ELSEVIER

www.elsevier.nl/locate/carres

Carbohydrate Research 319 (1999) 112–123

CARBOHYDRATE RESEARCH

A novel trisaccharide glycolipid biosurfactant containing trehalose bears ester-linked hexanoate, succinate, and acyloxyacyl moieties: NMR and MS characterization of the underivatized structure

Steven Wynn Esch [a,d], Martha D. Morton [b,*], Todd D. Williams [c], Clarence S. Buller [d]

[a] *Higuchi Biosciences Center, The University of Kansas, Lawrence, KS 66047, USA*
[b] *Nuclear Magnetic Resonance Laboratory, The University of Kansas, Lawrence, KS 66045, USA*
[c] *Mass Spectrometry Laboratory, The University of Kansas, Lawrence, KS 66045, USA*
[d] *Department of Molecular Biosciences, The University of Kansas, Lawrence, KS 66045, USA*

Received 3 February 1999; accepted 10 May 1999

Abstract

A Gram-positive actinomycete growing on *n*-hexadecane secreted a family of anionic glycolipid surfactant homologs. The major homolog, with a molecular weight of 1210.6347, had the formula $C_{58}H_{98}O_{26}$. Following mild alkaline saponification, ^{1}H and ^{13}C NMR spectroscopy were used to characterize the non-reducing trisaccharide backbone: β-Glc*p*-(1→3)-α-Glc*p*-(1↔1)-α-Glc*p* ('laminaratrehalose'). Hexanoate, succinate, 3-hydroxyoctanoate, and 3-hydroxydecanoate were found in 3:1:1:1 molar ratio using GC–EIMS analysis of fatty acid methyl esters (FAME) prepared by transesterification. We found that the β-hydroxy acids bore secondary hexanoate chains in 3-*O*-ester linkage, giving acyloxyacyl anions of appropriate *m*/*z* in FABMS and FABMS/MS spectra. COSY, HETCOR, HMBC, and HMQC NMR experiments established the acylation pattern: succinate at C-2 of the terminal α-glucopyranose ring; hexanoate at C-3″ of the β-glucopyranose ring; 3-hexanoyloxyoctanoate and 3-hexanoyloxydecanoate at the 2′- and 4-positions. In FABMS spectra, the homologs flanked the molecular ion by ±14 and ±28 amu, suggesting heterogeneity in acyl chain length.

Keywords: Trisaccharide, glucose; Trehalose oligosaccharide, laminaratrehalose; Biosurfactant, anionic; Glycolipid, succinylated trehalose; Lipooligosaccharide, acyltrehalose; Lipopolysaccharide, lipid A

1. Introduction

Hydrophobic organic compounds, such as petroleum hydrocarbons, can be used by a diversity of oil-consuming lifeforms. Plants and animals provide additional hydrocarbon compounds including the cuticular wax layer of foliar plant surfaces [1]. Utilization of these solid and liquid hydrocarbons by oil-degrading microorganisms requires access to a growth substrate, which is often virtually insoluble in the aqueous phase. 'Hydrocarbonoclastic' microorganisms [2] often make and secrete one or more surface-active agents (biosurfactants and/or bioemulsifiers) [3–5]. Any new or unexamined hydrocarbonoclastic environmental isolate offers a potential for the discovery of a new biosurfactant structure.

* Corresponding author. Tel.: +1-785-331-0284; fax: +1-785-864-5396.
E-mail address: mmorton@caco3.chem.ukans.edu (M.D. Morton)

0008-6215/99/$ - see front matter
PII: S0008-6215(99)00122-6

During a survey for wild-type microorganisms capable of degrading hydrophobic compounds, various isolates from soil were obtained by enrichment culture methods. One of these, 'isolate Q,' readily utilized liquid phase *n*-hexadecane as a carbon and energy source. Unconfirmed preliminary chemotaxonomic analysis placed this aerobic, orange-pigmented, non-sporulating, non-motile, apparently non-filamentous pleomorphic rod in the *Corynebacterium–Mycobacterium–Nocardia–Rhodococcus* (CMNR) group of related Gram-positive genera. Preliminary chemotaxonomic data suggest that this isolate is most probably a new species of *Rhodococcus*. A highly mucoid colony morphology and the display of unusual wetting action in peptone broth culture paralleled the strong wetting action characteristic of other biosurfactant producing species [6]. Furthermore, aliquots from cultures grown on *n*-hexadecane demonstrated strong emulsification activity in a kerosene/buffer shake assay. The surface tension of batch cultures incorporating kerosene dropped to a low value soon after inoculation [7].

Here we report the purification and characterization of an alkali-labile glycolipid biosurfactant ('lipid Q,' Fig. 1) produced by terrestrial isolate Q. The novel glycolipid structure incorporated the disaccharide trehalose as a subunit of its acylated trisaccharide backbone. This suggested a possible relationship to the set of known acyltrehalose-conjugated oligosaccharides produced by various CMNR bacteria [8]. The acylation pattern and carbohydrate sequence of lipid Q suggest a similarity to both the pentasaccharide lipid produced by *Nocarida corynebacteroides* SM1 in response to the presence of hydrocarbon substrates [9,10] and the trehalose tetraester (DSM4325-TT) of *Rhodococcus erythropolis* DSM4325 [11]. Unique to lipid Q with respect to all other known trehalose-containing glycolipids, and (to our knowledge) unprecedented among the lipids of Gram-positive eubacteria in general, was the presence of acyloxyacyl structures in the *O*-ester linkage to the carbohydrate. Comparable acyloxyacyl structures commonly occur in amide and/or ester linkage to various analogs of the 4′,1-diphosphorylated diglucosamine reducing terminus [β-D-Glc*p*N-(1→6)-α-D-Glc*p*N] of the lipopolysaccharide (LPS) and lipooligosaccharide (LOS) of Gram-negative eubacteria; an architecture conventionally known as 'lipid A' [12]. We adopted a nomenclature convention previously established for the acyloxyacyl anatomy of lipid A: directly sugar-linked β-hydroxy acids were identified as primary (1°) acyl chains; those esterified to the primary chains at the β-hydroxy position were designated secondary (2°) acyl chains [12,13].

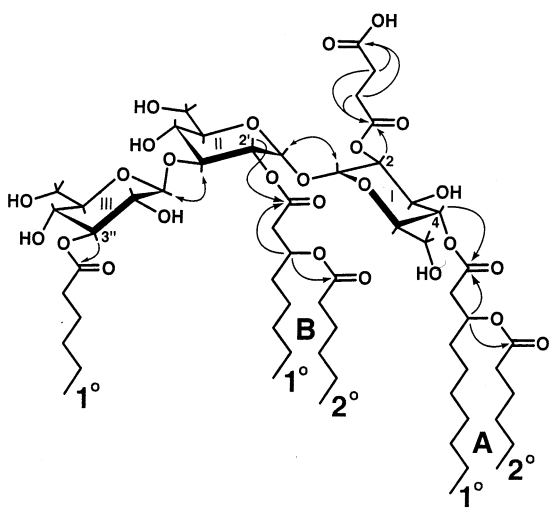

Fig. 1. Structure of the lipid Q. Roman numerals (I, II, III) identify the carbohydrate rings sequentially, beginning with the terminal glucose ring of the trehalose subunit as the 'reducing end' of the molecule. Arrows depict $^{1}H-^{13}C$ correlations for acyl groups, as established by HMBC data. Letters A and B correlate the data for acyloxyacyl groups presented in Table 3 with carbohydrate ring positions-4 and -2′, respectively.

2. Results

Coryneform Q grew rapidly on glucose in a chemically defined liquid medium containing biotin and thiamine supplements, but could also utilize *n*-hexadecane (somewhat more slowly) as an alternative source of carbon and energy as well. After nutrient depletion, the crude lipid extracts of whole cells and cell-free supernatant were prepared. Isolate Q produced significant amounts of a cell-associated glycolipid (CAQ) and a chromatographically indistinguishable extracellular glycolipid (lipid Q) in cultures containing *n*-hexadecane, but these same compounds were not evident after

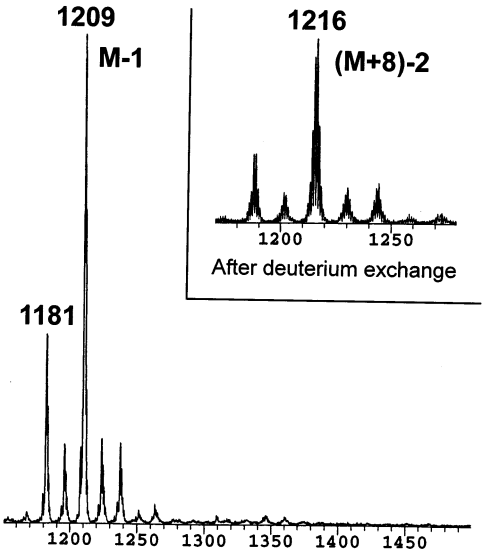

Fig. 2. The molecular-ion region of the negative-ion FABMS spectrum reveals five molecular species distributed ± 14 and ± 28 amu around a major homolog [M − 1] of m/z 1209. The inset shows the same region of the negative-ion FABMS spectrum following deuterium exchange with CD_3OD.

growth on glucose alone [7]. The extracellular glycolipid, produced in *n*-hexadecane cultures, was quantitatively precipitated from cell-free culture supernatant by acidification to pH 4. Thin-layer chromatography (TLC) of the crude lipid extract established that lipid Q lacked phosphorus or primary amino groups, but accounted for almost all of the carbohydrate in the crude lipid extract. Without further purification, the crude extract was subjected to mild acid hydrolysis over progressive time intervals in trifluoroacetic acid (TFA) and strong acid hydrolysis in HCl. Analytical TLC of the liberated carbohydrate indicated that lipid Q contained a structure that migrated in a manner consistent with a glucose trisaccharide, and a disaccharide that co-migrated with an authentic α,α-trehalose reference standard, and after prolonged hydrolysis, yielded only glucose [7]. An IR spectrum of the purified extract indicated the presence of ester bonds (1733 cm^{-1}).

FABMS of purified lipid Q in negative-ion mode showed a strong ion current at m/z 1209 [M − 1], flanked by lesser peaks ± 14 and ± 28 amu distant (Fig. 2). A matching series of ions appeared for sodiated gas-phase adducts of lipid Q in positive-ion mode. As a percentage of the total, the areas of the five peaks in Fig. 2 ranked as follows: m/z 1181 (17.13%), m/z 1195 (7.96%), m/z 1209 (54.91%), m/z 1223 (9.91%), and m/z 1237 (10.09%). These data suggested a family of five glycolipid homologs differing only by the number of methylene units in the fatty acyl portion.

Preliminary data from NMR afforded a carbon count of 58–60 and an additional estimate of 96–100 hydrogen atoms. ^{13}C NMR data identified 20 carbons in the chemical shift region for carbohydrate and also counted seven carbonyl groups. Since only glucose had been obtained after strong acid hydrolysis, 18 of the 20 resonance signals in the 60–110 ppm range were attributed to oligosaccharide. The remaining two resonance signals correlated with two hydroxylated carbons in the acyl side chains. No chemical shifts conclusive for amino sugars were detected, nor were there chemical shifts attributable to unsaturation of the acyl chains.

Exact-mass FABMS experiments established an empirical molecular weight of 1210.6347 for the major glycolipid homolog. Together with the insights supplied by NMR, and assuming that only carbon, hydrogen and oxygen were present, a computational matching program provided a molecular formula very close to the measured exact mass: $C_{58}H_{98}O_{26}$ (MW 1210.634635). This afforded a double-bond equivalence of 10, consistent with three closed hexose rings of a glucose trisaccharide and seven ester carbonyl groups. There appeared, however, to be three more oxygens in the molecule than the count required for direct acylation of the carbohydrate with seven *n*-acyl substituents. Although FABMS spectra in negative-ion mode gave a molecular ion [M − 1] of m/z 1209 for the major homolog, after proton/deuterium exchange in CD_3OD and FABMS analysis in deuterated matrix, the largest molecular ion peak shifted to m/z 1216 (Fig. 2). This implied that there were originally eight exchangeable protons on the fully protonated lipid Q, since M − 1 shifted to [M − 8H + 8D] − D.

In negative-ion FABMS/MS spectra of the m/z 1209 precursor, sizable ion current appeared at m/z 503, m/z 485, and m/z 467 (Fig. 3). Since the anhydrous mass of a neutral non-reducing glucose trisaccharide is 504 Da,

an m/z of 503 is consistent with the loss of one proton [14,15]. The NMR data for the deacylated carbohydrate confirmed that a trisaccharide was present (see below). The sister ions at m/z 485 and m/z 467 could be explained as evidence of fragmentation pathways resulting in the additional loss of one and two molecules of H_2O, respectively. Two more ions assignable to a deacylated disaccharide fragment were found at m/z 341 (ionized disaccharide) and m/z 323 (loss of H_2O). The naked trisaccharide peak at m/z 503 in negative-ion FABMS/MS spectra shifted to twin peaks at m/z 509 and m/z 510 in the FABMS spectrum following proton/deuteron exchange. We evaluated these peaks as $504-7H+7D-D$ and $504-7H+7D-H$, indicating that seven of the 11 hydroxyl positions on the trisaccharide were exchangeable, the remaining four being acylated. (Departure of each acyl group left a proton to balance the vacated oxygen.) Additional peaks in the deuterated spectrum were prominent at m/z 490 (e.g., $510-D_2O$; $509-HDO$), at m/z 489 (e.g., $509-D_2O$), and at m/z 470 (e.g., $509-D_2O-HDO$). Peaks correlated to disaccharide fragments were also shifted by appropriate amounts. Thus, the identity of the carbohydrate peaks in the protonated FABMS/MS spectrum (Fig. 3) was confirmed by correlation to peaks in the deuterated spectrum. These interpretations suggested that FABMS fragmentation of the underivatized lipid Q homologs produced ions mostly as a result of structurally uninformative side-chain losses (solid dots in Fig. 3), and not by cleavage of the carbohydrate backbone ([8] p. 221). It also indicated one exchangeable proton was not part of the carbohydrate moiety, but was associated with a free carboxylate. From NMR data, this exchangeable proton was later determined to be on the free carboxylate of succinate.

Fatty acid methyl esters (FAME) were prepared from the purified mixture of lipid Q homologs by transesterification with methanolic HCl, then analyzed by GC–EIMS (direct on-column splitless injection of the reaction mixture). MS analysis software (Bench Top PBM, Palisade Corp., Newfield, NY) identified the unknown FAME peaks as methyl hexanoate, dimethyl succinate, traces of

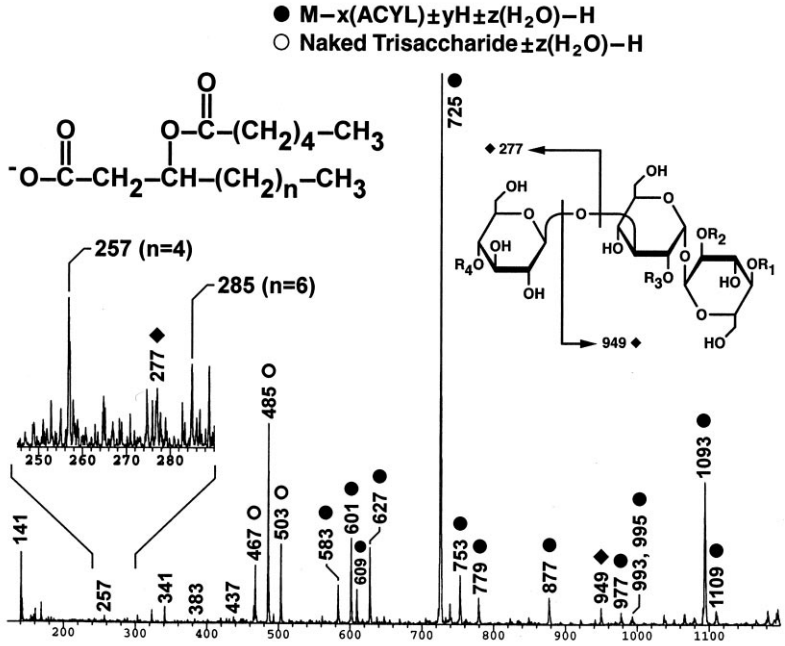

Fig. 3. A FABMS/MS experiment of the major homolog m/z 1209 [M − 1] precursor shows that loss of acyl groups was the principal type of fragmentation event (solid bullets, ●). Open bullets (○) designate the deacylated trisaccharide ion at m/z 503, with additional water losses at m/z 485 and m/z 467. Also shown in the expanded portion of the low-mass region of the spectrum are peaks at m/z 257 and m/z 285 corresponding to anions of intact acyloxyacyl structures. Ions designated with a solid diamond (◆) at m/z 277 and m/z 949 appear to correlate with cleavage of the β-(1→3) glycosidic linkage, as shown, where R_1 and R_3 correspond to acyloxyacyl groups, R_2 is succinate, and R_4 is hexanoate.

Table 1
Composition, retention times and normalized peak areas (integrated ion current) for GC–EIMS of FAME derived from a purified mixture of lipid Q homologs by transesterification

Composition	Retention time (s)	Normalized area
Me-C_6	1:50	2.789
Me_2-succinate	3:31	1.276
Me-C_8	5:54	0.136
Me-3-OH-C_8	8:39	1.000
Me-3-OH-C_9	10:41	0.072
Me-3-OH-C_{10}	12:36	0.536

methyl octanoate, and the methyl esters of three saturated β-hydroxy acids (Table 1). Chain lengths of the two major β-hydroxy acids could not be rigorously confirmed by the electron-impact mass spectral (EIMS) fragmentation data because the molecular ion current of their methyl esters was too weak to be distinguished reliably from background. However, their identities were indirectly established as β-hydroxyoctanoate and β-hydroxydecanoate by comparison of their GC retention times to those of an homologous series of straight-chain FAME, and to published equivalent-chain-length (ECL) data from analysis of β-hydroxy FAME on columns of similar polarity [7].

An unusual GC peak corresponding to β-hydroxynonanoate was observed with a normalized peak area amounting to only 4.5% of the total area of all three hydroxy acid peaks. The EIMS spectrum of this peak included a base-peak fragment ion at m/z 103, in common with the fragment ion profiles of the two flanking β-hydroxy acid peaks in the column elution sequence. The integrated GC–EIMS ion current data of Table 1 indicated that β-hydroxyoctanoate was roughly twice as plentiful as β-hydroxydecanoate in the mix of naturally occurring lipid Q homologs under the cultivation conditions used. This suggested that methyl hexanoate, dimethyl succinate, methyl β-hydroxyoctanoate, and methyl β-hydroxydecanoate might appear in a stoichiometric ratio of 3:1:1:1, respectively, for the major glycolipid homolog (MW 1210), and in a stoichiometric ratio of 3:1:2:0 for the next most abundant molecular homolog (MW 1182).

While considering a variety of hypothetical arrangements for lipid Q acyl groups, evidence for the anions of 3-hexanoyloxyoctanoate and 3-hexanoyloxydecanoate was discerned in the low-mass region of the negative-ion FABMS/MS spectrum (Fig. 3). The signal for 3-hexanoyloxyoctanoate at m/z 257 was almost twice that for 3-hexanoyloxydecanoate at m/z 285, in agreement with the apparent 2:1 ratio of β-hydroxyoctanoate to β-hydroxydecanoate observed by integration of ion current from GC–EIMS peaks. The same peaks again appeared at m/z 257 and m/z 285 in the deuterated FABMS spectrum, and their relative magnitudes showed the same approximate 2:1 ratio.

The deacylated trisaccharide backbone was prepared from lipid Q by mild alkaline saponification and then subjected to ^{1}H and ^{13}C NMR analysis (Table 2). The structure β-Glc*p*-(1→3)-α-Glc*p*-(1↔1)-α-Glc*p* was identified by comparison of ^{1}H NMR and ^{13}C NMR chemical shift data to literature values [16–18]. A heteronuclear multiple-bond correlation experiment (HMBC) confirmed these anomeric linkage configurations. ^{1}H NMR coupling constants revealed the configuration presiding at the anomeric carbons (β 7.5 Hz, α 3.7 Hz and α′ 3.7 Hz) [19].

Equipped with data for the naked trisaccharide, a comparison was made to NMR data for the carbohydrate region of the native molecule (Table 2). ^{13}C NMR spectroscopy confirmed the presence of seven acyl-group carbonyl carbons (Table 3), 20 carbons in the carbohydrate range between 110 and 60 ppm (Tables 2 and 3), and various additional acyl-group carbons (Table 3). These were resonances consistent with the structural components and molecular formula derived from GC–EIMS and exact-mass FABMS analysis. Specifically, two of the ^{13}C resonances (at 71.7 and 71.5 ppm) fell into the carbohydrate range, but were later identified as associated with the beta carbons (–$CH_2HCORCH_2$–) of the β-hydroxy acids. A similar situation was observed in the proton spectrum, where the beta protons of the β-hydroxy acids (–$CH_2HCORCH_2$–) displayed chemical shifts so far downfield that they overlapped the anomeric protons of the trehalose linkage (see Fig. 6). Furthermore, both ^{1}H and ^{13}C NMR data ruled out the possibil-

Table 2
^1H and ^{13}C NMR assignments for the carbohydrate region of the native lipid Q structure and for the naked trisaccharide [a]

	Lipid Q		Trisaccharide	
Atom	^1H (in CD_3OD)	^{13}C (in CD_3OD)	^1H (in D_2O)	^{13}C(in D_2O)
H(C-1)^{I}	5.27	93.0	5.38; $^3J_{1,2}$ 3.7(d) α	95.8
H(C-2)^{I}*	4.80	74.1	3.84	73.6
H(C-3)^{I}	4.10	70.2	4.07	75.1
H(C-4)^{I}*	4.90	72.6	3.62	72.1
H(C-5)^{I}	3.82	72.4	3.71	74.7
H(C-6)^{I}	3.78	62.4	4.02/4.04	63.2/63.0
H(C-1)^{II}	5.23	93.2	5.40; $^3J_{1,2}$ 3.7(d) α	95.6
H(C-2)^{II}*	4.88	73.3	4.02	74.5
H(C-3)^{II}	4.01	82.1	4.27; $^3J_{3,4}$ 9.4	84.2
H(C-4)^{II}	3.48	77.9	3.74	70.7
H(C-5)^{II}	3.78	74.4	4.03	74.7
H(C-6)^{II}	3.53	62.1	3.93, 4.04	63.2/63.0
H(C-1)^{III}	4.58	105.1	4.93; $^3J_{1,2}$ 7.91(d) β	105.2
H(C-2)^{III}	3.30	73.3	3.54	76.0
H(C-3)^{III}*	4.95	78.6	3.70	78.1
H(C-4)^{III}	3.45	69.8	3.63	72.2
H(C-5)^{III}	3.62	70.1	4.00	74.7
H(C-6)^{III}	4.00	62.4	4.04/4.05	63.2/63.0

[a] The 1-bond proton to carbon correlations for lipid Q were established with a HETCOR experiment, and those for the trisaccharide were established with a HMQC experiment. Selected J values are in Hz.

Table 3
Acyl group ^1H and ^{13}C NMR assignments for the native lipid Q glycolipid structure [a]

Atom	^1H (in CD_3OD)	^{13}C (in CD_3OD)
βCH of 1° chain of acyloxyacyl group A	5.28	71.7
βCH of 1° chain of acyloxyacyl group B	5.22	71.5
–CH_2– unit of succinate	2.70	30.6
–CH_2– unit of succinate	2.65	30.3
γCH_2 of 1° chain of an acyloxyacyl group	1.66	35.5
αCH_2 units of 2° chains of acyloxyacyl groups	2.33	35.4
αCH_2 of 1° hexanoate at 3″ ring position	2.39	35.3
γCH_2 of 1° chain of an acyloxyacyl group	1.65	35.3
Internal –CH_2– units of acyl chains	1.3	33.2–32.5
Internal –CH_2– units of acyl chains	1.3	30.7, 30.9
αCH_2 of 1° chain of acyloxyacyl group B	2.80	40.2
αCH_2 of 1° chain of acyloxyacyl group A	2.70	40.1
βCH_2 of 1° hexanoate at 3″ ring position	1.6	26.1–25.8
Internal –CH_2– units of acyl chains	1.3	26.1
Penultimate –CH_2– units of acyl chains	1.3	23.9–23.6
ωCH_3 units of acyl chains	0.9	14.6–14.5
Free acidic carbonyl carbon of succinate		175.3
2° carbonyl carbon of acyloxyacyl group B		175.3
Carbonyl of 1° hexanoate at 3″ ring position		175.0
2° carbonyl carbon of acyloxyacyl group A		174.9
Ester carbonyl carbon of succinate		173.8
1° carbonyl carbon of acyloxyacyl group B		171.8
1° carbonyl carbon of acyloxyacyl group A		171.3

[a] Lengths of the 1° β-hydroxy acid chains in groups A and B could not be ascertained by NMR data, but mass spectral evidence suggested that group A preferentially incorporated the longer 1° chain (see Section 3 and Figs. 1 and 6).

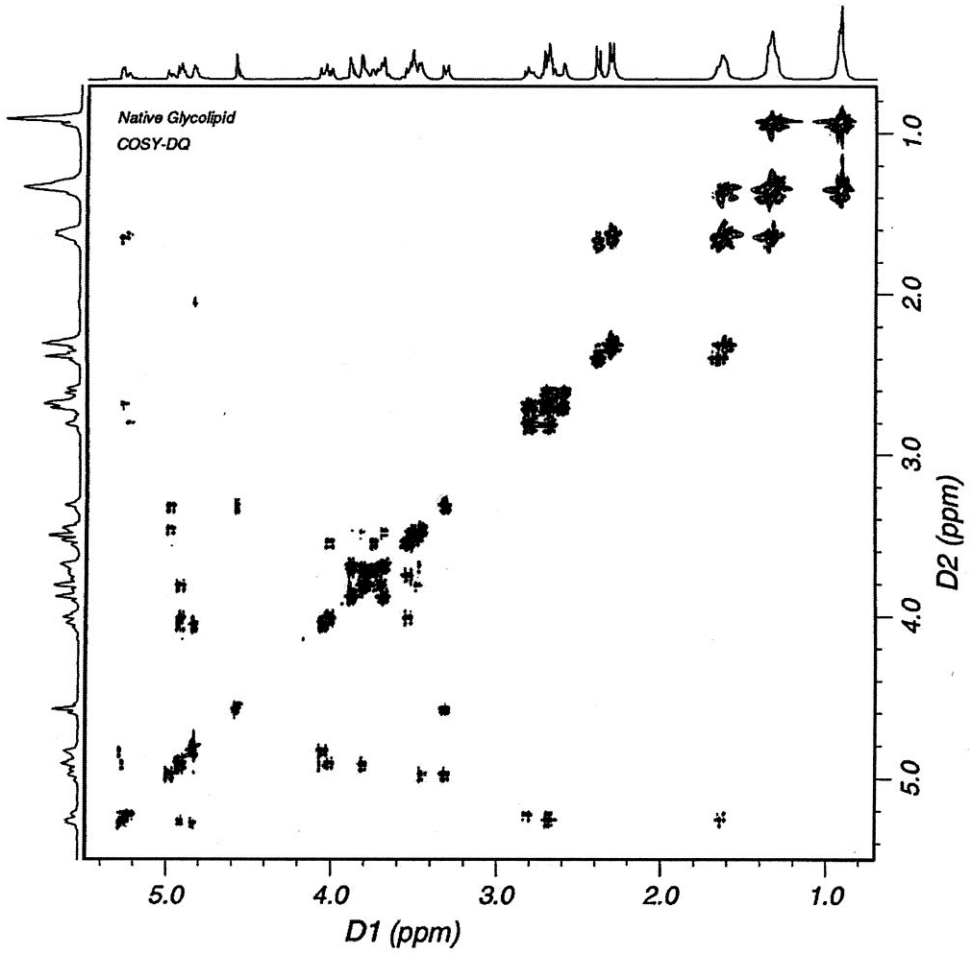

Fig. 4. $^{1}H-^{1}H$ COSY-DQF contour plot of lipid Q.

ity that an amino sugar such as glucosamine was present, consistent with the m/z 503 in the negative-ion FABMS spectrum.

With three distinct carbon resonances in the anomeric region of the ^{13}C NMR spectrum, we confirmed the presence of only three sugar moieties. A HETCOR experiment ($^{13}C-^{1}H$ heteronuclear correlation) permitted unambiguous assignment of carbon resonance values to all structurally significant protonated carbons in the spectrum. A $^{1}H-^{1}H$ COSY-DQF (correlation spectroscopy with double quantum filtering) allowed assignment of proton chemical shifts to the carbohydrate rings and furnished the early assessment that there were only four positions of acylation on the trisaccharide (Fig. 4). Specifically, by comparing the COSY-DQF data for the native lipid Q molecule (Fig. 4) with data for the naked trisaccharide (Fig. 5), the positions occupied by acyl groups could be conclusively established as localizing to carbons-2^{I}, -4^{I}, -2^{II} and -3^{III} of the carbohydrate backbone. The arrangement of fatty acids into a β-hydroxyacyloxyacyl structure was confirmed by the COSY-DQF correlation of two downfield protons that correlated to the lipid rather than the carbohydrate backbone (Fig. 6). This hypothesis was later supported by a HMBC experiment which showed a connectivity between H-2^{I} (the proton on carbon # 2 of ring I) and one of the carbonyl carbons of succinate (Fig. 1). Connectivites were also seen for carbohydrate H-4^{I} and H-2^{II} to the two primary-chain carbonyl carbons of the acyloxyacyl groups. The beta protons of these primary acyloxyacyl chains showed a further connectivity to the secondary chain hexanoate carbonyls. Finally, using the COSY-DQF to ensure the assignment of the protons on the

above carbons, it was possible to trace the connectivities around each sugar ring of the carbohydrate backbone. The HMBC data then confirmed the linkages of the anomeric carbons of the trisaccharide. FABMS/MS data provided circumstantial evidence for preferential placement of 3-hexanoyloxyoctanoate at the 2^{II}-position.

3. Discussion

Structural elucidation of lipid Q (Fig. 1) was achieved by a combination of MS and NMR technologies, without resorting to many of the chemical derivatization steps ordinarily necessary [8,20]. NMR analysis successfully placed acyl groups on the trisaccharide backbone as illustrated in Fig. 1. The only remaining ambiguity involved which acyloxyacyl structure should be assigned to the number 4^{I} ring hydroxyl group, and which belonged at the 2^{II}-position. Although the available data do not unequivocally establish preferential assignment of 3-hexanoyloxyoctanoate to the 2^{II}-position, we noticed that the peak at m/z 725 in the FABMS/MS spectrum of Fig. 3 showed much larger ion current than the peak at m/z 753. If it is valid to argue that the internal 2^{II} ring position is more sterically protected from collision-induced dissociation than terminal ring position-4^{I} on the molecule, then preferential loss of 3-hexanoyloxydecanoate from the 4-position would lead to a relative abundance of ions at m/z 725 that retain 3-hexanoyloxyoctanoate at the 2^{II}-position. Thus, we suggest placement of acyloxyacyl groups as shown in Fig. 1 for the most common lipid Q homolog, with β-hydroxydecanoate forming the 1° chain of acyloxyacyl group A at ring position-4.

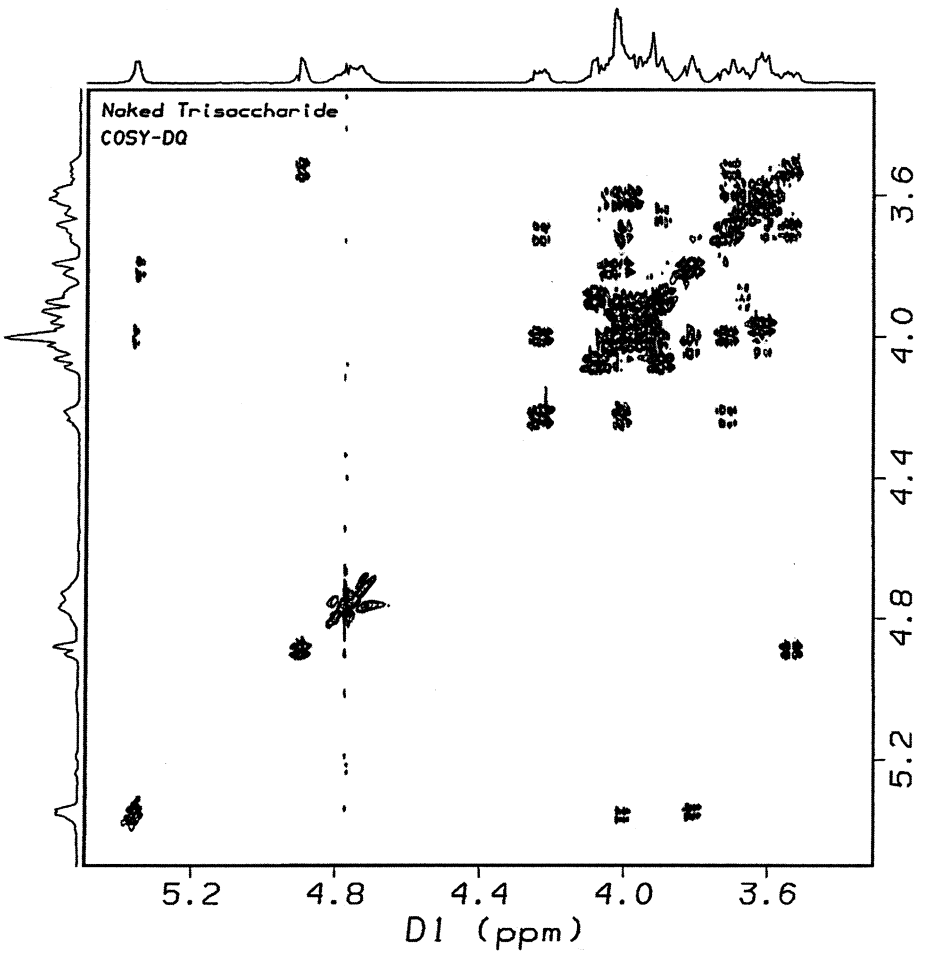

Fig. 5. $^{1}H-^{1}H$ COSY-DQF contour plot of the trisaccharide backbone.

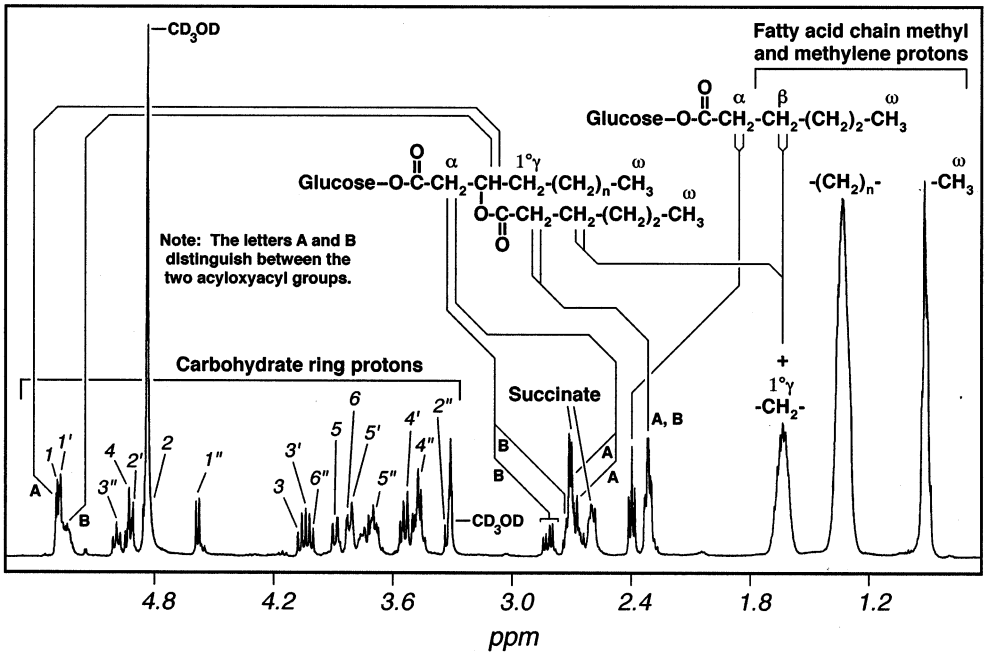

Fig. 6. Schematic of the proton spectrum illustrating both the carbohydrate and acyl group resonance assignments. Letters A and B distinguish between the two acyloxyacyl groups at carbohydrate ring positions-4 and -2′, respectively, as indicated in Fig. 1 and Table 3.

Leaving succinate out of account, the average chain length of hydrophobic acyl groups in the succinylated trehalose tetraester of *R. erythropolis* DSM4325 [11] is 9.3 (28 acyl carbons in three acyl groups). This latter value is characteristic of three other succinylated trehalose tetraesters reported in the literature [32–34]. By comparison, the average acyl chain length calculable for lipid Q is 7.2 (36 acyl carbons present in five acyl chains). Furthermore, the average acyl chain length in the succinylated pentasaccharide lipid of *N. corynebacteroides* SM1 [9,10] is only 4.4 (31 acyl carbons present in seven acyl chains). Thus, the average chain length of hydrophobic acyl groups in lipid Q lies at a point intermediate between the average acyl chain lengths common to several succinylated trehalose tetraesters and that of the SM1 pentasaccharide lipid. A trend toward shorter acyl chain lengths associated with larger carbohydrate backbones may represent a clue to the functional role played by this class of biological surfactants.

To our knowledge, lipid Q furnishes the first known example of an *O*-ester-linked acyloxyacyl motif among the glycolipids of Gram-positive eubacteria, and among the known acyltrehalose LOS of Gram-positive CMNR bacteria [8,20–24]. Yet acyloxyacyl structures are almost always observed in the form of amide-linked (and often also *O*-ester-linked) substituents of the LPS of Gram-negative bacteria [12]. A balanced discussion of the implications of such a finding exceeds the scope of this paper [S.W. Esch, manuscript in preparation].

4. Experimental

Cultivation.—Each liter of growth medium contained: 4.13 g KH_2PO_4; 1.26 g K_2HPO_4; 3.18 g Na_2HPO_4; 0.40 g NH_4Cl; 0.12 g $MgSO_4$; 0.010 mg biotin; 0.025 mg thiamine; 0.10% glucose (w/v) and 5 mL of trace metal salts solution. Each liter of trace metal salts solution contained 500 mg EDTA (free acid form), 3.42 mL 1.0 M NaOH, 200 mg $FeSO_4 \cdot 7H_2O$, 10 mg $ZnSO_4 \cdot 7H_2O$, 3 mg $MnCl_2 \cdot 5H_2O$, 30 mg H_3BO_3, 20 mg $CoCl_2 \cdot 6H_2O$, 1 mg $CuCl_2 \cdot 2H_2O$, 2 mg $NiCl_2 \cdot 6H_2O$, and 3 mg $NaMoO_4 \cdot 2H_2O$. Production cultures (400-mL) in 2000-mL baffled Erlenmeyer culture flasks

(Bellco, Vineland, NJ) each received a 1% (4 mL) inoculum from a 24-h 30-mL starter culture, followed by incubation on a rotating platen with 2.54 cm offset operated at 200 rpm, 30 °C ambient. At 48 h post inoculation each flask received 1.035 mL (= 0.20%, w/v) of filter-sterilized '99% pure' *n*-hexadecane (Sigma Chemical, St. Louis, MO). Hexadecane purity was verified by GC–EIMS, the only detectable analyte peak correlating with *n*-C_{16}. Incubation continued for an additional 48 h, at which time hexadecane depletion was established by visual inspection: absence of a surface pellicle of frozen hydrocarbon following centrifugation of a culture aliquot at 4 °C.

Preparation of crude lipid extract.—After centrifugation at 12,000*g*, 4 °C for 20 min, the culture supernatant was vacuum-filtered through nylon membranes of 0.45 and 0.22 μm pore size (MSI, Westboro, MA). The cell-free supernatant, containing a mixture of water, protein, carbohydrate, glycolipid, and soluble pigment, was acidified at 25 °C to pH 4.0 with 1 N H_2SO_4. It produced 2.6 g/L of orange-colored gel precipitate after centrifugation as above. The gel precipitate was extracted once into 1:2:0.8 $CHCl_3$–MeOH–ddH_2O [25], then twice more into 1:1 $CHCl_3$–MeOH. Extracts of pelleted whole cells were prepared in like manner, except that 2:1 $CHCl_3$–MeOH was used for the third extraction step. After adjusting combined extracts to 1:1:0.9 $CHCl_3$–MeOH–ddH_2O [25], the emulsified mixture was centrifuged at 2520*g*, 4 °C, for 2 h. The lower chloroform phase extract was rinsed with water, dried with anhydrous Na_2SO_4, reduced to residue by rotary evaporation, and taken up in $CHCl_3$ for storage at −20 °C over Na_2SO_4. As a general rule, desolvated dry lipid residue was taken up in 2 mL of fresh $CHCl_3$ per gram of wet cells or gel precipitate extracted.

Acid hydrolysis of crude lipid extract.—For strong acid hydrolysis following solvent removal, a residue of crude lipid extract containing approximately 900 μg of lipid Q was subjected to 1.0 mL of aq 1.0 N HCl under argon at 98 °C for intervals ranging from 30 to 240 min. After cooling, this was extracted four times successively with 1.0-mL portions of petroleum ether, evaporated to dryness under reduced pressure in the presence of KOH pellets, then reconstituted in 1.0 mL water. Mild acid hydrolysis was carried out on a residue of crude lipid extract containing approximately 2.7 mg lipid Q, using 2.0 mL of 2 N TFA under argon at 90 °C for incubation times of 10 min and 2 h. Hydrolysates were cooled, extracted twice with 1.0-mL portions of fresh $CHCl_3$, evaporated to dryness under a jet of N_2 at 25 °C, then reconstituted in 750 μL of 4:1 MeOH–water.

Analytical TLC.—All TLC plates were prewashed by a full-length ascent of 1:1 $CHCl_3$–MeOH prior to activation and loading. LK6 analytical plates and LK6D analytical plates (Whatman, Clifton, NJ) were used for analytical lipid chromatography. Plates were subjected to single-pass development in 65:24:4 $CHCl_3$–MeOH–ddH_2O in sandwich chamber mode, providing a solvent-saturated atmosphere. Dittmer–Lester phosphate stain [26] was used for detection of lipids incorporating phosphate groups. A 0.25% ninhydrin solution was prepared in water-saturated butanol for use in detection of primary amino groups [27]. Carbohydrate was detected by spraying plates with a 0.5% solution of α-naphthol in 1:1 MeOH–water, followed by spraying with concentrated H_2SO_4 and heating to 110 °C until intense blue–purple bands were visible [28].

K5 analytical plates (Whatman, Clifton, NJ) were used for carbohydrate TLC. Aqueous solutions of acid hydrolysates were spot-loaded to achieve a total application of 10–12 μg total lipid carbohydrate per 2–3 mm spot at the origin. Authentic standards prepared in MeOH–water (1:1 to 4:1) were applied to achieve 8 μg/spot. Plates were developed in 85:15 acetonitrile–water [29], then visualized with α-naphthol–H_2SO_4, as above.

Purification of glycolipid by preparative TLC.—The crude supernatant lipid extract was applied to prewashed glass-backed Silica Gel G Uniplate Taper Plates (Analtech, Newark, DE) as a band of 160 mm length at a uniform loading of 1.0 mg total lipid carbohydrate (2.4 mg glycolipid) per plate. Following single-pass development in 65:24:4 $CHCl_3$–MeOH–ddH_2O, lipid Q was visual-

ized non-destructively by spraying one edge of the preparative plate lightly with water, then exposing the moistened edge to a stream of filtered air delivered through a Pasteur pipette packed with iodine crystals. A band of silicic acid 2.5 cm wide was scraped from each plate at an upper boundary defined by the maximum R_f attained by lipid Q. From this material purified glycolipid was re-extracted and concentrated by the method employed for whole cells. Carbohydrate assay was conducted using the phenol–sulfuric acid reaction [30,31].

Mild alkaline deacylation for NMR studies.—Approximately 11 mg of purified lipid Q, representing 4.6 mg reducing sugar, was redissolved in 0.6 mL MeOH + 0.4 mL $CHCl_3$ + 1.0 mL 0.2 N methanolic KOH, then capped under argon and held at 25 °C for 30 min [28]. Addition of 0.4 mL MeOH + 1.6 mL $CHCl_3$ + 1.2 mL of 0.2 N HCl + 0.6 mL H_2O resulted in a saline biphasic mixture having a MeOH–water upper phase 0.01 N in excess HCl. After centrifugation at 850*g*, the lower $CHCl_3$ phase was discarded, and the upper phase washed three times successively with 2.0-mL aliquots of fresh $CHCl_3$ to assure quantitative removal of protonated fatty acids. Addition of 0.18 mL of 0.2 N KOH, achieved a final upper-phase excess acid concentration of 0.001 N in HCl. This solution was evaporated to dryness alongside a tray of fresh KOH pellets in a vacuum chamber at reduced pressure, then redissolved in 0.3 mL of D_2O. A second round of evaporation and uptake into D_2O achieved a final KCl concentration of 0.78 N and a trisaccharide concentration of 15 mg/mL.

Transesterification for GC–EIMS analysis of FAME.—Minimal loss of volatile short-chain methyl esters was desired. Therefore, 10 mg of lipid Q, representing 0.58 mg of fatty acids was dissolved in 4 mL of 0.5 N anhydrous methanolic HCl (Supelco, Bellefonte, PA), capped under argon, and incubated at 98 °C for 2 h. After cooling to room temperature, the reaction mixture was injected directly onto the GC column. GC–EIMS spectra were acquired using a Nermag (Paris, France) R10-10 instrument. Direct splitless injections of the transesterification reaction mixture were made onto a 15 m, 0.25 mm ID, DB-1 (dimethyl-polysiloxane) capillary column (J&W Scientific, Folsom, CA). Oven conditions were programmed to hold at 60 °C for 3 min, followed by a linear ramp to 250 °C at 8 °C/min.

FABMS and FABMS/MS.—FABMS was obtained using an AUTOSPEC-Q tandem hybrid mass spectrometer (EBEqQ; VG Analytical) equipped with an OPUS data system. FAB experiments employed a cesium gun operating at 20 keV energy and 2 μA emission. Sample in CH_2Cl_2 was added to *m*-nitrobenzyl alcohol (mNBA) as the matrix. Linked scans at constant B/E were performed with precursor ions attenuated 60% with Ar in the first field-free region gas cell. Collision-induced dissociation (CID) spectra were recorded in continuum mode under data system control. The scan range was 2000–5 amu acquired at 10 s/dec; multiple scans were integrated. The collision energy was 8 keV. Exact-mass FAB experiments were carried out at 1:10,000 resolution using linear voltage scans under data system control and collecting continuum data. Polyethylene glycol ions served as bracketing calibrant ions.

NMR spectroscopy.—NMR spectra were obtained using a Bruker AM-500 spectrometer (500 MHz ^1H operating frequency) and a General Electric QE-Plus spectrometer (300 MHz ^1H operating frequency) at 24 °C. The deuterated solvent CD_3OD was purchased from MSD Isotopes and Aldrich Chemical Co., Milwaukee, WI and $CDCl_3$ and D_2O were purchased from Cambridge Isotope Laboratories (Andover, MA). ^1H spectra were referenced against the residual proton impurity in the deuterated solvent. ^{13}C spectra were referenced externally against DSS in D_2O. Double quantum filtered ^1H–^1H correlation spectroscopy (COSY-DQF) was performed on the native glycolipid with solvent suppression of the residual water peak and double quantum filtered in a 256×2048 point matrix. ^{13}C–^1H heteronuclear correlation (HETCOR) was used to determine the one-bond carbon-to-proton correlations for the native glycolipid. An HMBC was used to determine the two- and three-bond connectivities of the side-chain carbons and/or protons to the trisaccharide backbone. This experiment was also used to confirm the glycosidic linkages of the substituted trisaccharide.

Deacylation of lipid Q was achieved by mild alkaline saponification, as described above. The identity of the deacylated trisaccharide was confirmed by three methods; ^{13}C NMR, ^{1}H NMR and a COSY-DQF at 35 °C. The 1-bond proton to carbon correlations for the naked trisaccharide were established with a heteronuclear multiple quantum filtered correlation (HMQC) experiment.

Acknowledgements

The authors thank Robert Drake and Homigol Biesiada for their efforts in acquiring the GC–EIMS and FAB spectra. The tandem mass spectrometer was purchased with the aid of a National Institutes of Health grant S10 RR0 6294-01 (TDW) and the University of Kansas.

References

[1] P.E. Kolattukudy (Ed.), *Chemistry and Biochemistry of Natural Waxes*, Elsevier, Amsterdam, 1976.
[2] R.M. Atlas, in J. Lederberg (Ed.), *Encyclopedia of Microbiology*, Vol. 3, Academic Press, San Diego, 1992, pp. 363–369, and references therein.
[3] C. Syldatk, S. Lang, F. Wagner, V. Wray, L. Witte, *Z. Naturforsch. C*: *Biosci.*, 40c (1985) 51–60.
[4] C. Syldatk, S. Lang, U. Matulovic, F. Wagner, *Z. Naturforsch. C*: *Biosci.*, 40c (1985) 61–67.
[5] P. Goswami, H.D. Singh, *Biotechnol. Bioeng.*, 37 (1991) 1–11.
[6] T. Matsuyama, K. Kaneda, I. Ishizuka, T. Toida, I. Yano, *J. Bacteriol.*, 172 (1990) 3015–3022.
[7] S.W. Esch, Ph.D. Dissertation, University of Kansas, Lawrence, Kansas, USA, 1998.
[8] G.S. Besra, P.J. Brennan, in C. Fenselau (Ed.), *Mass Spectrometry for the Characterization of Microorganisms* (ACS 204th National Meeting, Aug. 1992, Washington, DC), ACS Symp. Ser., 541, American Chemical Society, Washington, DC, 1994, pp. 203–232.
[9] M. Powalla, S. Lang, V. Wray, *Appl. Microbiol. Biotechnol.*, 31 (1989) 473–479.
[10] J.-S. Kim, M. Powalla, S. Lang, F. Wagner, H. Lünsdorf, V. Wray, *J. Biotechnol.*, 13 (1990) 257–266.
[11] E. Ristau, F. Wagner, *Biotechnol. Lett.*, 5 (1983) 95–100.
[12] U. Zähringer, B. Lindner, E.T. Rietschel, *Adv. Carbohydr. Chem. Biochem.*, 50 (1994) 211–276.
[13] A.L. Erwin, R.S. Munford, *J. Biol. Chem.*, 265 (1990) 16444–16449.
[14] C. Bosso, J. Defaye, A. Heyraud, J. Ulrich, *Carbohydr. Res.*, 125 (1984) 309–317.
[15] S.W. Hunter, I. Jardine, D.L. Yanagihara, P.J. Brennan, *Biochemistry*, 24 (1985) 2798–2805.
[16] B. Coxon, Developments in Food Carbohydrate, in C.K. Lee (Ed.), *Developments Series*, Vol. 2, Applied Science Publishers, London, 1980, pp. 351–390.
[17] J.H. Bradbury, G.A. Jenkins, *Carbohydr. Res.*, 126 (1984) 125–156.
[18] T. Usui, N. Yamaoka, K. Matsuda, K. Tuzimura, H. Sugiyama, S.S. Seto, *J. Chem. Soc., Perkin Trans. 1*, (1973) 2425–2432.
[19] S.T. Summerfelt, E.J.-M. Selosse, P.J. Reilly, W.S. Trahanovsky, *Carbohydr. Res.*, 203 (1990) 163–172.
[20] G.O. Aspinall, D. Chatterjee, P.J. Brennan, *Adv. Carbohydr. Chem. Biochem.*, 51 (1995) 169–242.
[21] S. Alugupalli, F. Portaels, L. Larsson, *J. Bacteriol.*, 176 (1994) 2962–2969.
[22] S. Alugupalli, M.-A. Lanéelle, L. Larsson, M. Daffé, *J. Bacteriol.*, 177 (1995) 4566–4570.
[23] K.-H. Khoo, R. Suzuki, H.R. Morris, A. Dell, P.J. Brennan, G.S. Besra, *Carbohydr. Res.*, 276 (1995) 449–455.
[24] M.-A. Lanéelle, G. Silve, L.M. López Marín, M. Daffé, *Eur. J. Biochem.*, 238 (1996) 270–279.
[25] E.G. Bligh, W.J. Dyer, *Can. J. Biochem. Physiol.*, 37 (1959) 911–917.
[26] J.C. Dittmer, R.L. Lester, *J. Lipid Res.*, 5 (1964) 126–127.
[27] W.W. Christie, *Lipid Analysis*, 2nd ed., Pergamon, New York, 1982.
[28] M. Kates, in R.H. Burdon, P.H. van Knippenberg (Eds.), *Laboratory Techniques in Biochemistry and Molecular Biology*, Vol. 3, Pt. 2, 2nd ed, Elsevier, Amsterdam, 1986.
[29] J.F. Robyt, B.J. White, *Biochemical Techniques: Theory and Practice*, Waveland Press, Prospect Heights, IL, 1987.
[30] S.C. Kushwaha, M. Kates, *Lipids*, 16 (1981) 372–373.
[31] M. Dubois, K.A. Gilles, J.K. Hamilton, P.A. Rebers, F. Smith, *Anal. Chem.*, 28 (1956) 350–356.
[32] M.J. Espuny, S. Egjido, M.E. Mercadè, A. Manresa, *Toxicol. Environ. Chem.*, 48 (1995) 83–88.
[33] A. Passeri, S. Lang, F. Wagner, V. Wray, *Z. Naturforsch. C*: *Biosci.*, 46c (1991) 204–209.
[34] S.G. Batrakov, B.V. Rozynov, T.V. Koronelli, L.D. Bergelson, *Chem. Phys. Lipids*, 29 (1981) 241–266.