Published in final edited form as:

Nat Rev Immunol. 2015 October ; 15(10): 643–654. doi:10.1038/nri3889.

Lipid and Small Molecule Display by CD1 and MR1

Ildiko Van Rhijn^{1,2}, Dale I. Godfrey^{2,3,4}, Jamie Rossjohn^{5,6,7}, D. Branch Moody¹

¹Division of Rheumatology, Immunology and Allergy, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, MA02115, USA

²Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands

³Department of Microbiology and Immunology, Peter Doherty Institute for Infection and Immunity, University of Melbourne, Parkville, Victoria 3010, Australia

⁴Australian Research Council Centre of Excellence for Advanced Molecular Imaging, University of Melbourne, Parkville, Victoria 3010, Australia

⁵The Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria 3800, Australia

⁶Institute of Infection and Immunity, Cardiff University School of Medicine, Heath Park, Cardiff CF14 4XN, UK

⁷Australian Research Council Centre of Excellence in Advanced Molecular Imaging, Monash University, Clayton, Victoria 3800, Australia

Abstract

The antigen-presenting molecules CD1 and MR1 display lipids and small molecules to T cells. The antigen display platforms in the four human CD1 proteins are laterally asymmetrical, so that the TCR binding surfaces are comprised of roofs and portals, rather than the long grooves seen in the MHC antigen-presenting molecules. T cell receptors can bind CD1 proteins with left- or right-sided footprints, creating unexpected modes of antigen recognition. The use of tetramers of human CD1a, CD1b CD1c or MR1 proteins now allow detailed analysis of the human T cell repertoire, and 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.

A primary activation signal for T cells occurs when $\alpha\beta$ 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 $\alpha\beta$ TCRs to respond to the same antigen, thereby forming highly distinct (or private) TCR repertoires. To date, nearly all technology that seeks to manipulate or detect human $\alpha\beta$ T cell responses is 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

Correspondence to D.B.M. bmoody@partners.org.

The authors have no competing interests as defined by Nature Publishing Group, or other interests that might be perceived to influence the interpretation of the article.

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 portion of the overall $\alpha\beta$ T cell repertoire recognizes non-peptide antigens presented by non-polymorphic antigen-presenting molecules that are encoded within the MHC locus (such as HLA-E⁴⁻⁶) or outside the MHC locus (such as CD1a^{7,8}, CD1b⁹, CD1c¹⁰, CD1d¹¹ and MHC class I-related protein (MR1)^{12,13}). Moreover, $\alpha\beta$ ¹⁴, $\gamma\delta$ ^{15,16} and $\delta/\alpha\beta$ 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 $\beta 2$ microglobulin-associated heavy chains, which were later named as MHC class I and CD1 proteins¹⁸. Human CD1 genes are located on chromosome 1¹⁹ and encode five CD1 isoforms, which were assigned to group 1 (CD1a, CD1b, CD1c and CD1e) or group 2 (CD1d) based sequence homology²⁰. A second reason for designating two groups is that, whereas CD1d is constitutively expressed²¹, the group 1 CD1 genes are inducible and coordinately regulated by myeloid cells^{22,23}. CD1 genes are present in all placental mammals, birds and marsupials²⁴ (Figure 1). The differing size and composition of CD1 loci in modern mammals probably reflect selective pressure owing to immune function²⁵⁻²⁹. 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 of each CD1 isoform. The exception is muroid rodents, which encode two copies of one isoform, CD1d1 and CD1d2.

Human CD1 proteins.

It is now clear that each type of CD1 protein has a distinct cell biological function³⁰. For example, CD1e is a soluble lipid transfer protein³¹, whereas CD1a, CD1b, CD1c and CD1d are membrane-bound antigen-presenting molecules. Each type of CD1 protein takes different routes through cells³² and has markedly different expression patterns on B cells (CD1c⁺CD1d⁺), myeloid dendritic cells (CD1a⁺CD1b⁺CD1c⁺CD1d⁺), epithelial cells (CD1d⁺) and Langerhans cells (CD1a⁺)²¹. The transcription of the group 1 and group 2 CD1 genes is differently induced by microbial stimuli, with bacterial stimuli selectively upregulating group 1 CD1 protein expression by myeloid DCs^{23,33}. Furthermore, each human CD1 protein has 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. Because mice lack group 1 CD1 proteins, most studies have focused on the recognition of CD1d by a population of CD1d-dependent $\alpha\beta$ 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}, CD1 transgenic mice³⁶ 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 the study 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 characteristic of mycobacteria⁴⁰. Many more types of lipid antigen have since been identified, including glycolipid, phospholipid, glycopospholipid, sulfolipid and lipopeptide antigens⁴¹. 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 (Figure 2). The head groups vary in size, ranging from the small carboxylate moiety in free mycolic acid to 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 cavities⁴². Lipid interactions with the interior of the groove are relatively non-specific, as one ligand can insert in different orientations, but in general, the head group positioning is precise⁴³.

Small molecules.

As an exception to the general rule that CD1-dependent T cell activation occurs in response to amphipathic lipids, phenyl pentamethyldihydrobenzofuran sulfonate (PPBF) is a synthetic, non-lipid small molecule that activates T cells via CD1d⁴⁴. The molecular mechanism of PPBF action is not fully understood and was difficult to imagine because it lacks the flexible aliphatic chains and a discrete hydrophilic head group (Figure 2). Furthermore, the TCR co-recognition model predicts 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 — an issue that was later highlighted through studies of skin oils and riboflavin-derivative antigens (discussed below).

Skin oils.

In 1989, T cells with autoreactivity to CD1a and CD1c proteins were discovered²². 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 $\alpha\beta$ T cell line (BC2) to isolate CD1a-binding antigens produced in human cells and tissues⁴⁵. CD1a-binding autoantigens preferentially accumulate in the skin as compared with other tissues⁴⁵. This finding is consistent with studies showing that CD1a-autoreactive T cells home to the skin and that skin expresses more CD1a than other organs⁷. Extraction of whole cells or lipid-CD1 complexes with chloroform, followed by mass spectrometry, identified the CD1a-presented autoantigens as extremely hydrophobic

skin oils — **wax esters, squalene** and triacylglycerides⁴⁵ (Figure 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 fit⁴⁸. For CD1a, CD1c and CD1d, the observed volumes of the antigen-binding clefts (1,280-1,780 Å³) 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 to 86 carbons, so individual antigens can diverge from the known volumes of CD1 grooves⁴⁹ (Figure 2). So, how can CD1 proteins bind lipids with such varied chain lengths?

Moreover, the cleft of CD1b (2,200 Å³) is nearly 50 percent larger than the cavities in other CD1 isoforms (Figure 3), but few correspondingly large self lipids (~76 carbon) 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 from CD1 proteins with smaller clefts⁵⁰. 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' tunnel⁴². Crystallographic 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 seated along with the added ligand^{51, 52}. Further, 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 sliding⁵¹. Along with similar studies of CD1d⁵³, 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 (Figure 3a). Mass spectrometry studies identified the endogenous chaperone lipids as diacylglycerides and deoxydihydroceramides (Figure 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 TCR⁵⁰.

Scaffold lipids are similar to **class II 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 when present in different cellular subcompartments⁴⁹. The loading of large ligands is expected to require the removal of the antigen and the scaffold lipid from the upper and lower chambers⁵¹. 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 (positions 80, 86). These effects relax the CD1b conformation, creating a larger portal for exogenous antigen entry and ligand exit from the cleft⁵⁵. By contrast, loading of small lipids is now thought to require emptying of the upper chamber only⁵⁰ and occurs efficiently at neutral pH, when the interdomain tethers are intact⁴². 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 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 CD1d⁵⁹, CD1c¹⁰ and CD1a⁶⁰ crystal structures show electron densities that correspond to the lipid antigen present only in the F' pocket with spacer lipids occupying the A' pocket (Figure 2). At the other end of the size spectrum, 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. CD1a and CD1c molecules have unnamed notches in the lateral wall of the F' pocket, which 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 β 2-microglobulin. For many years, MR1 molecules were known to mediate the activation of mouse and human **mucosal-associated invariant T (MAIT) cells** in response to certain microorganisms^{12, 63, 64}. 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 B9)⁶⁵. MR1-presented antigens are modified metabolites derived from the riboflavin (vitamin B2) pathway: 5-oxopropylideneamino-6-D-ribitylaminouracil (5-OP-RU) and 5-(2-oxoethylideneamino)-6-D-ribitylaminouracil (5-OE-RU) respectively⁶⁶. Ribityllumazines and ribitylpyrimidines lack aliphatic hydrocarbon chains (Figure 2), reinforcing the concept that T cells can respond to non-peptide, non-lipid, small molecules.

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^{65–68} (Figure 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 fraction of the riboflavin derivative is exposed for TCR recognition^{65, 69} (Figure 4).

Origin of MR1-presented antigens.

Riboflavin derivatives can be considered foreign antigens because they arise in certain bacteria and fungi. Yet, the most potent MAIT cell ligands derive from covalent conjugation to methylglyoxal or glyoxal and other host or microbial intermediates, so that the resulting product is a hybrid neo-antigen⁶⁶. A central question now is whether MAIT cells mainly recognize vitamin B derivatives or also see other types of antigen. The spectrum of bacterial and fungal organisms that activate MAIT cells^{12, 13, 63} corresponds well to the spectrum of organisms with riboflavin pathways⁶⁵. This observation, along with the detection of large numbers of MAIT cells by MR1 tetramers bound to one type of ribitylpyrimidine ligand (5-OP-RU), suggests that one known antigen broadly supports MAIT cell recognition⁶⁹. Furthermore, deletion of either of the two riboflavin biosynthetic genes (*ribA* and *ribG*) from *L. lactis*, or related genes in *Escherichia coli*, ablates T cell recognition, indicating that other antigens do not exist in these species^{66, 70}. Then again, crystal structures of 5-OP-RU–MR1 complexes show that the F' pocket of MR1 remains empty, so as yet undiscovered larger ligands might bind within MR1.

TCR recognition of peptides, lipids and metabolites

Ternary crystal structures comprised of TCRs bound to CD1⁷¹ or MR1 protein complexes⁷² can be compared with the 34 unique structures of TCR–MHC class I complexes⁷¹. TCR recognition of α -galactosylceramide (α GalCer)¹⁴, 5-OP-RU⁷² or an 11-mer peptide⁷³ 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 (Figure 4). However, recent structures of $\alpha\beta$, $\gamma\delta$ and $\delta/\alpha\beta$ TCRs have revealed clear and fundamental differences between CD1 and MHC antigen display. Studies detailed below describe unexpected modes of recognition that fall outside the co-recognition paradigm.

Left-right asymmetry of CD1 architecture.

A universal difference between MHC and CD1 proteins relates to the presence or absence of lateral symmetry in the antigen display platform (Figure 5). In MHC proteins, the TCR contact region can be imagined as a platform that is bisected by a plane that is perpendicular to the axis of the groove. The groove is accessible on both sides of the plane (Figure 5, **green**). By contrast, the four human CD1 molecules show a fundamental left-right asymmetry (Figure 5, **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 $\alpha 1$ and $\alpha 2$ helices to form the A' roof^{42, 61, 62, 74}. 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 (Figure 5, **green**). Thus, the CD1 display platform is dominated by the outer surface of the CD1 protein on the left side and the 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 (Figure 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 (Figure 5b).

Influence of asymmetry on antigen display.

With some notable exceptions^{75, 76}, many TCR footprints on MHC class I molecules are located near the centre of the platform⁷¹. The central location and extensive exposure of peptides on MHC, along with the relatively large size of TCR footprints (1,200-2,400 Å²), favours substantial contact of the TCR with exposed peptide⁷¹. By contrast, the asymmetric location 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 (Figures 4 and 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. 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, 77-79}. 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 center with some near the extreme left of right side of the platform, creating modes of antigen recognition that are not known from the MHC system⁷¹.

NKT cell TCR footprints on CD1d.

In contrast to the 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 cleft¹⁴ (Figure 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 antigens⁸⁰. By contrast, a variable (type II) NKT cell TCR known as XV19 showed an orthogonal rotation and left of centre footprint. Because it is docked over the A' roof, the XV19 TCR makes most contact with CD1d, but it also makes some contact with the sulfate head group

near the centre of the platform. Overall, these glycolipid-CD1d-TCR structures illustrate the left-right shift and the differing rotation of $\alpha\beta$ TCRs docking on CD1d molecules. Although CD1c has not been co-crystallized with TCRs, alanine scanning mutagenesis suggests that individual $\alpha\beta$ TCRs use different left- or right-sided footprints as well⁸¹.

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, 82} (Figure 2). Moreover, CD1a-autoreactive T cell clones lack the fine antigen specificity that is seen for recognition of glycolipids and peptides^{79, 83} and, instead cross-react with several hydrophobic molecules, including methyl fatty acids, triacylglycerides and wax esters⁴⁵. 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 – a mode of activation we have termed absence of interference.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^{84, 85}. 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 barrier and not within APCs. Accordingly, the absence of interference model predicts that Langerhans cells normally express a cohort 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 TCRs⁷. 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 interference theory has recently been tested through the study of ternary interactions between lipid, CD1a and TCR. Tetramers usually require loading with homogenous antigen so that multiple arms of the tetramer carry the same antigen, creating the avidity needed to bind an antigen-specific T cell⁸⁶. Therefore, it was surprising that a CD1a-autoreactive TCR known as BK6 can bind CD1a tetramers loaded with diverse self-lipids⁵⁸. Using the BK6 TCR to pull down bound lipid-CD1a complexes, lipidomic profiling detected hundreds of distinct lipids released from the lipid-CD1a-TCR complexes, which were therefore 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 pattern of 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 CD1a-lysophosphatidylcholine (LPC) complexes (Figure 6a, **red**). LPC does not fully nest within the cleft, but its head group leans rightward as it traverses the F' portal, exposing a large area on the A' roof to which the TCR can bind (Figure 6a–b). Importantly, the BK6 TCR contacts with the A' roof but not the LPC ligand. Thus, CD1a is the antigenic target, and LPC is a permissive ligand that permits TCR to contact CD1a without interference.

A second structure of the BK6 TCR binding to CD1a carrying diverse endogenous ligands (endog-CD1a) also illustrates absence of interference (Figure 6b). The endog-CD1a-BK6 TCR structure shows electron density in the F' pocket that corresponds to that of a fatty acid (Figure 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 CD1a⁵⁸. 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 sphingomyelin⁵⁸, sulfatide⁶² or the mycobactin-like lipopeptide dideoxymycobactin⁶⁰ (Figure 6c). In CD1a, the A' roof is formed in part by salt bridges between the residues Arg73, Arg76 and Glu154, which tether the $\alpha 1$ and $\alpha 2$ 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, the broader CD1a-reactive repertoire likely uses both mechanisms. Recognition of dideoxymycobactin-CD1a probably occurs via co-recognition (Figure 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, 87, 88}. For other known CD1a antigens, such as sphingolipids and phospholipids^{89–91}, the mechanism of recognition has not been solved structurally, but co-recognition is favoured on the basis of its large head group.

CD1 molecules bind $\gamma\delta$ TCRs

There are many subsets of $\gamma\delta$ T cells⁹². In humans, one subset expresses TRGV9 and TRDV2 and recognizes soluble antigens, such as alkyl phosphates and alkyl amines^{93–95},

probably indirectly through allosteric modification of the cell surface molecule butyrophilin 3A1^{96–98}. Other $\gamma\delta$ T cell subsets target surface molecules, such as endothelial protein C receptor⁹⁹, T10 and MHC class I polypeptide-related sequence A (MICA), directly without associated antigen⁹⁷.

The first report of CD1-mediated antigen presentation described a response by both $\alpha\beta$ and $\gamma\delta$ T cells²². With a few exceptions^{100, 101}, CD1 research focused on $\alpha\beta$ T cells, but recent tetramer¹⁰² and crystallographic^{15, 16, 103} studies have confirmed the $\gamma\delta$ TCR-CD1d interaction. The $\gamma\delta$ TCR 9C2 (TRGV5, TRDV1) recognizes α GalCer-CD1d complexes (Figure 4), whereas the $\gamma\delta$ TCR DP10.7 (TRGV4, TRDV1) binds sulfatide-CD1d complexes. Both structures show a left-sided TCR footprint, orthogonal rotation and reliance on binding of tryptophan residues from CDR1 δ to CD1d. One difference is that the TCR 9C2 uses the CDR3 γ loop to bind the protruding α GalCer head group, whereas the TCR DP10.7 uses the hypervariable residues in CDR3 δ for all interactions with sulfatide.

Several V δ genes can also be used by $\alpha\beta$ T cells, including TRDV4-TRDV8, and share V α -related names¹⁰⁴. In addition, it has been long known that TRDV1 and TRDV3 genes can join J α -C α genes^{105–107}, but it was unclear how these V δ gene products contribute to the overall specificity of the “hybrid” TCR. A recent study¹⁷ of such a hybrid TCR, composed of a TRDV domain fused to the J α -C α domain and paired with a TCR β -chain, showed that it binds α GalCer-CD1d using an orthogonal and left-aligned docking mode. Mirroring the $\gamma\delta$ TCR 9C2, the TRDV1 CDR1 δ loop bound mainly to CD1d, whereas the CDR3 β loop bound the galactosyl head group of α GalCer.

Role of CD1 ligands.

Although most work on $\gamma\delta$ T cells emphasizes direct TCR contact with the monomorphic surfaces of cell surface molecules⁹², crystal structures involving CD1d-lipid complexes suggest the possibility that $\gamma\delta$ TCRs recognize antigen that is physically displayed. This is supported by the observed contact of the CDR3 δ and CDR3 γ loops with sulfatide and α GalCer, 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 $\gamma\delta$ TCR and the CD1 molecule itself, rather than the bound ligand. Specifically, the $\gamma\delta$ 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 $\gamma\delta$ 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, 108–110}. The left-right shift hypothesis outlined for $\alpha\beta$ TCRs (Figure 5) might be relevant to $\gamma\delta$ TCRs as well (Figure 4, **right panel**). 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 CD1a³⁵, CD1b⁹, CD1c¹⁰ and MR1^{66, 69} 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 or MR1 proteins are almost identical among humans, so a single CD1 or MR1 tetramer reagent can be used for any human donor and quantitatively track antigen-specific T cells in disease states. Both invariant NKT cells^{111, 112} and MAIT cells¹¹³ were discovered by detecting T cells expressing similar $\alpha\beta$ TCRs. Only later were they found to recognize CD1d¹¹⁴ and MR1, respectively⁶³. Now, with the generation of 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 allows a 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^{67, 69}, as well as to the discovery of two previously unknown T cell types that recognize CD1b^{115, 116}.

The CD1b-specific TCR repertoire.

Both MAIT cells and NKT cells expand in high numbers and 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, 117, 118}. 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 with TRAV1-2 joined to TRAJ9 paired with an apparently biased population of TRBV6-2 chains¹¹⁶. This TCR pattern was seen among unrelated donors, indicating interdonor conservation. Such TCRs showed high affinity for GMM-CD1b complexes^{115, 116}. Based on the TCR structure and antigen specificity, such CD1b-reactive T cells were called germline-encoded mycolyl-specific (GEM) T cells¹¹⁶. 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 subsets¹¹⁶ (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 CD1b-GMM¹¹⁵. Among these T cell clones derived from patients with tuberculosis, TRB4-1 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 leprosy⁸³. 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 ligands⁵⁰. 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⁺ TCRs⁸¹ (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 donors¹¹⁹. 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 barrier, 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, 120}. These cells secrete cytokines that play central roles in host defense and tissue repair^{7, 115, 116, 120–122}. Lipid antigens such as α GalCer are bioavailable, and NKT cells can be activated and polarized in effector function by glycolipid antigens or altered glycolipid ligands^{123,124}. Although therapeutic outcomes have been limited α GalCer 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^{125,126–128}. 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 been recently discovered, so they have not entered trials, but they might now likewise be tested in humans to stimulate, block or detect T cell responses. Antigens that control human CD1- 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.

Acknowledgements

The authors thank R. Birkinshaw, S. Huang, T-Y Cheng, and K. Wucherpfennig for advice and graphical images. This work is supported by the Bill and Melinda Gates Foundation, NIAID (AI049313, AI111224, U19 111224), National Health and Medical Research Council 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).

GLOSSARY

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 MHC class I, CD1, MR1 proteins or other antigen-presenting molecules. Antigen-loaded tetramers bind antigen-specific T cell receptors with high avidity so that antigen-specific T cells can be directly counted or isolated by flow cytometry

Mucosal-associated invariant T cells

(MAIT cells). T cells that express a structurally conserved invariant T cell receptor and are selected by the MHC class I-related molecule MR1

Class II 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

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

Squalene

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

Wax esters

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

Spacer lipids

Hydrophobic compounds that bind together with and fill up part of the CD1 cleft that is not occupied by antigenic lipids

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

Secretory pathway

A series of protein transport reactions by which newly folded proteins transit from the endoplasmic reticulum to the Golgi apparatus and to the 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, which 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

References

1. Garboczi DN et al. Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature* 384, 134–141 (1996). [PubMed: 8906788]
2. Garcia KC et al. An alphabeta T cell receptor structure at 2.5 Å and its orientation in the TCR-MHC complex. *Science* 274, 209–219 (1996). [PubMed: 8824178]
3. Zinkernagel RM 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, 421–36 (1997). [PubMed: 9393624]
4. Salerno-Goncalves R, Fernandez-Vina M, Lewinsohn DM & Szein MB 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, 5852–62 (2004). [PubMed: 15494539]
5. Heinzl AS et al. HLA-E-dependent presentation of Mtb-derived antigen to human CD8+ T cells. *J Exp Med* 196, 1473–81 (2002). [PubMed: 12461082]
6. Adams EJ & Luoma AM The adaptable major histocompatibility complex (MHC) fold: structure and function of nonclassical and MHC class I-like molecules. *Annu Rev Immunol* 31, 529–61 (2013). [PubMed: 23298204]
7. de Jong A et al. CD1a-autoreactive T cells are a normal component of the human alphabeta T cell repertoire. *Nature immunology* 11, 1102–9 (2010). [PubMed: 21037579]
8. de Lalla C et al. High-frequency and adaptive-like dynamics of human CD1 self-reactive T cells. *European journal of immunology* 41, 602–10 (2011). [PubMed: 21246542]
9. Kasmar AG et al. CD1b tetramers bind {alpha}{beta} T cell receptors to identify a mycobacterial glycolipid-reactive T cell repertoire in humans. *The Journal of experimental medicine* (2011).
10. Ly D et al. CD1c tetramers detect ex vivo T cell responses to processed phosphomycoketide antigens. *The Journal of experimental medicine* 210, 729–41 (2013). [PubMed: 23530121]
11. Dellabona P et al. In vivo persistence of expanded clones specific for bacterial antigens within the human T cell receptor alpha/beta CD4–8- subset. *J.Exp.Med* 177, 1763–1771 (1993). [PubMed: 7684433]
12. Gold MC et al. Human mucosal associated invariant T cells detect bacterially infected cells. *PLoS Biology* 8, e1000407 (2010). [PubMed: 20613858] * Starting with a broad analysis of all human CD8+ T cells, this paper demonstrated 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, 701–8 (2010). [PubMed: 20581831]

14. Borg NA et al. CD1d-lipid-antigen recognition by the semi-invariant NKT T-cell receptor. *Nature* 448, 44–9 (2007). [PubMed: 17581592]
15. Luoma AM et al. Crystal structure of Vdelta1 T cell receptor in complex with CD1d-sulfatide shows MHC-like recognition of a self-lipid by human gammadelta T cells. *Immunity* 39, 1032–42 (2013). [PubMed: 24239091] * These papers reported the structures of Vδ1⁺ γδTCRs bound to lipid-CD1d complexes, providing detailed molecular mechanisms of CD1 recognition.
16. Uldrich AP et al. CD1d-lipid antigen recognition by the gammadelta TCR. *Nat Immunol* 14, 1137–45 (2013). [PubMed: 24076636] * These papers reported the structures of Vδ1⁺ γδTCRs bound to lipid-CD1d complexes, providing detailed molecular mechanisms of CD1 recognition.
17. Pellicci DG et al. The molecular bases of delta/alphabeta T cell-mediated antigen recognition. *J Exp Med* (2014).
18. McMichael AJ et al. A human thymocyte antigen defined by a hybrid myeloma monoclonal antibody. *Eur.J.Immunol* 9, 205–210 (1979). [PubMed: 376318]
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). [PubMed: 3093894]
20. Calabi F, Jarvis JM, Martin L & Milstein C Two classes of CD1 genes. *Eur.J.Immunol* 19, 285–292 (1989). [PubMed: 2467814]
21. Dougan SK, Kaser A & Blumberg RS CD1 expression on antigen-presenting cells. *Curr Top Microbiol Immunol* 314, 113–41 (2007). [PubMed: 17593659]
22. Porcelli S et al. Recognition of cluster of differentiation 1 antigens by human CD4-CD8- cytolytic T lymphocytes. *Nature* 341, 447–450 (1989). [PubMed: 2477705] ** This seminal paper demonstrated 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 γδTCR 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, 1758–66 (2005). [PubMed: 16034117]
24. Baker ML & Miller RD Evolution of mammalian CD1: marsupial CD1 is not orthologous to the eutherian isoforms and is a pseudogene in the opossum *Monodelphis domestica*. *Immunology* 121, 113–21 (2007). [PubMed: 17244156]
25. Dossa RG, Alperin DC, Hines MT & Hines SA The equine CD1 gene family is the largest and most diverse yet identified. *Immunogenetics* 66, 33–42 (2014). [PubMed: 24196432]
26. Hayes SM & Knight KL Group 1 CD1 genes in rabbit. *J Immunol* 166, 403–10 (2001). [PubMed: 11123318]
27. Van Rhijn I et al. The bovine CD1 family contains group 1 CD1 proteins, but no functional CD1d. *J Immunol* 176, 4888–93 (2006). [PubMed: 16585584]
28. Eguchi-Ogawa T et al. Analysis of the genomic structure of the porcine CD1 gene cluster. *Genomics* 89, 248–61 (2007). [PubMed: 17112699]
29. Looringh van Beeck FA et al. Two canine CD1a proteins are differentially expressed in skin. *Immunogenetics* 60, 315–24 (2008). [PubMed: 18488214]
30. Kasmar A, Van Rhijn I & Moody DB The evolved functions of CD1 during infection. *Curr Opin Immunol* 21, 397–403 (2009). [PubMed: 19541469]
31. de la Salle H et al. Assistance of microbial glycolipid antigen processing by CD1e. *Science*. 310, 1321–1324 (2005). [PubMed: 16311334]
32. Brigl M & Brenner MB CD1: antigen presentation and T cell function. *Annu.Rev.Immunol* 22:817-90, 817–890 (2004). [PubMed: 15032598]
33. Yakimchuk K et al. *Borrelia burgdorferi* infection regulates CD1 expression in human cells and tissues via IL1-beta. *European journal of immunology* 41, 694–705 (2011). [PubMed: 21246541]
34. Adams EJ Lipid presentation by human CD1 molecules and the diverse T cell populations that respond to them. *Curr Opin Immunol* 26, 1–6 (2014). [PubMed: 24556395]
35. Kasmar AG et al. Cutting Edge: CD1a tetramers and dextramers identify human lipopeptide-specific T cells ex vivo. *J Immunol* 191, 4499–503 (2013). [PubMed: 24089190]

36. Felio K et al. CD1-restricted adaptive immune responses to Mycobacteria in human group 1 CD1 transgenic mice. *The Journal of experimental medicine* 206, 2497–509 (2009). [PubMed: 19808251]
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). [PubMed: 12077262]
38. Dascher CC et al. Conservation of a CD1 multigene family in the guinea pig. *J.Immunol* 163, 5478–5488 (1999). [PubMed: 10553074]
39. Porcelli S, Morita CT & Brenner MB CD1b restricts the response of human CD4 – 8 - T lymphocytes to a microbial antigen. *Nature* 360, 593–597 (1992). [PubMed: 1281285]
40. Beckman EM et al. Recognition of a lipid antigen by CD1-restricted T cells. *Nature* 372, 691–694 (1994). [PubMed: 7527500]
41. Layre E, de Jong A & Moody DB Human T cells use CD1 and MR1 to recognize lipids and small molecules. *Curr Opin Chem Biol* 23c, 31–38 (2014).
42. Gadola SD et al. Structure of human CD1b with bound ligands at 2.3 Å, a maze for alkyl chains. *Nature Immunol.* 3, 721–726 (2002). [PubMed: 12118248] ** This CD1b structure work provided 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 burgdorferi* antigens by NKT cells. *Proc Natl Acad Sci U S A* 107, 1535–40 (2010). [PubMed: 20080535]
44. Van Rhijn I et al. CD1d-restricted T cell activation by nonlipidic small molecules. *Proc Natl Acad Sci U S A* 101, 13578–83 (2004). [PubMed: 15342907]
45. de Jong A et al. CD1a-autoreactive T cells recognize natural skin oils that function as headless antigens. *Nat Immunol* 15, 177–85 (2014). [PubMed: 24362891] ** This paper identified 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 approach.
46. Prigozy TI et al. Glycolipid antigen processing for presentation by CD1d molecules. *Science* 291, 664–667 (2001). [PubMed: 11158680]
47. Mattner J et al. Exogenous and endogenous glycolipid antigens activate NKT cells during microbial infections. *Nature* 434, 525–529 (2005). [PubMed: 15791258]
48. Cheng TY et al. Role of lipid trimming and CD1 groove size in cellular antigen presentation. *Embo J* 25, 2989–99 (2006). [PubMed: 16794581]
49. Moody DB et al. Lipid length controls antigen entry into endosomal and nonendosomal pathways for CD1b presentation. *Nature Immunol.* 3, 435–442 (2002). [PubMed: 11938350]
50. Huang S et al. Discovery of deoxyceramides and diacylglycerols as CD1b scaffold lipids among diverse groove-blocking lipids of the human CD1 system. *Proceedings of the National Academy of Sciences of the United States of America* 108, 19335–40 (2011). [PubMed: 22087000]
51. Garcia-Alles LF et al. Structural reorganization of the antigen-binding groove of human CD1b for presentation of mycobacterial sulfoglycolipids. *Proc Natl Acad Sci U S A* 108, 17755–60 (2011). [PubMed: 22006319] * This paper showed how CD1b ligands of different sizes can slide upwards towards the top of the groove, and it identified spacer lipids in the CD1b system.
52. Garcia-Alles LF et al. Endogenous phosphatidylcholine and a long spacer ligand stabilize the lipid-binding groove of CD1b. *Embo J* 25, 3684–92 (2006). [PubMed: 16874306]
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, 1131–44 (2007). [PubMed: 17485514]
54. Jackman RM et al. The tyrosine-containing cytoplasmic tail of CD1b is essential for its efficient presentation of bacterial lipid antigens. *Immunity* 8, 341–51 (1998). [PubMed: 9529151]
55. Relloso M et al. pH-dependent interdomain tethers of CD1b regulate its antigen capture. *Immunity* 28, 774–86 (2008). [PubMed: 18538591]
56. Im JS et al. Kinetics and cellular site of glycolipid loading control the outcome of natural killer T cell activation. *Immunity* 30, 888–98 (2009). [PubMed: 19538930]

57. Im JS et al. Direct measurement of antigen binding properties of CD1 proteins using fluorescent lipid probes. *J.Biol.Chem* 279, 299–310 (2004). [PubMed: 14551186]
58. Birkinshaw RW et al. alphabeta T cell antigen receptor recognition of CD1a presenting self lipid ligands. *Nat Immunol* 16, 258–66 (2015). [PubMed: 25642819]
59. Wun KS et al. A molecular basis for the exquisite CD1d-restricted antigen specificity and functional responses of natural killer T cells. *Immunity* 34, 327–39 (2011). [PubMed: 21376639]
60. Zajonc DM et al. Molecular mechanism of lipopeptide presentation by CD1a. *Immunity* 22, 209–219 (2005). [PubMed: 15723809]
61. Scharf L et al. The 2.5 Å structure of CD1c in complex with a mycobacterial lipid reveals an open groove ideally suited for diverse antigen presentation. *Immunity* 33, 853–62 (2010). [PubMed: 21167756] * This study analysed a crystal structure of CD1c bound to mannosyl-phosphomycoketide, revealing an large F' portal, a mechanism for capture of branched chain lipids and two accessory portals.
62. Zajonc DM, Elsliger MA, Teyton L & Wilson IA Crystal structure of CD1a in complex with a sulfatide self antigen at a resolution of 2.15 Å. *Nat.Immunol* 4, 808–815 (2003). [PubMed: 12833155]
63. Treiner E et al. Selection of evolutionarily conserved mucosal-associated invariant T cells by MR1. *Nature* 422, 164–9 (2003). [PubMed: 12634786]
64. Le Bourhis L et al. Antimicrobial activity of mucosal-associated invariant T cells. *Nature immunology* 11, 701–8 (2010). [PubMed: 20581831]
65. Kjer-Nielsen L et al. MR1 presents microbial vitamin B metabolites to MAIT cells. *Nature* 491, 717–23 (2012). [PubMed: 23051753] ** This paper provided several advances in the understanding of MAIT biology, including 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 identification of an unexpected mechanism of covalent attachment in generating stable metabolite-MR1 complexes at the surface of cells.
66. Corbett AJ et al. T-cell activation by transitory neo-antigens derived from distinct microbial pathways. *Nature* 509, 361–5 (2014). [PubMed: 24695216]
67. Eckle SB et al. A molecular basis underpinning the T cell receptor heterogeneity of mucosal-associated invariant T cells. *J Exp Med* 211, 1585–600 (2014). [PubMed: 25049336]
68. McWilliam HE, Birkinshaw RW, Villadangos JA, McCluskey J & Rossjohn J MR1 presentation of vitamin B-based metabolite ligands. *Curr Opin Immunol* 34C, 28–34 (2015).
69. Reantragoon R et al. Antigen-loaded MR1 tetramers define T cell receptor heterogeneity in mucosal-associated invariant T cells. *J Exp Med* 210, 2305–20 (2013). [PubMed: 24101382]
70. 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, 4641–9 (2015). [PubMed: 25870247]
71. Rossjohn J et al. T Cell Antigen Receptor Recognition of Antigen-Presenting Molecules. *Annu Rev Immunol* (2014).
72. Patel O et al. Recognition of vitamin B metabolites by mucosal-associated invariant T cells. *Nat Commun* 4, 2142 (2013). [PubMed: 23846752]
73. Tynan FE et al. A T cell receptor flattens a bulged antigenic peptide presented by a major histocompatibility complex class I molecule. *Nat Immunol* 8, 268–76 (2007). [PubMed: 17259989]
74. Zeng Z et al. Crystal structure of mouse CD1: an MHC-like fold with a large hydrophobic binding groove. *Science* 277, 339–345 (1997). [PubMed: 9219685]
75. Hahn M, Nicholson MJ, Pyrdol J & Wucherpfennig KW Unconventional topology of self peptide-major histocompatibility complex binding by a human autoimmune T cell receptor. *Nat Immunol* 6, 490–6 (2005). [PubMed: 15821740]
76. Liu YC et al. A molecular basis for the interplay between T cells, viral mutants, and human leukocyte antigen micropolymorphism. *J Biol Chem* 289, 16688–98 (2014). [PubMed: 24759101]
77. Brossay L et al. Mouse CD1-autoreactive T cells have diverse patterns of reactivity to CD1+ targets. *Journal of Immunology* 160, 3681–3688 (1998).

78. Mallevaey T et al. A molecular basis for NKT cell recognition of CD1d-self-antigen. *Immunity* 34, 315–26 (2011). [PubMed: 21376640]
79. Kawano T et al. CD1d-restricted and TCR-mediated activation of V α 14 NKT cells by glycosylceramides. *Science* 278, 1626–1629 (1997). [PubMed: 9374463]
80. Rossjohn J, Pellicci DG, Patel O, Gapin L & Godfrey DI Recognition of CD1d-restricted antigens by natural killer T cells. *Nature reviews. Immunology* 12, 845–57 (2012).
81. Roy S et al. Molecular basis of mycobacterial lipid antigen presentation by CD1c and its recognition by alphabeta T cells. *Proc Natl Acad Sci U S A* 111, E4648–57 (2014). [PubMed: 25298532]
82. Garboczi DN et al. Assembly, specific binding, and crystallization of a human TCR-alphabeta 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). [PubMed: 8955188]
83. Moody DB et al. Structural requirements for glycolipid antigen recognition by CD1b-restricted T cells. *Science* 278, 283–6 (1997). [PubMed: 9323206]
84. Manolova V et al. Functional CD1a is stabilized by exogenous lipids. *Eur. J. Immunol* (2006).
85. Sugita M et al. Separate pathways for antigen presentation by CD1 molecules. *Immunity*. 11, 743–752 (1999). [PubMed: 10626896]
86. Altman JD et al. Phenotypic analysis of antigen-specific T lymphocytes. *Science* 274, 94–96 (1996). [PubMed: 8810254]
87. Moody DB et al. T cell activation by lipopeptide antigens. *Science* 303, 527–31 (2004). [PubMed: 14739458]
88. Young DC et al. Synthesis of dideoxymycobactin antigens presented by CD1a reveals T cell fine specificity for natural lipopeptide structures. *J Biol Chem* 284, 25087–96 (2009). [PubMed: 19605355]
89. Shamshiev A et al. Presentation of the same glycolipid by different CD1 molecules. *J. Exp. Med* 195, 1013–1021 (2002). [PubMed: 11956292]
90. Agea E et al. Human CD1-restricted T cell recognition of lipids from pollens. *J Exp Med* 202, 295–308 (2005). [PubMed: 16009719]
91. Bourgeois EA et al. Bee venom processes human skin lipids for presentation by CD1a. *J Exp Med* 212, 149–63 (2015). [PubMed: 25584012]
92. Chien YH, Meyer C & Bonneville M gammadelta T cells: first line of defense and beyond. *Annu Rev Immunol* 32, 121–55 (2014). [PubMed: 24387714]
93. Tanaka Y, Morita CT, Nieves E, Brenner MB & Bloom BR Natural and synthetic non-peptide antigens recognized by human gamma delta T cells. *Nature* 375, 155–158 (1995). [PubMed: 7753173]
94. Bukowski JF, Morita CT & Brenner MB Human gamma delta T cells recognize alkylamines derived from microbes, edible plants, and tea: implications for innate immunity. *Immunity* 11, 57–65 (1999). [PubMed: 10435579]
95. Tanaka Y et al. Nonpeptide ligands for human gamma delta T cells. *Proc. Natl. Acad. Sci. U.S.A* 91, 8175–8179 (1994). [PubMed: 8058775]
96. Vavassori S et al. Butyrophilin 3A1 binds phosphorylated antigens and stimulates human gammadelta T cells. *Nat Immunol* 14, 908–16 (2013). [PubMed: 23872678]
97. Adams EJ, Gu S & Luoma AM Human gamma delta T cells: Evolution and ligand recognition. *Cell Immunol* (2015).
98. Sandstrom A et al. The intracellular B30.2 domain of butyrophilin 3A1 binds phosphoantigens to mediate activation of human Vgamma9Vdelta2 T cells. *Immunity* 40, 490–500 (2014). [PubMed: 24703779]
99. Willcox CR et al. Cytomegalovirus and tumor stress surveillance by binding of a human gammadelta T cell antigen receptor to endothelial protein C receptor. *Nat Immunol* 13, 872–9 (2012). [PubMed: 22885985]
100. Leslie DS et al. CD1-mediated gamma/delta T cell maturation of dendritic cells. *J. Exp. Med* 196, 1575–1584 (2002). [PubMed: 12486100]

101. Spada FM, Koezuka Y & Porcelli SA CD1d-restricted recognition of synthetic glycolipid antigens by human natural killer T cells. *J.Exp.Med* 188, 1529–1534 (1998). [PubMed: 9782130]
102. Bai L et al. The majority of CD1d-sulfatide-specific T cells in human blood use a semiinvariant Vdelta1 TCR. *Eur J Immunol* 42, 2505–10 (2012). [PubMed: 22829134]
103. Pellicci DG et al. The molecular bases of delta/alphabeta T cell-mediated antigen recognition. *J Exp Med* 211, 2599–615 (2014). [PubMed: 25452463]
104. Lefranc MP & Rabbitts TH Genetic organization of the human T-cell receptor gamma and delta loci. *Res Immunol* 141, 565–77 (1990). [PubMed: 1965674]
105. Miossec C et al. Further analysis of the T cell receptor gamma/delta+ peripheral lymphocyte subset. The V delta 1 gene segment is expressed with either C alpha or C delta. *J.Exp.Med* 171, 1171–1188 (1990). [PubMed: 2182762]
106. Miossec C et al. Molecular characterization of human T cell receptor alpha chains including a V delta 1-encoded variable segment. *Eur J Immunol* 21, 1061–4 (1991). [PubMed: 1826885]
107. Peyrat MA et al. Repertoire analysis of human peripheral blood lymphocytes using a human V delta 3 region-specific monoclonal antibody. Characterization of dual T cell receptor (TCR) delta-chain expressors and alpha beta T cells expressing V delta 3J alpha C alpha-encoded TCR chains. *J Immunol* 155, 3060–7 (1995). [PubMed: 7673722]
108. Agea E et al. Human CD1-restricted T cell recognition of lipids from pollens. *J.Exp.Med* 202, 295–308 (2005). [PubMed: 16009719]
109. Dieude M et al. Cardiolipin binds to CD1d and stimulates CD1d-restricted gammadelta T cells in the normal murine repertoire. *Journal of immunology* 186, 4771–81 (2011).
110. Mangan BA et al. Cutting edge: CD1d restriction and th1/th2/th17 cytokine secretion by human vdelta3 T cells. *J Immunol* 191, 30–4 (2013). [PubMed: 23740951]
111. Fowlkes BJ et al. A novel population of T-cell receptor alpha beta-bearing thymocytes which predominantly expresses a single V beta gene family. *Nature* 329, 251–254 (1987). [PubMed: 3114646]
112. Budd RC et al. Developmentally regulated expression of T cell receptor beta chain variable domains in immature thymocytes. *J.Exp.Med* 166, 577–582 (1987). [PubMed: 3496420]
113. Tilloy F et al. An invariant T cell receptor alpha chain defines a novel TAP- independent major histocompatibility complex class Ib-restricted alpha/beta T cell subpopulation in mammals. *J.Exp.Med* 189, 1907–1921 (1999). [PubMed: 10377186]
114. Bendelac A et al. CD1 recognition by mouse NK1+ T lymphocytes. *Science* 268, 863–865 (1995). [PubMed: 7538697]
115. Van Rhijn I et al. TCR bias and affinity define two compartments of the CD1b-glycolipid-specific T Cell repertoire. *J Immunol* 192, 4054–60 (2014). [PubMed: 24683194]
116. Van Rhijn I et al. A conserved human T cell population targets mycobacterial antigens presented by CD1b. *Nature immunology* 14, 706–13 (2013). [PubMed: 23727893] * Expanding the field of invariant T cells beyond NKT cells and MAIT cells, this paper discovered human germline-encoded mycolyl-reactive (GEM) T cells specific for CD1b.
117. Grant EP et al. Molecular recognition of lipid antigens by T cell receptors. *J.Exp.Med* 189, 195–205 (1999). [PubMed: 9874576]
118. Cohen NR, Garg S & Brenner MB Antigen Presentation by CD1 Lipids, T Cells, and NKT Cells in Microbial Immunity. *Adv Immunol* 102, 1–94 (2009). [PubMed: 19477319]
119. van Schaik B et al. Discovery of invariant T cells by next-generation sequencing of the human TCR alpha-chain repertoire. *J Immunol* 193, 5338–44 (2014). [PubMed: 25339678]
120. Le Bourhis L, Mburu YK & Lantz O MAIT cells, surveyors of a new class of antigen: development and functions. *Current opinion in immunology* 25, 174–80 (2013). [PubMed: 23422835]
121. Fernandez CS et al. MAIT cells are depleted early but retain functional cytokine expression in HIV infection. *Immunol Cell Biol* 93, 177–88 (2015). [PubMed: 25348935]
122. Gold MC et al. Human thymic MR1-restricted MAIT cells are innate pathogen-reactive effectors that adapt following thymic egress. *Mucosal immunology* 6, 35–44 (2013). [PubMed: 22692454]

123. Miyamoto K, Miyake S & Yamamura T A synthetic glycolipid prevents autoimmune encephalomyelitis by inducing TH2 bias of natural killer T cells. *Nature* 413, 531–534 (2001). [PubMed: 11586362]
124. Kim EY, Lynch L, Brennan PJ, Cohen NR & Brenner MB The transcriptional programs of iNKT cells. *Semin Immunol* 27, 26–32 (2015). [PubMed: 25841627]
125. Fujii S et al. NKT cells as an ideal anti-tumor immunotherapeutic. *Front Immunol* 4, 409 (2013). [PubMed: 24348476]
126. Tefit JN et al. Efficacy of ABX196, a new NKT agonist, in prophylactic human vaccination. *Vaccine* 32, 6138–45 (2014). [PubMed: 25218293]
127. Motohashi S et al. A phase I-II study of alpha-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, 2492–501 (2009). [PubMed: 19201905]
128. Schneiders FL et al. Clinical experience with alpha-galactosylceramide (KRN7000) in patients with advanced cancer and chronic hepatitis B/C infection. *Clin Immunol* 140, 130–41 (2011). [PubMed: 21169066]
129. 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 U S A* 102, 8668–73 (2005). [PubMed: 15939887]
130. Miller MM et al. Characterization of two avian MHC-like genes reveals an ancient origin of the CD1 family. *Proc Natl Acad Sci U S A* 102, 8674–9 (2005). [PubMed: 15939884]
131. Nguyen TKA et al. Expression patterns of bovine CD1 in vivo and assessment of the specificities of the anti-bovine CD1 antibodies. *PLoS One* doi:10.1371/journal.pone.0121923 (2015).
132. Dascher CC et al. Conservation of a CD1 multigene family in the guinea pig. *J Immunol* 163, 5478–88 (1999). [PubMed: 10553074]
133. Loorigh van Beeck FA 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, 1424–31 (2009). [PubMed: 19185921]

Figure 1 |. Mammalian CD1 genes.

The tree illustrates the origins of modern species after the key event of the bird-mammal split. The discovery of avian CD1 orthologues indicates the existence of ancestral CD1 genes prior to the bird-mammal split, and CD1 genes are universally or widely conserved in mammals. The number of CD1 genes in rabbits²⁶, mice, guinea pigs, primates, dogs²⁹, cattle¹³¹, pigs²⁸, horses²⁵ and chickens^{129, 130} differ. For guinea pigs, we discovered one CD1 gene in the updated genome, in addition to the published genes^{132, 133}, which we named after the closest human CD1 orthologue, which was CD1a. All three group 1 CD1 proteins are absent in mice, creating a need for additional experimental models to study CD1a, CD1b and CD1c.

In contrast to the short peptide ligands of MHC class I proteins, 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 (TCRs). 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 group⁴⁵. The MHC class I-related

protein (MR1)-presented antigen 5-oxopropylideneamino-6-D-ribitylaminouracil and the CD1d-presented antigen phenyl pentamethyldihydrobenzofuran 5-sulfonate⁴⁴ are small molecules that are neither peptides nor lipids⁶⁶.

Figure 3 | Scaffold and spacer lipids bind to CD1b.

a | A co-crystal study of CD1b bound to antigen (shown in yellow and red), a synthetic sulfoglycolipid analogue known as SGL12 with a combined lipid length of 33 carbons. This structure also detects a second electron density (shown in orange) positioned in the lower chamber of the CD1b groove, within the T' tunnel⁵¹. **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.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 4 | T cell receptor (TCRs) that function through co-recognition.

This figure shows examples of $\alpha\beta$ TCRs^{14, 72, 73} and a $\gamma\delta$ TCR¹⁶ 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 (α GalCer); the mucosal-associated invariant T (MAIT) cell TCR interacts with MHC class I-related protein (MR1) displaying 5-oxopropylideneamino-6-D-ribitylaminouracil; the ELS4 TCR interacts with the MHC class I molecule HLA-B35 displaying an Epstein-Barr virus (EBV)-derived peptide; and the 9C2 $\gamma\delta$ TCR interacts with CD1d displaying α GalCer. 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 mechanisms of TCR co-recognition.

Figure 6 | Direct recognition of CD1a molecules carrying permissive ligands.

a | The T cell receptor (TCR) BK6 forms a left-sided footprint (lower panel, red) 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. Other non-permissive ligands disrupt a triad of amino acids (R76, E154 and R73) within the A' roof⁵⁸.

Table 1 |

Human invariant T cell types that recognize non-polymorphic antigen-presenting molecules

	CD1b	CD1b	CD1c	CD1d	MR1	MHC
Designation	GEM T cells	LDN5-like cells		NKT cells	MAIT cells	
Antigen	Glucose monomycolate	Glucose monomycolate	Phospho-mycoketide	αGalCer & others	vitamin B metabolites	Peptides
TCR α chain Variable gene	Invariant TRAV1-2	Biased TRAV17	Diverse Variable	Invariant TRAV10	Invariant TRAV1-2	Diverse All
Joining gene CDR3 length	TRAJ9 Uniform	Uniform		TRAJ18 Uniform	TRAJ33 Uniform	All Variable
TCR β chain Variable gene CDR3 length	Biased TRBV6-2 Variable	Biased TRBV4-1 Variable	Biased TRBV7-9 TRBV7-8 Variable	Biased TRBV25-1 Variable	Biased TRBV20-1 TRBV6-1 Variable	Diverse All Variable
Intradonor conservation	High	High	High	High	High	Low
Interdonor conservation	High	High	High	High	High	Low