The EMBO Journal (2006) 25, 3684–3692 | © 2006 European Molecular Biology Organization | All Rights Reserved 0261-4189/06
www.embojournal.org

THE EMBO JOURNAL

Endogenous phosphatidylcholine and a long spacer ligand stabilize the lipid-binding groove of CD1b

Luis F Garcia-Alles[1,*], Kees Versluis[2], Laurent Maveyraud[1], Ana Tesouro Vallina[1], Sebastiano Sansano[3], Nana Fatimath Bello[1], Hans-Jürgen Gober[3], Valérie Guillet[1], Henri de la Salle[4,5,6], Germain Puzo[1], Lucia Mori[3], Albert JR Heck[2], Gennaro De Libero[3] and Lionel Mourey[1,*]

[1]CNRS, UMR5089, Département Mécanismes Moléculaires des Infections Mycobactériennes, Institut de Pharmacologie et de Biologie Structurale, Toulouse, France, [2]Department of Biomolecular Mass Spectrometry, Utrecht University, Utrecht, The Netherlands, [3]Experimental Immunology, Department of Research, Basel University Hospital, Basel, Switzerland, [4]INSERM, U 725 'Biologie des Cellules Dendritiques Humaines', Strasbourg, France, [5]Université Louis Pasteur, Strasbourg, France and [6]Etablissement Français du Sang-Alsace, Strasbourg, France

CD1 proteins present lipid antigens to T cells. The antigens are acquired in the endosomal compartments. This raises the question of how the large hydrophobic CD1 pockets are preserved between the moment of biosynthesis in the endoplasmic reticulum and arrival to the endosomes. To address this issue, the natural ligands associated with a soluble form of human CD1b have been investigated. Using isoelectric focusing, native mass spectrometry and resolving the crystal structure at 1.8 Å resolution, we found that human CD1b is simultaneously associated with endogenous phosphatidylcholine (PC) and a 41–44 carbon atoms-long spacer molecule. The two lipids appear to work in concert to stabilize the CD1b groove, their combined size slightly exceeding the maximal groove capacity. We propose that the spacer serves to prevent binding of ligands with long lipid tails, whereas short-chain lipids might still displace the PC, which is exposed at the groove entrance. The data presented herein explain how the CD1b groove is preserved, and provide a rationale for the *in vivo* antigen-binding properties of CD1b.

The EMBO Journal (2006) **25,** 3684–3692. doi:10.1038/sj.emboj.7601244; Published online 27 July 2006
Subject Categories: immunology; structural biology
Keywords: antigen; CD1 proteins; crystal structure; endogenous lipid; mass spectrometry

*Corresponding authors. LF Garcia-Alles, Département Mécanismes Moléculaires des Infections Mycobactériennes, Institut de Pharmacologie et de Biologie Structurale, CNRS UMR5089, 205 Route Narbonne, Toulouse, Haute-Garonne 31077, France.
Tel.: +33 5 61 17 54 40; Fax: +33 5 61 17 59 94;
E-mail: Luis-Fernando.Garcia-Alles@ipbs.fr or
L Mourey, Département Mécanismes Moléculaires des Infections Mycobactériennes, Institut de Pharmacologie et de Biologie Structurale, CNRS UMR5089, 205 Route Narbonne Toulouse, Haute-Garonne 31077, France. Tel.: +33 5 61 17 54 36; Fax: +33 5 61 17 59 94;
E-mail: Lionel.Mourey@ipbs.fr

Received: 16 March 2006; accepted: 26 June 2006; published online: 27 July 2006

Introduction

CD1 molecules, like MHC proteins, are specialized in the presentation of antigens to T-cell receptors (TCRs). MHC molecules display peptide antigens, whereas CD1 molecules present lipid antigens of both self and microbial origin (Brigl and Brenner, 2004; De Libero and Mori, 2005). In humans, the CD1 family includes group I CD1a, CD1b and CD1c, group II CD1d, and the CD1e protein. Reminiscent of the common antigen-presenting role, the structures of CD1 and MHC class-I molecules are similar (Moody *et al*, 2005). Nevertheless, major divergences are observed in the arrangement of residues that interact with the presented antigens. Thus, the antigen-binding grooves of CD1 molecules are deeper, and more voluminous and hydrophobic than those of MHC proteins.

One of the most intriguing aspects of the biology of CD1 molecules is their intracellular distribution and trafficking. After biosynthesis in the endoplasmic reticulum (ER) and transfer to the plasma membrane, CD1 molecules are internalized and delivered to endocytic compartments, before being recycled to the cell surface for presentation of specific antigens to T cells (Moody and Porcelli, 2003). For instance, the encounter of CD1b with microbial lipids occurs in the late endosomes–lysosomes. Lipid transfer proteins (LTPs) and hydrolytic enzymes present in these compartments have been shown to edit the CD1-bound lipids and promote the processing of some of the antigens, before presentation to T cells (Prigozy *et al*, 2001; Kang and Cresswell, 2004; Winau *et al*, 2004; Zhou *et al*, 2004; de la Salle *et al*, 2005).

Because antigen-loading occurs in the endocytic network, the question arises as to how the hydrophobic grooves of CD1 molecules are protected after biosynthesis. ER-resident LTPs, for example, microsomal triglyceride transfer protein (MTP) (Brozovic *et al*, 2004; Dougan *et al*, 2005), might transfer endogenous lipids to the nascent CD1 pockets. These lipids are required for maintaining the integrity of the grooves, but might also have functional relevance. Consistent with this, chemical inhibition of MTP or silencing of MTP expression was found to diminish the CD1d-mediated presentation of antigens to NKT cells (Dougan *et al*, 2005). Concerning the identity of the lipids that associate with nascent CD1 molecules, two studies pointed to phosphatidylinositol (PI) and glycosyl-PI (GPI) as candidates (Joyce *et al*, 1998; Park *et al*, 2004). However, the acyl chains of these phospholipids are too small to completely cover the enormous binding groove of CD1b, which can accommodate hydrophobic chains of about 70 carbons (Gadola *et al*, 2002; Batuwangala *et al*, 2004).

Understanding how CD1 molecules are assembled in cells will facilitate our comprehension of their ligand selectivity, loading kinetics and perhaps of their intracellular trafficking requirements. In this work, we have addressed the identification of the endogenous molecules present in the antigen-

binding groove of recombinant soluble human CD1b (shCD1b), the CD1 member with the largest binding groove. This is achieved using a combination of isoelectric-focusing electrophoresis (IEF), native mass spectrometry (MS) and by determining the protein crystal structure at high resolution.

Results

Non-charged natural ligand in the hCD1b pocket

The N-glycosylated extracellular domain of hCD1b was expressed together with human β2-microglobulin (β2m) in mouse cells and purified by anti-CD1b affinity chromatography (Figure 1A). In order to prevent damage to the protein or to its lipid content, elution was carried out with a citrate buffer (pH 5.0) supplemented with 50% ethylene glycol. Lipid binding by this recombinant soluble protein (shCD1b) was then studied by IEF, a useful technique for monitoring the formation of protein–lipid complexes, provided that the global isoelectric point changes upon binding of the ligand (Cantu *et al*, 2003; Zhou *et al*, 2004). The migration pattern of shCD1b could only be altered by the addition of negatively charged lipids, such as PI and a biotinylated derivative of phosphatidylethanolamine (PE-biotin) (Figure 1B). Phosphatidic acid or lipids with no net charge like phosphatidylcholine (PC) or phosphatidylethanolamine (PE) did not induce any band shift. The same behavior was observed for shCD1b expressed in medium supplemented with serum that had been delipidated with 1-butanol/diisopropylether (40:60, vol/vol) (Cham and Knowles, 1976), and for shCD1b deglycosylated with either the peptide-*N*-(*N*-acetyl-β-glucosaminyl) asparagine amidase F or with the endo-β-*N*-acetylglucosaminidase F_3 (endoF_3) (data not shown). These results seem incompatible with the presence of PI or GPI species in the purified protein, in contradiction of what has been proposed (Joyce *et al*, 1998; Park *et al*, 2004). Reported IEF data obtained on insect-expressed mouse CD1d also indicate that its binding groove is not loaded with PI or any other negatively charged lipid (Cantu *et al*, 2003; Zajonc *et al*, 2005b).

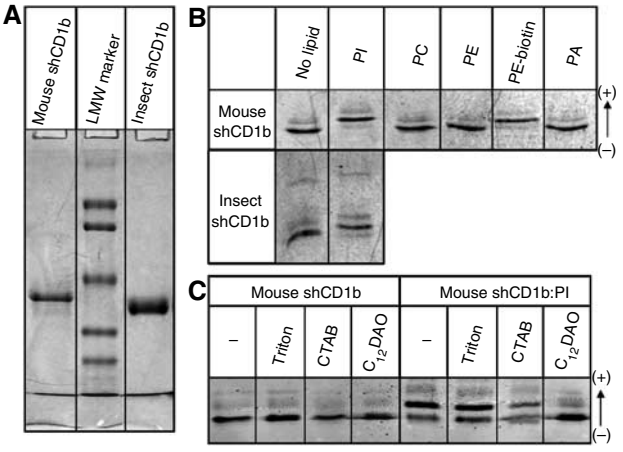

Figure 1 IEF analysis of shCD1b lipid binding. (**A**) SDS–PAGE gel (12%) of purified shCD1b expressed in mouse cells (first lane) or in insect cells (third lane). Note that shCD1b heavy chains are glycosylated. The middle lane contains the molecular weight marker: from top to bottom 108, 90, 50.7, 35.5, 28.6 and 21.2 kDa. (**B**) IEF gel of mouse-expressed shCD1b (top) and insect-expressed shCD1b (bottom) after incubation with 50 μM of the indicated lipids. (**C**) IEF gel after incubation of mouse-expressed shCD1b with no detergent (first lane), or with Triton X-100, CTAB or C_{12}DAO. The four lanes to the right are from a similar experiment using mouse-expressed shCD1b that had been preloaded with PI and purified by size-exclusion chromatography.

The possibility that charged lipids occupied the binding pocket was further ruled out because no IEF change was observed upon incubation of shCD1b with detergents, at doses above their critical micelle concentration (Figure 1C, first 4 lanes). Detergent-binding competition and/or transfer of the CD1b-bound lipid to the detergent micelles would have been expected to modify the IEF position of the protein. This is what we observed when the complex shCD1b:PI, prepared by incubation of shCD1b with PI and purified by size-exclusion chromatography, was incubated with the same detergents (Figure 1C, last four lanes). This complex could be reverted to the unloaded protein by addition of C_{12}DAO. Interestingly, the shCD1b:PI complex was stable enough as to persist in the presence of Triton X-100 (Figure 1C) or taurocholate (data not shown).

To confirm that our observations are independent of the expression system, a second shCD1b construct was expressed in *Drosophila melanogaster* S2 cells. The soluble protein was purified by Ni^{2+} affinity chromatography. This preparation, which looked pure in SDS gels (Figure 1A), resolved on native IEF gels as three protein bands (Figure 1B). Nevertheless, all the bands shifted towards the anode upon incubation with PI, similar to the mouse-expressed protein. This confirms that PI is not the endogenous lipid of shCD1b expressed in insect cells. That all bands are displaced in the same manner suggests that the heterogeneity of this preparation comes from the presence of several protein species (e.g. caused by limited proteolysis of the protein) rather than from the presence of different lipid cargoes in the binding groove.

The next experiment was performed to discard the possibility that PI was originally present in shCD1b but was lost through the purification protocol. Mouse-expressed shCD1b was purified and then incubated overnight with radioactive [^{3}H]PI. The protein–lipid complex was separated from the lipid excess by size-exclusion chromatography. After pooling the protein fractions, the specific radioactivity of the protein–lipid complex was estimated to be 9500 d.p.m./nmol. The pooled solution was then injected into the anti-CD1b affinity column. After elution of the protein using our standard conditions, the collected material still presented 6000 d.p.m./nmol, that is, 63% of the original specific radioactivity. Indeed, the recovery is probably higher, considering that the protein pool loaded into the anti-CD1b column was probably contaminated with small amounts of the [^{3}H]PI liposomes that elute from the size-exclusion column just before the protein peak.

Mass spectrometry analysis of shCD1b

To date, MS has been used thoroughly to identify the lipid cargo of CD1 molecules (Joyce *et al*, 1998; Park *et al*, 2004; Giabbai *et al*, 2005; Wu *et al*, 2006). The usual protocol consists of an organic extraction of purified CD1 preparations, followed by MS analysis of the lipid extract. A major concern of this approach is that the results obtained might not reflect the real lipid scenario, depending on (i) the presence of potential lipid-associated proteins contaminating the CD1 preparation, (ii) the solubility of the lipids in the

extraction solvent and (iii) the concordance between the ionization capabilities of the extracted lipids and the MS ionization technique selected. These problems can be circumvented by directly detecting the CD1b:lipid non-covalent complexes by native ESI-MS. No extraction is required, and the direct detection of the multi-ionized protein–lipid complexes ensures both the detection of all lipid ligands, independently of their ionization efficiencies (van den Heuvel and Heck, 2004; Benesch and Robinson, 2006), and the discrimination of signals originating from contaminating proteins. In the present work, we performed native MS measurements as an attempt to better characterize the endogenous lipid content of hCD1b.

The native ESI-mass spectrum for endoF_3-deglycosylated shCD1b with its endogenously loaded lipid cargo is shown in Figure 2A. From the charge distributions observed, at least four different species could be extracted, with average molecular masses (in Da) of 47 645 ± 7 (D), 47 047 ± 2 (C), 46 858 ± 4 (B) and 46 232 ± 2 (A). A peak at 11 733 ± 1 Da could be attributed to β2m (not shown), whose theoretical molecular weight is 11 731 Da. The species at 46 232 Da fit in with the calculated mass of 46 226 Da for the lipid-free shCD1b (i.e. the CD1b:β2m heterodimer), assuming that Fuc-α(1-6)-GlcNAc moieties, which are not removed by the endoF_3, are anchored to the three asparagines predicted to be N-glycosylated. The other ion signals are significantly broader, consistent with the heterogeneity of ligands present in these species (see below and Supplementary Figure S1). Nevertheless, the data presented hereunder permit to assign these signals to shCD1b:PC:UL (D), shCD1b:PC (C) and shCD1b:UL (B), with UL standing for unknown ligand.

The spectrum of Figure 2A was deconvoluted, averaged over all charged states, to a neutral mass spectrum (Figure 2B, black line). This spectrum was recorded using a relatively harsh desolvation energy at the interface of the ESI source, with the intention to sharpen the ion signals and record the masses with higher accuracy. These conditions may also lead to dissociation of non-covalent complexes during ionization. For this reason, the experiment was repeated at lower desolvation energy. This led to a deconvoluted neutral mass spectra with significantly broader peaks (Figure 2B, gray line), but higher relative abundance of the shCD1b:PC:UL complex. Reducing the desolvation also increased the peak masses by about 100–125 Da. At the lowest desolvation energy possible, only ions originating from this shCD1b:PC:UL were observed, albeit with very broad ion signals (data not shown). This indicates that the only species present in solution is the shCD1b:PC:UL complex, in a 1:1:1 stoichiometry. The same conclusion was reached for shCD1b in the glycosylated state (data not shown).

To further verify that the 47 645 Da molecular mass was from the shCD1b:PC:UL complex, the 14+ ions of this species were subjected to tandem MS (Figure 2C). The precursor ion at approximate *m/z* 3400 dissociates to give strong ion signals corresponding to (i) the neutral loss of around 620 Da (indicated in Figure 2C by −UL, 14+), (ii) the loss of a positively charged molecule of around 780 Da (−PC, 13+) and (iii) the loss of a singly charged mass of 1400 Da (−UL, −PC, 13+). In addition, the molecular ions from PC and UL could be directly observed in the low *m/z* region (inset to Figure 2C), but not the 886 *m/z* PI peaks observed in a previous study on hCD1b (Park *et al*, 2004). The data

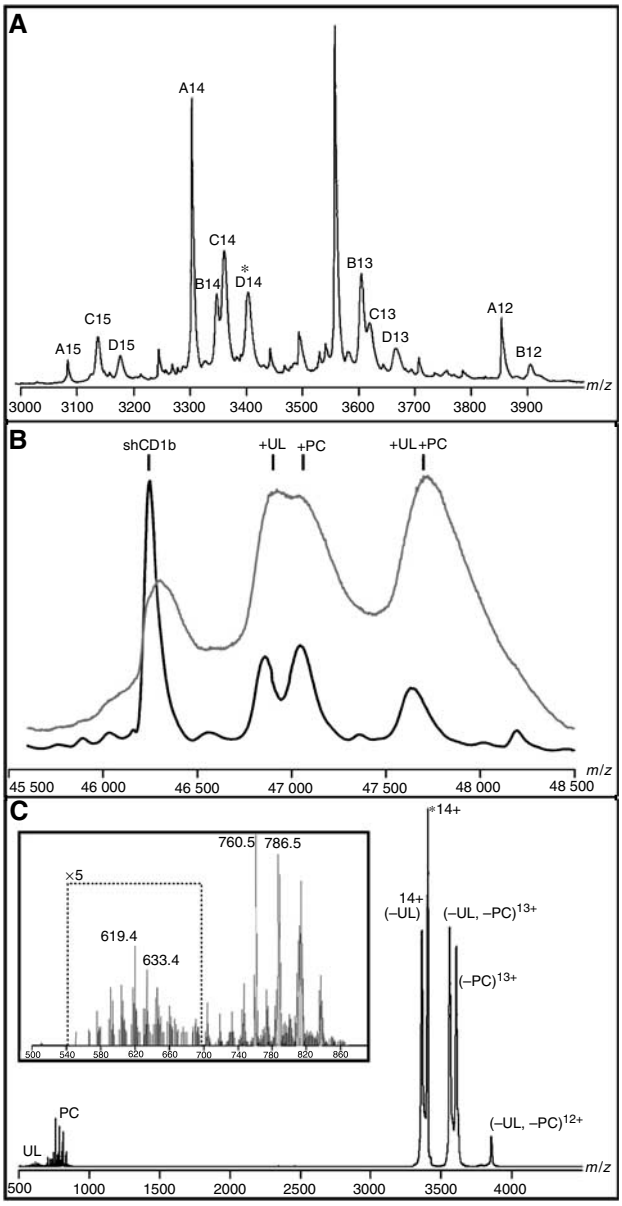

Figure 2 MS detection of natural ligands associated to shCD1b. (**A**) Positive-ion mode native ESI-MS from mouse-expressed endoF_3-deglycosylated shCD1b. Ion peaks arising from shCD1b, shCD1b:UL (unknown ligand), shCD1b:PC (phosphatidylcholine) and shCD1b:UL:PC species are labeled A, B, C and D, respectively. The labels are completed with a number indicative of their multi-charged state. (**B**) Neutral mass spectrum obtained by deconvolution, averaging over all charged states, of the spectrum of panel A. This spectrum was recorded using harsh desolvation energy (black line). A second deconvoluted spectrum was obtained by repeating the experiment at lower energy (gray line). The lower desolvation of the protein induces a broadening of the signals and a 100–150 Da displacement towards higher values. (**C**) The D14 precursor ion of panel A (indicated with an asterisk) was selected and subjected to tandem MS. Three other major species appear in the high-mass range, originating from the loss of either PC (−PC) or UL (−UL), or both (−UL, −PC). Signals corresponding to the UL and PC ligands were also detected in the low-mass range of this positive-ion mode ESI-mass spectrum. This mass region is enlarged in the inset. The *y*-axis scale in the mass range 540–700 *m/z* is zoomed five-fold in this inset.

shown in the inset to Figure 2C (see also Supplementary Figure S1B–C) reveal the heterogeneity of both the PC and the UL species captured by shCD1b. From the ion intensities,

average masses of 615 and 814 Da could be calculated for UL and PC, respectively.

The two major PC constituents have 760.5 and 786.5 masses. Recording MS/MS spectra on these ions result in the neutral loss of C16:0 and C18:1 fatty acids from the 760.5 ion and neutral loss of only C18:1 from the 786.5 ion, and in both cases an intense ion is observed at *m/z* 184.1, a PC marker fragment ion (data not shown). Concerning the UL molecule, in addition to the most intense molecular peak at 619.4 *m/z*, other major peaks are at 633.4, 617.4, 603.4 and 591.4 *m/z*. These series of molecular masses suggest that UL contains an acyl chain varying in length and unsaturation. The odd mass values indicate that either no nitrogen or an even number of nitrogens is present. This ligand dissociates from the shCD1b:PC:UL complex without causing any charge modification of the remaining part, indicative of the absence of easily ionizable group in its chemical structure. The peak intensities observed for UL are 10-fold lower than for PC, a likely consequence of its low ionization cross-section. This might explain why this ligand was not detected in previous studies. Unfortunately, the low abundance of the UL ions hampered the exploitation of tandem MS data.

The same type of ESI-MS measurements was performed on shCD1b expressed from insect cells. Analysis of the MS data also evidences the simultaneous association of shCD1b with one molecule of PC and one unknown ligand (data not shown). The PC and UL peak distribution differed slightly from those presented in Figure 2C for mouse-expressed shCD1b. In the case of insect-expressed shCD1b, the most intense PC signals are observed at *m/z* 730.5, 758.5 and 786.5, and for UL at *m/z* 587.4, 615.4 and 643.5.

Crystal structure of natively folded shCD1b

The results presented above indicate that shCD1b associates with PC and a second unknown molecule. None of the two ligands has ever been observed in the crystal structures reported for hCD1b (Gadola *et al*, 2002; Batuwangala *et al*, 2004), most likely because the protein was always prepared by *in vitro* refolding from the denatured *Escherichia coli*-derived protein. Moreover, refolding was always performed in the presence of externally added lipids and detergents, with the intention to form the corresponding protein–lipid complexes and to prevent protein aggregation. Thus, the native structure of CD1b and the original content of its lipid-binding groove remain unknown. This motivated our decision to attempt the crystallization of natively folded shCD1b.

Mouse-expressed shCD1b was purified as described above. To reduce the heterogeneity, and to increase the likelihood of crystallization, the protein was deglycosylated by treatment with endoF_3. The mass loss accompanying the deglycosylation was estimated by MS to be of ca. 5.2 kDa. The resulting deglycosylated shCD1b could be crystallized, and its structure was determined by molecular replacement and refined to a resolution of 1.8 Å.

The structure of natively folded shCD1b is essentially identical to the published structures of refolded CD1b molecules (Gadola *et al*, 2002; Batuwangala *et al*, 2004). The heavy chain folds into the α1, α2 and α3 domains, and is closely associated with β2m. Our hCD1b superimposes well with the structures of hCD1b in complex with glucose monomycolate (GMM), with PI and with the ganglioside GM2, with root-mean-square deviations (RMSD) on the position of backbone atoms of 0.7, 0.9 and 1.0 Å, respectively. Compared to CD1b:GMM, significant deviations are observed for the loops delineated by residues 55–56 (RMSD of 1.8 Å on backbone atoms) and 104–109 (1.5 Å). With regard to CD1b:PI and CD1b:GM2, important differences are found for the stretches of amino acids 20–44 (1.6 Å), 79–90 (1.6 Å), 104–109 (1.5 Å) and 125–151 (1.6 Å). The binding groove of shCD1b is slightly less voluminous (2430 $Å^3$) than in the case of CD1b:PI (2510 $Å^3$), CD1b:GM2 (2550 $Å^3$) and CD1b:GMM (2570 $Å^3$). This groove consists of three different pockets, previously called C′, A′ and F′, the last two being interconnected by the traversing T′ tunnel. The A′, F′ and C′ channels merge to form the major entry portal to the groove at the TCR recognition surface between the α1 and α2 helices. A second portal was observed at the opposite side of the C′ pocket of reported hCD1b structures, falling below the α2 helix. This portal, which is shown in Figure 3A for the case of the CD1b:PI structure, was proposed to provide a possible exit for hydrophobic tails longer than C16 carbons (Gadola *et al*, 2002). In our structure, however, this exit is either not present or in a closed state (Figure 3B). This is caused by the displacement of the backbone atoms of the Asn128–Ala129 loop residues relative to the α-helical stretch formed by Gln152–Leu161. For instance, the distance between the Cα atoms of Arg159 and Ala129 decreases from 6.6–6.9 Å in previous structures to 5.7 Å in our structure. This is unlikely to originate from crystal packing effects, as this protein region is not directly involved in establishing crystal contacts in the different structures.

Comparisons of the side-chain conformations of residues lining the lipid groove reveal minor displacements between our structure and previously reported structures (average RMSD values of 1.0 Å on 333 side-chain atoms from 82 residues). In addition to the movement of the loop that contains Ala129, which leaves the C′ portal closed, the most important conformational changes are observed for the Leu147 and Arg140 side chains, both contributing to the reduction of the pocket volume in the connecting region between the T′ and F′ channels. Interestingly, the side chain of Leu147 lies in close proximity to Phe84, a residue that has been proposed as candidate to adopt an alternate conformation that would close the F′ pocket in the absence of bound ligand (Gadola *et al*, 2002). This was presented as a potential mechanism to tailor the binding capacity of the CD1b groove to the ligand requirements. Despite the lipid-bound state of the structure presented here (see below) does not permit to confirm this hypothesis, the displacement of the Leu147 might evidence a relatively high plasticity of this protein region.

Presence of a long hydrophobic spacer in the shCD1b lipid groove

The electron density map of natively folded shCD1b shows a continuous tube of density filling the lipid-binding groove (Figure 4A). Such density, which cannot be accounted for by the presence of a single phospholipid molecule, is on the contrary consistent with the simultaneous presence of a PC molecule and a C41 linear molecule. The partial electron density of the phosphocholine polar head and the glycerol connectivity between the two fatty acid chains were taken as starting points to model a PC molecule. In this manner, C18 and C16 fatty acid chains could be fully accommodated in the

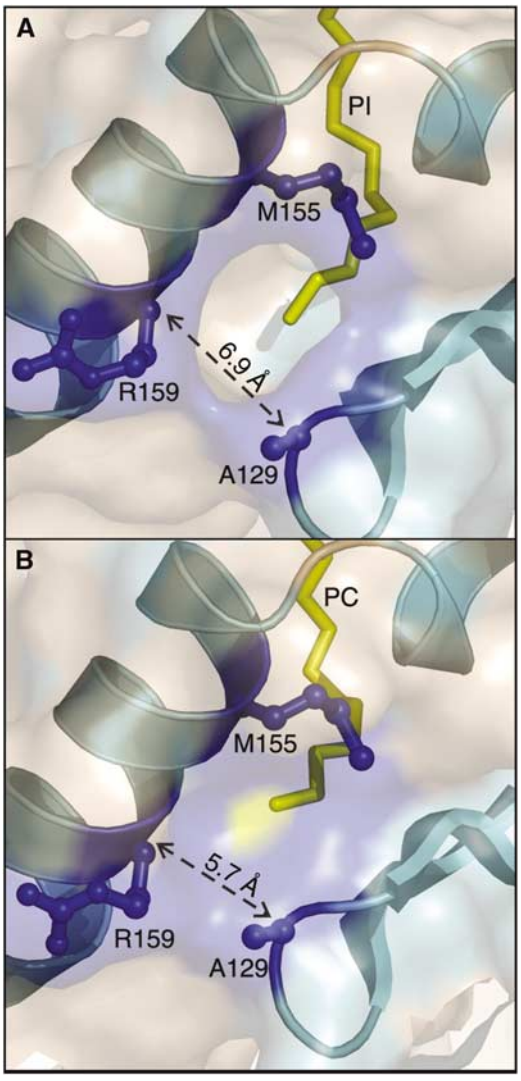

Figure 3 The C′ portal is closed in natively folded hCD1b. Representation of the protein surface at the entry portal to the C′ channel, as observed in the crystal structure of hCD1b in complex with PI (**A**), or in the structure of natively folded hCD1b (**B**). This portal was proposed to allow egress of the lipid (shown in yellow) from the interior of the protein. The Cα and side-chain atoms of residues lining the portal are drawn as balls and sticks and colored in blue. The separation between the Cα atoms of Arg159 and Ala129 is indicated.

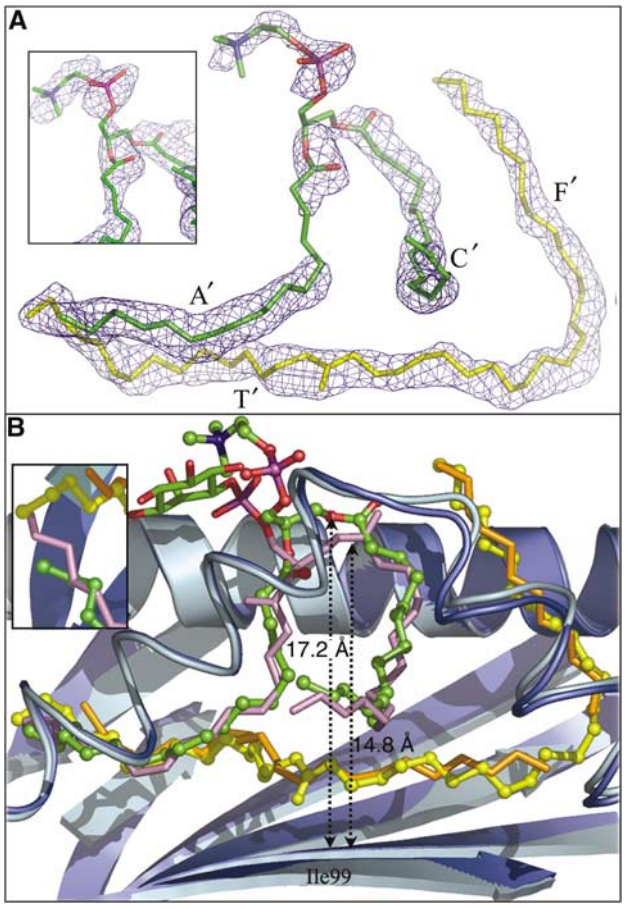

Figure 4 Natural ligands in the binding groove of shCD1b. (**A**) Electron density for the PC and the spacer molecule in the shCD1b crystal structure. Both lipids are drawn as sticks. The PC is colored by atom type (carbon green, oxygen red, phosphorous violet and nitrogen blue), the spacer molecule in yellow. The electron density is from the final $2F_o-F_c$ map contoured at 1.0σ and represented as a blue mesh. Letter labels indicate the groove channels occupied by the close portions of electron density. The inset shows the same map for the polar head region of PC, contoured at 0.6σ. (**B**) Disposition of ligands in the CD1b lipid groove. The crystal structures of shCD1b presented in this work (cyan cartoon representation) and of hCD1b in complex with PI (in blue) were superimposed. For clarity, the helices of the α2 domain are represented as loops. PC and UL ligands are represented as balls and sticks and colored as in panel A. The PI and the two detergents from the CD1b:PI structure are represented as sticks. The polar head of PI is colored by atom type, the fatty acid chains in violet and the two detergent molecules in orange. The distance between the Cα atom of Ile99 and the topmost atom of the glycerol moiety of PI or PC is indicated for both structures. The inset shows a top view of the region where the acyl chains of PC and PI stop, at the connection between the A′ channel and the T′ tunnel.

A′ and C′ pockets, respectively. The phosphocholine polar head is directed from the central groove opening towards the protein surface. Its associated electron density becomes connected to the glycerol moiety at a contour level of 0.6σ in the $2F_o-F_c$ map (inset to Figure 4A), indicative of a relatively high flexibility. This is consistent with the lack of interactions observed between the PC headgroup and the protein. A single hydrogen-bond interaction is established between the phosphate and the side-chain hydroxyl of Thr157.

A second very well-defined and continuous portion of electron density fully occupies the T′ channel and the F′ pocket of shCD1b (Figure 4A), protruding out towards the TCR recognition surface. A 41 carbon atoms-long hydrophobic molecule could be built into this density. Considering the most intense 619 Da MS signal observed for UL, a maximal 44 carbon-atom length is expected for this ligand. The combined observations therefore suggest that the UL molecule is a C41–C44-long linear molecule (e.g. a wax ester, or a C41-long 1,3-diacylglycerol). At most, three atoms would extend beyond the F′ pocket, being too mobile to be observed. As a proposition, the final structure has been modeled with a C41-long ester (a C16 fatty acid esterified to a C24 alcohol). However, the exact nature of this ligand remains to be determined.

Interestingly, the position of the PC relative to the β-sheet platform that encloses the α1–α2 domain changes considerably with regard to the lipids in CD1b structures obtained after *in vitro* refolding (Figure 4B). For comparison, the

glycerol moiety of PC in our structure is positioned 17.2 Å above the Cα atom of Ile99, whereas in the CD1b:PI structure the topmost atom of the PI glycerol is only 14.8 Å away and in the case of GMM an intermediate distance of 15.6 Å is found (not shown). This protruding of the PC molecule and its high mobility, evidenced also by a high average *B* value (Table I, atomic *B* values range from 50 to 90 $Å^2$, being highest for polar head atoms), indicate that this lipid is not fully shielded from the aqueous environment, and rather suggests that PC might be prone to exchange reactions. The PC displacement appears to mostly originate from a UL-induced outwards movement of the C18 acyl chain located in the A′ channel, compared to the modeled C17 chain of the CD1b:PI structure (inset to Figure 4B). Contrarily, the emplacement of the C16 chains in the C′ channel is similar in both structures, ruling out effects caused by the closed state of the C′ portal.

Discussion

In contrast to MHC molecules, the antigen-binding grooves of CD1 molecules are large and very hydrophobic, therefore adapted to present lipid antigens. Understanding the antigen-presentation process implies solving the question of how these large hydrophobic pockets are protected between protein biosynthesis in the ER and the encounter of the antigens in the endocytic compartments. ER-resident LTPs like MTP might transfer lipids to the nascent CD1 molecules during biosynthesis (Brozovic *et al*, 2004; Dougan *et al*, 2005). The inhibition of this process causes a diminution of the ulterior CD1d presentation of antigens to NKT cells. Natural lipid ligands are therefore expected to play chaperone-like and functional roles, akin to the function of invariant chains in the assembly of MHC class II molecules, and reminiscent of endogenous peptides that associate with MHC proteins (Anderson and Cresswell, 1994; Jardetzky *et al*, 1996).

Work has been conducted to identify these ligands. Joyce *et al* (1998) concluded that GPI represents >90% of the material extracted with organic solvents from purified mCD1d1. Their investigations were mostly carried out with membrane-anchored mCD1d1, which was recovered from the cell membranes by solubilization with detergents, a treatment that might modify the original lipid content of mCD1d1. More recently, the same group reported that PI, and not GPI, is the natural ligand associated to soluble forms of mCD1d1, hCD1d and hCD1b (Park *et al*, 2004). The discrepancy between the two reports was merely ascribed to the different purification conditions and to the MS ionization technique employed. Nevertheless, evidence has now been collected that contradicts PI or GPI as being the major endogenous lipids of CD1 molecules: (i) IEF-based lipid-binding studies on soluble forms of mCD1d (Cantu *et al*, 2003; Zajonc *et al*, 2005b) and hCD1b (this work) indicate that these CD1 molecules do not contain negatively charged molecules; (ii) the MS and crystallographic data presented in this work reveal the presence of PC and UL bound to shCD1b, but not of PI; and (iii) extra electron density found in the crystal structure of insect-produced mCD1d was recently attributed to PC (Giabbai *et al*, 2005).

The contradictory data concerning the identity of the endogenous CD1-lipids suggest that PC and PI might indeed exchange easily depending on the experimental conditions selected for CD1 production and purification. Regardless of the identity of this lipid, the combined size of the two hydrophobic chains of mammalian phospholipids is still insufficient to shield the voluminous groove of CD1b. Additional prebound molecules are required for optimal stabilization. The structural and MS data presented in this work confirm that shCD1b is expressed in cells in complex with a fully buried C41–C44 spacer lipid, in addition to a partially accessible C32–C36 PC, located at the groove entrance. The buried emplacement of the spacer suggests that the assembly occurs during biosynthesis. The combined ca. 80 carbon-atoms length of the hydrophobic chains of PC and UL is above the C65–C70 maximal capacity of the CD1b groove (Batuwangala *et al*, 2004). However, the C′ portal is still present in a closed state. As a consequence of this overfilling, the PC appears to be displaced towards the aqueous phase. The overall effect might be to render PC prone to lipid exchange processes, strongly suggesting that the endogenous spacer also plays a functional role.

In vivo experiments have proved that the hCD1b-mediated presentation of short lipids to T cells occurs with much faster kinetics than the presentation of long antigens, most likely because short lipids can directly load onto CD1b at the cell surface, whereas long antigens and CD1b must encounter in late endosomes (Shamshiev *et al*, 2000; Moody *et al*, 2002). Trafficking of CD1b through the endocytic pathway might be necessary for the removal of the spacer molecule, a process that would be mediated by specialized LTPs. This would release the space to host known antigens, that is mycobacterial GMM (Moody *et al*, 1997), mycolic acids (Beckman *et al*, 1994) or diacylsulfoglycolipids (Gilleron *et al*, 2004).

Table I Statistics for crystallographic data collection and refinement

Data collection[a]	
Resolution range (Å)	32.72–1.80 (1.90–1.80)
Completeness (%)	96.4 (85.5)
Number of unique reflections	43 311
Multiplicity	3.5 (2.3)
R_{sym} (%)[b]	6.8 (19.8)
I/σ	11.8 (3.5)
Refinement statistics[a]	
Resolution range (Å)	20.00–1.80 (1.85–1.80)
Number of reflection (work/free)	38 558/2516
R factor[c]	20.6 (25.2)
R_{free}[c]	24.1 (30.0)
Average B values ($Å^2$)[d]	
Protein	32.0 (3000)
Ligand (PC)	70.7 (52)
Ligand (UL)	52.4 (42)
Solvent	43.1 (281)
Ramachandran statistics (%)	
Most favored	93.2
Additionally allowed	6.2
Disallowed	0.6 (heavy chain Asp33 and Ser106)
RMSD from ideal geometry	
Bond length (Å)	0.016
Bond angles (deg)	1.608
ω torsion angle (deg)	6.221

[a]Numbers in parentheses apply to the highest resolution shell.
[b]$R_{sym} = (\sum_h \sum_i |I_i(h) - I(h)|)/(\sum_h \sum_i I_i(h))$.
[c]$R = (\sum_h |F_{obs}(h) - F_{calc}(h)|)/\sum_h |F_{obs}(h)|$.
[d]Number of atoms are in parentheses.

Alternatively, the disposition inside the groove of short-chain lipids (e.g. bacterial PI-mannosides or endogenous lipids) might be different after removal of the spacer, comparable to what is observed in the structure of CD1b:PI. Consistent with this, we have observed that shCD1b interacts *in vitro* with short phospholipids with fast kinetics (<10 min to reach saturation), whereas the branched and long diacylsulfoglycolipids load slowly (>12 h for saturation), and no loading is observed with mycobacterial mycolic acids (data not shown).

Similarly, the delivery of CD1d molecules to the late endosomes is known to be important for antigen presentation (Chiu *et al*, 2002). Extra electron density attributed to a C16-long fatty acid has been recently found in the binding pocket of hCD1d, accompanying a co-crystallized synthetic variant of α-GalCer (Zajonc *et al*, 2005a). The presence of this endogenous lipid also fits in with the much slower *in vitro* binding kinetics of α-GalCer to mCD1d, compared to a shorter synthetic analogue. The spacer would obstruct the insertion of the C26 fatty acid of α-GalCer into the A′ pocket, the presence of LTPs being required for the extraction of the preloaded molecule. On the contrary, the C8 fatty acid chain of the synthetic analogue perfectly accommodated into the A′ pocket in combination with the spacer molecule.

A major question remains to be solved concerning the identity of the hydrophobic spacer. Investigations are currently being conducted to address this issue. This molecule is expected to be non-charged, hydrophobic, non-ramified, 41–44 carbon atoms long and with a molecular mass of around 619 Da. These observations permit to discard triacylglycerides, 1,2-diacylglycerols, steryl esters or polyisoprenoids. Ester waxes or 1,3-diacylglycerols might comply with the expected physicochemical and structural features.

In conclusion, preloading CD1 molecules with such spacers is proposed to represent not only a general mechanism for the structural stabilization of these proteins, but also a means to modulate the lipid-binding properties of the protein and to ensure that antigen loading will take place at the right compartment. Efforts are now devoted to decipher the importance of such long spacer molecules for the stability and function of CD1b.

Materials and methods

Reagents

All lipids were purchased from Avanti Polar Lipids and detergents were from Fluka. Radioactive PI, L-α-[*myo*-inositol-2-^{3}H(N)] ([^{3}H]PI) was from Perkin-Elmer. Lipid liposomes at 0.1–1 mM concentrations were prepared by sonication using standard protocols. Anti-CD1b BCD1b3.1 mAb was covalently attached to *N*-hydroxysuccinimide-activated sepharose resin (Amersham Biosciences) following the provided protocol. Size-exclusion chromatography was performed using a Superdex 75 10/300 GL column (Amersham Biosciences), equilibrated with 10 mM Bis-Tris/50 mM NaCl/1 mM EDTA buffer (pH 6.5). The endoF_3 protein fusion to the maltose-binding protein (MBP) was prepared as described (Tarentino *et al*, 1999).

shCD1b expression in mouse cells and purification

The pBlueCD1B plasmid containing the full-length hCD1b cDNA (Manolova *et al*, 2003) was used as template for the amplification of the extracellular domain by PCR using Advantage-HF 2 DNA Polymerase (BD Clontech, CH), and a 5′ primer at the beginning of the leader sequence (5′hCD1B*Sal*I, 5′-TATAGTCGACATGCTGCTGCT GCCATTTCAAC) and a 3′ primer at the end of the predicted extracellular region (3′hCD1B*sol**Cla*I, 5′-CCATCGATGGGGTTTCTCC AGTAGAGGA). The PCR product was cloned into the pBluescript-Bir vector, which contains the sequence for the BirA peptide tag (GGGLNDIFEAQKIEWHE). The DNA fragment encoding shCD1b fused in-frame with the C-terminal BirA tag sequence was subcloned into BCMGSNeo, and supertransfected into J558 cells previously transfected with human β2m cDNA. Transfection, selection, screenings with BirA-peptide-specific mAb and cloning were performed as previously described (Nowbakht *et al*, 2005). The clone producing the highest level of shCD1b in the culture supernatant was expanded for large-scale production in RPMI 1640 medium supplemented with 3% FCS, 2 mM GlutaMAX I, MEM non-essential amino acids, 1 mM sodium pyruvate and 100 μg/ml kanamycin (all from Invitrogen, CH). Cell supernatants were harvested by centrifugation and filtration through 0.22 μm membranes. The solutions were supplemented with 2 mM EDTA and 1 mM PMSF, and stored at 4°C.

The protein was purified by affinity chromatography using the anti-CD1b (BCD1b3.1) column. Analysis of the loaded supernatant and of the collected flow-through by Western blot, developed with a biotinylated anti-BirA antibody, proved that all shCD1b had been retained inside the column. The column was washed with 50 ml of a PBS buffer containing 1 mM EDTA (pH 7.3), and shCD1b was eluted with 50 ml of a buffer containing 50 mM citrate, 1 mM EDTA and 50% ethylene glycol (pH 5.0). Pooling and concentration furnished about 1–2 mg of N-glycosylated hCD1b per liter of cell supernatant.

hCD1b expression and purification from insect cells

Recombinant shCD1b was produced in the S2 *Drosophila* cell line following the strategy described for recombinant soluble CD1e (de la Salle *et al*, 2005). The extracellular domain of CD1b, without the signal peptide (residues 17–302), was fused to the signal peptide of the Bip protein and a C-terminal V5-His tag (pMTbipV5HisA, Invitrogen). The β2 m cDNA was cloned into pMTV5MTHisA (Invitrogen). Then, a unique vector encoding β2m, CD1b and hygromycin phosphotransferase was constructed and transfected in S2 cells using the JetPEI transfection reagent (Polyplus transfection, Illkirch, France). Cells were selected in InsectXpress culture medium (Cambrex), supplemented with 10% FCS, 300 μg/ml hygromycin (Roche Applied Science) and gentamycin (Sigma). For production, cells were cultured to 2×10^7 cells/ml. Cells were washed and then transferred in InsectXpress medium without serum, plus 750 μM $CuSO_4$ for 5 days. Soluble hCD1b was purified by chromatography on Chelating Sepharose Fast Flow (Amersham Biosciences), followed by dialysis against 25 mM sodium acetate (pH 5.5) and purification by ion-exchange chromatography. This results in a preparation including only the CD1b heavy chain and β2m, as judged by SDS–PAGE analysis.

Isoelectric-focusing electrophoresis

The basic protocol to measure lipid loading onto CD1 molecules has been described (Cantu *et al*, 2003). Briefly, 10–20 μM of N-glycosylated shCD1b and 50 μM lipid were shaken in glassware tubes in a solution containing 50 mM sodium acetate/50 mM NaCl/1 mM DTT/1 mM EDTA, pH 5.0, at 30°C and 200 r.p.m. After 12 h, the solution was cooled on ice and 1 μl applied to the IEF gel. Incubation of shCD1b and purified shCD1b:PI with detergents was carried out for 10 min at room temperature, replacing the lipids with the following detergents (final concentration): Triton X-100 (1.5 mM), taurocholate (10 mM), CTAB (1 mM) or C_{12}DAO (3 mM). Electrophoresis was performed in a PhastGel system (Amersham Biosciences) for 500–700 accumulated volt hours (AVh). Proteins in IEF gel were detected by staining with Coomassie R 350.

Mass spectrometry

Before mass analysis, protein samples were buffer exchanged into 50 mM aqueous ammonium acetate (pH 6.7) by a three-step dilution-concentration using 5 kDa Millipore centrifuge filters. MS measurements were performed in the positive ion mode using an electrospray ionization quadrupole time-of-flight (ESI-Q-ToF 1) instrument (Micromass, Manchester, UK) equipped with a Z-spray nano-electrospray ionization source. Nanospray needles were made from borosilicate glass capillaries (Kwik-Fil, World Precision Instruments, Sarasota, FL) on a P-97 puller (Sutter Instruments, Novato, CA), coated with a thin gold layer by using an Edwards Scancoat (Edwards Laboratories, Milpitas, CA) six Pirani 501 sputter coater. To produce intact gas-phase ions from the large noncovalent protein–lipid complexes, the ions were cooled collisionally by increasing the pressure in the first vacuum stages of the

mass spectrometer to values ranging from 5.5 to 7 mbar (Krutchinsky *et al*, 2000). Nano-electrospray voltages were optimized for generation and transmission of the macromolecular protein complexes; the needle voltage varied between 1300 and 1500 V and the sample cone voltage was set at between 10 and 80 V. The pressure and applied desolvation voltages (i.e. sample cone voltage) were selected to facilitate preservation of non-covalent interactions and promote efficient ion desolvation in the interface region of the instrument. This, in turn, resulted in adequately sharp ion signals to allow reasonable accurate mass determination.

For tandem MS experiments, precursor ions were isolated in the quadrupole mass analyzer and accelerated into an argon-filled linear hexapole collision cell. Various collision energies were used, with argon at a pressure of 3.0×10^{-2} mbar. All mass spectra were mass calibrated by using an aqueous solution of cesium iodide (50 mg/ml).

Crystallization of deglycosylated shCD1b

Purified shCD1b (1 mg) expressed in mouse cells was treated with 0.1 mg of the MBP-endoF_3 fusion protein in 100 mM acetate buffer (final pH 5.1) for 4 h at 30°C. After pH neutralization by addition of 1 M HEPES (pH 7.5), the solution was cooled on ice, diluted five-fold with deionized water and loaded into an anion-exchange chromatography BioScale Q2 column (Bio-Rad). The protein eluted at 37% buffer B, in a linear gradient from buffer A (10 mM Bis-Tris/1 mM EDTA/pH 6.5) to buffer B (500 mM NaCl in buffer A). Protein fractions were pooled, concentrated and the NaCl concentration reduced to about 30–50 mM by dilution-concentration using 10 kDa Vivaspin centrifuge filters. The best crystals were grown after 1–3 days at 20°C, using the hanging-drop vapor-diffusion method, from drops containing 1 µl hCD1b (5–7 mg/ml) and 0.5 µl precipitant (1.5 M $(NH_4)_2SO_4$, 5% iPrOH, 0.1 M sodium citrate, pH 5.50).

Structure determination, analysis and presentation

Crystals were briefly immersed into the crystallization solution supplemented with 20% glycerol before being transferred into a gaseous nitrogen flux at 100 K. Diffracted intensities were collected from a single crystal at ESRF (Grenoble, France) on beamline ID14 EH4, tuned at 0.9756 Å, using an ADSC Q4 CCD detector. Diffraction images were processed using MOSFLM (Leslie, 1987), scaled with SCALA (Evans, 1993) and further processed with the CCP4 programs package (Collaborative Computational Project Number 4, 1994). Data processing statistics are summarized in Table I. Crystals belong to the orthorhombic space group $P2_12_12_1$ ($a = 39.83$ Å, $b = 103.90$ Å, $c = 114.83$ Å) and contain one heterodimer of CD1b per asymmetric unit.

The structure of the native hCD1b was solved by the molecular replacement method, with the program PHASER (Storoni *et al*, 2004), using both polypeptidic chains of unliganded hCD1b, as found in the crystal structure of hCD1b in complex with GM2 (PDB entry 1GZP; Gadola *et al*, 2002). A single unambiguous solution was identified with a *Z*-score of 18.6 for data between 20.0 and 4.0 Å resolution. This solution was refined using data between 20.0 and 1.8 Å resolution with REFMAC (Murshudov *et al*, 1997). Along refinement, the F_o-F_c electron density maps indicated that Asn20 and Asn57 on chain A were glycosylated, and Fuc-α(1-6)-GlcNAc disaccharides were added to the model. Solvent molecules were subsequently incorporated, as indicated by both $2F_o-F_c$ and F_o-F_c electron density maps. Two lipid molecules were included in the model in order to account for a tubular electron density that was clearly visible in the F_o-F_c map, at a contour level of 3.0σ, in the A′, T′, F′ and C′ channels. The final model includes 382 residues, four sugar moieties, two lipids (PC and UL), two sulfate anions, one glycerol molecule and 265 water molecules. Statistics for the refined structure are given in Table I. Protein stereochemistry was validated using PROCHECK (Laskowski *et al*, 1993).

Molecular surface areas and volumes were calculated using the CASTp server (Liang *et al*, 1998). A solvent probe of 1.4 Å was used to map internal cavities and pockets of the proteins. Figures 3 and 4 were generated with PyMOL (DeLano, 2002).

Supplementary data

Supplementary data are available at *The EMBO Journal* Online.

Acknowledgements

We acknowledge Patrick Van Roey (Wadsworth Center, New York, USA) for kindly providing the plasmid and instructions to prepare the MBP-endoF_3 fusion protein and Steven A Porcelli (Albert Einstein College of Medicine, New York, USA) for providing the BCD1b3.1 cells for production of anti-CD1b mAb. This work was benefited from the constant support of the CNRS and the Swiss National Foundation (grant 3100A0-109918/1). We are grateful to the staff of the ESRF (Grenoble, France) for excellent data collection facilities.
Coordinates and structure factors of shCD1b have been deposited in the Protein Data Bank (www.rcsb.org), accession code 2H26.

References

Anderson KS, Cresswell P (1994) A role for calnexin (IP90) in the assembly of class II MHC molecules. *EMBO J* **13:** 675–682

Batuwangala T, Shepherd D, Gadola SD, Gibson KJ, Zaccai NR, Fersht AR, Besra GS, Cerundolo V, Jones EY (2004) The crystal structure of human CD1b with a bound bacterial glycolipid. *J Immunol* **172:** 2382–2388

Beckman EM, Porcelli SA, Morita CT, Behar SM, Furlong ST, Brenner MB (1994) Recognition of a lipid antigen by CD1-restricted alpha beta+ T cells. *Nature* **372:** 691–694

Benesch JL, Robinson CV (2006) Mass spectrometry of macromolecular assemblies: preservation and dissociation. *Curr Opin Struct Biol* **16:** 245–251

Brigl M, Brenner MB (2004) CD1: antigen presentation and T cell function. *Annu Rev Immunol* **22:** 817–890

Brozovic S, Nagaishi T, Yoshida M, Betz S, Salas A, Chen D, Kaser A, Glickman J, Kuo T, Little A, Morrison J, Corazza N, Kim JY, Colgan SP, Young SG, Exley M, Blumberg RS (2004) CD1d function is regulated by microsomal triglyceride transfer protein. *Nat Med* **10:** 535–539

Cantu III C, Benlagha K, Savage PB, Bendelac A, Teyton L (2003) The paradox of immune molecular recognition of alpha-galactosylceramide: low affinity, low specificity for CD1d, high affinity for alpha beta TCRs. *J Immunol* **170:** 4673–4682

Cham BE, Knowles BR (1976) A solvent system for delipidation of plasma or serum without protein precipitation. *J Lipid Res* **17:** 176–181

Chiu YH, Park SH, Benlagha K, Forestier C, Jayawardena-Wolf J, Savage PB, Teyton L, Bendelac A (2002) Multiple defects in antigen presentation and T cell development by mice expressing cytoplasmic tail-truncated CD1d. *Nat Immunol* **3:** 55–60

Collaborative Computational Project Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D* **50:** 760–763

de la Salle H, Mariotti S, Angenieux C, Gilleron M, Garcia-Alles LF, Malm D, Berg T, Paoletti S, Maitre B, Mourey L, Salamero J, Cazenave JP, Hanau D, Mori L, Puzo G, De Libero G (2005) Assistance of microbial glycolipid antigen processing by CD1e. *Science* **310:** 1321–1324

De Libero G, Mori L (2005) Recognition of lipid antigens by T cells. *Nat Rev Immunol* **5:** 485–496

DeLano WL (2002) The PyMOL molecular graphics system http://www.pymol.org

Dougan SK, Salas A, Rava P, Agyemang A, Kaser A, Morrison J, Khurana A, Kronenberg M, Johnson C, Exley M, Hussain MM, Blumberg RS (2005) Microsomal triglyceride transfer protein lipidation and control of CD1d on antigen-presenting cells. *J Exp Med* **202:** 529–539

Evans PR (1993) Data reduction. In *Proceedings of the CCP4 Study Weekend: Data Collection and Processing*, Sawyer L, Isaacs N, Bailey S (eds) pp 114–122. Warrington, UK: Science and Engineering Research Council, Daresbury Laboratory

Gadola SD, Zaccai NR, Harlos K, Shepherd D, Castro-Palomino JC, Ritter G, Schmidt RR, Jones EY, Cerundolo V (2002) Structure of human CD1b with bound ligands at 2.3 Å, a maze for alkyl chains. *Nat Immunol* **3:** 721–726

Giabbai B, Sidobre S, Crispin MD, Sanchez-Ruiz Y, Bachi A, Kronenberg M, Wilson IA, Degano M (2005) Crystal structure of mouse CD1d bound to the self ligand phosphatidylcholine: a molecular basis for NKT cell activation. *J Immunol* **175:** 977–984

Gilleron M, Stenger S, Mazorra Z, Wittke F, Mariotti S, Bohmer G, Prandi J, Mori L, Puzo G, De Libero G (2004) Diacylated sulfoglycolipids are novel mycobacterial antigens stimulating CD1-restricted T cells during infection with *Mycobacterium tuberculosis*. *J Exp Med* **199:** 649–659

Jardetzky TS, Brown JH, Gorga JC, Stern LJ, Urban RG, Strominger JL, Wiley DC (1996) Crystallographic analysis of endogenous peptides associated with HLA-DR1 suggests a common, polyproline II-like conformation for bound peptides. *Proc Natl Acad Sci USA* **93:** 734–738

Joyce S, Woods AS, Yewdell JW, Bennink JR, De Silva AD, Boesteanu A, Balk SP, Cotter RJ, Brutkiewicz RR (1998) Natural ligand of mouse CD1d1: cellular glycosylphosphatidylinositol. *Science* **279:** 1541–1544

Kang SJ, Cresswell P (2004) Saposins facilitate CD1d-restricted presentation of an exogenous lipid antigen to T cells. *Nat Immunol* **5:** 175–181

Krutchinsky AN, Ayed A, Donald LJ, Ens W, Duckworth HW, Standing KG (2000) Studies of noncovalent complexes in an electrospray ionization/time-of-flight mass spectrometer. *Methods Mol Biol* **146:** 239–249

Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* **26:** 283–291

Leslie AGW (1987) Profile fitting. In *Proceedings of the Daresbury Study Weekend: Computational Aspects of Protein Crystal Data Analysis*, Helliwell JR, Machin PA, Papiz MZ (eds) pp 39–50. Warrington, UK: Science and Engineering Research Council, Daresbury Laboratory

Liang J, Edelsbrunner H, Woodward C (1998) Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. *Protein Sci* **7:** 1884–1897

Manolova V, Hirabayashi Y, Mori L, De Libero G (2003) CD1a and CD1b surface expression is independent from *de novo* synthesized glycosphingolipids. *Eur J Immunol* **33:** 29–37

Moody DB, Briken V, Cheng TY, Roura-Mir C, Guy MR, Geho DH, Tykocinski ML, Besra GS, Porcelli SA (2002) Lipid length controls antigen entry into endosomal and nonendosomal pathways for CD1b presentation. *Nat Immunol* **3:** 435–442

Moody DB, Porcelli SA (2003) Intracellular pathways of CD1 antigen presentation. *Nat Rev Immunol* **3:** 11–22

Moody DB, Reinhold BB, Guy MR, Beckman EM, Frederique DE, Furlong ST, Ye S, Reinhold VN, Sieling PA, Modlin RL, Besra GS, Porcelli SA (1997) Structural requirements for glycolipid antigen recognition by CD1b-restricted T cells. *Science* **278:** 283–286

Moody DB, Zajonc DM, Wilson IA (2005) Anatomy of CD1–lipid antigen complexes. *Nat Rev Immunol* **5:** 387–399

Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* **53:** 240–255

Nowbakht P, Ionescu MC, Rohner A, Kalberer CP, Rossy E, Mori L, Cosman D, De Libero G, Wodnar-Filipowicz A (2005) Ligands for natural killer cell-activating receptors are expressed upon the maturation of normal myelomonocytic cells but at low levels in acute myeloid leukemias. *Blood* **105:** 3615–3622

Park JJ, Kang SJ, De Silva AD, Stanic AK, Casorati G, Hachey DL, Cresswell P, Joyce S (2004) Lipid–protein interactions: biosynthetic assembly of CD1 with lipids in the endoplasmic reticulum is evolutionarily conserved. *Proc Natl Acad Sci USA* **101:** 1022–1026

Prigozy TI, Naidenko O, Qasba P, Elewaut D, Brossay L, Khurana A, Natori T, Koezuka Y, Kulkarni A, Kronenberg M (2001) Glycolipid antigen processing for presentation by CD1d molecules. *Science* **291:** 664–667

Shamshiev A, Donda A, Prigozy TI, Mori L, Chigorno V, Benedict CA, Kappos L, Sonnino S, Kronenberg M, De Libero G (2000) The alphabeta T cell response to self-glycolipids shows a novel mechanism of CD1b loading and a requirement for complex oligosaccharides. *Immunity* **13:** 255–264

Storoni LC, McCoy AJ, Read RJ (2004) Likelihood-enhanced fast rotation functions. *Acta Crystallogr D* **60:** 432–438

Tarentino A, Plummer Jr T, Quinones G (1999) An improved amplification system for the production of Endo F3. *Glycobiology* **9:** iii

van den Heuvel RH, Heck AJ (2004) Native protein mass spectrometry: from intact oligomers to functional machineries. *Curr Opin Chem Biol* **8:** 519–526

Winau F, Schwierzeck V, Hurwitz R, Remmel N, Sieling PA, Modlin RL, Porcelli SA, Brinkmann V, Sugita M, Sandhoff K, Kaufmann SH, Schaible UE (2004) Saposin C is required for lipid presentation by human CD1b. *Nat Immunol* **5:** 169–174

Wu D, Zajonc DM, Fujio M, Sullivan BA, Kinjo Y, Kronenberg M, Wilson IA, Wong CH (2006) Design of natural killer T cell activators: structure and function of a microbial glycosphingolipid bound to mouse CD1d. *Proc Natl Acad Sci USA* **103:** 3972–3977

Zajonc DM, Cantu III C, Mattner J, Zhou D, Savage PB, Bendelac A, Wilson IA, Teyton L (2005a) Structure and function of a potent agonist for the semi-invariant natural killer T cell receptor. *Nat Immunol* **6:** 810–818

Zajonc DM, Maricic I, Wu D, Halder R, Roy K, Wong CH, Kumar V, Wilson IA (2005b) Structural basis for CD1d presentation of a sulfatide derived from myelin and its implications for autoimmunity. *J Exp Med* **202:** 1517–1526

Zhou D, Cantu III C, Sagiv Y, Schrantz N, Kulkarni AB, Qi X, Mahuran DJ, Morales CR, Grabowski GA, Benlagha K, Savage P, Bendelac A, Teyton L (2004) Editing of CD1d-bound lipid antigens by endosomal lipid transfer proteins. *Science* **303:** 523–527

