Immunity, Vol. 22, 209–219, February, 2005, Copyright ©2005 by Elsevier Inc. DOI 10.1016/j.immuni.2004.12.009

Molecular Mechanism of Lipopeptide Presentation by CD1a

Dirk M. Zajonc,[1] M.D. Max Crispin,[1,6] Thomas A. Bowden,[1] David C. Young,[3,5] Tan-Yun Cheng,[3] Jingdan Hu,[4] Catherine E. Costello,[5] Pauline M. Rudd,[6] Raymond A. Dwek,[6] Marvin J. Miller,[4] Michael B. Brenner,[3] D. Branch Moody,[3] and Ian A. Wilson[1,2,*]

[1]Department of Molecular Biology
[2]Skaggs Institute for Chemical Biology
The Scripps Research Institute
10550 North Torrey Pines Road
La Jolla, California 92037

[3]Division of Rheumatology, Immunology and Allergy
Brigham and Women's Hospital
Harvard Medical School
1 Jimmy Fund Way
Boston, Massachusetts 02115

[4]Department of Chemistry and Biochemistry
University of Notre Dame
251 Nieuwland Science Hall
Notre Dame, Indiana 46556

[5]Mass Spectrometry Resource
Boston University School of Medicine
Boston, Massachusetts 02118

[6]Department of Biochemistry
Oxford Glycobiology Institute
University of Oxford
Oxford OX1 3QU
United Kingdom

Summary

CD1a is expressed on Langerhans cells (LCs) and dendritic cells (DCs), where it mediates T cell recognition of glycolipid and lipopeptide antigens that contain either one or two alkyl chains. We demonstrate here that CD1a-restricted T cells can discriminate the peptide component of didehydroxymycobactin lipopeptides. Structure analysis of CD1a cocrystallized with a synthetic mycobactin lipopeptide at 2.8 Å resolution further reveals that the single alkyl chain is inserted deep within the A′ pocket of the groove, whereas its two peptidic branches protrude along the F′ pocket to the outer, α-helical surface of CD1a for recognition by the TCR. Remarkably, the cyclized lysine branch of the peptide moiety lies in the shallow F′ pocket in a conformation that closely mimics that of the alkyl chain in the CD1a-sulfatide structure. Thus, this structural study illustrates how a single chain lipid can be presented by CD1 and that the peptide moiety of the lipopeptide is recognized by the TCR.

Introduction

The human CD1 family of nonpolymorphic, glycosylated antigen-presenting molecules consists of five proteins, CD1a, CD1b, CD1c, CD1d, and CD1e, which are expressed on professional antigen-presenting cells, such as B cells, LCs, and DCs (Porcelli, 1995). The functions of CD1 molecules are well studied in humans, mice, and guinea pigs, and CD1 proteins have been identified in all mammalian species studied to date (Dascher and Brenner, 2003). The human family of CD1 proteins can be separated into two groups based on their sequence similarity and immunological functions. In humans, group 1 CD1 consists of CD1a, CD1b, CD1c, and CD1e, whereas the only group 2 isoform is CD1d (Calabi et al., 1989). Whereas group 1 CD1 proteins are thought to participate in host defense by presenting microbial lipid antigens to cytotoxic T cells, group 2 proteins respond to endogenous antigens and carry out immunoregulatory functions (Gumperz and Brenner, 2001). Nevertheless, recent studies have established a role for group 2 CD1 in host defense (Fischer et al., 2004; Skold and Behar, 2003; Vincent et al., 2003).

CD1a differs from the other group 1 CD1 molecules in terms of cellular expression and intracellular localization. Firstly, CD1a is the only CD1 protein to be constitutively expressed at high levels on LCs, where Langerin, an LC-expressed C type lectin participates in glycolipid antigen capture and presentation to T cells (Hunger et al., 2004). Secondly, unlike other CD1 isoforms, which have endosomal targeting sequences in their cytoplasmic tails and recycle to LAMP1$^+$ late endosomes, the short cytoplasmic tail of CD1a lacks such motifs, so that this isoform localizes predominantly at the cell surface with only low levels in recycling endosomes.

CD1 exhibits a similar fold and architecture as major histocompatibility complex (MHC) class I molecules, but its binding groove has evolved into a narrow and deep hydrophobic cleft, more suited to binding lipids and glycolipids (Gadola et al., 2002; Zajonc et al., 2003; Zeng et al., 1997). This binding groove is formed by the two interconnecting pockets, A′ and F′, in mouse CD1d and human CD1a. The A′ pocket is completely buried deep inside the protein and connected to the solvent via the F′ pocket, which extends from the A′-F′ junction to the protein surface. In addition to the A′ and F′ pockets, CD1b bears a T′ tunnel that runs below the F′ pocket and a C′ portal, which is located underneath the α2 helix and connects the interior groove with the outer surface of CD1. These additional pockets are necessary for the accommodation of larger lipids, such as mycolic acids, which contain up to 80 carbon atoms.

The variety of different classes of lipid antigens that has been identified so far includes self-antigens, such as the common phosphoglycerolipids (De Silva et al., 2002; Gumperz et al., 2000; Joyce et al., 1998) and sphingolipids (Kawano et al., 1997; Shamshiev et al., 1999, 2002; Wu et al., 2003), and foreign lipids, such as lipoarabinomannan (Sieling et al., 1995), phosphatidylinositol mannoside (Fischer et al., 2004), mycolic acids (Beckman et al., 1994), glucose monomycolates (Moody et al., 1997), diacylated sulfoglycolipids (Gilleron et al., 2004),

*Correspondence: wilson@scripps.edu

and mannosyl phosphodolichols (Moody et al., 2000) from *Mycobacterium tuberculosis* and related pathogenic microbes. CD1a can present certain exogenously acquired mycobacterial lipid antigens to specific αβ T cell receptors (TCRs) for host cell defense (Rosat et al., 1999).

However, the natural diversity and biological relevance of these CD1 antigens has remained elusive until recently. The identification of mycobacterial lipopeptides as a new class of CD1-restricted antigens (Moody et al., 2004) provided the first example of a bacterial antigen specific to CD1a. These lipopeptides are called didehydroxymycobactin (DDMs) and belong to the N-aryl-capped mycobactin subfamily of the siderophores, which are iron-chelating molecules (Snow and White, 1969). Upon infection, mycobacteria scavenge iron from the host cell by expressing both cell-associated and secreted siderophores, which bind iron with a remarkable affinity and transport it to the cell wall (Wooldridge and Williams, 1993). The synthesis of N-aryl-capped mycobactins is carried out by a multimodular enzyme complex, which consists of nonribosomal peptide synthetases (NRPSs) and polyketide synthases (PKSs) (Crosa and Walsh, 2002; Quadri, 2000), whereby the peptide chain is first elongated and then acylated and hydroxylated to form the final product. The synthesis of mycobactins is essential for mycobacterial growth inside macrophages (De Voss et al., 2000). These natural products represent the first known lipopeptide antigens presented by CD1 proteins and suggest that CD1-restricted TCRs, like MHC-restricted TCRs, can also discriminate polypeptide sequences. Interestingly, although the CD1a binding groove is highly hydrophobic and relatively rigid, the repertoire of bound antigens is quite chemically diverse. The CD1a sulfatide self-antigen, a common sphingolipid, is composed of two alkyl chains, a sphingosine backbone and a fatty acid, and a sulfated galactose headgroup. On the contrary, DDM lipopeptide foreign antigens are composed of a single alkyl chain and a more complex headgroup formed by amino acids and organic acids.

Here, we report the crystal structure of CD1a in complex with a synthetic mycobactin lipopeptide JH-02215, which demonstrates how a single alkyl chain is anchored in the binding groove and how the peptidic portion protrudes from the groove for T cell recognition. Thus, lipopeptide presentation to T cells involves a molecular mode of recognition that has features of both CD1 and MHC molecules because it involves TCR recognition of the peptide, like MHC molecules, and lipid anchoring in the binding groove, as in CD1.

Results

T Cell Activation by CD1a with Bound Lipopeptides Is Specific for the Peptide Moiety

DDM lipopeptides are assembled from five organic acids, salicylic acid, methylserine, lysine, hydroxybutyrate, and lysine that form the peptidic headgroup. The ϵ amino group of the central lysine is then acylated with a C_{20}, monounsaturated fatty acid, and the terminal lysine is cyclized as a result of intramolecular attack of its ϵ amino group on its own terminal carboxylate (Figure 1B). Thus, the peptide moiety consists of two branches, the N-aryl branch composed of salicylmethylserine and the lysine branch composed of hydroxybutyryl-lysine, which radiate from the central acylated lysine (Figure 1B). Although the structural basis for recognition of these molecules by CD1a was unknown, we proposed from modeling studies that the fatty acyl unit of DDMs would be inserted deeply into the A′ pocket of the hydrophobic binding groove, which would leave the polypeptide exposed on the surface of CD1a (Moody et al., 2004).

To determine whether the CD1a-mediated T cell response is, indeed, specific for these organic and amino acids that form the polypeptide, DDM homologs with altered peptide structures were isolated. First, we used high pressure liquid chromatography (HPLC) to isolate spontaneous hydrolysis products of DDM cleaved at the internal ester group; the resulting product lacks the lysine branch, and the $[M + H]^+$ was detected by electrospray ionization mass spectrometry at m/z 642 (DDM-642). In addition, DDM was treated with trifluoroacetic acid (TFA) to modify the aryl branch through opening of the oxazoline ring formed from the salicylmethylserine linkage. Collision-induced dissociation mass spectrometric analysis (CID-MS) showed that intact DDM, $[M + H]^+$ m/z 838 (DDM-838), was converted to a product whose $[M + H]^+$ appears at m/z 856 (DDM-856) due to the addition of a single water molecule at the expected position. The CD1a-restricted reporter T cell line CD8-2 showed little or no response to both DDM-856 and DDM-642, providing evidence that the aryl and lysine branches both play a role in T cell activation (Figure 1A). Furthermore, the synthetic lipopeptide JH-02215 failed to stimulate the same T cell line, suggesting that the moieties added to the C-terminal lysine mask the antigenic T cell epitope (Figure 1B, right).

Structure Determination of the CD1a-Lipopeptide Complex

To directly study the molecular mechanism of this lipopeptide recognition by CD1, we sought to crystallize human CD1a with lipopeptides. Because DDMs from pathogenic mycobacteria can be produced only in nanogram to microgram quantities, synthetic DDM-like lipopeptide (JH-02215) was selected for crystallization, as it could be produced in adequate quantities and recapitulated most of the chemical structural features of DDM (Figure 1B). Compared to mycobacterial DDM-838 (C_{20}), JH-02215 has a shorter (C_{16}) saturated alkyl chain, lacks the α-methyl group on its serine moiety, and carries a hydroxyl group on the central lysine, as in mycobactin. In addition, JH-02215 was synthesized with its lysine attached to a silicon atom carrying two additional phenyl and one tertiary butyl group that improved chemical yield and was predicted not to be involved in CD1a binding. JH-02215 mimics all of the other chemical features of DDM (shown in black in Figure 1B) and binds to human CD1a.

Soluble CD1a protein was produced in the *Drosophila melanogaster* expression system (DES) and purified as previously described (Zajonc et al., 2003). After partial demannosylation of CD1a, the hexahistidine affinity tag was removed by carboxypeptidase A digestion. The

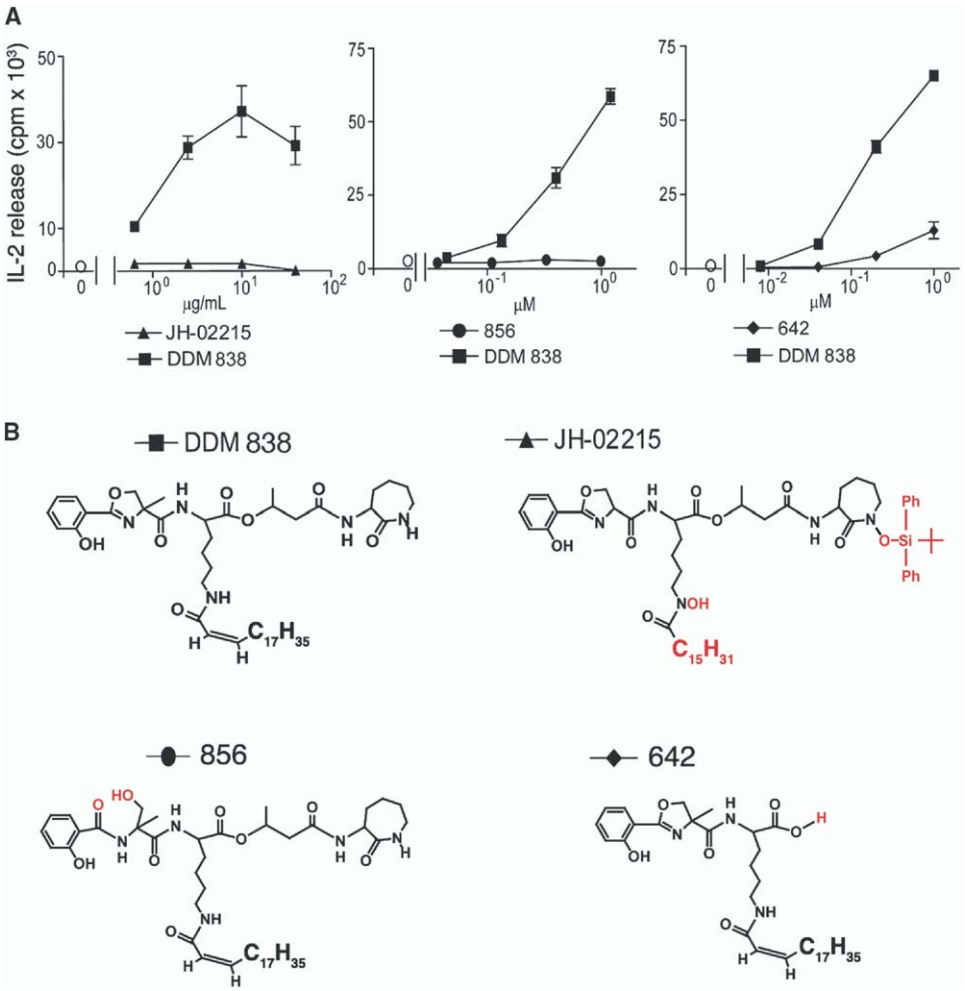

Figure 1. T Cell Stimulation by Lipopeptide Antigens

(A) The CD1a-restricted reporter T cell line CD8-2 was tested for its ability to release IL-2 in response to monocyte-derived dendritic cells to purified DDMs with a unsaturated C_{20} acyl chain (DDM-838) or to DDM-like lipopeptides (DDM-856, DDM-642, and synthetic JH-02215). Only DDM-838 supported significant release of IL-2. The data are shown as the mean ± SD and are representative of two or more experiments.
(B) Structures of studied lipopeptides with chemical differences are shown in red.

synthetic lipopeptide JH-02215 was loaded onto CD1a after forming mixed micelles with Tween 20 detergent. By using the sitting drop vapor diffusion method, we obtained high-quality diffracting crystals in a crystallization condition similar to that reported for the CD1a-sulfatide complex (Zajonc et al., 2003). As the crystal unit cell dimensions were almost identical to the CD1a-sulfatide crystals, we determined the CD1a-lipopeptide structure by molecular replacement using the CD1a-sulfatide structure (1ONQ). The complex structure was refined to a resolution of 2.8 Å (Table 1 and Figure 2A) and final R_{cryst} and R_{free} values of 21.7% and 27.7%, respectively. The asymmetric unit of the crystal contained two CD1-lipopeptide complexes (A and B), which were very similar due to tight noncrystallographic symmetry (NCS) restraints. However, the applied NCS averaging of the electron density did not improve the electron density for the ligand, which had weaker and overall less well-defined electron density, consistent with higher B values for the protein itself. Therefore, we describe here the structural features of the CD1a-ligand complex A, except where otherwise stated.

Interactions between the Lipopeptide and CD1a

As in the CD1a-sulfatide structure, the CD1a-β_2M heterodimer is composed of the three CD1a heavy chain domains, α1, α2, and α3, noncovalently associated with β_2M (Zajonc et al., 2003). The α1 and α2 helices form a narrow and deep antigen binding groove on top of a six-stranded β sheet platform. The antigen binding groove consists of the deeply buried A′ pocket and the shallow F′ pocket, which connects the A′ pocket with the outer surface of CD1a. When compared to the CD1a-sulfatide structure, no major conformational differences are observed in the global conformation of CD1a when bound to lipopeptide; key residues that form the binding groove are also not substantially al-

Table 1. Data Collection and Refinement Statistics for the CD1a-Lipopeptide Complex

Data Collection	
Resolution range (Å)[a]	50.0–2.8 (2.9–2.8)
Completeness (%)[a]	94.4 (85.5)
Number of unique reflections	24,273
Redundancy	3.0
R_{sym}[a,b] (%)	9.8 (44.4)
I/σ[a]	14.6 (2.2)
Refinement Statistics	
Number of reflections (F > 0)	22,766
Maximum resolution (Å)	2.8
R_{cryst}[c] (%)	21.7
R_{free}[d] (%)	27.7
Number of Atoms	
Protein	6,031
Lipopeptide ligand	101
N-linked carbohydrate	62
Water	19
Ramachandran Statistics (%)	
Most favored	90.3
Additional allowed	9.3
Generously allowed	0.2
Disallowed	0.3
Rmsd from Ideal Geometry	
Bond length (Å)	0.013
Bond angles (°)	1.48
Average B Values (Å^2)[e]	
Protein	54.5
Lipid ligand	74.9

[a] Numbers in parentheses refer to the highest resolution shell.
[b] $R_{sym} = (\Sigma_h\Sigma_i|I_i(h) - \langle I(h)\rangle|(\Sigma_h\Sigma_i I_i(h)) \times 100$, where $\langle I(h)\rangle$ is the average intensity of i symmetry-related observations with reflections with Bragg index h.
[c] $R_{cryst} = (\Sigma_{hkl}|F_o - F_c|/\Sigma_{hkl}|F_o|) \times 100$, where F_o and F_c are the observed and calculated structure factors, respectively, for all data.
[d] R_{free} was calculated as for R_{cryst} but on 5% of data excluded before refinement.
[e] B values were calculated with the CCP4 program TLSANL (CCP4, 1994; Howlin et al., 1993).

tered despite the differences in the number of alkyl chains (one versus two) and the chemical nature of the hydrophilic head groups of these two antigens. The root mean square deviation (rmsd) of the Cα atoms of the α1-α2 domain (A1-180) between the two structures is only 0.42 Å. Thus, the CD1a binding groove appears to be a rather rigid cavity in which two ligands of quite different chemical composition maneuver to optimally fit the available pockets.

As the central lysine residue of the peptidic headgroup is N amide-linked through its ϵ amino group to the C_{16} fatty acyl chain, the total length of the alkyl chain then extends to 21 carbons. The lipopeptide ligand is bound mainly through van der Waals contacts, with the C_{21} alkyl chain inserted so that its terminal methyl unit is located at the end of the A′ pocket (Figure 2B). Similar to the sulfatide, most of the dihedral angles of the alkyl chain of the lipopeptide in both molecules (A and B) are in preferred conformations, close to their minimum energy state. A few dihedral angles do not adopt these conformations as they are, together with the gauche conformers, necessary for the alkyl chain to adopt both the U-shape conformation around the central pole of the A′ pocket (Phe70, Val12) and the curved entry into the F′ pocket. Of the 36 dihedral angles, 21 are in *trans* (lowest energy) and 6 in gauche$^+$ or gauche$^-$ (6 kJ/mol higher energy).

The connecting lysine is situated at the A′-F′ intersection and anchors the branched headgroup at the bottom of the F′ pocket (Figure 2B). Both peptidic branches are folded up so as to form a U shape, whereas the ends of the branches, representing the tips of the U, protrude to the T cell recognition surface of CD1a (Figure 3B). The N-aryl branch (Figure 2C) is located on the medial side of the F′ pocket close to Arg73 and stabilized by hydrogen bonding. The lysine branch is on the lateral side of the F′ pocket, close to Tyr84. Compared to bacterial DDMs, the synthetic lipopeptide has additional phenyl groups and the ter-butyl group attached to the distal lysine, which are not ordered in the electron density maps and were, therefore, not included in the final structure (Figures 2C, 3A, and 3B). The B values for the ligand and the protein residues are almost identical in the A′ pocket and gradually increase as the ligand protrudes from the A′-F′ junction into the F′ pocket and up to the T cell recognition surface. This phenomenon of increasing B values and partial loss of electron density for the ligand at the T cell recognition surface of CD1a was also observed in CD1a-sulfatide structure (Zajonc et al., 2003) and is most likely due to the paucity of contacts for the headgroup, especially for those moieties protruding farthest from the CD1a surface.

Hydrogen bonding occurs mainly between the N-aryl and the lysine branches of the peptidic moiety of the ligand with Arg73 and Ser77 of the A′-F′ intersection and the F′ pocket, respectively (Figure 4). Arg73, located on the α1 helix, plays a crucial role in hydrogen bonding, as it simultaneously interacts with Glu154 and Thr158 of the α2 helix and with the oxazoline ring of the N-aryl branch of the lipopeptide. Consistent with a role of Arg73 in stabilizing the ligand for T cell recognition, hydrolysis of the oxazoline ring in DDM-838 to form DDM-856 results in loss of T cell recognition (Figure 1). On the lateral side of the F′ pocket opening, the lysine branch is stabilized by the backbone oxygen of Ser77, which forms the only hydrogen bond between the α1 helix and the peptide bond in the lysine branch.

Comparison of Sulfatide versus Synthetic Lipopeptide

In addition to the basic differences in the chemical composition of their hydrophilic head groups, CD1a bound sulfatides and lipopeptides differ in the number (two versus one) and overall length (C_{36} versus C_{21}) of their alkyl chains. A comparison of their conformations when bound in the CD1a groove provides some insight into the mechanisms by which CD1a can present chemically diverse ligands (Figure 5A). The distal ends of the alkyl chain of the lipopeptide ligand and the sphingosine base of the sulfatide ligand are both inserted into the A′ pocket such that they occupy a similar location and conformation. However, the proximal ends of these two alkyl chains take very different paths across the A′-F′ junction. In the sulfatide structure, the sphingosine chain bends laterally and upward to make a sharp

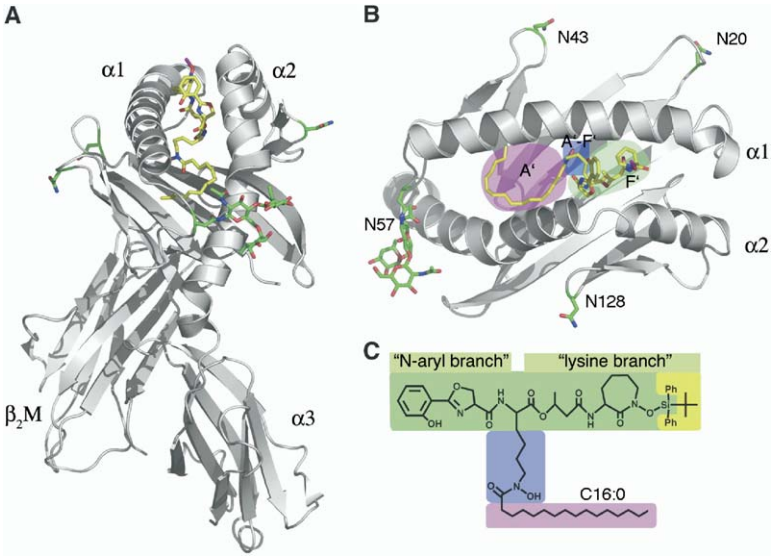

Figure 2. Overview of the CD1a-Lipopeptide Structure

(A) Cartoon representation (front view) of the CD1a (α1–α3 domains)-β_2M heterodimer (light gray) with bound lipopeptide JH-02215 (yellow). N-linked glycosylation sites and carbohydrates are in green. Atom colors for all structural representations: yellow, carbon; red, oxygen; blue, nitrogen; and pink, silicon.

(B) Top view looking into the CD1a binding groove. The individual pockets of the binding groove are color coded to highlight the different pockets (pink, A′; green, F′; and blue A′-F′ junction). N20, N43, N57, and N128 represent the four N-linked glycosylation sites. Only for N57 could two sugars be modeled, with presumed disorder for the sugars at the other glycosylation sites.

(C) Chemical structure of the synthetic lipopeptide JH-02215. The chemical elements of the ligand are colored to match the corresponding regions of the binding groove (B) in which the different moieties of the ligand (alkyl chain, acylated lysine, and peptide) are bound. Yellow-colored chemical groups are not present in the final structure due to a lack of any convincing electron density and, hence, are likely disordered. The terms N-aryl branch and lysine branch are used throughout the article to describe the two different segments of the peptidic headgroup of the ligand.

S-shaped curve, whereas the lysine moiety of the lipopeptide takes a relatively straight path across the A′-F′ junction (Figure 5A, right). The extra curve in the sulfatide ligand serves to position the sulfogalactosyl headgroup more laterally and superiorly in the F′ pocket, so that this hydrophilic moiety is located near the TCR contact surface. Also, the S-shaped kink buries a larger number of methylene units of the longer sulfatide chain within the groove, where the ligand makes extensive van der Waals contacts with the interior of the groove. Comparison of structures in the F′ pocket shows that the N-aryl branch of the lipopeptide is located in a position similar to the sulfogalactosyl moiety of the sulfatide ligand. Surprisingly, the lysine branch mimics the path and conformation of the corresponding fatty acyl chain of the sulfatide (Figure 5A, left). Many of the residues that are involved in hydrogen bonding with the galactose headgroup of the sulfatide to CD1a also function in stabilizing the N-aryl and cyclized lysine moieties of the lipopeptide.

Comparison of Natural (DDM-838) versus Synthetic (JH-02215) Lipopeptide

Overall, good agreement is found between the crystal structure of CD1a-JH-02215 and the previously proposed, energy-minimized model of CD1a bound to mycobacterial DDM-838 (Figure 5B) (Moody et al., 2004). In particular, the terminal portions of the alkyl chain are

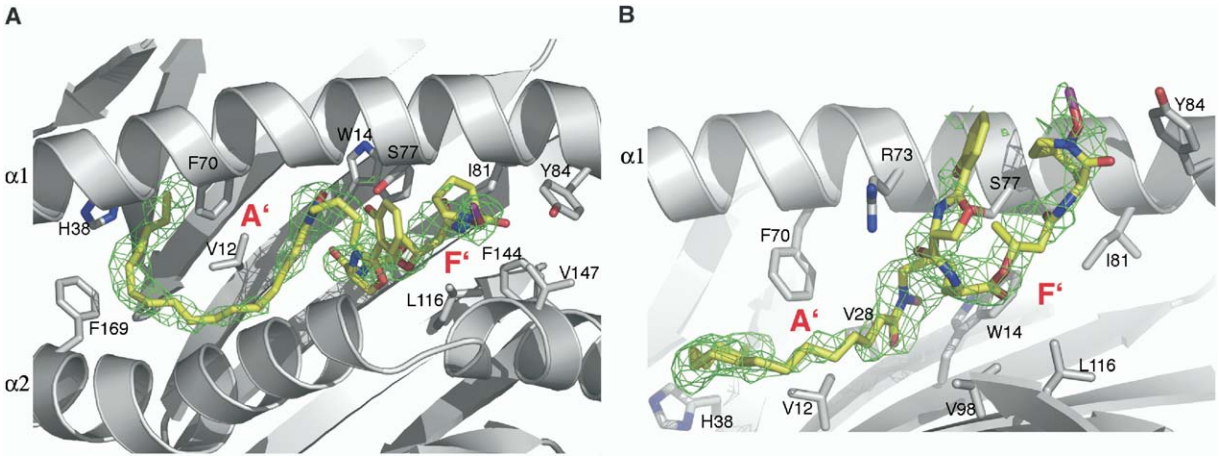

Figure 3. Conformation of the Lipopeptide JH-02215 in the CD1a Binding Groove

After omitting the lipopeptide ligand coordinates, a shake-omit $F_O - F_C$ map was calculated and contoured at 1.8 σ as a green mesh around the ligand (yellow). Several important contact residues of the A′ and F′ pocket (A′ and F′, respectively) are depicted and labeled.

(A) View looking down into the binding groove (TCR view).

(B) Side view, after removing the α2 helix for clarity.

The N-aryl-branch is less well ordered than the aliphatic backbone or the lysine-branch. No electron density for the ter-butyl and phenyl groups extending from the end of the lysine branch of the synthetic DDM (highlighted in yellow in Figure 2C) is visible.

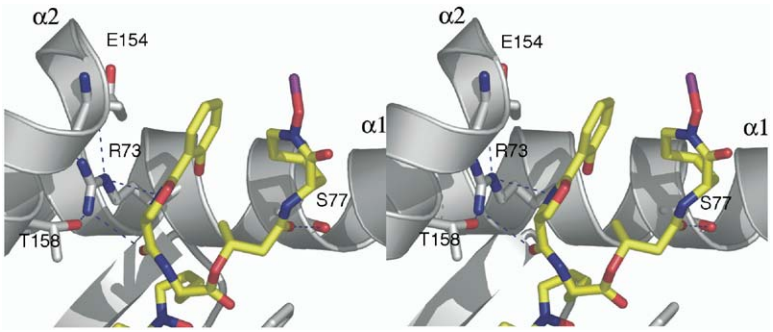

Figure 4. Stereoview of the Hydrogen-Bond Network Stabilizing the Peptidic Headgroup of the Ligand

The lipopeptide ligand bound in the CD1a binding groove is shown in a side view, where the first half of the α2 helix (residues 137–152) is omitted for clarity. Hydrogen bonds between the protein residues were assessed with the program HBPLUS (McDonald and Thornton, 1994) and drawn as a blue dotted line. Hydrogen bonds between the protein and the lipopeptide ligand were assessed with the program CONTACT as part of the CCP4 suite (CCP4, 1994) based on the distance and termed potential, as they do not strictly meet the angle requirements. Arg73 (R73) serves as the major hydrogen donor as it interacts with Glu154 (E154) and Thr158 (T158) of the α2 helix and partially stabilizes the N-aryl branch of the ligand.

located in the A′ pocket, and the relative orientation of the N-aryl and lysine branches on the medial and lateral walls of the F′ pocket are essentially the same in both structures. The main difference between the structures relates to the depth of tethering of the U-shaped lipopeptide moiety within the F′ pocket. This difference is best explained by the differences in length of the alkyl chain found in JH-02215 (C_{16}) and DDM-838 (C_{20}). The shorter fatty acid chain in the CD1a-JH-02215 structure lipid takes a direct path across the A′-F′ junction, whereas the DDM-838 model depicts the lipid in an S-shaped conformation, similar to the known path of the longer alkyl chain in the sulfatide. Furthermore, the shorter alkyl chain of JH-02215 tethers the peptidic headgroup more deeply within the F′ pocket, which results in more limited exposure of the peptide termini at the TCR contact surface (Figure 5B, left).

Model of the CD1a-Lipopeptide-TCR Trimolecular Complex

A model of the DDM-838 lipopeptide-specific αβ TCR (CD8-2) was constructed, by analogy to MHC class I-TCR complex structures, in order to facilitate understanding of the mechanism of CD1-ligand recognition

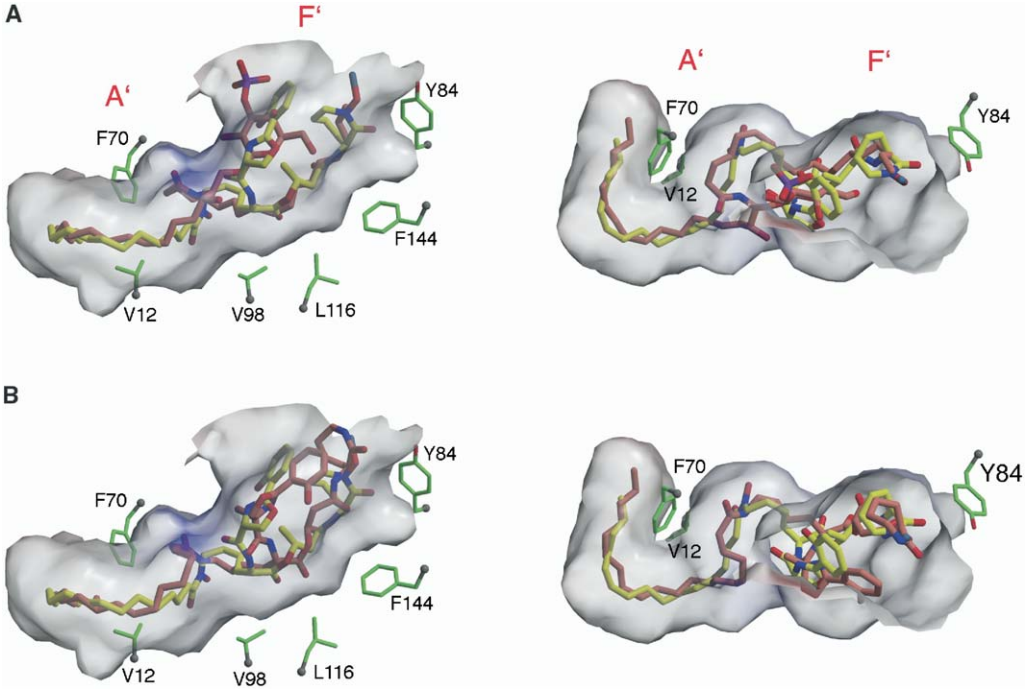

Figure 5. Comparative Binding Analysis of Different CD1a Ligands

Electrostatic surfaces of the CD1a binding groove are shown as transparent binding pockets with bound ligands. The electrostatic surface potential was calculated in GRASP (Nicholls et al., 1991) (–15 to +15 *kT*/e) and demonstrates the highly hydrophobic nature of the binding groove. Red is electronegative and blue is electropositive. Superimposition of the different ligands is shown in a side view (left) and in a top view looking into the groove (right). A′ and F′ pockets are labeled as such. The ligand only protrudes to the surface in the F′ pocket but is completely buried in the A′ pocket and A′-F′ junction.

(A) Superimposition of the sulfatide (salmon) with JH-02215 (yellow) using their crystal structure coordinates 1ONQ and 1XZ0.

(B) Superimposition of the modeled DDM-838 structure (pink) with the crystal structure JH-02215 (yellow).

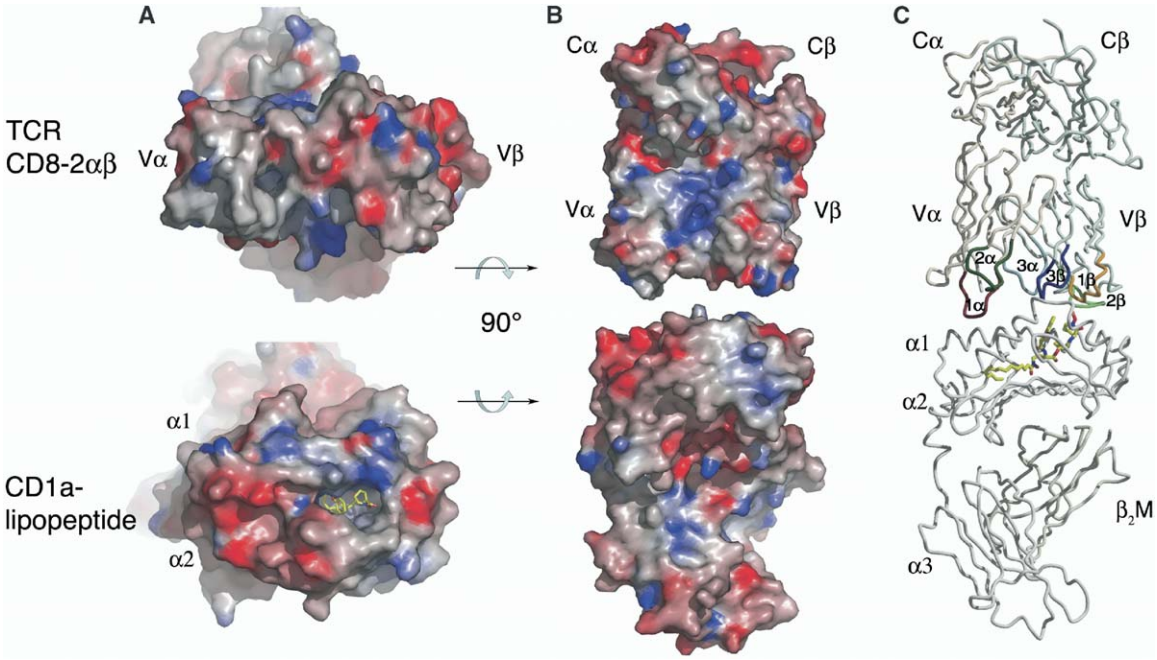

Figure 6. TCR-CD1a-Lipopeptide Complex Model

A model of a TCR CD1a-lipopeptide trimolecular complex was computed as described in the Experimental Procedures. The top panel shows a model of the CD1a-dideoxymycobactin specific αβ TCR CD8-2, whereas the bottom panel depicts the CD1a-JH-02215 lipopeptide crystal structure.

(A and B) Electrostatic surface representation of CD1a and TCR CD8-2 is shown. Surfaces and electrostatic potentials were calculated in GRASP (Nicholls et al., 1991). Red is electronegative and blue is electropositive (–15 to +15 *kT*/e). (A) View of the contact surfaces that are involved in TCR-CD1a complex formation. Top, view down onto the TCR CDR loops of the variable α and β domain (Vα and Vβ, respectively). Bottom, view looking down into the CD1a binding groove, with bound ligand in yellow. Note how little lipopeptide antigen is visible. (B) Front view of the modeled TCR-CD1a complex, with the TCR and CD1 slightly pulled apart.

(C) Front view of the TCR-CD1a complex interaction similar to (B). The Cα atom trace is shown in light gray (CD1a-β_2M and TCR α chain) and blue gray (TCR β chain). The complementarity determining regions (CDRs) are colored as follows: CDR1α, brown; CDR2α, green; CDR3α, light blue; CDR1β, orange; CDR2β, light green; and CDR3β, dark blue. The carbon atoms of the lipopeptide ligand JH-02215 are in yellow, oxygens are red, and nitrogens are blue. Note that the α chain contacts primarily CD1a residues, as the lipid below in the A′ pocket is completely buried. Vβ contacts both CD1a and the peptide portion of the ligand.

by CD1-restricted T cells. Compared to MHC class I and class II, in which the peptide ligands are exposed at the MHC surface along the α1-α2 domain, most of the CD1 ligands are occluded by the CD1 protein itself (Batuwangala et al., 2004; Gadola et al., 2002; Zajonc et al., 2003). The antigenic carbohydrate or peptidic epitopes of the glycolipids and lipopeptides, respectively, are the only components that protrude to any extent at the CD1 surface (Figures 6A and 6C), although one of the alkyl chains in the sulfatide bound to CD1a also reaches the surface. In the case of CD1a-restricted T cells, most of the complementarity determining regions (CDRs) only encounter the nonpolymorphic CD1 protein surface, rather than the ligand, especially for Vα recognition of the CD1 surface above the A′ pocket. The CDRs of the α chain (CDRα) are likely to interact primarily with the surface of CD1 on the N- and C-terminal side of the α1 and α2 helices, respectively, where the bound ligand is not exposed. The ligand is only exposed in the F′ pocket towards the C-terminal end of the α1 helix (Figures 6A and 6C), where the CDRs of the β chain (CDRβ) are well positioned to interact with the peptidic headgroup of the lipopeptide (Figure 6C). It is possible that CDR3α, if of appropriate length and conformation, could extend over to this side, as seen in the TCR (A6 and B7) interactions with HLA-A2 peptide complexes (Ding et al., 1998; Garboczi et al., 1996). Although specific interactions between individual residues of trimolecular complex cannot be quantitated at present, the charged residues that surround the entry to the binding groove in CD1a are likely to be located in close proximity to the CDR loops of the β chain, in particular to CDR3β but also to CDR3α, which would also confer specificity for this particular CD1 isotype. The possibility of other CD1-TCR docking orientations cannot be discounted, but it seems reasonable to assume that the CDRs of the TCR and the antigenic epitope of the CD1 bound ligand must come into close proximity for recognition events, similar to TCR docking onto MHC molecules. Slightly different docking orientations of the TCR on MHC class I and class II molecules have been observed (Rudolph and Wilson, 2002) that lead to a slightly different disposition of the CDRs with respect to the presented antigens, but such subtle changes cannot be predicted yet for TCR-CD1 docking models.

Discussion

MHC class I and II proteins present peptidic antigens for which many of the sequence-specific differences in peptide binding are accounted for by allelic polymorphisms

that lead to differences in the antigen binding groove structures to accommodate different ligands. CD1 proteins are nonpolymorphic, yet they present a diversity of antigenic structures including mycolates, diacylglycerols, isoprenoid lipids, sphingolipids, lipopeptides, and small hydrophobic molecules (Beckman et al., 1994; De Silva et al., 2002; Gumperz et al., 2000; Joyce et al., 1998; Kawano et al., 1997; Moody et al., 1997, 2000, 2004; Shamshiev et al., 1999, 2002; Sieling et al., 1995; Wu et al., 2003). Even the same CD1 isoform can present antigens of markedly different structures, as illustrated by CD1a, which presents self-sulfoglycolipids with two alkyl chains and foreign lipopeptides with one alkyl chain to T cells (Moody et al., 2004; Zajonc et al., 2003). Thus, an understanding of how antigen binding grooves of nearly invariant structure can specifically bind such chemically diverse antigens has emerged as a central problem in understanding the structural basis of antigen presentation by CD1. Here, the crystal structure of CD1a bound to a synthetic lipopeptide antigen has provided further insights into the molecular mechanism of antigen binding by showing how alkyl chains of a discrete length are captured via a conserved mechanism in the A′ pocket, whereas the F′ pocket binds the diverse chemical elements of each ligand in different and flexible ways.

Initially, it was not easy to comprehend how CD1a proteins could present both sulfatides that have two alkyl chains and lipopeptides with a single alkyl chain. The CD1a-JH-02215 structure provides a clear and surprising answer to this question by illustrating that the lateral wall of the F′ pocket can accommodate either the N-aryl branch of a lipopeptide or the acyl chain of sulfatides, where these diverse moieties follow nearly the same path (Figure 5A). Furthermore, the longer chain lengths of the alkyl chains in the sulfatide (C_{36}) are accommodated in part by the prolonged S-shaped diversion through the broader A′-F′ portal region, in contrast to the direct path taken by the shorter (C_{21}), single chain JH-02115. Moreover, the anchoring of each lipid terminus at the distal end of the A′ pocket leads to a more subtle mechanism in which the S-shaped conformation at the A′-F′ junction compensates for some variation in the alkyl chain length so as to facilitate correct positioning of the ligand headgroups for TCR contact within the F′ pocket. Antigens that vary from this restricted optimal length by more than the small amount of tolerance provided by the S-shaped diversion would make inadequate or inappropriate contacts with the hydrophobic binding groove, resulting in weaker and less efficient binding. Studies on CD1a presentation of DDM-lipopeptides have shown that DDMs that vary slightly in alkyl chain length and saturation state can all be recognized, but the presence of a single unsaturation at the C_{2-3} position or extension of the acyl chain length from C_{18} to C_{20} surprisingly increases antigenic potency (Moody et al., 2004). The molecular mechanism by which the C_{2-3} *cis* unsaturation increases DDM potency by 40-fold has not been elucidated, but its location at the A′-F′ junction suggests that it could favor an S-shaped configuration. This precise lipid length discrimination of DDM homologs has not been described for α-galactosyl ceramides presented by mCD1d; furthermore, hCD1b can present glucose monoycolates that vary substantially in length from C_{12-80}. These two isoforms lack the blunt-ended A′ pocket seen in CD1a and also have larger F′ pockets that connect back to the A′ pocket (CD1d) or the T′ tunnel (CD1b), which allows accommodation of lipid tails of varied and longer length (Batuwangala et al., 2004; Gadola et al., 2002; Zeng et al., 1997).

These studies also provide evidence that the TCR-mediated discrimination of DDM antigens is focused on both the N-aryl and lysine branches of the peptide and reveals that the termini of these branches are adjacent to the surface of CD1a that is predicted to serve as the TCR contact region (Figure 6). The hydrolysis reaction that renders DDM-856 unrecognizable to T cells involves the methylserine component that forms a hydrogen bond with Arg73 on the α helix so that this alteration likely affects the positioning of the peptide within the F′ pocket (Figures 1 and 5). More generally, the functional data show that the peptidic headgroup controls the T cell response and is consistent with the CD1a-JH-02215 structure where this moiety is exposed at the T cell recognition surface. Thus, a CD1a-restricted TCR, as for MHC-restricted TCRs, appears to be able to scan and respond in a sequence-specific way to peptide ligands. DDM lipopeptides are nonribosomally encoded and invariant in sequence (Crosa and Walsh, 2002). However, the ability of the F′ pocket to flexibly accommodate ligands of varied structure suggests the possibility that other lipopeptides, encoded by mammalian DNA or viral RNA and acylated during posttranslation processing, might constitute an even more diverse array of antigens that can be presented to human αβ T cells.

Experimental Procedures

Protein Expression, Purification, and Crystallization

The extracellular domain of CD1a was expressed, purified, and partially deglycosylated as previously reported (Zajonc et al., 2003). In addition, the C-terminal hexahistidine tag was truncated by carboxypeptidase A digestion. In brief, 8 mg of the heterodimeric CD1a-β_2M protein was incubated for 18 hr at room temperature in 100 mM Tris-HCl buffer at (pH 7.7) with carboxypeptidase A (1 U/mg CD1 protein; Sigma, C9268). An average loss of four histidine residues was observed by mass spectrometry analysis (data not shown). The CD1a-β_2M protein was purified from carboxypeptidase A by anion exchange chromatography on MonoQ in 10 mM Tris-HCl (pH 8.5). Lipopeptide loading was performed as follows: 200 μg of the lipopeptide antigen was dissolved in 40 μl chloroform:methanol (1:1) and mixed vigorously with 1 ml 1% Tween 20 detergent solution. The mixed lipopeptide-Tween 20 micelles were diluted with 4 ml CD1a-β_2M protein in 100 mM potassium phosphate buffer at (pH 6.0) to a final molar ratio of lipopeptide to CD1a of 3:1 and incubated for 18 hr at room temperature under gentle agitation. The CD1a-lipopeptide mix was concentrated to 2 ml by using a 4 ml ultrafiltration device (Amicon, 30 kDa MWCO) and subjected to size exclusion chromatography (SEC) on Superdex S200 HR16/60. Fractions containing monomeric CD1a-β_2M protein were pooled and concentrated to 8.5 mg/ml in 10 mM HEPES at pH 7.5 and 30 mM NaCl. A second batch of protein was prepared by using this protocol and concentrated to 15 mg/ml. Nanodrop crystallization trials (100 nl drops) were performed by a crystallization robot (Syrrx), and initial crystallization conditions were reproduced and optimized manually by using the sitting drop vapor diffusion method. The best crystals were obtained at room temperature by mixing 1 μl protein (15 mg/ml) with 1 μl precipitant solution (24%

polyethylene glycol monomethylether [MPEG] 2000 and 0.1 M Tris-HCl [pH 8.0]).

Structure Determination

Crystals were flash cooled at a temperature of 100 K in mother liquor containing 20% glycerol. Diffraction data from a single crystal were collected at Beamline 8.3.1 of the Advanced Light Source, Berkeley and processed to 2.8 Å with the Denzo-Scalepack suite (Otwinowski and Minor, 1997) in spacegroup $P2_1$ (unit cell dimensions: a = 55.96 Å; b = 43.23 Å; c = 209.94 Å; β = 91.04°). As the unit cell dimensions are almost identical to the CD1a-sulfatide crystal, with two CD1a-β_2M heterodimers per asymmetric unit (molecules A and B), molecular replacement was straightforward using the CD1a-sulfatide structure (1ONQ) as the initial model. Subsequent rigid-body refinement in CNS version 1.1 (Brünger et al., 1998) to a resolution of 3.5 Å resulted in an R_{cryst} of 36.8% (R_{free} of 37.2%). The initial refinement included several rounds of simulated annealing starting at 2000 K, conjugate gradient minimization, and restrained temperature-factor refinement against the maximum likelihood target (Pannu and Read, 1996). Tight NCS restraints were maintained between the two molecules in the asymmetric unit throughout the refinement. Furthermore, lowering the NCS weights at the final stages of the refinement resulted in an increase in R_{free} and was, therefore, not used. The refinement progress was judged by monitoring the R_{free} for crossvalidation (Brünger, 1992). The model was rebuilt into σ_A-weighted *2Fo – Fc* and *Fo – Fc* difference electron density maps by using the program O (Jones et al., 1991). At a later stage of refinement, N-linked carbohydrates were built at two out of the eight total Asn-X-Thr/Ser motifs in molecules A and B. Water molecules were assigned during the refinement in CNS for >3 σ peaks in an *Fo – Fc* map and retained if they satisfied hydrogen bonding criteria and returned *2Fo – Fc* density >1 σ after refinement. Starting coordinates for the lipopeptide ligand were obtained with the molecular modelling system INSIGHT II (Accelrys, Inc.) and then energy minimized for 100 cycles with Discover. The lipopeptide library for REFMAC (Murshudov et al., 1997) was created by using the CCP4 program suite 5 (CCP4, 1994; Potterton et al., 2003). Final refinement steps were performed by using the TLS procedure in REFMAC (Winn et al., 2001) with a total of four anisotropic domains (domains 1 and 3, α1-α2 domain of CD1a including carbohydrates and lipopeptide ligand of molecules A and B; domains 2 and 4, α3 domain and β_2M of molecules A and B) and resulted in improved electron density maps for the lipopeptide ligand and a further drop in R_{free}. To reduce phase bias, a "shake-omit map" was calculated as a difference map by using CCP4. After omission of the lipopeptide ligand, the remaining coordinates were perturbated in Moleman2 (Kleywegt, 1997) to a final rmsd of 0.2 Å. In addition, random shifts of up to 20 $Å^2$ and up to 0.05 were applied to the B values and the occupancies, respectively. The CD1a-lipopeptide structure has a final R_{cryst} = 21.7% and R_{free} = 27.7%. The quality of the model was assessed with the program Molprobity (Lovell et al., 2003).

TCR Modeling

A homology model of the CD1a-dideoxymycobactin restricted $\alpha\beta$ T Cell clone CD8-2 was calculated with the Swiss Model Server (Schwede et al., 2003). The coordinates of the model and the CD1a-lipopeptide structure were superimposed onto the corresponding regions of a HLA-A2 TCR complex (1AO7) (Garboczi et al., 1996). To compensate for steric clashes between both CD1a and TCR model, the TCR coordinates were manually tilted 16° along the vertical axis, as determined with the program LSQKAB (Kabsch, 1976) as part of the CCP4 suite (CCP4, 1994). The final orientation showed a maximal interaction between the CDRs of the TCR and the T cell recognition surface and the lipopeptide ligand of CD1a.

Lipopeptide Synthesis, Purification, and T Cell Assay

The lipopeptide compound JH-02215 was obtained as an intermediate during the synthesis of Mycobactin S (Hu and Miller, 1997). The compound was diluted to a 50 μM solution in 1:1 methanol: water and analyzed by CID-MS/MS to confirm its structure. *M. tuberculosis* strain H37Ra was purchased from Difco (Detroit, MI). Total lipids were purified by extracting the lyophilized bacteria with $CHCl_3$:MeOH 1:2 and then $CHCl_3$:MeOH 2:1. The extractable lipids were separated by centrifugation. The total lipids were dried and mixed vigorously in cold acetone and left on ice for 1 hr, followed by centrifugation. The acetone soluble fraction contained the enriched DDMs. To purify individual molecular species of mycobactin, including DDM-838 and DDM-642, total *M. tuberculosis* lipids were dissolved in 50 mL of 100:25 hexane:chloroform and loaded onto a solid phase extraction silica gel column (10 × g silica, Alltech, Waukegan, IL) followed by elution with 40 mL aliquots of hexane: chloroform:2-propanol:acetic acid in the following ratios: 100:25:5:0, 100:25:10:0, 100:25:15:0 (three times), and 100:25:15:1 (two times). DDM-838 and related compounds eluted in the 100:25:15:1 eluent fraction. The lipids in the DDM-containing fraction were then separated further by using reversed phase high-performance liquid chromatography (RP-HPLC) on a C18 column and eluted with a binary gradient (0 min 20% B, 4 min 20% B, 35 min 60% B, 45 min 60% B, and 50 min 20% B) from 20% to 60% B at 0.7 ml/min (A = 50:30:20:0.02 methanol:acetonitrile:water:trifluoroacetic acid, B = 93:7:0.02 2-propanol:hexane:trifluoroacetic acid). The LC-MS system used was a Thermo Finnigan LCQ Advantage quadrupole ion trap mass spectrometer with a split interface to facilitate fraction collection. The $[M + H]^+$ of DDM-642, m/z 642, was subjected to CID-MS/MS analysis, which yielded collision products corresponding to a form of DDM in which the β-hydroxybuturate-lysine moiety was absent. DDM-856 was generated by hydrolysis of DDM-838 in acetonitrile:water (80:20) with 0.1% TFA for 24 hr, yielding a product with $[M + Na]^+$ m/z 878, and the expected collisionally induced product ions were observed at m/z 740 (loss of salicylic acid accompanied by hydrogen transfer) and m/z 682 (cleavage of the central ester moiety), which indicated that the oxazoline ring had been hydrolyzed.

Monocyte-derived DCs were prepared from peripheral blood by centrifugation over Ficoll-Hypaque (Amersham, Piscataway, NJ), adherence to plastic tissue-culture flask (Falcon, Franklin Lakes, NJ), culture of adherent cells with 300 U/ml of GM-CSF and 200 U/ml IL-4 for 72–96 hr, followed by γ-irradiation (5000 rad) as described. IL-2 release from J.RT-3, transfected with the $\alpha\beta$ chains of CD8-2 T cells (10^5/well with 10 ng/ml of PMA) was measured by culturing T cells and 5 × 10^4 DCs plus antigen in 200 μl/well in 96-well plates. After 24 hr, 50 μl of supernatant was transfered to wells containing 125 μl of media and 10^4 IL-2-dependent HT-2 cells. Cells were cultured for 24 hr before adding 1 μCi [^{3}H] thymidine for an additional 6–24 hr of culture, followed by harvesting and counting β emissions.

Structure Presentation

The program PyMOL was used to prepare Figures 2, 3, and 4 (DeLano, 2002). The programs Molscript (Kraulis, 1991), GRASP (Nicholls et al., 1991), and Raster3D (Merritt and Bacon, 1997) were used to prepare Figures 5 and 6.

Supplemental Data

Supplemental Data including five supplemental figures are available online with this article online at http://www.immunity.com/cgi/content/full/22/2/209/DC1/.

Acknowledgments

We thank the staff of the Advanced Light Source BL 8.3.1 for support with data collection. This study was supported by National Institutes of Health grants GM62116, CA58896 (I.A.W.), AI49313, AR48632 (D.B.M), AI28973 (M.B.B), AI30988, AI054193 (M.J.M.), RR10888 (C.E.C), and a postdoctoral fellowship from the Skaggs Institute for Chemical Biology (D.M.Z.). M.D.M.C was supported by a joint scholarship between the Oxford Glycobiology Institute and The Scripps Research Institute. D.B.M is supported by grants from the Pew Foundation Scholars in the Biomedical Sciences, The Mizutani Foundation for Glycoscience, and the Cancer Research Institute. This is manuscript number 16971-MB of The Scripps Research Institute.

Received: November 11, 2004
Revised: December 16, 2004
Accepted: December 22, 2004
Published: February 22, 2005

References

Batuwangala, T., Shepherd, D., Gadola, S.D., Gibson, K.J., Zaccai, N.R., Fersht, A.R., Besra, G.S., Cerundolo, V., and Jones, E.Y. (2004). The crystal structure of human CD1b with a bound bacterial glycolipid. J. Immunol. *172*, 2382–2388.

Beckman, E.M., Porcelli, S.A., Morita, C.T., Behar, S.M., Furlong, S.T., and Brenner, M.B. (1994). Recognition of a lipid antigen by CD1-restricted $\alpha\beta^+$ T cells. Nature *372*, 691–694.

Brünger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature *355*, 472–475.

Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, R., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. *D54*, 905–921.

Calabi, F., Jarvis, J.M., Martin, L., and Milstein, C. (1989). Two classes of CD1 genes. Eur. J. Immunol. *19*, 285–292.

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

Crosa, J.H., and Walsh, C.T. (2002). Genetics and assembly line enzymology of siderophore biosynthesis in bacteria. Microbiol. Mol. Biol. Rev. *66*, 223–249.

Dascher, C.C., and Brenner, M.B. (2003). Evolutionary constraints on CD1 structure: insights from comparative genomic analysis. Trends Immunol. *24*, 412–418.

De Silva, A.D., Park, J.J., Matsuki, N., Stanic, A.K., Brutkiewicz, R.R., Medof, M.E., and Joyce, S. (2002). Lipid protein interactions: the assembly of CD1d1 with cellular phospholipids occurs in the endoplasmic reticulum. J. Immunol. *168*, 723–733.

De Voss, J.J., Rutter, K., Schroeder, B.G., Su, H., Zhu, Y., and Barry, C.E., 3rd. (2000). The salicylate-derived mycobactin siderophores of *Mycobacterium tuberculosis* are essential for growth in macrophages. Proc. Natl. Acad. Sci. USA *97*, 1252–1257.

DeLano, W.L. (2002). The PyMOL Molecular Graphics System (San Carlos, CA: DeLano Scientific).

Ding, Y.H., Smith, K.J., Garboczi, D.N., Utz, U., Biddison, W.E., and Wiley, D.C. (1998). Two human T cell receptors bind in a similar diagonal mode to the HLA-A2/Tax peptide complex using different TCR amino acids. Immunity *8*, 403–411.

Fischer, K., Scotet, E., Niemeyer, M., Koebernick, H., Zerrahn, J., Maillet, S., Hurwitz, R., Kursar, M., Bonneville, M., Kaufmann, S.H., and Schaible, U.E. (2004). Mycobacterial phosphatidylinositol mannoside is a natural antigen for CD1d-restricted T cells. Proc. Natl. Acad. Sci. USA *101*, 10685–10690.

Gadola, S.D., Zaccai, N.R., Harlos, K., Shepherd, D., Castro-Palomino, J.C., Ritter, G., Schmidt, R.R., Jones, E.Y., and Cerundolo, V. (2002). Structure of human CD1b with bound ligands at 2.3 Å, a maze for alkyl chains. Nat. Immunol. *3*, 721–726.

Garboczi, D.N., Ghosh, P., Utz, U., Fan, Q.R., Biddison, W.E., and Wiley, D.C. (1996). Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. Nature *384*, 134–141.

Gilleron, M., Stenger, S., Mazorra, Z., Wittke, F., Mariotti, S., Bohmer, G., Prandi, J., Mori, L., Puzo, G., and 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.

Gumperz, J.E., and Brenner, M.B. (2001). CD1-specific T cells in microbial immunity. Curr. Opin. Immunol. *13*, 471–478.

Gumperz, J.E., Roy, C., Makowska, A., Lum, D., Sugita, M., Podrebarac, T., Koezuka, Y., Porcelli, S.A., Cardell, S., Brenner, M.B., and Behar, S.M. (2000). Murine CD1d-restricted T cell recognition of cellular lipids. Immunity *12*, 211–221.

Howlin, B., Butler, D.S., Moss, D.S., Harris, G.W., and Driessen, H.P.C. (1993). TLSANL: TLS parameter analysis program for segmented anisotropic refinement of macromolecular structures. J. Appl. Crystallogr. *26*, 622–624.

Hu, J., and Miller, M.J. (1997). Total synthesis of a mycobactin S, a siderophore and growth promotor of *Mycobacterium smegmatis*, and determination of its growth inhibitory activity against *Mycobacterium tuberculosis*. J. Am. Chem. Soc. *119*, 3462–3468.

Hunger, R.E., Sieling, P.A., Ochoa, M.T., Sugaya, M., Burdick, A.E., Rea, T.H., Brennan, P.J., Belisle, J.T., Blauvelt, A., Porcelli, S.A., and Modlin, R.L. (2004). Langerhans cells utilize CD1a and langerin to efficiently present nonpeptide antigens to T cells. J. Clin. Invest. *113*, 701–708.

Jones, T.A., Cowan, S., Zou, J.Y., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. *A47*, 110–119.

Joyce, S., Woods, A.S., Yewdell, J.W., Bennink, J.R., De Silva, A.D., Boesteanu, A., Balk, S.P., Cotter, R.J., and Brutkiewicz, R.R. (1998). Natural ligand of mouse CD1d1: cellular glycosylphosphatidylinositol. Science *279*, 1541–1544.

Kabsch, W. (1976). A solution for the best rotation to relate two sets of vectors. Acta Crystallogr. *A32*, 922–923.

Kawano, T., Cui, J., Koezuka, Y., Toura, I., Kaneko, Y., Motoki, K., Ueno, H., Nakagawa, R., Sato, H., Kondo, E., et al. (1997). CD1d-restricted and TCR-mediated activation of vα14 NKT cells by glycosylceramides. Science *278*, 1626–1629.

Kleywegt, G.J. (1997). Validation of protein models from Cα coordinates alone. J. Mol. Biol. *273*, 371–376.

Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of proteins. J. Appl. Crystallogr. *24*, 946–950.

Lovell, S.C., Davis, I.W., Arendall, W.B., 3rd, de Bakker, P.I., Word, J.M., Prisant, M.G., Richardson, J.S., and Richardson, D.C. (2003). Structure validation by Cα geometry: ϕ ψ and Cβ deviation. Proteins *50*, 437–450.

McDonald, I.K., and Thornton, J.M. (1994). Satisfying hydrogen bonding potential in proteins. J. Mol. Biol. *238*, 777–793.

Merritt, E.A., and Bacon, D.J. (1997). Raster3D: photorealistic molecular graphics. Meth Enzymology *277*, 505–524.

Moody, D.B., Reinhold, B.B., Guy, M.R., Beckman, E.M., Frederique, D.E., Furlong, S.T., Ye, S., Reinhold, V.N., Sieling, P.A., Modlin, R.L., et al. (1997). Structural requirements for glycolipid antigen recognition by CD1b-restricted T cells. Science *278*, 283–286.

Moody, D.B., Ulrichs, T., Muhlecker, W., Young, D.C., Gurcha, S.S., Grant, E., Rosat, J.P., Brenner, M.B., Costello, C.E., Besra, G.S., and Porcelli, S.A. (2000). CD1c-mediated T-cell recognition of isoprenoid glycolipids in *Mycobacterium tuberculosis* infection. Nature *404*, 884–888.

Moody, D.B., Young, D.C., Cheng, T.Y., Rosat, J.P., Roura-Mir, C., O'Connor, P.B., Zajonc, D.M., Walz, A., Miller, M.J., Levery, S.B., et al. (2004). T cell activation by lipopeptide antigens. Science *303*, 527–531.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum likelihood method. Acta Crystallogr. *D53*, 240–255.

Nicholls, A., Sharp, K.A., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins *11*, 281–296.

Otwinowski, Z., and Minor, W. (1997). HKL: Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Pannu, N.S., and Read, R.J. (1996). Improved structure refinement through maximum likelihood. Acta Crystallogr. *A52*, 659–668.

Porcelli, S.A. (1995). The CD1 family: a third lineage of antigen-presenting molecules. Adv. Immunol. *59*, 1–98.

Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003). A graphical user interface to the *CCP4* program suite. Acta Crystallogr. *D59*, 1131–1137.

Quadri, L.E. (2000). Assembly of aryl-capped siderophores by modular peptide synthetases and polyketide synthases. Mol. Microbiol. *37*, 1–12.

Rosat, J.P., Grant, E.P., Beckman, E.M., Dascher, C.C., Sieling, P.A., Frederique, D., Modlin, R.L., Porcelli, S.A., Furlong, S.T., and Brenner, M.B. (1999). CD1-restricted microbial lipid antigen-specific recognition found in the $CD8^+$ $\alpha\beta$ T cell pool. J. Immunol. *162*, 366–371.

Rudolph, M.G., and Wilson, I.A. (2002). The specificity of TCR/pMHC interaction. Curr. Opin. Immunol. *14*, 52–65.

Schwede, T., Kopp, J., Guex, N., and Peitsch, M.C. (2003). SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. *31*, 3381–3385.

Shamshiev, A., Donda, A., Carena, I., Mori, L., Kappos, L., and De Libero, G. (1999). Self glycolipids as T-cell autoantigens. Eur. J. Immunol. *29*, 1667–1675.

Shamshiev, A., Gober, H.J., Donda, A., Mazorra, Z., Mori, L., and De Libero, G. (2002). Presentation of the same glycolipid by different CD1 molecules. J. Exp. Med. *195*, 1013–1021.

Sieling, P.A., Chatterjee, D., Porcelli, S.A., Prigozy, T.I., Mazzaccaro, R.J., Soriano, T., Bloom, B.R., Brenner, M.B., Kronenberg, M., Brennan, P.J., et al. (1995). CD1-restricted T cell recognition of microbial lipoglycan antigens. Science *269*, 227–230.

Skold, M., and Behar, S.M. (2003). Role of CD1d-restricted NKT cells in microbial immunity. Infect. Immun. *71*, 5447–5455.

Snow, G.A., and White, A.J. (1969). Chemical and biological properties of mycobactins isolated from various mycobacteria. Biochem. J. *115*, 1031–1050.

Vincent, M.S., Gumperz, J.E., and Brenner, M.B. (2003). Understanding the function of CD1-restricted T cells. Nat. Immunol. *4*, 517–523.

Winn, M.D., Isupov, M.N., and Murshudov, G.N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. Acta Crystallogr. *D57*, 122–133.

Wooldridge, K.G., and Williams, P.H. (1993). Iron uptake mechanisms of pathogenic bacteria. FEMS Microbiol. Rev. *12*, 325–348.

Wu, D.Y., Segal, N.H., Sidobre, S., Kronenberg, M., and Chapman, P.B. (2003). Cross-presentation of disialoganglioside GD3 to natural killer T cells. J. Exp. Med. *198*, 173–181.

Zajonc, D.M., Elsliger, M.A., Teyton, L., and Wilson, I.A. (2003). Crystal structure of CD1a in complex with a sulfatide self antigen at a resolution of 2.15 Å. Nat. Immunol. *4*, 808–815.

Zeng, Z., Castano, A.R., Segelke, B.W., Stura, E.A., Peterson, P.A., and Wilson, I.A. (1997). Crystal structure of mouse CD1: An MHC-like fold with a large hydrophobic binding groove. Science *277*, 339–345.

Accession Numbers

Coordinates and structure factors for the CD1a-lipopeptide complex have been deposited in the Protein Data Bank under accession code 1XZ0.