Use of ^{1}H NMR ROESY for Structural Determination of *O*-Glycosylated Amino Acids from a Serine-Containing Glycopeptidolipid Antigen†

Michel Rivière and Germain Puzo*

Centre de Recherche de Biochimie et Génétique Cellulaires, Centre National de la Recherche Scientifique, 118 Route de Narbonne, 31062 Toulouse Cedex, France

Received August 1, 1991; Revised Manuscript Received December 23, 1991

ABSTRACT: A serine-containing glycopeptidolipid antigen isolated from *Mycobacterium xenopi* typified a new class of mycobacterial glycopeptidolipid antigens devoid of the C-mycoside core structure [Rivière, M., & Puzo, G. (1991) *J. Biol. Chem. 266*, 9057–9063]. The lipopeptide core assigned to C_{12}-Ser-Ser-Phe-alloThr-OCH_3 exhibits three potential sites of glycosylation. The carbohydrate parts are composed of 3-*O*-methyl-6-deoxy-α-L-talopyranosyl and 2,3,4-tri-*O*-methyl-L-rhamnopyranosyl(α1→3)-2-*O*-lauroyl-L-rhamnopyranosyl(α1→3)-L-rhamnopyranosyl(α1→3)-2,4-di-*O*-(acetyl, lauroyl)-6-deoxy-α-L-glucopyranosyl appendages. In the present work, the carbohydrate attachment sites were successfully determined by ROESY experiments on the native glycopeptidolipid using chloroform as solvent. From the NOE contacts, we unambiguously established that the acylated serine is glycosylated by the 3-*O*-methyl-6-deoxy-α-L-talopyranosyl appendage while the 2,3,4-tri-*O*-methyl-L-rhamnopyranosyl(α1→3)-2-*O*-lauroyl-L-rhamnopyranosyl(α1→3)-L-rhamnopyranosyl(α1→3)-2,4-di-*O*-(acetyl, lauroyl)-6-deoxy-α-L-glucopyranosyl appendage is bound to the C-terminal alloThr-OCH_3. From these data, the acetyl and lauroyl residues on the C-2 and C-4 of the basal monosaccharide unit were successfully localized. Furthermore, the "L" absolute configuration for the serines and the phenylalanine residues and the "D" configuration for the allothreonine were established. The primary structure of this novel type of mycobacterial antigen, a serine-containing glycopeptidolipid, has now been fully established.

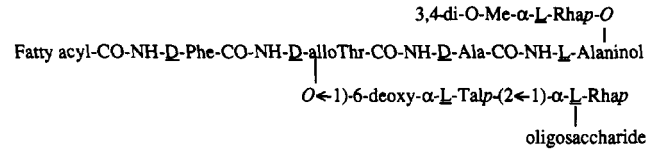
The increase in cases of atypical human tuberculosis caused by nontuberculous mycobacteria such as *Mycobacterium avium* complex, *Mycobacterium kansasii*, and *Mycobacterium xenopi* in immunocompromised hosts, and particularly in AIDS patients, has changed the epidemiology of these diseases.

The search for species-specific antigens remains an important goal in the development of better tools for the diagnosis of human mycobacterial diseases but also in the understanding mycobacteria pathogenicity at a molecular level. From the literature data [for reviews, see Brennan (1984) and Puzo (1990)], it emerges that the species-specific mycobacterial immunoreactive components are glycolipids contained in the cell walls of most of the nontuberculous mycobacteria. According to their chemical structure, these glycolipids are classified into three types: phenolic glycolipids (Phe Gl), trehalose-containing lipooligosaccharides (LOS), and C-mycoside glycopeptidolipids (GPL). Their species specificity is conferred by small oligosaccharides or monosaccharides, the structure of which is unique in nature.

The general structure of the GPLs, which are also called polar C-mycosides, is summarized by Chart I. They share a common C-mycoside core structure, differing only in the nature of the monosaccharide at the nonreducing end of the oligosaccharidic appendage. The C-mycoside GPL antigens are mainly restricted to the *M. avium* complex where they characterize the 31 serovars which compose this complex.

The intramacrophagic survival of *M. avium*, and the other pathogenic mycobacteria, can be correlated with the presence of a capsule surrounding the cell wall which may be observed when the baccillus is phagocytosed by the macrophage (Frehel et al., 1986). It was demonstrated by immunocytochemistry

Chart I



that GPLs are the major components of the *M. avium* capsule, suggesting that these molecules are involved in *M. avium* intramacrophagic survival (Terelensky & Barrow., 1983). However, recently it was observed that rough *M. avium* variants, which are devoid of GPLs, also synthesize capsule (Rastogi et al., 1989). Thus the protecting role of GPLs against macrophage microbicidal activity is not yet clearly established.

Recently, we isolated a major immunoreactive glycopeptidolipid, named GPL X-I, from the cell wall of *M. xenopi*, a pathogenic nontuberculous mycobacteria (Rivière & Puzo, 1991). The structure of this antigen was partially established and is the first example of a glycopeptidolipid mycobacterial antigen devoid of a C-mycoside core. Moreover, despite the presence of a tetrasaccharidic appendage, its polarity in silicic acid TLC[1] is lower than that of the C-mycosides, and thus the GPL X-I structure typifies a new class of mycobacterial antigens, namely, serine-containing GPLs.

The unusual chromatographic behavior of GPL X-I, compared to the known GPLs, is related to the presence of two C_{12} fatty acids bound to the tetrasaccharidic appendage. To localize these alkali-labile groups and to establish its carbohydrate structure, the native and peracetylated GPL X-I derivative have been successfully analyzed by 2D NMR homo-

† This work was supported by the Region Midi-Pyrénées (Ref. RECH 8900712) and La Ligue Nationale Contre le Cancer.

* Address correspondence to this author at the Centre de Recherche de Biochimie et Génétique Cellulaires, CNRS, 118 Rte. de Narbonne, 31062 Toulouse Cedex, France.

[1] Abbreviations: alloThr, allothreonine; Me, methyl; Talp, talopyranosyl; Rhap, rhamnopyranosyl; HPLC, high-performance liquid chromatography; TLC, thin layer chromatography; COSY, correlated spectroscopy; DCI, desorption chemical ionization.

0006-2960/92/0431-3575$03.00/0

nuclear correlation spectroscopy, mass spectrometry, and classical carbohydrate analysis. The carbohydrate part was found to be composed of the tetrasaccharide 2,3,4-tri-*O*-Me-L-Rhap(α1→3)-2-*O*-lauroyl-L-Rhap(α1→3)-L-Rhap(α1→3)-2,4-di-*O*-(acetyl,lauroyl)-6-deoxy-α-L-Glcp and a 3-*O*-Me-6-deoxy-α-L-Talp monosaccharidic unit.

The unexpected presence of two serine residues, instead of the alanine and alaninol residues found in the known GPLs, was revealed, after hydrolysis, by reverse-phase HPLC analysis of the phenyl thiocarbamyl amino acid derivatives. The tetrapeptidic sequence Ser-Ser-Phe-alloThr-OCH_3 was established from the diagnostic acylium fragment ions observed in the pyrolysis electron impact mass spectrum. The determination of the glycosylation sites was a major analytical problem due to the heterogeneity of the carbohydrate parts (a tetrasaccharide and a monosaccharide) and to the presence of three potential sites of glycosylation within the peptide. From the DCI mass spectrum of native GPL X-I, we have tentatively proposed that the tetrasaccharide binds the C-terminal alloThr-OCH_3 while the monosaccharide glycosylates the N-terminal Ser residue.

The goal of the present report is the development of a new analytical strategy, based on two-dimensional ^{1}H NMR rotating frame NOE (2D ROESY) analysis of native GPL X-I, allowing the unambiguous characterization of the glycosylation sites, the peptide sequence, and finally the exact localization of the acetyl and lauroyl groups at the C-2 and C-4 of the 6-deoxy-α-L-Glcp. This approach could be generally used for the characterization of the glycosylation sites of glycopeptides produced from the glycoprotein degradation.

Materials and Methods

Amino Acid Absolute Configuration. GPL X-I (1 mg) was hydrolyzed as previously described (Rivière & Puzo, 1991), and the amino acid absolute configurations were established by chiral TLC (Chiralplate, Macherey-Nagel, Germany) according to Günther (1988). Plates (10 × 10 cm) were revealed by 0.2% ninhydrin solution in ethanol. Amino acid standards were from Sigma. Phenylalanine and allothreonine enantiomers were resolved with methanol–water–acetonitrile, 50:50:200 (v/v/v), while D- and L-serine were separated with methanol–water, 10:80 (v/v).

NMR Spectroscopy. GPL X-I glycopeptidolipid was purified by HPLC as previously described. Samples (10 mg) were solubilized in C^2HCl_3, degassed in the NMR tube by repeated evacuation, and sealed under argon.

^{1}H NMR was performed on a Bruker AM-300 WB spectrometer operating at 300 MHz with an Aspect 2000 computer. Proton spectra were recorded at 40 °C, and chemical shifts were referenced indirectly to external tetramethylsilane by setting the ^{1}H signal of residual $CHCl_3$ at 7.258 ppm. The 1D spectrum was recorded by using 40 ° pulses with a recycle delay of 0.3 s and a 2.7-s acquisition time. A total of 64 scans were accumulated. The spectral width used was 3 kHz, which, with a data memory of 16K expanded to 32K by zero filling, gave a digital resolution of 0.18 Hz.

The standard 2D COSY (Aue et al., 1976) and relayed coherence transfer COSY (RCT COSY) (Eich et al., 1982) pulse sequences supplied by Bruker were employed [for theoretical aspects, see Ernst et al. (1987)]. A total of 256 free induction decays (FID) of 16 scans for COSY and 64 scans for RCOSY (recycle delays of 1 s) were accumulated. The RCOSY was obtained with a mixing time (D2) of 25 ms.

2D spectra are shown in the absolute value representation as contour plots. They were obtained over a spectral width of 2252 Hz with 256 × 1024 matrix data points, and expanded to 1024 × 2048 by zero filling, which give a final digital resolution of 2.1 Hz. The spectra were multiplied by a nonshifted sine-bell digital filtering function in both dimensions to enhance the resolution.

The two-dimensional rotating frame NOE (2D ROESY) (Bothner-By et al., 1984; Bax & Davis, 1985) spectra were recorded in the phase-sensitive mode using the TPPI method, i.e., time-proportional phase incrementation (Marion & Wüthrich, 1983). The rf carrier frequency was placed about 1 kHz downfield relative to the center of proton resonances during the spin-lock time and at the center during evolution and acquisition. Two different spin-lock times were used: 150 and 250 ms with a continuous field of about 2500 Hz. The spectral width was 2252 Hz in both dimensions, and the spectra were recorded with 256 experiments (FID) of 64 scans each (recycle delays of 2 s). The 256 × 1024 data point matrices were zero-filled to 1024 × 2048 and multiplied with a nonshifted sine bell along F_2 and a $\pi/3$ phase-shifted sine bell along F_1 prior to phase-sensitive Fourier transformation.

Results

The absolute configuration of the amino acids was established by TLC on chiral plates after hydrolysis. The two serines belong to the L series as does phenylalanine, while the alloThr belongs to the D series. Compared to the D–Phe, D–alloThr, and D-Ala tripeptidic core of the GPLs from the *M. avium* complex, only the allothreonine has the same configuration in GPL X-I.

^{1}H NMR Correlated Spectroscopy. The spin systems of each amino acid residue were determined on the basis of the connectivity plots observed in the ^{1}H-^{1}H COSY and RCOSY spectra (Figure 1). In all these experiments GPL X-I was dissolved in deuterated chloroform instead of acetone to enhance the resolution of the four amide protons in the 1D ^{1}H NMR spectrum.

The cross-peak pattern showing a connectivity plot between the β protons at 3.94 ppm and γ methyl protons at 1.145 ppm unambiguously characterized the allothreonine. From the β-proton connectivity, the α proton was localized at 4.6 ppm, and finally the allothreonine amide proton resonance was assigned to the doublet at δ = 7.29 ppm. Starting from the high-field NH proton resonance at δ = 6.47 ppm, an A_2X spin system was found, supported by the RCOSY spectrum (Figure 1b), which is assigned to serine residue 1 (Marion & Wüthrich, 1983). The unambiguous proton assignments of the two remaining amino acid residues (Ser and Phe) required the analysis of the peracetylated derivative. From the COSY and RCOSY spectra of the native compound, the two amide proton resonances (δ = 7.395 ppm, δ = 6.855 ppm) show a similar connectivity pattern with the α-proton resonances at 4.461 and 4.764 ppm and the β protons at 3.878 and 3.732 ppm and 3.235 and 3.018 ppm, respectively. From the literature chemical shift data ($2.7 < \delta_{H\beta\text{-PHe}} < 3.3$ ppm; $3.8 < \delta_{H\beta\text{-Ser}} < 4.8$ ppm) (Wüthrich, 1986), the lower field β-proton resonances can be assigned to the second Ser residue. This attribution is supported by the shift of the β protons from 3.878 and 3.732 ppm to 4.35 and 4.22 ppm in the spectrum of the peracetylated GPL X-I derivative (Table I).

The complete assignment of the carbohydrate protons of the GPL X-I dissolved in deuterated acetone has been reported previously (Rivière & Puzo, 1991). However, as mentioned above, the chemical shift of the amino acid protons was established from a solution of GPL X-I in deuterated chloroform. Since an exact knowledge of the anomeric proton chemical shifts is required for further localization of the tetrasaccharide appendage and the monosaccharide residue on the peptide core,

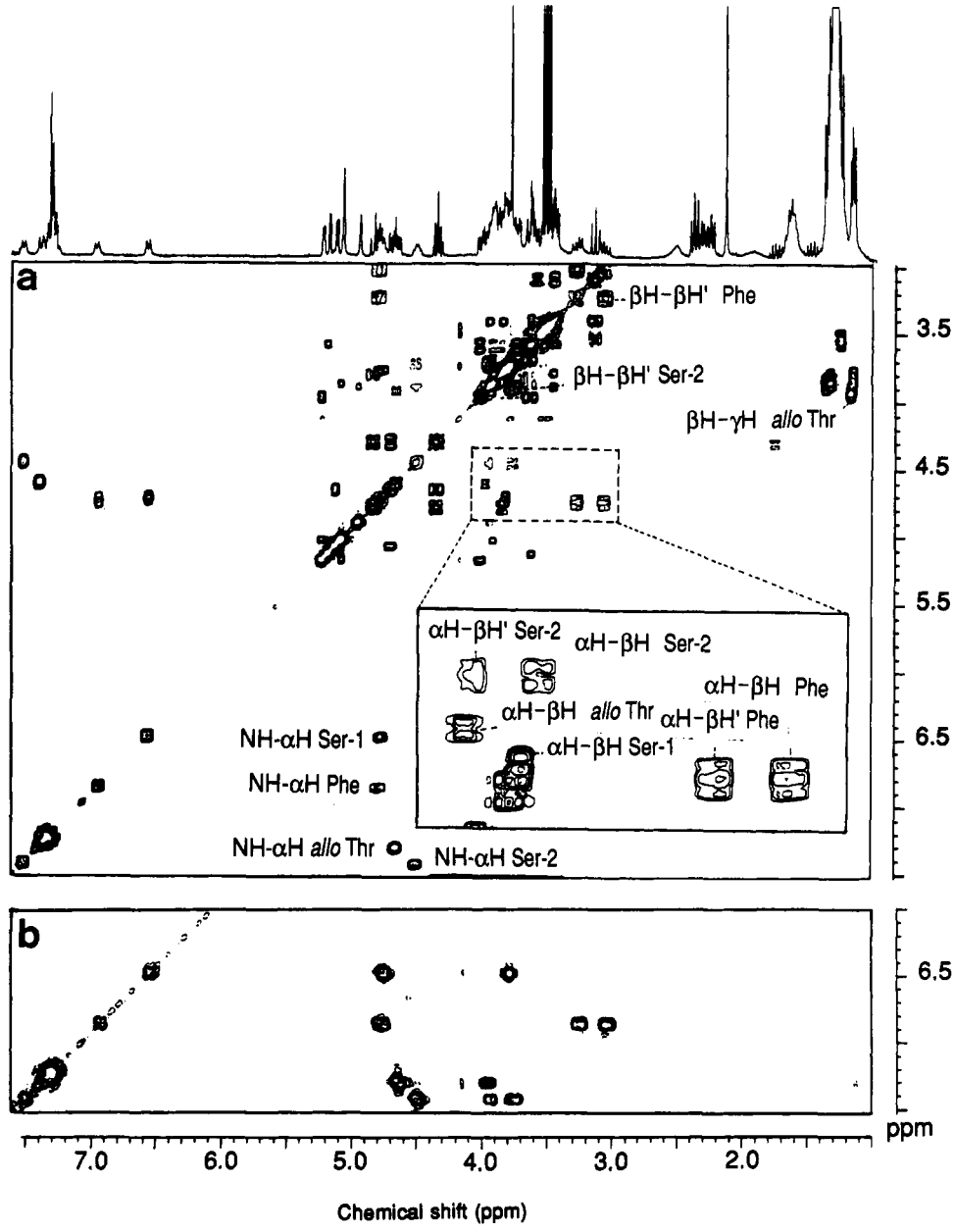


FIGURE 1: 300-MHz ^{1}H NMR spectra of the native GPL X-I (5 mg) recorded in C^2HCl_3. The complete 1D spectrum is at the top of the figure showing the four characteristic doublet signals of the amide protons between 6.5 and 7.5 ppm. (a) Standard COSY 90 spectrum with an enlargement showing the identification of the connectivity plot corresponding to the amino acid protons. (b) Partial relayed COSY spectrum showing the correlation spots between the amide and the α and β amino acid protons.

the ring proton resonances were assigned (Table I). This assignment, based on the sequential connectivities of cross peaks in the 2D COSY and RCOSY spectra of the native and peracetylated GPL X-I, is straightforward and in excellent agreement with those previously obtained. The connectivity plot systems of the 2,4-di-*O*-(acetyl, lauroyl)-6-α-L-Glcp, 3-*O*-Me-6-deoxy-α-L-Talp, and 2,3,4-tri-*O*-Me-α-L-Rhap are well resolved and allow their unambiguous assignments to be made. Moreover, in spite of the overlapping of their respective anomeric proton signals at 5.17 ppm, the two remaining rhamnopyranosyl unit spin systems were descriminated by the presence of an acyl group on one of them, which induces a characteristic shift of the *gem* H-2 proton signal at δ = 5.17 ppm.

^{1}H ROESY NMR Experiments. Proton NMR 2D NOESY experiments were attempted to establish the sequence of the tetrapeptide core and to localize the glycosidic appendages by dipolar magnetization transfer. However, this approach was unsuccessful despite altering the temperature from 25 to 50 °C or using different solvents such as $(C^2H_3)_2CO$ and C^2HCl_3 in order to modify the correlation time τ_c and thus enhance NOE exchanges. We therefore selected ROESY experiments. The expanded region of the 2D ROESY spectrum presented in Figure 2 shows the magnetization transfer cross peaks involving the amide protons. The upper spectrum shows a cross peak between the methylene protons (2.17 ppm) of a C_{12} fatty acyl residue and the NH proton of serine 1 (6.47 ppm) indicating that the lipid moiety of the core is bound to serine 1 residue. The lower spectrum shows the NOE contacts between amide protons and their neighboring amino acid α protons and allows the sequence of the tetrapeptidic core to be unambiguously established. The cross peak between the NH proton (δ = 7.395 ppm) of serine 2 and the α proton (δ = 4.743 ppm) of serine 1 indicates that the two serine residues are linked together. In the same way, cross-relaxation exchanges are observed between the Hα at 4.461 ppm of serine 2 and the NH proton (6.855 ppm) of phenylalanine. Finally, from the NOE contact between the phenylalanine Hα (4.764

Table I: ^{1}H NMR Assignment of Native and Peracetylated GPL X-I in C^2HCl_3

residue	proton	^{1}H chemical shifts (ppm) native	acetylated
serine 1	NH	6.47	6.27
	Hα	4.743	4.43
	Hβ	3.771	3.8
serine 2	NH	7.395	7.085
	Hα	4.461	4.53
	Hβ	3.878, 3.732	4.35, 4.22
allothreonine	NH	7.29	6.82
	Hα	4.605	4.62
	Hβ	3.94	3.85
	γCH_3	1.145	–[a]
phenylalanine	NH	6.855	6.91
	Hα	4.764	4.68
	Hβ	3.235, 3.018	3.23, 3.05
	C_6H_5	7.245	7.25
2,3,4-tri-*O*-Me-α-L-Rhap	H-1	5.137	4.875
	H-2	3.589	3.71
	H-3	3.41	3.32
	H-4	3.097	3.064
	H-5	3.537	–
	CH_3	1.21	–
2-*O*-lauroyl-α-L-Rhap	H-1	5.017	4.842
	H-2	5.174	4.935
	H-3	3.981	3.93
	H-4	3.592	5.024
	H-5	3.846	–
	CH_3	1.335	–
α-L-Rhap	H-1	5.017	4.808
	H-2	3.868	4.935
	H-3	3.716	3.93
	H-4	3.59	4.991
	H-5	3.868	–
	CH_3	1.3	–
2,4-di-*O*-(acetyl,lauroyl)-6-deoxy-α-L-Glcp	H-1	5.086	4.901
	H-2	4.652	4.662
	H-3	4.318	3.995
	H-4	4.785	4.787
	H-5	3.797	–
	CH_3	1.12	–
3-*O*-Me-6-deoxy-α-L-Talp	H-1	4.896	4.875
	H-2	3.902	5.08
	H-3	3.418	3.57
	H-4	3.802	5.229
	H-5	3.868	–
	CH_3	1.29	–

[a] –, unlocalized.

ppm) and the amide proton (7.29 ppm) of the allothreonine, the sequence: fatty acyl → Ser-1 →Ser-2 → Phe → alloThr-OCH_3 is established and is in complete agreement with that previously proposed.

The partial 2D ROESY spectrum (Figure 3) shows some of the exchange cross peaks involving carbohydrate ring protons. The cross peaks due to chemical exchange transfer by the homonuclear Hartmann–Hahn effect (dotted line) are descriminated from NOE cross-relaxation transfer peaks because they appear in opposite phase. The NOE cross peaks found in this region are mainly due to interactions with anomeric carbohydrate protons. Two cross peaks (indicated by arrows), both well defined, are decisive for the localization of the saccharidic appendages. The strong NOE exchange between the anomeric proton of the 3-*O*-Me-6-deoxy-α-L-Talp (4.896 ppm) and the β protons (3.77 ppm) of the acylated serine unit indicates that it is this monosaccharide that glycosylates the amino acid. In a similar fashion the glycosylation of the allothreonine by the tetrasaccharide moiety is deduced from the NOE contact between the anomeric proton 5.086 ppm of the basal 2,4-di-*O*-(acetyl, lauroyl)-6-deoxy-α-L-Glcp monosaccharide unit and the β proton of the allothreonine (3.94 ppm).

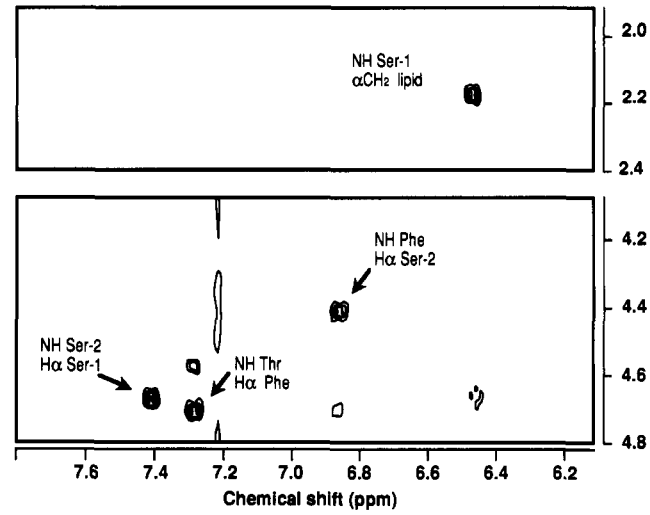


FIGURE 2: Partial 2D ROESY spectrum of native GPL X-I showing the magnetization exchange cross peaks involving the amino acid amide protons. Positive and negative contour levels are plotted without distinction. The spin-lock time was 250 ms.

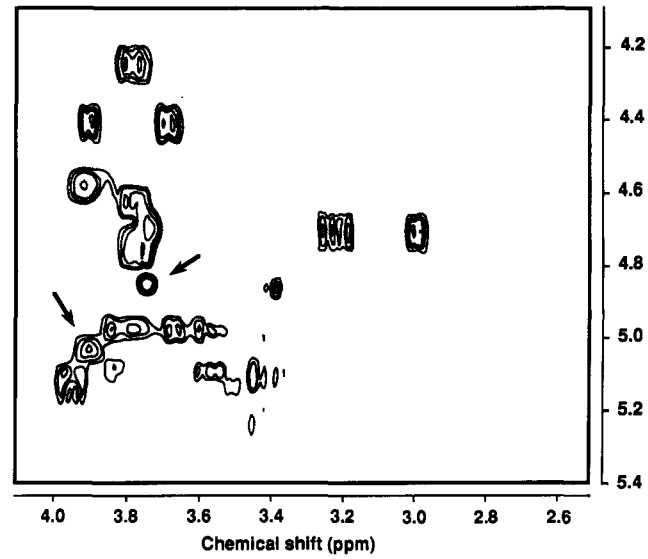


FIGURE 3: Partial 2D ROESY spectrum of native GPL X-I corresponding to exchange cross peaks of carbohydrate protons. Connectivity cross peaks involving magnetization exchange between carbohydrate anomeric protons and β amino acid protons are indicated by an arrow. Diagonal in-phase contour levels are plotted with a dotted line while anti-phase levels are with a solid line. The spin-lock time was 250 ms.

By EI mass spectrometry, ^{1}H–^{1}H COSY, and RCOSY experiments, the C_{12} fatty acyl and the acetyl residues were localized on the C-4 and C-2 of the 6-deoxy-α-L-Glcp, but their respective positions were unresolved. However, in the 2D ROESY experiment, a NOE contact (not shown) is observed between the H-5 of the 6-deoxy-α-L-Glcp and the signal at $\delta = 2.28$ ppm assigned to the β methylene protons of the C_{12} fatty acyl residue. From molecular modeling, this dipolar exchange is in agreement with the C_{12} residue acylating the C-4 of 6-deoxy-α-L-Glcp. Therefore, the structure 2-*O*-acetyl-4-*O*-lauroyl-6-deoxy-α-L-Glcp can be proposed for the basal monosaccharide of the tetrasaccharidic moiety. The complete structure of the GPL X-I is established from the NOE contacts reported in Figure 4.

Discussion

Due to their serological importance, the structure of the *M. avium* complex GPLs has been extensively studied, and it

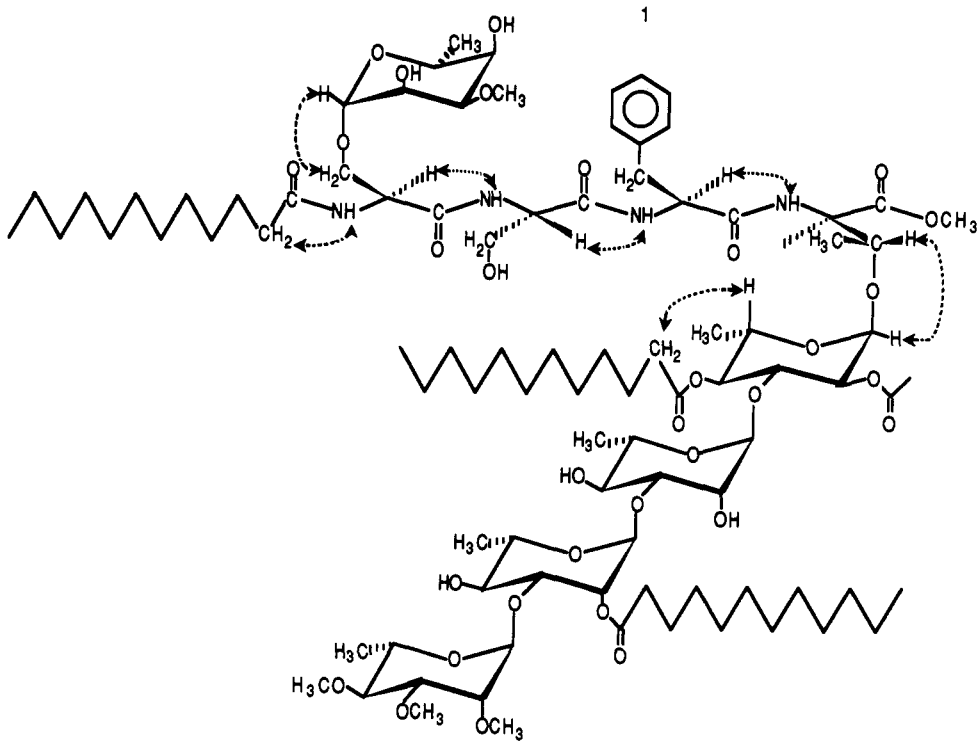


FIGURE 4: Complete structure of glycopeptidolipid GPL X-I from *M. xenopi* showing NOE exchange transfers observed with 2D ROESY experiments.

emerges that their serotypic specificities lie in the nature and arrangement of the sugars at the nonreducing end of the oligosaccharide moiety (Chart I). Most of these studies have therefore focused on the structural elucidation of this part of the molecule, and, for this reason, the analytical strategies routinely used rely on the specific release of the oligosaccharide linked to allothreonine by reductive alkalinolysis [for example, see Chatterjee et al., (1987)]. This procedure, which does not affect the alaninol-linked monosaccharidic unit, allows the recovering of the oligosaccharide alditol for its structural determination by conventional carbohydrate analytical tools. However, the drawback of this approach is that it requires considerable quantities of purified compound which are not always available, but also that it does not allow the recovering of alkali labile groups borne by carbohydrates and which are of prior importance in antibody binding reaction (Rivoire et al., 1989). In the case of the *M. xenopi* GPL X-I, such conventional strategy could not be applied due to (i) the nonspecificity of alkalinolysis which affects the two saccharide moiety, (ii) the presence of fatty acyl and acetyl groups bound to the tetrasaccharide which are lost during alkalinolysis, and (iii) the methyl ester form of the carboxy-terminal amino acid.

Taking into account these specific structural features of GPL X-I and the small amount of available material (25 mg) for conventional analysis route, we developed a new analytical approach for the elucidation of the peptidic glycosylation sites as well as the localization of fatty acyl and acetyl groups bound to the tetrasaccharide. This method is based on the direct analysis of the native molecule, by combination of two-dimensional homonuclear correlated spectroscopy and two-dimensional rotating frame NOE (2D ROESY) experiments. The limiting step of this strategy is the identification and the attribution of each amino acid and carbohydrate spin system involved in the linkage between the peptidic and the carbohydrate parts. These can be elucidated by correlated spectroscopy (COSY, RCOSY, or TOCSY) experiments.

In a previous report, the structure of the carbohydrate parts of GPL X-I was determined by ^{1}H NMR experiments but due to the overlapping of the amide proton resonances we were unable to assign the amino acid residues. To overcome this analytical problem, different solvents were investigated and finally chloroform was found to be the most suitable in spite of the overlapping of the two rhamnopyranosyl anomeric proton resonances. The proton chemical shifts of GPL X-I in chloroform were in agreement with those previously found for the GPL X-I carbohydrate part, and with those published by Marion and Wüthrich (1983) for the amino acids. The amino acid spin systems were assigned on the basis of the sequential connectivity patterns and the chemical shift of the protons. By this approach, one serine and the allothreonine were unambiguously identified. The two remaining amino acids, a serine and a phenylalanine showing similar connectivity patterns in RCOSY, were tentatively characterized from the downfield resonance of the β proton of the serine compared to that of the phenylalanine. This assumption was supported by the downfield shift of the β proton resonances of the serine in the spectrum of peracetylated GPL X-I. The two serines exhibit different connectivity patterns which can be correlated to the presence of one glycosylated serine. However, this information was not decisive since the glycosylated serine must be localized in the tetrapeptide and its glycosyl residue characterized.

Sequencing of the native peptide core by 2D NOESY experiments was unsuccessful. No significant cross relaxation via dipolar exchange was observed even by changing the mixing time, the temperature, or the solvent. This was probably due to the inadequate relationship between the Larmor frequency and the rotational correlation time τ_c of the molecule since NOE exchanges are too small or equal to zero. To overcome this problem the ROESY sequence was selected. The spectra were recorded with two different mixing times (150 and 250 ms), but no significant differences were observed. Unlike NOESY, rotating frame NOE experiments provided spatial proximity informations. The sequence of the tetrapeptide was easily established from the expected cross peaks between the amide protons and the neighboring amino acid α protons.

Moreover, an additional dipolar cross-relaxation exchange is observed between the NH proton of serine 1 and aliphatic protons of a fatty acyl residue, allowing its localization at the NH_2 terminus of the tetrapeptide. The sequence deduced from these NOE contacts is in complete agreement with that previously established from EI mass spectrometry analysis of the native GPL X-I. The strong NOE exchange between the anomeric proton of 2-*O*-acetyl-4-*O*-lauroyl-6-deoxy-α-L-Glcp and the β proton of the allothreonine indicates the glycosylation of the allothreonine by the tetrasaccharide. Similarly, it was established that the 3-*O*-Me-6-deoxy-α-L-Talp glycosylated the acylated serine 1. Thus ^{1}H NMR ROESY experiments, which can be applied to the milligram range of native glycopeptidolipid, is a powerful analytical tool allowing the sequencing of the peptide and the determination of the glycosylation sites.

In the GPL X-I structure proposed previously, the respective location of the C_{12} fatty acyl and acetyl residues linked to the 6-deoxy-α-L-Glcp was undetermined. The downfield chemical shifts of the H-2 and H-4 of this sugar unit were consistent with the presence of *gem* carboxy groups. Chemical procedures such as selective deacetylation by mild alkalinolysis treatment with aminopropyl alcohol (Haines, 1990) were unsuccessful. In order to localize acyl substituents on the peptide or oligosaccharide, ^{1}H–^{13}C heteronuclear multiple bound correlation (HMBC) experiments have been successfully applied (Summers et al., 1986; Lerner et al., 1987; Van Halbeek, 1990). However, in the GPL X-I ^{13}C spectrum (75 MHz), the carbonyl resonances are not resolved enough to allow their unambiguous assignment and therefore determination of the respective location of the acetyl and acyl substituents by this method. Nevertheless, the ROE connectivity cross peak between the H-5 of the 6-deoxy-α-L-Glcp and an aliphatic methylene proton signal of a lauroyl residue strongly suggests that the lauroyl fatty acid esterifies the hydroxyl group of the C-4 while the acetyl is linked to the C-2 hydroxyl group.

As previously mentioned, the 31 serovars of the *M. avium* complex differ by the structure of the distal sugar epitope of the oligosaccharide moiety of the GPL antigens. Using monoclonal antibodies generated against various *M. avium* serovars associated with the HIV, it was observed that the GPLs purified after smooth alkalinolysis from serovars 1, 4, and 9 do not bind their respective monoclonals while the native GPLs present in the crude lipidic fraction are recognized (Rivoire et al., 1989). These experiments reveal that alkali treatment, proposed to simplify GPL purification, leads to the loss of alkali labile residues bound to the sugars which are determinant for mycobacteriophage D4 or antibody binding (Dhariwal et al., 1986). Moreover, because the oligosaccharidic hapten structures were established after reductive alkalinolysis of the whole GPLs, it can be assumed that some of these structures are incomplete. Their structural reinvestigation (Rivoire et al., 1989; Jardine et al., 1989; McNeil et al., 1987) revealed the presence of *O*-acetyl groups bound to monosaccharide localized at the nonreducing end of the oligosaccharide. A similar phenomenon has been previously described for the phenolic glycolipid epitope 2,6-dideoxy-4-*O*-Me-α-D-arabino-hexopyranosyl(1→3)-4-*O*-Ac-2-*O*-Me-α-L-Fucp of *M. kansasii* (Gilleron et al., 1990). These observations highlight the importance of the development of new analytical strategies allowing the structural determination of complex glycolipid antigens in their native form. The analytical approach developed here for the structural determination of the GPL X-I is applicable both to the reinvestigation of the *M. avium* GPLs and to the analysis of glycopeptides arising from enzymatic or chemical hydrolysis of complex glycoprotein.

ACKNOWLEDGMENTS

We thank P. Servin for recording NMR spectra and for helpful discussions and P. Winterton for correcting the English version of the manuscript.

REFERENCES

Aue, W. P., Bartholdi, E., & Ernst, R. R. (1976) *J. Chem. Phys. 64*, 2229–2246.

Barrow, W. W., Ullom, B. P., & Brennan, P. J. (1980) *J. Bacteriol. 144*, 814–822.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 63*, 207–213.

Bothner-By, A. A., Stephens, R. L., Lee, J., Warren, C. D., & Jeanloz, R. W. (1984) *J. Am. Chem. Soc. 106*, 811–813.

Brennan, P. J. (1984) *The Mycobacteria: A Sourcebook* (Kubica, G. P. & Wayne, G. L., Eds) Part A, pp 467–490, New York.

Brennan, P. J., & Goren, M. B. (1979) *J. Biol. Chem. 254*, 4205–4211.

Chatterjee, D., Aspinall, G. O., & Brennan, P. J. (1987) *J. Biol. Chem. 262*, 3528–3533.

Dhariwal, K. R., Liav, A., Vatter, A. E., Dhariwal, G., & Goren, M. B. (1986) *J. Bacteriol. 168*, 283–293.

Draper, P., & Rees, R. J. W. (1970) *Nature 228*, 860–861.

Eich, G., Bodenhauser, G., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 3731–3732.

Ernst, R. R., Bodenhauser, G., & Wokaun, A. (1987) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Oxford University Press (Clarendon), London and New York.

Frehel, C., Ryter, A., Rastogi, N., & David, H. L. (1986) *Ann. Inst. Pasteur/Microbiol. 137B*, 239–257.

Gilleron, M., Venisse, A., Fournié, J. J., Rivière, M., Dupont, M. A., Gas, N., & Puzo, G. (1990) *Eur. J. Biochem. 189*, 167–173.

Günther, K. (1988) *J. Chromatogr. 448*, 11–30.

Haines, H. A., Konowicz, P. A., & Jones, H. F. (1990) *Carbohydr. Res. 205*, 406–409.

Jardine, I., Scanlan, G., McNeil, M., & Brennan, P. J. (1989) *Anal. Chem. 61*, 416–422.

Kumar, A., Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1–15.

Lerner, L., & Bax, A. (1987) *Carbohydr. Res. 166*, 35–46.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–970.

McNeil, M., Tsang, Y. A., & Brennan, P. J. (1987) *J. Biol. Chem. 262*, 2630–2635.

Puzo, G. (1990) *Crit. Rev. Microbiol. 17*, 305–327.

Rastogi, N., Levy-Frebault, V., Blom-Potar, M.-C., & David, H. L. (1989) *Zbl. Bakt. Hyg. A 270*, 345–360.

Rivière, M., & Puzo, G. (1991) *J. Biol. Chem. 266*, 9057–9063.

Rivoire, B., Ranchoff, B. J., Chatterjee, D., Gaylord, H., Tsang, A. Y., Kolk, A. H. J., Aspinall, G. O., & Brennan, P. J. (1989) *Infect. Immun. 57*, 3147–3158.

Summers, M. F., Marzilli, L. G., & Bax, A. (1986) *J. Am. Chem. Soc. 108*, 4285–4294.

Terelensky, M. J., & Barrow, W. W. (1983) *Infect. Immun. 41*, 1312–1321.

Van Halbeek, H. (1990) in *Frontiers of NMR in Molecular Biology*, pp 195–213, Alan R. Liss, Inc., New York.

Woodbury, J. L., & Barrow, W. W., (1989) *J. Gen. Microbiol. 135*, 1875–1884.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, pp 13–26, John Wiley & Sons, New York.