Differing roles of CD1d2 and CD1d1 proteins in type I natural killer T cell development and function

Srinivasan Sundararajan, Jingjing Zhang, S. Harsha Krovi, Romain Bedel, Kathryn D. Tuttle, Natacha Veerapen, Gurdyal S. Besra, Yogesh Khandokar, Praveena, Jérôme Le Nours, Jennifer L. Matsuda, Jamie Rossjohn, Laurent Gapin

CD1 molecules are glycoproteins that adopt an MHC class I-like structure (1). In contrast to the polymorphic MHC molecules that present peptides, CD1 molecules exhibit limited polymorphism and are ideally suited to present lipid-based antigens (2). Based on sequence homology, function, and tissue distribution, the CD1 family has been classified into three distinct groups: group I, includes CD1a, -b, and -c; group II consists of CD1d, and group III is composed of CD1e (3). The number of CD1 isoforms expressed in any given mammal varies considerably, however. For example, humans express all five CD1 isoforms, rabbits lack CD1c, and mice express only the CD1d protein. However, the basis for such varied usage of genes within the CD1 locus across mammals is largely unknown, but presumably reflects important functional differences related to pathogen-driven evolutionary pressures.

Lipid antigen-loaded CD1d molecules are recognized by a population of innate-like T lymphocytes, the natural killer T (NKT) cells. Two broad classes of NKT cells, termed type I and type II, have been defined on the basis of T cell antigen receptor (TCR) expression and antigen reactivity (4). Type I NKT cells or natural killer T (NKT) cells are selected by CD1d molecules in the thymus. Two homologous genes in the mouse genome encode CD1d molecules, which are expressed primarily by cells of the hematopoietic lineage, including B and T cells, macrophages, and dendritic cells (21, 22). At the transcriptional level, CD1d1 expression is approximately 95% sequence homology with each other, and are likely the product of gene duplication (2, 20). Mouse CD1d molecules are expressed by cells of the hematopoietic lineage, including B and T cells, macrophages, and dendritic cells (21, 22). At the transcriptional level, CD1d1 is expressed at higher levels than CD1d2 in all tissues tested except the thymus, where the mRNA levels appear to be equal for both genes (19). Of the 17 amino acid differences between CD1d1 and CD1d2 proteins, 9 occur within the antigen-binding cleft, which includes all of these are responsive to the prototypical antigen α-galactosylceramide (αGalCer) when presented by CD1d molecules, irrespective of their TCRβ repertoire diversity (8–10). How variation in CD1d itself impacts on type I NKT cell repertoire and function remains unclear, however.

Upon antigenic stimulation, type I NKT cells respond rapidly by producing Th1, Th2, and Th17 cytokines, and can also be cytotoxic (11, 12). Type I NKT cells can be activated by bacterial ligands, and sense changes in host lipid metabolism through recognition of transiently expressed endogenous lipids (13–15). Chronic stimulation by these “autoantigens” leads to the product of gene duplication (20). Mouse CD1d molecules are expressed by cells of the hematopoietic lineage, including B and T cells, macrophages, and dendritic cells (21, 22). At the transcriptional level, CD1d1 is expressed at higher levels than CD1d2 in all tissues tested except the thymus, where the mRNA levels appear to be equal for both genes (19). Of the 17 amino acid differences between CD1d1 and CD1d2 proteins, 9 occur within the antigen-binding cleft, which includes all of these are responsive to the prototypical antigen α-galactosylceramide (αGalCer) when presented by CD1d molecules, irrespective of their TCRβ repertoire diversity (8–10). How variation in CD1d itself impacts on type I NKT cell repertoire and function remains unclear, however.

Upon antigenic stimulation, type I NKT cells respond rapidly by producing Th1, Th2, and Th17 cytokines, and can also be cytotoxic (11, 12). Type I NKT cells can be activated by bacterial ligands, and sense changes in host lipid metabolism through recognition of transiently expressed endogenous lipids (13–15). Chronic stimulation by these “autoantigens” leads to the product of gene duplication (20). Mouse CD1d molecules are expressed by cells of the hematopoietic lineage, including B and T cells, macrophages, and dendritic cells (21, 22). At the transcriptional level, CD1d1 is expressed at higher levels than CD1d2 in all tissues tested except the thymus, where the mRNA levels appear to be equal for both genes (19). Of the 17 amino acid differences between CD1d1 and CD1d2 proteins, 9 occur within the antigen-binding cleft, which includes all of these are responsive to the prototypical antigen α-galactosylceramide (αGalCer) when presented by CD1d molecules, irrespective of their TCRβ repertoire diversity (8–10). How variation in CD1d itself impacts on type I NKT cell repertoire and function remains unclear, however.

Upon antigenic stimulation, type I NKT cells respond rapidly by producing Th1, Th2, and Th17 cytokines, and can also be cytotoxic (11, 12). Type I NKT cells can be activated by bacterial ligands, and sense changes in host lipid metabolism through recognition of transiently expressed endogenous lipids (13–15). Chronic stimulation by these “autoantigens” leads to the product of gene duplication (20). Mouse CD1d molecules are expressed by cells of the hematopoietic lineage, including B and T cells, macrophages, and dendritic cells (21, 22). At the transcriptional level, CD1d1 is expressed at higher levels than CD1d2 in all tissues tested except the thymus, where the mRNA levels appear to be equal for both genes (19). Of the 17 amino acid differences between CD1d1 and CD1d2 proteins, 9 occur within the antigen-binding cleft, which includes all of these are responsive to the prototypical antigen α-galactosylceramide (αGalCer) when presented by CD1d molecules, irrespective of their TCRβ repertoire diversity (8–10). How variation in CD1d itself impacts on type I NKT cell repertoire and function remains unclear, however.
Cys168→Trp168 mutation. The cysteine 168 residue participates in an intradomain disulfide bond that is highly conserved in classical and nonclassical MHC I molecules (1, 19, 23), and is thought to be critical for the folding of MHC-I (24, 25). The impact of this mutation on CD1d2 structure and function is unknown. Interestingly, the bovine CD1d gene (26) and the primordial CD1-encoding gene found in chickens (27) also contain an amino acid change at position 168. This change was previously shown to affect the nature of the antigens presented by these molecules (2). Intriguingly, rats possess only one CD1D gene, which, based on sequence homology, more resemble mouse CD1DI than CD1D2, with a cysteine at position 168 (28). It remains unclear what advantage, if any, the gene duplication event that led to the existence of CD1D2 in the mouse might be.

The development of type I NKT cells requires expression of CD1d by cortical thymocytes (29) and is impaired in mice lacking CD1d expression. Interestingly, even though some targeting strategies led to the disruption of both CD1D genes (30, 31), one particular strain of CD1D-deficient mice, which was generated by using embryonic stem cells of 129/Sv origin and has been largely distributed in the scientific community and used in multiple studies, disrupted only the CD1DI gene (32). Furthermore, sequencing of the CD1D2 gene from C57BL/6 mice revealed a frame-shuffling mutation at the beginning of the fourth exon encoding the α3 domain, thereby abolishing CD1d2 protein expression in this strain (22). Altogether, these results contributed to the assumption that CD1d2 molecules do not play any role in the development of type I NKT cells. Earlier studies that identified type I NKT cells by using costaining for TCRβ chains and the NK1.1 marker concurred with this notion (33). However, it remains unknown whether the CD1D2 gene is mutated in other strains of mice and whether its gene product might play a role in the selection and function of iNKT cells.

Here we report that CD1d2 molecules are expressed in the thymus of CD1DI1−/−-deficient mice (32), where they select functional type I NKT cells. High-throughput sequencing of CD1D gene transcripts from the thymus of BALB/c mice showed a transcript ratio of CD1DI:CD1D2 close to 1:1, supporting the possibility for a role of CD1d2 in shaping the development of iNKT cells. The CD1D2 crystal structures revealed that CD1d2 adopts an overall architecture similar to CD1d1 and that the Cys168Trp change was found to appear different for CD4+ and CD8+ T cells. To ensure that CD4+ T cells from CD1D1D2−/−CD4−/− mice (34, 35). An alternative explanation could be provided if some residual iNKT cells were still present in the sample of the sorted CD4+ T cells from CD1D1−/−TCRα−/− mice. Indeed, upon careful flow cytometry analysis involving the acquisition of several millions of events, we clearly detected a population of PBS57-CD1d1 tetramer−/− cells in the thymus, spleen, and liver of 6- to 8-wk-old C57BL/6, CD1DI−/−, CD1D1D2−/−, and Jα18−/− mice stained with CD1d1 tetramers loaded with PBS57 (PBS57-CD1d), and TCRβ mAbs and analyzed by flow cytometry (data representative of n ≥ 3 per group across two experiments). The percentage of iNKT cells in each sample is shown. (C) Summary of the data shown in B with percentage of iNKT cells in thymus, spleen, and liver of 6- to 8-wk-old C57BL/6, CD1DI−/−, CD1D1D2−/−, and Jα18−/− mice (P < 0.05, **P < 0.01, ***P < 0.001, and ****P < 0.0001; ns, not significant).

Results

**Vα14-Jα18 TCR Usage in Conventional CD4**+ **and CD8**+ **T Cells from CD1DI−/− Mice.** We wanted to determine the frequency of the Vα14-Jα18 rearrangement within conventional CD4+ and CD8+ TCRβ+ T cells. To ensure that CD4+ and CD8+ T cell populations would be devoid of any potentially contaminating iNKT cells, we sorted cells from the spleens of C57BL/6 CD1DI−/−TCRα−/− mice by flow cytometry. The frequency of gene usage was measured by quantitative PCR and normalized to the levels found in PBS57-CD1d1 tetramer+ TCRβ+ cells sorted from the spleen of CD1DI−/−TCRα−/− mice. The strategy of amplification is depicted in SI Appendix, Fig. S1. As seen in Fig. 1A, 2-3% of conventional T cells use the Vα14 gene segment. Conventional splenic CD8+ T cells that use the Vα14 gene segment appear to only very rarely rearrange with the Jα18 gene segment, as revealed by the use of a Jα18-specific probe. Surprisingly, the situation appeared different for CD4+ T cells, in which the results suggested that all Vα4 segments were rearranged with Jα18, like in iNKT cells. Sequencing of the PCR product showed that it corresponded to the canonical Vα14-Jα18 iNKT rearrangement. These perplexing results suggested that, within the CD4+ T cells derived from CD1DI−/−-deficient mice, Vα4 might have preferentially rearranged with Jα18, contradicting previous findings showing that Vα14 can rearrange with all Jα segments (34, 35). An alternative explanation could be provided if some residual iNKT cells were still present in the sample of the sorted CD4+ T cells from CD1DI−/−TCRα−/− mice. Indeed, upon careful flow cytometry analysis involving the acquisition of several millions of events, we clearly detected a population of PBS57-CD1d1 tetramer−/− cells in the thymus, spleen, and liver of CD1DI−/− mice (Fig. 1B). This population was absent from the organs of CD1DI/CD1D2−/− mice and greatly decreased in Jα18−/− mice, which lack the invariant Tα18 gene segment (36, 37). The few
cells that were labeled in Jα18<sup>−/−</sup> mice likely correspond to the atypical iNKT cells previously identified in that strain (57, 38). These results contrast with previous experiments that examined the presence of NKT cells in CD1<sup>D1</sup>/− mice by using TCRβ<sup>+</sup> and NK1.1<sup>+</sup> staining and concluded that CD1d2 expression could not rescue NKT cell development (33). Instead, our results demonstrate that, even though the proportion of iNKT cells in CD1<sup>D1</sup>/− mice is only 1−4% of the iNKT cells found in the organs of C57BL/6 (B6) mice (Fig. 1C), CD1<sup>D1</sup>/− mice should not be considered as iNKT cell-deficient, and CD1d2 molecules can potentially select for iNKT cells.

The CD1D2 Sequence from CD1D1<sup>−/−</sup> Mice and the WT BALB/c Strain Is Functional. B6 mice harbor a point mutation in exon 4 of CD1D2, which encodes the α<sub>3</sub> domain (22). This frame-shift mutation introduces a stop codon, abolishing surface expression of the molecule. In the 129/sv strain, however, CD1D2 does not contain this mutation (22). To confirm that CD1D2 in the CD1D1<sup>−/−</sup> mice (32) is not the pseudogene found in the B6 mice, we designed primers that span part of exon 4 and the adjoining intronic area. Genomic DNA from CD1D1<sup>−/−</sup> mice, B6 mice, and another WT strain, BALB/c, was purified and used for amplification of the exon 4 fragment. Sequencing results confirmed that, whereas the B6-derived sequence contains a stop codon in exon 4, BALB/c and 129/sv strains had a functional CD1D2 sequence. These results demonstrate that no cross-over of the CD1 locus ever occurred during the backcross of the CD1D1<sup>−/−</sup> mice to B6 mice and that the CD1 locus is of 129/sv origin in these mice. Furthermore, these results suggest that CD1d2 molecules have the potential to be expressed in BALB/c mice.

CD1d2 Protein Is Expressed in the Thymus but Not the Periphery of CD1D1<sup>−/−</sup> Mice. Because no serological reagent currently exists to discriminate CD1d1 from CD1d2 proteins, we first examined CD1d2 expression on cells derived from B6 CD1D1<sup>−/−</sup> mice. Controls were provided by staining cells from double-deficient CD1D1/CD1D2 mice (30, 33) and heterozygous CD1D1<sup>+/−</sup> mice generated by crossing CD1D1<sup>−/−</sup> mice with B6 mice. In this latter strain, the CD1d protein levels represent the combined gene activity of a single copy of CD1D1 B6 allele (because CD1D2 is mutated in B6) and a single copy of the functional CD1D2 allele from the KO mice (because CD1D1 has been knocked out in these mice and the CD1D2 allele is of 129/sv origin). As seen in Fig. 2, we observed a low level of CD1d expression in the thymus of CD1D1<sup>−/−</sup> mice, attributable to CD1d2 expression (Fig. 2A). In particular, whereas CD1D1<sup>+/−</sup> thymocytes displayed haploinsufficient patterns, expressing half the amount of CD1d found on thymocytes from WT mice [as measured by geometric mean fluorescence intensity (gMFI)], the CD1D1<sup>−/−</sup> thymocytes expressed only 1/25 of the level expressed in WT mice (Fig. 2C).

Interestingly, this expression was limited to the thymus, in agreement with previous results (33). In the periphery, we observed no protein expression of CD1d2 on whole splenocytes or individual splenic populations (Fig. 2B and D).

Early experiments with the use of RNase protection assays suggested that CD1D2-encoding transcripts were equally abundant compared with CD1D1 transcripts in BALB/c thymocytes (19). To verify these findings, we generated cDNA from B6 and BALB/c DP thymocytes and amplified a portion of exon 2 by using primers targeting identical sequences between CD1D1 and CD1D2 genes. This segment of DNA, however, contains eight nucleotide polymorphisms between the two genes. PCR products were sequenced by using high-throughput sequencing, and the relative expression of CD1D1 and CD1D2 transcripts was determined in each strain. The transcript ratio of CD1D1 to CD1D2 was overwhelmingly skewed toward CD1D1 in B6, with an expression ratio close to 5:1. This likely reflects the degradation of CD1D2-encoding mRNA as a result of nonsense-mediated mRNA decay (39). In BALB/c, however, the ratio of CD1D1:CD1D2 is close to 1:1 (Fig. 2E). Altogether, these results demonstrate that CD1d2 molecules can be expressed in the thymus of mice with a functional CD1D2 gene.

Development and Repertoire of CD1d2-Selected iNKT Cells. We enriched PBS57−CD1d1+ cells from the thymus of B6 and CD1D1<sup>−/−</sup> mice by using magnetic beads and analyzed the maturation of iNKT cells by using the cell-surface markers CD44 and NK1.1 to follow their maturation progress from stages 1 to 3 (40, 41). CD1d2-selected iNKT cells in CD1D1<sup>−/−</sup> mice appeared to have a mature compared with CD1d1-selected iNKT cells from B6 mice (Fig. 3A). Indeed, the majority of iNKT cells in the CD1D1<sup>−/−</sup> mice were of the more immature stage 1 (CD44<sup>−</sup>NK1.1<sup>−</sup>) or stage 2 (CD44<sup>+</sup>NK1.1<sup>+</sup>) phenotype compared with WT mice, in which approximately 80% of cells are the most mature stage 3 (CD44<sup>+</sup>NK1.1<sup>−</sup>) phenotype (Fig. 3C). However, the absolute number of iNKT cells stages detected in CD1D1<sup>−/−</sup> thymi was significantly lower than that observed in control thymi as a result of reduced total iNKT cell numbers in the CD1D1<sup>−/−</sup> mice. Because analysis of iNKT cells by staging is insufficient, as populations within each stage are heterogeneous (16, 42), we also characterized the CD1d2-selected iNKT cells by their expression of the transcription factors PLZF, RORγt, and Tbet to define terminally differentiated effector cells. Compared with the iNKT subset makeup found in the B6 thymus, iNKT cells in

Sundararaj et al.  PNAS Early Edition | 3 of 10
the thymus of C1D1−/− mice showed a trending increase in the iNKT2 (PLZF<sup>hi</sup>, Tbet<sup>hi</sup>, RORγ<sup>hi</sup>) and iNKT17 (PLZF<sup>mid</sup>. Tbet<sup>lo</sup>, RORγ<sup>lo</sup>) populations and a concomitant decrease in iNKT1 (PLZF<sup>lo</sup>, Tbet<sup>lo</sup>, RORγ<sup>lo</sup>) cells (Fig. 3 B and C). Interestingly, even though the vast majority of the CD1d2-selected iNKT cells are of the iNKT1 phenotype (i.e., PLZF<sup>lo</sup>, Tbet<sup>lo</sup>), these cells are mostly NK1.1<sup>−</sup> and lack expression of the NK receptors Ly49C/I, in striking contrast with the same population of iNKT1 cells in B6 mice (Fig. 3F). This might have contributed to their lack of detection in previous studies (33).

Next, we examined the V<sub>j</sub> repertoire expressed by iNKT cells selected on CD1d1 in B6 mice or on CD1d2 in C1D1−/− mice.
by flow cytometry. Similar to what has been reported previously for the CD1d-selected iNKT repertoire (43, 44), CD1d2-selected iNKT cells were enriched for Vβ8 and Vβ7 cells compared with conventional T cells. However, a small but consistent increase in the proportion of cells using the Vβ gene segment, concomitant to a reduction in Vβ7 gene usage, was observed for CD1d2-selected iNKT cells compared with CD1d1-selected iNKT cells (Fig. 3D). To further compare the TCR repertoires of CD1d1-selected and CD1d2-selected iNKT cells, we sorted PBS57-CD1d1 tetramer+ cells from the thymi of two C57BL/6 mice and more than 30 pooled CD1D1+ thymi. Total RNA was extracted, and cDNA was synthesized. We amplified Vβ8.1/8.2 rearrangements by using specific Vβ8.1/2 and Cβ primers, and the PCR fragments were sequenced by using next-generation sequencing. Fig. 3E presents a visual depiction of all sequences found in the B6 and CD1D1−/− Vβ8 iNKT repertoires, where every dot represents a unique CDR3β sequence, and the size of the dot represents how many times this sequence was found, and the color represents which Jβ was used. As expected, CD1d1-selected iNKT cells expressed a polyclonal repertoire with no overrepresentation of any particular sequence (45). Although the CD1d2-selected iNKT cell repertoire was also polyclonal, it appeared to be much more limited, with fewer clones and fewer Jβs represented. In fact, when comparing the overlap between the two repertoires, we observed that only 9% of the sequences from CD1D1−/− iNKT cells are found in the B6 repertoire. Thus, the majority of Vβ8.1/8.2 TCR sequences found in the iNKT cells from CD1D1−/− mice appear unique to these iNKT cells. To further explore the antigen reactivity of CD1d1 vs. CD1d2-selected iNKT TCR repertoires, we cloned libraries of Vβ8.1/8.2 TCR sequences into retroviral plasmids. We then transduced a TCR-negative T hybridoma with the Vα14-Jα18 TCRα chain and the two different Vβ libraries derived from B6 or CD1D1−/− iNKT cells. The hybridoma was further transduced with a NFAT-GFP construct, allowing for measurement of TCR stimulation on a per-cell basis. Hybridomas expressing Vβ8 TCR repertoires derived from B6 or CD1D1−/− mice were cultivated overnight in the presence of the CHB lymphoma that had been previously transfected (or not) with CD1d1- or CD1d2-expressing constructs. As seen in Fig. 3F, approximately 2% of the hybridoma expressing the B6 iNKT cell repertoire expressed GFP upon stimulation by CD1d-expressing APCs. The proportion of autoreactive cells was increased to 4–5% of the cells when the repertoire derived from CD1D1−/− iNKT cells was interrogated in a similar fashion (Fig. 3F). Cell sorting of these GFP+ cells and subsequent restimulation with CD1d-expressing APCs demonstrated the autoreactivity of these cells toward CD1d-expressing APCs. Altogether, these data suggest that the TCR repertoire of iNKT cells found in CD1D1−/− mice is more autoreactive than the iNKT cells. The hybridoma was further transduced with a NFAT-GFP construct, allowing for measurement of TCR stimulation on a per-cell basis. Hybridomas expressing Vβ8 TCR repertoires derived from B6 or CD1D1−/− mice were cultivated overnight in the presence of the CHB lymphoma that had been previously transfected (or not) with CD1d1- or CD1d2-expressing constructs. As seen in Fig. 3F, approximately 2% of the hybridoma expressing the B6 iNKT cell repertoire expressed GFP upon stimulation by CD1d-expressing APCs. The proportion of autoreactive cells was increased to 4–5% of the cells when the repertoire derived from CD1D1−/− iNKT cells was interrogated in a similar fashion (Fig. 3F). Cell sorting of these GFP+ cells and subsequent restimulation with CD1d-expressing APCs demonstrated the autoreactivity of these cells toward CD1d-expressing APCs. Altogether, these data suggest that the TCR repertoire of iNKT cells found in CD1D1−/− mice is more autoreactive than the TCR repertoire expressed by B6-derived iNKT cells. This might be a result of the nature of the CD1d2 molecules themselves, alterations in the self-lipid repertoire, and/or their low levels of expression in the thymus of CD1D1−/− mice (Fig. 2).

**CD1d2-Selected iNKT Cells Are Functional.** To test whether CD1d2-selected iNKT cells were functional in vivo, we injected 2 μg of the prototypic antigen αGC, which has a 26-carbon-long fatty acid chain, i.v. to B6 and CD1D1−/− mice and harvested livers from the animals 90 min later to interrogate IL-4 and IFN-γ production within iNKT cells (44). Unlike B6 controls, in which 90% of iNKT cells were activated as detected by IFN-γ intracellular cytokine staining, no CD1D1−/−-derived iNKT cells were responsive to injected antigen, resembling PBS solution-injected controls. These results probably reflect the absence of the CD1d molecule expression in the peripheral tissues required to present the antigen. However, iNKT cells can also be activated in a TCR-independent manner in response to inflammatory cytokines secreted by activated antigen-presenting cells (46, 47). Therefore, we injected 40 μg of LPS or PBS solution control to B6 or CD1D1−/− mice and harvested their livers 7 h after injection. Cells were stained intracellularly for IFN-γ directly ex vivo without further stimulation and/or addition of Brefeldin A. Although the activation of iNKT cells through LPS is less robust than through TCR engagement by the αGC/CD1d complex, the results clearly demonstrated that CD1d1-selected as well as CD1d2-selected iNKT cells could produce IFN-γ under these conditions (Fig. 3G and H). These results demonstrate that CD1d2-selected iNKT cells are functional and capable of secreting cytokines upon stimulation. Furthermore, these results support the idea that the maintenance of iNKT cell effector function is TCR independent and does not require tonic signaling mediated by CD1d molecules.

**Overall Architecture of CD1d2.** To gain insight into the detailed architecture of the CD1d2 molecule, we expressed and purified CD1d2 from the supernatant of insect cells and subsequently determined its structure. CD1d2 adopted a standard MHC-I fold, like CD1d1 (Fig. 4 and SI Appendix, Table S1). Similar to the overall architecture of other CD1d molecules (48–51), the α1- and α2-helices are positioned atop of a six antiparallel β-strands platform to form a deep hydrophobic antigen-binding groove (~1,300 Å²) that accommodates lipid-based antigens (Fig. 4A). Following the nomenclature adopted for mouse CD1d3 (5), the antigen-binding cleft is composed of two distinct A′- and F′-pockets (Fig. 4A and B) that differ in size and architecture. The contour of the A′-portal (~550 Å²) is delineated by a number of bulky hydrophobic residues (W40, F70, Y73, F77, F116, and W168) with residues Y73 and F77 separating the A′-pocket from the F′-pocket (Fig. 4B). The F′-pocket (750 Å²) is shaped by an extensive network of hydrophilic residues. Here, residues L100, V118, F120, A125, F128, and L143 form the floor of the pocket and a number of key hydrophobic residues from the α1- and α2-helices (Y73, F77, S181, L84, V85, M88, W142, L143, P146, I147, L150, and T156) define the overall shape of the F′-pocket (Fig. 4B, Right). The CD1d2 crystal structure also revealed clear separate unaccounted and continuous electron density within the A′- and F′-portals to protrude out of the binding cleft (SI Appendix, Fig. S2). This density suggested the presence of an endogenous bound lipid containing a head group that was acquired from the HIS insect cells during the recombinant expression of CD1d2. Although the precise identity of the endogenous ligand is undetermined, we modeled in the unbiased electron density the common self-membrane phospholipid phosphatidic acid (PA). This enabled us to evaluate the approximate length of the acyl chains that the endogenous lipid may contain. PA consisting of two acyl chains composed of 10 and 18 carbons could accommodate the continuous electron density that was visible in the A′- and F′-pockets, respectively (SI Appendix, Fig. S2).

To formally demonstrate that CD1d2 presents lipid antigens of shorter acyl-chain length, we conducted lipid loading experiments with a truncated analog of αGC in which the acyl chain was 10 carbon atoms in length [αGC (C10)]. Although we could not load recombinant CD1d2 with αGC, the αGC (C10) form readily complexed with CD1d2. We then subsequently determined the structure of CD1d2-αGC (C10, SI Appendix, Table S1), in which the C10 acyl chain was clearly observed within the A′-pocket of CD1d2. The structure of CD1d2-αGC (C10) was very similar to that of CD1d2 bound to endogenous lipids, and the mode of αGC (C10) presentation by CD1d2 was very similar to that of CD1d1-αGC. Accordingly, CD1d2 can present αGC (C10) in a manner analogous to CD1d1 presentation of α-GC.

**Comparison of the CD1d2 and CD1d1 Crystal Structures.** CD1d2 and CD1d1 share high amino acid sequence identity (93%; SI Appendix, Fig. S3) that is reflected by a very similar α1- and α2-domain architecture with an rmsd of 0.9 Å over 178 Co atoms (Fig. 4D). Thus, the presence of the Cys168Trp polymorphism did not lead to
gross conformational changes in the CD1d2 architecture. The majority of the polymorphisms between CD1d1 and CD1d2 are evenly distributed over the α2-helix and the six β-sheet platform, whereas the α1-helix sequence is conserved between CD1d1 and CD1d2 (Fig. 4D and SI Appendix, Fig. S3). Ten of these polymorphic residues are located on the molecular surface of CD1d2 (SI Appendix, Fig. S4A). The remaining six residues (I30, T102, F116, W168, and A172 in CD1d2) are predominantly buried within the A′-pocket of CD1d2 (Fig. 4D and SI Appendix, Fig. S4B). Interestingly, residues I30, T102, F116, and W168 are directly implicated in remolding the overall architecture of the A′-portal relative to CD1d1 (Figs. 4E and 5A and SI Appendix, Fig. S4A). The substitution of V30, L116, and C168 in CD1d1 by the bulkier hydrophobic I30, F116, and W168 residues in CD1d2, respectively, contributes to a reduction of the A′-pocket volume of CD1d2 by ~400 Å³ compared with CD1d1, and also serves to prevent the collapse of the A′-pocket as a result of the lack of the intermolecular disulfide bond. Here, to accommodate the endogenous lipid and αGC (C10), the side chain of F70 was repositioned in close vicinity of I30 and W168 (~4 Å distance), and consequently this triad of residues contributed to restrict the size of the A′-portal (Figs. 4E and 5A). Accordingly, the structural rearrangement incurred by these polymorphic residues may impact the overall repertoire of lipids that can be presented by CD1d2 compared with CD1d1. Modeling of longer acyl chains in the A′-pocket was sterically disfavored by the architecture of the CD1d2 groove (Fig. 5C and D). Indeed, the A′-pocket of CD1d2 favors the loading of shorter sized lipids compared with that of CD1d1, as evidenced by the CD1d2-αGC(C10) structure. In the CD1d2 A′-pocket, residues T102 and F116 also impacted the overall position of the acyl chain of the endogenous lipid (Fig. 5B). By contrast, the F′-pockets of CD1d2 and CD1d1 are highly conserved (Fig. 4D) whereby only A125 (CD1d2) was replaced by V125 (CD1d1; Fig. 4D and E). Thus, molecular insights into the antigen-binding cleft of CD1d2 show that lipid antigens with long acyl chains are not as efficiently presented compared with CD1d1 molecules.

Fig. 4. CD1d2 crystal structure. (A) Cartoon representation of the CD1d2 crystal structure. CD1d2 and β2m are colored in gray and light orange, respectively. (B) Close-up views of the A′- (Left) and F′-pockets (Right) of CD1d2. Both pockets are colored in pale green and shown as surface representations. (C, Left) Unbiased Fo-Fc electron density map (magenta) contoured at 2.0σ level of the αGC (C10) lipid Ag. (C, Right) 2Fo-Fc electron density map (marine) contoured at 0.8σ level of the αGC (C10) lipid Ag. The lipid Ag is shown as black sticks. For clarity, only the α1- and α2-helices of CD1d2 are colored in light gray and shown as cartoon representations. (D) Superimposition of the α1- and α2-domains of the crystal structures of CD1d2 (gray) and CD1d1 (PDB ID code 1CD1; light blue) (48). The polymorphic residues between CD1d2 and CD1d1 are shown as sticks. For clarity, only the residues located inside the binding groove are shown. (E) Surface representation of the binding pocket of CD1d2 (light green) and CD1d1 (pale yellow). A superimposition of the α1- and α2-domains of CD1d2 (gray) and CD1d1 (light blue) is also shown.
chains) derived from B6 thymic CD1D2 cells. When the gene is functional, and, in BALB/c mice, is B might hybridoma cells was assessed after 18 h. As seen in Fig. 6, CD1d2-expressing CHB cells poorly stimulated the hybridomas compared with CD1d1-expressing cells when cultured in the presence of glycolipid antigens with long acyl chains, including the prototypical iNKT antigen αGC (C26), the αGC variant with a shorter sphingosine chain OCH9 (C24), and the Spingomonas-derived antigen GSL-1 (C14). In striking contrast, a version of αGC with a 10-carbon acyl chain, or the α-PS-sphingosine antigen, which lack an acyl chain (52, 53), equivalently stimulated the hybridoma when cultivated with CD1d1- or CD1d2-expressing cells. The β-anomeric form of psychosine did not stimulate the hybridoma when cultured with either APCs. CHB cells expressing CD1d1 molecules with the Cys168→Trp168 mutation were also poor presenters of αGC (C26) but could present αGC (C10) as efficiently as WT CD1d1 molecules (Fig. 6). Altogether, these results suggest that CD1d2 molecules are poor presenters of long acyl chain glycolipid antigens but can present certain antigens with short acyl chains as efficiently as CD1d1 molecules. These results demonstrate that both CD1d1 molecules in mice are not equivalent and can selectively present different lipid antigens to iNKT cells.

CD1d2-αGC-C10 Tetramer+ cells. Having established that CD1d2 molecules preferentially present glycolipid antigens with short acyl chains, we loaded soluble CD1d2 monomers with αGC(C10) and generated CD1d2 tetramers. Thymocytes from BALB/c, B, CD1D1−/−, and Jα18− mice were independently stained by using the PBS57-CD1d1 or the αGC(C10)-CD1d2 tetramers (Fig. 7A). The percentage of cells stained in each thymus was essentially identical between tetramers, even though αGC(C10)-CD1d2 tetramers stained with lower intensity compared with PBS57-CD1d1 tetramers. Containing with saturating concentrations of CD1d1 and CD1d2 tetramers loaded with PBS57 or αGC(C10), respectively, were initially unsuccessful because no CD1d2 signal was detected in these instances (data not shown). However, sequential staining with the two tetramers could be achieved if CD1d1 tetramers were used first. With the use of this strategy, we did not detect a significant population of αGC(C10)-CD1d2-specific cells (Fig. 7B). Instead, a fraction of iNKT cells was stained with both tetramers, whereas some PBS57-CD1d1 tetramer+ cells were not labeled with the αGC(C10)-CD1d2 tetramer. Interestingly, cells labeled with the αGC(C10)-CD1d2 tetramer expressed the highest levels of TCR on their surface (Fig. 7B), suggesting that the αGC(C10)-CD1d2 binds only to the highest avidity iNKT TCRs.

Discussion

Peptide-presenting major histocompatibility molecules are encoded by the most polymorphic genes in the genome. Most of this polymorphism involves residues that are lining the antigen-binding groove of MHC molecules (54). At the population level, these polymorphisms provide various possible binding solutions to the myriad of antigenic peptides that need to be presented for immunosurveillance by T cells. By contrast, lipid-binding CD1 molecules are essentially monomorphic; with the number and complexity of CD1-encoding genes varying greatly between species. The functional consequence of expressing multiple CD1 proteins presumably reflects pathogen-driven evolutionary pressure. The mouse genome contains two genes, CD1D1 and CD1D2, which encode for the CD1d molecules. A collection of observations had contributed to the misleading idea that CD1D2 might not play any relevant role in the biology of lipid-reactive T cells. By contrast, we now show that, in the 129/Sv and BALB/c mouse strains, the CD1D2 gene is functional, and, in BALB/c mice, is expressed equivalently to the CD1D1 gene at the mRNA level in the thymus. The reason why CD1d2 protein expression was detected in the thymus of CD1D1−/− mice (albeit at low levels) but not in splenocytes is uncertain. Early experiments examining the pattern of CD1d expression in whole organs or in cell lines suggested that it was controlled, at least in part, at the transcriptional level (19, 55). Our current understanding of CD1D...
CD1D1 mice were somewhat enriched for the iNKT cells that express half the TCR of the highest avidity, which would otherwise be negatively selected in WT animals. Alternatively, the presentation of a different set of self-antigen(s) responsible for the positive selection of iNKT cells between CD1d1 and CD1d2 could contribute to these differences in repertoire and phenotype. These two possibilities are not exclusive. Future experiments aimed at manipulating the expression level of CD1d2 to match the levels observed for CD1d1 should be informative in that matter.

One aspect that determines the function of a CD1 protein is the size and shape of the antigen-binding groove. For example, human CD1b has a large antigen-binding groove that can accommodate ligands of as much as 80 carbon atoms in length (49). In contrast, the human CD1a isoform has a unique architecture that appears to favor “headless” antigens, with a restricted A′-pocket that prefers lipids of shorter acyl-chain length (C18–C20) (50, 58). Our work resonates with the CD1a molecule in that CD1d2 favors the presentation of shorter length lipid ligands in comparison with CD1d1. Despite high homology between the two CD1d molecules, changes in amino acids at position I30, F116, and especially W168 restrict the size of the A-pocket compared with CD1d1. The Trp-168 blocks part of the A′-pocket so that long fatty acid chain glycolipids will not fit. In agreement with the structural data, APCs engineered to express the CD1d2 molecules or the C168W CD1d1 mutant were poor presenters of common glycosphingolipids, including the prototypical antigen αGC, which consist of fatty acid with a length of 24–26 carbons. Interestingly, comparable observations had been previously reported whereby the equivalent cysteine residue in the bovine CD1d and the ancestral chicken CD1-d2 are replaced by a tryptophan and phenylalanine, respectively (26, 27). These restricted the size of the A′-pocket of the two isoforms and reduced their capacity to bind long-chain lipid-based antigens (26, 27). Similarly, CD1d2 molecules efficiently presented only glycolipid antigens with short or no acyl chains. Of interest, α-psychosine, which contains a single lipid chain and has been proposed to possibly play a role in the selection of iNKT cells (52), was presented well by both CD1d1 and CD1d2. Although the lipid portions of antigens are buried in the CD1d groove, the composition of the fatty acids, including the length, the degree of unsaturation, and the presence of other modifications, strongly influences antigenic potency (59–61). As such, it is reasonable to hypothesize that the balance between various antigen species that might be presented by CD1d1 and/or CD1d2 in the thymus possibly affects the tuning of the iNKT TCR repertoire to different ligands. CD1d2 might also be important in shaping the repertoire of other CD1d-restricted T cells.

**Methods**

**Mice.** The CD1d1+/− (32), CD1d1−/−CD1d2+/− (30), and Jα18(neo) (36) mice have been described previously and were backcrossed >10 times on the C57BL/6 background. C57BL/6 and BALB/c were purchased from Jackson Laboratories. All mice were used between 6 and 12 wk and were age-matched.

Fig. 7. αGC(C10)-CD1d2 tetramer staining. (A) Thymocytes from BALB/c, C57BL/6, CD1d1−/− and Jα18−/− mice were stained with CD1d1 tetramers loaded (or not) with PBS57 or CD1d2 tetramers loaded (or not) with αGC (C10) and TCRβ mAbs. The percentage of tetramer−/− is shown. Data are representative of two independent experiments. (B) Thymocytes from the aforementioned mouse strains were first stained with αGC(C10)-CD1d2 tetramers, washed, and subsequently stained with PBS57-CD1d1 tetramers. Data are representative of three independent experiments. Simultaneous staining with both tetramers did not allow for the detection of the fluorophore associated with the CD1d2 tetramer.
for each experiment. All mice were raised in a specific pathogen-free environment at the Biology Resource Center at National Jewish Health (NJIH) and the Office of Laboratory Animal Research at the University of Colorado Denver (UCD). All animal procedures were approved by the NHU (protocol AS2780-10-16) and UC-D [protocol 06-4314(05)1E] Institutional Animal Care and Use Committees and were carried out in accordance with the approved guidelines.

**Lymphocyte Isolation.** Single-cell suspensions were prepared from the thymus, spleen, and liver by manual disruption by using a syringe plunger. The liver was perfused with PBS solution and then cut into small pieces, disrupted with a syringe plunger, and liver lymphocytes were isolated by centrifugation by using a 33% (vol/vol) Percoll gradient (GE Healthcare).

**High-Throughput Sequencing.** CD69 (“CD4” “CD8”) preselection double-positive thymocytes were sorted by flow cytometry and washed twice in ice-cold PBS solution. Total RNA was extracted by using the RNeasy Mini Kit (Qiagen). cDNA was made by using SuperScript III Reverse Transcriptase (Invitrogen). The Vβ-Cp region was amplified by using specific Vβ 5′-ATTATCTCGAGATCACC-3′ and Cj 5′-CCATGTGGGAGTAGACATTTCACTTC-3′ primers with required specific Ion Torrent tags. Purified PCR products were sent for high-throughput sequencing by using the Ion Torrent platform. Sequence analysis was done with in-house software, and gene identity was assigned on the basis of sequence alignment with published sequences (International ImMunoGeneTics Information System).

Alternatively, exon 2 of CD1d genes was amplified by PCR with Phusion high-fidelity polymerase (NEB) using specific primers (forward, 5′-TACACTCTGCCGTGTCGTTCTG-3′; reverse, 5′-AGCTCTGCAGTGCTGAGAAAC-3′) with required specific Ion Torrent tags. Purified PCR products were sequenced and buffer was exchanged again in 96-well round-bottomed plates. GFP expression on the hybridomas was analyzed by flow cytometry.

**CHB Transfectants.** cDNA encoding CD1d1 and CD1d2 molecules were amplified from B6 or CD1d1-/- spleenocytes by using the primers 5′-ATTATCTCGAGACCACTACGCACTTAC-3′ and 5′-ATTAAAGAATTCCTCAGAGGGTCTGTTGAAGAAGC-3′ and cloned into retroviral plasmids. Plasmids were sequence-verified, and retroviruses were prepared as previously described (8) and used to transduce CD1d-negative CHB cell lymphomas (64). Stably transduced cells were sorted for similar CD1d cell surface expression by flow cytometry by using an Aria cell sorter (BD Bioscience). The C168W CD1d1 mutant was generated by using QuikChange II Site-Directed Mutagenesis Kit following the manufacturer’s instructions. CHB cells expressing the C168W mutant were similarly generated.

**Cryocrystallization.** After data collection, the crystal structures were refined using CNS (www.cns.sbg.ac.at) (56). The model of CD1d2 was refined until convergence was reached for all components and all the non-crystallographic symmetry constraints were satisfied. The model was then submitted to an additional 10 cycles of conjugate gradient refinement. The final model of CD1d2 was used for calculations. The coordinates relating to the data reported in this study are deposited in the PDB database with ID codes 6BMH and 6BMK (CD1d2-αGalCer C10 structure, endogenous).

**Bioblot.** For the purpose of biotinylation, the purified CD1d2 was desalted and incubated eight parts of CD1d2 with one part of buffer A (100 mM Bicine, pH 8.3) and one part of buffer B (100 mM MgOAc, 500 μM o-biotin) along with 2 μL of protein ligase (1 mg/ml). The biotinylation reaction was carried out at room temperature overnight. The excess of free biotin was removed by size-exclusion chromatography by using a Superdex 200 10/30 gel filtration column.


