Structural determinants in a glucose-containing lipopolysaccharide from *Mycobacterium tuberculosis* critical for inducing a subset of protective T cells

Prithwiraj De\(^1\), Michael McNeil\(^1\), Mei Xia\(^3\), Claudia M. Boot\(^2\), Danny C. Hesser\(^1\), Karolien Deneef\(^2\), Christopher Rithner\(^2\), Tyler Sours\(^2\), Karen M. Dobos\(^1\), Daniel Hoft\(^3\) and Delphi Chatterjee\(^1\)*

From the \(^1\)Mycobacteria Research Laboratories, Department of Microbiology, Immunology and Pathology, Colorado State University, Fort Collins, CO 80523, USA.

\(^2\)Central Instrument Facility, Department of Chemistry, Fort Collins, CO 80523, USA.

\(^3\)Departments of Internal Medicine, Saint Louis University, St. Louis, Missouri 63104

Running Title: *Structural and functional diversity in mGLP*

*To whom correspondence should be addressed: Delphi Chatterjee* \(^1\)Mycobacteria Research Laboratories, Department of Microbiology, Immunology and Pathology, Colorado State University, Fort Collins, CO 80523, USA.

Delphi.chatterjee@colostate.edu; Tel (970) 491-7495; FAX (970)491-1815

Keywords: lipoglycan, gamma delta T cells, tuberculosis, structure analysis, vaccine
Abstract

Mycobacteria synthesize intracellular, methylated glucose–containing lipopolysaccharides (mGLP) proposed to modulate bacterial fatty acid metabolism. Recently, it has been shown that Mycobacterium tuberculosis mGLP specifically induces a specific subset of protective γδ T cells. Mild base treatment, which removes all the base-labile groups, reduces the specific activity of mGLP required for induction of these T cells, suggesting that acylation of the saccharide moieties is required for γδ T-cell activation. On the basis of this premise, we used analytical LC/MS and NMR methods to identify and locate the acyl functions on the mGLP saccharides. We found that mGLP is heterogeneous with respect to acyl functions and contains acetyl, isobutyroyl, succinyl, and octanoyl groups and that all acylations in mGLP, except for succinyl and octanoyl residues, reside on the glucosyl residues immediately following the terminal 3-O-methyl glucose. Our analyses also indicated that the octanoyl residue resides at position 2 of an internal glucose toward the reducing end. LC/MS analysis of the residual product obtained by digesting the mGLP with pancreatic alpha amylase revealed that the product is an oligosaccharide terminated by α-(1→4)-linked 6-O-methyl-D-glucosyl residues. This oligosaccharide retained none of the acyl groups, except for the octanoyl group, and was unable to induce protective γδ T cells. This observation confirmed that mGLP induces γδ T cells and indicated that the acylated glucosyl residues at the nonreducing terminus of mGLP are required for this activity.

Mycobacterium spp. produces many exotic lipids and glycolipids that have demanded exploration into their biological functions. Many of the glycans among these glycolipids are naturally methylated (1). Among these, two classes of methylated polysaccharides: 3-O-methylmannoplyasaccharide (MMP) and 6-O-methylglucose-containing lipopolysaccharides (mGLP) have been implicated in regulation of fatty acid synthesis (2-4). However, this has been contradicted by the observation that Rv3032 and MSMEG_5084 knock out mutants of M. tuberculosis H37Rv and M. smegmatis, known to be impaired in mGLP synthesis, displayed wild-type fatty acid contents (5). mGLPs have been found in several Nocardia species and in M. phlei, M. smegmatis, M. bovis BCG, M. tuberculosis, M. leprae and M. xenopi (6-10). The characterization and biosynthesis of mycobacterial mGLP have been reported in details over the last two decades (11-14). Current knowledge on biosynthesis of these methylated glucans suggested that acylation with acetyl, propionyl, isobutyryl, octanoyl and succinyl groups from their respective acetyl CoA and methylation occurs simultaneously after the saccharide moiety has been assembled. Kamisango et al (15) have proposed that mGLP is synthesized from the reducing end towards the nonreducing end through sequential glucosylation and methylation reactions. The major function associated with mGLP in many previous reports has been fatty acid metabolism. However, an account of γδ T cell activating biological activity has been recently recognized (16). Several forms of mGLP were identified in the total mixture using ESI LC/MS: mGLP with 1 to 3 acetyl and an octanoyl; mGLP with isobutyryl, acetyl, and octanoyl; and mGLP with octanoyl, isobutyryl and 3 acetyl + 1 succinyl residue. The location of the acyl functions was determined by 2D NMR and QToF LC/MS/MS experiments. Previous work has shown that in M. phlei the non-succinyl acyl substituents except octanoyl are on the terminal 3-O-Me Glc and with some evidence that it may also have two acyl groups at 4 and 6 positions (6). In this work, we have shown that a C18 reverse phase (RP)-based nanoLC-nanoESI-MS/MS analysis of methylated acylated glycans in negative ion mode is possible and that diagnostic fragment ions can allow determination of the location of the acyl functions with some confidence. Although these molecules have complex structures and molecular weights that range from 3500 to 4000, the spectra are completely interpretable and are consistent with the previous structural assignments. In addition,
they reveal subtle features which were not apparent from earlier studies (6,17) and contribute to novel biological activity (16).

Results

1H PRESAT NMR of deacylated mGLP (mGP) and its effect on γδ2 T cells

Purified mGLP was treated with mild base and the resulting product was desalted using a Bio-Gel (P-2) column and examined by NMR (Fig.1A) (4). The resulting mGP revealed well resolved anomic protons in 1H NMR. A clear doublet at δ 4.89 ppm (J = 3.76 Hz) was attributed to the anomic proton of α-Glcp-(1→2)-glyceric acid. The two overlapping doublets at δ 4.77 ppm (J = 7.9 Hz) and δ 4.78 ppm (J = 8.2 Hz) confirmed the presence of two β-Glcp-(1→3)-Glcp linked residues. The presence of one α-Glcp-(1→6)-α-Glcp linkage, was evident at δ 5.05 ppm (J = 3.84 Hz). The overlapping cluster of peaks between δ 5.25–δ 5.38 integrated to 16 protons, and were assigned to α-Glcp-(1→4)-α-glycosyl backbone (4,8). A distinct AB2 pattern of the peaks at δ 4.12 (1H, t, J = 9.1 Hz, >CH-), and a set of two overlapping doublets, at δ 4.07 (2H, dd, J = 9.4 Hz, J = -3.2 Hz, >CH2-) were identified (18). Notably, the negative coupling constant was attributed to the two-bond (2J) coupling between the geminal diastereotopic methylene protons of the glyceric acid residue (19,20). The NMR analysis suggested that mGP from M. tuberculosis shared a carbohydrate back bone very similar if not identical to that of other species (5,7,10,21).

The effect of mGP on γδ2 T cells

This mGP was not able to activate γδ2 T cells efficiently when compared with the parent mGLP indicating that the acyl modifications of the saccharide backbone were necessary for biological activities (Fig. 7B in the Discussion illustrates T-cell activity of mGP along with other derivatives). Thus, more detailed biochemical analyses of mGLP were pursued to further identify the nature and modification driving the biological activity.

Monosaccharide Composition of mGLP

The monosaccharide analysis by GC/MS of the G-50 purified mGLP showed three distinct hexoses, 6-O-Me-Glc, Glc and a 3-O-Me Glc in the ratio of 11:8:1 (Supporting Information Fig S-1). Surprisingly, our mGLP preparation did not contain the 2-N-acetyl-2,6-dideoxy-β-glucopyranose as in M. bovis, BCG (10).

Identification of Acyl Functions in G50 purified native mGLP by 1H NMR

Native mGLP was analyzed first by 1D-1H NMR spectroscopy (Fig. 1B). The resonances for the acyl substitutions were all evident between δ 0.5 to 2.6 ppm except in this experiment succinyl resonances could not be unambiguously identified due to overlap issues. The chemical shifts at δ 2.37 ppm (2H, m, H6), δ 1.52 ppm (2H, m, H5), δ 1.26 ppm (8H, m, H4), and δ 0.83 ppm (3H, t, J = 6.64 Hz, -Me terminal) provided the evidence for the presence of the octanoyl residue. The chemical shift at δ 2.55 ppm (>CH-) and a set of two overlapping doublets at δ 1.05 ppm (3H, d, J = 7.0 Hz) and δ 1.06 ppm (3H, d, J = 7.0 Hz) were attributed to an isobutyryl residue. In addition, three differentially located acetyl groups were at δ 1.99, 2.00 and 2.03 ppm (3xs, 3xCH3). When the spectrum was integrated with reference to eight protons at δ 1.26 ppm (8H, m, Hγ), the anomic region accounted for ~twenty protons, suggesting that all the mGLP-isomers contained the octanoyl residue. However, the integral value (2.7) exceeds well over two protons at δ 2.37 ppm (2H, m, Hα) indicating possible overlap from other acyl residues. The relative integral values, (the non overlapping β-protons (2H) of octanoyl group at δ 1.52 ppm was the reference integral) revealed an approximate acyl variation ratio (with reference to the octanoyl residue) as octanoyl: isobutyryl: acetyl: succinyl = 1: 0.5: 0.67: 0.35 in the mixture of differentially acylated mGLPs.

Heteronuclear Single Quantum Coherence (HSQC; 1H-13C correlation NMR spectroscopy) and Total Correlation Spectroscopy (TOCSY; through bond 1H-1H correlation NMR spectroscopy) for confirmation of acyl groups.

The HSQC experiment of the G-50 purified mGLP revealed (Fig. 2) the 13C resonances at δ
35.0 ppm (Cₐ), δ 25.8 ppm (Cβ), δ 29.5 ppm (Cγ), δ 32.2 ppm (Cₐ₋₆), δ 22.5 ppm (C₇) and δ 14.0 ppm (Me) for the octanoyl chain. The methylene carbons identified in the HSQC experiment (Fig. 2) at δ 32.0 ppm (Cβ) and at δ 30.0 ppm (Cₐ) were correlated with H² spin system in the TOCSY experiment (Supporting Information; Fig. S-2). As the respective methylene proton’s chemical shifts were at δ 2.55 ppm (Hₐ) and δ 2.37 ppm (Hβ), we assigned this methylene system to a possible succinyl residue which was also confirmed by mass spectrometry. The HSQC experiment also revealed the methine (>CH-) proton of the isobutyryl residue at δ 2.55 ppm (m, Hₐ) is overlapping with the methylene proton peak of a succinyl residue. A clear spin system correlation was observed in the TOCSY experiment (Supporting Information; Fig. S-2) between the peak (>CH- proton at δ 2.55 ppm) and a set of two overlapping doublets at δ 1.05 ppm (3H, d, ²J = 7.0 Hz) and δ 1.06 ppm (3H, d, ²J = 7.0 Hz). In the HSQC experiment the ¹³C-for the acetyl isobutyl carbons was confirmed at δ 20.0 ppm. As for the glyceric acid residue, the HSQC experiment showed distinct methylene carbon at δ 62.2 ppm (Cβ) with the corresponding protons at δ 4.15 and δ 4.00 ppm respectively. The carbon centered at δ 79.8 ppm (δ 4.17 ppm (Hₐ) was attributed to Cₐ of the glyceric acid. The different chemical shifts for the diastereotopic protons (Hₐ) (giving rise to an ABC pattern; Hₐ and 2 X Hβ) for the glyceric acid has been reported earlier (18).

The HSQC spectra of the mGLP revealed five distinct sets of methylene protons (in red; Fig. 2). The protons at δ 3.92-3.60 ppm with corresponding ¹³C chemical shift at δ 60.0-61.0 ppm was attributed to the H6/C6 of α-D-Glcp unit/s linked through (1→4) glycosidic bonds (22). The H6/C6 of the α-D-6-OMe Glcp units, linked through (1→4) glycosidic bonds, were assigned to δ 3.71-3.38 ppm with the corresponding ¹³C chemical shift at δ 70.0-71.0 ppm (23). The protons between δ 3.82-3.67 with the corresponding ¹³C chemical shift at δ 62.5 ppm was attributed to the H6/C6 β Glcp- units (24). The ¹³C at δ 66.0 ppm showing two different proton chemical shifts at δ 3.84 ppm and δ 3.63 ppm is possibly due to the diastereotopic relationship, and was attributed to the H6/C6 of the α-(1→6)-Glcp-(1→2) glyceric acid-unit (25,26). We could assign the ¹³C of the methylene peaks at δ 63.8 ppm with the proton chemical shifts between 64.42-4.22 ppm to the H6/C6 of the Glcp units that are acyl substituted. The HSQC experiment revealed a proton at δ 4.8 ppm (>CH-in blue) with a ¹³C chemical shift at 69.5 ppm. We assigned this to the H-2 of a Glcp that is likely to be acylated. The only other possibility, i.e. H-3 (since H-4 is perhaps glycosyl linked)- would have resulted in a more downfield shift (23,27) if acylated.

The Nuclear Overhauser Effect Spectroscopy (NOESY; ¹H-¹H correlation NMR spectroscopy) allowing for information on acyl substitution on the carbohydrate backbone.

In the NOESY experiment however, we did not see any through space coupling of the acetyl or isobutyryl residues with either ring protons (δ 3.0-4.5 ppm) or the anomeric protons (δ 4.6-5.7 ppm) (Fig. 3). This suggested that the acetyl residues and the isobutyryl residue were perhaps attached to the primary carbon (C-6) position of Glcp units. However, the octanoyl residue showed through space correlation not only with Glcp ring protons at δ 3.53-3.72 ppm, but also with the anomeric protons at δ 5.02-5.12 ppm (α-(1→4)-Glcp) and δ 4.90 ppm α-(1→6)-Glcp indicating that the octanoyl chain is located on the C-2 position of α-Glcp at the reducing end. Further, the α-protons of the octanoyl residue showed specific NOE correlations with protons at δ 2.55 ppm (possibly Hₐ of a succinyl residue (correlation with isobutyryl was ruled out because the pattern would have been different) and acetylated methylene (CH₂-) protons at δ 4.20 ppm in addition to anomeric protons at δ 5.02-5.12 ppm (α-(1→4)-Glcp). This suggested that the octanoyl residue is possibly attached to the C-2 of a glucosyl residue with a succinyl substitution closely as in 6-substituted β (1→3) Glcp (5,28,29) and not on the glyceric acid as has been reported in the past (4). Characterization of all the acyl groups, present in mGLP, are listed in Table 1.

**LC/MS Analysis of G50 purified Native mGLP**

At first, a direct infusion of G-50 purified native mGLP in the ESI/MS negative mode showed
singly charged cluster of ions at m/z 1033, identified as a contaminant and [M-2H]^2 at m/z 1918.2 and triply charged dominant species at m/z 1278.5 [M-3H]^3 which agreed with the calculated molecular weight of mGLP based on published structural studies (4) at 3838.61. This MW of mGLP corresponded to twelve O-methylated Glc, eight Glc, one Glyceric acid and acyl groups comprising three acetyl, one octyl, one isobutyryl. The spectra of the mGLP also revealed that the preparation was contaminated with trace amounts of lysocephatidyl dimannoside (lyso-PIM2) m/z 895.39 and m/z 987.44 (30).

We reasoned that the mGLP could be further resolved into uniform acetylated forms and these separated forms could then be analyzed with MS/MS for the precise location of the acyl groups. The G-50 purified mGLP was subjected to LC/MS with ammonium acetate and acetoniitrile linear gradient and readily yielded the doubly and triply deprotonated anions [M – 2H]^2 and [M – 3H]^3. Overall, fourteen differentially functionalized mGLP isoforms were identified (Fig. 4, Table 2). Among these fourteen forms, seven were non-succinylated (i-vii) and the same seven were found to be modified with a single succinyl residue (ii-viia; MW+100 Da). The major acylforms (iiii & iiiia) of mGLP, as evident from relatively higher abundance of ions (Fig. 4), were found to have three acetyl, one isobutyryl, one octanoyl with and without a succinyl residue. The other acylforms were missing either an acetyl or isobutyryl group, and/or an extra methyl group as detailed in Table 2.

Glycomic and Acylation Profile of diacylated mGLP (MW 3795) by QToF ESI-MS/MS

Identification of the site of acyl modification in mGLP was sought using tandem MS of native mGLP. The triply charged ions at m/z 1264.1, with a mass corresponding to a composition of one 3-O-Me Glc, eleven 6-O-Me Glc, eight Glc, two acetyl, one isobutyryl, one octanoyl and one glyceric acid residues was subjected to the MS/MS fragmentation in the negative ion (ESI) mode using collision induced dissociation (CID) (the ions and structures are presented in Fig. 5 and the mass spectrum in Supporting Information Fig. S-4A-E). Collision energy was optimized to be 80 eV and 100 eV. A charge reduced mass fragmentation pattern was observed whereby singly charged (M-H) product ions were obtained (31).

The non-reducing end m/z 263.1 (C1) fragment and the corresponding m/z 245.1 (B1) fragment account for a 3-O-Me glucosyl residue substituted with one isobutryl residue. The m/z 467.1 (C2) fragment (see Fig 5), has the correct mass of a diglucosyl with one methyl, one isobutyryl and one acetyl residues, signifying that one acetyl function is located on the second Glc unit of the non-reducing end. Additional information on the sequence at the non-reducing end of the molecule came from the ions produced by double glycosidic cleavages. Such cleavages are found in non-derivatized oligosaccharides, when subjected to negative ESI tandem MS, and are labeled with D in Figs 5 and Supporting Information S-4A-E. The subscripts indicate the two cleavage glycosidic bonds counting from non-reducing end (32). The double cleavage ions where the 3-O-Me glucosyl residue (residue 1) is lost proved particularly informative. The m/z 1273.4 (D8,1), 1097.3 (D7,1), 921.30 (D6,1), 745.2 (D5,1), 569.1 (D4,1) fragments all contain three non-methylated glucosyl units with four, three, two and zero 6-O-Me glucosyl residues. The m/z 569.1 (D4,1) fragment corresponded to three non-methylated glucosyl residues with two acetyl groups. Given the fact that the second Glc has an acetyl unit, the additional acetyl must be either on the third or fourth Glc unit. The ions at m/z 379.1 (D3,1)' and 555.1 (D3,2)' show that the second acetyl is on the fourth glucosyl residue. The ions at m/z 365.1 (D3,2) and m/z 717.2 (D6,2)' are consistent with this assignment. These fragments (along with the NMR analysis showing the acetyl groups are on the 6-position) account for a sequence of octaglucose of the non-reducing end of mGLP as α-D-Glcβp(3Me)(6Isobutyryl)-(1→ 4)-α-D-Glcαp(6Ac)-(1→ 4)-α-D-Glcβp(6Me)-(1→ 4)-α-D-Glcβp(6Me)-(1→ 4)-α-D-Glcβp(6Me).

The reducing end Z fragments (Z and X ions are numbered from the reducing end of the molecule)
from m/z 2607.9 (Z12) to m/z down to m/z 1199.4 in 176 Da shows the presence of eight additional unbranched 6-O-methyl glucosyl units towards the reducing end from the non-reducing octaglucoside just described. If the next 6-O-methyl glucoside was also unbranched an ion at m/z 1023.3 should be present however, this ion is missing. Instead the Z5 ion is present 861.3 showing that the reducing end 6-O-methyl glucosyl unit (the fourth main chain glucosyl from the reducing end) is branched with a single glucosyl unit. These leave a reducing end with four non-O-methylated glucosyl residues, one octanoyl and one gliceric acid residue. The ions at of m/z 449.1(D17,16) and m/z 347.1 (\( \text{O}_{2}X_{1} \) -gliceric acid) contain the octanoyl group but not the gliceric acid group and thus show the octanoyl group cannot be attached to the gliceric acid. The HSQC and NOESY NMR experiments showed that the octanoyl group is attached to 2-position of a glucosyl unit (Fig. 2) further substantiated by ion at m/z 435.1 (\( \text{O}_{2}X_{1} \)). The double cleave ion at m/z 1051 (\( \text{O}_{2}A_{2,2,4} \) which contains the octanoyl group along with the branched 6-O-methyl glucosyl residue, both β-glucosyl residues, a linear glucosyl residue and the O-3, O-4, O-5 and O-6 oxygen of the glucosyl residue at the reducing end (but not O-2) rules out the possibility of the octanoyl group being attached to the first glucosyl residue. This conclusion is confirmed by the ion at m/z 147 (\( \text{O}_{2}X_{1} \)) which contain O-2 of the first glucosyl residue but no octanoyl group.

Therefore the entire mGLP follows the following sequence- α-D-GlcP(3Me)(6Isobutyl)- (1 → 4)-α-D-GlcP(6Ac)-(1 → 4)-α-D-GlcP -(1 → 4)-α-D- Glcp(6Ac)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D- Glcp(6Me)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-β-D- Glcp(1 → 3)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D- Glcp(6Me)-(1 → 2) Gliceric acid.

Exhaustive digestion of native mGLP with porcine α-amylase yielded one enzyme major limit product E1 (Fig 6), which was characterized by subsequent LC/MS revealing m/z 1489.06 [M-2H]²⁺ (Supporting Information Fig. S-5). This mass was attributed to the sequence: α-D- Glcp(6Me)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D- Glcp(6Me)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D- Glcp(6Me)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D-GlcP(6Me)(1 → 4)-α-D-GlcP(6Me)[β-D-GlcP- (1 → 3)](1 → 4)-α-D-GlcP(6Me)-(1 → 4)-α-D- Glcp[β-D-GlcP- (1 → 3)](2octanoyl)-(1 → 6)-α-D-GlcP -(1 → 2) Gliceric acid. This indicated that 6-O-Me Glc units are responsible for enzyme resistance. Notably, the succinyll residue along with all the acyls but octanoyl were lost during the digestion. The product E1 was purified from the digest mixture and was tested for its ability to activate γ9δ2 T cells. In comparison to native mGLP, a ~90% loss in γ9δ2 T cells expansion ability was associated with E1 (Fig 7A).

**Discussion**

Bacillus \( \text{Calmette–Guérin} \) (BCG) vaccine is primarily used against tuberculosis particularly in the endemic countries. BCG vaccines comprised of attenuated \( M. \) \( \text{bovis} \) strains, are the only vaccines known to induce protective immunity even when given to infants at birth. Despite these potent stimulatory capacities of \( M. \) \( \text{tuberculosis} \) and BCG, a third of the world is latently infected with \( M. \) \( \text{tuberculosis} \) and ~1.5 million people die annually from TB disease complications. Recent TB vaccine development efforts have focused almost exclusively on the induction of αβ T cells specific for immunodominant peptide epitopes, perhaps only enhancing the evolutionary advantage for the pathogen. It certainly remains possible that induction of the right combination of αβ T cell antigen specificity and effector phenotypes will result in more successful TB vaccines. However, the current goals also involve development of novel vaccine strategies and targets, in tandem with further characterization of nonconventional T cells (33) including protective immune subsets such as γ9δ2.
T cells. The γδ T cells (also termed Vγ9Vδ2) provide a natural bridge between innate and adaptive immunity, rapidly and potently respond to pathogen infection in mucosal tissues, and are prominently induced by both TB infection and BCG vaccination. Therefore, these cells may serve as potent targets for TB immunotherapy. Recent work has demonstrated important TB protective effects of γδ T cells against intracellular replication of mycobacteria in both nonhuman primates (NHP) and human systems (34,35). It has been shown that NHP develop protective memory γδ T cells after BCG vaccination, and has greatly contributed to our knowledge of how these TB protective γδ T cells function in vivo and protect against primary TB and other infections (36-42). Our group has shown that BCG vaccination in humans induces γδ memory T cells (35,43,44) and that these γδ T cells develop pathogen specificity (43,44) can potent intracellular mycobacterial growth (44), and utilize a novel protective mechanism to inhibit intracellular M. tuberculosis (45). In addition, we have now described that mGLP from M. tuberculosis, and not simple phosphoantigens previously shown to activate all γδ T cells, can expand a protective subset of γδ T cells. The mGLP-induced T cell subset expresses a restricted subset of receptor (TCR) sequences. Unlike αβ T cells, γδ T cells do not require antigen processing prior to recognition of certain bacterial, lipid, and tumor antigens. Recognition of these T cells by mGLP is a novel finding and the mechanism is unclear. mGLP is not a major component of M. tuberculosis, in fact we obtained ~200 mg of purified mGLP from 2,500 g of biomass (0.01%). Despite the presence of number of glycosyl residues the molecule is somewhat hydrophobic and acylated making it behave like a lipid rather than a glycan. The octanoyl group in mGLP has been suggested to have a specific role in stabilizing the polysaccharide in helical conformation, in providing it with further discriminatory capability when binding fatty-CoAs and in anchoring mGLP intermediates to the cytoplasmic membrane during the elongation steps (6,46). Due to stereochemical constraints arising from the α-(1→4) linkages in the main chain, mGLP likely adopts a helical conformation in solution, with the methyl groups facing the inner cylindrical hydrophobic cavity (5).

In an effort to establish a structure to function relativity, we took a stepdown approach in dissecting mGLP and testing for biological potency. Thus far, we were able to show that Smith degradation of mGLP leading to the formation of a polylol was unable to expand γδ T cells (Fig. 7B). Following this, the product obtained after mild alkali treatment (mGP: NMR showing the absence of any acyl functions) could only weakly expand T cells indicating that in some fashion, the acyl groups contributed to the biological interplay of mGLP with γδ T cells. Next, we treated mGLP with porcine α-amylase. The major acidic product isolated after size exclusion chromatography was analyzed and yielded one major doubly charged product with m/z 1489.06 [M-2H]2- confirming it to be an oligomer comprised of eleven methylGlc:five Glc:one Glyceric acid and one octanoyl residue (Fig. 6). This enzyme resistant product was also unable to induce γδ T cells. These results prompted us to conclude that all or some of the small acyl residues located at the non-reducing end of the molecule prior to the assembly of the 6-O-methyl Glc containing glycan segment are responsible for the TB-specific γδ T cell inducing biologic activity. Because of the involvement of the acyl groups in the specific biological activity we set out to identify the acyl group at each position of esterification. It has been shown clearly that acyl functions were located at the terminal 3-O-methyl Glc end and the glyceric acid and the octanoyl group are at the reducing end of the molecule (6).

For current analyses, we relied on 2D-NMR and ESI tandem MS. The questions we sought were: a) if the five small acyl groups in mGLP (three acetyl, one isobutyryl, one succinyl) influence the T cell recognition and b) how are these distributed throughout the molecule and c) what is the degree of heterogeneity. Due to the complexity and heterogeneity in the molecule, we took one dominant LC purified homogeneous isofrom MW 3795 (m/z 1264.1 [M-3H]3), composition: diacylated mGLP) for a detailed MS/MS experiment by sensitive ultra-high resolution QToF MS/MS. A detailed
fragmentation analysis indicated the two acetates to be on the 6- position of the non-reducing end Glcp(s), and NMR and mass spectrometry analyses suggest that in the *M. tuberculosis* mGLP, the octanoyl group is at the C2 position of the second Glcp from the reducing end.

To summarize, from the structure/function relationship studies, it is now possible to build upon this knowledge such that synthetic intermediates can be generated in large amounts for in vivo application.

**Experimental Procedure**

**Materials**

Sephadex G-50 (fine) was obtained from Sigma-Aldrich. All reagents for biochemical separations were obtained from Acros Organics (silica gel 60) and Sigma-Aldrich (chloroform and methanol).

**Isolation of mGLP**

Large scale Mid-log mycobacterial cultures (*M. tuberculosis*, H37Rv) were harvested and washed twice with sterile H2O. The pellet was lyophilized and extracted with chloroform/methanol/H2O (10/10/3 vol/vol/vol) twice at room temperature. The organic phase was dried under N2 and stored at −20 °C. Total lipid extracts were fractionated over silica gel-60 (EM Science, Fort Washington, PA, USA) using increasing amounts of methanol in chloroform. The 100 percent methanol eluent was found to be enriched with mGLP (16) and taken for further resolution.

**Purification of mGLP**

100% CH3OH fraction (12 mg of carbohydrate) was dissolved in water (0.5 mL), and applied onto the size exclusion column (Sephadex G-50; 114 x 0.75 cm), and eluted with water. The flow rate was maintained at 0.55 mL/min. Fractions (120; each 2.5 mL per fraction/5.0 min) were collected. A quick α-naphthol charring TLC assay was performed to identify the carbohydrate-enriched (25th to 43th; 62.5 to 107 mL) fractions. Every three consecutive fractions were then pooled, and the monosaccharide composition was determined after derivatization using GC/MS. The analysis revealed that the 62.5 to 70 mL fraction had the enriched mGLP (without detectable impurities by 1H-NMR). Overall purification and characterization strategy has been presented as flow sheet in the Supporting Information Fig. S-6.

**Deacylation of mGLP**

G-50 purified mGLP (1.0mg) was dissolved in 0.2 N NaOH (1mL) and allowed to remain at 55° for 2 hours. It was then neutralized with acetic acid and the solution dried and applied to a Bio-Gel P-2 column (0.5x50) in water for desalting. The deacylated product (mGP) was checked by 1H NMR to ensure completion of deacylation.

**Porcine α-Amylase treatment**

mGP (1.0 mg) and mGLP (3.0 mg) in phosphate buffered saline (PBS, pH 7.2) were treated simultaneously with porcine-pancreatic-α-amylase (Sigma, 3 and 6 units) for 24 and 72 hours respectively at 37°C. Thin-layer chromatography (TLC) on Silica 60 gel plates (Merck) with a solvent system composed of chloroform/methanol/water/ (56:38:10 vol/vol) was used to monitor enzyme activity (formation of new products and utilization of starting material) visualized by spraying with α-naphthol-sulfuric acid solution followed by charring at 120°C. Whereas mGP was digested fully, only partial change was observed with mGLP. Enzyme was deactivated and the digestion mixture desalted on a Bio-Gel P-2 column (0.5 x 30) followed by a G-50 column. The α-naphthol positive fractions were used for downstream analyses.

**Monosaccharide composition**

Aliquots of G-50 column eluents were hydrolyzed with 2M trifluoroacetic acid, converted to alditol acetates, and analyzed using GC/MS performed as described previously (47).

**1D and 2D NMR analysis**

All PRESAT 1H NMR were recorded in D2O on a 400-MHz Innova (Varian) and 2D (TOCSY, C2HSQC and NOESY) on a 500MHz Innova (Varian) instrument at 25 °C. All chemical shifts are based on the reference to the HOD peak at δ4.64 ppm. In-built Varian parameters were used for recording spectra.

**Liquid chromatography Time-of-Flight mass spectrometry**

**Accurate mass LC/MS analyses**

LC/MS were performed on an Agilent 1260 Infinity series HPLC in-line with a 6224 time of...

*Structural and functional diversity in mGLP*
flight (TOF) MS equipped with a multimode ESI/APCI ionization source operated in negative ESI mode. Gradient separation of a 0.4 mg/mL solution of mGLP in water was performed over an HPLC column (Waters X-Bridge C18, 2.1 x 150 mm, 3.5 µm particle size) held at 40 °C, with a consistent 0.32 mL/min flow rate. Injections were 2 µL and all solvents were LC/MS grade (Fisher Optima). Starting conditions, 90% solvent A (H2O with 10 mM ammonium acetate), 10% solvent B (acetonitrile with 10 mM ammonium acetate), were held for 5 minutes and then increased to 70% B over 10 minutes in a linear gradient, followed by an increase to 100% B over 1 minute, then held for 4 minutes as a wash step. MS instrument parameter settings were: gas temperature = 310 °C, vaporizer temperature = 200 °C, gas flow = 10 mL/min, nebulizer pressure = 45 psig, and charging voltage = 2000 V. MS source parameter settings were set as: capillary voltage = 2500 V, fragmentor = 40, skimmer1 = 60.

**Liquid chromatography quadrupole Time-of-Flight mass spectrometry**

Structural elucidation of the mGLP was carried out by ultra-performance liquid chromatography (UPLC) on a Waters Acquity UPLC H-Class system coupled to a Bruker MaXis Plus QTOF MS instrument. Separation was performed in gradient mode with a Waters Acquity UPLC BEH C18 1.7 um column (2.1 x 50 mm) at 40 °C. Mobile phase components were 10 mM ammonium acetate in water (A) and acetonitrile (B). The flow rate was 0.4 mL/min. The proportion of acetonitrile was increased from 10% to 70% in 3 min, to 100% in 3.4 min, and held at 100% for 1.4 min. The post-time was 2 min and injection volume 3 µL.

Internal instrument mass-scale calibration was performed in enhanced quadratic mode during chromatographic dead time by infusing Agilent ESI-L low concentration tuning mix.

Data acquisition in the negative electrospray ion mode with a mass-to-charge ratio (m/z) range of 110-4000 at 1Hz was performed in the full MS scan mode for the first 2 minutes, during which the internal calibrant was introduced into the LC flow. This was followed by a one-minute MRM scan mode alternating collision-induced dissociation (CID) energies of 6eV and 40eV on m/z 1264.1 (M-3H)³ parent ion with m/z width of 0 and 6 respectively. The final 3.8 minutes were in full MS scan mode. Source settings for all time segments were: capillary voltage 2400V, end-plate offset 500V, nebulizer gas pressure 3 bar, drying gas flow 10 L/min, and drying temperature 300 °C. Instrument controls were performed via the Bruker HyStar v4.1 software package. Data was processed using Bruker Compass 2.0 Data Analysis 4.4 software.

**γδT cell stimulatory activity**

The assay was performed as previously described (16). Briefly, to expand γδT cells, isolated PBMCs (1x10⁶) were cultured with novel antigen fractions or controls (medium rested and 20ug/ml MtBL). On day 7, the PBMCs were harvested, counted, stained with anti-γδTCR (clone 11F2), anti-αβTCR (clone B3) and anti-CD3 PerCP (clone SK7). Absolute numbers (AN) of γδT cells were computed by multiplying the flow cytometric percentages times the numbers of viable cells present after expansion. Expansion indices (EI) were calculated as the fold expansion of the absolute number of γδT cells after stimulation with treated lysates compared to the absolute number of γδT cells after rest in medium.

**Acknowledgements**

This project has been funded in whole or in part with federal funds from the National Institutes of Health, 2R01AI048391-09 (to Daniel F. Hoft), and by the Bill & Melinda Gates Foundation, Grant OPP1118659 (to Daniel F. Hoft). We thank MJ Riches and Milena Vaselinovic for technical help with the LC/MS analysis.

**Conflict of Interest**

The authors declare that they have no conflicts of interest with the contents of the article.
References

1. Chatterjee, D., and Brennan, P. J. (2009) Glycosylated components of the mycobacterial cell wall; structure and function. in Microbial Glycobiology: Structures, Relevance and Applications (Holst, O., Brennan, P. J., and Itzstein, v. M. eds.), First Ed., Academic Press, Oxford, UK. pp 147-167
2. Yabusaki, K. K., and Ballou, C. E. (1979) Effect of polymethylpolysaccharides on the hydrolysis of palmitoyl coenzyme A by a thioesterase from Mycobacterium smegmatis. J Biol Chem 254, 12314-12317
3. Bloch, K. (1977) Control mechanisms for fatty acid synthesis in Mycobacterium smegmatis. Adv Enzymol Relat Areas Mol Biol 45, 1-84
4. Forsberg, L. S., Dell, A., Walton, D. J., and Ballou, C. E. (1982) Revised structure for the 6-O-methylglucose polysaccharide of Mycobacterium smegmatis. J Biol Chem 257, 3555-3563
5. Jackson, M., and Brennan, P. J. (2009) Polymethylated polysaccharides from Mycobacterium species revisited. J Biol Chem 284, 1949-1953
6. Smith, W. L., and Ballou, C. E. (1973) The 6-O-methylglucose-containing lipopolysaccharides of Mycobacterium phlei. Locations of the neutral and acidic acyl groups. J Biol Chem 248, 7118-7125
7. Lee, Y. C., and Ballou, C. E. (1964) 6-O-Methyl-D-Glucose from Mycobacteria. J Biol Chem 239, PC3602-3603
8. Lee, Y. C. (1966) Isolation and characterization of lipopolysaccharides containing 6-O-methyl-D-glucose from Mycobacterium species. J Biol Chem 241, 1899-1908
9. Pommier, M. T., and Michel, G. (1986) Isolation and characterization of an O-methylglucose-containing lipopolysaccharide produced by Nocardia otitidis-caviarum. J Gen Microbiol 132, 2433-2441
10. Tuffal, G., Albigot, R., Riviere, M., and Puzo, G. (1998) Newly found 2-N-acetyl-2,6-dideoxy-beta-glucopyranose containing methyl glucose polysaccharides in M.bovis BCG: revised structure of the mycobacterial methyl glucose lipopolysaccharides. Glycobiology 8, 675-684
11. Sambou, T., Dinadayala, P., Stadthagen, G., Barilone, N., Bordat, Y., Constant, P., Levillain, F., Neyrolles, O., Gicquel, B., Lemassu, A., Daffe, M., and Jackson, M. (2008) Capsular glucan and intracellular glycogen of Mycobacterium tuberculosis: biosynthesis and impact on the persistence in mice. Mol Microbiol 70, 762-774
12. Stadthagen, G., Sambou, T., Guerin, M., Barilone, N., Boudou, F., Kordulakova, J., Charles, P., Alzari, P. M., Lemassu, A., Daffe, M., Puzo, G., Gicquel, B., Riviere, M., and Jackson, M. (2007) Genetic basis for the biosynthesis of methylglucose lipopolysaccharides in Mycobacterium tuberculosis. J Biol Chem 282, 27270-27276
13. Gest, P., Kaur, D., Pham, H. T., van der Woerd, M., Hansen, E., Brennan, P. J., Jackson, M., and Guerin, M. E. (2008) Preliminary crystallographic analysis of GpgS, a key glucosyltransferase involved in methylglucose lipopolysaccharide biosynthesis in Mycobacterium tuberculosis. Acta Crystallogr Sect F Struct Biol Cryst Commun 64, 1121-1124
Structural and functional diversity in mGLP

Kaur, D., Pham, H., Larrouy-Maumus, G., Riviere, M., Vissa, V., Guerin, M. E., Puzo, G., Brennan, P. J., and Jackson, M. (2009) Initiation of methylglucose lipopolysaccharide biosynthesis in mycobacteria. *PLoS One* **4**, e5447

Kamisango, K., Dell, A., and Ballou, C. E. (1987) Biosynthesis of the mycobacterial O-methylglucose lipopolysaccharide. Characterization of putative intermediates in the initiation, elongation, and termination reactions. *J Biol Chem* **262**, 4580-4586

Xia, M., Hesser, D. C., De, P., Sakala, I. G., Spencer, C. T., Kirkwood, J. S., Abate, G., Chatterjee, D., Dobos, K. M., and Hoft, D. F. (2016) A Subset of Protective gamma9delta2 T cells is Activated by Novel Mycobacterial Lipid Components. *Infect Immun* **84**, 2449-2462

Dell, A., and Ballou, C. E. (1983) Fast-Atom-Bombardment, Negative-Ion Mass-Spectrometry of the Mycobacterial O-Methyl-Deuterium-Glucose Polysaccharide and Lipopolysaccharides. *Carbohydrate Research* **120**, 95-111

Sawyer, D. T., and Brannan, J. R. (1966) Proton Nuclear Magnetic Resonance Studies of Several Polyalcohols, Hydroxy Acids, and Derivatives of D-Gluconic Acid. *Anal Chem.* **38**, 192-198

Hill, R. E., Iwanow, A., Sayer, B. G., Wysocka, W., and Spenser, I. D. (1987) The regiochemistry and stereochemistry of the biosynthesis of vitamin B6 from triose units. *J Biol Chem* **262**, 7463-7471

Anet, F. A. L., and Park, J. (1992) Proton chemical shift assignments in citrate and trimethyl citrate in chiral media. *J. Am. Chem. Soc.* **114**, 411-416

Tuffal, G., Albigot, R., Monsarrat, B., Ponthus, C., Picard, C., Rivière, M., and Puzo, G. (1995) Purification and LSIMS analysis of methyl glucose polysaccharides from *Mycobacterium xenopi*, a slow growing mycobacterium. *J.Carbohydr.Chem.* **14**, 631-642

Dinadayala, P., Lemassu, A., Granovski, P., Cerantola, S., Winter, N., and Daffe, M. (2004) Revisiting the structure of the anti-neoplastic glucans of Mycobacterium bovis Bacille Calmette-Guerin. Structural analysis of the extracellular and boiling water extract-derived glucans of the vaccine substrains. *J Biol Chem* **279**, 12369-12378

Thiem, J., and Matwiejuk, M. (2011) Defining oxyanion reactivities in base-promoted glycosylations. *Chemical Communications* **47**, 8379-8381

Gastebois, A., Mouyna, I., Simenel, C., Clavaud, C., Coddeville, B., Delepierre, M., Latge, J. P., and Fontaine, T. (2010) Characterization of a new beta(1-3)-glucan branching activity of Aspergillus fumigatus. *J Biol Chem* **285**, 2386-2396

Usui, T., Kazuo Matsuda, N. Y., Tazimura, K., Sugiyama, H., and, and Seto, S. (2014) 13C and 1H Nuclear Magnetic Resonance for Structural Investigation of some Glucans. *Journal of Agricultural and Biological Chemistry* **39**, 1071-1076

Meng, X., Pijning, T., Dobruchowska, J. M., Yin, H., Gerwig, G. J., and Dijkhuizen, L. (2016) Structural determinants of alternating (alpha1 --> 4) and (alpha1 --> 6) linkage specificity in reuteransucrase of Lactobacillus reuteri. *Sci Rep* **6**, 35261

Alalouf, O., Balazs, Y., Volkinshtein, M., Grimpel, Y., Shoham, G., and Shoham, Y. (2011) A new family of carbohydrate esterases is represented by a GDSL hydrolase/acetylxylan esterase from Geobacillus stearothermophilus. *J Biol Chem* **286**, 41993-42001
28. Maranha, A., Moynihan, P., Miranda, V., Correia Lourenço, E., Nunes-Costa, D., Fraga JS, P., J. B. P., Macedo-Ribeiro S, Ventura, M. R., Clarke, A. J., and Empadinhas, N. (2015) Octanoylation of early intermediates of mycobacterial methylglucose lipopolysaccharides. *Sci Rep* 5, 13610

29. Nunes-Costa, D., Maranha, A., Costa, M., Alarico, S., and Empadinhas, N. (2017) Glucosylglycerate metabolism, bioversatility and mycobacterial survival. *Glycobiology* 27, 213-227

30. Hsu, F.-F., Turk, J., Owens, R. M., Rhoades, E. R., and Russell, D. G. (2007) Structural Characterization of Phosphatidyl-myoinositol Mannosides from Mycobacterium bovis Bacillus Calmette Guérin by Multiple-Stage Quadrupole Ion-Trap Mass Spectrometry with Electrospray Ionization. I. PIMs and Lyso-PIMs. *Journal of the American Society for Mass Spectrometry* 18, 466-478

31. Hopper, J. T. S., Kleitos, S., and J., O. N. (2012) Charge state and adduct reduction in electrospray ionization – mass spectrometry using solvent vapor exposure. *Analytical Biochemistry* 421, 788-790

32. Zhang, Z., and Linhardt, R. J. (2009) Sequence Analysis of Native Oligosaccharides Using Negative ESI Tandem MS. *Current Analytical Chemistry* 5, 225-237

33. Godfrey, D. I., Uldrich, A. P., McCluskey, J., Rossjohn, J., and Moody, D. B. (2015) The burgeoning family of unconventional T cells. *Nat Immunol* 16, 1114-1123

34. Chen, Z. W. (2016) Protective immune responses of major Vgamma2Vdelta2 T-cell subset in M. tuberculosis infection. *Curr Opin Immunol* 42, 105-112

35. Abate, G., Spencer, C. T., Hamzabegovic, F., Blazevic, A., Xia, M., and Hoft, D. F. (2015) Mycobacterium-Specific gamma9delta2 T Cells Mediate Both Pathogen-Inhibitory and CD40 Ligand-Dependent Antigen Presentation Effects Important for Tuberculosis Immunity. *Infect Immun* 84, 580-589

36. Chen, C. Y., Yao, S., Huang, D., Wei, H., Sicard, H., Zeng, G., Jomaa, H., Larsen, M. H., Jacobs, W. R., Jr., Wang, R., Letvin, N., Shen, Y., Qiu, L., Shen, L., and Chen, Z. W. (2013) Phosphoantigen/IL2 expansion and differentiation of Vgamma2Vdelta2 T cells increase resistance to tuberculosis in nonhuman primates. *PLoS Pathog* 9, e1003501

37. Chen, Z. W. (2013) Diverse immunological roles of gammadelta T cells. *Cell Mol Immunol* 10, 1

38. Chen, Z. W. (2013) Multifunctional immune responses of HMBPP-specific Vgamma2Vdelta2 T cells in M. tuberculosis and other infections. *Cell Mol Immunol* 10, 58-64

39. Chen, Z. W., and Letvin, N. L. (2003) Vgamma2Vdelta2+ T cells and anti-microbial immune responses. *Microbes Infect* 5, 491-498

40. Chen, Z. W., and Letvin, N. L. (2003) Adaptive immune response of Vgamma2Vdelta2 T cells: a new paradigm. *Trends Immunol* 24, 213-219

41. Shen, L., Shen, Y., Huang, D., Qiu, L., Sehgal, P., Du, G. Z., Miller, M. D., Letvin, N. L., and Chen, Z. W. (2004) Development of Vgamma2Vdelta2+ T cell responses during active mycobacterial coinfection of simian immunodeficiency virus-infected macaques requires control of viral infection and immune competence of CD4+ T cells. *J Infect Dis* 190, 1438-1447

42. Shen, Y., Zhou, D., Qiu, L., Lai, X., Simon, M., Shen, L., Kou, Z., Wang, Q., Jiang, L., Estep, J., Hunt, R., Clagett, M., Sehgal, P. K., Li, Y., Zeng, X., Morita, C. T., Brenner, M. B.,
Letvin, N. L., and Chen, Z. W. (2002) Adaptive immune response of Vgamma2Vdelta2+ T cells during mycobacterial infections. *Science* **295**, 2255-2258

43. Spencer, C. T., Abate, G., Blazevic, A., and Hoft, D. F. (2008) Only a subset of phosphoantigen-responsive gamma9delta2 T cells mediate protective tuberculosis immunity. *J Immunol* **181**, 4471-4484

44. Worku, S., and Hoft, D. F. (2003) Differential effects of control and antigen-specific T cells on intracellular mycobacterial growth. *Infect Immun* **71**, 1763-1773

45. Spencer, C. T., Abate, G., Sakala, I. G., Xia, M., Truscott, S. M., Eickhoff, C. S., Linn, R., Blazevic, A., Metkar, S. S., Peng, G., Froelich, C. J., and Hoft, D. F. (2013) Granzyme A produced by gamma9delta2 T cells induces human macrophages to inhibit growth of an intracellular pathogen. *PLoS Pathog* **9**, e1003119

46. Hindsgaul, O., and Ballou, C. E. (1984) Affinity purification of mycobacterial polymethyl polysaccharides and a study of polysaccharide-lipid interactions by 1H NMR. *Biochemistry* **23**, 577-584

47. Chatterjee, D., Hunter, S. W., McNeil, M., and Brennan, P. J. (1992) Lipoarabinomannan. Multiglycosylated form of the mycobacterial mannosylphophatidylinositol. *J. Biol. Chem.* **267**, 6228-6233

Abbreviations: mGLP, 6-O-methylglucolipopolysaccharide; mGP, 6-O-methylglucopolysaccharide; lc-ms, liquid chromatography-mass spectrometry; ESI, Electro spray ionization; NMR, Nuclear Magnetic Resonance.
Table 1: Characterization of acyl functions of mGLP

| Acyl Groups | <sup>1</sup>H Chemical Shift (δ ppm)# | <sup>13</sup>C Chemical Shift (δ ppm) | <sup>1</sup>H Multiplicity | Proton Connectivity | Connectivity confirmation |
|-------------|-------------------------------------|-------------------------------------|---------------------------|-------------------|--------------------------|
| Isobutyryl  | 2.55                                | 33.8                                | Overlap with -COCH(Me)_2 | 2X CH<sub>3</sub> | HSQC, MS/MS               |
| -COCH(Me)<sub>2</sub> |                                      |                                     |                           |                   |                          |
| Acetyl      | 1.99, 2.00, 2.03                     | 20.0                                | 3X singlets              | -                 | HSQC, TOCSY               |
| Succinyl    | C-6; β-Glc                           |                                     |                           |                   |                          |
| -COCH<sub>2</sub>COOH | 2.55                                | 30.0                                | Overlap with -COCH<sub>2</sub>-Succinyl | -CH<sub>2</sub>- | HSQC, TOCSY               |
| Octanoyl    | C-2, Gc                              |                                     |                           |                   |                          |
| -COCH<sub>2</sub>(CH<sub>2</sub>)<sub>3</sub>CH<sub>3</sub> | 2.37                                | 35.0                                | Overlap with -CH<sub>3</sub>- | -CH<sub>3</sub>- | HSQC, TOCSY               |
| -COCH<sub>2</sub>(CH<sub>2</sub>)<sub>3</sub>CH<sub>3</sub> | 1.52                                | 25.8                                | multiplet 2X -CH<sub>3</sub>- | (2.37, 1.26 ppm) | HSQC, TOCSY               |
| 3X -COCH<sub>3</sub> |                                      |                                      |                           |                   |                          |
| Succinyl    | C-6, Gc                              |                                     |                           |                   |                          |
| Acetyl      | C-6; 3-OMe Glc                       |                                     |                           |                   |                          |

HSQC, <sup>1</sup>H-<sup>13</sup>C correlation NMR spectroscopy; TOCSY, Total <sup>1</sup>H-<sup>1</sup>H Correlation NMR Spectroscopy; NOESY, Nuclear Overhauser Effect NMR Spectroscopy; Through space <sup>1</sup>H-<sup>1</sup>H correlation; MS/MS, Tandem Mass Spectroscopy; *Ref: J Biol Chem 284, 1949-1953, (2009), Scientific Reports 5, 13610 (2015). # Proton reference, HOD δ 4.64 ppm
Table 2: Acyl modifications and Heterogeneity on the carbohydrate backbone—Analysis by LC/MS

| 3-O-Me Glc: 6-O-Me Glc: Glc | 1:11:8 | Remarks |
|-------------------------------|--------|---------|
| Extra-O Methyl (14 Da)       |        |         |
| Backbone variation           |        |         |
| 0 (m/z 1756.6; M-2)          | +1     | -1      |
| (m/z 1170.7; M-3)            | (m/z 1175.4; M-4) | (m/z 1166.0; M-3) |
| Remarks                      |        |         |
| Non-succinyl                 |        |         |
| Octanoyl (127 Da)            | 1      | 1       |
| Isobutyryl (72 Da)           | 0      | 1       |
| Acetyl (42 Da)               | 3      | 2       |
| m/z (M-2)                    | 1882.7 | 1917.7  |
| m/z (M-3)                    | 1254.8 | 1278.1  |
| mGLP Isoforms                | ia     | iii     |
| m/z (M-2)                    | 1932.7 | 1967.7  |
| m/z (M-3)                    | 1288.1 | 1311.4  |
| Succinyl (+100 Da)           |        |         |

*Supporting Information Fig. S-4
Fig. 1: $^1$H NMR of A. deacylated mGLP (mGP): The deacylation was achieved by mild base hydrolysis followed by desalting. The anomic region ($\delta$ 4.7-5.5 ppm) revealed the type of glycosidic linkages present in mGLP backbone and B. mGLP: the aliphatic region (between $\delta$ 0.5-3.0 ppm) revealed the different acyl groups present in mGLP. Succinyl group was confirmed with HSQC NMR (in Fig.2)
Fig. 2: HSQC NMR (H-C 2D correlation) spectrum of G-50 purified native mGLP: NMR was performed in D$_2$O at room temperature. Red contour peaks correspond to methylene (-CH$_2$-) groups and blue contour peaks correspond to methyl (-CH$_3$) and methine (>CH-) groups. GA: glyceric acid.
Fig. 3: The Nuclear Overhauser Effect correlation spectroscopy (NOESY) of native mGLP (D$_2$O, no spin, MW ~3800, mixing time 0.3 sec): Through space correlation of all of protons in the octanoyl residue with the Glcp ring-protons; succinyl residue (Inset: magnified $\delta$ 4.0-5.2 ppm); acylated (possibly succinylated) methylene protons; $\alpha(1\rightarrow4)$ and $\alpha(1\rightarrow6)$ anomeric protons signify octanoyl as a ring substitution other than C-6 of Glcp.
Structural and functional diversity in mGLP

Fig. 4: LC/MS (negative ionization) of native mGLP: Reverse phase C-18 column with NH4OAc:CH3CN gradient was used to resolve the isoforms. Each ion cluster corresponds to one isoform of mGLP (altogether fourteen isoforms; seven non-succinylated and seven succinylated); upper panel represents (M-2H)-2 ions and lower panel represents corresponding (M-3H)-3 ions.
Fig. 5: Tandem mass-spectroscopic analysis of mGLP. ESI-Collision Induced Dissociation; negative ion; 80 eV, 100 eV fragment ions (M-H)^- of LC-pure native mGLP isoform (MW 3795; m/z 1264.1 (M-3H)^3). Zi and Xi ions correspond to the number of glycosidic-linkage from the reducing end; Ci, Di and Ai ions correspond to the number of glycosidic-linkage from the non-reducing end. The fragment ions suggested the locations of isobutyryl, acetyl, octanoyl, glyceric acid residues and β-D-Glcp-(1→3) branches on mGLP skeleton.
Fig. 6: Representation of native mGLP and its enzymatic porcine α (1→4) amylase digestion product. The mass spectrometry of the major product (Supplementary Information Fig S-5) isolated corresponded to the above structure drawn of the reducing end of mGLP (with glyceric acid and octanoyl groups intact) arising after the enzymatic cleavage of three Glcp plus one Glcp(3Me) units from the non-reducing end carrying isobutyryl and acetyl residues.
Fig 7: γδ2 T-cell expansion profile of mGLP derivatives with different human PBMC volunteers. A. Concentration (0.01, 0.1 and 1.0 µg/mL) wise γδ2 T-cell expansion profile of mGLP derivatives (Absolute Numbers of expanded T-cells with 3 volunteers). △ Media rested (MR) + Interleukin 2 (IL2) is the baseline control in the absence of any antigen. □ Native mGLP + IL2 showed the best expansion ability at 0.1 µg/mL; saturation of biological response may be responsible for a dip in expanded T-cell numbers at 1.0 µg/mL. The enzyme digested product ○ mGLP E1+IL2 (lost four non-reducing end hexoses, two-three acetyl and one isobutyryl group of mGLP) showed inability for T-cell expansion at 0.01 or 0.1 µg/mL but a very weak expansion at 1.0 µg/mL. B. Concentration (0.01, 0.1 and 1.0 µg/mL) wise γδ2 T-cell expansion profile of mGLP derivatives (Absolute Numbers of expanded T-cells with two volunteers). △ Media rested (MR) + Interleukin 2 (IL2) is the. □ Native mGLP + IL2 showed the best expansion ability at 0.1 µg/mL. X The deacetylated mGLP (mGP +IL-2) showed inability for T-cell expansion at 0.01 or 0.1 µg/mL but a very weak expansion at 1.0 µg/mL. ◊ The Smith degraded product from mGLP +IL2 showed inability for T-cell expansion at 0.01 or 0.1 µg/mL but a very weak expansion at 1.0 µg/mL.
Structural determinants in a glucose-containing lipopolysaccharide from Mycobacterium tuberculosis critical for inducing a subset of protective T cells
Prithviraj De, Michael McNeil, Mei Xia, Claudia M. Boot, Danny C. Hesser, Karolien Denef, Christopher Rithner, Tyler Sours, Karen M. Dobos, Daniel Hoft and Delphi Chatterjee

J. Biol. Chem. published online May 1, 2018

Access the most updated version of this article at doi: 10.1074/jbc.RA118.002582

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts