Structural Characterization of Mycobacterial Phosphatidylinositol Mannoside Binding to Mouse CD1d

Dirk M. Zajonc, Gary D. Ainge, Gavin F. Painter, Wayne B. Severn and Ian A. Wilson

*J Immunol* 2006; 177:4577-4583; doi: 10.4049/jimmunol.177.7.4577
http://www.jimmunol.org/content/177/7/4577

**References**  This article cites 54 articles, 18 of which you can access for free at: http://www.jimmunol.org/content/177/7/4577.full#ref-list-1

**Subscription**  Information about subscribing to *The Journal of Immunology* is online at: http://jimmunol.org/subscription

**Permissions**  Submit copyright permission requests at: http://www.aai.org/About/Publications/JI/copyright.html

**Email Alerts**  Receive free email-alerts when new articles cite this article. Sign up at: http://jimmunol.org/alerts
Mycobacterial phosphatidylinositol mannolipids (PIMs) are major components of the outer leaflet of the plasma membrane of mycobacteria and form the glycolipid core of the cell wall associated lipids lipomannan (LM) and lipoarabinomannan (LAM) (1). These glycolipids, together with many other lipids that are unique to mycobacteria (reviewed in Refs. 1 and 2), form the foundation of the mycobacterial cell envelope and provide a special lipid barrier that is responsible for the high resistance to external factors, such as cytotoxic oxygen radicals that are generated in the human macrophage upon infection.

In addition to their structural role in the cell envelope, many of these lipids are potent immunomodulators. LAM is a virulence factor that not only scavenges toxic oxygen-free radicals (3), but also inhibits IFN-γ-mediated activation of macrophages (4). Furthermore, LAM, LM, and PIM can all induce transcription of cytokine mRNA and elicit cytokine production, which indicates that most of their biological activity is associated with the PIM core, rather than with the complex carbohydrate cap (5). PIMs and LAM can be taken up directly by the mannose receptor (CD206) on myeloid dendritic cells (6) and are sorted into endosomes, where lipid transfer proteins and the acidic pH environment promote the loading of these glycolipids onto CD1 molecules for subsequent presentation to T cells (7).

CD1 receptors are a family of lipid-Ag presenting molecules that are structurally related to peptide-presenting MHC class I molecules. However, in contrast to the MHC-lineage, CD1 proteins are nonpolymorphic and feature a deep and hydrophobic binding groove, ideally suited for binding self- and foreign lipids, glycolipids, lipopeptides and other hydrophobic molecules (8). In humans, the various CD1 isoforms are divided into group 1 (CD1a, CD1b, CD1c, and CD1e) and group 2 (CD1d) based on sequence similarities (9), but mice only express group 2 CD1d. Whereas group 1 CD1 are recognized by conventional αβ+ or γδ+ T cells, the main population of group 2 CD1 molecules activate NKT cells (10), which express both an αβ+ TCR and typical markers, such as NK1.1. NKT cells are important in suppressing tumor growth and in regulating a variety of microbial, allergic and autoimmune conditions by rapidly secreting Th1 and Th2 cytokines (11).

Murine and human NKT cell populations are heterogeneous, but the majority of NKT cells express TCRs with an invariant α-chain (iNKT cells, Vα14-Jα18-Vβ8, Vα14-Jα18-Vβ7, or Vα14-Jα18-Vβ2 in mouse and Vα24-Jα18-Vβ11 in human). Various foreign ligands can activate iNKT cells, such as the nonmammalian glycolipid, α-galactosylceramide (α-GalCer) (12), microbial α-glycuronosylceramides (13, 14) (containing either glucuronic acid or galacturonic acid (GaLa-GSL)), mycobacterial PIM4 (15), as well as the self-ligand isoglobotrihexosylceramide (iGb3) (16). Crystal structures of CD1d in complex with α-GalCer or GaLa-GSL (17–19) have revealed the distinctive hydrogen-bonding network between the α-linked carbohydrate headgroup and CD1d, which has not
been seen in any other CD1 complex, including human CD1a-sulfatide (20), CD1a-lipopeptide (21), CD1b-P (22), CD1b-glucosemonoycylate (23), or mouse CD1d-phosphatidylcholine (CD1d-PC) (24).

Mycobacterial PIMs (PIM2, PIM4, and PIM6), and the more complex LAMs induce a complex immune response, some of which are initiated through presentation of these mycobacterial glycolipids by CD1b and CD1d (15, 25, 26). To address the question of how a natural bacterial Ag, PIM2, is presented by CD1d in comparison to other self-glycolipids, such as PC, and foreign and self-sphingolipids, such as αGalCer and sulfate, and how PIM4 is recognized by iNK T cells, we crystallized the complex of synthetic PIM2 bound to mouse CD1d and determined its three-dimensional structure to 2.6 Å resolution.

Materials and Methods

Synthesis of 2,6-(Di-O-α-D-mannopyranosyl)-1-O-(1’-2’-di-O-hexadecanoyl-s-n-glycero-3-phosphoryl)-n-myo-inositol (PIM2)

A synthetic sample of PIM2 was prepared from 1-0-allyl-3,4,5-tri-O-benzyl-n-myo-inositol (27) that was converted into the advanced intermediate, 3,4,5-tri-O-benzyl-2,6-di-O-(2,3,4,6-tetra-O-benzyl-α-D-mannopyranosyl)n-myo-inositol, by synthetic methods described elsewhere (28).

Triethylammonium 3,4,5-tri-O-benzyl-2,6-di-O-(2,3,4,6-tetra-O-benzyl-α-D-mannopyranosyl)-1-O-(1’-2’-di-O-hexadecanoyl-s-n-glycero-3-phosphoryl)-n-myo-inositol

A mixture of 3,4,5-tri-O-benzyl-2,6-di-O-(2,3,4,6-tetra-O-benzyl-α-D-mannopyranosyl)n-myo-inositol (230 mg, 0.154 mmol) and triethylammonium 1,2-di-O-hexadecanoyl-sn-glycero-3-phosphonate (223 mg, 0.304 mmol) were dried by evaporation with pyridine (2 ml) and redissolved in pyridine (5 ml) under argon. Pivaloyl chloride (230 mg, 1.56 mmol) was added and the reaction stirred for 30 min. when a freshly prepared solution of iodine (150 mg, 0.580 mmol) in pyridine (5 ml) was added and the mixture stirred for an additional 10 min. CHCl3 (100 ml) was added and the mixture stirred for an additional 10 min.

The organic layer was washed with Na2S2O3 (10% aqueous, 100 ml) and CHCl3 (100 ml) was added and the mixture stirred for additional 10 min. The organic layer was washed with Na2S2O3 (10% aqueous, 100 ml) and CHCl3 (100 ml) was added and the mixture stirred for an additional 10 min. The organic layer was washed with Na2S2O3 (10% aqueous, 100 ml) and CHCl3 (100 ml) was added and the mixture stirred for an additional 10 min.

The mixture was filtered through Celite and purified on a silica gel eluting with CHCl3 and CHCl3/MethOH/H2O (70:4:0.4) to afford PIM2 (32 mg, 0.028 mmol, 67%) as a white powder after lyophilization from H2O. α-pH = +35.2 (c. 0.25, CDCl3/CDOD2/D2O 70:40:6); δ H NMR (300 MHz, CDCl3, δ 7.42–6.92 (m, 55H), 5.88 (br s, 1H), 5.64 (br s, 1H), 5.25 (br s, 1H), 4.86–3.75 (m, 48H), 3.37–3.12 (m, 48H), 2.80–2.40 (br, 48H), 2.16–2.08 (m, 48H), 1.53–1.40 (m, 48H), 1.32–1.21 (m, 48H), 0.89–0.80 (m, 48H); δ 31P NMR (CDCl3, 121.5 MHz) δ = 0.08; MS-ES (M+H)+ calculated for C130H164O23P: 2124.1407, found 2124.2895.

Structure determination

Crystals were flash-cooled at 100 K in mother liquor containing 25% glycerol. Diffraction data from a single crystal were collected at Beamline 8.2.1 of the Advanced Light Source and processed to 2.6 Å with the Denso- Scalepack suite (30) in spacegroup P2221 (unit cell dimensions: a = 41.71 Å, b = 40.06 Å, c = 107.07 Å). One CD1-lipid complex occupies the asymmetric unit with an estimated solvent content of 52.2% based on a Matthews’ coefficient (Vm) of 2.57 Å3/Da. Molecular replacement in P212221 was carried out in CCP4 (31) using the program MOLREP (32) and the CD1d-sulfatide structure (2AKR) as the search model, with the ligand removed, and resulted in a Rfree of 43.5% and a correlation coefficient of 0.55. Subsequent rigid-body refinement in REFMAC 5.2 produced an Rcryst of 37.5%.

The initial refinement included several rounds of restrained refinement against the maximum likelihood target in REFMAC 5.2. At a later stage of refinement, carbohydrates were built at all three N-linked glyco-sylation sites in CD1. Refinement progress was judged by monitoring the Rfree for cross-validation (33). The model was rebuilt into σA-weighted 2Fo-Fc and Fc − Fo difference electron density maps using the program O (34). Water molecules and additional detail were added during refinement using the water ARD module for >3σ peaks in an Fo − Fe map and retained if they satisfied hydrogen-bonding criteria and returned 2Fo − Fc > 1σ after refinement. Starting coordinates for the PIM2 ligand were obtained from the CD1b-Pi structure (1GQZ) and modified accordingly, using the molecular modeling system INSIGHT II (Accelrys). The PIM2 library for REFMAC (35) was created using the Dundee PRODRG2 server (36). Final refinement steps were performed using the translation, libration, and screw rotation displacement procedure in REFMAC (37) with a total of three anisotropic domains (α1–α3 domains and β-M) and resulted in improved electron density maps for the glycolipid ligand and an additional drop in Rfree. The higher than usual Rfree (14.3%) is mainly due to the weakness of the overall intensity of the diffraction data (I/σI) from a very thin needle-like crystal. However, the electron density maps were clearly interpretable and the structure refinement was straightforward.

The CD1d-PIM2 structure has a final Rcryst = 21.2% and Rfree = 28.0%, and the quality of the model (Table I) was excellent as assessed with the program Molprobity (38).

Structure presentation

The program PyMol (http://pymol.sourceforge.net/) was used to prepare all of the figures. The PDB2PQR server (39) and the program APBS (40) were used to calculate the electrostatic surface potentials. The Chimera program was used to prepare all chemical structures.

Results

Structure determination of the CD1d-PIM2 complex

Soluble mouse CD1d-β-M protein (residues 1–279 H chain and 1–99 β-M) was expressed by Spodoptera frugiperda (SF9) cells and purified to homogeneity, as described in the methods section. Synthetic PIM2 (2,6-[Di-O-α-D-mannopyranosyl]-1-O-[1,2-di-O-palmitoyl-sn-glycero-3-phosphoryl]-n-myo-inositol) was loaded onto CD1d by incubation with a 7-fold molar excess of glycolipid without detergent. The CD1d-PIM2 complex was further purified from free lipid and protein aggregates by size exclusion chromatography. The CD1d-PIM2 complex was crystallized by sitting drop vapor diffusion and its three-dimensional structure determined by molecular replacement using the protein coordinates of...
This cavity is lined mainly by hydrophobic residues (Fig. 2), which represent the intensity of i symmetry-related observations for reflections with Bragg index h.

Specific interactions between CD1d and PIM2

PIM2 has a more elaborate headgroup than the ligands of previously crystallized CD1d complexes and contains two additional mannoses that are connected by α-1–2 and α-1–6 linkages to the inositol of PI. Surprisingly, the entire headgroup is ordered in the crystal structure as a result of seven specific hydrogen bond interactions between CD1d and the polar groups of PIM2 (Fig. 4). Among these CD1d residues, Asp153, Thr156, Arg79, and Ser76 are also involved in stabilizing other CD1d ligands, such as short-chain α-GalCer, GalA-GSL, and sulfatide (17, 18, 29) but, in this case, Thr159 confers additional specificity for the 2’-mannose of PIM2. Interestingly Asp80, which is usually important for defining the orientation of the sphingolipid backbone, due to formation of specific hydrogen bonds with the 3’-OH group of GalA-GSL and the 3’-OH and 4’-OH groups of α-GalCer, does not interact here with the carbonyl group of the sn2-linked palmitic acid. Arg79 neutralizes the phosphate, while simultaneously stabilizing the 6’-mannose of PIM2 (Figs. 4 and 5A). The phosphate is pulled toward the α1 helix by Arg79 in such a way that the inositol flips back toward the α2 helix to interact with Asp153, while still remaining centered in the binding groove (Figs. 4 and 5A).

Comparison of different CD1d ligands

The complex inositol-dimannoside headgroup of PIM2 extends from the center of the binding groove toward the N terminus of the α1 helix, so as to sit above the entrance to the A’ pocket (Fig. 5A). The headgroup is further aligned between CD1d sulfatide, in which the β-linked galactose is bound by Asp153 in a similar general orientation to the inositol in PIM2 (Fig. 5B). The 2’ mannose sits above the entrance to the A’ pocket, thereby blocking any access by the TCR to the underlying alkyl chain. This situation contrasts with the CD1d-sulfatide structure, in which the fatty acid chain is accessible for recognition by specific NKT cell TCRs. Such recognition was confirmed recently from a specific, sulfatide-restricted, NKT cell hybridoma that can distinguish between lysosulfatide (without any
fatty acid) or sulfatide (with a C16-C24 fatty acid) (29). The head-group of α-linked glycosphingolipids, such as short-chain αGal-Cer (Fig. 5C) or GalA-GSL (data not shown) mainly interacts with the α2 helix, whereas the more complex PIM2 is bound in a completely different orientation (Fig. 5, A and D), such that the 6'-mannose interacts closely with the α1 helix and the 2'-mannose is in a pocket between the α1 and α2 helices. Interestingly, all three ligands appear to activate Vα14T1 iNKT cells (12, 14, 15).

The overall binding of PIM2 and PC to CD1d is very different. Although the central, negatively charged phosphate of the two phosphoglycerolipids, is effectively neutralized by Arg79, PIM2 sits deeper in the CD1d binding groove and, in addition, is rotated 180°, compared with PC in the binding groove (Fig. 5E). As a result, PIM2 inserts its sn1-fatty acid into the A pocket and sn2-fatty acid into the F' pocket in contrast to the PC structure (24), but the glycerol backbone is oriented differently. The alkyl chain at sn1 position of the PIM2 glycerol is inserted into the A' pocket, whereas the sn2-alkyl chain (sn2) fully fills the F' pocket.

**FIGURE 1.** Overview of the CD1d-PIM2 structure. A, Schematic representation (front view) of the CD1d (α1-α3 domains in gray)-β2M heterodimer (blue-gray) with β-strands and α-helices highlighted and with bound phosphatidylinositol-dimannoside ligand in stick representation (yellow). N-linked carbohydrates are depicted as gray stick models attached to N20, N42, and N165. Atom colors for all structural representations for Figs. 1–4 are yellow/orange/cyan, carbon; red, oxygen; blue, nitrogen; orange, sulfate; purple, phosphate. B, Chemical structure of the synthetic PIM2 used in this study and an abundant species of natural PIM4 (56) for comparison. As the PIM2 phosphate has a pK_{a} ~2, it is negatively charged in the crystal structure at pH 5.5 and at physiological pH. Carbon atoms 1 (sn1) and 2 (sn2) of glycerol that are esterified with fatty acyl moieties are indicated with the length of the individual alkyl chains given as the number of carbon atoms (C16 palmitate acid; CH3-C18 tuberculostearate). The linkages between the individual mannose residues (α1-2 and α1-6) are labeled. Note that the natural tuberculostearic acid of PIM4 has a methyl substituent at carbon 10 of the sn1-linked fatty acid. The arrow in A indicates where the additional mannoses of PIM4 would be connected to PIM2.

**FIGURE 2.** Conformation of PIM2 in the CD1d binding groove. A side view is shown along the CD1d binding groove with the α2 helix removed for clarity. After excluding the ligand coordinates, a shake-omit map was calculated as a Fo-Fc electron density map as previously reported (21) and contoured at 2σ as a blue mesh around the ligand (yellow). Several key contact residues that are involved in ligand binding are depicted and labeled. Note how Arg79 (R79) is positioned to neutralize the negative charge of the phosphate group of PIM2. The alkyl chain at sn1 position of the glycerol (sn1) is inserted into the A' pocket of the binding groove, whereas the sn2-alkyl chain (sn2) fully fills the F' pocket.

**FIGURE 3.** Modeling of natural PIM4 ligand binding. Natural PIM4 (yellow sticks) was modeled into the mouse CD1d binding groove (for electrostatic surface representation, see Fig. 5). A, The high degree of exposure of the PIM4 headgroup is shown. B, Tuberculostearic acid (R-10-methyloctadecanoic acid) can easily be accommodated in the A' pocket. Man, mannose residues; ino, inositol. CD1d residues Asp153 (D153) and Arg79 (R79) are labeled to provide orientation for comparison with Fig. 5D.
where the sn1-fatty acid occupies the F′ pocket. However, PIM2 cannot insert its sn1-fatty into the F′ pocket as it sits deeper in the groove, due to the extra carbon of the glycerol, and would exceed the maximum length of C16 that can optimally fit in this pocket. In the CD1d-PC structure, only a C12 fatty acid was observed in the F′/H11032 pocket. On the other hand, the CD1d residues that contact these disparate PIM2 and PC ligands adopt similar orientations.

Next, we compared the binding of PIM2 as a glycerolipid to glycosphingolipids, such as sulfatide (Fig. 5F). As expected, the additional phosphate group between the diacylglycerol backbone of PIM2 slightly raises the inositol-dimannoside headgroup above the binding groove, while it retains the same lateral disposition as the galactose of the sulfatide. The lipid backbones of these ligands do not superimpose well, mainly as a result of the different geometry of the glycerol in PIM2 vs the N-amide linkage of the fatty acid with the sphingosine of sulfatide or α-GalCer. In conclusion, three conserved polar interactions between the sphingolipid backbones of α-GalCer and sulfatide with CD1 residues Arg79, Asp80, and Thr156 orient the glycosphingolipids in a similar manner to each other in the CD1d binding groove (17, 18, 29), whereas the glycerolipid backbones of PC and PIM2 seem to be more flexible and differentially bound by CD1d. As a result, the five hydrogen bond interactions between PIM2 and CD1d (Fig. 4) dominate the binding orientation of the glycolipid in the CD1 binding groove and are responsible for the fine-positioning of the headgroup (Fig. 4).

![FIGURE 4.](http://www.jimmunol.org/) Stereoview of the hydrogen-bond network between PIM2 and CD1d. The PIM2 ligand bound in the CD1d binding groove is shown in a rear view from the C-terminal end of the α1 helix. Hydrogen bonds between the protein and the ligand residues are depicted as blue dashed lines and range from 2.5 to 3.5 Å. The α1 helix residues Arg79 (R79) and Ser76 (S76) form hydrogen bonds with the phosphate moiety of the ligand and Arg79 also interacts with the 6′-mannose. Asp153 (D153) of the α2 helix hydrogen bonds to the central inositol ring, while Thr156 (T156) stabilizes the 2′-mannose. Thr156 (T156) forms the only obvious hydrogen-bond interactions with the glycerolipid backbone, whereas Asp80 (D80) does not seem to interact with the ligand, as it does with α-GalCer, GalA-GSL, or sulfatide (24, 25, 32), although it does orient and stabilize the Arg79 conformation.

![FIGURE 5.](http://www.jimmunol.org/) Comparative binding analysis of different CD1d ligands. The CD1d binding pocket is shown as a molecular surface with electrostatic potential (electronegative in red and electropositive in blue from −30kT/e to 30kT/e) in A–D, in a top view, with bound PIM2 (A), with sulfatide (PDB code 2AKR) (B), or with the short-chain α-GalCer variant PBS-25 (PDB code 1ZSL) (C). D. For PIM2, the 2′-mannose packs nicely against the α2 helix, whereas the 6′-mannose sits above Arg79. In PBS-25, the galactose orients parallel to the surface of the α1 and α2 helices (C), whereas sulfatide (B) and PIM2 (A and D) are more exposed, with PIM2 having the most accessible surface area for T cell recognition. E and F, Schematic comparison of the ligand binding site of CD1d in a side view, with the α2 helix removed for better visualization. E, Binding of PIM2 (ligand yellow, CD1d gray) vs PC (ligand green, CD1d gray-blue). F, Comparison of PIM2 vs sulfatide (ligand cyan, CD1d dark green). This panel is slightly rotated, compared with E, to better visualize the ligand headgroups.
Discussion

Tuberculosis is a global health threat, with a total of 3 billion infected people and 2 million attributed deaths each year (44). Current vaccination strategies fail to protect adults effectively from pulmonary tuberculosis and the causative agent Mycobacterium tuberculosis is often unresponsive to many antibiotics due to its waxy and impenetrable cell envelope. Upon phagocytosis of the microbe by macrophages, cell wall lipids, such as PIMs and LAMs, are released into the lumen of phagosomes and subsequently trafficked into downstream endosomes and finally into lysosomes (45). Along this route, these lipids can be detected by the immune system as they interact with CD1 molecules (46), which constantly survey the lipid content of APCs (47, 48) for subsequent activation of T cells.

PIM4 was identified recently as an agonist of a subset of human and mouse NKT cells, when presented by CD1d (15). CD1d tetramer-hapten-staining experiments show that only 0.26% of all mouse liver lymphocytes react to PIM4, while 32% of all liver lymphocytes can be stained with CD1d/α-GalCer tetramers. However, 25% of α-GalCer reactive, mouse spleen NKT cells can be activated by PIM4 and subsequently secrete similar levels of IFN-γ (49), and the low number of activated NKT cells in mice, the biological relevance for PIM4 as a mouse NKT cell agonist requires further investigation.

Comparison of the binding and exposure of the various ligand headgroups in the CD1d binding groove allows several predictions to be made regarding interaction between ligand and its specific TCR. With the assumption that the iNKT cell TCR engagement to CD1d will involve direct contacts with the invariant α-chain, particularly through the variable loop of the CDR 3 (CDR3α) (Fig. 5). This proposed interaction correlates strongly with the observed Vα14 restriction, whereas the TCR β-chain can vary without greatly compromising ligand recognition (11). However, the β-chain, especially CDR3β, would be ideally situated to interact with the additional two α–6-linked mannoeieties in PIM4 and, hence, is likely to be involved in discrimination among the different PIM species (PIM2–PIM6).

NKT cells have an interspecific cross-reactivity, as mouse NKT cells can recognize α-GalCer when presented by mouse CD1d or human CD1d and vice versa. When the crystal structures of α-GalCer bound to either mouse or human CD1d are compared, it is apparent that the highly restricted iNKT cell TCR must have some plasticity that compensates for the slightly different presentation of α-GalCer in either species (52). The crystal structures of four human Vα24 TCRs were recently determined (53, 54) and reveal a large surface cavity between the CDR3 loops of the α- and β-chains. Depending on the degree of TCR plasticity, this cavity could change in size and shape to accommodate the different headgroups of the CD1d ligands. In the case of PIM4 or iGb3, these changes must be substantial as their respective headgroups differ greatly in size (α-GalCer, one sugar; iGb3, three sugars; and PIM4, four sugars and one inositol).

For CD1a and CD1b, the stability of the CD1-lipid complex was previously correlated with the extensive number of nonpolar van der Waals’ interactions between the two alkyl chains of the glycolipid and the hydrophobic residues in the CD1 binding groove (lipid anchoring), rather than with specific hydrogen-bond interactions to the glycolipid headgroup (20–23). For mouse and human CD1d, all of the complex crystal structures so far have revealed a number of specific polar interactions and an extensive hydrogen-bond network between the headgroup and conserved CD1 residues that likely increase binding specificity and stability of the complex (17–19, 29). In light of these findings, it seems plausible that, at least for PIMs, only one alkyl chain is necessary to recruit NKT cells (55). The electron density of the CD1d-PIM2 complex structure reveals that the headgroup is much more ordered, compared with those in glycolipids or lipopeptides bound to CD1b or CD1a, respectively (20–23). Thus, we could speculate that, in addition to the observed lipid anchoring, the polar interaction between CD1d and the headgroup take on increasing importance for forming a stable glycolipid complex and in the fine-positioning of the antigenic epitopes for TCR recognition, which could be of significance for development of novel immunomodulatory agents.

Acknowledgments

We thank the staff of the Advanced Light Source, BL 8.2.1, for support with data collection and Julie Vanhnasy for technical assistance. This is manuscript 18058-MB of The Scripps Research Institute.

Disclosures

The authors have no financial conflict of interest.

References


