# Supplementary Material

# Invariant NKT cells recognize glycolipids from pathogenic Gram-positive bacteria

Yuki Kinjo[1, 2,13], Petr Illarionov[3,13], José Luis Vela[1,13], Bo Pei[1], Enrico Girardi[4], Xiangming Li[5], Yali Li[4], Masakazu Imamura[6], Yukihiro Kaneko[2], Akiko Okawara[2], Yoshitsugu Miyazaki[2], Anaximandro Gómez-Velasco[7], Paul Rogers[8], Samira Dahesh[9], Satoshi Uchiyama[9], Archana Khurana[1], Kazuyoshi Kawahara[10], Hasan Yesilkaya[11], Peter W. Andrew[11], Chi-Huey Wong[6], Kazuyoshi Kawakami[12], Victor Nizet[9], Gurdyal S. Besra[3], Moriya Tsuji[5], Dirk M. Zajonc[4], Mitchell Kronenberg[1]

[1] Division of Developmental Immunology, La Jolla Institute for Allergy & Immunology, La Jolla, California 92037, USA.

[2] Department of Chemotherapy and Mycoses, National Institute of Infectious Diseases, Tokyo, 162-8640, Japan.

[3] School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom.

[4] Division of Cell Biology, La Jolla Institute for Allergy & Immunology, La Jolla, California 92037, USA.

Nature Immunology doi:10.1038/ni.2096

[5] HIV and Malaria Vaccine Program, Aaron Diamond AIDS Research Center, 455 First Avenue, New York, NY 10016, USA.

[6] Department of Chemistry and Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California 92037, USA.

[7] Division of Infectious Diseases, Department of Medicine, University of British Columbia, 2733 Heather Street, Vancouver, V5Z 3J5, Canada

[8] Kyowa Hakko Kirin California, Inc., La Jolla, California 92037, USA

[9] Department of Pediatrics and Skaggs School of Pharmacy & Pharmaceutical Sciences, University of California, San Diego, La Jolla, California 92093, USA.

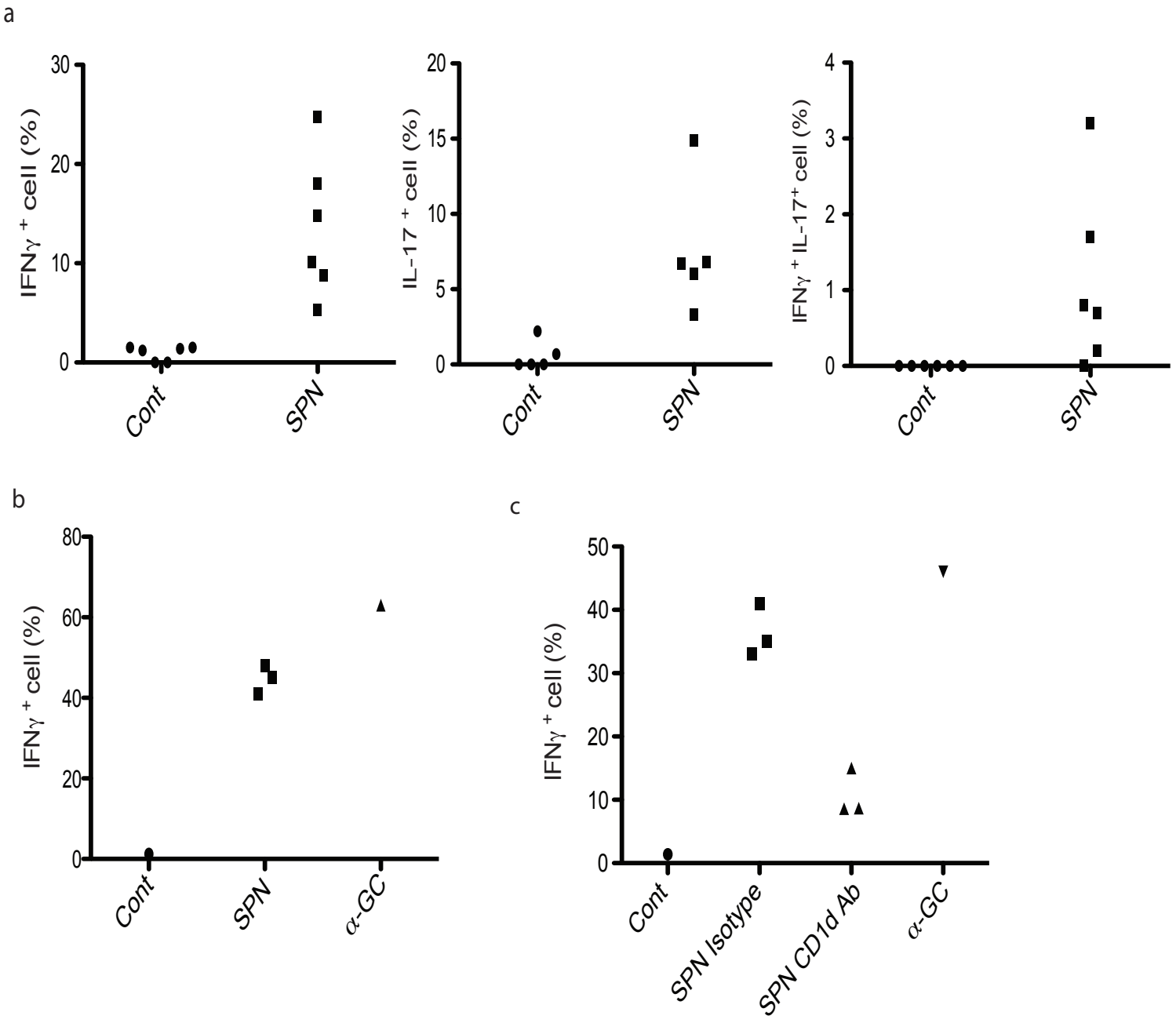
[10] Department of Applied Material and Life Science, College of Engineering, Kanto Gakuin University, Yokohama, Kanagawa 236-8501, Japan.

[11] Department of Infection, Immunity and Inflammation, University of Leicester, Leicester, POB 138, University Road, Leicester, LE1 9HN, UK.

[12] Department of Medical Microbiology, Mycology and Immunology, Tohoku University Graduate School of Medicine, Sendai, Miyagi 980-8575, Japan.

[13]YK, PI and JLV contributed equally to this work.

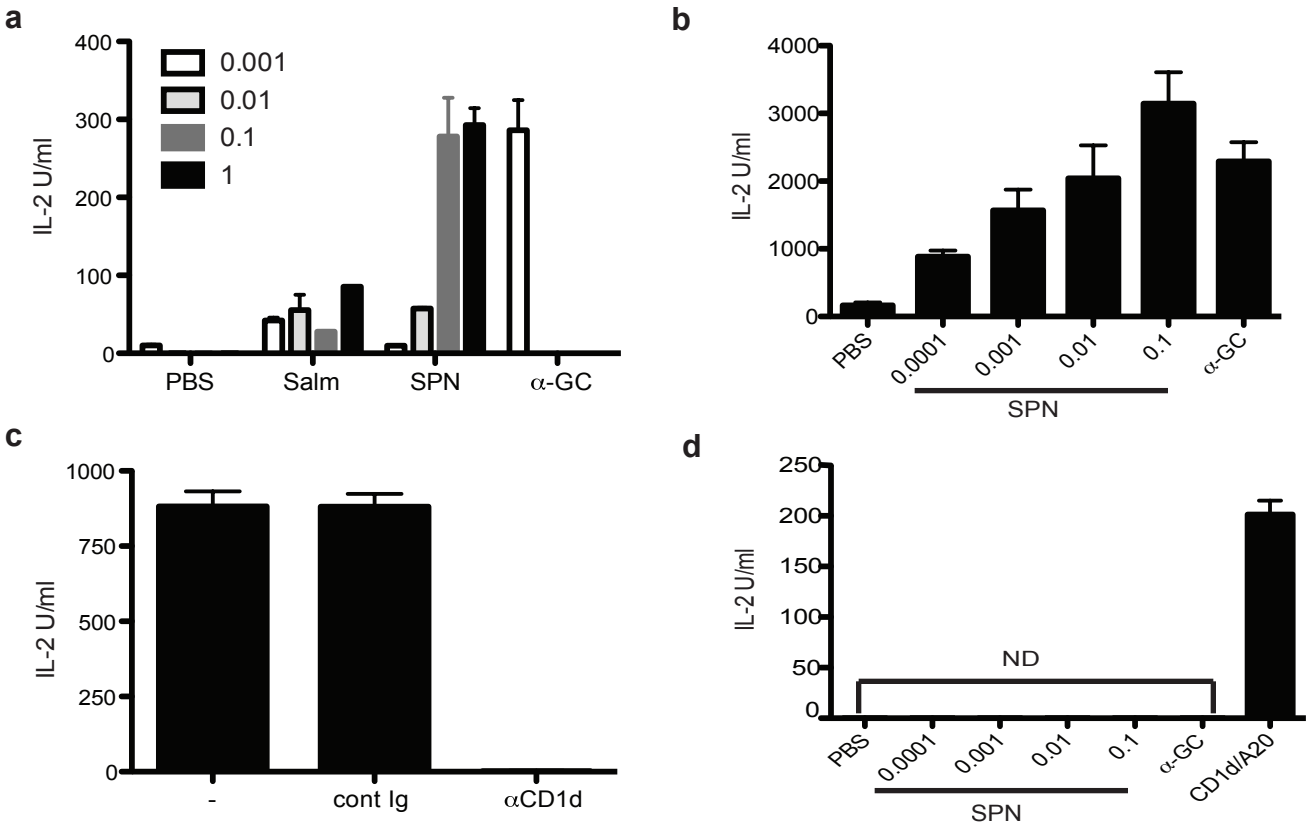
Correspondence should be addressed to M.K. (mitch@liai.org)




**Supplementary Figure 1. In vivo responses of Vα14 *i*NKT cells to *S. pneumoniae*.**

(a) The percentage of intracellular cytokine positive cells among α-GalCer loaded CD1d tetramer positive cells in lung measured at 13h after intratracheal *S. pneumoniae* infection. SPN; *S. pneumoniae*, Each symbol represents cells combined from at least five mice from one experiment.

(b, c) The percentage of intracellular cytokine positive cells among α-GalCer loaded CD1d tetramer positive cells in spleen measured at 6h after intravenous *S. pneumoniae* infection. Each symbol represents individual mouse from one representative experiment. SPN; *S. pneumoniae*, α-GC; α-GalCer, Experiments were done two times.




**Supplementary Figure 2. Responses of CD1d reactive hybridomas to CD1d coated plates incubated with sonciates of *S. pneumoniae* strain URF 918**.

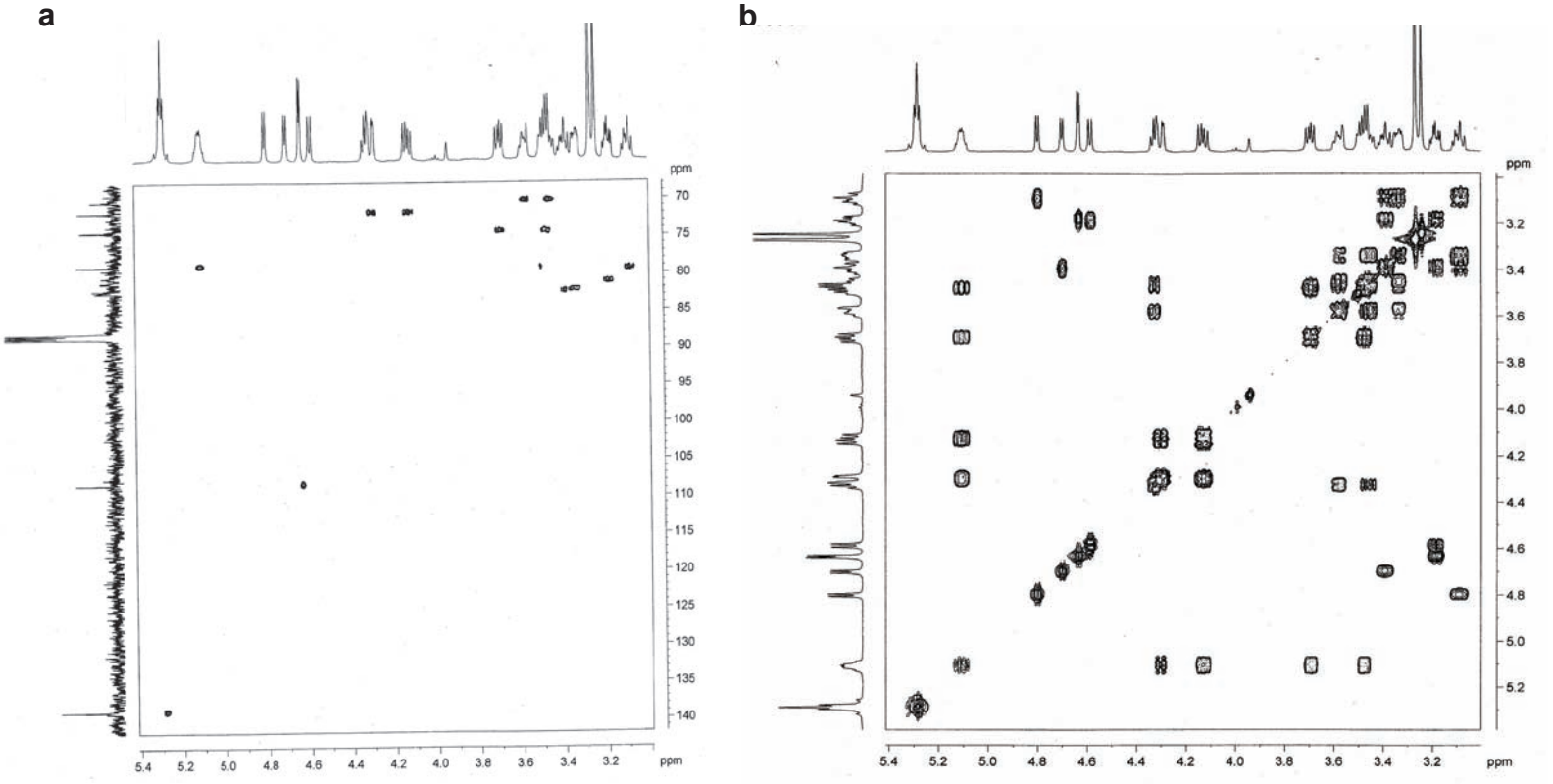
(a) IL-2 release by Vβ8.2/Vα14 *i*NKT cell hybridoma 1.2 to the indicated fold dilutions of sonicates of *S. pneumoniae* (SPN) or *Salmonella typhimurium* (Salm). Positive controls were CD1d coated plates loaded with α-GalCer (α-GC) at 4ng/well and negative controls were CD1d coated plates loaded with phosphate buffered saline (PBS). The amount of sonicates 0.001, 0.01, 0.1 and 1 was equivalent to material from 10^5, 10^6, 10^7, 10^8 bacteria/well, respectively. Each bar shows mean ± SEM from duplicate wells.

(b) The response of Vβ8.2/Vα14 *i*NKT cell hybridoma 2C12 to sonicates of SPN is indicated as for hybridomas 1.2. Each bar shows mean ± SD from triplicate wells.

(c) Blockade of the 2C12 response to the SPN sonicate with addition of the anti-CD1d monoclonal antibody 1B1. Each bar shows mean ± SEM from triplicate wells.

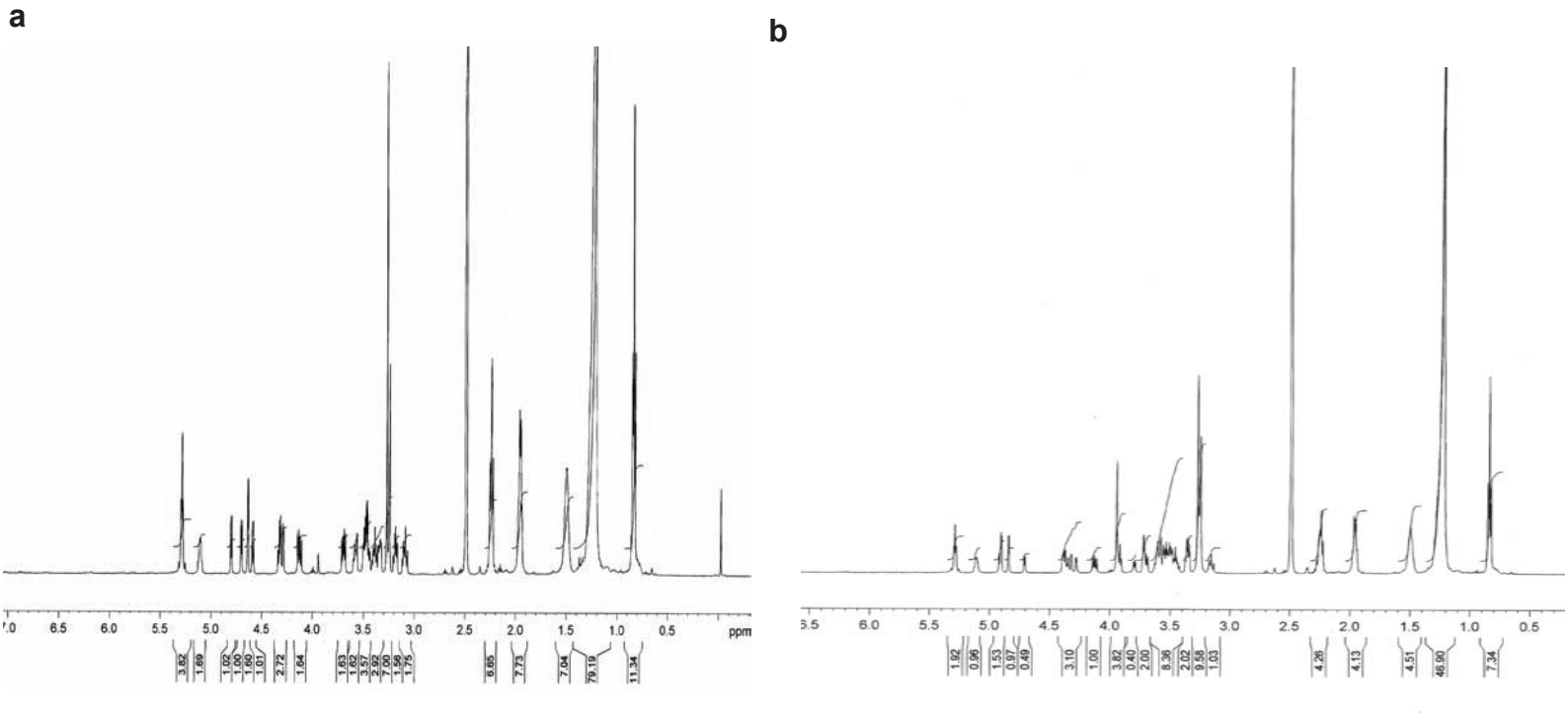
(d) Hybridoma 19, a CD1d reactive cell that does not have the invariant Vα14 TCR does not react to the SPN sonicate or αGalCer, but reacts to A20 cells transfected to express CD1d. Each bar shows mean ± SD from triplicate wells.

All experiments were repeated at least two times.




**Supplementary Figure 3. Two-dimensional NMR spectra of monoglucosyl glycolipid purified from *S. pneumoniae*.**

(**a**) Two-dimensional 1H -13C HSQC NMR spectra of SPN-Glc-DAG recorded in $d_6$-$Me_2SO$-$CDCl_3$ (5:1, v/v) at 300K.
(**b**) Two-dimensional 1H -1H COSY NMR spectra of SPN-Glc-DAG recorded in $d_6$-$Me_2SO$-$CDCl_3$ (5:1, v/v) at 300K.

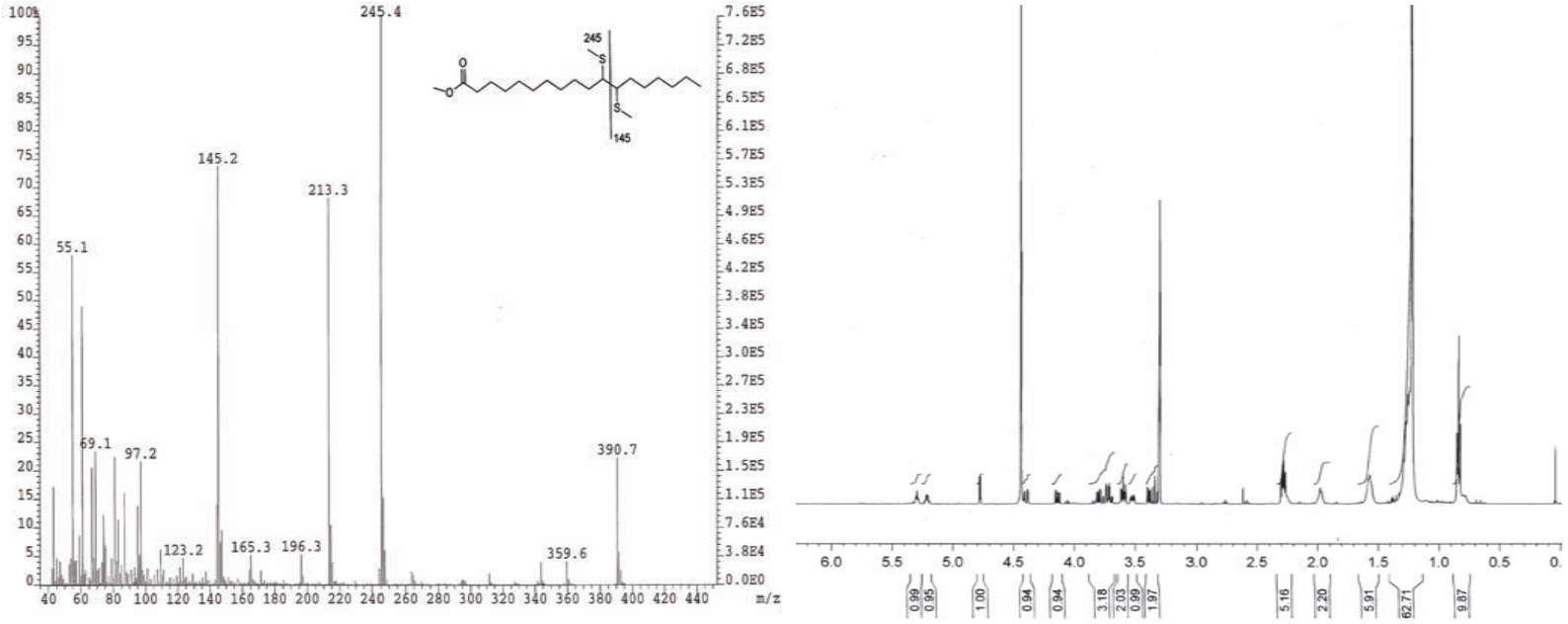



**Supplementary Figure 4. NMR analysis of *S. pneumoniae* glycolipids**.

(**a**) 1H -NMR spectra of SPN-Glc-DAG recorded in $d_6$-$Me_2SO$-$CDCl_3$ (5:1, v/v) at 300K.
(**b**) 1H -NMR spectra of SPN-Gal-Glc-DAG recorded in $d_6$-$Me_2SO$-$CDCl_3$ (5:1, v/v) at 300K.

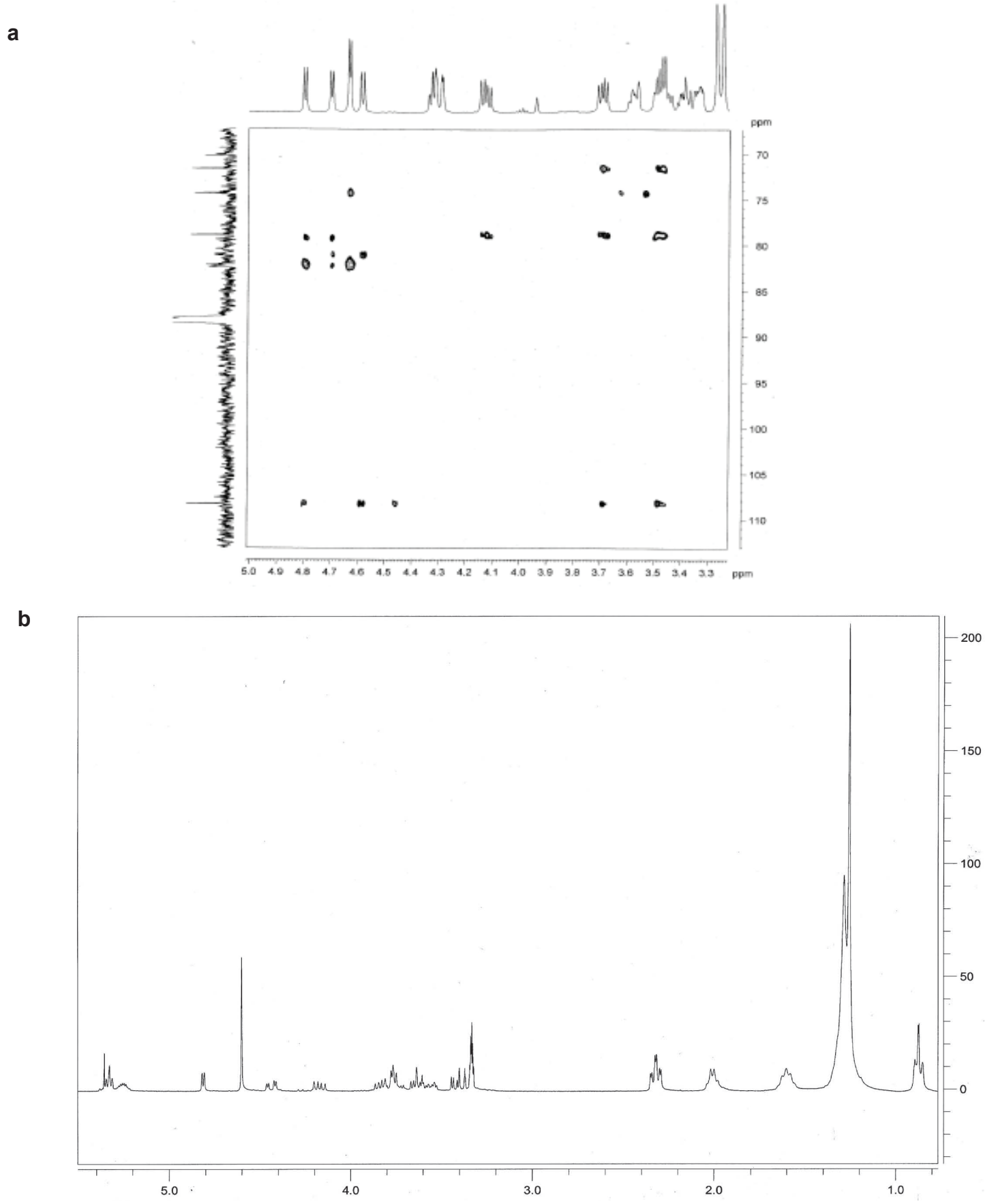
a

b



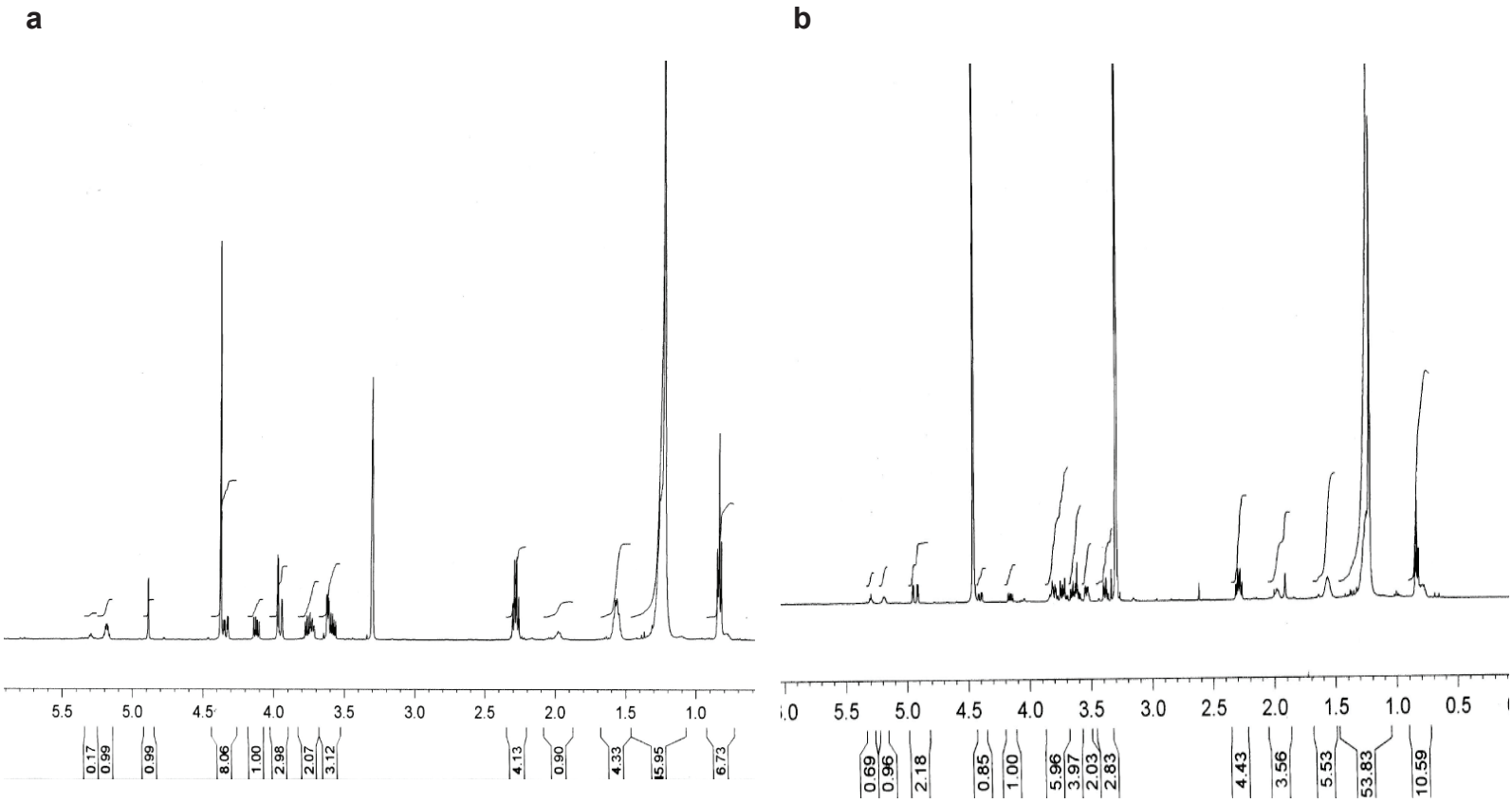

**Supplementary Figure 5. Structural analysis of glycolipids purified from *S. pneumoniae* and GBS.**

(a) Gas chromatography-electron impact ionization mass spectra of dimethyl disulfide derivative of SPN-Glc-DAG cis-vaccenic acid methyl ester.
(b) 1H -NMR spectra of GBS-Glc-DAG recorded in CDCl3-CD3OD (5:3, v/v) at 300K.



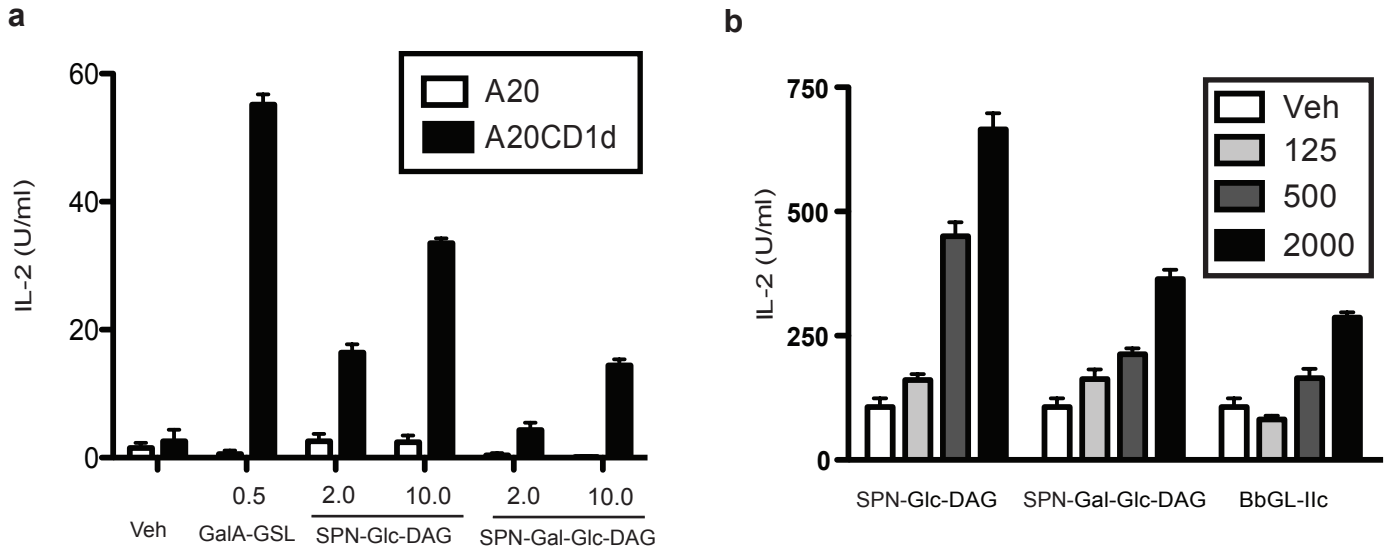

**Supplementary Figure 6. Structural analysis of purified and synthetic *S. pneumoniae* glycolipids.**

(**a**) Two-dimensional 1H -13C HMBC NMR spectra of SPN-Glc-DAG recorded in $d_6$-$Me_2SO$-$CDCl_3$ (5:1, v/v) at 300K.
(**b**) 1H -NMR spectra of Glc-DAG-s2 recorded on a Bruker AV300 (300 MHz) in $CDCl_3$-$CD_3OD$ (5:3, v/v) at 300K.




## Supplementary Figure 7. Structural analysis of glycolipids purified from *S. suis* and GBS.

(**a**) 1H -NMR spectra of SSu-Man-DAG recorded in CDCl3-CD3OD (5:3, v/v) at 300K.
(**b**) 1H -NMR spectra of GBS12-Glc-Glc-DAG recorded in CDCl3-CD3OD (5:3, v/v) at 300K.



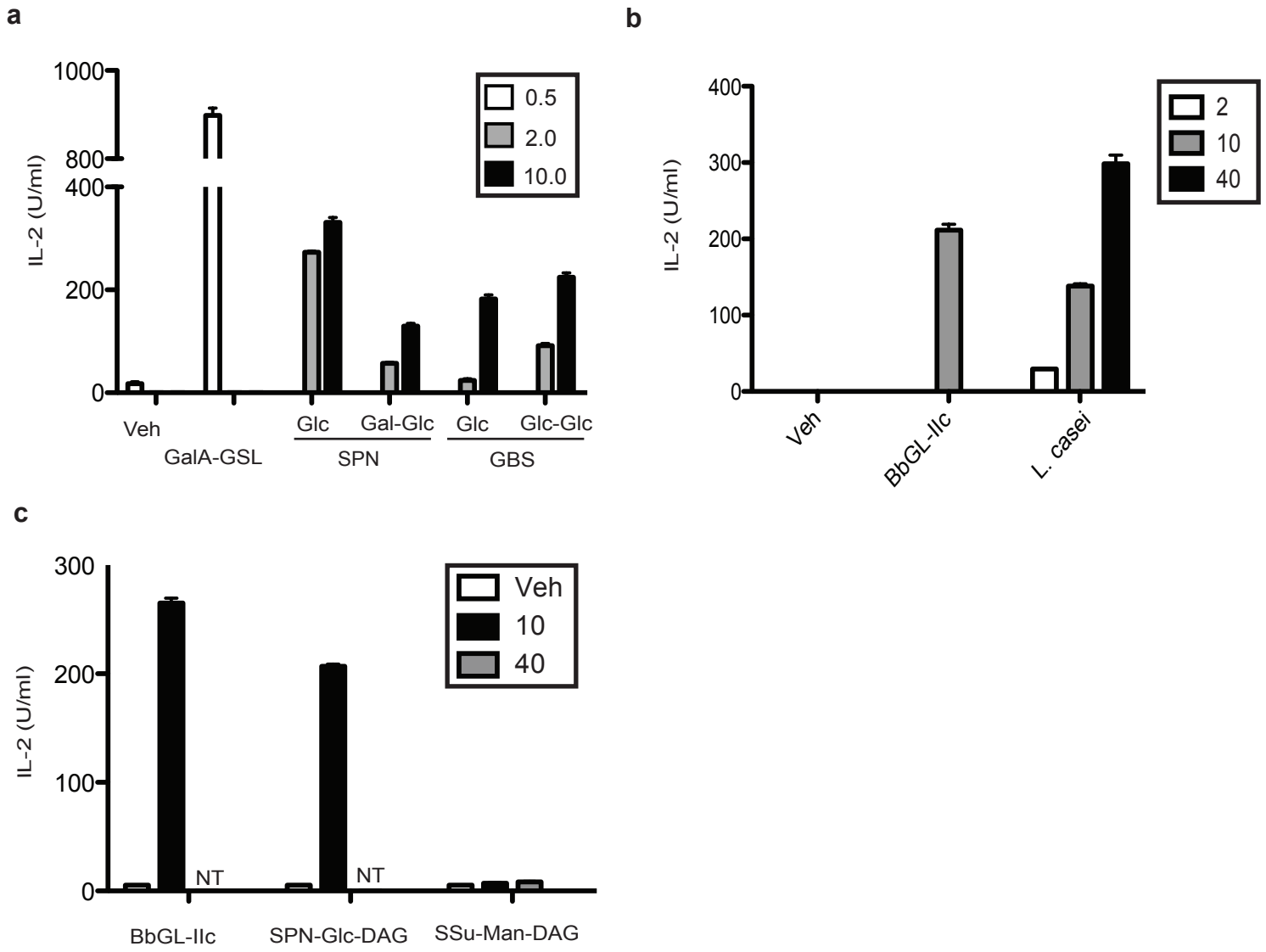

## Supplementary Figure 8. IL-2 response of a Vβ10/Vα14 *i*NKT cell hybridoma to purified *S. pneumoniae* glycolipids.

IL-2 response of Vβ10/Vα14 *i*NKT cell hybridoma 1.4 to A20 CD1d transfectants (**a**) or CD1d coated plates (**b**) was measured in the presence of the indicated glycolipids.

(a) A20 CD1d transfectants were pulsed with the indicated concentrations in μg/ml of the *S. pneumoniae* DAG antigens or the synthetic *Sphingomonas* antigen GalA-GSL.

(**b**) Coated plates were incubated with the indicated concentration in ng/well of the *S. pneumoniae* DAG antigens or the synthetic *B. burgdorferi* antigen BbGL-IIc.

Each bar shows mean ± SEM from triplicate wells. Experiments were performed at least two times.



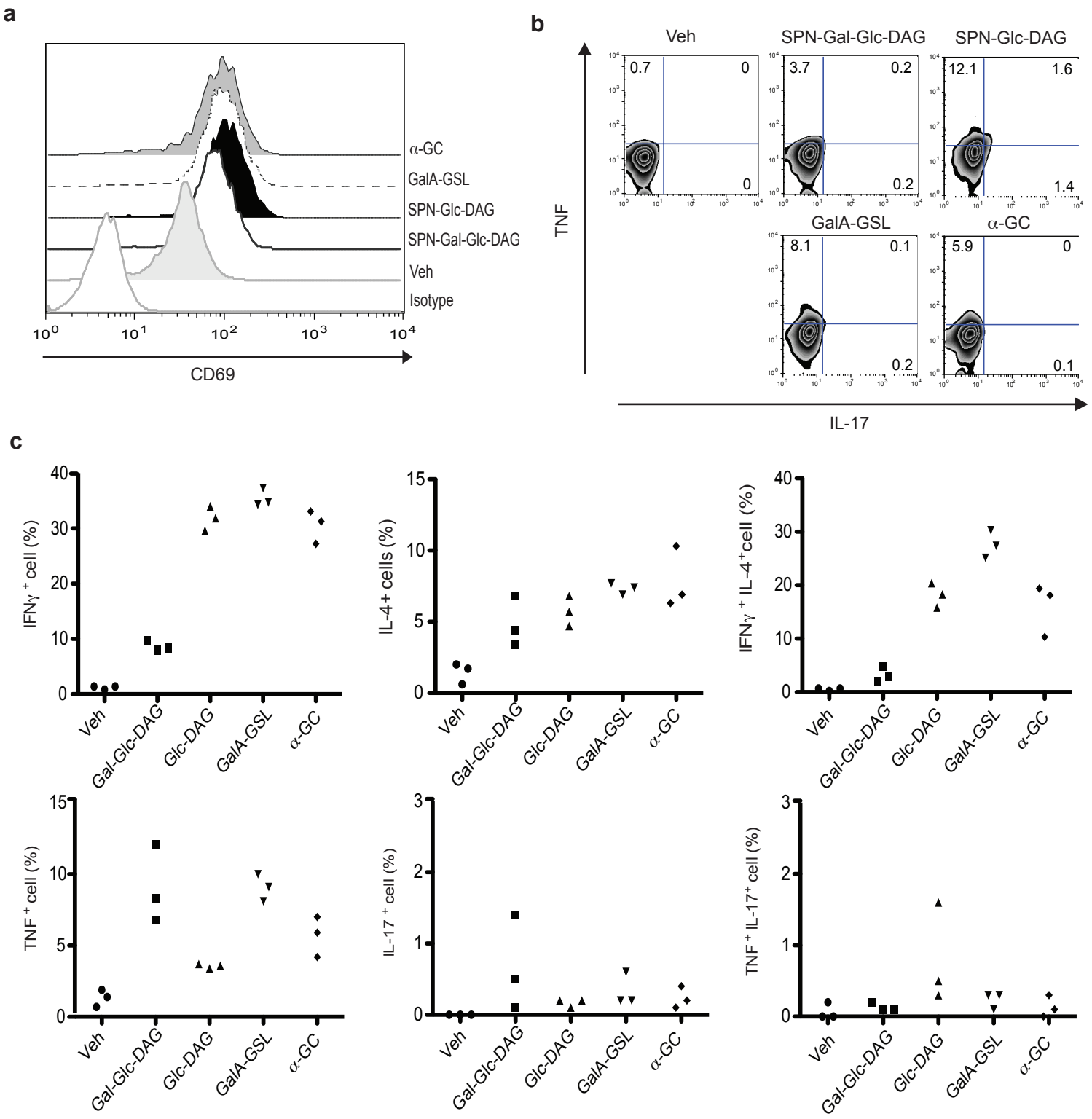

**Supplementary Figure 9. Vα14 *i*NKT cell hybridomas respond to GBS antigens but not to the purified mannosyl DAG antigen from *S. suis* and the IL-2 response of a Vα14 *i*NKT cell hybridoma to purified *Lactobacillus* glycolipids.**

(a) Hybridoma 1.4 was tested by culture with A20 CD1d transfectants pulsed with the indicated concentrations (μg/ml) of the synthetic *Sphingomonas* antigen GalA-GSL or purified DAG glycolipids from *S. pneumoniae* (SPN) or group B *Streptococcus* (GBS).

(b) IL-2 response of Vα14 iNKT cell hybridoma 1.2 to A20 CD1d transfectants that were pulsed with the indicated concentrations in μg/ml of the compounds shown. Each bar shows mean ± SEM from triplicate wells. Experiments were performed two times.

(c) Hybridoma 1.2 was tested by culture with A20 CD1d transfectants pulsed with the indicated concentrations (μg/ml) of the synthetic *B. burgdorferi* antigen or purified monoglucosyl DAG glycolipid from *S. pneumoniae* (SPN) or mannosyl DAG glycolipid from *S. suis* (SSu). NT; not tested.

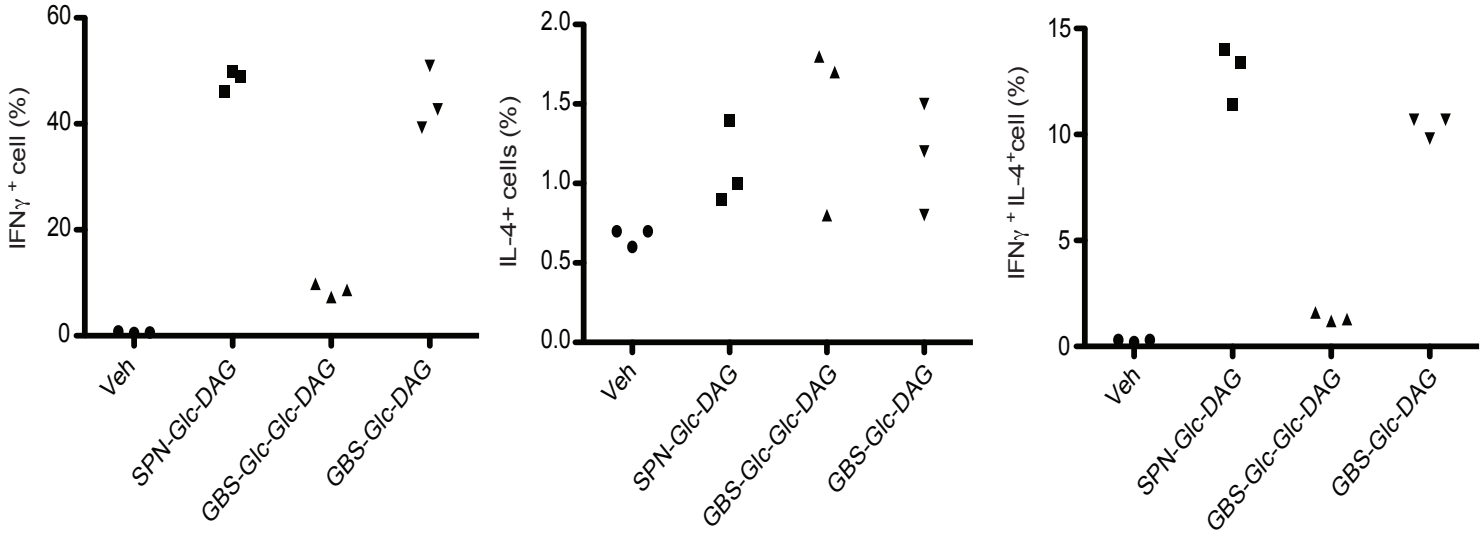
Each bar shows mean ± SEM of IL-2 release from triplicate wells. The experiment was performed twice.




## Supplementary Figure 10. In vivo responses of Vα14 *i*NKT cells to DCs pulsed with purified *S. pneumoniae* glycolipids.

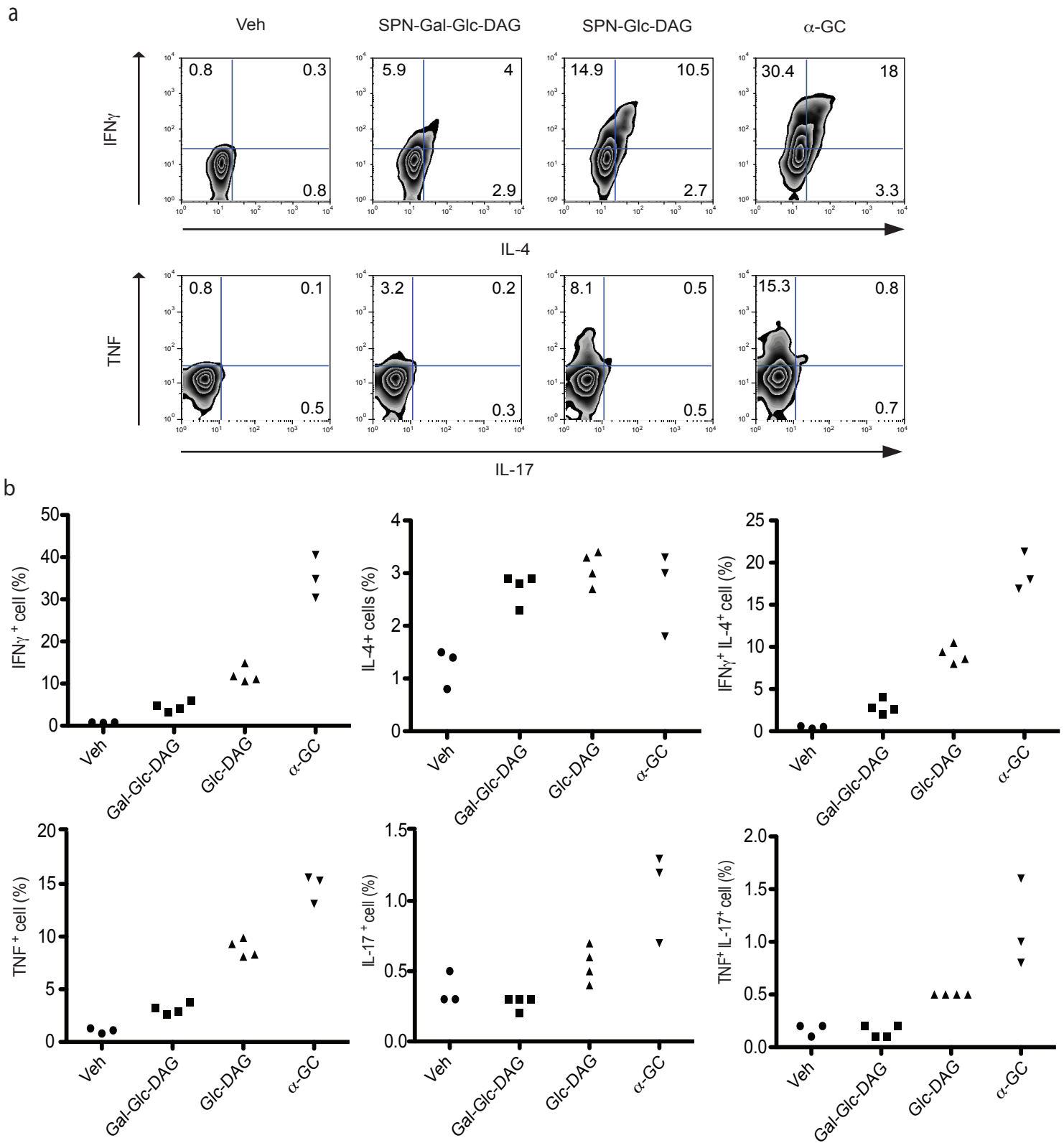
(a) Induction of expression of the activation marker CD69 by gated Vα14*i* NKT cells at 14h after injection of DC pulsed with either α-GalCer (α-GC), synthetic *Sphingomonas* antigen GalA-GSL, purified *S. pneumoniae* mono- (Glc-DAG) or disaccharide (Gal-Glc-DAG) containing antigens (0.1, 10, 20 or 20 μg/ml, respectively), vehicle control. Isotype control staining for CD69 expression is indicated in the bottom histogram.
(b) Intracellular TNF and IL-17 expression by Vα14*i*NKT cells (α-GalCer loaded CD1d tetramer positive liver MNC) at 14h after transfer of DCs as described above. The representative data of 3 mice are shown.
(c) The percentage of intracellular cytokine positive cells among Vα14*i*NKT cells are shown. Each symbol represents individual mouse. Gal-Glc-DAG;SPN-Gal-Glc-DAG, Glc-DAG;SPN-Glc-DAG.
The experiments were performed at least three times.

Nature Immunology doi:10.1038/ni.2096




## Supplementary Figure 11. In vivo responses of Vα14 *i*NKT cells to purified GBS glycolipids.

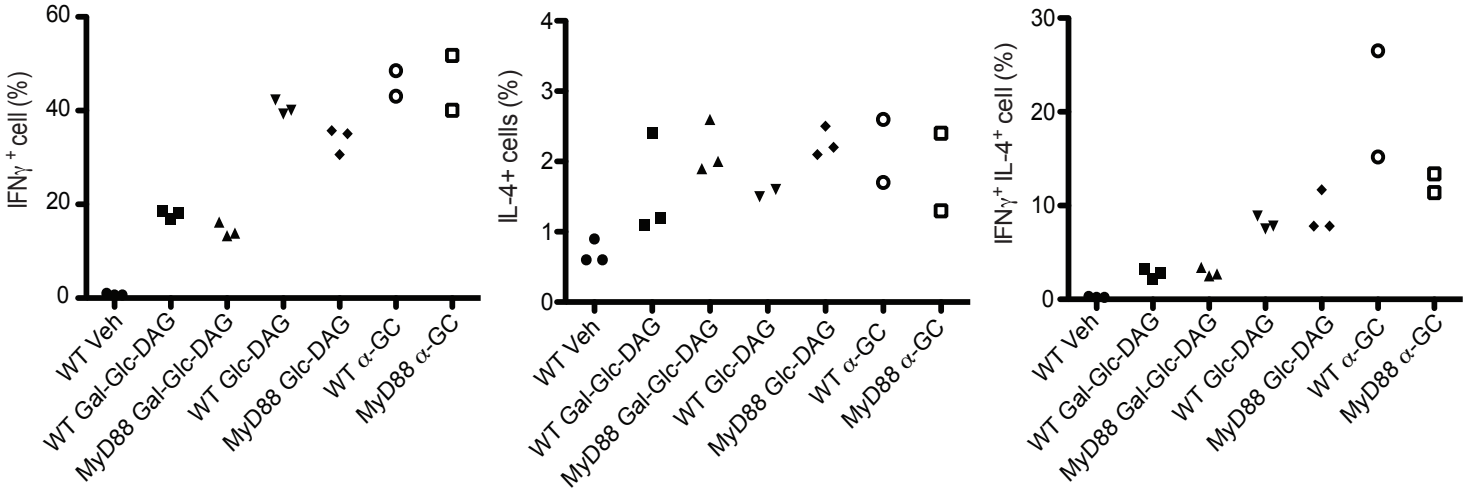
The percentage of intracellular cytokine positive cells among tetramer+ liver MNC at 14h after transfer of DCs that had been pulsed with vehicle (veh), SPN-Glc-DAG, GBS-Glc-DAG or GBS-Glc-Glc-DAG (20 μg/ml). Each symbol represents individual mouse from one experiment with three mice. Experiments were performed at least two times.




**Supplementary Figure 12. Early in vivo responses of Vα14 *i*NKT cells to purified *S. pneumoniae* glycolipids.**

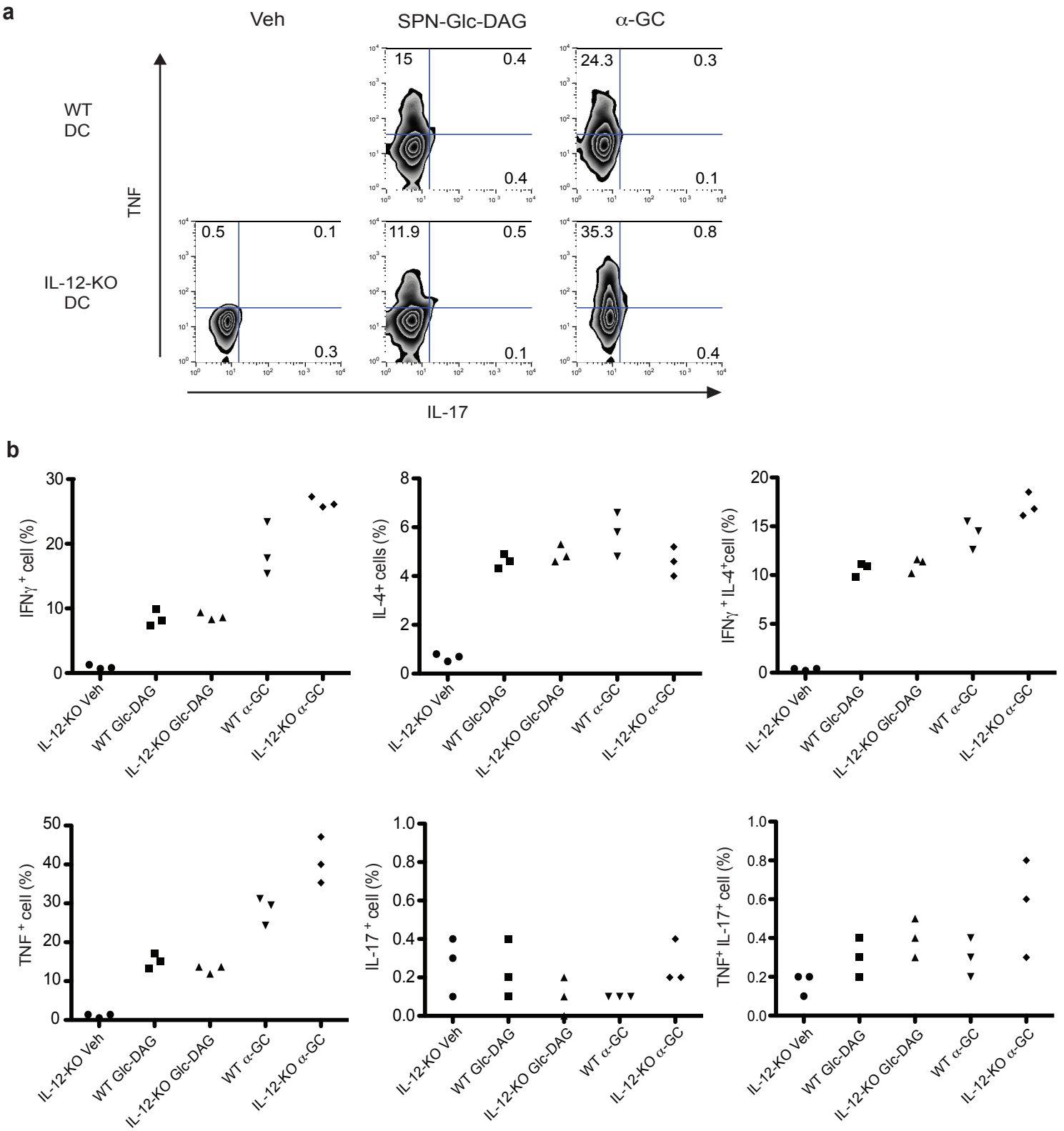
(a) Expression of intracellular cytokines by α-GalCer loaded CD1d tetramer positive cells in liver MNC at 4h after transfer of DCs that had been pulsed with purified SPN-Gal-Glc-DAG, SPN-Glc-DAG, or α-GalCer (20, 20 or 0.1 μg/ml, respectively). The representative data of 3-4 mice are shown. (**b**) The percentage of intracellular cytokine positive cells among α-GalCer loaded CD1d tetramer positive liver MNC is shown. Gal-Glc-DAG; SPN-Gal-Glc-DAG, Glc-DAG; SPN-Glc-DAG. Each symbol represents individual mouse from one representative experiment. Experiments were performed at least two times.

Nature Immunology doi:10.1038/ni.2096




## Supplementary Figure 13. In vivo responses of Vα14 *i*NKT cells to purified *S. pneumoniae* glycolipids when innate immune activation is impaired.

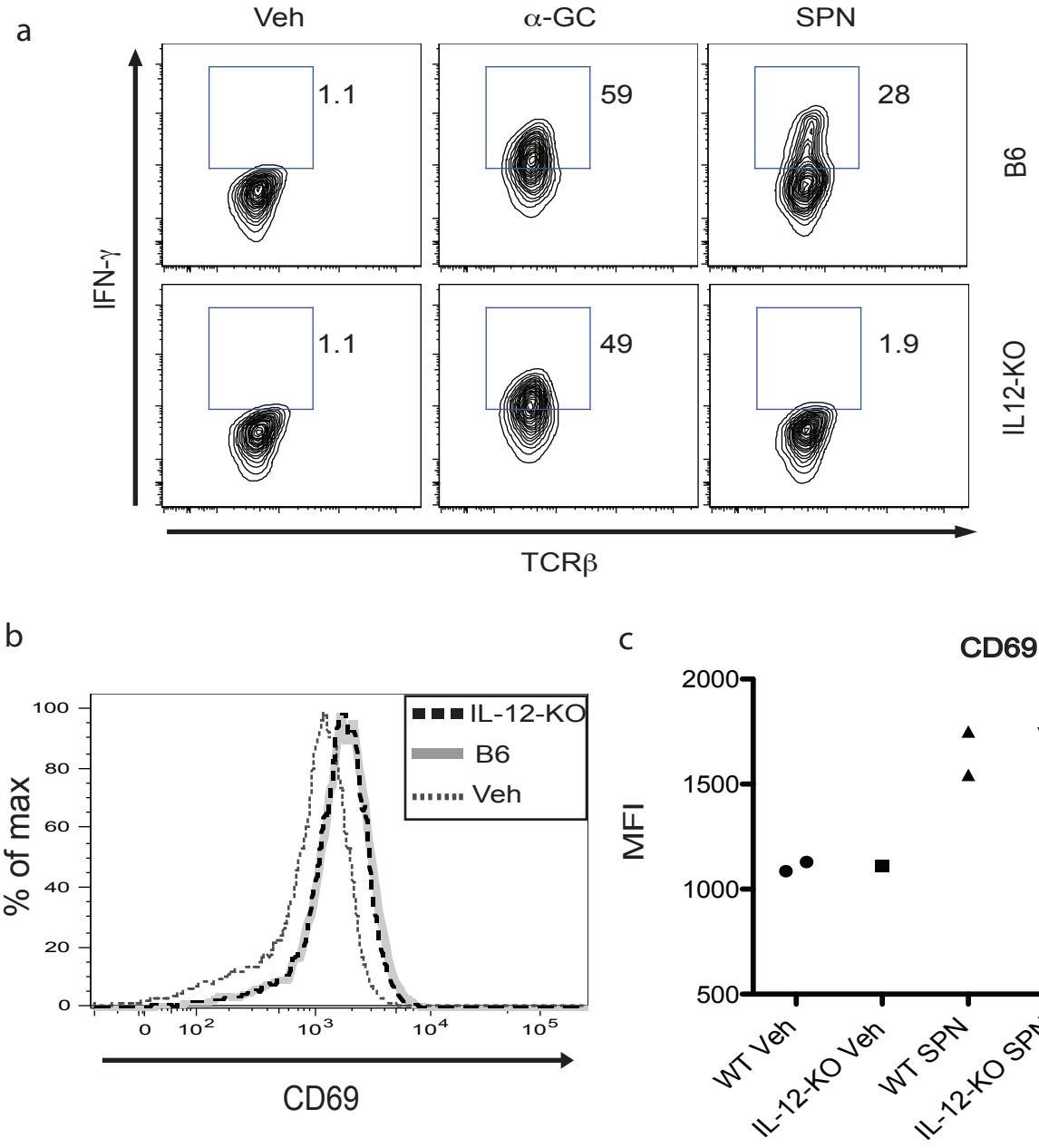
Strong and equivalent responses were induced by transfer of antigen loaded DCs from MyD88-/-/ Triflps2/lps2 mice into wild type or MyD88-/- mice. The percentage of intracellular cytokine positive cells among α-GalCer loaded CD1d tetramer positive liver MNC at 14h after transfer of DCs pulsed with the indicated purified *S. pneumoniae* antigens (20 μg/ml) or positive control antigen α-GalCer (α-GC, 0.1 μg/ml). Each symbol represents individual mouse from one experiment with 2 (α-GC) or 3 (others) mice. Experiments were performed at least two times.




## Supplementary Figure 14. In vivo responses of Vα14 *i*NKT cells to purified *S. pneumoniae* glycolipids when IL-12 is absent.

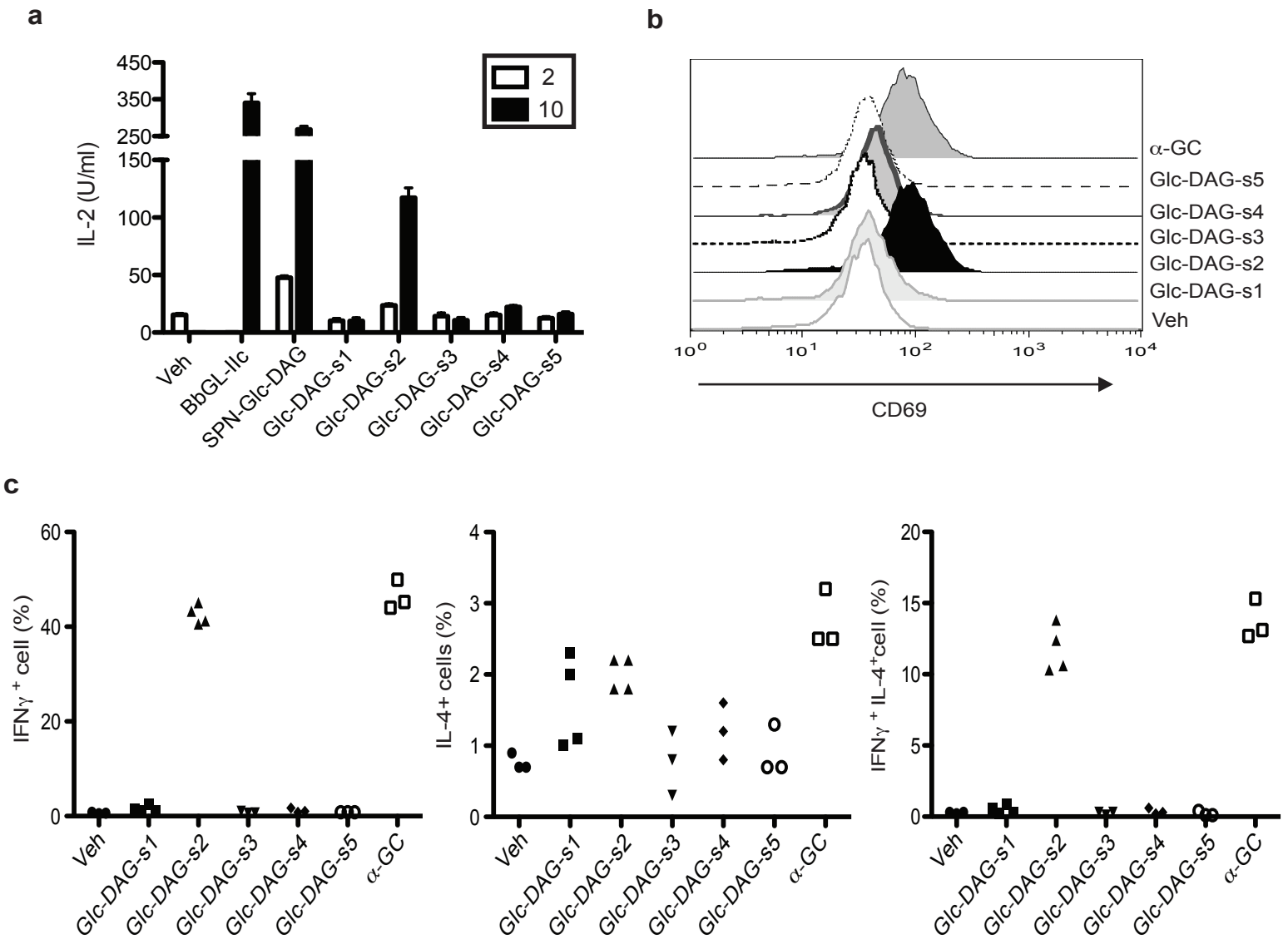
(**a, b**) IL-12 independent activation of *i*NKT cells by SPN-Glc-DAG. (a) Expression of intracellular cytokines by α-GalCer loaded CD1d tetramer positive cells in liver MNC was measured at 4h after transfer of IL-12-KO DCs or WT DCs that had been pulsed with purified SPN-Glc-DAG, or α-GalCer (20 or 0.1 μg/ml, respectively) into IL-12-KO mice. The representative data of 3 mice are shown. (**b**) The percentage of intracellular cytokine positive cells among α-GalCer loaded CD1d tetramer positive liver MNC is shown. Each symbol represents individual mouse from one experiment with three mice are shown. Experiments were performed two times.

Nature Immunology doi:10.1038/ni.2096




**Supplementary Figure 15. In vivo responses of Vα14 *i*NKT cells to *S. pneumoniae* when IL-12 is absent.**

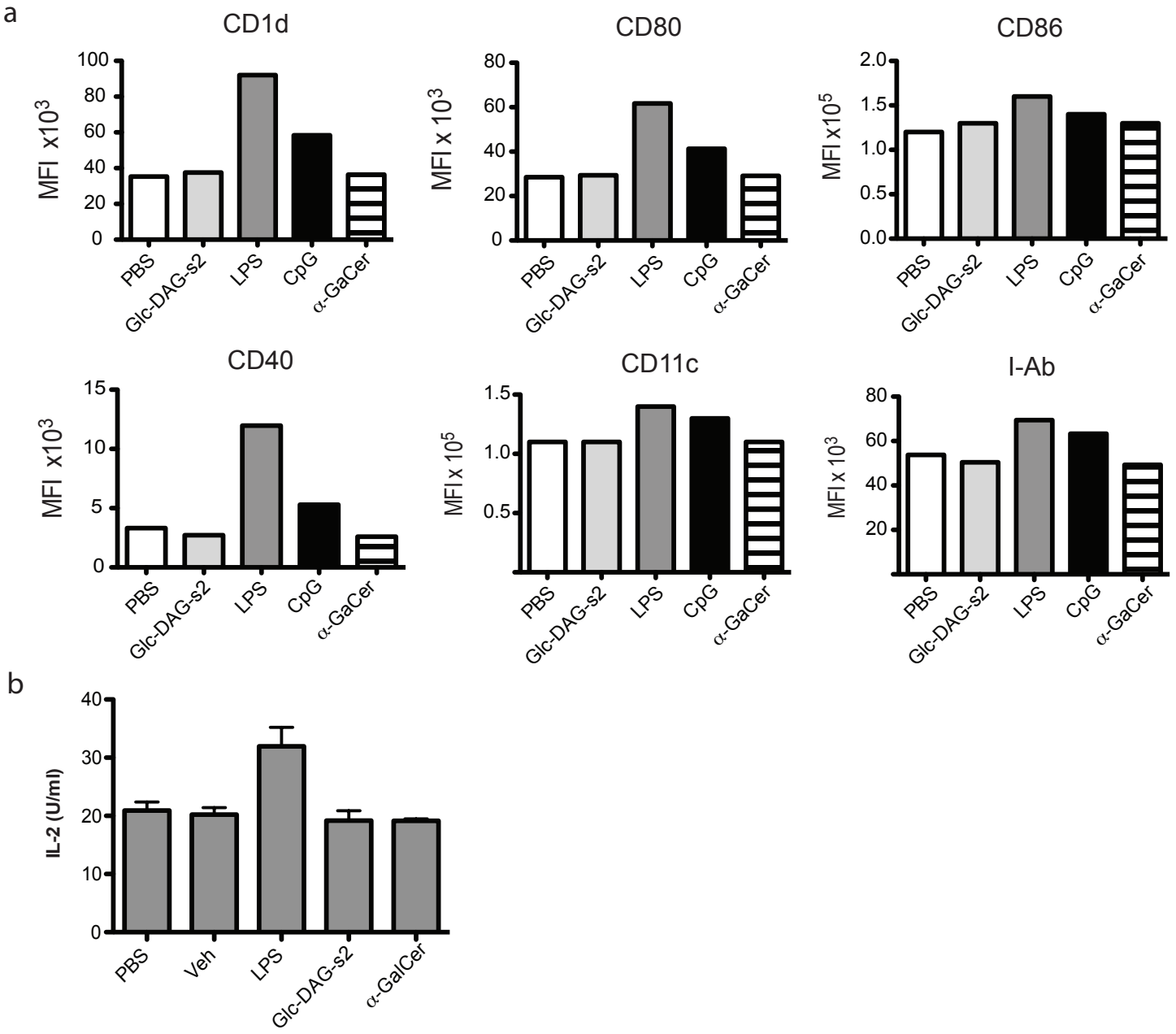
(a,b) Expression of intracellular IFN-γ (a) and CD69 (b) by α-GalCer loaded CD1d tetramer positive spleen cells were measured at 6h after *S. pneumoniae* infection. The representative data of one or two mice are shown. Please note that histograms of B6 and IL-12-KO are overlapped (b).
(c) The mean Fluoresence intensity (MFI) of CD69 expression by α-GalCer loaded CD1d tetramer positive spleen cells are shown. Each symbol represents individual mouse.




## Supplementary Figure 16. *i*NKT cell responses to synthetic *S. pneumoniae* antigens are highly dependent on the fatty acid structure.

(**a**) IL-2 response of the Vα14*i*NKT cell hybridoma 1.2 to CD1d transfected A20 cells incubated with concentrations (μg/ml) of the indicated synthetic glycolipids.

(**b, c**) Expression of CD69 (b) and intracelluar cytokines (c) by α-GalCer loaded CD1d tetramer positive liver MNC from wild type mice at 14h after transfer of DCs that had been pulsed with vehicle (veh), synthetic variants of the *S. pneumoniae* Glc-DAG antigens or α-GalCer (20 or 0.1 μg/ml). (b) Representative data of 3-4 mice are shown. (c) The percentage of intracellular cytokine positive cells among CD1d tetramer positive cells are shown. Each symbol represents individual mice. The Similar results were obtained from 2-3 experiments.

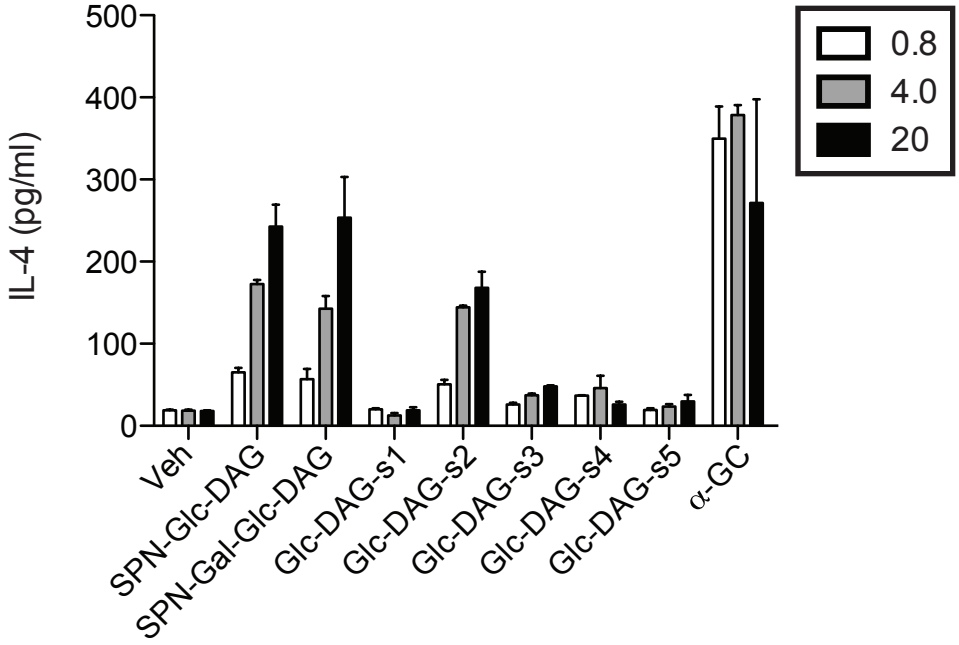



## Supplementary Figure 17. *S. pneumoniae* glycolipid does not stimulate APC.

(**a**) *S. pneumoniae* glycolipid does not simulate BMDCs. BMDCs were cultured with either Glc-DAG-s2, LPS, CpG, or α-GalCer (20, 0.2, 10 or 0.1 μg/ml) for 24h. Induction of indicated cell surface markers were shown as mean fluoresence intensity (MFI).

(**b**) *S. pneumoniae* glycolipid does not alter autoreactive response of *i*NKT cell. IL-2 production by CD1d autoreactive non-Vα14 NKT hybridoma stimulated with CD1d-A20 cells that had been pulsed with either Glc-DAG-s2, LPS, α-GalCer (10, 0.2, 0.1 μg/ml). Each bar shows mean ± SEM from triplicate wells.

Experiments were performed at least two times.




**Supplementary Figure 18. Human Vα24iNKT cells recognize purified and synthetic *S. pneumoniae* glycolipids and produce IL-4.**

Human *i*NKT cell lines (3 x 10^4) were cultured with human CD1d transfected Hela cells (3 x 10^4) for 24h in the presence of the indicated glycolipids (μg/ml). Each bar shows mean ± SD from triplicate wells. The representative data from one of 5 V$\alpha$24*i*NKT cell lines are shown.

**Supplementary Table 1. Data collection and refinement statistics**

| **Data collection statistics** | |
|---|---|
| Resolution range (Å)[a] | 50-1.70 (1.76-1.70) |
| Unique reflections | 56157 |
| $R_{sym}$ (%)[a, b] | 4.6 (53.6) |
| Multiplicity[a] | 3.9 (3.8) |
| I/σI[a] | 25.5 (2.2) |
| Completeness (%)[a] | 98.8 (97.1) |
| | |
| **Refinement statistics** | |
| Protein atoms | 3177 |
| Ligand | 53 |
| Carbohydrates | 56 |
| Waters | 382 |
| Ramachandran favored (%) | 99.1 |
| Ramachandran allowed (%) | 0.9 |
| Rmsd bonds (Å) | 0.013 |
| Rmsd angles (°) | 1.489 |
| $R_{cryst}$[c] | 18.7 |
| $R_{free}$[d] | 21.9 |
| Average B values protein ($Å^2$)[e] | 26.7 |

| | |
|---|---|
| Average B values ligand ($Å^2$) [e] | 64.7 |
| Average B values carbohydrates ($Å^2$) [e] | 44.5 |
| Average B values waters ($Å^2$) [e] | 23.4 |

[a]Numbers in brackets refer to the highest resolution shell.

[b]$R_{sym}=(\Sigma_h\Sigma_i \left| I_i(h)-<I(h)>I \right| (\Sigma_h\Sigma_I I_i(h))$ x 100, where <I(h)> is the average intensity of 'i' symmetry-related observations for reflections with Bragg index h.

[c]$R_{cryst}=(\Sigma_{hkl} \left| F_o-F_c \right| / \Sigma_{hkl} \left| F_o \right|)$ x 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 3% of data excluded before refinement.

[e]B values were calculated with the CCP4 program TLSANL[1].

**Supplementary Results**

**Immunology assays and characterization of bacterial glycolipids**

As described previously [2], the crude lipid extracts of *S. pneumoniae* (SPN) R6 (sample **1)**, two clinical isolates of group B *Streptococci* (GBS) types 12 (sample **2)** and 17(sample **3**), GBS A909 (sample **4**), *S.suis* (SSu) (sample **5**) were fractioned on DEAE-cellolose, acetate form. The final purification of the individual glycolipids was achieved by silica gel (Fluka, 60 mesh) column chromatography using a stepwise gradient of chloroform/methanol (20:1 to 4:1, v/v).

Two glycolipids fractions were isolated:

| Sample | Fraction 1, G-DAG ($R_f$ 0.71, TLC, chloroform-methanol-water 65:25:4, v/v) | Fraction 2, DG-DAG ($R_f$ 0.56, TLC, chloroform-methanol-water 65:25:4, v/v) |
|---|---|---|
| **1** | SPN-Glc-DAG | SPN-Gal-Glc-DAG |
| **2** | GBS12-Glc-DAG | GBS12-Glc-Glc-DAG |
| **3** | GBS17-Glc-DAG | GBS17-Glc-Glc-DAG |
| **4** | GBSA909-Glc-DAG | GBSA909-Glc-Glc-DAG |
| **5** | SSu-Man-DAG | SSu-Glc-Glc-DAG |

Freeze-dried R6 *S. pneumoniae* (2.91 g) gave 98 mg of whole lipid extract. After chromatographic purifications, a pure SPN-Glc-DAG (11.27 mg) and SPN-Gal-Glc-DAG (30.87 mg) has been isolated. SPN-Glc-DAG represents 11.5% of the total lipids. SPN-Gal-Glc-DAG represents 31.5% of the total lipids. Combined together, these two glycolipids represent 43% of

the total isolated lipids. However, some capsulated clinical isolates yield four times less of these lipids (data not shown).

| | Total lipid | SPN-Glc-GAG | SPN-Gal-Glc-DAG |
|---|---|---|---|
| Content (mg) | 98 | 11.27 | 30.87 |
| Percentage (%) | 100 | 11.5 | 31.5 |

The bacterial glycolipids were subsequently analyzed in positive mode ES-MS:

| Glycolipid | SPN-Glc-DAG | SPN- Gal-Glc-DAG | GBS12-Glc-DAG | GBS12-Glc-Glc-DAG | GBS17-Glc-DAG | GBS17-Glc-Glc-DAG |
|---|---|---|---|---|---|---|
| $[M+Na^+]^+$ *m/z* | 751.7; 779.7; 805.7 | 859.5; 885.6 913.5; 941.5 | 697.5; 725.6; 751.6; 779.6 | 859.6; 887.7; 913.7; 941.7 | 697.5; 725.6; 751.6; 779.7 | 859.7; 887.7; 913.7; 941.7 |

| Glycolipid | SSu-Man-DAG | SSu-Glc-Glc-DAG | GBSA909-Glc-DAG | GBSA909-Glc-Glc-DAG |
|---|---|---|---|---|
| $[M+Na^+]^+$ *m/z* | 697.5; 725.6 | 859.5; 887.5; 913.6; 941.6 | 697.7; 725.8; 751.8; 779.8 | 859.5; 887.5; 913.6; 941.6 |

Difference in 26 or 28 Da between signals in the glycolipid MS-spectra corresponded with different (with a double bond or extra $CH_2$-$CH_2$ unit, respectively) fatty acid compositions. On mass spectral analysis the fraction 1 and fraction 2 glycolipids displayed nearly the same pattern of signal distribution with difference in 162 Da, which corresponds to one hexose unit. $^{1}H$ and $^{13}C$ NMR spectra recorded in *$d_6$*-DMSO- $CDCl_3$ (5:1, v/v) or $CDCl_3$-$CD_3OD$ (5:3, v/v) were in agreement with a glycosyl diacylglycerol (G-DAG) (Fraction 1) and diglycosyl diacylglycerol (DG-DAG) (Fraction 2) structures. Besides, most G-DAGs (except SSu-Man-DAG) have nearly identical NMR spectra, as well as do DG-DAGs (except for SPN-Gal-Glc-DAG). For detailed analysis, we used data for SPN-Glc-DAG, SPN-Gal-Glc-DAG, GBS12-Glc-Glc-DAG, SSu-Man-DAG. Indeed, the $^{1}H$ -$^{13}C$ HSQC NMR spectra of SPN-Glc-DAG (Supplementary Fig. 3a) showed one anomeric resonance at $\delta_{H1/C1}$ 4.64/ 108.3 (for SSu-Man-DAG $\delta_{H1/C1}$ is 4. 89/ 109.0). The $^{1}H$ -$^{13}C$ HSQC NMR spectra of GBS12-Glc-Glc-DAG showed two anomeric resonances at $\delta_{H1/C1}$ 4.95/ 104.21 and 4.91/ 104.89 (for SPN-Gal-Glc-DAG these ones are 4.92/ 105.19 and 4.84/ 105.97) (data not shown).

The diacylated glycerol unit was identified by two-dimensional $^{1}H$ -$^{1}H$ correlation spectroscopy (COSY) NMR spectra from its deshielded H-2 proton resonance at 5.16 ppm that correlated with H-1, H-1', H-3 and H-3' resonances at 4.13, 4.32, 3.46 and 3.70 ppm, respectively (Supplementary Fig. 3b). The presence of one unsaturated fatty acid was identified in SPN-Glc-DAG (Supplementary Fig. 4a), SPN-Gal-Glc-DAG (Supplementary Fig. 4b) by $^{1}H$ NMR spectra that showed two protons resonance at 5.31 ppm. In other samples there was variation in the amount of unsaturated fatty acid. The fatty acid methyl ester analysis of SPN-Glc-DAG and SPN-Gal-Glc-DAG by GC/MS showed two main fatty acids: hexadecanoic ($C_{16}$) and octadecenoic acid ($C_{18:1}$) in agreement with the ES-MS spectra ($C_{16}$/ $C_{18:1}$ G-DAG $[M+Na^+]^+$ is

779.7, $C_{16}$/ $C_{18:1}$ DG-DAG $[M+Na^+]^+$ is 941.6). It has been reported [2] that in GBS glycolipids the unsaturated fatty acid esterifies *sn*-2 position of glycerol.

The structure of octadecenoic acid was identified as *cis*-vaccenic (11-octadecenoic acid) by analysis of fragmentation of its bis-disulfide derivative [3]. Indeed, GC-MS (electron impact) (Supplementary Fig. 5a) showed two characteristic fragments with m/z 245 and 145, which is in agreement with fragmentation of a control *cis*-vaccenic acid methyl ester derivative. Based on the assignment of $^1$H-NMR GBS-Glc-DAG (Supplementary Fig. 5b), as well as SPN-Glc-DAG $^1$H -$^1$H COSY NMR (Supplementary Fig. 3b), $^1$H -$^{13}$C HMBC (Supplementary Fig. 6a) and proton coupling constants (Supplementary Fig. 4a) G-DAGs was shown to be α-glucopyranosyl derivative as follows. The small $J_{1,2}$ coupling constant (3.6 Hz) of H-1 at 4.64 ppm indicated an α-anomeric configuration. The large coupling constant of H-2 (*dd*) at 3.18 ppm $J_{2,3}$ (8.9 Hz), H-3 (*t*) at 3.39 ppm $J_{3,4}$ (8.9 Hz) correlates with *gluco*-configuration of sugar. Altogether, these data indicate a 1,2-di-*O*- acyl- (α-glucopyranosyl)- (1→3)-glycerol structure, which is confirmed by $^1$H-NMR data of chemically synthesized analog Glc-DAG-s2 (Supplementary Fig. 6b). This is in agreement with the published data about the presence of α-glucosyl diacylglycerol in different *Streptococci* [4] and its biosynthetic role as anchor for lipoteichoic acid [5].

The $^1$H and $^{13}$C NMR spectra of the lipid SSu-Man-DAG showed the presence of α-mannopyranosyl residue. The very small $J_{1,2}$ coupling constant (0.7 Hz) of H-1 at 4.89 ppm (Supplementary Fig. 7a) as well as the large H1-C1 coupling constant $J_{C1,H1}$ (172.0 Hz) of C-1 at 109.0 ppm (data not shown) indicated an α-anomeric configuration of mannose. All other sugar proton signals are in agreement with the reported ones for α-mannopyranoside [6]. These data, as well as ES-MS, indicate a 1,2-di-*O*- acyl-(α-mannopyranosyl)- (1→3)-glycerol structure of SSu-

Man-DAG, in which $C_{14}$/ $C_{16}$ are the major fatty acids. This is a first example of mannosylated diacylglycerol from *Streptococci.*

Analysis of DG-DAGs from fraction 2 by $^{1}H$ and $^{13}C$ NMR experiments revealed α-glucosyl diacylglycerol with attached second hexose as a common structural motif for all DG-DAGs. The small $J_{1,2}$ coupling constant (3.7 Hz) of H-1 at 4.91 ppm of the DG-DAGs second sugar as well as $J_{1,2}$ coupling constant (3.6 Hz) of H-1 at 4.84 ppm of the SPN-Gal-Glc-DAG indicated it is in an α-anomeric configuration (Supplementary Fig. 4b). The linkage between two different glycosyl units was established from $^{1}H$ -$^{1}H$ COSY NMR, $^{1}H$ -$^{13}C$ HMBC, $^{1}H$ -$^{13}C$ HMQC NMR that revealed correlation between C-1 of hexose at 104.89 ppm and glucosyl H-2 at 3.52 ppm of glucose (for GBS12-Glc-Glc-DAG), as well as C-1 of hexose at 105.97 ppm and at glucosyl H-2 at 3.35 ppm for SPN-Gal-Glc-DAG. Taken together, these data indicate that the hexose is alpha linked at O-2 of α-glucosyl diacylglycerol. $^{1}H$ and $^{13}C$ NMR spin systems of GBS12-Glc-Glc-DAG hexose (Supplementary Fig. 7b and data not shown) corresponded to α-glucopyranose, while these ones for SPN-Gal-Glc-DAG (Supplementary Fig. 4b and data not shown) corresponded to α-galactopyranose.

Altogether, these data indicate a 1,2-di-*O*- acyl- (α-galactopyranosyl)- (1→2)- (α-glucopyranosyl)- (1→3)-glycerol structure of SPN-Gal-Glc-DAG and 1,2-di-*O*- acyl- (α-glucopyranosyl)- (1→2)- (α-glucopyranosyl)- (1→3)-glycerol structure of GBS12-Glc-Glc-DAG and the rest of DG-DAGs.

Gal-Glc-DAG has been isolated [4] from *S. pneumoniae* before, while Glc-Glc-DAG is a characteristic glycolipid for group B as well as other *Streptococci* [2].

**Supplementary Methods**

**NMR analysis**. Deuterated solvents were from Aldrich. Glycolipid NMR spectra were recorded in $d_6$-$Me_2SO$ - $CDCl_3$ (5:1, v/v) or $CDCl_3$ - $CD_3OD$ (5:3, v/v). NMR spectra were recorded on a Bruker DRX500 operating at 500.13 MHz for $^1H$ and 125.77 MHz for $^{13}C$. The instrument was equipped with a 5-mm $^1H$-{$^{13}C$}-{X}-triple broadband inverse triple resonance *z*-gradient probe head, and all spectra were run at 300 K. $d_6$-$Me_2SO$ was used as lock, and the residual solvent was used as internal reference (2.49 ppm for $^1H$ and 39.5 ppm for $^{13}C$). Data were acquired and processed using XWINNMR version 2.6 software on a Silicon Graphics work station. The $^{13}C$ NMR data were acquired with $^1H$ decoupling and NOE. All two-dimensional NMR data were acquired nonspinning. Data points (2048) were used in acquisition for the fast domain (F2), and 512 points were used in the incremented domain (F1). The F1 dimension was zero-filled to 1024 data points in processing. The HSQC, HMBC NMR experiments (all performed using pulsed field gradients) were acquired in phase-sensitive mode using the time-proportional phase increment. In all three experiments, a Qsine window function shifted by $\pi/2$ was used for both frequency domains in processing. GARP decoupling was employed in F1 for the HSQC experiment.

**Fatty Acid Analysis.** Glycolipids (1 mg) were *O*-deacylated with 0.23 ml of 0.25 M $NaOCH_3$ in dichloromethane –methanol (2:1, v/v) at 24°C. On completion (monitored by TLC) the reaction mixture was neutralized with HCl, diluted with diethyl ether (3 ml) and extracted with water. The organic fraction was concentrated and fatty acid methyl esters were analyzed by Hewlett-

Packard gas–liquid chromatography (GLC-MS), with HP-5 glass capillary column, programmed at 8°C per min, from 125–250°C.

**Double Bond Localization.** The double bond in unsaturated fatty acids methyl esters was localized by GC-MS (electron impact) after dimethyl disulfide derivatization [3], and its configuration was determined by using methyl ester of *cis*-vaccenic acid as reference.

**Mass spectrometic analysis.** Glycolipids were dissolved in dichloromethane-methanol (2:1, v/v) and analyzed by Electrospray ionization mass spectrometry (Micronos LCT) in positive mode.

**Mouse CD1d-Glc-DAG-s2 crystallization and structure determination**

**Protein expression and purification.** The recombinant soluble mouse CD1d-β2m (CD1d-β2m) heterodimeric molecule was produced using the baculovirus expression system as previously described [7]. The protein was purified by anion-exchange chromatography on a MonoQ 5/50 column (GE Healthcare) equilibrated with 10 mM Tris-HCl pH 8.0 and eluted with a linear gradient of 0-350 mM NaCl. The fractions containing the CD1d-β2m heterodimer were pooled together and incubated with a six-fold molar excess of Glc-DAG-s2 dissolved in DMSO (final concentration of DMSO 1.1%) at 30°C for 4 hours in the presence of 100 mM Tris-HCl pH 7. The CD1d-glycolipid complex was then purified by gel filtration on a Superdex200 10/300 column (GE Healthcare) equilibrated in 50 mM HEPES pH 7.5, 150 mM NaCl. The fractions containing the complex were then concentrated to approximately 9 mg/ml in 10 mM HEPES pH 7.5, 30 mM NaCl for subsequent crystallization.

**Protein crystallization.** Initial screens (PEG/Ion I & II, Wizard I & III, Hampton Research) were set up at 4°C and 23°C with the help of a nanoliter dispensing crystallization robot (Phoenix, Art Robbins Instruments). Several conditions resulted in the formation of needle-like crystals and the most promising ones among those were optimized manually. The best crystals were obtained in 8% tacsimate pH 7.5, 17% PEG3350 by sitting drop vapor diffusion at 4°C, mixing 0.4 μl of complex at 9 mg/ml with 0.4 μl of precipitant.

**Structure determination and refinement.** Crystals were flash-cooled at 100K in crystallization solution containing 20% glycerol. Data collection was performed at beamline 7.1 of the Stanford Synchrotron Radiation Laboratory and processed to a resolution of 1.7 Å with the HKL2000 software [8]. The protein crystallized in space group $P2_12_12_1$ with cell parameters a = 42.191 Å b = 107.381 Å c = 110.305 Å $\alpha = 90°$ $\beta = 90°$ $\gamma = 90°$ and with one CD1d-lipid complex in the asymmetric unit (estimated solvent content of 56.5%).

Molecular replacement was performed using Molrep [9, 10], and the crystal structure of the CD1d-iGb3 complex (PDB ID 2Q7Y, [11]) with the ligand removed as a search model. The initial phases derived from this model were refined using maximum-likelihood restrained refinement coupled with TLS refinement in Refmac [1, 9]. Three TLS groups were defined containing the α1, α2 domain and the glycolipid ligand (group 1), the α3 domain (group 2) and the β2m chain (group 3), respectively. A set of 3% of the reflections were set aside for the calculation of $R_{free}$ to monitor refinement progress. 2Fo-Fc and Fo-Fc maps were used in the program Coot [12] for modeling and manual building. Figure 6 was generated with the programs Pymol (http://pymol.sourceforge.net), the PDB2QR server [13] and APBS [14]. Electrostatic surfaces are shown as electronegative in red and electropositive in blue from -30 *k*T/e to 30 *k*T/e. Potential hydrogen bonds were considered of below a distance cutoff of 3.5 Å. The quality of the final

model was assessed with Molprobity [15] and the validation tools available in Coot. Structure factors and coordinates were deposited in the PDB database with code 3T1F.

**Surface Plasmon Resonance studies**

Surface Plasmon Resonance studies using a refolded and biotinylated Vα14Vβ8.2 TCR were carried out as previously reported for the borrelial diacylglycerolipid BbGL-IIc [16]. Seven dilutions of mCD1d-Glc-DAG-s2 complexes were injected with increasing concentration (0.3125-20μM) over a streptavidin sensorchip that has been immobilized with 300RU of biotinylated TCR. The experiment was performed twice with two different TCR preparations.

**Supplementary References**


1. Winn, M.D., Isupov, M.N. & Murshudov, G.N. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr.* **D57**, 122-133. (2001).
2. Fischer, W. The polar lipids of group B Streptococci. I. Glucosylated diphosphatidylglycerol, a novel glycopholipid. *Biochim Biophys Acta* **487**, 74-88 (1977).
3. Scribe, P., Guezennec, J., Dagaut, J., Pepe, C. & Saliot, A. Identification of the position and the stereochemistry of the double bond in monounsaturated fatty acid methyl esters by gas chromatography/mass spectrometry of dimethyl disulfide derivatives. *Anal. Chem.* **60**, 928–931 (1988).
4. Brundish, D.E., Shaw, N. & Baddiley, J. The glycolipids from the non-capsulated strain of Pneumococcus I-192R, A.T.C.C. 12213. *Biochem J* **97**, 158-165 (1965).
5. Seo, H.S., Cartee, R.T., Pritchard, D.G. & Nahm, M.H. A new model of pneumococcal lipoteichoic acid structure resolves biochemical, biosynthetic, and serologic inconsistencies of the current model. *J Bacteriol* **190**, 2379-2387 (2008).
6. Meldal, M., Christensen, M.K. & Bock, K. Large-scale synthesis of d-mannose 6-phosphate and other hexose 6-phosphates. *Carbohydrat. Res.* **235**, 115-127 (1992).
7. Zajonc, D.M. *et al.* Structural basis for CD1d presentation of a sulfatide derived from myelin and its implications for autoimmunity. *J Exp Med* **202**, 1517-1526 (2005).
8. Otwinowski, Z. & Minor, W. HKL: Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326 (1997).
9. CCP4 Collaborative Computational Project, Number 4. The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr.* **D50**, 760-763 (1994).
10. Vagin, A.A. & Teplyakov, A. MOLREP:an automated programm for molecular replacement. *J. Appl. Cryst.* **30**, 1022-1025 (1997).
11. Zajonc, D.M., Savage, P.B., Bendelac, A., Wilson, I.A. & Teyton, L. Crystal structures of mouse CD1d-iGb3 complex and its cognate Valpha14 T cell receptor suggest a model for dual recognition of foreign and self glycolipids. *J Mol Biol* **377**, 1104-1116 (2008).
12. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).
13. Dolinsky, T.J., Nielsen, J.E., McCammon, J.A. & Baker, N.A. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res* **32**, W665-667 (2004).
14. Baker, N.A., Sept, D., Joseph, S., Holst, M.J. & McCammon, J.A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* **98**, 10037-10041 (2001).
15. Lovell, S.C. *et al.* Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins* **50**, 437-450 (2003).
16. Wang, J. *et al.* Lipid binding orientation within CD1d affects recognition of Borrelia burgorferi antigens by NKT cells. *Proc Natl Acad Sci U S A* **107**, 1535-1540 (2010).