Extraction, isolation and NMR data of the tetraether lipid calditoglycerocaldarchaeol (GDNT) from *Sulfolobus metallicus* harvested from a bioleaching reactor

Authors: Bode ML (CSIR Biosciences); Buddoo SR (CSIR Biosciences); Minnaar SH (BHP Billiton); Du Plessis CA (BHP Billiton)

Abstract
The successful extraction and isolation of the hydrolysed tetraether lipid calditoglycerocaldarchaeol (GDNT) from *Sulfolobus metallicus*, a key thermophilic bioleaching archaeon, is described. The archaeal biomass was recovered directly from a thermophilic (68^{o}C) bioleaching tank reactor used to extract nickel from a pentlandite mineral concentrate. The initial Soxhlet extraction method employed was scaled to a bench-scale extraction procedure suitable for the preparation of gram-scale quantities of GDNT. The GDNT so obtained was analysed by 1D- and 2D-NMR techniques, providing the first complete ^{13}C and 2-D NMR data-set for GDNT, including that for the intact underivatised calditol moiety. The study demonstrates the feasibility of recovering high-quality GDNT from thermophilic archaeal-mediated bioleaching reactors. The recovery of these lipids at relatively low cost, as a by-product from bioleaching reactors used in the metals processing industry, has important implications for future tetraether lipid availability and costs.

1. Introduction

The membranes of Archaea are unique in that they are constructed from tetraether lipids instead of the ester lipids found in the membranes of Bacteria and Eucarya (De Rosa et al., 1977a). The backbone of these tetraether lipids takes one of two forms, either caldarchaeol or glycerol dialkyl glycerol tetraether (GDGT), with symmetrical headgroups or calditoglycerocaldarchaeol, also known as glycerol dialkyl nonitol tetraether (GDNT), with unsymmetrical headgroups (Gliozzi et al., 2002; Kanichay et al., 2003 and Benvegnu et al., 2004). The alkyl chains connecting the headgroups may contain up to four cyclopentane rings per alkyl chain (Gliozzi et al., 2002). The role of these lipids in the archaeal cell membrane is critical in protecting the cell and the neutral pH cytoplasm from the solution environment of its natural habitat, with its extreme acidity and metals concentration, while remaining stable at high temperatures and facilitating the interaction with membrane-associated enzymes. The unique properties of tetraether lipids include: mono-layer self-assembly (Strobl et al., 1985 and Dote et al., 1990), hydrophobicity, impermeability, ultra-low electrical conductivity, optical transparency and stability under autoclave, ultra-violet and gamma radiation conditions (Bakowsky et al., 2000). Extracted lipids can also be functionalized in order to facilitate covalent bonding to surfaces (Frant et al., 2006) with subsequent formation of surface coatings. The unique properties of tetraether lipids have led to a wide variety of applications including drug delivery liposomes (Bagatolli et al., 2000; Bernina Biosystems GmbH, 2001 and Sprott et al., 2003), boundary lubricants (United States Of America Navy, 1992), biomedical nano-coatings (Kermamed Medizintechnik GmbH et al., 2002) and S-layer supported lipid bilayers (Wetzer et al., 1997).

Despite the unique properties of tetraether lipids, and their eminent suitability in various biomaterial applications, commercial implementation of technology

applications has been limited. The main reason for this limitation has been the prohibitively high cost of such lipids (i.e. exceeding US$ 5000 per gram) derived from cultured archaeal biomass, and the complexity of synthesizing chemical analogues (Benvegnu et al., 2004). Lipid yields from archaeal biomass are low, resulting in high final product costs when archaeal biomass is cultured for the sole purpose of lipid production.

The use of thermophilic Archaea in agitated tank bioleaching reactors is a well-established hydrometallurgical technology (Batty and Rorke, 2006). S*ulfolobus metallicus,* and related species, dominate the archaea population in such reactors (Norris et al., 2000). The archaea act as catalysts in the oxidation of iron and sulphur resulting in the dissolution or beneficiation of the metal of interest, typically copper, cobalt, nickel or gold (Rawlings et al., 2003). Bioleaching reactors used for this purpose are large (i.e. 1300 m^3) and produce archaeal biomass in the order of tons per day. Harvesting of this archaeal biomass can be readily achieved by relatively simple centrifugation methods, without negatively impacting on the mineral extraction efficiency or operation of the reactors. The aim of this investigation was to determine whether high-quality tetraether lipids, suitable for use in the technology applications discussed above, could be recovered from the biomass harvested directly from a continuous culture bioleaching tank reactor.

2. Experimental procedures

2.1. Preparation of biomass

A 200 L agitated continuous culture tank reactor was operated for a period of six weeks. A mineral concentrate slurry was pumped into the reactor at a solids concentration of 15% (w/w). The mineral concentrate contained 22% (w/w) pentlandite, $(Fe,Ni)_9S_8$. The particle size distribution of the mineral concentrate was 30 – 80 μm. The air flow sparged into the reactor was supplemented with both oxygen (85 % v/v) and carbon dioxide (1% v/v). The dissolved oxygen concentration was maintained in the range 1-4 mg L^{-1}. The hydraulic residence time was 48 hours, the pH 1.5, and the redox potential 700 mV (Ag vs. Ag/Cl in 3M KCl). Phosphate and ammonia were supplemented to maintain a concentration of 10 and 50 mg L^{-1} respectively in solution. Other nutrients were derived from impurities associated with the mineral concentrated and were not added to the reactors. The suspended cells concentration was approximately 10^9 cells mL^{-1}. Although non-aseptic conditions were used, the microbial population in the steady state reactor comprised >90% *Sulfolobus metallicus*, reflecting the selective growth conditions imposed by the reactor operating conditions. The reactor effluent, containing both mineral and biomass suspended solids, was subjected to two-stage centrifugation. Mineral solids were removed during centrifugation at 500 g for 5 minutes, while the archaeal biomass was recovered at 2000 g for 5 minutes in a Sharples centrifuge. Freeze-drying of the archaeal biomass was carried out over a 48 h period in a Virtis Genesis freeze-dryer.

2.2. Laboratory-scale lipid extraction

2.2.1. Ashing

Extractions were carried out on freeze-dried biomass material obtained as described above. In order to ascertain accurately the lipid yield from each extraction, it was necessary to determine the amount of biomass present in each batch of the solid starting material, defined as the total mass of carbonaceous material present. Percentage biomass was determined by accurately weighing a sample of approximately 2 g of freeze-dried material and ashing overnight in a furnace at 650°C to remove all biomass present. The same sample was subjected a second time to overnight heating at 650°C to ensure that a constant mass had been reached. Carbonaceous material is volatilised at this temperature and all that remains after ashing is inorganic matter. In this way the percentage biomass, from which lipid could be extracted, was determined. ICP analysis was performed on the non-carbonaceous material remaining after ashing.

2.2.2. Extraction

The initial procedure followed was based on that described for *Sulfolobus acidocaldarius* (Lo et al., 1989), in which first a neutral and then an acidic Soxhlet extraction is carried out on freeze-dried material. This Soxhlet extraction was found to be suitable for use with freeze-dried material ranging in mass from 5-70 g and is illustrated here for a 25 g sample. Freeze-dried material (25 g) was extracted overnight in a Soxhlet apparatus with $CHCl_3$: MeOH (1:1) (300 ml) and the extraction solution (I) was put aside. The solid material was then stirred for 2 h in 280 ml of a mixture of $CHCl_3$: MeOH: 5% trichloroacetic acid (TCA) (1:2:0.8). The slurry was centrifuged and the supernatant added to the original extract (I). The solid precipitate was washed with 100 ml of the acid extraction medium used above. This washing was also added to the original extract (I). Chloroform (750 ml), methanol (150 ml) and water (300 ml) were added to extract solution I to make a final solution of $CHCl_3$: MeOH: H_2O of roughly 8:4:3. The two phases were mixed and separated into the upper phase (II) and the lower phase (III). The upper phase (II) was extracted with an equal volume (800 ml) of $CHCl_3$. This chloroform layer (IV) was set aside for evaporation. The lower phase from the original separation (III) was washed with an equal volume (1000 ml) of a $CHCl_3$: MeOH: H_2O (3:48:47) mixture. The lower phase here (V) was washed with the upper phase (II) from the original separation. The lower phase after this final wash was added to the chloroform layer previously set aside (IV) and this was evaporated.

A number of variations on this general extraction procedure were tested with respect to ease of extraction and final lipid yield: wet paste instead of freeze-dried material was extracted directly using the Soxhlet method described above; neutralisation of wet paste in order to lyse cells was carried out prior to lyophilisation and extraction (the pH of the wet paste was adjusted to 7 with 25% NH_4OH prior to drying) and finally a separate work-up of the acidic and neutral extractions was carried out.

2.2.3. Hydrolysis

To the material obtained after evaporation was added 30 ml of 1 M methanolic HCl (3.41 ml 32% HCl in 30 ml MeOH) and this was heated overnight at reflux (75ºC). The pH was adjusted to pH 14 with 8M KOH, after addition of water (20 ml). This was heated at 80°C for 1 h. The pH was adjusted to 3 with conc. HCl and $CHCl_3$ (60 ml) was added. The chloroform layer was separated and then washed with 60 ml of a $CHCl_3$: MeOH: H_2O (3:48:47) mixture. The chloroform layer was separated, dried over $MgSO_4$ and evaporated.

2.3. Development of a bench-scale lipid extraction process

Soxhlet extractions were found to be suitable for amounts of freeze-dried starting material of up to approximately 70 g, but for larger amounts these extractions were found to be impractical. Thus, a method was developed that would be suitable for amounts upwards of 500 g of starting material.

2.3.1. Extraction and hydrolysis

2.3.1.1. Combined neutral-acid extraction

Freeze-dried biomass (25 g) was transferred to a 1 L Büchi rotary evaporator flask and to this was added 400 ml of a $CHCl_3$: MeOH: 5% TCA mixture (1:2:1). This mixture was rotated on the rotary evaporator at atmospheric pressure for 2 hours at a gentle reflux temperature of 60°C. The mixture was then cooled to 25°C and filtered. The filter cake was transferred back to the 1 L Büchi flask to which was added a fresh mixture of solvent with the same composition as before. This procedure was continued for 3 more cycles. In total, 4 extractions were conducted. At the end of the fourth cycle the filtrates were combined and separated into two layers on standing. The upper layer was a green colour, probably due to the presence of nickel salts liberated during the bioleaching process, whereas the lower layer was a clear, brown colour. A white/yellow precipitate was noticeable at the interface, probably denatured protein material. The upper layer was extracted with a mixture of MeOH: H_2O (1:1, 800 ml). The combined chloroform extracts were distilled under vacuum to leave a dark brown oily residue.

1M Methanolic hydrochloric acid (100 ml) was added to the residue and this was heated at 80°C for 16 hours. The reaction mixture was then cooled to 25°C and water (100 ml) added. The pH of the resultant mixture was adjusted to 14 using 8M KOH. The mixture was then subjected to base hydrolysis at 80°C for 1 hour. The mixture was then cooled to 25°C and adjusted to pH 3 with 32% HCl. The resultant mixture was extracted with $CHCl_3$ (3 x 200ml). The chloroform extract was separated, dried with magnesium sulphate and distilled under vacuum to yield a brown lipid residue.

2.3.1.2. Neutral followed by acidic extraction

Freeze-dried biomass (100g) was transferred to a 2 L Labmax® reactor and to this was added 1 L of a $CHCl_3$: MeOH mixture (4:1). The mixture was stirred for 2 hours at 300 rpm and 60°C. The mixture was then cooled to 30°C and filtered. The slightly wet filter cake was transferred back to the reactor and fresh solvent added with the same composition as before. The extraction was continued as per the first extraction. The process was repeated once more resulting in a total of three neutral extractions.

After the third extraction, the filter cake was transferred back to the reactor and extracted with 800 ml of a $CHCl_3$: MeOH: 5% TCA mixture (1:2:1) for 2 hours at 60°C and 300 rpm. The mixture was then cooled to 30°C and filtered. The neutral extracts were combined to give a mixture of $CHCl_3$: MeOH (2400: 600 ml). To this was added MeOH: H_2O (600:900 ml) to make a final mixture of 8:4:3 of $CHCl_3$: MeOH: H_2O. The lower chloroform layer was drained and the upper layer was extracted with chloroform (1000 ml). Chloroform: H_2O (600: 100 ml) was added to the acidic filtrate to make a final mixture of 8:4:3 of $CHCl_3$: MeOH: H_2O. The lower chloroform layer was drained and the upper phase was extracted with chloroform (800 ml). All the chloroform layers (from both neutral and acid extracts) were combined and distilled under vacuum at 50°C to yield a dark brown residue.

The above residue was hydrolysed with 1M methanolic HCl for 16 hours at 80°C. The mixture was then cooled to 25°C and water (100 ml) added. The pH was then adjusted to 12 with 8M KOH. This mixture was hydrolysed for 1 hour at 80°C. The pH was then adjusted to 3 with 32% HCl. Chloroform: methanol (270: 30 ml) was then added to make a final mixture of 8:4:3 of $CHCl_3$: MeOH: H_2O. The lower chloroform phase was drained and the upper phase was extracted with chloroform (3 x 300 ml). The chloroform was distilled under vacuum at 50°C to yield a crude lipid residue.

2.3.1.3. Bench-scale extraction, neutral followed by acidic extraction

The process described above (2.3.1.2) was scaled five times. The extraction of 500 g of freeze-dried biomass was carried out in a 15 L Büchi GR15 reactor and the solvent make-up, separations and extractions were carried out in a Büchi CR26 reactor. Distillation of chloroform was carried out in a 20 L rotary evaporator. Hydrolysis of extracted material was carried out as described previously.

2.4. Isolation

In each case the material obtained from hydrolysis of the total lipid extract was subjected to gravity silica gel column chromatography. The eluents used were $CHCl_3$ followed by $CHCl_3$: diethyl ether (8:2) and finally $CHCl_3$: MeOH (8:2) (after Lo et al., 1989). To illustrate: for an 8 g hydrolysate sample (from extraction of 500 g solid material), a mass of 300 g silica gel 60 (0.063-0.200 mm) was used and fractions were eluted with 1.5 L of each eluent. Thin-layer chromatography (TLC) on silica gel 60 F_{254} plates, using phosphomolybdate reagent (Stahl, 1969) for compound visualization, was used to monitor elution of fractions. ^{1}H-NMR (200 MHz, Varian) and ^{13}C-NMR (50 or 100 MHz, Varian) spectroscopy was used to identify the eluted fractions, where possible. Final purification of the hydrolysed polar lipid (GDNT) was performed using silica gel flash chromatography, with elution using $CHCl_3$: MeOH (9:1). To illustrate, for a 2.7 g crude hydrolysed tetraether lipid sample, 250 g silica gel 60 (0.040-0.063 mm) was used and approximately 1.5 L of solvent was required for sample elution. Lipid material obtained was analyzed by mass spectrometry and ^{1}H and ^{13}C-NMR spectroscopy.

2.5. Analysis methods

High resolution electrospray mass spectrometry for accurate mass determination was performed on a Waters API Q-TOF Ultima instrument operating in ES+ ionisation mode. Chemical ionisation mass spectrometry was performed on a Micromass VG-70SEQ, operating at 1000 kV. All proton NMR spectra were run on a Varian Gemini 2000 NMR spectrometer at 200 MHz. ^{13}C NMR spectra were run on a Varian Unity Plus NMR spectrometer at 100 MHz or on a Varian Gemini 2000 NMR spectrometer at 50 MHz. The COSY, HMQC and HMBC experiments were run on a Varian Unity Plus where ^{13}C is observed at 100 MHz and proton at 400 MHz. Samples were made up in $CDCl_3$ and chloroform peaks referenced at 77.00 ppm for ^{13}C and 7.26 ppm for ^{1}H.

3. Results and Discussion

3.1. Laboratory-scale extraction and hydrolysis

Results of all laboratory-scale extractions performed are summarized in Table 1. For all the extractions the crude lipid yield as a percentage of biomass was 0.4-0.5%. Direct extraction of wet paste gave similar results (adjudged by TLC) to that of freeze-dried material but the amount of wet paste extractable at one time was very low compared to what was possible on dried material. In the case where neutralization was performed prior to extraction, results were comparable with those obtained without neutralization so there did not appear to be any advantage associated with this approach. Separation of the work-up for acidic and neutral extractions was problematic in that the neutral extraction resulted in an emulsion that could only be broken using centrifugation. Apart from this, the separate neutral extraction gave a yield comparable to the best acid-base combined extraction. Results from analysis of the residue obtained after ashing showed this material to be composed predominantly of Fe and Ni, a result consistent with the conditions inside the bioleaching reactor.

3.2 Development of a bench-scale process

The laboratory-scale procedure for the extraction of dry biomass requires the use of a Soxhlet extraction. While this type of extraction is suitable for small amounts of solids it is not scalable if larger amounts of material need to be extracted. Furthermore, it is a lengthy procedure and the extracted material is kept at the boiling point of the solvent for this entire period. If the extracted material is not stable, decomposition could be accelerated. The extraction process itself may not be efficient since it is a static procedure and channels may be formed in the solid, which means that some portions of the solid may not be exposed to the solvent. Therefore, alternative procedures were investigated for the extraction of biomass that would be suitable for scale-up. The crude tetraether lipid yield was used as an indication of the effectiveness of the various procedures. The results of these experiments are given in Table 2.

In the first extraction procedure (extraction A), biomass was extracted using a rotary evaporator in reflux mode. In this procedure the freeze-dried biomass was rotated at 150 rpm while being in intimate contact with warm solvent. The neutral and acidic extractions were combined and carried out as one unit operation. The crude lipid yield was 0.5%, comparable to that obtained using the Soxhlet apparatus (0.4-0.5%).

During the downstream processing (DSP) of the extracts, it was found that a significant amount of gelatinous material was extracted and this made filtration and separation of the phases difficult. Furthermore, ultimate scale-up of this process would be limited by rotary evaporator size.

In the second procedure (extraction B), the extraction was carried out at gentle reflux in a 2 L Labmax® reactor fitted with an anchor stirrer, recommended for slurry reactions. The biomass was first extracted with $CHCl_3$: MeOH (80: 20%, ~azeotropic mixture, 3 times) followed by one acidic extraction. The chloroform extracts from both treatments were combined and subsequent hydrolyses and purifications were carried out on the combined residues. The overall yield was 0.5%, once again comparable to the Soxhlet extraction. The DSP of the neutral extracts improved significantly since the solvent extracts filtered easily and phase separations were quick and clear, without emulsion formation. However, the acidic extraction produced significant amounts of gelatinous material, making DSP difficult and slow.

The Labmax® procedure above was scaled and carried out in a 15L Büchi GR15 reactor, using 500 g of dry biomass. The agitator in this reactor is a paddle-type stirrer and stirring is very efficient, even for slurry reactions. The crude lipid yield was 0.7%, representing an improvement over the previous protocols used. As with the previous reaction, the DSP of the neutral extracts was significantly better than for the acidic extract. Thus, bench-scale extractions of GDNT proved to be at least as good as, or better than laboratory-scale procedures.

3.3. Lipid purification

The first fractions eluted from gravity silica column chromatography with chloroform contained material with a ^{1}H-NMR spectrum closely resembling that of the tetraether lipid fractions but with a significantly higher integral ratio of alkyl protons to protons attached to oxygen-carrying carbons and OH protons (30:1) than expected for the hydrolysed tetraether lipids GDNT or GDGT. This fraction most likely contained free alkyl chains, either extracted from the biomass or resulting from cleavage of tetraether lipid material during extraction and hydrolysis. Elution with $CHCl_3$: diethyl ether yielded the less polar hydrolysed lipid fraction (GDGT), identified using ^{1}H and ^{13}C-NMR spectroscopy (Figure 1) and CI mass spectrometry. Elution with $CHCl_3$: MeOH gave a fraction containing some of the less polar lipid (GDGT) as well as the hydrolysed polar lipid fraction (GDNT), identified using ^{1}H and ^{13}C-NMR spectroscopy and high resolution electrospray mass spectrometry. This dark brown fraction containing both tetraether lipids as well as small amounts of other contaminants is what is referred to as the crude tetraether lipid fraction (Tables 1 and 2). Thin layer chromatography of this fraction on silica gel plates (silica gel 60 F_{254}) shows that most of the material is GDNT (around 70%) with an $R_f = 0.28$ in $CHCl_3$: MeOH (9:1); GDGT is also present, with an $R_f = 0.96$ in $CHCl_3$: MeOH (9:1). Results from chromatography show that the amount of GDNT extracted is roughly 5-10 times higher than that of GDGT. De Rosa et al. (1977a) report isolation of two hydrolysed tetraether lipids from thermoacidophilic bacteria with characteristic R_f values on TLC plates in $CHCl_3$: MeOH (9:1) of 0.95 and 0.35, which correlates with our findings. These were identified as glycerol dialkyl glycerol tetraether (GDGT) and glycerol dialkyl nonitol tetraether (GDNT), respectively. Lo et al. (1989) report an R_f value of 0.45 for GDNT in $CHCl_3$: MeOH (9:1), prior to final purification.

The crude tetraether lipid fraction was further purified by flash chromatography on silica gel [elution $CHCl_3$: MeOH (9:1)] to give purified GDNT, one spot on TLC and with ^{1}H and ^{13}C-NMR spectra matching the GDNT structure (Figure 2). Flash chromatography removes much of the dark colour of the crude tetraether lipid fraction, the coloured material elutes first from the column, followed by GDNT, resulting in a purified honey-coloured GDNT (Figure 3). This method of purification is considerably simpler than that described by Lo et al. (1989), where initial silica column chromatography is followed by three different preparative thin layer chromatography steps and then by acetone precipitation. Our method involves initial purification by gravity silica column chromatography followed by one, or at most two, purifications by flash silica chromatography to give pure GDNT. In addition, our method is readily scalable, giving rise to gram-quantities of GDNT.

The best yield of crude tetraether lipid (0.7% based on biomass) was obtained from the extraction of biomass on a 500 g scale (extraction C), see Table 2. This batch also gave the best yield of purified GDNT, 0.37% based on biomass. Thus from 500 g of freeze-dried starting material, with a biomass content of 78% was obtained 2.7 g of crude lipid and 1.44 g pure GDNT.

3.4 Analysis of lipids

3.4.1. Mass spectrometry

High resolution electrospray mass spectrometry of the purified GDNT fraction showed the presence of three compounds. The conditions under which the mass spectrometry was performed led to the base peaks being visible as Na-adducts, with the M+1 ion as well as the K-adducts being seen (Figure 4, above). The isotopes of the first compound interfered with the accurate mass determination of the second and third compounds in the sample and thus isotope modelling was used to assist in the assignment of formulae for the second and third compounds (Figure 4, below). The first three panels in this figure (Figure 4, below) show the individual isotope models for the three different compounds and the fourth panel in this figure shows how these combine to give a pattern that corresponds to the spectrum actually observed (Figure 4, above). The three compounds appear to be present in the GDNT fraction in very similar concentrations, as estimated by the isotope patterns (Figure 4, below). The first compound gave a base peak with an accurate mass of 1478.297 (Figure 5), with the calculated value for $C_{92}H_{174}O_{11}Na$ being 1478.295. The accurate mass of the M+1 ion was 1456.314 (Figure 5), with the calculated value for $C_{92}H_{175}O_{11}$ being 1456.313. This gives a compound with the formula $C_{92}H_{174}O_{11}$ which corresponds to a GDNT structure containing 4 cyclopentane rings (two per chain), shown in Figure 6. The second and third compounds differed by mass units of two and four and correspond to the formulae $C_{92}H_{176}O_{11}$ (three cyclopentane rings) and $C_{92}H_{178}O_{11}$ (two cyclopentane rings), respectively. Isolated GDNT most often occurs as a mixture of closely related compounds differing only in the number of cyclopentane rings present in the alkyl chains, the exact number of cyclisations being dependent on the temperature at which the organism grows. The positions of the cyclisations shown in Figure 6 are based on compounds reported by De Rosa et al. (1983), as well as on the NMR data. The structure of the calditol moiety is after Sugai et al. (1995), who synthesised all possible isomers of calditol to determine the correct stereochemistry.

Low resolution chemical ionization mass spectrometry was performed on GDGT as it has a lower molecular mass than GDNT, for which this technique was unsuccessful. The M^+ peak (in the region of 1290) was observed and fragmentation of the molecule was observed as the loss of isoprene units from the parent compound.

3.4.2 ^{1}H and ^{13}C-NMR spectroscopy

^{1}H-NMR spectra of the two hydrolysed tetraether lipids GDGT and GDNT are very similar and show predominantly signals appearing between 0.7 and 1.9 ppm, which correspond to alkyl chain protons and signals appearing between 3 and 4 ppm, which correspond to protons on carbon atoms attached directly to oxygen or OH protons. The ratio of the integral for the alkyl protons to the integral for the protons on oxygen-bearing carbon atoms and OH protons was found to be useful in determining which lipid was present. The less polar lipid (GDGT) gave a ratio of around 10:1, while a lower ratio of approximately 5:1 was observed for the polar lipid (GDNT).

^{13}C-NMR analysis of purified GDNT at 100 MHz, gave the requisite number of signals (16) for the C-O region between 55 and 90 ppm, see Table 3 and Figure 7. This is the region of the spectrum where signals resulting from the headgroups appear. From the DEPT spectrum (Figure 8) 15 proton-bearing carbons were observed (9 x CH_2 and 6 x CH), leaving one quaternary carbon. De Rosa and Gambacorta (1988) observed only 8 signals in the C-O region for GDNT. They assigned these 8 and indicated that a ninth carbon gave an overlapping signal. They did not observe any carbon signals for the calditol portion of the molecule, but gave no explanation for their absence. De Rosa et al. (1980) were only able to observe ^{13}C signals corresponding to the calditol carbons as the fully acetylated derivative of GDNT, but these signals were assigned to the incorrect open-chain nonitol structure. The correct calditol structure was determined by Sugai et al. (1995) and they reported ^{13}C chemical shifts for the free calditol moiety, cleaved from GDNT, but made assignments only for the acetylated calditol. To our knowledge, this is the first report of all 16 expected signals of underivatised GDNT being observed in the C-O region. Our assignments for signals in the C-O region of native GDNT (Table 3) are made based on heteronuclear multiple quantum coherence (HMQC, showing carbon-proton correlations for directly attached protons, Figure 9), heteronuclear multiple bond connectivity (HMBC, showing 2-3 bond carbon-proton correlations, Figure 10) and distortionless enhancement by polarisation transfer (DEPT, distinguishing between carbons carrying one, two or three protons, Figure 8) experiments, together with chemical shift calculations. Comparisons have also been made with some of the original assignments of De Rosa et al. (1988). These authors found that wherever cyclisation occurred at C-3 of the alkyl chain (starting from either end), the ^{13}C signal for the first carbon of that chain shifted to values higher by 1.1-1.2 ppm, compared to where cyclisation occurred at C-7 (as depicted in Figure 6) or where no cyclisation occurred. Our results for carbons designated **d**, **e**, **f** and **g** (Figure 6) are in agreement with those of De Rosa et al. for alkyl chains having no cyclisation at position C-3, but with possible cyclisation at C-7 or no cyclisation. Mass spectrometry results indicate that cyclisations are present in our compounds and thus they must be placed at position C-7, as depicted in Figure 6. The presence of cyclisation is also confirmed by carbon shifts present in the alkyl region, further discussed below. Thus, for carbons designated **a**-**g** (Table 3 and Figure 6) our results confirm those obtained by De Rosa

et al. (1988) but they were unable to obtain a discrete signal for carbon **h** and were unable to find any signals corresponding to carbons **j**-**p**. Our assignments for carbons **l**-**n** are in agreement with results obtained on some synthetic derivatives of calditol prepared by Bleriot et al. (2002), who established the absolute configuration of calditol from *Sulfolobus*. A COSY experiment was also performed but did not prove to be particularly useful because of overlapping signals that result from the large number of protons present in the molecule and the small area of the spectrum over which the signals appear.

Interestingly, our first ^{13}C NMR spectrum of GDNT, acquired from a slightly less pure sample, showed only 8 signals in the C-O region. Increasing the relaxation delay time from d_1=0 s to d_1=10 s caused peaks corresponding to the calditol moiety to start emerging from the baseline of the spectrum. It appears that in our slightly less pure sample of GDNT, something present in the sample was delaying the relaxation of these carbon atoms, possibly chelation of metals with the oxygen-rich calditol. When the sample was further purified by flash chromatography, the calditol peaks emerged without the relaxation delay time having to be adjusted and these peaks were fully visible at d_1=0 s. Fine splitting of three peaks in the C-O region of the ^{13}C spectrum was observed at 100 MHz that wasn't observed at 50 MHz. This splitting is possibly the result of the purified GDNT being present as a number of closely related compounds, differing only in the number of cyclopentane rings in the alkyl chains.

The region of the GDNT ^{13}C spectrum between 15 and 45 ppm is shown in Figure 11. This region contains signals from the alkyl chains stretching between the two headgroups. Assignments are shown on the spectrum (Figure 11) for the methine and methyl carbon signals. One of the three GDNT compounds has been chosen for illustrative purposes but obviously the principles also apply to the other two GDNT components. Our results are largely in agreement with those obtained by De Rosa et al. (1977b) on spectra of the isolated alkyl chains that they obtained by cleavage of the tetraethers with HI. Slight differences might be expected, particularly towards the chain ends, as our spectra originate from intact GDNT while theirs are from alkyl groups cleaved from their headgroups. One notable difference is the shift of the carbons at position **r**, where our signals are observed at around 29 ppm, while De Rosa et al. report a shift of 34 ppm. The signals at 44 ppm are diagnostic for the presence of cyclisations in the chain and originate from carbons designated **t**. The other methine carbon of the cyclopentane ring appears at 39 ppm (**s**). The methyl carbons immediately adjacent to the cyclopentane rings give signals around 17 ppm (designated **v**), while the methine carbons attached to these methyls give shifts of 38 ppm (**u**). In contrast to this, those methyl carbons not immediately adjacent to cyclopentane rings give signals at around 19 ppm (designated **q**), with the attached methine carbon signals appearing at around 33 ppm for those in the centre of the chain (**w**) or 29 ppm for those near the chain termini (**r**). Certain of the CH_2 signals that are discrete may be assigned, for example shifts of the eight methylene carbons that are bordered on each side by other methylene groups appear at 24 and 25 ppm, but as might be expected from a large molecule such as GDNT containing long alkyl chains, there are a large number of overlapping CH_2 signals, particularly around 36-37 ppm, which makes individual assignments of these signals impossible at 100 MHz. However, the ^{13}C spectrum in this alkyl region (15-45 ppm) is in fact simpler than might be expected from a compound containing 76 carbons giving signals in this region and this can be explained by the fact that the two alkyl chains in GDNT are at

least pseudo-symmetrical, a lack of total symmetry being a result of the different headgroups: glycerol and calditol.

^{13}C-NMR analysis of GDGT gave the expected signals in the region between 55 and 90 ppm (Figure 12) where signals resulting from the headgroups are found. The two headgroups (glycerol moieties) in GDGT are equivalent and as such, only 5 signals are present in this region as a result of the symmetry of this compound.

4. Conclusion

The bench scale extraction and isolation of high-purity GDNT, harvested from bioleaching reactors, has successfully been demonstrated in this study. The isolation procedure described to obtain pure GDNT is significantly simpler than previously described methods. In addition, previously reported methods have not been scalable and were performed at milligram scale only. The identity of the isolated material has been confirmed by mass spectrometry and NMR. All expected signals in the C-O region of the ^{13}C NMR spectrum of native GDNT have been observed for the first time. This study paves the way for scale-up of the extraction and purification process to be used in tandem with industrial scale hydrometallurgical bioleaching reactors. The large-scale and low incremental cost of biomass harvesting is likely to create the commercial possibility of producing a high-quality GDNT lipid at a substantially reduced cost. Such lipids, reliably produced in relatively large quantities, could stimulate the wide variety of tetraether lipid applications based upon their unique properties. In turn, a market demand for these lipids could provide a substantial additional revenue stream and thus influence the competitive advantage of thermophilic tank bioleaching technology.

Acknowledgements

The authors wish to thank Dr Madelyn Bekker and Dr Marietjie Stander for acquisition of NMR and mass spectra, respectively.

References

Bagatolli, L., Gratton, E., Khan, T.K. and Chong, P.L-G. 2000. Two-phyton fluorescence microscopy studies of bipolar tetraether giant liposomes from thermoacidophilic archaebacteria *Sulfolobus acidocaldarius*. Biophysical Journal 79, 416-425.

Bakowsky, U., Rothe, U., Antonopoulos, E., Martini, T., Henkel, L. and Freisleben, H-J. 2000. Monomolecular organization of the main tetraether lipid from *Thermoplasma acidophilum* at the water-air interface. Chemistry and Physics of Lipids 105, 31-42.

Batty, J.D. and Rorke. G.V. 2006. Development and commercial demonstration of the BioCOP™ thermophile process. Hydrometallurgy 83, 83-89.

Benvegnu, T., Brard, M. and Plusquellec, D. 2004. Archaeabacteria bipolar lipid analogues: structure, synthesis and lyotropic properties. Current Opinion in Colloid and Interface Science 8, 469-479.

Bernina Biosystems GmbH, 2001. Tetraether lipid derivatives and liposomes and lipid agglomerates containing tetraether lipid derivatives, and use thereof. United States Patent 6,316,260 B1. 13-11-2001.

Blériot, Y., Untersteller, E., Fritz, B. and Sinaÿ, P. 2002. Total synthesis of calditol: structural clarification of this typical component of Archaea order *Sulfolobales*. Chem. Eur. J. 8, 240-246.

De Rosa, M., De Rosa, S., Gambacorta, A. and Bu'Lock, J.D. 1977a. Lipid structures in the Caldariella group of extreme thermoacidophile bacteria. J. Chem. Soc., Chem. Comm. 514-515.

De Rosa, M., De Rosa, S. and Gambacorta, A. 1977b. ^{13}C-NMR assignments and biosynthetic data for the ether lipids of *Caldariella*. Phytochemistry, 16, 1909-1912.

De Rosa, M., De Rosa, S., Gambacorta, A. and Bu'Lock, J.D. 1980. Structure of calditol, a new branched-chain nonitol, and of the derived tetraether lipids in the thermoacidophile archaebacteria of the *Caldariella* group. Phytochemistry 19, 249-254.

De Rosa, M., Gambacorta, A., Nicolaus, B., Chappe, B. and Albrecht, P. 1983. Isoprenoid Ethers; backbone of complex lipids of the Archaebacterium *Sulfolobus solfataricus*. Biochimica et Biophysica Acta 753, 249-256.

De Rosa, M. and Gambacorta, A. 1988. The lipids of Archaebacteria. Prog. Lipid. Res. 27, 153-175.

Dote, J.L., Barger, W. R., Behroozi, F., Chang, E. L., Lo, S. –L., Montague, C. E. and Nagumo, M. 1990. Monomolecular film behaviour of tetraether lipids from a thermoacidophilic archaebacterium at the air/water interface. Langmuir 6, 1017-1023.

Frant, M., Stenstad, P., Johnsen, H., Dölling, K., Rothe, U., Schmid, R. and Liefeith, K. 2006. Anti-infective surfaces based on tetraether lipids for peritoneal dialysis catheter systems. Materialwissenschaft und Werkstofftechnik, 37, 6, 538 – 545.

Gliozzi, A., Relini, A. and Chong, P.L.-G. 2002. Structure and permeability properties of biomimetic membranes of bolaform archaeal tetraether lipids. Journal of Membrane Science 206, 131-147.

Kanichay, R., Boni, L. T., Cooke, P.H., Khan, T.K. and Chong, P.L.-G. 2003. Calcium-induced aggregation of archaeal bipolar tetraether liposomes derived from the thermoacidophilic archaeon *Sulfolobus acidocaldarius*. Archaea 1, 175-183.

Kermamed Medizintechnik GmbH, Institut Für Bioprozess- Un Analysenmesstechnik E.V., Surface Interface Technologies GmbH, 2002. Biologically-functionalised, metabolically-inductive implant surfaces. World Intellectual Property Organization PCT Patent Application WO 02/085250 A2. 31-10-2002

Lo, S.-L.; Montague, C.E. and Chang, E.L. 1989. Purification of glycerol dialkyl nonitol tetraether from *Sulfolobus acidocaldarius*. Journal of Lipid Research 30, 944-949.

Norris, P.R., Burton, N.P. and Foulis, A.M. 2000. Acidophiles in bioreactor mineral processing. Extremophiles 4, 71-76.

Rawlings, D.E., Dew, D.W. and du Plessis, C.A. 2003. Biomineralization of metal-containing ores and concentrates. TRENDS in Biotechnology 21, 38-44.

Ryley, S., Chyla, A.T. and Peterson, I.R. 2000. An air-stable biomimetic Langmuir-Blodgett bilayer. Thin Solid Films 370, 294-298.

Sprott, G.D., Sad, S., Fleming, L.P., Dicaire, C.J., Patel, G.B. and Krishnan L. 2003. Archaeosomes varying in lipid composition differ in receptor-mediated endocytosis and differentially adjuvant immune responses to entrapped antigen. Archaea 1, 151-164.

Stahl, E. 1969. Thin-Layer chromatography A laboratory handbook, Springer-Verlag, Berlin.
Strobl, C., Six, L., Heckmann, K., Henkel, B. and Ring, K. 1985. Physicochemical characterization of tetraether lipids from Thermoplasma acidophilum. Z. Naturforsch. 40c, 219-222.
Sugai, A., Sakuma, R., Fukuda, I., Kurosawa, N., Itoh, Y.H., Kon, K., Ando, S. and Itoh, T. 1995. The structure of the core polyol of the ether lipids from *Sulfolobus acidocaldarius*. Lipids 30, 339-246.
United States Of America Navy, 1992. New class of lubricants derived from archaebacteria lipids. United States Patent 5,098,588. 24-03-1992.
Wetzer, B., Pum, D. and Sleytr, U.B. 1997. S-layer stabilized solid supported lipid bilayers. Journal of Structural Biology 119, 123-128.

Tables

Table 1. Yields from laboratory-scale extractions

Extraction	Mass extracted/g	Percentage biomass	Crude tetraether lipid[†] yield /mg	Crude tetraether lipid yield/ (% biomass)
Freeze-dried	25	71	74	0.4
Freeze-dried	10	71	37	0.5
Wet paste	16	22	Not done	~0.4[#]
Neutralized prior to extraction	6	71	18	0.4
Neutral extraction only	70	74	243	0.5

[†]mostly GDNT with some GDGT
[#]estimated from thin layer chromatography

Table 2. Yields from scale-up extractions performed

Extraction	Mass extracted/g	Percentage biomass	Crude tetraether lipid yield/mg	Crude tetraether lipid yield/ (% biomass)	Purified GDNT yield/ (% biomass)
A. Combined neutral-acid extraction	25	74	88	0.5	0.29
B. Neutral followed by acidic extraction	100	78	420	0.5	0.22

C. Scale-up of neutral followed by acidic extraction	500	78	2700	0.7	0.37

Table 3. Assignment of the ^{13}C NMR signals in the C-O region of GDNT

Atom	^{13}C shift/ppm
a	62.98 (t)
b	78.44 (d)
c and h	71.10, 71.07 (t) and 71.62 (t)
d and g	68.59 (t) and 68.82, 68.86 (t)
e and f	70.05 (t) and 70.14 (t)
i	77.84 (d)
j	70.54, 70.59 (t)
k	90.86 (d)
l[#]	81.24 (d)
m	75.03 (d)
n	76.15 (d)
o	80.46 (s)
p	63.04 (t)

[#] after Blériot et al., 2002

Figures

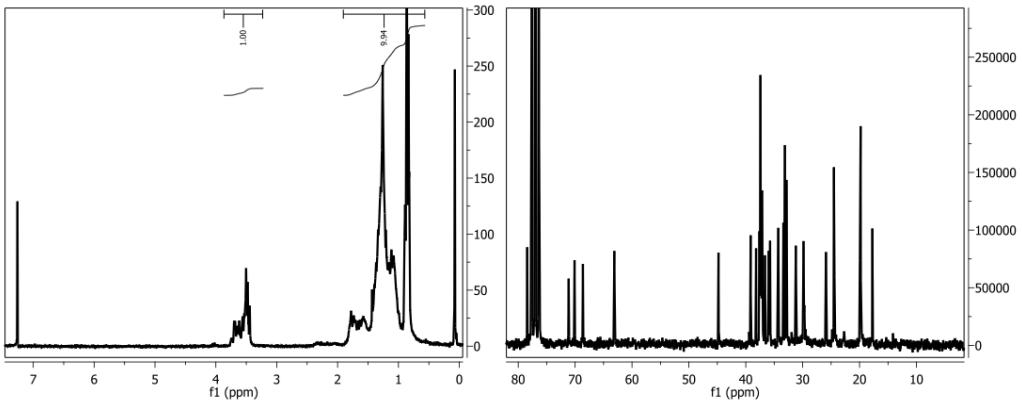


Figure 1. ^{1}H NMR (left) and ^{13}C NMR (right) spectra of GDGT

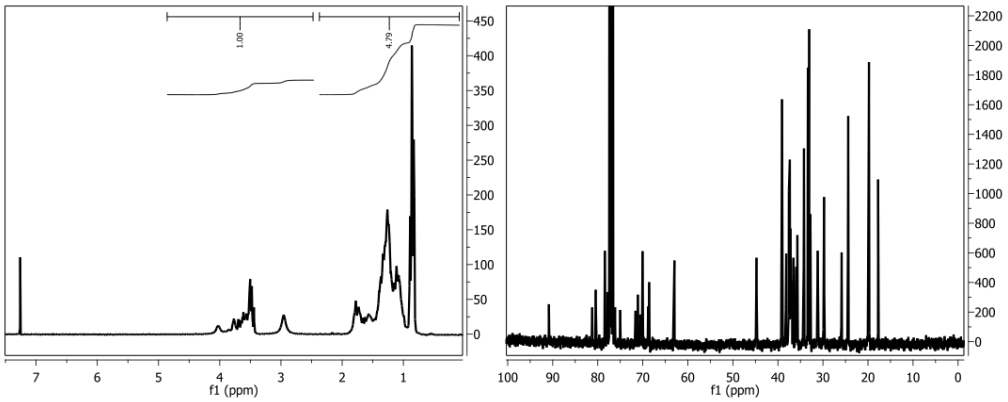


Figure 2. ^{1}H NMR (left) and ^{13}C NMR (right) spectra of GDNT

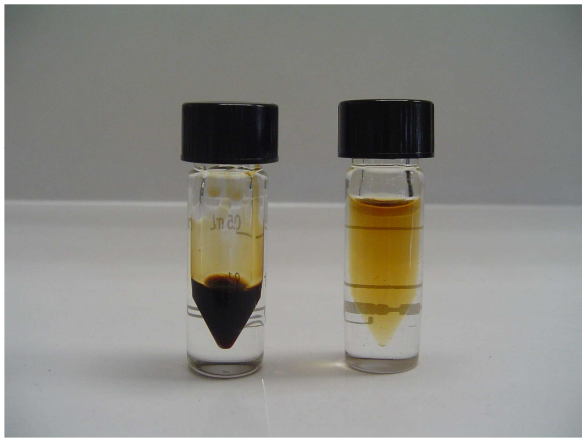
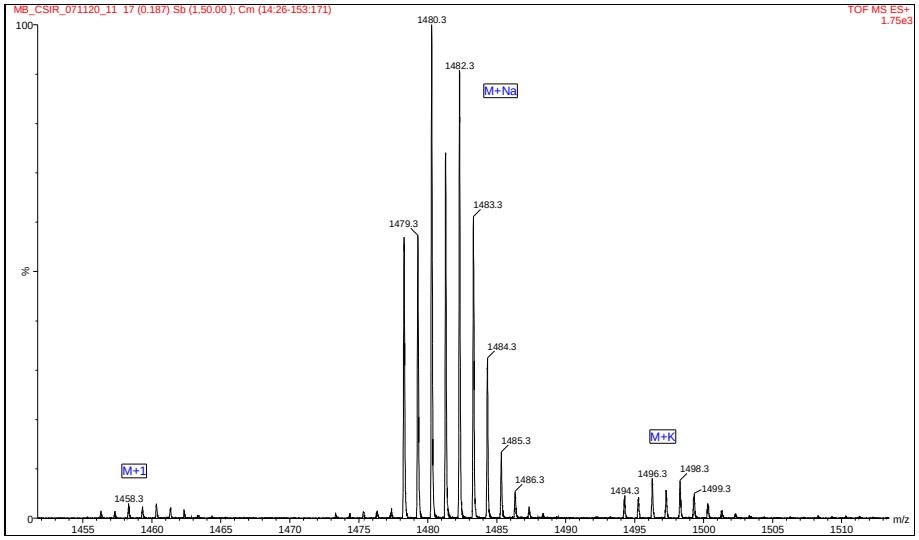


Figure 3. Crude (left) and purified (right) GDNT.



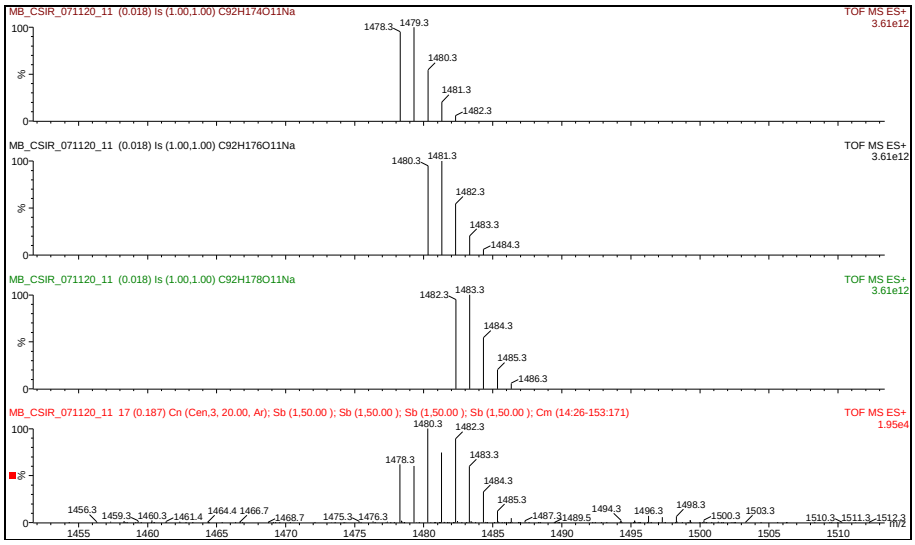


Figure 4. Electrospray mass spectrometry results for GDNT (above), showing isotope models of the three compounds present (top three panels, below) and how they combine to give a pattern (fourth panel, below) that corresponds to the observed mass spectrum (above).

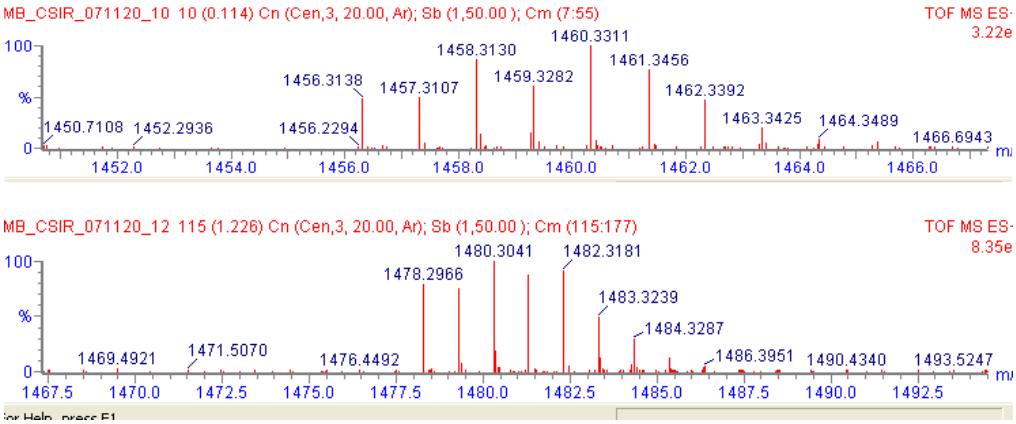


Figure 5. Accurate mass data for GDNT with the top panel showing the M+1 ion and the lower panel showing M+NaCl for $C_{92}H_{174}O_{11}$

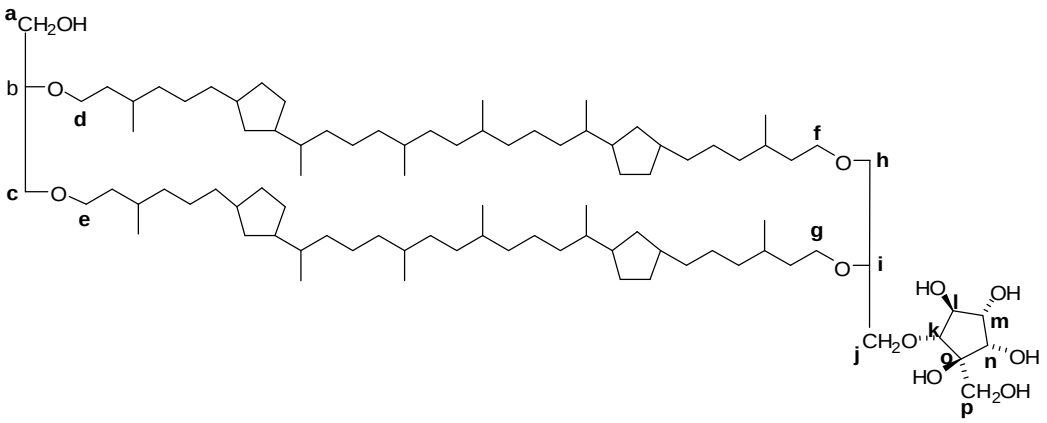


Figure 6. Backbone structure for GDNT ($C_{92}H_{174}O_{11}$) from *Sulfolobus metallicus*

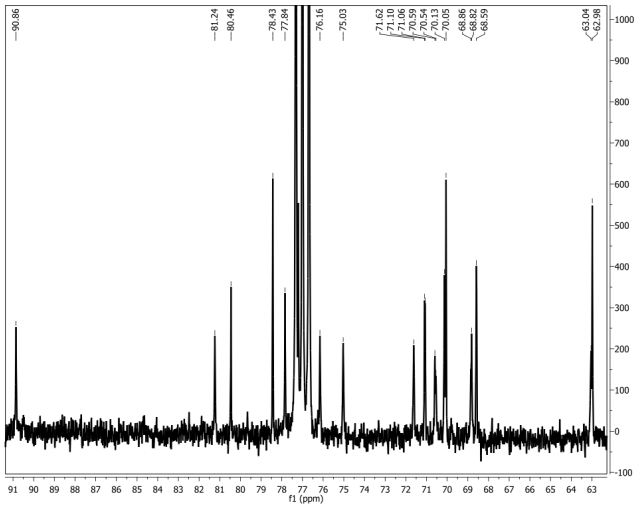


Figure 7. Expanded C-O region of the ^{13}C NMR spectrum of GDNT

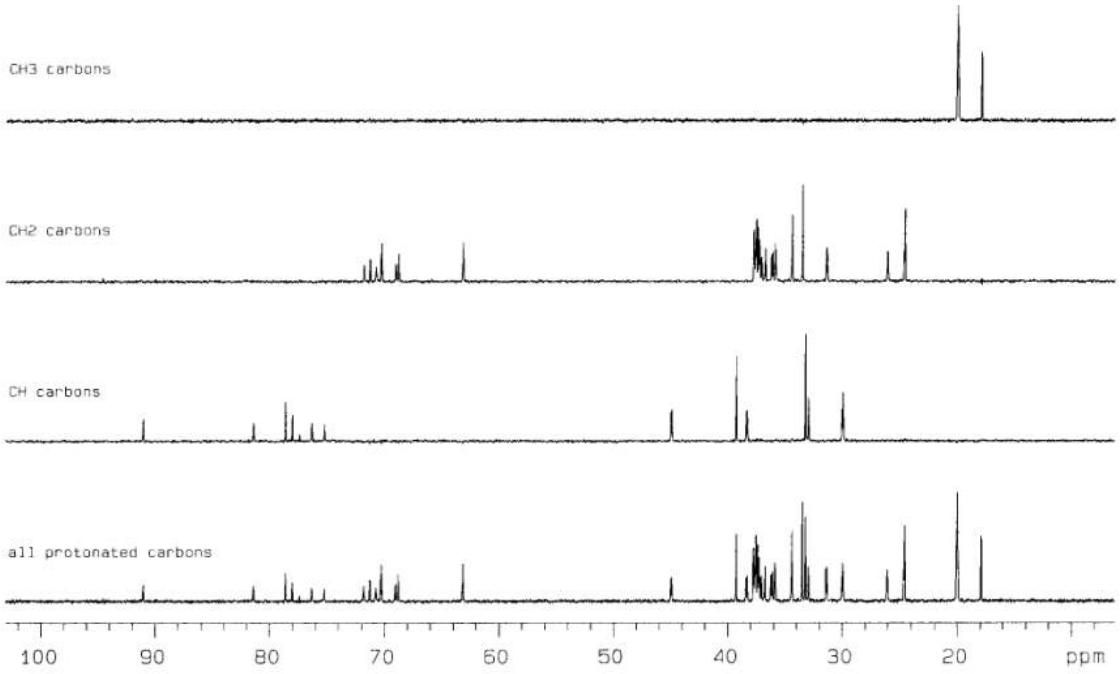


Figure 8. DEPT experiment showing peak multiplicities

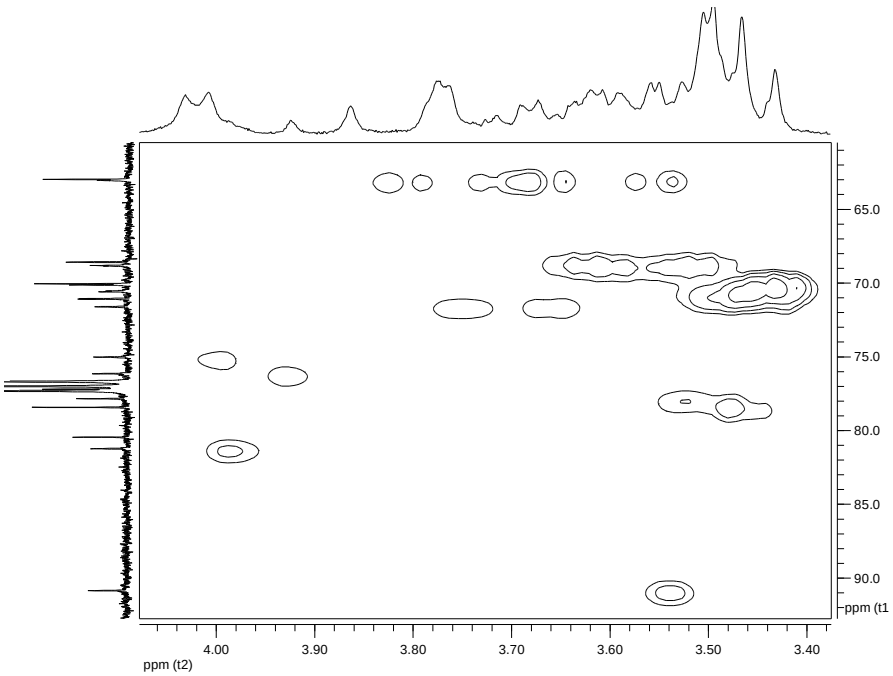


Figure 9. HMQC spectrum of the C-O region of GDNT

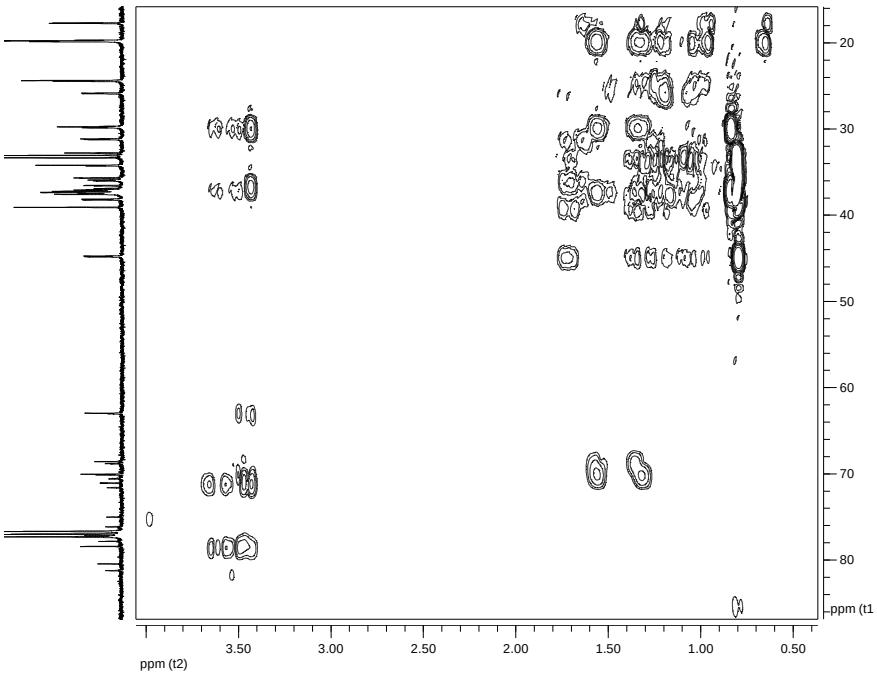


Figure 10. HMBC spectrum of GDNT

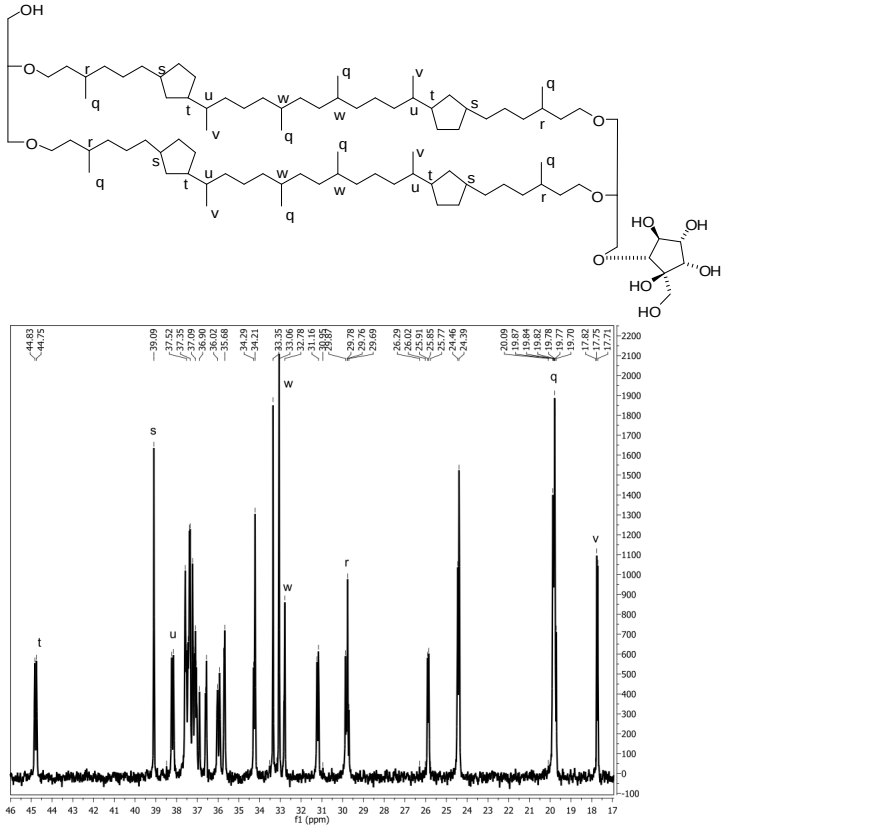


Figure 11. Expanded alkyl region of the ^{13}C NMR spectrum of GDNT, showing assignments of methine and methyl carbons

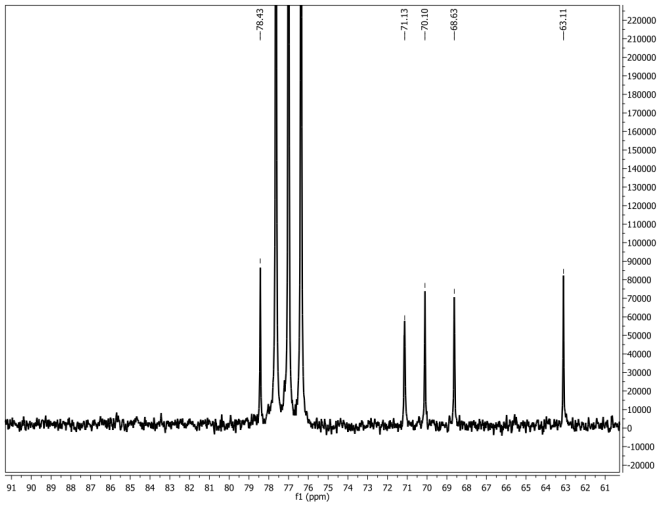


Figure 12. Expanded C-O region of the ^{13}C NMR spectrum of GDGT