Lipid and small-molecule display by CD1 and MR1

Ildiko Van Rhijn[1,2], Dale I. Godfrey[3,4], Jamie Rossjohn[5–7] and D. Branch Moody[1]

Abstract | The antigen-presenting molecules CD1 and MHC class I-related protein (MR1) display lipids and small molecules to T cells. The antigen display platforms in the four CD1 proteins are laterally asymmetrical, so that the T cell receptor (TCR)-binding surfaces are comprised of roofs and portals, rather than the long grooves seen in the MHC antigen-presenting molecules. TCRs can bind CD1 proteins with left-sided or right-sided footprints, creating unexpected modes of antigen recognition. The use of tetramers of human CD1a, CD1b, CD1c or MR1 proteins now allows detailed analysis of the human T cell repertoire, which has revealed new invariant TCRs that bind CD1b molecules and are different from those that define natural killer T cells and mucosal-associated invariant T cells.

[1]*Division of Rheumatology, Immunology and Allergy, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA.*
[2]*Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands.*
[3]*Department of Microbiology and Immunology, Peter Doherty Institute for Infection and Immunity, University of Melbourne, Parkville, Victoria 3010, Australia.*
[4]*Australian Research Council Centre of Excellence for Advanced Molecular Imaging, University of Melbourne, Parkville, Victoria 3010, Australia.*
[5]*Infection and Immunity Program and The Department of Biochemistry and Molecular Biology, Biomedicine Discovery Institute, Monash University, Clayton, Victoria 3800, Australia.*
[6]*Institute of Infection and Immunity, Cardiff University School of Medicine, Heath Park, Cardiff CF14 4XN, UK.*
[7]*Australian Research Council Centre of Excellence in Advanced Molecular Imaging, Monash University, Clayton, Victoria 3800, Australia.*
Correspondence to D.B.M.
e-mail: bmoody@partners.org
doi:10.1038/nri3889
Published online
21 September 2015

A primary activation signal for T cells occurs when αβ T cell receptors (TCRs) contact peptide antigens bound to MHC molecules. The co-recognition of peptide and MHC proteins is among the most influential biological discoveries of the twentieth century[1–3]. This model explained that the functional specificity of T cells for peptides and MHC proteins results from direct contact of a TCR with a hybrid surface formed by a peptide and an MHC molecule. Furthermore, the high level of polymorphism in MHC Ia genes explains, at least in part, why individuals typically use different αβ TCRs to respond to the same antigen, thereby forming highly distinct (or private) TCR repertoires. To date, nearly all technologies that seek to manipulate or detect human αβ T cell responses are based on the principles of peptide–MHC co-recognition. For example, vaccine subunits are derived from proteins, and antigenic epitopes that control T cells during infection, vaccination, cancer and autoimmune diseases are mapped through peptide sequencing. However, peptide–MHC complexes are not the sole targets of human T cell responses. It is increasingly appreciated that a substantial proportion of the overall αβ T cell repertoire recognizes antigens presented by non-polymorphic antigen-presenting molecules that are encoded within the MHC locus (such as HLA-E[4–6]) or outside the MHC locus (such as CD1a[7,8], CD1b[9], CD1c[10], CD1d[11] and MHC class I-related protein (MR1)[12,13]). Moreover, αβ[14], γδ[15,16] and δ/αβ[17] T cells can all recognize CD1 proteins. Here, we highlight recent studies of CD1 antigen display that provide clear exceptions to the principles of peptide–MHC co-recognition. We focus on the asymmetric nature of human CD1 antigen display platforms and propose that T cell activation can occur by an unexpected mechanism of absence of interference with an approaching TCR. New evidence shows that T cells recognizing CD1 or MR1 proteins are abundant in humans, supporting the use of lipid and small-molecule antigens as a new approach to therapy.

The CD1 genes

In early work, antibodies were used to identify two β2-microglobulin-associated heavy chains; these were later named MHC class I and CD1 proteins[18]. Human CD1 genes are located on chromosome 1 (REF. 19) and encode five CD1 isoforms, which were assigned to group 1 (CD1a, CD1b, CD1c and CD1e) or group 2 (CD1d) on the basis of sequence homology[20]. A second reason for designating two groups is that, whereas CD1d is constitutively expressed[21], the group 1 CD1 genes are inducible and coordinately regulated primarily by myeloid cells[22,23]. CD1 genes are present in all placental mammals, birds and marsupials[24] (FIG. 1). The differing size and composition of CD1 loci in modern mammals probably reflect selective pressure owing to immune function[25–29]. Moreover, the retention of many CD1 genes, with some species encoding more than ten CD1 genes, is consistent with the existence of non-redundant functions for each CD1 isoform. The exception is muroid rodents, which encode two copies of one isoform: namely, CD1d1 and CD1d2.

Human CD1 proteins. It is now clear that each type of CD1 protein has a distinct biological function[30]. For example, CD1e is a soluble lipid transfer protein[31], whereas CD1a, CD1b, CD1c and CD1d are membrane-bound

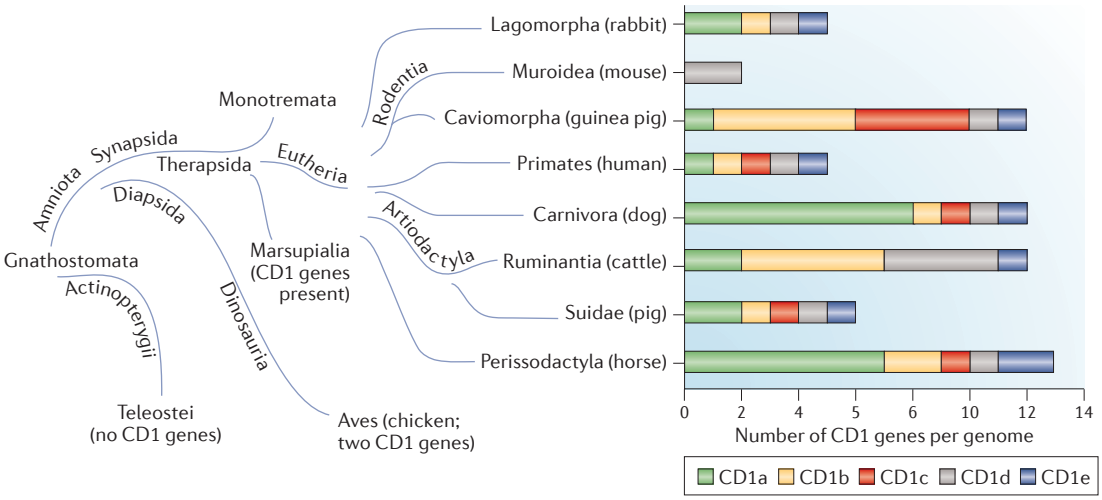

Figure 1 | **Mammalian CD1 genes.** The tree illustrates the origins of modern species after the bird–mammal split. The discovery of avian CD1 orthologues indicates the existence of ancestral CD1 genes before the bird–mammal split, and CD1 genes are universally or widely conserved in mammals. The number of CD1 genes differs in rabbits[26], mice, guinea pigs, primates, dogs[29], cattle[126], pigs[28], horses[25] and chickens[127,128]. For guinea pigs, we discovered one CD1 gene in the updated genome, in addition to the published genes[38,129], which we named after the closest human CD1 orthologue (CD1a). All group 1 CD1 proteins are absent in mice, creating a need for additional experimental models to study CD1a, CD1b and CD1c.

antigen-presenting molecules. Each type of CD1 protein takes different routes through cells[32] and has different expression patterns — B cells express CD1c and CD1d; myeloid dendritic cells (DCs) express CD1a, CD1b, CD1c and CD1d; epithelial cells express CD1d; and Langerhans cells express CD1a[21]. The transcription of the group 1 and group 2 CD1 genes is differently induced by microbial stimuli, with bacterial stimuli selectively upregulating the expression of group 1 CD1 proteins by myeloid DCs[23,33]. Furthermore, each human CD1 protein has a different antigen-binding cleft architecture, with differing numbers of pockets (known as A′, C′, F′ and T′ pockets) and accessory portals (known as C′ and D′/E′ portals)[34]. As mice lack group 1 CD1 proteins, most studies have focused on the recognition of CD1d by a population of CD1d-dependent αβ T cells known as natural killer T (NKT) cells. The growing appreciation of functional divergence of CD1 isoforms provides a clear rationale for the development of new tools to study CD1a, CD1b and CD1c proteins *in vivo* or *ex vivo*, including human group 1 CD1 tetramers[9,10,35] and CD1-transgenic mice[36] or small animals, such as guinea pigs, that naturally express group 1 CD1 proteins[37,38].

CD1-presented antigens

Lipids. T cell responses to CD1 molecules were discovered during studies of *Mycobacterium tuberculosis*[39]. The peptide–MHC co-recognition model predicted that this pathogen would generate peptide antigens, but studies of the whole bacteria showed that antigens could be extracted into organic solvents that exclude proteins. Indeed, in 1994, CD1b was reported to present free mycolic acid, a long-chain α-branched, β-hydroxy fatty acid that is characteristic of mycobacteria[40]. Many more types of lipid antigen have since been identified, including glycolipid, phospholipid, glycophospholipid, sulfolipid and lipopeptide antigens[41]. Most antigens are amphipathic lipids that contain one, two or three aliphatic hydrocarbon chains and a hydrophilic head group comprised of polar or charged moieties (FIG. 2). The head groups vary in size, ranging from the small carboxylate moiety in free mycolic acid to the large polysaccharides in gangliosides. Head groups protrude from the CD1 cavity to bind TCRs, whereas the long and flexible alkyl chains can insert deeply and bend to match the shape of the CD1 cavities[42]. Lipid interactions with the interior of the groove are relatively nonspecific, as one ligand can insert in different orientations, but in general, the head group positioning is more precise[43].

Small molecules. As an exception to the general rule that CD1-dependent T cell activation occurs in response to amphipathic lipids, phenyl 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonate (PPBF) is a synthetic, non-lipid small molecule that activates T cells via CD1d[44]. The molecular mechanism of PPBF action is not fully understood and was difficult to predict because it lacks flexible aliphatic chains and a discrete hydrophilic head group (FIG. 2). Furthermore, the TCR co-recognition model proposes that antigens must exceed the CD1 cleft volume so that they can protrude for direct TCR contact. However, PPBF is less than half the mass of most antigens and is smaller than the volume of the CD1d cleft. Thus, PPBF raised questions about the role of particularly small or non-lipid antigens in the activation of T cells; this issue was later highlighted through studies of skin oils and riboflavin-derivative antigens (discussed below).

Accessory portals
Small gaps present in the side or bottom of the clefts present in CD1b (C′ portal) and CD1c (D′ and E′ portals). Whereas the main F′ portal is present in all CD1 proteins and allows antigen contact with T cell receptors (TCRs), accessory portals probably have a separate sizing function that allows lipids to partially escape from the interior of the cleft at a site distant from TCR contact.

Tetramers
Reagents comprised of a fluorophore-conjugated core surrounded by four antigen-presenting molecules (for example, MHC class I, CD1 or MR1). Antigen-loaded tetramers bind antigen-specific T cell receptors with sufficient avidity so that antigen-specific T cells can be directly counted or isolated by flow cytometry.

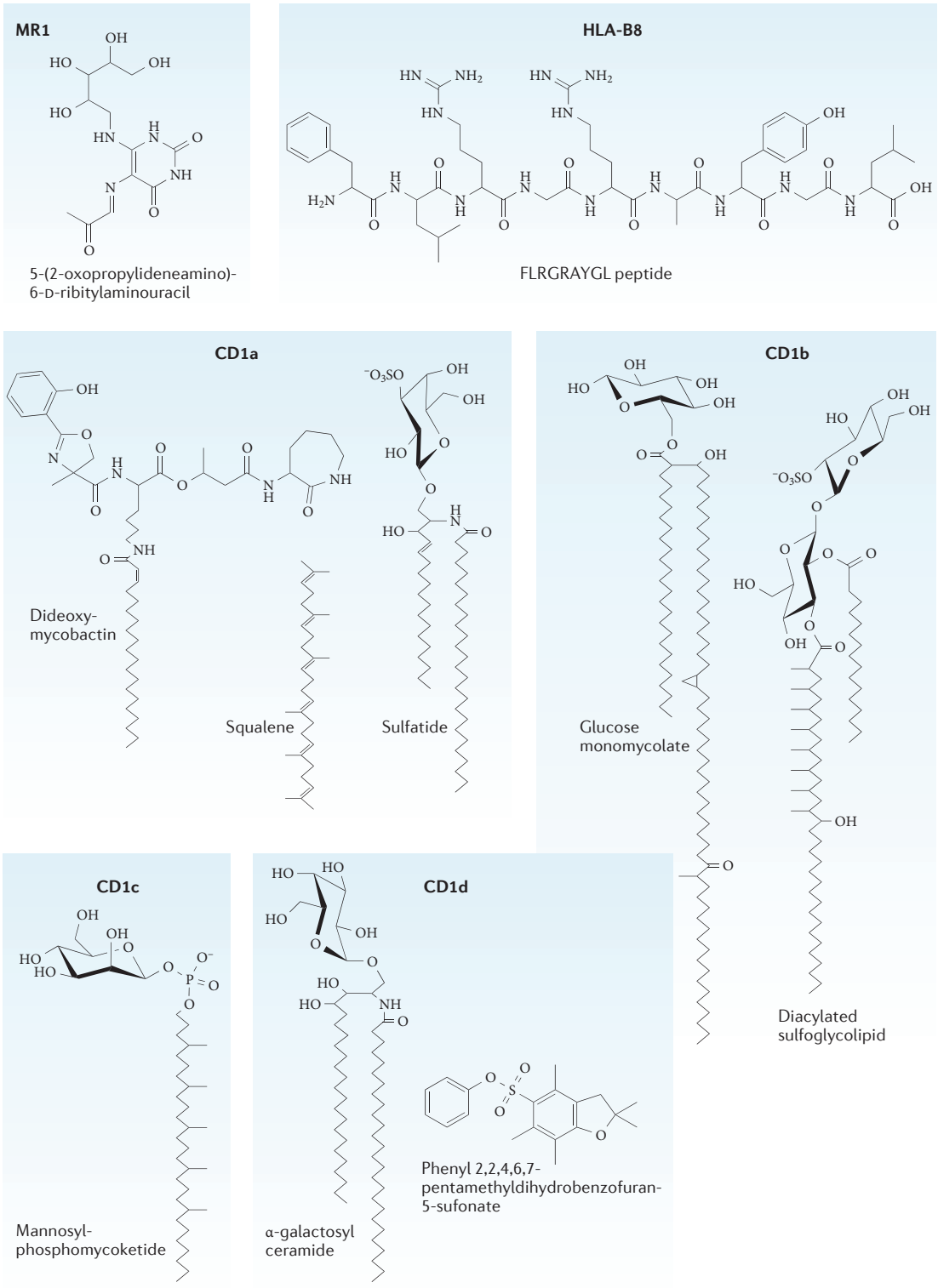

Figure 2 | **Antigens presented by CD1 and MR1.** In contrast to the short peptide ligands of MHC class I proteins (for example, HLA-B8), CD1 proteins typically present amphipathic lipids derived from self or foreign sources. Most CD1-presented antigens contain two distinct components. The hydrophilic head groups are comprised of carbohydrate, peptide or inorganic esters that protrude from CD1 molecules to contact T cell receptors. The flexible aliphatic hydrocarbon chains anchor the ligands within the CD1 cleft. However, not all CD1-presented antigens are amphipathic lipids. Skin oils, including squalene, lack a discernable hydrophilic head group[45]. The MHC class I-related protein (MR1)-presented antigen 5-(2-oxopropylideneamino)-6-D-ribitylaminouracil and the CD1d-presented antigen phenyl 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonate[44] are small molecules that are neither peptides nor lipids[65].

Skin oils. In 1989, T cells with autoreactivity to CD1a and CD1c proteins were discovered[22]. Based on the co-recognition model, it was assumed that this autoreactivity derives from TCR contact with defined lipid autoantigens bound in the CD1 cleft. Recently, this assumption was tested using a human CD1a-autoreactive αβ T cell line (BC2) to isolate CD1a-binding antigens produced in human cells and tissues[45]. CD1a-binding autoantigens preferentially accumulate in the skin[45], and this is consistent with studies showing that CD1a-autoreactive T cells home to the skin and that CD1a expression is higher in the skin than in other organs[7]. Extraction of whole cells or lipid–CD1 complexes with chloroform, followed by mass spectrometry, identified the CD1a-presented autoantigens as extremely hydrophobic skin oils: namely, wax esters, squalene and triacylglycerides[45] (FIG. 2). Unlike amphipathic lipids, oils lack hydrophilic head groups composed of sugars or other polar elements. Similar to PPBF, the small molecular volume of oils raises questions about how or whether they could protrude above the CD1 presentation platform to contact TCRs.

Scaffold lipids. Enzymes trim the carbohydrate moieties of glycolipids in antigen-presenting cells (APCs)[10,46,47], and nearly all antigens that bind MHC class Ia molecules are trimmed to fit the groove. However, the alkane lipid moieties of CD1-presented antigens are chemically unreactive and are not trimmed to fit[48]. For CD1a, CD1c and CD1d, the observed volumes of the antigen-binding clefts (1,280–1,780 Å^3) roughly match the combined lipid length of the two aliphatic hydrocarbon chains present in common cellular sphingolipids and diacylglycerides (36–46 carbons). However, CD1 ligands can have one, two or three alkyl chains with a length of 12–86 carbons, so individual antigens can diverge from the known volumes of CD1 grooves[49] (FIG. 2). So, how can CD1 proteins bind lipids with such varied chain lengths?

Moreover, the cleft of CD1b (2,200 Å^3) is nearly 50% larger than the cavities in other CD1 isoforms (FIG. 3), but few correspondingly large self lipids (~76 carbons) are present in mammalian membranes. When lipid ligands were eluted from CD1a, CD1b, CD1c and CD1d produced in human cells, the average mass of ligands released from CD1b was not larger than that released from CD1 proteins with smaller clefts[50]. A straightforward explanation would be that two or more lipids bind CD1b concomitantly. An early study showed that phosphatidylinositol bound in the 'upper chamber' of CD1b and that two unknown ligands (possibly detergents) were present in the T′ tunnel[42]. Crystal structures of lipid–CD1b complexes containing phosphatidylcholine or sulfoglycolipid later showed that electron densities corresponding to ligands in the cleft were larger than the known size of added ligands, implying the existence of chaperone lipids that bind CD1b alongside the added ligand[51,52]. Furthermore, by comparing crystal structures of CD1b molecules bound to various ligands, the positioning of CD1b residues near the TCR contact surface was influenced by ligand size, a finding that was interpreted as ligand sliding[51]. Along with similar studies of CD1d[53], these studies provide an explicit structural mechanism by which the size of the lipid within the groove could alter TCR contact sites on the outer surface of the CD1 complex (FIG. 3a). Mass spectrometry studies identified the endogenous chaperone lipids as diacylglycerides and deoxydihydroceramides[50,51] (FIG. 3b). These lipids were designated as scaffold lipids to emphasize that they bind below the antigen and can be thought of as pushing the antigen towards the TCR[50].

Scaffold lipids are similar to class II-associated invariant chain peptide (CLIP), except that scaffold lipids only partially occupy the CD1b cleft and so function as a sizing mechanism rather than blocking all ligand exchange. These insights might explain the differing selectivity of CD1b for small (32-carbon) and large (80-carbon) ligands

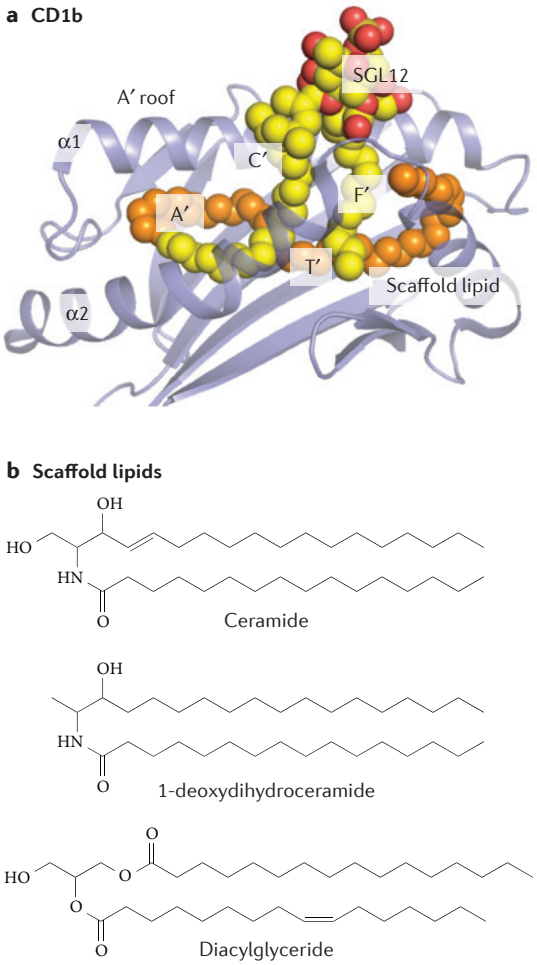

Figure 3 | **Scaffold and spacer lipids bind to CD1b.** **a** | A co-crystal study of CD1b bound to antigen (shown in yellow and red). The antigen is a synthetic sulfoglycolipid analogue known as SGL12 with a combined lipid length of 33 carbons. This structure also detects a second electron density (corresponding to a scaffold lipid, shown in orange) positioned in the lower chamber of the CD1b groove, within the T′ tunnel[51]. **b** | CD1b-binding natural scaffold lipids — diacylglycerides and deoxydihydroceramides — are unusually hydrophobic self lipids. The term 'scaffold' refers to the location of one particular kind of spacer lipid, which is located at the bottom of the CD1b groove and lifts of the sulfoglycolipid antigen towards the surface.

Wax esters
Fatty acids linked to an alcohol to form hydrophobic lipids, including those that accumulate on the skin surface.

Squalene
An abundant, organ-specific polyunsaturated branched chain lipid with 30 carbons that accumulates in the skin and activates T cells via CD1a.

Scaffold lipids
Specialized types of spacer lipid that are located within the lower section (T′ tunnel) of the CD1b cleft. Scaffold is an analogy to building scaffolds, which provide upwards-directed support to larger objects, which in this case is the antigen.

Class II-associated invariant chain peptide
(CLIP). A short amino acid sequence in the invariant chain that binds within the MHC class II groove shortly after translation so that it functions to block loading of self peptides during the early stages of MHC class II exit from the endoplasmic reticulum and Golgi apparatus.

when present in different cellular subcompartments[49]. The loading of large ligands is expected to require the removal of the antigen and the scaffold lipid from the upper and lower chambers[51]. This lipid exchange process is promoted by acid pH in lysosomes[49,54], which uncouples tethering amino acid side chains located on the top and side of the CD1b groove (at positions 80 and 86, respectively). These effects relax the CD1b conformation, creating a larger portal for exogenous antigen entry and ligand exit from the cleft[55]. By contrast, loading of small lipids is now thought to require emptying of the upper chamber only[50] and occurs efficiently at neutral pH, when the interdomain tethers are intact[42]. Direct measurements of lipid exchange within CD1 are limited[55–58]. However, these findings suggest a working model in which short self lipids can be loaded together with scaffold lipids at neutral pH in the secretory pathway, followed by the capture of larger foreign lipids during endosomal recycling.

Spacer lipids. The more general problem of lipid sizing for all CD1 isoforms is accomplished through spacer lipids and accessory portals. Unlike the upward push of scaffold lipids in CD1b, 'spacer lipid' is a more general term for any lipid that occupies the groove together with antigens. For example, certain CD1d[59], CD1c[10] and CD1a[60] crystal structures show electron densities that correspond to the lipid antigen present only in the F′ pocket with spacer lipids occupying the A′ pocket (FIG. 2). At the other end of the size range, large lipids probably protrude through small gaps in the lateral walls of CD1b and CD1c, known as the C′ and D′/E′ portals, respectively[42,61]. Unlike the F′ portal, which allows lipids to protrude onto the antigen display platform, these accessory portals are located beneath the α-helices and so allow the lipid to escape from a position that is distant from the TCR. Also, CD1a and CD1c molecules have unnamed notches in the lateral wall of the F′ pocket that could allow lateral escape of bound lipids[61,62]. In summary, whereas the MHC antigen display system trims antigens to fit, the CD1 molecules use scaffold lipids, spacer lipids, accessory portals and notches to fine-tune lipid ligands to cleft volume.

MR1-presented antigens

Vitamin B derivatives. Similar to CD1 molecules, MR1 molecules are comprised of non-polymorphic heavy chains bound to β2-microglobulin. For many years, MR1 molecules were known to mediate the activation of mouse and human mucosal-associated invariant T cells (MAIT cells) in response to certain microorganisms[12,13,63]. Using recombinant MR1 proteins to capture ligands derived from culture media and bacteria, the ligands were recently identified as a photodegradation product (6-formylpterin (6-FP)) of folic acid (vitamin B_9)[64]. MR1-presented antigens are modified metabolites derived from the riboflavin (vitamin B_2) pathway: 5-(2-oxopropylideneamino)-6-D-ribitylaminouracil (5-OP-RU) and 5-(2-oxoethylideneamino)-6-D-ribitylaminouracil (5-OE-RU), respectively[65]. Ribityl lumazines and ribityl pyrimidines lack aliphatic hydrocarbon chains (FIG. 2), reinforcing the concept that T cells can respond to small molecules that are neither peptides nor lipids.

MR1 antigen display. Crystal structures of 5-OP-RU–MR1 and 6-FP–MR1 complexes show an 'aromatic cradle' within the A′ pocket of the MR1 antigen-binding cleft that is well suited for binding the derivatives[64–67] (FIG. 4). At the base of this hydrophobic A′ pocket is a lysine residue (Lys43) that forms a Schiff base with 6-FP, 5-OP-RU and related ligands. In contrast to the substantially protruding head groups of CD1-presented antigens, these vitamin B derivatives are closely sequestered within the MR1 cleft, such that only a tiny proportion of the riboflavin derivative is exposed for TCR recognition[64,68] (FIG. 4).

Origin of MR1-presented antigens. Riboflavin derivatives can be considered as foreign antigens because they arise in certain bacteria and fungi. However, the most potent MAIT cell ligands derive from covalent conjugation of a riboflavin precursor to methylglyoxal or glyoxal and other host or microbial intermediates, so that the resulting product is a hybrid neo-antigen[65]. A central question now is whether MAIT cells mainly recognize vitamin B derivatives or also see other types of antigen. The range of bacterial and fungal organisms that activate MAIT cells[12,13,63] corresponds well to the range of organisms with riboflavin pathways[64]. This observation, along with the detection of large numbers of MAIT cells by MR1 tetramers bound to one type of ribityl pyrimidine ligand (5-OP-RU), suggests that one known antigen broadly supports recognition by MAIT cells. Furthermore, deletion of either of the two riboflavin biosynthetic genes (*ribA* and *ribG*) from *Lactococcus lactis*, or related genes in *Escherichia coli*, ablates T cell recognition, indicating that other antigens do not exist in these species[65,69]. However, crystal structures of 5-OP-RU–MR1 complexes show that the F′ pocket of MR1 remains empty, so larger ligands that are yet to be discovered might also bind within MR1.

TCR recognition of peptides, lipids and metabolites

Ternary crystal structures comprised of TCRs bound to CD1 (REF. 70) or MR1 protein complexes[71] can be compared with the 34 unique structures of TCR–MHC class I complexes[70]. TCR recognition of α-galactosylceramide (αGalCer)[14], 5-OP-RU[71] or an 11-mer peptide[72] involves prototypical co-recognition interactions: that is, the approaching TCR contacts a hybrid surface formed by the α-helical regions of the antigen-presenting molecule and the ligand, sitting in the cleft between the helices (FIG. 4). However, recent structures of αβ, γδ and δ/αβ TCRs have revealed clear and fundamental differences between CD1 and MHC antigen display. The studies detailed below describe unexpected modes of recognition that fall outside the co-recognition paradigm.

Left–right asymmetry of CD1 architecture. The key difference between MHC and CD1 proteins relates to the absence or presence of lateral symmetry in the antigen display platform (FIG. 5a). In MHC proteins, the TCR contact region can be imagined as a platform bisected by a plane that is perpendicular to the axis of the groove. The groove is accessible on both sides of the

Secretory pathway
A series of protein transport reactions in which newly folded proteins transit from the endoplasmic reticulum to the Golgi apparatus and to the cell surface. For CD1, this pathway provides self lipids that are loaded onto CD1 proteins at neutral pH.

Endosomal recycling
A process by which CD1 proteins shuttle from the cell surface to the endosomal network and back. CD1b, CD1c and CD1d proteins contain tyrosine-containing motifs in their cytoplasmic tails that mediate binding to adaptor proteins and transport to endosomes and lysosomes, where lipids derived from outside the antigen-presenting cell bind CD1 proteins at neutral or acidic pH.

Spacer lipids
Hydrophobic compounds that bind alongside antigenic lipids and fill up part of the CD1 cleft that is not occupied.

Mucosal-associated invariant T cells
(MAIT cells). T cells that express a structurally conserved invariant T cell receptor and are selected by MR1.

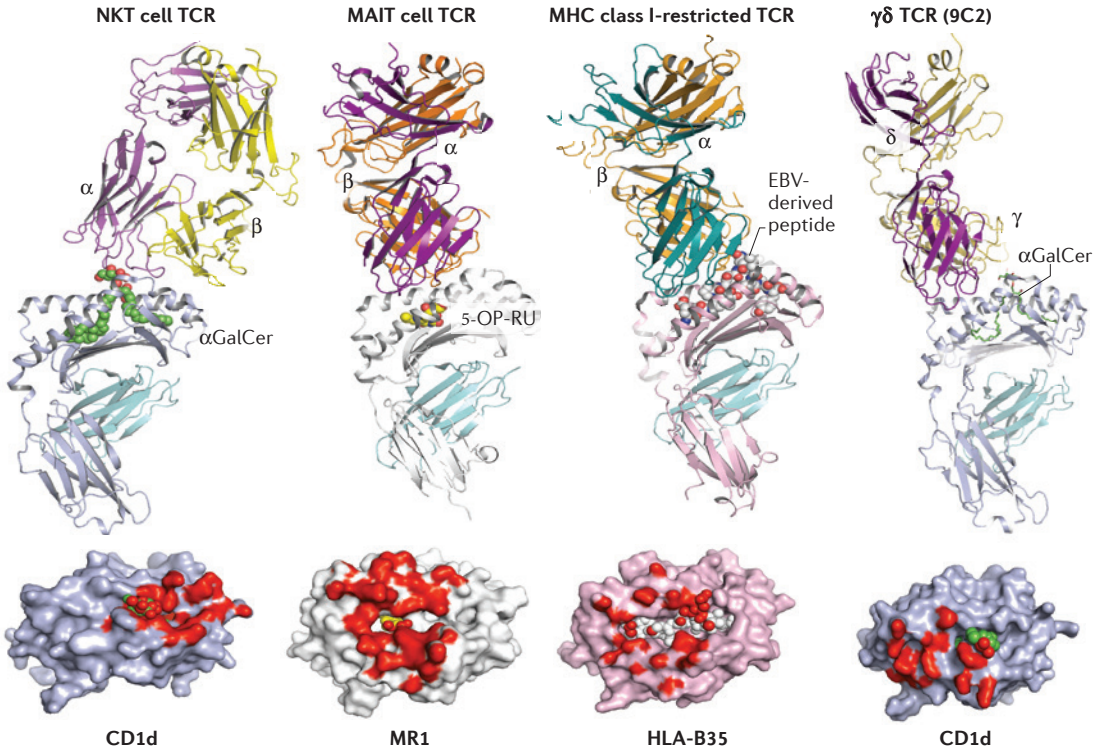

Figure 4 | **T cell receptors that function through co-recognition.** This figure shows examples of αβ T cell receptors (TCRs)[14,71,72] and a γδ TCR[16] that make direct contact with a hybrid surface formed by the antigen-presenting molecule and the antigen as it protrudes from the groove. The natural killer T (NKT) cell TCR interacts with CD1d displaying α-galactosylceramide (αGalCer); the mucosal-associated invariant T (MAIT) cell TCR interacts with MHC class I-related protein (MR1) displaying 5-(2-oxopropylideneamino)-6-D-ribitylaminouracil (5-OP-RU); the MHC class I-restricted TCR (ELS4) interacts with the MHC class I molecule HLA-B35 displaying an Epstein–Barr virus (EBV)-derived peptide; and the type II (9C2) γδ TCR interacts with CD1d displaying αGalCer. The TCR footprint is shown in red in the lower panels. In these four examples, the TCR contacts both the antigen-presenting molecule and the bound ligand, so these represent the mechanism of TCR co-recognition.

plane (FIG. 5a, green). By contrast, the four types of CD1 molecules show a fundamental left–right asymmetry (FIG. 5a, red and green). At one end of the CD1 cleft, located above the A′ pocket, amino acid side chains (interdomain tethers) reach across the space between the α1 and α2 helices to form the A′ roof[42,61,62,73]. Above the F′ pocket and on the right-hand side of the CD1 molecule, the hydrophilic head groups of the bound antigen protrude upwards to contact the TCR (FIG. 5a, green). Thus, the CD1 display platform is dominated by the outer surface of the CD1 protein on the left side and the lipid ligand on the right side.

A second general point of contrast between MHC and CD1 proteins is the extent to which antigens are exposed to the outer surface. Whereas MHC class I grooves are exposed to solvent across their lateral dimension for more than 20 Å, as measured from the interior cusp of the A′ pocket to the F′ pocket, the equivalent opening in CD1 is only ~10–13 Å wide (FIG. 5a). On the basis of its small size and rounded shape, this opening is known as the F′ portal. Whereas the term 'groove' accurately describes the long, laterally oriented and uncovered nature of the antigen-binding cleft in MHC molecules, CD1 clefts are much less like grooves because the A′ roof covers much of the top of the antigen-binding cleft. The cavity looks like an empty boot viewed from the side: broad at the bottom but tapering to a narrower opening on the right (FIG. 5b).

Influence of asymmetry on antigen display. With some notable exceptions[74,75], many TCR footprints on MHC class I molecules are located near the centre of the platform[70]. The central location and extensive exposure of peptides on MHC, along with the relatively large size of TCR footprints (1,200–2,400 Å^2), favours substantial contact of the TCR with exposed peptide[70]. By contrast, the asymmetrical locations of the F′ portals on CD1 molecules mean that the contribution of the ligand is pushed towards the edge of the platform. The portals in CD1 proteins are smaller than the grooves in MHC proteins in all cases reported to date, but the head groups of CD1 ligands vary in size. After traversing the F′ portal, head groups can lean leftwards and occupy a central location on CD1 or lean rightwards to reside at the extreme right edge of CD1 (FIGS 4,5). Therefore, individual TCRs could preferentially contact the CD1 protein on the left or instead take a central or right-sided approach to mainly contact the lipid ligand.

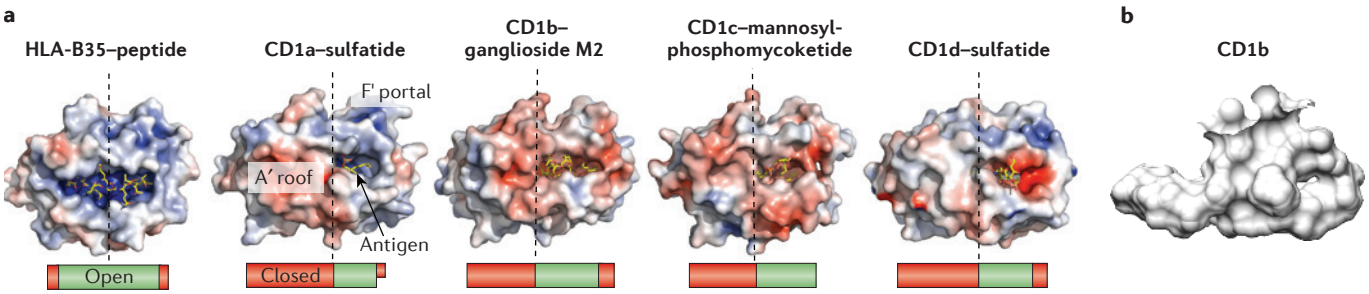

Figure 5 | **CD1 proteins have laterally asymmetrical antigen display platforms.** **a** | The grooves of MHC class I proteins (represented by HLA-B35 here; left) and MHC class II (not shown) proteins are broadly open to solvent, creating a situation in which peptide is presented on both sides of the bisecting plane (dashed line). Most interactions of the peptide with the T cell receptor (TCR) occur near the centre of the platform. By contrast, all four human CD1 antigen-presenting molecules use an asymmetrical antigen display platform. In human CD1 proteins, the α1 and α2 helices connect to form A′ roofs, which cover approximately half of the cleft. These roof structures prevent direct TCR binding to antigen positioned on the left side of the platform. Antigens such as sulfatide[62], ganglioside M2 (REF. 42) and phosphomycoketide[61] protrude from the open end of the cleft through a rounded opening known as the F′ portal. The hydrophilic head groups of individual antigens vary in size so that they can protrude minimally or extensively. The site of antigen protrusion is on the right side of the platform, but antigens with large head groups can pivot either to the left, creating an epitope near the centre of CD1, or to the right, creating an ectopic epitope at the extreme right. **b** | The transparent surface of the CD1b 'groove' demonstrates that it is not a groove-like structure that is broadly exposed to solvent. Instead, the shape of CD1 grooves resembles a boot viewed from the side because the A′ roof creates a constricted, portal-like connection with the TCR contact surface.

This left–right shift hypothesis is attractive because it might explain why so many CD1-reactive T cells show mixed antigen dependency and autoreactivity[7,8,22,76–78]. Furthermore, lateral shifts in TCR footprints might explain the varied dependence on antigen for individual TCRs or the predominant reliance on TCR α- or TCR β-chains for antigen recognition. Although there are still only a few crystal structures available, the TCR footprints on CD1 complexes are positioned off-centre with some near the extreme left or right side of the platform, creating modes of antigen recognition that are not known from the MHC system[70].

NKT cell TCR footprints on CD1d. In contrast to the typically diagonal–orthogonal orientation of TCRs on MHC class I complexes, the first structure of a lipid–CD1d–TCR complex revealed that the TCR is oriented in parallel with the axis of the cleft[14] (FIG. 4). The galactose head group of the ligand αGalCer rests on the centre-right side of CD1d. The TCR is markedly right-shifted such that the TCR α-chain is positioned above the galactose moiety and the TCR β-chain is positioned to the far right such that it makes minimal contact with CD1d and no contact with the antigen. This extreme right-sided footprint explains why the TCR α-chain is dominant for the antigen specificity of invariant NKT cells. Many other type I NKT cell TCRs also assume this right-sided footprint when recognizing other α- or β-linked hexosylceramide or isogloboside antigens[79]. By contrast, a variable (type II) NKT cell TCR known as XV19 showed an orthogonal docking and left-of-centre footprint. As it is docked over the A′ roof, the XV19 TCR makes most contact with CD1d, but it also makes some contact with the sulfatide head group near the centre of the platform. Overall, these glycolipid–CD1d–TCR structures illustrate the left–right shift and the differing rotation of αβ TCRs docking on CD1d molecules. Although CD1c has not been co-crystallized with TCRs, alanine scanning mutagenesis suggests that individual αβ TCRs also use different left-sided or right-sided footprints[80].

Absence of interference. The examples discussed above show how TCRs bind to exposed head groups. However, CD1a-presented skin oils lack obvious hydrophilic head groups for the usual hydrogen bonding or charge–charge interactions between TCR and antigen[1,2,81] (FIG. 2). Moreover, CD1a-autoreactive T cell clones lack the fine antigen specificity that is seen for recognition of glycolipids and peptides[78,82] and, instead they cross-react with several hydrophobic molecules, including methyl fatty acids, triacylglycerides and wax esters[45]. These findings suggest that the stimulatory compounds do not directly contact the TCR. Indeed, the small size and hydrophobic nature of squalene and related lipids might allow them to nest inside the CD1a groove, so that they do not interfere with TCR binding to the outer surface of CD1a; we have termed this mode of activation 'absence of interference' (REF. 45).

Supporting this hypothesis, CD1a proteins acquire antigens over time, and when emerging at the cell surface, they predominantly carry antigens with large hydrophilic head groups, including sphingomyelin. In fact, sphingomyelin and other non-permissive ligands block CD1a autoreactivity *in vitro*[45]. Unlike other CD1 isoforms, which extensively recycle to lysosomes, CD1a molecules mainly reside at the surface of epidermal Langerhans cells and capture exogenous ligands[83,84]. Wax esters accumulate in the cornified epithelium, and squalene is a major lipid in sebum, so both of these self ligands accumulate within or outside of the epidermis and not within APCs. Accordingly, the absence of interference model predicts

Sebum
Mixed hydrophobic oils that are produced by glands in the hair follicles and released to form an outer protective barrier on the skin surface.

that Langerhans cells normally express a variety of CD1a proteins that carry ligands that are non-permissive for TCR binding. After a breach in the surface of the skin, non-permissive self ligands might be exchanged for exogenous permissive ligands, such as squalene, unmasking the surface of CD1a complexes for direct recognition by TCRs[7]. The nested oils might stabilize the interior of the cleft and change the outer shape of the CD1a complex, permitting allosteric changes that favour TCR binding. Alternatively, permissive ligands might function by displacing larger non-permissive lipids to create space on the outer surface of CD1a for the approaching TCR.

TCR recognition of CD1a

Permissive ligands. The absence of interference theory has recently been tested through the study of ternary interactions between lipid, CD1a and TCR. Tetramers usually require loading with homogeneous antigen so that multiple arms of the tetramer carry the same antigen, creating the avidity needed to bind an antigen-specific T cell[85]. Therefore, it was surprising that a CD1a-autoreactive TCR known as BK6 could bind CD1a tetramers loaded with diverse self lipids[58]. Using the BK6 TCR to pull down bound lipid–CD1a complexes, lipidomic profiling detected hundreds of distinct lipids. Because many chemically distinct lipid ligands permitted TCR binding to CD1a, they were termed permissive ligands. Subtractive analysis showed that sphingomyelin was specifically excluded from lipid–CD1a–TCR complexes and blocked tetramer staining of BK6 TCR$^+$ cells, indicating that sphingomyelin is a non-permissive ligand[45,58].

This unusual TCR response to most but not all lipid ligands was explained by the ternary structure of BK6 TCR–lipid–CD1a. The BK6 TCR has an extreme left-sided footprint over the A′ roof of lysophosphatidylcholine (LPC)–CD1a complexes (FIG. 6a, red). LPC does not fully nest within the cleft, but its head group leans to the right as it traverses the F′ portal, exposing a large area on the A′ roof to which the TCR can bind (FIG. 6a,b). Importantly, the BK6 TCR contacts the A′ roof but not the LPC ligand. Thus, CD1a is the antigenic target, and LPC is a permissive ligand that allows the TCR to contact CD1a in the absence of interference.

A second structure of the BK6 TCR binding to CD1a carrying diverse endogenous ligands (endog–CD1a) also illustrates absence of interference (FIG. 6b). The endog–CD1a–BK6 TCR structure shows electron density in the F′ pocket that corresponds to that of a fatty acid (FIG. 6b, left). Unlike the LPC–CD1a structure, but consistent with the nesting hypothesis, the observed electron density does not protrude through the F′ portal to the surface. In fact, free fatty acids could be extracted from CD1a–TCR complexes, suggesting that some ligands are small enough to nest within CD1a[58]. Overall, permissive ligands function by two related mechanisms that fit the concept of absence of interference: nesting within CD1a or exiting the portal and turning to the right to become positioned away from the A′ roof.

Non-permissive ligands. The molecular mechanisms by which non-permissive ligands disrupt TCR binding are illustrated in binary structures of CD1a bound to sphingomyelin[58], sulfatide[62] or the mycobactin-like lipopeptide dideoxymycobactin[60] (FIG. 6c). In CD1a, the A′ roof is formed in part by salt bridges between the residues Arg73, Arg76 and Glu154, which tether the α1 and α2 helices. Sphingomyelin, sulfatide and dideoxymycobactin disrupt the intrinsic structure of the A′ roof. Thus, large ligands can interfere with TCR contact by inserting themselves within the intrinsic structure of the A′ roof or by protruding through the F′ portal to a position on top of the A′ roof.

Dual mechanisms. Any single CD1a-restricted TCR either does or does not contact lipid, so absence of interference and co-recognition are mutually exclusive mechanisms for individual TCRs. However, it is likely that the broader CD1a-reactive repertoire uses both mechanisms. Recognition of dideoxymycobactin–CD1a probably occurs via co-recognition (FIG. 6c), as loading of the lipopeptide into CD1a is necessary for TCR binding and this ligand does protrude from the cavity to alter the surface above the F′ portal[35,60,86,87]. For other known CD1a antigens, such as sphingolipids and phospholipids[88–90], the mechanism of recognition has not been solved structurally, but co-recognition is favoured on the basis of their large head groups.

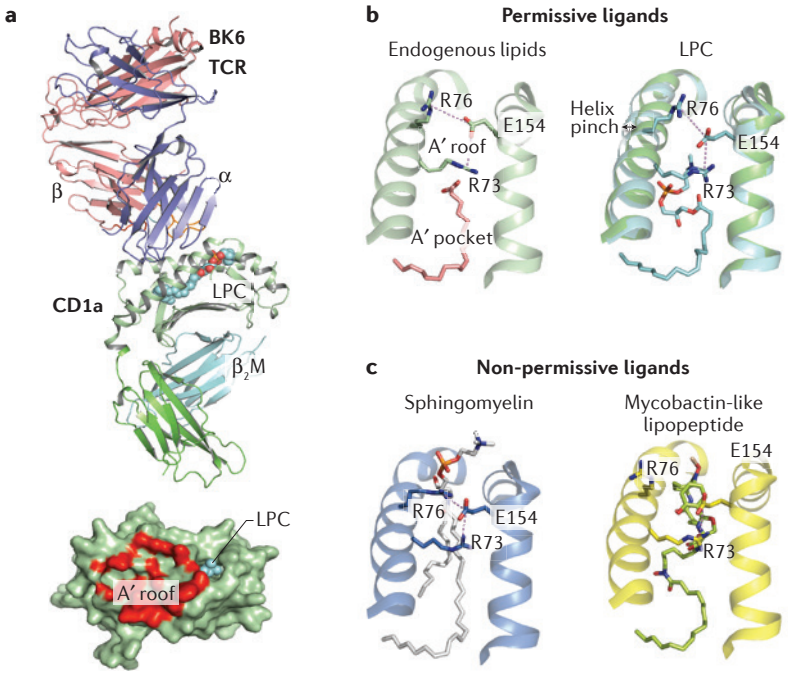

Figure 6 | **Direct recognition of CD1a molecules carrying permissive ligands.** **a** | The BK6 T cell receptor (TCR) forms a left-sided footprint (shown in red in the lower panel) that contacts CD1a itself rather than the lipid ligand lysophosphatidylcholine (LPC)[58]. **b** | Permissive lipid ligands act through absence of interference using two proposed mechanisms. Some endogenous lipids have small head groups that are predicted not to protrude substantially from the groove, whereas LPC does protrude from the groove but takes a rightward position on the surface so that it does not contact a left-binding TCR. **c** | Based on binary TCR structures[60,62], non-permissive ligands — such as sphingomyelin, sulfatide and the mycobactin-like lipopeptide — bind in the groove and are thought to have an anti-antigenic function by blocking TCR docking to CD1a. Non-permissive ligands can disrupt a triad of amino acids (R76, E154 and R73) within the A′ roof[58]. β2M, β2-microglobulin.

CD1 molecules bind γδ TCRs

There are many subsets of γδ T cells[91]. In humans, one subset expresses TCRγ variable chain 9 (TRGV9) and TCRδ variable chain 2 (TRDV2) and recognizes soluble antigens, such as alkyl phosphates and alkyl amines[92–94], probably indirectly through allosteric modification of the cell surface molecule butyrophilin 3A1 (REFS 95–97). Other γδ T cell subsets directly target surface molecules— such as endothelial protein C receptor[98], T10 (also known as CD38 and ADPRC1) and MHC class I polypeptide-related sequence A (MICA)[96].

The first report of CD1-mediated antigen presentation described a response by both αβ and γδ T cells[22]. With a few exceptions[99,100], CD1 research focused on αβ T cells, but recent tetramer[101] and crystallography[15–17] studies have confirmed the γδ TCR–CD1d interaction. The γδ TCR 9C2 (which is composed of TRGV5 and TRDV1) recognizes αGalCer–CD1d complexes (FIG. 4), whereas the γδ TCR DP10.7 (which is composed of TRGV4 and TRDV1) binds sulfatide–CD1d complexes. Both structures show a left-sided TCR footprint, orthogonal rotation and reliance on binding of tryptophan residues from complementarity-determining region 1δ (CDR1δ) to CD1d. One difference is that the TCR 9C2 uses the CDR3γ loop to bind the protruding αGalCer head group, whereas the TCR DP10.7 uses the hypervariable residues in CDR3δ for all interactions with sulfatide.

Several TRDV genes can also be used by αβ T cells, including *TRDV4–TRDV8*, and share TCRα variable chain (TRAV)-related names[102]. In addition, it has been long known that *TRDV1* and *TRDV3* genes can join TRAJ and TCRα constant chain (TRAC) genes[103–105], but it was unclear how these TRDV gene products contribute to the overall specificity of the 'hybrid' TCRs. A recent study[17] of such a hybrid TCR, composed of a TRDV1 domain fused to the TRAJ–TRAC domain and paired with a TCR β-chain, showed that it binds αGalCer–CD1d using an orthogonal and left-aligned docking mode. Mirroring the γδ TCR 9C2, the TRDV1 CDR1δ loop bound mainly to CD1d, whereas the CDR3β loop bound the galactosyl head group of αGalCer.

Role of CD1 ligands. Although most work on γδ T cells emphasizes direct TCR contact with the monomorphic surfaces of cell surface molecules[91], crystal structures involving lipid–CD1d complexes suggest that γδ TCRs recognize antigens that are physically displayed. This is supported by the observed contact of the CDR3δ and CDR3γ loops with sulfatide and αGalCer, respectively[15,16]. Indeed, the highly diverse CDR3 junctional residues might mediate non-cross-reactive recognition of other lipid antigens through direct TCR contact with antigen. An alternative view suggests that the most important interactions are between the γδ TCR and the CD1 molecule itself, rather than between the γδ TCR and the bound ligand. Specifically, the γδ TCR might indirectly detect the presence of any bound ligand in CD1 or the absence of an interfering ligand. Current evidence suggests that both scenarios occur because some γδ T cells seem to be highly dependent on the bound lipid antigen, whereas others seem to be tolerant to many different bound antigens[15,16,89,106,107]. The left–right shift hypothesis outlined for αβ TCRs (FIG. 5) might also be relevant to γδ TCRs (right panel of FIG. 4). The two known TRDV1 TCR footprints are left-shifted, and most interactions involve the A′ roof of the CD1 molecule. The bound ligands contribute in a small way to TCR contact, which is consistent with the observed partial dependence on ligand for TCR binding.

CD1 and MR1 tetramers

Mouse and human CD1d tetramers have been in use for 15 years, whereas tetramers of human CD1a[35], CD1b[9], CD1c[10] and MR1 (REFS 65,68) molecules were developed recently. *Ex vivo* tetramer studies enable researchers to avoid biases that may emerge during long-term *in vitro* culturing of T cells. Unlike MHC Ia molecules, CD1 proteins are almost identical among humans, as are MR1 proteins, so a single CD1 or MR1 tetramer reagent can be used for any human donor and used to quantitatively track antigen-specific T cells in disease states. Both invariant NKT cells[108,109] and MAIT cells[110] were discovered by detecting T cells expressing similar αβ TCRs. Only later were they were found to recognize CD1d[111] and MR1, respectively[63]. Now, with the generation of CD1 and MR1 tetramers, responding T cells can be tracked according to antigen specificity rather than TCR expression. By removing TCR gene usage as the means of detection, ligand-loaded tetramers allow broader and unbiased study of all TCRs that recognize a given antigen complex. This approach is already leading to a broadening of the types of TCR that meet the definition of MAIT cells[66,68], and it also led to the discovery of two previously unknown T cell types that recognize CD1b[112,113].

The CD1b-specific TCR repertoire. Both MAIT cells and NKT cells exist as large T cell populations that express nearly identical (invariant) but non-clonal TCRs in a process that occurs among genetically unrelated donors. Thus, the two defining features of these TCRs are intradonor and interdonor TCR conservation, whereas MHC-reactive TCRs generally lack these features. These conserved TCR patterns in NKT cells and MAIT cells derive from germline-encoded variable and joining genes with limited N-region additions to yield stringently conserved α-chains (TABLE 1).

Group 1 CD1 tetramers are now allowing the analysis of CD1a-, CD1b- and CD1c-reactive TCR repertoires. The earliest studies failed to detect intradonor or interdonor conservation in the group 1 CD1-reactive TCR repertoire[7,114,115]. These studies found that individual TCRs recognizing CD1a, CD1b or CD1c in combination with various antigens expressed differing variable, joining and diversity segments, suggesting that the group 1 CD1-restricted TCR repertoire is diverse. This finding was often offered as a point of contrast with the stereotyped nature and innate functions of NKT cells and MAIT cells. However, the non-polymorphic group 1 CD1 proteins might be expected to activate similar TCRs present among unrelated individuals. Indeed, when comparisons were simplified to assess the diversity of

human TCRs that recognize one CD1 protein (CD1b) paired with one antigen (mycobacterial glucose monomycolate (GMM)), intradonor and interdonor TCR conservation was readily identified (TABLE 1).

Fulfilling the criterion of intradonor conservation, polyclonal T cells expressed TCRs with highly similar TCR α-chains composed of TRAV1-2 joined to TCRα joining 9 (TRAJ9) paired with an apparently biased population of TCRβ variable chain 6-2 (TRBV6-2) chains[113]. This TCR pattern was seen among unrelated donors, indicating interdonor conservation. Such TCRs showed high affinity for GMM–CD1b complexes[112,113]. Based on the TCR structure and antigen specificity, such CD1b-reactive T cells were called germline-encoded mycolyl-specific (GEM) T cells[113]. GEM T cells are less frequent in human blood than NKT cells or MAIT cells, but the degree and pattern of TCR conservation is equivalent to that found in these other TCR-defined T cell subsets[113] (TABLE 1).

A separate population of GMM-reactive T cells with intermediate affinity for CD1b also showed interdonor TCR conservation but with a different TCR pattern and lower affinity for GMM–CD1b[112]. Among these T cell clones derived from patients with tuberculosis, TRBV4-1 was the most frequently used β-chain, and some clones expressed TRAV17. This pattern was seen nearly two decades ago in a clone (known as LDN5) derived from a patient with leprosy[82]. Thus, the newly discovered, polyclonal T cells were designated LDN5-like T cells. This name indicates that this TCR is not a unique or private TCR, as previously thought, but is instead an *in vivo*-expanded T cell type (TABLE 1). Thus, TCR bias and affinity define two compartments of the CD1b-reactive repertoire.

GMM is the first antigen to be systematically investigated for TCR diversity in the group 1 CD1 system, and it revealed two invariant T cell types. At a minimum, human CD1 proteins bind hundreds of ligands[50]. Thus, considering all available lipid–CD1 combinations, it is possible that the CD1 system supports a network of interdonor-conserved TCRs. Supporting this idea, recent studies of human TCRs isolated using CD1c tetramers loaded with mycobacterial phosphomycoketides show frequent expression of TRBV7-8$^+$ and TRBV7-9$^+$ TCRs[80] (TABLE 1). Sequencing of the TRAV1-2$^+$ T cell repertoire identified 16 additional TCR α-chains that do not match the known MAIT cell and GEM T cell TCR motifs but are conserved among the majority of human donors[116]. If conserved TCR patterns could be traced back to disease-related antigens (such as mycobacterial GMM), invariant TCRs might be used for immunodiagnosis.

Towards therapy

The immunodominant peptides for any pathogen or autoimmune disease differ according to the MHC haplotypes of the individual patients. Therefore, peptide-based immunomodulation is not broadly practiced in humans. However, the non-polymorphic nature of CD1 and MR1 proteins removes this key obstacle, so antigen-based T cell activation or polarization could be harnessed for therapy. Recent reviews document broad evidence that MAIT cells, CD1a-autoreactive T cells and NKT cells circulate in the blood and enter tissues in high numbers[8,12,117]. These cells secrete cytokines that have central roles in host defence and tissue repair[7,112,113,117–119]. Lipid antigens such as αGalCer are bioavailable, and

Table 1 | **Human invariant T cell types that recognize non-polymorphic antigen-presenting molecules**

Features of invariant T cells	CD1b	CD1b	CD1c	CD1d	MR1	MHC
Designation	GEM T cells	LDN5-like cells	ND	NKT cells	MAIT cells	αβ T cells
Antigen	Glucose monomycolate	Glucose monomycolate	Phosphomycoketide	αGalCer and others	Vitamin B metabolites	Peptides
TCR α-chain	Invariant	Biased	Diverse	Invariant	Invariant	Diverse
TCR α-chain variable gene	TRAV1-2	TRAV17	Diverse	TRAV10	TRAV1-2	All
TCR α-chain joining gene	TRAJ9	Diverse	Diverse	TRAJ18	TRAJ12, TRAJ20 and TRAJ33	All
TCR α-chain CDR3 length	Uniform	Uniform	Variable	Uniform	Uniform	Variable
TCR β-chain	Biased	Biased	Biased	Biased	Biased	Diverse
TCR β-chain variable gene	TRBV6-2	TRBV4-1	TRBV7-8 and TRBV7-9	TRBV25-1	TRBV6-1 and TRBV20-1	All
TCR β-chain CDR3 length	Variable	Variable	Variable	Variable	Variable	Variable
Intradonor conservation	High	High	High	High	High	Low
Interdonor conservation	High	High	High	High	High	Low

CDR3, complementarity-determining region 3; GEM, germline-encoded mycolyl-specific; MAIT, mucosal-associated invariant T; MR1, MHC class I-related protein; ND, not determined; NKT, natural killer T; TRAV, TCRα variable chain; TRAJ, TCRα joining chain; TRBV, TCRβ variable chain.

NKT cells can be activated and polarized in effector function by glycolipid antigens or altered glycolipid ligands[120,121]. Although therapeutic outcomes have been limited, αGalCer induces consistent T cell responses in clinical trials, regardless of patient genetic background, with some encouraging results from recent small-scale cancer and vaccine trials[122,123–125]. This early-stage work provides proof of principle to support more directed study of antigen formulation and administration. Vitamin B derivatives, group 1 CD1-reactive antigens or non-permissive ligands have only been discovered recently, and so they have not entered clinical trials but might now also be tested in humans to stimulate, block or detect T cell responses. Antigens that control human CD1-reactive or MR1-reactive T cells offer promise for the development of new, 'one-size-fits-all' T cell immunotherapy approaches that are not possible with polymorphic antigen-presenting molecules.

1. Garboczi, D. N. *et al.* Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature* **384**, 134–141 (1996).
2. Garcia, K. C. *et al.* An αβ T cell receptor structure at 2.5Å and its orientation in the TCR–MHC complex. *Science* **274**, 209–219 (1996).
3. Zinkernagel, R. M. The Nobel Lectures in Immunology. The Nobel Prize for Physiology or Medicine, 1996. Cellular immune recognition and the biological role of major transplantation antigens. *Scand. J. Immunol.* **46**, 421–436 (1997).
4. Salerno-Goncalves, R., Fernandez-Vina, M., Lewinsohn, D. M. & Sztein, M. B. Identification of a human HLA-E-restricted CD8$^+$ T cell subset in volunteers immunized with *Salmonella enterica* serovar Typhi strain Ty21a typhoid vaccine. *J. Immunol.* **173**, 5852–5862 (2004).
5. Heinzel, A. S. *et al.* HLA-E-dependent presentation of Mtb-derived antigen to human CD8$^+$ T cells. *J. Exp. Med.* **196**, 1473–1481 (2002).
6. Adams, E. J. & Luoma, A. M. The adaptable major histocompatibility complex (MHC) fold: structure and function of nonclassical and MHC class I-like molecules. *Annu. Rev. Immunol.* **31**, 529–561 (2013).
7. de Jong, A. *et al.* CD1a-autoreactive T cells are a normal component of the human αβ T cell repertoire. *Nat. Immunol.* **11**, 1102–1109 (2010).
8. de Lalla, C. *et al.* High-frequency and adaptive-like dynamics of human CD1 self-reactive T cells. *Eur. J. Immunol.* **41**, 602–610 (2011).
9. Kasmar, A. G. *et al.* CD1b tetramers bind αβ T cell receptors to identify a mycobacterial glycolipid-reactive T cell repertoire in humans. *J. Exp. Med.* **208**, 1741–1747 (2011).
10. Ly, D. *et al.* CD1c tetramers detect *ex vivo* T cell responses to processed phosphomycoketide antigens. *J. Exp. Med.* **210**, 729–741 (2013).
11. Dellabona, P. *et al. In vivo* persistence of expanded clones specific for bacterial antigens within the human T cell receptor α/β CD4$^-$8$^-$ subset. *J. Exp. Med.* **177**, 1763–1771 (1993).
12. Gold, M. C. *et al.* Human mucosal associated invariant T cells detect bacterially infected cells. *PLoS Biol.* **8**, e1000407 (2010).
Starting with a broad analysis of all human CD8$^+$ T cells, this paper demonstrates an *in vivo* response of human MAIT cells during natural *M. tuberculosis* infection.
13. Le Bourhis, L. *et al.* Antimicrobial activity of mucosal-associated invariant T cells. *Nat. Immunol.* **11**, 701–708 (2010).
14. Borg, N. A. *et al.* CD1d–lipid-antigen recognition by the semi-invariant NKT T-cell receptor. *Nature* **448**, 44–49 (2007).
15. Luoma, A. M. *et al.* Crystal structure of Vδ1 T cell receptor in complex with CD1d-sulfatide shows MHC-like recognition of a self-lipid by human γδ T cells. *Immunity* **39**, 1032–1042 (2013).
16. Uldrich, A. P. *et al.* CD1d–lipid antigen recognition by the γδ TCR. *Nat. Immunol.* **14**, 1137–1145 (2013).
References 15 and 16 report structures of TRDV1$^+$ γδ TCRs bound to lipid–CD1d complexes, providing detailed molecular mechanisms of CD1 recognition.
17. Pellicci, D. G. *et al.* The molecular bases of δ/αβ T cell-mediated antigen recognition. *J. Exp. Med.* **211**, 2599–2615 (2014).
18. McMichael, A. J. *et al.* A human thymocyte antigen defined by a hybrid myeloma monoclonal antibody. *Eur. J. Immunol.* **9**, 205–210 (1979).
19. Calabi, F. & Milstein, C. A novel family of human major histocompatibility complex-related genes not mapping to chromosome 6. *Nature* **323**, 540–543 (1986).
20. Calabi, F., Jarvis, J. M., Martin, L. & Milstein, C. Two classes of CD1 genes. *Eur. J. Immunol.* **19**, 285–292 (1989).
21. Dougan, S. K., Kaser, A. & Blumberg, R. S. CD1 expression on antigen-presenting cells. *Curr. Top. Microbiol. Immunol.* **314**, 113–141 (2007).
22. Porcelli, S. *et al.* Recognition of cluster of differentiation 1 antigens by human CD4$^-$CD8$^-$ cytolytic T lymphocytes. *Nature* **341**, 447–450 (1989).
This seminal paper demonstrates that CD1 proteins activate both αβ and γδ T cells. Whereas most work in the field has concentrated on αβ T cells, structural proof of γδ TCR binding to CD1d was finally published in 2013.
23. Roura-Mir, C. *et al. Mycobacterium tuberculosis* regulates CD1 antigen presentation pathways through TLR-2. *J. Immunol.* **175**, 1758–1766 (2005).
24. Baker, M. L. & Miller, R. D. Evolution of mammalian CD1: marsupial CD1 is not orthologous to the eutherian isoforms and is a pseudogene in the opossum *Monodelphis domestica*. *Immunology* **121**, 113–121 (2007).
25. Dossa, R. G., Alperin, D. C., Hines, M. T. & Hines, S. A. The equine CD1 gene family is the largest and most diverse yet identified. *Immunogenetics* **66**, 33–42 (2014).
26. Hayes, S. M. & Knight, K. L. Group 1 CD1 genes in rabbit. *J. Immunol.* **166**, 403–410 (2001).
27. Van Rhijn, I. *et al.* The bovine CD1 family contains group 1 CD1 proteins, but no functional CD1d. *J. Immunol.* **176**, 4888–4893 (2006).
28. Eguchi-Ogawa, T. *et al.* Analysis of the genomic structure of the porcine CD1 gene cluster. *Genomics* **89**, 248–261 (2007).
29. Looringh van Beeck, F. A. *et al.* Two canine CD1a proteins are differentially expressed in skin. *Immunogenetics* **60**, 315–324 (2008).
30. Kasmar, A., Van Rhijn, I. & Moody, D. B. The evolved functions of CD1 during infection. *Curr. Opin. Immunol.* **21**, 397–403 (2009).
31. de la Salle, H. *et al.* Assistance of microbial glycolipid antigen processing by CD1e. *Science* **310**, 1321–1324 (2005).
32. Brigl, M. & Brenner, M. B. CD1: antigen presentation and T cell function. *Annu. Rev. Immunol.* **22**, 817–890 (2004).
33. Yakimchuk, K. *et al. Borrelia burgdorferi* infection regulates CD1 expression in human cells and tissues via IL1-β. *Eur. J. Immunol.* **41**, 694–705 (2011).
34. Adams, E. J. Lipid presentation by human CD1 molecules and the diverse T cell populations that respond to them. *Curr. Opin. Immunol.* **26**, 1–6 (2014).
35. Kasmar, A. G. *et al.* Cutting edge: CD1a tetramers and dextramers identify human lipopeptide-specific T cells *ex vivo*. *J. Immunol.* **191**, 4499–4503 (2013).
36. 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).
37. Hiromatsu, K. *et al.* Induction of CD1-restricted immune responses in guinea pigs by immunization with mycobacterial lipid antigens. *J. Immunol.* **169**, 330–339 (2002).
38. Dascher, C. C. *et al.* Conservation of a CD1 multigene family in the guinea pig. *J. Immunol.* **163**, 5478–5488 (1999).
39. Porcelli, S., Morita, C. T. & Brenner, M. B. CD1b restricts the response of human CD4$^-$8$^-$ T lymphoyctes to a microbial antigen. *Nature* **360**, 593–597 (1992).
40. Beckman, E. M. *et al.* Recognition of a lipid antigen by CD1-restricted αβ$^+$ T cells. *Nature* **372**, 691–694 (1994).
41. Layre, E., de Jong, A. & Moody, D. B. Human T cells use CD1 and MR1 to recognize lipids and small molecules. *Curr. Opin. Chem. Biol.* **23**, 31–38 (2014).
42. Gadola, S. D. *et al.* Structure of human CD1b with bound ligands at 2.3Å, a maze for alkyl chains. *Nat. Immunol.* **3**, 721–726 (2002).
This CD1b structure work provides the earliest structural insights into lipid–CD1 complex formation and initiated functional work on ligand size, accessory portals, spacer lipids and the two-compartment antigen loading model.
43. Wang, J. *et al.* Lipid binding orientation within CD1d affects recognition of *Borrelia burgorferi* antigens by NKT cells. *Proc. Natl Acad. Sci. USA* **107**, 1535–1540 (2010).
44. Van Rhijn, I. *et al.* CD1d-restricted T cell activation by nonlipidic small molecules. *Proc. Natl Acad. Sci. USA* **101**, 13578–13583 (2004).
45. de Jong, A. *et al.* CD1a-autoreactive T cells recognize natural skin oils that function as headless antigens. *Nat. Immunol.* **15**, 177–185 (2014).
This paper identifies natural, tissue-specific autoantigens for the CD1a system. Surprisingly, these lipid autoantigens were small in size and lacked hydrophilic head groups, which supported a model in which ligands act by absence of interference with the TCR binding.
46. Prigozy, T. I. *et al.* Glycolipid antigen processing for presentation by CD1d molecules. *Science* **291**, 664–667 (2001).
47. Mattner, J. *et al.* Exogenous and endogenous glycolipid antigens activate NKT cells during microbial infections. *Nature* **434**, 525–529 (2005).
48. Cheng, T. Y. *et al.* Role of lipid trimming and CD1 groove size in cellular antigen presentation. *EMBO J.* **25**, 2989–2999 (2006).
49. Moody, D. B. *et al.* Lipid length controls antigen entry into endosomal and nonendosomal pathways for CD1b presentation. *Nat. Immunol.* **3**, 435–442 (2002).
50. Huang, S. *et al.* Discovery of deoxyceramides and diacylglycerols as CD1b scaffold lipids among diverse groove-blocking lipids of the human CD1 system. *Proc. Natl Acad. Sci. USA* **108**, 19335–19340 (2011).
51. Garcia-Alles, L. F. *et al.* Structural reorganization of the antigen-binding groove of human CD1b for presentation of mycobacterial sulfoglycolipids. *Proc. Natl Acad. Sci. USA* **108**, 17755–17760 (2011).
This paper shows how CD1b ligands of different sizes can slide upwards towards the top of the groove, and it identifies spacer lipids in the CD1b system.
52. Garcia-Alles, L. F. *et al.* Endogenous phosphatidylcholine and a long spacer ligand stabilize the lipid-binding groove of CD1b. *EMBO J.* **25**, 3684–3692 (2006).
53. McCarthy, C. *et al.* The length of lipids bound to human CD1d molecules modulates the affinity of NKT cell TCR and the threshold of NKT cell activation. *J. Exp. Med.* **204**, 1131–1144 (2007).
54. Jackman, R. M. *et al.* The tyrosine-containing cytoplasmic tail of CD1b is essential for its efficient presentation of bacterial lipid antigens. *Immunity* **8**, 341–351 (1998).
55. Relloso, M. *et al.* pH-dependent interdomain tethers of CD1b regulate its antigen capture. *Immunity* **28**, 774–786 (2008).
56. Im, J. S. *et al.* Kinetics and cellular site of glycolipid loading control the outcome of natural killer T cell activation. *Immunity* **30**, 888–898 (2009).
57. Im, J. S. *et al.* Direct measurement of antigen binding properties of CD1 proteins using fluorescent lipid probes. *J. Biol. Chem.* **279**, 299–310 (2004).
58. Birkinshaw, R. W. *et al.* αβ T cell antigen receptor recognition of CD1a presenting self lipid ligands. *Nat. Immunol.* **16**, 258–266 (2015).
59. Wun, K. S. *et al.* A molecular basis for the exquisite CD1d-restricted antigen specificity and functional responses of natural killer T cells. *Immunity* **34**, 327–339 (2011).
60. Zajonc, D. M. *et al.* Molecular mechanism of lipopeptide presentation by CD1a. *Immunity* **22**, 209–219 (2005).

61. Scharf, L. *et al.* The 2.5Å structure of CD1c in complex with a mycobacterial lipid reveals an open groove ideally suited for diverse antigen presentation. *Immunity* **33**, 853–862 (2010).
This study analyses a crystal structure of CD1c bound to mannosyl-phosphomycoketide, revealing a large F′ portal, a mechanism for capture of branched chain lipids and two accessory portals.
62. Zajonc, D. M., Elsliger, M. A., Teyton, L. & Wilson, I. A. Crystal structure of CD1a in complex with a sulfatide self antigen at a resolution of 2.15Å. *Nat. Immunol.* **4**, 808–815 (2003).
63. Treiner, E. *et al.* Selection of evolutionarily conserved mucosal-associated invariant T cells by MR1. *Nature* **422**, 164–169 (2003).
64. Kjer-Nielsen, L. *et al.* MR1 presents microbial vitamin B metabolites to MAIT cells. *Nature* **491**, 717–723 (2012).
This paper provides several advances in the understanding of MAIT cell biology, including the identification of the chemical structures of MR1 antigens as vitamin B metabolites that matched the range of pathogens that activate MAIT cells, as well as the identification of an unexpected mechanism of covalent attachment in generating stable metabolite–MR1 complexes at the surface of cells.
65. Corbett, A. J. *et al.* T-cell activation by transitory neo-antigens derived from distinct microbial pathways. *Nature* **509**, 361–365 (2014).
66. Eckle, S. B. *et al.* A molecular basis underpinning the T cell receptor heterogeneity of mucosal-associated invariant T cells. *J. Exp. Med.* **211**, 1585–1600 (2014).
67. McWilliam, H. E., Birkinshaw, R. W., Villadangos, J. A., McCluskey, J. & Rossjohn, J. MR1 presentation of vitamin B-based metabolite ligands. *Curr. Opin. Immunol.* **34**, 28–34 (2015).
68. Reantragoon, R. *et al.* Antigen-loaded MR1 tetramers define T cell receptor heterogeneity in mucosal-associated invariant T cells. *J. Exp. Med.* **210**, 2305–2320 (2013).
69. Soudais, C. *et al. In vitro* and *in vivo* analysis of the Gram-negative bacteria-derived riboflavin precursor derivatives activating mouse MAIT cells. *J. Immunol.* **194**, 4641–4649 (2015).
70. Rossjohn, J. *et al.* T cell antigen receptor recognition of antigen-presenting molecules. *Annu. Rev. Immunol.* **33**, 169–200 (2015).
71. Patel, O. *et al.* Recognition of vitamin B metabolites by mucosal-associated invariant T cells. *Nat. Commun.* **4**, 2142 (2013).
72. 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).
73. Zeng, Z. *et al.* Crystal structure of mouse CD1: an MHC-like fold with a large hydrophobic binding groove. *Science* **277**, 339–345 (1997).
74. Hahn, M., Nicholson, M. J., Pyrdol, J. & Wucherpfennig, K. W. Unconventional topology of self peptide–major histocompatibility complex binding by a human autoimmune T cell receptor. *Nat. Immunol.* **6**, 490–496 (2005).
75. Liu, Y. C. *et al.* A molecular basis for the interplay between T cells, viral mutants, and human leukocyte antigen micropolymorphism. *J. Biol. Chem.* **289**, 16688–16698 (2014).
76. Brossay, L. *et al.* Mouse CD1-autoreactive T cells have diverse patterns of reactivity to CD1$^+$ targets. *J. Immunol.* **160**, 3681–3688 (1998).
77. Mallevaey, T. *et al.* A molecular basis for NKT cell recognition of CD1d–self-antigen. *Immunity* **34**, 315–326 (2011).
78. Kawano, T. *et al.* CD1d-restricted and TCR-mediated activation of Vα14 NKT cells by glycosylceramides. *Science* **278**, 1626–1629 (1997).
79. Rossjohn, J., Pellicci, D. G., Patel, O., Gapin, L. & Godfrey, D. I. Recognition of CD1d-restricted antigens by natural killer T cells. *Nat. Rev. Immunol.* **12**, 845–857 (2012).
80. Roy, S. *et al.* Molecular basis of mycobacterial lipid antigen presentation by CD1c and its recognition by αβ T cells. *Proc. Natl Acad. Sci. USA* **111**, E4648–E4657 (2014).
81. Garboczi, D. N. *et al.* Assembly, specific binding, and crystallization of a human TCR-αβ with an antigenic Tax peptide from human T lymphotropic virus type 1 and the class I MHC molecule HLA-A2. *J. Immunol.* **157**, 5403–5410 (1996).
82. Moody, D. B. *et al.* Structural requirements for glycolipid antigen recognition by CD1b-restricted T cells. *Science* **278**, 283–286 (1997).
83. Manolova, V. *et al.* Functional CD1a is stabilized by exogenous lipids. *Eur. J. Immunol.* **36**, 1083–1092 (2006).
84. Sugita, M. *et al.* Separate pathways for antigen presentation by CD1 molecules. *Immunity* **11**, 743–752 (1999).
85. Altman, J. D. *et al.* Phenotypic analysis of antigen-specific T lymphocytes. *Science* **274**, 94–96 (1996).
86. Moody, D. B. *et al.* T cell activation by lipopeptide antigens. *Science* **303**, 527–531 (2004).
87. Young, D. C. *et al.* Synthesis of dideoxymycobactin antigens presented by CD1a reveals T cell fine specificity for natural lipopeptide structures. *J. Biol. Chem.* **284**, 25087–25096 (2009).
88. Shamshiev, A. *et al.* Presentation of the same glycolipid by different CD1 molecules. *J. Exp. Med.* **195**, 1013–1021 (2002).
89. Agea, E. *et al.* Human CD1-restricted T cell recognition of lipids from pollens. *J. Exp. Med.* **202**, 295–308 (2005).
90. Bourgeois, E. A. *et al.* Bee venom processes human skin lipids for presentation by CD1a. *J. Exp. Med.* **212**, 149–163 (2015).
91. Chien, Y. H., Meyer, C. & Bonneville, M. γδ T cells: first line of defense and beyond. *Annu. Rev. Immunol.* **32**, 121–155 (2014).
92. Tanaka, Y., Morita, C. T., Nieves, E., Brenner, M. B. & Bloom, B. R. Natural and synthetic non-peptide antigens recognized by human γδ T cells. *Nature* **375**, 155–158 (1995).
93. Bukowski, J. F., Morita, C. T. & Brenner, M. B. Human γδ T cells recognize alkylamines derived from microbes, edible plants, and tea: implications for innate immunity. *Immunity* **11**, 57–65 (1999).
94. Tanaka, Y. *et al.* Nonpeptide ligands for human γδ T cells. *Proc. Natl Acad. Sci. USA* **91**, 8175–8179 (1994).
95. Vavassori, S. *et al.* Butyrophilin 3A1 binds phosphorylated antigens and stimulates human γδ T cells. *Nat. Immunol.* **14**, 908–916 (2013).
96. Adams, E. J., Gu, S. & Luoma, A. M. Human γδ T cells: evolution and ligand recognition. *Cell. Immunol.* **296**, 31–40 (2015).
97. Sandstrom, A. *et al.* The intracellular B30.2 domain of butyrophilin 3A1 binds phosphoantigens to mediate activation of human Vγ9Vδ2 T cells. *Immunity* **40**, 490–500 (2014).
98. Willcox, C. R. *et al.* Cytomegalovirus and tumor stress surveillance by binding of a human γδ T cell antigen receptor to endothelial protein C receptor. *Nat. Immunol.* **13**, 872–879 (2012).
99. Leslie, D. S. *et al.* CD1-mediated γ/δ T cell maturation of dendritic cells. *J. Exp. Med.* **196**, 1575–1584 (2002).
100. Spada, F. M., Koezuka, Y. & Porcelli, S. A. CD1d-restricted recognition of synthetic glycolipid antigens by human natural killer T cells. *J. Exp. Med.* **188**, 1529–1534 (1998).
101. Bai, L. *et al.* The majority of CD1d-sulfatide-specific T cells in human blood use a semiinvariant Vδ1 TCR. *Eur. J. Immunol.* **42**, 2505–2510 (2012).
102. Lefranc, M. P. & Rabbitts, T. H. Genetic organization of the human T-cell receptor gamma and delta loci. *Res. Immunol.* **141**, 565–577 (1990).
103. Miossec, C. *et al.* Further analysis of the T cell receptor γ/δ$^+$ peripheral lymphocyte subset. The Vδ1 gene segment is expressed with either Cα or Cδ. *J. Exp. Med.* **171**, 1171–1188 (1990).
104. Miossec, C. *et al.* Molecular characterization of human T cell receptor α chains including a Vδ1-encoded variable segment. *Eur. J. Immunol.* **21**, 1061–1064 (1991).
105. Peyrat, M. A. *et al.* Repertoire analysis of human peripheral blood lymphocytes using a human Vδ3 region-specific monoclonal antibody. Characterization of dual T cell receptor (TCR) δ-chain expressors and αβ T cells expressing Vδ3JαCα-encoded TCR chains. *J. Immunol.* **155**, 3060–3067 (1995).
106. Dieude, M. *et al.* Cardiolipin binds to CD1d and stimulates CD1d-restricted γδ T cells in the normal murine repertoire. *J. Immunol.* **186**, 4771–4781 (2011).
107. Mangan, B. A. *et al.* Cutting edge: CD1d restriction and Th1/Th2/Th17 cytokine secretion by human vδ3 T cells. *J. Immunol.* **191**, 30–34 (2013).
108. Fowlkes, B. J. *et al.* A novel population of T-cell receptor αβ-bearing thymocytes which predominantly expresses a single Vβ gene family. *Nature* **329**, 251–254 (1987).
109. Budd, R. C. *et al.* Developmentally regulated expression of T cell receptor β chain variable domains in immature thymocytes. *J. Exp. Med.* **166**, 577–582 (1987).
110. Tilloy, F. *et al.* An invariant T cell receptor α chain defines a novel TAP- independent major histocompatibility complex class Ib-restricted α/β T cell subpopulation in mammals. *J. Exp. Med.* **189**, 1907–1921 (1999).
111. Bendelac, A. *et al.* CD1 recognition by mouse NK1$^+$ T lymphocytes. *Science* **268**, 863–865 (1995).
112. Van Rhijn, I. *et al.* TCR bias and affinity define two compartments of the CD1b–glycolipid-specific T cell repertoire. *J. Immunol.* **192**, 4054–4060 (2014).
113. Van Rhijn, I. *et al.* A conserved human T cell population targets mycobacterial antigens presented by CD1b. *Nat. Immunol.* **14**, 706–713 (2013).
Expanding the field of invariant T cells beyond NKT cells and MAIT cells, this paper discovered human GEM T cells specific for CD1b.
114. Grant, E. P. *et al.* Molecular recognition of lipid antigens by T cell receptors. *J. Exp. Med.* **189**, 195–205 (1999).
115. Cohen, N. R., Garg, S. & Brenner, M. B. Antigen presentation by CD1 lipids, T cells, and NKT cells in microbial immunity. *Adv. Immunol.* **102**, 1–94 (2009).
116. van Schaik, B. *et al.* Discovery of invariant T cells by next-generation sequencing of the human TCR α-chain repertoire. *J. Immunol.* **193**, 5338–5344 (2014).
117. Le Bourhis, L., Mburu, Y. K. & Lantz, O. MAIT cells, surveyors of a new class of antigen: development and functions. *Curr. Opin. Immunol.* **25**, 174–180 (2013).
118. Fernandez, C. S. *et al.* MAIT cells are depleted early but retain functional cytokine expression in HIV infection. *Immunol. Cell Biol.* **93**, 177–188 (2015).
119. Gold, M. C. *et al.* Human thymic MR1-restricted MAIT cells are innate pathogen-reactive effectors that adapt following thymic egress. *Mucosal Immunol.* **6**, 35–44 (2013).
120. Miyamoto, K., Miyake, S. & Yamamura, T. A synthetic glycolipid prevents autoimmune encephalomyelitis by inducing T_H2 bias of natural killer T cells. *Nature* **413**, 531–534 (2001).
121. Kim, E. Y., Lynch, L., Brennan, P. J., Cohen, N. R. & Brenner, M. B. The transcriptional programs of iNKT cells. *Semin. Immunol.* **27**, 26–32 (2015).
122. Fujii, S. *et al.* NKT cells as an ideal anti-tumor immunotherapeutic. *Front. Immunol.* **4**, 409 (2013).
123. Tefit, J. N. *et al.* Efficacy of ABX196, a new NKT agonist, in prophylactic human vaccination. *Vaccine* **32**, 6138–6145 (2014).
124. Motohashi, S. *et al.* A Phase I–II study of α-galactosylceramide-pulsed IL-2/GM-CSF-cultured peripheral blood mononuclear cells in patients with advanced and recurrent non-small cell lung cancer. *J. Immunol.* **182**, 2492–2501 (2009).
125. Schneiders, F. L. *et al.* Clinical experience with α-galactosylceramide (KRN7000) in patients with advanced cancer and chronic hepatitis B/C infection. *Clin. Immunol.* **140**, 130–141 (2011).
126. Salomonsen, J. *et al.* Two CD1 genes map to the chicken MHC, indicating that CD1 genes are ancient and likely to have been present in the primordial MHC. *Proc. Natl Acad. Sci. USA* **102**, 8668–8673 (2005).
127. Miller, M. M. *et al.* Characterization of two avian MHC-like genes reveals an ancient origin of the CD1 family. *Proc. Natl Acad. Sci. USA* **102**, 8674–8679 (2005).
128. Nguyen, T. K. A. *et al.* Expression patterns of bovine CD1 *in vivo* and assessment of the specificities of the anti-bovine CD1 antibodies. *PLoS ONE* **10**, e129755 (2015).
129. Looringh van Beeck, F. A. *et al.* Functional CD1d and/or NKT cell invariant chain transcript in horse, pig, African elephant and guinea pig, but not in ruminants. *Mol. Immunol.* **46**, 1424–1431 (2009).

Acknowledgements
The authors thank R. Birkinshaw, J. Le Nours, S. Huang, T.-Y. Cheng and K. Wucherpfennig for advice and graphical images. This work is supported by the Bill and Melinda Gates Foundation, National Institute of Allergy and Infectious Diseases (NIAID; grants AI049313, AI111224 and U19 111224), National Health and Medical Research Council (NHMRC) of Australia (1013667 and 1083942), Australian Research Council (DP140100977 and CE140100011), Cancer Council of Victoria (1042866), NHMRC Senior Principal Research Fellowship (to D.I.G.; 1020770) and NHMRC Australia Fellowship (to J.R.; AF50). All figures, including figure 1 and figure 5b, are original.

Competing interests statement
The authors declare no competing interests.