Published in final edited form as:
*J Am Soc Mass Spectrom*. 2011 December ; 22(12): 2160–2170. doi:10.1007/s13361-011-0240-7.

# Structural definition of trehalose 6-monomycolates and trehalose 6,6'-dimycolates from the pathogen *Rhodococcus equi* by multiple-stage linear ion-trap mass spectrometry with electrospray ionization

**Fong-Fu Hsu**[*,1], **Jens Wohlmann**[2], **John Turk**[1], and **Albert Haas**[2]

[1]Mass Spectrometry Resource, Division of Endocrinology, Diabetes, Metabolism, and Lipid research, Department of Internal Medicine, Box 8127, Washington University School of Medicine, St. Louis, MO 63110

[2]Institute for Cell Biology, University of Bonn, Ulrich-Haberland-Str. 61a 53121 Bonn, Germany

## Abstract

The cell wall of the pathogenic bacterium *Rhodococcus equi* (*R. equi*) contains abundant trehalose monomycolate (TMM) and trehalose dimycolate (TDM), the glycolipids bearing mycolic acids. Here, we describe multiple-stage ($MS^n$) linear ion-trap (LIT) mass spectrometric approaches toward structural characterization of TMM and TDM desorbed as $[M + Alk]^+$ (Alk = Na, Li) and as $[M + X]^-$ (X = $CH_3CO_2$, $HCO_2$) ions by electrospray ionization (ESI). Upon $MS^n$ (n=2,3,4) on the $[M + Alk]^+$ or the $[M + X]^-$ adduct ions of TMM and TDM, abundant structurally informative fragmentations are readily available, permitting fast assignment of the length of the meromycolate chain and of the α-branch on the mycolyl residues. In this way, structures of TMM and TDM isolated from pathogenic *R. equi* strain 103 can be determined. Our results indicate that the major TMM and TDM molecules possess 6, and/or 6'-mycolyl groups that consist of mainly C14 and C16 α-branches with meromycolate branches ranging from C18 to C28, similar to the structures of the unbound mycolic acids found in the cell envelope. Up to 60 isobaric isomers varying in chain length of the α-branch and of the meromycolate backbone were observed for some of the TDM species in the mixture. This mass spectrometric approach provides a direct method that affords identification of various TMM and TDM isomers in a mixture of which the complexity of this lipid class has not been previously reported using other analytical methods.

### Keywords

Cord factor; Trehalose dimycolate; Trehalose monomycolate; mycolic acids; glycolipids; *Rhodococcus equi*.; internal glucose loss; tandem mass spectrometry; ESI

## Introduction

*Rhodococcus equi* (*R. equi*) is a gram-positive intracellular pathogen that can cause severe bronchopneumonia in foals and in immuno-compromised individuals such as patients with AIDS . It is one of the major causes of lung disease in foals between 1 and 6 months of age [1-3]. The cell envelope contains many lipid species with unusual structures, including mycolic acids, trehalose monomycolate (TMM) and trehalose dimycolate (TDM) [4, 5].

[*]To whom the correspondence should be addressed: Dr. Fong-Fu Hsu, Box 8127, Washington University School of Medicine, 660 S Euclid, St. Louis, MO 63110. Tel: 314-362-0056, fhsu@im.wustl.edu..

TDM (or called cord factor) and TMM consist of a trehalose core to which two or one mycolic acid residue was esterified at 6,6' or at 6 position to form trehalose 6,6'-dimycolate (TDM) or trehalose 6-monomycolate (TMM), respectively (see Scheme 1 and 2). Mycolic acids are long-chain α-alkyl-β-hydroxy fatty acids produced by the mycolata including the genera *Corynebacterium*, *Mycobacterium*, and *Nocardia*, *Rhodococcus* The chain length ranges from 20 (shortest chains in corynebacteria) to more than 80 carbons (longest ones in mycobacteria), depending on the producing species. For example, *R. equi* strain 103 contains a homologous series of mycolic acids having chain length ranging from C30 to C50 with 0 to 2 double bonds [6], while mycolic acids from other strains have chain length between C24 and C48 with 0 to 4 double bonds [4]. During growth of *Mycobacterium smegmatis* in biofilms, TDM in cell envelope is hydrolyzed by a TDM-specific esterase to release free mycolic acids [7]. Mycolic acids, TDM, and TMM, together with phospholipids such as cardiolipin, phosphatidylethanolamine, and phosphatidylinositol as well as glycolipids such as phosphatidylinositol mannosides perform filler roles in completing the outer leaflet of the asymmetric lipid bilayer [8]..

The biological activities of TDM and TMM in infection with pathogenic mycolata , including immunomodulation [9], granulomagenic activity [10] and the participation of TDM in the inhibition of phagosome-lysosome fusion have been well documented [11-13]. Pro-inflammatory cytokine production, granuloma formation, cachexia, and mortality, can also be induced by TDM [10, 14].

The traditional method for characterization of these complex lipids has been a difficult task, requiring laborious separation, purification, and chemical reaction, followed by spectroscopic analyses using IR, proton and carbon NMR, and GC/MS [15-19]. Recently, a MALDI-TOF mass spectrometric approach has been used to determine the molecular masses of intact TMM [20] and TDM [21], and of the masses of the mycolic acid moieties as methyl esters following their release from TMM and TDM by hydrolysis. This approach requires TLC separation of the released mycolic acids into subclasses and does not provide structural information [20, 21]. Here, we report a simple LIT ESI-$MS^n$ method towards direct characterization of TMM and TDM isolated from the cell envelope of pathogenic *Rhodococcus equi*, revealing the numerous structures including the various isomers for each of the lipid species.

## Method

### Sample preparation

*R. equi* strain 103, a virulent, plasmid-containing strain whose chromosomal DNA sequence has recently been reported was originally isolated from the lung of a pneumonic foal in Ontario, Canada [22], and was grown in 50 ml brain heart infusion broth (BHI) at 37°C, with shaking at 200 rpm to an optical density (600 nm) of 1.0 and autoclaved at 121°C under 2.15 kPa for 20 min. Total lipids were extracted as previously described [6]. For further purification, 10 mg of the crude lipid extracts in 300ul $CHCl_3/CH_3OH$ (2:1; v/v), were loaded to a 3ml/200mg Macherey-Nagel amino Chromabond Sep-Pak column (Duren, Germany). The column was first washed with 2 ml EtOAc:Hexane (15:85 v/v), followed by 1.5 ml diisopropyl ether:HOAc (98:2; v/v), and then eluted with 2 ml Acetone/Methanol (9:1.35, v/v) (by gravity) to a vial. The eluant containing TDM and TMM was dried under a stream of nitrogen. The dried sample was redissolved in $CHCl_3/CH_3OH$ (1/2) before ESI-MS analysis.

### Mass Spectrometry

Low-energy CID $MS^n$ experiments were conducted on a Thermo Finnigan (San Jose, CA) linear ion-trap (LIT) mass spectrometer with Xcalibur (version 2.01) operating system. High resolution (R=100,000 at *m/z* 400) mass measurements on the $[M + Na]^+$ ions of the TMM and TDM molecules and their subsequent $MS^n$ fragment ions were performed on a Thermo LTQ Orbitrap Velos. TMM and TDM (100 *u*L) from *R. equi* were dissolved in chloroform/methanol (1/2), and $CH_3CO_2Na$ or $CH_3CO_2Li$ (2 *u*m) was added before infusion (2 μL/min) into the ESI source, where the skimmer was set at ground potential, the electrospray needle was set at 4.5 kV, and temperature of the heated capillary was 300 °C. The automatic gain control of the ion trap was set to $5 \times 10^4$, with a maximum injection time of 200 ms. Helium was used as the buffer and collision gas at a pressure of $1x10^{-3}$ mbar (0.75 mTorr). The mass resolution of the instrument was tuned to 0.6 Da at half peak height. The $MS^n$ experiments were carried out with an optimized relative collision energy ranging from 18-25% and with an activation q value at 0.25, and the activation time at 30-50 ms. Mass spectra were accumulated in the profile mode, typically for 3-10 min for $MS^n$ (n = 2, 3, 4, and 5) spectra.

### Nomenclature

The abbreviations previously used for mycolic acids were incorporated for designation of TMM and TDM. For example, the 2-tetradecyl-3-hydroxy-eicosanoic acid containing a $C_{18}$-meromycolate chain and a $C_{16}$ α-branch was designated as 18:0/16:0-mycolic acid [6]. Accordingly, trehalose monomycolate (TMM) having 18:0/16:0-mycolic acid attached to 6 (or 6') of the trehalose core, a 6-(2-tetradecyl-3-hydroxyeicosanoyl) trehalose is designated as 18:0/16:0-TMM. The (6)-2-tetradecyl-3-hydroxyeicosanyl (6')-2-dodecyl-3-hydroxy-docosanyl-trehalose, which is a trehalose dimycolate (TDM) containing a 18:0/16:0- and a 20:0/14:0-mycolyl groups at 6 and 6', respectively, is designated as (18:0/16:0-20:0/14:0)-TDM. There is no distinction among isomers in which the mycolyl groups attached to 6 or 6' are exchanged. Hence, for example, the structures of (18:0/16:0-20:0/14:0)-TDM and (20:0/14:0-18:0/16:0)-TDM are not distinguishable. There is no distinction between a cyclic chain and a double bond in the structural assignment. Hence, the designation of 18:1/16:0-TMM, for example, only signifies the unsaturated state of the $C_{18}$-meromycolate chain is one.

## Results and Discussion

TMM and TDM formed intense $[M + Na]^+$ or $[M + Li]^+$ ions when subjected to ESI in the presence of $Na^+$ or $Li^+$ in positive-ion mode. In the negative-ion mode in the presence of anions such as $CH_3CO_2^-$ or $HCO_2^-$, TMM and TDM formed $[M + CH_3CO_2]^-$ or $[M + HCO_2]^-$ ions. As shown in Figure 1, the profiles of the ESI-MS spectra of the $[M + Na]^+$ ions of TMM (panel a) and TDM (panel b) isolated from *R. equi*, are nearly identical to those of the $[M + CH_3CO_2]^-$ ions of the corresponding TMM (see supplemental material, Figure S1a) and TDM (Figure S1b), respectively, indicating that the above adduct ions are applicable for profiling the TMM and TDM lipids. High-resolution mass measurements and structural characterization with LIT $MS^n$ indicated that ion series with various numbers of unsaturated bonds are present and many isomers were identified for all the TMM and TDM species in the mixture (see supplementary material Table S1 and Table S2). The $MS^n$ mass spectrometric approaches and the fragmentation processes leading to the structural characterization of TMM and TDM are described below.

### Characterization of TMM as $[M + Na]^+$ or $[M + Li]^+$ ions

Upon CID, the $[M + Na]^+$ ions of TMM yielded fragment ions corresponding to loss of the meromycolate chain (designated as a ion) (Scheme 1), loss of mycolic acid substituent as a ketene (b), as well as the sodiated ions of mycolylhexose (c) and of 6-acyl-glucose (mainly,

6-hexadecanoylglucose and 6-tetradecanoylglucose) (d). For example, the LIT $MS^2$ spectrum of the base ion of *m/z* 871 (Figure 2a) is dominated by the ions at *m/z* 709 (loss of 162) (c), together with the ion at *m/z* 691 (loss of 180) (e) arising from loss of a glucose residue. The spectrum also contained the ions at *m/z* 603, and 575 (a ions) arising from elimination of an eicosanal ($C_{19}H_{39}CHO$) and a doeicosanal ($C_{21}H_{43}CHO$) residues from the meromycolate chain, respectively. This is consistent with the observation of the ions at *m/z* 441 and 413 (d ions), representing the sodiated ions of hexadecanoylglucose (16:0-acyl glucose) and of tetradecanoylglucose (14:0-acyl glucose) arising from further loss of a glucose residue from *m/z* 603 and 575, respectively. The fragmentation processes were supported by the LIT $MS^3$ spectra of the ions of *m/z* 603 (871 → 603, Figure 2b) and of *m/z* 575 (871 → 575, Figure S2a). The former spectrum is dominated by the ion at *m/z* 441, together with the sodiated trehalose ions at *m/z* 365 and 347 arising from losses of 16:0-fatty acid as a ketene and as an acid, respectively, and the sodiated glucose ion at *m/z* 203. The latter spectrum (Figure S2a) contained the analogous ions at *m/z* 413, along with ions at *m/z* 365, 347 and 203. Further dissociation of the ion of *m/z* 709 (871 → 709, Figure 2c) gave rise to ions at *m/z* 649, 619 and 589, arising from bond cleavages across the hexose ring bearing the 6-mycolyl group (Scheme 1). The spectrum also contained the sodiated ions of 6-hexadecanoylglucose at *m/z* 441 and of 6-tetradecanoylglucose at *m/z* 413, arising from losses of an octadecanal ($C_{17}H_{35}CHO$) and an eicosanal ($C_{19}H_{39}CHO$) residues, respectively. The results indicate that the ion at *m/z* 709 represents both a sodiated ions of 6-(2-tetradecyl-3-hydroxy-eicosanyl)-glucose (18:0/16:0-Glc) and of 6-(2-dodecyl-3-hydroxy-docosanyl)-glucose (20:0/14:0-Glc). The results indicate that the ion of *m/z* 871 mainly represents a 6-(2-tetradecyl-3-hydroxy-doeicosanoyl) trehalose (18:0/16:0-TMM) and a 6-(2-dodecyl-3-hydroxy-tetraeicosanoyl) trehalose (20:0/14:0-TMM).

In Figure 2a, a minor ion at *m/z* 547 arising from loss of tetraeicosanal is also present. This ion together with the minor ion of *m/z* 385 representing a sodiated 12:0-acyl glucose cation seen in Figure 2c indicates that the ion of *m/z* 871 also represents a minor 22:0/12:0-TMM isomer. The $MS^3$ spectrum of the ion of *m/z* 547 (871 → 547, not shown) is dominated by the ion at *m/z* 385, consistent with the presence of the 22:0/12:0-TMM isomer.

The above structural assignment is further confirmed by LIT $MS^2$ on the corresponding $[M + Li]^+$ ion at *m/z* 855 (Figure S2b), which yielded prominent ion at *m/z* 693 (loss of 162), together with ions at *m/z* 675 (loss of 180), 603, 587, 559, 531 and 349. These ions are 16 Da lighter than those seen in Figure 2a, and are the lithiated analogs arising from the same fragmentation processes. The profile of the $MS^3$ spectrum of the ion at *m/z* 693 (855 → 693; Figure S2c) is similar to that seen in Figure 2c, and the ions at *m/z* 675, 633, 603, 573, 425, 397 and 369 are also16 Da lighter than the sodiated analogs. The results further support the fragmentation processes (Scheme 1) and the assignment of the major isomers of 18:0/16:0-TMM and 20:0/14:0-TMM and the minor 22:0/12:0-TMM isomer.

Structural characterization of the $[M + Na]^+$ ions of TMM with unsaturated bond(s) is exemplified by $MS^n$ on the ion of *m/z* 1009, which consisted of one unsaturated bond (Table S1). The $MS^2$ spectrum of the ion of *m/z* 1009 (see supplemental material, Figure S3a) contained the prominent ion at *m/z* 847 (loss of 162), and the ion of *m/z* 829 (loss of 180), arising from loss of a glucose residue. The ions at *m/z* 603 and 575 arise from loss of unsaturated octaeicosenal and tricosenal, respectively, indicating that the unsaturated bond is located at the $C_{27}$- or $C_{29}$-meromycolate chain. The ion at *m/z* 603 was formed together with the ions at *m/z* 441 (loss of 162) and 423 (loss of 180), arising from the 28:1/16:0-TMM isomer; while the ions of *m/z* 575, 413 and 395 arose from the analogous losses from the 30:1/14:0-TMM isomer. The assignments were further supported by $MS^3$ on the ion of *m/z* 847 (1009 → 847) (Figure S3b), which yielded ions at *m/z* 829, 787, 757, and 727, arising from cleavages of the glucose ring, and the ions at *m/z* 443 and *m/z* 413, arising from further

losses of $C_{27}H_{53}CHO$ and $C_{29}H_{57}CHO$, respectively. The results support that the ion of *m/z* 1009 represents both a 28:1/16:0-TMM and 30:1/14:0-TMM isomers, which possess unsaturated meromycolate chain.

The sodiated TMM ions containing two unsaturated bonds were seen at *m/z* 923, 951, 979, 1007, and 1035 (Table S1). The profiles of the $MS^2$ spectrum of the ion at *m/z* 979 (Figure S3c) and of the $MS^3$ spectrum of the ion at *m/z* 817 (979 → 817; Figure S3d) are similar to those seen in Figure S3a and S3b. However, the major sodiated acylglucose ion is seen at *m/z* 439 (Figure S3d), representing a sodiated 16:1-acyl glucose ion arising from loss of $C_{25}H_{49}CHO$. The results indicate that the ion represents mainly a 26:1/16:1-TMM isomer, which possesses an unsaturated bond in both the meromycolate and α-alkyl chains. This structure assignment is also consistent with the observation of the ion at *m/z* 601 in Figure S3c, arising from loss of $C_{25}H_{49}CHO$ residue. In Figure S3d, minor ions at *m/z* 441 (16:0-acyl glucose) and 413 (14:0-acyl glucose) arising from losses of $C_{25}H_{47}CHO$ and $C_{27}H_{51}CHO$, respectively, are also present. The presence of these ions are consistent with the observation of the ions of of *m/z* 575 and 603, arising from loss of $C_{27}H_{51}CHO$ and $C_{25}H_{47}CHO$, respectively. The results indicate the presence of the minor isomers of 26:2/16:0-TMM and 28:2/14:0-TMM, of which the unsaturated bonds are located at the meromycolate chain and the α-alkyl group is saturated.

The above results demonstrated the utility of $MS^n$ in the identification of TMM molecules including the information of distribution of double bond(s) in α-alkyl or meromycolate chain.

## Characterization of TDM as $[M + Na]^+$ ions

The isomeric structures grew immensely for TDM because the presence of an additional mycolyl group as compared to TMM. Scheme 2 summarized the multiple-stage mass spectrometric approaches toward the structural elucidation of TDM as the $[M + Na]^+$ ions in this study. The LIT $MS^n$ on the $[M + Na]^+$ ion also yielded fragment ions corresponding to loss of the meromycolate chain (designated as a and a‘ ion) (Scheme 3), loss of the mycolic acid substituent as a ketene (b and b’), together with the sodiated ions of mycolylhexose (c and c’) and of acyl-hexose (mainly hexadecanoyl- or tetradecanoylhexose) (d and d’) similar to those seen for TMM.

For example, the LIT $MS^2$ spectrum of the ion at *m/z* 1378 (Figure 3a) contained major ions at *m/z* 1109 (a) and 1081 (a’) arising from loss of an octadecanal and an eicosanal residues from the meromycolate chain, respectively, along with the ions at *m/z* 899 (b) and 871 ( b’), arising from losses of 2-dodecyl-3-hydroxy-eicosanoic acid and 2-tetradecyl-3-hydroxy-eicosanoic acid (or 2-dodecyl-3-hydroxy-docosanoic acid) as ketenes, respectively; while the ions at *m/z* 709, 737, and 681 represent the sodiated ions of mycolylglucose. Further dissociation of the ion of *m/z* 1109 (1378 → 1109, Figure 3b) led to ions at *m/z* 927, 899 (b) and 871 (b’), which are equivalent to the sodiated TMM ions arising from further losses of 12:0-, 14:0-, and 16:0-fatty acyl acids as ketenes, respectively. The ions at *m/z* 841, 813 and 785 arise from elimination of an octadecanal ($C_{17}H_{35}CHO$), eicosanal ($C_{19}H_{39}CHO$) or a docosanal ($C_{21}H_{43}CHO$) residue from the remaining mycolyl group and represent mainly the sodiated 6,6′-dihexadecanoyl-trehalose, 6-hexadecanoly-6′-tetradecanoyl-trehalose, and 6,6′-ditetradecanoyl-trehalose, respectively. These diacyl-trehalose structures were confirmed by observation of the prominent acylglucose ions at *m/z* 441 and 413 in the $MS^4$ spectrum of the ion of *m/z* 813 (1378 → 1109 → 813; Figure 3c) and the ions of *m/z* 441 and 413, respectively, in the $MS^4$ spectra of the ions of *m/z* 841 and of 785 (data not shown). The ions at *m/z* 709, 737, and 765 (Figure 3b) arose mainly from further loss of a glucose residue (loss of 162 Da) from ions of *m/z* 871, 899 and 927, respectively. Both the $MS^3$ spectrum of the ion of *m/z* 871 (1378 → 871) and the $MS^4$ spectrum of the ion of *m/z* 871

(1378 → 1109 → 871) (not shown) are identical to the $MS^2$ spectrum of the TMM ion of *m/z* 871 shown earlier (Figure 2a), consistent with the notion that the ions of *m/z* 871, 927 and 899 seen in Figure 3b are equivalent to the sodiated TMM ions.

The LIT $MS^4$ spectrum of the ion of *m/z* 709 (1378 → 1109 → 709) and the $MS^5$ spectrum of the ion at *m/z* 709 (1378 → 1109 → 871 → 709) (not shown) are similar to the spectrum shown in Figure 2c, in agreement with the notion that the ion of *m/z* 709 arises from further loss of a glucose residue from *m/z* 871. The results indicate that *m/z* 709 represents both a sodiated ions of 20:0/14:0-Glc and of 18:0/16:0-Glc. The combined structural information shows that the ion at *m/z* 1378 mainly represents a (6)-2-tetradecyl-3-hydroxy-eicosanyl (6')-2-dodecyl-3-hydroxy-docosanyl-trehalose ((18:0/16:0-20:0/14:0)-TDM) (i.e., $R_1$ = $C_{17}H_{35}$, $R_2$ = $C_{14}H_{29}$, $R_1'$ = $C_{19}H_{39}$, $R_2'$ = $C_{12}H_{25}$) and a (6)-2-tetradecyl-3-hydroxy-eicosanyl, (6')-2-tetradecyl-3-hydroxy-eicosanyl-trehalose (18:0/16:0-18:0/16:0)-TDM (i.e., $R_1$ = $C_{17}H_{35}$, $R_2$ = $C_{14}H_{29}$, $R_1'$ = $C_{17}H_{35}$, $R_2'$ = $C_{14}H_{29}$) (Scheme 3).

The LIT $MS^5$ spectrum of the ion of *m/z* 737 (1378 → 1109 → 899 → 737) (Figure 3d) contained the ions at *m/z* 719, 677, 647 and 617, arising from bond cleavages across the hexose ring bearing the 6-mycolyl group, and the sodiated acylglucose ions at *m/z* 441 and 413, indicating that the *m/z* 737 ion represents both a sodiated 2-tetradecyl-3-hydroxy-docosanyl-6'-glucose (20:0/16:0-Glc) and dodecyl-3-hydroxy-tetracosanyl-6'-glucose (22:0/14:0-Glc). The combined information demonstrates that the ion of *m/z* 1378 also represents a (6)-2-dodecyl-3-hydroxy-eicosanyl, (6')-2-tetradecyl-3-hydroxy-docosanyltrehalose ((18:0/14:0-20:0/16:0)-TDM) and a minor (18:0/14:0-22:0/14:0)-TDM isomer. The LIT $MS^5$ spectrum of the ion of *m/z* 765 (1378 → 1109 → 927 → 765; not shown) contained ion at *m/z* 441, suggesting the presence of a minor isomer of (18:0/12:0-22:0/16:0)-TDM.

In Figure 3b, the ion at *m/z* 947 corresponding to loss of a glucose residue (loss of 162 Da) is also present, indicating a glucose residue had been cleaved from *m/z* 1109. High resolution mass spectrometry confirms that this loss is indeed a hexose residue (Table S3). The $MS^3$ spectrum of *m/z* 947 (1378 → 1109 → 947; data not shown) contained ions at *m/z* 651 (loss of $C_{19}H_{39}CHO$) and 691 (loss of 16:0 fatty acid), indicating that the ion of *m/z* 947 contains a 20:0/14:0-mycolyl and a 16:0-fatty acyl groups attached to a glucose residue. Similar internal loss of glucose also resulted in the formation of the ions of *m/z* 919 (Figure 3e), and *m/z* 891 (Figure 3f) seen in the later analogous $MS^3$ spectra; and this unusual hexose loss has been previously reported [23-27].

Further dissociation of the ion of *m/z* 1081 (1378 → 1081, Figure 3e) gave rise to ions at *m/z* 899, 871 and 843, arising from losses of 12:0-, 14:0-, and 16:0-fatty acyl ketenes, respectively. The ion at *m/z* 919 arose from internal loss of glucose (loss of 162 Da) as seen earlier; while ions at *m/z* 813 and 785 arose from further losses of a $C_{17}H_{35}CHO$ or a $C_{19}H_{39}CHO$ residue from meromycolate chain of the remaining mycolyl group. The spectrum also contained ions at *m/z* 737 and 709, and 681, arising from further loss of a glucose residue (loss of 162 Da) from ions at *m/z* 899, 871 and 843, respectively. The LIT $MS^5$ spectra of the ions of *m/z* 709 (1378 → 1081 → 871 → 709) and of 737 (1378 → 1081 → 899 → 737) (data not shown) are identical to those shown earlier (Figure 2c and Figure 3d). The combined structural information from the $MS^3$ spectrum of *m/z* 1081, $MS^4$ spectrum of *m/z* 899 and $MS^5$ spectrum of *m/z* 737 led to assignment of the (20:0/12:0-20:0/16:0)-TDM isomer (1081-899-737 series); while structural information from $MS^n$ on 1081, 871, and 709 (1081-871-709 series) gave assignment of isomers of (20:0/14:0-18:0/16:0)-TDM and (20:0/14:0-20:0/14:0)-TDM. Similarly, the LIT $MS^5$ spectrum of the ion of *m/z* 681 (1378 → 1081 → 843 → 681) (data not shown) contained ions at *m/z* 651, 621, 591 and 561, arising from cleavages of the glucose ring and ions at *m/z*

413 (sodiated 14:0-acyl glucose) and 385 (sodiated 12:0-acyl glucose), from losses of $C_{17}H_{35}CHO$ and $C_{19}H_{39}CHO$, suggesting that the ion at *m/z* 681 represents both a sodiated 18:0/14:0-Glc and 20:0/12:0-Glc ions. These structural information (1081-841-681 series) confirms the formerly assigned (20:0/16:0-18:0/14:0)-TDM and 20:0/16:0-20:0/12:0)-TDM structures.

$MS^3$ on the ion of *m/z* 1053 (1378 → 1053) (Figure 3f) yielded ions at *m/z* 871, 843, and 815 by losses of 12:0-, 14:0- and 16:0-fatty acyl ketene, respectively; and the ions at *m/z* 709, 681, and 653 arose from further loss of a glucose residue. The ions of *m/z* 709 and 681 consists of the same structures defined earlier and the ion at *m/z* 653 represents a sodiated 18:0/12:0-Glc. These structural information identifies (22:0/12:0-18:0/16:0)-TDM, (22:0/12:0-20:0/14:0)-TDM (1053-871-709 series); (22:0/14:0-18:0/14:0)-TDM, (22:0/14:0-20:0/12:0)-TDM (1053-843-681 series); and the (22:0/16:0-18:0/12:0)-TDM (1053-815-653 series). The presence of the (22:0/14:0-18:0/14:0)-TDM and (22:0/16:0-18:0/12:0)-TDM isomers are in agreement with the assignment of 18:0/14:0-22:0/14:0-and 18:0/12:0-22:0/16:0-TDM isomers as described earlier.

Further dissociation of the ion of *m/z* 1025 (1378 → 1025) (data not shown) gave rise to ions at *m/z* 843 and 815, which yielded ions of *m/z* 681 and 653, respectively, by loss of glucose (loss of 162 Da). Structural information from $MS^5$ on the ions of *m/z* 681 (1378 → 1025 → 843 → 681) and of *m/z* 653 (1378 → 1025 → 815 → 653) (not shown) readily gave structural assignments of (24:0/12:0-18:0/14:0)-, and (24:0/14:0-18:0/12:0)-TDM isomers.

Collectively, the above results demonstrate that the ion at *m/z* 1378 represents major isomers of (18:0/16:0-20:0/14:0)-, (18:0/16:0-18:0/16:0)-, and (20:0/14:0-20:0/14:0)-TDM as well as minor (18:0/14:0-20:0/16:0)-, (18:0/12:0-22:0/16:0)-, (18:0/14:0-22:0/14:0)-, (20:0/12:0-20:0/16:0)-, (22:0/12:0-18:0/16:0)-, (22:0/12:0-20:0/14:0)-, (22:0/14:0-18:0/14:0)-, (22:0/14:0-20:0/12:0)-, (24:0/12:0-18:0/14:0)-, and of (24:0/14:0-18:0/12:0)-TDM isomers.

The number of the isomeric structures became even more enormous for the species that possesses unsaturated bond(s). For example, more than 24 isomers were identified for the ion of m/z 1404 (Table S2). These isomers varied in the chain lengths of the meromycolate and α-branch residues, as well as in the location of unsaturated bond at meromycolate or α-branch. The mass spectra obtained by LIT $MS^n$ and the description of structural assignments of the ion of *m/z* 1404 are appendixed in supplementary material (Figure S4 and the text entitled “Structural assignments of the $[M + Na]^+$ ion of m/z 1404 by LIT $MS^n$”).

## Characterization of TMM as the $[M + CH_3CO_2]^-$ or $[M + HCO_2]^-$ ions

The $[M + CH_3CO_2]^-$ adduct ion of TMM at *m/z* 907 (Figure S1a) and the $[M + HCO_2]^-$ adduct ion at *m/z* 893 correspond to the $[M + Na]^+$ ion at *m/z* 871 (Figure 1a). The $MS^2$ spectra of the ions at *m/z* 907 (Figure 4a) and at *m/z* 893 (data not shown) are nearly identical and contained the ion at *m/z* 847 representing a $[M - H]^-$ ion arising from loss of the $CH_3CO_2H$ or $HCO_2H$ component. The spectrum (Figure 4a) is dominated by the 16:0- and 14:0-acyltrehalose anions at *m/z* 579 and 551, arising from loss of $CH_3CO_2H$, followed by elimination of a octadecanal or an eicosanal residue; while the ions at *m/z* 341 and 323 represent the deprotonated trehalose and $[trehalose - H_2O]^-$ anions, respectively. The results are in agreement with the earlier structure assignment of 18:0/16:0-TMM and 20:0/14:0-TMM deduced from the $[M + Na]^+$ ion of *m/z* 907. The ion at *m/z* 829 arose from loss of $H_2O$ and gave rise to the dehydrated mycolic acid anion of *m/z* 505 by loss of trehalose (loss as [trehalose – $H_2O$]; 324 Da). The $MS^3$ spectrum of the ion at *m/z* 829 (907 → 829; Figure 4b) contained the ions at *m/z* 505 and 323, consistent with the fragmentation process. Further dissociation of the ion of *m/z* 579 (907 → 579; Figure 4c) yielded ions at *m/z* 323

([trehalose – $H_2O$]$^-$) and *m/z* 255 (16:0-carboxylate anion), confirming that the ion at *m/z* 579 represents a deprotonated 16:0-acyltrehalose anion. The results are in agreement with the assignment of the 18:0/16:0-TMM isomer. The $MS^3$ spectrum of the ion at *m/z* 551 (907 → 551; data not shown) contained ions at *m/z* 323 ([trehalose – $H_2O$]$^-$) and 227 (14:0-carboxylate anion), consistent with the presence of 20:0/14:0-TMM.. In Figure 4a, a minor ion at *m/z* 523 is also present. The $MS^3$ spectrum of the ion of *m/z* 523 (907 → 523; not shown), contained ions at *m/z* 323 ([trehalose – $H_2O$]$^-$) and 199 (12:0-carboxylate anion), suggesting the presence of the minor 22:0/12:0-TMM isomer. The observation of the 16:0- and 14:0-carboxylate anions at *m/z* 255 and 227 as seen in the in the $MS^2$ (Figure 4a) and $MS^3$ (Figure 4c) spectra provides useful information for assignment of the α-branch. In contrast, the fragment ions of sodiated acylglucose (i.e., ions at *m/z* 441 and 413) that defined the α-branch are of low abundance in the $MS^n$ spectra of the $[M + Na]^+$ ions of TMM and TDM..

## Characterization of TDM as $[M + CH_3CO_2]^-$ ions (the related figures were presented as supplementary material Figure S5)

The LIT $MS^2$ spectra of the $[M + CH_3CO_2]^-$ ion at *m/z* 1414.1 (Figure S5a) and of the $[M + HCO_2]^-$ ion at *m/z* 1399.9 (not shown) are nearly identical. The spectrum (Figure S5a) contained ions at *m/z* 1354 arising from loss of $CH_3CO_2H$, at *m/z* 1085 arising from further loss of an octadecanal residue, and at *m/z* 1067 (1085 – $H_2O$) arising from additional loss of $H_2O$. The ions at *m/z* 1057 (1354 – $C_{19}H_{39}CHO$) and 1039 (1057 – $H_2O$) arose from the analogous losses involving an eicosadecanal residue. Further dissociation of the ion at *m/z* 1067 (1414 → 1067; Figure S5b) yielded ions at *m/z* 839 (loss of 14:0-fatty acid) and 811 (loss of 16:0-fatty acid), which gave rise to ions of *m/z* 533 and 505, respectively, by loss of the sugar moiety (loss as [trehalose – 2 $H_2O$]). These fragmentation processes were supported by $MS^4$ on the ion at *m/z* 839 (not shown), which yielded prominent ion at *m/z* 533, and by $MS^4$ on the ion at *m/z* 811 (1414 → 1067 → 811; Figure S5c), which is dominated by the ion at *m/z* 505. The results indicate the molecule consists of both a 18:0/14:0- and a 18:0/16:0-mycolyl groups. The ion of *m/z* 533 (551 – $H_2O$) represents a deprotonated anion of the dehydrated mycolic acid of *m/z* 551, which consists of both a 20/16:0- and a 22:0/14:0-mycolyl groups, as revealed by the $MS^3$ spectrum of the ion at *m/z* 551 (1414 → 551; data not shown). These results led to assignment of the structures of 18:0/14:0-20:0/16:0 and 18:0/14:0-22:0/14:0-TDM (1413-1067-839-533 series). The ion at *m/z* 505 (523 – $H_2O$) also represents a dehydrated mycolic acid anion of *m/z* 523 arising from 18:0/16:0- and 20:0/14:0-mycolic acids. These information gave assignment of 18:0/16:0-18:0/16-TDM and 18:0/16:0-20:0/14:0-TDM (1413-1067-811-505 series).

Further dissociation of the ion of *m/z* 1039 (1414 → 1039; Figure S5d) gave rise to ions at *m/z* 839, 811, and 783, arising from losses of 12:0-, 14:0-, and 16:0-fatty acids, respectively. These results indicate the molecule consists of 20:0/12:0-, 20:0/14:0-, and 20:0/16:0-mycolyl residues. The ions of *m/z* 533 (551- $H_2O$), 505 (523- $H_2O$), and 477 (495- $H_2O$) are the dehydrated mycolic acid anions arising from ions at *m/z* 839, 811, and 783, respectively, by loss of trehalose residue (loss of [trehalose – 2 $H_2O$]) (data not sown). These ions bear the similar isomeric structures as described earlier. These structural information gave assignment of 20:0/12:0-20:0/16:0 and 20:0/12:0-22:0/14:0-TDM (1414-1039-839-533 series), 20:0/14:0-18:0/16:0, and 20:0/14:0-20:0/14:0 (1414-1039-811-505 series), and 20:0/16:0-18:0/14:0, 20:0/16:0-16:0/16:0, and 20:0/16:0-20:0/12:0 (1414-1039-783-477 series). The observation of the 18:0/14:0-22:0/14:0-TDM and 20:0/12:0-22:0/14:0-TDM isomers is also consistent with the presence of the ion at *m/z* 1011 arising from elimination of the 22:0-meromycolate chain, in the $MS^2$ spectrum of *m/z* 1414 (Figure S5a). The observation of the 18:0/14:0-20:0/16:0-TDM (1414-1067-839-533 series) and 20:0/12:0-20:0/16:0-TDM (1414-1039-839-533 series) is also consistent with the

assignment of 20:0/16:0-20:0/12:0-TDM (1414-1039-783-477 series) and 20:0/16:0-18:0/14:0-TDM (1414-1039-783-477 series) as described earlier.

The ions at *m/z* 817, 789, 761, and 733 (Figure S5a) are the deprotonated 6,6'-diacyltrehalose anions arising from cleavages of the meroaldehyde residues on both the mycolyl groups. Further dissociation of ions of *m/z* 817, 789, 761 and 733 (data not shown) confirms that these ions represent deprotonated 6,6′-dihexadeanoyltrehalose, 6-hexadecanoyl-6′-tetradecanoyltrehalose, 6,6′-ditetradecanoyltrehalose (or 6-hexadecanoyl-6′-dodecanoyltrehalose), and 6-tetradecanoyl-6′-dodecanolytrehalose anions, respectively. The results are consistent with the notion that the α-branch of the mycolic acid substituents mainly consist of 16:0-, 14:0, and 12:0-fatty acid moieties. The structural assignments, again, are in agreement with earlier assignment derived from $MS^n$ on the corresponding $[M + Na]^+$ adduct ion of *m/z* 1378.

## Conclusions

We employed LIT $MS^n$ to define the structures of TMM and TDM in the cell envelope of *R. equi* without extensive chromatographic separation and chemical reaction steps. We unveiled the structural diversity of the *R. equi* TMMs and TDMs seen by the presence of whole array of the homologous masses (*m/z*) of which each mass contains numerous isomers arising from the variation and permutation of the mycolyl groups on both 6 and 6' positions of the trehalose core (Table S1 and S2), not including the positional isomers that may arise from the differences in the location of double bond(s) along the meromycolate or α-alkyl chain. Thus, hundreds of TMM and TDM structures are present in the cell envelope of *R. equi*. We also observed the structural similarity among the mycolic acid, TMM and TDM. For example, the most prominent TMM ion of *m/z* 871, and the most prominent TDM ion of *m/z* 1378 all contained mainly the 18:0/16:0- and 20:0/14:0-mycolyl residues. These findings are in accord with our previous report that 18:0/16:0- and 20:0/14:0-mycolic acids (seen at *m/z* 523 as $[M - H]^-$) are the most prominent mycolic acid found in *R. equi*. [6], and consistent with the notion that mycolic acids are released by enzymatic hydrolysis of TDM [7]. While ESI LIT $MS^n$ on the $[M + Na]^+$ (or $[M + Li]^+$ ) ions is readily applicable for structural identification and profiling TMM and TDM, negative-ion $MS^n$ on the $[M + CH_3CO_2]^-$ (or $[M + HCO_2]^-$) is also useful for structure assignments and profiling these complex lipids.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This research was supported by US Public Health Service Grants P41-RR-00954, R37-DK-34388, P60-DK-20579, P01-HL-57278, P30-DK56341 and Deutsche Forschungsgemeinschaft (SFB 670 [A.H.])

## References

1. Giguère S, Prescott JF. Clinical Manifestations, Diagnosis, Treatment and Prevention of Rhodococcus Equi Infections in Foals. Vet. Microbiol. 1997; 56:313–334. [PubMed: 9226845]
2. Prescott JF. Rhodococcus Equi: An Animal and Human Pathogen. Clin Microbiol Rev. 1991; 4:20–34. [PubMed: 2004346]
3. von Bargen K, Haas A. Molecular and Infection Biology of the Horse Pathogen Rhodococcus Equi. FEMS Microbiol Rev. 2009; 33:870–891. [PubMed: 19453748]
4. Sutcliffe IC. Cell Envelope Composition and Organisation in the Genus Rhodococcus. Antonie van Leeuwenhoek. 1998; 74:49–58. [PubMed: 10068788]

5. Sutcliffe I, Brown A, Dover L. The Rhodococcal Cell Envelope: Composition, Organisation and Biosynthesis. Journal. 2010; 16:29–71.
6. Hsu FF, Soehl K, Turk J, Haas A. Characterization of Mycolic Acids from the Pathogen Rhodococcus Equi by Tandem Mass Spectrometry with Electrospray Ionization. Anal. Biochem. 2011; 409:112–122. [PubMed: 20946862]
7. Ojha AK, Trivelli X, Guerardel Y, Kremer L, Hatfull GF. Enzymatic Hydrolysis of Trehalose Dimycolate Releases Free Mycolic Acids During Mycobacterial Growth in Biofilms. J. Biol. Chem. 2010; 285:17380–17389. [PubMed: 20375425]
8. Brennan PJ, Nikaido H. The Envelope of Mycobacteria. Annu. Rev. Biochem. 1995; 64:29–63. [PubMed: 7574484]
9. Lemaire G, Tenu JP, Petit JF, Lederer E. Natural and Synthetic Trehalose Diesters as Immunomodulators. Med Res Rev. 1986; 6:243–274. [PubMed: 3526051]
10. Rhoades ER, Geisel RE, Butcher BA, McDonough S, Russell DG. Cell Wall Lipids from Mycobacterium Bovis Bcg Are Inflammatory When Inoculated within a Gel Matrix: Characterization of a New Model of the Granulomatous Response to Mycobacterial Components. Tuberculosis. 2005; 85:159–176. [PubMed: 15850754]
11. Axelrod S, Oschkinat H, Enders J, Schlegel B, Brinkmann V, Kaufmann SH, Haas A, Schaible UE. Delay of Phagosome Maturation by a Mycobacterial Lipid Is Reversed by Nitric Oxide. Cell Microbiol. 2008; 10:1530–1545. [PubMed: 18363878]
12. Indrigo J, Hunter RL Jr. Actor JK. Cord Factor Trehalose 6,6′-Dimycolate (Tdm) Mediates Trafficking Events During Mycobacterial Infection of Murine Macrophages. Microbiology. 2003; 149:2049–2059. [PubMed: 12904545]
13. Spargo BJ, Crowe LM, Ioneda T, Beaman BL, Crowe JH. Cord Factor (Alpha,Alpha-Trehalose 6,6′-Dimycolate) Inhibits Fusion between Phospholipid Vesicles. Proc. Natl. Acad. Sci. U. S. A. 1991; 88:737–740. [PubMed: 1992465]
14. Geisel RE, Sakamoto K, Russell DG, Rhoades ER. In Vivo Activity of Released Cell Wall Lipids of Mycobacterium Bovis Bacillus Calmette-Guerin Is Due Principally to Trehalose Mycolates. J. Immunol. 2005; 174:5007–5015. [PubMed: 15814731]
15. Kretschmer A, Bock H, Wagner F. Chemical and Physical Characterization of Interfacial-Active Lipids from Rhodococcus Erythropolis Grown on N-Alkanes. Appl Environ Microbiol. 1982; 44:864–870. [PubMed: 16346110]
16. Asselineau J, Lederer E. Structure of the Mycolic Acids of Mycobacteria. Nature. 1950; 166:782–783. [PubMed: 14780245]
17. Asselineau J, Lederer E. Sur La Constitution Chimique Des Acides Mycoliques De Deux Souches Humaines Virulentes De Mycobacterium Tuberculosis. Biochimica et Biophysica Acta. 1951; 7:126–145. [PubMed: 14848079]
18. Wong MY, Steck PA, Gray GR. The Major Mycolic Acids of Mycobacterium Smegmatis. Characterization of Their Homologous Series. J. Biol. Chem. 1979; 254:5734–5740. [PubMed: 447680]
19. Rapp P, Bock H, Wray V, Wagner F. Formation, Isolation and Characterization of Trehalose Dimycolates from Rhodococcus Erythropolis Grown on N-Alkanes. J. General Microbiol. 1979; 115:491–503.
20. Fujita Y, Naka T, Doi T, Yano I. Direct Molecular Mass Determination of Trehalose Monomycolate from 11 Species of Mycobacteria by Maldi-Tof Mass Spectrometry. Microbiology. 2005; 151:1443–1452. [PubMed: 15870454]
21. Fujita Y, Naka T, McNeil MR, Yano I. Intact Molecular Characterization of Cord Factor (Trehalose 6,6′-Dimycolate) from Nine Species of Mycobacteria by Maldi-Tof Mass Spectrometry. Microbiology. 2005; 151:3403–3416. [PubMed: 16207922]
22. De la Pena-Moctezuma A, Prescott JF. Association with Hela Cells by Rhodococcus Equi with and without the Virulence Plasmid. Vet. Microbiol. 1995; 46:383–392. [PubMed: 8560735]
23. Harvey DJ, Mattu TS, Wormald MR, Royle L, Dwek RA, Rudd PM. Internal Residue Loss" Rearrangement During the Fragmentation of Carbohydrates Derivatized at the Reducing Terminus. Anal. Chem. 2002; 74:734–740. [PubMed: 11866052]

24. Ma YL, Vedernikova I, Van den Heuvel H, Claeys M. Internal Glucose Residue Loss in Protonated O-Diglycosyl Flavonoids Upon Low-Energy Collision-Induced Dissociation. J. Am. Soc. Mass Spectrom. 2000; 11:136–144. [PubMed: 10689666]

25. Brull LP, Kovacik V, Thomas-Oates J, Heerma W, Haverkamp J. Sodium-Cationized Oligosaccharides Do Not Appear to Undergo “Internal Residue Loss” Rearrangement Processes on Tandem Mass Spectrometry. Rapid Commun. Mass Spectrom. 1998; 12:1520–1532. [PubMed: 9796537]

26. Warrack BM, Hail ME, Triolo A, Animati F, Seraglia R, Traldi P. Observation of Internal Monosaccharide Losses in the Collisionally Activated Dissociation Mass Spectra of Anthracycline Aminodisaccharides. J. Am. Soc. Mass Spectrom. 1998; 9:710–715.

27. Hsu FF, Turk J. Studies on Sulfatides by Quadrupole Ion-Trap Mass Spectrometry with Electrospray Ionization: Structural Characterization and the Fragmentation Processes That Include an Unusual Internal Galactose Residue Loss and the Classical Charge-Remote Fragmentation. J. Am. Soc. Mass Spectrom. 2004; 15:536–546. [PubMed: 15047058]

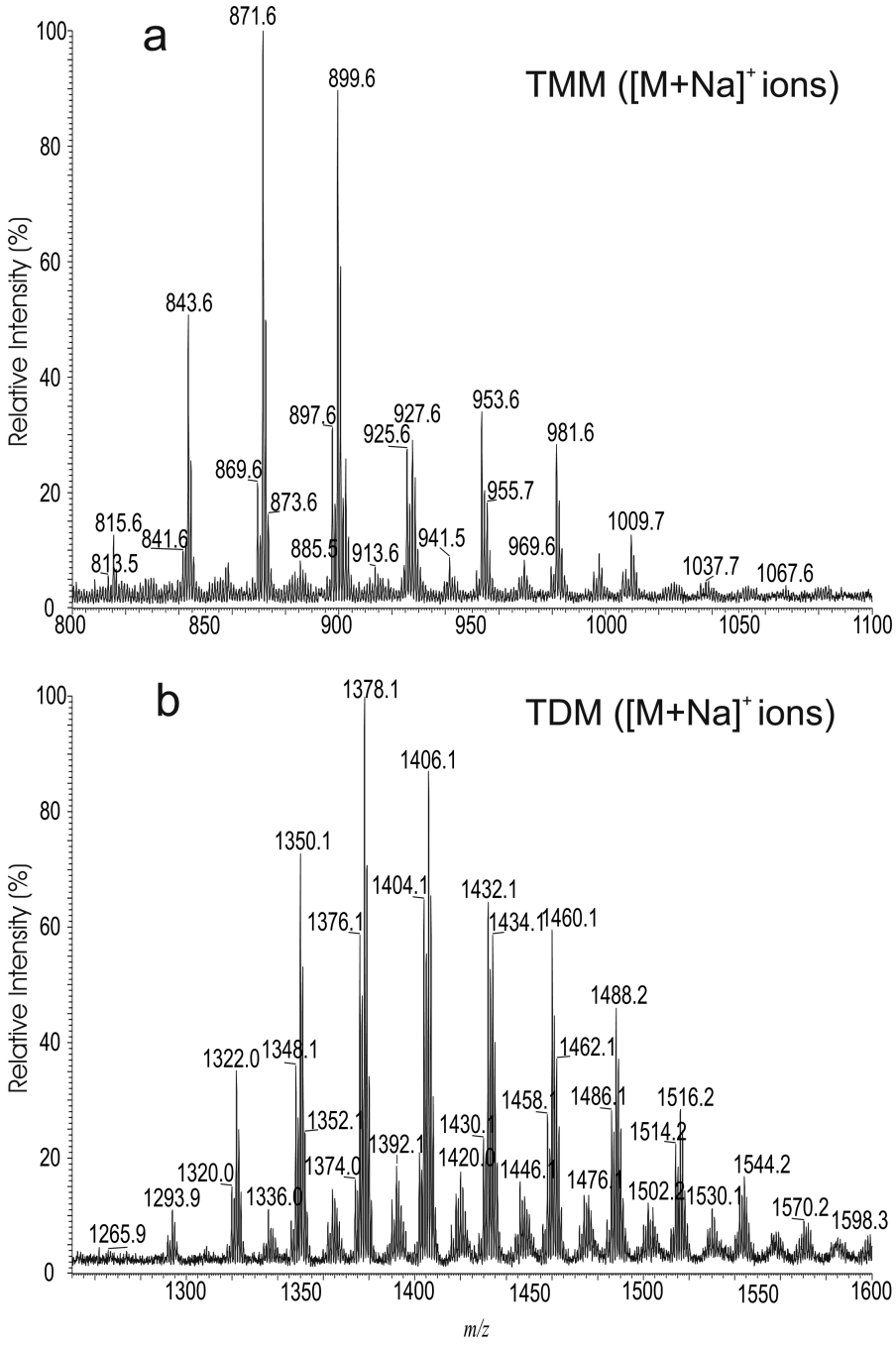



**Figure 1.**

The ESI-MS spectra of the $[M + Na]^+$ ions of TMM (a) and of TDM (b). These profiles are similar to their $[M + CH_3CO_2]^-$ adduct ions (see supplemental material Figure S1).

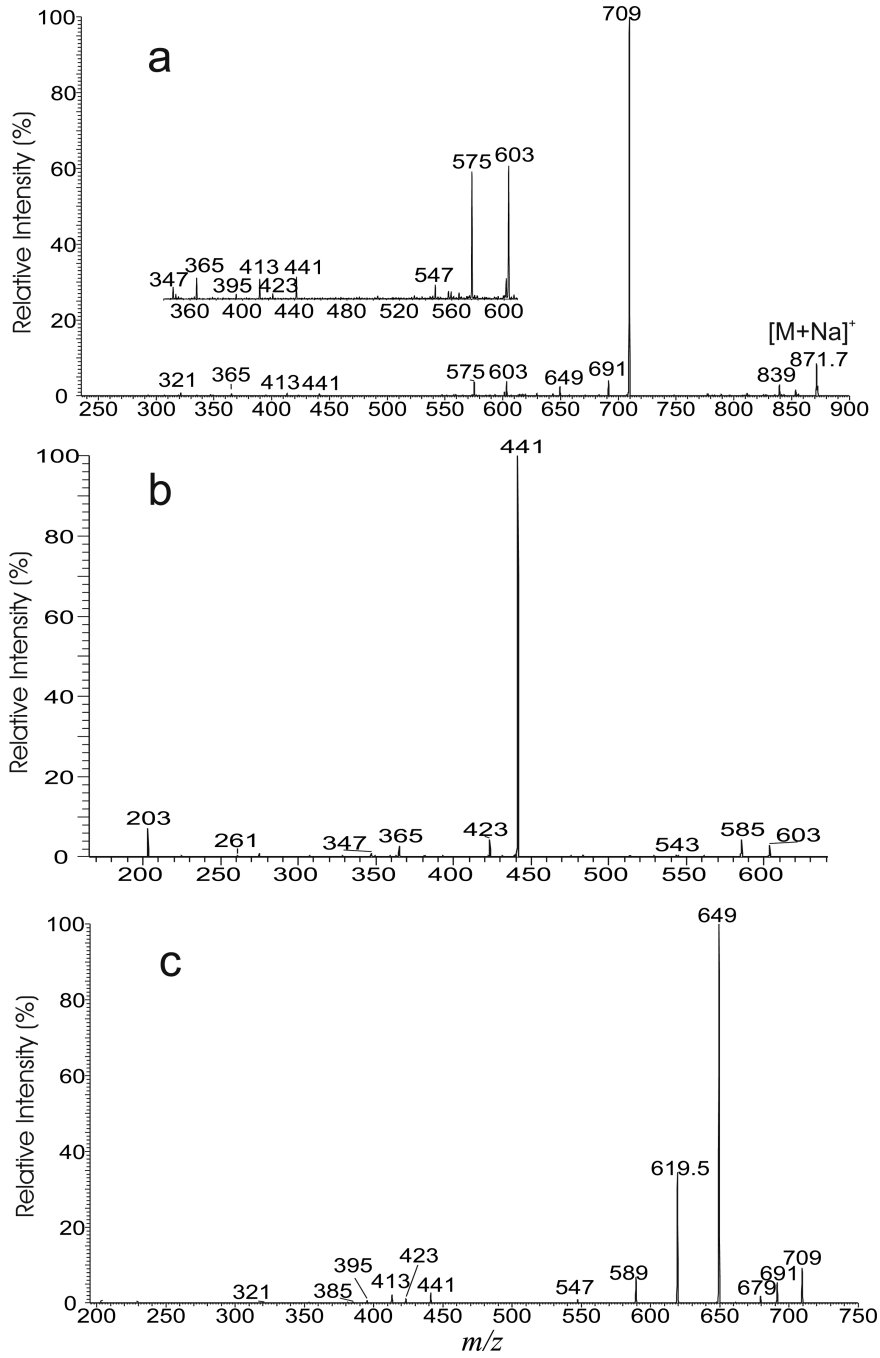



**Figure 2.**

The LIT $MS^2$ spectrum of the $[M + Na]^+$ ion of TMM at *m/z* 871 (a), its $MS^3$ spectra at *m/z* 603 (871 → 603) (b), and at *m/z* 709 (871 → 709) (c).

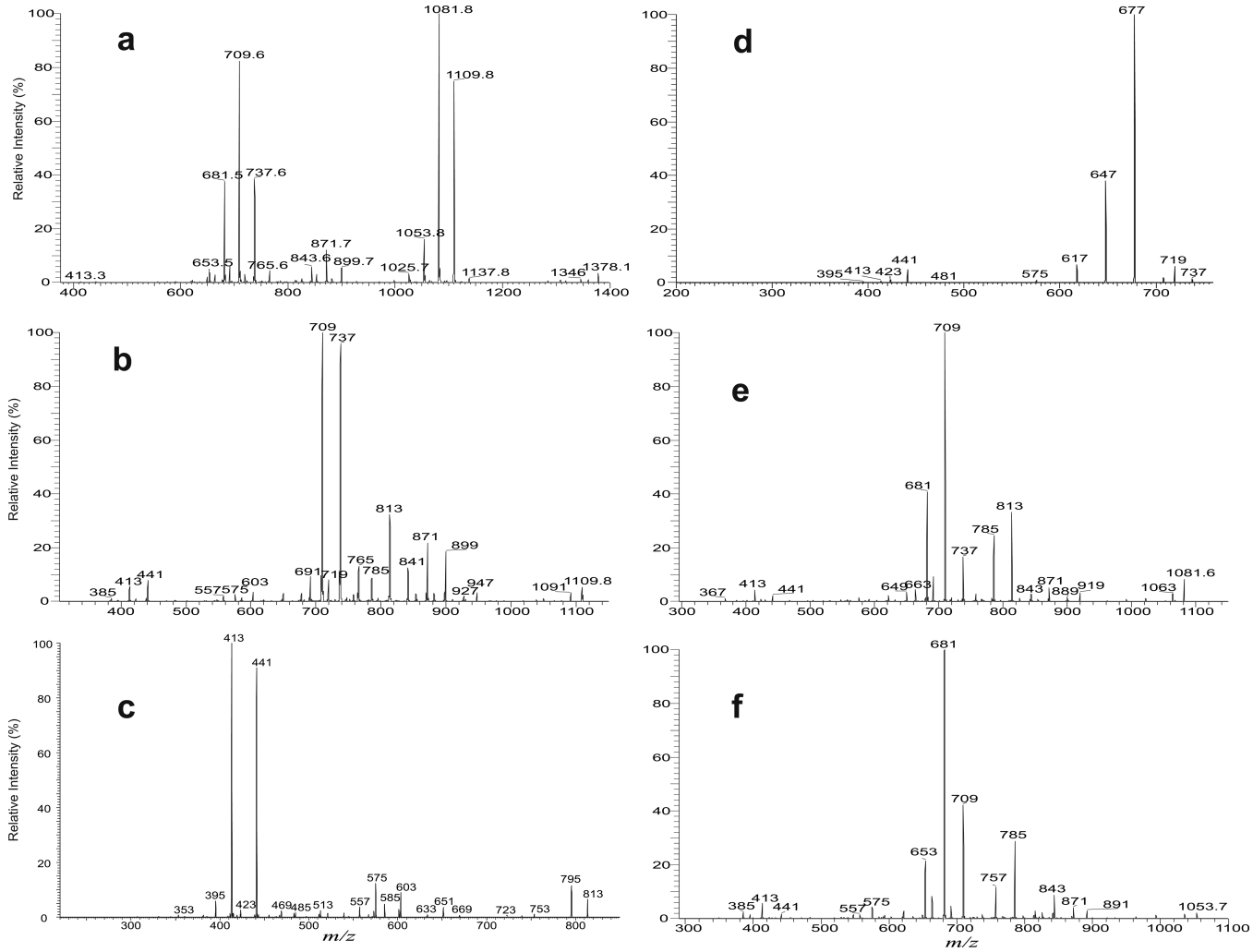



**Figure 3.**

The LIT $MS^2$ spectrum of the $[M + Na]^+$ ion of TDM at *m/z* 1378 (a), its $MS^3$ spectra at *m/z* 1109 (1378 → 1109) (b), its $MS^4$ spectra at *m/z* 813 (1378 → 1109 → 813) (c), and at *m/z* 737 (1378 → 1109 → 737) (d); the $MS^3$ spectra of the ions of *m/z* 1081 (1378 → 1081) (e), and of *m/z* 1053 (1378 → 1053) (f)

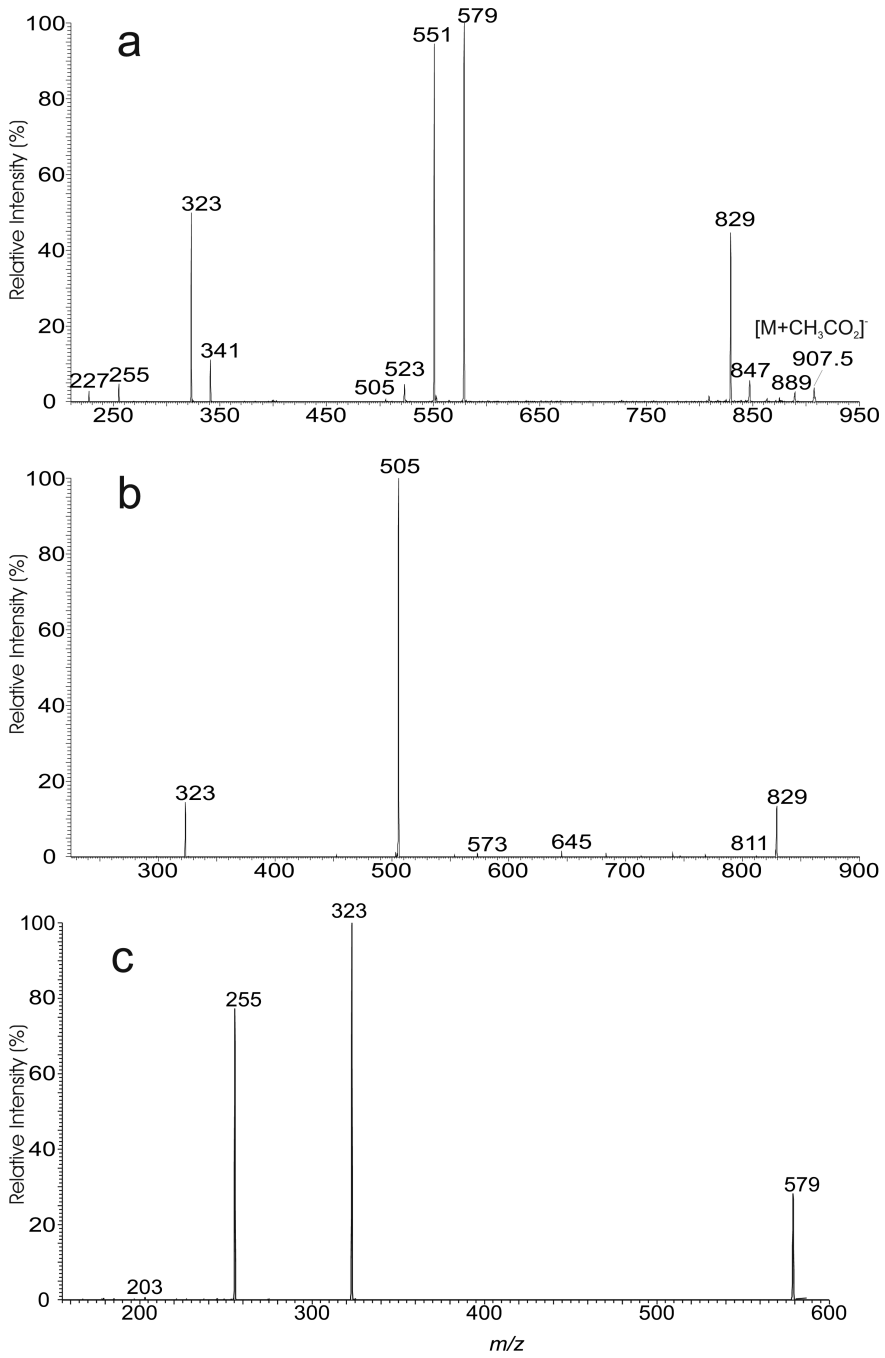
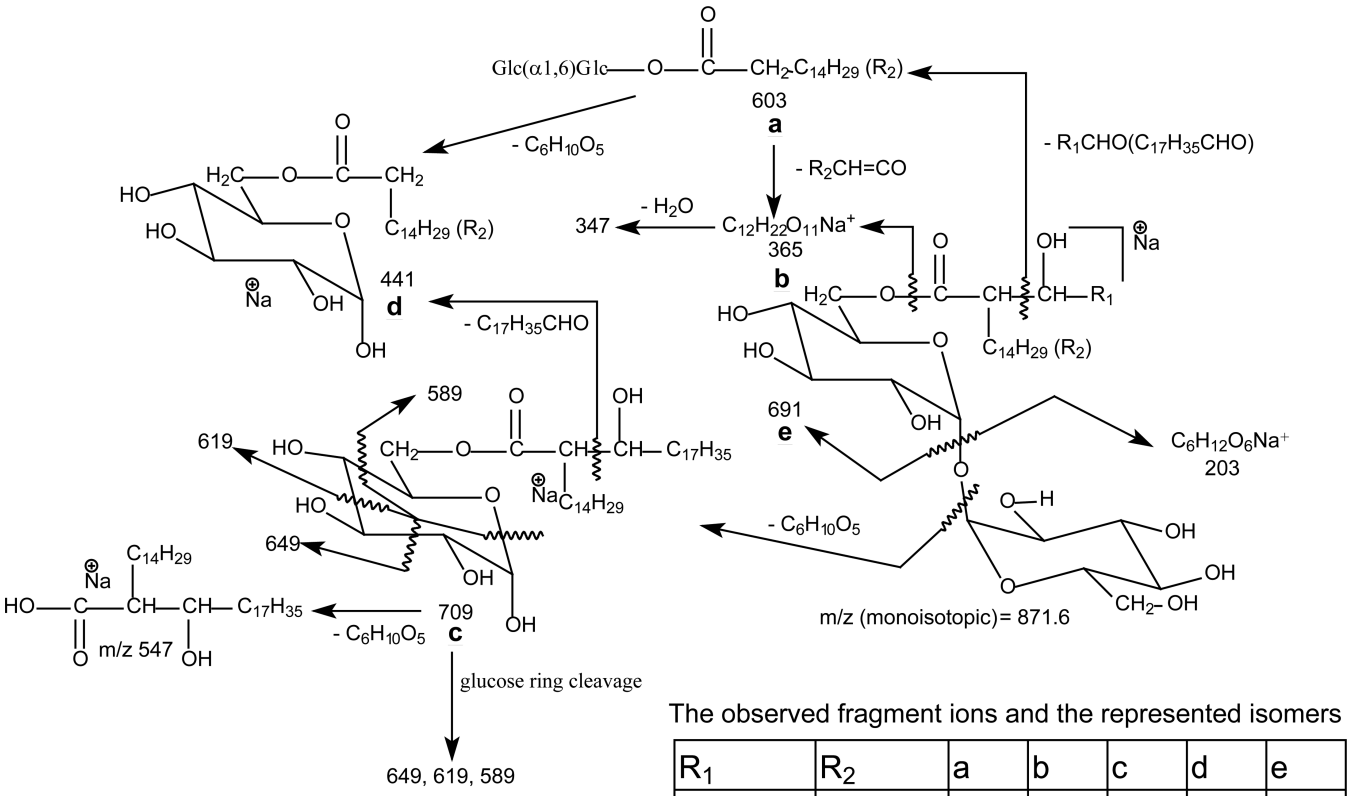



**Figure 4.**

The LIT $MS^2$ spectra of the $[M + CH_3CO_2]^-$ ion of TMM at *m/z* 907 (a), its $MS^3$ spectra at *m/z* 829 (907 → 829) (b), and at *m/z* 579 (907 → 579) (c).

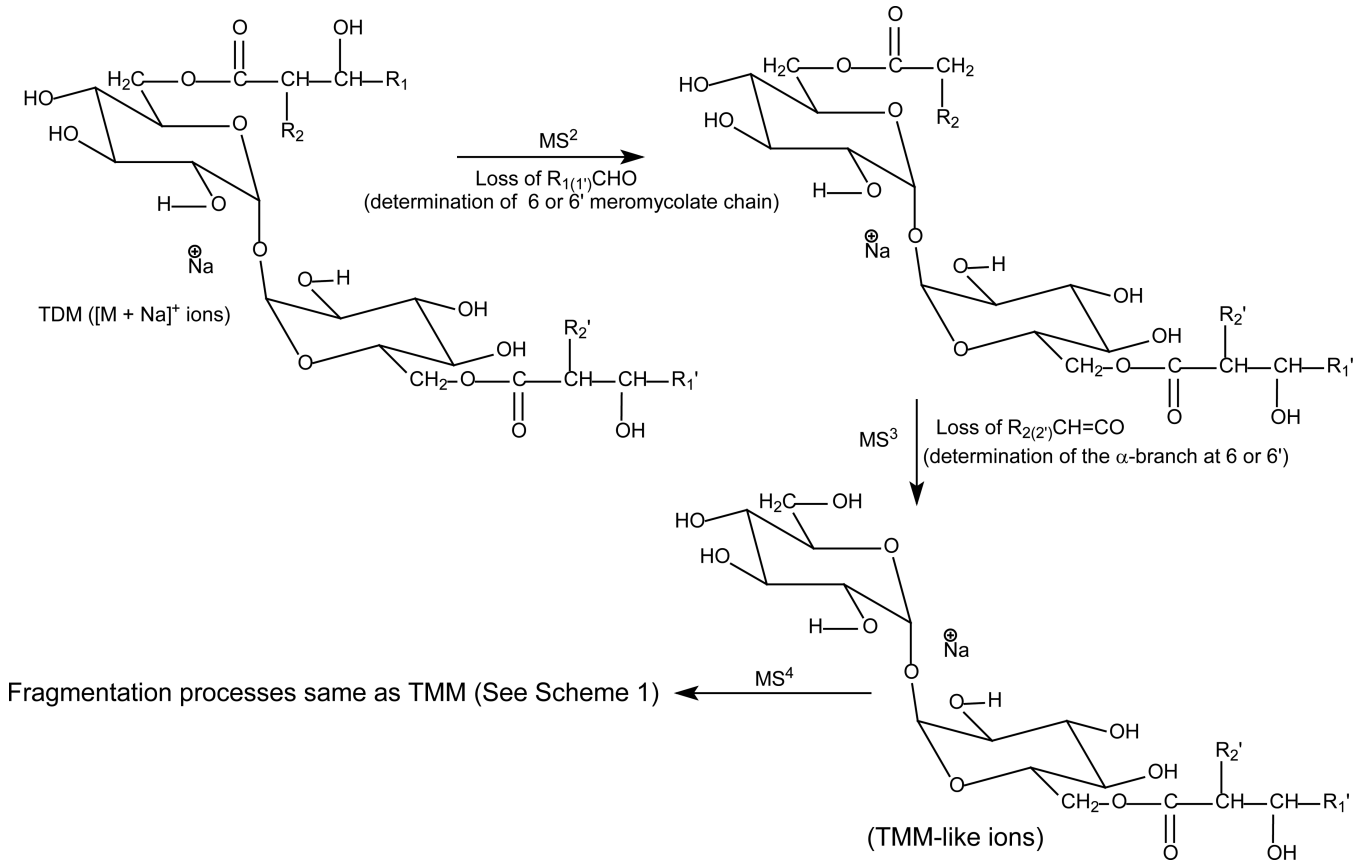



The observed fragment ions and the represented isomers

| $R_1$ | $R_2$ | a | b | c | d | e |
|---|---|---|---|---|---|---|
| $C_{17}H_{35}$ | $C_{14}H_{29}$ | 603 | 365 | 709 | 441 | 691 |
| $C_{19}H_{39}$ | $C_{12}H_{25}$ | 575 | 365 | 709 | 413 | 691 |
| $C_{21}H_{43}$ | $C_{10}H_{21}$ | 547 | 365 | 709 | 385 | 691 |

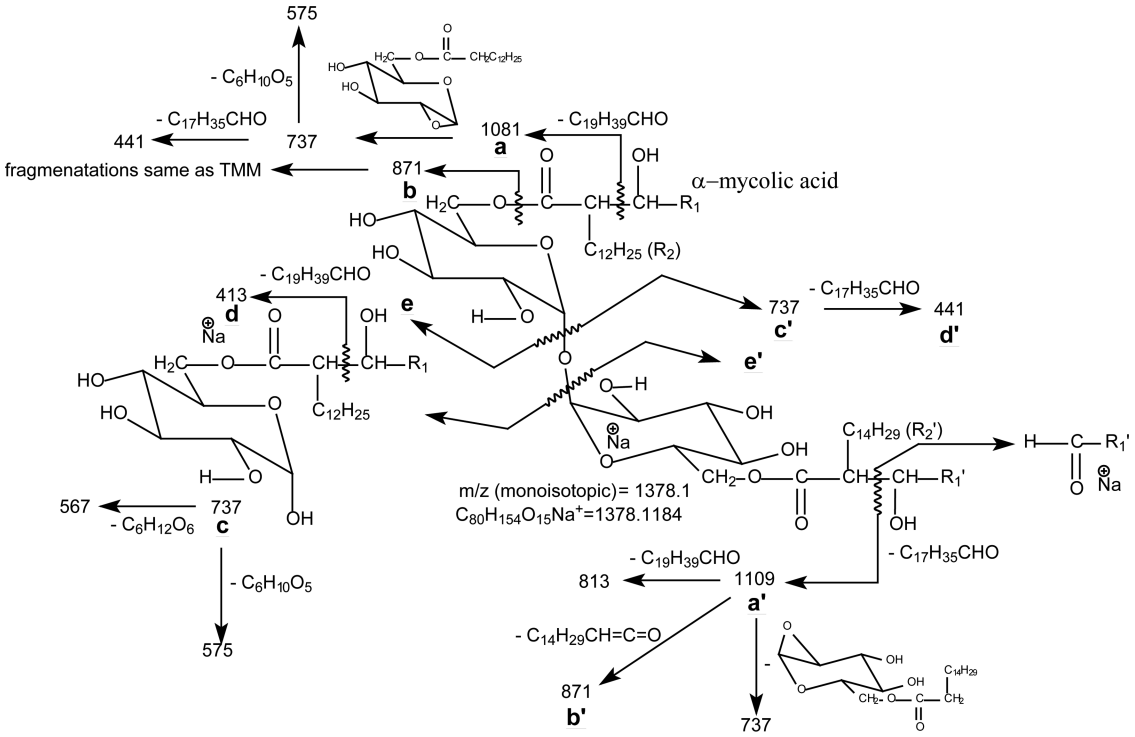
**Scheme 1.**

The fragmenation pathways proposed for the $[M + Na]^+$ ions of 6-mycolyl-α,α'-D-trehalose (TMM) (the indicated m/z values are ions seen for18:0/16:0-TMM, which is one of the 3 isomers that give rise to the $[M+Na]^+$ of m/z 871)




**Scheme 2.**

The fragmentation tree applying multiple-stage mass spectrometry ($MS^n$) for structural assignment of the $[M + Na]^+$ ions of 6,6'-dimycolyl-α,α'-D-trehalose (TDM).




The major composition of meromycolate and α-branches seen for ion of 1378

| $R_1$ | $R_2$ | $R_1'$ | $R_2'$ |
|---|---|---|---|
| $C_{19}H_{39}$ | $C_{12}H_{25}$ | $C_{17}H_{35}$ | $C_{14}H_{29}$ |
| $C_{17}H_{35}$ | $C_{14}H_{29}$ | $C_{17}H_{35}$ | $C_{14}H_{29}$ |
| $C_{19}H_{39}$ | $C_{12}H_{25}$ | $C_{19}H_{39}$ | $C_{12}H_{25}$ |

**Scheme 3.**

The fragmentation pathways proposed for the $[M + Na]^+$ ions of 6,6'-dimycolyl-α,α'-D-trehalose (TDM). The indicated m/z values are fragment ions seen for (20:0/14:0-18:0/16:0-TDM isomer, which is one of the major isomers that represent the sodiated molecular species of m/z 1378.1

## Table 1

Structures of the sodiated mycolylglucoside ions revealed by $MS^n$

| $[M + Na]^+$ | Mycolic acid compositions | Isomeric structures (meramycolate/α-branch) | |
|---|---|---|---|
| m/z | HC chain length:# of unsaturation | major isomers | minor isomers |
| 901 | 48:2 | 32:1/16:1 | |
| 875 | 46:1 | 30:1/16:0; 32:1/14:0 | |
| 873 | 46:2 | 30:1/16:1; 30:2/16:0 | 32:2/14:0; 28:2/18:0 |
| 863 | 45:0 | 29:0/16:0; 31:0/14:0 | |
| 861 | 45:1 | 31:1/14:0; 29:1/16:0 | |
| 847 | 44:1 | 28:1/16:0; 30:1/14:0 | |
| 845 | 44:2 | 28:1/16:0 | |
| 835 | 43:0 | 27:0/16:0 | 26:0/17:0 |
| 819 | 42:1 | 26:1/16:0; 28:1/14:0 | |
| 817 | 42:2 | 28:1/14:1; 26:1/16:1 | 28:2/14:0; 26:2/18:0 |
| 807 | 41:0 | 25:0/16:0; 26:0/15:0 | 24:0/17:0 |
| 805 | 41:1 | 26:1/15:0 | |
| 793 | 40:0 | 24:0/16:0; 26:0/14:0 | |
| 789 | 40:2 | 24:1/16:1; 26:1/14:1 | |
| 791 | 40:1 | 24:1/16:0; 26:1/14:0 | 28:1/12:0 |
| 779 | 39:0 | 24:0/15:0 | 23:0/16:0 |
| 777 | 39:1 | 24:1/15:0 | |
| 765 | 38:0 | 22:0/16:0; 24:0/14:0 | 20:0/18:0 |
| 763 | 38:1 | 22:1/16:0; 24:1/14:0 | |
| 761 | 38:2 | 24:1/14:1; 22:1/16:1 | |
| 751 | 37:0 | 21:0/16:0; 20:0/17:0 | 22:0/15:0 |
| 749 | 37:1 | 22:1/15:0; 20:1/17:0 | 24:1/13:0 |
| 737 | 36:0 | 20:0/16:0; 22:0/14:0 | 18:0/18:0 |
| 735 | 36:1 | 20:1/16:0; 22:1/14:0 | 20:0/16:1; 18:0/18:1 |
| 723 | 35:0 | 20:0/15:0; 18:0/17:0 | 19:0/16:0 |
| 721 | 35:1 | 20:0/15:1 | |
| 709 | 34:0 | 20:0/14:0; 18:0/16:0 | |
| 707 | 34:1 | 20:1/14:0; 18:0/16:1 | 20:0/14:1; 18:1/16:0; 22:1/12:0 |
| 705 | 34:2 | 20:1/14:1; 19:2/15:0 | |
| 695 | 33:0 | 18:0/15:0; 20:0/13:0 | |
| 693 | 33:1 | 18:1/15:0 | |
| 681 | 32:0 | 18:0/14:0; 20:0/12:0 | |
| 679 | 32:1 | 18:0/14:1 | |
| 677 | 32:2 | 18:1/14:1 | |
| 667 | 31:0 | 18:0/13:0 | |
| 653 | 30:0 | 18:0/12:0 | |

| $[M + Na]^+$ | Mycolic acid compositions | Isomeric structures (meramycolate/α-branch) | |
|---|---|---|---|
| m/z | HC chain length:# of unsaturation | major isomers | minor isomers |
| 651 | 30:1 | 18:0/12:1 | |