Cholesteryl esters stabilise human CD1c conformations for recognition by self-reactive T cells

Salah Mansour[a,b,1], Anna S Tocheva[a], Chris Cave-Ayland[c], Moritz M Machelett[b,d], Barbara Sander[c], Nikolai M Lissin[e], Peter E Molloy[e], Mark S Baird[f], Gunthard Stübs[g], Nicolas WJ Schröder[h], Ralf R Schumann[i], Jörg Rademann[j], Anthony D Postle[a], Bent K Jakobsen[e], Ben G Marshall[a,b], Rajendra Gosain[c], Paul Elkington[a,b], Tim Elliott[b,k], Chris-Kriton Skylaris[b,c], Jonathan W Essex[b,c], Ivo Tews[b,d], and Stephan D Gadola[a,b,l,1]

[a]*Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK*

[b]*Institute for Life Sciences, University of Southampton, Southampton, UK*

[c]*School of Chemistry, University of Southampton, Southampton, UK*

[d]*Centre for Biological Sciences, University of Southampton, Southampton, UK*

[e]*Immunocore Ltd, Abingdon, UK*

[f]*School of Chemistry, Bangor University, Bangor, Wales, UK*

[g]*Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany*

[h]*Institute for Pathology, Otto-von-Guericke University Magdeburg, Magdeburg, Germany*

[i]*Institute for Microbiology and Hygiene, Charité University Medical Center, Berlin, Germany*

[j]*Division of Medicinal Chemistry, Institute of Pharmacy, Freie Universität Berlin, Berlin, Germany*

[k]*Cancer Sciences Unit, Faculty of Medicine, University of Southampton, Southampton, UK*

[l]*Novartis Institutes of Biomedical Research, Basel, Switzerland*

[1] To whom correspondence should be addressed: stephan.gadola@Novartis.com; (or s.mansour@soton.ac.uk)

Key words

CD1 group 1, lipid antigen presentation, self-reactive T cells, cholesteryl esters

Abstract

CD1c-dependent self-reactive T cells are abundant in human blood, but self-antigens presented by CD1c to the T cell receptors of these cells are unknown. Here we present a crystal structure of CD1c determined at 2.4 Å revealing an extended ligand binding potential of the antigen groove and a substantially different conformation compared to known CD1c structures. Computational simulations exploring different occupancy states of the groove re-enacted these different CD1c conformations, and suggested cholesteryl esters (CE) and acylated steryl glycosides (ASG) as new ligand classes for CD1c. Confirming this, we show that binding of CE and ASG to CD1c enables the binding of human CD1c self-reactive T cell receptors. Hence, human CD1c adopts different conformations dependent on ligand occupancy of its groove, with CE and ASG stabilising CD1c conformations that provide a footprint for binding of CD1c self-reactive T cell receptors.

Significance statement

T cells autoreactive to the antigen presenting molecule CD1c are abundant in human blood but lipid antigens recognised by these T cells remained unknown. A new ligand-filled 2.4 Å structure of CD1c and computational simulations thereof indicated substantial conformational plasticity of CD1, revealing new structural features, as well as the potential of CD1c to present acylated-sterols to T cells. We confirmed these predictions by demonstrating loading onto CD1c and direct biophysical interaction of CD1c-Lipid complexes with self-reactive human T cell receptors for two lipid classes: Cholesteryl-esters similar to those that accumulate in foamy macrophages (e.g. in atherosclerosis) and acylated steryl-glycosides from *B.burgdorferi*. These findings differentiate CD1c from other CD1 family members and indicate new roles of CD1 in human immunity.

\body

Introduction

Cluster of differentiation 1 (CD1) proteins are a family of major histocompatibility complex (MHC) class I-like glycoproteins that present lipid antigens to T cells. CD1 restricted T cells are abundant in humans and play important roles in host defence and immune regulation. Human CD1 proteins comprise five CD1 isoforms, CD1a, CD1b, CD1c, CD1d, and Cd1e, which exhibit different intracellular trafficking behaviours and ligand binding preferences (1). Structurally, the main differences between these CD1 isoforms lie in the architecture of their lipophilic ligand binding grooves. While all CD1 isoforms share a highly conserved A' channel (or pocket) for binding C18-C26 acyl chains, specialisation is provided by further connecting channels (2-7). In CD1a, the A' channel is "fused" to a wide and shallow F' channel enabling binding of lipopeptides such as mycobacterial

didehydroxymycobactin (DDM) (8). CD1b features a unique T' tunnel that connects A' and F', thereby forming a "superchannel" for accommodating very long acyl chains, e.g. mycobacterial mycolates (2, 4). CD1d, the only isoform also conserved in rodents, exhibits a two-branched ligand binding groove with two linear channels A' and F' connected near the main portal into the groove, known as the F' portal. A similar two-branched arrangement of A' and F' is seen in CD1e, the only CD1 isoform not expressed on the cell surface. Compared to CD1d, CD1a and CD1b, the portal into the groove in CD1e is widely exposed, consistent with its known role in lipid transfer processes inside lysosomes (6).

CD1c presents foreign- (9, 10) as well as self- lipid antigens to T cells (11). Two recent crystal structures of human CD1c revealed a similar two-branched design as CD1d and CD1e, with two channels A' and F' connecting near the groove portal. In these structures, a mycobacterial phosphomycoketide (PM) or mannosyl-β1-phosphomycoketide (MPM) occupied the A' channel, while an undefined short spacer ligand was present in the F' channel (7, 12). The spatial arrangement of these ligands in the CD1c groove was very similar and virtually overlapping in 3D comparisons with that of alpha-galactosylceramide (αGC) in human CD1d (Fig S1*A* and S1*B*). As CD1c and CD1d are known to traffic to the same intracellular compartments for antigen sampling (13), these CD1c-PM and CD1c-MPM structures did not readily explain how CD1c and CD1d could functionally differentiate. Furthermore, the F' channel in both CD1c-PM and CD1c-MPM was widely open to solvent, which was strikingly different to known structures of CD1a, CD1b and CD1d, and reminiscent of CD1e (7, 12). Based on these facts we hypothesized that human CD1c might undergo substantial conformational transformations in the F' channel region upon binding of more optimal ligands, with relevance for T cell receptor binding.

Results

Structural features of human CD1c with extended-ligand binding and closure of the F' channel roof

Based on our experience with human CD1b and CD1d (2, 3), we first attempted to refold the extracellular α1-α3 domains of human CD1c ("wild type CD1c", $CD1c^{wt}$) together with β2m, using detergents as surrogate ligands. Although $CD1c^{wt}$/β2m complexes of correct size and stoichiometric composition could be purified, these complexes rapidly disintegrated after buffer exchange at 4°C. This suggested an unstable quaternary conformation of the complex with loss of the non-covalent association between $CD1c^{wt}$ and β2m. Thus, we decided to employ a similar strategy to Scharf *et al*, by grafting the α3 domain of human CD1b onto the antigen binding α1-α2 domain of human CD1c ($CD1c^{bα3}$) (7). Using $CD1c^{bα3}$ instead of $CD1c^{wt}$, with otherwise identical refolding conditions as

before, resulted in stable soluble CD1c$^{b\alpha3}$/β2m complexes that ultimately yielded diffracting protein crystals. The structure was determined by molecular replacement and refined with the use of data to 2.4 Å (Table 1).

The structure shows the typical MHC class I-like domain arrangement that is common to all known CD1/β2m complexes (Fig. 1*A*). Clear electron density is seen for three hydrocarbon chain spacer ligands (SL) filling the A' and F' channels (Fig. 1*B*). We refer to this new ligand occupied state as CD1c-SL. Similar to other known CD1/ligand structures, including CD1c-PM and CD1c-MPM, a single C18 stearic acid ligand saturates the A' channel of CD1c-SL and protrudes slightly into the F' portal (Fig. 1*A*-*C*). In striking difference to previous CD1/ligand structures, inc

luding CD1c-PM and CD1c-MPM, the F' channel of CD1c-SL is occupied by two C12 spacer ligands, likely corresponding to lauric acids present in the refolding mixture, that are stacked in a parallel manner (Fig. 1*A*-*C*). These differences in ligand saturation between the previously solved CD1c structures and CD1c-SL are paralleled by radical differences in the conformation of the F' channel. In contrast to either CD1c-PM or CD1c-MPM, the F' channel in CD1c-SL is shielded from solvent by an F' channel roof (F' roof). Tyr155 and Leu150 on α2 as well as Glu83 and His87 on α1 provide tethering interactions between the two alpha-helices which result in the formation of this F' roof above the F' channel (Fig. 2*A* and Fig. S2*A*). Consequently, a new portal into the side of CD1c, which we propose to call the G' portal, can be clearly defined in CD1c-SL (Fig. 2*A* and Fig. S2*A*). Furthermore, the E' portal (7), which is an open conduit from the F' channel to the exterior of the protein described in CD1c-PM and CD1c-MPM, is virtually absent in CD1c-SL (Fig. S2*B*).

Molecular dynamics simulations of CD1c-SL

The above differences between CD1c-SL and the previously determined CD1c-PM and CD1c-MPM structures (7, 12) indicated that ligand occupancy of the F' channel is a major determinant of human CD1c protein conformation. To further address this hypothesis we performed molecular dynamics (MD) simulations in the presence and absence of CD1c ligands at neutral pH. As expected, MD simulations performed on CD1c-SL with all observed bound ligands were fully consistent with a stable CD1c protein complex (Fig. 3*A*, and video S1). In contrast, removal of all bound ligands from CD1c-SL led to a rapid inward motion of both alpha-helices with progression to a complete collapse of both A' and F' channels (Fig. 3*A* and 3*B*, video S2). Next, we examined the behaviour of the complex with more hydrophilic ligands. For this we chose to exchange the two aliphatic lauric acids within F' with polyethylene glycol (PEG) molecules that were present in the crystallization buffer. In

these MD simulations, the PEG molecules rapidly evacuated the F' channel (video S3**)**. In consequence, the antigen binding domain went through a rapid succession of changes from the initial closed F' roof conformation to a transiently open conformation (Fig. 3*C*-*E*) before collapse of the F' channel. Notably, the MD simulations with PEG ligands closely re-enacted the observed conformational differences between CD1c-SL and CD1c-MPM in the region of the F' channel roof. Rapid disengagement of α1-α2 roof-tethering interactions upon F' channel ligand evacuation greatly increased the flexibility of these residues, at times leading to the adoption of configurations highly similar to CD1c-PM and CD1c-MPM (Fig. 3*C* and 3*D*). Therefore, human CD1c can adopt both open and closed F' roof conformations depending on F' channel ligand occupancy.

The CD1c$^{b\alpha3}$/β2m complex used for CD1c-SL is recognised by human CD1c self-reactive T cells

In light of the major conformational differences between CD1c-SL and either CD1c-PM or CD1c-MPM, we aimed to confirm the functional validity of the soluble CD1c$^{b\alpha3}$/β2m complexes that were used for crystallisation and structure determination of CD1c-SL. Fluorescent-conjugated CD1c$^{b\alpha3}$/β2m tetramers produced from the same protein batch that was used for crystallisation (CD1c-SL tetramers) were employed to generate CD1c-SL tetramer-positive TCR αβ+ T cell lines and clones from human blood by FACS sorting (Fig. 4*A*). While these T cells were brightly stained with CD1c-SL tetramers, they failed to bind either CD1b- or CD1d-tetramers (Fig. 4*A*). Conversely, CD1c-SL tetramers failed to bind to CD1d-restricted human invariant Natural Killer T cells (iNKT) (Fig. S3*A*). In cellular assays, these T cells exhibited strong CD1c-dependent cytokine secretion in the absence of added exogenous ligands (Fig. 4*B*). To further examine the specific binding of CD1c-SL tetramers to CD1c-restricted TCRs, we generated a human Jurkat T cell line with stable expression of both TRAV22 and TRBV6.2 TCR chains from a CD1c self-reactive T cell clone (clone NM4) (Fig. S3*B* and Table S1). These Jurkat-NM4 cells brightly stained with CD1c-SL tetramers, whereas CD1c-SL tetramers failed to stain CD8-1 Jurkat cells expressing the mycoketide specific CD1c-restricted CD8-1 TCR (14), or other Jurkat cell lines expressing CD1a-, CD1b-, and CD1d-restricted TCRs (Fig. 4*C*; Fig. S4*A* and S4*B*). These results were highly consistent with a physiologically relevant and functionally differentiated state of CD1c-SL, and thus they suggested that the 3D conformation exhibited by CD1c-SL represents a valid model to interrogate the ligand binding potential of the F' channel of human CD1c.

Ligand binding potential of human CD1c

To start exploring the potential spectrum of CD1c bound ligands, we carried out molecular docking simulations employing the antigen binding cavity of CD1c-SL as the template (Fig. 5*A* and Fig. 5*B*).

In initial experiments, steroids such as cholesterol, and cholesterol-like detergents such as CHAPS showed favourable docking poses within the F' channel of CD1c-SL, suggestive of their potential as CD1c ligands. Following from these results, we envisaged cholesteryl esters (CE) and acylated steryl glycosides (ASG) as possible groove stabilising ligands for CD1c as they might simultaneously engage both CD1c's A' and F' channels via binding of their fatty acid and cholesteryl moieties, respectively. Indeed, molecular docking simulations produced favourable docking poses for different natural CE and ASG, thereby supporting the notion that human CD1c could present these lipids to T cells (Fig. 5*B*). To address this hypothesis, we refolded CD1c proteins with acylated cholesteryl ß-**D**-galactoside (ACGal), an ASG of the human pathogen *Borrelia burgdorferi sensu lato* which is soluble in aqueous buffers (15, 16). Both $CD1c^{b\alpha3}$ and $CD1c^{wt}$ proteins could be successfully refolded in the presence of ACGal, enabling the generation of CD1c-ACGal tetramers which specifically stained Jurkat T cells expressing the CD1c self-reactive NM4-TCR (Fig. 6*A*). To further characterise the molecular interaction between CD1c and NM4-TCR, we produced soluble recombinant NM4-TCR (Fig. S5) and measured its binding to both CD1c-SL and $CD1c^{wt}$-ACGal in Surface Plasmon Resonance (SPR) (Fig. 6*B*). Consistent with the results obtained with Jurkat-NM4 T cells, soluble recombinant NM4-TCR exhibited similar binding at equilibrium to CD1c-SL (Kd 6.05 ± 0.4 μM) and CD1c-ACGal (Kd 6.13 ± 0.4 μM) (Fig. 6*B*).

Cholesteryl esters (CE) and acylated steryl glycosides (ASG) are CD1c stabilising ligands

The above results indicated that CE and ASG could provide CD1c with conformational stability, thereby enabling recognition by CD1c autoreactive TCRs. To address this hypothesis we devised a cell-free assay for direct testing of the ligand-dependent binding of NM4-TCR to CD1c. We first produced soluble fluorescent NM4-TCR tetramers and confirmed their CD1c-dependent binding to CD1c-expressing T2 lymphoblasts (T2-CD1c) (Fig. 7*A*). Binding of NM4-TCR TCR tetramers to T2-CD1c lymphoblasts was effectively blocked by an anti-CD1c blocking antibody (Fig. 7*A*). Consistent with the autoreactive nature of the interaction, binding of NM4-TCR tetramers to T2-CD1c cells did not require the addition of antigens to T2-CD1c cultures. Next, we established a bead-based assay to measure the binding of NM4-TCR tetramers to CD1c after *in situ* ligand exchange. First, CD1c-SL-coated MACSi beads (CD1c-beads), but not unconjugated MACSi beads stained brightly with NM4-TCR tetramers (Fig. 7*B*). We then tested several ligand-stripping procedures, and found that Triton X-100 was most effective in abolishing NM4-TCR tetramer binding to the CD1c-beads (Fig. S6*A*). Finally, we assessed different lipids for their ability to reconstitute NM4-TCR tetramer binding to ligand-stripped CD1c-beads. Vehicle (DMSO), mycolic acid or gangliosides did not reconstitute NM4-TCR tetramer binding (Fig. S6*B*). In contrast, all tested acylated steryl ligands, including ACGal, the CE 5-

cholestene 3-palmitate and the plant sterol acylated β-sitosteryl glucoside (ASGlu) clearly reconstituted binding of NM4-TCR tetramers to ligand-stripped CD1c-beads (Fig. 7*C*).

Discussion

CD1c-dependent self-reactive T cells are abundant in the blood of healthy neonates and adults (17, 18), but the endogenous lipid antigens that are presented by CD1c to these T cells have remained unknown. Guided by the new CD1c-SL structure presented here, we now find that CD1c can bind CE and ASG, and that both these ligand classes enable the binding of self-reactive T cell receptors to CD1c. Two previous CD1c structures, CD1c-PM and CD1c-MPM, had revealed how CD1c binds and presents methylated monoalkyl chain ligands such as mycobacterial mycoketides (7, 12). In both CD1c-PM and CD1c-MPM a single mycoketide molecule was bound to the A' channel, in analogy to the arrangement seen for the stearic acid in CD1c-SL. Together CD1c-SL, CD1c-PM and CD1c-MPM thus illustrate a certain promiscuity of the A' channel, which is the most conserved region of the CD1 groove for ligand binding.

CD1c-PM and CD1c-MPM complexes are exclusively recognized by mycoketide-specific human T cells, but not by CD1c self-reactive T cells (12). The specificity of this interaction was shown to be determined by subtle structural features of the mycoketide (12, 19). Conversely, it had remained unclear whether and how ligands bound to the F' channel of CD1c could influence T cell receptor binding. In CD1c-PM and CD1c-MPM, a small undefined spacer ligand was observed in the F' channel. In striking difference to other known antigen presenting CD1 proteins, the F' channel in these structures showed a widely open conformation, thus potentially exposing its contents to solvent (12). In stark contrast, two lauric acid molecules, stacked on top of each other, fill the F' channel in the new CD1c structure presented here, CD1c-SL, where tethering interactions between the two alpha-helices provide for a roof over the F' channel, thereby shielding the bound ligands from solvent. Based on its position above the F' channel, this new roof structure in CD1c-SL is called the F' roof. Below the F' roof, a well-defined large portal, here proposed to be called the G' portal, provides an open conduit from the exterior into the F' channel of CD1c-SL. The clear conformational differences between CD1c-SL and the previous CD1c structures suggested that the F' channel of CD1c behaves similarly to a venus fly trap, dramatically adapting its conformation with closure of the roof over the channel in response to a sufficient degree of ligand occupancy. The results of our MD simulations of CD1c-SL supported this model. In these MD simulations, ligand evacuation from the F' channel rapidly induced a seamless transition from a closed F' roof conformation as seen in CD1c-SL,

to an open F' groove conformation, as seen in CD1c-PM or CD1c-MPM. Together, these findings identify an F' roof and G' portal in CD1c-SL as new structural features of the CD1 protein family.

Prompted by the observed extended-ligand binding in CD1c-SL, computational docking simulations with CD1c-SL as a template indicated the possibility that CD1c may present two new ligand classes, CE and ASG, to T cells. Indeed, the docking simulations suggested that both CE and ASG could engage the A' and F' channel of CD1c simultaneously via their acyl chains and steryl moieties, respectively. Using several complementary experimental approaches we confirmed that both CE and ASG can be presented by CD1c to human T cells. First, we generated recombinant CD1c-ASG complexes by employing a synthetic ASG from *B. burgdorferi*, ACGal (7, 15), which could be used in aqueous refolding conditions. We used these CD1c-ASG complexes in tetramer-based flow cytometry experiments as well as in SPR aided measurements of CD1c-TCR binding affinity, showing their specific binding to a CD1c self-reactive T cell receptor, NM4-TCR. To extend these findings to other ASG and to CE, and to assess whether ASG and CE support CD1c recognition by the CD1c self-reactive NM4-TCR, we devised a novel cell-free, MACSi bead-aided, flow cytometry-based assay. Starting from CD1c-SL coated MACSi beads we found that NM4-TCR tetramer binding to these beads was lost upon washing with the non-ionic surfactant Triton-X, but could be reconstituted after pulsing of the same beads with either ACGal or CE. These findings demonstrate that ASG and CE are ligands for CD1c, and, together with the above discussed structural findings in CD1c-SL and CD1c-PM/MPM, they suggest a model for CD1c recognition by self-reactive TCRs. In this model, endogenous ligands, such as CE that fill both the A' and F' channels of CD1c, induce a closed F' roof conformation of CD1c providing a stable footprint for self-reactive TCR binding.

Although the galactosyl headgroup in ACGal had no measurable influence on CD1c self-reactive NM4-TCR binding in our studies, it is possible that other TCRs can differentiate between either ACGal or CE, between different ASG, or even between different chemically modified CE.

CD1c may indeed be involved in immune regulation by different bacterial steryl glycosides, including ACGal from *B. burgdorferi*, the causative agent of Lyme disease (16), and α-cholesteryl acylated glucosides (α-ACGlu) from *H. pylori* (20), the causative agent of gastric ulcers. In docking simulations using CD1c-SL as template, α-ACGlu showed similar favourable poses as ACGal and CE (Fig. 5*B*). In support of a possible role of CD1c in Lyme disease, infection of human skin with *B. burgdorferi* was previously shown to strongly upregulate CD1c expression on myeloid dendritic cells (21).

Several lines of evidence suggest a role for CD1c and CD1c self-reactive T cells in human auto-inflammatory and autoimmune diseases. For example, CD1c self-reactive T cells were found to be elevated in autoimmune thyroid tissues and in systemic lupus erythematosus (SLE) (22, 23). In SLE, these cells induce immunoglobulin (Ig) class switching to IgG and increase Ig secretion in CD1c+ B cells (23). Conversely, CD1c^{high} myeloid dendritic cells (CD1c+mDC) infiltrate inflamed tissues in different autoimmune conditions, including e.g. rheumatoid arthritis (RA), vitiligo or autoimmune thyroiditis (22, 24, 25). In RA, these CD1c+mDC induce proliferation and cytokine secretion of autologous CD4+ T cells (26). Furthermore, CD1c is strongly induced in foam cell macrophages (FCM) which are characterised by their strong intracellular accumulation of CE (27, 28). FCM are typically seen in the inflammatory lesions of atherosclerosis, but are also present in other chronic inflammatory and infectious conditions, including e.g. tuberculosis (29). As CD1c+ FCM have full antigen presenting capabilities it is intriguing to speculate that they promote tissue inflammation in atherosclerosis or other chronic inflammatory conditions via CD1c mediated presentation of CE to self-reactive T cells. Known mechanisms which could induce CE accumulation in CD1c+ macrophages or dendritic cells in such conditions include the induction of Acyl-CoA:Cholesterol AcylTransferase (ACAT-1) mediated CE synthesis via toll-like receptor stimulation, or the increased cellular CE uptake via CD36 that can be induced by RA plasma (30, 31).

In conclusion, ASG and CE stabilise human CD1c proteins for the specific interaction with T cell receptors from human CD1c-reactive T cells, with possible roles in infection and inflammation. The extended ligand binding potential of CD1c, revealed by the new structure presented here, CD1c-SL, and the identification of ASG and CE as new ligand classes for CD1c, complement our understanding of how the five human non-polymorphic CD1 isoforms differentiate in their function as lipid binding and T cell regulating proteins.

Materials and Methods

CD1 cloning and recombinant proteins

CD1c constructs: Two CD1c constructs were generated for these studies: 1) Wild-type CD1c ($CD1c^{wt}$), encoding the extracellular α1-α3 domains of human CD1c; and 2) a CD1c/CD1b hybrid construct ($CD1c^{b\alpha3}$) encoding α1 and α2 domains of human CD1c and the α3 domain of human CD1b. Primers for $CD1c^{wt}$: forward, 5'–ATGGGCAACGCGGATGCGTCCCAG-'3; reverse, 5'-AGCTTAATGCCATTCGATTTTCTGAGCTT-'3. $CD1c^{b\alpha3}$ was generated by constructing a wild type CD1c (containing residues 18-296) splice variant by introducing an ApaL1 site at V200 (GTA to GTG) by site directed mutagenesis and subsequently cloned CD1b α3 into this construct using the primers: CD1b α3 forward, 5'–TTCTGTGCACTAATCATACAATATCAAGG-'3; CD1b α3 reverse, 5'-CACCGGATCCCGGACCCCAGTAGAGGATGATGTCC-'3.

CD1 protein production and in vitro refolding: Plasmids encoding the extracellular domains of human CD1b, human CD1d (3), $CD1c^{wt}$, $CD1c^{b\alpha3}$ and human β2m were separately cloned into the prokaryotic expression vector pET23d (Novagen), and recombinant proteins were generated separately as inclusion bodies in *E. coli* Rosetta strain (Novagen). Inclusion bodies were thoroughly washed and fully denatured and reduced in 6 M guanidine-HCL and 20 mM DTT prior to *in vitro* refolding. Refolding of $CD1c^{wt}$/β2m, $CD1c^{b\alpha3}$/β2m, human CD1b/β2m and human CD1d/β2m was carried out by oxidative *in vitro* refolding as previously described(32, 33), in the presence of the following detergents and lipids: Sorbitan stearate (SPAN60), CHAPS hydrate (both Sigma), α-galactosylceramide KRN7000 (Avanti Polar Lipids). Glucose monomycolate (GMM) (34) and acylated cholesteryl ß-**D**-galactoside (ACGal) were synthesized as previously described (15). Correctly folded proteins were purified by repeated FPLC (Pharmacia) size-exclusion chromatography using preparatory grade SD75 26/60 and analytical grade SD75 GL 10/300 gel filtration columns (GE Healthcare).

CD1 tetramers: Refolded CD1 complexes (CD1b, CD1c and CD1d) were biotinylated via an engineered BirA motif at the C-terminus, repurified by size exclusion chromatography, and used to generate fluorescent-labelled CD1 tetramers (33) by conjugating them to phycoerythrin (PE)-streptavidin (Sigma).

Soluble TCR and TCR tetramers: TCR heterodimers were generated as described previously (32, 35, 36). Briefly, the extracellular region of each TCR chain was produced as inclusion bodies from *E. coli* Rosetta following cloning into the bacterial expression vector pGMT7. To produce stable, disulfide-linked heterodimers, cysteines were incorporated into the TCR α- and β-chain constant regions, replacing residues Thr48 and Ser57 respectively. Expression, refolding and purification of the disulfide-linked NM4-TCR αβ heterodimers were carried out as previously described (37). Purified refolded TCR proteins was assessed by both reducing and non-reducing SDS-PAGE analysis (Fig. S5). NM4-TCR tetramers were produced using modified TCRβ chains, containing a C-terminus BirA-tag motif, which was specifically biotinylated. Subsequently, biotinylated TCRs were purified by size exclusion chromatography before conjugation to PE-streptavidin (Sigma) to generate fluorescently labelled TCR tetramers.

CD1c-Beads: Biotinylated CD1c-SL complexes were conjugated to anti-biotin MACSiBeads (Miltenyi). Stripping of bead conjugated CD1c was achieved by incubating beads with PBS containing 1-5% Triton X-100 detergent (Acros Organics) for 2 hours at 37°C. Beads were then washed thoroughly with PBS and then pulsed overnight with lipid antigens dissolved in DMSO (Sigma). Lipids include cholesteryl ester (CE) 5-cholestene 3-palmitate (Sigma) and the plant sterol acylated β-sitosteryl glucoside (ASGlu) (Matreya). Beads were then extensively washed with PBS before staining with NM4-TCR tetramer and acquired on a FACSCalibur (Becton Dickinson).

Protein crystallography

Proteins (in 20mM Tris-HCL pH 7.5 and 50mM NaCl) at 8 mg/ml concentration were crystallised using sitting-drop vapor diffusion in 96 well plates at 20°C, with (0.2M Magnesium chloride, 0.1M Tris pH 8, 10% PEG 8000) as precipitant (1:1 protein to precipitant ratio), using an ARI Gryphon nanodrop dispenser (Art Robbins Instruments). Crystals were harvested in mother liquor containing 20% glycerol as cryoprotectant and flash frozen in liquid nitrogen. Crystals were tested at the Southampton Diffraction Centre and data collected at the Diamond Light Source beamline I04 (DLS, Oxford) at the cryogenic temperature of 100°C and a wavelength of 0.9795 Å. Data reduction, molecular replacement and refinement were carried out with CCP4 (38). The structure was solved by molecular replacement with CD1c-MPM (PDB 3OV6) and iteratively built with Coot (39), using automated water structure building in ArpWarp(40) and refinement in Refmac5, resulting in good Ramachandran statistics (96.3% favored, 3.7% allowed, no outliers). Residues of the F' channel roof

are highly ordered in the structure, with His87 being modelled in double conformation. Pymol (41) was used to create the figures shown in this publication.

Molecular docking

Three acylated steryl ligands, including two ASG (ACGal, α-ACGlu) and 1 CE (3-O-acylated cholesterol) were selected (Fig. 5*B*). Starting from SMILES string representations of the ligands, protonation states were assigned and low energy 3D conformations were generated with CORINA (42) Ligands were docked into the CD1c-SL crystal structure, using GOLD 5.0 (43). The binding cavity was defined as follows: based on the ligands presented in the binding pocket of CD1c-SL, all protein residues with heavy atoms within a radius of 6 Å were selected and the atom numbers were saved in a cavity file using the graphical visualization and analysis program Hermes from the GOLD suite. This cavity file was defined in the GOLD configuration file as the area into which the ligands are docked. Additionally, the "do_cavity" option in GOLD was set to 1 to allow the restriction of the binding site to concave, solvent accessible surfaces. Ligands were treated as *flexible* while the protein structure was kept *rigid*. For the genetic algorithm (GA), automatic settings were used. The search efficiency was set to 1 using the autoscale flag to allow GOLD to apply optimal settings for each ligand. Predicted poses were assessed using the ChemPLP, an empirical scoring function optimised for binding pose prediction. ChemPLP has been shown to be the most effective scoring function implemented in GOLD for both pose prediction and virtual screening (44). For each ligand, 25 independent docking runs were conducted with ten GA runs each, resulting in a total of 250 poses per ligand and snapshot, keeping the top-scoring pose for analysis to ensure sufficient sampling and reproducibility of the docking runs.

Molecular Dynamics simulations

To prepare the CD1c-SL 3D structure, hydrogen atoms were added using the PROTANATE3D module of the MOE software package (45), and all water molecules were retained. The structure was further solvated with the TIP3P water model in a box with a minimum distance from protein of 8 Å. Sodium cations were added to neutralise the overall charge of the system, giving a total system size of 50,000 atoms. Three different simulation setups were created: (I) the deposited crystallographic model with stearic acid bound in the A' channel, and two lauric acid molecules bound to the F' channel; (II) CD1c with all ligand molecules removed; (III) CD1c with stearic acid in the A' channel,

and three further polyethylene glycol short chain moieties. Of these one was located within the A' channel and two more within the F' channels.

Simulations were carried out using the AMBER 12 package (46, 47), with the FF99SB forcefield. Parameters for ligand molecules were provided by the GAFF forcefield (48) and applied using the ANTECHAMBER module. Partial atomic charges were assigned using the AM1-BCC method (49). The equilibration protocol included a series of successive minimisations, gradually releasing restraints on the heavy atoms of the system. Heavy atoms of the protein were then restrained with a force constant of 1000 kcal.mol^{-1}.Å^{-2} and the system was gradually heated to 300 K over 200 ps, and equilibrated to 1 atm pressure for 200 ps under the isothermal-isobaric ensemble. The system was then cooled and the procedure repeated with protein restraints removed. Production runs were carried out under constant volume and temperature dynamics. Temperature control was achieved using the Langevin thermostat, with a collision constant of 3.0 ps-1. Pressure regulation employed the Berendsen barostat with a relaxation time of 2.0 ps. Dynamics were carried out using a 2 fs timestep for integration of the equations of motion and a cut off distance of 8 Å for non-bonded interactions. PME was used to correct for long-range electrostatics. The simulations were carried out on the Emerald GPU cluster, using the CUDA implementation of the PMEMD module (50). Three repeats of each simulation setup were carried out to give a total of 9 molecular dynamics trajectories, each 200 ns in length. Pocket volumes during MD trajectories were calculated using the softwaretool PocketAnalyzer (51)

DCs, T cells and T cell assays

DC Generation: Peripheral blood mononuclear cells (PBMC) were isolated from blood by density gradient centrifugation (Ficoll-Hypaque; GE Healthcare). CD1c expressing monocyte-derived dendritic cells (mo-DC) were differentiated from CD14+ monocytes that were purified by positive selection using the CD14+ MACSibeads kit from Miltenyi Biotec and cultured in RPMI in the presence of 25ng/ml IL-4 and 10ng/ml GM-CSF (Immunotools) for 5 days. CD1c expression was confirmed by FACS analysis using APC-conjugated anti-CD1c antibody (clone AD5-8E7) (Miltenyi).

T Cell lines and clones: To generate CD1c restricted T cell lines, CD14- T cell fractions were cultured with mo-DC for 7 days before the addition of 400 IU/ml IL-2 (Proleukin, Chiron). PE-CD1c-SL tetramer+/ CD3+ lines and clones were sorted into 96-well round-bottom plates by FACSAria (BD Biosciences) and re-stimulated with 1 ug/ml phytohaemagglutinin (PHA) (Sigma) in the presence of 5

x 10^4 autologous γ-irradiated (35Gy) PBMCs. Cells were grown in T cell growth media (RPMI 1640, 2% human AB serum [Sigma], 10% fetal bovine serum [FBS], 0.1 mg/ml kanamycin, 1 mM sodium pyruvate, 1% non-essential amino acids, 1% L-glutamax, and 50 mM 2-mercaptoethanol [Sigma] and IL-2).

Cytokine assays: T cells, 1 x 10^5 per well, were stimulated for 24 hours in the presence of 1 x 10^5 T2 lymphoblasts or T2-CD1c lymphoblasts in 96 well plates. Culture supernatants were then analysed for cytokine concentrations for IFN-γ, TNF-α, IL-4, IL-10 using a Bio-Plex 200 system (Bio-Rad laboratories).

Transgenic TCR expressing cell lines: TCR alpha and beta chains from the following T cell clones were cloned into the third generation pELNS lentivector (kindly provided by University of Pennsylvania): Clones CD8-2 (CD1a/DDM specific), LDN5 (CD1b/GMM specific) CD8-1 (CD1c/Phosphomycoketide specific) were kindly provided by Dr Branch Moody; iNKT clone 1369 (32); and NM4 T cell clone (CD1c restricted; Table S1). CD1c was also cloned into pELNS lentivector. Primers used for cloning TCR alpha and beta chains into pELNS lentivector were: FL-CD8-1-α, 5'-GCGCGCTAGCCGCCACCATGCTCCTGCTGCTCGTCCC-'3; FL-LDN5-α, 5'-GCGCGCTAGCCGCCACCATGGAAACTCTCCTGGGAGTGTC-'3; FL-CD8-2-α, 5'-GCGCGCTAGCCGCCACCATGGCCTCTGCACCCATCT-'3; FL-NM4-α, 5'-GCGCGCTAGCCGCCACCATGAAGAGGATATTGGGAGC-'3;FL-α-REV, 5'-GCGCGAGATCTGCTTCTCTTGGCCCGAGAGCCGCTGGACCACAGCCGCAGCGT-'3, FL-CD8-1-β, 5'-GCGCCCTAGGATGCTGCTGCTTCTGCTGCT-'3; FL-LDN5-β, 5'-GCGCCCTAGGATGGGCTGCAGGCTGCTCTG-'3; FL-CD8-2-β, 5'-GCGCCCTAGGATGGGCTGCAGGCTCCTCTG-'3; FL-NM4-β, 5'-GCGCCCTAGGATGAGCCTCGGGCTCCTGTGCTGTGGGGCC-'3; FL-β-REV, 5'-TATGGATCCGGAGCTAGCCTCTGGAATCCTTTC-'3. Primers used to clone CD1c were: CD1cFLFwd, 5'-GCGCCCTAGGCGCCACCATGCTGTTTCTGCAGTTTCTGCTGC-'3; CD1cFLRev, 5'-GCGCGCGTCGACTCACAGGATGTCCTGATATGAGC-'3.

Lentiviruses encoding different TCR alpha and beta chains were generated in HEK293TN cells after co-transfection with engineered pELNS lentivector (2.5 μg) and three accessory plasmids, pCMV-VSV-G (1.5 μg), pRSV.REV (3 μg) and pMDL.pg.RRE (3 μg)(52). Lentiviral particles were collected, filtered and used for direct transduction of T2 lymphoblasts and Jurkat T cell lines. Transduced cells were sorted by flow cytometry on a FACSAria and maintained in complete RPMI (10% FBS, 1% non-

essential amino acids, 1% L-Glutamax, 1mM sodium pyruvate, 100IU penicillin and 100ug/ml streptomycin).

Flow cytometry

The following fluorescent reagents were used: PE-conjugated tetramers (tet): PE-CD1c^{wt}-ACGal-tet, PE-CD1c$^{b\alpha3}$-SL or -ACGal-tet; PE-CD1b-GMM-tet; PE-CD1d-αGC-tet; PE-NM4-TCR-tet, FITC-conjugated anti-human CD4, and APC-conjugated anti-human CD3 (Immunoltools). PE-conjugated anti-human TCR alpha-beta (Miltenyi). After addition of staining reagents, cells were incubated at 4°C for 45 min, washed twice in ice-cold PBS/1%FBS, and acquired on a four-colour FACSCalibur flow cytometer (Becton Dickinson). Propidium iodide (Sigma) was used to exclude dead cells. Data were processed using CellQuest Pro software (BD Biosciences).

Surface plasmon resonance

Streptavidin (~5000RU) was amino-coupled to a Biacore CM-5 chip (BIAcore AB, UK) and 50 μg/mL biotinylated lipid-CD1c complexes or control proteins (CD1d-αGC, CD1b-GMM, and HLA-A2*01-NY-Eso-1(157–165) complex) were loaded on individual flow cells until the response measured ~1000RU. Recombinant TCRs were serially diluted and flowed over the protein-loaded flow cells at a rate of 5 or 50 μL/min for determination of equilibrium binding or kinetics. Responses were recorded in real time on a Biacore 3000 machine at 25°C, and data were analysed using the BIAevaluation software (Biacore) as described previously (32).

Acknowledgements

This work was supported by HEFCE (SDG, SM), BBSRC (AST), EPSRC (CCA; grant code EP/G03690X/1), the Wessex innovation grant fund (SDG, SM). The authors thank the following people for their invaluable support of this work: Stuart Findlow and Chris Holes at the Macromolecular Crystallisation Facility at the Centre for Biological Sciences, Pete Horton and Simon Coles at the Southampton Diffraction Centre for support, the staff at the Diamond Light Source for excellent user support, Richard Jewell and Carolann McGuire (FACS facility, Faculty of Medicine, University of Southampton). Special thanks to Branch D. Moody for donating CD1 group 1 restricted T cell lines

and clones (CD8-1, CD8-2, LDN5), the *chemical computing group* for a MOE licence, and the Emerald High Performance Computing facility.

References

1. Barral DC & Brenner MB (2007) CD1 antigen presentation: how it works. *Nature reviews. Immunology* 7(12):929-941.
2. Gadola SD*, et al.* (2002) Structure of human CD1b with bound ligands at 2.3 A, a maze for alkyl chains. *Nat Immunol* 3(8):721-726.
3. Koch M*, et al.* (2005) The crystal structure of human CD1d with and without alpha-galactosylceramide. *Nat Immunol* 6(8):819-826.
4. Batuwangala T*, et al.* (2004) The crystal structure of human CD1b with a bound bacterial glycolipid. *J Immunol* 172(4):2382-2388.
5. Zajonc DM, Elsliger MA, Teyton L, & Wilson IA (2003) Crystal structure of CD1a in complex with a sulfatide self antigen at a resolution of 2.15 A. *Nat Immunol* 4(8):808-815.
6. Garcia-Alles LF*, et al.* (2011) Crystal structure of human CD1e reveals a groove suited for lipid-exchange processes. *Proc Natl Acad Sci U S A* 108(32):13230-13235.
7. Scharf L*, et al.* (2010) The 2.5 A structure of CD1c in complex with a mycobacterial lipid reveals an open groove ideally suited for diverse antigen presentation. *Immunity* 33(6):853-862.
8. Zajonc DM*, et al.* (2005) Molecular mechanism of lipopeptide presentation by CD1a. *Immunity* 22(2):209-219.
9. Matsunaga I*, et al.* (2004) Mycobacterium tuberculosis pks12 produces a novel polyketide presented by CD1c to T cells. *J Exp Med* 200(12):1559-1569.
10. Moody DB*, et al.* (2000) CD1c-mediated T-cell recognition of isoprenoid glycolipids in Mycobacterium tuberculosis infection. *Nature* 404(6780):884-888.
11. Lepore M*, et al.* (2014) A novel self-lipid antigen targets human T cells against CD1c(+) leukemias. *J Exp Med* 211(7):1363-1377.
12. Roy S*, et al.* (2014) Molecular basis of mycobacterial lipid antigen presentation by CD1c and its recognition by alphabeta T cells. *Proc Natl Acad Sci U S A* 111(43):E4648-4657.
13. Staquet MJ, Jacquet C, Dezutter-Dambuyant C, & Schmitt D (1997) Fibronectin upregulates in vitro generation of dendritic Langerhans cells from human cord blood CD34+ progenitors. *J Invest Dermatol* 109(6):738-743.
14. Rosat JP*, et al.* (1999) CD1-restricted microbial lipid antigen-specific recognition found in the CD8+ alpha beta T cell pool. *J Immunol* 162(1):366-371.
15. Stubs G, Rupp B, Schumann RR, Schroder NW, & Rademann J (2010) Chemoenzymatic synthesis of a glycolipid library and elucidation of the antigenic epitope for construction of a vaccine against Lyme disease. *Chemistry* 16(11):3536-3544.
16. Stubs G*, et al.* (2009) Acylated cholesteryl galactosides are specific antigens of borrelia causing lyme disease and frequently induce antibodies in late stages of disease. *J Biol Chem* 284(20):13326-13334.
17. de Lalla C*, et al.* (2011) High-frequency and adaptive-like dynamics of human CD1 self-reactive T cells. *Eur J Immunol* 41(3):602-610.
18. de Jong A*, et al.* (2010) CD1a-autoreactive T cells are a normal component of the human alphabeta T cell repertoire. *Nat Immunol* 11(12):1102-1109.
19. Ly D*, et al.* (2013) CD1c tetramers detect ex vivo T cell responses to processed phosphomycoketide antigens. *J Exp Med* 210(4):729-741.

20. Tannaes T, Grav HJ, & Bukholm G (2000) Lipid profiles of Helicobacter pylori colony variants. *APMIS : acta pathologica, microbiologica, et immunologica Scandinavica* 108(5):349-356.
21. Yakimchuk K*, et al.* (2011) Borrelia burgdorferi infection regulates CD1 expression in human cells and tissues via IL1-beta. *Eur J Immunol* 41(3):694-705.
22. Roura-Mir C*, et al.* (2005) CD1a and CD1c activate intrathyroidal T cells during Graves' disease and Hashimoto's thyroiditis. *J Immunol* 174(6):3773-3780.
23. Sieling PA*, et al.* (2000) Human double-negative T cells in systemic lupus erythematosus provide help for IgG and are restricted by CD1c. *J Immunol* 165(9):5338-5344.
24. Lebre MC*, et al.* (2008) Rheumatoid arthritis synovium contains two subsets of CD83-DC-LAMP- dendritic cells with distinct cytokine profiles. *Am J Pathol* 172(4):940-950.
25. Wang CQ*, et al.* (2011) Th17 cells and activated dendritic cells are increased in vitiligo lesions. *PLoS One* 6(4):e18907.
26. Moret FM*, et al.* (2013) Intra-articular CD1c-expressing myeloid dendritic cells from rheumatoid arthritis patients express a unique set of T cell-attracting chemokines and spontaneously induce Th1, Th17 and Th2 cell activity. *Arthritis Res Ther* 15(5):R155.
27. Sekiya M, Osuga J, Igarashi M, Okazaki H, & Ishibashi S (2011) The role of neutral cholesterol ester hydrolysis in macrophage foam cells. *Journal of atherosclerosis and thrombosis* 18(5):359-364.
28. Chang TY*, et al.* (2001) Roles of acyl-coenzyme A:cholesterol acyltransferase-1 and -2. *Current opinion in lipidology* 12(3):289-296.
29. Russell DG, Cardona PJ, Kim MJ, Allain S, & Altare F (2009) Foamy macrophages and the progression of the human tuberculosis granuloma. *Nat Immunol* 10(9):943-948.
30. Yin YW*, et al.* (2014) TLR4-mediated inflammation promotes foam cell formation of vascular smooth muscle cell by upregulating ACAT1 expression. *Cell death & disease* 5:e1574.
31. Voloshyna I*, et al.* (2013) Plasma from rheumatoid arthritis patients promotes pro-atherogenic cholesterol transport gene expression in THP-1 human macrophages. *Experimental biology and medicine* 238(10):1192-1197.
32. Matulis G*, et al.* (2010) Innate-like control of human iNKT cell autoreactivity via the hypervariable CDR3beta loop. *PLoS biology* 8(6):e1000402.
33. Gadola SD*, et al.* (2003) Generation of CD1 tetramers as a tool to monitor glycolipid-specific T cells. *Philosophical transactions of the Royal Society of London* 358(1433):875-877.
34. Sahb MM ADJ, Baird MS (2015) Glucose monomycolates based on single synthetic mycolic acids. *Chemistry and physics of lipids*.
35. Sanderson JP*, et al.* (2012) Natural variations at position 93 of the invariant Valpha24-Jalpha18 alpha chain of human iNKT-cell TCRs strongly impact on CD1d binding. *Eur J Immunol* 42(1):248-255.
36. Sanderson JP*, et al.* (2013) CD1d protein structure determines species-selective antigenicity of isoglobotrihexosylceramide (iGb3) to invariant NKT cells. *Eur J Immunol* 43(3):815-825.
37. Gadola SD*, et al.* (2006) Structure and binding kinetics of three different human CD1d-alpha-galactosylceramide-specific T cell receptors. *J Exp Med* 203(3):699-710.
38. Winn MD*, et al.* (2011) Overview of the CCP4 suite and current developments. *Acta. Cryst.* D67:235-242.
39. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr* D60:2126-2132.
40. Langer G, Cohen S, Lamzin V, & Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* 3:1171-1179.
41. Schrodinger L (2010) The AxPyMOL Molecular Graphics Plugin for Microsoft PowerPoint, Version 1.0.
42. Sadowski J, Gasteiger J, & Klebe G (1994) Comparison of Automatic Three-Dimensional Model Builders Using 639 X-ray Structures. *Journal of chemical information and computer sciences* 34(4):1000-1008.

43. Jones G, Willett P, & Glen RC (1995) Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J Mol Biol* 245(1):43-53.
44. Liebeschuetz JW, Cole JC, & Korb O (2012) Pose prediction and virtual screening performance of GOLD scoring functions in a standardized test. *Journal of computer-aided molecular design* 26(6):737-748.
45. Inc CCG (2013) 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7. Molecular Operating Environment (MOE), 2013.10.
46. Hornak V*, et al.* (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* 65(3):712-725.
47. Case DA*, et al.* (2012) AMBER 12.
48. Wang J, Wolf RM, Caldwell JW, Kollman PA, & Case DA (2004) Development and testing of a general amber force field. *Journal of computational chemistry* 25(9):1157-1174.
49. Jakalian A, Jack DB, & Bayly CI (2002) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *Journal of computational chemistry* 23(16):1623-1641.
50. Salomon-Ferrer R, Götz AW, Poole D, Le Grand S, & Walker RC (2013) Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. *J. Chem. Theory Comput.* 9(9):3878-3888.
51. Craig IR, Pfleger C, Gohlke H, Essex JW, & Spiegel K (2011) Pocket-space maps to identify novel binding-site conformations in proteins. *Journal of chemical information and modeling* 51(10):2666-2679.
52. Dull T*, et al.* (1998) A third-generation lentivirus vector with a conditional packaging system. *J Virol* 72(11):8463-8471.

Figure Legends

Figure 1 Overview of human CD1c structure with aliphatic lipid spacers. (*A*) Cartoon representation of the CD1c-SL structure (α1-α3 domains in blue; β2m orange) in two orientations with bound lipid spacers represented as Van Der Waals spheres (yellow in A' channel; pink in F' channel, oxygen atoms shown in red). (*B*) Ligands bound in the CD1c antigen binding cavity shown with the F_O-F_C electron density calculated from an omit map and contoured at 1.5σ (grey mesh). (*C*) Chemical structure of bound ligands in CD1c-SL, C12 lauric acid and C18 stearic acid.

Figure 2 Comparisons between the structures of CD1c-SL and CD1c-MPM. (*A*) A pair of orthogonal views is shown for the molecular surface of CD1c-SL. (*B*) Equivalent views for CD1c-MPM. The ligands are represented as sticks and colour-coded as in **Fig. 1**. Key residues forming the F' channel roof in CD1c-SL are shown in silver-blue (CD1c-SL) and green (CD1c-MPM). These residues are shown in stick representation in the magnified views (right inset).

Figure 3 Molecular dynamics modelling of CD1c-SL. (*A*) Pocket volumes for different simulation setups against time, protein without ligand (red), protein with PEG ligands (blue), and protein with lauric acid ligands (green). (*B*) Comparison of CD1c-SL (light blue) versus a configuration from MD with empty channels (blue). Significant displacement of the α2 helix leads to collapse of the A' and F' channel in CD1c with empty channels. (*C* and *D*) Configurations of the CD1c-SL 40 ns trajectory with PEG with its F' roof residues shown in green and F' roof residues of CD1c-MPM in red. (*C*) Configuration at 0 ns of MD shows the tethered arrangement of the roof residues as seen in CD1c-SL (green). (*D*) As simulation progresses, at 40 ns the structure of the F' roof has been lost, giving configurations that more closely resemble CD1c-MPM (compare red and green). (*E*) Root mean square deviation (RMSD) of the four F' roof side chains with respect to their starting tethered arrangement. Small RMSD values indicate configurations highly similar to the CD1c-SL, whilst larger values indicate loss of this configuration. Data are representative of 3 independent experiments.

Figure 4 CD1c-SL tetramers bind human CD1c self-reactive αβ T cells. (*A*) Sorting of the T cell clone NM4 from a paternal T cell line using CD1c-SL tetramers. NM4 T cells are not stained by CD1b-GMM tetramers or CD1d-αGC tetramers. (*B*) NM4 T cells secrete TNF-α and IL-10 in response to CD1c-positive lymphoblasts (T2-CD1c) but not to parental CD1c-negative T2 lymphoblasts (T2). (**C**) Binding of CD1c-SL tetramers to different Jurkat lines stably expressing different CD1-restricted TCRs: NM4, CD1c-restricted self-reactive TCR; CD8-1, CD1c-restricted mycoketide-specific TCR; CD8-2, CD1a-restricted dideoxymycobactin-specific TCR; LDN5, CD1b-restricted GMM-specific TCR; iNKT, CD1d-

restricted TCR. Data in (*B*) are representative of at least 3 experiments (mean and s.d. of triplicate measurements).

Figure 5 Molecular docking into CD1c-SL binding cavity. (*A*) Surface representations of the binding cavities of CD1c-SL (top) and CD1c-MPM (bottom; PDB 3OV6) shown in different orientations. (*B*) Docking simulations for binding of cholesteryl 6-O-oleoyl β-D-galactopyranoside (ACGal), cholesteryl 6-O-tetradecanoyl α-D-glucopyranoside (α-ACGlu), and cholesteryl-oleate (CE) into the cavity of CD1c. Acyl chains and cholesterol moieties are accommodated within the A' and the F' channel, respectively. Ligand chemical structures are shown above the corresponding docking pose.

Figure 6 CD1c-ACGal complex binding to NM4-TCR. (*A*) Flow cytometry density plots of NM4-Jurkat and parental Jurkat T cell lines stained with $CD1c^{wt}$-ACGal (left two panels) and $CD1c^{ba3}$-ACGal (right two panels) tetramers. (*B*) Surface plasmon resonance measurements (BiaCore) for binding of NM4-TCR to immobilised $CD1c^{wt}$-ACGal (left) and $CD1c^{ba3}$-SL (right) complexes at equilibrium. K_D, calculated dissociation constant; RU, response units. Data in (*B*) are representative of 2 independent experiments.

Figure 7 Acylated steryl ligands reconstitute TCR-tetramer binding to CD1c. (*A*) Staining of T2-CD1c and parental T2 lymphoblasts with anti-CD1c antibody (clone AD5-8E7) (left panel) or NM4-TCR tetramer (right panel). NM4-TCR tetramer binding to T2-CD1c is blocked by anti-CD1c antibody (right panel). (*B*) NM4-TCR tetramer binding to unconjugated beads (left panel) and CD1c-conjugated beads (middle and right panels), before (middle) and after (right) treatment of beads with Triton-X. (*C*) NM4-TCR tetramer binding to Triton-X-treated CD1c-conjugated beads after incubation ("loading") with DMSO (vehicle), the acylated steryl glycosides ACGal and β-sitosteryl glucoside (ASGlu), or the cholesteryl ester 5-cholestene 3-palmitate (CE). Numbers shown in top right quadrant of dot plots in (*B*) and (*C*) show the mean fluorescence intensity (MFI) of NM4-TCR tetramer staining. Data in (*B*) and (*C*) are representative of at least 3 independent experiments.