3
1.0
TCRα WT R107A WT R107A
TCRβ WT WT D114A D114A
TCRα-TCRβ combinations

Figure 5. Ternary complex structure of Y-50 TCR-TMM-CD1b.

A) Overall structure of the Y-50 TCR-TMM-CD1b complex. The main chains of TCRα, TCRβ, and CD1b were shown as indicated. TMM was presented as yellow sphere.

B) Superimposition of the structure of Y-50 TCR alone (PDB)(pink) and Y-50 TCR bound to TMM-CD1b (PDB)(blue)(upper panel). CDR3β regions (boxed) were magnified (lower panels). C) Close-up view of TMM (*R,R*) and the side chain of R114 within CDR3α. β-hydroxy group of TMM is shown in red.

D) Y-50 reporter cells were stimulated with natural configuration of synthetic TMM (*R,R*) or non-natural stereoisomers, (*S,S*) or (*S,R*+*R,S*), in the presence of CD1b-DC2.4 and analyzed for GFP and CD69 expression. Structures of stereoisomers are shown below (*R*, red; *S*, black).

E) Close-up view of TMM (*R,R*) and the side chain of R107 (CDR3α) and D114 (CDR3β). Hydroxy groups of TMM that formed hydrogen bonds to the TCR side chains are shown in red.

F) Reporter cells expressing indicated combination of Y-50 TCRαβ mutants were stimulated with TMM on CD1b-DC2.4 and percentages of $GFP^{+}CD69^{+}$ cells are shown.

Data are shown as mean ± SD of triplicates (D and F) and representative results from two independent experiments are shown.

Fig. 6

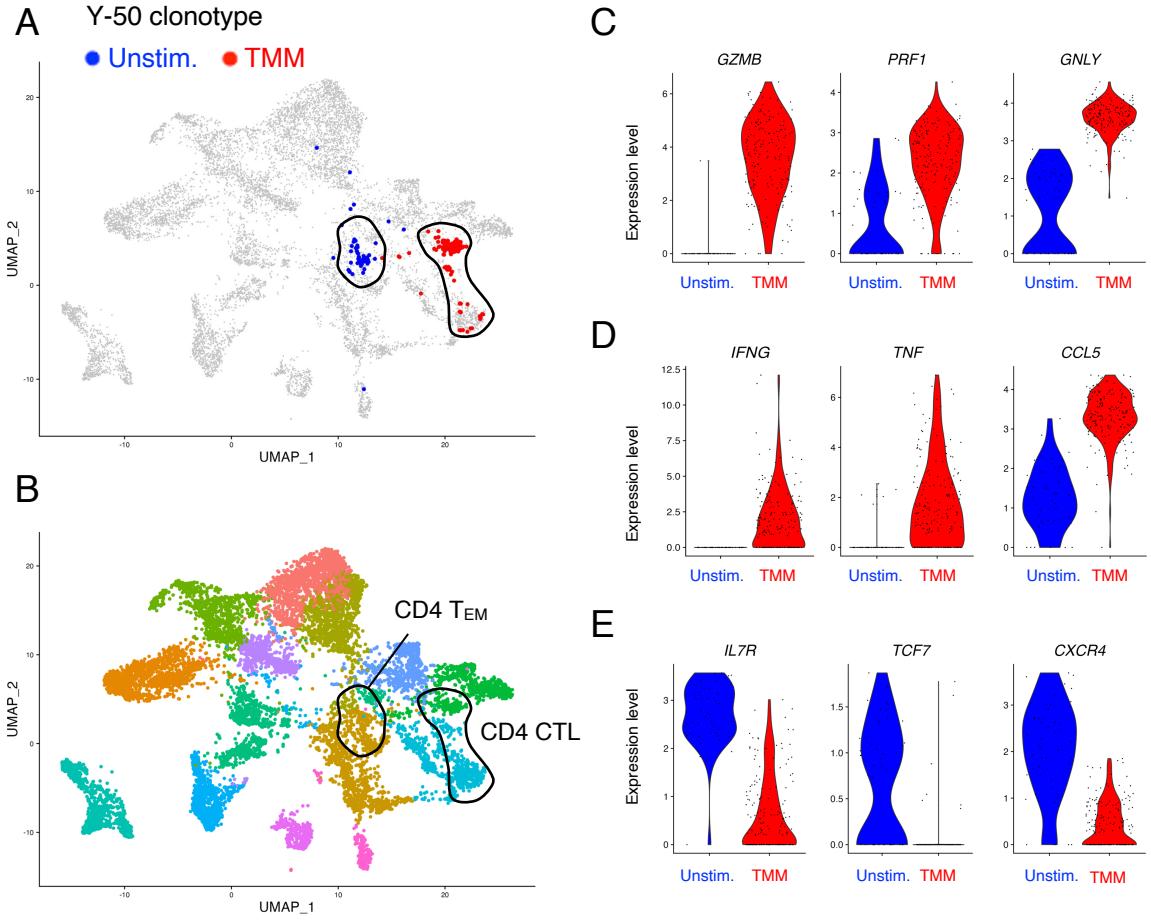


Figure 6. Functional maturation of TMM-reactive T cells upon TMM stimulation.
A-B) Cluster shift of Y-50 clonotype before and after TMM stimulation. T cells expressing Y-50 clonotype defined by sc-TCR-RNA-seq were overlayed (A) on UMAP plot of T cells in PBMCs from donors including the donor used in Fig. 1 (B).
C-E) Differentially expressed genes in Y-50 T cells upon TMM stimulation. The expression of representative genes encoding cytotoxic effector molecules (C), pro-inflammatory cytokines and chemokines (D) and stemness-related molecules (E) are shown in violin plots.

Fig. 7

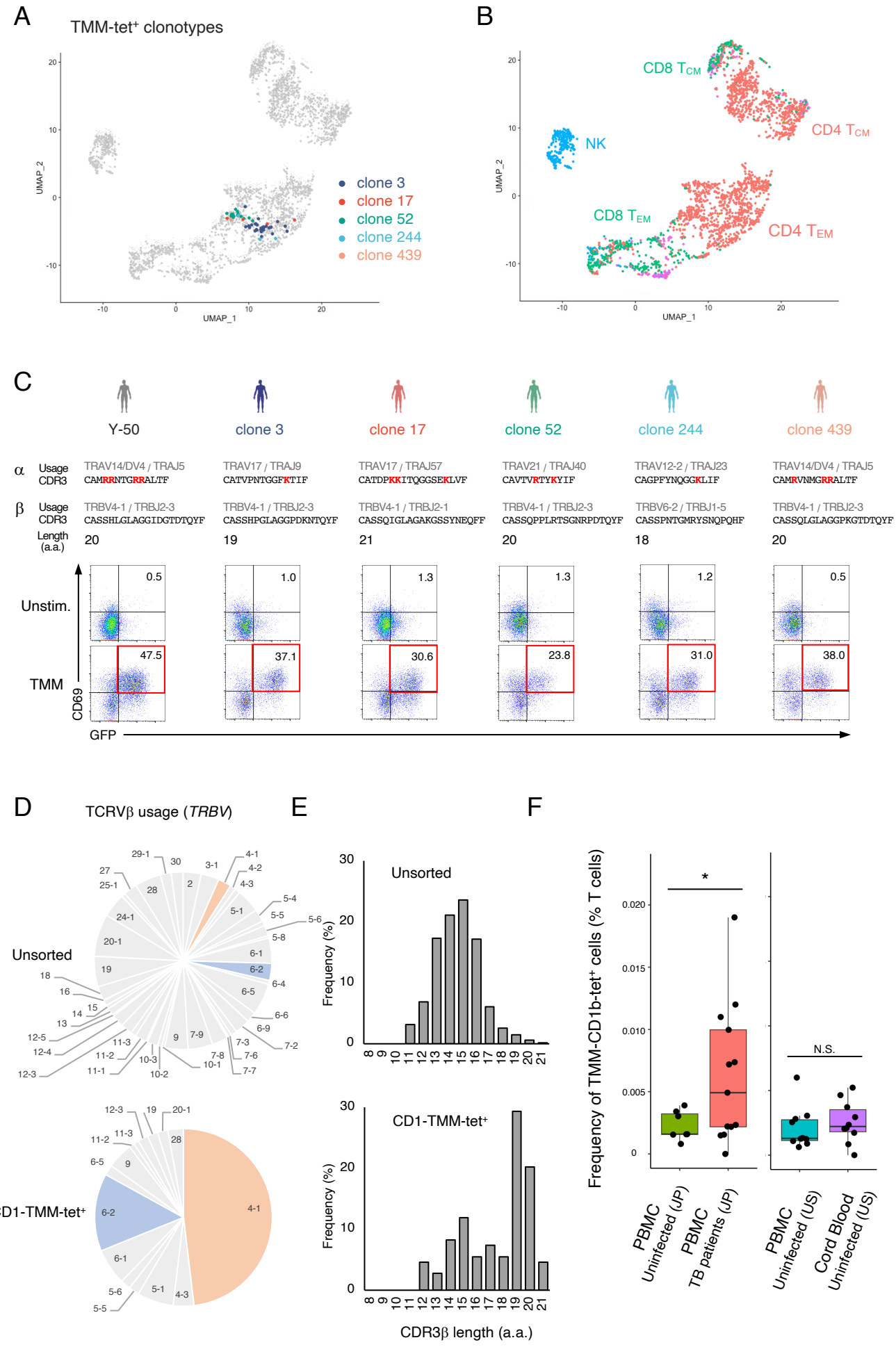

A
TMM-tet+ clonotypes
clone 3
clone 17
clone 52
clone 244
clone 439
UMAP_1
UMAP_2
B
CD8 TCM
CD4 TCM
NK
CD8 TEM
CD4 TEM
C
Y-50
clone 3
clone 17
clone 52
clone 244
clone 439
α Usage TRAV14/DV4 / TRAJ5 | TRAV17 / TRAJ9 | TRAV17 / TRAJ57 | TRAV21 / TRAJ40 | TRAV12-2 / TRAJ23 | TRAV14/DV4 / TRAJ5
CDR3 CAMRRNTGRRALTF | CATVPNTGGFKTIF | CATDPKKITQGGSEKLVF | CAVTVRTYKYIF | CAGPFYNQGGKLIF | CAMRVNMGRRALTF
β Usage TRBV4-1 / TRBJ2-3 | TRBV4-1 / TRBJ2-3 | TRBV4-1 / TRBJ2-1 | TRBV4-1 / TRBJ2-3 | TRBV6-2 / TRBJ1-5 | TRBV4-1 / TRBJ2-3
CDR3 CASSHLGLAGGIDGTDTQYF | CASSHPGLAGGPDKNTQYF | CASSQIGLAGAKGSSYNEQFF | CASSQPPLRTSGNRPDTQYF | CASSPNTGMRYSNQPQHF | CASSQLGLAGGPKGTDTQYF
Length (a.a.) 20 | 19 | 21 | 20 | 18 | 20
Unstim.
TMM
CD69
GFP
0.5
47.5
1.0
37.1
1.3
30.6
1.3
23.8
1.2
31.0
0.5
38.0
D
TCRVβ usage (TRBV)
Unsorted
CD1-TMM-tet+
E
Unsorted
CD1-TMM-tet+
Frequency (%)
CDR3β length (a.a.)
F
Frequency of TMM-CD1b-tet+ cells (% T cells)
*
N.S.
PBMC Uninfected (JP)
PBMC TB patients (JP)
PBMC Uninfected (US)
Cord Blood Uninfected (US)

Figure 7. TMM-specific T cells with similar characteristics are shared among individuals.

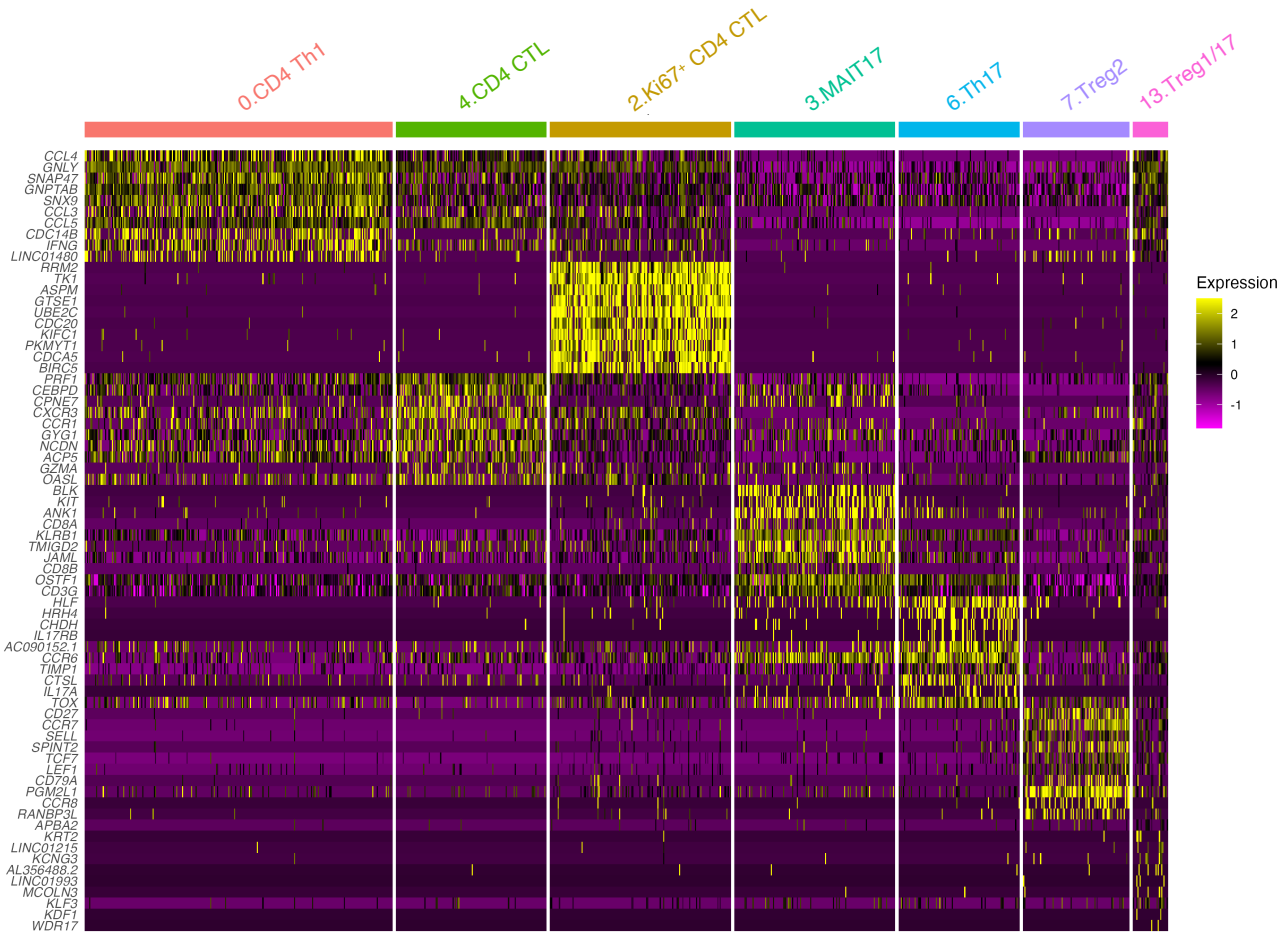
A-B) Highly-enriched TMM-reactive clonotypes identified by TMM-CD1b-tet sorting and sc-TCR-RNA-seq were overlayed (A) on UMAP plot of all $CD3^+$ T cells from 13 healthy donors (B). CD4 T_{EM}, $CD4^+$ effector memory T cells. CD4 T_{CM}, $CD4^+$ central memory T cells.

C) TCR usages, CDR3 sequences and length of CDR3β region of clonotypes detected in (A). Each clonotype was reconstituted into reporter cells and analyzed for TMM reactivity using CD1b-DC2.4 as APCs. GFP^+CD69^+ population is boxed. Y-50 TCR is shown as a control.

D-E) Frequency of TCRVβ usage (D) and length of CDR3β region (E) of unsorted or top 27 TMM-CD1b-$tetramer^+$ T cell clonotypes.

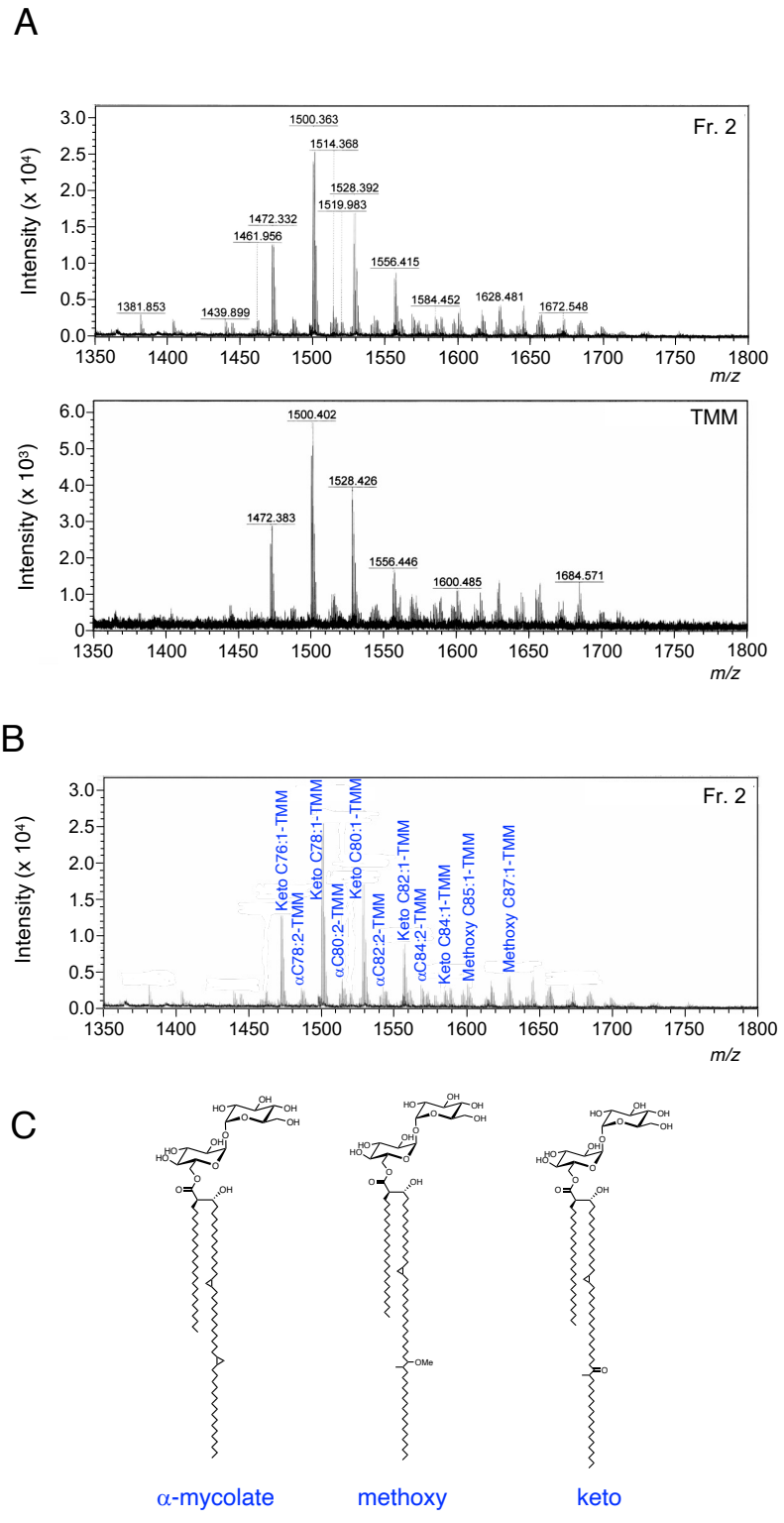
F) PBMCs from Japanese donors (uninfected donors, n = 7; TB patients, n = 13) or PBMC (n = 10) and cord blood cells (n = 10) from uninfected US donors were stained with PE-conjugated TMM-loaded CD1b tetramer, APC-conjugated CD1b-endo tetramer and anti-CD3 antibody. The percentages of TMM-CD1b tetramer positive within endo-CD1b tetramer negative population in $CD3^+$ T cells are shown (TMM-CD1b-tet^+). Medians are indicated with horizontal bars. P values are based on an exact two-sided Welch rank-sum test; $*P < 0.05$.

Extended Data Fig. 1



Extended Data Figure 1. Characteristic genes in each cluster of UMAP plot in Fig. 1D.
Log-normalized expression of the top 10 genes for each T cell cluster in Fig. 1D are shown in heatmap. The colors of annotated clusters are consistent with the dots shown in Fig. 1D.

Extended Data Fig. 2

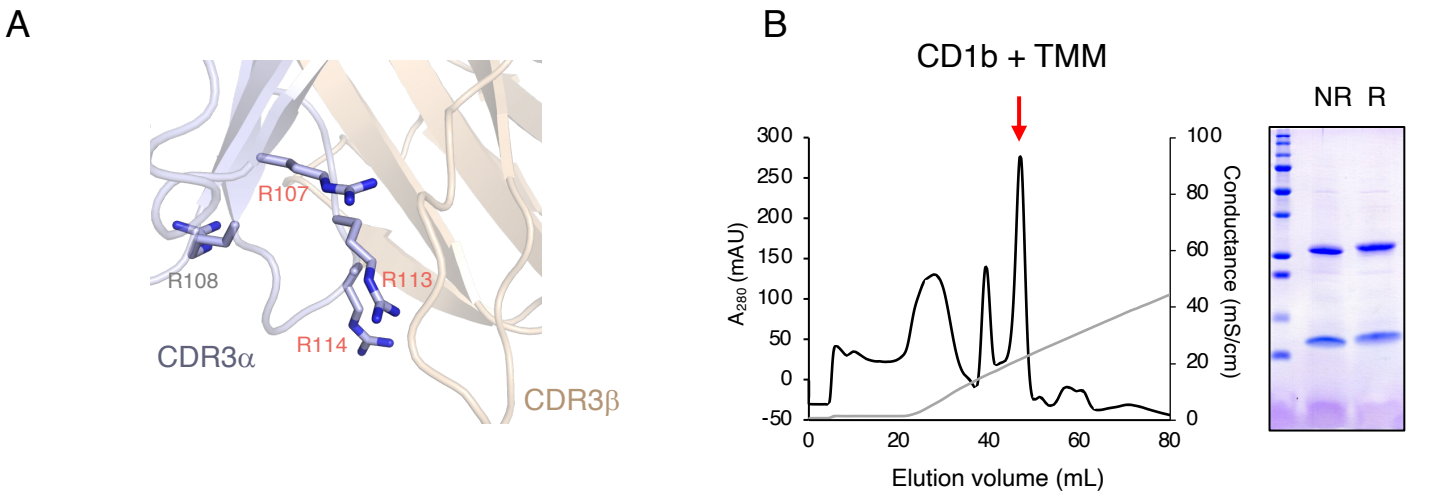


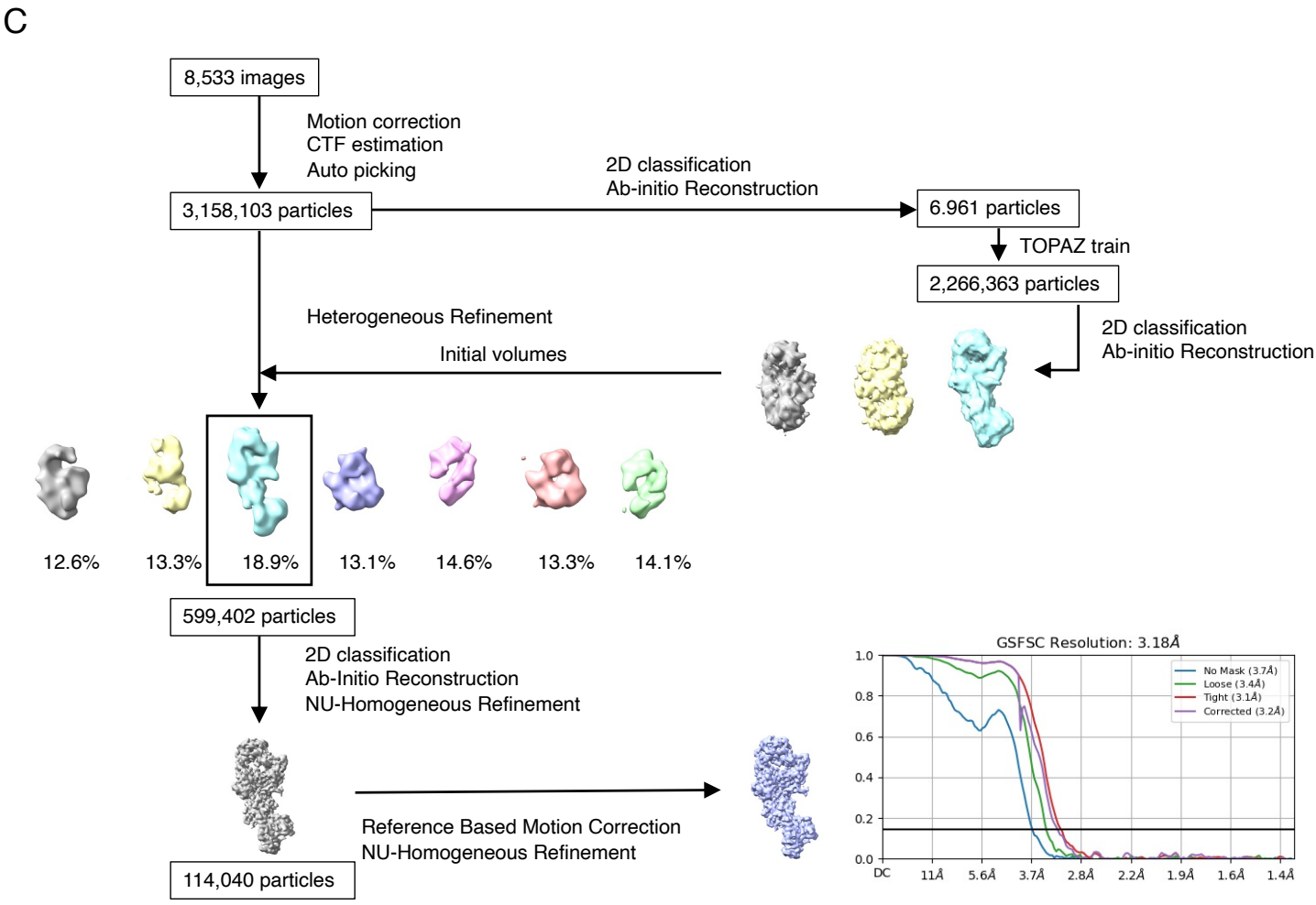
Extended Data Figure 2. Identification of antigen recognized by Y-50.

A-B) MALDI-TOF MS spectrum of lipid fraction 2 (Fr2) and the reference spectrum of *M. tuberculosis* H37Rv TMM (lower) (A). The subclass of mycolate and chain length annotated based on the detected *m/z* matching the structural formula were indicated on upper panel of (A) (B). Related to Fig. 2D.

C) Structural characteristics of TMM composed of three mycolic acid subclasses.

Extended Data Fig. 3





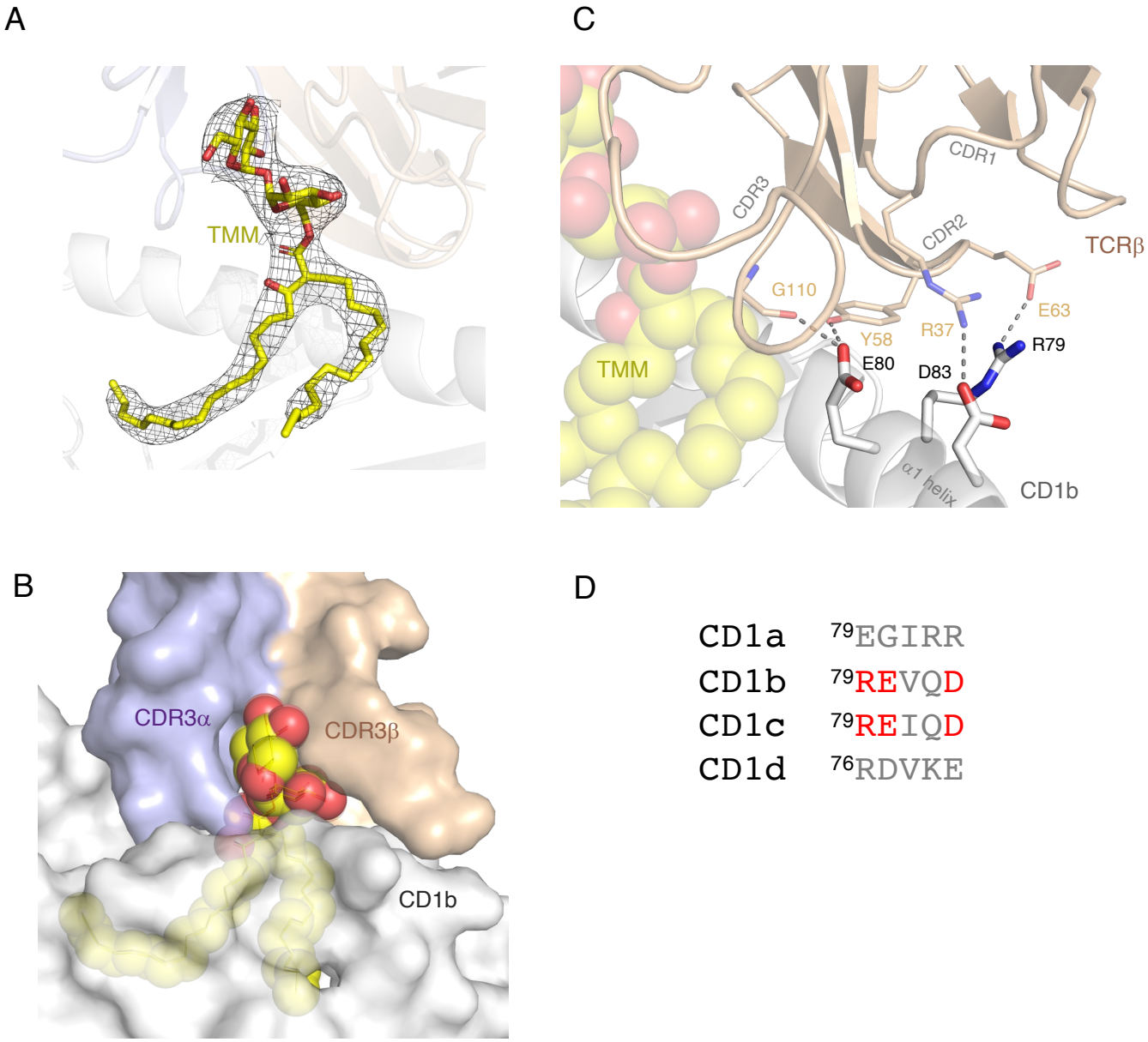
Extended data Fig. 3. Structural analysis of Y-50 TCR-TMM-CD1b ternary complex.

A) Close-up view of the side chains of CDR3α arginine residues (R107, R108, R113 and R114) in the crystal structure of Y-50 (PDB 8XUB).

B) Refolding of recombinant CD1b (CD1b) with synthetic TMM (TMM). Red arrow indicates refolded CD1b-TMM-β2m complex separated by anion-exchange chromatography (left), which was confirmed by SDS-PAGE followed by CBB staining (right). NR, non-reducing; R, reducing.

C) Data processing workflow of cryo-EM analysis. Flow chart of cryo-EM image processing. The following 3D variability analysis is described in methods. FSC curves of the final map and final 3D reconstructions of Y-50-TMM-CD1b complex color-coded according to local resolution.

Extended Data Fig. 4



Extended data Fig. 4. Structural basis of TMM-CD1b recognition by Y-50 TCR.

A) Cryo-EM density map (gray mesh) and a model structure TMM (yellow).

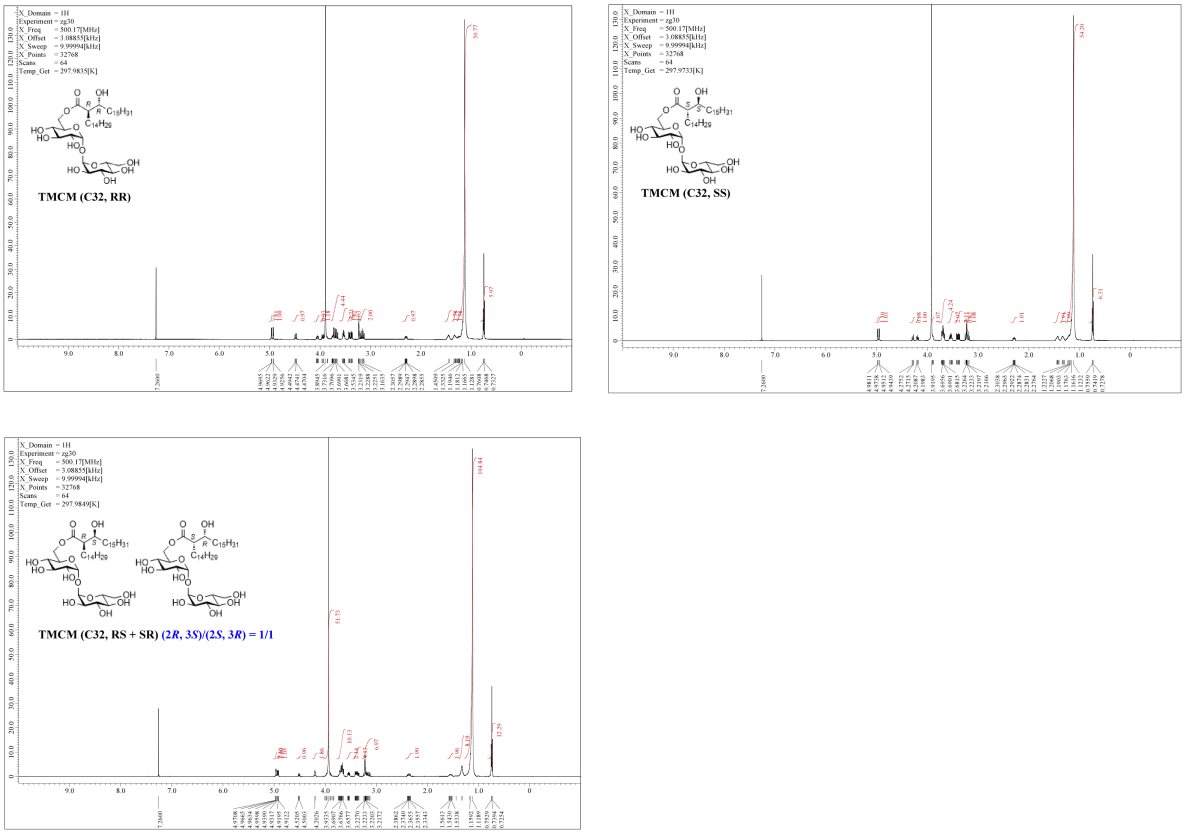
B) TMM (yellow sphere) and surface representation of the CDR3 (violet), CDR3β (brown) loops and CD1b (gray) are shown. O atoms of TMM are shown in red.

C) Close-up view of the side chains of R79, E80 and D83 in CD1b that interact with the side chains of R37 (CDR1β), Y58 and E63 (CDR2β) and G110 (CDR3β).

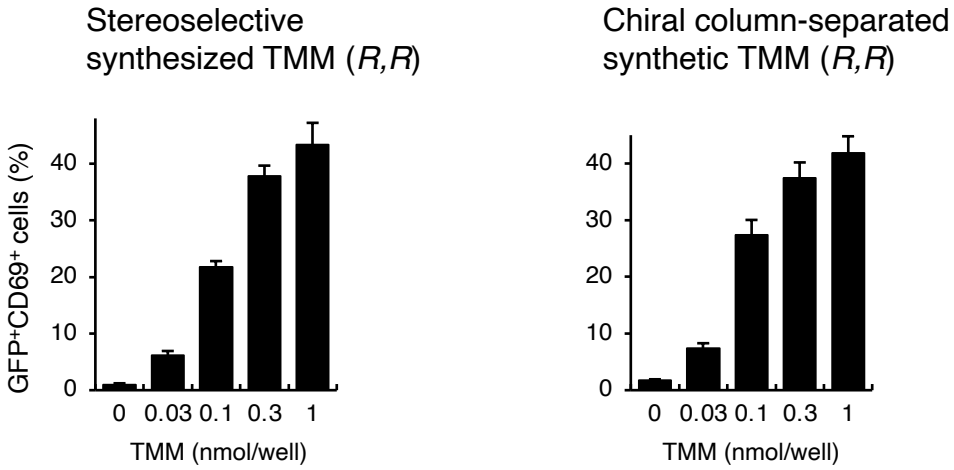
D) Conservation of RExxD motif in human CD1b and CD1c. Amino acid sequence of CD1a (NP_001307581), CD1b (NP_001755.1), CD1c (NP_001756.2) and CD1d (NP_001306074) are aligned. Numbers indicate the amino acid position of mature peptide (excluding signal peptide).

Extended Data Fig. 5

A



B

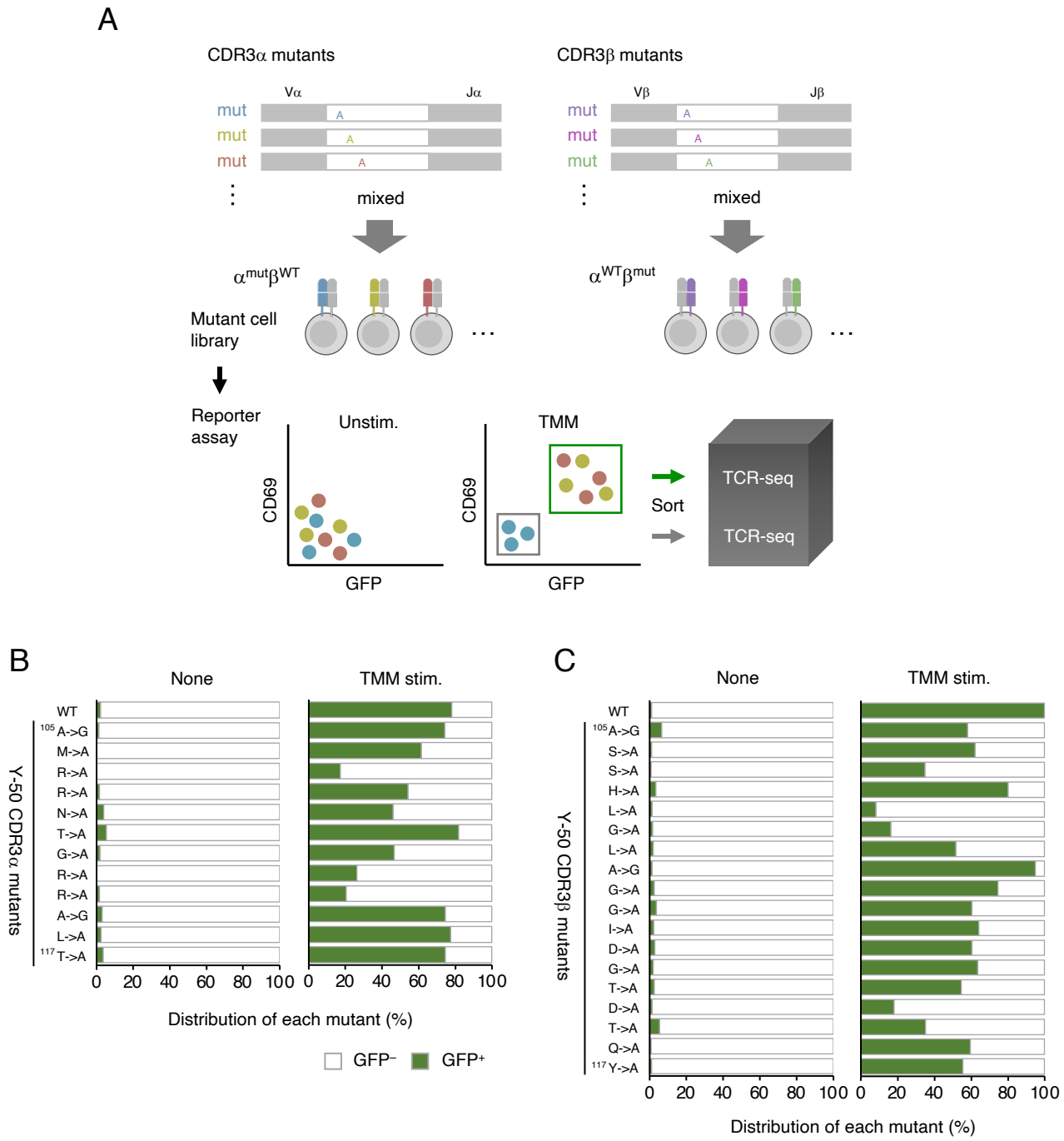


Extended Data Figure 5. Synthesis of TMM.

A) NMR chart of synthetic TMMs ((*R,R*), (*S,S*) and (*R,S* + *S,R*)).

B) Antigenic activity of synthetic TMM. Y-50 reporter cells were stimulated with stereoselective synthesized (left) or chiral column-separated synthetic (right) TMM (*R,R*) in the presence of CD1b-DC2.4 and analyzed for GFP and CD69 expression. Data show mean ± SD for triplicate assays and representative results from two independent experiments.

Extended Data Fig. 6

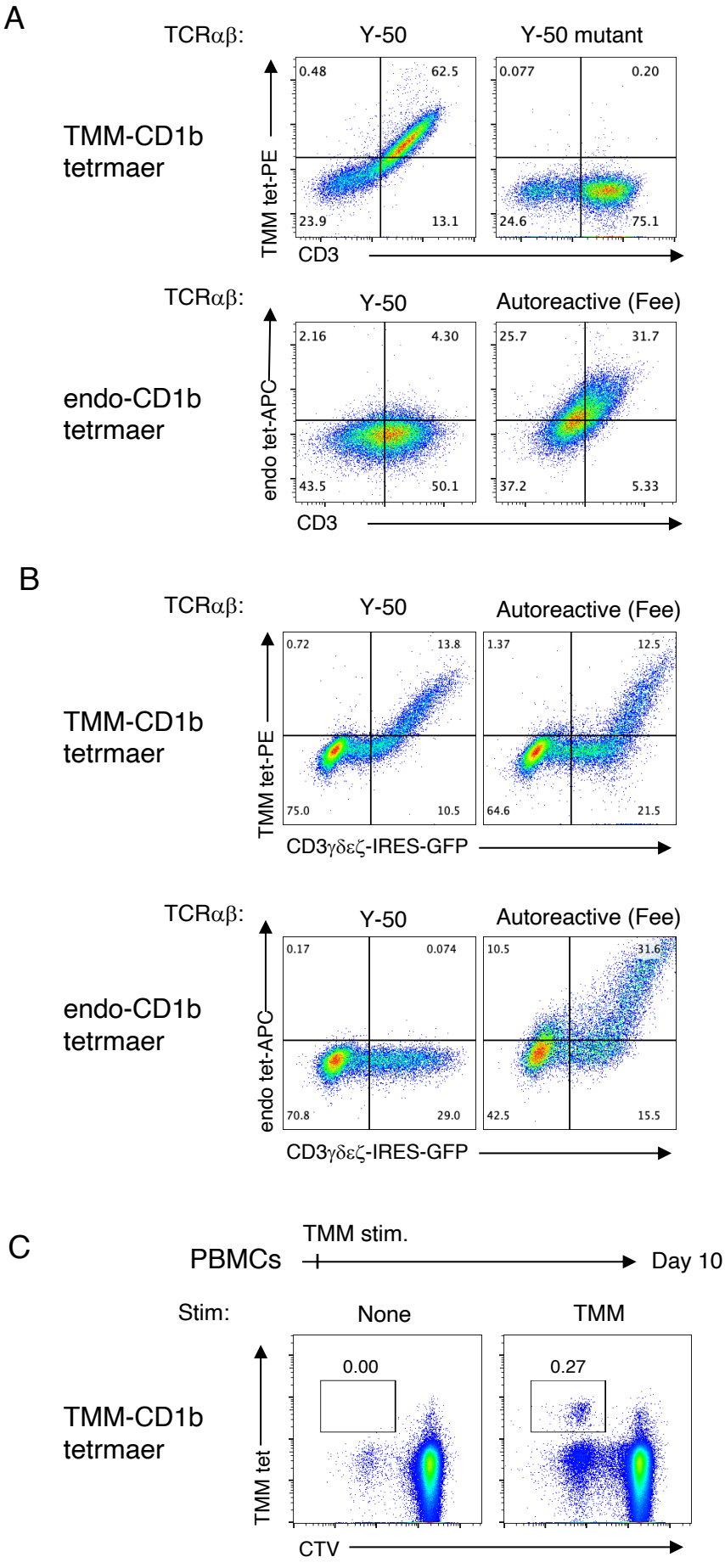


Extended Data Figure 6. NGS-based mutagenesis scanning.

A) Schematic procedure of NGS-based mutagenesis scanning.

B-C) Distribution of reporter cells expressing each mutant in GFP$^+$ and GFP$^-$ populations with or without TMM stimulation. The ratio of each CDR3α mutant (B) and CDR3β mutant (C) are shown as percentages. The number of amino acids were shown in accordance with the ImMunoGeneTics (IMGT) definition.

Extended Data Fig. 7



Extended Data Figure 7. Validation of TMM-loaded CD1b tetramers.

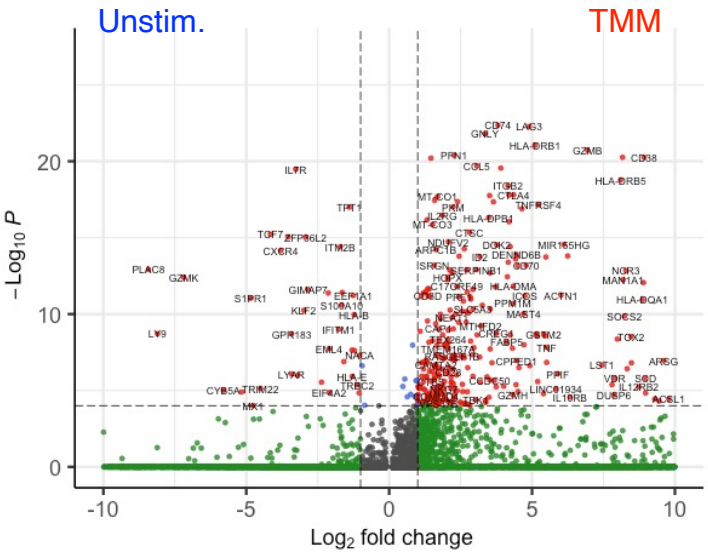
A) Reporter cells expressing Y-50 and mutant Y-50 TCR shown in Fig. 3B (Δ12) were stained with PE-conjugated TMM-loaded CD1b tetramer and anti-CD3 antibody. As a control tetramer, CD1b biotinylated monomer conjugated by SA-APC (endo-CD1b tet) was prepared. Endo-CD1b tet reactivity was confirmed by Y-50 (negative control) and autoreactive Fee TCR (Gherardin et al., 2021) (positive control).

B) HEK 293T cells were transfected with plasmids encoding Fee TCR and human CD3s-IRES-GFP and stained with PE-conjugated TMM-CD1b or APC-conjugated endo-CD1b tetramers.

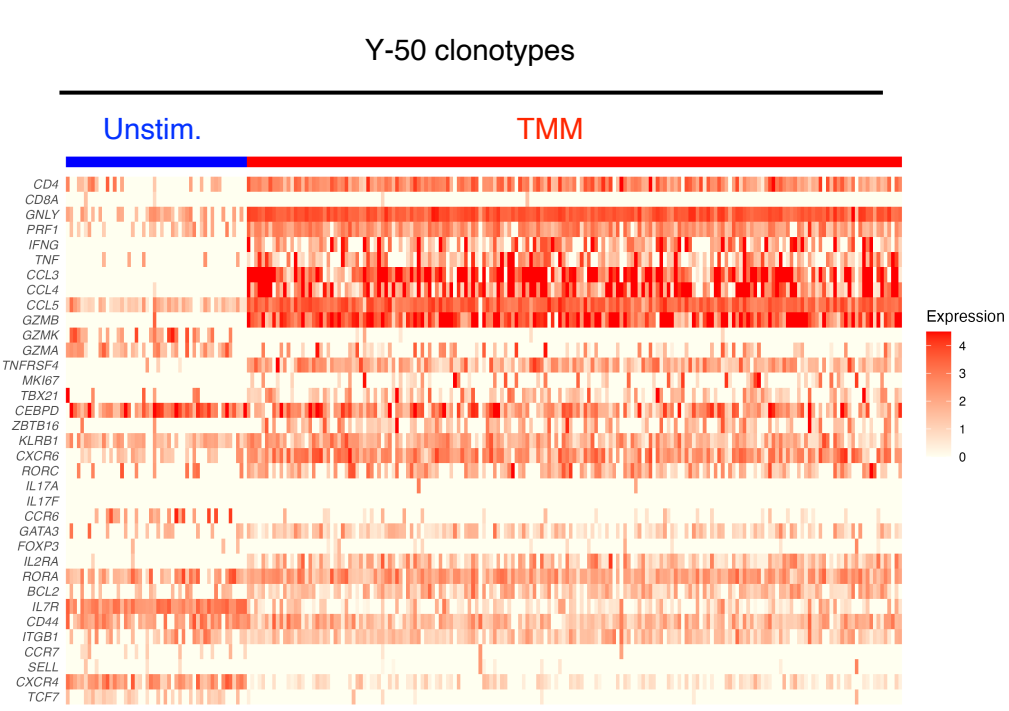
C) CTV-labeled PBMCs were left unstimulated (None) or stimulated with plate coated TMM (TMM) for 10 days and stained with PE-conjugated TMM-CD1b tetramer.

Extended Data Fig. 8

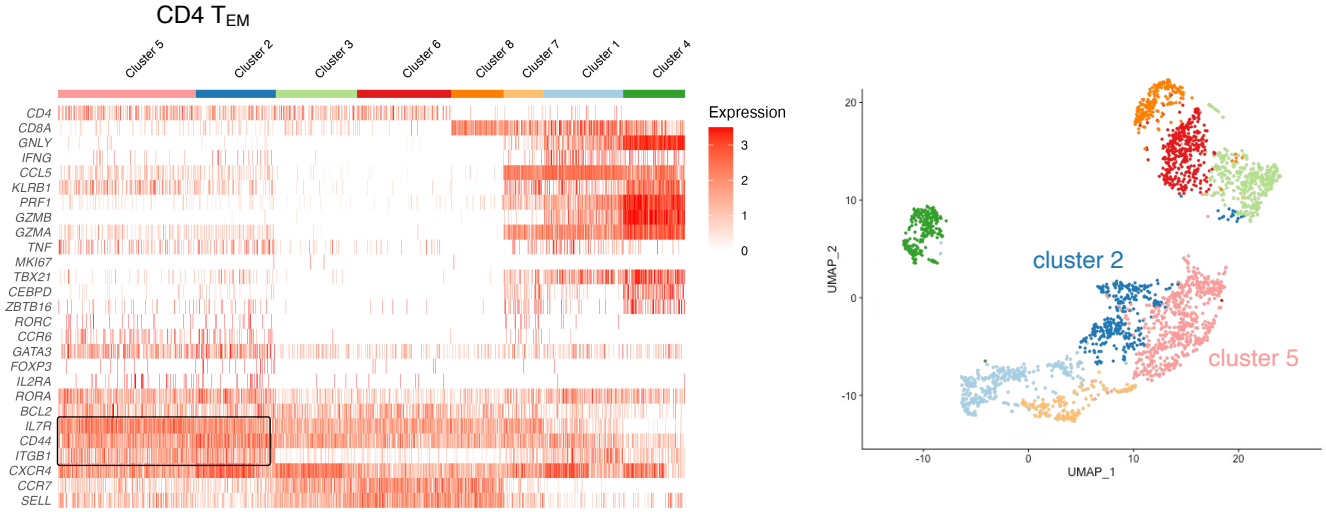
A



B



C



Extended Data Figure 8. Gene expression profile of TMM-specific T cells with or without TMM stimulation.

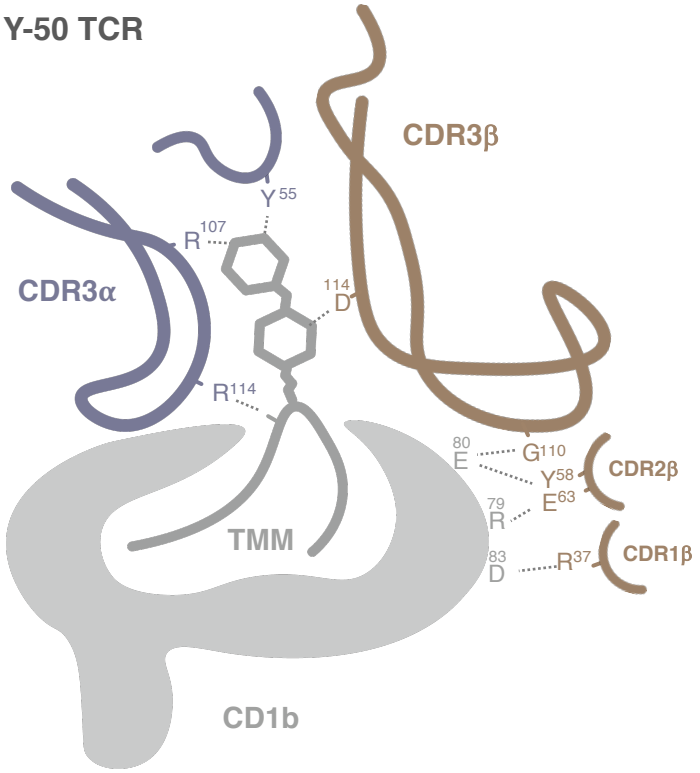
A) Volcano plot showing differentially expressed genes of Y-50 T cells in response to TMM stimulation.

B) Heatmap showing log-normalized expression of marker genes in individual Y-50 cells in the presence or absence of TMM stimulation.

C) Heatmap showing log-normalized expression of marker genes for each $CD3^+$ T clusters shown in Fig. 7A-B. Typical signature of $CD4^+$ effector memory T cells were boxed.

Extended Data Fig. 9

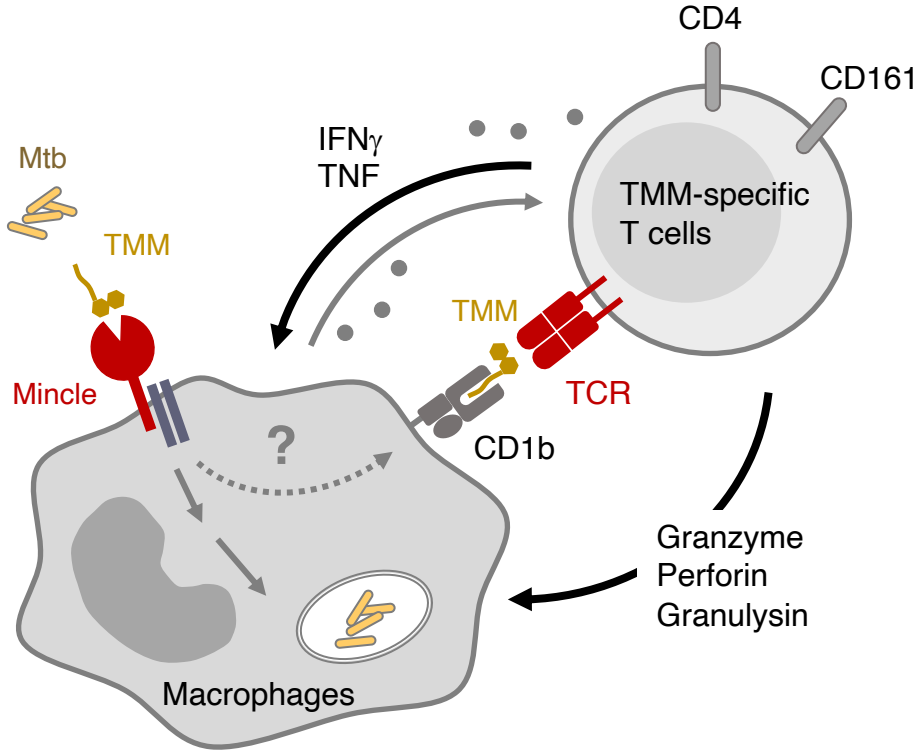
A



B

	TMM-specific T cells	GMM-specific T cells
Restriction	CD1b	CD1b
Antigen	TMM	GMM
TCRα chain	Variant	Invariant, Biased
V gene		TRAV17, TRAV1-2
J gene		TRAJ9
CDR3	Cationic	
TCRβ chain	Biased	Biased
V gene	TRBV4-1, TRBV6-2	TRBV4-1, TRBV6-2
J gene		
CDR3 length	Extended	Variable
Cell type	$CD4^+$ T_{EM}	$CD4^+$ or DN

C



Extended Data Figure 9. Characterization of TMM-specific TCR.
A) Schematic representation of CD1b-restricted glycolipid TMM recognition by Y-50 TCR (PDB 8ZOX). CDR loops of TCRαβ chains, CD1b and TMM are illustrated. Dotted lines indicate hydrogen bonds or salt bridges.
B) Characteristics of TMM- and GMM-specific T cells.
C) Dual roles of mycobacterial TMM in T cells and pattern recognition receptor activation. Mycobacterial TMM is recognized by Mincle and activates macrophages. Mycobacteria-infected macrophages present TMM on CD1b. Upon TMM recognition, TMM-specific T cells produce cytotoxic effector molecules and pro-inflammatory cytokines including IFNγ and TNF.

Supplementary Table 1. Mycobacterial lipid-reactive candidate T cells in human PBMCs.

Ex#	clone ID	TRAV	CDR3α	TRAJ	TRBV	CDR3β	TRBJ
U	2	1-2	CAAMDSNYQLIW	33	6-4	CASSDGGGTDTQYF	2-3
	14	8-3	CAVGGGGYQKVTF	13	11-2	CASSLVSSGALTEQYF	2-7
	15	1-2	CAVVDSNYQLIW	33	6-5	CASSPGSGVAEQYF	2-7
	53	20	CAVRVYGQNFVF	26	20-1	CSALQPGLAGPQITYEQYF	2-7
	59	1-2	CAVRDSNYQLIW	33	6-1	CASSPSTGGSSPLHF	1-6
	86	1-2	CAVRDGDYKLSF	20	4-3	CASSQDGSSGANVLTF	2-6
	104	1-2	CAVRDSNYQLIW	33	4-3	CASSLATEPSYEQYF	2-7
	221	1-2	CAAMDSNYQLIW	33	20-1	CSAKGPAGADTGELFF	2-2
V	1	1-2	CAVVDSNYQLIW	33	6-1	CASSSSGSAADTQYF	2-3
	2	27	CAGGNSGNTPLVF	29	4-1	CASRREGENIQYF	2-4
	3	8-6	CAVSDGRRDDKIIF	30	7-6	CASSPLQRGTYEQYF	2-7
	153	29/DV5	CAASTYGGATNKLIF	32	30	CAWSRRLQVLNGYTF	1-2
	569	5	CAEIPYTGGGNKLTF	10	4-1	CASSQGGGATGNTIYF	1-3
	570	8-4	CAVSVGLGFGNVLHC	35	6-5	CASSYSPGVWPQHF	1-5
	573	8-1	CAVNAPHGEKLVF	8	5-1	CASSPGTSGVYNEQFF	2-1
	597	19	CALSERSTGNQFYF	49	6-5	CASSYSPGVWPQHF	1-5
	606	17	CATVPNNDYKLSF	20	6-1	CASTYSTMNTEAFF	1-1
	648	1-2	CAVKDNYGQNFVF	26	30	CAHRTRGETQYF	2-5
	828	41	CAAFGNEKLTF	48	30	CAWSVSSGWPLHF	1-6
	835	17	CATDRAGANNLFF	36	4-1	CASSQDPGNEKLFF	1-4
	1030	14/DV4	CAMPAGTGRRALTF	5	11-3	CASSLTSGGYGGTDTQYF	2-3
	1176	1-2	CAVKDSNYQLIW	33	4-2	CASSVERGAGANVLTF	2-6
X	23	12-2	CAANSGGSNYKLTF	53	30	CAWTRETGQPQHF	1-5
	34	23/DV6	CAALHGSSNTGKLIF	37	11-2	CASSDWTSSSGSYEQYF	2-7
	84	8-4	CAVRPAAGNKLTF	17	28	CASMADYNEQFF	2-1
	291	3	CAVRVLSGGYNKLIF	4	14	CASSRRPYSGSTDTQYF	2-3
	309	29/DV5	CAASAKTSGSRLTF	58	7-2	CASSLDPGLAKNIQYF	2-4
	605	5	CAESIWGSQGNLIF	42	19	CASSIDGRQLGSFF	1-1
	700	1-2	CAVSALYGQNFVF	26	11-2	CASSDWTSSSGSYEQYF	2-7
Y	0	8-2	CVVRLYQKVTF	13	14	CASSQGYRGSSYNEQFF	2-1
	5	2	CAVPSNTGKLIF	37	7-8	CASSLAQGTGNNSPLHF	1-6
	10	38-1	CAFTPNTGNQFYF	49	4-1	CASSQEFRDTEAFF	1-1
	11	5	CAESIGTGGFKTIF	9	30	CAWSVLAGRGETQYF	2-5
	13	12-2	CAVNGGGFKTIF	9	7-3	CASSLTPKRSNTGELFF	2-2
	14	14/DV4	CAMREGMGGNMLTF	39	20-1	CSARHSASSYEQYF	2-7
	16	12-3	CAKGFNDYKLSF	20	7-8	CASSLRGHLQETQYF	2-5
	17	17	CATLRNTNAGKSTF	27	20-1	CSAVLTDTQYF	2-3
	18	14/DV4	CAMRGSNTGKLIF	37	7-9	CASSLAGGQGLAYTF	1-2
	22	8-3	CAVGNSGGYQKVTF	13	7-2	CASSPGAYEQYF	2-7
	24	17	CATKVTSGGSYIPTF	6	20-1	CSALTSGRPDTQYF	2-3
	28	9-2	CALSETDSWGKLQF	24	20-1	CSARGLAGGQYF	2-7
	29	2	CAVERDFNKFYF	21	20-1	CSAPLAGGQFF	2-1
	36	1-2	CAVMDSNYQLIW	33	6-1	CASSDRGHSPLHF	1-6
	50	14/DV4	CAMRRNTGRRALTF	5	4-1	CASSHLGLAGGIDGTDTQYF	2-3
	56	8-4	CAVSPMIYNQGGKLIF	23	14	CASSQRGQVKGNTIYF	1-3
	62	14/DV4	CAILYNFNKFYF	21	12-4	CASSLAQIAKNIQYF	2-4
	66	20	CAVQARTQGGSEKLVF	57	12-3	CASRDYLSTDTQYF	2-3
	83	22	CAVVLWGGADGLTF	45	10-1	CASSDGREQYF	2-7
	87	12-2	CAVSGSSNTGKLIF	37	4-2	CASSQEPQQETQYF	2-5
	92	1-2	CAVRDRDYKLSF	20	6-2	CASSYSDFNEQFF	2-1
	114	17	CATAGRIGARLMF	31	28	CASRPTSGRAYGETQYF	2-5
	151	8-6	CAVSTFSSGSARQLTF	22	11-2	CASSLASGRPTDTQYF	2-3

Supplementary Table 2. Data collection and refinement statistics of the crystallographic analysis of Y-50.

Crystal name	**Y-50**
PDB ID	**8XUB**
Data collection statistics	
Space group	***P*** **2 2_1 2_1**
Cell constants (Å)	a = 68.1 b = 168.7 c = 173.1
Resolution (Å)[a]	44.03 - 2.50 (2.59 - 2.50)
R-merge (%)[a]	2.62(42.8)
Completeness (%)[a]	99.9 (100.0)
$<I/\sigma I>$[a]	17.1 (1.6)
Multiplicity[a]	2.0 (2.0)
Refinement statistics	
R (%)	*22.6*
R_{free} (%)	*29.9*
Root mean square deviations from ideal values	
Bond length (Å)	0.009
Bond angle (°)	1.07
Ramachandran plot	
Favored (%)	94.5
Outlier (%)	0
Rotamer outlier (%)	2.84

Statistics for the highest-resolution shell are shown in parentheses.

Supplementary Table 3. Cryo-EM data collection, refinement and validation statistics.

PDB ID	**8ZOX**
EMDB ID	**EMD-60321**
Data collection and processing	
Magnification	103,703
Voltage (kV)	300
Electron exposure (e^{-}/Å^2)	60
Defocus range (μm)	-0.6 to -1.8
Pixel size (Å)	0.675
Symmetry imposed	*C*1
Initial particle image	4,075,153
Final particle image	114,111
Map resolution (Å)	3.18
FSC threshold	0.143
Map sharpening *B* factor (Å^2)	118.4
Refinement	
Model composition	
Non-hydrogen atoms	6,590
Protein residues	813
Ligands	5
B factor (Å^2)	
Protein	125.4
Ligand	114.0
Root-mean-square deviations	
Bond lengths (Å)	0.005
Bond angles (°)	0.604
Validation	
MolProbity score	1.91
Clashscore	11.9
Poor rotamers (%)	0.71
Ramachandran plot	
Favored (%)	95.4
Allowed (%)	4.5
Disallowed (%)	0.1

Supplementary Table 4. Donor information.

Donor ID	Age (years)	Gender	Definition
F507	26	M	Active
F508	49	M	Active
F509	83	F	Active
F510	55	M	Active
F511	64	M	Active
F514	79	M	Active
F515	39	M	Active
F516	60	M	Active
F517	90	F	Active
F518	34	F	Active
F519	19	M	Active
F520	78	M	Active
F521	63	M	Active
MI001	27	M	Uninfected
MI002	25	M	Uninfected
MI003	24	F	Uninfected
MI004	27	M	Uninfected
MI011	60	F	Uninfected
MI012	27	F	Uninfected
MI013	36	F	Uninfected
NA001	49	M	Uninfected
NA002	31	M	Uninfected
NA003	44	M	Uninfected
NA004	22	F	Uninfected
NA005	33	F	Uninfected
NA006	28	M	Uninfected
NA007	27	M	Uninfected
NA008	22	F	Uninfected
NA009	40	F	Uninfected
NA010	56	F	Uninfected

Active: 56.8 + 22.6 y.o., male = 77%
Uninfected (JPN): 32.3 + 12.8 y.o., male = 23%
Uninfected (US): 35.2 + 11.6 y.o., male = 38%