J|A|C|S ARTICLES

Published on Web 10/21/2010

Structural Analysis of an Unusual Bioactive *N*-Acylated Lipo-Oligosaccharide LOS-IV in *Mycobacterium marinum*

Yoann Rombouts,[†,‡] Elisabeth Elass,[†,‡] Christophe Biot,[†,‡] Emmanuel Maes,[†,‡] Bernadette Coddeville,[†,‡] Adeline Burguière,[§] Caroline Tokarski,[⊥] Eric Buisine,[#] Xavier Trivelli,[†,‡] Laurent Kremer,[§,||] and Yann Guérardel*[,†,‡]

Université de Lille 1, Unité de Glycobiologie Structurale et Fonctionnelle, UGSF, F-59650 Villeneuve d'Ascq, France, CNRS, UMR 8576, F-59650 Villeneuve d'Ascq, France, Laboratoire de Dynamique des Interactions Membranaires Normales et Pathologiques, Université de Montpellier II et I, CNRS UMR 5235, case 107, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France, INSERM, DIMNP, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France, Miniaturisation pour l'Analyse, la Synthèse & la Protéomique (MSAP), USR CNRS 3290, IFR 147, Université de Lille 1 Sciences et Technologies, 59655 Villeneuve d'Ascq Cedex, France, and Ecole Nationale Supérieure de Chimie de Lille, Bâtiment C7, Avenue Mendeleïev - B.P. 90108, 59652 Villeneuve d'Ascq Cedex, France

Received July 1, 2010; E-mail: yann.guerardel@univ-lille1.fr

Abstract: Although lipo-oligosaccharides (LOSs) are recognized as major parietal components in many mycobacterial species, their involvement in the host−pathogen interactions have been scarcely documented. In particular, the biological implications arising from the high degree of structural species-specificity of these glycolipids remain largely unknown. Growing recognition of the *Mycobacterium marinum*−*Danio rerio* as a specific host−pathogen model devoted to the study of the physiopathology of mycobacterial infections prompted us to elucidate the structure-to-function relationships of the elusive end-product, LOS-IV, of the LOS biosynthetic pathway in *M. marinum*. Combination of physicochemical and molecular modeling methods established that LOS-IV resulted from the differential transfer on the caryophyllose-containing LOS-III of a family of very unusual *N*-acylated monosaccharides, naturally present as different diastereoisomers. In agreement with the partial loss of pathogenecity previously reported in a LOS-IV-deficient *M. marinum* mutant, we demonstrated that this terminal monosaccharide conferred to LOS-IV important biological functions, including macrophage activating properties.

Introduction

Nontuberculous mycobacteria, including *Mycobacterium marinum*, are responsible for opportunistic infections primarily in immune-deficient patients such as AIDS patients.[1] The incapacity to eradicate these diseases can, at least partly, be attributed to the presence of the specific and unusual mycobacterial cell envelope known to play an important role in the pathogenesis of mycobacterial infections and to confer resistance to many antibiotics and bactericidal treatments.[2,3] The mycobacterial cell wall "core" structure consists of long-chain mycolic acids attached to the arabinogalactan/peptidoglycan backbone.[4,5] This mycolyl-arabinogalactan-peptidoglycan structure serves as an anchor for numerous free lipids/glycolipids that are ubiquitous or species-specific,[6] which are important in directing host−pathogen interactions and contributing to modulation of the host immune system.[7] The latter include phthiocerol dimycocerosates,[8] phenolic glycolipids (PGL),[8] phosphatidylinositol mannosides,[3] glycopeptidolipids,[9,10] and trehalose-containing glycolipids such as trehalose dimycolate,[11] sulfolipids,[12,13] multiacylated trehalose,[14] and lipo-oligosaccharides (LOSs).[15]

LOSs are cell surface glycolipids described in at least 10 mycobacterial species including *Mycobacterium kansasii*, the strain Canetti of *Mycobacterium tuberculosis* and *Mycobacterium marinum*.[16−18] Although LOSs are commonly described as highly antigenic glycoconjugates exposed to the cell surface, very little is known with respect to their general functions during the mycobacterial infection.[17,19] All LOSs contain an α,α′-

† Université de Lille 1.
‡ CNRS, UMR 8576.
§ Université de Montpellier II et I.
|| CNRS, UMR 5235, Laboratoire de Dynamique des Interactions Membranaires Normales et Pathologiques.
⊥ Miniaturisation pour l'Analyse, la Synthèse & la Protéomique.
Ecole Nationale Supérieure de Chimie de Lille.

(1) Adle-Biassette, H.; Huerre, M.; Breton, G.; Ruimy, R.; Carbonnelle, A.; Trophilme, D.; Yacoub, M.; Regnier, B.; Yeni, P.; Vilde, J. L.; Henin, D. *Ann. Pathol.* **2003**, *23*, 216–35.
(2) Jarlier, V.; Nikaido, H. *J. Bacteriol.* **1990**, *172*, 1418–23.
(3) Daffe, M.; Draper, P. *Adv. Microbiol. Physiol.* **1998**, *39*, 131–203.
(4) Kremer, L., Baulard, A. R., Besra, G. S. Genetics of mycolic acid biosynthesis. In *Molecular Genetics of Mycobacteria*; Hatfull, G. F., Jacobs, W. R., Jr., Eds.; ASM Press: Washington DC, 2000; pp 173−190.
(5) Bhowruth, V.; Alderwick, L. J.; Brown, A. K.; Bhatt, A.; Besra, G. S. *Biochem. Soc. Trans.* **2008**, *36*, 555–65.
(6) Brennan, P. J.; Nikaido, H. *Annu. Rev. Biochem.* **1995**, *64*, 29–63.
(7) Karakousis, P. C.; Bishai, W. R.; Dorman, S. E. *Cell. Microbiol.* **2004**, *6*, 105–16.
(8) Minnikin, D. E.; Kremer, L.; Dover, L. G.; Besra, G. S. *Chem. Biol.* **2002**, *9*, 545–53.
(9) Chatterjee, D.; Khoo, K. H. *Cell. Mol. Life Sci.* **2001**, *58*, 2018–42.
(10) Schorey, J. S.; Sweet, L. *Glycobiology* **2008**, *18*, 832–41.

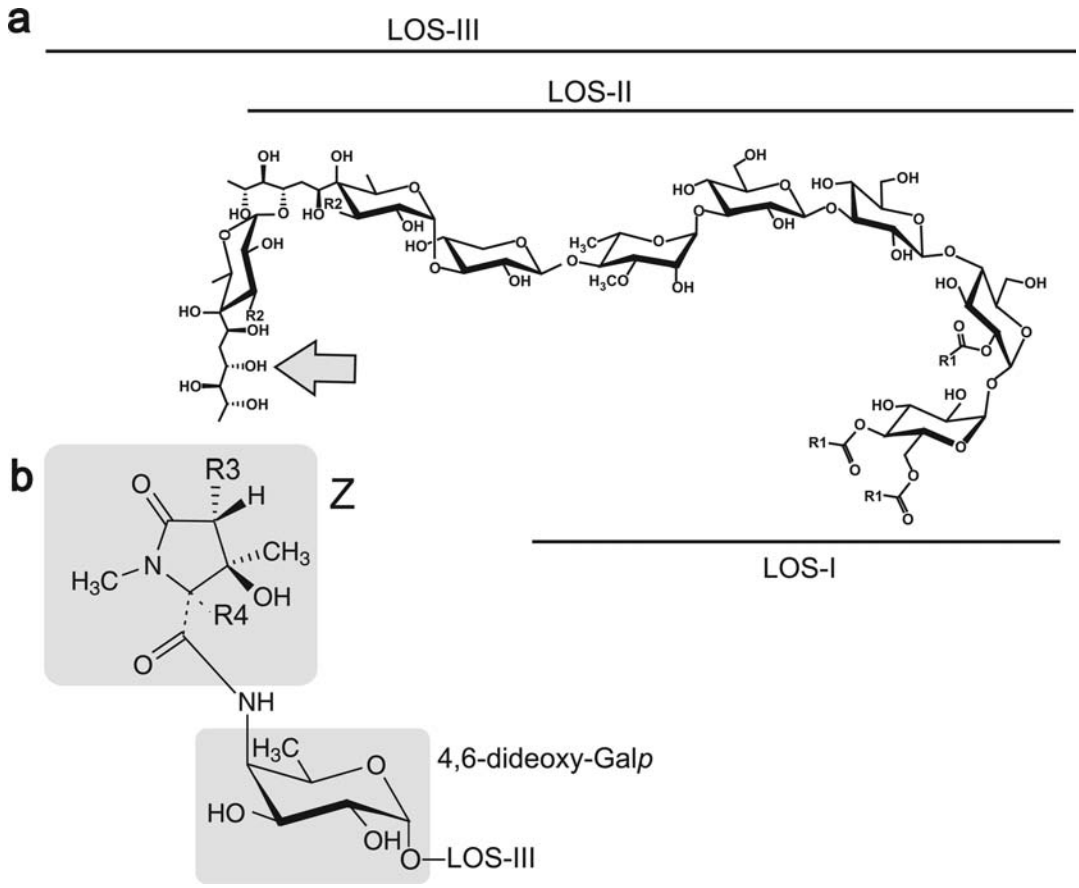

Figure 1. Structure of LOSs from *M. marinum*. (a) All LOS are constituted by a polyacylated βGlc*p*-1→3-βGlc*p*-1→4-αGlc*p*-1→1-αGlc*p* core differentially substituted by 3-O-Me-Rha and Car derivative residues. R1 stands for 2,4-dimethyl-branched fatty acid chains and R2 for either −H or −OH groups. LOS-III is substituted by a so far undescribed substituent on its nonreducing terminal Car residue, as indicated by the arrow. Full structure of each molecule is described in Table S1 of the Supporting Information. (b) Novel family of *N*-acetylated dideoxy-galactoses substituting LOS-III, which structures are described in the present report. R3 stands for either −H or $-OCH_3$ and R4 for either −H or −COOH.

trehalose moiety that is substituted by different species-specific glycan sequences. In *M. marinum*, four major lipooligosaccharides, named LOS-I to LOS-IV (Figure 1 and Table S1 of the Supporting Information), have been previously identified.[20] A fifth molecule, LOS-II*, has been subsequently described as a biosynthetic intermediate between LOS-I and LOS-II.[21] Fine structural determination of their oligosaccharide structures not only allowed to confirm the common core sequence of these molecules but also to unravel two rare monosaccharides belonging to the caryophyllose (Car) family in LOS-II, LOS-III, and LOS-IV (Table S1 of the Supporting Information).[18] However, the exact nature of the unusual nonreducing terminal monosaccharide that typifies the LOS-IV end-product remained uncharacterized. Although this compound was tentatively identified as *N*-acylated α-amino-hexopyranose, the fine structure of its substituent group remained elusive.[18]

In our previous work, we have demonstrated that one particular *M. marinum* isolate, named Mma7, produced an altered LOS profile.[18] Mma7 failed to produce the more polar LOS-IV but instead accumulated large amounts of the LOS-III precursor. Local genomic comparison of the LOS biosynthetic cluster indicated the presence of a highly disorganized region in Mma7 compared to the standard M strain that are likely to be responsible for the defect LOS-IV production in Mma7. In particular, a deletion of the 5′-end of the *losA* gene was identified. Earlier studies emphasized the role of LosA in the synthesis of LOS-IV, since inactivation of LosA was associated with impairment of LOS-IV production and the concomitant

(11) Hunter, R. L.; Olsen, M. R.; Jagannath, C.; Actor, J. K. *Ann. Clin. Lab. Sci.* **2006**, *36*, 371–86.
(12) Mougous, J. D.; Petzold, C. J.; Senaratne, R. H.; Lee, D. H.; Akey, D. L.; Lin, F. L.; Munchel, S. E.; Pratt, M. R.; Riley, L. W.; Leary, J. A.; Berger, J. M.; Bertozzi, C. R. *Nat. Struct. Mol. Biol.* **2004**, *11*, 721–9.
(13) Kumar, P.; Schelle, M. W.; Jain, M.; Lin, F. L.; Petzold, C. J.; Leavell, M. D.; Leary, J. A.; Cox, J. S.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 11221–6.
(14) Hatzios, S. K.; Schelle, M. W.; Holsclaw, C. M.; Behrens, C. R.; Botyanszki, Z.; Lin, F. L.; Carlson, B. L.; Kumar, P.; Leary, J. A.; Bertozzi, C. R. *J. Biol. Chem.* **2009**, *284*, 12745–51.
(15) McNeil, M.; Chatterjee, D.; Hunter, S. W.; Brennan, P. J. *Methods Enzymol.* **1989**, *179*, 215–42.
(16) Hunter, S. W.; Murphy, R. C.; Clay, K.; Goren, M. B.; Brennan, P. J. *J. Biol. Chem.* **1983**, *258*, 10481–7.
(17) Daffe, M.; McNeil, M.; Brennan, P. J. *Biochemistry* **1991**, *30*, 378–88.
(18) Rombouts, Y.; Burguiere, A.; Maes, E.; Coddeville, B.; Elass, E.; Guerardel, Y.; Kremer, L. *J. Biol. Chem.* **2009**, *284*, 20975–88.
(19) Ortalo-Magne, A.; Lemassu, A.; Laneelle, M. A.; Bardou, F.; Silve, G.; Gounon, P.; Marchal, G.; Daffe, M. *J. Bacteriol.* **1996**, *178*, 456–61.
(20) Burguiere, A.; Hitchen, P. G.; Dover, L. G.; Kremer, L.; Ridell, M.; Alexander, D. C.; Liu, J.; Morris, H. R.; Minnikin, D. E.; Dell, A.; Besra, G. S. *J. Biol. Chem.* **2005**, *280*, 42124–33.
(21) Ren, H.; Dover, L. G.; Islam, S. T.; Alexander, D. C.; Chen, J. M.; Besra, G. S.; Liu, J. *Mol. Microbiol.* **2007**, *63*, 1345–59.

accumulation of the biosynthetic precursor LOS-III.[20] Interestingly, a *M. marinum losA* deficient mutant strain was also found to be less efficient to enter macrophages in vitro.[22] Despite the fact that the *losA* product was attributed to the glycosyltransferase responsible for the transfer of terminal monosaccharide on LOS-III, complementation of Mma7 with a functional *losA* failed to restore LOS-IV production.[18] Therefore, to apprehend to which extent LOS-IV participates in the pathogenesis of *M. marinum* and to delineate LOS-IV structure−function relationship activities, it appeared crucial to establish the fine structure of the yet uncharacterized LOS-IV molecule. From a functional point of view, purified LOSs from *M. marinum* were found to inhibit the release of TNF-α in LPS-stimulated macrophages.[18] This unexpected finding suggested that LOSs may represent key effectors capable to interfere with the host immune proinflammatory response, which is central to the physiopathological events that culminate with granuloma formation.[23] However, this process is not only mediated by proinflammatory cytokines secretion but relies also on macrophage cell surface antigen expression. In this context, both the intercellular adhesion molecule-1 antigen (ICAM-1) and the marker of activation CD40 have been reported to promote cell−cell or cell−matrix interactions and T-cell recruitment/activation into granuloma.[24−32]

As a first step toward the elucidation of the biological functions of LOSs in mycobacterial pathogenicity, not only in *M. marinum* but also in other LOS-producing mycobacterial species such as the Canetti strain, we have undertaken a thorough structural description of LOS-IV from *M. marinum*. Thanks to the combination of a vast panel of methods including mass spectrometry, nuclear magnetic resonance, and molecular modeling, we clearly established that LOS-IV was capped by a family of very unusual *N*-acylated 4-amino-galactopyranose residues. These residues were found to bear three different *N*-acyl groups, collectively called "Z" that generate a structural microheterogeneity of LOS-IV (Figure 1, Table S1 of the Supporting Information). We then investigated the role of LOSs in the inflammatory response, by assessing the effect of LOS-IV on the induction of the macrophage cell surface markers and cytokine secretion.

Materials and Methods

***M. marinum* Strains and Growth Culture Conditions.** *M. marinum* strain M, used in this study, was first isolated from a human patient as described previously.[33] Bacteria were grown at 30 °C on plates containing Middlebrook 7H10 supplemented with oleic acid/albumin/dextrose/catalase enrichment or in Sauton's broth medium.

Purification of Neutral and Acidic LOS-IV. Extraction and purification of LOSs were performed as previously described.[18] Briefly, extracted polar lipids were dissolved in chloroform/methanol (2:1, v/v) and applied on a DEAE cellulose column for purification. LOS-I to LOS-III and neutral LOS-IV were eluted by chloroform/methanol (2:1, v/v) from the DEAE column. Acidic LOS-IV was eluted from DEAE by 40−60 mM ammonium acetate in chloroform/methanol (2:1, v/v). Presence of LOS in the eluted fractions was monitored by 1D-HP thin layer chromatography (TLC) on a silica gel 60 glass back TLC plate (20 × 20 cm; Merck) run in chloroform/acetic acid/methanol/water (40:25:3:6, v/v/v/v). LOSs were visualized by spraying the plates with orcinol/sulphuric acid reagent followed by heating. The neutral and acidic LOS-IV were further purified into individual species by preparative TLC on HPTLC plates of silica Gel 60 (Merck) using chloroform/acetic acid/methanol/water (40:25:3:6, v/v/v/v) as the running solvent. The glycolipids and lipids were revealed by iodine vapors. Following detection, plates were dried overnight, and the corresponding glycolipid were scraped from the plates. Glycolipids were extracted from the silica gel using chloroform/methanol (2:1, v/v) and purified on Sep-pack C18 cartridge (Waters, Milford, CT, USA).

Chemical Procedures. Glycolipids were de-*O*-acylated in chloroform/methanol (2:1, v/v) and 0.1 mM NaOH at 37 °C for 2 h. The reagents were evaporated, and the resulting oligosaccharides dissolved in water were purified on a carbograph column (Alltech carbograph SPE column). Oligosaccharides were de-*N*-acylated in 4 M KOH at 100 °C for 16 h, followed by neutralization with HCl. De-*N*-acylated oligosaccharides were further purified on a carbograph column. Carboxylic acid groups of oligosaccharides were converted into methyl esters by incubation in dry dimethylsulfoxide (100 μL) and iodomethane (100 μL) for 2 h at room temperature. After lyophylization, compounds were dissolved in water and purified on a carbograph column.

Matrix-Assisted Laser Desorption Ionization Mass Spectrometry (MALDI-MS) Analysis. The molecular masses of oligosaccharides were measured by MALDI-TOF on a Voyager DE STR reflectron mass spectrometer (PerSeptive Biosystems, Framingham, MA, USA), equipped with a 337 nm UV laser. Samples were prepared by mixing directly on the target 1 μL of water diluted oligosaccharide solution and 1 μL of 2,5-dihydroxybenzoic acid matrix solution (10 mg/mL dissolved in methanol−water).

Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FTICR-MS) Analysis. The sample fractions were diluted with water/methanol (v/v), acidified with 0.1% formic acid to obtain 500 fmol/μL final concentrations. The samples were analyzed with an Apex Qe 9.4 T fourier transform ion cyclotron resonance mass spectrometer (Bruker Daltonics, Bremen, Germany). The FT-ICR mass spectrometer is equipped with a nanoelectrospray source. Detection was carried out in positive mode. A potential of 1.1 kV was applied on the needle (PicoTip Emitter, New Objective, Woburn, MA, USA). The detection parameters were as follows: broadband detection, 512 K acquisition size, start mass at *m/z* 500. Ions were accumulated in the storage hexapole during 1 s and in the second hexapole during 0.01 s. Spectra were calibrated using external calibration based on prior analysis of Angiotensin I (sequence DRVYIHPFHL, $MH^+ = 648.846$ amu) via DataAnalysis software (Bruker Daltonics, Bremen, Germany).

Purification of the Terminal Monosaccharide of Acidic OS-IV. Oligosaccharides were hydrolyzed in 4 M trifluoroacetic acid solution at 100 °C for 4 h. Prior to purification, trifluoroacetic acid was removed by successive coevaporation with methanol under nitrogen flow. Then, the terminal monosaccharide of acidic OS-IV was purified by high performance liquid chromatography (HPLC)

(22) Alexander, D. C.; Jones, J. R.; Tan, T.; Chen, J. M.; Liu, J. *J. Biol. Chem.* **2004**, *279*, 18824–33.
(23) Clay, H.; Volkman, H. E.; Ramakrishnan, L. *Immunity* **2008**, *29*, 283–94.
(24) Ma, J.; Chen, T.; Mandelin, J.; Ceponis, A.; Miller, N. E.; Hukkanen, M.; Ma, G. F.; Konttinen, Y. T. *Cell. Mol. Life Sci.* **2003**, *60*, 2334–46.
(25) Lukacs, N. W.; Chensue, S. W.; Strieter, R. M.; Warmington, K.; Kunkel, S. L. *J. Immunol.* **1994**, *152*, 5883–9.
(26) Lebedeva, T.; Dustin, M. L.; Sykulev, Y. *Curr. Opin. Immunol.* **2005**, *17*, 251–8.
(27) Lopez Ramirez, G. M.; Rom, W. N.; Ciotoli, C.; Talbot, A.; Martiniuk, F.; Cronstein, B.; Reibman, J. *Infect. Immun.* **1994**, *62*, 2515–20.
(28) DesJardin, L. E.; Kaufman, T. M.; Potts, B.; Kutzbach, B.; Yi, H.; Schlesinger, L. S. *Microbiology* **2002**, *148*, 3161–71.
(29) Lazarevic, V.; Myers, A. J.; Scanga, C. A.; Flynn, J. L. *Immunity* **2003**, *19*, 823–35.
(30) Ordway, D.; Henao-Tamayo, M.; Orme, I. M.; Gonzalez-Juarrero, M. *J. Immunol.* **2005**, *175*, 3873–81.
(31) Ordway, D.; Harton, M.; Henao-Tamayo, M.; Montoya, R.; Orme, I. M.; Gonzalez-Juarrero, M. *J. Immunol.* **2006**, *176*, 4931–9.
(32) Hogan, L. H.; Markofski, W.; Bock, A.; Barger, B.; Morrissey, J. D.; Sandor, M. *Infect. Immun.* **2001**, *69*, 2596–603.
(33) Stinear, T. P.; et al. *Genome Res.* **2008**, *18*, 729−41.

on a AMINEX ion exclusion HPX-87H column (7.8 × 300 mm, BIO-RAD, Richmond, VA, USA) using H_2O with a flow rate of 1 mL/min. Monosaccharide was detected by UV spectroscopy at 206 nm.

NMR Analysis. NMR experiments have been performed at 300 K on Bruker Avance 400, Avance 600, AvanceII 800, and AvanceIII 900 spectrometers equipped with a 5 mm broad-band inverse probe, a 5 mm triple resonance cryoprobe, a 3 mm broad-band inverse probe, and a 5 mm triple-resonance cryoprobe, respectively. Prior to NMR spectroscopic analyses into deuterium oxide, mono- and oligosaccharides were repeatedly exchanged in ^2H_2O (99.97% ^2H, Euriso-top, Saint-Aubin, France) with intermediate freeze-drying and finally dissolved in ^2H_2O and transferred in standard or Shigemi (Allison Park, USA) tubes. Samples used in light water contained 10% ^2H_2O. Chemical shifts (ppm) were calibrated taking methyl group from internal acetone at δ^1H 2.225 and δ^{13}C 31.55. Nitrogen-15 was referenced indirectly as described by Wishart et al.[34] All experiments were recorded without sample spinning. The COSY, ROESY, NOESY, ^{13}C-HSQC, ^{13}C-HMBC, ^{15}N-HSQC, and ^{15}N-HMBC experiments were performed using the Bruker standard sequences.

Molecular Modeling. Model structures suitable for computational studies of the aglycon and the first sugar were considered to mimic the enantiomers (2*S*, 3*S*, 4*R*) and (2*R*, 3*S*, 4*R*) of the major aglycon molecule Zc. The conformational space for Zc (2*S*, 3*S*, 4*R*) and Zc (2*R*, 3*S*, 4*R*) was systematically explored using the HF/3-21 g(d) ab initio method as implemented in Gaussian03.[35] The effect of water was evaluated using the integral equation formalism (polarizable continuum) model (IEF-PCM).[36] Finally, geometry optimizations were then performed using density functional theory (DFT) at the B3LYP level[37] with the 6-31 g(d) basis set.

Cell Culture and Flow Cytometry Analysis. Human promonocytic leukemia THP-1 cells (ECACC no. 88081201) were differentiated into macrophages in the presence of 50 nM 1,25-dihydroxy-vitamin D3 (Calbiochem, Darmstadt, Germany) for 72 h. To investigate the effect of LOSs on ICAM-1 expression, differentiated THP-1 cells were seeded at a density of 3.5×10^5 cells/well, in RPMI 1640 supplemented with 2% FCS and L-glutamine, together with the glycolipids. Purified LOSs (LOS-I, LOS-II, and LOS-IV) were resuspended in ethanol, coated on plates and dried at 37 °C, as previously described.[18] Control wells were layered with solvent without glycolipids. The LOS-IV oligosaccharidic moiety (OS-IV) and LPS from *Salmonella typhimurium* (Alexis Corporation), used as positive control were dissolved in apyrogen water and sonicated prior to addition to the cells. Eventual endotoxin contamination of LOS samples was evaluated with the *Limulus* amoebocyte lysate assay kit (QCL1000; BioWhittaker, Walkersville, MD, USA). After 20 h of incubation, expression on THP-1 cells of human ICAM-1 (CD54) and CD40 was analyzed by flow cytometry. Briefly, 250 000 cells were incubated 20 min at 4 °C with 20 μg/mL human IgG (Sigma), washed three times, and incubated for 40 min with 10 μL of PE-conjugated anti-ICAM-1 (CD54) or FITC-conjugated anti-CD40 mouse monoclonal $IgG_{1\kappa}$ antibodies (BD Biosciences) in PBS containing 0.04% NaN_3 and 0.05% BSA. Both PE- and FITC-conjugated mouse isotype control IgG (BD Biosciences) were used as negative controls. Cells were washed and analyzed using a Becton Dickinson FACScalibur flow cytometer and gated for forward- and side-angle light scatters. Approximately 8 000 particles of the gated population were analyzed. The fluorescence channels were set on a logarithmic scale, and the mean fluorescence intensity was determined.

Results

Structural Analyses. Identification of Two LOS-IV Isoforms. *M. marinum* possesses four major LOS subtypes (LOS-I to -IV) that can be distinguished in one-dimensional thin-layer chromatography (TLC).[18] During the purification process, LOS-IV was purified from the other LOS subtypes by anion-exchange chromatography. Indeed, whereas LOS-I to -III were rapidly eluted from the DEAE-cellulose column in the equilibrating solvent, LOS-IV was retained on the column and eluted in the presence of ammonium acetate. Surprisingly, TLC analysis revealed that, in addition to LOS-I, LOS-II, and LOS-III, the nonretained fractions contained an additional unidentified glycolipid exhibiting a chromatographic behavior similar to LOS-IV (Figure S1 of the Supporting Information). This glycolipid was originally assigned to as LOS-IV, presuming that LOS-IV was incompletely retained by anion-exchange chromatography. However, successive recycling of this product on DEAE gel chromatography demonstrated that it had no affinity for anion-exchange gel chromatography, thus suggesting that this compound differed from the LOS-IV retained on the DEAE column. Both fractions, referred to as neutral LOS-IV and acidic LOS-IV, were further purified by preparative HPTLC and studied separately in their intact (LOS) and deacylated forms generated by alkaline hydrolysis (OS). Relative quantification of both fractions established that acidic LOS-IV represented about 95% of total LOS-IV. Then, the fine structures of these compounds were solved by a combination of mass spectrometry (MS), nuclear magnetic resonance (NMR), and molecular modeling. As described below, our structural analysis led to the identification of three different glycolipids structurally related to LOS-IV: neutral LOS-IV consisting of LOS-IVa and LOS-IVb and acidic LOS-IV consisting of a single molecule, LOS-IVc. Although proven to be different, LOS-IVa and LOS-IVb could not be individually separated and therefore were further studied as a mixture.

Comparison of Oligosaccharide Sequences. First, major structural features of neutral and acidic OSs were compared. In a previous study, we established the nature of the monosaccharides and linkage patterns of the oligosaccharide core common to all LOSs from *M. marinum*.[18] Two dimensional ^{1}H/^{1}H-TOCSY and ^{1}H/^{13}C-HSQC spectra permitted to identify both ^{1}H and ^{13}C spin systems of all glycosyl residues. The homonuclear vicinal coupling constant ($^3J_{H-H}$) observed by COSY experiments established the configuration of the monosaccharides. Finally, the heteronuclear scalar coupling correlation ($^3J_{H-C}$) observed in the ^{1}H/^{13}C-HMBC spectrum established the glycosidic sequence of all LOSs. In particular, the neutral OS-IV glycosidic sequence was identified as a α-4-amino-4,6-dideoxy-Gal*p*-(1→c)-α-Car-(1→c)-α-Car-(1→4)-*β*-Xyl*p*-(1→4)-α-3-O-Me-Rha*p*-(1→3)-*β*-Glc*p*-(1→3)-*β*-Glc*p*-(1→4)-α-Glc*p*-(1→1)-α-Glc*p* (Table S2 of the Supporting Information). Comparison of ^{1}H and ^{13}C NMR parameters of neutral and acidic OS-IV showed that they exhibited very similar chemical shifts values and spin systems, which established that both fractions contained identical oligosaccharidic cores, characterized by the presence of nine monosaccharides (Table S3 of the Supporting Information). Furthermore, distinctive chemical shifts of IX-H4/C4 at δ 4.38/55.6 and 4.40/55.6 for neutral LOS-IV and δ 4.28/56.3 for acidic LOS-IV demonstrated that amino-dideoxy-Gal residues (IX) were further substituted in their C-4

(34) Wishart, D. S.; Bigam, C. G.; Yao, J.; Abildgaard, F.; Dyson, H. J.; Oldfield, E.; Markley, J. L.; Sykes, B. D. *J. Biomol. NMR* **1995**, *6*, 135–40.
(35) Frisch, M. J.; et al. *GAUSSIAN 03*, Revision A; Gaussian, Inc.: Wallingford, CT, 2003.
(36) Cammi, R.; Mennucci, B.; Tomasi, J. *J. Phys. Chem. A* **2000**, *104*, 5631–5637.
(37) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.

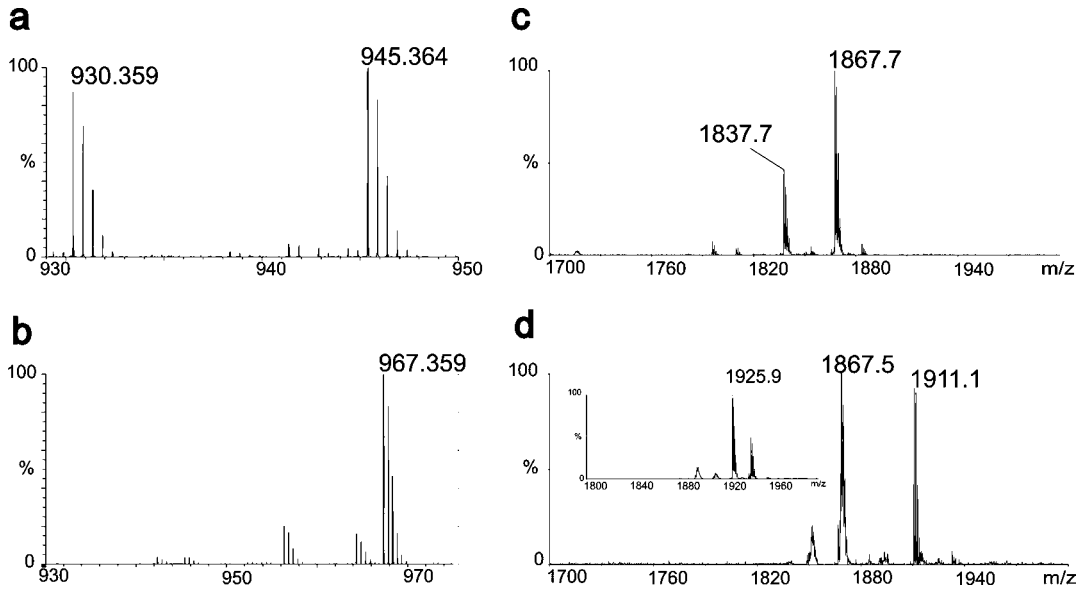

Figure 2. Mass spectrometry analysis of neutral and acidic OS-IV. Left panel: NanoESI-Qh-FT-ICR analysis of (a) neutral and (b) acidic OS-IV generated by *O*-deacylation of LOS-IV fractions. The doubly charged ions $[M + 2Na]^{2+}$ at $m_{exp} = 945.364$ ($m_{th} = 945.366$, $\Delta m = 2{,}11$ ppm) and $m_{exp} = 930.361$ ($m_{th} = 930.359$, $\Delta m = 2{,}15$ ppm) established that neutral OS-IV had $C_{74}O_{50}H_{128}N_2$ and $C_{73}O_{49}H_{126}N_2$ compositions.The doubly charged ion at $m_{exp} = 967.359$ ($m_{th} = 967.361$, $\Delta m = 2{,}06$ ppm) demonstrated that neutral OS-IV had $C_{75}O_{52}H_{128}N_2$ composition. Right panel, MALDI-MS analysis of $[M + Na]^+$ adducts of (c) neutral OS-IV and (d) acidic OS-IV showing in the inset, carboxymethylated acidic OS-IV (*m*/*z* 1925.9).

positions via an amide bond by yet unidentified *N*-acyl groups (Z). Considering that both LOS-IV fractions exhibited different chromatographic behaviors on an anion-exchange column, we postulated that neutral OS-IV differed from acidic OS-IV solely by the nature of their respective *N*-acyl substituents Z.

The exact atomic compositions of the two OS-IV fractions were then established by high resolution mass spectrometry using a nanoelectrospray-fourier transform ion cyclotron resonance mass spectrometer (nanoESI-FTICR MS). NanoESI-Qh-FT-ICR MS analysis of the neutral OS-IV fraction generated two doubly charged signals at *m*/*z* 945.364 $[M + 2Na]^{2+}$ and *m*/*z* 930.359 $[M + 2Na]^{2+}$ (Figure 2a), thus supporting the existence of two different compounds OS-IVa and OS-IVb with calculated exact compositions of $C_{74}O_{50}H_{128}N_2$ and $C_{73}O_{49}H_{126}N_2$, respectively. Consistently, MALDI-MS analysis of the same preparation generated two distinct signals at *m*/*z* 1867.7 $[M + Na]^+$ and 1837.7 $[M + Na]^+$, confirming the presence of two distinct molecules (Figure 2c). The fact that a single oligosaccharide core ($C_{66}O_{46}H_{116}N_1$) was observed by NMR in the neutral OS-IV fraction strongly suggested that OS-IVa and OS-IVb only differed in their respective *N*-acyl substituents Za and Zb. On this basis, the compositions of the Za and Zb were established as $C_8O_4H_{12}N_1$ and $C_7O_3H_{10}N_1$, respectively. In comparison, nanoESI-Qh-FT-ICR MS analysis of acidic OS-IV showed a single doubly charged signal at *m*/*z* 967.359 $[M + 2Na]^{2+}$, which permitted to establish OS-IVc molecular composition as $C_{75}O_{52}H_{128}N_2$ (Figure 2b). Based on the known core structure, the molecular composition of Zc was calculated as $C_9O_6H_{12}N_1$. Surprisingly, MALDI-MS analysis of native acidic OS-IV generated an additional $[M + Na]^+$ signal at *m*/*z* 1867.5 suggesting a molecular mass of 1844 that differed by 44 amu from the intact molecule observed as a $[M + Na]^+$ adduct at *m*/*z* 1911.1 (Figure 2d). This apparent loss of 44 amu was tentatively attributed to a partial decarboxylation of the *N*-acyl group Zc upon laser desorption. Stabilization of the carboxyl group of Zc by methyl esterification prior to MALDI-MS analysis allowed to generate a single $[M + Na]^+$ signal at *m*/*z* 1925.9, consistent with the presence of a −COOH group on the native molecule (Figure 2d).

Altogether, comparison of both neutral and acidic LOS-IV fractions conclusively revealed the presence of three structurally related molecules sharing identical oligosaccharidic sequences but substituted in the C4 position at the terminal nonreducing amino-dideoxy-Gal residue (IX) by three distinct *N*-acylated groups Za, Zb, and Zc. Retrospectively, the presence of two different substituents Za and Zb in neutral LOS-IV fraction is in agreement with the observation on NMR spectra of a slight heterogeneity of IX H2/C2 and H4/C4 NMR parameters (Table S2 of the Supporting Information). This prompted us to (i) inquire about the structure of neutral LOS-IV Za and Zb and (ii) establish the structure of the more complex Zc substituent characterizing acidic LOS-IV.

NMR Identification of the *N*-Acyl Groups Za and Zb. Structure of Za and Zb were established by NMR from the neutral OS-IV fraction. Their respective NMR signals were easily distinguished from each other owing to their different quantities, Za accounting for about 80% of total molecules. Based on $^1H-^{13}C$ HSQC (Figure 3) and HMBC (Figure S2 of the Supporting Information) NMR experiments, Za atoms were assigned to the following: $-CH_3$ (δ 1.42/21.6), $N-CH_3$ (δ 2.79/30.0), $O-CH_3$ (δ 3.61/60.8), CH−N (δ 4.14/72.4), CH−O (δ 4.30/84.7), C═O (δ 173.8) and C−OH (δ 77.7). The sequence assignment $-CH(OCH_3)-C(CH_3)(OH)-CHR-N(CH3)-CO-$ was deduced from the numerous heteronuclear multiple bond correlations ($^xJ_{H-C}$) observed on the HMBC spectrum, as depicted on Figure S2 of the Supporting Information. In particular, the observation of $^3J_{H-C}$ correlations between $N-CH_3$ (δ 2.79) and C═O (δ 173.8)/CH−R (δ 72.4) clearly established the sequence $CH-N(CH_3)-CO$. Then, $^3J_{H-C}$ correlation between $CH(OCH_3)$ (δ 4.30) and C═O (δ 173.8) demonstrated that Za exhibits a cyclic structure. Accordingly, on the COSY spectrum, the total absence of $^3J_{H-H}$ correlation originating from CH_3 (δ 1.42) confirmed that this group is carried by the oxygenated quaternary carbon (C−OH) at 77.7 ppm. Finally,

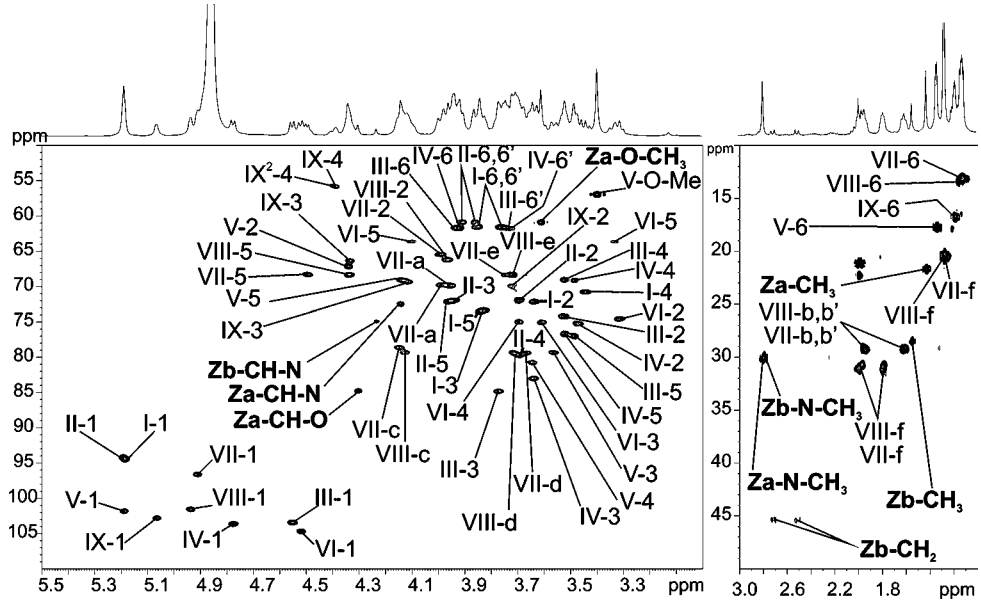

Figure 3. $^{1}H-^{13}C$ HSQC NMR analysis of neutral OS-IV fraction. Details of the anomer and bulk regions (left panel) and the methyl and methylene resonances containing regions (right panel) of the $^{1}H-^{13}C$ HSQC NMR spectrum of the neutral OS-IV. I to IX correspond to the nine monosaccharides of the neutral OS-IV. The two aglycon groups substituting the monosaccharide IX are named Za and Zb. ^{1}H and ^{13}C chemical shift values are indicated in Table S2 of the Supporting Information.

based on this sequence, the linkage position of the *N*-acyl group on Za was assigned to the remaining cycle carbon CH−N at δ 72.4. It was however not possible to unambiguously observe on the HMBC spectrum a clear correlation between IX-H4 and the remaining C═O group associated to the *N*-acyl group linking Za to the α-4-amino-4,6-dideoxy-Gal*p* residue although this correlation could be observed in the major compound derived from the acidic LOS-IV fraction. Surprisingly, the only correlation observed on the COSY experiment is a rare $^{5}J_{H-H}$ long distance correlation between the protons of $N-CH_3$ and the proton of CH(OMe). The veracity of such unusually long distance correlation was however confirmed by performing the NMR analysis of a product comparable to Za, the *N*-methyl-2-pyrrolidone (NMP) (Figure S3 of the Supporting Information). Altogether, NMR and MS analyses permitted the identification of the *N*-acyl substituent Za as a 3-hydroxy-4-methoxy-1,3-dimethyl-5-oxopyrrolidine-2-carboxylic acid (also named *N*-methyl-3-hydroxy-3-methyl-4-methoxy-5-oxoproline). Consequently, a α-4-(2-carbamoyl-3-hydroxy-4-methoxy-1,3-dimethyl-5-oxopyrrolidine)-4,6-dideoxy-galactopyranose sequence could be inferred as the terminal monosaccharide of OS-IVa.

The relative configurations of the asymmetric carbons from *N*-methyl-3-hydroxy-3-methyl-4-methoxy-5-oxoproline Za were established by the observation of the nuclear Overhauser effect (NOE) on the $^{1}H-^{1}H$ NOESY NMR spectrum. On this basis, the three asymmetric carbons labeled 2*, 3*, and 4* could be assigned as (2*S*, 3*S*, 4*R*) (Figure 4a). Starting from 2*, the small NOE connectivities from protons H-2 and H-6 of the α-4-amino-4,6-dideoxy-galactopyranose residue to the protons CH_3-N and CH−O groups of Za, respectively, supported the conformation (2*S*). Then, observation of an intense NOE correlation between CH_3 (δ 1.42) and CH−N (δ 4.14) indicated that the carbon 3* had also a (*S*) configuration. Finally, correlation between $-CH_3$ and $O-CH_3$ (δ 3.61) established that 4* has a (*R*) configuration. Lack of NOE correlation between CH−N and CH−O also strongly supported the (4*R*) configuration. Altogether, NMR data established that 3-hydroxy-4-methoxy-1,3-dimethyl-5-oxopyrrolidine-2-carboxylic is in the (2*S*, 3*S*, 4*R*) configuration (Figure 4a).

In addition to signals assigned to Za, a set of minor signals associated to another *N*-acylated substituent Zb was observed on $^{1}H-^{13}C$ HSQC (Figure 3) and HMBC (Figure S2 of the Supporting Information) NMR spectra. According to their $^{1}H/^{13}C$ chemical shifts, these signals were assigned to $-CH_3$ (δ 1.54/28.4), $N-CH_3$ (δ 2.79/29.9), $-CH_2-$ (δ 2.71, 2.50/45.4) CH−N group (δ 4.23/74.8), C═O (δ 177.3) and an hydroxylated quaternary carbon (δ 73.1). Similarly to Za, $^{1}H-^{13}C$ HMBC-NMR correlations established that Zb was a 3-hydroxy-1,3-dimethyl-5-oxopyrrolidine-2-carboxylic acid (also named *N*-methyl-3-hydroxy-3-methyl-5-oxoproline), only differing to Za by the absence of a $O-CH_3$ group. Then, intense $^{1}H/^{1}H$ NOE correlations between $-CH_3$ and CH−N, established that Zb exhibited either (*R*, *R*) or (*S*, *S*) configurations (NOE) (Figure 4b). However, lack of obvious NOE contacts between the Zb and 4,6-dideoxy galactopyranose, as observed for Za, prevented to unambigously distinguish the two configurations. Altogether, NMR experiments allowed the establishment of the nature of the terminal monosaccharide of the minor OS-IVb as an α-4-(2-carbamoyl-3-hydroxy-1,3-dimethyl-5-oxopyrrolidine)-4,6-dideoxy-galactopyranose (Figure 4b).

NMR Identification of the *N*-Acyl Group Zc. Mass spectrometry analysis of acidic OS-IV fraction established the presence of a single Zc ($C_9O_6H_{12}N_1$) substituting the 4,6-dideoxy-galactopyranose residue. However, subsequent NMR analyses demonstrated that this compound occurred as two distinct diastereoisomers (2*S*, 3*S*, 4*R*) and (2*R*, 3*S*, 4*R*) that exhibited slightly different ^{1}H and ^{13}C chemical shifts (Table S3 of the Supporting Information). As for Za and Zb, $^{1}H-^{13}C$ HSQC (Figure S4 of the Supporting Information) and HMBC (Figure S5 of the Supporting Information) experiments permitted the identification of CH_3 (δ 1.32/19.6 and 1.63/23.5), $N-CH_3$ (δ 2.80/31.0 and 2.81/31.0), $O-CH_3$ (δ 3.60/61.1 and 3.65/61.9), CH−O (δ 4.20/85.3 and 3.96/84.2), two C═O (δ 170.3−175.4

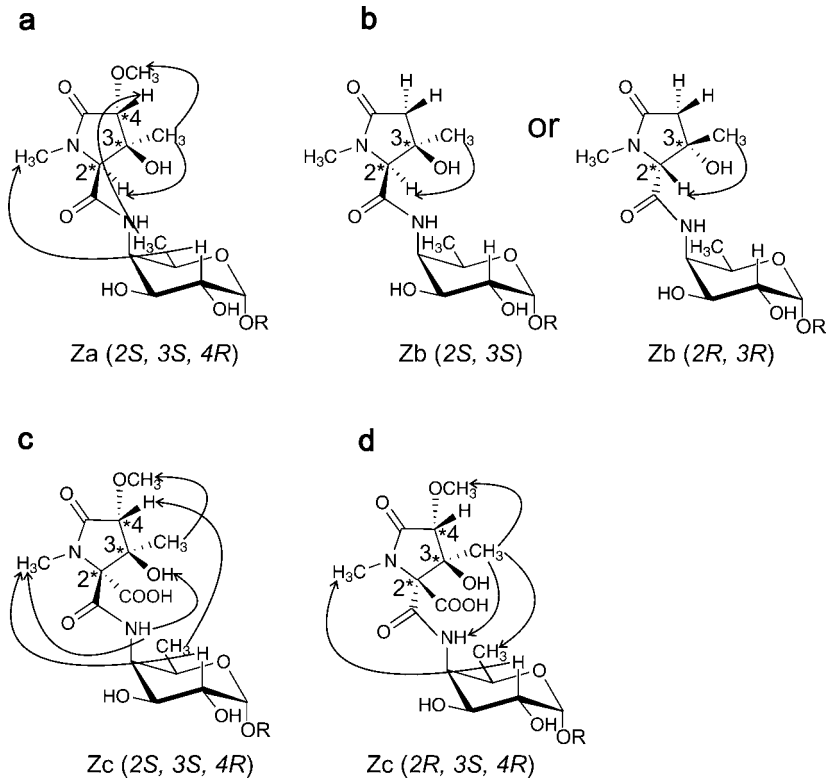

Figure 4. Structures and conformations of aglycon structure Za and Zb or Zc substituting the terminal 4-amino-4,6-dideoxyGal*p* residue of neutral or acidic LOS-IV, respectively. (a) (2*S*, 3*S*, 4*R*) α-4-(2-carbamoyl-3-hydroxy-4-methoxy-1,3-dimethyl-5-oxopyrrolidine)-4,6-dideoxy-Gal*p*, (b) α-4-(2-carbamoyl-3-hydroxy-1,3-dimethyl-5-oxopyrrolidine)-4,6-dideoxy-Gal*p*, (c) α-4-((2*S*, 3*S*, 4*R*) 2-carbamoyl-3-hydroxy-4-methoxy-1,3-dimethyl-5-oxopyrrolidine-2-carboxylic acid)-4,6-dideoxy-Gal*p*, and (d) α-4-((2*R*, 3*S*, 4*R*) 2-carbamoyl-3-hydroxy-4-methoxy-1,3-dimethyl-5-oxopyrrolidine-2-carboxylic acid)-4,6-dideoxy-Gal*p*. Configurations of asymmetric carbons of aglycon groups were deduced from nuclear Overhauser effect (full arrow) NMR experiments.

and 170.6−176.4), as well as two distinct quaternary carbons (δ 80.5−81.9 and 77.3−83.4). In agreement with the atomic composition of Zc, an additional −COOH group was identified in both diastereoisomers at 173.2 and 172.3 ppm, from the ^{13}C NMR spectrum (Figure S6 of the Supporting Information). Finally, NMR analysis of OS-IV in H_2O permitted the identification of protons of two CO−NH groups at 10.23 and 10.27 ppm in agreement with the presence of two diastereoisomers substituting the terminal monosaccharide (Table S3 of the Supporting Information). The surprisingly strong upfield resonance of these protons compared to a typical CO−NH group was tentatively attributed to the strong effect of the neighboring the −COOH group.

As for Za and Zb, multiple $^{x}J_{H-C}$ observed on the HMBC spectrum (Figure S5 of the Supporting Information) permitted the unambiguous establishment of the structure of Zc as a cyclic compound −CH(OCH_3)−C(CH_3)(OH)−C−N(CH_3)−CO− presenting an overall identical sequence as Za. However, Zc differed from Za by the replacement of the −CH−N− group by a quaternary carbon further substituted by the additional carboxyl group, as demonstrated by the large downfield chemical shift of this carbon from δ 72.4 to δ 81.9 and 83.4 in both diastereoisomers. Failure to observe any $^{x}J_{H-C}$ correlations in the HMBC experiment due to absence of neighboring protons strongly supported the proposed location of the carboxyl group. Finally, the $^{3}J_{H-C}$ correlation from the H4 of the α-4-amino-4,6-dideoxy-Gal*p* to carbonyl groups at 170.3 ppm unambiguously demonstrated that −NH−CO was linking IX-C4 to the quaternary carbon C−N of Zc. Thus, these experiments firmly established the structure of Zc as 3-hydroxy-4-methoxy-1,3-dimethyl-5-oxopyrrolidine-2,2-dicarboxylic acid. The nature of Zc and its linkage to LOS-IV were also confirmed by studying the structures of intact isolated *N*-acylated α-4-amino-4,6-dideoxy-Gal*p* and de-*N*-acylated acidic OS-IV moiety (Figure S7 and S8 of the Supporting Information).

Analysis of intraresidual NOE correlations on the ROESY spectrum permitted the identification of two diastereoisomers of Zc, (2*S*, 3*S*, 4*R*) and (2*R*, 3*S*, 4*R*), easily distinguished owing to different signal integrations (70% for the former and 30% for the last). NOE correlation between $-CH_3$ and $O-CH_3$ demonstrated that carbons 3 and 4 were always in opposite configurations, as established for Za. However, the extra-residual NOE connectivities from IX-H2 to $N-CH_3$ and from IX-H6 to $CH-OCH_3$ established that the major form of Zc exhibited a (2*S*, 3*S*, 4*R*) configuration (Figure 4c). This was further confirmed by NOE correlations from CO−NH (δ 10.23) to −OH (δ 6.70) and to $N-CH_3$ (δ 2.80). The minor form of Zc was characterized by a very different pattern of extra-residual correlations including strong NOE correlations from $Zc-CH_3$ to CO−NH and IX-H6 that permitted to characterize its configuration as (2*R*, 3*S*, 4*R*) (Figure 4d).

Taken collectively, these results indicate that the terminal monosaccharide of acidic OS-IV consisted of a mixture of two diastereoisomers α-4-((2*S*, 3*S*, 4*R*) 2-carbamoyl-3-hydroxy-4-methoxy-1,3-dimethyl-5-oxopyrrolidine-2-carboxylic acid)-4,6-dideoxy-galactopyranose and α-4-((2*R*, 3*S*, 4*R*) 2-carbamoyl-3-hydroxy-4-methoxy-1,3-dimethyl-5-oxopyrrolidine-2-carboxylic acid)-4,6-dideoxy-galactopyranose.

Molecular Modeling of *N*-Acylated Dideoxy-Galactose. Based on the deduced configuration of individual carbons of Zc, we have modeled the spatial conformation of the very unusual terminal α-4-(2-carbamoyl-3-hydroxy-4-methoxy-1,3-dimethyl-

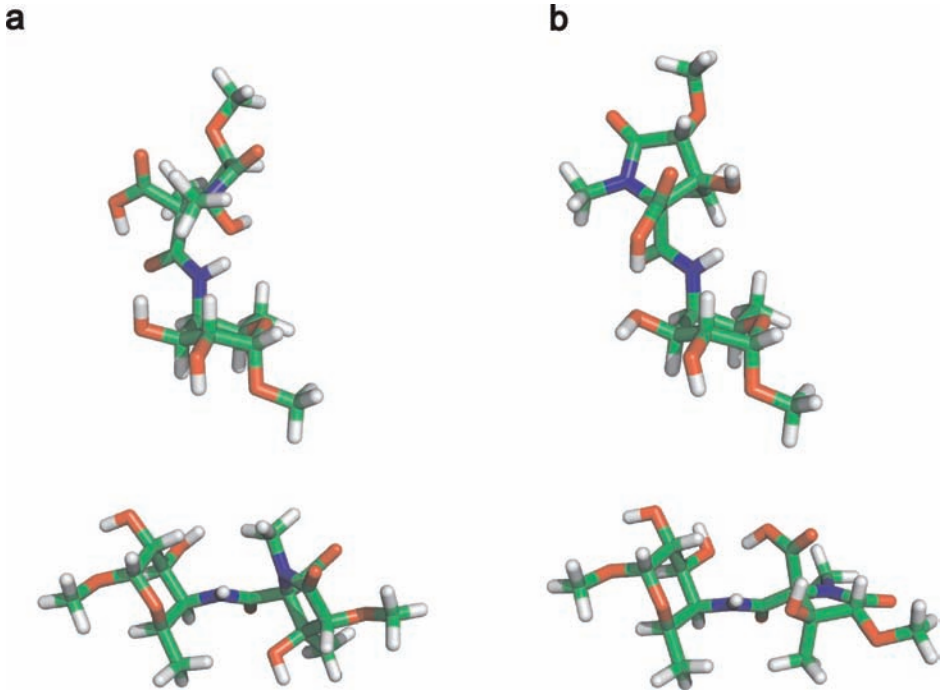

Figure 5. Spatial conformations of the two diastereoisomers of α-4-(2-carbamoyl-3-hydroxy-4-methoxy-1,3-dimethyl-5-oxopyrrolidine-2-carboxylic acid)-4,6-dideoxy-Gal*p* from major acidic LOS-IV. Lowest B3LYP/6-31 g(d) IEF-PCM energy conformations of (a) (2*S*, 3*S*, 4*R*) and (b) (2*R*, 3*S*, 4*R*) diastereoisomers. Atoms are colored green (carbon), white (hydrogen), red (oxygen), and blue (nitrogen).

5-oxopyrrolidine-2-carboxylic acid)-4,6-dideoxy-galactopyranose residue of LOS-IV. This allowed us to compare the electronic and geometrical details of the two Zc (2*S*, 3*S*, 4*R*) and (2*R*, 3*S*, 4*R*) diastereoisomers substituted monosaccharides (Figure 5). Under the investigated experimental conditions (pH = 6.51), compound Zc is mainly protonated, so only the carboxylic form was considered in the calculations. Proposed structural models are in agreement with the experimental data based on the NOE connectivity patterns, thus allowing direct structure comparisons by spectroscopic method and three-dimensional reconstruction. These results indicated that compound (2*S*, 3*S*, 4*R*) Zc can thus adopt a folded conformation whereas compound (2*R*, 3*S*, 4*R*) Zc can adopt an extended conformation (Figure S9 of the Supporting Information).

Induction of Macrophage Cell Surface Antigens. Previous studies have reported that microbial infection modulates the macrophage cell activation process by interfering with the expression of cell surface antigens.[7,24] In the case of mycobacterial infections, cell surface antigens such as the adhesion molecule ICAM-1 and the CD40 activation marker are necessary for granuloma formation and bacterial containment burden.[29,32,38−40] The ability of purified LOSs to trigger the expression of these molecules in human differentiated THP-1 macrophages was therefore investigated. To this aim, LOS-I, LOS-II, and LOS-IV were used to evaluate the influence of the terminal α-3-O-Met-Rha, α-Car, and *N*-acylated α-4-amino-4,6-dideoxy-Gal*p* residues on the activity of the intact and de-*O*-acylated glycolipids.

Flow cytometry analysis using either anti-ICAM-1 or anti-CD40 antibodies revealed that these cell surface antigens were expressed at low levels on the surface of unstimulated THP-1 cells (Figure 6). In contrast, LOS-IV, along with LPS, a known ICAM-1 stimulator, strongly induced ICAM-1 expression after 20 h incubation (Figure 6a). Considering the high prevalence of acidic LOS-IV in the mycobacteria cell wall compared to neutral LOS-IV (20 times more), only acidic LOS-IV was used as stimulating factors. Similarly, stimulation of the cells with acidic LOS-IV also caused CD40 expression (Figure 6b). Interestingly, in contrast to LOS-IV, LOS-I and LOS-II failed to induce ICAM-1 or CD40 significant expression on the macrophage cell surface. Because these glycolipids only differ from LOS-IV molecules by the nonreducing terminus of their glycan moieties, the loss of activity of LOS-I and LOS-II can be associated with the absence of the unusual terminal nonreducing monosaccharide. In order to define the relative contribution of the lipidic or carbohydrate domains of LOS-IV in the cell surface antigen-inducing activity, the ability of LOS-IV devoid of fatty acids to modulate ICAM-1 and CD40 expression was assessed. As shown in Figure 6c and d, the oligosaccharide moiety of LOS-IV alone failed to stimulate ICAM-1 and CD40 expressions. Together, these results support the view that both the lipidic aglycone anchor and the specific carbohydrate domains of LOS-IV are necessary to trigger ICAM-1 and CD-40 inductions.

Discussion

The mycobacterial cell envelope contains a vast array of glycolipids characterized by an extraordinary species-specific structural diversity. The structure of individual glycolipids varies from one species to the other in terms of sequence, monosaccharide composition, and/or acylation status. In particular, the structures of many complex and rare, sometimes unique, monosaccharides found in mycobacterial glycolipids have been

(38) Johnson, C. M.; Cooper, A. M.; Frank, A. A.; Orme, I. M. *Infect. Immun.* **1998**, *66*, 1666–70.
(39) Sullivan, L.; Sano, S.; Pirmez, C.; Salgame, P.; Mueller, C.; Hofman, F.; Uyemura, K.; Rea, T. H.; Bloom, B. R.; Modlin, R. L. *Infect. Immun.* **1991**, *59*, 4154–60.
(40) Okamoto Yoshida, Y.; Umemura, M.; Yahagi, A.; O'Brien, R. L.; Ikuta, K.; Kishihara, K.; Hara, H.; Nakae, S.; Iwakura, Y.; Matsuzaki, G. *J. Immunol.* **2010**, *184*, 4414–22.

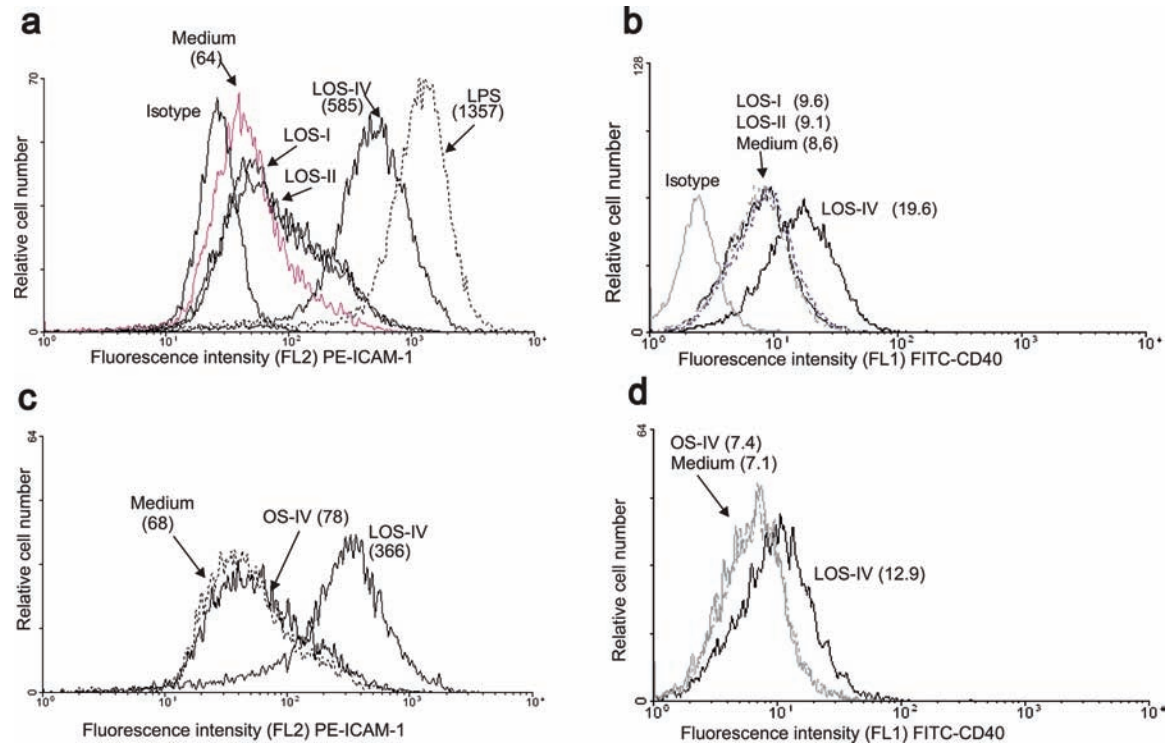

Figure 6. Modulation of cell surface antigen expression on macrophages by LOSs and oligosaccharide generated from Los-IV. Differentiated THP-1 cells were incubated with 20 μg/mL of various LOSs (LOS-I, LOS-II, LOS-IV), 50 ng/mL of LPS (a, b) or 20 μg/mL of oligosaccharide generated from deacylated LOS-IV (c, d). After 20 h incubation, analysis of ICAM-1 (CD54) and CD40 expression on unstimulated (medium) and activated cells was performed by flow cytometry, using PE-conjugated anti-CD54 or FITC-conjugated anti-CD40, respectively. Cells were exposed also with irrelevant antibodies (PE- or FITC-conjugated mouse isotype controls). Results are shown as linear-log scale fluorescence histograms. The mean fluorescence intensity is shown in parentheses. The data are representative of three independent experiments.

determined, for instance, polymethylated and acetylated monosaccharides in *M. kansasii* phenolic glycolipids,[41,42] branched monosaccharides in *M. gastri*[43] and *M. marinum* LOSs,[18] as well as *N*-acylated monosaccharides in *M. kansasii* LOSs[44] and *M. intracellulare* glycopeptidolipids.[45,46] Although cell wall associated glycolipids exhibit immunomodulatory properties and participate in the pathogenic processes by manipulating host−pathogen interactions,[3,7] only little information is available regarding the biological functions of individual molecules and the relevance of their structural diversity. Herein, we have isolated and characterized three members of a new family of dicyclic *N*-acylated monosaccharides. They all share a 4,6-dideoxy-Gal*p* substituted by a 3-hydroxy-3-methylated-pyrrolidone cycle. Although absolute configuration of this monosaccharide was not definitively established by total synthesis, it was tentatively assigned as D-configuration based on comparison with other bacterial strains and biosynthetic data. Indeed, every 4-acetamido-4,6-dideoxy-Gal*p* residue so far observed in bacterial polysaccharides were shown to exhibit a D-configuration, irrespective of the bacterial species they originate from.[47−49] Exhaustive analysis of the nucleotide sugar pathways involved in the biosynthesis of enterobacterial common antigens (ECA) further confirmed that dTDP-4-acetamido-4,6-dideoxy-D-Gal*p* originated from dTDP-D-Glc through a dTDP-6-deoxy-D-xylo-4-hexulose.[50] A similar pathway seems to be conserved in *M. marinum* (Figure 7).

Although dicyclic *N*-acylated monosaccharides isolated from LOSs share a common structural core, they differ by the substitutions of their respective cycle by either methoxy and/or carboxy groups. This heterogeneity generated two neutral and one acidic glycolipids referred to as LOS-IVa, LOS-IVb, and LOS-IVc, respectively. The complete form of LOS-IVc substituted by both carboxy and methoxy groups, presumably corresponding to the end-product of the LOS biosynthetic pathway, represents around 95% of the total LOS-IV content. Interestingly, the molecular constraints that originate from the additional pyrrolidone cycle stabilized the glycolipids into distinct conformations, providing an unique opportunity to establish spatial conformations of natural carbohydrate structures in solution both by analysis the NOE connectivity pattern and by modeling their minimum energy conformers. This demonstrated that, whereas the (2*S*, 3*S*, 4*R*) diastereoisomer presented a folded structure with the two cycles facing each others, the

(41) Gilleron, M.; Venisse, A.; Riviere, M.; Servin, P.; Puzo, G. *Eur. J. Biochem.* **1990**, *193*, 449–57.
(42) Watanabe, M.; Aoyagi, Y.; Ohta, A.; Minnikin, D. E. *Eur. J. Biochem.* **1997**, *248*, 93–8.
(43) Gilleron, M.; Vercauteren, J.; Puzo, G. *Biochemistry* **1994**, *33*, 1930–7.
(44) Hunter, S. W.; Fujiwara, T.; Murphy, R. C.; Brennan, P. J. *J. Biol. Chem.* **1984**, *259*, 9729–34.
(45) Fujiwara, N.; Nakata, N.; Maeda, S.; Naka, T.; Doe, M.; Yano, I.; Kobayashi, K. *J. Bacteriol.* **2007**, *189*, 1099–108.
(46) Fujiwara, N.; Nakata, N.; Naka, T.; Yano, I.; Doe, M.; Chatterjee, D.; McNeil, M.; Brennan, P. J.; Kobayashi, K.; Makino, M.; Matsumoto, S.; Ogura, H.; Maeda, S. *J. Bacteriol.* **2008**, *190*, 3613–21.
(47) Lugowski, C.; Kulakowska, M.; Romanowska, E. *Infect. Immun.* **1983**, *42*, 1086–91.
(48) Vinogradov, E. V.; Knirel, Y. A.; Thomas-Oates, J. E.; Shashkov, A. S.; L'Vov V, L. *Carbohydr. Res.* **1994**, *258*, 223–32.
(49) Kilcoyne, M.; Shashkov, A. S.; Perepelov, A. V.; Nazarenko, E. L.; Gorshkova, R. P.; Ivanova, E. P.; Widmalm, G.; Savage, A. V. *Carbohydr. Res.* **2005**, *340*, 1557–61.
(50) Samuel, G.; Reeves, P. *Carbohydr. Res.* **2003**, *338*, 2503–19.

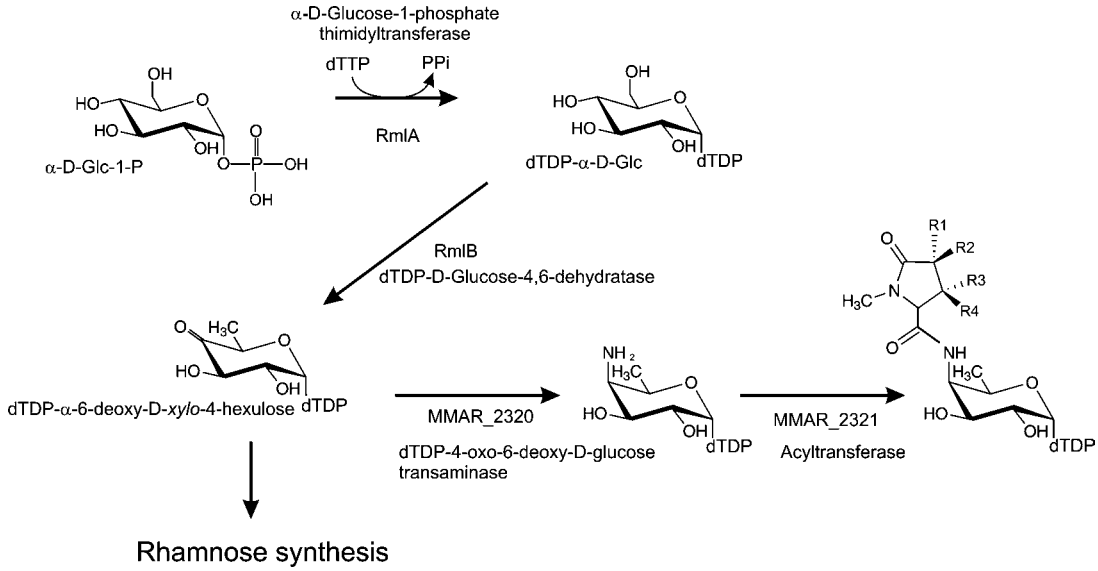

Figure 7. Proposed partial biosynthetic pathway of the *N*-acylated 4-amino-4,6-dideoxy-Gal*p* identified in LOS-IV of *M. marinum*.

(2*R*, 3*S*, 4*R*) diastereoisomer exhibits an extended shape. The almost orthogonal positioning of the pyrrolidone cycle compared with the hexopyranoside cycle observed in last isomer is greatly stabilized by the COO−H···OH−C3′ H bond between the two cycles.

Importantly, structural delineation of LOSs offers the possibility to assign potential biological functions to various enzymes encoded within the LOS biosynthetic gene cluster. Several genes involved in the LOS-II biosynthesis pathway have been previously described.[21] In particular, *MMAR_2309* was found to encode a UDP-glucose dehydrogenase (UdgL) involved in the first step of D-xylose synthesis, while at least three other genes (*MMAR_2310*, *MMAR_2311*, and *MMAR_2312*), cotranscribed with *MMAR_2309*, were proposed to be involved in the last step of D-xylose synthesis and transfer to LOS-I. Transposon insertion within *MMAR_2332* gene resulted in the accumulation of LOS-II*, a LOS-II biosynthetic precursor possessing the β-xylose but lacking the terminal caryophyllose residue capping LOS-II.[21] Moreover, although a gene coding for a potential glycosyltransferase implicated in the transfer of the LOS-IV terminal monosaccharide was identified as LosA (*MMAR_2313*),[20] lack of structural information prevented further identification of biosynthetic genes. Despite the fact that the proposed first steps of biosynthesis of 4-amino-4,6-dideoxy-galactose residue are likely to be similar to those of the rhamnose biosynthetic pathway, requiring an α-D-glucose-1-phosphate thimidylyltransferase and a dTDP-D-Glc 4,6-dehydratase to catalyze the formation of dTDP-4-keto-6-deoxy-D-Glc (Figure 7),[51,52] no genes encoding these enzymes were found within the LOS biosynthetic gene cluster. However, both an α-D-Glc-1-phosphate thimidylyltransferase and a dTDP-D-Glu-4,6-dehydratase have previously been identified in mycobacteria, as RmlA and RmlB participating in the biogenesis of the dissacharide linker unit that connects together arabinogalactan and peptidoglycan.[53] It is therefore tempting to speculate that *rmlA* and *rmlB* may also participate to the early steps of 4,6-dideoxy-Gal synthesis. In the next step, dTDP-4-keto-6-deoxy-D-Glc is converted into dTDP-4-amino-4,6-dideoxy-D-Gal by a TDP-4-oxo-6-deoxy-D-Glc transaminase enzyme,[54] an enzyme predicted to be encoded by *MMAR_2320* present in the LOS cluster. This is supported by the high degree of conservation between MMAR_2320 and the TDP-4-oxo-6-deoxy-D-Glc transaminase WecE from *Escherichia coli* (Table S4 of the Supporting Information).[54]

With respect to the biosynthesis/transfer of the aglycone moiety, we identified a putative acyltransferase encoded by *MMAR_2321* that is probably cotranscribed with *MMAR_2320*. Thus, MMAR_2321 may represent a potent candidate for the transfer of the pyrrolidone cycle onto the dideoxy-sugar through an amide bond. Such an acyl-transferase involved into the transfer of *N*-acylated aglycon moiety of glycopeptidolipids have been previously identified in *M. avium/M. intracellulare*, although no significant homology was found with *MMAR_2321*.[45,46] BLAST analysis revealed that MMAR_2321 shares 44% identity with a putative transferase (GenBank: GU576498.1) from the O-antigen biosynthesis gene cluster of *Vibrio cholerae* O:5, that catalyzes the transfer of a pyrrolidone derivative onto the polysaccharide moiety of LPS.[55] This cross-species conservation strongly suggests that this *N*-acyltransferases family shows a high degree of specificity for pyroglutamate substituents. Overall, based on the studies conducted by McNeil and collaborators,[56,57] we propose that the early biosynthesis steps of LOS-IV terminal *N*-acylated monosaccharide are common to those of rhamnose biosynthesis in mycobacteria whereas the specific enzymes required for its full synthesis are coded by the genes contained within the LOS gene cluster (Figure 7).

A 4-amino-4,6-dideoxy-Gal substituted by a yet undescribed *N*-acylated group was previously also observed in terminal

(51) Trefzer, A.; Salas, J. A.; Bechthold, A. *Nat. Prod. Rep.* **1999**, *16*, 283–99.
(52) Erbel, P. J.; Barr, K.; Gao, N.; Gerwig, G. J.; Rick, P. D.; Gardner, K. H. *J. Bacteriol.* **2003**, *185*, 1995–2004.
(53) Li, W.; Xin, Y.; McNeil, M. R.; Ma, Y. *Biochem. Biophys. Res. Commun.* **2006**, *342*, 170–8.
(54) Hwang, B. Y.; Lee, H. J.; Yang, Y. H.; Joo, H. S.; Kim, B. G. *Chem. Biol.* **2004**, *11*, 915–25.
(55) Hermansson, K.; Jansson, P. E.; Holme, T.; Gustavsson, B. *Carbohydr. Res.* **1993**, *248*, 199–211.
(56) Ma, Y.; Mills, J. A.; Belisle, J. T.; Vissa, V.; Howell, M.; Bowlin, K.; Scherman, M. S.; McNeil, M. *Microbiology* **1997**, *143*, 937–45.
(57) Stern, R. J.; Lee, T. Y.; Lee, T. J.; Yan, W.; Scherman, M. S.; Vissa, V. D.; Kim, S. K.; Wanner, B. L.; McNeil, M. R. *Microbiology* **1999**, *145*, 663–71.

position of LOS-II from the strain Canetti of *M. tuberculosis*.[17] Interestingly, the unknown aglycon structure of *N*-acylated 4-amino-4,6-dideoxy-Gal terminal monosaccharide of LOS-II from the Canetti strain possesses the same mass as Za from neutral LOS-IV of *M. marinum*,[17] suggesting the presence of a similar/conserved structure in this species. Previously, the gene *Rv1500* of *M. tuberculosis* H37Rv (conserved in the Canetti strain) was identified as an homologue of MMAR_2313/*LosA* (Table S4 of the Supporting Information).[20] BLAST analysis also revealed that the putative MMAR_2321/acyltransferase displayed 90% of identity with the *Rv1505c*-encoded protein of *M. tuberculosis* H37Rv, also found in the Canetti strain. In contrast, although *MMAR_2320*/aminotransferase appears fully conserved in the Canetti genome, its orthologue in *M. tuberculosis* H37Rv is separated in two genes (*Rv1503c* and *Rv1504c*) by a stop codon. Whether this interruption leads to unproductive enzymes unable to synthesize LOS-IV remains however to be addressed. Altogether, structural and genomic data suggest that the aglycone group terminally substituting the LOS-II from the Canetti strain is also a pyroglutamate derivative, as identified in *M. marinum*.

In our previous study, we demonstrated that LOSs purified from *M. marinum* not only failed to induce significant TNF-α secretion in human macrophage-like differentiated THP-1 cells but also inhibited its secretion.[18] The molecular mechanisms by which LOS-IV exerts this inhibitory effect are not clearly understood. Since it is well documented that IL-10 negatively regulates TNF-α synthesis, we investigated the production of IL-10 in LOS-IV stimulated THP-1 cells. However, no IL-10 secretion was detected in response to LOS-IV (data not shown). Recent studies demonstrated that lack or inhibition of TNF-α production during mycobacterial infection lead to accelerated intracellular bacterial growth and dissemination through increasing formation of granulomas.[23,58] In this context, LOSs, along with other glycolipids such as PGL that are known to inhibit TNF-α secretion, may play an important role in granulomas formation and bacterial dissemination.[59−61] Expression of cell surface antigens on activated monocytes/macrophages and dendritic cells represents also a critical step in the immune system activation process and in the control of tuberculosis dissemination. Indeed, through cell−cell and cell−matrix interactions, cell surface markers mediate both the recruitment of cells during granuloma formation and the presentation of antigens to the T-cell receptor.[24,26] Previous studies have reported that whole *M. tuberculosis* cells as well as purified mycobacterial antigens increased the expression of both the adhesion molecules ICAM-1 and the activation marker CD40 in primary macrophages or THP-1 cell line.[27,28,62−65] Not only ICAM-1 and CD40 are expressed in macrophages and dendritic cells within lung granuloma of mice infected by *M. tuberculosis* or *M. bovis* BCG,[30,31,40,66,67] but expression of both markers are also required for the formation of mature granulomas and control of the bacterial burden.[29,32,38−40]

The present work indicates that LOS-IV, but none of other LOSs tested, induced expression of both ICAM-1 and CD40 on the surface of THP-1 macrophages. In line with this result, we have also observed that only LOS-IV stimulates IL-8 secretion from THP-1 cells (data not shown), a chemokine that plays a central role in leukocyte recruitment during granuloma formation. These differential responses between the LOS subtypes are very likely to result from structural variations in their glycan moiety. In particular, presence of the terminal *N*-acylated dideoxygalactose, that is unique to LOS-IV, appears essential to the inducing activity of the LOS family whereas the substitution by a Car derivative does not confer modulating activity. Moreover, the fact that deacylation of LOS-IV completely abrogated the cell surface antigen-inducing activity clearly indicated that the sole oligosaccharidic sequence was not sufficient to trigger the biological response of LOS-IV as both the specific oligosaccharidic core and the lipid anchor are required. This is reminiscent with other studies demonstrating that the fatty acid chains of the mannosyl-phosphatidyl inositol anchor of lipomannan and lipoarabinomannan are required for ligation to either CD14 or LPS binding protein.[68] These results are also consistent with the fact that the immunomodulatory functions of mycobacterial glycolipids/lipoglycans required the integrity of their acyl chains to mediate the formation of micelles and facilitate the multivalency/presentation to their ligands.[69] Therefore, by analogy with other glycolipids, we propose that biological activity of LOS-IV is dependent on its specific carbohydrate domain whereas its lipid moiety plays a crucial role for an efficient presentation to cellular receptors.

Macrophages isolated from ICAM-1-knockout mice exhibit decreased phagocytic activity, indicating an additional role of this molecule in phagocytosis.[70] Therefore, the enhanced regulation of ICAM-1 may impact the early phases of mycobacterial infection. In this context, it is noteworthy that *M. marinum* LOS-IV mutant carrying a transposon in *losA* was less efficient in entering/invading macrophages relative to the parental strain,[22] although it cannot be inferred whether this defect was directly linked to an altered phagocytic activity due to down-regulation of ICAM-1 expression or whether LOS-IV acts directly a ligand to a yet unidentified macrophage receptor that facilitates invasion.

In conclusion, this work represents a pioneering study devoted to the immunomodulatory role of *M. marinum* LOSs, so far the less studied mycobacterial glycolipids from biosynthetic and functional points of view. Further studies are now required to investigate whether this family of glycolipids, especially LOS-IV, plays a role in the physiopathological features such as granuloma formation during the course of infection. The

(58) Lin, P. L.; Myers, A.; Smith, L.; Bigbee, C.; Bigbee, M.; Fuhrman, C.; Grieser, H.; Chiosea, I.; Voitenek, N. N.; Capuano, S. V.; Klein, E.; Flynn, J. L. *Arthritis Rheum.* **2010**, *62*, 340–50.
(59) Reed, M. B.; Domenech, P.; Manca, C.; Su, H.; Barczak, A. K.; Kreiswirth, B. N.; Kaplan, G.; Barry, C. E., III *Nature* **2004**, *431*, 84–7.
(60) Okamoto, Y.; Fujita, Y.; Naka, T.; Hirai, M.; Tomiyasu, I.; Yano, I. *Microb. Pathog.* **2006**, *40*, 245–53.
(61) Lee, K. S.; Dubey, V. S.; Kolattukudy, P. E.; Song, C. H.; Shin, A. R.; Jung, S. B.; Yang, C. S.; Kim, S. Y.; Jo, E. K.; Park, J. K.; Kim, H. J. *FEMS Microbiol. Lett.* **2007**, *267*, 121–8.
(62) Ghosh, S.; Saxena, R. K. *Exp. Mol. Med.* **2004**, *36*, 387–95.
(63) Giacomini, E.; Iona, E.; Ferroni, L.; Miettinen, M.; Fattorini, L.; Orefici, G.; Julkunen, I.; Coccia, E. M. *J. Immunol.* **2001**, *166*, 7033–41.
(64) Xu, Y.; Liu, W.; Shen, H.; Yan, J.; Yang, E.; Wang, H. *Microbes Infect.* **2010**, *12*, 683–689.
(65) Scandurra, G. M.; Williams, R. B.; Triccas, J. A.; Pinto, R.; Gicquel, B.; Slobedman, B.; Cunningham, A.; Britton, W. J. *Microbes Infect.* **2007**, *9*, 87–95.
(66) Gonzalez-Juarrero, M.; Orme, I. M. *Infect. Immun.* **2001**, *69*, 1127–33.
(67) Mogga, S. J.; Mustafa, T.; Sviland, L.; Nilsen, R. *Scand. J. Immunol.* **2003**, *58*, 327–34.
(68) Elass, E.; Coddeville, B.; Guerardel, Y.; Kremer, L.; Maes, E.; Mazurier, J.; Legrand, D. *FEBS Lett.* **2007**, *581*, 1383–90.
(69) Sidobre, S.; Puzo, G.; Riviere, M. *Biochem. J.* **2002**, *365*, 89–97.
(70) Paine, R., III; Morris, S. B.; Jin, H.; Baleeiro, C. E.; Wilcoxen, S. E. *Am. J. Physiol. Lung. Cell. Mol. Physiol.* **2002**, *283*, L180–7.

expected results may be of high biological significance to understand the pathogenesis of other LOS-producing mycobacterial species, such as the Canetti strain, a smooth variant of *M. tuberculosis*.

Acknowledgment. This study was funded by a grant from the french National Research Agency (ANR-05-MIIM-025) to Y.G. and L.K. and a grant from the Ministère de l'Enseignement Supérieur to Y.R. The 800 and 900 MHz spectrometers were funded by Région Nord-Pas de Calais, European Union (FEDER), Ministère Français de la Recherche, Université Lille1-Sciences et Technologies, and CNRS. The 600 MHz facility used in this study was funded by the European Union, Région Nord-Pas de Calais, CNRS, and Institut Pasteur de Lille. Financial support from the TGE RMN THC Fr3050 for conducting the research on the 800 and 900 MHz spectrometers is gratefully acknowledged. The 400 MHz facility was funded by the Centre Commun de Mesure RMN.

Supporting Information Available: Figure S1, purification and TLC analysis of polar glycolipids; Figures S2−S8, supporting structural data; Table S1, summary of known LOSs sequences; Tables S2 and S3, NMR parameters of neutral and acidic OS-IV; Table S4, ORF and predicted proteins of *M. marinum* LOS cluster; and complete refs 33 and 35. This material is available free of charge via the Internet at http://pubs.acs.org.

JA105807S