Mass spectrometric studies of ether lipids in Archaea and sediments

By

Christopher Steven Knappy

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at the University of York

University of York
Department of Chemistry

October 2010

Abstract

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) has been implemented as a means to separate and detect tetraether lipid cores derived from the complex lipids of Archaea. Distinct dissociation pathways during tandem mass spectrometry were noted for the lipid cores, providing information regarding their structure on-line. Analysis of cellular material from species of *Methanothermobacter* and *Sulfolobus* revealed tetraether lipid cores which contain up to four cyclopentyl rings per etherified alkyl chain, including structures identified previously in each genera. Identical structures were similarly identified in novel isolates from New Zealand hot springs. Product ions in the MS/MS spectra of the lipid cores include those formed from individual losses of both ring-containing C_{40} alkyl chains, allowing the reported structures to be verified with respect to the distribution of the rings within the two chains. A number of additional, hitherto unreported isomers and higher homologues of the ring-containing structures were resolved, both chromatographically and/or by characteristic product ions in MS/MS. Structures in which the two chains appear to be conjoined by a covalent link were also identified in *Ignisphaera aggregans*, the first such identifications in a *Euryarchaeote*. The array of structures revealed highlights both the complexity of the archaeal lipidome, which is more extensive than has been attributed previously, and the potential of LC-MS/MS as a powerful tool for probing tetraether lipid core structure. Ether lipid cores extracted from ancient aquatic sediments and contemporary soil were used to investigate the scope of LC-MS/MS for profiling of extremely complex distributions sourced from ecological communities as opposed to single organisms. Over 100 ether lipid components in total were identified during the studies, the vast majority of which represent novel structures. These include isoprenoid lipid cores of known archaeal origin and structures which may represent their transformation products; triolic structures in which one of the two capping glycerol moieties has been lost and chain or glycerol methylated higher homologues. A wealth of non-isoprenoid lipid cores were similarly identified, with inferred structures suggestive of a eubacterial or mixed eubacterial/archaeal origin. The components, once constrained to more specific origins, may be of chemotaxonomic value for use in modern environmental profiling or in palaeoecological reconstructions made using fossilised lipid cores.

Author's Declaration

I hereby declare that the work described in this thesis is my own, except where otherwise acknowledged, and has not been submitted previously for a degree at this or any other university.

……

Contents

List of tables

List of schemes

List of figures

List of definitions and abbreviations

$\pm 1\sigma$	One standard deviation from the mean
16S rRNA	A small sub-unit of ribosomal RNA
$[a,b,c_d]$	The number of simultaneous losses of each type of molecule during formation of a product ion during CID, where a = water, b = C_3H_4O and c_d = biphytadiene containing "c" carbon atoms and "d" cyclopentyl rings
AcOH	acetic acid
Acrolein	propenal
Aeolian	By wind
Allochthonous	Formed in a place other than where they are found
Amplicon	Section of DNA formed from an amplification process
Anammox	Anaerobic ammonium oxidation; the conversion of ammonium and nitrite to nitrogen gas
ANME	Anaerobic methane oxidising (also methanotrophic)
Anteiso	Hydrocarbons methyl-branched at the ω3 chain position
AOM	Anaerobic oxidation of methane
APCI	Atmospheric pressure chemical ionisation
Archaeol	GDD lipid containing two ether-bound phytanyl chains
Authigenic	Formed where they are found
Autochthonous	Formed in the same place as they are found
Autotroph	Produces organic compounds from CO_2
Axenic	Not contaminated by or associated with any other organisms; pure culture
a_z	Dimensionless stability parameter; axial solution to the Matthieu equation

Benthic	Live in, on or near to the sea bed (the benthic zone)
Biphytanyl	3,7,11,15,18,22,26,30-octamethyldotriacontanyl; also used to describe a ring-containing derivative thereof
BIT index	Branched and isoprenoid tetraether index
Bolaamphiphilic	Molecule containing two polar termini at each end of a long hydrophobic chain
BPC	Base peak chromatogram
BSTFA	bis(trimethylsilyl) trifluoroacetamide
Caldarchaeol	GDGT lipid containing two ether-bound biphytanyl chains; used here specifically to describe GDGT **2a**
Calditol	A polyol utilised in GDCT lipids (see Fig. 1.4.); apparently exclusive to the members of the *Sulfolobales*
CARD-FISH	Catalysed reporter deposition FISH; a FISH process which utilises horseradish peroxidase-labelled oligonucleotides as tags and fluorescently labelled tyramide as reporter molecules
CBT	Cyclisation index of branched tetraethers
CE	% of maximum collision energy
Ch	cyclohexyl
Chemolithoautotrophic	Obtain energy from chemical reactions, utilise inorganic substances as electron donors and utilise carbon dioxide as a carbon source
CI	Chemical ionisation
CID	Collision induced dissociation
Cobalamins	Cobalt-containing compounds which contain the substituted ring and nucleotide structure of vitamin B12
Cp	cyclopentyl
Cren	*Crenarchaeote*

Crenarchaeol	GDGT lipid containing one biphytanyl chain containing two Cp rings and one chain containing two Cp and one Ch ring; the characteristic lipid core for Group I *Crenarchaeota*
Crenarchaeota	A kingdom of the Archaea
$\delta^{13}C$	A measure of the ratio of stable isotopes of carbon, C_{13}:C_{12}, typically quoted in parts per thousand (permil)
Da	daltons
DC	Direct current
DCM	dichloromethane
Decapod	An order of the crustaceans
Diabolic	The alkyl chain found in diabolic acid; 15,16-dimethyltriacontane
Diagenesis	Cumulative chemical, physical and/or biological alteration of chemical structure prior to or during the earlier stages of sedimentary burial
Dihydroxyarchaeol	GDD lipid containing two 3-hydroxy-3,7,11,15-tetramethylhexadecanyl chains
DNA	deoxyribonucleic acid
e	Elementary charge
EAOG	European Association of Organic Geochemists
EI	Electron ionisation
ELSD	Evaporative light scattering detection
Epipelagic	Living in surface waters (~0-200 m depth)
ESI	Electrospray ionisation
Eur	*Euryarchaeote*
Euryarchaeota	A kingdom of the Archaea
eV	Electron volt

Extreme halophile	Organisms which typically require a salinity of greater than ~2 M (NaCl) in order to survive
f	Frequency of the alternating potential applied to the end-cap electrodes
Facultative	Can survive under exposure to or using
FID	Flame ionisation detector
FISH	Fluorescence in-situ hybridisation; a process which hybridises specific DNA sequences with a fluorescent tag to enable detection and quantification of the sequence
FT-ICR	Fourier transform-ion cyclotron resonance
GC	Gas chromatography
GC-MS	Gas chromatography-mass spectrometry
GDCT	Glycerol dialkyl calditol tetraether
GDD	Glycerol dialkyl diether
GDD_M	Chain methylated structural variant of an isoprenoid GDD
GDGT	Glycerol dialkyl glycerol tetraether
GDGT*	An isobar of a known GDGT lipid which co-elutes during chromatographic separation
GDGT’	An isobar of a known GDGT lipid which elutes earlier during chromatographic separation
GDGT’’	An isobar of a known GDGT lipid which elutes later during chromatographic separation
$GDGT_M$	A structure which represents a methylated or polymethylated variant of a known isoprenoid GDGT, with the additional carbon atom(s) added to one or both of the etherified alkyl chains

$GDGT_M$*	An isobar of a major $GDGT_M$ lipid which co-elutes during chromatographic separation
$GDGT_{maj}$	Major GDGT isomers (as defined in the main text)
$GDGT_{min}$	Minor GDGT isomers (as defined in the main text)
$GDGT[N(1)_{P(1)}, N(2)_{P(2)}]$	GDGT lipid core where N(1) and N(2) are the number of carbon atoms and P(1) and P(2) the number of cyclopentyl rings in the first and second ether-bound chains, respectively (designation of first and second chains is arbitrary)
GDHT	Glycerol dialkyl homoglycerol tetraether
$GDHT[N(1)_{P(1)}, N(2)_{P(2)}]$	GDHT lipid core where N(1) and N(2) are the number of carbon atoms and P(1) and P(2) the number of cyclopentyl rings in the first and second ether-bound chains, respectively (designation of first and second chains is arbitrary)
GDH_2T	Glycerol dialkyl homoglycerol tetraether in which the homoglycerol group contains two additional carbon atoms by comparison to glycerol
GDNT	Glycerol dialkyl nonitol tetraether
GDT	Glycerol dialkanol triol
Glycerol-1-P	Glycerol 1-phosphate
GMD	Glycerol monoalkyl diether
GMT	Glycerol monoalkanol triol
GMGT	Glycerol monoalkyl glycerol tetraether
GMM	Glycerol monoalkyl monoether
GTGT	Glycerol trialkyl glycerol tetraether
$GTGT_M$	Chain-methylated structural variant of an isoprenoid GTGT
GTHT	Glycerol trialkyl homoglycerol tetraether

h	Hours
HCT	Bruker HCTultra ETD II ion trap instrument
Heterocystous	Contain differentiated heterocysts, thick-walled cells which serve as a site for N_2 fixation, within the cyanobacterial filaments
Heterotroph	Uses organic carbon derived from other organisms for nutrition
HI	Hydroiodic acid
Homocaldarchaeol	GDGT lipid containing one biphytanyl and one 3,7,11,13,15,18,22,26,30-nonamethyldotriacontanyl chain
HPLC	High performance liquid chromatography
H-shaped caldarchaeol	GMGT lipid containing two covalently-conjoined biphytanyl chains
Hydroxyarchaeol	GDD lipid containing one phytanyl and one 3-hydroxy-3,7,11,15-tetramethylhexadecanyl chain
Hyperthermophile	Organism with optimal growth temperature typically above 80^oC
ID	Internal diameter
IPL	Intact polar lipid; a lipid which contains at least one phosphatidyl or glycosyl head group
Iso	Hydrocarbons methyl-branched at the ω2 chain position
Isocaldarchaeol	An isomer of caldarchaeol in which the two capping glycerol groups are aligned *syn* to one-another as opposed to *anti*
Isohomocaldarchaeol	An isomer of homocaldarchaeol in which the two capping glycerol groups are aligned *syn* to one-another as opposed to *anti*

Isotopomers	Molecules with the same molecular formula but which contain different isotopes of some of the atoms
IPA	isopropanol
Korarchaeota	A kingdom of the Archaea
Ladderane	Hydrocarbon containing two or more linearly concatenated cyclobutyl rings
Lateral gene transfer	Transfer of genetic material from one organism to another, non-ancestral organism
LC	(High performance) liquid chromatography
LCQ	Finnigan MAT LCQ ion trap instrument
LC-MS	Liquid chromatography-mass spectrometry
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
$LC\text{-}MS^n$	Liquid chromatography-multistage tandem mass spectrometry
$LiAlH_4$	lithium aluminium hydride
LMCO	Low mass cut-off
LOD	Limit of detection
m	Mass of the ion
$M^{+\bullet}$	Molecular ion
$[M+H]^+$	Protonated molecule
$[M+H+1]^+$	Protonated molecule of the +1 higher isotopomer
$[M+H+2]^+$	Protonated molecule of the +2 higher isotopomer
$[M-H]^-$	Deprotonated molecule
$[M+K]^+$	Potassiated molecule
$[M+Na]^+$	Sodiated molecule
$[M+NH_4]^+$	Ammoniated molecule
$[M+OCHO]^-$	Formylated molecule
MAAT	Mean annual air temperature

MALDI	Matrix assisted laser desorption ionisation
MBT index	Methylation index of branched tetraethers
Metagenomic	Relating to genetic material recovered directly from the environment
Mesophile	An organism which inhabits environments where moderate temperatures and pH values prevail
Min	Minute
M	Molar
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MS^2	Tandem mass spectrometry
MS^n	Multistage tandem mass spectrometry
MTH(ΔH)	*Methanothermobacter thermautotrophicus* Strain ΔH
m/z	Mass-to-charge
n	Number of analytical replicates
Nanoarchaeota	A proposed kingdom of the Archaea
NMR	Nuclear magnetic resonance
$[N_P]$	An alkyl group containing N carbon atoms and P cyclopentyl rings encapsulated within the chain
$[N_{P+H}]$	An alkyl group containing N carbon atoms and P cyclopentyl plus H cyclohexyl rings encapsulated within the chain
$[N_P]^L$	Neutral loss of an $[N_P]$ chain as a diene, with L the formal number of molecules of water simultaneously lost
NZ	New Zealand
oTOF	Orthogonal time of flight
PAN MS/MS	Panoramic MS/MS
ppm	Parts per million

Pelagic	Of or relating to the open ocean
Phytanyl	3,7,11,15-tetramethylhexadecanyl
Plasmalogens	A glyceride containing one alkyl chain bound as an acyl group (at *sn*-1) and one unsaturated alkyl chain bound as a vinyl ether group (at *sn*-2)
PMI	Pentamethylicosane
psi	Pounds per square inch
q_{eject}	q_z value at which ions are axially destabilised beyond the dimensions of the ion trap
Qh-FT-ICR	Hybrid quadrupole-hexapole-FT-ICR instrument
q_z	Dimensionless stability parameter; axial solution to the Matthieu equation
$q_{z,precursor}$	q_z value of the precursor ion during CID
r_0	Internal radius of the ring electrode
RF	Radio frequency
RIC	Reconstructed ion chromatogram
RNA	ribonucleic acid
rpm	Revolutions per minute
σ	Standard deviation
sCID	In-source CID
Sesterterpanyl	3,7,11,15,19-pentamethylicosane
Single ion monitoring	An operational mode of a mass spectrometer in which its mass analyser targets an individual, specific ion (or handful of specific ions)
sn-1	One of the primary carbinols in glycerol or calditol moieties
sn-2	The secondary carbinol in glycerol or calditol moieties
sn-3	One of the primary carbinols in glycerol or calditol moieties

sn-1,2	The *sn*-1 and *sn*-2 carbinols of glycerol/calditol have been modified
sn-2,3	The *sn*-2 and *sn*-3 carbinols of glycerol/calditol have been modified
SST	Sea surface temperature
τ, τ_1, τ_2	Time intervals across the full peak width for GDGT **16** in the base peak chromatogram obtained following LC-MS/MS analysis of the extract from VDG IV-1.9 sediment
TEX_{86} index	Tetraether index of tetraethers consisting of 86 carbon atoms
Thaumarchaeota	A proposed kingdom of the Archaea
Thermoacidophile	Organisms which exhibit optimal growth typically between temperatures of 70-80^{o}C and pH values of 2-4
Thermophile	Organism with optimal growth temperature typically above 45^{o}C
TIC	Total ion current
TLC	Thin layer chromatography
TMS	trimethylsilyl
t_R	Retention time
Turbidite	A deep ocean sedimentray deposit formed by turbidity currents
Type strain	The strain designated as the nomenclatural type to which all other novel strains are compared for assignment of species
U	Amplitude of the DC potential (applied to end-caps)
UK	United Kingdom
U^{k}_{37}	Unsaturation ratio of methyl-ketones containing 37 carbon atoms

USA	United States of America
UV/Vis	Ultra violet/visible
V	Amplitude of the RF potential (applied to ring electrode); not to be confused with its use as a unit of electrical potential, the volt, also given as V
VDG	Vena Del Gesso
V_{max}	Maximum value of V that can be attained by the ion trap instrument
ω	Oscillatory frequency of the RF potential
ω_z	The axial component of the secular frequency
w:v	Weight to volume ratio
Xenobiotic	A chemical not naturally synthesised by an organism to which it is exposed
z	Charge number of the ion
z_0	Half of the axial distance between the end-cap electrodes

Acknowledgements

I would like to take this opportunity to thank the University of York for providing the funding necessary to allow for this research. The EAOG are also thanked for provision of a travel grant which facilitated a two month research stint in New Zealand, with the fruit of those labours forming part of this thesis. Hugh Morgan deserves a special mention for his role in this trip, serving as perhaps the greatest host in the Antipodes.

I am extremely grateful to Brendan Keely for his advice and direction throughout the supervision of the project and, where the need arose, for being my partner in quiz and personal chef. I would also like to thank the members of the Keely group both past and present for tolerating me so politely and making my stay in York so enjoyable. In particular; to Phil, Debs, Jayne, Matt, Haslina, Suleman, Neung, Angela, Andy, Denize, Kim, Yao and Ryan – thank you.

Special thanks must go to Matt Pickering, Jayne Kirk, Trevor Dransfield and Ed Bergström for training me in the noble arts of liquid chromatography and mass spectrometry. Matt is particularly thanked for taking me under his wing, being an encouragable drinking buddy and for providing the soundtrack to this thesis; a soundtrack that drove me to the writing room and pulled me into the lab in equal measures. Angela Ballantyne and Muhammad Suleman were also "godsends" during production of this thesis. Ed Bergström requires additional thanks for discussion regarding mass spectrometric instrumentation and for his help in operating the Qh-FT-ICR instrument. Paul de Blacquiere provided a minor, but essential role by extracting the Oxford Clay sediment. A number of people from a variety of institutions provided the archaeal cellular biomass used in these studies. Without the help of James Chong and Daniella Barilla (York), Charlotte Nunn and Mike Danson (Bath) and Hugh Morgan and Christine Tan (Waikato, NZ), much of the work described would not have been possible.

Finally, I would like to thank my Mum and Dad (Jan and Rob), brother (Adam) and grandparents (Betty, Cliff and Vera) for their love and patience during what has been a challenging chapter of my life and, of course, for letting me dig up their gardens.

Chapter 1

Introduction

1.1. Archaea

In the classification system for life proposed by Woese *et al.* (1977; 1987; 1990), the Archaea represent one of the three domains, evolutionarily distinct from both the eubacteria and eukaryota (Fig. 1.1). The archaeal domain is extremely antiquated, having diverged from the ancestral line of the eubacteria in the primaeval biosphere (Fox *et al.*, 1980; Woese, 1987). The slow rates of evolution identified for some archaeal species indicate that contemporary archaea within these lineages may represent some of the most primitive forms of extant life (Woese, 1987). More conservative ages, however, have also been proposed recently (Blank, 2009). Within the archaeal domain, four phylogenetically distinct kingdoms have been identified: *Euryarchaeota*, *Crenarchaeota*, *Korarchaeota* and *Nanoarchaeota*[†]. The second pair are rare, with the *Korarchaeota* being defined largely on the basis of distinct archaeal 16S rRNA sequences observed in environmental settings (Barns *et al.*, 1996; Elkins *et al.*, 2008) and the *Nanoarchaeota* represented by a sole known symbiont (Huber *et al.*, 2002), the phylogenetic placement of which has been questioned (Brochier *et al.*, 2005). *Euryarchaeota* and *Crenarchaeota* (Woese *et al.*, 1990), on the other hand, are of widespread occurrence in the environment. Prior to 1992, archaea were believed to be restricted to distinctive and extreme ecological niches, such as the rumen of ruminents and other anoxic environments (Zeikus, 1977), hypersaline environments (Grant and Ross, 1986), volcanic hot springs (e.g. Brock *et al.*, 1972) and hydrothermal vents (e.g. Jones *et al.*, 1983). As such, they were typically categorised into three phenotypically distinct groups, comprising extreme halophiles, methanogens and thermophiles (Woese, 1987). More recently, however, archaea have been identified as major contributors to the pelagic biomass of the world's oceans (up to 20% of all picoplankton), indicating that archaea can survive and actively compete in mesophilic conditions (Fuhrman *et al.*, 1992; DeLong, 1992; Karner *et al.*, 2001; DeLong, 2006). Furthermore, their high abundances in natural marine environments indicate that they play important roles in the major cycles of elements, particularly the carbon and nitrogen cycles (Francis *et al.*, 2007).

[†] A fifth distinct kingdom, the *Thaumarchaeota*, which segregates the mesophilic members of the *Crenarchaeota*, has also recently been proposed (Brochier-Armanet *et al.*, 2008; Spang *et al.*, 2010).

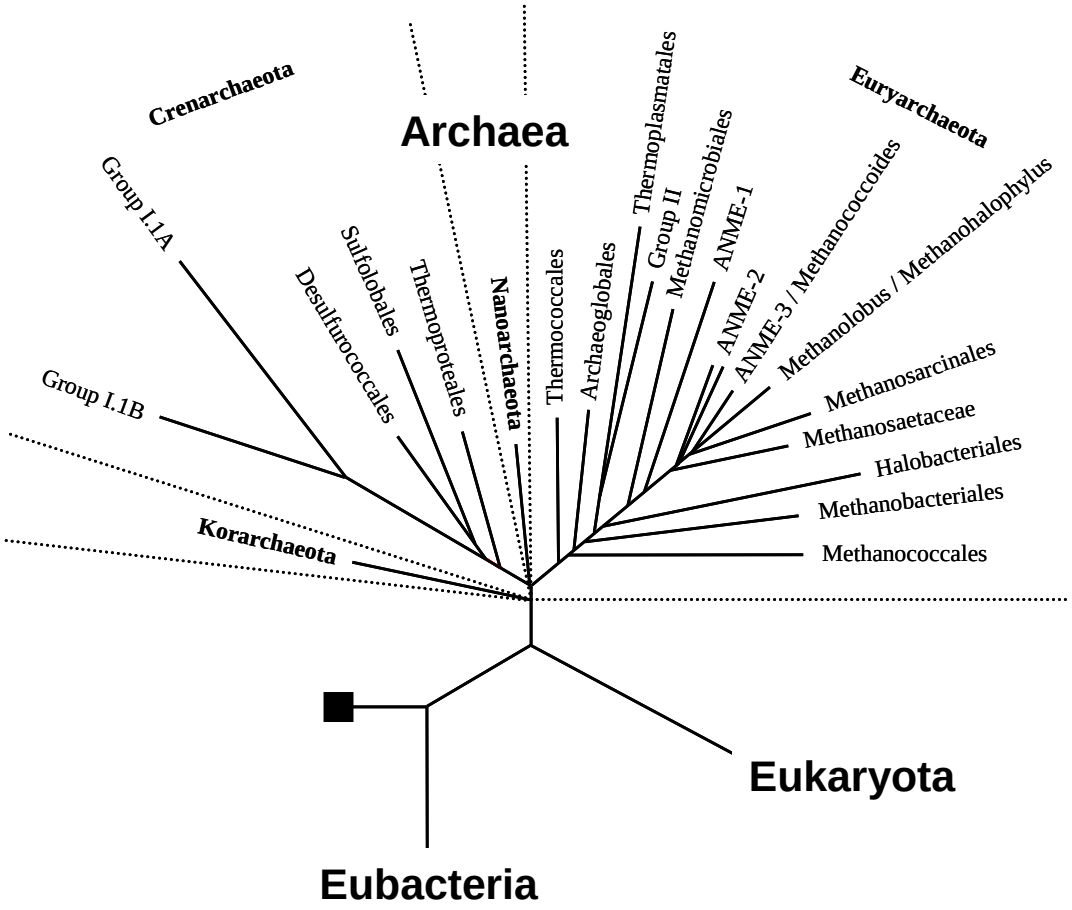


Fig. 1.1. The universal phylogenetic tree of life based on 16S rRNA sequences (Woese *et al.*, 1990), with the phylogeny of the archaeal branch shown in detail (adapted from Schleper *et al.*, 2005).

Several distinct clades of archaea are now known to contribute to oceanic ecosystems, including Marine Group I *Crenarchaeota* and Group II *Euryarchaeota*, which represent the major pelagic species (DeLong, 2006; 2007 and references therein). Group I *Crenarchaeota* can live at variable depths within the water column (e.g. Damsté *et al.*, 2002a) and appear to be chemolithoautotrophic, aerobically oxidising ammonia to nitrite as a source of energy and utilising bicarbonate or carbon dioxide derived therefrom as a carbon source (Pearson *et al.*, 2001; Wuchter *et al.*, 2003; Könneke *et al.*, 2005). Less is known about the Group II *Euryarchaeot*a, largely due to the lack of a representative isolate from the clade (Turich *et al.*, 2007). Molecular biomarker approaches (both genetic and lipid-based) have identified a significant community of viable archaea living within the marine benthos and in

deep sub-surface sediments (see Teske and Sørensen, 2008 for an extensive review of the topic). Other studies have indicated that methanotrophic (ANME) archaea, which live in syntrophic consortia with sulfate-reducing eubacteria, are found to be abundant in sediments with a substantial methane through-flux, such as cold seep carbonates and mud volcanoes (Hoehler *et al.*, 1994; Hinrichs *et al.*, 1999; 2000; Boetius *et al.*, 2000; Orphan *et al.*, 2001; 2002; Pancost *et al.*, 2001b). These organisms are believed to be responsible for large scale anaerobic oxidation of methane, also known as AOM (Orphan *et al.*, 2002 and references therein), a process responsible for significant turnover of the methane released into oceanic environments (Reeburgh, 2007).

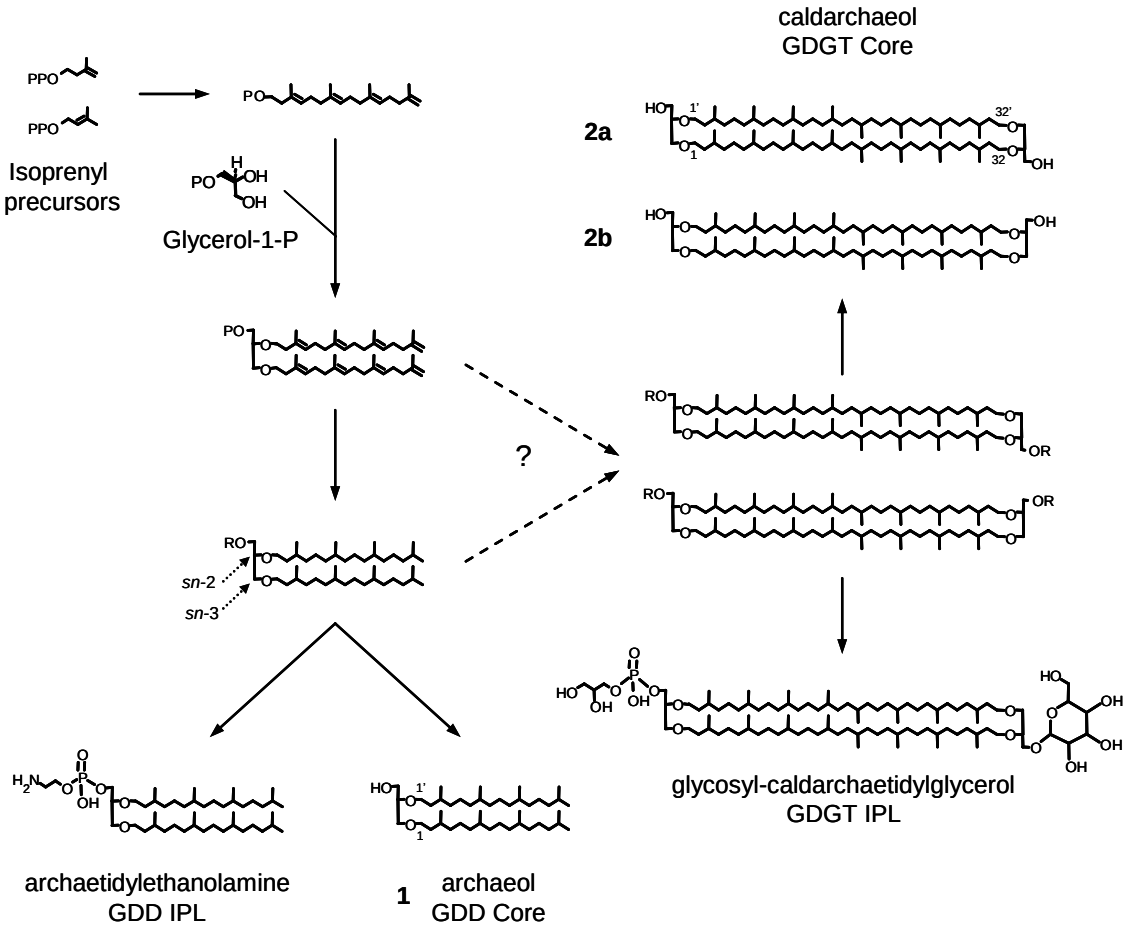
In addition, archaea have been identified in a number of non-marine environments in which moderate conditions prevail, such as freshwater lakes (Powers *et al.*, 2010 and references therein), peat bogs (e.g. Schouten *et al.*, 2000; Weijers *et al.*, 2004; 2006a; Liu *et al.*, 2010) and soils (e.g. Leininger *et al.*, 2006; Weijers *et al.*, 2006b).

1.2. Lipids of the Archaea

1.2.1. Typical features and biosynthesis

All members of the Archaea synthesise complex polyisoprenoid ether lipids, unique to this domain, which are used within the cellular membranes of the organisms to confer structure. The complex ether lipids, often referred to as intact polar lipids (IPLs), typically comprise an ether lipid core derived from glycerol which is capped by modified phosphate and/or glycosyl head groups (De Rosa and Gambacorta, 1988a; Sturt *et al.*, 2004). Interestingly, the stereochemistry of the glycerol moieties established during biosynthesis of the lipids is also unique, with the hydrophobic chains in the structures exclusively bound to the *sn*-2 and *sn*-3 carbinols of the polar groups (De Rosa and Gambacorta, 1988a). By contrast, eukaryotic and eubacterial lipids typically comprise mainly saturated or unsaturated straight chain fatty acids, ester-bound to glycerol at the *sn*-1 and *sn*-2 positions (see Section 1.3 for contrary examples of non-archaeal ether lipids). The simplest ether lipid core is archaeol (**1**; see Scheme 1.1), a glycerol dialkyl diether (GDD) lipid which comprises two C_{20}

phytanyl (i.e. 3,7,11,15-tetramethylhexadecanyl) chains ether-linked to a glycerol group exhibiting the unique *sn*-2,3 stereochemistry (Kates, 1972). The isoprenoid chains in this core are biosynthesised *via* the classical mevalonate pathway and, as phosphorylated geranylgeraniol units, are condensed onto a glycerol-1-phosphate template (Koga, and Morii 2007; Scheme 1.1). Subsequent additions of head groups followed by hydrogenation of the geranylgeranyl chains yield complex lipids containing **1** as the core.



Scheme 1.1. A condensed synthetic route to formation of archaeal GDD and GDGT lipids, including both IPL and core structures, of which examples are given. P represents a phosphate group. R represents a generic polar headgroup.

In some archaea, two complex lipids based on **1** can undergo an intermolecular dimerisation, whereby the terminal carbons of the phytanyl chains are coupled *via* formation of two new C-C bonds, establishing a pair of C_{40} biphytanyl (i.e.

3,7,11,15,18,22,26,30-octamethyldotriacontanyl) hydrocarbons (Nemoto *et al.*, 2003). The chemical mechanism through which this coupling occurs is currently unknown, although the process is inhibited in *Thermoplasma acidophilum* by addition of a squalene epoxidase inhibitor to the inoculated growth medium (Kon *et al.*, 2002; Nemoto *et al.*, 2003). On a more fundamental level, there is debate as to whether the geranylgeranyl chains in the cores of the lipids involved in this coupling are fully saturated or contain remnant double bonds (Koga and Morii, 2007; Eguchi *et al.*, 2000; 2003). One feature of the coupling that is more extensively understood is the apparent lack of specificity for the orientation of the GDD precursors prior to condensation; it appears that coupling of the precursors is equally favourable whether the two glycerol groups are initially aligned parallel or antiparallel to one-another (Gräther and Arigoni, 1995). As such, subsequent dephosphorylation of the dimerisation products yields *anti* and *syn* isomers of the tetraether lipid caldarchaeol[†] (**2a** and **2b**, respectively), the simplest examples of glycerol dialkyl glycerol tetraether (GDGT) lipid cores (Scheme 1.1).

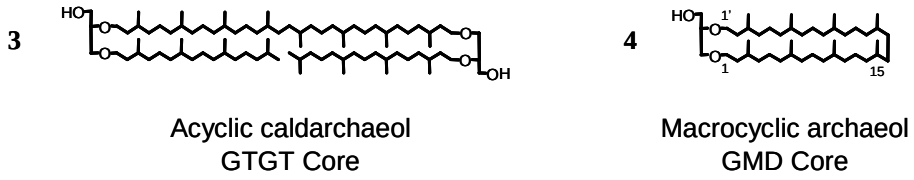


Fig. 1.2. Structures of the archaeal lipids GTGT **3** and GMD **4**.

Acyclic caldarchaeol (**3**; Fig. 1.2), a glycerol trialkyl glycerol tetraether (GTGT) lipid, has also been observed in a number of thermophilic and mesophilic archaea (De Rosa and Gambacorta, 1988a; Uda *et al.*, 2000; Hopmans *et al.*, 2000; Schouten *et al.*, 2008b; 2008c; de la Torre *et al.*, 2008; Pitcher *et al.*, 2010) and aquatic sediments (Hopmans *et al.*, 2000; Schouten *et al.*, 2000; Pancost *et al.*, 2001b; 2006). The lipid is presumably formed when the intermolecular cross coupling does not proceed to completion following formation of the first C-C bond (de la Torre *et al.*, 2008). Interestingly, macrocyclic archaeol (**4**), denoted here as a glycerol

[†] This term was first used by Nishihara *et al.* (1987b) to describe GDGT **2a** in *Methanobacterium thermoautotrophicum* and, in a more general sense, to provide a systematic descriptor for other dialkyl and trialkyl tetraether lipids that had been found previously in *Caldariella acidophila* (e.g. De Rosa *et al.*, 1977b). Here, it is used in a restricted fashion, solely as a formal name for lipid **2a**.

monoalkyl diether (GMD) lipid, has also been identified in *Methanocaldococcus jannaschii* (formerly *Methanococcus jannaschii*), an isolate from a deep-sea hydrothermal vent (Comita *et al.*, 1983; 1984; Sprott *et al.*, 1991), tentatively in two other members of the *Methanococcales* (Koga *et al.*, 1998) and in cold seep settings (Birgel and Peckmann, 2008a; De Boever *et al.*, 2009). As *M. jannaschii* also synthesises GDD **1** (Comita *et al.*, 1983; 1984; Sprott *et al.*, 1991), it is conceivable that GMD **4** may be formed *via* an intramolecular coupling between two terminal methyl groups on the opposing phytanyl chains of a single GDD lipid. A direct biosynthetic link between **1** and **4**, however, has yet to be established conclusively.

1.2.2. Structural variation

Other GDD lipids to have been identified (Fig. 1.3) include structures containing one phytanyl ether-bound chain and one sesterterpanyl (i.e. C_{25}; 3,7,11,15,19-pentamethylicosane) ether-bound chain (**5**; De Rosa *et al.*, 1982; 1983a; 1986; 1988a; Upasani *et al.*, 1994; Qiu *et al.*, 1999; Romano *et al.*, 2007; Wang *et al.*, 2007; Lipp *et al.*, 2008; 2009) or, otherwise, containing two ether-bound sesterterpanyls (**6**; De Rosa *et al.*, 1983a; 1988a; Sako *et al.*, 1996; Morii *et al.*, 1999; Tachibana *et al.*, 2000). The *sn*-2 and *sn*-3 hydroxyarchaeols **7** and **8**, respectively, which contain both a phytanyl and a 3-hydroxy-3,7,11,15-tetramethylhexadecanyl chain, have been identified in a number of archaea (Ferrante *et al.*, 1988; 1989; Sprott *et al.*, 1990; 1993; 1997; 1999; Nishihara and Koga, 1991; Swain *et al.*, 1997; Koga *et al.*, 1998) and in environmental settings which harbour archaeal organisms (e.g. Hinrichs *et al.*, 1999; 2000; Pancost *et al.*, 2000; 2001a; 2001b; Blumenberg *et al.*, 2004; Schubert *et al.*, 2006; Thiel *et al.*, 2007a; 2007b; Niemann and Elvert, 2008; Stadnitskaia *et al.*, 2005; 2008). Dihydroxyarchaeols containing two hydroxylated phytanyl chains (e.g. **9**; Orphan *et al.*, 2002; Kelley *et al.*, 2005; Thiel *et al.*, 2007a; Rossel *et al.*, 2008; Bradley *et al.*, 2009b) and an extended hydroxyarchaeol (**10**; Stadnitskaia *et al.*, 2008; Chevalier *et al.*, 2010), in which the chain-bound hydroxyl group is located at the C-7 position of a sesterterpanyl chain (i.e. yielding a 7-hydroxy-3,7,11,15,19-pentamethylhexadecanyl chain), have also been identified in marine microbial mats and in carbonates from cold seeps. In addition, polyunsaturated archaeols (**11**), hydroxyarchaeols (**12**) and

sesterterpanyl GDDs (precise structures unknown), most likely precursors of the saturated diethers, have been identified in a variety of archaea (e.g. Nichols *et al.*, 1992; 2004; Hafenbradl *et al.*, 1993; Upasani *et al.*, 1994; Qiu *et al.*, 1998; 2000; Gonthier *et al.*, 2001; Nishihara *et al.*, 2002; Gibson *et al.*, 2005; de Souza *et al.*, 2009). A monounsaturated archaeol structure identified in Mediterranean mud dome sediments (Pancost *et al.*, 2001b) and belonging to this lipid set has been suggested as a possible transformation product of hydroxyarchaeols (Ekiel and Sprott, 1992; Nichols *et al.*, 1993). A *bis*-phytanyl diether tetritol (**13**), containing an unusual 1,2,3,4-tetrahydroxybutane group in place of glycerol, has been identified in a methanogen (De Rosa *et al.*, 1986; Trincone *et al.*, 1988), although a biosynthetic origin for this component has been disputed (see Koga *et al.*, 1993b).

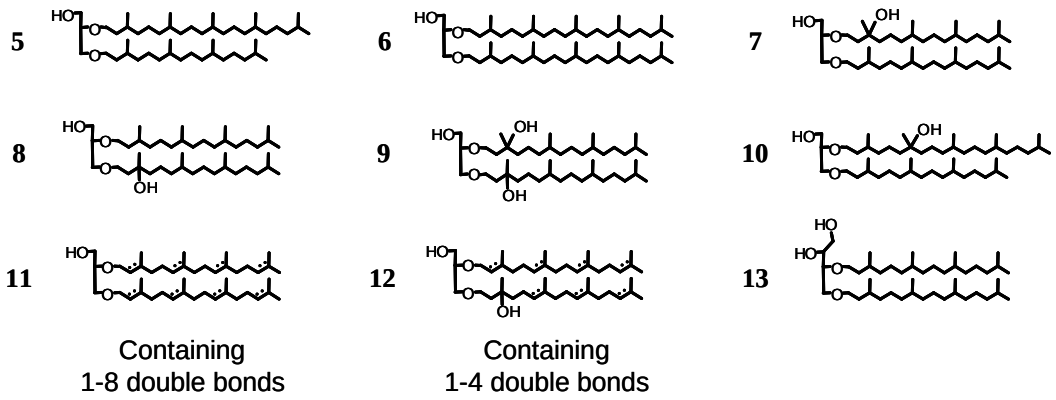


Fig. 1.3. The variety of archaeal GDD lipid cores identified to date.

Variation in GDGT lipid core structure is often provided by the incorporation of cyclopentyl (Cp) moieties within the aliphatic chains of the tetraether lipid cores (e.g. **14**-**21**; Fig. 1.4). Mesophilic and thermophilic members of both the *Euryarchaeota* and *Crenarchaeota* are known to synthesise tetraether lipids containing between 0 and 8 Cp rings in total, with up to 4 rings per biphytanyl chain (De Rosa and Gambacorta, 1988a; DeLong *et al.*, 1998; Damsté *et al.*, 2002b). Interestingly, despite the large number of potential GDGT structures conceivable from pairwise combinations of the known ring-containing biphytanes, only a handful of lipids have been identified to date, both in environmental samples and in archaeal cultures. This restriction appears to be due, in part, to controls on the positions in which the rings are incorporated within the alkyl chains (De Rosa *et al.*, 1980c).

Further limitation is attributable to the small number of combinations of the alkyl chains that appear to be adopted, with all such tetraether molecules definitively identified to date exhibiting biphytanyl pairs that differ by, at most, two Cp rings (De Rosa and Gambacorta, 1988a; Thurl and Schäfer, 1988; Schouten *et al.*, 2000; Pancost *et al.*, 2001b). GTGT lipids containing up to four Cp rings have been identified in a thermophilic *Crenarchaeote* (de la Torre *et al.*, 2008), with a GTGT structure containing one Cp ring located on a biphytanyl chain similarly identified in mud domes associated with methane seeps (Hopmans *et al.*, 2000; Pancost *et al.*, 2001b). GMD lipids containing one or two Cp rings (**22** and **23**, respectively; Fig. 1.4) have also been identified in carbonates from cold seep environments (Stadnitskaia *et al.*, 2003; 2005; van Dongen *et al.*, 2007; Birgel and Peckmann, 2008a; De Boever *et al.*, 2009). The apparent absence of ring-containing relatives of **1**, and occurrence of Cp ring-containing derivatives of **2**-**4** suggests that ring formation occurs either simultaneous to, or following generation of C_{40} isoprenoid chains during ether lipid biosynthesis. A mechanism involving α methyl-δ methylene cyclisation within a coupled biphytanyl chain has been proposed (Weijers *et al.*, 2006a), although this has yet to be confirmed either *in vitro* or *in vivo*. Cobalamins have, however, been proposed as speculative, but plausible, catalytic agents which may mediate the cyclisation processes (Galliker *et al.*, 1998). A novel pair of Cp ring-containing GDGT lipids, crenarchaeol and its regioisomer (**24** and **24''**, respectively; Fig. 1.4), which also contain a cyclohexyl (Ch) ring in addition to four Cp rings, have been identified in cultures of *Crenarchaeota* (Damsté *et al.*, 2002b; Schouten *et al.*, 2008c; de la Torre *et al.*, 2008; Pitcher *et al.*, 2010) and within the environment (e.g. Schouten *et al.*, 2000; 2002; Pearson *et al.*, 2004; Powers *et al.*, 2004; Weijers *et al.*, 2004; 2006a; 2006b; Pitcher *et al.*, 2009b). An additional GDGT core, **25** (Fig. 1.4), containing one Ch and five Cp rings has also been identified in a moderately themophilic Group I.1b *Crenarchaeote* (Pitcher *et al.*, 2010). Homocaldarchaeol (**26a**; Fig. 1.4), a GDGT core which contains an additional branching methyl group at the C-13 chain position of one of the two ether-bound biphytanyl chains by comparison with **2**, has been identified along with two *syn* isomers (e.g. **26b**) in a thermophilic methanogen (Galliker *et al.*, 1998). A GMD lipid core containing a similar α,ω-dietherified C_{41} isoprenoid chain was tentatively identified in a hydrothermal sulfide chimney (Blumenberg *et al.*, 2007). The location

of the non-isoprenoid methyl group within the chain was different to that of the C_{41} chain in homocaldarchaeol (Blumenberg *et al.*, 2007).

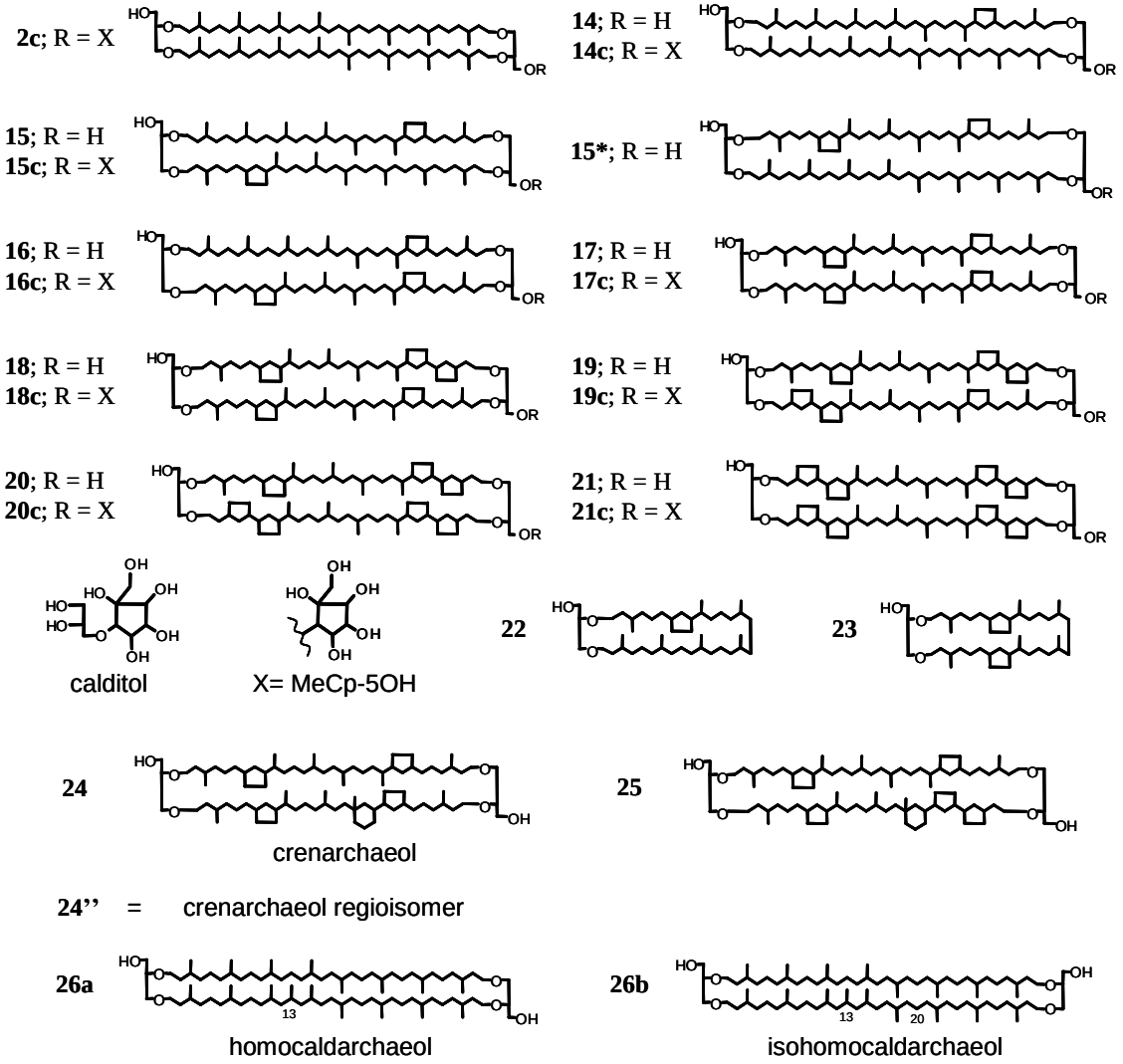


Fig. 1.4. The variety of archaeal GDGT, GDCT and GMD lipid cores identified to date. A previously identified isomer of **26a** and **26b**, in which the additional methyl group with respect to **2a** and **2b** is located at the C-20 as opposed to C-13 chain position (Galliker *et al.*, 1998), is not shown, nor is a GMD lipid core containing an α,ω-bound C_{41} chain (Blumenberg *et al.*, 2007).

A series of isoprenoid lipid cores (**2c** and **14c**-**21c**; Fig. 1.4), containing 0-8 Cp rings and of greater polarity than the similarly cyclised GDGT lipid cores, have also been identified in members of the archaeal order *Sulfolobales* (De Rosa *et al.*, 1977a; 1980c; 1983b; 1988a), in several aquatic sediments (Sturt *et al.*, 2004; Biddle *et al.*,

2006; Lipp *et al.*, 2008; 2009) and in a peat bog (Liu *et al.*, 2010). The structures of the lipids closely resemble those of GDGT structures, sharing the same isoprenoid hydrocarbons, but are furnished with a polyol as opposed to a glycerol moieity at one end of the structure. The other end of each structure retains the glycerol group found in the analogous GDGT lipids. The polyol was originally classified as nonitol and was described as an open chain (4-hydroxymethyl)-1,2,3,4,5,6,7,8-octahydroxyoctane group (De Rosa *et al.*, 1977a; 1980a). As a consequence, lipid cores containing this polyol were denoted as glycerol dialkyl nonitol tetraether (GDNT) lipids (De Rosa and Gambacorta, 1988a). Subsequent structural characterisation has, however, confirmed that the polyol is calditol (see Fig. 1.4), which contains seven hydroxyl groups and a cyclic five-membered saturated ring of carbon atoms (Koga and Morii, 2005 and references therein). A reclassification of lipids which contain the polyol under the nomenclature glycerol dialkyl calditol tetraether (GDCT) should, therefore, prove to be more appropriate.

Recently, another new family comprising glycerol monoalkyl tetraether (GMGT) lipids, which contain C_{80} isoprenoid hydrocarbons and are often referred to as H-shaped lipids, have been identified in thermophilic representatives of the *Euryarchaeota* (**27**-**31**; Fig. 1.5). A structural relative of **2**, H-shaped caldarchaeol (**27**), in which the two biphytanyl chains are conjoined by a C-C covalent link, was identified in the thermophilic methanogen *Methanothermus fervidus,* although the precise position of the link was only assigned tentatively (Morii *et al.*, 1998). This lipid was also identified in the thermophilic methanogen, *Methanothermobacter thermautotrophicus* (Morii *et al.*, 1998) and subsequently in four strains of the *Euryarchaeotal* order *Thermococcales* grown at 85^{o}C (Sugai *et al.*, 2004). Lipid **27** has also been identified in sediments in a number of non-thermophilic environments, suggesting that GMGT lipids are not exclusive to thermophilic archaea (Schouten *et al.*, 2008a). Additional structural variation in monoalkyl tetraether lipids has been observed in *Aciduliprofundum boonei*, a thermophilic *Euryarchaeote* that inhabits deep sea hydrothermal vents (Reysenbach *et al.*, 2006). This archaeon was shown to synthesise GDGT lipids, lipid **27** and a series of structures which were assigned as GMGT lipids containing 1-4 Cp rings (**28**-**31**), although the structures of the GMGT lipids were not confirmed fully (Schouten *et al.*, 2008b). On the basis of these

assignments, it was noted that the ratio of GDGT to GMGT lipids containing the same number of rings in *A. boonei* was approximately constant, suggesting that the GMGT lipids are synthesised directly from their GDGT counterparts. Consequently, it is most likely that the Cp rings in the C_{80} isoprenoid hydrocarbons in the GMGT lipids occur at analogous positions to those in individual C_{40} chains in GDGT lipids.

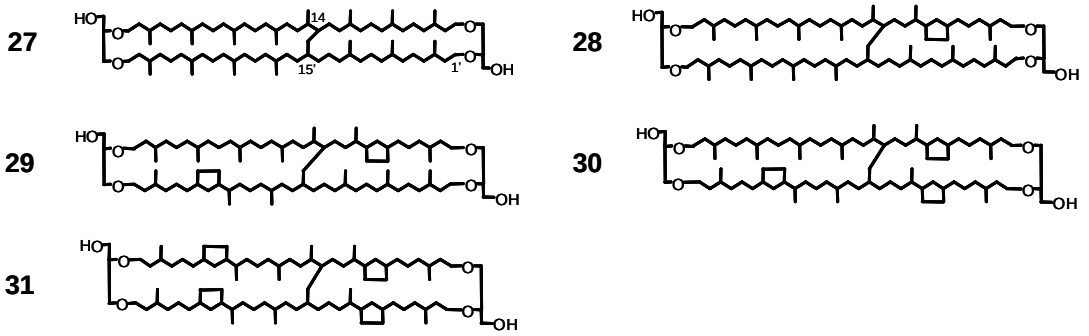


Fig. 1.5. The variety of archaeal GMGT lipid cores identified to date. The position of the covalent link between the isoprenoid chains in GMGT lipids is not known conclusively. Consequently, its placement is tentative based on previous reports of the structures (Morii *et al.*, 1998; Schouten *et al.*, 2008a; 2008b).

Glycerol monoalkyl monoether (GMM) lipid cores, containing either an *sn*-2 or an *sn*-3 glycerol-bound phytanyl chain (**32** and **33**, respectively; Fig. 1.6), have also been identified previously in archaea (Joo *et al.*, 1968; De Rosa *et al.*, 1986; 1988a; Sprott *et al.*, 1990) and in environments in which these organisms live (Jahnke *et al.*, 2001; Orphan *et al.*, 2001; Pancost *et al.*, 2001b; Oba *et al.*, 2006; Chevalier *et al.*, 2010). Notably, however, acidic degradation of hydroxyarchaeols can yield the GMM cores (Sprott *et al.*, 1990; Ekiel and Sprott, 1992), hence these structures may represent transformation products as opposed to authentic biosynthetic products.

Fig. 1.6. The structures of archaeal GMM lipid cores containing ether-bound phytanyl chains.

1.3. Eubacterial ether lipids

1.3.1. General features

Although acyl lipids are the most common components of eukaryotic and eubacterial membrane envelopes, some species of eubacteria are known also to produce membrane lipids containing ether linkages (e.g. Caillon *et al.*, 1983; Langworthy *et al.*, 1983; De Rosa *et al.*, 1988b; Huber *et al*. 1992; 1996; Wagner *et al.*, 2000; Jahnke *et al.*, 2001; Rütters *et al.*, 2001; Damsté *et al.*, 2002c; 2004; 2005; 2007; Paściak *et al.*, 2003; Ring *et al.*, 2006; Sturt *et al.*, 2004; Boumann *et al.*, 2006; Rattray *et al.*, 2008). Notably, in all cases where stereochemical assignments were made, an *sn*-1,2 glycerol stereochemistry was established for these components; a key structural feature which distinguishes them from archaeal-derived lipids. An additional discriminatory feature of the eubacterial ether lipids is the inclusion of non-isoprenoid hydrocarbons as the hydrophobic component of the molecule (*cf.* isoprenoid hydrocarbons found in archaeal lipids). The ether lipids typically contain straight chain or branched chain hydrocarbons and/or "ladderanes", the latter being alkyl chains which contain linearly concatenated cyclobutyl rings (i.e. **l**, **m**, **p** and **q**; Fig. 1.7 and Table 1.1). Many of the unusual non-isoprenoid alkyl chains represented in the ether lipids have also been identified in eubacteria as the related fatty acids (see Kaneda, 1991; Damsté *et al.*, 2002c).

1.3.2. Structural variation

A remarkable variety of non-isoprenoid ether lipids and mixed ether/ester lipids have subsequently been identified in natural environments, matching or complementing those found in the eubacterial cultures (Fig. 1.7 and Table 1.1). Pancost *et al.* (2001a) reported three series of non-isoprenoid GDD lipid cores containing ether-bound normal, *iso* or *anteiso* alkyl chains exhibiting 14-17 carbon atoms and 0-1 cyclopropyl or Ch rings in a cold seep carbonate. Subsequently, the known variety of GDD lipid cores has been extended following analysis of lipid extracts from hydrothermal vent chimneys, with a total of over 40 structures being assigned tentatively (Bradley *et al.*, 2009a). These structures include lipids containing 13-19

carbons within each ether-bound chain. Additional GDD structures containing C_{20} and C_{21} alkyl chains have also been identified in hot spring microbial communities (Jahnke *et al.*, 2001), while GDD cores containing unusual combinations of shorter alkyl chains have been identified in hot spring silica sinters (Pancost *et al.*, 2006) and in a hydrothermal sulfide (Blumenberg *et al.*, 2007). Furthermore, at least ten ladderane-containing GDD lipid cores were identified in enrichment cultures of *Planctomycetes* from wastewater treatment plants, with a similar variety of glycerol monoether/monoester structures containing ladderane chains also reported (Damsté *et al.*, 2002c; 2004; 2005; Boumann *et al.*, 2006; Rattray *et al.*, 2008). An extensive variety of non-isoprenoid GMM lipids containing identical or very similar alkyl chains have also been observed (Caillon *et al.*, 1983; Langworthy *et al.*, 1983; Huber *et al*., 1992; 1996; Hinrichs *et al.*, 2000; Orphan *et al.*, 2001; Damsté *et al.*, 2002c; 2005; 2007; Pape *et al.*, 2005; Pancost *et al.*, 2005; 2006; Boumann *et al.*, 2006; Rattray *et al.*, 2008; Bouloubassi *et al.*, 2009; Bradley *et al.*, 2009b; Jagersma *et al.*, 2009; Chevalier *et al.*, 2010; Kinnaman *et al.*, 2010; Opperman *et al.*, 2010). It has, however, been suggested that the GMM lipids may represent the lyso forms of glycerol monoether/monoester lipid cores (Langworthy *et al.*, 1983; Rütters *et al.*, 2001).

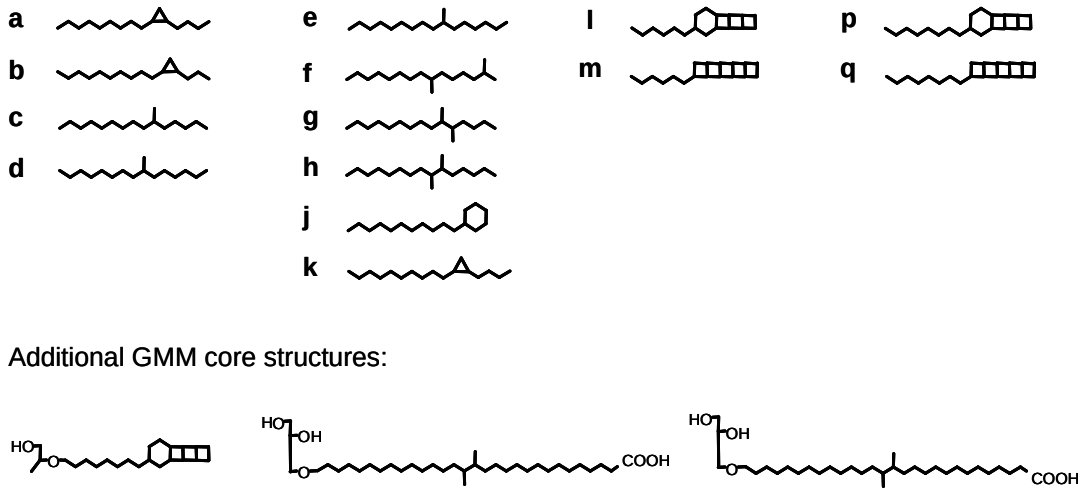


Fig. 1.7. Some of the unusual hydrocarbons found in the eubacterial GMM and GDD lipids shown in Table 1.1. Three additional monoether lipids, containing either an unusual polar terminus or membrane-spanning alkyl chains, are also shown.

In addition to ether lipids containing a single glycerol substrate, a series of non-isoprenoid GDGT lipids have also been identified, including **34**-**42** (Fig. 1.8), which are found predominantly in terrestrial or coastal environments (Damsté *et al.*, 2000; Schouten *et al.*, 2000; Weijers *et al.*, 2006a; 2006b; 2010). These lipids contain 13,16-dimethyloctacosanyl and/or 5,13,16-trimethyloctacosanyl chains (i.e. **34**-**36**) which, in some of the lipids (i.e. **37**-**42**), have been modified to include a Cp ring. An *sn*-1,2 stereochemistry was confirmed for the glycerol moieties in **35**, precluding an origin for this lipid from archaea (Weijers *et al.*, 2006a). It has been suggested that these GDGT lipids may be formed *via* coupling of two non-isoprenoid GDD lipids, mirroring the process for tetraether lipid formation in the Archaea (Damsté *et al.*, 2000; Weijers *et al.*, 2006a). Formation of **34** by this route could occur, for example, *via* the coupling of a pair of GDD molecules which each contain *i*-C_{15} chains bound to both *sn*-1 and *sn*-2 glycerol carbinols. Notably, the carbon termini of the *iso* chains are almost identical to those found in the phytanyl chains in GDD **1**, which are cross-coupled during formation of GDGT **2**. As discussed above, GDD lipids containing *i*-C_{15} chains are, themselves, indicative of a eubacterial origin, strongly suggesting that GDGTs **34**-**42** are also synthesised by eubacteria. Furthermore, an α,ω-13,16-dimethyloctacosanoic diacid, which contains an identical alkyl chain to the C_{30} chain found in GDGT **34**, has been identified in several species belonging to the eubacterial genus *Thermoanaerobacter* (Jung *et al.*, 1994; Lee *et al.*, 2002; Balk *et al.*, 2009) and is utilised in *bis*-glycerol tetraester lipids (Jung *et al.*, 1994; Lee *et al.*, 2002). GMD lipids containing C_{30}-C_{35} methyl-branched, non-isoprenoid alkyl chains have also been identified in hot spring (Pancost *et al.*, 2005; 2006) and cold seep (Blumenberg *et al.*, 2007) environments and in marine iron sulfide nodules (van Dongen *et al.*, 2007). Blumenberg *et al.* (2007) determined the C_{30} and C_{31} chains in two of these lipids (e.g. **43**; Fig. 1.8) to be identical to those found in GDGTs **34**-**36** and proposed that the C_{32}-C_{35} chains in the remaining lipids were more heavily methyl-branched variants of these structures.

Table 1.1. The variety of GMM and GDD lipid cores of confirmed or suspected eubacterial origin identified to date[a].

		R2[b]	H	C_{13}	C_{14}		C_{15}				C_{16}								
				n	n	i	?	n	i	ai	?	n	i	ai	**a**	**b**	D	**c**	**d**
R1[b]																			
H					+	+		+	+	+	+	+	+				+		
C_{13}	ai									+									
C_{14}	n				+			+					+		+	+				
	i				+										+	+				
	ai														+	+				
	D																+			
C_{15}	n			+[c]	+			+				+			+	+				
	i				+				+			+			+	+		+	+	
	ai				+	+			+	+				+	+	+				
C_{16}	?										+									
	n				+							+								
	ai														+	+				
	a											+			+	+				
	b											+			+	+				
	D																+			
C_{17}	?																			
	n																			
	ai																			
	e																			
	k				+			+				+			+	+				
C_{18}	?						+													
	n																			
C_{19}	?						+													
	n																			
C_{20}	n																			
	p		+		+	+		+	+			+	+					+	+	

[a] Compiled from Liefkins *et al*., 1979; Langworthy *et al*., 1983; Huber *et al*., 1992; 1996; Hinrichs *et al*., 2000; Jahnke *et al*., 2001; Orphan *et al*., 2001; Pancost *et al*., 2001a; 2006; Damsté *et al*., 2005; Rattray *et al*., 2008; Bradley *et al*., 2009a; Opperman *et al*., 2010 and references therein.

[b] R1 is the *sn*-2 ether-bound chain, R2 the *sn*-1 ether-bound chain; H = hydrogen; n = normal isomer; i = iso-branched isomer; ai = anteiso-branched isomer; ? = saturated chain (undisclosed structure); D = chain which contains one double bond equivalent (undisclosed structure). Letters in bold face represent unusual chain isomers (structures for which are shown in Fig. 1.7).

[c] (+) = Lipid identified previously. In some cases, the orientation of chains within the lipid core is not known and, consequently, R1 and R2 may be interchanged from the orientations shown.

C_{17}										C_{18}							C_{19}				C_{20}				C_{21}
?	n	i	ai	**e**	**f**	**g**	**h**	**j**	**k**	?	D	n	i	ai	**l**	**m**	D	n	i	ai	D	n	**p**	**q**	D
	+	+	+	+					+	+	+	+	+	+				+	+	+	+	+			+
				+		+	+																		
									+																
+												+					+							
	+										+														
		+									+														
				+							+														
+										+															
	+	+																							
			+																						
				+																					
								+																	
										+								+				+			+
											+														
																									+
																						+			
	+	+		+	+										+	+							+	+	

Table 1.1. (cont.)

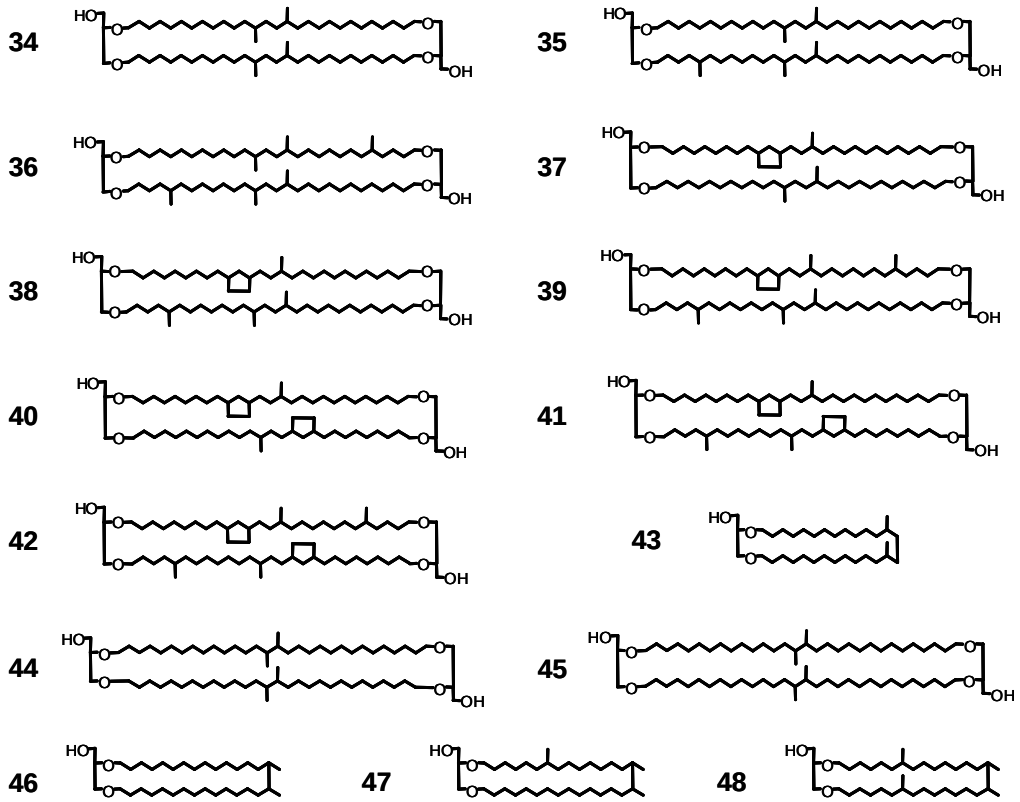


Fig. 1.8. The variety of GDGT lipid cores, confirmed or suspected to be eubacterial in origin, identified to date. Four of the GMD lipid cores which are believed to originate from eubacteria are also shown. An isomer of GDGT **36**, tentatively identified by Weijers *et al*. (2010), is not shown.

A study into the lipids of species belonging to the eubacterial order *Thermotogales* led to identification of a different set of non-isoprenoid ether lipids (Damsté *et al.*, 2007). These included GMM lipids (see Table 1.1), mixed ether/ester lipids (i.e. monoether/monoester, monoether/triester, diether/diester and triether/monoester) and GDGT lipids containing diabolic acid chains (see Clarke *et al.*, 1980; Carballeira *et al.*, 1997 for the origin of the nomenclature) or the saturated derivatives thereof. More specifically, the GDGT lipids contained either one saturated diabolic (i.e. 15,16-dimethyltriacontanyl) chain and one 13,14-dimethyloctacosanyl chain (**44**; Fig. 1.8), or otherwise two saturated diabolic chains (**45**) as the ether-bound moieties (Damsté *et al.*, 2007). GMD lipids containing unmethylated (**46**), monomethylated (**47**) or dimethylated (**48**) saturated diabolic chains (Fig. 1.8) were among cores of this class identified in authigenic carbonates from sulfur mines (Baudrand *et al.*, 2010).

Two lipids, one a GDD core and the other a GDGT core (**49** and **50**, respectively; representative structures shown in Fig. 1.9), which appear to show structural features that are a hybrid between those typical for archaeal ether lipids and those typical for their eubacterial equivalents have been identified in aquatic sediments (Thiel *et al.*, 1999; Schouten *et al.*, 2000). GDD **49** was tentatively assigned to contain an ether-bound phytanyl chain, typical of archaea, and an ether-bound *n*-C_{16} chain, common in eubacterial lipids (Thiel *et al.*, 1999). Unfortunately, the positions of the alkyl chains with respect to the glycerol carbinols was not determined. GDGT **50**, on the other hand, was shown to contain one ether-bound 3,7,11,15,18-pentamethyltriacontanyl chain and one ether-bound 3,7,11,15,18-pentamethyldotriacontanyl chain, although the orientation of the chains within the lipid was not determined (Schouten *et al.*, 2000). These two chains can be considered to comprise a phytanyl chain condensed with an *i*-C_{15} or an *i*-C_{17} chain, respectively. Both of these *iso* chains are typical products of eubacterial biosynthesis (see Table 1.1). The apparent paradoxical observations of domain-distinct hydrocarbons in the same lipids, and the ensuing question of their biogenic origin, are yet to be resolved.

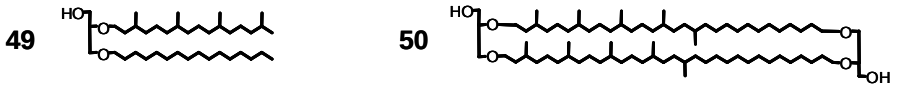


Fig. 1.9. Representative structures for unusual ether lipids containing alkyl chains which represent a hybrid between typical archaeal and eubacterial hydrocarbons. The distribution and orientation of the chains is not known for either lipid; consequently, the orientations shown are arbitrary choices made solely for illustrative purposes.

Although some eukaryotes also produce glycerol ether lipids, including plasmalogens (a major compound class which are also found in some eubacteria), these components are not further discussed.

1.4. Ether lipids as biomarkers

1.4.1. Isoprenoid lipids of archaeal origin

Important analytical developments which revolutionised the identification of archaea in natural environments include the invention of the polymerase chain reaction (Saiki *et al.*, 1988 and references therein) and the subsequent developments of fluorescence in-situ hybridisation (FISH; see Levsky and Singer, 2003) and a more sensitive extension of the latter technique, catalysed reporter deposition FISH (CARD-FISH; see Pernthaler *et al.*, 2002; Ishii *et al.*, 2004). These techniques have facilitated the amplification or fluorescent labelling of archaeal-specific DNA or 16S rRNA sequences found in environmental samples, allowing identification of the archaea, including those species which prove resistant to cultivation, that are living *in situ* (e.g. Fuhrman *et al.*, 1992; DeLong *et al.*, 1992; 1999; Karner *et al.*, 2001; Teira *et al.*, 2004; Herndl *et al.*, 2005; Rossel *et al.*, 2008; Jagersma *et al.*, 2009). A complementary approach to use of genetic markers for the identification of archaea has been to profile ether lipids in the environment. As ether-linked isoprenoid IPLs and their naked ether lipid cores are specific to the Archaea, both classes of compound serve as suitable biomarkers for the domain. It has been widely accepted that, following cell death, the polar head groups in IPLs from the cell membrane are rapidly hydrolysed over the order of a few days (Harvey *et al.*, 1986), leaving only the remnant lipid cores. Consequently, ether-linked isoprenoid IPLs present in the environment have been used as biomarkers for living archaeal cells, whereas core lipids have typically been used to indicate a dead or fossilised archaeal signature (e.g. Lipp *et al.*, 2008; 2009; Pitcher *et al.*, 2009a; Huguet *et al.*, 2010). This assertion has, however, been challenged recently following a theoretical prediction of the rates of decay of phosphatidyl and glycosyl lipids in marine environments (Schouten *et al.*, 2010). Following these calculations, the authors suggested that, although the phosphatidyl-capped IPLs are probably hydrolysed over a short timescale, the glycosyl-capped IPLs may persist over much longer time periods and be encapsulated into the developing sediment. This appears to be supported by the identification of fossilised glycolipids of heterocystous cyanobacterial origin in 49 million year old sediments (Bauersachs *et al.*, 2010).

Table 1.2. Distributions of isoprenoid lipid cores within the different phyla of the Archaea as reported to date[a].

		Crenarchaeota					***Euryarchaeota***										
Lipid(s)		*Thermoproteales*	*Desulfurococcales*	*Sulfolobales*	*Group I*	*Other*	*Thermococcales*	*Archaeoglobales*	*Methanococcales*	*Methanobacteriales*	*Thermoplasmatales*	*Halobacteriales*	*Methanomicrobiales*	*ANME*	*Methanosarcinales*	*Other*	Apparent Specificity
GDD	**1**	+[b]	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
	5-6		+									+			+[c]		
	7-10			+					+	+	+			+	+	+	
	11-12						+					+			+[c]	+	
GMD	**4**								+					+			*Methanocaldococcaceae* / ANME
	22-23													+			Methanotrophs
GTGT	**3**			+	+						+					+	
	+ Cp rings				+												
GDGT	**2**	+	+	+	+	+	+	+	+	+	+		+	+	+	+	
	14-17	+	+	+	+		+				+			+		+	
	18-21	+		+			+				+						
	24, **24''**, **25**				+												Group I *Crenarchaeota*
	26									+							*M. marburgensis*
GDCT	**2c**, **14c-21c**			+													*Sulfolobales*
GMGT	**27**						+									+	*Euryarchaeota*?
	28-31															+	*A. boonei*

[a] Compiled from De Rosa and Gambacorta, 1988a; Sprott *et al.*, 1990; Upasani *et al.*, 1994; Galliker *et al.*, 1998; Uda *et al.*, 2000; Gonthier *et al.*, 2001; Stadnitskaia *et al.*, 2003; Blumenberg *et al.*, 2004; Koga and Morii, 2005; Schouten *et al.*, 2007c; 2008b; 2008c; de la Torre *et al.*, 2008; De Boever *et al.*, 2009; Pitcher *et al.*, 2010 and references therein.

[b] (+) indicates at least one species from the phylum produces at least one of the categorised lipids.

[c] Sesterterpanyl and unsaturated GDDs noted for some *Methanosarcinales* (De Rosa and Gambacorta, 1988a; Nichols *et al.*, 1992) may have been confused with degradates of GDD **7** (Sprott *et al.*, 1990; Ekiel and Sprott, 1992).

Organic matter of biogenic origin preserved in aquatic sediments can also provide a record of planktonic productivity in the overlying waters, allowing the ecology, and often the *in situ* conditions present during deposition, to be determined. In marine environments, the major flux of archaeal GDGT lipid cores into the developing sediment appears to be sourced from epipelagic archaea, probably *Crenarchaeota*, inhabiting the shallower depths (<100 m) of the water column (Wuchter *et al.*, 2005; 2006). IPLs detected in deep sub-surface sediments could, therefore, either be representative of a fossilised signal from these pelagic archaea or, as has been suggested (Lipp *et al.*, 2008; 2009), reflect the viable community of archaea living within the sediments. In all likelihood, however, the IPL signal is attributable to a mixed origin from the two sources (Schouten *et al.*, 2010). Fossilised GDGT lipid cores, on the other hand, have been identified in ancient sediments from as old as the Upper Jurassic (Ventura *et al.*, 2007 and references therein), suggesting that the lipid cores can be preserved over geological timescales. Their long term survival in such sediments is believed to be a consequence of the stability of the ether linkages to chemolysis, thermolysis and enzymatic breakdown (Pease *et al.*, 1998 and references therein). Consequently, in cases where the sediment is stratified and in which the different strata can be reliably dated, changes in the archaeal population within the body of water overlying the sediment can be reconstructed as a function of time *via* interpretation of the stratigraphy of the lipid core distributions.

Some ether lipid cores are produced by a myriad of different archaea (Table 1.2) and, consequently, whilst they serve as domain-specific biomarkers, their identification does not allow classification of the species present to more specific phyla. For instance, GDD **1** is found in the vast majority of the Archaea for which lipid profiles are available (e.g. De Rosa and Gambacorta, 1988a; Koga and Morii, 2005; Schouten *et al.*, 2008c), probably reflecting its status as a key intermediate in the biosynthesis of other types of lipid core (e.g. GTGT, GDGT and GMT cores; see Section 1.2.1). In addition, **1** has been identified in a number of disparate environments including hydrothermal environments (e.g. Blumenberg *et al.*, 2007), terrestrial soils (e.g. Gattinger *et al.*, 2003) and methane cold seeps, in the latter case probably originating from ANME organisms (e.g. Blumenberg *et al.*, 2004; Niemann *et al.*, 2008; Rossel *et al.*, 2008). Similarly, GDGT **2** and Cp ring-containing GDGT

lipid cores **14-21** can originate from a variety of source organisms and, thus, are non-specific (e.g. De Rosa and Gambacorta, 1988a; Pancost *et al.*, 2001b; Damsté *et al.*, 2002b; Schouten *et al.*, 2007c). Several of the known ether lipids, on the other hand, are apparently restricted to a select group of archaea and, consequently, serve as more specific biomarkers (Table 1.2). Crenarchaeol (**24**) and its regioisomer (**24''**), for example, appear to be restricted to species of Group I *Crenarchaeota* involved in aerobic oxidation of ammonia, including both mesophilic (Damsté *et al.*, 2002b; Schouten *et al.*, 2008c) and thermophilic (de la Torre *et al.*, 2008; Pitcher *et al.*, 2009b; 2010) organisms.

1.4.2. Non-isoprenoid lipids of eubacterial origin

Eubacterial ether lipid cores present in natural environments can also serve as biomarkers, on occasion with reasonable taxonomic specificity. For example, GDD and GMM lipids exhibiting chains containing 14-18 carbons have been used as markers for sulfate-reducing eubacteria within AOM consortia in marine environments (e.g. Hinrichs *et al.*, 2000; Pancost *et al.*, 2001a; Niemann and Elvert, 2008). Other GDD and GMM lipids, with chains containing 15-21 carbons, have been used to reveal the presence of hyperthermophilic eubacteria in hot spring silica sinters and microbial mats (Jahnke *et al.*, 2001; Pancost *et al.*, 2005; 2006; Kaur *et al.*, 2008). A ladderane GMM, present either as an IPL structure or as its lipid core, has been used as a biomarker for organisms involved in anaerobic oxidation of ammonia (Anammox) in marine environments (e.g. Kuypers *et al.*, 2003; Wakeham *et al.*, 2007; Jaeschke *et al.*, 2009).

In spite of their "orphan" status (i.e. the fact that no biological progenitor for these lipids has been identified to date; Weijers *et al.*, 2009), some of the most widely used eubacterial ether lipid biomarkers are GDGT lipids **34-42**. Evidence is, however, beginning to accumulate to suggest that the eubacterial species which produce the lipids are anaerobes or facultative aerobes (Weijers *et al.*, 2006a), are heterotrophic (Oppermann *et al.*, 2010) and reside within the *Acidobacteria* (Weijers *et al.*, 2009; Peterse *et al.*, 2010) or, possibly, within the *Alphaproteobacteria* or *Actinobacteria* (Liu *et al.*, 2010). Structures **34-42** are often found to be the dominant GDGT lipids

in continental environments such as soils (Hopmans *et al.*, 2004; Weijers *et al.*, 2006b; 2007c), peats (Schouten *et al.*, 2000; Weijers *et al.*, 2006a), rivers (Herfort *et al.*, 2006b; Kim *et al.*, 2006; 2007) and lakes (Schouten *et al.*, 2000; Blaga *et al.*, 2009; 2010; Damsté *et al.*, 2009). By contrast, they typically represent a much smaller proportion of the total GDGT profiles of marine sediments, particularly those deposited in stretches of open ocean (Schouten *et al.*, 2000; Hopmans *et al.*, 2004). This has led to the suggestion that the organisms that produce **34**-**42** are indigenous to terrestrial soils and are transported to aquatic environments *via* fluvial processes such as surface run-off or erosion (Hopmans *et al.*, 2004). The identification of the lipids in marine and lake sediments can, therefore, indicate an allochthonous input of terrestrial material into the site. In order to quantify the effect, Hopmans *et al.* (2004) proposed the BIT index (Equation 1.1), an index relating the relative proportions of **34**-**36** to the proportion of crenarchaeol (**24**), as a proxy for terrestrial soil input into aquatic environments.

$$\text{BIT} = \frac{[\mathbf{34}] + [\mathbf{35}] + [\mathbf{36}]}{[\mathbf{34}] + [\mathbf{35}] + [\mathbf{36}] + [\mathbf{24}]} \quad (1.1)$$

Although **24** is found in soils, the proportion of this lipid relative to **34**-**36** is typically very low in such samples (Hopmans *et al.*, 2004; Weijers *et al.*, 2006b). Consequently, the vast majority of **24** identified in an aquatic sample is likely to have been produced within the water column, allowing **24** to be used as a fully autochthonous end-member. Likewise, as **34**-**36** are assumed to be formed solely in soils, they represent, as a lipid set, a fully allochthonous end-member within the marine environment. The BIT index, therefore, provides an approximate measure of the extent of admixing between the two end-members; a BIT value of 0 implies that all lipids have an autochthonous origin, whereas a BIT value of 1 implies that all lipids are terrigenous. Between these extremes, the closer the calculated BIT value to 1, the greater the inferred input of continental organic carbon. The proxy has subsequently been used to trace inputs of exogenous terrestrial organic carbon both in contemporary studies and in geochemical studies implemented to reconstruct

ancient aquatic environments (e.g. Menot *et al.*, 2006; van Dongen *et al.*, 2008; Yamamoto *et al.*, 2008; Kim *et al.*, 2009a; 2009b; Smith *et al.*, 2010).

Several factors can, however, potentially complicate interpretations made using the BIT index regarding transport of organic matter. In particular, the assumptions regarding the nature of the end-members used in the index have had to be relaxed. Although crenarchaeol levels in most soils are typically low compared to the levels of non-isoprenoid lipids, some soils show more significant levels (Weijers *et al.*, 2006b; Leininger *et al.*, 2006). Flux of terrestrially-derived **24** into aquatic environments would augment the pool of this lipid that was derived *in situ*, leading to an, albeit small, reduction in calculated BIT and, consequently, an underestimation of the extent of fluvial transport of terrestrial organic matter into the site. In addition, other studies have suggested that non-isoprenoid GDGT lipids may be produced *in situ* in lacustrine (Damsté *et al.*, 2009; Tierney *et al.*, 2009; 2010; Bechtel *et al.*, 2010; Blaga *et al.*, 2010) and marine (Peterse *et al.*, 2009a) environments. As the autochthonous lipids are not discriminated from those derived from soil, their quantification *via* use of the BIT index could lead to an overestimation of the transported organic matter. Furthermore, whilst the BIT index does appear to correlate to fluvially transported input of soil-derived organic carbon, it does not account for aeolian transport (Hopmans *et al.*, 2004) or for other forms of terrestrial organic matter, such as plant debris (Walsh *et al.*, 2008). As such, calculated BIT values may underestimate the total terrestrial organic matter input, particularly in areas without a readily available soil or peat as an exogenous source (Walsh *et al.*, 2008).

Diagenetic effects also appear to have an influence on the BIT values determined for marine sediments. Huguet *et al*. (2008; 2009) examined turbidites from the Madeira Abyssal Plain and determined that crenarchaeol is more substantially degraded than the non-isoprenoid GDGT lipids following prolonged exposure to oxic conditions. This leads to substantial differences in BIT values between the oxidised and unoxidised sections of the turbidites, despite both sections of the sediment originating from the same depositional environment (see Huguet *et al.*, 2008 and references therein). Consequently, whilst it appears that the BIT index can quantify

the degree of terrestrial organic matter preserved in such sediments, it does not necessarily reflect the quantity of organic matter actually transported from the continent (Huguet *et al.*, 2008; 2009).

An international study in which several laboratories partook in analysis of the same two sediment lipid extracts indicated that, whilst measurements made for BIT showed, in general, good reproducibility at each laboratory, the interlaboratory variation in the determined BIT values was much larger (Schouten *et al.*, 2009). This was attributed to differential relative response factors for eubacterial lipids **34**-**36** to archaeal lipid **24** at the different laboratories. Experimental factors can, therefore, also introduce bias into the determination of BIT values.

1.5. Homeoviscous adaptation of the archaeal membrane

Whereas diether lipids are incorporated into a bilayer within the archaeal cell membrane (Fig. 1.10a; De Rosa *et al.*, 1982; 1983a), tetraether lipids are bolaamphiphilic and, consequently, span the cell membranes (Fig. 1.10b; Langworthy *et al.*, 1982; Gliozzi *et al.*, 2002; Albers *et al.*, 2006). The relative proportions of GDD, GMD, GTGT and/or GDGT lipids expressed by an archaeon are, to some extent, dependent on the kingdom and phylum to which the organism belongs. Halophilic archaea, for example, contain GDD **1** as their dominant membrane lipid whereas methanogenic archaea often show significant levels of GTGT **3** and GDGT **2** in addition to lipid **1** (De Rosa and Gambacorta, 1988a; Koga *et al.*, 1993a; 1993b). Apart from phylogenetic differences, adaptive changes can also lead to variations in the ratio of lipids expressed by a particular organism. For example, varying the growth temperature of cultures of *M. jannaschii* results in differences in the average degree of archaeol cross-coupling observed, with more **2** and **4** formed from **1** when the organism is stressed at increasingly high temperatures (Sprott *et al.*, 1991). Equivalent increases in the ratio of **2** to **1** were observed in both *Archaeoglobus fulgidus* (Lai *et al.*, 2008) and *Thermococcus kodakaraensis* (Matsuno *et al.*, 2009) grown at incrementally higher temperatures. Such variations in lipid distributions are believed to reflect cellular responses that reduce membrane fluidity at elevated temperatures, effected by incorporating higher proportions of the

more rigid membrane-spanning lipids. Similarly, archaea can also show differing proportions of expressed diether and tetraether lipids during different phases of growth (Kramer and Sauer, 1991; Sprott *et al.*, 1991; Morii and Koga, 1993; Matsuno *et al.*, 2009). In one study, the proportion of GDGT lipid (as a percentage of the total membrane lipid) expressed by a set of archaea was found to correlate more strongly with the optimal pH than with the optimal temperature for growth of each organism (Macalady *et al.*, 2004). This led to the suggestion that tetraether lipid expression in archaea is predominantly a mechanism which allows survival of the organisms in acidic conditions (i.e. at decreased pH).

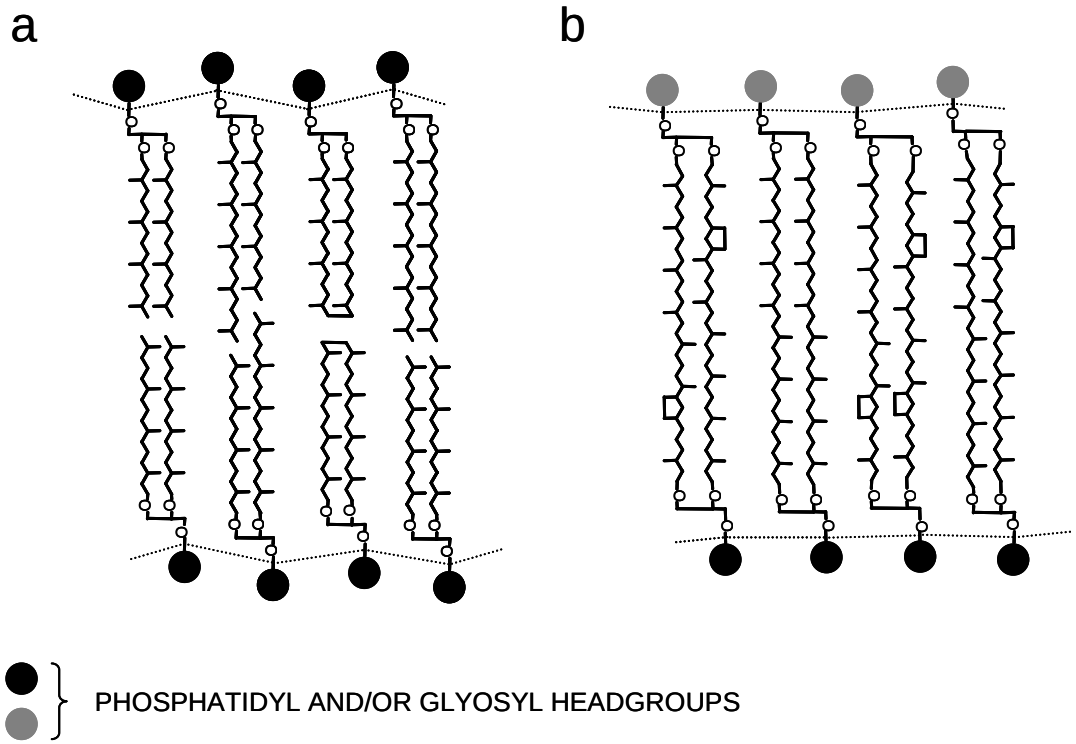


Fig. 1.10. Hypothetical archaeal cell membrane envelopes comprising: a) diether lipid bilayer; b) tetraether lipid monolayer. Adapted from De Rosa *et al*. (1982; 1983a) with alterations made following consideration of Gulik *et al*. (1988). No known archaeon expresses all of the diether lipids shown in the composite bilayer shown in a). Archaeal membranes which are predominantly of b) character can contain an a) component and *vice versa*.

In addition to changes in the type of lipid core utilised in the cellular membrane, archaea can also adapt the isoprenoid chains expressed within the lipid cores to counteract changes in physiological conditions. The expression of sesterterpanyl chains in the diether lipids of some species of *Halobacteriaceae* is related to the salinity of the medium used for growth, although whether the proportion of sesterterpanyl-containing lipids expressed increases or decreases under elevated salinity is strain-dependent (Morth and Tindall, 1985). The average degree of unsaturation observed within the GDD lipids of the psycrophilic archaeon *Methanococcoides burtonii* increases upon growth of the organism at lower temperature (Nichols *et al.*, 2004). On the other hand, the average number of Cp rings in the GDGT lipids of thermophilic archaea increases with increasing growth temperature (De Rosa *et al.*, 1980b; Uda *et al.*, 2001; 2004; Lai *et al.*, 2008; Shimada *et al.*, 2008). Similarly, an increase in the temperature of growth of *Sulfolobus solfataricus* and a second *Sulfolobus* species (genetically similar to *Sulfolobus shibatae*) leads to an increase in the average number of Cp rings incorporated into the GDCT lipids (De Rosa *et al.*, 1980b; Murae *et al.*, 2001; 2002). Ether lipid membranes consisting of structures containing more Cp rings are more tightly packed and have higher gel-to-liquid crystalline transition temperatures (Gliozzi *et al.*, 2002 and references therein). Accordingly, adaptation of the proportion of Cp ring-containing lipids utilised in the membrane probably regulates its behaviour, perhaps altering the fluidity or proton permeability (Gabriel and Chong, 2000; Gliozzi *et al.*, 2002). This adaptive behaviour appears to extend to species of mesophilic *Crenarchaeota*, which, at higher temperatures, also increase the number of rings incorporated into the GDGT isoprenoid chains (Damsté *et al.*, 2002b; Wuchter *et al.*, 2004; Schouten *et al.*, 2007a). Temperature is, however, not the only physiological stress that leads to changes in the proportions of ring-containing tetraether lipids expressed by archaea: ambient pH (Shimada *et al.*, 2008) and the presence (or absence) of oxygen during growth (Trincone *et al.*, 1989) are also known to influence the degree of cyclisation observed in some organisms. In addition, one archaeon, *Thermoproteus tenax*, appears to vary the number of Cp rings in its membrane lipid cores depending on whether the organism is grown heterotrophically or autotrophically (Thurl and Schäfer, 1988). *Sulfolobus*

acidocaldarius, on the other hand, shows similar core lipid profiles irrespective of the nutritional status adopted by the organism during growth (Langworthy, 1977).

The polar head groups in archaeal IPL structures are also subject to variation in response to changes in physiological conditions. The average number of glycosyl units in the polyglycosylated phospholipids of *T. acidophilum*, for example, increases with both increasing growth temperature and reducing pH (Shimada *et al.*, 2008).

1.6. Molecular proxies for palaeotemperature

A number of proxies have been devised for correlating parameters derived from molecular abundances to marine environmental temperature. These calibrations can then be applied to interpret similar measurements made for geological samples in order to aid in reconstructions of past oceanic temperatures. Many of the established proxies are based on inorganic measurements, such as isotopic ($^{18}O/^{16}O$; Urey, 1947; Epstein and Mayeda, 1953; Erez and Luz, 1983) or elemental (Mg/Ca; Elderfield and Ganssen, 2000; Sr/Ca; Beck *et al.*, 1992) ratios of carbonates found in the fossilised shells of planktonic foraminifera or in corals. An organic palaeothermometer proxy, the U^k_{37} proxy, based upon fossilised unsaturated alkenones originating from haptophyte algae, has also been developed (Brassell *et al.*, 1986; Prahl *et al.*, 1988). The alteration of the membrane lipids of mesophilic archaea and eubacteria in response to environmental temperature (Wuchter *et al.*, 2004; Schouten *et al.*, 2007a; Weijers *et al.*, 2007c) has allowed GDGT lipid cores to be used as the basis set for two alternative organic molecular proxies for palaeotemperature; the TEX_{86} index and the MBT/CBT index for reconstruction of terrestrial mean annual air temperature (MAAT).

1.6.1. TEX_{86} index

The identification and quantification of the crenarchaeol regioisomers (**24** and **24''**) and tetraether lipids containing 1-3 Cp rings (**14**-**16**) in sediments from both marine and lacustrine environments has led to the development of a proxy for temperature,

the TEX_{86} index (Equation 1.2), which shows positive correlation to increasing surface temperature of the water body at the time of sediment deposition (Schouten *et al.*, 2002; Powers *et al.*, 2004).

$$TEX_{86} = \frac{[\mathbf{15}] + [\mathbf{16}] + [\mathbf{24''}]}{[\mathbf{14}] + [\mathbf{15}] + [\mathbf{16}] + [\mathbf{24''}]} \quad (1.2)$$

$$SST = (56.2 * TEX_{86}) - 10.78 \quad (1.3)$$

The initial calibration, determined for contemporary environments using 44 marine surface sediments and their associated average overlying sea surface water temperatures (SSTs), employed a linear regression model (Schouten *et al.*, 2002). A similar TEX_{86} calibration established using surface sediments from large lakes has also allowed extension of the proxy for use in reconstruction of continental freshwater palaeoenvironments (Powers *et al.*, 2004; 2010). Sluijs *et al.* (2006) suggested that an alternative index, the TEX_{86}' index, in which the contribution of [**16**] to the denominator of the TEX_{86} index is removed, is more appropriate for reconstruction of subtropical arctic ocean temperatures. More recently, a larger data set, comprising 223 core-top sediments from a variety of worldwide locations, has allowed the proxy to be calibrated for use globally, with the correlation between TEX_{86} and SST (Equation 1.3) remaining linear for SSTs between 5-30^{o}C (Kim *et al.*, 2008). Using the same data set, Liu *et al.* (2009) independently suggested a non-linear regression model for conversion of calculated TEX_{86} into SST. In a more extensive study, Kim *et al.* (2010) proposed two new logarithmic indices, TEX_{86}^{L} (Equation 1.4) and TEX_{86}^{H} (Equation 1.5) with each index correlated to SST using linear regression (Equations 1.6-1.7). The new indices provide the best correlations to the global SST data set, with TEX_{86}^{H} recommended for palaeoreconstructions of SSTs above 15^{o}C and TEX_{86}^{L} recommended for reconstructions in which SST is below 15^{o}C (Kim *et al.*, 2010). For reconstructions in which temperatures both above and below 15^{o}C may be expected, such as polar oceans in past greenhouse environments, TEX_{86}^{L} was suggested as the more reliable proxy.

$$TEX_{86}^{L} = \log\left\{\frac{[\mathbf{15}]}{[\mathbf{14}] + [\mathbf{15}] + [\mathbf{16}]}\right\} \quad (1.4)$$

$$TEX_{86}^{H} = \log\left\{\frac{[\mathbf{15}] + [\mathbf{16}] + [\mathbf{24''}]}{[\mathbf{14}] + [\mathbf{15}] + [\mathbf{16}] + [\mathbf{24''}]}\right\} \quad (1.5)$$

$$SST = (67.5 * TEX_{86}^{L}) + 46.9 \quad (1.6)$$

$$SST = (68.4 * TEX_{86}^{H}) + 38.6 \quad (1.7)$$

Whereas changes in oceanic salinity can alter oxygen isotope ratios in foraminifera shells and, consequently, distort temperature reconstructions based upon $^{18}O/^{16}O$ ratios in the fossilised shell carbonates (Jaffrés *et al.*, 2007 and references therein), the TEX_{86} proxy appears to be both salinity and nutrient independent (Wuchter *et al.*, 2004). Furthermore, given the excellent preservation characteristics of GDGT core lipids in ancient sediments, the proxy can be used in SST reconstructions from sediments of ages well in excess of the maximum age afforded from use of the U_k^{37} index (~6 Ma; Schouten *et al.*, 2002). Consequently, the SST calibrations have allowed widespread reconstruction of palaeotemperatures from as old as the Cretaceous *via* determination of the TEX_{86} indices of ancient sediments (e.g. Schouten *et al.*, 2003b; Jenkyns *et al.*, 2004; Dumitrescu *et al.*, 2006; Hofmann *et al.*, 2008; Forster *et al.*, 2007).

Despite the extensive usage and general success of the TEX_{86} palaeothermometer, a number of reports have identified factors which may bias TEX_{86} values and/or compromise the accuracy of temperatures reconstructed. These can include: changes in water column pH (Pearson *et al.*, 2008), correlation of the sedimentary signal with subsurface water temperature as opposed to SST (Menzel *et al.*, 2006; Huguet *et al.*, 2007; Lee *et al.*, 2008), seasonal variation in epipelagic *Crenarchaeota* productivity and/or transport of the lipids produced by these organisms to the sediment (Wuchter *et al.*, 2005; 2006; Herfort *et al.*, 2006a; Huguet *et al.*, 2007; Blaga *et al.*, 2009; Powers *et al.*, 2010), an unusual ecology of *Crenarchaeota* within the water column (Trommer *et al.*, 2009), variable expression of the crenarchaeol regioisomer by

different species of marine *Crenarchaeota* or as a function of unidentified changes in growth conditions (Wuchter *et al.*, 2004; Schouten *et al.*, 2007a; Schouten *et al.*, 2008c; Pitcher *et al.*, 2010), unusually high proportions of individual Cp-ring containing lipids in the sediment (Sluijs *et al.*, 2006; Damsté *et al.*, 2009), lateral or upward transport of archaeal lipid signatures within the water column (Herfort *et al.*, 2006a; Lee *et al.*, 2008), long-term oxic degradation of GDGT lipids (Huguet *et al.*, 2009), thermal maturation of the sediment (Schouten *et al.*, 2004), contributions of lipids from *Euryarchaeota* or active benthic archaea to the sedimentary pool (Shah *et al.*, 2008; Blaga *et al.*, 2009; Lipp *et al.*, 2009; Powers *et al.*, 2010), allochthonous input of terrestrially-derived isoprenoid lipids (Sluijs *et al.*, 2006; Weijers *et al.*, 2006b; Blaga *et al.*, 2009; Huguet *et al.*, 2009; Powers *et al.*, 2010), a more appropriate local, as opposed to global calibration (Trommer *et al.*, 2009; Bechtel *et al.*, 2010) and laboratory or analytical instrument-related inaccuracies (Schouten *et al.*, 2009; Escala *et al.*, 2009). The influence of such factors must be considered carefully before palaeotemperature reconstructions using the TEX_{86} index are attempted.

1.6.2. MBT/CBT index

Weijers *et al.* (2007c) identified apparent correlations between the relative proportions of branched non-isoprenoid GDGT lipids found in terrestrial soils and the *in situ* environmental MAAT and soil pH. Specifically, two indices calculated using the concentrations of **34**-**42** in the soils proved to be of particular use, the first relating to the degree of methyl branching in the lipids (the methylation index of branched tetraethers, MBT; Equation 1.8) and the second the logarithm of a parameter accounting for the degree of Cp ring expression within the lipids (the cyclisation index of branched tetraethers, CBT; Equation 1.9).

$$\text{MBT} = \frac{[\mathbf{34}] + [\mathbf{35}] + [\mathbf{36}]}{[\mathbf{34}] + [\mathbf{37}] + [\mathbf{40}] + [\mathbf{35}] + [\mathbf{38}] + [\mathbf{41}] + [\mathbf{36}] + [\mathbf{39}] + [\mathbf{42}])} \quad (1.8)$$

$$\text{CBT} = -\log \left\{ \frac{[\mathbf{37}] + [\mathbf{38}]}{[\mathbf{34}] + [\mathbf{35}]} \right\} \quad (1.9)$$

$$\text{MBT} = 0.122 + (0.187 * \text{CBT}) + (0.020 * \text{MAAT}) \quad (1.10)$$

$$\text{CBT} = 3.33 - (0.38 * \text{pH}) \quad (1.11)$$

Principal components analysis of data obtained from an array of globally distributed soil samples indicated that changes in soil pH were largely responsible for variations in the values of CBT, whereas both MAAT and pH contributed to variations in MBT values (Weijers *et al.*, 2007c). This led to the development of empirical calibrations for palaeoreconstruction of both continental soil pH and MAAT (Weijers *et al.*, 2007c). These calibrations (Equations 1.10-1.11) have subsequently been used to reconstruct ancient terrestrial MAATs *via* interpretation of branched non-isoprenoid GDGTs present in marine sediment cores, whose presence therein was attributed to fluvial transport from the continent (Weijers *et al.*, 2007a; 2007b; Schouten *et al.*, 2008d; Donders *et al.*, 2009).

Whilst the relationships between the CBT and MBT indices and pH or MAAT have been qualitatively verified in other studies using contemporary soils (Damsté *et al.*, 2008; Peterse *et al.*, 2009a; 2009b; 2009c; 2010) and lake (Tierney *et al.*, 2010; Zink *et al.*, 2010) or marine (Rueda *et al.*, 2009) sediments, several authors have suggested that local calibrations may be more accurate predictors of MAAT than the global calibration (Damsté *et al.*, 2008; Peterse *et al.*, 2009b; Tierney *et al.*, 2010; Zink *et al.*, 2010). As discussed above (Section 1.4.2), evidence is emerging to suggest that, in addition to their production in soils, non-isoprenoid GDGT lipids **34-42** may also be produced in lakes and in the ocean (Damsté *et al.*, 2009; Peterse *et al.*, 2009a; Tierney *et al.*, 2009; 2010; Bechtel *et al.*, 2010; Blaga *et al.*, 2010). Both allochthonous and autochthonous non-isoprenoid GDGT lipid cores would be

preserved in the aquatic sediments under this regime, with lipids from the latter source more likely to reflect lake or sea water temperature than MAAT. As a consequence, a degree of restraint in using such sediments for reconstruction of continental MAAT has been advocated unless the origin of the non-isoprenoid GDGT lipid signal is clear (Tierney *et al.*, 2009; Damsté *et al.*, 2009; Peterse *et al.*, 2009a; Bechtel *et al.*, 2010; Blaga *et al.*, 2010).

1.7. Instrumentation

1.7.1. Gas chromatography

Chromatography is a term used to refer to any process which can separate a mixture of components on the basis of their differential transfer equilibria between a mobile phase, used to carry the components, and an immobilised or stationary phase, through which they migrate (Skoog *et al.*, 1998). One of the most widely used instrumental chromatographic techniques is gas chromatography (GC), which can facilitate the separation, detection and quantification of individual components in extremely complex mixtures. The physicochemical principle which underpins GC separation is the partitioning of analytes between a mobile gaseous phase and a liquid stationary phase at temperatures above their boiling points (James and Martin, 1952). Notably, this requires that analytes are in a gaseous form at the onset of the separation.

In modern GC instruments, samples are introduced by injection of a small plug (typically 1μL) of solution containing the mixture of components, whereby both the solvent and the dissolved analytes are flash volatilised. This can be mediated by injection into a vaporisation chamber, with subsequent carriage of the gaseous molecules onto a heated chromatographic column by the gaseous mobile phase (split/splitless injection). Alternatively, sample introduction can occur *via* injection directly onto the column followed by *in situ* vaporisation (on-column injection). The column is typically an open capillary tube (length 1-50 m) made of an inert, thermally conductive material (often fused silica) which is coated on its internal face with a thin film (0.1-5 μm thickness) of an immobilised thermostable liquid of low

volatility, which acts as the stationary phase. The column is seated in a thermostat-regulated oven, which can be heated to temperatures of up to ~400^{o}C as required. The temperature employed during the separation can either be fixed or gradually varied *via* temperature programming. An inert carrier gas (typically H_2 or He) operated at a constant flow rate acts as the eluent, carrying the vaporised analytes through the column. The retention of each analyte on the column is foremost dependent on its volatility. Highly volatile components (for instance molecules of the solvent) are sparingly soluble in the liquid film and, as the gaseous phase is mobile, these components rapidly migrate through the column and are eluted after only a short period of time. Components of lower volatility, on the other hand, have a greater propensity for dissolution and reside for longer periods on the column. The magnitude of this residence period, more commonly known as the retention time, increases with reducing analyte volatility, often leading to different retention times for molecules with distinct structures and, consequently, to their separation. Provided the analytical conditions remain constant, the retention times observed for a given analyte should be constant and, as such, this measurement proves to be of characteristic value. Factors related to instrument setup can affect absolute retention times including the column length and diameter, the film thickness, the carrier gas composition and its flow rate and the temperature program. The composition of the liquid film can influence not only the absolute retention times, but also the elution order of the analytes. A number of film chemistries are available, many of which are polyalkyl siloxanes, with polydimethyl siloxane the least polar option. Replacement of a proportion of the methyl substituents in the phases with other, more polar (e.g. phenyl) groups leads to greater retention of polar analytes.

As molecules elute from the end of the column, they can be detected using one of a variety of different detectors, two of the most widely used of which are the flame ionisation detector (FID) and mass spectrometer (see Section 1.7.3). During FID detection, the eluting gaseous analyte stream is mixed with H_2 and air and subsequently ignited. This generates ions which, over a given interval of time, provide a detectable current with a magnitude roughly proportional to the mass of reduced carbon that has been pyrolysed. As such, a plot of current vs. time generates a chromatogram which exhibits a series of peaks relating to eluting components.

Integration of the areas of these peaks, coupled to comparison with standards can allow for quantification of each analyte.

1.7.2. High performance liquid chromatography

One limitation of GC separation is the requirement for analytes to be in the gaseous state, which excludes analysis of non-thermally labile (i.e. those with a high polarity, large mass or containing long aliphatic chains) or thermally unstable components. An alternative instrumental approach for separation of mixtures is by high performance liquid chromatography (HPLC; often further abbreviated to LC). The retention of analytes which facilitates separation *via* HPLC occurs by transfer between solid and liquid phases (Skoog *et al.*, 1998). As the sole requirement in this case is that analytes are soluble, HPLC can facilitate separation of components that are not GC-amenable.

A liquid chromatograph used for HPLC typically comprises a high pressure pump, a solvent degasser, a solvent proportioning chamber, a sample injection system, a column oven, a detector and solvent transfer lines to connect the various components. The separation itself occurs on an interchangeable stainless steel column (length 5-30 cm) packed with a series of small, highly uniform particles of diameter of 2, 3, 5 or 10 μm, typically composed of a modified or non-modified silica. The column temperature can be regulated if required. Up to four separate solvents can be used to make up the eluent in some chromatographs, with these solvents degassed, mixed in appropriate proportions and pumped through the column at high pressure (typically ranging from 50-1000 bar). The analyte mixture is dissolved in a solvent chosen to closely match the initial eluent composition and injected as a small volume (usually 1-20 μL) into the pre-column flow. The analytes are carried onto the column and migrate along its length until they are eluted and detected, either *via* a spectroscopic technique or otherwise, a spectrometric technique (see Section 1.7.3). The rate of migration along the column and, consequently, the retention time of each analyte is governed by the column length and diameter, the column temperature and the eluent flow rate. The most influential factor, however, to affect the migration and separation of the analytes is the partitioning equilibrium of

each between the eluent and the stationary phase; as was discussed for GC, analytes which are partitioned into or absorbed onto the stationary phase more effectively will take longer to elute. The implicit interplay between the two phases provides the possibility to alter both analyte retention times and their elution order by modifying the characteristics of one or both of the phases. The simplest way in which analyte retention time can be altered is to change the polarity of the mobile phase; either by altering the solvents used during isocratic elution, in which the elutropic strength of the mobile phase does not vary, or by utilising gradient elution, in which the elutropic strength is gradually increased as a function of time. An alternative and more potent way of altering chromatographic behaviour is to utilise a different stationary phase and choose eluents which are appropriate for the chemistry of that phase. Such choices lead to chromatographic separations which fall into two main categories; normal- and reversed-phase HPLC.

1.7.2.1. Normal-phase liquid chromatography

In normal-phase HPLC, the stationary phase has a polar chemistry, with the column typically packed with silica or silica modified to bear bonded amino or cyano functionality. The mobile phase is of low to medium polarity, consisting of apolar solvents, usually *n*-hexane or *n*-heptane, mixed with small amounts of polar modifiers, such as ethanol, propanol, isopropanol or ethyl acetate. The mechanism of retention in normal-phase HPLC is believed to be through adsorption of analytes onto the stationary phase *via* hydrogen bonding and dipole-dipole type interactions with the silanol or modified silanol groups on the sorbent (see Snyder *et al.*, 1980; 1982). Given the nature of the adsorptive interactions, the Lewis acidity/basicity and molecular polarisability of an analyte each contribute to its retention; in general, the more polar an analyte, the later its elution from the column. In competition with adsorption is the partitioning of the analyte into the mobile phase or displacement of the analyte by a solvent molecule, both of which are favoured if the polarity of the eluent more closely resembles that of the stationary phase. As such, gradient elution in normal-phase HPLC proceeds by increasing the polarity of the mobile phase as a function of time to invoke elution of more strongly retained analytes.

1.7.2.2. Reversed-phase liquid chromatography

In reversed-phase HPLC, the stationary phase has an apolar chemistry, comprising a modified silica, bonded typically with octyl (C_8) or octadecyl (C_{18}) alkyl chains. The mobile phase consists of a polar eluent, usually an aqueous-based buffer with proportions of methanol, acetonitrile or isopropanol added. By contrast with normal-phase HPLC, less polar analytes are retained for longer during reversed-phase HPLC, although the mechanism by which retention occurs is not fully understood. Nevertheless, two models for the retention process have been proposed. In the first, the partition model, the hydrocarbon chains in the bonded silanol groups are perceived to act as a bulk homogeneous liquid (Vailaya and Horváth, 1998). In this regime, analytes partition between this immobile hydrophobic liquid and the more polar mobile phase. As a result of both hydrophobic interactions with the polar eluent and lipophilic interactions with the bonded phase, analytes which are more apolar are partitioned more effectively into the hydrocarbon layer and elute later. In the second model, retention is assumed to occur *via* physical adsorption of analytes onto the alkyl chains, either at the interfacial surface with the solvent or deeper within the hydrocarbon layer (Nikitas *et al.*, 2004). In this model, lipophilic molecules are adsorbed more strongly *via* increased Van der Waals interaction and, consequently, are retained for longer. As neither model can fully explain all of the experimental observations made regarding the phenomenon of reversed-phase separation, it has been suggested that both mechanisms may be in operation simultaneously (Rafferty *et al.*, 2007). Furthermore, particles in reversed-phase columns typically contain some residual unmodified silanol groups, which can interact with hydrophilic groups in the analytes and affect retention (Welsch *et al.*, 1990). Irrespective of the retention mechanism, a reduction in eluent polarity facilitates earlier elution of analytes. As such, gradient elution in reversed-phase HPLC typically starts with the eluent at high polarity, which is subsequently reduced to elute strongly retained, apolar components.

1.7.3. Mass spectrometry and hyphenated techniques

Mass spectrometry (MS) was developed in the early part of the 20th century (see de Hoffmann and Stroobant, 2002) as a technique to measure the mass-to-charge (*m/z*) ratios of ionic species in the gas phase. Subsequent instrumental developments have led to the widespread use of this technique as a tool for characterising organic molecules, even in samples where they are present in trace quantities. A typical mass spectrometer consists of a number of key parts: an inlet, through which a sample can be introduced; an ionisation source, used to produce gaseous ions from molecules; a mass analyser, which separates ions on the basis of *m/z* ratios; a detector, used to detect ions; a data processor, used to quantify and interpret the ion current and a vacuum pump (or series of pumps), used to evacuate charge-neutral gas molecules from the instrument. The information provided by the spectrometer is not restricted solely to the *m/z* charge ratio of the ionised species, but can also include the *m/z* ratios of its isotopomers (i.e. molecules with the same molecular formula but which contain different isotopes of some of the atoms) or of fragment ions formed from the species during or subsequent to ionisation. Together, these measurements can be used as characteristic fingerprints to confirm the presence of previously reported analytes or be used to infer the structure of novel molecules.

The coupling of chromatographic techniques to mass spectrometry, to form so-called "hyphenated" techniques, has led to a revolution in analytical chemistry. One of the drawbacks of GC with FID detection is that the detector response is indiscriminate of the type of organic analyte pyrolysed. Gas chromatography-mass spectrometry (GC-MS), on the other hand, provides information regarding the mass and structure of each separated component. Similarly, spectroscopic detectors used to detect analytes during HPLC, with the exception of evaporative light-scattering detection (ELSD), require some form of chromophore within the molecules. Liquid chromatography-mass spectrometry (LC-MS) is not hampered by this restriction.

The coupling of GC to MS was reasonably facile as both techniques require analytes to be in the gaseous phase. As such, GC-MS instruments often use classical ionisation sources, such as electron ionisation sources. The challenge in coupling

HPLC to MS was far more substantial; the large quantities of solvent which also elute along with the analyte in HPLC could either compromise the vacuum in the ionisation source or lead to trapping of the analyte *via* freezing of the solvent under the reduced pressure. As such, establishment of robust LC-MS relied upon the development of sources which facilitate ionisation at atmospheric pressure; namely electrospray ionisation and atmospheric pressure chemical ionisation. In both GC-MS and LC-MS, a variety of different mass analysers have been employed, including magnetic sector, time of flight and ion trap analysers.

1.7.3.1. Electron ionisation

Electron ionisation (EI; de Hoffmann and Stroobant, 2002) involves introduction of a gaseous sample into an ionisation chamber in a direction perpendicular to an accelerated electron beam (with electrons typically at 70 eV) formed by thermionic emission (Fig. 1.11). As the accelerated electrons collide with the gas molecules, they knock out electrons residing in the molecular outer electron shells, forming positive molecular ions ($M^{+\bullet}$). These odd-electron ions are often metastable, and undergo fragmentation to generate both odd- and even-electron ions typically of lower *m/z*. Upon exiting the ionisation chamber, the $M^{+\bullet}$ and fragment ions are focused and accelerated into the mass analyser. Subsequent detection of these ions yields an information rich mass spectrum which can be used to reconstruct the structure of the original analyte.

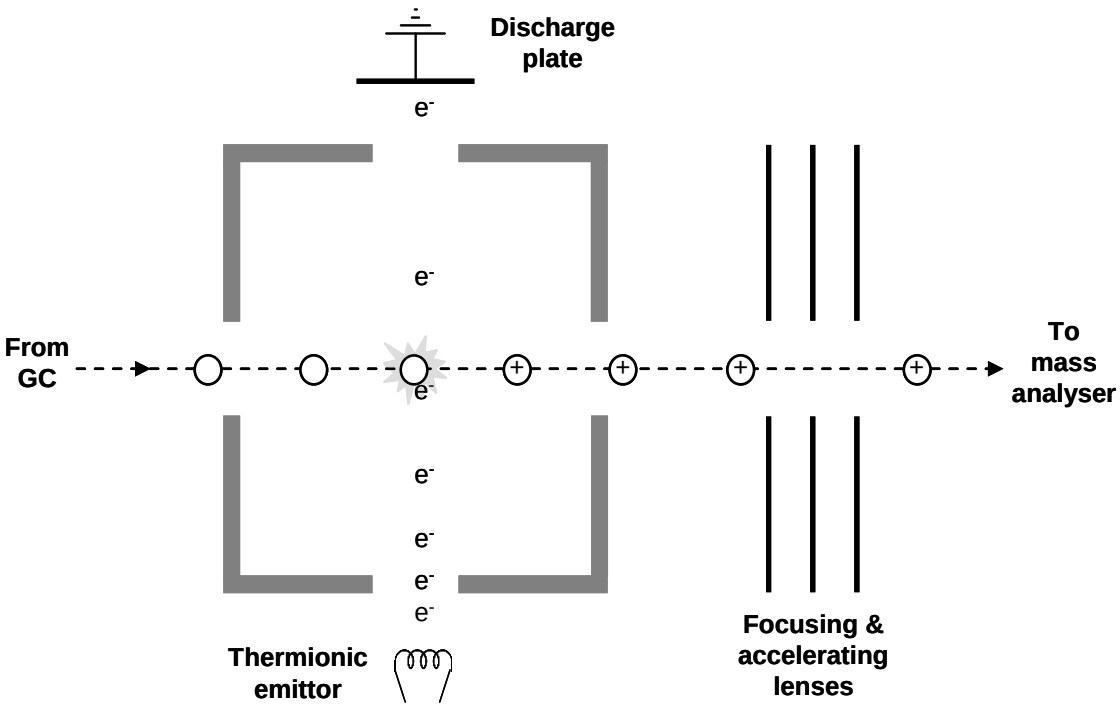


Fig. 1.11. Diagram of an EI source (adapted from de Hoffmann and Stroobant, 2002).

1.7.3.2. Electrospray ionisation

Although the principle physical technique had been developed much earlier (Dole *et al.*, 1968), electrospray ionisation (ESI) was fashioned for use in mass spectrometry in the mid-1980s, primarily as a tool to effect ionisation of macromolecules (see Fenn, 2002). It has henceforth blossomed into one of the most frequent choices of source for ionisation from liquid samples in modern MS instruments, on account of the significant analytical advantages that it provides. ESI facilitates ionisation of involatile and/or thermally labile components which are not amenable to EI. It is also deemed a "soft" ionisation technique as the ions formed have low residual energy and undergo little fragmentation post-formation. In addition, ESI can lead to formation of highly charged ions. Since *m/z* is the measurable quantity in MS, multiple charging of large molecules can generate ions with an *m/z* value within the operational range of available mass analysers. Consequently, macromolecules containing multiple ionisation sites, such as proteins, can be both ionised and detected intact by ESI MS.

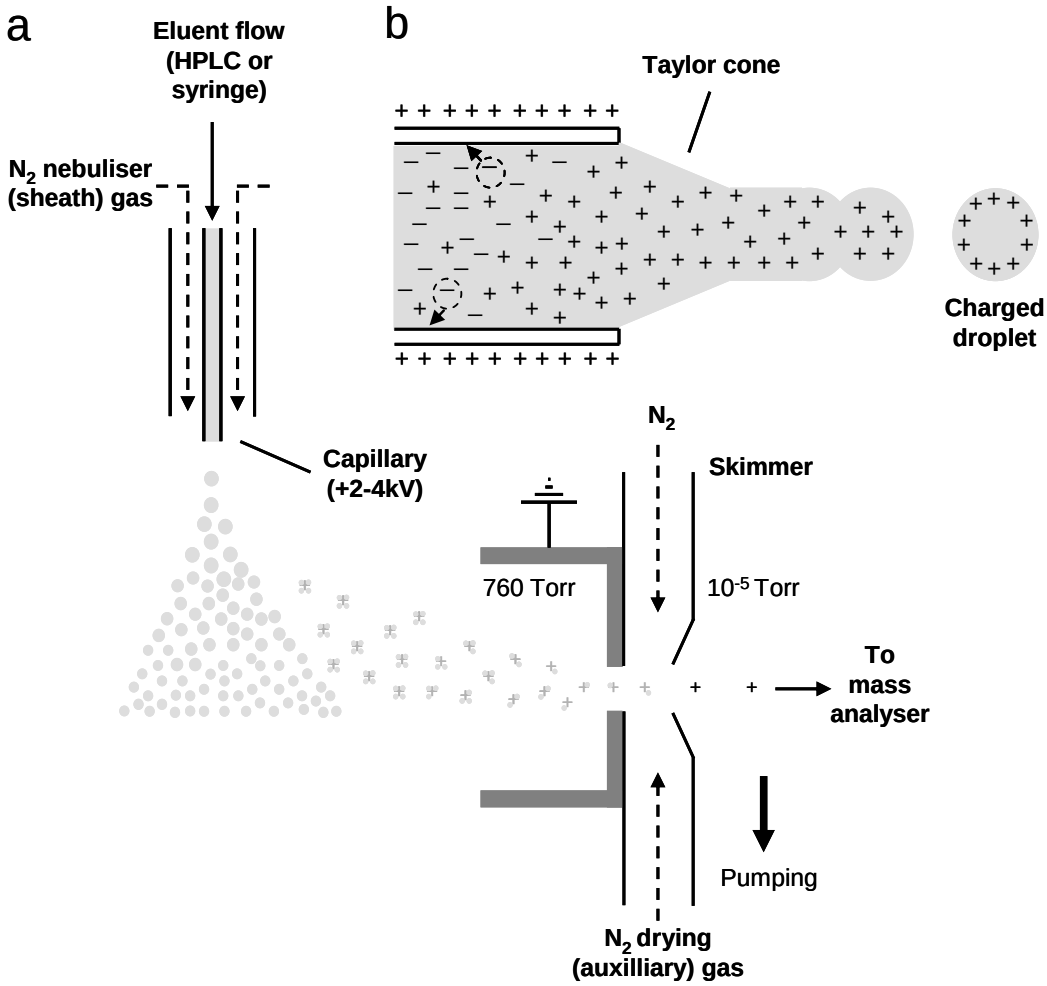


Fig. 1.12. a) Diagram of an orthogonal ESI source, operated in positive ion mode (modified from de Hoffmann and Stroobant, 2002; Voyksner and Lee, 1999); b) magnification of the capillary tip, showing formation of the spray.

During ESI, a solvent flow containing the analytes, originating either as the eluent from HPLC or introduced directly *via* syringe injection, enters the source *via* a thin conductive capillary. This capillary can be oriented parallel or orthogonal (as shown in Fig. 1.12a) to the inlet to the mass analyser depending on the design of the source. ESI can be operated either in positive ion mode (i.e. generating positive ions) or in negative ion mode (i.e. generating negative ions). In the first case, the capillary is held at a large positive potential and, in the latter case, at a large negative potential, typically of magnitudes ± 2-4 kV relative to a counter electrode located a few centimetres away. As solvent and encapsulated analyte molecules exit the capillary tip, they become charged by the field to a polarity identical to that of the capillary. In

the locale of the tip, a cone-shaped meniscus known as the Taylor cone is initially formed due to the surface tension of the liquid and the applied electric field (Fig. 1.12b). A flow of nebulising (sheath) nitrogen gas aids in formation of a spray consisting of small, charged droplets (Bruins, 1998). Solvent evaporation processes subsequently lead to reduction in the size of the droplets, with analytes eventually desolvated as free gaseous species. Once formed, the free ions are drawn towards a transfer capillary by both electrostatic attraction and by the pressure gradient generated by differential vacuum pumping. The ions traverse a curtain (also known as auxiliary or drying) nitrogen gas flow, used to remove any residual solvent molecules, before being directed *via* a series of ion optics to the mass analyser.

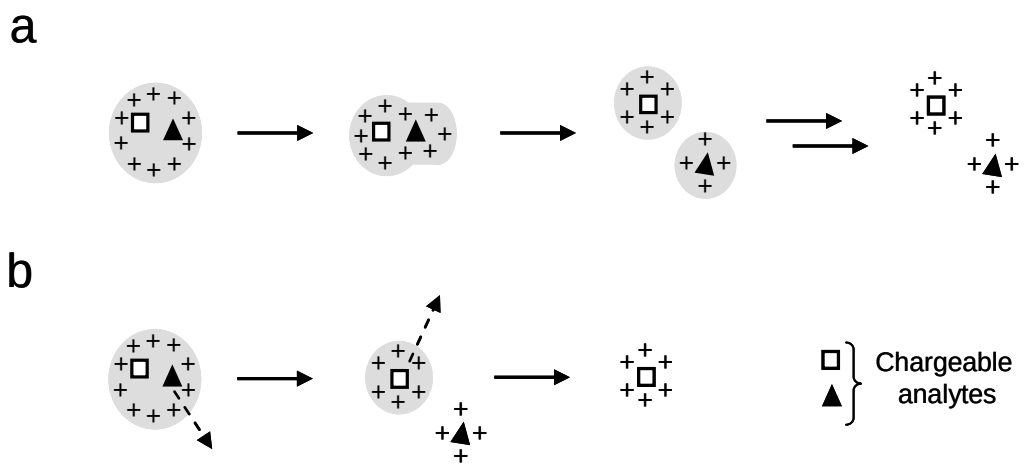


Fig. 1.13. Ionisation models for ESI: a) Charge residue model (Dole *et al.*, 1968); b) ion evaporation model (Iribarne and Thomson, 1976).

Two models have been proposed for how the droplet evaporation/desolvation processes may occur (Fig. 1.13). In the charge residue model (Dole *et al.*, 1968), the reduction in droplet size leads to an increase in internal coulombic repulsion between charges, tending towards the Rayleigh limit. At this limit, the surface tension of the droplet is insufficient to maintain its integrity and Coulombic explosion occurs, generating a series of smaller charged droplets (Fig. 1.13a). By way of iterative fissions of this nature, the droplet size diminishes until a single charged analyte remains in the droplet. Evaporation of the remaining solvent yields the free ionic species in the gas phase, bearing the maximum charge supportable by the molecule

and afforded by the droplet. In the ion evaporation model (Iribarne and Thomson, 1976), the droplet size is reduced as before. As ions are distributed towards the surface of the droplets, this leads to an increase in the surficial electric field. This increase can, in turn, facilitate direct desorption of charged analytes into the gaseous phase (Fig. 1.13b). In reality, the true mechanism may lie somewhere in between these two idealisms (Cole, 2000).

Singly charged ions formed in positive mode ESI are typically protonated molecules ($[M+H]^+$) or cationised molecules, such as sodiated ($[M+Na]^+$), potassiated ($[M+K]^+$) or ammoniated ($[M+NH_4]^+$) species. In each case, the ions which afford the charge are provided by the solvent in the eluent. Negative mode ESI leads to formation of deprotonated ($[M-H]^-$) or anionised species (e.g. formylated molecules; $[M+OCHO]^-$). Multiply charged ions can comprise a mixture of cationic or anionic species.

1.7.3.3. Atmospheric pressure chemical ionisation

In atmospheric pressure chemical ionisation (APCI), the flow of eluent containing the analytes is delivered into the source *via* a heated capillary, which can be aligned either parallel (as shown in Fig. 1.14) or orthogonal to the transfer capillary to the mass analyser in much the same fashion as discussed for the different source designs for ESI. As the solution ejects from the capillary it is nebulised, using a sheath nitrogen gas flow, into a vaporisation chamber where both solvent and analyte molecules are evaporated. On leaving the chamber, the vapour flow passes a high voltage (5-6 kV) corona discharge needle, which facilitates ionisation.

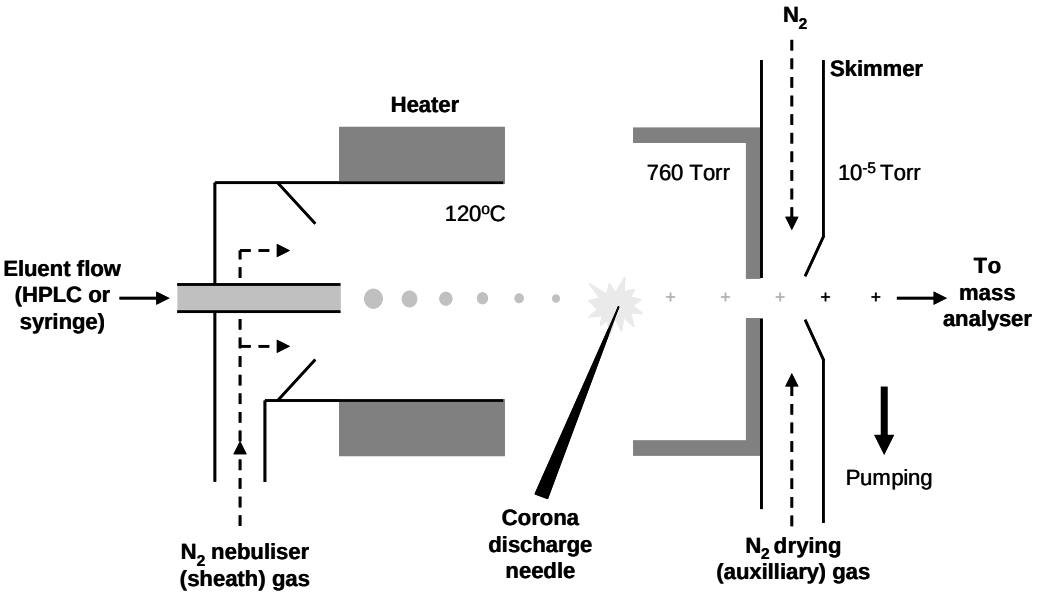


Fig. 1.14. Diagram of an APCI source (adapted from de Hoffmann and Stroobant, 2002).

When operated in positive ion mode, the corona needle discharges and ionises molecules of the nebulising N_2 gas to create primary ions (Good *et al.*, 1970; Horning *et al.*, 1973; Equations 1.12-1.13). At atmospheric pressure, the primary ions undergo numerous collisions with solvent molecules from the eluent flow and the resultant gas phase reactions generate secondary ions (shown for H_2O only; Equations 1.14-1.17). These ions, in turn, also participate in further gas phase reactions, transferring protons to other solvent molecules and to the analyte molecules, in the latter case generating $[M+H]^+$ species (Equation 1.18). The $[M+H]^+$ are then transmitted to the mass analyser as described for ESI in Section 1.7.3.2.

PRIMARY IONS:	$N_2 + e^- \rightarrow N_2^{+\bullet} + 2e^-$	(1.12)
	$N_2^{+\bullet} + 2N_2 \rightarrow N_4^{+\bullet} + N_2$	(1.13)
SECONDARY IONS:	$N_4^{+\bullet} + 2H_2O \rightarrow H_2O^{+\bullet} + 2N_2$	(1.14)
	$H_2O^{+\bullet} + H_2O \rightarrow H_3O^+ + HO^\bullet$	(1.15)
	$H_3O^+ + H_2O + N_2 \rightarrow H^+(H_2O)_2 + N_2$	(1.16)
	$H^+(H_2O)_{n-1} + H_2O + N_2 \rightarrow H^+(H_2O)_n + N_2$	(1.17)
SAMPLE IONISATION:	$M + H^+(H_2O)_n \rightarrow [M+H]^+ + nH_2O$	(1.18)

Negative ion mode APCI involves entirely different mechanisms of gas phase ion formation, including electron capture, proton abstraction and anion adduction (McEwen and Larsen, 2009). As with ESI, APCI in either mode typically leads to little fragmentation of the positive or negative species formed. On account of the higher temperatures used in APCI than in ESI, the technique is less suitable for ionisation of components with low thermal stability.

1.7.3.4. Ion trap mass spectrometry

The first mass analysers separated ions through their differing parabolic motions in a static magnetic field (see de Hoffmann and Stroobant, 2002). Much later, designs for mass analysers which selected for different ions using electrodynamic fields were patented, including the Paul trap, also known as an ion trap (Paul and Steinwedel, 1960). In its initial mode of operation, this device allowed stabilisation of ions of a single *m/z* ratio within a three-dimensional evacuated space. In the early 1980s, an alternative mode of operation for the device, know as the mass instability mode, was proposed (Stafford *et al.*, 1984). In this mode, ions of differing *m/z* ratio could be trapped simultaneously before being ejected and detected, providing a means to generate mass spectra.

The ion trap consists of three individual hyperbolic electrodes; two curved end-cap electrodes, positioned opposing one-another at a distance of $2z_0$, and a ring electrode of radius r_0, aligned on an axis perpendicular to that between the end-caps (Fig. 1.15). One of the end-cap electrodes contains an inlet aperture allowing ions to enter the trap, while the other contains several small perforations which allow ions to exit the trap and be transferred to a detector. During operation, an alternating potential at a radio frequency (RF) is applied to the ring electrode and the end-cap electrodes are grounded, creating a quadrupolar field (see March, 1997). Ions admitted into this field are subject to electrodynamic forces and undergo oscillatory motion at a frequency unique to each *m/z*, known as the secular frequency. The ion trajectories are complex due to the fluxionality of the potential well in which each ion resides but approximate to figure-of-eight (Lissajous) shapes (March, 1997). Ions which oscillate with stable trajectories are retained within the trap, whilst ions whose

trajectories become sufficiently radially destabilised (i.e. towards the ring electrode) collide with the electrode and are discharged. Ions whose trajectories become destabilised to a sufficient degree in the axial direction (i.e. towards the end-cap electrodes) can be ejected from the trap through the apertures in either the inlet or the exit electrode, with 50% of the ions ejected leaving through each. As such, mass spectra can be generated by purposeful axial destabilisation of ions, initiated by variations in the applied field, followed by their subsequent detection. Ion-ion repulsion leads to natural, as opposed to invoked, destabilisation of ion trajectories and, consequently, a constant pressure of helium (10^{-3} Torr) is incorporated into the trap as a damping gas. Through interactions with this gas, ions are collisionally "cooled" to lower kinetic energies, aiding their focusing towards the centre of the trap (Stafford *et al.*, 1984).

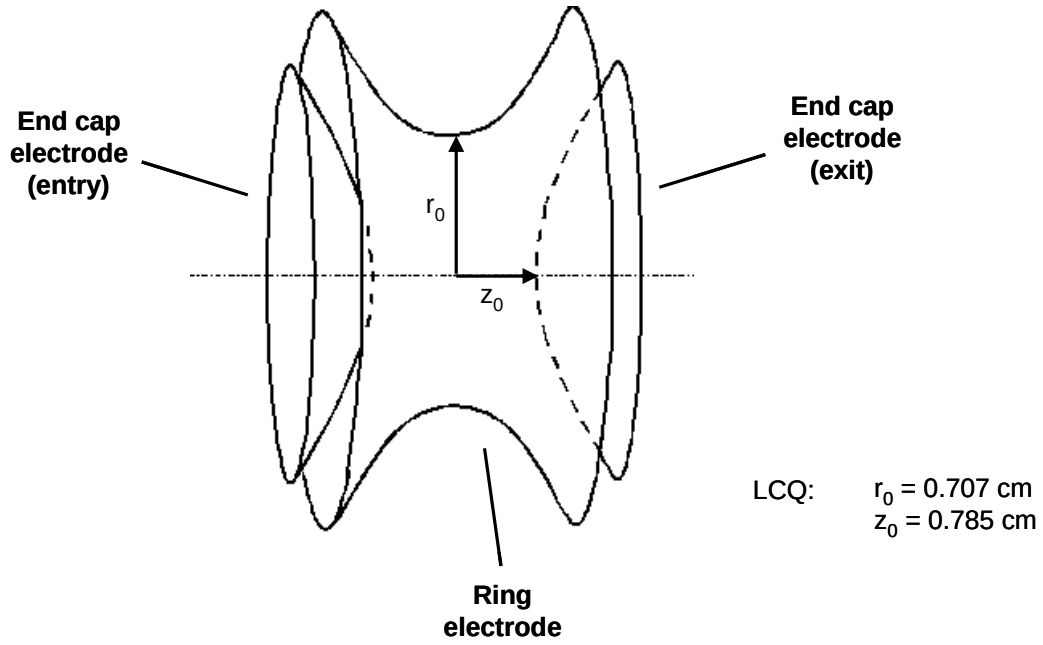


Fig. 1.15. Diagram of an ion trap analyser (apertures in end cap electrodes not shown). Dimensions for the Finnigan MAT LCQ ion trap spectrometer are also provided.

The trajectories of ions within the trap are governed by a Newtonian differential equation which can be reframed into the classic Mathieu equation (see Jonscher and Yates, 1997; March, 1997). This equation has dimensionless parametric solutions, both for the axial direction, given by a_z and q_z (Equations 1.19 and 1.20; taken from

March, 1997), and for the radial direction (similar, but not shown). Notably, a_z and q_z are functions of the amplitude of the RF potential applied to the ring electrode (V), and, more importantly, to the inverse of the *m/z* ratio of the ion in question.

$$a_z = \frac{-16zeU}{m(r_0^2 + 2z_0^2)\omega^2} \quad (1.19)$$

$$q_z = \frac{8zeV}{m(r_0^2 + 2z_0^2)\omega^2} \quad (1.20)$$

a_z, q_z	=	Dimensionless stability parameters; axial solutions to the Mathieu equation
z	=	Charge number of the ion
e	=	Elementary charge
U	=	Amplitude of the DC potential (applied to end-caps)
V	=	Amplitude of the RF potential (applied to ring electrode)
m	=	Mass of the ion
r_0	=	Internal radius of the ring electrode
z_0	=	Half of the axial distance between the end-cap electrodes
ω	=	Oscillatory frequency of the RF potential

A plot of a_z against q_z (for a fixed r_0 and z_0) generates a generic stability diagram (Fig. 1.16), in which regions of differing radial and axial stability resulting from the combination of the two axial parameters (as governed by selection of U, V and ω) can be described. Projections of the a_z and q_z values of ions of given *m/z* onto the diagram can either lie in regions representative of trajectories which are simultaneously stable in both axial and radial directions (i.e. for positive ions, the region shaded grey in Fig. 1.16) or, otherwise, which are unstable in at least one direction. It is apparent that, along the line $a_z = 0$, ions of all *m/z* values have stable radial trajectories, but also have stable axial trajectories provided the associated q_z is below a threshold value (i.e. $q_{eject} = 0.908$). It is for this reason that modern ion trap instruments operate with end-cap electrodes grounded (i.e. U = 0); provided the

initial choice for the magnitude of applied V is conservative, this facilitates the preliminary trapping of ions encompassing almost the entire *m/z* range. During operation, the magnitude of the RF potential applied to the ring electrode is gradually increased from its initial value, leading to a concomitant increase in the q_z value of each ion. Once the q_z value of a given ion exceeds q_{eject}, the trajectory of that ion is no longer axially stable (i.e. the axial or z-component of its motion exceeds the dimensions of the trap) and, consequently, the ion is ejected. Since ions of low *m/z* surpass this threshold value at lower values of applied V, ions are ejected in order of increasing *m/z* ratio during the scanning procedure.

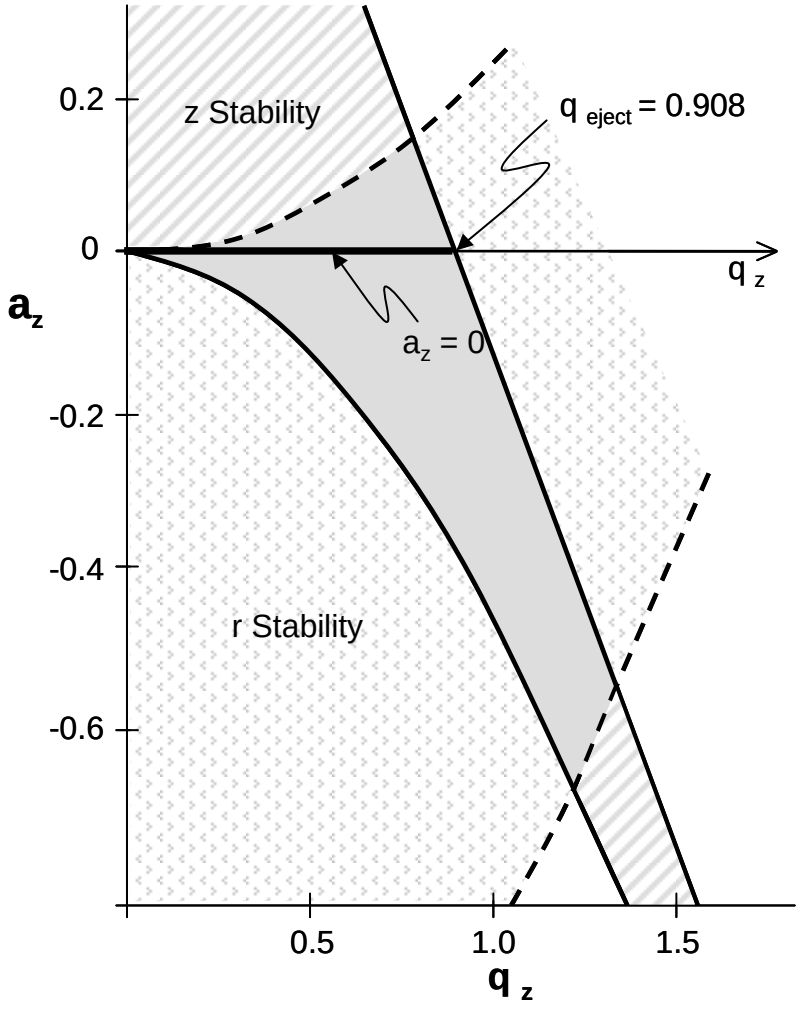


Fig. 1.16. Generic stability diagram for positive ions in a quadrupolar field (adapted from Jonscher and Yates, 1997).

One operational restriction in the axial ejection of ions solely through RF potential ramping relates to a limitation in the maximum value of this potential (V_{max}) that can be attained by typical instruments. As a result of this upper limit, ions with large *m/z* ratios can not be ejected solely by modification of the RF potential. To extend the operational range, ion trap instruments often utilise the process of resonance ejection. This is achieved by application of a supplementary alternating potential of a fixed frequency (f) to the end-cap electrodes. The axial component of the secular frequency (ω_z) of an ion of given *m/z* is proportional to ω (i.e. the frequency of the RF potential applied to the ring electrode) but also increases with increasing q_z value (Jonscher and Yates, 1997). Thus, RF ramping will result in each ion gradually being brought into resonance (i.e. $\omega_z = f$) with the applied supplementary potential. Through power absorption, the z-component of the ion trajectory is increased under resonance and, provided the amplitude of the end-cap potential is sufficiently large, eventually exceeds the dimensions of the trap. Consequently, resonance ejection allows ejection of the ions at lower values of V, including ions of large *m/z*, which can be ejected at $V < V_{max}$.

1.7.3.5. Multistage tandem mass spectrometry

As discussed in Sections 1.7.3.2 and 1.7.3.3, formation of $[M+H]^+$ by APCI and ESI leads to little associated fragmentation. While this maintains a high ion current which, in turn, aids detection of the ion, no information is provided as to the structure of the molecule. Consequently, dissociation of the ion and measurement of the dissociation products by tandem mass spectrometry (MS/MS; or MS^2) is often required to further probe structure (see de Hoffmann and Stroobant, 2002). Ion trap instruments are ideally designed for this purpose, allowing the necessary isolation, dissociation and scanning processes to occur sequentially in the same space.

During a typical MS/MS experiment, precursor ions of a single *m/z* value are selected and isolated in the ion trap *via* resonance ejection of all ions of different *m/z* (March, 2000). In an operational sense, this is effected by ejection of ions with lower *m/z* as described in Section 1.7.3.4, followed by application of a notched waveform to the end-cap electrodes at fixed V, which leads to ejection of ions of higher *m/z*. As

this waveform encompasses a full range of frequencies but excludes the frequency equal to that of the ω_z of the precursor ions, these ions are not ejected. The isolated ions are allowed to relax into the middle of the trap (typically with V set to give a q_z of 0.2-0.3) before they are subjected to collision induced dissociation (CID). In CID, an oscillatory potential with a frequency equal to the ω_z of the ions is applied to the end-cap electrodes. As the amplitude of this potential is small (explaining its frequent label as a "tickle" voltage), resonance excitation as opposed to ejection occurs. This excitation leads to an increase in the kinetic energy of the ions, with the final magnitude of energy dependent on the q_z value of the precursor ($q_{z,precursor}$), the amplitude of the resonance voltage and the period for which this voltage is applied (see Evans *et al.*, 2000). These three settings are often combined into a single parameter, the "collision energy", reported as a percentage of the maximum excitation energy that the instrument can provide. The excited ions undergo more energetic collisions with the molecules of helium gas in the trap and the excess kinetic energy is converted to internal energy, leading to dissociation of the precursor ions and yielding product ions. Where a number of different dissociations are possible, a statistical distribution of different product ions may be formed from the multiple precursor ions that occupy the trap. As the product ions have a different *m/z* ratio and, therefore, a different ω_z to that of the precursor ions, they are unaffected by the applied resonance potential. As such, they do not undergo further CID but are collisionally cooled into the centre of the trap. Subsequent ejection of the product ions as described in Section 1.7.3.4 leads to acquisition of an MS/MS spectrum. Alternatively, product ions of a particular *m/z* can be isolated as described above and used as the precursors for a subsequent round of CID and so on. This process, known as multistage tandem mass spectrometry (MS^n), can allow a detailed picture of the fragmentation pathways of organic ions to be established. Although multistage spectra up to MS^{12} have been obtained in some cases (Louris *et al.*, 1990), the diminishing ion currents inherently observed after each round of CID/ion ejection usually prevent dissociation beyond MS^5 or MS^6. As ion trap instruments are compatible with HPLC, MS^n spectra can be recorded online for components separated from mixtures (i.e. LC-MS/MS and LC-MS^n analysis is possible using such instruments).

1.7.3.6. Low mass cut-off

An inherent problem with tandem mass spectrometry when performed in an ion trap is the issue of product ion stability following CID. Although the precursor ions have stabilised trajectories at the q_z value used during the process, product ions with *m/z* below a certain threshold value, the low mass cut-off (LMCO; typically ~25% of the precursor *m/z* value), will have q_z values (at the applied V) which are above q_{eject}. Consequently, the ions at *m/z* below the LMCO are lost from the trap as soon as they form and are not detected during the MS/MS or MS^n scan. Several approaches have been developed to circumvent the problem of LMCO, including high amplitude short time excitation CID and pseudo MS^3 (Cunningham *et al.*, 2006), pulsed q dissociation (Meany *et al.*, 2007) and q_z value optimisation (Yang *et al.*, 2009). Each method is underpinned by a key principle; while the magnitude of V can be high during resonance excitation of the precursor ions, it must be lowered before the increased internal energies of these ions invoke their dissociation if product ions with *m/z* below the LMCO are to be trapped.

The HCT ion trap (Bruker) has a feature called Panoramic (PAN) MS/MS, which uses the *m/z* values of the selected precursor ion and of the smallest product ion required to be visualised to perform CID in a manner most likely to give an information rich spectrum. Depending on these values, the precursor ion can be dissociated by conventional CID, reduced LMCO CID or fast excitation dissociation (Lubeck *et al.*, 2007). In essence, the LMCO now approaches the value set for the minimum *m/z* which needs to be observed, which can be much lower than the value of the LMCO for conventional operation. As such, product ions as low as *m/z* 60 can be observed using this process, even from precursors with *m/z* > 1000.

1.7.3.7. In-source CID

As the only essential requirements for CID are an excitation voltage and an inert gas as a collisional buffer, the process is not restricted to the mass analyser and can also be invoked in the ionisation source of the mass spectrometer. More specifically, in-source CID (sCID) occurs in the interfacial head space between the APCI/ESI source

and the mass analyser (Harrison, 1999). The excitation potential is typically applied to the skimmer (shown in Fig. 1.12a and Fig. 1.14), with the N_2 curtain/drying flow serving as a collision gas. By contrast with trap-based CID, sCID does not include a precursor selection step and, as a consequence, ions of all *m/z* formed during ionisation are simultaneously excited. Those ions which are susceptible subsequently dissociate and a sCID MS spectrum can be recorded.

1.7.3.8. Other ionisation sources and mass analysers

Three other fundamental mass analyser designs were utilised in addition to ion trap instruments to obtain mass spectral data during the studies described herein; an orthogonal time of flight (oTOF) analyser, a quadrupole analyser and a Fourier transform-ion cyclotron resonance (FT-ICR) analyser. Similarly, another ionisation technique, matrix assisted laser desorption ionisation (MALDI), was also employed. Since the use of the analysers and/or ionisation source was restricted to a handful of experiments, discussion of the mode of operation of each analyser and of MALDI is beyond the scope of the thesis. Should further details be required, the principles of MALDI MS are covered in a number of literature reviews, including that of Zenobi and Knochenmuss (1998), which provides a concise introduction to the technique. For details regarding oTOF MS, an excellent review of the technique by Guilhaus *et al.* (2000) is recommended. Similarly, a review of FT-ICR MS by Marshall *et al.* (1998) is both comprehensive and accessible. Quadrupole mass analysers utilise similar physical principles to discriminate on the basis of *m/z* to those discussed for ion trap mass spectrometry in Section 1.7.3.4. More specific functional features of quadrupole analysers were adequately covered by Dawson (1986).

1.7.4. Previous instrumental approaches for identification of ether lipids

A number of analytical methods have been utilised to detect ether lipids, either as IPLs, as the remnant lipid cores (formed following purposeful degradation of the IPLs by acidic methanolysis or base hydrolysis, or by natural degradation of these components in the environment) or as further degradates. The earliest approaches focused on separation of IPL lipids using thin layer chromatography followed by

degradative analysis of each isolated component (e.g. Nishihara and Koga, 1987b). The stereochemistry of the glycerol groups could be elucidated from the optical rotation of the lipid cores (e.g. Langworthy *et al.*, 1983) or selective degradates (Kates *et al.*, 1967) or, in some cases, by ^{19}F NMR of the (*R*)- and (*S*)-Mosher ester derivatives (Dale and Mosher, 1973) of the full lipid core (Weijers *et al.*, 2006a). The alkyl chains within the structures, on the other hand, were typically characterised by GC and GC-MS as either alkyl halides, liberated from the parent structures following ether cleavage of the lipid cores with HI or BCl_3 (or alternatively, BBr_3) or as hydrocarbons following subsequent reduction of these halides with $LiAlH_4$ or Zn/AcOH (see Panganamala *et al.*, 1971; Koga and Morii, 2006). Where possible, ^{1}H and ^{13}C NMR were also employed in lipid identification (e.g. Comita *et al.*, 1984; De Rosa *et al.*, 1977b; 1977c). Later, mass spectrometric methods were developed to identify the IPLs using fast atom bombardment MS (Sprott *et al.*, 1997) or ESI MS (Murae *et al.*, 2001; 2002), forgoing the use of a separation step. LC-MS/MS methodologies are now available to separate IPLs from complex mixtures on the basis of their differing polar head groups and determine the identity of these head groups online (e.g. Sturt *et al.*, 2004).

Time of flight-secondary ion MS has been used to identify ether lipid cores by direct analysis of environmental samples (Thiel *et al.*, 2007a; 2007b), whereas MALDI MS (Macaladay *et al.*, 2004) and ESI MS (Bode *et al.*, 2008) have been used to identify the cores in lipid extracts from cellular material. Several chromatographic methods have also been developed to separate and detect ether lipid cores in their native form, including HPLC coupled with Fourier transform infrared spectroscopy (Mancuso *et al.*, 1986) or ELSD (Shimada *et al.*, 2002; 2008), supercritical fluid chromatography (DeLuca *et al.*, 1986; Hedrick *et al.*, 1991) or high temperature GC (Nichols *et al.*, 1993; Weijers *et al.*, 2006a; Pancost *et al.*, 2008). Other methods have utilised a preliminary chemical derivatisation procedure which furnishes the lipid cores with an active fluorophore or chromophore. In this manner, the 9-anthroyl (Bai and Zelles, 1997) or dinitrobenzoyl (Demizu *et al.*, 1992) derivatives of ether lipid cores can be detected by HPLC coupled to a fluorescence or UV/Vis detector, respectively. Recently, GC-MS of diether lipid cores as their silylated derivatives has been the preferred methodological choice for their detection and characterisation (e.g. Pancost

et al., 2001a; Stadnitskaia *et al.*, 2003; Blumenberg *et al.*, 2007), although LC-MS has also been used to detect the unmodified structures in some studies (Turich *et al.*, 2007; Reigstad *et al.*, 2008; Weijers *et al.*, 2009). Tetraether lipid cores, on the other hand, are now almost exclusively detected *via* normal-phase LC-MS, which allows structures containing differing numbers of Cp or Ch rings to be separated, ionised *via* APCI and detected as the $[M+H]^+$ species (Hopmans *et al.*, 2000; Schouten *et al.*, 2007b; Escala *et al.*, 2007). Whilst LC-MS analysis performed in this manner is sufficient to identify known tetraether lipid cores, NMR and/or degradative analyses are still required to ascertain structural information for novel tetraether core components (e.g. Damsté *et al.*, 2002b).

Quantitation of the relative abundances of tetraether lipids, a necessary requirement for calculation of TEX_{86}, BIT, MBT and CBT indices, is typically achieved *via* integration of LC-MS peak areas in ion chromatograms for individual lipid components (Schouten *et al.*, 2002; Hopmans *et al.*, 2004; Weijers *et al.*, 2007c). Absolute quantitation, on the other hand, can be achieved either *via* use of external standard curves (e.g. Weijers *et al.*, 2007c) or use of a synthetic C_{46} GTGT lipid (Patwardhan and Thompson, 1999) as an internal standard (Huguet *et al.*, 2006b), with the latter technique affording the greater degree of accuracy (Huguet *et al.*, 2006b). As both the structures and molecular masses of the different isoprenoid GDGT lipid cores are, in general, relatively similar, the differences in their mass spectrometric ionisation efficiencies are assumed to be small, or at the very least, are largely resilient to differences in experimental conditions (Schouten *et al.*, 2009). The relative ionisation efficiencies of isoprenoid compared to non-isoprenoid lipid cores, on the other hand, appear to be much more susceptible to interlaboratory variation, probably reflecting the much larger difference in masses between the two compound classes (Escala *et al.*, 2009; Schouten *et al.*, 2009).

1.8. Summary and Aims

Archaea are cosmopolitan organisms in the contemporary biosphere, living on land and in the ocean, over temperatures ranging from subzero through to the hottest temperatures known to support life. The cores of the membrane lipids of these organisms are thermally and chemically robust on account of ether linkages within the structures. They can be preserved over geological timescales, providing a record of past productivity of the Archaea. Due to physicochemical controls on the lipids formed in some environments, archaeal ether lipids can also be employed in palaeotemperature reconstructions. A whole host of ether lipid cores have been identified previously and serve as useful biomarkers for the archaeal domain. In general, however, most of the cores are common to several phylogenetically distinct organisms within the domain and, consequently, they are of little use as marker compounds for specific archaeal taxa. Ether lipid cores consistent with a eubacterial origin have also been identified although, apart from a few select cases, little is known about the phylogeny of the progenitor species for these lipids. Clearly, the identification of new biomarkers which may offer greater specificity would provide a distinct advantage in both determination of ecology in modern environments and in palaeoenvironmental reconstructions. The necessity for degradative analysis of novel structures proves a bottleneck in the identification and assignment of ether lipids.

The main aim of the work presented in this thesis was to extend current LC-MS methodology for the analysis of ether lipid cores to provide online acquisition of MS^n spectra and assess the use of this technique for rapid, yet detailed, profiling of complex mixtures of such components. The work described focuses solely on lipid core structural analysis, predominantly describing tetraether lipid cores of the Archaea. Chapter 2 focuses on aspects of LC-MS/MS method development using lipid extracts from a model archaeon. Chapters 3 and 4 provide an assessment of the scope of the technique for profiling archaeal membrane lipids from cultured cell material, including the validation of the method against previous reports for several organisms. Novel structures and their potential value as biomarkers are also considered in these chapters. To conclude, Chapter 5 investigates the use of the

methodology for the identification of both known and novel structures in highly complex lipid extracts from environmental soils and sediments.

Chapter 2

LC-MS/MS methodology for structural elucidation of tetraether lipid cores

Publications resulting from work discussed in this chapter:

Knappy, C. S.; Chong, J. P. J.; Keely, B. J.; 2009, Rapid discrimination of archaeal tetraether lipid cores by liquid chromatography-tandem mass spectrometry, *J. Am. Soc. Mass Spectrom.*, **20**, 1, 51-59.

2.1. Introduction

2.1.1. Limitations in previous methods for detection and structural elucidation of tetraether lipid cores

From the first identification of archaeol, GDD **1**, in the halophilic archaeaon *H. cutirubrum* (Sehgal *et al.*, 1962; Kates *et al.*, 1965), to the recent identification of GMGTs **27**-**31** (Morii *et al.*, 1998; Sugai *et al.*, 2004; Schouten *et al.*, 2008b), ether lipid cores have proven to be challenging molecules to assign structurally. They have relatively large masses, are significantly polar, contain a myriad of stereocentres and often exhibit unusual structural motifs, such as isoprenoid chains containing cyclopentyl rings (De Rosa and Gambacorta, 1988a), all of which test conventional analytical methodologies. A number of analytical techniques have been employed to date in order to identify and characterise tetraether lipid structures in both an intact form, with attached polar head groups, and as core structures, in which the head groups have been removed (see Section 1.7.4). Each technique used to accomplish this, however, has its own caveats: for example, 2D TLC-NMR based approaches each have a sizeable limit of detection (LOD), restricting use to samples in which lipid concentrations are reasonably high. Chemical degradation procedures for examination of the hydrocarbon chains by GC and GC-MS analysis (De Rosa and Gambacorta, 1988a; Koga and Morii, 2006), are cumbersome and, consequently, difficult to apply to complex mixtures of lipids. In addition, degradative approaches destroy information regarding the particular combination of the chains within the full lipid structure. A significant advantage of use of APCI LC-MS for analysis of ether lipid cores is that the distinct cores of a sample can be separated from one another and analysed in an intact form, even if present at very low levels (typical LOD ~ 0.4 ng on column; Schouten *et al.*, 2007b). Furthermore, the only preparative step required before analysis is chromatographic isolation of the polar lipid fraction on alumina. Coupled to an analytical throughput faster than one sample per hour, LC-MS leads to significantly greater sample turnover than can be achieved with other methods. One limitation of the current LC-MS methodology arises from use of only a single stage of mass spectrometric analysis to identify the tetraether lipids. Thus, the lipids are assigned on the basis of their retention times, the *m/z* of the protonated

molecule and observable fragment ions formed from losses of 18 ($-H_2O$) and 74 ($-C_3H_6O_2$) *m/z* units in the mass spectra (Hopmans *et al.*, 2000). These losses both arise from the glycerol moiety, a common biosynthetic precursor that is not exclusive to tetraether lipids. Futhermore, the neutral fragments lost do not include any of the core hydrocarbon architecture, making it impossible to differentiate between isobaric GDGT lipids that have different constituent biphytanyl partners. The potential for discovery of new tetraether lipids in archaea, particularly in novel strains, is significant. Consequently, a rapid method that is able to identify tetraether core lipids at low concentration in archaeal cellular material or environmental matrices and simultaneously provide a means to discriminate them on the basis of their structure would prove to be highly beneficial.

2.1.2. Aims

The aims of the work presented in this chapter were to extend current LC-MS methodology for tetraether core lipid detection to include online aquisition of tandem mass spectra and to investigate the use of this approach in the structural characterisation of lipid cores during chromatographic separation.

2.2. Results and Discussion

2.2.1. Tetraether lipid core samples

In order to provide tetraether core lipid material for analysis, aerobically harvested cells from a pure culture of *Methanothermobacter thermautotrophicus* strain ΔH (previously known as *Methanobacterium thermoautotrophicum*; Zeikus and Wolfe, 1972; Wasserfallen *et al.*, 2000), grown at 45^{o}C or 70^{o}C, were refluxed in acidic methanol to extract polar lipids and simultaneously cleave the phosphatidyl and glycosyl head groups of the component structures (Hopmans *et al.*, 2000). *M. thermautotrophicus* ΔH (MTH(ΔH)) was chosen as a model organism for this study on account of the extensive body of previous literature which describes the core lipids of the organism (Grant *et al.*, 1985; Nishihara *et al.*, 1987a; 1987b; 1989), all of which suggest that it expresses GDGT **2a** as the dominant tetraether lipid core. The simplicity of the MTH(ΔH) lipid core profile makes the organism ideal as a candidate for method development. Details regarding the organism, and a more extensive discussion regarding its lipid cores, are provided in Chapter 3.

2.2.2. Ion trap mass spectrometry

Initial studies focused on analysis of the lipid core material by mass spectrometry, tandem mass spectrometry and liquid chromatography-tandem mass spectrometry using a Finnigan MAT LCQ ion trap mass spectrometer (henceforth denoted LCQ). Following acquisition of a higher energy, higher capacity Bruker HCTultra ETD II ion trap instrument (henceforth denoted HCT), subsequent analyses were also performed using this second mass spectrometer. As the MS and MS/MS spectra recorded on the two instruments show significant differences, both are reported for the MTH(ΔH) extract, described in detail in this chapter for comparison. In subsequent chapters, it has often been sufficient to describe the spectra as recorded solely on a particular mass spectrometer.

2.2.3. Direct infusion MS (LCQ and HCT)

2.2.3.1. MS spectra

Direct infusion of a portion of the 45°C MTH(ΔH) extract into the LCQ, operated in positive ion mode using an APCI source, generated a mass spectrum dominated by a peak at *m/z* 1302.2 (Fig. 2.1a). Similarly, the HCT, operated in an analogous fashion, generated a mass spectrum dominated by a peak at *m/z* 1302.5 (Fig. 2.1b). An ion with approximately the same *m/z* has been observed previously during LC-MS studies of archaeal lipid extracts (Hopmans *et al.*, 2000; Schouten *et al.*, 2000), where it was assigned as the protonated molecule of **2a**. As both ion trap instruments have modest optimal mass accuracies (LCQ ~0.5 *m/z* units; HCT ~0.15 *m/z* units), the base peak ion observed in each case is consistent with that expected for this lipid ($C_{86}H_{173}O_6^+$; theoretical *m/z* 1302.3). Furthermore, the known dominance of GDGT **2a** in MTH(ΔH) allows the base peak ion observed in the mass spectrum of the lipid extract to be assigned as the $[M+H]^+$ of **2a**. Notably, the presence or absence of isobaric lipid **2b** cannot be confirmed from the mass spectrum. No ions consistent with the sodiated, ammoniated or potassiated species were observed in either spectrum, confirming, as noted previously, that tetraether lipid cores ionise exclusively as the $[M+H]^+$ species in APCI (Hopmans *et al.*, 2000). Ions at *m/z* 1284.2 (LCQ) and 1284.4 (HCT) were also observed in the spectra, as were ions with very low signal strengths at *m/z* 1228.2 (LCQ) and 1228.4 (HCT). The ions correspond to fragment ions formed during ionisation *via* losses of a molecule of water and $C_3H_6O_2$ from the $[M+H]^+$ of **2a**. The two fragment ions have been observed previously in the mass spectrum of **2a**, during LC-MS analysis using a single quadrupole mass spectrometer (Hopmans *et al.*, 2000; Schouten *et al.*, 2000), where they were denoted as characteristic ions for tetraether core structures. Both fragment ions are less prominent in spectra recorded using LCQ and HCT mass spectrometers and, consequently, their use as characteristic ions may be restricted to mass spectra of concentrated samples. Two additional ions, found at *m/z* 1316.1 and *m/z* 653.4 on the LCQ and at *m/z* 1316.5 and *m/z* 653.7 on the HCT, were also detected. These ions relate to the $[M+H]^+$ of homocaldarchaeol (**26**) and archaeol (**1**), respectively, and will be discussed in more detail in Section 2.2.4.2.5 and Chapter 3.

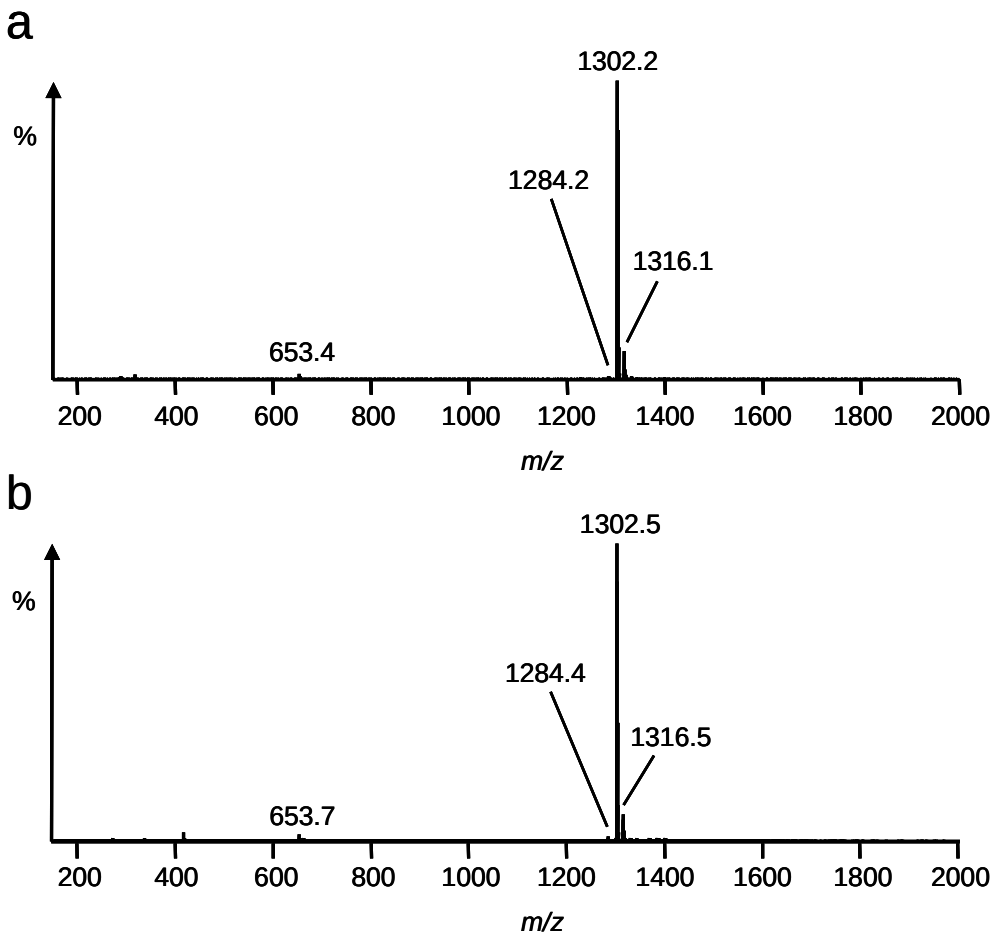


Fig. 2.1. Direct infusion mass spectra for the lipid extract from MTH(ΔH) grown at 45^{o}C, as recorded on: a) LCQ; b) HCT instruments.

2.2.3.2. Nominal *m/z* values

On account of the large mass sufficiency in tetraether lipids, the *m/z* values for the $[M+H]^+$ of **2**, as measured on both spectrometers, are significantly larger than the nominal value of a species with molecular formula $C_{86}H_{173}O_6$ (1301 *m/z* units). This explains why the base peak ion, which contains zero nitrogen atoms, is measured at an approximately even *m/z* value and so does not appear to adhere to the formal nitrogen rule for even-electron ions (McLafferty and Turečček, 1993). Previous reports of the MS and LC-MS analyses of tetraether lipids have often quoted the measured *m/z* values of the tetraether lipids to their nearest integer value (Hopmans *et al.*, 2000; Schouten *et al.*, 2000; Huguet *et al.*, 2006b; Escala *et al.*, 2007; Schouten *et al.*, 2007b). In order to avoid the possibility of confusion, particularly in

more complex samples, it is important to state whether the *m/z* value quoted for the $[M+H]^+$ of a tetraether lipid represents the value *measured* on the instrument used in its detection or the *nominal* value for the species that the ion represents.

2.2.3.3. Use of MTH(ΔH) as a tuning standard

The dominance of **2** in the MTH(ΔH) 45°C and 70°C extracts allows the samples to be used directly for instrument tuning, avoiding the need for preparation of a pure tetraether standard. Thus, use of either extract allowed the vaporisation temperature, nitrogen gas flows and corona discharge parameters to be optimised for tetraether analysis on both mass spectrometers (see Section 7.5.1). More importantly, by tuning the chosen spectrometer to give optimal signal for the ion measured at approximate *m/z* 1302 (nominal *m/z* 1301), this approach allows the ion optics of the instrument to be optimised for tetraether lipid core detection prior to LC-MS/MS analyses.

2.2.4. LC-MS/MS (LCQ)

Following the success of the direct infusion MS studies, the 70°C MTH(ΔH) extract was analysed by LC-MS/MS on the LCQ using a narrow mass spectral scan range (*m/z* 1280-1320). Lipids were separated using method A, derived from Hopmans *et. al.* (2000), which uses a normal phase amino column chemistry and binary mobile phase comprising hexane and isopropanol (IPA). The initial isocratic portion of the gradient program (1% IPA composition; 5 min) was retained, but limitations in the degree of control in solvent proportioning afforded by the liquid chromatograph employed, which could only proportion solvents to integer values, forced the final composition to be set to 2% IPA (see Section 7.6.1 for details). The total analysis time was extended from the 50 min method used by Hopmans *et al*. (2000) to 61.25 min in order to ensure that an equivalent gradient program up to 1.8% IPA was maintained.

2.2.4.1. LC-MS

2.2.4.1.1. Base peak chromatograms

The MS base peak chromatogram of the extract of MTH(ΔH) contains four peaks, with retention times (t_R) of 7, 16, 23 and 35 min and having base peak ions at, in turn, *m/z* 1316.0, 1303.9, 1302.0 and 1299.9 (average values over 12 analyses). Reconstructed mass chromatograms for the four distinct base peak ions (Fig. 2.2a) revealed an additional species with $[M+H]^+$ at *m/z* 1316.0 (t_R = 22 min) that had been concealed by partial co-elution with that having $[M+H]^+$ at *m/z* 1302.0 in the base peak chromatogram. The latter corresponds to the protonated molecule of GDGT **2a**, which has been observed previously by LC-MS with a similar retention time (Hopmans *et al.*, 2000). Likewise, an early eluting peak with $[M+H]^+$ at *m/z* 1303.9, corresponding to the protonated molecule of GTGT **3**, has been observed previously under similar conditions (Hopmans *et al.*, 2000). LC-MS detection of **2b** has not been reported, nor has the occurrence of this lipid been confirmed in the ΔH strain of *Methanothermobacter thermautotrophicus,* although it has been identified in two other strains of *M. thermautotrophicus* (strains Hveragerdi and Marburg[†]; Gräther and Arigoni, 1995). *M. marburgensis* shares a close phylogenetic relationship to MTH(ΔH), with the two type strains of each species being approximately 98% similar on the basis of their 16S rRNA sequences (Wasserfallen *et al.*, 2000). Given that **2b** is expressed by close genetic relatives to MTH(ΔH) and that the peak in the *m/z* 1301.5-1302.5 mass chromatogram for the extract from this organism shows the greatest asymmetry of all of the peaks identified (asymmetry factor = 2.9, calculated as the ratio of the peak median to peak tail and peak median to peak front distances), it is highly likely that **2b** is synthesised by the organism and that it co-elutes with **2a** under the chromatographic conditions employed. The component with an $[M+H]^+$ at *m/z* 1299.9 exhibits a very different retention time to that of an isobaric GDGT containing one cyclopentane ring in its core structure (i.e. GDGT **14**, observed to elute approximately 1 min later than **2a** in many LC-MS studies including Hopmans *et al.*, 2000 and those described in Chapters 3 and 5). GMGT **27** has previously been

[†] Both more recently reclassified as strains of the species *Methanothermobacter marburgensis* (Wasserfallen *et al.*, 2000; Ding *et al.*, 2010)

identified as a minor component in the TLC analysis of the tetraether lipids in this archaeon (Morii *et al.*, 1998) and was observed to elute late in LC-MS analyses of the tetraether lipids of a thermoacidophile (Schouten *et al.*, 2008b) and of lipids extracted from marine and lacustrine sediments (Schouten *et al.*, 2008a). Accordingly, the latest eluting component in the base peak chromatogram is assigned as GMGT **27**. The two lipids observed to have protonated molecules at *m/z* 1316.0 (**51**, t_R = 7 min and **26**, t_R = 22 min) have not, it seems, been previously identified *via* LC-MS analysis.

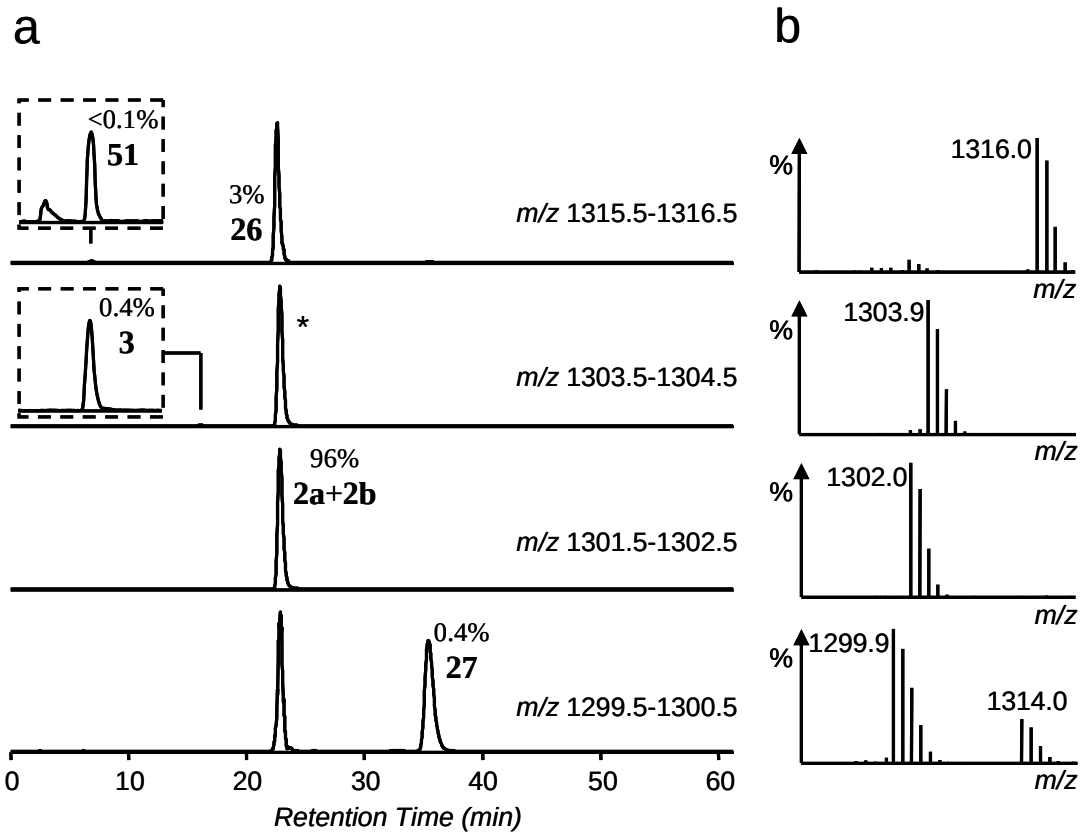


Fig. 2.2. a) Reconstructed mass chromatograms for the four distinct base peak ions identified in the LC-MS base peak chromatogram of the MTH 70°C extract. Insets show magnifications of small peaks. The peak labelled * corresponds to the $[M+H+2]^+$ higher isotopomers of GDGTs **2a** and **2b**; b) Representative mass spectra for lipid **2**, **3**, **26** and **27**. Lipid **51** gave a mass spectrum identical to that of **26**.

2.2.4.1.2. MS spectra

Representative mass spectra for tetraether lipids **2**, **3**, **26** and **27** (Fig. 2.2b) show that these species all exhibit base peak ions corresponding to the $[M+H]^+$, as noted during direct infusion MS of the MTH(ΔH) 45^oC extract (Section 2.2.3). The *m/z* values quoted on the spectra in Fig 2.2b are averages recorded over 12 LC-MS/MS analyses, during which the $[M+H]^+$ of the species were measured at *m/z* values ranging from 0.8-1.2 units greater than the nominal *m/z* value in each case. The isotopic profiles in the MS spectra of each lipid were reproducible throughout the analyses, although they were determined to be slightly different, on average, to the theoretical isotope pattern predicted for each species. For example, the $[M+H+1]^+$ and $[M+H+2]^+$ isotopic peaks of **2**, which are predicted to occur in 95% and 46% abundance relative to the $[M+H]^+$ peak, respectively occur in 83±3% and 35±1% (n = 12) relative abundance in the mass spectra of **2**. A similar discrepancy between the theoretical (46%) and measured (33%) abundance of the $[M+H+2]^+$ isotopic peak, relative to the $[M+H]^+$ peak, was also noted for crenarchaeol (**24**) during LC-MS analyses of lipids extracted from marine sediments (Weijers *et al.*, 2004). Despite the discrepancy, the robustness of the isotope peak ratios in the replicate analyses allowed determination of the relative abundances of the five tetraether lipids in the extract by integration of peak areas in the relevant reconstructed mass chromatogram. Given the structural similarity of the ionisable glycerol termini in each lipid core, identical ionisation efficiencies were assumed for each component. The lipid extract is dominated by caldarchaeol isomers **2a** and **2b**, which contribute 96% of the total tetraether component. Lipid **26** contributes approximately 3% of the total extract and lipids **3**, **27** and **51** contribute only minor amounts to the total lipid mass (<1% combined). The mass spectrum obtained over the elution time of lipid **27** highlights the presence of a co-eluting species that has a protonated molecule at *m/z* 1314.0. The isotopic profile of this species matches that observed for lipids **2-3**, **26-27** and **51**, consistent with a tetraether structure. As this compound was subordinate to co-eluting lipid **27**, an MS/MS spectrum was not generated under the operational scan range chosen. Consequently, discussion of the likely structure of the species is deferred to Chapter 3.

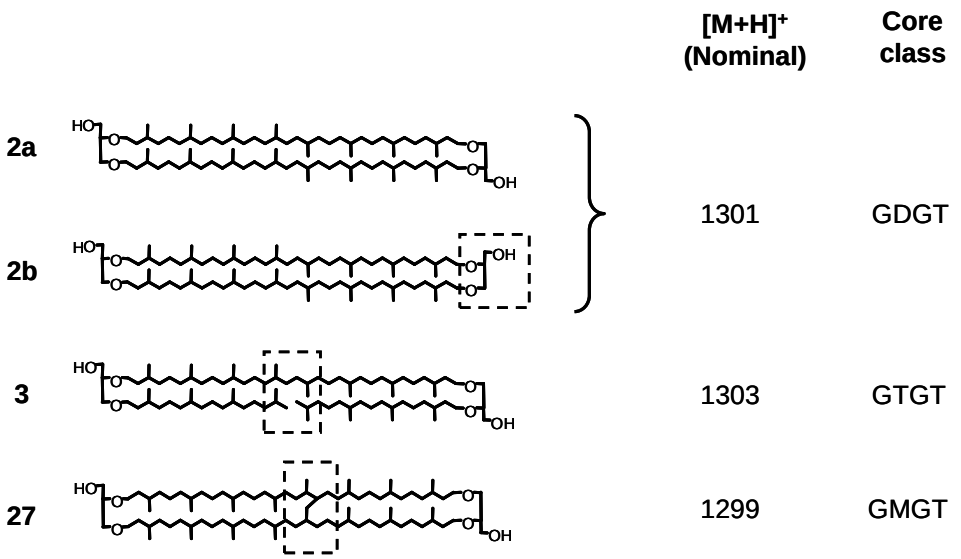


Fig. 2.3. The three classes of tetraether core identified in MTH(ΔH), with the nominal *m/z* value of the protonated molecule in each case; **2a** = caldarchaeol, **2b** = isocaldarchaeol, **3** = acyclic caldarchaeol, **27** = H-shaped caldarchaeol (structure from Morii *et al.*, 1998; Schouten *et al.*, 2008a; 2008b). Dashed boxes enclose regions that vary from structure **2a**.

2.2.4.1.3. Retention time drift

Although normal phase LC-MS/MS provides a high resolution method for separating tetraether lipids, the retention times are generally not reproducible. Drifts in retention time of up to several minutes were observed, without consistency in direction, in all 12 repeat analyses, most noticeably between analyses recorded on different days. The large difference in the polarities of the two components of the mobile phase has caused similar problems in previous attempts to purify dinosterol (Smittenberg and Sachs, 2007), suggesting that this problem may be inherent to liquid chromatography-based lipid analyses that use hexane/isopropanol mixes as mobile phase. Several unsuccessful attempts were made to reduce the size of the retention time drifts, including fixing the column temperature (30^{o}C) throughout the analyses, changing from amino to cyano column chemistry, incorporating identical column conditioning (exactly 15 min) and backflushing (exactly 10 min) periods during successive analyses, flushing the column with dichloromethane (DCM) between analyses and substituting DCM for part or all of the IPA in the mobile phase.

Since absolute retention times were not robust during the LC-MS/MS analyses, cholesterol, chosen largely on account of its ready availability and solubility in hexane/IPA, was introduced to the MTH(ΔH) 70^{o}C extract as an internal standard in the hope that it would act as a reproducible retention time marker. Unfortunately, however, the retention times of tetraether components **2-3**, **26-27** and **51** relative to that of cholesterol varied between successive LC-MS/MS analyses, indicating that this approach was also inadequate.

Of the tetraether lipid cores, GDGT **2a** is undoubtedly the most widely occurring, being expressed by the majority of archaeal species known to biosynthesise tetraether lipids (De Rosa and Gambacorta, 1988a; Schouten *et al.*, 2007c; Table 1.2) and found in almost all aquatic sediments (e.g. Hopmans *et al.*, 2000; Schouten *et al.*, 2000; Pearson *et al.*, 2004; Powers *et al.*, 2004), soils (Gattinger *et al.*, 2003; Leininger *et al.*, 2006; Weijers *et al.*, 2006b) and peats (Weijers *et al.*, 2004; 2006a) profiled for GDGT lipids to date. The frequency with which **2a** appears in a variety of matrices indicates that it should serve as a suitable retention time marker for tetraether lipids. Consequently, novel tetraether components reported in the subsequent chapters are assigned on the basis of their elution order and approximate retention time relative to **2a**.

2.2.4.2. Tandem MS Analysis

2.2.4.2.1. Generalities

Online MS/MS spectra were generated following CID of the most abundant ion observed in each mass spectral scan. Due to the presence of some common structural features in lipids **2-3** and **27** (Fig. 2.3), the product ions generated in the ion trap from CID of the protonated molecules show distinct similarities, albeit reflecting the 2 Da differences in the precursor protonated molecules. Most notable is the presence of three distinct regions in the MS/MS spectra of lipids **2-3** and **27** (Figure 2.4). Region 1 (*m/z* > 1150) results from loss of small neutral molecules (assigned as formal losses of water and C_3H_4O), Region 2 (*m/z* 900-1025) from formal loss of one phytanyl group with accompanying loss of small neutral molecules and Region 3

(*m/z* 550-850) from formal losses of either one biphytanyl group or two phytanyl groups together with small neutral molecules. The number of these regions in which product ions are observed in the MS/MS spectrum of the tetraether lipid analysed is indicative of the specific core structural type. The average *m/z* values measured for each product ion over 12 analyses are also shown in Fig. 2.4. Owing to the large mass defect in the product ions, these values can be up to 1.4 *m/z* units greater than the nominal *m/z* values for the ions. In the interests of clarity, only ions which could reliably be assigned to formal molecular losses of water, C_3H_4O, 1-phytene or 1,31-biphytadiene are reported for the LCQ spectra. Although all product ions reported were observed consistently in the repeat analyses, the relative abundances of the ions within a given MS/MS spectrum varied significantly between analyses. Accordingly, interpretation of the favourability of a particular neutral loss over another during CID from the relative abundances of the two product ions should not be attempted unless the ratio of the signals of the ions is either very large or highly reproducible in repeat analyses. The MS/MS spectra shown in Fig. 2.4 are representative examples of typical product ion distributions. The CID spectra of lipids **2-3** and **27** are discussed individually and summarised (Table 2.1).

*2.2.4.2.2. MS/MS of GDGT **2***

On account of the co-elution of the isobaric GDGTs **2a** and **2b**, the $[M+H]^+$ of each is indiscriminantly trapped and subjected to CID, unavoidably producing overlap of the product ions from both species in the resultant MS/MS spectrum (Fig. 2.4a). In this initial discussion of the dissociation of protonated caldarchaeol, it is assumed that the $[M+H]^+$ of **2a** and **2b** undergo identical dissociation; in the interests of clarity, the two lipids are accordingly discussed together as GDGT **2**. The MS/MS spectrum of **2** shows two distinct groups of product ions: Regions 1 and 3. The former includes a number of product ions resulting from small neutral losses from one of the terminal glycerol moieties in **2**. As has been reported previously for fragment ions of the $[M+H]^+$ of GDGT **2** (Hopmans *et al.*, 2000), the product ions at *m/z* 1283 (-18; H_2O) and *m/z* 1227 (-74; $C_3H_6O_2$) result from loss from/of this terminal functionality.

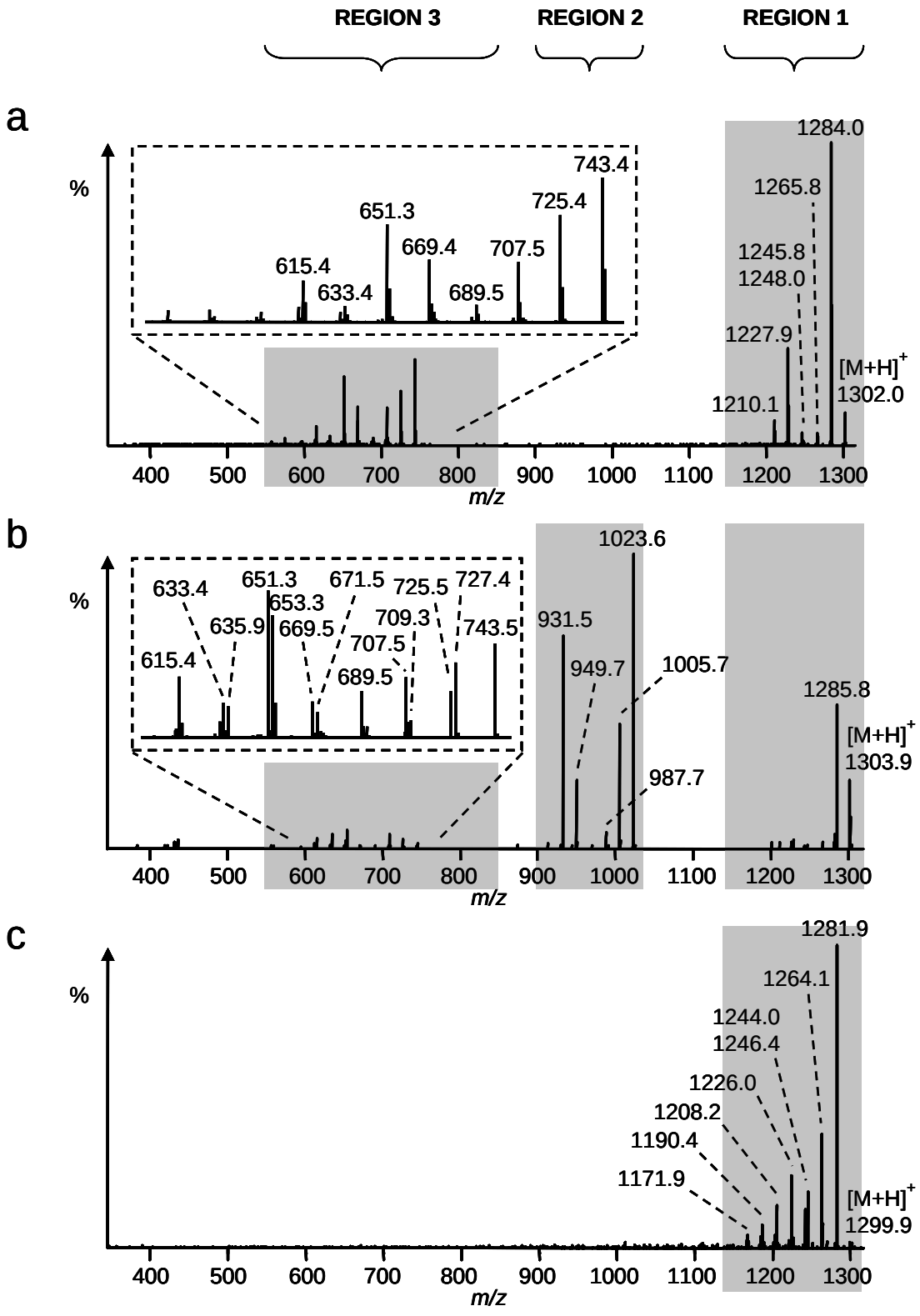


Fig. 2.4. MS/MS spectra (LCQ) for lipids **2-3** and **27** in MTH(ΔH). Insets show expansions of Region 3 of the spectra where appropriate. a) GDGT **2**; b) GTGT **3**; c) GMGT **27**. Product ions are labelled with the average *m/z* value measured (n = 12).

A number of other product ions were observed, arising from additional formal losses of one or two further molecules of water from the $[M+H]^+$ of **2**. Interestingly, formation of a product ion at *m/z* 1245 is attributed to loss of C_3H_4O from one end of the protonated caldarchaeol molecule. The existence of this product ion suggests that the ion at *m/z* 1227 in the MS/MS spectrum of **2** may be formed in two stages, with separate losses of water and C_3H_4O, as opposed to being formed in one stage *via* the loss of a $C_3H_6O_2$ unit (see Section 2.2.6 for more detailed discussion).

Region 3 of the MS/MS spectrum shows product ions arising from loss of one of the two biphytanyl hydrocarbon chains in **2**. The product ion at *m/z* 743 is formed *via* loss of 1,31-biphytadiene from the $[M+H]^+$ of **2**, with double hydrogen transfer back to the charge-retaining species. Further product ions in this region also form as a result of loss of biphytadiene, but show concomitant losses of 1-4 molecules of water and/or loss of C_3H_4O.

*2.2.4.2.3. MS/MS of GTGT **3***

The MS/MS spectrum of GTGT **3** (Fig. 2.4b) shows similarities to that of **2**. In this case, however, product ions are observed in Regions 1, 2 and 3. Several ions in Regions 1 and 2 correspond to fragment ions observed in single-stage mass spectrometry of **3** (Hopmans *et al.*, 2000).

Region 1 of the MS/MS spectrum is dominated by a product ion at *m/z* 1285, corresponding to a species in which a molecule of water has been lost from the $[M+H]^+$ of **3**. Region 2 of the spectrum reveals product ions in which one of the two constituent phytanyl chains of **3** has been lost during CID. The peak at *m/z* 1023 represents a product ion formed by loss of 1-phytene ($C_{20}H_{40}$) with proton transfer to the charge-retaining species. Other product ions in this region have also lost phytene, but have suffered additional losses of C_3H_4O and/or 1-2 molecules of water. Region 3 of the MS/MS spectrum shows a series of product ions in which the second phytanyl chain has also been lost from the $[M+H]^+$ of **3**, again *via* elimination of $C_{20}H_{40}$ with proton transfer back to the charge-retaining species. Loss of this second alkyl chain generates product ions that are structurally identical to those produced

when a biphytanyl chain is lost during CID of the $[M+H]^+$ of **2** and, as such, this spectral region exhibits essentially the same ions as are observed in the equivalent region of the MS/MS spectrum of **2**. Notably, however, a second related series of product ions at higher *m/z* (+2 units) is also apparent in this region. These ions are consistent with loss of a molecule of biphytadiene (*cf.* the mechanism of loss of biphytadiene from the $[M+H]^+$ of **2**), with or without additional loss of C_3H_4O or water, though such losses would be expected to show a concomitant loss of a phytanyl group. The precise mechanisms for formation of the product ions within this series are unclear, but could be explained by a common rearrangement in which a phytanyl group has been transposed to elsewhere on the architecture prior to loss of biphytadiene (see Section 2.2.11.7).

*2.2.4.2.4. MS/MS of GMGT **27***

The MS/MS spectrum of GMGT **27** (Fig. 2.4c) reveals product ions that are formed *via* the same small molecule losses that generated the product ions seen in Region 1 of the MS/MS spectra of tetraethers **2** and **3**. The overall degree of dissociation for **27** is far less pronounced than for the other tetraethers, with product ions being present only in spectral Region 1 on account of the covalent link between the two hydrocarbon chains in this molecule. Provided that this C-C bond remains intact during dissociation, any product ion generated will maintain a C_{80} hydrocarbon core, resulting only in loss of small neutral molecules. This tethering of the biphytanyl chains, however, allows more extensive loss of functionality from both glycerol groups and consequently, product ions relating to losses of C_3H_4O and/or 1-4 molecules of water are observed in Region 1 of the MS/MS spectrum of **27**.

Table 2.1. Product ions generated in each of the three distinct regions of the MS/MS spectra following CID of the $[M+H]^+$ ions of tetraether lipids **2-3** and **27** on the LCQ spectrometer.

	Mass Loss (Da)[a]	Number of neutral fragments lost $[H_2O, C_3H_4O, C_{20}H_{40}, C_{40}H_{78}]$[b]	*m/z* of corresponding product ion generated from $[M+H]^+$ of [c]		
			2	**3**	**27**
$[M+H]^+$	-0	[0,0,0,0]	1301	1303	1299
REGION 1	-18	[1,0,0,0]	**1283**	1285	**1281**
	-36	[2,0,0,0]	1265	1267	1263
	-54	[3,0,0,0]	1247		1245
	-56	[0,1,0,0]	1245		1243
	-74	[1,1,0,0]	1227		1225
	-92	[2,1,0,0]	1209		1207
	-110	[3,1,0,0]			1189
	-128	[4,1,0,0]			1171
REGION 2	-280	[0,0,1,0]		**1023**	
	-298	[1,0,1,0]		1005	
	-316	[2,0,1,0]		987	
	-354	[1,1,1,0]		949	
	-372	[2,1,1,0]		931	
REGION 3	-558	[0,0,0,1]	743		
	-560	[0,0,2,0]		743	
	-576	[1,0,0,1]	725	727	
	-578	[1,0,2,0]		725	
	-594	[2,0,0,1]	707	709	
	-596	[2,0,2,0]		707	
	-612	[3,0,0,1]	689		
	-614	[3,0,2,0]		689	
	-632	[1,1,0,1] or [4,0,2,0]	669	671	
	-634	[1,1,2,0]		669	
	-650	[2,1,0,1]	651	653	
	-652	[2,1,2,0]		651	
	-668	[3,1,0,1]	633	635	
	-670	[3,1,2,0]		633	
	-686	[4,1,0,1]	615		
	-688	[4,1,2,0]		615	

[a] Mass loss from $[M+H]^+$. Values quoted are the nominal masses of the neutral molecules lost.

[b] Number of each of four types of neutral loss occurring; [water, propenal?, phytene, 1,31-biphytadiene].

[c] Values quoted are the nominal *m/z* values for the product ion observed. The base peak ion in MS/MS is shown in bold face.

*2.2.4.2.5 MS/MS of tetraether lipids **26** and **51***

From the full-MS base peak chromatogram of unknown lipids **26** and **51** it is evident that they both have protonated molecules with nominal *m/z* 1315. The 14 Da increase relative to GDGT **2** suggests that these lipids might be higher homologues of **2** (i.e. $C_{87}H_{174}O_6$ tetraether lipids), although the nature of this homologation, the position of the additional carbon and the core lipid structural type, can not be ascertained from single-stage mass spectral analysis.

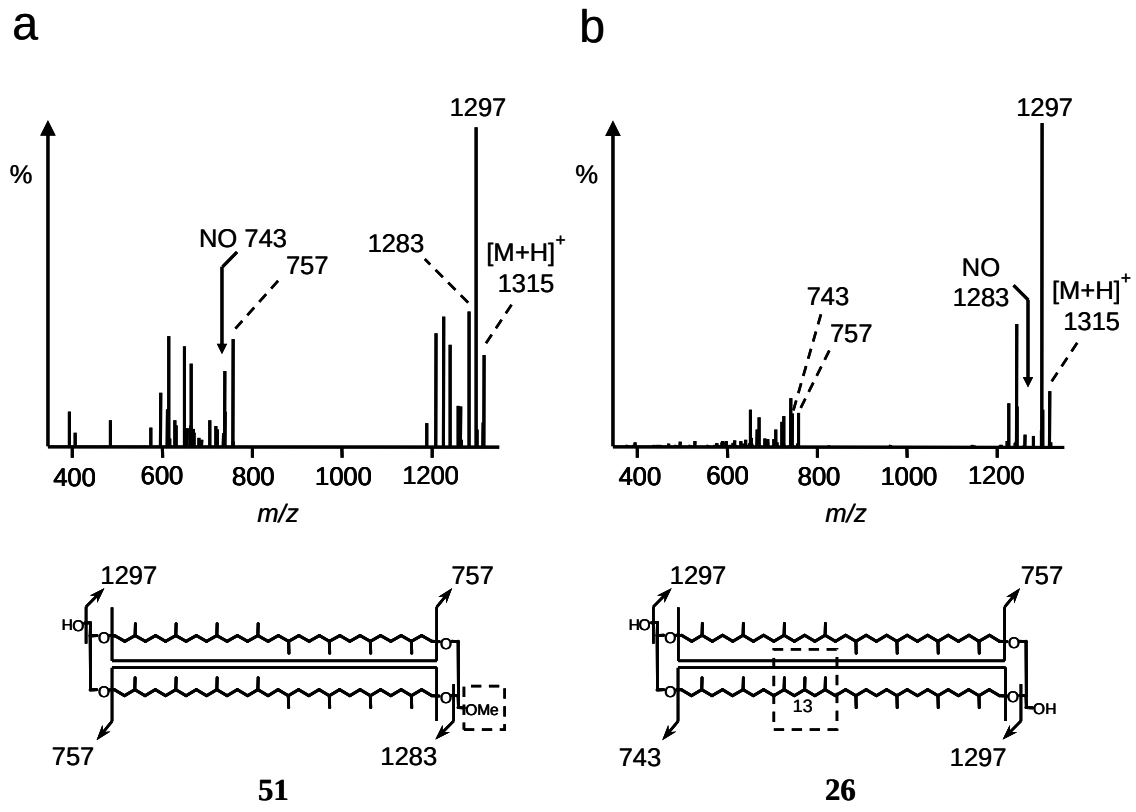


Fig. 2.5. MS/MS spectra (LCQ) of tetraether cores having protonated molecules at nominal *m/z* 1315, showing the characteristic product ions and dissociations responsible for the generation of these ions (protonation not shown). Dashed boxes highlight the position of the additional carbon atom (*cf.* **2a**) in each case; a) **51**. The absence in this spectrum of the product ion at *m/z* 743 is highlighted; b) **26**. The absence in this spectrum of the product ion at *m/z* 1283 is highlighted.

The MS/MS spectra generated from CID of the $[M+H]^+$ of **26** and **51** (Fig. 2.5) only exhibit product ions in Regions 1 and 3, suggesting that both structures are dialkyl tetraether lipids. The MS/MS spectral Region 3, relating to product ions in which an alkyl chain has been lost, provides significant information about the location of the additional carbon atom in **26** and **51**. The protonated molecules of both lipids generate a product ion at *m/z* 757, corresponding to loss of biphytadiene ($C_{40}H_{78}$) during CID. Notably, only **26** exhibits a product ion at *m/z* 743 arising from loss of $C_{41}H_{80}$, indicating that **26** is a structure based on caldarchaeol in which one of the alkyl chains has an additional carbon. Previous studies exploring the biosynthetic pathways of the tetraethers in *M. marburgensis* revealed production of an unexpected $C_{87}H_{174}O_6$ lipid homologue of **2**, which was named homocaldarchaeol (Galliker *et al.*, 1998). Isotopic enrichment ^{13}C-NMR studies determined that the additional carbon in this lipid was a methyl group situated on the C-13 position of one of the two biphytanyl chains of **2** (structure shown in Fig. 2.5). As such, given the close phylogenetic relationship between *M. marburgensis* and MTH(ΔH) (see Section 2.2.4.1.1), it is likely that lipid **26** determined in these studies corresponds to homocaldarchaeol. This lipid is discussed in greater detail in Chapter 3.

The absence of a product ion at *m/z* 743 in the MS/MS spectrum of **51** is consistent with a structure in which only $C_{40}H_{78}$ alkyl groups are present, indicating that the additional carbon atom in this lipid must be incorporated into one of the two capping glycerol groups. The MS/MS spectral Region 1, in which only terminal functionality is lost from the $[M+H]^+$ of **26** and **51**, supports both of these assertions. Product ions at *m/z* 1297 and 1279, relating to loss of one or two molecules of water are observed in the MS/MS spectrum of **26**, wholly consistent with losses from the glycerol functionality observed during CID of the $[M+H]^+$ of **2**. The $[M+H]^+$ of **51**, on the other hand, dissociates to give a product ion at *m/z* 1297 relating to loss of one molecule of water, but does not undergo a subsequent loss of a second molecule of water. It does, however, generate a product ion at *m/z* 1283, consistent with a loss of methanol. As such, it is highly likely that **51** represents a caldarchaeol molecule methyl etherified at the primary hydroxyl of one of its glycerol head groups. All other product ions observed in the MS/MS spectrum of **51** are consistent with this assignment, as is the earlier elution of the component by comparison to **2**. Lipid **51** is

likely to have been generated as an artifact during the acidic methanolysis of the MTH(ΔH) culture material. Strong acidic methanolysis of purified *sn*-2 hydroxyarchaeol has been shown to genererate an artificial methyl ether (Sprott *et al.*, 1990; Ekiel and Sprott, 1992), although the etherification occurred exclusively at the chain hydroxyl and not on the *sn*-1 glycerol hydroxyl group for this lipid. Similarly, an acidic methanolytic extract of the archaeon *Methanothrix concilii*, which produces *sn*-3 hydroxyarchaeol as a major core lipid, also contained chain-etherified *sn*-3 methoxyarchaeol, determined to be an artifact of the extraction procedure by its absence in extracts obtained using milder methanolytic conditions (Ferrante *et al.*, 1988). In addition, previous studies of the products of methanolysis of bacteriochlorophylls have shown the methyl etherification of a secondary hydroxyl, suggesting a precedent for nucleophilic substitution of free hydroxyl groups during acidic methanolysis (Wilson, 2004). To test this assertion, MTH(ΔH) cell material was extracted *via* ethanolytic as opposed to methanolytic cleavage. **51** was no longer observed in the MS base peak chromatogram of the ethanolytic lipid extract although the expected ethyl ether was, unfortunately, not observed. As such, it is clear that **51** is generated from **2** during the extraction of MTH(ΔH) and does not occur naturally in this archaeon.

2.2.5. MS/MS (HCT)

2.2.5.1. Generalities

The MS/MS spectra recorded for lipids **2-3** and **27** on the HCT instrument (Fig. 2.6) during LC-MS/MS are qualitatively similar to those recorded on the LCQ instrument. In particular, the number of distinct regions in which product ions are observed in each MS/MS spectrum correlates to the number of easily dissociable alkyl chains in the tetraether structures, as was noted for the LCQ spectra. Likewise, almost all of the product ions formed from a given lipid during analysis on the LCQ instrument are also observed following CID of the same lipid using the HCT instrument. In addition, the base peak ion observed in the LCQ spectrum of each lipid is also base peak ion in the corresponding HCT spectrum.

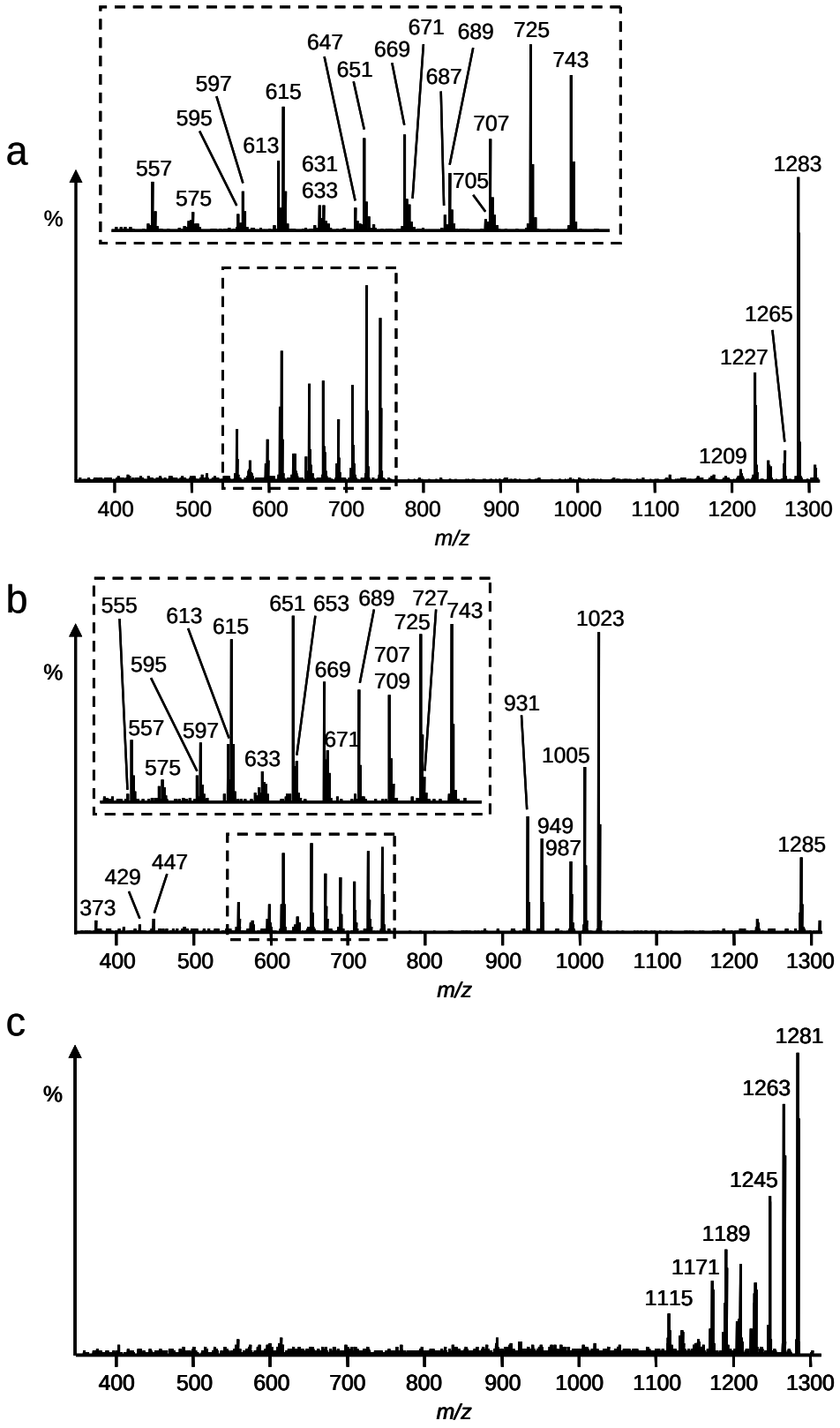


Fig. 2.6. MS/MS spectra (HCT) for: a) GDGT **2** ($[M+H]^+$ *m/z* = 1302.5); b) GTGT **3** ($[M+H]^+$ *m/z* = 1304.0); c) GMGT **27** ($[M+H]^+$ *m/z* = 1300.4). Insets show expansions of the spectra. Product ions are labelled with their nominal *m/z* value.

Due to the higher operational energy of the HCT ion trap and its higher ion capacity, several new product ions were identified for each lipid (Tables 2.2-2.4) that were not easily observed or were not reproducible in MS/MS spectra recorded in successive analyses using the LCQ ion trap. For lipids **2** and **3**, this is, in part, attributable to the greater prominence of Region 3 ions in the HCT spectra compared with the analogous LCQ spectra. Whereas some Region 3 product ions generated using the LCQ were detected in too low frequency to be reliably interpreted, these same ions were much more prominent in the HCT spectra and were reproducible between analyses. Additional characteristic ions are observed in Region 1 of the HCT spectrum of GMGT **27** which are not present in the LCQ spectrum. The HCT MS/MS spectra of lipids **2**-**3** and **27** are addressed individually.

2.2.5.2. MS/MS of GDGT 2

The MS/MS spectrum of **2** acquired on the HCT instrument shows product ions in two distinct spectral regions, as expected (Fig. 2.6a; Table 2.2). Product ions in Region 1 of the spectrum were identical to those observed in the analogous MS/MS spectrum recorded on the LCQ. Product ions in Region 3 at *m/z* 705, 687, 647, 631, 613, 597, 595, 575 and 557 can be observed at very low relative abundance in the LCQ MS/MS spectrum of GDGT **2** (Fig. 2.4a). The small ion currents detected for these ions mean that they are often too weak to be reliably detected and/or assigned. In the HCT MS/MS spectrum, on the other hand, these ions are much more prominent and, consequently, they can be used reliably for characterisation of GDGT **2** when analysed using the HCT instrument. Ions at *m/z* 573 and 555 are also observed in the HCT spectrum, but were never observed in LCQ spectra. The ion at *m/z* 595 is tentatively assigned as a protonated molecule of 1-hydroxybiphytan-32-ol. All ions in the HCT spectrum that were not assigned during the LCQ analyses are not attributable to formation *via* any combination of losses of molecules of water, C_3H_4O and 1,31-biphytadiene. As such, they must form *via* alternative dissociation processes, with potential processes responsible discussed in detail in Section 2.2.11.

Table 2.2. Product ions generated in each of the three distinct regions of the MS/MS spectra following CID of the $[M+H]^+$ of GDGT lipid **2** on the HCT spectrometer.

	Mass Loss (Da)[a]	Number of neutral fragments lost [H_2O, C_3H_4O, $C_{20}H_{40}$, $C_{40}H_{78}$][b]	*m/z* of corresponding product ion generated from $[M+H]^+$ of **2** [c]
$[M+H]^+$	-0	[0,0,0,0]	1301
REGION 1	-18	[1,0,0,0]	**1283**
	-36	[2,0,0,0]	1265
	-54	[3,0,0,0]	1247
	-56	[0,1,0,0]	1245
	-74	[1,1,0,0]	1227
	-92	[2,1,0,0]	1209
REGION 2			
REGION 3	-558	[0,0,0,1]	743
	-576	[1,0,0,1]	725
	-594	[2,0,0,1]	707
	-596	?	705
	-612	[3,0,0,1]	689
	-614	?	687
	-630	?	671
	-632	[1,1,0,1]	669
	-650	[2,1,0,1]	651
	-654	?	647
	-668	[3,1,0,1]	633
	-670	?	631
	-686	[4,1,0,1]	615
	-688	?	613
	-704	?	597
	-706	[2,2,0,1]	595
	-726	?	575
	-728	?	573
	-744	?	557
	-746	?	555

[a] Mass loss from $[M+H]^+$. Values quoted are the nominal masses of the neutral molecules lost.

[b] Number of each of four types of neutral loss occurring; [water, propenal?, phytene, 1,31-biphytadiene].

[c] Values quoted are the nominal *m/z* values for the product ion observed. The base peak ion in MS/MS is shown in bold face.

All product ions reported in the HCT and LCQ MS/MS spectra were accompanied by a less prominent neighbouring ion, shifted by +1 *m/z* units. This second series of ions could potentially represent odd-electron species, formed from tetraether lipids *via* homoloytic bond cleavages (see Section 2.2.11.3). It is, however, more likely that they represent product ions formed from dissociation of the $[M+H+1]^+$ isotope peak of each lipid. The 3 *m/z* unit isolation width used for precursor selection during CID will result in indiscriminate trapping of both $[M+H]^+$ and $[M+H+1]^+$ and consequently, product ions may be formed from both. Similarly, a series of product ions shifted +2 *m/z* units from the assigned ions are likely to be formed by CID of the $[M+H+2]^+$ isotopic peak.

2.2.5.3. MS/MS of GTGT 3

The HCT MS/MS spectrum of **3** (Fig. 2.6b; Table 2.3), as expected, shows product ions in the three major designated regions of the spectrum and includes all product ions observed in the LCQ MS/MS spectrum, apart from the ions at *m/z* 1267 and 653, which were observed consistently in LCQ spectra but not in HCT spectra. Region 3 ions at *m/z* 631, 629, 613, 597, 595, 575, 573, 557 and 555, all of which were weak or not reproducible in LCQ spectra, were observed in the HCT spectra. Of these ions, however, only that at *m/z* 595 can be assigned to formal losses of 1,31-biphytadiene, water and C_3H_4O (Table 2.3). Three other product ions (*m/z* 447, 429 and 373) were also observed following CID of the $[M+H]^+$ of **3** on the HCT instrument. These ions lie outside of the boundaries for Region 3, as classified in Section 2.2.4.2. The ion at *m/z* 373 is likely to be a protonated glycerol monophytanyl ether formed from simultaneous losses of molecules of water, C_3H_4O, 1-phytene and 1,31-biphytadiene from the $[M+H]^+$ of **3** during CID. This ion is also observed following loss of phytene from the $[M+H]^+$ of GDD **1** during CID (see Chapter 3 or Gattinger *et al.*, 2003; Turich *et al.*, 2007). The ions at *m/z* 447 and 429 are difficult to ascribe to any combination of losses of molecules of water, C_3H_4O or isoprenoid chains alone. Consequently, the two ions may have formed *via* cleavages involving C-C bond scission along one of the three composite alkyl chains. Alternatively, given the *m/z* difference of the ions at *m/z* 447 and 429 to the ion at *m/z* 373 (+74 and +56 *m/z* units, respectively), the unassigned ions may have formed

via a rearrangement during CID which transposed a $C_3H_6O_2$ or C_3H_4O from one glycerol in the $[M+H]^+$ of **3** onto either the other glycerol group or one of the C_{20} phytanyl chains (see Section 2.2.11.7 for detailed discussion).

2.2.5.4. MS/MS of GMGT 27

The MS/MS spectrum of **27** recorded on the HCT instrument shows product ions in spectral Region 1, including all product ions derived from losses of small molecules as discussed in Section 2.2.4.2.4, but did not show ions in spectral Regions 2 or 3, as expected. Several additional ions were observed in the HCT MS/MS spectrum of **27** (Fig. 2.6c; Table 2.4) that were either absent or too weak to be assigned in its LCQ spectrum. Some of these ions have *m/z* values just below the defined boundaries for Region 1 stated in Section 2.2.4.2. For example, an ion observed at *m/z* 1115, formed from loss of 184 Da is most likely a fully deoxygenated C_{80} carbocation containing three terminal double bonds. Ions at *m/z* 1187, 1169 and 1133 may have been formed by loss of C_3H_4O units from both ends of the molecule with simultaneous losses of one or three molecules of water for the last two, respectively. An ion at *m/z* 1227 most probably forms *via* loss of four molecules of water from the $[M+H]^+$ of **27**. Other product ions formed during the higher energy CID of the $[M+H]^+$ of **27** (*m/z* 1221, 1205, 1203, 1185, 1153, 1131 and 1129) are less easily assigned to formal losses of water or C_3H_4O and must form *via* alternative dissociation processes, which may include those outlined in Section 2.2.11.4 or alternatively, may represent dissociations through which part of the C_{80} hydrocarbon core is lost from the structure.

Table 2.3. Product ions generated in each of the three distinct regions of the MS/MS spectra and in an additional low *m/z* region following CID of the $[M+H]^+$ of GTGT lipid **3** on the HCT spectrometer.

	Mass Loss (Da)[a]	Number of neutral fragments lost [H_2O, C_3H_4O, $C_{20}H_{40}$, $C_{40}H_{78}$][b]	*m/z* of corresponding product ion generated from $[M+H]^+$ of **3** [c]
$[M+H]^+$	-0	[0,0,0,0]	1303
REGION 1	-18	[1,0,0,0]	1285
REGION 2	-280	[0,0,1,0]	**1023**
	-298	[1,0,1,0]	1005
	-316	[2,0,1,0]	987
	-354	[1,1,1,0]	949
	-372	[2,1,1,0]	931
REGION 3	-560	[0,0,2,0]	743
	-576	[1,0,0,1]?	727
	-578	[1,0,2,0]	725
	-594	[2,0,0,1]?	709
	-596	[2,0,2,0]	707
	-614	[3,0,2,0]	689
	-632	[4,0,2,0]	671
	-634	[1,1,2,0]	669
	-650	[2,1,0,1]?	653
	-652	[2,1,2,0]	651
	-670	[3,1,2,0]	633
	-672	?	631
	-674	?	629
	-688	[4,1,2,0]	615
	-690	?	613
	-706	?	597
	-708	[2,2,2,0]	595
	-728	?	575
	-730	?	573
	-546	?	557
	-548	?	555

LOW *m/z* IONS	-856	?	447
	-874	?	429
	-930	[2,1,1,1]	373

Table 2.3. (cont.)

[a] Mass loss from $[M+H]^+$. Values quoted are the nominal masses of the neutral molecules lost.

[b] Number of each of four types of neutral loss occurring; [water, propenal?, phytene, 1,31-biphytadiene].

[c] Values quoted are the nominal *m/z* values for the product ion observed. The base peak ion in MS/MS is shown in bold face.

Table 2.4. Product ions generated in the three distinct MS/MS spectral regions following CID of the $[M+H]^+$ of GMGT lipid **27** on the HCT spectrometer.

	Mass Loss (Da)[a]	Number of neutral fragments lost [H_2O, C_3H_4O, $C_{20}H_{40}$, $C_{40}H_{78}$][b]	*m/z* of corresponding product ion generated from $[M+H]^+$ of **27** [c]
$[M+H]^+$	-0	[0,0,0,0]	1299
REGION 1	-18	[1,0,0,0]	**1281**
	-36	[2,0,0,0]	1263
	-54	[3,0,0,0]	1245
	-56	[0,1,0,0]	1243
	-72	[4,0,0,0]	1227
	-74	[1,1,0,0]	1225
	-78	?	1221
	-92	[2,1,0,0]	1207
	-94	?	1205
	-96	?	1203
	-110	[3,1,0,0]	1189
	-112	[0,2,0,0]	1187
	-114	?	1185
	-128	[4,1,0,0]	1171
	-130	[1,2,0,0]	1169
	-146	?	1153
	-166	[3,2,0,0]	1133
	-168	?	1131
	-170	?	1129
	-184	[4,2,0,0]	1115
REGION 2			
REGION 3			

[a] Mass loss from $[M+H]^+$. Values quoted are the nominal masses of the neutral molecules lost.

[b] Number of each of four types of neutral loss occurring; [water, propenal?, phytene, 1,31-biphytadiene].

[c] Values quoted are the nominal *m/z* values for the product ion observed. The base peak ion in MS/MS is shown in bold face.

2.2.6. MS^n analysis (LCQ)

Although MS/MS analysis of tetraether lipids leads to formation of a number of product ions from each lipid, multi-stage mass spectrometry (MS^n), whereby individual product ions formed in MS/MS can be isolated and subjected to CID multiple times, could facilitate enhanced characterisation of the lipids. To this effect, the $[M+H]^+$ of GDGT **2** was dissociated in MS/MS during direct infusion of an MTH(ΔH) 70°C extract and individual product ions were subjected to a second round of CID to generate MS^3 spectra. On the whole, ion currents observed in MS^3 were typically more than 30 times lower than those observed during MS/MS and often showed low signal stability. Analysis to MS^4 and beyond was not possible, even for sample concentrations where the ion current for the $[M+H]^+$ of **2** in MS was large. Nevertheless, by selecting individual product ions at *m/z* 1283, 1245, 1227 or 1209 in MS/MS and subjecting these ions to CID, several product ions could be observed in the resulting MS^3 spectra (Fig. 2.7). The product ions reported for the MS^3 spectra are restricted to those which had *m/z* values matching those of product ions observed during MS/MS, indicating that, in spite of the erratic signals, these ions were genuine product ions formed by similar dissociation pathways to those operating during MS/MS. Other ions were observed, but were erratic and not easily explained and, as such, these ions may not be authentic product ions of **2**. Consequently, they are not reported.

The presence of some of the MS^3 product ions suggest that the formal losses of water and biphytadiene reported from the $[M+H]^+$ of **2** in Section 2.2.4.2 may occur *via* a cascade of discrete molecular losses, as opposed to occurring *via* loss of a single, larger molecule. By way of example, the MS/MS product ion at *m/z* 1283, formed from loss of water from the $[M+H]^+$ of **2**, dissociates in MS^3 to give a product ion at *m/z* 725, formed *via* loss of 1,31-biphtyadiene (Fig. 2.7a). This might suggest that the product ion at *m/z* 725 observed in MS/MS (see Fig. 2.4a and Table 2.1) is formed *via* initial loss of water followed by a loss of biphytadiene, or *vice versa*. An alternative explanation, whereby the 576 Da loss which generates the ion at *m/z* 725 from **2** is of an individual molecule, perhaps of biphyt-1-en-32-ol, cannot, however, be ruled out. In all probability, both processes may be in operation during CID.

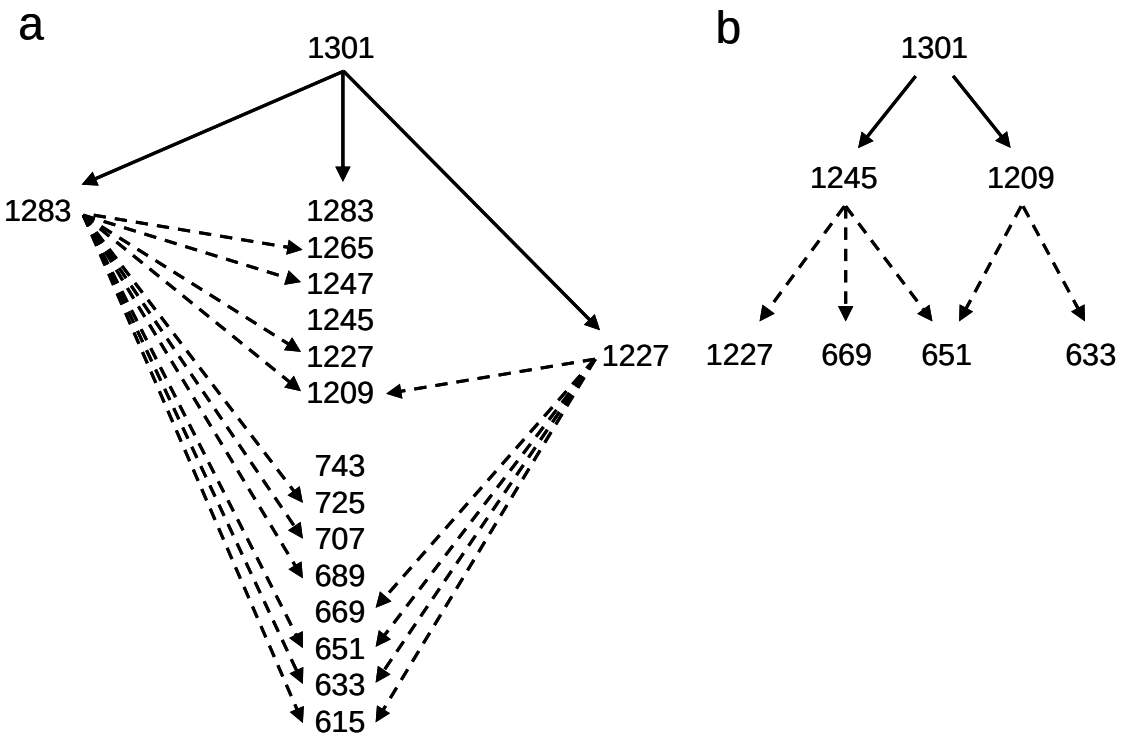


Fig. 2.7. Ions observed in the MS^2 spectrum of lipid **2** (CE = 41%; solid lines) and in the MS^3 spectra obtained following CID (variable CE; dashed lines) of: a) MS^2 ions at *m/z* 1283 (CE = 38%) and *m/z* 1227 (CE = 35%); b) MS^2 ions at *m/z* 1245 (CE = 38%) and *m/z* 1209 (CE = 30%). All values shown are the nominal *m/z* values of ions observed. CE is the % of maximum collision energy.

The presence of an MS^3 product ion at *m/z* 1227, obtained from the precursor at *m/z* 1283 (Fig. 2.7a) suggests that the $[M+H]^+$ of **2** can dissociate *via* a two step process involving loss of water followed by subsequent loss of C_3H_4O. Conversely, the formation of an MS^3 product ion at *m/z* 1227 from the precursor at *m/z* 1245 (Fig. 2.7b) indicates that the molecules can also be lost in the reverse order. These observations provide strong support for the argument that the ion at *m/z* 1227 in the MS/MS spectrum of **2** may form *via* a two step process as opposed to loss of $C_3H_6O_2$ (Section 2.2.4.2.2).

The absence of an ion at *m/z* 743 following MS^3 of the precursor ion at *m/z* 1283 supports the interpretation of *m/z* 743 as having retained all of the oxygen atoms present in **2** (Section 2.2.4.2.2 and Table 2.1). Similarly, ions at *m/z* 725, 707 and

689 could not be observed following MS^3 of the ion at *m/z* 1227 (Fig. 2.7a), consistent with the assignment of the three ions as products of dehydration of **2**, in which the C_3 backbone of both glycerol groups is still intact (Section 2.2.4.2.2 and Table 2.1).

It is important to remember that, although MS^3 analysis proves useful in these select cases, it was found in general that the MS^3 spectra were of significantly lower signal strength than the MS/MS spectra and often gave erratic ion currents. Furthermore, most of the MS^3 spectra did not contain product ions that were not also observed in MS/MS and consequently, routine MS^3 analysis is likely to be redundant for characterisation of tetraether cores.

2.2.7. In-source CID (LCQ)

2.2.7.1. Advantages of sCID in tetraether core lipid analysis

One of the major drawbacks of CID on ion trap instrumentation is the problem of the low mass cut-off (LMCO), where ions with *m/z* below a certain value, typically 1/3 to 1/4 of the *m/z* of the precursor ion, are not detected on account of their instability within the ion trap at the operational q_z value typically used during the process (see Section 1.7.3.6). As such, there invariably remains a question as to whether any of the precursor ion current is sequestered into "invisible" ions with *m/z* values below the LMCO following CID. In-source CID (sCID) can be used to circumvent the LMCO problem, providing dissociation of ionised analytes outside of the ion trap itself (see Section 1.7.3.7). Consequently, application of sCID allowed a cursory investigation into whether **2** may generate product ions below the ion trap LMCO during CID.

2.2.7.2. Direct infusion sCID MS of an MTH(ΔH) extract

The LCQ MS spectrum of a direct infusion of the MTH(ΔH) 45°C extract, generated with sCID activated at 50% maximum collision energy (Fig. 2.8b), contains an ion at *m/z* 1302.0 corresponding to the $[M+H]^+$ of **2**. Other ions were observed at nominal

m/z values identical to those of product ions observed in the LCQ MS/MS spectrum of **2** (reproduced in Fig. 2.8a for comparison), indicating that in-source CID leads to similar dissociation of the precursor as is observed during CID within the ion trap itself. Ions at *m/z* 613, 597, 575 and 557 were also observed in the sCID MS spectrum, as reported in the HCT MS/MS spectrum, but not in the LCQ spectrum, of lipid **2**.

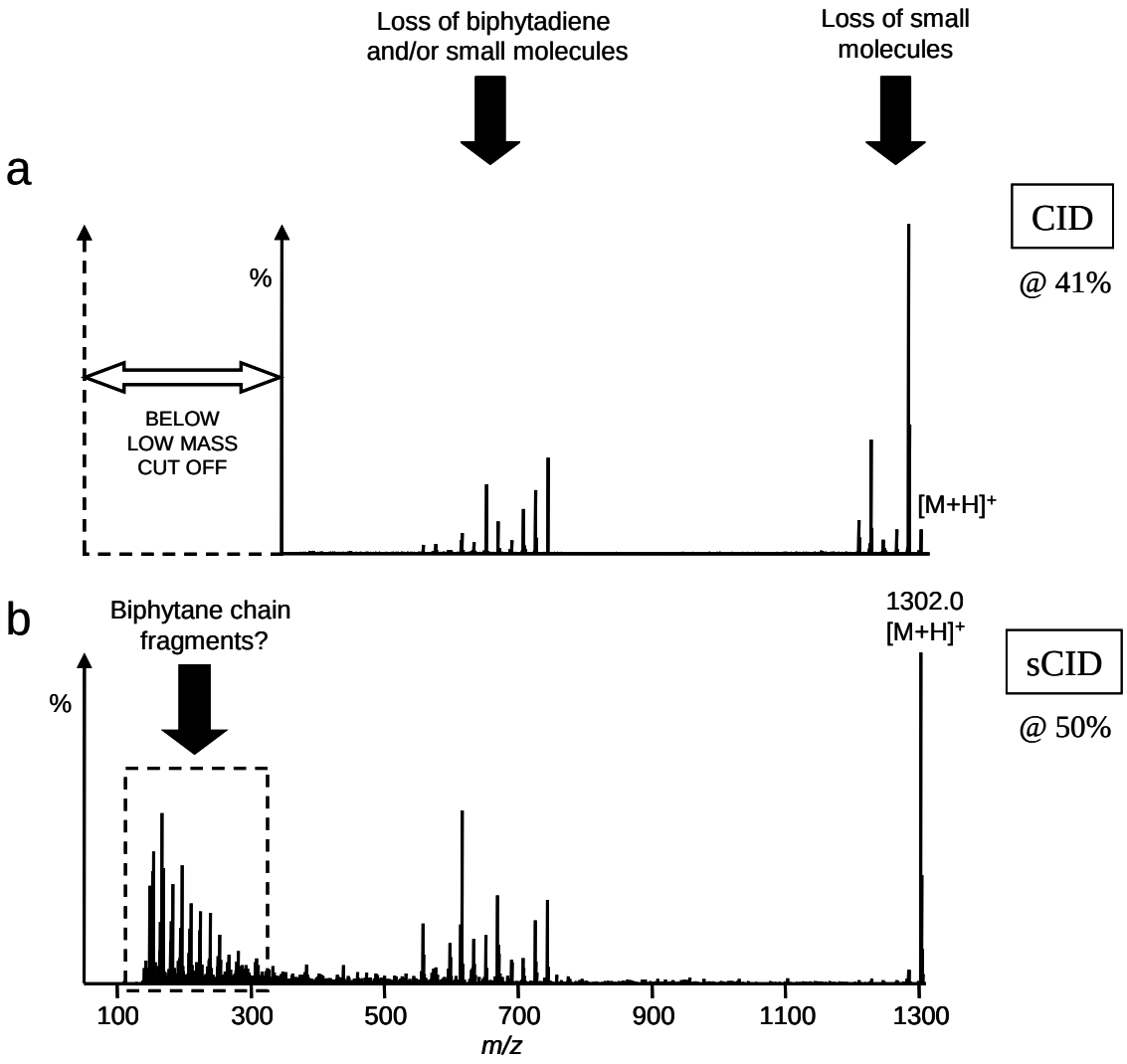


Fig. 2.8. a) MS/MS spectrum (LCQ) of GDGT **2**, reproduced for comparison; b) the *m/z* 50-1310 region of the sCID mass spectrum (LCQ) of the MTH(ΔH) 45°C extract.

As there is no precursor selection during sCID, there is no associated LMCO for the process and, consequently, ions in the sCID mass spectrum were detected across the

full operational MS scan range of the instrument (between *m/z* 50-2000). Intriguingly, a number of ions with *m/z* values below the LMCO (*m/z* 360) for conventional trap-based CID of the $[M+H]^+$ of **2** were observed in the sCID mass spectrum using the LCQ. Furthermore, these ions were not observed in the MS spectrum of the MTH(ΔH) 45^{o}C extract obtained when sCID was not activated (Fig. 2.1a), indicating that they are product ions formed during sCID. Since **2** is the dominant lipid in the MTH(ΔH) 45^{o}C extract, it is most probable that these ions are low *m/z* product ions formed from cleavage of bonds within the biphytane chains present in **2**. Unfortunately, the lack of precursor selection during sCID prevents definitive association of these ions to a particular source. It is conceivable that they may have originated from lipids other than **2** which occur in the MTH(ΔH) extract, or alternatively, from the delivery or makeup solvents used during the direct infusion. As such, despite providing encouraging evidence that there are product ions missed by conventional trap-based CID, the low *m/z* ions cannot be used in any diagnostic fashion.

2.2.7.3. Direct infusion sCID MS/MS of 2

A peculiar effect of applying sCID during direct infusion of the MTH(ΔH) extract, dissolved in DCM and injected into a flow of MeOH, is that it appears to enhance the total MS/MS ion current obtained during subsequent conventional trap-based CID of the $[M+H]^+$ of **2** compared with that obtained in the absence of sCID. This effect was optimal when the sCID was set to 48% of the maximum collision energy, after which the total MS/MS ion current obtained following CID was increased approximately 20 fold relative to the MS/MS current obtained with sCID off (i.e. at 0% maximum collision energy). This enhancement appears to be linked to an increase in the ion currents for characteristic product ions, evidenced by a similar 20 fold relative difference in ion currents for the base peak ion at nominal *m/z* 1283 in the MS/MS spectrum obtained with sCID either on or off. It is difficult to see why activating sCID would modify MS/MS spectra obtained from CID. Even if sCID raises the internal energy of the $[M+H]^+$ of **2** so that it resides in an excited, non-dissociative state which is subsequently more susceptible to CID, this energy would be expected to be dissipated during its low-energy collisions with He gas molecules

during the ion cooling process prior to CID (see Sections 1.7.3.4 and 1.7.3.5). Furthermore, CID of the $[M+H]^+$ of **2** performed in the absence of sCID destroys all but a small proportion of the precursor. CID with sCID activated, on the other hand, leads to a higher proportion of remnant precursor ion, indicating that the degree of CID of the $[M+H]^+$ of **2** is actually more extensive with sCID off. One possible explanation for enhancement in MS/MS ion current would be that some core GDGT **2** signal is present in a non-mass spectrometrically amenable form, locked within a molecule with an *m/z* value above the maximum value of the operational scan range employed (i.e. above *m/z* 2000). sCID could conceivably lead to dissociation of this species, liberating additional protonated molecules of **2** which, in turn, might increase the mass spectral signal for the precursor ion in CID and, consequently, lead to a larger MS/MS ion current. This possibility could be tested by comparison of the ion current for the $[M+H]^+$ of **2** in the MS spectra obtained with or without sCID activated. With sCID turned off, the ion current at *m/z* 1302.0, relating to the $[M+H]^+$ of **2**, was approximately five times greater than in the spectrum obtained when sCID (48%) was activated. Not only does this refute the presence of non-mass spectrometrically amenable ions containing **2**, it also indicates that the conversion of *m/z* 1302.0 ion current to MS/MS ion current is approximately 100 times more efficient when sCID is turned on.

2.2.7.4. Solvent dependence

All of the observations regarding the relationship between application of sCID and MS and MS/MS ion currents were reproduced following a repeat infusion of the MTH(ΔH) extract in DCM. Infusions made using 99:1 hexane:isopropanol as both delivery and makeup solvent also showed enhancement of total MS/MS ion current with sCID on as opposed to sCID off. In this case, the optimal sCID collision energy setting was found to be at 42% of the maximum collision energy, generating an order of magnitude increase in total MS/MS ion current compared with that obtained with sCID off. Irrespective of how the effect occurs, sCID appears to be a useful tool for increasing MS/MS ion current. This may prove to be an important modification to the MS/MS methodology when attempting to record the spectra of tetraether lipids in directly infused, dilute samples. The fact that the enhancement appears to be solvent

dependent may mean that the collision energy settings need to be optimised prior to analysis of such samples.

2.2.8. Panoramic MS/MS (HCT)

2.2.8.1. Advantages of PAN MS/MS for tetraether core lipid analysis

Low *m/z* ions observed in the sCID mass spectrum of the MTH(ΔH) 45^{o}C extract recorded on the LCQ instrument, suggest that the $[M+H]^+$ of GDGT **2** may dissociate to generate product ions at *m/z* values below the LMCO of the ion trap instrument. These ions are absent from MS/MS spectra of **2** generated using conventional CID within the ion trap, suggesting that product ions of potential characteristic value are possibly being overlooked on account of the operational constraints of the trapping and dissociation process. As PAN can circumvent LMCO (see Section 1.7.3.6), this technique was implemented to attempt to visualise the low *m/z* ions suspected to form from the $[M+H]^+$ of **2** during CID.

2.2.8.2. PAN MS/MS of GDGT 2

In order to maximise the likelihood of seeing low *m/z* ions in the PAN MS/MS spectra, the smallest fragment setting was fixed at the lowest allowed value (*m/z* 60) and the tickle level was set to the maximum value (1000%). Direct infusion of the MTH(ΔH) 45^{o}C extract was performed with PAN MS/MS of the $[M+H]^+$ of **2** operated with these two settings applied, allowing the other PAN parameters to be optimised within these constraints (see Section 7.5.5). The resulting PAN MS/MS spectrum obtained for **2** contains the majority of the product ions observed in the HCT MS/MS spectrum obtained using conventional CID, although product ions above *m/z* 1250 observed in the latter spectrum are absent in the PAN MS/MS spectrum (Fig. 2.9a). An ion at *m/z* 519 is observed in the PAN MS/MS spectrum which was absent in the conventional HCT MS/MS spectrum.

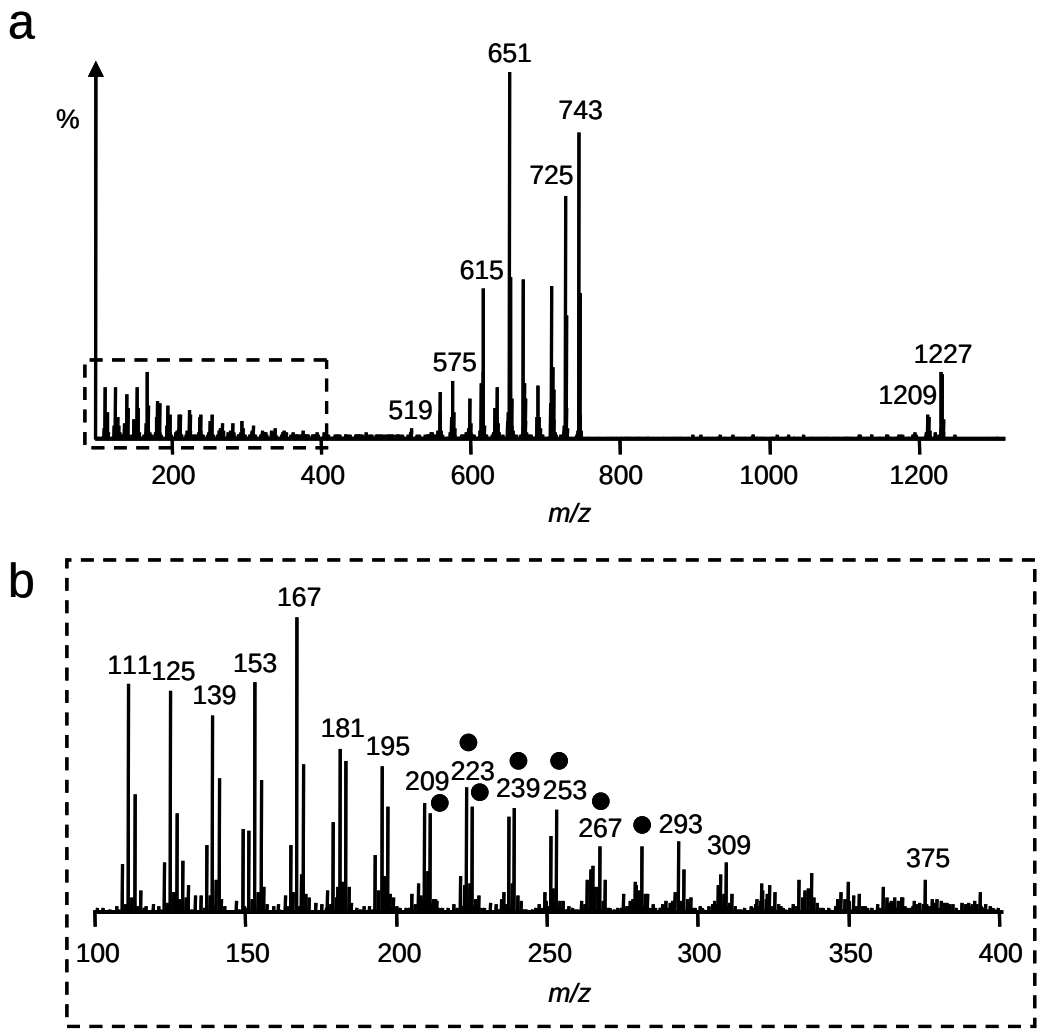


Fig. 2.9. a) The PAN MS/MS spectrum (HCT) of GDGT **2**; b) Expansion of the *m/z* 100-400 region of the spectrum. Product ions are labelled with their nominal *m/z* value. Ions labelled with filled circles were also observed in the Qh-FT-ICR MS/MS spectrum of **2**.

A number of product ions were observed in the PAN MS/MS spectrum at *m/z* values below the LMCO of the HCT MS/MS spectrum (*m/z* 351). Many of the ions at low *m/z* (*m/z* 100-400; Fig. 2.9b) match those observed in the sCID MS spectrum of the MTH(ΔH) 45^{o}C extract, recorded on the LCQ spectrometer (Table 2.5). The ions can be segregated into four homologous series, with *m/z* = 109+14m, 111+14m, 113+14m and 115+14m (m = 0-14). Other ions possibly representing higher members (m > 14) of these homologoues series, such as that at *m/z* 375, which may represent *m/z* 109+(14x19), are occasionally observed but are, in general, too weak

to be wholly reliable in the PAN MS/MS spectra. The low *m/z* values of ions within these homologous series indicate that they must necessarily be formed by cleavages along the biphytanyl chains of **2**. Ions comprising the *m/z* 111+14m and 113+14m series are discussed in detail in Section 2.2.9.6. Ions within the other two series cannot currently be assigned. Irrespective of the structures of all four ion series, they are consistently observed in PAN MS/MS spectra and, consequently, should serve as characteristic product ions for **2**. An ion at *m/z* 149, which does not lie in any of the homologous series, is also reproducibly observed in PAN MS/MS spectra.

2.2.8.3. PAN LC-MS/MS

To extend the application of PAN MS/MS in order to facilitate its use in characterisation of complex mixtures of tetraether lipids, PAN CID was used in place of conventional CID during LC-MS/MS analyses of an MTH(ΔH) 70^{o}C extract using the HCT instrument, with MS and PAN MS/MS spectra continually recorded online. PAN MS/MS spectra of **2** recorded in this manner were similar to the spectra recorded following direct infusion of the MTH(ΔH) 45^{o}C extract, with two notable distinctions. Firstly, the product ions at *m/z* 1245, 1227 and 1209 in the PAN MS/MS spectrum of **2** obtained during the direct infusion were absent in the spectra obtained during PAN LC-MS/MS analyses. More importantly, although the product ions with *m/z* < 400 were the same in both spectra, the abundances of these ions relative to one-another differed. To ensure that the fluctuation in ion abundances related to the PAN LC-MS/MS procedure and not use of MTH(ΔH) samples which had been grown at different temperatures, the PAN LC-MS/MS procedure was repeated three more times using the MTH(ΔH) 70^{o}C extract. In each case, the relative ion abundances were different, indicating that the reproducibility of PAN MS/MS spectra, generated from the same lipid found within the same sample, is poor.

PAN MS/MS spectra of GTGT **3** and GMGT **27** were also obtained during the LC-MS/MS analyses. These lipids dissociate in PAN MS/MS to give a series of ions below *m/z* 400, with many of these ions at *m/z* values matching those of ions observed below *m/z* 400 in the spectrum of **2**. More information regarding PAN MS/MS of GMGT lipids is provided in Chapter 4.

Table 2.5. Product ions in the PAN MS/MS spectrum (HCT) of **2** with *m/z* < 400, with correlations to those found in the sCID MS spectrum (LCQ) of an MTH(ΔH) 45°C extract and in the Qh-FT-ICR MS/MS spectrum of **2** also shown.

Mass loss (Da)[a]	*m/z* of corresponding product ion generated [b]	Series	m	Also observed in		Putative assignment
				sCID MS (LCQ)	MS/MS (Qh-FT-ICR)	
-992	309	113 + 14m	14			$C_{22}H_{45}^+$
-994	307	111 + 14m	14			$C_{22}H_{43}^+$
-1006	295	113 + 14m	13	✓		$C_{21}H_{43}^+$
-1008	293	111 + 14m	13	✓		$C_{21}H_{41}^+$
-1020	281	113 + 14m	12	✓	✓	$C_{20}H_{41}^+$
-1022	279	111 + 14m	12			$C_{20}H_{39}^+$
-1034	267	113 + 14m	11	✓	✓	$C_{19}H_{39}^+$
-1036	265	111 + 14m	11			$C_{19}H_{37}^+$
-1048	253	113 + 14m	10	✓	✓	$C_{18}H_{37}^+$
-1050	251	111 + 14m	10			$C_{18}H_{35}^+$
-1062	239	113 + 14m	9	✓	✓	$C_{17}H_{35}^+$
-1064	237	111 + 14m	9	✓		$C_{17}H_{33}^+$
-1076	225	113 + 14m	8	✓	✓	$C_{16}H_{33}^+$
-1078	223	111 + 14m	8	✓	✓	$C_{16}H_{31}^+$
-1080	221	109 + 14m	8			?
-1090	211	113 + 14m	7	✓	✓	$C_{15}H_{31}^+$
-1092	209	111 + 14m	7	✓		$C_{15}H_{29}^+$
-1094	207	109 + 14m	7			?
-1104	197	113 + 14m	6	✓		$C_{14}H_{29}^+$
-1106	195	111 + 14m	6	✓		$C_{14}H_{27}^+$
-1108	193	109 + 14m	6			?
-1118	183	113 + 14m	5	✓		$C_{13}H_{27}^+$
-1120	181	111 + 14m	5	✓		$C_{13}H_{25}^+$
-1122	179	109 + 14m	5	✓		?
-1132	169	113 + 14m	4	✓		$C_{12}H_{25}^+$
-1134	167	111 + 14m	4	✓		$C_{12}H_{23}^+$
-1136	165	109 + 14m	4	✓		?
-1146	155	113 + 14m	3	✓		$C_{11}H_{23}^+$
-1148	153	111 + 14m	3	✓		$C_{11}H_{21}^+$
-1150	151	109 + 14m	3	✓		?
-1152	149	?	-	✓		?
-1160	141	113 + 14m	2			$C_{10}H_{21}^+$
-1162	139	111 + 14m	2			$C_{10}H_{19}^+$
-1164	137	109 + 14m	2			?
-1172	129	115 + 14m	1			?
-1174	127	113 + 14m	1			$C_9H_{19}^+$
-1176	125	111 + 14m	1			$C_9H_{17}^+$

-1178	123	109 + 14m	1	?
-1186	115	115 + 14m	0	?
-1188	113	111 + 14m	0	$C_8H_{17}^+$
-1190	111	111 + 14m	0	$C_8H_{15}^+$
-1192	109	109 + 14m	0	?

Table 2.5. (cont.)

[a] Mass loss from $[M+H]^+$. Values quoted are the nominal masses of the neutral molecules lost.

[b] Values quoted are the nominal *m/z* values for the product ion observed.

2.2.9. High mass accuracy mass spectrometry

2.2.9.1. Qh-FT-ICR instrumentation and advantages of high mass accuracy MS

Although the assignments of product ions in the ion trap MS/MS spectra of **2**-**3** and **27** in Sections 2.2.4.2 and 2.2.5 represent the most likely structures based on simple elimination of water, C_3H_4O and/or isoprenoid chains from each lipid, the difficulty of obtaining MS^n spectra, coupled to the relatively low mass accuracy on the instruments used means that the assignments are, at best, tentative. Furthermore, some product ions observed in MS/MS and PAN MS/MS spectra could not be assigned solely on the basis of the *m/z* values determined on account of the large number of possible ions which could potentially have an identical nominal mass-to-charge ratio. Fourier transform-ion cyclotron resonance (FT-ICR) MS is both a high resolution and high mass accuracy technique able to measure ions to sub-ppm mass accuracies (Marshall and Hendrickson, 2002). As such, the instrument provides a route to determination of the elemental compositions of ions detected. To this end, a hybrid quadrupole-hexapole FT-ICR (Qh-FT-ICR) instrument comprising an online quadrupole analyser, a hexapole collision cell (containing argon gas) and an ion funnel, incorporated onto the front end of the instrument, was used. As the quadrupole/collision cell combination can be operated in transmission-only mode (i.e. with the quadrupole operated in RF-only mode and collision energy in the hexapole ~ 0 eV) or in CID mode (i.e. with a precursor ion selected in the quadrupole and dissociated in the hexapole), both high mass accuracy MS and MS/MS analyses are possible on the instrument. As such, CID of the $[M+H]^+$ of **2** using the Qh-FT-ICR would allow product ions to be better characterised by assignment to specific compositions or, failing that, by constriction to a limited range of compositions.

2.2.9.2. Direct infusion high mass accuracy MS

An MTH 45^oC extract was introduced into the Qh-FT-ICR instrument *via* direct infusion positive-mode ESI. The FT-ICR mass spectrum obtained in transmission-only mode was dominated by the $[M+Na]^+$ (*m/z* 1324.3059) and $[M+NH_4]^+$ (*m/z*

1319.3510) of **2**, although an ion relating to the $[M+H]^+$ (*m/z* 1302.3228) of **2** was also observed. The sodiated molecule of **2** was similarly observed to dominate in a previous positive ESI MS analysis of tetraether core lipids from *Thermococcus kodakaraensis* (Matsuno *et al.*, 2009). As was noted during ion trap MS (Section 2.2.4.1.2), the abundances of the $[M+H+1]^+$ (92%) and $[M+H+2]^+$ (42%) higher isotopic species relative to the $[M+H]^+$ of **2** were similarly depleted by comparison to the theoretical values (95% and 46%, respectively) in the Qh-FT-ICR mass spectrum, although here to a lesser extent.

2.2.9.3. High mass accuracy MS/MS of 2

CID of the ion at *m/z* 1302.3228 gave rise to a series of product ions in the Qh-FT-ICR MS/MS spectrum (Fig. 2.10a), many of which were observed in the ion trap MS/MS spectra (both LCQ and HCT) obtained following CID of a precursor ion with the same nominal *m/z*. Application of the SmartFormula software program, which generates best-fit elemental compositions to ions on the basis of their accurate *m/z* values, with an additional constraint forbidding nitrogen atoms in any product ion, allowed 48 even-electron ions (including some remnant precursor at *m/z* 1302.3228) to be characterised (Table 2.6). No odd-electron ions were identified. All of the even-electron ions had *m/z* values within 0.3 ppm of the theoretical value of the molecular formula assigned to them. Many had nominal *m/z* values identical to ions observed in the MS/MS spectra of **2** recorded on the ion trap instruments (Tables 2.1 and 2.2), whereas others were not observed in the ion trap spectra.

2.2.9.4. Elemental compositions of LCQ- and HCT-assigned ions

The elemental compositions, determined on the Qh-FT-ICR, of all ions for which specific neutral losses were assigned in Sections 2.2.4.2.2 and 2.2.5.2 are consistent with those losses. Some ions assigned in the ion trap MS/MS spectra of **2**, specifically those at nominal *m/z* 1265, 1247, 1245 and 1209, were not observed in the Qh-FT-ICR MS/MS spectrum of this lipid. As such, their assignments could not be verified.

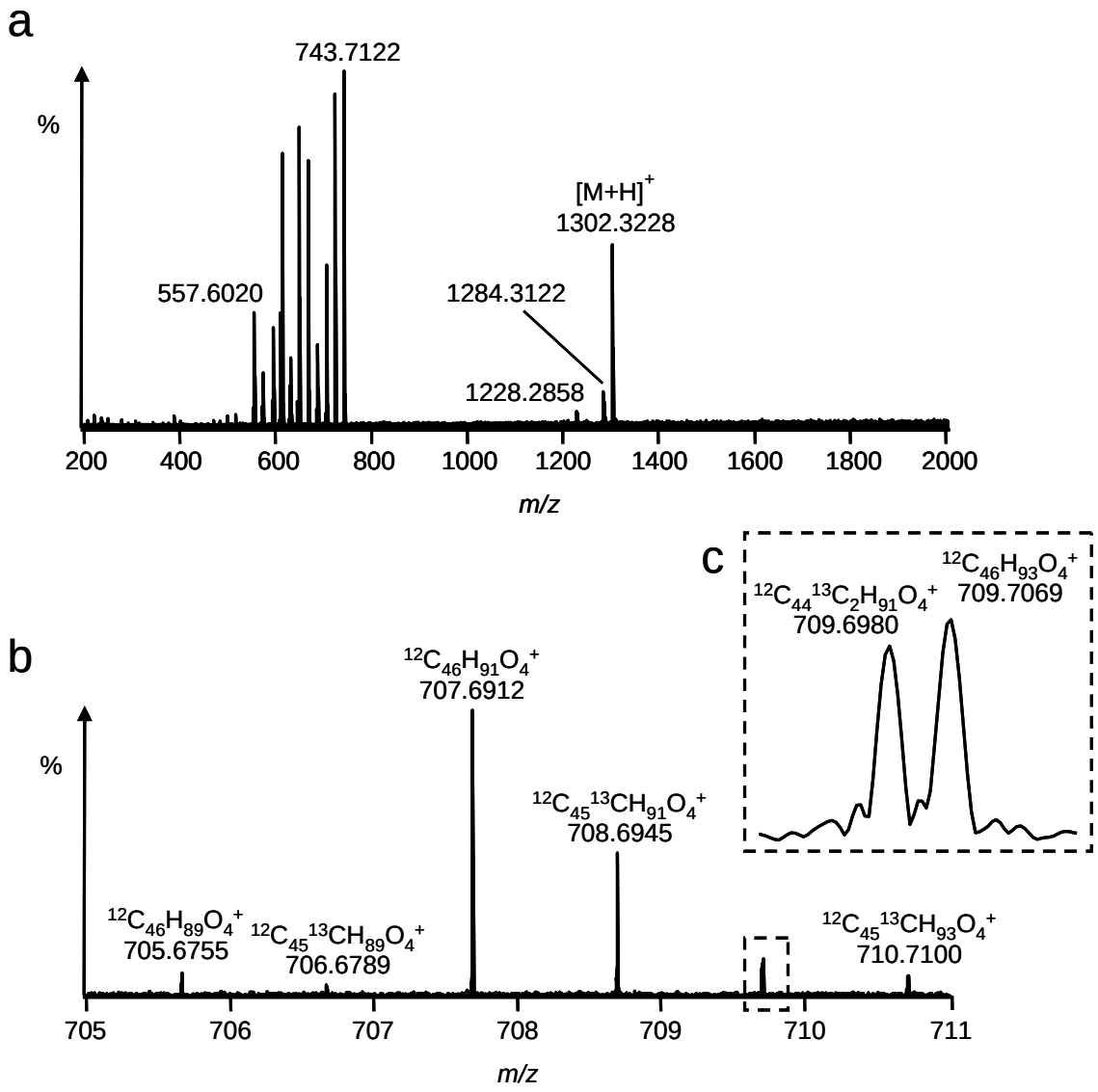


Fig. 2.10. a) MS/MS spectrum of **2**, recorded on a Qh-FT-ICR mass spectrometer; b) magnification of the *m/z* 705-711 spectral region; c) further magnification of peaks with nominal *m/z* 709.

2.2.9.5. Elemental compositions of ions not assigned in ion trap MS/MS spectra

A number of other ions were observed in the Qh-FT-ICR MS/MS spectrum of **2**, some of which were observed in the HCT MS/MS spectrum but could not be assigned to a specific neutral loss (see Table 2.2), whilst others were not reported in the ion trap MS/MS spectra at all. These include ions which are shifted by approximately -2 *m/z* units from an ion trap-assigned ion, with ionic formulae

confirming that they contain two fewer hydrogen atoms. Other ions are shifted by approximately +2 *m/z* units from the ion trap-assigned ions, and contain two additional hydrogen atoms. Additionally, ions were observed at approximately +1 and, on occasion, +2 *m/z* units both from ions assigned on the ion trap instruments and those that could not be assigned in Sections 2.2.4.2 and 2.2.5. The accurate *m/z* values of these ions suggest that they represent the singly and doubly ^{13}C-containing species, respectively. By way of example, the ion at *m/z* 707.6912 ($^{12}C_{46}H_{91}O_4^+$), formally assigned as a product ion arising from loss of two molecules of water and one molecule of 1,31-biphytadiene from the $[M+H]^+$ of **2** (Tables 2.1 and 2.2), is accompanied by ions at *m/z* 705.6755 ($^{12}C_{46}H_{89}O_4^+$), containing two fewer hydrogen atoms, and at *m/z* 709.7069 ($^{12}C_{46}H_{93}O_4^+$) containing two additional hydrogen atoms (Fig. 2.10b). Ions at *m/z* 708.6945 ($^{12}C_{45}{}^{13}CH_{91}O_4^+$), 706.6789 ($^{12}C_{45}{}^{13}CH_{89}O_4^+$) and 710.7100 ($^{12}C_{45}{}^{13}CH_{93}O_4^+$) were also observed and represent singly ^{13}C substituted counterparts of each of the monoisotopic ions. An ion at 709.6980 ($^{12}C_{44}{}^{13}C_2H_{91}O_4^+$) represents the doubly ^{13}C substituted higher isotopic counterpart of the ion at *m/z* 707.6912. The ion at *m/z* 709.6980 is clearly resolved from one at *m/z* 709.7069 by the Qh-FT-ICR, which was not the case with the LCQ and HCT ion trap instruments where a single peak at nominal *m/z* 709 was observed, assigned as the ^{13}C-containing counterpart of $^{12}C_{46}H_{91}O_4^+$ only. The ^{13}C-containing species in the Qh-FT-ICR spectrum are most likely formed from the mono- and di-^{13}C isotopomers of **2**, the $[M+H]^+$ of which will be indiscriminantly trapped and simultaneously subjected to CID along with **2** as a consequence of the wide window (20 *m/z* units) used during the precursor selection on the Qh-FT-ICR. Encouragingly, the accurate mass evidence supports the conclusions drawn in Section 2.2.5.2 regarding +1 and +2 *m/z* ions series observed in the ion trap MS/MS spectra of **2**. Ions at *m/z* 647.6336 ($^{12}C_{43}H_{83}O_3^+$), 629.6230 ($^{12}C_{43}H_{81}O_2^+$) and 611.6124 ($^{12}C_{43}H_{79}O^+$) in the Qh-FT-ICR MS/MS spectrum of **2** are shifted by -4 *m/z* units from ion trap-assigned ions and contain four fewer hydrogen atoms in each case. Since all of the assigned ions for which -4 *m/z* unit counterparts are observed have formally lost 1,31-biphytadiene and $C_3H_6O_2$ (with one, two or three additional losses of water, respectively), the dissociation of glycerol is implicated in the -4 proton difference.

Table 2.6. Comparisons of the measured *m/z* values of product ions in the Qh-FT-ICR MS/MS spectrum of **2** and the theoretical values of the molecular formulae assigned for the ions using the SmartFormula program.

Measured *m/z*	Assigned ionic formula	Theoretical *m/z*	*m/z* error (ppm)[a]	Assigned neutral loss	Consistent with ion trap assignments?
1302.3228	$C_{86}H_{173}O_6^+$	1302.3227	0.1	-	-
1284.3122	$C_{86}H_{171}O_5^+$	1284.3121	0.1	H_2O	✓
1228.2858	$C_{83}H_{167}O_4^+$	1228.2859	0.1	$C_3H_6O_2$	✓
743.7122	$C_{46}H_{95}O_6^+$	743.7123	0.1	$C_{40}H_{78}$	✓
727.7174	$C_{46}H_{95}O_5^+$	727.7174	0.0	$C_{40}H_{78}O$	n.r
725.7017	$C_{46}H_{93}O_5^+$	725.7018	0.1	$C_{40}H_{80}O$	✓
709.7069	$C_{46}H_{93}O_4^+$	709.7068	0.0	$C_{40}H_{80}O_2$	n.r
707.6912	$C_{46}H_{91}O_4^+$	707.6912	0.0	$C_{40}H_{82}O_2$	✓
705.6755	$C_{46}H_{89}O_4^+$	705.6755	0.1	$C_{40}H_{84}O_2$	n.a
689.6806	$C_{46}H_{89}O_3^+$	689.6806	0.0	$C_{40}H_{84}O_3$	✓
687.6651	$C_{46}H_{87}O_3^+$	687.6650	0.1	$C_{40}H_{86}O_3$	n.a
671.6700	$C_{46}H_{87}O_2^+$	671.6701	0.1	$C_{40}H_{86}O_4$	n.r
669.6755	$C_{43}H_{89}O_4^+$	669.6755	0.1	$C_{43}H_{84}O_2$	✓
653.6594	$C_{46}H_{85}O^+$	653.6595	0.1	$C_{40}H_{88}O_5$	n.r
651.6649	$C_{43}H_{87}O_3^+$	651.6650	0.1	$C_{43}H_{86}O_3$	✓
647.6336	$C_{43}H_{83}O_3^+$	647.6337	0.1	$C_{43}H_{90}O_3$	n.a
635.6700	$C_{43}H_{87}O_2^+$	635.6701	0.1	$C_{43}H_{86}O_2$	n.r
633.6544	$C_{43}H_{85}O_2^+$	633.6544	0.0	$C_{43}H_{88}O_4$	✓
631.6387	$C_{43}H_{83}O_2^+$	631.6388	0.1	$C_{43}H_{90}O_4$	n.a
629.6230	$C_{43}H_{81}O_2^+$	629.6231	0.2	$C_{43}H_{92}O_4$	n.r
615.6438	$C_{43}H_{83}O^+$	615.6438	0.1	$C_{43}H_{90}O_5$	✓
613.6281	$C_{43}H_{81}O^+$	613.6282	0.1	$C_{43}H_{92}O_5$	n.a
611.6124	$C_{43}H_{79}O^+$	611.6125	0.2	$C_{43}H_{94}O_5$	n.r
597.6332	$C_{43}H_{81}^+$	597.6333	0.2	$C_{43}H_{92}O_6$	n.a
595.6387	$C_{40}H_{83}O_2^+$	595.6388	0.2	$C_{46}H_{90}O_4$	✓
595.6176	$C_{43}H_{79}^+$	595.6176	0.0	$C_{43}H_{94}O_6$	n.r
575.6125	$C_{40}H_{79}O^+$	575.6125	0.1	$C_{46}H_{94}O_5$	n.a
573.5968	$C_{40}H_{77}O^+$	573.5969	0.1	$C_{46}H_{96}O_5$	n.a
559.5812	$C_{39}H_{75}O^+$	559.5812	0.0	$C_{47}H_{98}O_5$	n.r
557.6020	$C_{40}H_{77}^+$	557.6020	0.0	$C_{46}H_{96}O_6$	n.a
555.5863	$C_{40}H_{75}^+$	555.5863	0.1	$C_{46}H_{98}O_6$	n.a
519.5501	$C_{36}H_{71}O^+$	519.5499	0.2	$C_{50}H_{102}O_5$	n.r
487.5237	$C_{35}H_{67}^+$	487.5237	0.0	$C_{51}H_{106}O_6$	n.r
475.4874	$C_{33}H_{63}O^+$	475.4873	0.2	$C_{53}H_{110}O_5$	n.r
459.4926	$C_{33}H_{63}^+$	459.4924	0.3	$C_{53}H_{110}O_6$	n.r
431.4612	$C_{31}H_{59}^+$	431.4611	0.1	$C_{55}H_{114}O_6$	n.r
417.4455	$C_{30}H_{57}^+$	417.4455	0.1	$C_{56}H_{116}O_6$	n.r

391.2843	$C_{24}H_{39}O_4^+$	391.2843	0.0	$C_{62}H_{134}O_2$	n.r
367.3935	$C_{25}H_{51}O^+$	367.3934	0.2	$C_{61}H_{122}O_5$	n.r
347.3672	$C_{25}H_{47}^+$	347.3672	0.2	$C_{61}H_{126}O_6$	< LMCO
311.3308	$C_{21}H_{43}O^+$	311.3308	0.2	$C_{65}H_{130}O_5$	< LMCO
281.3203	$C_{20}H_{41}^+$	281.3203	0.1	$C_{66}H_{132}O_6$	< LMCO
267.3047	$C_{19}H_{39}^+$	267.3046	0.1	$C_{67}H_{134}O_6$	< LMCO
253.2890	$C_{18}H_{37}^+$	253.2890	0.0	$C_{68}H_{136}O_6$	< LMCO
239.2733	$C_{17}H_{35}^+$	239.2733	0.2	$C_{69}H_{138}O_6$	< LMCO
225.2577	$C_{16}H_{33}^+$	225.2577	0.1	$C_{70}H_{140}O_6$	< LMCO
223.2420	$C_{16}H_{31}^+$	223.2420	0.0	$C_{70}H_{142}O_6$	< LMCO
211.2420	$C_{15}H_{31}^+$	211.2420	0.0	$C_{71}H_{142}O_6$	< LMCO

Table 2.6. (cont.)

[a] ppm = parts per million.

n.r = not reported in the ion trap MS/MS spectra.

n.a = reported, but not assigned in the HCT MS/MS spectra.

< LMCO = have *m/z* below the low mass cut-off in CID of **2**, performed using the ion trap instruments.

A particularly interesting pair of product ions is observed at *m/z* 597.6332 and 595.6176 in the Qh-FT-ICR MS/MS spectrum, determined as $^{12}C_{43}H_{81}^{+}$ and $^{12}C_{43}H_{79}^{+}$ ions, respectively. Both ions are fully deoxygenated, yet they contain more carbon atoms than are present in a biphytanyl chain. Since each biphytanyl chain in **2** is tethered to the three carbon atoms in each glycerol unit *via* an ether oxygen and oxygen is no longer present in the C_{43}-bearing product ions, it is apparent that the two product ions are formed *via* a rearrangement in which the propionic unit from a glycerol moiety is transposed onto one of the C_{40} biphytanyl chains (see Section 2.2.11.7). The ion observed at nominal *m/z* 595 in the HCT ion trap MS/MS spectrum of **2** was suggested to form *via* loss of two molecules of water, two molecules of C_3H_4O and a molecule of 1,31-biphytadiene. An ion at *m/z* 595.6387 ($^{12}C_{40}H_{83}O_2^{+}$) in the Qh-FT-ICR spectrum (Table 2.6) confirms that **2** can undergo these losses. Given the resolution of the additional species at *m/z* 595.6176 ($^{12}C_{43}H_{79}^{+}$) in the Qh-FT-ICR spectrum (Table 2.6), the peak at nominal *m/z* 595 in the HCT spectrum most likely represents both of these product ions.

2.2.9.6. Elemental compositions of ions observed in the PAN MS/MS spectrum

Included among the series of ions generated on the Qh-FT-ICR were several ions with *m/z* values which fall below the LMCOs of the ion trap instruments used. Some of these ions were also observed in the PAN MS/MS spectrum of **2**, where they were assigned to two homologous series with *m/z* 111+14m and 113+14m (see Section 2.2.8.2). Ions at *m/z* 211.2420, 225.2577, 239.2733, 253.2890, 267.3047 and 281.3203 in the Qh-FT-ICR MS/MS spectrum have elemental compositions in the series $C_nH_{2n+1}^{+}$ (n = 15-20), representing carbocations containing no double bond equivalents (i.e. double bonds or rings). These ions have nominal *m/z* values matching those of ions within the *m/z* 113+14m series observed in the PAN MS/MS spectrum of **2**, allowing the full PAN series, by proxy, to be tentatively assigned as $C_nH_{2n+1}^{+}$ (n = 8-22). An ion at *m/z* 223.2420, was assigned an elemental composition within the series $C_nH_{2n-1}^{+}$ (n = 16), indicative of a carbocation containing one double bond equivalent. An ion at nominal *m/z* 223 in the *m/z* 111+14m series was similarly observed in the PAN MS/MS spectrum of **2**, and can now be assigned as $C_{16}H_{31}^{+}$. Again, by association, this implies that the *m/z* 111+14m homologous series can be

tentatively assigned as $C_nH_{2n-1}^+$ (n = 8-22). As discussed in Section 2.2.8.2, both series of ions must be formed by direct cleavage of the biphytanyl chains in **2**, with generation of carbocations of various lengths. The mechanistic implications of this are discussed in Section 2.2.11.8.

Ions within the *m/z* 109+14m and 115+14m homologous series in the PAN MS/MS spectrum of **2** were not identified in the analogous Qh-FT-ICR spectrum. As such, the structural nature of these ions remains unclear.

2.2.10. Derivatisation

2.2.10.1. Necessity

Although the product ions generated from tetraether lipid **2** during CID are assigned to particular neutral losses, there remains ambiguity about how these losses occur. For example, three molecules of water can be lost from the $[M+H]^+$ of **2** during CID, with a fourth also lost in combination with expulsion of other neutral molecules. Whether these losses result from dissociation of the primary hydroxyl groups in the structure or *via* cleavage of one or both of the distinct ether functionalities (i.e. *sn*-2 and *sn*-3 C-O bonds) is not discernable on the basis of the MS/MS spectra of the native structures. Two chemical modifications, silylation and acetylation, were employed in order to modify the terminal hydroxyls and thus explore the losses of water during CID, allowing conclusions regarding the mechanism of water loss from **2** to be drawn. Silicon is less electronegative than hydrogen and consequently, TMS groups added to **2** should act as mild electron donors by comparison with the proton in the native structure. By contrast, acetate groups are electron withdrawing and, consequently, may weaken the ether linkages within the tetraether core. As such, a combination of both derivatisation protocols was used to probe the mechanisms in operation during CID of tetraether core lipids.

2.2.10.2. Silylation

In order to attempt to protect **2** as its di-TMS ether derivative, an MTH(ΔH) 45^{o}C extract was silylated using bis(trimethylsilyl) trifluoroacetamide (BSTFA), trimethylsilyl chloride (TMS chloride) and pyridine. The mass spectrum of a direct infusion of the silylation reaction products contained a number of peaks, including those at *m/z* 1302.5, 1374.5 and 1446.6, relating, in turn, to the protonated molecules of **2** and its mono- and di-TMS ether derivatives (Fig. 2.11a). Importantly, the monoisotopic peaks of the mono- (**52**) and disilylated (**53**) derivatives are not the most abundant isotopic species present in the spectrum, with the +1 isotopomers at *m/z* 1375.5 and 1447.5 present in greatest abundance in both cases (Fig. 2.11a, inset). This is largely consistent with the theoretical isotopic peak pattern predicted for the protonated molecules of both monosilylated ($C_{89}H_{181}O_6Si^+$; monoisotopic peak predicted at 96% of the +1 higher isotopic peak) and disilylated ($C_{92}H_{189}O_6Si_2^+$; monoisotopic peak predicted at 89% of the +1 higher isotopic peak) derivatives of **2**, although the relative abundances measured (99% and 86%, respectively) were slightly askew. Nevertheless, by using the summed abundances of the first four isotopic peaks for **2** and for its two silylated derivatives, and the assumption that **2**, **52** and **53** all have the same ionisation efficiency in MS, the ratio of **2**:**52**:**53** in the silylation reaction mixture was estimated to be 3:50:47. That the silylation reaction had not proceeded to completion was fortuitous, as it allowed derivatives of **2** in which one or both of the terminal hydroxyls had been modified to be investigated in CID.

The MS/MS spectrum of **52** (Fig. 2.11b) contains two regions of product ions, as expected for a GDGT core, and exhibits product ions at *m/z* values either matching those of product ions found in the MS/MS spectrum of **2**, or shifted from these ions by +72 *m/z* units (Table 2.7), reflecting TMS-bound analogues. The MS/MS spectrum of **53** (Fig. 2.11c) also contains two regions of product ions, with the ions observed largely exhibiting product ions either matching those in the MS/MS spectrum of **2**, or shifted by +72 (TMS-bound analogues) or +144 (*bis*-TMS bound analogues) *m/z* units from these ions (Table 2.7). Notably, however, ions at nominal *m/z* 1297, 1229, 781, 671, 653 and 649 in the spectrum of **53** are not consistent with

either of the above categories. On account of the +1 isotope peaks (i.e. $[M+H+1]^+$) of **52** and **53** representing the base peak ions in MS, these ions were selected for MS/MS. The majority of the product ions formed from these precursors appear as predominantly monoisotopic species in the MS/MS spectrum. The only exceptions are the ion at nominal *m/z* 1355 (measured at *m/z* 1356.4) in the MS/MS spectrum of **52** and the ion at nominal *m/z* 1389 (measured at *m/z* 1390.5) in the MS/MS spectrum of **53**, which are less prominent than the corresponding +1 isotope species at *m/z* 1357.5 and 1391.5, respectively. Typically, $[M+H+1]^+$ species dissociate to monoisotopic product ions when the heavy isotope, in this case a ^{29}Si or ^{13}C atom, is contained within the neutral molecule lost. Otherwise, where the neutral molecule lost is monoisotopic, +1 higher isotopomeric product ions are formed. The probability that a monoisotopic product ion is formed from the $[M+H+1]^+$ species is directly correlated to the number of Si and/or C atoms in the neutral molecule dissociated during formation of that ion; the greater the number of atoms in the molecule lost, the greater the chance that the heavy isotope is among them. This accounts for the observation that it is only the product ions formed from the least massive neutral losses from **52** and **53** that show more prominent +1 isotopic species than monoisotopic species in the MS/MS spectra. An additional complication arises from the fact that, on account of the 3 *m/z* unit isolation width used during the precursor selection stage of CID, the $[M+H]^+$ species of **52** and **53** are indiscriminately trapped along with the targeted $[M+H+1]^+$ species, and subsequently dissociated at the same time. As the $[M+H]^+$ species contain no heavy isotope, their dissociation may also account for the prominence of monoisotopic species in the MS/MS spectra.

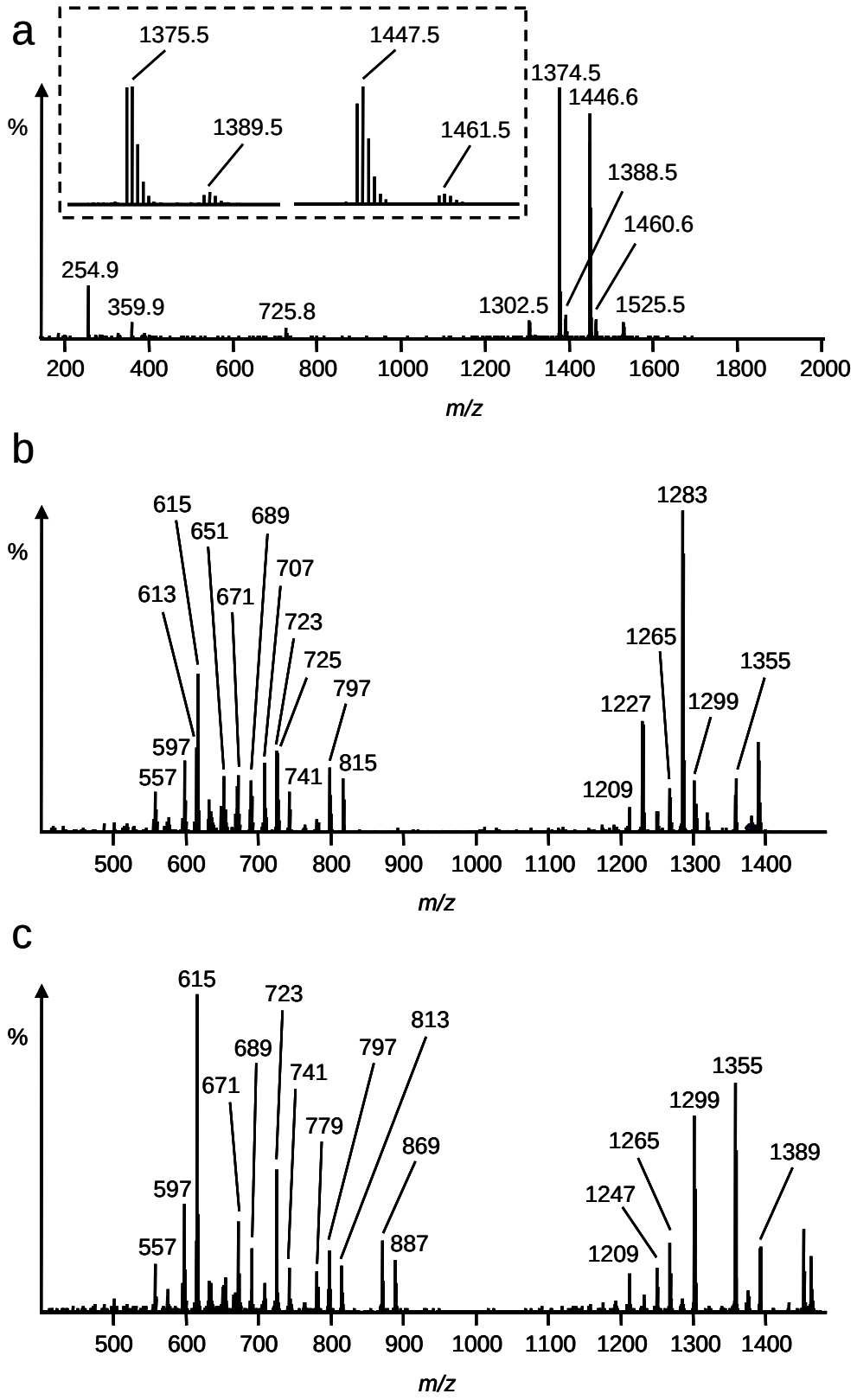
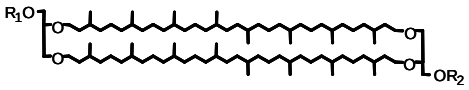
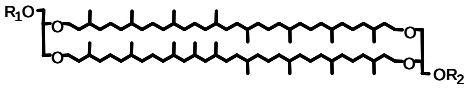


Fig. 2.11. a) Mass spectrum (HCT) of the silylation products of an MTH(ΔH) 45°C extract. Inset shows expansion of regions of the spectrum; MS/MS spectra (HCT) of: b) **52** ($[M+H+1]^+$ *m/z* = 1375.5); c) **53** ($[M+H+1]^+$ *m/z* = 1447.5).

Table 2.7. Structures of and MS and MS/MS (HCT) spectral data for silylated (**52**-**55**) and acetylated (**56**-**57**) derivatives of tetraether lipids **2** and **26**.



52: R1 = H; R2 = $SiMe_3$
53; R1 = R2 = $SiMe_3$
56; R1 = R2 = Ac



54: R1[a] = H; R2[a] = $SiMe_3$
55; R1 = R2 = $SiMe_3$
57; R1 = R2 = Ac

Lipid	Precursor [b]	Product ions observed in MS/MS [c, d]
52	1375 (1374.5)	1355,1317,1301,1299,**1283**,1265,1247,1227,1209; 815,797,779,741,725,723,709,707,689,687,671,669,667,651,647,633,631,615,613,597, 595,575,557,555
53	1447.5 (1446.6)	1389,1373,1371,1355,1337,1319,1299,1297,1283,1281,1265,1247,1229,1227,1209; 887,869,813,797,795,781,779,761,741,723,707,689,671,653,651,649,633,631,**615**,613, 597,595,575,557
54	1389.5 (1388.5)	1369,1355,1351,1339,1331,1313,**1297**,1279,1261,1241,1223; 829,815,811,797,755,741,739,737,725,723,721,707,703,701,689,687,685,683,681,671, 669,665,653,651,649,647,645,643,633,631,629,627,615,613,611,609,597,595,591,589, 575,573,571,557,555
55	1461.5 (1460.6)	1441,1403,1387,1385,1373,**1369**,1351,1313,1297,1279,1261,1223,1205,1203; 901,887,883,869,827,813,811,809,797,795,793,781,779,755,741,737,723,721,703,689, 687,685,681,671,667,665,653,651,649,647,645,633,631,629,627,615,613,611,609,597, 573,571,557
56	1386.6	1367,1343,1329,1325,1307,1289,1269,1267,1265,1251,1249,1247,1229,1227,1209; 809,791,767,749,735,733,731,715,713,693,691,**675**,673,671,653,633,631,615,613,597, 595,575,573,557,555
57	1400.5	**1381**,1363,1357,1343,1339,1321,1303,1283,1281,1279,1265,1263,1261,1243; 823,809,781,763,749,747,745,735,707,705,693,691,689,687,629,615,611,597

[a] An isomer of this lipid in which R1 and R2 are interchanged is also expected amongst the silylation products and would be dissociated with **54** during CID.

[b] For **52**-**55**, the $[M+H+1]^+$ species was selected as the precursor ($[M+H]^+$ shown in brackets); For **56**-**57**, the $[M+H]^+$ was selected as the precursor. Measured *m/z* values quoted in each case.

[c] *m/z* values quoted are the nominal values for the monoisotopic product ions.

[d] Product ions listed in bold face represent the base peak in MS/MS.

A product ion observed at *m/z* 1355 in the MS/MS spectrum of **52** confirms that an individual molecule of water can be lost from the $[M+H]^+$ of this lipid during CID. By contrast, no product ion was observed at *m/z* 1427 in the MS/MS spectrum of **53**, indicating that the $[M+H]^+$ of this lipid does not lose a discrete molecule of water during CID. Likewise, ions at *m/z* 815, 797 and 779 in the MS/MS spectrum of **52**

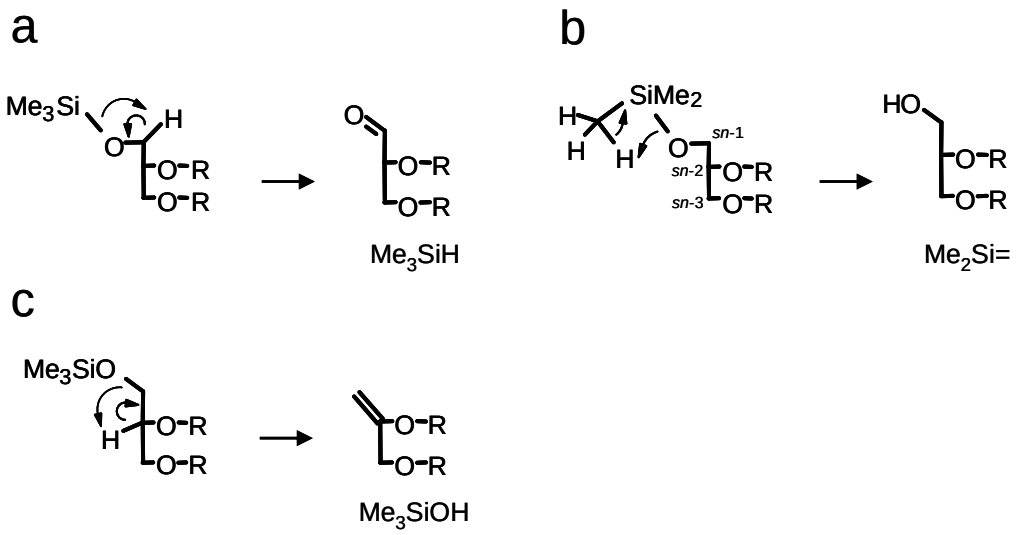
indicate loss of biphytadiene with zero, one and two additional losses of water. On the other hand, peaks at *m/z* 887 and 869, corresponding to ions formed *via* loss of biphytadiene without or with an additional molecule of water, were observed in the MS/MS spectrum of **53**, but no ion was observed at *m/z* 851, as would be expected if the loss of biphytadiene was accompanied by loss of two additional molecules of water. The preference for loss of water from **52**, which contains a terminal hydroxyl, compared to **53**, which is devoid of such functionality, indicates that terminal hydroxyl groups in tetraether lipids are eliminated during CID. Importantly, the absence of a product ion formed from loss of water from the $[M+H]^+$ of **53** does not indicate that the ether oxygen atoms cannot be lost as water; at least some of these oxygen atoms must be labile if losses of three molecules of water from the $[M+H]^+$ of **2** and four molecules of water from the $[M+H]^+$ of **27** during CID are to be explained. The lack of a product ion formed from loss of water from the $[M+H]^+$ of **53** may, therefore, be a consequence of the different electronic demands of terminal hydroxyl groups and TMS ether groups, particularly if the presence of the latter alters the susceptibility for ether cleavage.

Other product ions formed from the $[M+H]^+$ of **52** and **53** include those formed *via* losses of 56 and 74 Da from each. During CID of the $[M+H]^+$ of **2**, these nominal losses were attributed to loss of C_3H_4O and $C_3H_6O_2$, respectively (Tables 2.1 and 2.2). The remnant non-derivatised glycerol group in **52** means that these neutral losses are possible from this lipid, but it is much more difficult to envisage similar losses occurring from **53**, in which both glycerol groups are modified. Alternatively, the 74 Da loss from the $[M+H]^+$ of **53** could relate to loss of trimethylsilane, eliminated from **53** (and also from **52**) *via* transfer of a proton to the silicon in a TMS group (possible pathway indicated in Scheme 2.1a). The 56 Da loss from the $[M+H]^+$ of **53** cannot currently be assigned.

Losses of 72 and 90 Da, which were not observed during CID of the $[M+H]^+$ of **2**, were observed during CID of the $[M+H]^+$ of **52** and **53**. These losses can be attributed to elimination of the TMS groups in **52** and **53** *via* transfer of one of two distinct β-hydrogens. The 72 Da loss is likely to represent a loss of silene, proceeding *via* transfer of a proton from one of the methyl groups in a TMS moiety

to the neighbouring oxygen of an *sn*-1 glycerol carbinol (Scheme 2.1b). The 90 Da loss, on the other hand, probably represents a loss of trimethylsilanol *via* transfer of the proton on an *sn*-2 glycerol carbon to the neighbouring oxygen atom bound to the *sn*-1 glycerol carbon (Scheme 2.1c).

Ions at *m/z* 1388.5 and 1460.6 in the mass spectrum of the direct infusion of the silylation reaction products relate to the protonated molecules of mono- (**54**) and disilylated (**55**) derivatives of homocaldarchaeol (**26**). The MS/MS spectra of the two components (Table 2.7) are consistent with the assigned structure of **26** and with the mechanisms proposed above for CID of the $[M+H]^+$ of **52** and **53**. Although the $[M+H]^+$ of **55** does lose a molecule of water during CID, the product ion formed *via* this loss (*m/z* 1441) is less prominent in the MS/MS spectrum than the corresponding ion at *m/z* 1369 in the MS/MS spectrum of **54**, again implicating terminal hydroxyl groups as sources of water loss during CID of tetraether lipids. The ion at *m/z* 725.8 in the MS spectrum of the reaction products is likely to represent the TMS ether of archaeol, whilst the nature of ions at *m/z* 1525.5, 254.9 and 359.9 is unclear.



Scheme 2.1. Possible pathways for dissociation of the silyl ether group in **52** and **53** (ionising proton not shown): a) elimination of trimethylsilane; b) elimination of trimethylsilanol; c) elimination of dimethylsilene.

2.2.10.3. Acetylation

In order to attempt to form the diacetate of **2**, the 45^{o}C MTH(ΔH) lipid extract was acetylated using acetic anhydride and pyridine. The MS spectrum obtained following direct infusion of the resulting acetylates was dominated by a major component, **56**, which exhibited an $[M+H]^+$ at *m/z* 1386.6 (Fig. 2.12a), consistent with that expected for diacetylated **2**. No remnant ion with nominal *m/z* 1301 was observed in the spectrum, suggesting that the reaction had proceeded to completion. The MS/MS spectrum of **56** (Fig. 2.12b and Table 2.7) shows a prominent product ion at *m/z* 1367, generated by loss of one molecule of water. Since **56** no longer possesses either of the free terminal hydroxyls found in **2**, and given that loss of water from ester moieties is not typically observed during CID (Sigsby *et al.*, 1979a), this implicates the ether oxygen atoms as sources for loss of water during CID.

Other neutral losses observed from the $[M+H]^+$ of **56** include 42 (*m/z* 1343), 60 (*m/z* 1325) and 116 Da (*m/z* 1269), corresponding, in turn, to molecules of ketene, acetic acid and acetylated glycerol ($C_5H_8O_3$) (Scheme 2.2). In addition, an ion observed at *m/z* 1265 can be rationalised as having formed *via* loss of two molecules of acetic acid. As these four losses account for complete elimination of the acetyl groups in **56**, they should prove to be characteristic for glycerol-acetylated tetraether lipids. Interestingly, no product ions formed from dissociation of 98 Da ($C_5H_6O_2$) were observed, suggesting that at least one of the ether oxygens must necessarily be lost from the $[M+H]^+$ of **56** when a glycerol group dissociates during CID. Loss of 56 Da is also observed from the $[M+H]^+$ of **56**. Given that both protons on the terminal hydroxyls of **2** have been replaced *via* conversion to **56**, it is difficult to reconcile the 56 Da loss from the $[M+H]^+$ of **56** with the proposed mechanism for loss of C_3H_4O during CID of the $[M+H]^+$ of **2**, as discussed in Section 2.2.11.6. As such, the loss from the $[M+H]^+$ of **56** can not currently be assigned.

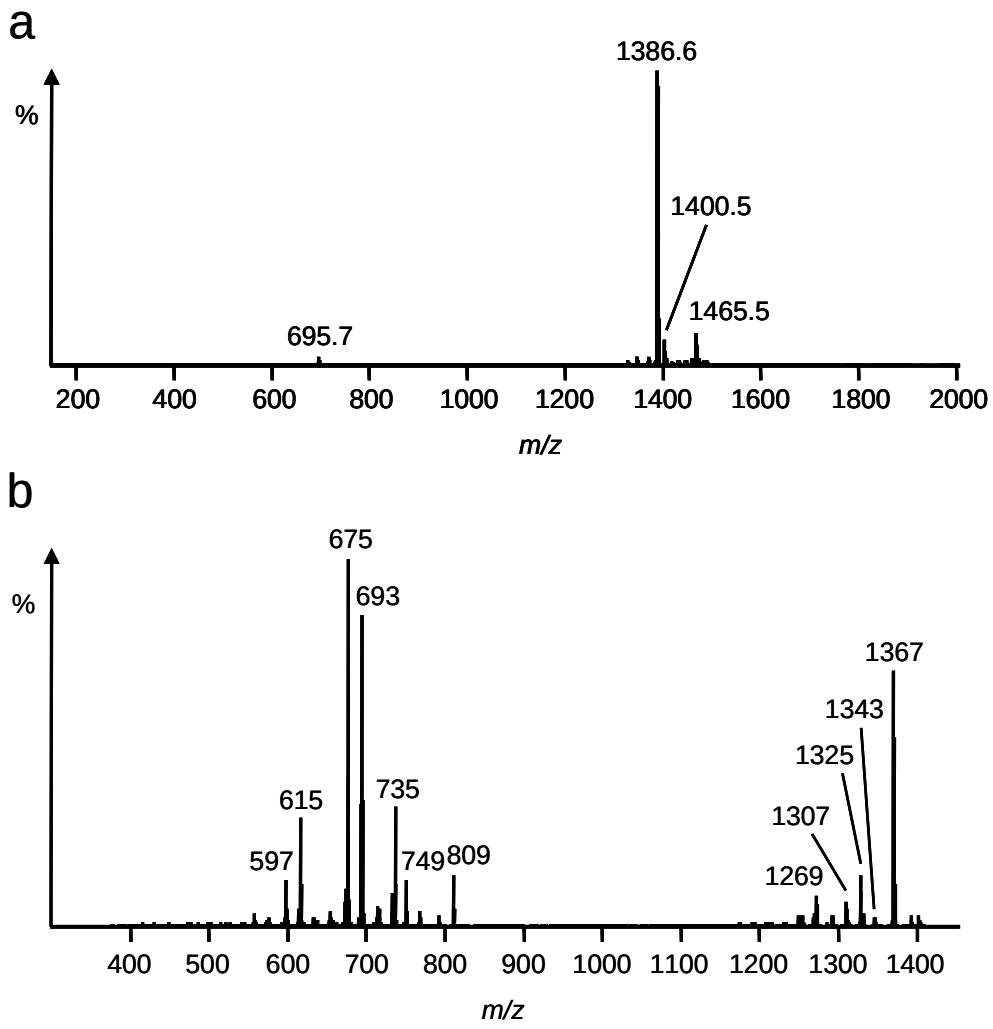
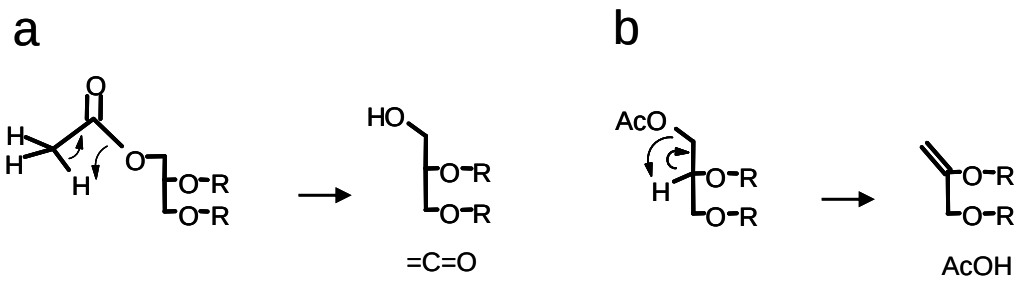


Fig. 2.12. a) MS spectrum (HCT) of the acetylation products of an MTH(ΔH) 45°C extract; b) MS/MS spectrum (HCT) of **56** ($[M+H]^+$ *m/z* = 1386.6).



Scheme 2.2. Possible pathways for dissociation of the acetate group in **56** based on Sigsby *et al.,* 1979a (ionising proton not shown): a) elimination of ketene; b) elimination of acetic acid.

Dissociation of the $[M+H]^+$ of **56** in CID also gives rise to an abundant series of product ions in which an isoprenoid chain has, in some form, been lost from the structure. From the relative ion abundances in the MS/MS spectrum of **56**, it is clear that some of the product ions in which this structural motif has been lost are formed more favourably than those formed from individual losses of the small molecules (i.e. water, ketene, acetate) discussed above. This contrasts with the predominance of ions arising from loss of a single molecule of water during CID of the $[M+H]^+$ of **2** and reflects a greater susceptibility for loss of a biphytanyl chain (with additional loss of functionality) from the $[M+H]^+$ of **56** than from the $[M+H]^+$ of **2**. By contrast with the MS/MS spectrum of **2**, no product ion consistent with loss of 1,31-biphytadiene (558 Da) could be observed in the spectrum of **56**. Instead, the smallest isoprenoid-bearing neutral loss has a mass of 576 Da, relating formally to biphytadiene plus a molecule of water. It is possible that the electronic demands of the terminal acetate groups affect the *sn*-2 and *sn*-3 ether linkages to differing extents, influencing which C-O bond in each ether cleaves. If these groups cleave selectively in opposite directions, 1-ene-biphytan-32-ol as opposed to 1,31-biphytadiene would be the smallest isoprenoid-bearing loss, consistent with the MS/MS spectral data obtained for **56**. An alternative explanation involves ether cleavage by a nucleophilic displacement involving a neighbouring acetate group. This mechanism is discussed further in Section 2.2.11.7.

The ion at *m/z* 1400.5 in the MS spectrum of the acetylation products relates to the $[M+H]^+$ of the diacetate derivative (**57**) of homocaldarchaeol (**26**). Key product ions at *m/z* 1381 (formed by loss of water) and at *m/z* 823 and 801 (formed *via* losses of a molecule of water and either a C_{40} or C_{41} diene) in the MS/MS spectrum of **57** support the conclusions drawn from MS/MS of **56** (Table 2.7). The ion at *m/z* 695.7 in the MS spectrum most likely represents the $[M+H]^+$ of acetylated archaeol.

2.2.11. CID mechanistic considerations

2.2.11.1. Necessity for understanding CID mechanisms

In the preceding sections of this chapter, the product ions formed from tetraether lipids **2**-**3** and **27** have been described and the differences between ion distributions obtained on different mass spectrometers have been discussed. Mechanistic detail, on the other hand, has been kept to a minimum, reserved until all data can be examined as a unified set. With this data set now to hand, the mechanisms by which CID of tetraether lipids may occur can finally be addressed. Understanding the underlying mechanisms active during CID is essential if MS/MS and LC-MS/MS are to be used to determine the structures of novel tetraethers, some of which are discussed in Chapters 3-5. In this section, potential mechanisms for dissociation of tetraether lipids **2**-**3** and **27** are discussed in light of the MS/MS spectra obtained for each.

2.2.11.2. Protonation of tetraethers in APCI

Tetraether lipid cores typically contain six oxygen atoms, providing six potential sites for protonation during ionisation. It is perhaps surprising, therefore, that only singly charged $[M+H]^+$ ions form during APCI of tetraethers. Protonation of the ether oxygens is likely to be preferred to protonation of the alcohol oxygens on account of their greater basicity (Lin *et al.*, 1993), although it remains probable that the total $[M+H]^+$ signal generated during ionisation of a tetraether lipid comprises a statistical mixture of all six possible mono-protonated species. A similar distribution of $[M+H]^+$ species, with protons bound to different ether oxygens, was observed following chemical ionisation of isomeric forms of mixed diethers (Morlender-Vais and Mandelbaum, 1997a). If the proton is assumed to be localised on one oxygen atom only, the $[M+H]^+$ forms of a tetraether lipid can fall into three categories; *sn*-1 hydroxyl-protonated, *sn*-2 ether-protonated or *sn*-3 ether-protonated (Fig. 2.13a). Each form could, in principle, be dissociated during CID, contributing to the product ion distributions observed in the MS/MS spectra of the tetraether lipid. A similar combination of product ions originating from methyl ether-protonated and ethyl-ether protonated ions was observed following CID of protonated diether molecules

containing both functionalities (Morlender-Vais and Mandelbaum, 1997a). To complicate matters, however, thermodynamic and *ab initio* calculations of proton affinities for gas phase polyethers have led to the suggestion that intramolecular hydrogen bonding is likely to be an important non-covalent interaction in the stabilisation of their protonated molecules (Meot-Ner, 1983; Sharma *et al.*, 1984; Lin *et al.*, 1993; Wasada *et al.*, 1996). Typically, this would involve bridging of the ionising proton between two ether oxygen atoms, with additional stability provided by a third and fourth ether oxygen provided that the ionic conformation allows these groups to be in close proximity to the hydrogen-bonded proton (Sharma *et al.*, 1984). Tautomeric processes could also lead to interconversion between different protonated forms. As such, the nature of the $[M+H]^+$ of a tetraether lipid is likely to be intermediate to the six possible monoprotonated forms, with the proton charge being stablised by more than one oxygen atom (e.g. Fig 2.13b). Which oxygens can be bridged by a proton will be heavily dependent on the conformational flexibility of the protonated tetraether lipid in question. For example, the acyclic nature of protonated GTGT lipids should allow for ionic conformations in which hydrogen bonds between oxygen atoms in either the same or different glycerol groups can form. Protonated GMGT lipids, on other hand, are likely to be conformationally restricted by the covalent bond between the isoprenoid chains and may not be able to form hydrogen bonds between two oxygens found at opposite ends of the ion. Protonated GDGT lipids most probably have conformational flexibilities that lie somewhere between those of their GTGT and GMGT counterparts.

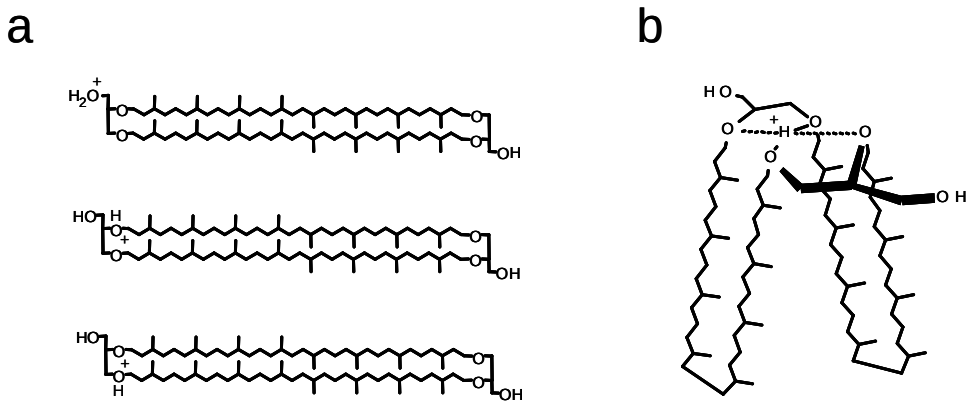
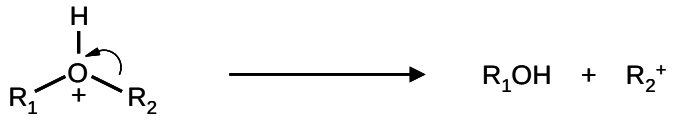


Fig. 2.13. Potential forms of the $[M+H]^+$ of **2**: a) species formed from protonation at a single site; b) a putative hydrogen bond-stablised form.

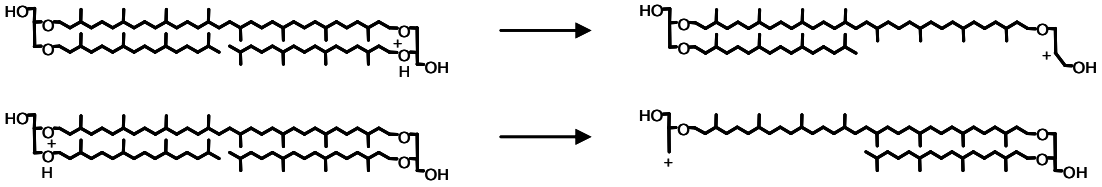
2.1.11.3. Mechanisms for preliminary ether cleavage

Many of the ions that form from tetraether lipids **2**-**3** during CID involve loss of one or more of the composite alkyl chains within the structures. These losses must involve at least one cleavage of an ether bond. Furthermore, this initial ether cleavage is almost certainly a charge-directed process (Morlender-Vais and Mandelbaum, 1997a), implicating ether-protonated forms of the $[M+H]^+$ (probably stabilised *via* hydrogen-bonding) in the process. Several mechanisms have been proposed for charge-directed cleavage of C-O ether bonds following ionisation or during CID (Sisgby *et al.*, 1979b; Morlender-Vais and Mandelbaum, 1997b; Shvily *et al.*, 1997; Ben Ari *et al.*, 2003; 2006). The first process involves an inductive cleavage at the protonated ether, leading to a dissociated carbocation and an alcohol (Sigsby *et al.*, 1979b; Scheme 2.3). The favourability of this process was intrinsically linked to the stability of the carbocation produced. For ethers where tertiary or secondary carbocations could be formed, this process was prominent, whereas formation of primary carbocations was observed to be fairly unfavourable. For protonated tetraether lipids, this process may be in operation but is possibly restricted to ethers bound to *sn*-2 glycerol carbons on account of the stability of the resultant secondary carbocation, although primary carbocations formed from cleavage of *sn*-3 C-O bonds may be stabilised sufficiently by the vicinal *sn*-2 oxygen lone pair to be competitively formed. Loss of 298 Da from the $[M+H]^+$ of **3** during

CID could be explained by dissociation of 1-phytanol by inductive cleavage of an *sn*-2 or *sn*-3 C-O bond (Scheme 2.4).

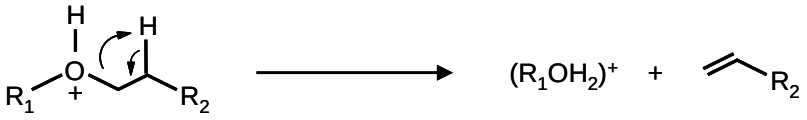


Scheme 2.3. CID of a generic protonated ether *via* inductive cleavage.

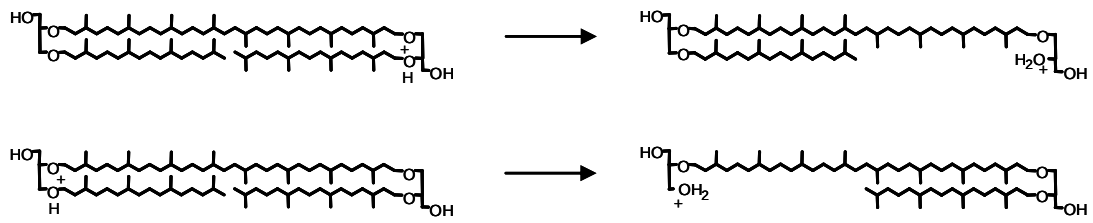


Scheme 2.4. Putative routes for the formation of product ions at nominal *m/z* 1005 from the $[M+H]^+$ of **3** *via* inductive cleavage during CID.

The second process implicated in cleavage of protonated ether bonds involves a four-centred hydrogen transfer from a β-CH or -CH_2 group to the ether oxygen, with elimination of an alkene and formation of a protonated alcohol (Sigsby *et al.*, 1979b; Scheme 2.5). As a consequence of the thermodynamic stability of substituted over unsubstituted alkenes, this process was more favourable for transfer of a secondary proton over a primary proton when both were available in the same molecule. All eight ether C-O bonds in protonated tetraethers have a β-proton, indicating that they could all conceivably be cleaved by this process. Thus, loss of 1-phytene from the $[M+H]^+$ of **3** during CID could occur *via* an elimination of this nature (Scheme 2.6).

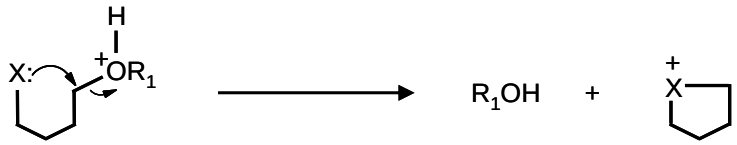


Scheme 2.5. CID of a generic protonated ether *via* elimination to an alkene and protonated alcohol.



Scheme 2.6. Putative routes for the formation of product ions at nominal *m/z* 1023 from the $[M+H]^+$ of **3** *via* eliminatory cleavage during CID.

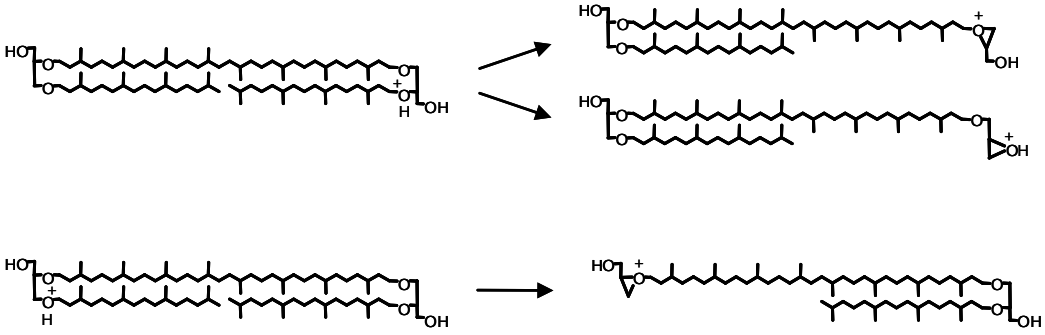
Several reports (Shvily *et al.*, 1997; Ben Ari *et al.*, 2006) have proposed an additional mechanism for dissociation of alcohols from protonated diethers during CID, whereby the alcohol is ejected from the proton-bearing ether by anchimeric assistance from another, nucleophilic ether group found elsewhere in the structure (generalised in Scheme 2.7).



Scheme 2.7. CID of a generic protonated ether *via* anchimerically-assisted displacement of an alcohol. X = nucleophilic atom or group.

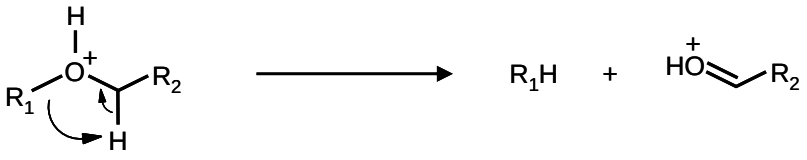
Neighbouring alcohol groups have also been implicated in dissociation of protonated polyethers by anchimeric assitance (Lin *et al.*, 1993). In *sn*-2 ether-protonated forms of tetraether lipids this could equate to nucleophilic attack of an *sn*-1 hydroxyl or *sn*-3 ether oxygen at the *sn*-2 carbon, forming a three-membered cyclic oxonium ion and ejecting an alcohol (Scheme 2.8). In *sn*-3 ether-protonated forms the reverse processes might occur, whereby the *sn*-1 or *sn*-2 oxygen atoms act as the nucleophilic entities in anchimeric dissociation of the *sn*-3 ether linkage (e.g. Scheme 2.8). Furthermore, if the conformation of the ion places the oxygen atoms of differing glycerol groups in close enough proximity to one-another, the oxygen atoms of the neutral glycerol moiety could also attack as nucleophiles, cleaving the protonated ether groups. Regardless of which oxygen atoms are involved as the

nucleophilic partner, anchimeric assistance could provide an alternative explanation for loss of 1-phytanol from the $[M+H]^+$ of **3** during CID.



Scheme 2.8. Putative routes for the formation of product ions at nominal *m/z* 1005 from the $[M+H]^+$ of **3** *via* anchimeric assistance during CID, generating three-membered oxonium ions.

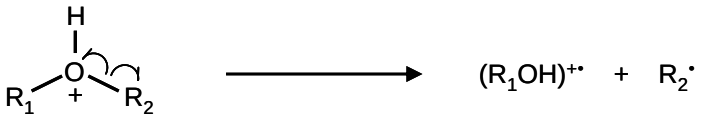
A fourth process suggested for ether cleavage involves a four-centred transfer of an α-proton to an ether-bound alkyl group, resulting in formation of an alkane and a protonated carbonyl (Sigsby *et al.*, 1979b; Scheme 2.9). The protonated ether group in protonated tetraether cores could cleave in either direction *via* this process, resulting in a saturated propionic unit in a glycerol group or in a mono-etherified fully saturated biphytanyl. Given that no loss of phytane was observed during CID of the $[M+H]^+$ of GTGT **3**, the initial ether-cleavage in protonated tetraether lipids is unlikely to occur *via* this process.



Scheme 2.9. CID of a generic protonated ether *via* an α-proton transfer, generating an alkane and protonated carbonyl.

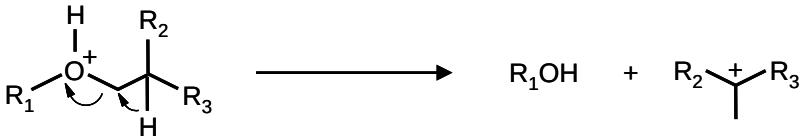
Two further mechanisms have been proposed for cleavage of protonated ether functional groups during CID (Sigsby *et al.*, 1979b; Morlender-Vais and Mandelbaum, 1997b; Ben Ari *et al.*, 2003). The first is a single-electron process,

involving homolytic cleavage of a C-O bond, leading to an alcoholic molecular ion and a hydrocarbon radical (Sigsby *et al.*, 1979b; Scheme 2.10). This process, however, proved to be largely unfavourable in CID of simple protonated ethers (Sigsby *et al.*, 1979b) and, since no odd-electron ions were identified by the SmartFormula program in the Qh-FT-ICR MS/MS spectrum of **2** (Table 2.6), this process, at least as a single-stage cleavage, can be ruled out for protonated tetraether lipids.



Scheme 2.10. CID of a generic protonated ether *via* homolytic cleavage.

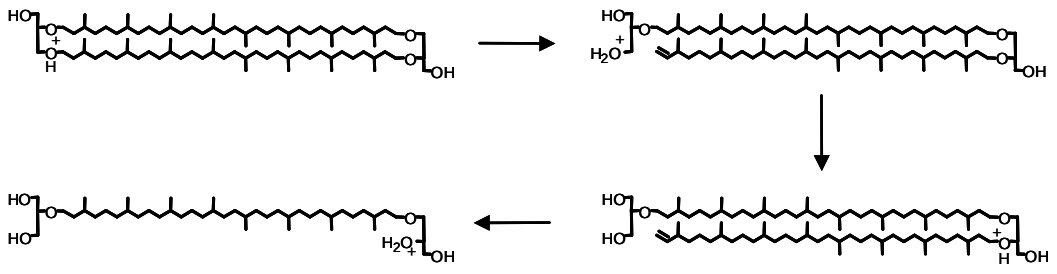
The other process suggested for ether cleavage, occuring in ions where a β-CH_3 group neighbours the protonated ether moiety, involves loss of an alcohol, facilitated by a 1,2 hydride shift, which serves to localise the resultant cationic charge onto the tertiary centre (Morlender-Vais and Mandelbaum, 1997b; Ben Ari *et al.*, 2003; Scheme 2.11). On account of the lack of methyl groups in positions β to any of the oxygen atoms within tetraether lipids **2-3** and **27**, this type of process is unlikely to operate during CID of these lipids, but could be important should tetraether core lipids containing β-methyl groups be identified in future.



Scheme 2.11. CID of a generic protonated ether *via* a 1,2-hydride shift.

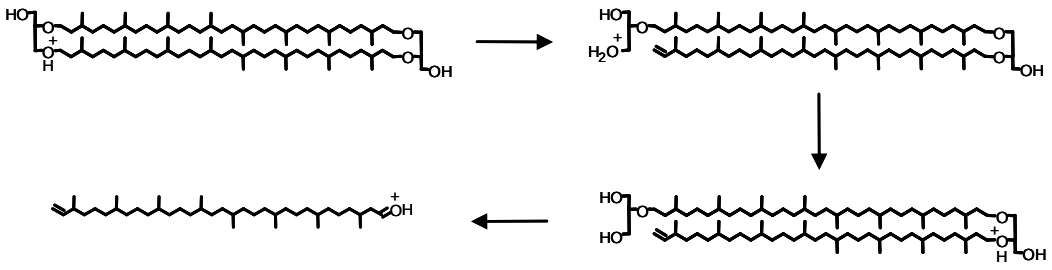
2.2.11.4. Mechanisms for a second ether cleavage

In protonated GDGT lipids, a single ether cleavage step is not sufficient to explain the loss of a biphytanyl chain from the structure during CID: a second ether cleavage is necessary if these chains are to be lost. By way of example, the loss of 1,31-biphytadiene from the $[M+H]^+$ of **2** must involve elimination of the ether at one end of the biphytanyl chain followed by a second elimination at the other end of the chain. If intramolecular hydrogen bonding is ignored, only one of these ether groups can be protonated at a given time, indicating that both charge-directed and charge-remote cleavage processes may be available to tetraether lipids during CID. Alternatively, if, following the initial charge-directed cleavage, the ionising proton was transferred from the protonated hydroxyl group formed during the cleavage to the ether oxygen located at the opposite end of the biphytanyl chain, the second cleavage could also occur *via* a charge-directed process (Scheme 2.12). This would be a likely scenario if the two oxygen atoms were held in close proximity by an initial hydrogen bond prior to the first ether cleavage. This would not, however, be a necessary condition for proton transfer to occur. Given that the initial cleavage destroys the macrocylic nature of the GDGT core, the intermediate ion formed after this cleavage would be expected to be much more confomationally mobile than the original protonated tetraether. As such, a conformational rearrangement could allow transfer of the ionising proton from the hydroxyl group to the ether oxygen prior to the second cleavage step.



Scheme 2.12. Putative route for the formation of a product ion at nominal *m/z* 743 from the $[M+H]^+$ of **2** *via* consecutive eliminations during CID.

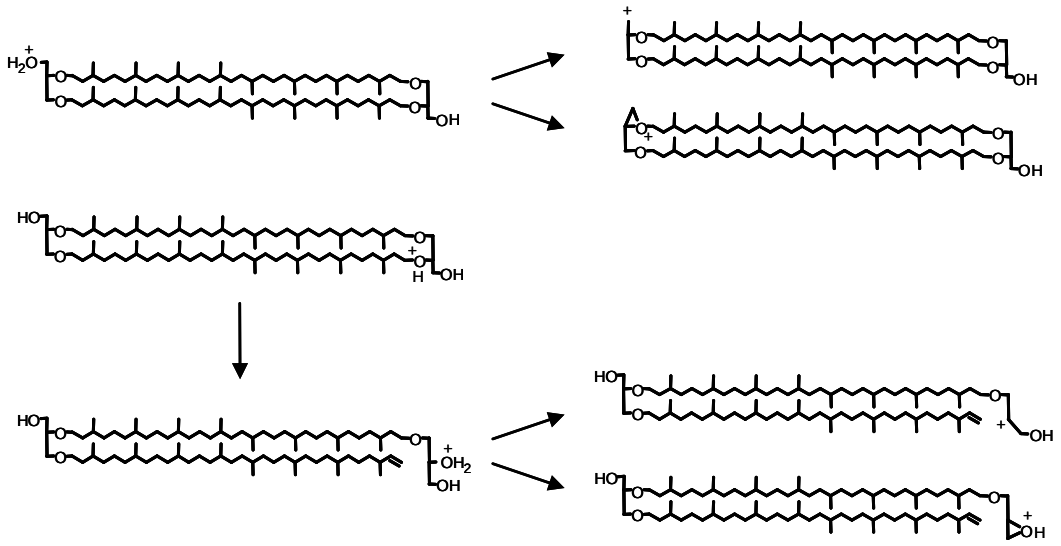
Elimination of a second ether linkage, following preliminary cleavage of an ether linkage by either inductive, anchimerically assisted or elimination processes, can explain almost all ions reported in Region 3 of the LCQ spectrum of **2** (exceptions being ions at *m/z* 689 and 615) and the additional ion at *m/z* 595 in the HCT spectrum of **2**. Furthermore, sequential ether cleavage steps of these types explain many of the ions at the same *m/z* observed in Region 3 of both LCQ and HCT spectra of GTGT **3**. Ether cleavages which result in formation of an alkane and a protonated carbonyl could, in principle, lead to product ions shifted from ions assigned in Region 3 of the MS/MS spectra of **2** and **3** by ±2 or -4 *m/z* units, depending on whether one or both cleavages occurred *via* this process. Ions shifted by these magnitudes were observed in the HCT and Qh-FT-ICR MS/MS spectra of **2** and in the HCT and LCQ spectra of **3**. Ions shifted by the same magnitude from ions in Region 1 of the HCT MS/MS spectrum of **27** were also observed. It is possible that these ions are formed by the alkyl/carbonyl cleavage process, either occurring during one (±2 *m/z* units) or both (-4 *m/z* units) ether cleavages. For example, the ion at nominal *m/z* 575 in the MS/MS spectrum of **2** could be protonated 1-ene-biphytadian-32-al formed by one elimination and one alkane/carbonyl ether cleavage at the two ends of the same biphytanyl chain (Scheme 2.13).



Scheme 2.13. Putative route for the formation of a product ion at nominal *m/z* 575 from the $[M+H]^+$ of **2** *via* one elimination and one alkane/carbonyl cleavage during CID.

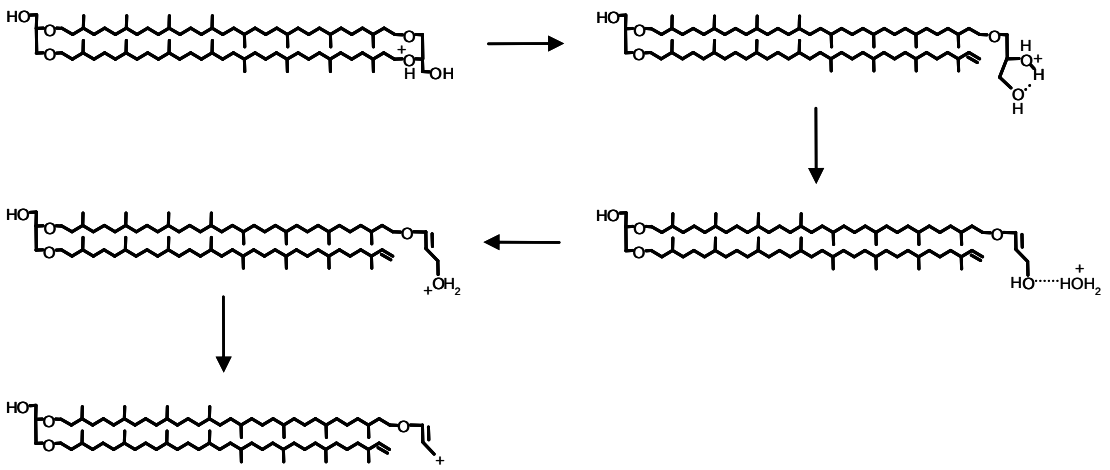
2.2.11.5. Mechanisms for loss of water

A previous study in which short chain aliphatic alcohols were protonated using atmospheric pressure ionisation and subjected to CID indicated that hydroxyl groups in alcohols can be lost as H_2O by inductive cleavage of the C-O bond generating a carbocation, or lost as the hydronium ion (H_3O^+) by elimination to generate a neutral alkene (Karpas *et al.*, 1994). Protonated hydroxyl groups may also lose water *via* anchimeric assistance from nucleophilic (e.g. ether or other hydroxyl) moieties found elsewhere within the same molecule (Shvily *et al.*, 1997). By association, it is easy to envisage how an individual molecule of water might be lost from *sn*-1 hydroxyl-protonated forms of tetraether lipid **2**. The *sn*-1 hydroxyl-protonated ion could lose water *via* induction or anchimeric assistance in a single cleavage step (Scheme 2.14). Alternatively, *sn*-2 and *sn*-3 ether-protonated ions could lose water *via* a two-step process involving initial charge-directed scission of the ether bond followed by inductive or anchimerically assisted cleavage of the hydroxyl group formed (Scheme 2.14). In either case, elimination of a hydronium ion (H_3O^+) could also occur, but would lead to a neutral alkene, undetectable in MS/MS.



Scheme 2.14. Putative routes for the formation of product ions at nominal *m/z* 1283 from the $[M+H]^+$ of **2** *via* loss of water during CID.

GDGT **2** was observed to lose up to three molecules of water without loss of any of its carbon-based architecture. GMGT **27**, on the other hand, loses four individual molecules of water without accompanying neutral losses containing carbon atoms. GTGT **3** can only lose two individual molecules of water at most before other functionality is lost. The losses from **3** suggest that both terminal hydroxyls are lost during CID of tetraether lipids. The additional molecule of water lost from GDGT **2** is most likely provided following generation of a third free hydroxyl group following ether cleavage. Similarly, the third and fourth molecules of water lost from GMGT **27** are likely to be provided following generation of free hydroxyls following two successive cleavages of ether linkages of the same type within the structure (i.e. when both *sn*-2 ether groups or both *sn*-3 ether groups, but not a mixture of the two, are cleaved). Interestingly, the first hydroxyl group lost from each structure during CID, if dissociated by an inductive or anchimerically assisted cleavage, would take the transferable proton with it. It would then be difficult to see how the additional hydroxyl groups could be lost by charge-directed processes. This either implies that neutral hydroxyl groups can be lost *via* charge-remote elimination processes or, more likely indicates that hydrogen bonding is present in the protonated molecules of all three tetraether core types. The latter would allow for elimination of the first molecule of water, with the liberated hydronium ion remaining tethered to the structure by a hydrogen bond to another oxygen atom (Scheme 2.15). A proton transfer from the hydronium ion to the oxygen would dissociate water and preserve the ionising proton within an intermediate ion. A subsequent elimination step with associated proton transfer could liberate a second molecule of water from **2**-**3** and **27**, with **27** eliminating a third molecule of water in the same manner. The final molecule of water lost from either **2** or **27** could be liberated *via* inductive or anchimerically assisted cleavage, taking the ionising proton with it and leaving a carbocation or oxonium ion (e.g. Scheme 2.15).

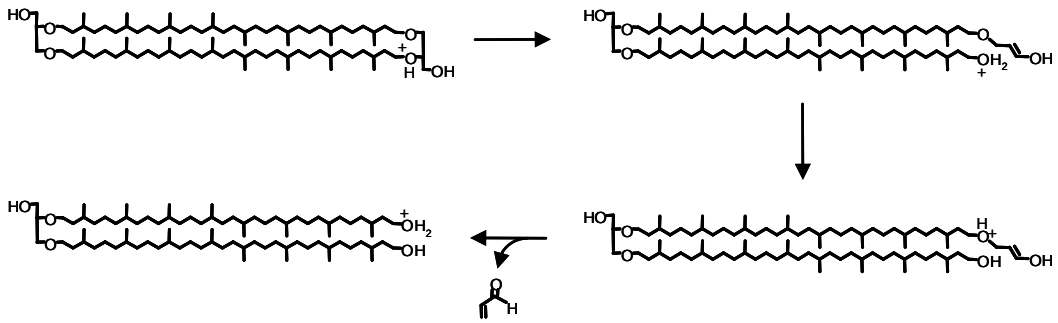


Scheme 2.15. Putative route for formation of a product ion at nominal *m/z* 1265 from the $[M+H]^+$ of **2** *via* consecutive losses of water during CID.

The product ions at *m/z* 689 and 615 in the MS/MS spectra of **2** necessarily require loss of an individual molecule of water in addition to any possible neutral loss bearing a biphytanyl chain that can be liberated by one or two ether cleavage steps. As such, other product ions formed from the $[M+H]^+$ of **2** attributed to loss of a single molecule might instead be formed by loss of 1,31-biphytadiene and subsequent losses of individual molecules of water.

2.2.11.6. Mechanism for loss of C_3H_4O and probable identity

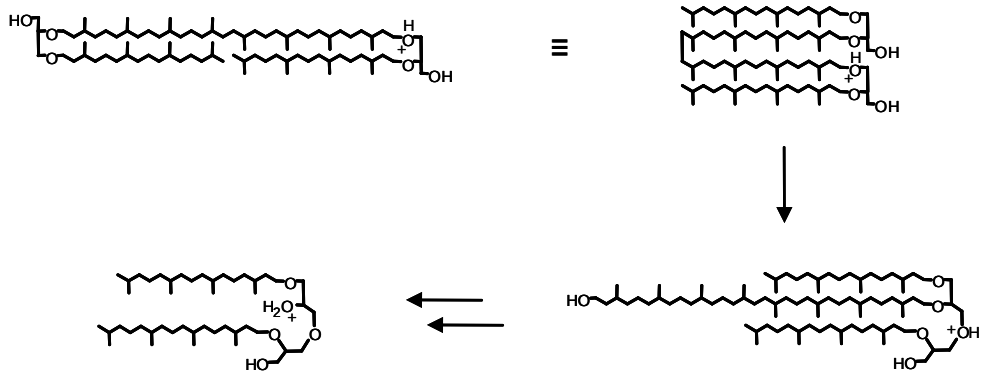
Loss of C_3H_4O from the $[M+H]^+$ of **2** and **27** during CID is unusual, and cannot be explained by ether cleavages *via* any consecutive pair of processes discussed in Section 2.2.11.3. The most likely explanation for formation of the ion at *m/z* 1245 from the $[M+H]^+$ of **2** is *via* an initial elimination of an *sn*-2 ether linkage followed by a retro-Michael addition in which a proton on the neighbouring *sn*-1 hydroxyl group is transferred to the *sn*-3 ether oxygen and a molecule of propenal is eliminated (Scheme 2.16). Propenal, also known as acrolein, is a known dehydration product of glycerol (Izard and Libermann, 1978; Stevens and Maier, 2008) and of glycerides (Stevens and Maier, 2008), lending further support to the assignment for tetraether lipids. On account of the covalent tether between the biphytanyl chains in the structure, two molecules of propenal can be lost from the $[M+H]^+$ of **27** during CID without loss of other neutral molecules occurring.



Scheme 2.16. Putative route for formation of a product ion at nominal *m/z* 1245 from the $[M+H]^+$ of **2** *via* retro-Michael addition during CID.

2.2.11.7. Mechanisms involving rearrangement

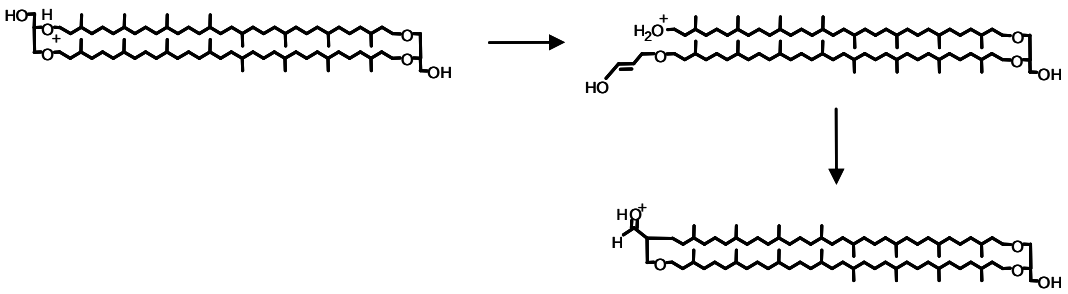
A series of ions in Region 3 of the MS/MS spectrum of **3** appear to have been formed *via* loss of a biphytanyl chain. The nature of the structure of **3** would suggest that this loss, if it were to occur *via* inductive cleavage or elimination of an ether, would be accompanied by loss of a phytanyl chain and a glycerol group. A rearrangement must, therefore, occur at some stage prior to the dissociation of the biphytanyl chain which tethers the glycerol group and phytanyl chain to a carbon retained in the product ion. This could be envisaged to occur *via* nucleophilic attack of the *sn*-1 hydroxyl group on one of the glycerol moieities at the *sn*-2 carbon of the other glycerol group in the structure, resulting in cleavage of the *sn*-2 C-O bond to the biphytanyl chain. Alternatively, the attack from an *sn*-1 hydroxyl could occur at the other end of the biphytanyl chain, cleaving the *sn*-3 ether linkage to this chain (as shown in Scheme 2.17). Rearrangements of this nature would provide the necessary tethering of one glycerol group to the other, and facilitate loss of 1-ene-biphytan-32-ol *via* a subsequent elimination step, forming the ion at *m/z* 727 (Scheme 2.17). The presence of ions at *m/z* 429 and 447 in the MS/MS spectrum of **3** can be explained by loss of 1-ene-biphytan-32-ol in this fashion, followed by loss of 1-phytene or 1-phytenol, respectively. Encouragingly, this mechanism would also account for the absence of loss of an individual molecule of 1,31-biphytadiene from the $[M+H]^+$ of **3** (i.e. no product ion at *m/z* 745 is observed), with at least one oxygen necessarily being lost during the rearrangement and subsequent cleavage.



Scheme 2.17. Putative route for formation of a product ion at nominal *m/z* 727 from the $[M+H]^+$ of **3** *via* a rearrangement and subsequent elimination during CID.

Anchimeric assistance of an acetate group in the ether cleavage processes in operation during CID of the $[M+H]^+$ of **56** (Section 2.2.10.3) may explain why 1-ene-biphytan-32-ol is lost, but 1,31-biphytadiene is not lost from this ion. Similar displacement of an alcohol group from a protonated ether by anchimeric assistance of a neighbouring acyl group has been reported previously for ions formed in gaseous phase ion-molecule reactions (Moraes and Eberlin, 2001).

Ions at nominal *m/z* 597 and 595 in the Qh-FT-ICR MS/MS spectra of the $[M+H]^+$ of **2** appear to contain 43 carbons and no oxygens, implying a rearrangement which has tethered a propionic unit from a glycerol group onto a biphytanyl chain at some point during CID (see Section 2.2.9.5). Elimination of an ether linkage can result in formation of an enol group within a terminal glycerol unit, which could potentially act as a carbon nucleophile (Scheme 2.18). If the enol carbon attacks at the terminal carbon in the biphytanyl chain to cleave an additional ether linkage, the requisite tethering of propionic and biphytanyl chains could occur.



Scheme 2.18. Putative route for the rearrangement of the $[M+H]^+$ of **2** during CID in which the propionic unit in a glycerol moiety is tethered to a biphytanyl chain.

2.2.11.8. Mechanisms involving cleavage of alkyl chains

Ions of low *m/z* (< *m/z* 400) in the Qh-FT-ICR and PAN MS/MS spectra of **2** are formed by cleavage of the biphytanyl chains. Two homologous series of ions, C_nH_{2n+1} and C_nH_{2n-1}, are implicated in Section 2.2.9.6 and, given that these ions occur in an almost continuous series of carbon numbers between n = 8 and n = 22, their formation from **2** suggests that the biphytanyl chains in this structure are susceptible to cleavage at numerous positions. It is unclear, however, whether the dissociation of the alkyl chains occurs whilst they are still bound to the rest of the tetraether architecture, or whether an entire cationic chain is first liberated from the structure and undergoes subsequent dissociation.

2.3. Conclusions

The MS and LC-MS/MS studies reported in this chapter have demonstrated that tetraether lipids can be detected, both during direct infusion and following chromatographic separation, as their protonated molecules in APCI. Although ESI of caldarchaeol (**2**) was also discussed, this technique leads to distribution of the ion current for the analyte into protonated, ammoniated and sodiated molecules, the latter of which dominate the spectra but do not undergo facile CID. As such, APCI is the recommended technique for generating $[M+H]^+$ of tetraether lipids for use in MS and MS/MS analyses. Separation of tetraether core lipid classes by LC-MS/MS appears to be achieved with acceptable resolution, but a significant problem of retention time drift can occur, in particular between analyses made on different days.

Use of GDGT **2** as a retention time marker should allow the elution order and approximate relative retention times of other tetraether lipid cores to be determined in relation to this lipid.

The CID of $[M+H]^+$ of the lipid cores has been achieved on three different types of mass spectrometer: a Qh-FT-ICR instrument and two ion trap instruments. In principle, other instruments capable of CID should also allow product ion formation from tetraether lipid cores. Although ion trap mass spectrometry has been used previously to detect tetraether lipids online during LC-MS (Escala *et al.*, 2007; 2009), this represents the first report of the CID of tetraether cores using this type of mass spectrometer. Ion trap MS/MS spectra of tetraether lipid cores are information rich and contain a myriad of product ions. These include characteristic ions which can be assigned to formal losses of water, C_3H_4O (most likely propenal), 1-phytene and 1,31-biphytadiene. These ions were often accompanied by peaks shifted by +1 or +2 *m/z* units, representing isotopomers formed from $[M+H+1]^+$ and $[M+H+2]^+$ precursors which were unavoidably trapped at the same time as the $[M+H]^+$ ion. Additional ions, which were particularly prominent in the HCT MS/MS spectra, were observed but cannot currently be assigned to particular molecular losses. An awareness of these ions is important when considering the likely structure of tetraether lipid cores on the basis of product ion distributions but, at present, specific conclusions regarding structure should be restricted to those drawn from product ions generated by loss of the four molecules discussed, or closely related structural variants containing clear-cut variations such as additional carbon atoms, hydroxyl groups, double bonds or rings. The product ions formed from loss of 1,31-biphtadiene from the $[M+H]^+$ of **2** or 1-phytene from the $[M+H]^+$ of **3** are of particular significance as they allow, for the first time, the precise mass of the alkyl chains in intact tetraether lipid cores to be deduced. In ion trap MS/MS, loss of the diene from the $[M+H]^+$ of **2** also represents the "cleanest" product ion peak in which the alkyl chain has been lost from the structure, with only the +1 and +2 *m/z* unit isotopomer peaks within 14 *m/z* units of the monoisotopic species. This feature makes ions of this type the most suitable for determining the mass of the chain or chains lost during CID. Interpretation of other product ions in which the alkyl chain has been lost, for example those in which 1-3 molecules of water have also been

expelled, is complicated by the presence of monoisotopic product ions shifted by +2 or -2 *m/z* units. As such, these ions can be used to support assignments made on the basis of the individual diene loss but, where possible, should not be interpreted independently.

A useful feature of the MS/MS spectra of tetraether lipid cores is that the number of distinct regions in which product ions can be observed relates directly to the number of labile alkyl chains within the structure. GMGT, GDGT and GTGT lipids, with one, two and three labile alkyl chains, respectively, dissociate to give product ions in one, two or three regions of the MS/MS spectrum. This allows the rapid classification of tetraether core types online during LC-MS/MS in under 90 min, far faster than alternative techniques which could be used to determine core type. LC-MS/MS using a narrow mass spectral scan range allowed six lipid cores to be initially identified in an MTH(ΔH) extract. Lipids **2a**, **3** and **27** have previously been identified in this archaeon, with lipids **2b** and **26** also implicated on the basis of their presence in close genetic relatives to MTH(ΔH). As lipid **2b** co-elutes with **2a** and is isobaric, it is difficult to determine the presence of both in routine analyses. This is also likely to be the case for other tetraether lipid cores that may be present as *syn* and *anti* isomeric pairs, although in the case of lipids **3**, **26** and **27**, this explanation would not account for the lower peak asymmetries observed for these lipids than for **2a**/**2b** during LC-MS. Nevertheless, for simplicity, tetraether lipids are reported solely as the *anti* isomers in subsequent chapters. This is most likely to be an over simplification of the lipid compositions of some of the samples investigated, with *syn* isomers probably being overlooked. The sixth lipid core in MTH(ΔH), **51**, was shown to be an artifact of the extraction procedure. As this lipid represents less than 0.1% of the total tetraether core lipid in the extract, artifact generation is not of major concern and should not lead to significant distortion of the natural lipid proportions present in archaeal organisms prior to extraction.

The Qh-FT-ICR spectrometer generates the greatest number of product ions of any of the instruments during conventional CID and provides accurate mass-to-charge ratios for ions in both MS and MS/MS. MS^3 studies using the LCQ ion trap instrument, on the other hand, allowed confirmation that glycerol can be lost from

tetraether cores *via* sequential losses of C_3H_4O and water. While both techniques proved to be useful in assignment of product ions observed in the ion trap mass spectra, they are both of limited use for routine LC-MS/MS. At the present time, the Qh-FT-ICR instrument used in the accurate mass studies has no interface for an APCI source and, consequently, can not be used for LC-MS/MS analysis of tetraether lipids cores. This instrument is, therefore, only likely to be of value for obtaining elemental compositions of simple lipid mixtures by direct infusion. MS^3 analysis is erratic and provides little additional information to that provided by MS/MS. Since total MS^3 ion currents are significantly less than those observed in MS/MS, the use of multistage mass spectrometry to MS^3 or beyond is redundant in structural analysis of tetraether cores.

PAN MS/MS of tetraether lipids allowed detection of product ions not observed from the same lipids using conventional ion trap MS/MS on account of the *m/z* values of these ions being below the conventional CID LMCO. More specifically, four series of ions with *m/z* < 400 could be identified in the PAN MS/MS spectrum of **2**. It was initially hoped that these low *m/z* ions, formed by cleavages along the biphytanyl chains, could be used to elucidate the structure of the chains in novel tetraether lipid cores in a manner analogous to that used to interpret EI mass spectra. EI spectral interpretation relies on use of the relative abundances of ions within the spectrum to determine which fragment ions are formed *via* more favourable cleavages that lead to greater charge stabilisation within the fragment ions (McLafferty and Tureĉek, 1993). The parent structure can subsequently be reconstructed by piecing together each fragment ion into a consistent whole. Attempts to use the *m/z* < 400 product ions in an analogous fashion to aid in the structural elucidation of novel tetraether species is almost certain to be hampered by the lack of reproducibility between spectra in replicate PAN LC-MS/MS analyses. An additional downside to the use of PAN MS/MS is that product ions observed in Region 1 of the conventional MS/MS spectra are not detected. Since these ions are important for interpretation of the likely structure of the polar termini in the tetraether lipid cores, LC-MS/MS, as opposed to PAN LC-MS/MS, is a more suitable and reliable technique for routine analysis of samples containing such lipids. Importantly, however, since PAN provides unique, low *m/z* ions, it should serve as a

complementary technique in some cases. sCID was also used to overcome the problem of LMCO in the ion trap instruments and generated the same low *m/z* ions from an MTH(ΔH) extract as those observed in the PAN MS/MS spectrum of **2**. The lack of precursor selection during sCID, prevents use of these ions for any characteristic purpose. An additional property of sCID is that, for reasons unknown, it appears to enhance MS/MS ion current in subsequent trap-based CID experiments. This enhancement could be of particular importance to improve the quality of weak MS/MS spectra, particularly for samples of low concentration where the associated precursor ion current is small.

Speculative mechanisms for CID of tetraether lipids have been provided and discussed in the context of the MS/MS data obtained for **2**-**3** and **27**. The dissociation processes presented can explain the formation of most of the product ions which originate from these lipids. Some ions in the MS/MS spectra of **2**, recorded on the different instruments, cannot be assigned on the basis of any of the listed processes, indicating that CID of the lipid may proceed by additional, more elaborate mechanisms. Silylation and acetylation of **2**, followed by CID, allowed some mechanistic conclusions to be drawn but more subtle or more varied structural manipulations are needed if the mechanisms of tetraether dissociation in CID are to be understood fully. Incidentally, silylation and acetylation are typical processes used to derivatise lipid extracts prior to analysis of diether core lipids using gas chromatography (Blau and Halket, 1993). The fact that silylated and acetylated tetraethers ionise and dissociate well during APCI and CID, respectively, may allow preparation of total lipid extracts by a single derivatisation, which makes the extract amenable to both GC and ion trap MS/MS analysis.

Chapter 3

Tetraether lipid core profiles of thermophilic archaea

3.1. Introduction

3.1.1. Thermophilic and hyperthermophilic archaea

Since the landmark discovery of living communities of micro-organisms in the hot springs of Yellowstone National Park (Brock and Brock, 1966), there has been active interest in the ecology of geothermal aquatic environments. A number of thermophilic archaea, with optimal growth temperature typically above 45^{o}C, and hyperthermophilic archaea, with optimal growth temperature above 80^{o}C, have subsequently been isolated from environments of this nature and are available in culture. In order to survive the extreme conditions they inhabit, (hyper)thermophilic archaea biosynthesise a number of unusual biomolecules that facilitate cellular functioning at elevated temperatures. Many of these molecules have been found to be of biotechnological value, particularly thermostable enzymes, such as DNA polymerases used in the polymerase chain reaction (reviewed in Egorova and Antranikian, 2005). The membrane lipids of (hyper)thermophiles also have the potential to fulfil important biotechnological functions on account of their temperature and pH stability, showing promising signs as self-adjuvanting delivery systems for vaccines (Krishnan and Sprott, 2008), as sterile coatings for catheter tubing (Frant *et al.,* 2006) and as thermostable lubricants (Chang, 1992). Thermophilic and hyperthermophilic archaea often express tetraether lipids containing cyclopentyl rings within their core structures (De Rosa and Gambacorta, 1988a; Schouten *et al.*, 2007c), with ring incorporation thought to be a physiologically-induced response to elevated temperature and possibly to increasing pH (see Sections 1.4.1 and 1.5). Whilst cyclopentyl ring-containing tetraether core lipids are indicative of an archaeal origin, mesophilic archaea can also express core lipids containing up to four cyclopentyl rings (Schouten *et al.*, 2008c). Consequently, the core lipids, when identified in environmental samples, can usually only serve as non-specific biomarkers for the archaeal domain unless the current or past conditions in the environmental setting are well understood. The search for lipid biomarkers which are more specific to a particular species or genus of the Archaea is, consequently, of the utmost importance if more detailed geochemical interpretations are to be made.

3.1.2. Aims

The primary aim of the work presented in this chapter was to extend the application of the LC-MS/MS methodology outlined in Chapter 2 for use in detailed profiling of the membrane lipids of thermophilic archaea. Specific objectives were to: determine if lipid profiles were consistent with those obtained previously for the organisms using alternative analytical methods, thereby allowing the LC-MS/MS methodology to be verified against a greater array of structures and to identify and confirm novel tetraether lipid cores using the potential of LC-MS/MS for detection at very low levels.

3.2. Results and Discussion

3.2.1. Organisms

Four species of thermophilic archaea were selected for lipid core profiling as examples of the two major kingdoms, *Crenarchaeota* and *Euryarchaeota*. These species were of particular appeal on account of the ready availability of cellular biomass at UK-based institutions, *Sulfolobus* species having been grown by Charlotte Nunn (University of Bath) and the *Methanothermobacter* species by James Chong (University of York). The conditions used in each case were selected to be close to those known to provide optimal growth in order to produce sufficient cellular material for each organism (Table 3.1). As discussed in Section 1.5, the core lipid profiles of archaea can be strongly dependent on the conditions utilised during growth, with temperature and pH, in particular, known to induce physiological changes in the core lipids expressed by some organisms. As such, the ambient growth temperature and pH utilised during growth is explicitly stated for each organism. Lipid cores were isolated from each organism by acidic methanolysis of a cell pellet, followed by purification of a polar fraction by flash column chromatography (Section 7.2).

Table 3.1. Organisms used in the studies, with the physiological temperature and pH values used for growth of each.

Organism	Strain	Growth Conditions		References
		Temperature / ^{o}C	pH	
Methanothermobacter thermautotrophicus	ΔH	45 or 70	6	Zeikus *et al.*, 1972; Guy *et al.*, 2004
Sulfolobus acidocaldarius	MR31	75	3-3.5	Reilly *et al.*, 2001
Sulfolobus solfataricus	P2	~75	3.5	De Rosa *et al.*,1975; Zillig *et al.*, 1980
Sulfolobus shibatae	B12	~75	3.5	Grogan *et al.*, 1990

3.2.2. LC-MS/MS

3.2.2.1. Extended and specific MS scan ranges

LC-MS/MS of the lipid extracts of the thermophilic archaea were performed using an analytical method similar to that described in Chapter 2. In order to facilitate a complete profile, as opposed to cursory investigation of a limited number of tetraether cores, the scan range was extended from *m/z* 1280-1320 (as used in Chapter 2) to *m/z* 900-1500 to allow detection of a wider range of components. Due to the nature of precursor selection made by the software on the ion trap instruments (typically set to pick the base peak ion in MS), where lipids are observed to co-elute fully, only the major lipid is typically selected for CID. Where acquisition of the MS/MS spectra of minor lipids was required, the lipid extract was re-analysed using a narrow mass spectral scan range, truncated to avoid the *m/z* value of the major species, allowing the minor species to be detected as the base peak ion and, consequently, selected for CID. In one case where a single narrow scan range was inadequate, precursor selection of the minor components was best achieved using a variable scan range, manually altered throughout the repeat analysis.

Some of the extracts were also analysed using a further extension of the mass spectral scan range (*m/z* 50-2000) in order to screen for the presence of diether lipid cores and/or components with *m/z* > 1500.

3.2.2.2. Eluent polarity and monitoring of column backflush

LC-MS analyses described for tetraether lipid cores extracted from archaea and from aquatic sediments, with separation performed on an amino column chemistry, have typically used an elution program which reaches a maximum elutropic strength at a mobile phase composition of <2% IPA (results in Chapter 2 and Hopmans *et al*., 2000; Schouten *et al*., 2000; Pearson *et al*., 2004; Schouten *et al*., 2007b). Although use of a less retentive cyano column chemistry in tandem with smaller, more powerfully-resolving column particle size has facilitated earlier elution of tetraether core lipids (Schouten *et al*., 2007b), a low maximum solvent polarity (1.8% IPA)

was still retained. In addition to a move to cyano column chemistry, two studies have also increased the final percentage IPA in the mobile phase to 10%, but this modification was only implemented in order to clean the column prior to subsequent analyses (Leininger *et al*., 2006; Reigstad *et al*., 2008). Other studies have used a more polar mobile phase modifier, *n*-propanol, in place of IPA, though the inclusion of this solvent has usually been to reduce retention times and increase sample throughput (Escala *et al*., 2007; 2009).

In order to investigate the expression of more polar components in archaeal organisms, LC-MS/MS method B (in which an additional polar flush to 15% IPA at 90 min was incorporated onto the end of LC-MS/MS method A) was used for analysis of the tetraether lipid core material extracted from each of the thermophilic organisms described in this chapter. Furthermore, in order to probe whether any potential tetraether components were retained on the analytical column following the analyses, MS spectra (*m/z* 50-2000) were recorded during the column backflushing procedure used to elute highly polar components from the column post-analysis.

3.2.2.3. Structural assignment

MS and MS/MS spectra for the tetraether cores in the lipid extracts from the thermophiles were typically recorded using both HCT and LCQ instruments. On account of the greater prominence of some of the ions of characteristic value in the HCT MS/MS spectra than in the LCQ spectra, spectra shown are typically those generated on the HCT instrument. Nevertheless, ion chromatograms for the extracts, were qualitatively similar on both instruments. The higher ion capacity of the HCT allows more ions to be trapped per MS scan and may explain why additional minor components were observed in some of the analyses performed on this instrument that were not identified in the analyses performed on the LCQ ion trap. The tetraether lipid cores described all exhibit characteristic product ions in the MS/MS spectra arising from sequential losses from the protonated molecule of water (-18 daltons; Da), C_3H_4O (-56 Da) and/or alkyl chains (various losses) as dienes, with other neutral losses specific to particular lipids also being observed. The tetraether lipid core type was determined from the number of distinct MS/MS spectral regions in

which product ions could be observed, as described in Chapter 2. Specific structural assignments for both known and novel lipids were determined on the basis of their retention times relative to lipid **2**, the *m/z* value of the $[M+H]^+$ in the MS spectrum and from characteristic product ions observed in the MS/MS spectra.

3.2.2.4. Nomenclature

Variants of GDGT and GMGT lipids in which an isoprenoid chain has been methylated are referred to as $GDGT_M$ and $GMGT_M$ lipid cores, respectively. Lipids designated GDGT' or GDGT'' represent earlier or later eluting isobars of known GDGT lipids, respectively. Lipids designated GDGT* denote isobars that co-elute with known GDGT lipids under the chromatographic conditions employed.

Alkyl groups are denoted $[N_P]$ throughout, where N is the number of carbon atoms within the hydrocarbon chain and P the number of cyclopentyl rings encapsulated within the chain. Losses of the chains as dienes in MS/MS are denoted $[N_P]^L$, where L represents the formal number of molecules of water lost simultaneously during formation of the product ion under discussion. Complete tetraether lipid cores are denoted GDGT$[N(1)_{P(1)}, N(2)_{P(2)}]$, where N(1) and N(2) are the number of carbons and P(1) and P(2) the number of cyclopentyl rings in the first and second alkyl chains within the structure, respectively. The orientation of these chains within the tetraether structures can not be determined by LC-MS/MS analysis. As such, full lipid core structures shown are either based on literature precedent or are arbitrary choices, provided in the interests of discussion only.

3.2.3. *Methanothermobacter thermautotrophicus*

3.2.3.1. Background

Methanobacterium thermoautotrophicus, a member of the *Euryarchaeota*, was first isolated from fermenting sludge taken from a sewage treatment plant (Zeikus and Wolfe, 1972) and has more recently been reclassified as *Methanothermobacter thermautotrophicus* strain ΔH (MTH(ΔH); Wasserfallen *et al*., 2000), with the full

genome sequence of the organism also now available (Smith *et al*., 1997). The organism is a strictly anaerobic, autotrophic methanogen which grows on hydrogen and carbon dioxide over the approximate temperature range 40-75^{o}C (optimally between 65-70^{o}C) and optimally at near neutral pHs of 7.2-7.8 (Zeikus and Wolfe, 1972). A cursory investigation of the core lipids of MTH(ΔH) was provided in Chapter 2. In this chapter, a more detailed lipid profile is elaborated upon, including the revelation of a host of minor components of potential taxonomic value which were either not detected using the methodology outlined in Chapter 2 or, otherwise, were not discussed.

3.2.3.2. Ion chromatograms

LC-MS analysis of an extract of methanolysed MTH(ΔH) cells grown at 70^{o}C was described in brief in Section 2.2.4. The base peak chromatogram generated for this extract using the extended method outlined in Section 3.2.2.1 (i.e. 90 min, *m/z* 900-1500) gave similar results, with a peak for GDGT lipid **2** observed to dominate and GTGT **3** and GMGT **27** also identified (Fig. 3.1a). A number of other minor components were observed in the base peak chromatogram, including a late eluting species **58** (*m/z* 1245.8), which only eluted once the percentage of IPA in the mobile phase was significantly greater than 2%. As was noted in Section 2.2.4.1, GDGT **2** was accompanied by a partially co-eluting lipid, **26** (*m/z* 1315.8), the presence of this lipid being more clearly highlighted in a reconstructed ion chromatogram for the range *m/z* 1310-1500 (Fig. 3.1b). As the signal for **2** is absent from this ion chromatogram, many of the minor components also observed in the base peak chromatogram are more easily visualised, including lipids **59** (*m/z* 1343.8), **60** (*m/z* 1381.7) and **61** (*m/z* 1329.7). Interestingly, identification of component **62**, which has an *m/z* 1313.8 and fully co-elutes with GMGT **27**, explains the origin of an ion at approximate *m/z* 1314 in the mass spectrum of **27** described in Section 2.2.4.1.2. An additional peak (**X**; *m/z* 1325.2), relating to an early eluting species, was also observed in the ion chromatogram. This component was determined to be an impurity found in the hexane solvent used during the extraction and purification procedures (see Section 7.1.1). In order to screen MTH(ΔH) for the presence of diether lipid cores, the extract was also analysed using a mass spectral scan range of

m/z 50-2000. The base peak chromatogram (not shown) was complicated by the presence of peaks relating to components which are not ether lipids. Notably, however, a reconstructed ion chromatogram for the range *m/z* 653-654, allowed an early eluting component with $[M+H]^+$ at *m/z* 653.5, consistent with that of lipid **1**, to be identified (Fig. 3.1c).

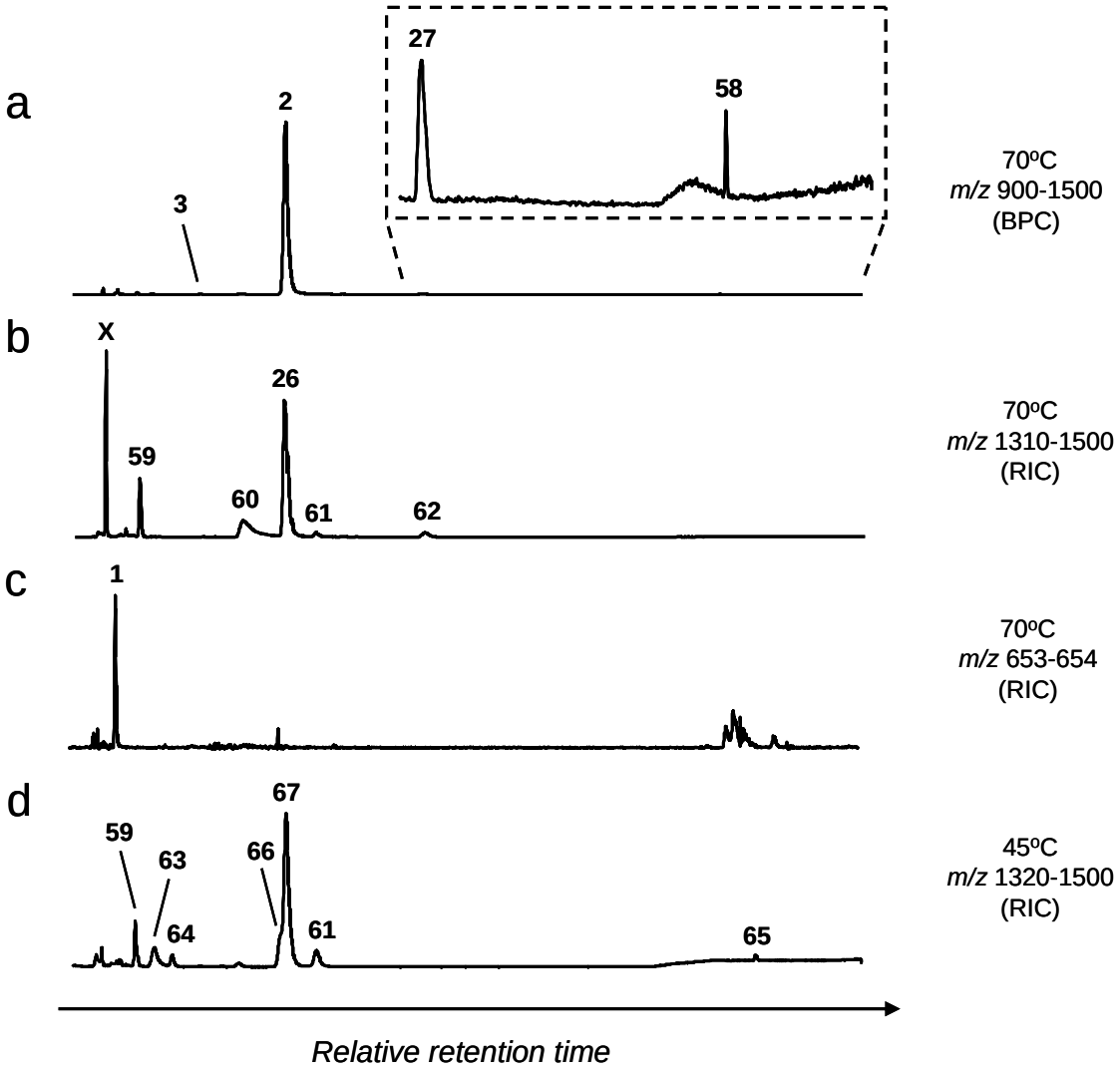


Fig. 3.1. HCT LC-MS ion chromatograms for lipid extracts from MTH(ΔH) 70°C (a-c) and 45°C (d) samples: a) base peak chromatogram (BPC) during full scan analysis (*m/z* 900-1500); b) reconstructed ion chromatogram (RIC) for *m/z* 1310-1500 during full scan analysis (*m/z* 900-1500); c) RIC (*m/z* 653-654) during extended scan analysis (*m/z* 50-2000); d) RIC (*m/z* 1320-1500) during full scan analysis (*m/z* 900-1500).

The base peak chromatogram of a lipid extract of MTH(ΔH) grown at 45^{o}C (not shown) was qualitatively similar to that for the 70^{o}C extract, with **2** being the dominant lipid and many of the same minor components observed, albeit in different proportions relative to **2**. Lipid **60** was not detected in the 45^{o}C sample, although additional components were observed that were absent from the chromatogram for the 70^{o}C sample. The additional components, including early eluting lipids **63** (*m/z* 1441.8) and **64** (*m/z* 1415.8), and late eluting lipid **65** (*m/z* 1337.7), are most easily visualised in an ion chromatogram generated for the mass spectral range *m/z* 1320-1500 (Fig. 3.1d). Lipid **66** (*m/z* 1329.8), which partially co-elutes with lipids **2** and **26**, and lipid **67** (*m/z* 1384.8), which fully co-elutes with these components, were also observed.

The MS spectra recorded during the column backflushing procedure implemented following LC-MS/MS analysis of both the 45^{o}C and 70^{o}C MTH(ΔH) extracts exhibit no peaks that could potentially relate to tetraether lipid cores.

3.2.3.3. Structural assignments

The HCT MS/MS spectra for lipids **2-3** and **27** were described in Section 2.2.5 and are consistent with the known structures of these lipids (De Rosa and Gambacorta, 1988a; Morii *et al*., 1998). On the basis of the MS/MS spectra obtained for **58-60** (Table 3.2), these appear to be structurally modified variants of GDGT **2**. The $[M+H]^+$ of lipid **59** loses water (*m/z* 1325), $C_3H_6O_2$ (*m/z* 1307), ketene (*m/z* 1301), acetic acid (*m/z* 1283) and 1-en-biphytan-32-ol (*m/z* 767) during CID, generating product ions in two regions of the MS/MS spectrum. The neutral losses observed from **59** are comparable to those observed following CID of acetylated tetraether lipid **56** (see Section 2.2.10.3). In addition, the +42 *m/z* unit increase in the $[M+H]^+$ of **59** compared to that of **2**, coupled with the earlier elution of the heavier lipid, suggest that **59** is the monoacetylated derivative of **2**. This was confirmed following direct infusion MS/MS (LCQ) of a synthetic monoacetylated derivative of **2** (measured *m/z* 1344.0), formed following an incomplete acetylation reaction using pyridine:acetic anhydride. The product ions observed for the synthetic acetylate matched those observed from **59**. Lipid **60** also appears to be derived from a GDGT

lipid core on the basis of product ions appearing in two regions of the MS/MS spectrum. They include ions formed from loss of water (*m/z* 1363), $C_3H_6O_2$ (*m/z* 1307) and 1,31-biphytadiene (*m/z* 823). Product ions at *m/z* 1283 and *m/z* 725 in the MS/MS spectrum of **60** are also observed in the MS/MS spectrum of **2**, where they are formed *via* loss of a hydroxyl group, independent of or simultaneous to loss of an alkyl chain, respectively. These ions suggest that the structure of lipid **60** closely resembles that of **2**, differing only in the nature of the terminal hydroxyl group, which bears an 81 Da group in place of an alcoholic hydrogen found in the latter structure. A possible candidate structure for **60** is as the phosphatidic acid derivative of **2**, a known phosphatidyl lipid component of MTH(ΔH) (Nishihara and Koga, 1990) which has a $[M+H]^+$ *m/z* ratio consistent with that observed in the base peak chromatogram. Notably, however, previous LC-MS analysis of the lipids of *Methanocaldococcus jannaschii*, indicated that the phosphatidic acid derivative of **1** elutes later than **1** under an alternative normal-phase chromatographic separation, suggesting that addition of a phosphate group increases lipid polarity (Sturt *et al*., 2004). This appears to be inconsistent with the elution of **60** earlier than **2** in the LC-MS/MS analyses of MTH(ΔH). As such, the nature of the structure of **60** remains unclear. Discussion of the likely nature of lipid **58** is reserved for Section 5.2.3.2.

Table 3.2. MS and MS/MS spectral data (HCT) for minor lipid components identified in lipid extracts from MTH(ΔH) grown at 45°C and 70°C.

Lipid	$[M+H]^+$	Product ions observed in MS/MS [a]
58	1245.8	1227,1209; **669**,651,633,615,613,557
59	1343.8	**1325**,1307,1301,1283,1269,1265,1251,1247,1227,1209; 767,749,731,725,707,693,691,689,675,671,669,651,647,633,615,613,597,575,557
60	1381.7	1363,1345,1325,1323,1307,1305,1283,1281,1265,1263,1247,1245,1227,1209,1207,1205; 823,805,791,789,787,771,769,749,747,**731**,729,725,713,707,705,689,687,675,671,669,657, 653,651,635,633,631,629,627,615,613,597,595,575,557
61	1329.7	**1311**,1293,1275,1255,1237,1219; 753,743,735,725,717,707,699,697,689,679,671,661,659,643, 641,625,623,615,613,605,603,601,587,585,583,569,567; 381
63	1441.8	1423,**1385**,1367,1349,1227,1209; 883,865,847,809,791,735,671,669,651,633,615,613,597,575,557
64	1415.8	**1397**,1383,1379,1365,1341,1327,1301,1283,1265,1247,1227; 839,821,807,789,765,763,747,733,725,707,689,677,671,659,615,597,575,557
65	1337.7	**1319**,1301,1277,1263,1245; 743,725,707,687,669,651,631,615
66/66*	1329.8	**1311**,1293,1275,1273,1255,1237; 771,757,753,739,735,723,721,703,701,697,683,665,661,651, 647,645,643,641,629,627,615,611,597,589,587,585
67	1384.8	1365,1341,1325,1315,**1301**,1283,1265,1247,1227,1209; 743,725,709,707,689,687,675,671,669,653,651,647,633,631

[a] Product ions listed in bold face represent the base peak ion in MS/MS.

Re-analysis of the MTH(ΔH) 70°C extract using the MS scan range *m/z* 1310-1340 allowed the signals for lipids **2** and **27** to be removed in MS, facilitating selection of lipids **26** and **62** for CID. The HCT MS/MS spectrum of lipid **26** (Fig. 3.2a) is consistent with that recorded on the LCQ instrument (Section 2.2.4.2.5). Product ions arising from losses of water (*m/z* 1297), C_3H_4O (*m/z* 1259) and concerted loss of both (*m/z* 1241) were observed, confirming a tetraether structure for the lipid. In addition, product ions were observed in two distinct regions of the MS/MS spectrum, confirming the lipid to have a dialkyl-type core structure. Product ions at *m/z* 757 (-558 Da) and *m/z* 743 (-572 Da), corresponding to losses of an acyclic $C_{40}H_{78}$ diene (i.e. $[40_0]^0$) or an acyclic $C_{41}H_{80}$ diene (i.e. $[41_0]^0$), were also observed in the MS/MS spectrum of **26**. The 14 *m/z* unit greater $[M+H]^+$ and marginally earlier retention time

observed for **26** relative to **2** during LC-MS are both consistent with a chain-methylated structural variant of caldarchaeol. Consequently, **26** is confirmed as $GDGT_M$ homocaldarchaeol, as noted previously in *M. marburgensis* (Galliker *et al*., 1998). Lipid **62** fully co-elutes with GMGT **27** during LC-MS and has an $[M+H]^+$ of 14 *m/z* units greater. The MS/MS spectrum of **62** (Fig. 3.2b) contains product ions in only one spectral region confirming that it has only a single alkyl chain within its structure. The product ions include losses of 1-3 molecules of water and/or C_3H_4O, with each ion shifted +14 *m/z* units relative to an ion in the MS/MS spectrum of **27**. On the basis of these observations it is apparent that **27** and **62** represent an homologous pair. Although MS/MS analysis is not able to fully confirm the nature of the additional carbon atom in **62**, it seems likely that the homologation between **27** and **62** would mirror that observed between dialkyl tetraether lipids **2** and **26**, with the extra carbon in **62** present in the form of a methyl group attached to the C-13 position on one of two biphytanyl chains. Interestingly, a series of C_{81} tetraacids, isolated from crude oil naphthenate deposits and believed to be of archaeal origin, show similar C-13 position homologation in relation to C_{80} tetraacids bearing 5-8 Cp rings, with the rings located in the same positions in the higher homologues (Lutnaes *et al*., 2007). As such, **62** is tentatively assigned as H-shaped homocaldarchaeol, a $GMGT_M$ lipid containing a C-13 methyl group as a modification to the basic isoprenoid structure of **27**. On the basis of its MS/MS spectrum (Table 3.2), lipid **61** most likely has a GDGT structure and contains at least one biphytadianyl chain, evidenced by a product ion formed *via* loss of 1-en-biphytan-32-ol (*m/z* 753). Other product ions formed from this lipid are more difficult to assign and, consequently, the structure cannot be elucidated further at the current time.

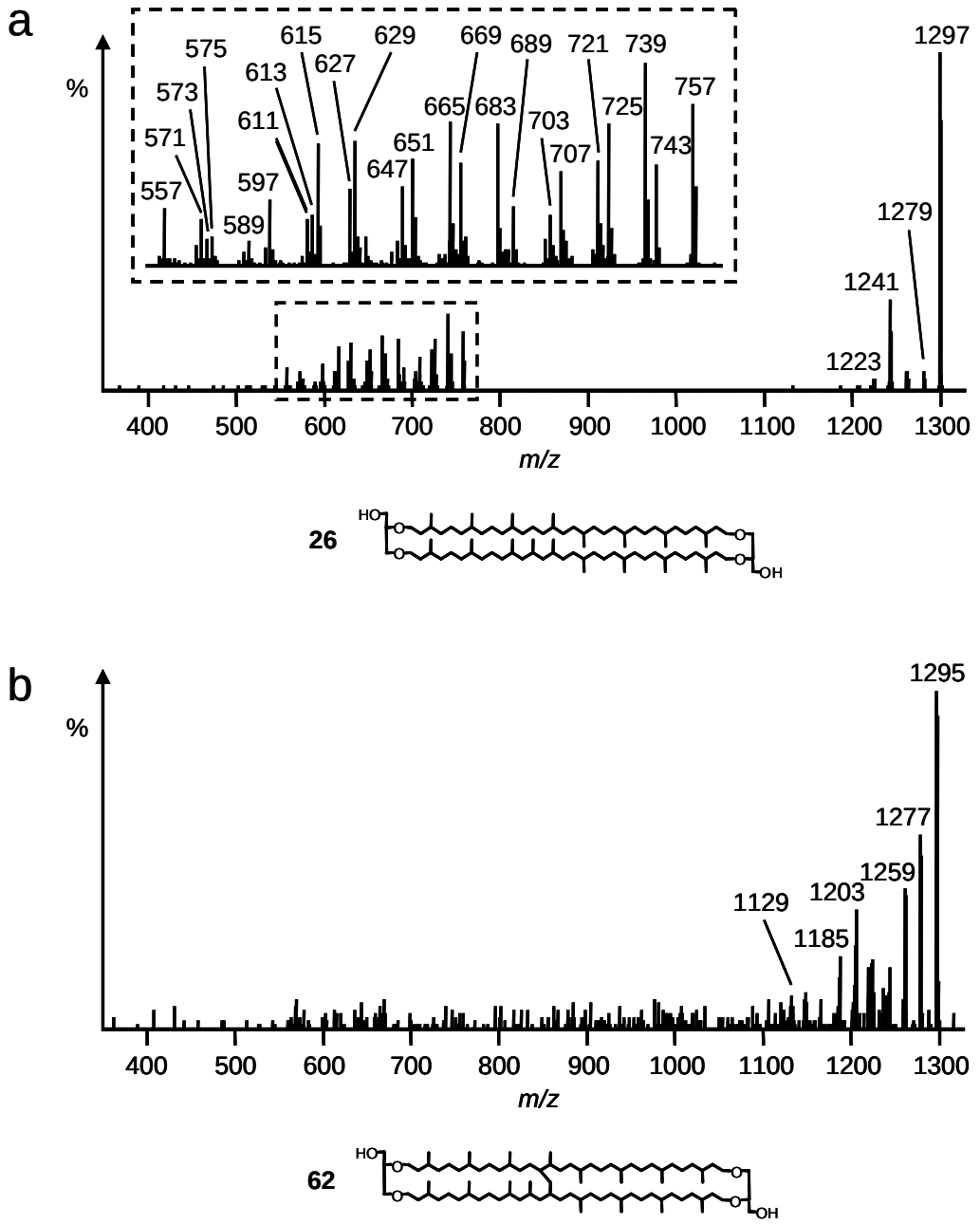


Fig. 3.2. MS/MS spectra (HCT) of: a) $GDGT_M$ **26** ($[M+H]^+$ *m/z* = 1315.8). Inset shows expansion of a region of the spectrum; b) $GMGT_M$ **62** ($[M+H]^+$ *m/z* = 1313.8). As the position of the covalent link between the isoprenoid chains and placement of the additional methyl group are not known, the structure shown for **62** is based upon isoprenoid lipids containing C_{41}, C_{80} and C_{81} hydrocarbons reported previously (Morii *et al*., 1998; Galliker *et al*., 1998; Lutnaes *et al*., 2007), but remains tentative nevertheless.

In order to generate MS/MS data for some of the components found in the MTH(ΔH) 45°C extract, the sample was re-analysed using a mass spectral scan range *m/z* 1320-1500, which removed signals for lipids **2** and **26** in MS and allowed for CID of co-eluting components **66** and **67** along with several other components. On the basis of MS/MS evidence, lipid **66** is a GDGT lipid containing two unmodified glycerol groups (Table 3.2). Product ions observed at *m/z* 771 and *m/z* 757 in the MS/MS spectrum could potentially be formed by loss of $[40_0]^0$ and $[41_0]^0$, respectively. This combination of alkyl chains within an individual tetraether structure is not consistent with the *m/z* of the $[M+H]^+$ of **66**, which is 28 *m/z* units greater than that of **2**, suggestive of a structure that contains two additional carbon atoms. A low intensity ion at *m/z* 743, possibly relating to loss of $[42_0]^0$ from the precursor, was observed in the spectrum but was too weak to provide conclusive assignment. As such, it is possible that **66** (denoted GDGT[41_0,41_0]) co-elutes with an isomeric form, **66*** (GDGT[40_0,42_0]), with loss of [42_0] from the latter during CID being a disfavoured process. Further work, possibly using a more concentrated sample of the lipid extract, is required in order to test this assertion. Whilst it is possible to assign lipids **63**, **64**, **65** and **67** as dialkyl tetraether lipids, the nature of the full structures of these lipids cannot be assigned at the present time. Nevertheless, on the basis of the products ions observed during CID of **63** (Table 3.2), it is most likely to be a structural variant of **2** in which one of the terminal glycerols has been modified. Similarly, lipid **64** loses methanol (*m/z* 1383) during CID (Table 3.2), suggesting that it may be a tetraether core which has been methyl etherified at one or both ends, structural modifications which would also account for its early elution. Lipid **67** may be a variant of **2** in which one of the hydrogen atoms within a terminal glycerol group is replaced by an 83 Da moiety which is simultaneously lost when an alkyl chain dissociates from the structure.

In the analysis performed using the scan range *m/z* 50-2000, the *m/z* value of the $[M+H]^+$ of the early eluting lipid, assigned as **1** during LC-MS, is consistent with the molecular formula of archaeol ($C_{43}H_{88}O_3$). Furthermore, the observed retention time of the component relative to that of **2** is also consistent with that observed for **1** in previous LC-MS analyses of samples which contained both of these lipids (Mancuso *et al*., 1986; Turich *et al*., 2007; Reigstad *et al*., 2008; Weijers *et al*., 2009). The

identity of this component was confirmed as GDD **1** following CID. The resulting MS/MS spectrum (Fig. 3.3) is dominated by a product ion at *m/z* 373 arising from loss of phytene ($C_{20}H_{40}$) from the precursor with proton transfer back to the charge-retaining species. This matches the observed loss of 280 Da in previous MS/MS spectra generated from **1** (Turich *et al*., 2007; Gattinger *et al*., 2003; de Souza *et al*., 2009) and to losses in the MS (Hopmans *et al*., 2000) and MS/MS (Section 2.2.5.3) spectra of GTGT **3**, which also contains ether-bound phytanyl hydrocarbon chains.

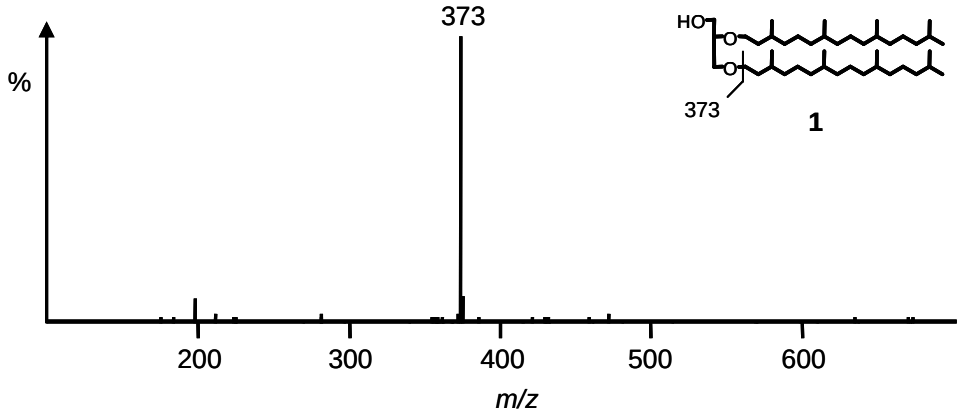


Fig. 3.3. MS/MS spectrum (HCT) of GDD **1** ($[M+H]^+$ *m/z* = 653.5), with neutral loss of 1-phytene (280 Da) illustrated.

3.2.3.4. PAN LC-MS/MS

By using a restricted *m/z* scan range (*m/z* 1310-1340), which removed the signal for co-eluting GDGT **2**, PAN LC-MS/MS of the MTH(ΔH) 45°C extract allowed a PAN MS/MS spectrum of $GDGT_M$ **26** to be obtained. The *m/z* 100-400 region of the spectrum (Fig. 3.4) is remarkably similar to the equivalent region of the PAN MS/MS spectrum of lipid **2** (Fig. 2.9b), with all of the product ions formed from the $[M+H]^+$ of **2** also being observed from that of **26**. Although some of the ions are more prominent in the PAN MS/MS spectrum recorded for **26** than they are in the spectrum for **2**, the relative proportions of product ions in **2** were observed to vary between analytical replicates (see Section 2.2.8.3), suggesting that the differences between the spectra of **26** and **2** would also be subject to variation in repeat analyses. Consequently, the product ions at *m/z* < 400 are of little characteristic value in differentiating **26** from **2**.

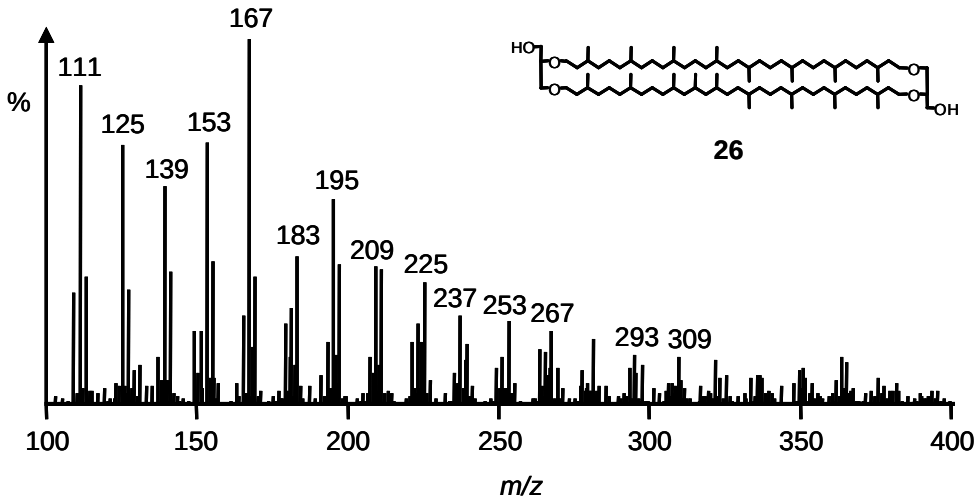


Fig. 3.4. *m/z* 100-400 region of the PAN MS/MS spectrum (HCT) of $GDGT_M$ **26**.

3.2.3.5. GC and GC-MS of isoprenoid hydrocarbons

Reductive ether cleavage of the core lipids of archaea, followed by GC and GC-MS analysis of liberated isoprenoid hydrocarbons has proven to be a useful method for confirming the structure of diether and tetraether lipids of archaeal origin (De Rosa and Gambacorta, 1988a; Hoefs *et al.*, 1997; DeLong *et al.*, 1998; Schouten *et al.*, 1998; Burhan *et al.*, 2002; Blumenberg *et al.*, 2007). To attempt to support the assignment of $GDGT_M$ **26**, lipid extracts from MTH(ΔH) cellular material, grown at both 45°C and 70°C, were treated with HI to generate alkyl iodides and diiodides, which were subsequently reduced to the hydrocarbons using either $LiAlH_4$ or zinc powder in acetic acid. The GC chromatograms for the hydrocarbon fraction from each extract (Fig. 3.5) show three major peaks, eluting at 11.7, 52.3 and 53.2 min. An additional major peak at 9.6 min, and several minor peaks eluting between 10-50 min, were also observed in the chromatogram of the hydrocarbons generated from the MTH(ΔH) 45°C extract (Fig. 3.5a). Similarly, an additional minor component which partially elutes with that at 11.7 min was noted in the chromatogram for the MTH(ΔH) 70°C extract (Fig. 3.5b).

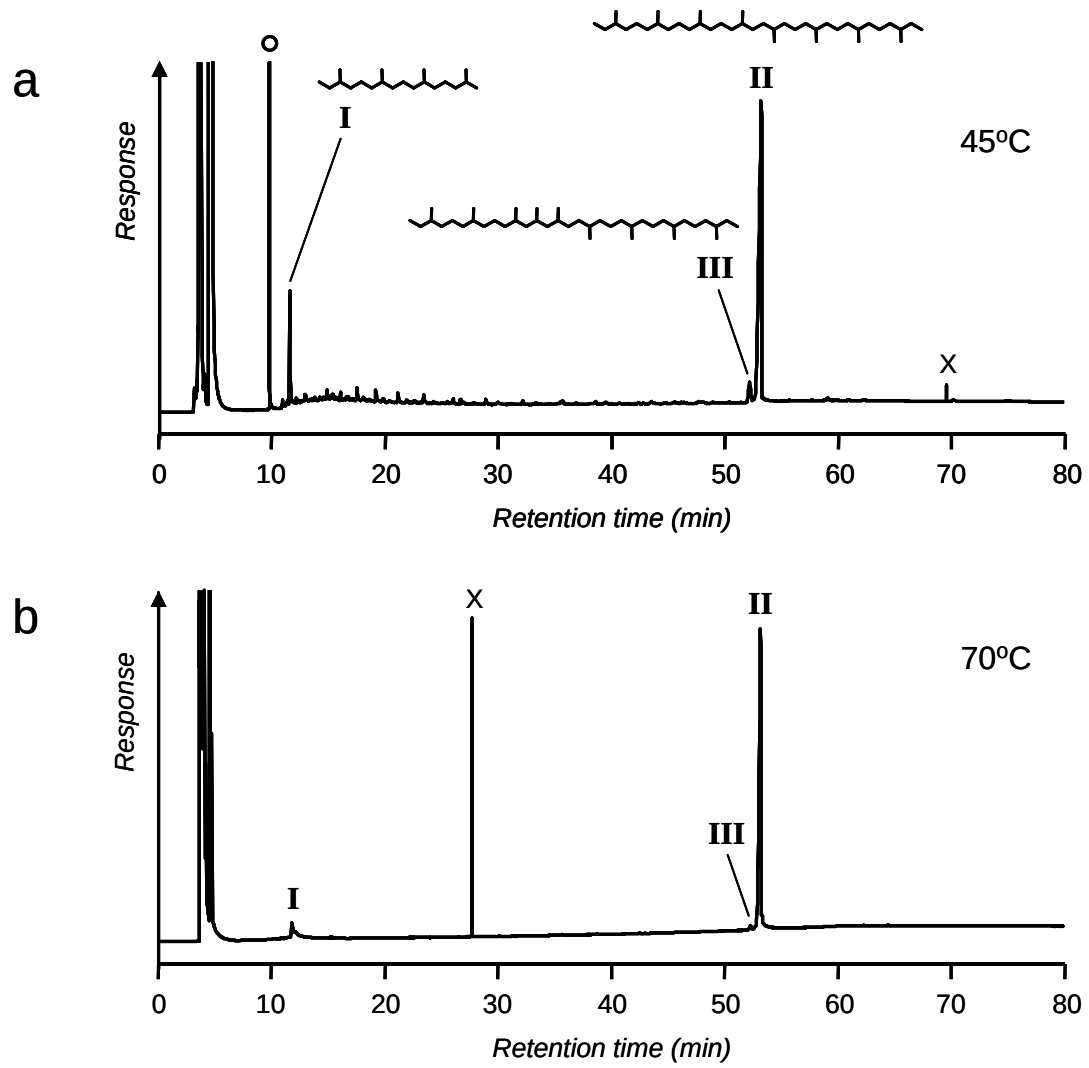


Fig. 3.5. GC-FID traces for hydrocarbons liberated following ether cleavage (HI-$LiAlH_4$) of lipid extracts from MTH(ΔH) cells, grown at: a) 45°C; b) 70°C. The peak labelled with an open circle is butylated hydroxytoluene. Peaks labelled with a cross represent electrical spikes/noise.

In order to identify the major components, the hydrocarbon fraction prepared from the MTH(ΔH) 45°C extract was also analysed by GC-MS using 70 eV EI. The component eluting at 9.6 min was identified as butylated hydroxytoluene, a known anti-oxidising stabiliser used in plastics and rubbers (Fries and Puttmann, 2002) and in organic solvents (Middleditch, 1989). Its presence in the extracts is attributed to contaminantion during preparation of the hydrocarbons. The component with retention time at 11.7 min (**I**) generated an EI spectrum which exhibited an $M^{+\bullet}$ at *m/z* 282 and prominent fragment ions at *m/z* 112/113, 126/127, 182/183, 196/197 and

252/253 (Fig. 3.6a). The molecular and fragment ions observed are consistent with those observed previously for phytane (De Rosa and Gambacorta, 1988a), the C_{20} isoprenoid chain found in GDD **1** and GTGT **3**. The $M^{+\bullet}$ of the component with retention time at 53.2 min (**II**) appears to be absent, making assignment of the hydrocarbon difficult. Prominent fragment ions were observed at *m/z* 126/127, 196/197, 266/267, 280/281, 322/323, 392/393 and 462/463 (Fig. 3.6b). Previous reports of the EI spectrum of biphytane showed identical prominent fragment ions (De Rosa *et al*., 1977b; 1988a; DeLong *et al*., 1998; Blumenberg *et al*., 2007; Ventura *et al*., 2008), although an $M^{+\bullet}$ at *m/z* 562 was observed in some cases (De Rosa *et al*., 1977b; 1988a; Blumenberg *et al*., 2007; Ventura *et al*., 2008). On the basis of the mass spectral evidence and the dominance in the lipid core extract of GDGT **2**, which contains two biphytanyl chains, component **II** is assigned as biphytane. The EI spectrum of the component with retention time at 52.3 min (**III**) was also apparently devoid of an $M^{+\bullet}$ ion, but showed prominent fragment ions at *m/z* 238/239, 336/337 and 364/365 (Fig. 3.6c) as were noted in the EI spectrum of 13-methylbiphytane, as reported previously (Blumenberg *et al*., 2007 and references therein). Additional prominent fragment ions, observed at *m/z* 126/127, 196/197, 280/281, 322/323 and 406/407 are also consistent with the structure of 13-methylbiphytane, in support of the assignment of **III** as this hydrocarbon. The earlier elution of **III** relative to **II** on the 100% dimethyl-polysiloxane column is unusual given the additional carbon atom in the 13-methylbiphytane, especially considering 11-methylbiphytane eluted later than **II** during GC analysis (5% phenyl-arylene column chemistry) of the hydrocarbons generated following ether cleavage of lipids extracted from a hydrothermal field sulfide located in the Mid-Atlantic Ridge (Blumenberg *et al*., 2007). Notably, however, a methylated isoprenoid containing two cyclopentyl rings was observed to elute earlier than non-methylated isoprenoid chains containing the same number of rings during GC analyses (5% diphenyl column chemistry) of HI-$LiAlH_4$ treated lipid extracts from the Be'eri sulfur deposit (Burhan *et al*., 2002). The additional methyl in this case was identified to be located somewhere between the C-11 and C-22 positions of the isoprenoid chain.

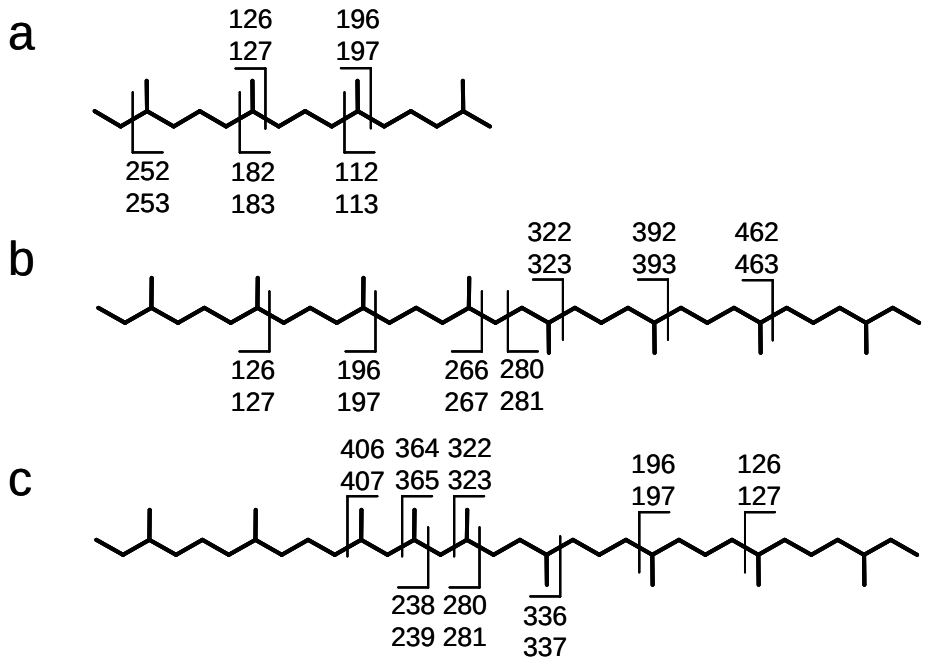


Fig. 3.6. Prominent ions in the EI mass spectra of hydrocarbons liberated following ether cleavage (HI-$LiAlH_4$) of lipid extracts from MTH(ΔH) cells grown at 45°C and their associated fragmentation: a) **I**, phytane; b) **II**, biphytane; c) **III**, 13-methylbiphytane.

3.2.3.6. Distribution and significance

In addition to known isoprenoid lipids **1**-**3** and **27**, several other components were identified in MTH(ΔH) including two lipids, **26** and **62**, which were not evident in a conventional base peak chromatogram. Significantly, both of these species were concealed by co-elution with a known tetraether lipid exhibiting an $[M+H]^+$ at 14 *m/z* units less. Lipid **26** partially co-elutes with **2** and lipid **62** fully co-elutes with **27**, suggesting that the concealed lipids may be homologues of caldarchaeol and H-shaped caldarchaeol, respectively. Lipid **26** was confirmed, by LC-MS/MS, as homocaldarchaeol, a C-13 position chain-methylated variant of **2**, fully characterised previously in *M. marburgensis* (Galliker *et al*., 1998). The position of the extra carbon in **26** is supported by the EI spectrum obtained following GC-MS of hydrocarbons formed by ether cleavage of total lipid extracts. Although the most logical position for methylation in **62** would mirror that observed in **26** (i.e. at the C-13 chain position), it should be noted that the placement of the methyl group in the

tetraether cannot be determined using the current MS/MS method. As such, the precise structure of **62** needs to be verified.

The abundance of **26** relative to **2** in MTH(ΔH) grown at 70^{o}C, as calculated from peak areas in the HCT LC-MS ion chromatograms is 2.3% (Table 3.3). This value is marginally smaller than that calculated from the LCQ LC-MS data for this sample (see Section 2.2.4.1.2), indicating that there may be a small degree of variability between relative abundances determined using the two instruments. Interestingly, the ratio of **62** to **27** calculated using the HCT data is 17.9% (Table 3.3). The substantial differences in the relative proportions of these lipids, coupled to the absence of methylated derivatives of other lipids of MTH(ΔH), suggest that the methylation is not coincidental nor does it occur *via* uncontrolled chemical conversion. It is likely that the methylation step is effected enzymatically by one or more methyl transferase enzymes, which show apparent selectivity for methylation of GMGT lipids over GDGT lipids, with methionine implicated previously as the most likely methyl source (Galliker *et al*., 1998). Enzymatic C-methylation is an important process in the biosynthesis of phytosterol lipids in plants and algae (Nes *et al*., 2000; 2003), although in these molecules, the methylation occurs at unsaturated carbon centres. Despite debate as to whether GDGT **2** forms *via* coupling of saturated (Nemoto *et al*., 2003) or unsaturated (Eguchi *et al*., 2000; 2003) GDD lipid cores (see Section 1.2.1), none of the studies into tetraether lipid biosynthesis undertaken to date have suggested that the C-13 positions in the isoprenoid chains of **2**, or in proposed diether precursors, are unsaturated at any stage during the biosynthesis (Koga and Morii, 2007). Furthermore, it has been noted that labelled biosynthetic precursors in which a C-14 to C-15 double bond is absent, are still incorporated into **2** and **26** when added into *M. marburgensis* fermentation broth (Galliker *et al*., 1998). This indicates the coupling of saturated diether lipids during the formation of tetraether lipid cores and would suggest that the key methylation step in the formation of **26** occurs at a fully saturated, non-allylic carbon. The absence of a higher homologue of GDD **1** in *M. marburgensis* indicated that the additional methyl in **26** is incorporated either during or after the coupling of the diether lipids in this archaeon (Galliker *et al*., 1998). Similarly, a higher homologue of GDD **1** could not be observed in the LC-MS analyses of MTH(ΔH) extracts grown at either temperature in the present study, fully supporting a late-stage methylation. In addition, no higher homologue of

GTGT **3** could be detected in the base peak chromatogram for extracts of cellular material grown at either temperature. The absence of the higher homologues of GDD and GTGT core lipids in MTH(ΔH) is unlikely to be an effect relating to the low natural abundances of **1** and **3** as these lipids occur in similar abundance to lipid **27**, a higher homologue of which (i.e. **62**) was observed. GTGT **3** is most likely produced when the coupling of diether lipids does not proceed to completion (see Section 1.2.1 and de la Torre *et al*., 2008), with only the first pair of phytanyl chains linked. Consequently, if the methylation were to occur during the cross-coupling process, a chain-methylated GTGT lipid homologue would also be expected. Since a lipid of this nature is not detected in MTH(ΔH), it would appear that the methylation step occurs after macrocycle formation.

Table 3.3. Proportions of some of the lipid cores found in MTH(ΔH), calculated from HCT LC-MS peak areas and, unless otherwise stated, quoted as a percentage relative to GDGT **2**.

Lipid	Growth temperature	
	45°C	70°C [a]
GTGT **3**	0.47	0.32 ± 0.01 [b]
$GDGT_M$ **26**	9.58	2.33 ± 0.01
GMGT **27**	0.04	0.65 ± 0.04
$GMGT_M$ **62**	0.02	0.12 ± 0.01
% Ratio **62**:**27**	55.9	17.9 ± 0.3

[a] Calculated as the average of analytical duplicates.

[b] Errors represent ±1σ from the mean value.

Galliker *et al.* (1998) reported that formation of lipid **26** was stimulated by the addition of a detergent (Tween-20), used to enable uptake of large biosynthetic precursors during *in vivo* experiments. Since no such detergent was present during growth of the culture employed in the present study, this cannot be the sole explanation for the formation of methylated tetraether lipids by archaea. It remains likely, however, that cellular stress plays a role in the degree of tetraether methylation expressed in archaeal lipid membranes, particularly as the proportion of **26** relative to **2** expressed by MTH(ΔH) when grown at 45^{o}C (9.6%) appears to be greater than when the organism is grown at 70^{o}C (Table 3.3). This apparent increase is supported by integration of peak areas in the GC chromatogram of the HI-$LiAlH_4$ liberated hydrocarbons, with the proportion of **III** relative to **II** (calculated assuming identical FID response per mole for each hydrocarbon) found to be greater for the lipid extract from MTH(ΔH) grown at 45^{o}C (4.6%) than for the extract from cells grown at 70^{o}C (1.1%). Interestingly, distinct changes are observed in cell morphology when MTH(ΔH) is grown at the extremes of its viable temperature range as opposed to within its optimal range of 65-70^{o}C (Zeikus and Wolfe, 1973). In particular, cells have elongated rod shapes when grown optimally, but show folded, corkscrew-like shapes when grown at 45^{o}C or 75^{o}C. Unfortunately, whether an increase in the proportion of **26** and the observed morphological changes are directly related is impossible to confirm on the basis of the limited data obtained for the organism so far.

The ratios of lipids **27** and **62** relative to **2** also appear to be temperature dependent in MTH(ΔH), with more of the monoalkyl cores observed in the 70^{o}C sample compared to in the 45^{o}C sample (Table 3.3). Furthermore, the proportion of phytane relative to biphytane expressed in the lipid cores reduces with increasing growth temperature (see Fig. 3.5), suggesting more substantial conversion of GDD **1** to tetraether lipids at elevated temperature. This agrees favourably with the decrease in the expression of **1** relative to **2** and **4** observed during growth of *M. jannaschii* at increasing temperatures (Sprott *et al.*, 1991), where increases in the levels of the more rigid GMD and GDGT lipid cores were thought to be an adaptation to reduce membrane fluidity (see Section 1.5). It is likely that the similar behaviour in MTH(ΔH), that is to say dimerisation of diether to tetraether lipid cores, under

elevated growth temperature is also a homeoviscous adaptation that restricts membrane fluidity. The formation of a covalent cross-link between isoprenoid chains in dialkyl tetraether lipids to form monoalkyl tetraether lipids also appears to be promoted by increasing ambient growth temperature. It should be noted that in both the 70^{o}C and the 45^{o}C MTH(ΔH) extracts lipids **2** and **26** contribute over 97% to the total tetraether core lipids, other components contributing less than 1% to the total core lipid found in the extracts from cells grown at each temperature. Consequently, the elevation of the levels of lipids **27** and **62** at high temperature is unlikely to effect a significant reduction in overall membrane fluidity.

The nature of the structures of other minor components detected in MTH(ΔH) have been delimited by the MS/MS spectral data obtained in each case, although only lipid **59** has been unambiguously assigned as a monoacetylated derivative of **2**. It is surprising that this lipid survives the methanolysis procedure used during lipid extraction and yet, it is similarly difficult to see how it could be generated artificially during extraction given the solvents and conditions used. If **59** is not an artifact, it must represent a very minor component of the total MTH(ΔH) lipid core pool, suggesting either that it may be an accidental by-product produced during lipid biosynthesis or, alternatively, may fulfil a very specific role within the cellular membrane of MTH(ΔH). The same is to be said of the other minor components detected in MTH(ΔH). Despite being minor contributors to the lipid profile of the organism, these unusual components could potentially serve as far more specific biomarkers for *Methanothermobacter* in environmental samples than the more major components (e.g. **2**), which can originate from a number of archaeal sources.

The core lipids reported for MTH(ΔH) represent structures identified following several screenings *via* LC-MS/MS analyses using a series of different MS scan ranges. These allow determination of some, but not all, of the co-eluting components observed in the base peak chromatograms. As such, other minor contributors to the lipid profile of the archaeon may have been overlooked.

3.2.4. *Sulfolobus* species

3.2.4.1. Background

Within the *Crenarchaeota*, the genus *Sulfolobus* comprises a number of organisms found in the aerobic zones of globally distributed acidic hot springs (Berkner and Lipps, 2008) and their surrounding thermal soils (Brock *et al.*, 1972). The first isolated member of the genus was *Sulfolobus acidocaldarius* (Brock *et al.*, 1972) and subsequently a number of other close genetic relatives have been identified and cultivated. These include *S. solfataricus* (formerly named "*Caldariella acidophila"*, De Rosa *et al.*, 1975; Zillig *et al.*, 1980) and *S. shibatae* (Grogan *et al.*, 1990). All three species discussed are aerobic, facultative autotrophs that grow on elemental sulfur and/or a variety of carbon sources (Brock *et al.*, 1972; De Rosa *et al.*, 1975; Grogan *et al.*, 1990). Furthermore, each organism is (hyper)thermoacidophilic, with optimal growth temperature and pH typically found between 70-87^{o}C and 2.0-5.4, respectively, depending on species and strain (Brock *et al.*, 1972; De Rosa *et al.*, 1975; Zillig *et al.*, 1980; Grogan *et al.*, 1989; 1990). Studies described herein focus on readily available strains *S. acidocaldarius* MR31, *S. solfataricus* P2 and *S. shibatae* B12. There has been some confusion (Zillig, 1993) regarding both the nomenclature used for *Sulfolobus* strains and the deposition of pure cultures of these strains to culture collections (such as the Deutsche Sammlung von Mikroorganismen und Zellkulturen, DSMZ, and the American Type Culture Collection, ATCC). Of the strains used in this study, only the *S. solfataricus* strain has ambiguous descriptors in the literature, where it was initially described as *Caldariella acidophila* strain MT-3 (De Rosa *et al.*, 1975) but is now commonly known as *S. solfataricus* strain P2 (Zillig *et al.*, 1980), available under this nomenclature in both major culture collections (culture numbers DSM1617 and ATCC35092). The complete genome sequence of *S. solfataricus* P2 is available (She *et al.*, 2001), as is the sequence for the type strain of *S. acidocaldarius* (Chen *et al.*, 2005).

Previous reports have suggested that all three *Sulfolobus* species (although not necessarily the same strains as used in the present study) can synthesise GDD **1**, GDGTs **2** and **14-20** and GDCTs **2c** and **14c-20c** (De Rosa *et al.*, 1983b; 1988a;

Hopmans *et al.*, 2000; Sturt *et al.*, 2004; Schouten *et al.*, 2007c). In addition, *S. solfataricus* and *S. acidocaldarius* are known to synthesise GDGT **21** and GDCT **21c** (De Rosa *et al.*, 1980c; 1983b; 1988a; Hopmans *et al.*, 2000). GTGT **3** and isomers of GDGTs **17** and **18** (**17''** and **18''**, respectively) have also been observed as minor components in core lipid extracts of *S. solfataricus* (De Rosa *et al.*, 1983b; Hopmans *et al.*, 2000; Ellen *et al.*, 2009; Pitcher *et al.*, 2010). Hydroxyarchaeol lipid cores have been identified as minor components in *S. acidocaldarius* (Sprott *et al.*, 1997), although whether the cores are *sn*-2 (i.e. **7**) or *sn*-3 (i.e. **8**) chain hydroxylated, or whether both structural types are simultaneously expressed by the organism, has not been determined. The effect of temperature on the number of Cp rings expressed in both GDGT and GDCT core lipids has been described for one strain of *S. solfataricus* (MT-4; De Rosa *et al.*, 1980b), where the average number of rings incorporated into the tetraether core lipids increased with increasing growth temperature. The extensive body of literature describing the lipid cores of these organisms coupled to the complexity of their lipid profiles makes them ideal targets for verification of the LC-MS/MS methodology described in Chapter 2, accounting for their selection in this study.

3.2.4.2. Ion chromatograms

The lipid extracts from the three species of *Sulfolobus* were initially analysed using a mass spectral scan range of *m/z* 900-1500. Among the peaks observed were several exhibiting base peak ions in the MS spectra consistent with the protonated molecules of known isoprenoid lipids **2-3** and **14-21**, most easily visualised in reconstructed ion chromatograms for *m/z* 1280-1320 (Fig. 3.7). The retention times of each component, relative to that of **2**, were similar to those observed in previous LC-MS analyses of the core lipids of *S. solfataricus* MT-4 (Hopmans *et al.*, 2000), allowing GTGT **3** and GDGTs **2** and **14-19** to be identified in the base peak chromatogram of *S. acidocaldarius* (Fig. 3.7a). Each of these lipids was also identified in the base peak chromatogram of the *S. shibatae* extract, along with two additional GDGT cores, **20** and **21** (Fig. 3.7c). Similarly, GDGTs **16-21** were identified in the base peak chromatogram of the *S. solfataricus* extract (Fig. 3.7d). The protonated molecule of each component was selected for CID during analysis of each extract.

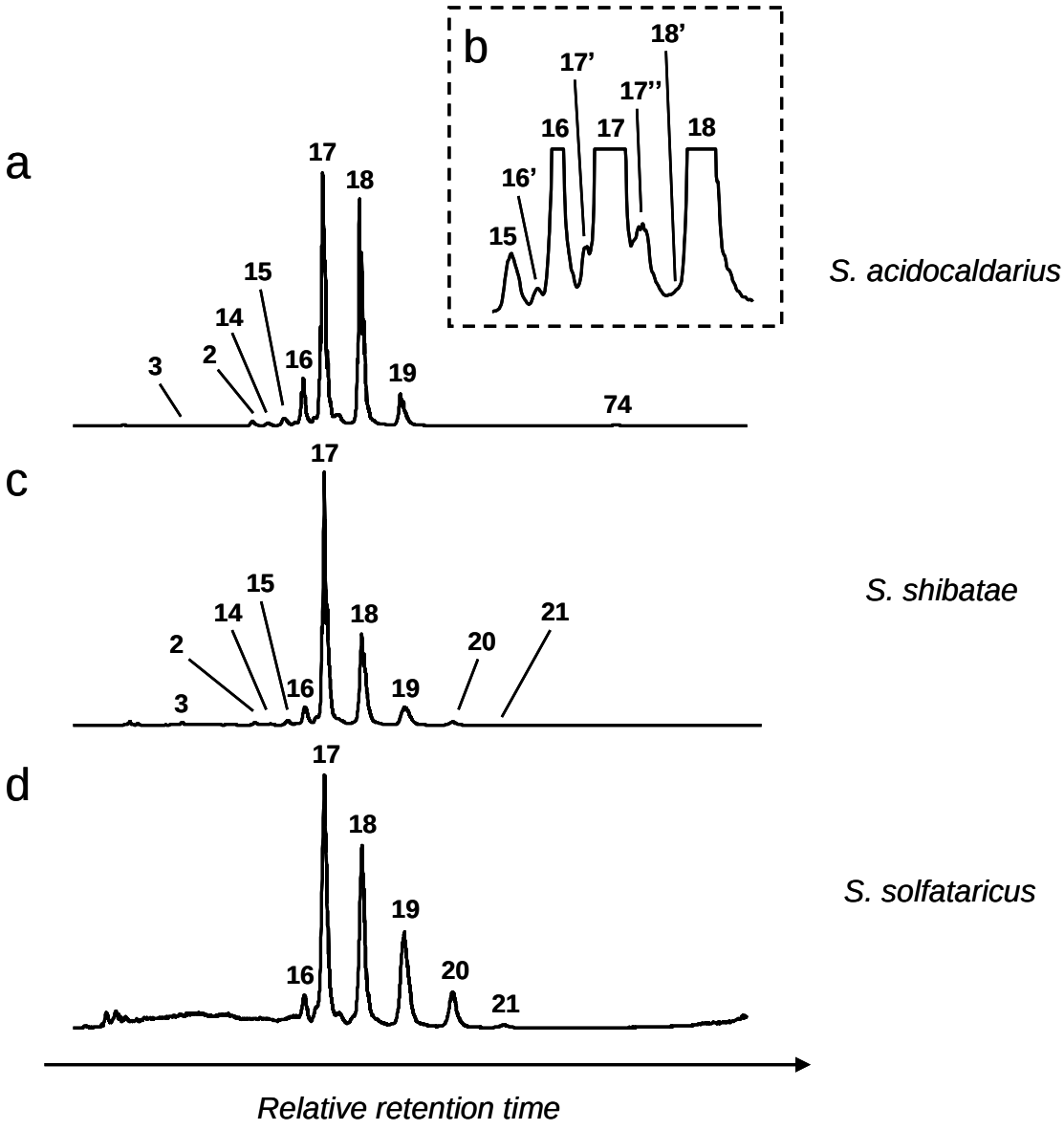


Fig. 3.7. HCT LC-MS ion chromatograms (*m/z* 1280-1320) for lipid extracts from species of *Sulfolobus*: a) *S. acidocaldarius*; b) expansion of part of the *S. acidocaldarius* chromatogram, highlighting the isomers of known isoprenoid lipid components **16**, **17** and **18**; c) *S. shibatae*; d) *S. solfataricus*. Isomers which co-elute with **15**, **17''** and **19** in each case are not labelled.

GDD **1**, expected to be a lipid core in each of the three organisms, is not observed in the base peak chromatogram for any of the species of *Sulfolobus* on account of the $[M+H]^+$ of this lipid lying outside of the scan range used during the analysis (*m/z* 900-1500). Similarly, hydroxydiether lipids **7** and/or **8**, expected as components of *S. acidocaldarius*, each have an $[M+H]^+$ outside of the scan range used. No attempts,

via re-analyses using an extended MS scan range (*m/z* 50-2000), were made in order to identify the GDD lipids in any of the *Sulfolobus* lipid extracts. The absence of peaks relating to GDGTs **2** and **14-15** in the base peak chromatogram for *S. solfataricus* is attributable to the high baseline observed during the period of the LC-MS separation in which these components elute. Individual ion chromatograms for lipids **2** (*m/z* 1301.5-1302.5), **14** (*m/z* 1299.5-1300.5) and **15** (*m/z* 1297.5-1298.5) filter out the background noise present in the base peak chromatogram and allow each of these components to be confirmed for *S. solfataricus* (Fig. 3.8). As the ion relating to the $[M+H]^+$ of each lipid is never the base peak ion in any MS scan, no MS/MS spectra were recorded for these lipids during the analysis of the extract from *S. solfataricus*.

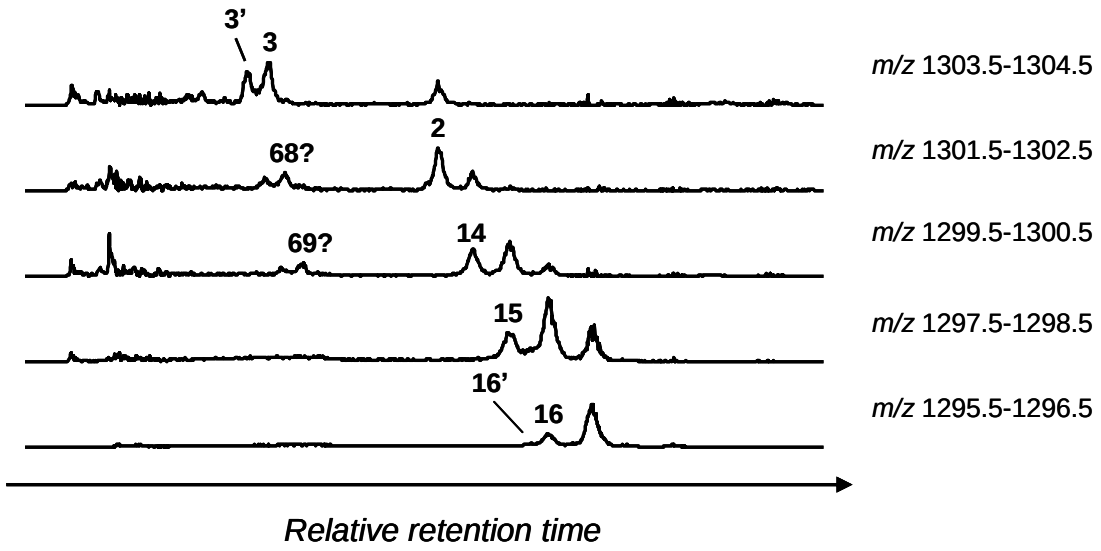


Fig. 3.8. HCT LC-MS ion chromatograms for the lipid extract from *S. solfataricus*, highlighting the presence of GTGT **3** and GDGTs **2** and **14-15**, which are concealed in the base peak chromatogram. An ion chromatogram for GDGTs **3.16** and **3.16'** is included for comparison.

Although GDGTs **20** and **21** were not identified in the base peak chromatogram for the extract from *S. acidocaldarius*, a reconstructed ion chromatogram for the MS range *m/z* 1287.5-1288.5 (Fig. 3.9) showed a peak at retention time, relative to **2** and **19**, which is approximately the same as that observed for lipid **20** in the base peak chromatograms for *S. shibatae* and *S. solfataricus*. No peak consistent with lipid **21**

was observed in an ion chromatogram reconstructed for the MS range *m/z* 1285.5-1286.5. As such, although no MS/MS data could be obtained to support the assignment made on the basis of LC-MS retention time, it appears that *S. acidocaldarius* does produce GDGT **20**, but at low levels when grown at 75°C. It does not, however, appear to produce GDGT **21** at any appreciable level at this growth temperature, although this lipid may conceivably be present in the extract at a concentration below the limit of detection for the LC-MS method used.

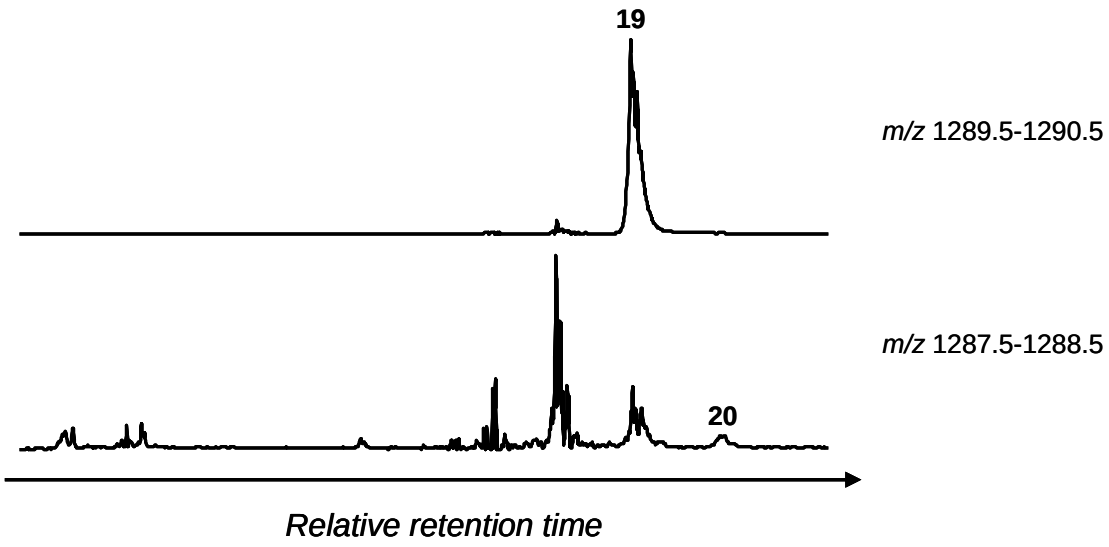


Fig. 3.9. HCT LC-MS ion chromatograms for the lipid extract from *S. acidocaldarius*, highlighting the probable presence of GDGT **20**. An ion chromatogram for GDGT **19** is included for comparison.

Components **16'** (*m/z* 1296.4), **17'** (*m/z* 1294.4), **17''** (*m/z* 1294.4) and **18'** (*m/z* 1292.4) were also identified in the base peak chromatogram of *S. acidocaldarius* (Fig. 3.7b), and in the chromatograms for *S. shibatae* and *S. solfataricus*. Each of the lipids shares an $[M+H]^+$ *m/z* value with a more abundant ring-containing GDGT lipid and elutes immediately before (GDGT') or immediately after (GDGT'') the major component. As such, lipids **16'**, **17'** and **18'** are likely to be less polar isomers of lipids **16**, **17** and **18**, respectively. Lipid **17''** is a more polar isomer of lipid **17** and has been observed in previous LC-MS analyses of core lipid extracts from strains of *S. solfataricus* (MT-4, grown at 87°C; Hopmans *et al.*, 2000 and P2, grown at 80°C; Pitcher *et al.*, 2010). Individual ion chromatograms for **14** (*m/z* 1299.5-1300.5) and

15 (*m/z* 1297.5-1298.5) from the lipid extract of *S. acidocaldarius* indicate that earlier eluting isomers of each of these lipids (**14'** and **15'**, respectively) are also produced by *S. acidocaldarius* at low levels (Fig. 3.10). Due to the low abundance of the two components, MS/MS spectra were not of sufficient signal strength for structural elucidations to be made. Structures **14'** and **15'** could not be detected in equivalent ion chromatograms for *S. shibatae* and *S. solfataricus*. A later eluting isomer of GDGT **18** was identified in previous LC-MS analysis of the lipids of a different strain of *S. solfataricus* (strain MT-4, grown at 87°C; Hopmans *et al.*, 2000), but could not be identified conclusively in the present study, either in the base peak chromatograms or in individual ion chromatograms, for any of the strains of *Sulfolobus*.

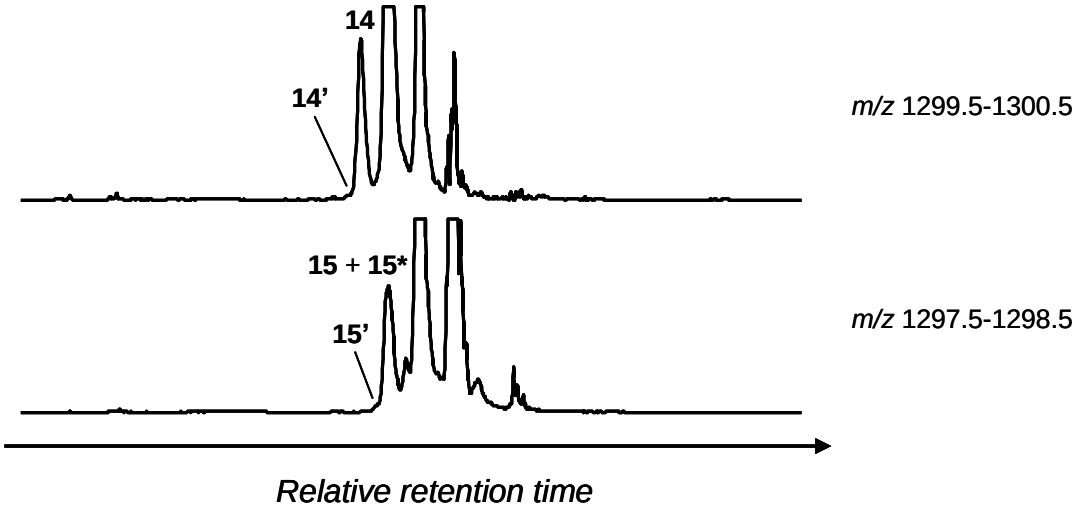


Fig. 3.10. HCT LC-MS ion chromatograms for the lipid extract from *S. acidocaldarius*, highlighting the presence of GDGTs **14'** and **15'**.

The presence of GTGT **3** in *S. acidocaldarius* is unprecedented, but was confirmed from its retention time relative to **2**, the *m/z* ratio of the $[M+H]^+$ and by MS/MS spectral data obtained for the lipid, all of which match those of the authenticated GTGT found in MTH(ΔH). Similarly, the component was confirmed in *S. shibatae* on the basis of chromatographic and mass spectral data. Although **3** could not be observed in the base peak chromatogram for *S. solfataricus*, a component with retention time and $[M+H]^+$ consistent with this lipid could be observed in a reconstructed ion chromatogram for the MS scan range *m/z* 1303.5-1304.5 (Fig. 3.8),

along with a second, slightly earlier eluting isobaric component (designated **3'**). Interestingly, components eluting marginally later than **3** and exhibiting $[M+H]^+$ at *m/z* 1302.3 (**68**) and *m/z* 1300.3 (**69**), were observed in the relevant ion chromatograms for both *S. shibatae* (Fig. 3.11) and *S. solfataricus* (Fig. 3.8) extracts. They were not, however, identified in similar chromatograms generated for *S. acidocaldarius*. Lipids **68** and **69** have $[M+H]^+$ *m/z* ratios and retention times (relative to both **2** and **3**) consistent with GTGT lipids containing one or two cyclopentyl chains, respectively. Components of this nature have been determined during LC-MS analyses of the core lipids of a thermophilic, ammonia oxidising crenarchaeote (de la Torre *et al.*, 2008), of lipid extracts from a mud volcano (Hopmans *et al.*, 2000) and of extracts from Mediterranean mud dome sediments (Pancost *et al.*, 2001b). As no MS/MS spectra were obtained for **68** or **69** during analysis of the *S. shibatae* or *S. solfataricus* extracts, the assignments remain tentative.

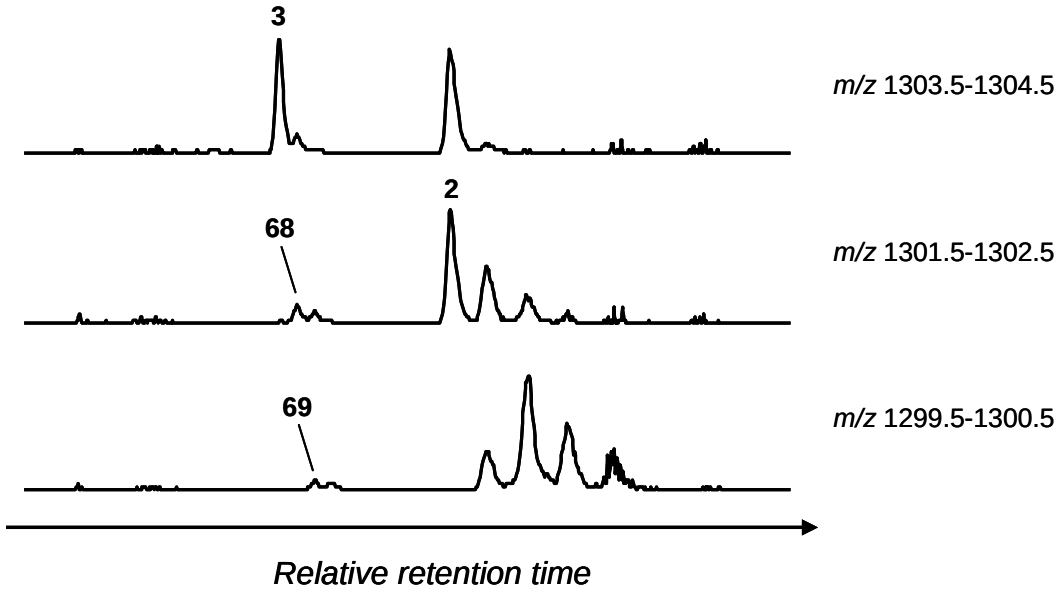


Fig. 3.11. HCT LC-MS ion chromatograms for the lipid extract from *S. shibatae*, highlighting the presence of tentatively assigned GTGTs **68** and **69**. An ion chromatogram for GTGT **3** is included for comparison.

The mass spectra of tetraethers **2** and **16**-**19** in the extract of *S. acidocaldarius* each show an additional minor peak (1.0-4.8% relative to base peak) at 14 *m/z* units higher than the base peak (e.g. Fig. 3.12c). The presence of the minor component is more clearly seen in reconstructed mass chromatograms encompassing two ranges, *m/z* 1289.5-1302.5 (Fig. 3.12a) and *m/z* 1303.5-1316.5 (Fig. 3.12b), where each GDGT lipid (**2** + **16**-**19**) is seen to co-elute with a homologous $GDGT_M$ partner (**26** + **70**-**73**). In each case, the retention time exhibited by the higher homologue is slightly less than that of its counterpart, which compares favourably with the differences in retention times observed for GDGT **2** and $GDGT_M$ **26** in MTH(ΔH). The peak observed for lipid **26** in Fig. 3.12b is amplified by a contribution from ions at *m/z* 1304.3, representative of the $[M+H+2]^+$ higher isotopic species of co-eluting lipid **2**. Interestingly, higher homologues of lipids **16'**, **17'** and **18'** may also be synthesised by *S. acidocaldarius* given the presence of **70'**, **71'** and **72'**, which each have a $[M+H]^+$ at +14 *m/z* units greater and co-elute with one of these three components (Fig. 3.12b). Examination of an ion chromatogram generated for the range *m/z* 1303.5-1316.5 for the extract from *S. shibatae* (not shown) indicated that peaks for **26** and **70'** are potentially present, whereas a similar chromatogram for *S. solfataricus* (not shown) indicated that a peak for **70'** may be present. The absence of other $GDGT_M$ lipids in the two organisms may suggest that **70'** is not a $GDGT_M$ lipid, but represents a different component, of lipid nature or otherwise, found in each of the *Sulfolobus* strains. As both **26** and **70'** were obscured by more abundant, co-eluting components (i.e. **2** and **16'**, respectively) in the initial analyses of *S. shibatae* and *S. solfataricus*, no MS/MS data was obtained for either structure. Consequently, since re-analyses using MS scan ranges which excluded signals for **2** and **16'** were not performed, currently neither **26** nor **70'** can be assigned with confidence in *S. shibatae* and *S. solfataricus*. Furthermore, since no reliable MS/MS spectrum was obtained for **70'** during analyses of *S. acidocaldarius* (see Section 3.2.4.4), the assignment of this lipid as a $GDGT_M$ derivative of **16'** is reserved until further spectral data can be collected. Similarly, **71'** and **72'** are only tentatively assigned as $GDGT_M$ derivatives of **17'** and **18'**, respectively.

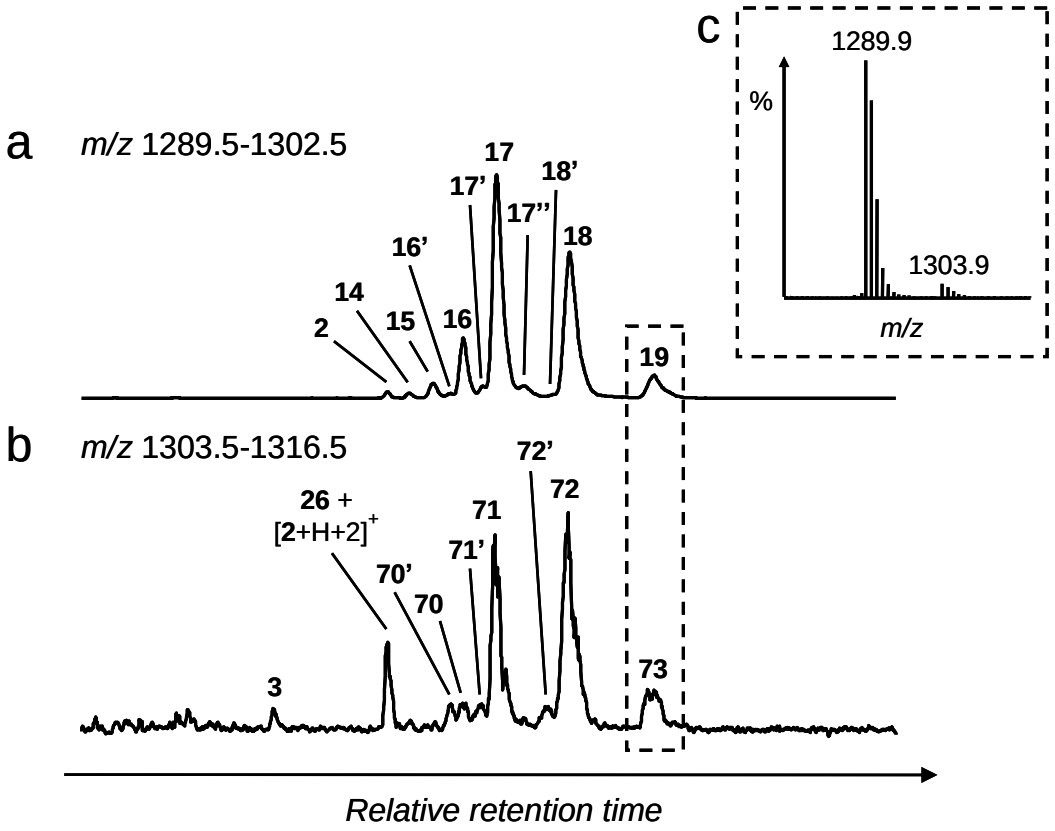


Fig. 3.12. LCQ LC-MS ion chromatograms for the lipid extract from *S. acidocaldarius*: a) *m/z* 1289.5-1302.5, highlighting GDGT lipids; b) *m/z* 1303.5-1316.5, highlighting $GDGT_M$ lipids. The peak for $GDGT_M$ **26** is obscured under a peak for the $[M+H+2]^+$ higher isotopic species of **2**. Co-eluting isomers of some of the GDGT and $GDGT_M$ lipids, while present, are not labelled; c) shows the average mass spectrum across the peak for lipids **19** and **73**.

A number of other components were evident in the base peak chromatograms (*m/z* 900-1500) of the lipid extracts. Most of these additional components are most easily visualised in ion chromatograms which exclude contributions from ions within the range *m/z* 1280-1320 (i.e. reconstructed ion chromatograms for *m/z* 900-1280 + 1320-1500). In particular, this range excludes abundant lipids **2**-**3** and **14**-**21**, thereby allowing some of the underlying minor components to be clearly visualised. Generation of an ion chromatogram of this nature for the *S. acidocaldarius* extract shows a series of peaks relating to lipids which only elute once the composition of IPA in the mobile phase has exceeded 2% (Fig. 3.13a). These include lipids **74**, **58**, **75**-**79** and **80** none of which appears to have been reported previously in archaea.

The late elution of these components suggests that they are significantly more polar than GDGTs **2** and **14**-**21**. A series of poorly resolved, early eluting lipids, each exhibiting $[M+H]^+$ with *m/z* ratios above *m/z* 1400, was also observed (Fig 3.13a). These components may represent derivatives of ring-containing tetraether lipids in which a terminal hydroxyl group has been capped with an apolar moiety. An additional early eluting component, **81** (*m/z* 1336.3), was also identified.

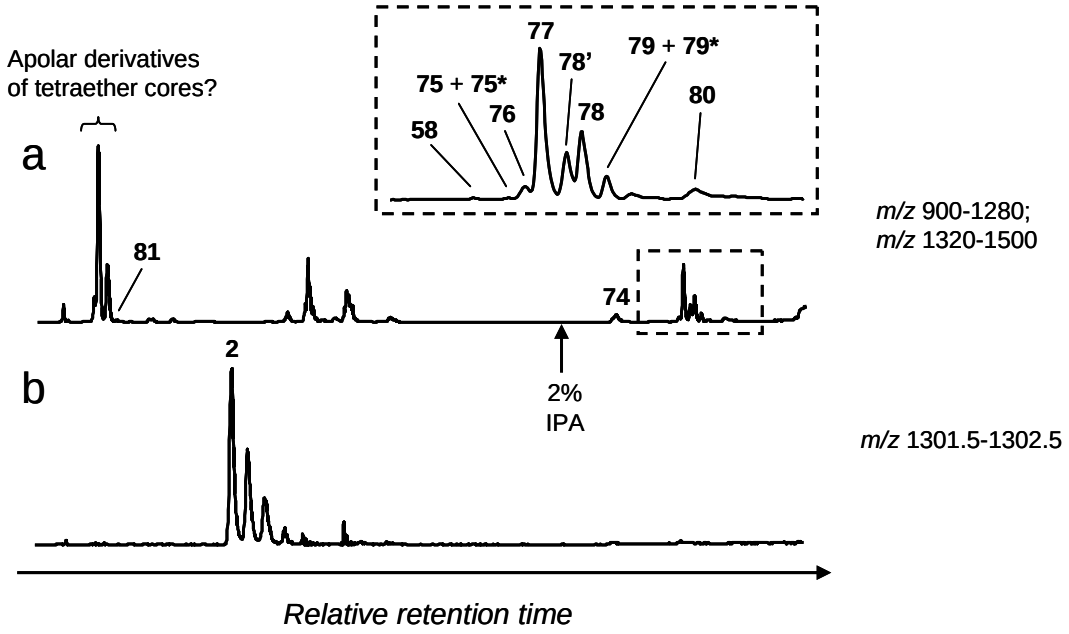


Fig. 3.13. HCT LC-MS ion chromatograms for the lipid extract from *S. acidocaldarius*: a) Base peak chromatogram, with ions in the range *m/z* 1280-1320 excluded; b) ion chromatogram for GDGT **2**, included for comparison.

An analogous chromatogram for *S. shibatae* (not shown) exhibits a number of similarities, including early eluting components with $[M+H]^+$ at *m/z* > 1400, many of which are identical to those observed in the chromatograms for *S. acidocaldarius*, and peaks for lipids **77**-**81**. Similarly, the equivalent chromatogram for the *S. solfataricus* extract shows many of the same early eluting peaks (Fig. 3.14) along with an unresolved, late eluting cluster of peaks which may represent polar derivatives of tetraether lipids.

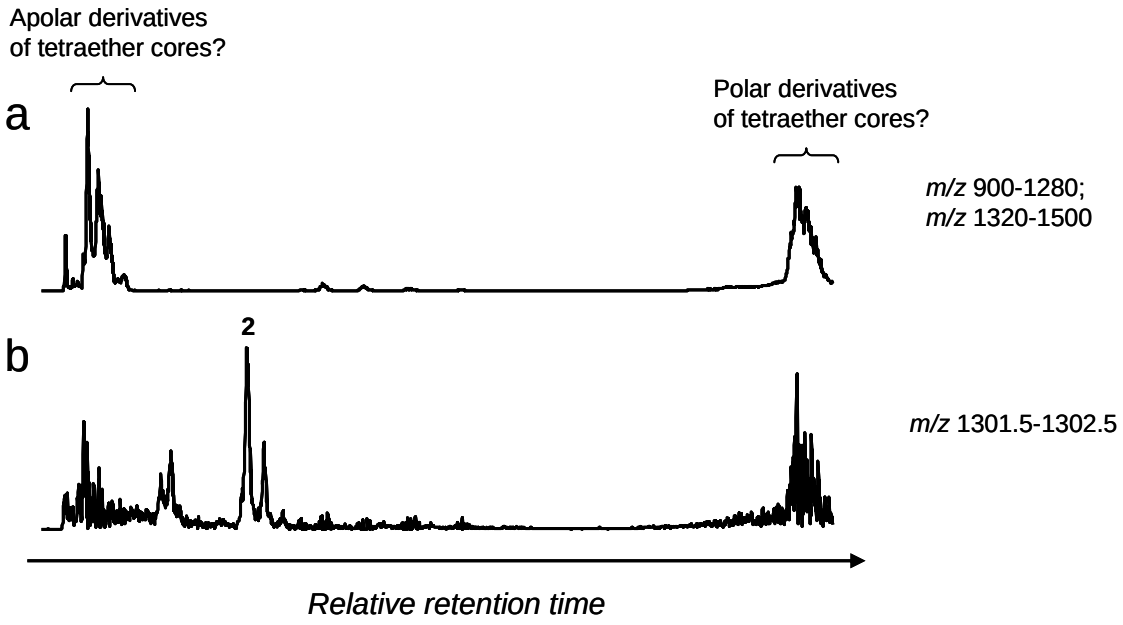


Fig. 3.14. HCT LC-MS ion chromatograms for the lipid extract from *S. solfataricus*: a) Base peak chromatogram, with ions in the range *m/z* 1280-1320 excluded; b) ion chromatogram for GDGT **2**, included for comparison.

3.2.4.3. Mass spectra obtained during the column backflush

GDCT lipids **2c** and **14c**-**21c** were not detected during the LC-MS/MS analyses of any of the *Sulfolobus* extracts. On account of their significant hydrophilicity, this is likely to be a result of retention of these components on-column for the entire duration of each analysis. The average MS spectrum recorded during the column backflushing procedure following LC-MS/MS analysis of each *Sulfolobus* lipid extract was dominated by a cluster of ions with *m/z* 1452.6-1462.4, distributed around a base peak ion at *m/z* 1456.4 (e.g. Fig. 3.15). The *m/z* of the base peak ion corresponds to that predicted for the $[M+H]^+$ of GDCT **17c** on the basis of its ionic formula, $C_{92}H_{175}O_{11}^+$ (theoretical *m/z* 1456.313). An ion at *m/z* 1456.314 was observed as a minor ion in a previous positive ion electrospray-time of flight MS analysis of a GDCT fraction isolated from *Sulfolobus metallicus* and assigned by the authors as the $[M+H]^+$ of **17c** (Bode *et al.*, 2008). The sodiated ($[M+Na]^+$; *m/z* 1478.297) and potassiated ($[M+K]^+$; *m/z* approximately 1494.3) molecules of **17c** were observed to be formed in preference to the $[M+H]^+$ species under the electrospray conditions employed in the study, explaining the low relative abundance

of *m/z* 1456.314 in the mass spectrum. By contrast with electrospray ionisation, positive mode APCI ionisation of lipids typically leads to a preference for formation of $[M+H]^+$ species (Byrdwell, 2001). This may explain the dominance of *m/z* 1456.4, and absence of sodiated and potassiated molecules, in the APCI MS spectra recorded during backflushing following analysis of extracts of the *Sulfolobus* species used in the present study. Furthermore, the dominant GDCT core lipid in the type strain of *S. acidocaldarius*, grown at an identical temperature to the culture of the MR31 strain used in the present studies (75^{o}C), was determined to be **17c** in negative ion halide adduct-MS (Murae *et al.*, 2002). Consequently, the ion with *m/z* 1456.4 in the backflush MS of each species is assigned to the $[M+H]^+$ of **17c**. Other neighbouring peaks in the ion cluster may potentially relate to GDCT lipids containing between 2-6 Cp rings. Additional ion clusters were observed in the backflush mass spectrum of each organism, including clusters containing ions shifted by -18, -36, -54, -72, -90, -162, -180, -218 and -236 *m/z* units from ions in the cluster surrounding the base peak ion (Fig. 3.15). These ions probably relate to fragment ions generated from GDCT lipids *via* losses of one (-18), two (-36), three (-54), four (-72) or five (-90) molecules of water, loss of the terminal pentacyclic ring in the calditol moiety *via* ether cleavage (-162 and -180) or loss of this ring with simultaneous loss of C_3H_4O from the glycerol terminus (-218 or -236). In each spectrum, a cluster of ions shifted by +14 *m/z* from ions in the cluster containing the base peak ion was also observed. These components could conceivably be higher homologues of GDCT lipids, although further work will be required in order to confirm their identities.

Signals for protonated molecules of GDCTs persisted at high levels for the full 10 minute period of the column backflush and were often only reduced to a baseline level if the column was subsequently backflushed with mobile phase comprising 40% IPA for up to a further 30 minutes. The difficulty in removing GDCT lipids from the analytical column may suggest that the backflushing protocol used is inadequate for total core lipid extracts containing these components, namely those taken from organisms within the order *Sulfolobales*. Consequently, it is recommended that, in future, GDCT lipids are separated from other ether lipids by flash column chromatography (Lo *et al.*, 1989; Nicolaus *et al.*, 1990) prior to normal phase LC-MS/MS analysis in order to reduce the risk of column deterioration.

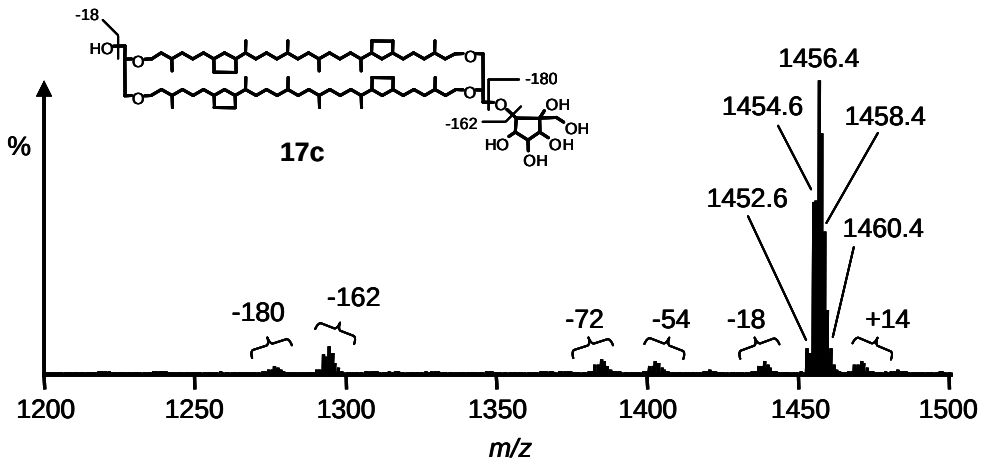


Fig. 3.15. Average mass spectrum (HCT) recorded during the column backflushing procedure following LC-MS/MS analysis of *S. acidocaldarius*. A cluster of dominant ions surrounding the base peak ion (*m/z* 1456.4) are labelled, with other ion clusters labelled with the difference in *m/z* to ions within the major cluster. Suspected fragmentations responsible for the some of the ion clusters are typified for GDCT **17c**.

3.2.4.4. Interpretation of MS/MS spectra of ring-containing dialkyl tetraether lipids

Tandem MS of lipids **14-21** in the *Sulfolobus* extracts generated product ions relating to losses of water, C_3H_4O and previously recognised biphytanyl chains [40_0] (-558 Da), [40_1] (-556 Da), [40_2] (-554 Da), [40_3] (-552 Da) and [40_4] (-550 Da) as dienes. Product ions were identified in two distinct spectral regions for each lipid (e.g. Fig. 3.16a), confirming that each has a GDGT core structure. For example, the MS/MS spectrum of GDGT **16** shows product ions at *m/z* 741 and *m/z* 739, formed from loss of $[40_2]^0$ and $[40_1]^0$, respectively (Fig. 3.16b). Product ions at *m/z* 723, 721, 705 and 703 also prove to be characteristic of the alkyl chains in the structure, and are formed from losses of one or two molecules of water in addition to loss of each alkyl chain. Notably, ions at *m/z* 707 and 701 in the spectrum can confuse structural assignment. At first sight, these ions could potentially be attributed to losses of $[40_3]^2$ and $[40_0]^2$, respectively. In reality, they are almost certainly formed *via* losses of isoprenoid chains containing one or two rings, as a result of an ether cleavage which does not generate an alkene and protonated alcohol (*cf.* ions at *m/z* 705 and 709 in the HCT

MS/MS spectrum of **2**; Section 2.2.5.2). This example highlights the need to consider first the product ions formed from losses of $[40_P]^0$ (P = 0-4) in order to assign the alkyl chains within ring-containing GDGT structures, with the losses of $[40_P]^1$ and $[40_P]^2$ used as subsequent supporting evidence, where necessary.

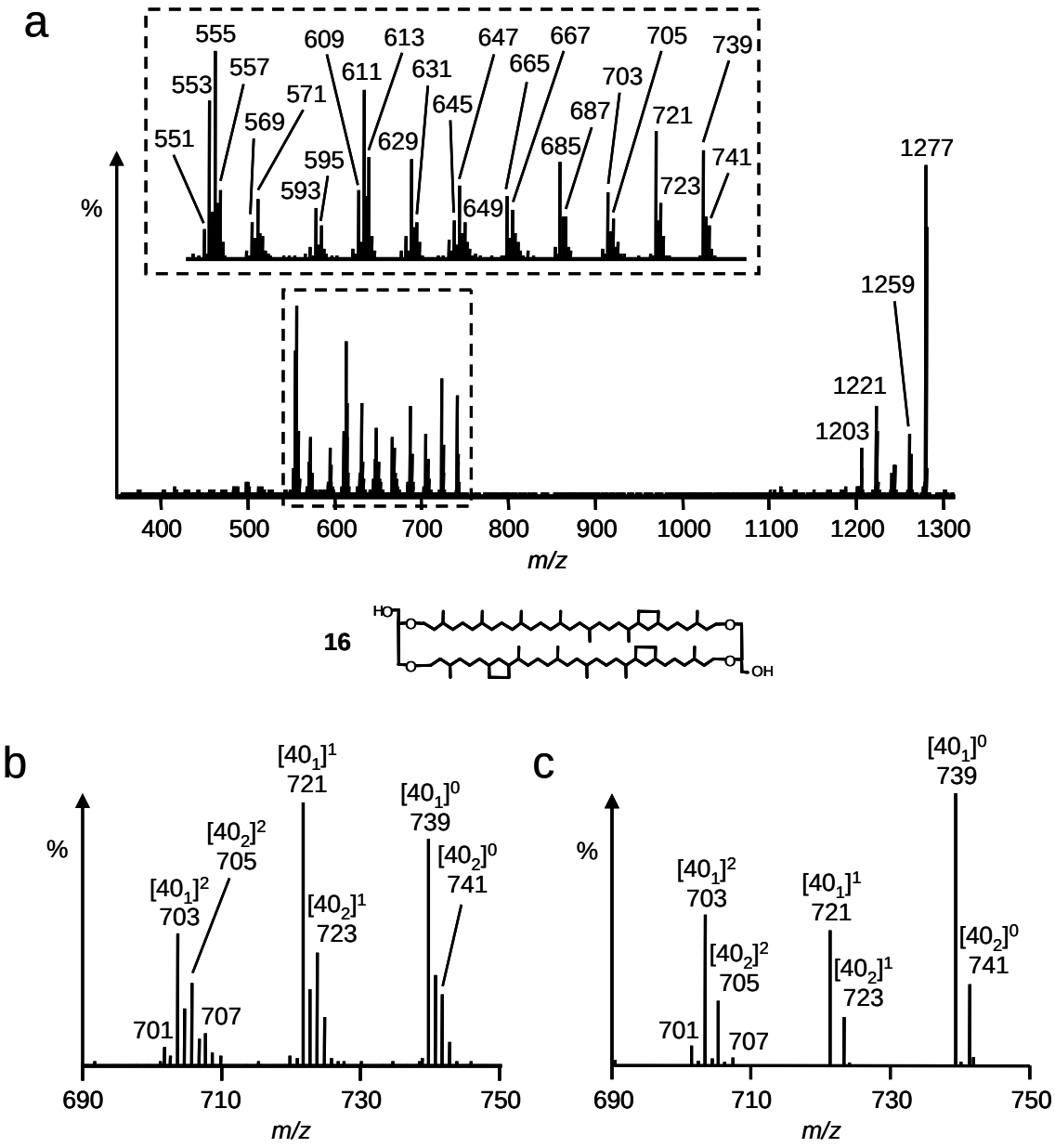


Fig. 3.16. MS/MS spectrum of GDGT **16** ($[M+H]^+$ *m/z* = 1296.4) from *S. acidocaldarius*, as recorded on: a) the HCT instrument. b) shows the magnification of the *m/z* 690-750 region of the spectrum; c) shows a similar magnification of the *m/z* 690-750 region of the spectrum acquired using the LCQ instrument.

Product ions in the HCT MS/MS spectra are often accompanied by ions at +1 *m/z* unit greater, relating to the higher isotopomer formed from CID of the $[M+H+1]^+$ isotopic precursor (e.g. the product ion at *m/z* 740 in Fig. 3.16b). Similarly, a second, relatively small peak at +2 *m/z* units, relating to the +2 isotopomer, is often also observed. Because the +2 species could be confused with loss of a cyclised biphytanyl chain having one less degree of unsaturation, additional care must be taken in interpreting the MS/MS spectra of ring-containing GDGT lipids. In this regard, the absence of an ion shifted by +3 *m/z* units usually indicates that the ion shifted +2 *m/z* units results from an isotopomer: almost without exception, an isotopic peak at +1 *m/z* is observed for all product ions formed from loss of $[40_P]^0$. A benefit of recording the MS/MS spectra of ring-containing GDGT lipids using the LCQ instrument is that higher isotopomers of the characteristic product ions are much less prominent in the spectra (e.g. Fig. 3.16c), enabling clearer distinction of product ions formed from losses of $[40_P]^0$. In the MS/MS spectra of **16**, the product ion formed from loss of $[40_1]^0$ is consistently observed to be more prominent than the ion formed from loss of $[40_2]^0$, suggesting that the chain containing fewer rings is lost more readily from the $[M+H]^+$ of the lipid during CID. This proves to be a general feature for the Cp ring-containing cores identified in the extracts from *Sulfolobus*; preference for loss of the heavier of the two isoprenoid chains during CID was noted for all lipids in which the two chains are different.

During acquisition of the MS spectrum of **16** in the *S. solfataricus* extract (Fig. 3.17a), the base peak ion ($[M+H]^+$; *m/z* 1295.8) was only slightly more prominent than several ions of lower *m/z* which were responsible for the high level of background ion current and, consequently, high baseline in the early part of the LC-MS chromatogram for the extract (Fig. 3.7d). Following CID, it was apparent that the MS/MS spectrum obtained was contaminated by product ions formed from the background ions in addition to those formed from the target precursor (i.e. $[M+H]^+$ of **16**). Whereas the pure spectrum of **16** shown in Fig. 3.16a exhibited a product ion at *m/z* 1277, formed from loss of water, the contaminated spectrum showed an additional ion at *m/z* 1275, suggestive of loss of water from a precursor with nominal *m/z* 1293 (Fig 3.17b). Similarly, additional ions were observed at *m/z* 737 and 719 in the contaminated spectrum (Fig. 3.17c) that were absent in the pure spectrum. Again,

without careful interpretation of the spectra, the additional ions could be attributed an origin from the $[M+H]^+$ of **16**, leading to an incorrect structural assignment. Survival beyond the precursor selection stage and subsequent dissociation of ions at 1-2 *m/z* units less than the target precursor would account for the undesired product ions generated during CID. Low precursor selectivity may prove to be particularly problematic in cases where fully or partially co-eluting lipid components in LC-MS have $[M+H]^+$ which differ by only 2 *m/z* units. Consequently, it is essential to check that the precursor selection step has been selective for the intended precursor before interpretations regarding lipid structure are made from the resulting MS/MS spectra. This is most easily achieved by examination of the product ion correlating to loss of water. Spectra in which this ion is accompanied by a prominent additional peak at -2 *m/z* units less should not be used for structural assignment.

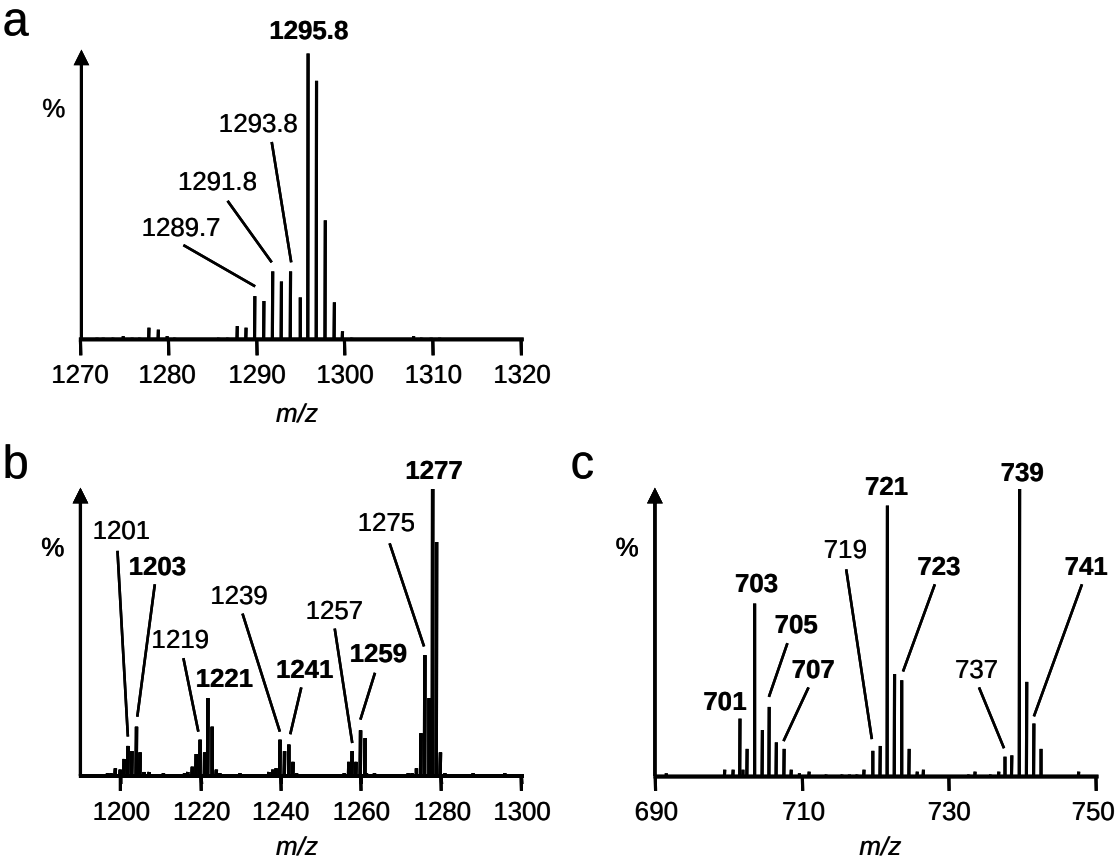


Fig. 3.17. a) Partial MS and b-c) partial MS/MS spectra (HCT) of GDGT **16**, recorded during analysis of the lipid extract from *S. solfataricus*. The target precursor and product ions derived thereof are shown in bold face. Contaminant ions are shown in standard font.

3.2.4.5. Structural assignments based upon MS/MS spectra

The MS/MS spectra of lipids **2** and **3** in the extracts of the *Sulfolobus* species exhibit identical product ions to those generated by CID of these lipids in the MTH(ΔH) extract. In most cases, the alkyl group losses observed in the MS/MS spectra of ring-containing GDGTs **14**-**21** are consistent with the structures of these lipids as previously reported for *Sulfolobus* species (De Rosa and Gambacorta, 1988a). For example, the most abundant ring-containing GDGT lipid in each extract, **17**, is known to contain isoprenoid chains which each have two incorporated cyclopentyl rings. The HCT MS/MS spectrum of this lipid (Fig. 3.18a) shows a product ion at *m/z* 739, formed from loss of $[40_2]^0$, entirely consistent with the reported structure. Ions formed from loss of $[40_2]^1$ (*m/z* 721) and $[40_2]^2$ (*m/z* 703) provide further corroboration. The key characteristic product ions observed in the MS/MS spectra of the other ring-containing GDGT lipids are summarised in Table 3.4. It has been widely reported that the GDGT lipids of *Sulfolobus* contain two biphytanyl chains that differ at most by one in the number of Cp rings incorporated in each of the two chains (De Rosa *et al.*, 1980c; 1983b; 1988a). The LC-MS/MS analysis of the GDGT lipids in the three *Sulfolobus* species suggest that biphytanyl combinations which differ by two rings are also synthesised by these organisms. For example, CID of the $[M+H]^+$ of lipid **19** generates product ions at *m/z* 739, 737 and 735, consistent with losses from the precursor of, in turn, $[40_4]^0$, $[40_3]^0$ and $[40_2]^0$ (Fig. 3.18b). The losses indicate that a novel isobar, GDGT[40_2,40_4] (denoted **19***), accompanies the known tetraether GDGT[40_3,40_3] (denoted **19**) in all three *Sulfolobus* extracts. It is noteworthy that analysis of tetraacids in crude oil naphthenate deposits revealed a tetraacid with a C_{80} H-shaped isoprenoid comprising cross-linked $[40_2]$ and $[40_4]$ chains, providing precedent for the occurrence of such a combination in archaeal lipids (Lutnaes *et al.*, 2007). Other novel or rarely reported lipids of archaea, assigned in the *Sulfolobus* species from the LC-MS/MS analyses and which contain unusual isoprenoid chain combinations are GDGT[40_0,40_2] (**15***; Thurl and Schäfer, 1988; Schouten *et al.*, 2000; Pancost *et al.*, 2001b) and GDGT[40_1,40_3] (**17''***; Hopmans *et al.*, 2000).

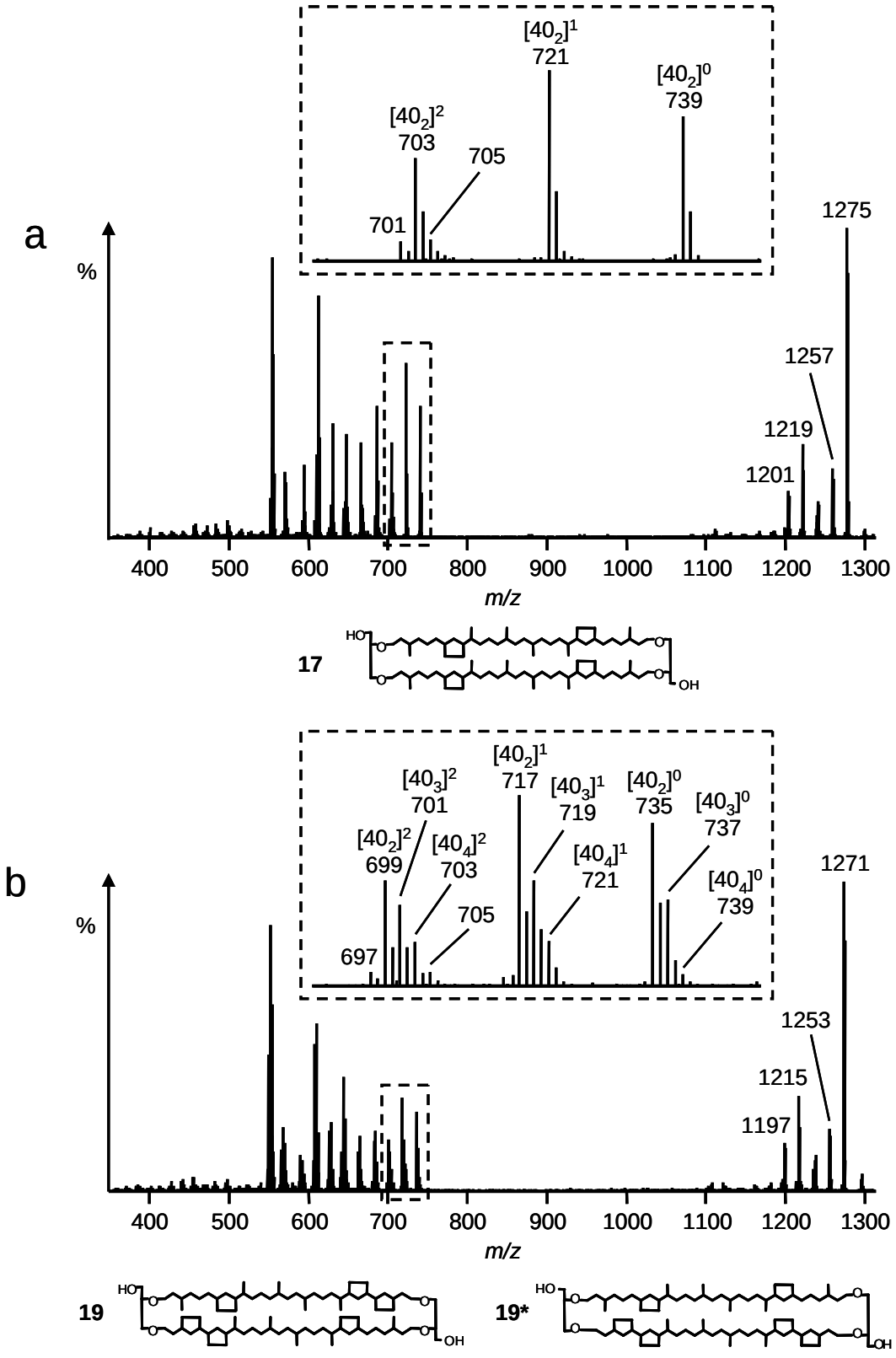


Fig. 3.18. MS/MS spectra (HCT) of: a) GDGT **17** ($[M+H]^+$ *m/z* = 1294.4); b) GDGT **19**/**19*** ($[M+H]^+$ *m/z* = 1290.3) from *S. acidocaldarius*. Insets show expansions of regions of the spectra, where appropriate.

Table 3.4. HCT MS data (measured $[M+H]^+$ *m/z* values) and partial MS/MS data (nominal *m/z* values of product ions) for GDGT and GDGT' lipid cores from species of *Sulfolobus*.

	No. of rings	0	1	2	3	4	4	5	6	7	8
Loss (Da)	Molecules Lost [a]	**2**	**14**	**15** + **15***	**16** or **16'**	**17** or **17'**	**17''** + **17''***	**18** or **18'**	**19** + **19***	**20**	**21**
-0	$[M+H]^+$	1302.4	1300.4	1298.4	1296.4	1294.4	1294.4	1292.4	1290.3	1288.2	1286.2
-18	[1,0,0]	**1283**[b]	**1281**	**1279**	**1277**	**1275**	**1275**	**1273**	**1271**	**1269**	1267
-36	[2,0,0]	1265	1263	1261	1259	1257	1257	1255	1253	1251	1249
-54	[3,0,0]	1247	1245	1243	1241	1239	1239	1237	1235	1233	1231
-56	[0,1,0]	1245	1243	1241	1239	1237	1237	1235	1233	1231	1229
-74	[1,1,0]	1227	1225	1223	1221	1219	1219	1217	1215	1213	1211
-92	[2,1,0]	1209	1207	1205	1203	1201	1201	1199	1197	1195	1193
-550	$[0,0,40_4]$								739	737	735
-552	$[0,0,40_3]$						741	739	737	735	
-554	$[0,0,40_2]$			743	741	739	739	737	735		
-556	$[0,0,40_1]$		743	741	739		737				
-558	$[0,0,40_0]$	743	741	739							
-568	$[1,0,40_4]$								721	719	717
-570	$[1,0,40_3]$						723	721	719	717	
-572	$[1,0,40_2]$			725	723	721	721	719	717		
-574	$[1,0,40_1]$		725	723	721		719				
-576	$[1,0,40_0]$	725	723	721							
-586	$[2,0,40_4]$								703	701	699
-588	$[2,0,40_3]$						705	703	701	699	
-590	$[2,0,40_2]$			707	705	703	703	701	699		
-592	$[2,0,40_1]$		707	705	703		701				
-594	$[2,0,40_0]$	707	705	703							
-624	$[1,1,40_4]$								665	663	661
-626	$[1,1,40_3]$						667	665	663	661	
-628	$[1,1,40_2]$			669	667	665	665	663	661		
-630	$[1,1,40_1]$		669	667	665		663				
-632	$[1,1,40_0]$	669	667	665							

[a] The number of simultaneous losses of each type of molecule during formation of the ion are listed as $[a,b,c_d]$, where a = water; b = C_3H_4O; c_d = biphytadiene containing "c" carbon atoms and "d" cyclopentyl rings (d = 0-4).

[b] Product ions listed in bold face represent the base peak in MS/MS. The base peak ion noted in the spectra of **21** does not correspond to any of the ions listed.

In order to obtain MS/MS spectra for $GDGT_M$ lipids in *S. acidocaldarius*, the extract was re-analysed using a dynamic narrow scan range, set at *m/z* 1306-1320 until **26** had eluted fully, after which the range was manually altered to *m/z* 1303-1320 for the remainder of the analysis. During the preliminary scan range, the $[M+H]^+$ of **2** and

its higher isotopomers were removed, allowing for selection of **26** for CID. Following manual alteration to the wider scan range, peaks for the $[M+H]^+$ of **16**-**19** were removed, allowing generation of MS/MS spectra for lipids **70**-**73**. Using this method, three previously concealed $GDGT_M$ species (**71**, **72** and **73**) generated MS/MS spectra with sufficient signal strength for reliable interpretations to be made. The MS/MS spectra of other $GDGT_M$ components (**26**, **70** and **70'**-**72'**) were too weak to be used for structural characterisation. Components **71**-**73** gave product ions in two distinct regions of the MS/MS spectra, confirming them to contain two discrete alkyl chains. The MS/MS spectrum of **71** (Fig. 3.19a) contains product ions formed *via* losses of $[40_2]^0$ (*m/z* 753) and $[41_2]^0$ (*m/z* 739), indicating that the lipid possesses a GDGT[40_2,41_2] structure. The presence of additional product ions, formed from losses of $[40_2]^1$ (*m/z* 735), $[41_2]^1$ (*m/z* 721) and $[40_2]^2$ (*m/z* 717), corroborate the assignment. The MS/MS spectrum of **72** (Fig. 3.19b) contains product ions arising from losses of $[40_2]^0$ (*m/z* 751) and $[41_2]^0$ (*m/z* 737), and additional ions formed from losses of $[40_3]^0$ (*m/z* 753) and $[41_3]^0$ (*m/z* 739). Other product ions observed in the MS/MS spectrum are consistent with loss of the same four alkyl chains with simultaneous loss of additional molecules of water, suggesting that the chromatographic peak comprises a mixture of GDGT[40_2,41_3] (denoted **72**) and GDGT[40_3,41_2] (denoted **72***) isomers. Although CID of the $[M+H]^+$ of **73** generated a weak MS/MS spectrum (Fig. 3.19c), it does, nevertheless, provide some information about the structure of the lipid core. Product ions at *m/z* 737 and 719 suggest loss of $[41_3]^0$ and $[41_3]^1$ from the precursor, respectively. This implies that one chain in **73** is a $[41_3]$ biphytanyl. The ions at *m/z* 737 and 719 in the MS/MS spectrum of **19** both contain a biphytanyl chain with three Cp rings (Fig. 3.18b). By implication, the product ions observed at the same nominal *m/z* for **73** probably also contain the same biphytanyl moiety. A combination of $[41_3]$ and $[40_3]$ chains in a GDGT-type core structure would account for the $[M+H]^+$ observed for **73** (*m/z* 1303.9). This interpretation could not be confirmed since ions relating to the expected losses of $[40_3]^0$ or $[40_3]^1$ in the MS/MS spectrum of **73** were not observed. In spite of the limited nature of the MS/MS evidence, **73** is tentatively assigned as a GDGT[40_3,41_3] structure. Other product ions in the MS/MS spectrum of **73** are not consistent with this structure and indicate that an isomeric structure, **73***, may also be synthesised by *S. acidocaldarius*. For example, the product ion at *m/z* 753 is

consistent with loss of $[40_4]^0$ from the precursor. A product ion at *m/z* 735 could relate to loss of either $[41_2]^0$ or $[40_4]^1$. Similarly, an ion at *m/z* 717 could be formed *via* loss of either $[41_2]^1$ or $[40_4]^2$. Thus it appears likely that lipid **73*** has a GDGT[40_4,41_2] structure and co-elutes with **73** under the chromatographic conditions employed. Although this observation is unprecedented, it is consistent with the observation of unusual product ions formed following CID of the $[M+H]^+$ of GDGT **19**, as discussed above.

Lipid **74** exhibits a $[M+H]^+$ at *m/z* 1320.4 in the MS spectrum and gives an MS/MS spectrum (Fig. 3.20) resembling that of GTGT **3**, allowing its classification as a GTGT lipid. The relatively late elution of this lipid coupled with the 16 *m/z* unit difference to the $[M+H]^+$ of GTGT **3** is consistent with **74** being a structural variant of **3** that contains an additional hydroxyl group. A product ion at *m/z* 1039 in the MS/MS spectrum of **74** indicates a loss of 1-phytene, as was observed for **3**. In addition, an ion at *m/z* 1023 (-296 Da) shares the same *m/z* as the base peak ion in the MS/MS spectrum of lipid **3**. This indicates that loss of 296 Da from the $[M+H]^+$ of **74** during CID generates a product ion that is, most likely, identical to that formed by loss of phytene from that of lipid **3**. The 296 Da loss most likely corresponds to a hydroxylated phytene moiety. Interestingly, this loss would be an expected loss from *sn*-2 and *sn*-3 hydroxyarchaeols (GDDs **7** and **8**), at least one of which has been identified previously as a core component of *S. acidocaldarius* (Sprott *et al.*, 1997). As such, **74** is assigned as acyclic hydroxycaldarchaeol. A product ion at *m/z* 743 for **74**, is attributed to sequential losses of 1-phytene and hydroxyphyt-1-ene. Ions reflecting protonated phytanyl and hydroxyphytanyl monoethers, at *m/z* 373 and 389, respectively, provide further support for the assignment of **74**. It is important to note that although the additional hydroxyl group in **74** has been confirmed to reside somewhere along a C_{20} phytanyl chain as opposed to on the C_{40} lipid-spanning biphytanyl chain, the precise position of this group, both in terms of its location along the C_{20} chain and whether this chain is ether bound to an *sn*-2 or an *sn*-3 carbinol, cannot be determined on the basis of the MS/MS evidence obtained. Given that the additional hydroxyl group in both *sn*-2 and *sn*-3 hydroxyarchaeols is located at the C-3 position of a phytanyl chain, the additional hydroxyl in **74** is tentatively placed at the same position.

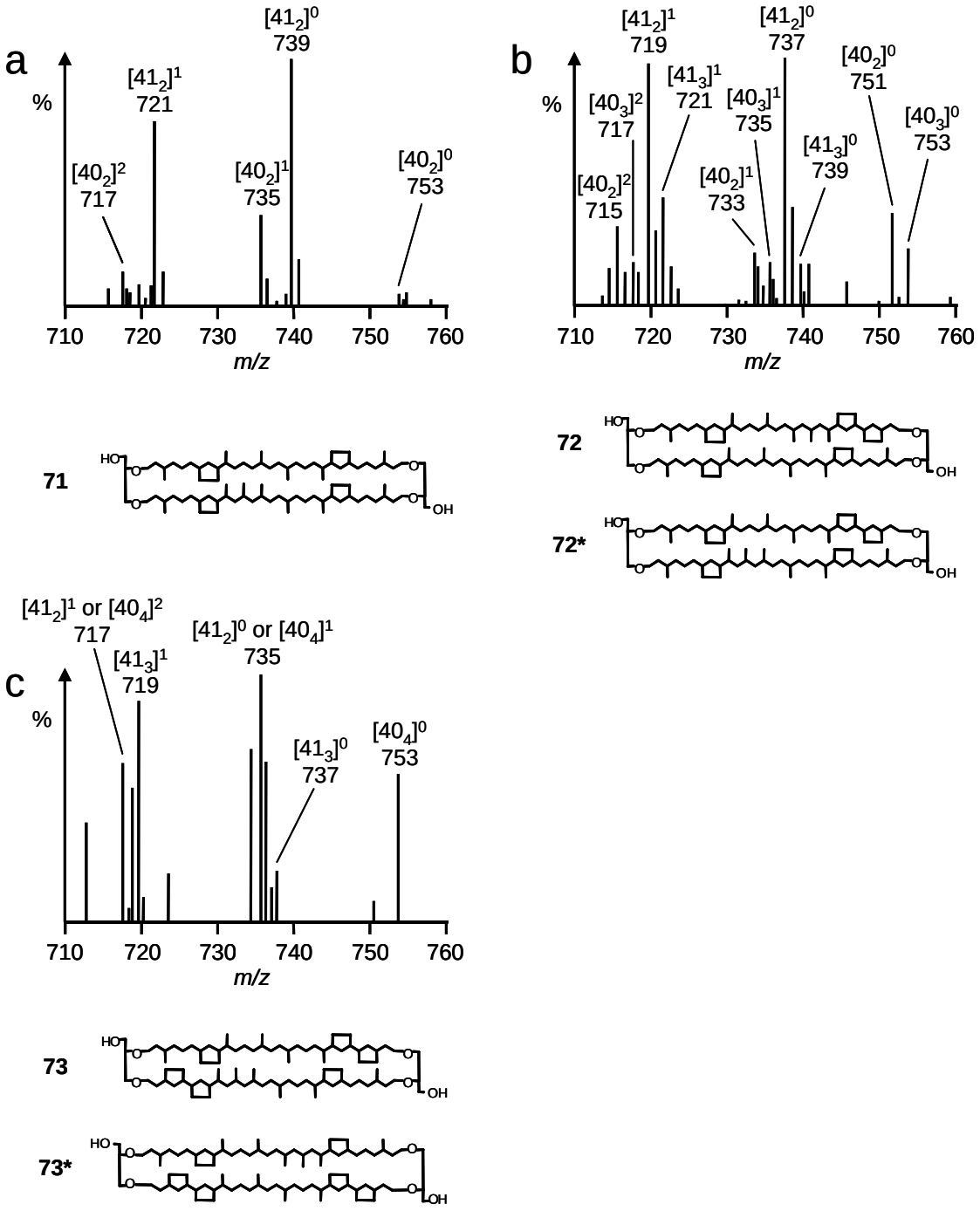


Fig. 3.19. Partial MS/MS spectra (HCT) of: a) $GDGT_M$ **71** ($[M+H]^+$ *m/z* = 1308.3); b) $GDGT_M$ **72**/**72*** ($[M+H]^+$ *m/z* = 1306.2) and c) $GDGT_M$ **73**/**73*** ($[M+H]^+$ *m/z* = 1304.3) from *S. acidocaldarius*. $GDGT_M$ structures shown are representative structures; the additional methyl group in each is tentatively placed at the C-13 chain position in accordance with the known structure of $GDGT_M$ **26** (Galliker *et al.*, 1998).

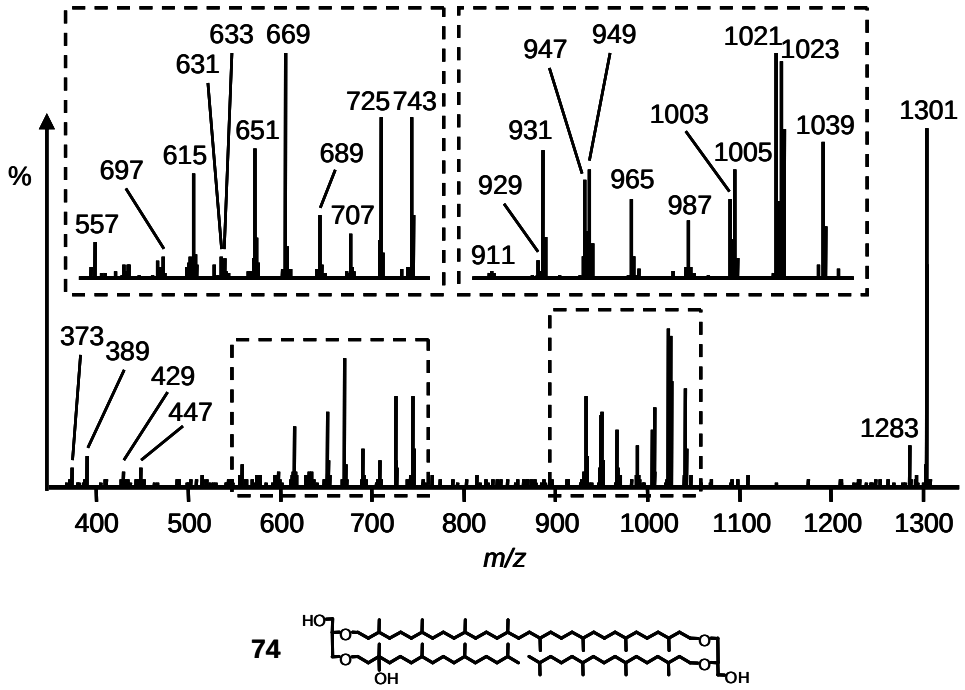


Fig. 3.20. MS/MS spectrum (HCT) of GTGT **74** ($[M+H]^+$ *m/z* = 1320.4) from *S. acidocaldarius*. Inset shows expansion of a region of the spectrum. The structure shown is a representative structure; the hydroxyl group is tentatively placed at the C-3 position of a phytanyl chain (*sn*-3 bound chain chosen arbitrarily) in accordance with its placement in GDDs **7**-**8**.

The HCT MS/MS spectrum of lipid **58** is weak, but contains some of the product ions observed in the spectrum for this component obtained during LC-MS/MS of an MTH(ΔH) extract (Section 3.2.3.3). The MS/MS spectrum of **75** (*m/z* $[M+H]^+$ = 1242.3) was too weak to be assigned reliably. The MS/MS spectra of lipids **76**-**79** and **78'** (Table 3.5), on the other hand, are qualitatively similar to that of **58** and suggest that these components are Cp ring-containing variants of **58**. Discussion of the probable nature of each of these components is presented in Section 5.2.3.2. The MS/MS spectrum of **81** (Table 3.5) exhibits product ions at *m/z* 1293, 1275 and 763, assigned as arising from loss of ketene, acetic acid and a 1-hydroxybiphyt-31-ene containing two Cp rings, respectively. Given the neutral losses observed in MS/MS, the early chromatographic elution of this lipid and the nominal 42 *m/z* unit difference between the $[M+H]^+$ of the lipid (*m/z* 1336.3) to that of **17**, **81** is assigned as the monoacetylated derivative of **17**. The MS/MS spectrum of lipid **80** (Table 3.5) shows product ions at *m/z* 1311 and 1255, formed from losses of water and $C_3H_6O_2$,

and product ions in two regions of the spectrum, indicative of a GDGT-type core structure. The exact nature of this structure remains unclear.

Table 3.5. MS and MS/MS spectral data (HCT) for minor lipid components identified in *Sulfolobus* extracts.

Lipid	$[M+H]^+$	Product ions observed in MS/MS [a]
76	1240.3	1221; 667,**665**,649,647,631,629,613,611,609,557,555
77	1238.3	1219,1201; **665**,647,629,611,609,593,591,571,569,555,553
78'	1236.3	1217,1199,1161; 665,**663**,647,645,643,629,627,611,609,607,591,571,569,555,553,551
78	1236.3	1217,1199; 665,**663**,647,645,643,629,627,611,609,593,591,571,569,555,553,551
79/79*	1234.3	**1215**,1197; 665,663,661,647,645,643,641,629,627,625,611,609,607,591,589,571,569,567, 565,555,553,551,549
80	1330.3	1311,1293,1275,1255,1237; **739**,721,703,685,683,665,647,645,629,611,609,589,571,569,555,553,551
81	1336.3	1317,1299,1293,1275,1257,1239,1219,1199; **763**,747,745,721,703,689,687,685,671,651,629,611,609,607,593,571,555,553

[a] Product ions listed in bold face represent the base peak ion in MS/MS.

Lipids with $[M+H]^+$ at *m/z* > 1400, including both early- and late-eluting components described for the *Sulfolobus* species, exhibit poor chromatographic resolution. As a result, the base peak ion in the mass spectrum of each component is usually accompanied by neighbouring ions relating to other lipids. Although the MS/MS spectra obtained for these components are suggestive of tetraether core lipid structures, inspection of product ions formed from loss of water indicated that, in all cases, the MS/MS spectrum obtained was contaminated by product ions arising from selection and CID of neighbouring lipids with similar *m/z* ratios. A similar problem, resulting from contamination of the spectrum with product ions formed from other GDCT lipids, was encountered during acquisition of the MS/MS spectrum of GDCT **17c** during the post-analysis backflushing procedure following LC-MS/MS of each

Sulfolobus extract. Consequently, the structures of each of these components cannot be assigned or, in the case of **17c**, verified by MS/MS.

3.2.4.6. PAN LC-MS/MS

In order to investigate whether ions below the low mass cut off in conventional ion trap LC-MS/MS analyses could be used to augment the characterisation of lipids from *Sulfolobus*, an analysis of the *S. solfataricus* extract was performed using PAN LC-MS/MS. The PAN MS/MS spectra for GDGTs **16**-**21**, **17'** and **17''**/**17''*** each exhibited product ions at $m/z < 400$, many of which had the same m/z ratio as an ion observed in the PAN MS/MS spectrum of **2** (Fig. 2.9b). Unfortunately, although ions with low m/z that were observed in the PAN MS/MS spectra of some lipids were absent in the spectra of others, most of the ions appeared consistently in the spectra of each, albeit in different relative abundances. Since the relative abundances of ions in PAN MS/MS spectra are variable between replicate analyses (Section 2.2.8.3), the proportion of the MS/MS TIC provided by a particular ion cannot be correlated with the probability of that ion being formed during CID. As such, it is impossible to make any specific assignments on the basis of the low m/z ions observed in the PAN MS/MS spectra of the lipids of *S. solfataricus* at the present time.

3.2.4.7. Distributions and significance

The major lipid cores identified in *Sulfolobus* include those reported previously for each species (De Rosa and Gambacorta, 1988a; Sturt *et al.*, 2004), with the MS/MS spectra of many of these components being consistent with the known structures of the lipid cores. A number of isobaric species were also resolved, either chromatographically or, in the case of co-eluting lipids, *via* product ions in the MS/MS spectra, indicating a significant degree of isomerism within the tetraether lipid cores of *Sulfolobus*. Although MS/MS cannot provide full elucidation of the structures of these isomers, it does allow the rings within each structure to be located to particular isoprenoid chains. The isomers observed include minor isobars ($GDGT_{min}$ = **14'**-**18'**, **17''** and **17''***) to the major GDGT lipid cores ($GDGT_{maj}$ = **2**, **14**-**21**, **15*** and **19***) that contain the same number of cyclopentyl rings. The rings in

GDGT' lipids are distributed between the two composite isoprenoid chains in an equivalent manner to the GDGT isomers, whereas the rings in GDGT* lipids are distributed differently between the two moieties by comparison to the GDGT lipids. GDGT'' isomers cannot be as easily categorised. Although some of the isomers, specifically **15*** and **17''**, have been observed previously (Thurl and Schäfer, 1988; Hopmans *et al.*, 2000; Schouten *et al.*, 2000; Pancost *et al.*, 2001b; Damsté *et al.*, 2002b), several appear to be novel structures.

The presence of GDGTs **15*** and **19*** as isomeric forms of known lipids **15** and **19** in all three *Sulfolobus* strains investigated is particularly surprising given that degradative analysis of the lipids of *S. solfataricus* strain MT-4 (De Rosa *et al.*, 1980c; 1983b) did not identify either GDGT* lipid. The profiling of the MT-4 strain included isolation of GDGT lipids containing 2 or 6 rings, followed by ether cleavage and GC-MS/^{13}C-NMR analysis of the resulting isoprenoid hydrocarbons. In both cases, only one isoprenoid hydrocarbon was liberated, with a monocyclic biphytane released from the two-ringed GDGT and a tricyclic biphytane released from the six-ringed GDGT (De Rosa *et al.*, 1980c). The possibility that the inferred neutral losses of [40_0] and [40_2] dienes from **15*** and [40_2] and [40_4] dienes from **19*** actually relate to losses of one ringed or three ringed monoenes or trienes from co-eluting isomers **15** or **19**, respectively, was considered. Since losses of [40_0], [40_2] and [40_4] chains as dienes were observed from GDGTs **2**, **14**, **16**, **18** and **20**-**21** as expected, this possibility can be rejected. As such, it seems likely that the GDGT* lipids are authentic components of *S. solfataricus* P2, *S. acidocaldarius* MR31 and *S. shibatae* B12. It is currently unclear if close genetic relative to these three species, *S. solfataricus* MT-4, does not express GDGT* lipids or if they were overlooked during lipid profiling of the organism. It is notable, however, that the viable growth range of the MT-4 strain (63-89^{o}C; De Rosa *et al.*, 1975) has both a higher maximum and minimum value than the other strains, possibly reflecting its status as an outlier.

GDGT **17''** was identified as a core lipid component of *S. solfataricus* strains MT-4 and P2 (Hopmans *et al.*, 2000; Ellen *et al.*, 2009; Pitcher *et al.*, 2010) and tentatively suggested to have a GDGT[40_1,40_3] structure (Hopmans *et al.*, 2000). Purification of the component by preparative LC-MS followed by ^{1}H and ^{13}C NMR analysis led to a

revision of the structure, with the authors suggesting it to be an isomer of lipid **17** in which two cyclopentyl rings were located in the same positions on each chain, but with the two glycerol groups now arranged *syn* as opposed to *anti* to one-another (Damsté *et al.*, 2002b). The MS/MS analyses reported here, however, indicate the presence of both **17''** (GDGT[40_2,40_2]) and co-eluting **17''*** (GDGT[40_1,40_3]), the latter not being noted during the NMR studies. A late eluting isomer, **18''**, that was also observed in LC-MS analysis of the core lipids of *S. solfataricus* strain MT-4 (Hopmans *et al.*, 2000), was not observed in the extracts of the three *Sulfolobus* strains described above. The authors proposed that this lipid may comprise a GDGT[40_1,40_4] structure. LC-MS/MS of the lipids of the three *Sulfolobus* species in the present study did not provide any evidence for combinations of isoprenoid chains in GDGT lipids in which the numbers of cyclopentyl rings of the two alkyl chains differ by more than two. As such, while a GDGT[40_1,40_4] structure would seem unlikely, the nature of **18''** could almost certainly be clarified by LC-MS/MS analysis of the *S. solfataricus* MT-4 lipid extract.

Although the isomerism exhibited by GDGT' lipids, which contain the same number of rings on each isoprenoid chain as their related GDGT counterparts, could relate to any number of structural variations, there is literature precedent to suggest that positional isomerism of rings within the isoprenoid chains may be responsible. DeLuca *et al.* (1986) reported that ether cleavage of the lipid cores of *Thermoplasma acidophilum* not only yielded isoprenoid hydrocarbons containing zero, one and three Cp rings, located in the typical positions reported for the chains in GDGT lipids expressed by *S. solfataricus* strains P2 and MT-4 (De Rosa *et al.*, 1977b; 1980) but also yielded a pair of isomeric hydrocarbons containing two Cp rings (Fig. 3.21). The rings in the first isomer were located in the same positions as reported for *S. solfataricus* (De Rosa *et al.*, 1977b; 1980c), that is at the opposite ends of the chain. The second isomer, on the other hand, had both rings located at the same end of the chain. Identical hydrocarbon isomers were reported following degradation of the core lipids of two species of *Picrophilus* (Uda *et al.*, 2004). Ether cleavage of lipids from a microbial mat collected from a hot spring in Yellowstone National Park, USA generated three isomeric hydrocarbons suggested to contain two cyclopentyl rings (Ward *et al.*, 1985). Similar ether cleavage of the polar lipids

isolated from sandstone from the Be'eri sulfur deposit, Israel, resulted in the identification of two gas chromatographically-resolved isomers of isoprenoid hydrocarbons containing two cyclopentyl rings, which generated almost identical EI-MS spectra (Burhan *et al.*, 2002). In the latter case, the authors confirmed that the latest eluting hydrocarbon was identical to that found in the lipids of *S. solfataricus* and attributed the similarity of the mass spectrum of the earlier eluting component to stereoisomerism. During the present study, attempts to profile the isoprenoid hydrocarbons of the lipids of *S. acidocaldarius via* treatment of a lipid extract with HI-$LiAlH_4$ were unsuccessful, probably on account of the very low amounts of extract available for the degradation procedure following completion of the LC-MS/MS analyses. As such, the nature of the isoprenoid chains in $GDGT_{min}$ isomers remains unclear and consequently, their structures cannot be assigned fully.

Fig. 3.21. Isomeric isoprenoid hydrocarbons containing two Cp rings identified following HI-$LiAlH_4$ treatment of the lipids from *T. acidophilum* (DeLuca *et al.*, 1986) and two species of *Picrophilus* (Uda *et al.*, 2004).

Each core lipid was calculated as a proportion of the total core lipid using the LC-MS peak areas of GTGT lipids **3** and **68**-**69**, GDGT lipids **2** and **14**-**21**, GDGT' lipids **14'**-**18'**, GDGT'' lipid **17''**, $GDGT_M$ lipids **26** and **70**-**73** and $GDGT_M$' lipids **70'**-**72'** lipids for each species of *Sulfolobus*. Integration of peaks for all $GDGT_{maj}$ or $GDGT_{min}$ components that comprise a mixture of GDGT and GDGT* lipids included contributions from each of the co-eluting lipids. Other lipids determined in the ion chromatograms for each species were excluded from the calculations. In each extract, the most abundant tetraether core was the $GDGT_{maj}$ lipid containing four cyclopentyl rings (Fig. 3.22a-b and Fig. 3.23a). This compares favourably with previously reported profiles of the lipids of *S. solfataricus* strains MT-4 and P2, grown at 87^{o}C and 80^{o}C, respectively (Trincone *et al.*, 1989; Hopmans *et al.*, 2000; Ellen *et al.*, 2009). A *Sulfolobus* species 99.9% similar to *S. shibatae* on the basis of

16S rRNA sequences also expressed lipid **17** as the major GDGT core (Murae *et al.*, 2001).

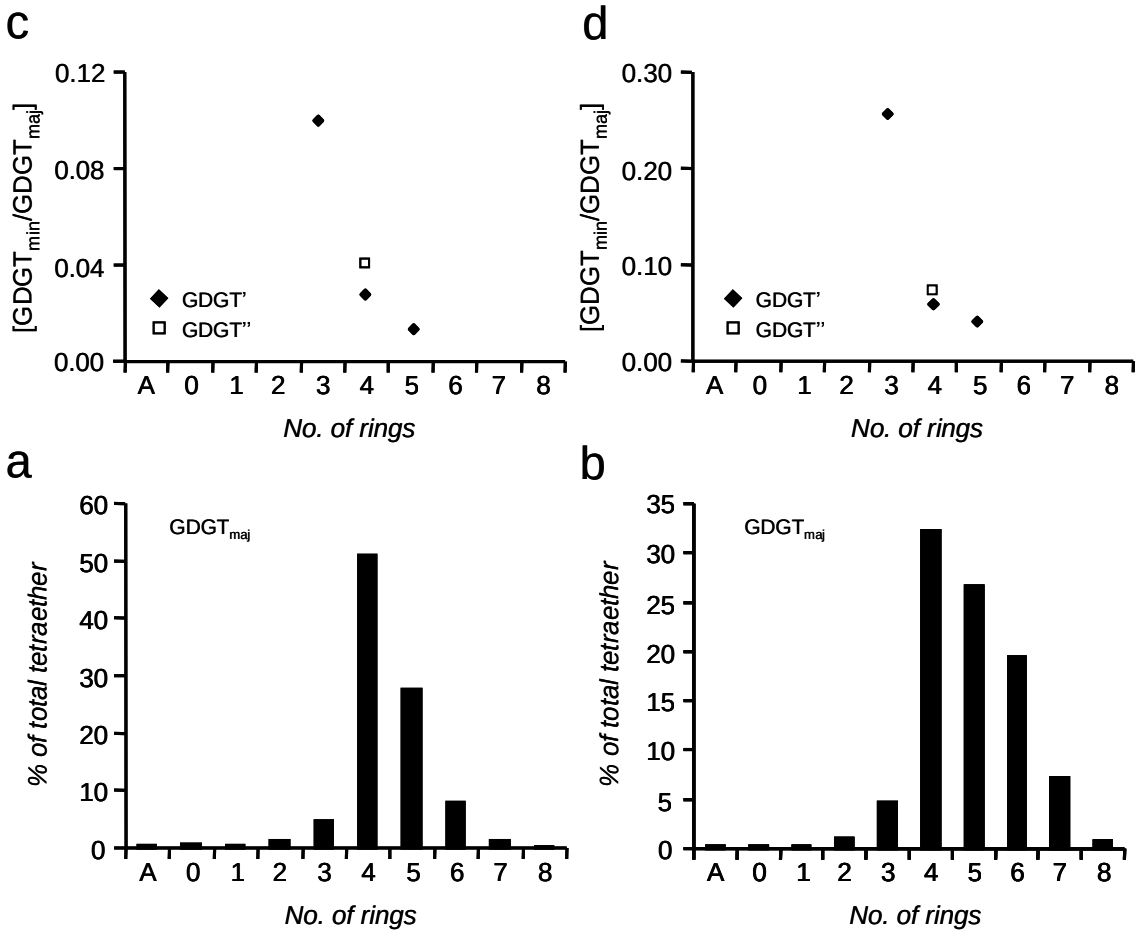
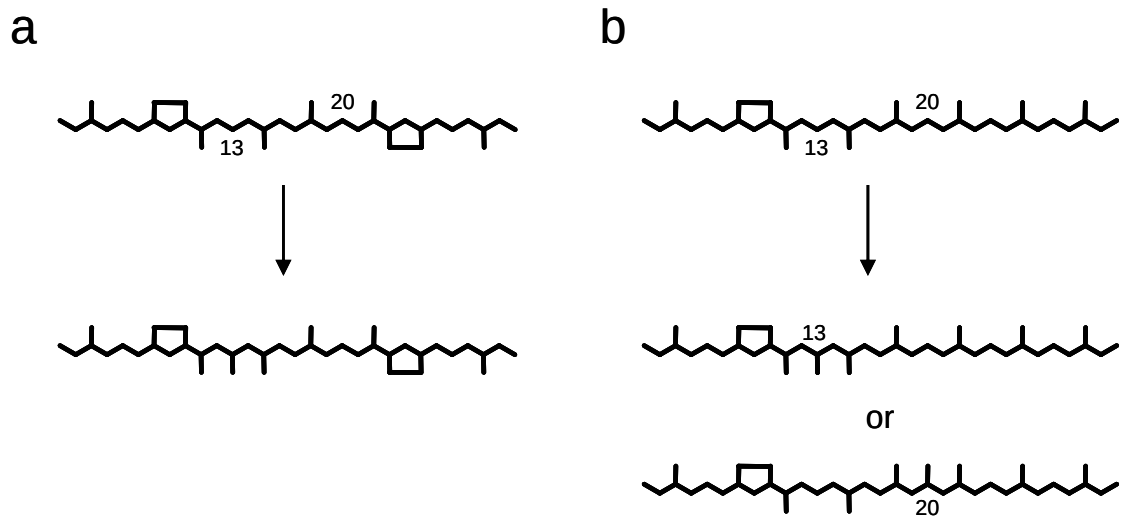


Fig. 3.22. Relative abundances of GTGT **3** (labelled A), major GDGT isomers ($GDGT_{maj}$ = GDGT + GDGT*) and minor GDGT isomers ($GDGT_{min}$ = GDGT' and GDGT'') containing the same number of cyclopentyl rings, displayed as the proportion of each $GDGT_{maj}$ component as a percentage of the total tetraether lipid or as the ratio of $GDGT_{min}$ to $GDGT_{maj}$ in *S. shibatae* (a and c) and *S. solfataricus* (b and d). Peak areas were recorded using the HCT instrument.

The degree of cyclisation in each extract was calculated as the weighted average of the number of rings in either the $GDGT_{maj}$ lipids or in both $GDGT_{maj}$ and $GDGT_{min}$. The values obtained using only the $GDGT_{maj}$ data set were found to be 4.30 ± 0.01 (n = 3) for *S. acidocaldarius*, 4.40 for *S. shibatae* and 4.88 for *S. solfataricus*. The values calculated with $GDGT_{min}$ lipids included are marginally less in each case. The

degree of cyclisation of the lipids for *S. solfataricus* P2, grown here at approximately 75°C, is higher than that determined previously for this strain (4.06) following growth at 80°C (Ellen *et al.*, 2009). This appears to be at odds with the decrease in the degree of cyclisation observed for several other organisms (Uda *et al.*, 2001; 2004; Lai *et al.*, 2008; Shimada *et al.*, 2008), including *S. solfataricus* strain MT-4 (De Rosa *et al.*, 1980b), when grown at lower temperature. Plots of the ratio of $GDGT_{min}$ to $GDGT_{maj}$ lipids, containing the same number of cyclopentyl rings (Fig. 3.22c-d and Fig. 3.23c) indicate that, for all three organisms, the largest ratio is observed for lipids containing three cyclopentyl rings. In each extract, GDGT **17''** is present in greater abundance than GDGT **17'**.



Scheme 3.1. Putative methylation of ether-bound biphytanyl chains containing: a) an even number of Cp rings and chemically equivalent C-13 and C-20 chain positions; b) an odd number of Cp rings and chemically inequivalent C-13 and C-20 positions. In the interests of clarity, the glycerol moieties and second alkyl chain within each GDGT or $GDGT_M$ core are not shown.

The structures of three $GDGT_M$ lipid cores were probed by MS/MS analysis of *S. acidocaldarius*, revealing them to contain one typical ($[40_2]$-$[40_4]$) and one methylated ($[41_2]$-$[41_3]$) Cp ring-containing biphytanyl chain. Other components are tentatively assigned as $GDGT_M$ lipids on the basis of their MS spectra and retention times relative to known isoprenoid GDGT lipids. The presence of $GDGT_M$ **26** in the

extract of *S. acidocaldarius*, coupled with the consistently shorter retention times of $GDGT_M$ lipids than their GDGT counterparts suggests that the C_{41} alkyl chains in **70**-**73** may also be methylated at the C-13 chain position, although this cannot be determined unequivocally by MS/MS analysis. Further support for this interpretation comes from the identification of a C_{41} alkane containing two Cp rings in the products of chemical degradation of a lipid extract from the Be'eri sulfur mine, Israel (Burhan *et al.*, 2002). Although *S. acidocaldarius* grows on sulfur, the geology of the Be'eri deposit indicates that conditions suitable for growth of thermophiles, namely high temperatures, were never *in situ* at the site (Nissenbaum and Kaplan, 1966; Druckman *et al.*, 1994). Consequently, the C_{41} hydrocarbon must have originated from an organism which can grow at a temperature below the known viable growth range of *S. acidocaldarius* (Burhan *et al.*, 2002). The additional carbon atom in this compound was located between the C-11 and C-22 positions in the alkyl chain and the two Cp rings were confirmed to be in positions identical to those of the hydrocarbons found in GDGTs **16**-**18**. Furthermore, as discussed above, C_{81} isoprenoid tetraacids are methylated at the C-13 position and contain Cp rings at equivalent positions to those observed in the C_{40} biphytanyl chains of GDGTs **20**-**21** (Lutnaes *et al.*, 2007). The previous identifications provide strong indications of the most likely position in the lipids of *S. acidocaldarius* of the additional carbon atom in the C_{41} alkyl chains which contain an even number of rings (i.e. zero, two or four), as both C-13 and C-20 chain positions are chemically equivalent (Scheme 3.1a). C_{41} chains containing an odd number of rings (i.e. one or three), where the C-13 and C-20 chain positions are no longer equivalent, cannot be similarly assigned as methylation at either site would give rise to structurally distinct isobars (Scheme 3.1b). A further complication also becomes evident when considering how the C_{41} chains may be arranged and/or orientated within a $GDGT_M$ lipid core. In all archaeal GDGT lipids for which the structure has been determined conclusively, the isopranyl chains are orientated such that the maximum number of Cp rings are proximal to the primary ether linkages (i.e. glycerol *sn*-3 or -3' positions) as opposed to the secondary ether linkages (*sn*-2 and -2') (De Rosa *et al.*, 1983b). Similarly, the additional methyl in the C_{41} alkyl chain of homocaldarchaeol (**26a**) was observed to be incorporated exclusively at a position closer to a primary ether linkage than to a secondary ether linkage (Galliker *et al.*, 1998). By contrast, homoisocaldarchaeol

(**26b**), a regioisomer of **26a** in which the two capping glycerols are arranged in a *syn* arrangement, shows selectivity, but not specificity, for incorporation of the additional carbon at the C-13 position over the C-20 position (Galliker *et al.*, 1998). As such, the possibility that $GDGT_M$ lipids **70**-**73**, as expressed by *S. acidocaldarius*, are comprised of several distinct isomeric forms cannot be discounted. Further characterisation of the $GDGT_M$ lipids, however, is required before the possibility of isomerism can be confirmed or dismissed.

The small reduction in retention times in all $GDGT_M$ lipids with respect to their co-eluting GDGT counterparts appears to be characteristic of these compounds, although this relationship does not appear to extend to $GMGT_M$ **62**, which elutes at approximately the same time as GMGT **27**. The minimal nature of the retention time shift caused by adding an additional methyl group indicates that methylation has only a small effect in increasing the hydrophobicity of GDGT lipids; much smaller than the change caused by reducing the number of Cp rings (Hopmans *et al.*, 2000). This suggests that tetraether methylation is unlikely to confer significant changes in membrane hydrophobicity to the Archaea, though it may provide an additional mechanism to govern membrane fluidity. As discussed in Section 1.3.2, some currently undescribed species of eubacteria produce GDGT lipids that contain C_{30} and/or C_{31} non-isoprenoid hydrocarbons containing 0-2 Cp rings (i.e. **34**-**42**). It has been observed that the average number of methyl branches observed in eubacterial tetraether lipid cores found in soils differs as a function of mean air temperature and soil pH, with the degree of methylation increasing with decreasing temperature and increasing pH (Section 1.6.2 and Weijers *et al.*, 2007c). This most likely represents an adaptation of eubacteria to changes in the environmental conditions to which they are subjected. As such, tetraether methylation in archaea may reflect a similar physiological response to changes in ambient temperature or pH, akin to the shifts from diether to tetraether lipids (Sprott *et al.*, 1991; Lai *et al.*, 2008; Matsuno *et al.*, 2009) or to more heavily cyclised GDGT lipids (De Rosa *et al.*, 1980b; Uda *et al.*, 2001; 2004) already known for various species of archaea (see Section 1.5).

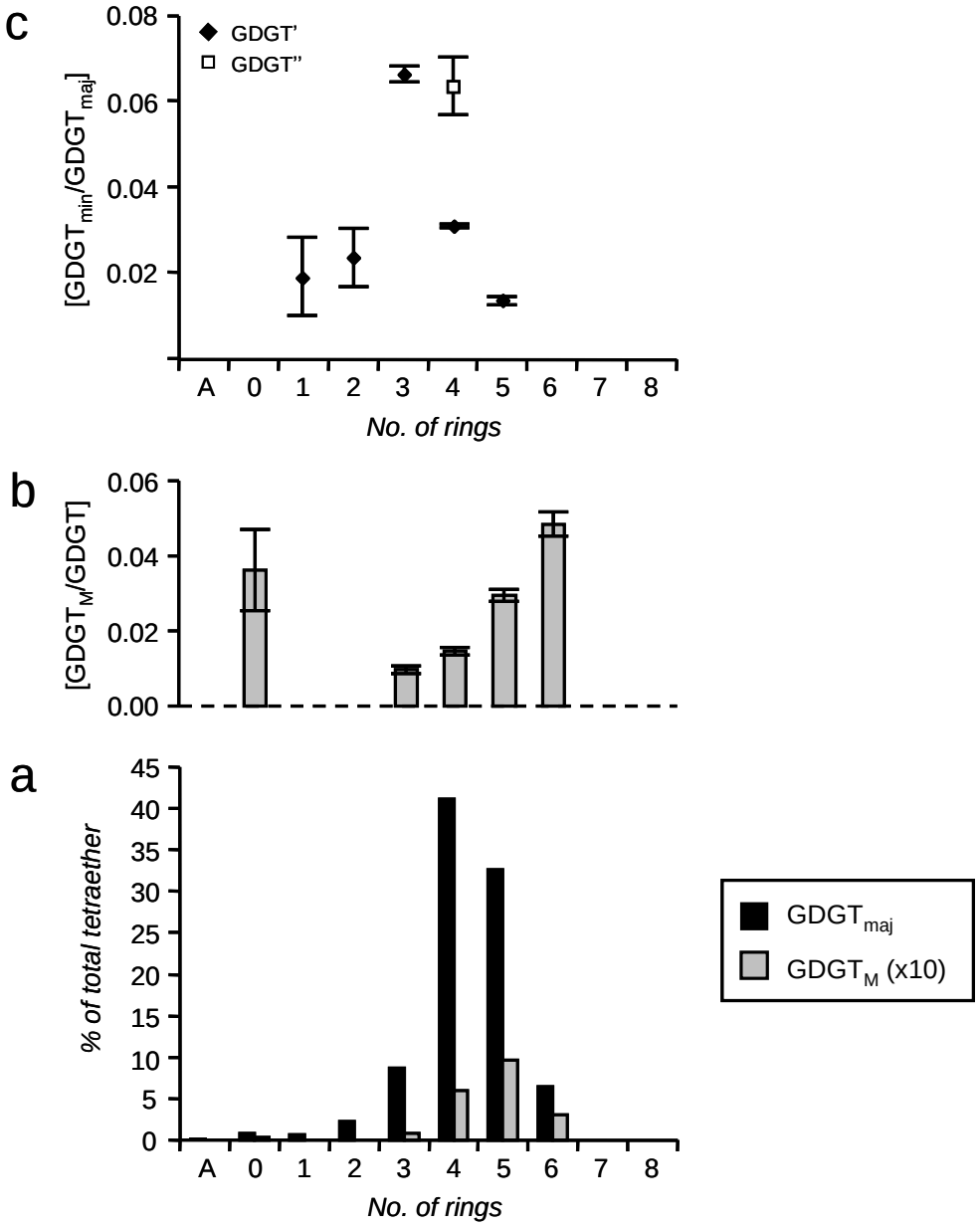
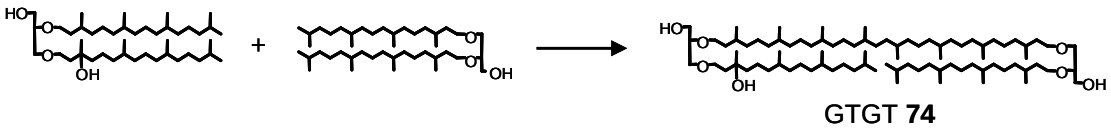


Fig. 3.23. Relative abundances of GTGT **3** (labelled A), major GDGT isomers ($GDGT_{maj}$ = GDGT + GDGT*), minor GDGT isomers ($GDGT_{min}$ = GDGT' and GDGT'') and methylated GDGTs ($GDGT_M$) containing the same number of cyclopentyl rings in *S. acidocaldarius*: a) proportion of each $GDGT_{maj}$ or $GDGT_M$ component as a percentage of the total tetraether lipid. $GDGT_M$ lipid proportions are magnified by a factor of 10 for clarity; b) ratios of $GDGT_M$ to GDGT; c) ratios of $GDGT_{min}$ to $GDGT_{maj}$. Error bars show ±1σ from the mean of analytical triplicates (LCQ).

The proportions of $GDGT_M$ lipids in *S. acidocaldarius*, calculated from peak areas in individual ion chromatograms, show an increased average cyclisation ratio of 4.70 ± 0.03 (n = 3) by comparison with the ratio observed in the GDGT lipids. Similarly, the most abundant methylated lipid is **72**, containing 5 Cp rings in total, one more than the most prominent GDGT lipid (Fig. 3.23a). By monitoring the ratio of $GDGT_M$:GDGT for tetraether species containing the same number of Cp rings in total, it is clear that the relative degree of methylation also increases as the number of Cp rings increases (Fig. 3.23b). This may indicate that more hydrophilic GDGTs (i.e. those containing more Cp rings) are methylated to a greater extent than less hydrophilic lipids, a postulate consistent with the high relative levels of methylation of hydrophilic GMGT **27** observed in MTH(ΔH), as described in Section 3.2.3.6. This may also explain why C_{81} and C_{82} homologues of C_{80} tetraacids were observed for the most heavily cyclised structures, containing 7 or 8 Cp rings, but were not apparent for less heavily cyclised structures containing 4-6 Cp rings (Smith *et al.*, 2007; Lutnaes *et al.*, 2007). The one exception to the trend of increasing $GDGT_M$:GDGT ratio with increasing ring number is that between GDGT **2** and $GDGT_M$ **26**, which show the second highest GDGT:$GDGT_M$ relative ratio in *S. acidocaldarius*, despite their structures containing no Cp rings. Lipid **2**, or an unsaturated precursor to this lipid, is suspected to be an intermediate in the formation of Cp ring-containing GDGT lipids (De Rosa *et al.*, 1980c; 1983b; Weijers *et al.*, 2006a) and in the formation of $GDGT_M$ **26** (Section 3.2.3.6 and Galliker *et al.*, 1998). The high relative abundance of **26** could, perhaps, reflect a route to the lipid which requires only a single biosynthetic modification to **2** (i.e. methylation) as opposed to the two successive modifications (i.e. methylation and Cp ring formation or *vice versa*) required for **70**-**73**.

The detection of ring-containing $GDGT_M$ lipids in *S. acidocaldarius* may also explain the presence of a triad of unidentified tetraether lipids with $[M+H]^+$ measured at *m/z* 1308, 1306 and 1304 or $[M+Na]^+$ measured at *m/z* 1330.6 in members of the *Euryarchaeotal* order *Thermoplasmatales* (Macalady *et al.*, 2004; Shimada *et al.*, 2008), most probably **71**, **72** and **73**. If the identity of the lipids in these species is confirmed, the presence of Cp ring-containing $GDGT_M$ lipids in *Euryarchaeota* would indicate that tetraether methylation occurs widely throughout

the archaeal domain. Whether this relates to retention of the biosynthetic pathways responsible for the lipid methylation following the divergent evolution of the kingdoms *Crenarchaeota* and *Euryarchaeota*, represents a convergent evolutionary adaptation or is a consequence of lateral gene transfer between the two domains at a later date is unclear. It is of note, however, that the genome sequence of *Thermoplasma acidophilum* suggests an extensive history of lateral gene transfer from *S. solfataricus* inhabiting the same thermal environments (Ruepp *et al.*, 2000).



Scheme 3.2. Putative route to formation of GTGT **74** during tetraether biosynthesis. Although saturated diether precursors are shown, the precursors may be unsaturated and/or phosphorylated. Similarly, the position of the hydroxyl group located on the phytanyl chain in both the hydroxylated GDD precursor and GTGT product is tentative and provided for illustration only.

GTGT **3** is thought to form *via* coupling of two molecules of GDD **1** *en route* to formation of GDGT **2** (Section 1.2.1 and de la Torre *et al.*, 2008). The tentative assignment, on the basis of LC-MS data only (i.e. with no supporting MS/MS data), of ring-containing GTGT lipids **68** and **69** in *S. solfataricus* and *S. shibatae* indicates that cyclopentyl ring-incorporation into tetraether lipid cores can occur even if the macrocycle has not been completed, as indicated previously (de la Torre *et al.*, 2008). The identification of GTGT **74** also reveals important information regarding the biosynthesis of tetraether lipids. As *S. acidocaldarius* is known to form C_{20} hydroxyarchaeol cores (Sprott *et al.*, 1997), it is likely that GTGT **74** is formed when a molecule of the hydroxyarchaeol (most probably **7** or **8**) has been scrambled into this process, replacing one of the two molecules of **1** involved in the coupling (Scheme 3.2). This would suggest that, apart from **1**, other functionalised GDD lipids can be coupled by the Archaea to form tetraether lipids. One potential caveat to the proposed route to **74** is the apparent absence of a biphytanyl chain-hydroxylated derivative of GTGT **2** in *S. acidocaldarius*, expected if the bond conjoining the two GDD lipids was formed between a hydroxyphytanyl and a

phytanyl chain as opposed to between two phytanyl chains. Similarly, no component consistent with a chain-hydroxylated derivative of GDGT **2** was observed in the base peak chromatogram of the lipid extract from this organism, nor was one observed in the MS spectrum obtained during the post-analysis column backflushing procedure. Given the large relative ratio of the fully-formed macrocycle **2** to the intermediary lipid **3** (approximately 11:1) in *S. acidocaldarius*, formation of **74** *via* simple random scrambling of hydroxyarchaeol into the tetraether biosynthetic pathway would be expected to generate equivalent, or even greater levels of hydroxycaldarchaeol. Although it is tempting to suggest that these observations may indicate that hydroxyphytanyl chains cannot form new C-C bonds during tetraether biosynthesis, with hydroxyarchaeol carried through the process on account of the ability of phytanyl chains to couple, there is some preliminary evidence to suggest that hydroxycaldarchaeols ($[M+H]^+$ *m/z* = 1317) can be formed by unidentified archaea (discussed briefly in Section 5.2.3.4). As such, the presence of **74** as the sole identifiable hydroxylated tetraether lipid in *S. acidocaldarius* may not be incidental and may be representative of regulated formation of this particular lipid, especially considering that it is present in greater abundance than GTGT **3** in the extract from the organism (approximately 6.8:1 relative ratio on the basis of peak area, assuming identical molar response). Interestingly, both *sn*-2 (**7**) and *sn*-3 (**8**) hydroxyarchaeols are susceptible to degradation under strong acidic conditions, degrading largely to glycerol monophytanyl monoethers, chain-etherified methoxyarchaeols or monounsaturated archaeols following treatment with 2.5% methanolic HCl at 50^{o}C for 2 h (Ferrante *et al.*, 1988; Sprott *et al.*, 1990; Ekiel and Sprott, 1992). Since more harsh acidic conditions were used during extraction of the *S. acidocaldarius* strain in the present study, it is possible that some, but not all, of GTGT **74** may have been destroyed during the extraction procedure. It is notable that the degradates which may be expected to be formed from **74** given the known degradation pathways of the hydroxyarchaeols, such as unsaturated or methoxylated derivatives of acyclic caldarchaeol, were not identified during the LC-MS/MS analyses. Nevertheless, the quantification of the lipid from the LC-MS ion chromatograms may be an under-representation of the true proportion of **74** expressed by the organism.

A number of previous reports have used the theoretical isotope ratios of the lipids to deconvolute ion clusters relating to GDCT lipids containing different numbers of rings in positive ion (Bode *et al.*, 2008) and negative ion (Murae *et al.*, 2001; 2002) time of flight mass spectra. Although a similar approach could be applied to the average mass spectra of GDCT lipid cores eluted during the post-analysis column backflushing procedure for the *Sulfolobus* species, the fact that some GDCT lipid material is still retained on column even after 10 minutes of backflushing suggests that the spectra may be distorted in favour of less well retained GDCT lipids of lower polarity (i.e. the less heavily cyclised GDCT lipids may be over-represented in the spectrum). As such, quantification of the components in this manner may not accurately reflect the true GDCT lipid profile and, consequently, is not provided.

3.3. Conclusions

Detailed profiles of the tetraether lipid cores of *Methanothermobacter thermautotrophicus* and of three species of *Sulfolobus* have been characterised by LC-MS/MS analysis and include a complex series of isoprenoid ether lipids. The majority of the components have been identified previously, with structures showing highly conserved features including glycerol backbones, a regular number of carbon atoms in the isoprenoid chain and a certain degree of symmetry regarding the location of cyclopentyl rings. MS/MS spectra of the components are consistent with the structures assigned, supporting the feasibility for use of the technique in future studies to profile novel strains of archaea. New structures have, however, also been unveiled during the work indicating that the lipid signatures of archaea may be far more complex than believed previously. The novel structures provide evidence for significant isomerism of the known tetraether core lipids. Furthermore, their presence in the archaea examined highlights the operation of unusual biosynthetic pathways in some of the species.

A series of methylated tetraether lipids have been identified in organisms from separate archaeal kingdoms, suggesting that lipid methylation, despite being a minor process, may be a widespread modification utilised throughout the domain. Whether this process has an early lineage in archaeal evolution, is the result of a more recent

horizontal gene transfer between the two kingdoms or represents independent evolution of the process in each kingdom is unclear. As, however, the genome sequences of MTH(ΔH) and *S. acidocaldarius*, which both express significant levels of methylated lipids, are available, this issue will hopefully be clarified in future as more becomes known about the genes encoding for the functional enzymes specific to tetraether lipid formation and subsequent methylation. It remains to be seen whether members of the other proposed archaeal kingdoms, the *Korarchaeota* (Barns *et al.*, 1996; Elkins *et al.*, 2008), *Nanoarchaeota* (Huber *et al.*, 2002) and *Thaumarchaeota* (Brochier-Armanet *et al.*, 2008; Spang *et al.*, 2010), will also be shown to synthesise $GDGT_M$ or $GMGT_M$ lipids. Since this family of compounds have been overlooked in previous archaeal lipid profiles, it may be necessary to assess which other species reported exhibit these lipids before their biological significance and capacity as biomarkers can be determined. As such, the procedural screening for these lipids in all novel cultures identified in the future is advocated.

Although evidence has been presented to suggest that the methylation step in tetraether biosynthesis occurs after completion of the macrocycle, the extent of methylation of GDGT lipids appears to differ with the number of cyclopentyl rings expressed. Likewise, GDGT and GMGT lipids appear to be methylated to different degrees, suggesting that tetraether methylation may depend on the hydrophilicity of the lipid precursor. Alternatively, if the methylation step precedes cyclopentyl ring incorporation during biosynthesis, occurring specifically on caldarchaeol or an unsaturated precursor of this lipid, the observed trend may, instead, be the result of more extensive ring incorporation into the methylated structure. Studies as to whether the physiological stresses known to govern biosynthesis of GDGT lipids (e.g. temperature or pH) may also invoke subtle changes in the degree of lipid methylation and, consequently, composition of the cellular membrane will be key if the roles of $GDGT_M$ and $GMGT_M$ lipids within membrane are to be understood.

Other minor components found in the organisms include lipids which appear to be structurally-modified derivatives of tetraether lipids **2** or **15-19**. Some of the derivatives elute at elutropic strengths higher than the maximum strength used in many previous LC-MS and LC-MS/MS analyses of tetraether core lipids (Hopmans *et al.*, 2000; Escala *et al.*, 2007; Schouten *et al.*, 2007b). As such, they may represent

biomarkers that have been overlooked in organisms and environmental matrices profiled previously using these methodologies. Alternatively, the minor components may be artifacts of the extraction and/or analytical procedures used in the studies. It is imperative that any structural modifications which may alter tetraether lipid profiles, be they naturally or artificially created, are properly assessed if accurate conclusions regarding lipid abundances are to be reached. In *Methanothermobacter* and *Sulfolobus* species, these components were typically found in levels that are largely inconsequential and, as a result, they have little bearing on the conclusions drawn. Their presence in the base peak chromatograms does, however, emphasise the need for use of higher final polarities in LC-MS/MS elution programs during tetraether lipid analysis. LC-MS/MS analyses of *Sulfolobus* lipid extracts have shown that GDCT lipids are retained on-column even if a final eluent composition of 15% IPA is used. As such, this mobile phase composition may be of insufficient elutropic strength to elute all tetraether lipid components in core lipid extracts, particularly if highly polar components are present. Until an improved chromatographic elution program is devised, monitoring of the column backflushing procedure by MS to screen for such lipids is imperative.

To summarise, LC-MS/MS appears to be a very useful technique for profiling the lipid cores of thermophilic archaeal species, providing information regarding the number of carbon atoms and number of cyclopentyl rings in the isoprenoid chains. No other analytical technique is currently able to provide comparable information, for complex mixtures of intact core lipids without prior extraction.

Chapter 4

Lipid core profiles of hyperthermophilic isolates from New Zealand hot springs

4.1. Introduction

4.1.1. Isolates from New Zealand hot springs

Niederberger *et al.* (2006; 2008) recently isolated several novel strains of *Crenarchaeota* from hot springs situated within the Taupo Volcanic Zone at Rotorua and Tokaanu, New Zealand. The first of these organisms to be fully described, *Ignisphaera aggregans* strain AQ1.S1^T, represents a strictly anaerobic, hyperthermophilic heterotroph that is viable at temperatures between 85-98^{o}C (optimal at 92-95^{o}C) and pH values between 5.4-7.0 (optimum pH 6.4). Other close phylogenetic relatives to the type strain (i.e. AQ1.S1^T) of *I. aggregans* were also isolated from the hot springs, including *Ignisphaera* sp. strain Tok37.S1 (Niederberger *et al.*, 2006). On the basis of the 16S rRNA gene sequences of AQ1.S1^T and Tok37.S1, it appears likely that the *Ignisphaera* possibly represent a novel genus within the order *Desulfurococcales*, although a second 16S rRNA amplicon from Tok37.S1 indicates that this strain may comprise a mixture of two different coccal organisms (Niederberger *et al.*, 2006). Tok37.S1 grows over a slightly extended temperature range relative to AQ1.S1^T (Morgan, 2009; personal communication). No profiles of the core lipids of either strain have been reported. A draft genome for the type strain of *I. aggregans* is available (Department of Energy Joint Genome Institute).

During the studies discussed above, a *Pyrobaculum* species (strain AQ1.S2) was isolated from a hot spring at Rotorua, where it was the dominant organism (Niederberger *et al.*, 2008). The organism is 98.9% similar to *Pyrobaculum calidifontis* on the basis of 16S rRNA gene sequences, although *P. calidifontis* is aerophilic whereas *Pyrobaculum* AQ1.S2 does not appear to grow under even trace amounts of oxygen (Niederberger *et al.*, 2008). Although no lipid profile is available for *Pyrobaculum* strain AQ1.S2, profiling of other members of the genus has shown that three close phylogenetic relatives (*P. aerophilum*, *P. islandicum* and *P. organotrophum*) can express GDD **1** and GDGTs **2** and **14**-**17** (Trincone *et al.*, 1992; Volkl *et al.*, 1993).

A *Desulfurococcus* species (strain Ket10.S1), isolated from a hot spring at Ketatahi, New Zealand and assigned to the genus on the basis of phylotypic observations as opposed to molecular genetic typing, has also been reported (Patel *et al.*, 1986). The organism is heterotrophic, with a viable temperature range of 75-98^{o}C (optimal 90^{o}C) and a viable pH range of 4.5-8.0 (optimal pH 6.5). The core lipids of the organism have not been previously reported. *Desulfurococcus mobilis* is the only member of the genus currently to have been profiled for core lipids, with this organism having been found to produce GDD **1** and GDGT **2** (Lanzotti *et al.*, 1987).

4.1.2. Aims

The major aim of the work described in this chapter was to provide detailed profiles of the lipid cores of the hyperthermophilic hot spring isolates *via* analysis of lipid extracts using LC-MS/MS. It was envisaged that the lipid profiles taken from the cultivated organisms would allow assessment of possible biomarkers for each species. This could, in turn, potentially allow each archaeal species to be identified on the basis of lipid cores found in the hot spring environments from which they were isolated.

4.2. Results and Discussion

4.2.1. Culture conditions, lipid extraction and analysis

Each hot spring isolate was grown in culture by either Hugh Morgan or Christine Tan (University of Waikato, NZ) using physiological conditions (Table 4.1) close to the known optimal conditions for growth in order to generate sufficient cellular material for analysis. The lipid cores were extracted from each organism *via* acidic methanolysis of a cell pellet and analysed *via* LC-MS/MS using method B. Classification of observed components as dialkyl and monoalkyl tetraether lipids was based on the MS/MS spectrum of each lipid: GDGT and $GDGT_M$ lipids show abundant ions in the range *m/z* 550 to 760 that do not feature prominently in the spectra of GMGT lipids, as described in Chapter 2. The nomenclature used for tetraether lipids and neutral losses from these lipids during CID are as described in Section 3.2.2.4.

Table 4.1. Organisms used in the studies, with the physiological temperature and pH values used for growth of each.

Organism	Strain	Growth Conditions		References
		Temperature / ^oC	pH	
Ignisphaera aggregans	$AQ1.S1^T$	95	6	Niederberger *et al*., 2006; 2008
Ignisphaera sp.	Tok37.S1	90	6	Niederberger *et al*., 2006
Pyrobaculum sp.	AQ1.S2	90	6	Niederberger *et al*., 2008
Desulfurococcus sp.	Ket10.S1	90	6	Patel *et al*., 1986

4.2.2. *Ignisphaera aggregans* AQ1.S1^{T}

4.2.2.1. Ether lipid profile

Following LC-MS/MS analysis, five major and three minor GDGT core lipids were identified in the ion chromatogram for the *I. aggregans* AQ1.S1^{T} extract (Fig. 4.1). The major lipids (**2** and **14**-**17**) were identified by comparison of MS/MS spectra with those of known lipids present in extracts from *Sulfolobus* (Section 3.2.4.5 and Table 3.4). The minor lipids (**15'**, **16'** and **17'**) were observed to elute slightly earlier than their corresponding major isobar (i.e. **15**, **16** and **17**) and, as with the species of *Sulfolobus*, represent regioisomers of the major lipids. In addition, minor lipid **14'**, which could not be observed clearly in the base peak chromatogram, was identified in a reconstructed ion chromatogram for *m/z* 1299.5-1300.5 (Fig. 4.1b). The MS/MS spectra of **16'**-**17'** exhibit product ions at *m/z* values identical to those observed for these components in the *Sulfolobus* extracts (Table 3.4). On the basis of the $[M+H]^{+}$ and MS/MS spectrum of lipid **14'**, the component appears to be an isomer of **14** (Table 4.2). CID of the major GDGT lipid with $[M+H]^{+}$ at *m/z* 1297.8 generates product ions formed from loss of alkyl chains containing one Cp ring (-556 Da; *m/z* 741), as is expected for GDGT **15**. Interestingly, however, additional product ions formed from losses of alkyl chains containing zero (-558 Da; *m/z* 739) and two (-554 Da; *m/z* 743) Cp rings are also observed in the resulting MS/MS spectrum (Fig. 4.2). A GDGT lipid containing such a combination of alkyl chains has been reported in *T. tenax* (Thurl and Schäfer, 1988) and also in a number of aquatic sediments following chemical degradation of isolated tetraether fractions to the composite isoprenoid chains (Schouten *et al*., 2000; Pancost *et al*., 2001b). As such, *I. aggregans* most likely produces a structural isomer to **15**, **15***, which has a zero ring-two ring isoprenoid hydrocarbon pairing and co-elutes under the chromatographic conditions employed. These lipids are consistent with those observed in the *Sulfolobus* species (Section 3.2.4.5). The MS/MS spectrum of **15'** in *I. aggregans* similarly suggests that it may also be accompanied by a co-eluting isomer, **15'***, and that the two isomers comprise one structure in which one of the two Cp rings is present in each of the isoprenoid chains and a second in which one chain contains both rings and the other contains none (Table 4.2).

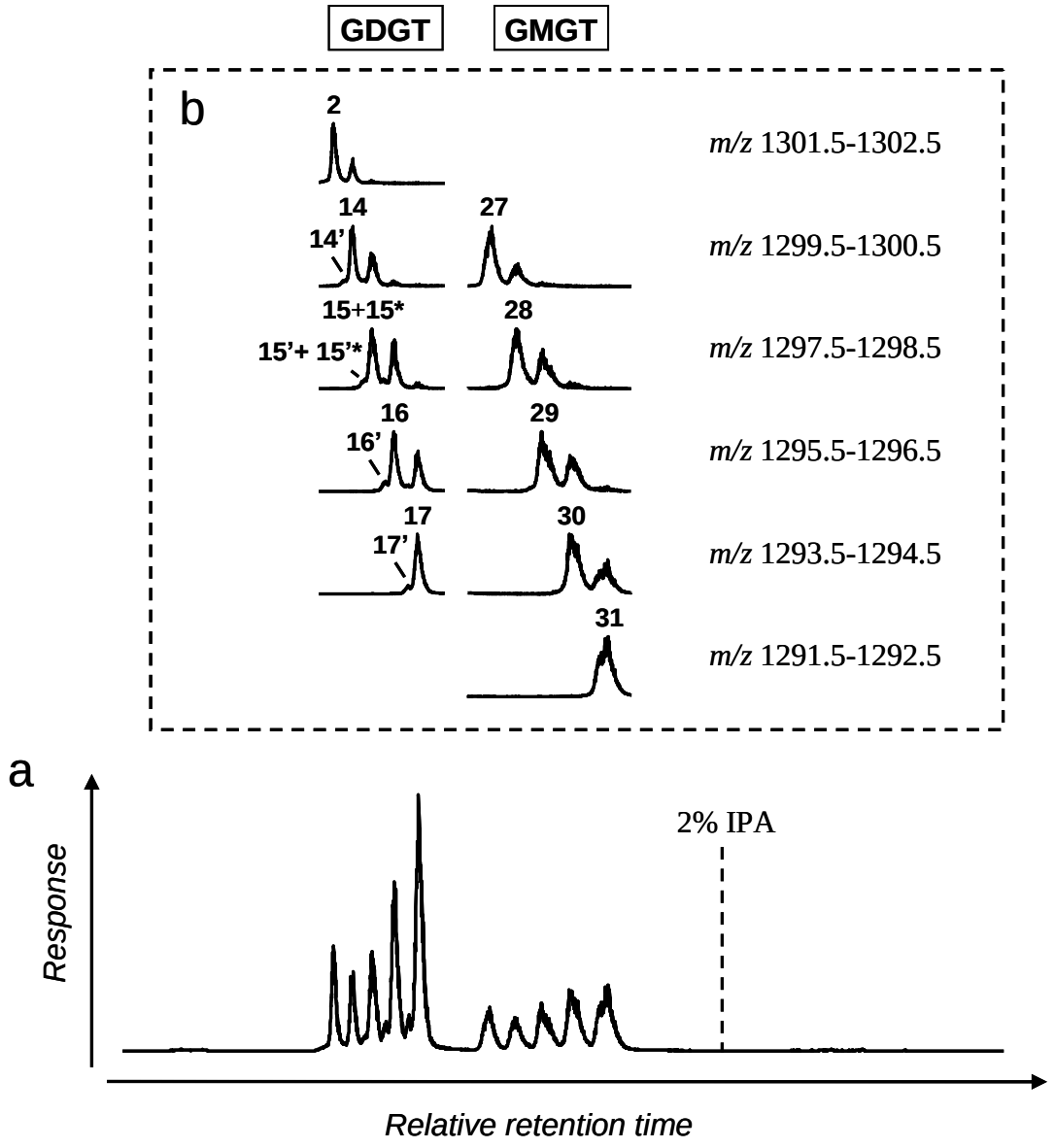


Fig. 4.1. HCT LC-MS ion chromatograms for a lipid extract from *I. aggregans* AQ1.S1^{T}: a) base peak chromatogram; b) individual ion chromatograms for tetraether lipids. The retention time relating to an eluent composition of 2% IPA is highlighted.

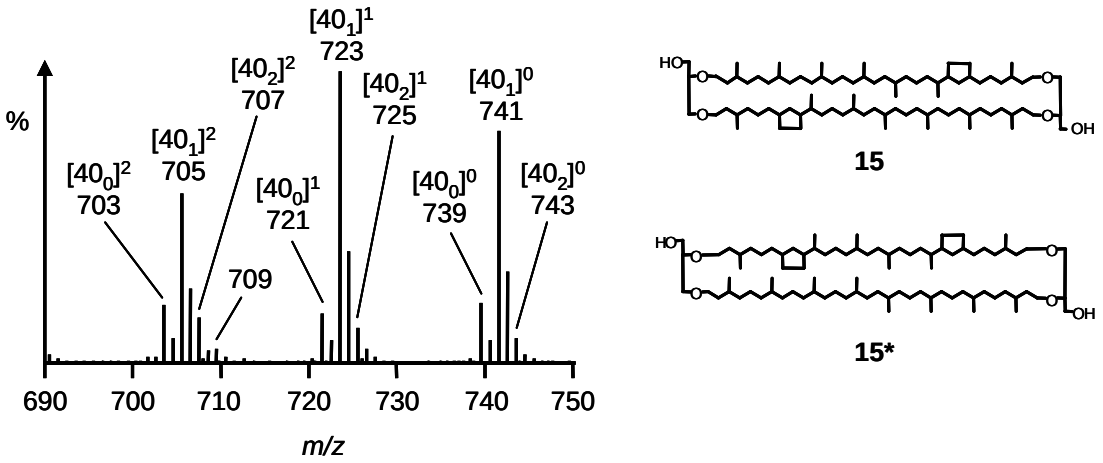


Fig. 4.2. Partial MS/MS spectrum (HCT) of GDGT **15**/**15*** ($[M+H]^+$ *m/z* = 1297.8) from *I. aggregans* AQ1.S1^T.

Previously, GMGTs **28**-**31** have only been assigned tentatively in the thermoacidophilic *Euryarchaeote A. boonei* (Schouten *et al*., 2008b). By contrast, GMGT **27** has been identified in a number of (hyper)thermophilic archaea (see Sections 1.4.1 and 3.2.3 and Morii *et al*., 1998; Sugai *et al*., 2004; Schouten *et al*., 2008b) and in sediments from non-thermophilic environments (Schouten *et al*., 2008a). The MS/MS spectrum of the latest eluting of the three lipids with $[M+H]^+$ at *m/z* 1299.8 in *I. aggregans* shows product ions to be formed exclusively from losses of water (up to four molecules) and/or C_3H_4O (up to two molecules), confirming this species to be a GMGT lipid. In addition, all of the product ions are identical to those in the MS/MS spectrum of authentic GMGT **27**, found in MTH(ΔH). In order to verify the presence of **27** in *I. aggregans* with greater confidence, the MTH(ΔH) extract was analysed before and after spiking with a small amount of the *I. aggregans* extract. The doped sample showed an enhancement of the peak for **27** relative to the un-doped sample, thereby confirming the presence of **27** in *I. aggregans*. CID of the protonated molecules of all other lipids assigned as GMGTs (i.e. **28**-**31**) generated MS/MS spectra analogous to that obtained for **27**, with product ions only being observed in the *m/z* range 1050 to 1300 (Table 4.2). Furthermore, expulsion of molecules of water and/or C_3H_4O from **27** and **28**-**31** to yield product ions within this range appears to be more favourable than the equivalent losses from GDGT lipids. For example, product ions at *m/z* 1257 (-36 Da; $2H_2O$) and *m/z* 1201 (-92 Da; $2H_2O$ + C_3H_4O) in the MS/MS spectrum of GMGT **30** (Fig. 4.3) are more

prominent than those arising from the same losses from GDGT **16** (*m/z* 1259 and *m/z* 1203; see Fig. 3.16). This difference in relative abundance most likely reflects the more limited dissociation pathways available for GMGT lipids, which cannot easily lose an isoprenoid chain, than for GDGT lipids. Unfortunately, the absence of product ions arising from losses of isoprenoid chains from GMGT lipids limits the structural information provided for these lipids by MS/MS. As such, co-eluting GMGT structural isomers, if present, are not distinguishable using this method.

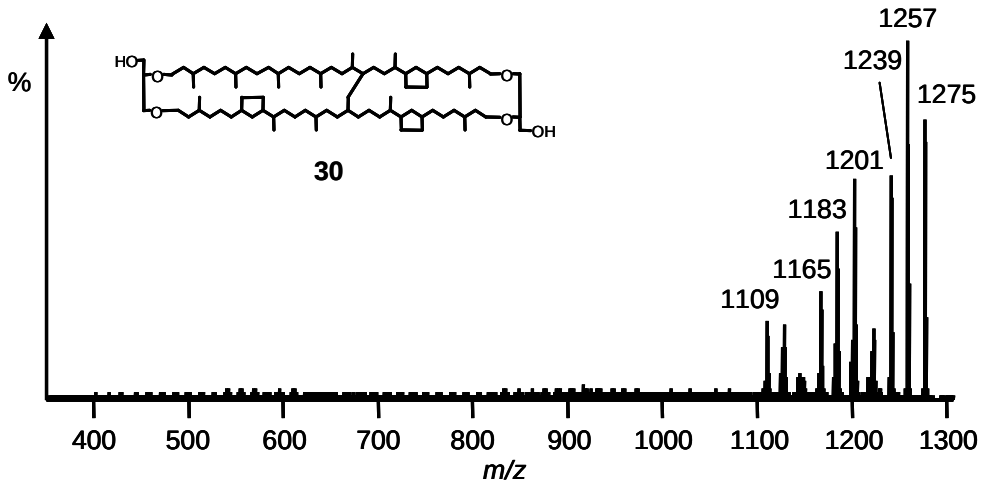


Fig. 4.3. MS/MS spectrum (HCT) of GMGT **30** ($[M+H]^+$ *m/z* = 1293.8) from *I. aggregans* AQ1.S1^T. The structure shown is a tentative assignment (Schouten *et al.*, 2008b); the position of the covalent cross-link may be different to that shown.

GTGT **3** could not be identified definitively either in the base peak chromatogram or in the appropriate reconstructed mass chromatogram (*m/z* 1303.5-1304.5) for the extract from *I. aggregans*. In order to screen for diether core lipids, which exhibit $[M+H]^+$ at *m/z* values outside the detection range used in the initial analyses, an extended mass spectral scan range (*m/z* 50-2000) was employed during repeat LC-MS/MS analyses of each extract. GDD **1** was identified in *I. aggregans* by examination of an ion chromatogram for *m/z* 653-654.5, which exhibited an early eluting component with retention time relative to **2**, $[M+H]^+$ (*m/z* 653.5) and MS/MS spectrum (containing a product ion at *m/z* 373) consistent with this lipid core.

Table 4.2. HCT MS data (measured $[M+H]^+$ *m/z* values) and partial MS/MS data (nominal *m/z* values of product ions) for GDGT' and GMGT lipid cores[a] from *I. aggregans* AQ1.S1^T.

	No. of Rings	1	2	0	1	2	3	4
Loss (Da)	Molecules Lost [b]	**14'**	**15'+15'***	**27**	**28**	**29**	**30**	**31**
-0	$[M+H]^+$	1299.8	1297.8	1299.8	1297.8	1295.8	1293.8	1291.7
-18	[1,0,0]	**1281**	**1279**	**1281**	**1279**	1277	1275	1273
-36	[2,0,0]	1263	1261	1263	1261	**1259**	**1257**	**1255**
-54	[3,0,0]	1245	1243	1245	1243	1241	1239	1237
-56	[0,1,0]	1243	1241	1243	1241	1239	1237	1235
-72	[4,0,0]			1227	1225	1223	1221	1219
-74	[1,1,0]	1225	1223	1225	1223	1221	1219	1217
-78	?			1221	1219	1217	1215	1213
-92	[2,1,0]	1207	1205	1207	1205	1203	1201	1199
-94	?			1205	1203	1201	1199	1197
-96	?			1203	1201	1199	1197	1195
-110	[3,1,0]			1189	1187	1185	1183	1181
-112	[0,2,0]			1187	1185	1183	1181	1179
-114	?			1185	1183	1181	1179	1177
-128	[4,1,0]			1171	1169	1167	1165	1163
-130	[1,2,0]			1169	1167	1165	1163	1161
-146	?			1153	1151	1149	1147	1145
-166	[3,2,0]			1133	1131	1129	1127	1125
-168	?			1131	1129	1127	1125	1123
-170	?			1129	1127	1125	1123	1121
-184	[4,2,0]			1115	1113	1111	1109	1107
-554	$[0,0,40_2]$		743					
-556	$[0,0,40_1]$	743	741					
-558	$[0,0,40_0]$	741	739					
-572	$[1,0,40_2]$		725					
-574	$[1,0,40_1]$	725	723					
-576	$[1,0,40_0]$	723	721					
-590	$[2,0,40_2]$		707					
-592	$[2,0,40_1]$	707	705					
-594	$[2,0,40_0]$	705	703					
-628	$[1,1,40_2]$		669					
-630	$[1,1,40_1]$	669	667					
-632	$[1,1,40_0]$	667	665					

[a] Data for GDGT and GDGT' lipid cores that were also identified in *Sulfolobus* are reported in **Table 3.4**.

[b] The number of simultaneous losses of each type of molecule during formation of the ion are listed as $[a,b,c_d]$, where a = water; b = C_3H_4O; c_d = biphytadiene containing "c" carbon atoms and "d" cyclopentyl rings (d = 0-4).

[c] Product ions listed in bold face represent the base peak in MS/MS.

4.2.2.2. PAN LC-MS/MS

In an attempt to obtain more information regarding the possible structures of GMGT lipids **27** and **28**-**31**, the *I. aggregans* extract was also analysed *via* PAN LC-MS/MS, allowing PAN MS/MS spectra of each GDGT and GMGT in the extract to be obtained. The PAN MS/MS spectra of the GDGT lipids were similar to those obtained for the lipids of the extract of *S. solfataricus* (Section 3.2.4.6). The PAN MS/MS spectra of GMGT lipids (e.g. **27**, Fig. 4.4) also exhibited product ions at *m/*z < 400, typically missed during conventional LC-MS/MS on account of the low mass cut-off. Unfortunately, all of the ions below *m/z* 400 were, in effect, consistent among the spectra of the GMGT and the GDGT lipid groups. Given that the abundance of each product ion within a PAN MS/MS spectrum cannot be used as a quantitative measure of the probability of formation of that ion during CID (see Section 2.2.8.3), the differences in relative ion abundances that were noted for these product ions in the spectra of each lipid cannot be interpreted effectively. As was noted for GDGT lipids, therefore, PAN MS/MS appears to provide little additional information of use in elucidation of the structures of GMGT lipids.

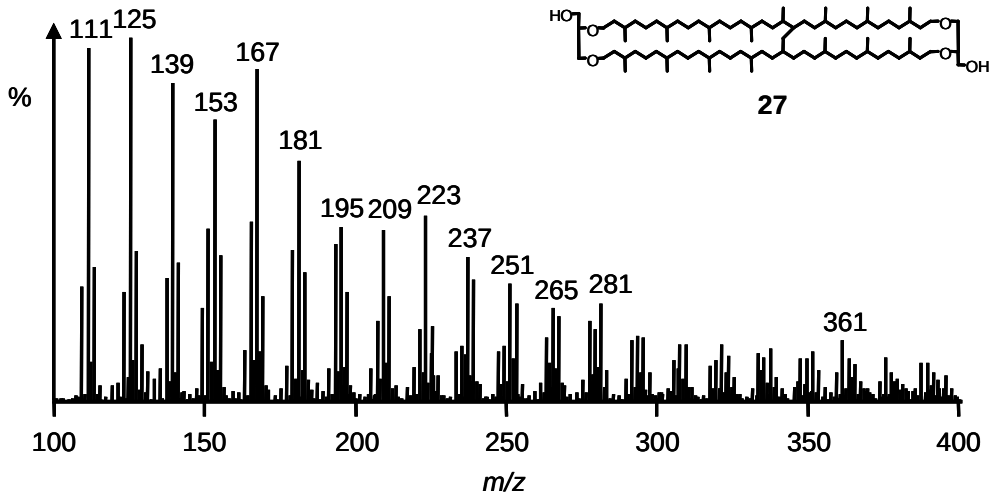


Fig. 4.4. *m/z* 100-400 region of the PAN MS/MS spectrum (HCT) of GMGT **27** ($[M+H]^+$ *m/z* = 1299.8) from *I. aggregans* AQ1.S1^T.

4.2.2.3. Distributions and significance

I. aggregans incorporates up to 4 Cp rings within its major GDGT lipid structures ($GDGT_{maj}$ = **2**, **14**, **15**+**15***, **16** and **17**), with minor, earlier eluting, structural isomers of each ring-containing lipid also observed ($GDGT_{min}$ = **14'**, **15'**+**15'***, **16'** and **17'**). It is likely that ring-incorporation into GDGT lipids is used to control membrane fluidity at the elevated temperatures under which *I. aggregans* grows, as noted for several other thermophilic archaea (De Rosa *et al*., 1980b; Uda *et al*., 2001; 2004; Shimada *et al*., 2008; Lai *et al*., 2008). Although the locations of the rings in lipids **15'** and **15'*** have been delimited to particular isoprenoid chains, full structural determination for these components was not obtained. Similarly, elucidation of the structures of **14'**, **16'** and **17'** will require further work. GDD **1** was also identified as a minor core lipid in the organism, present in approximately 0.4% abundance relative to GDGT **2** (assuming the ionisation efficiencies of each lipid are equivalent). GTGT **3**, on the other hand, could not be identified in *I. aggregans*, although this does not preclude the possibility that the organism may synthesise this lipid when grown under different culture conditions.

In addition to GDD and GDGT lipids, *I. aggregans* also produces GMGT lipids containing 0-4 Cp rings (**27** and **28**-**31**). Structural isomerism of the GMGT lipids, as was observed for the GDGT lipids, was not evident in any of the chromatograms. This is partly due to limitations in the information that LC-MS/MS can provide on these structures, namely the lack of specific information regarding the isoprenoid chain compositions. In particular, it seems likely that GMGT **29** would be accompanied by an isomer in which the two rings within the structure are located in the same isoprenoid chain (i.e. a zero-two ring chain combination), mirroring the isomerism observed in GDGTs **15** and **15***. Further structural profiling is required in order to confirm this and, consequently, **29** is reported as a single isomer. The structures shown for GMGT lipid cores in Fig. 1.5 are tentative assignments based on the structures of **27** (Morii *et al*., 1998; Schouten *et al*., 2008b) and **28**-**31** (Schouten *et al*., 2008b) reported previously. In these structures, the positions of the Cp rings are assumed to match those found in GDGT lipids containing the same number of rings. The covalent bond joining the isoprenoid chains is ascribed to a

link between a C-15' methyl on one chain and a C-14 methylene on the other chain. It is important to note that these previous assignments are founded on the best structural fit to the spectroscopic and spectrometric data obtained for **27** (Morii *et al*., 1998) but are not conclusive. As discussed briefly in Section 3.2.3.3, a series of C_{80} tetraacids containing 4-8 Cp rings, potential transformation products of GMGT lipids, have been identified in naphthenate deposits in crude oil pipelines (Lutnaes *et al*., 2006) and have been fully characterised by 2D NMR analysis (Lutnaes *et al*., 2007). In particular, the tetraacid containing four rings has been shown to possess the rings in positions along the isoprenoid chains that match those observed in fully characterised GDGT **17** (De Rosa and Gambacorta, 1988a; Lutnaes *et al*., 2007). Furthermore, the covalent link between the isoprenoid chains in the tetraacid was found to be between a C-15' methyl and a C-15 methyl group (Fig. 4.5a). Given that the C-15' methyl group in one GDD lipid couples intermolecularly to the C-15 methyl in a different GDD lipid during formation of GDGT lipids (Nemoto *et al*., 2003; Koga and Morii, 2007) and C-15' and C-15 methyl groups may couple intramolecularly during formation of GMD lipids, it is plausible that C-15' and C-15 methyl groups in a GDGT lipid may also couple intramolecularly to give a GMGT lipid. As such, GMGT **31** may have a structure resembling the four-ringed tetraacid as opposed to the structure shown in Fig. 1.5 and Fig. 4.5b. By association, a C-15' methyl to C-15 methyl bond may be the chain-spanning covalent link in GMGTs **27** and **28**-**30**. In either case, more rigorous structural analysis of the GMGT lipids is neccesary.

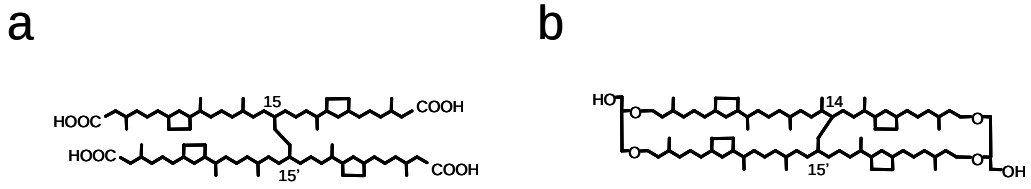


Fig. 4.5. a) C_{80} tetraacid containing four Cp rings (Lutnaes *et al*., 2007); b) tentative structure for GMGT **31** proposed by Schouten *et al.* (2008b).

To date, GMGT lipids have only been identified in cultured members of the *Euryarchaeota* (Chapter 3 and Morii *et al*., 1998; Sugai *et al*., 2004; Schouten *et al*., 2008b) or in marine and lacustrine sediments (Schouten *et al*., 2008a), prompting a suggestion that these lipids may act as biomarkers for *Euryarchaeota* in the environment (Schouten *et al*., 2008a). The identification of lipids **27** and **28**-**31** in *I. aggregans* indicates that this assumption must be re-assessed. Members of both the *Euryarchaeota* and the *Crenarchaeota* are able to incorporate isoprenoid cross-linked tetraether lipids into their cell membranes and, consequently, **27** and **28**-**31** can not be considered as kingdom-specific biomarkers in the interpretation of environmental lipid profiles.

GMGT lipids contribute nearly 39% to the total core lipid of *I. aggregans*, as calculated from the measured abundances of known lipids **1**, **2**, **14-17** and their minor isomers and **27-31**. This appears to represent one of the largest expressions of GMGT lipids in cultivated archaea reported to date, with only *Thermococcus waiotapuensis*, grown at 85^{o}C, reported to produce a higher proportion within its core lipids (48.9%; Sugai *et al*., 2004). It has been suggested that isoprenoid cross-linking in archaeal tetraether lipids may confer more rigidity to the cell membrane, preventing lysis even at very high temperatures (Schouten *et al*., 2008b). As such, the high levels of GMGT lipids expressed may indicate an adaptation of lipid biosynthetic pathways in this organism in order to allow survival under the extreme thermal stresses present in the hot springs from which it originates. It is unlikely, however, that this is the sole reason for the elevated GMGT levels observed in this archaeon. Different members of the order *Thermococcales* show variable proportions of expressed GDGT and GMGT lipids despite being grown at the same temperature (85^{o}C; Sugai *et al*., 2004). Furthermore, the percentage composition of GMGT cores observed in each of the eight (hyper)thermophilic organisms in which conclusive identifications have been presented does not appear to correlate with either the viable temperature range or maximum viable growth temperature of the organisms (Table 4.3). In addition, GMGT **27** has also been found in a number of sediments from low temperature marine and lacustrine environments (<35^{o}C; Schouten *et al*., 2008a). These observations suggest that taxon-specific differences among archaea may also place constraints on the extent of isoprenoid cross-linking in their membrane lipids.

Table 4.3. Organisms in which GMGT lipids have been identified to date, including the proportion of GMGT lipid as a percentage of the total core lipid found in each case.

Organism	Kingdom [a]	Growth temp. (°C)	Viable growth Range (°C)	GMGTs	% GMGT	References
Methanothermus fervidus	*Eury*	84	65-97	**27**	31.0	Stetter *et al.*, 1981; Koga *et al.*, 1993a; Morii *et al.*, 1998
Pyrococcus horikoshii	*Eury*	85	80-102	**27**	34.0	Sugai *et al.*, 2004
Thermococcus celer	*Eury*	85	75-97	**27**	15.5	Sugai *et al.*, 2004
Thermococcus guaymasensis	*Eury*	85	56-90	**27**	12.0	Sugai *et al.*, 2004
Thermococcus waiotapuensis	*Eury*	85	60-90	**27**	48.9	Sugai *et al.*, 2004
Aciduliprofundum boonei	*Eury*	70	55-75	**27-31**	n.r.	Reysenbach *et al.*, 2008; Schouten *et al.*, 2008b
Methanothermobacter thermautotrophicus	*Eury*	70	40-75	**27**	0.4[b]	Chapter 3; Morii *et al.*, 1998; Zeikus, 1972;
Ignisphaera aggregans	*Cren*	95	85-98	**27-31**	38.8	Current study; Niederberger *et al.*, 2006
Aquatic sediments	*Eury* *Cren?*	<35	n/a	**27**	0-6.5[b]	Schouten *et al.*, 2008a

n.r. = not reported

[a] *Eury = Euryarchaeote*; *Cren = Crenarchaeote*

[b] Diether lipids were not quantified during the lipid profiling undertaken in these studies. As such, the values quoted in each case are the percentages of the total tetraether lipid only.

The distribution of major GDGT lipids in *I. aggregans* grown at 95°C indicates that GDGT structures containing higher numbers of cyclopentyl rings are synthesised in preference to less heavily cyclised structures (Fig. 4.6a). The one exception is GDGT **2**, which is found in greater abundance in *I. aggregans* than GDGT **14** despite containing fewer (i.e. zero) Cp rings. Notably, however, **2** is suspected to be an intermediate in the formation of ring-containing lipids **14**-**17** (Weijers *et al.*, 2006a). As such, this discrepancy may represent the incomplete conversion of **2** to **14**-**17** during lipid biosynthesis. A similar trend is observed for $GDGT_{min}$ isomers **14'-17'**, which are also found in increasing abundance as ring number increases (Fig. 4.6a). Interestingly, however, the proportion of each minor GDGT to its major structural

isomer is not constant, increasing with ring number up to structures with three rings, followed by a decrease in the ratio for the structures containing four rings (Fig. 4.6b). The non-homogeneity of these ratios indicates that the organism may actively control the proportions of the different $GDGT_{min}$ isomers expressed, although it should be noted that the differences in the ratios, where statistically significant, are small.

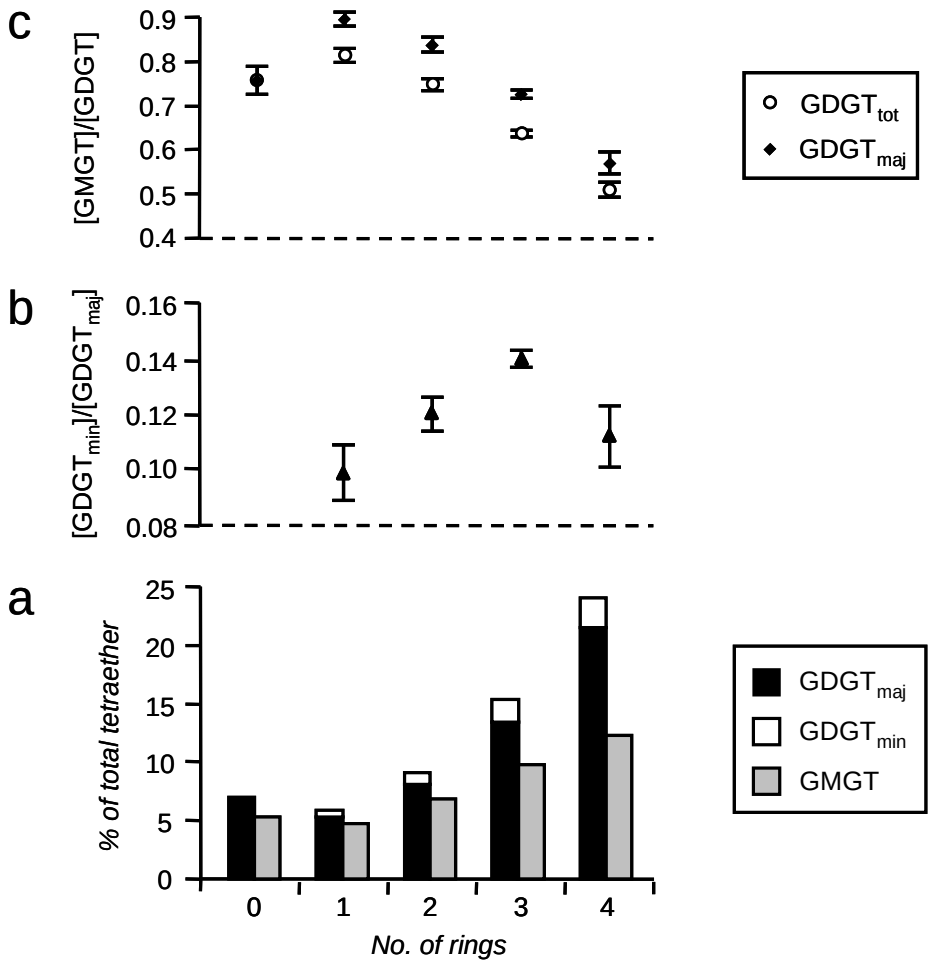


Fig. 4.6. Relative abundances of major GDGT isomers ($GDGT_{maj}$), minor GDGT isomers ($GDGT_{min}$ = GDGT') and GMGTs containing the same number of cyclopentyl rings in *Ignisphaera aggregans* AQ1.S1^T: a) proportion of each component as a percentage of the total tetraether lipid; b) ratios of $GDGT_{min}$ to $GDGT_{maj}$; c) ratio of GMGT to $GDGT_{maj}$ or total GDGT lipid ($GDGT_{tot}$ = $GDGT_{maj}$ + $GDGT_{min}$). Error bars show ±1σ from the mean of analytical triplicates (HCT).

The distribution of ring-containing GMGT lipids in the *I. aggregans* extracts generally mirrors the trend observed for the GDGT lipids, with the more heavily cyclised structures produced in greater relative abundance (Fig. 4.6a). Again, the only exception is for the lipid containing zero rings, **27**, which is found in greater relative abundance than **28**, which contains one ring. The close analogy between the abundances of GDGT and GMGT lipids suggests that their biosynthesis may be intrinsically linked. Similarly, a lipid profile of *A. boonei* reported previously showed that GMGT:GDGT ratios for lipids containing the same number of rings were approximately constant regardless of the ring number (Schouten *et al.*, 2008b). This led the authors to suggest that GDGT lipids may be the biosynthetic precursors to the GMGT lipids (Schouten *et al.*, 2008b). The absence of macrocyclic diether **4** in both *I. aggregans* and *M. thermautotrophicus* (Comita *et al.*, 1983; 1984), organisms which both express lipids **1**, **2** and **27**, implies that **27** is almost certainly formed directly from **2** (or an unsaturated precursor thereof) and not *via* cross-linking of a molecule of **4** with **1**. If the isoprenoid chains in ring-containing GDGT lipids **14**-**17** were ligated *via* a similar biosynthetic pathway, GMGT lipids **28**-**31** would be formed. If, however, this were the case then minor structural isomers of GMGTs **28**-**31**, formed from **14'**-**17'**, would also be expected in *I. aggregans* and such structures were not observed. It remains possible that the isomers are present but obscured by co-elution with the major GMGT lipids reported above.

The relationship between the GMGT and GDGT lipids has the potential to reveal biosynthetic relationships between the two core lipid families. The ratios of GMGT to $GDGT_{maj}$ and to total GDGT lipid ($GDGT_{tot} = GDGT_{maj} + GDGT_{min}$) for each ring number in *I. aggregans* (Fig. 4.6c) cover both possibilities: either that minor GMGT isomers are absent or that they are masked by co-elution. With the exception of **27**:**2,** the magnitude of GMGT:GDGT for both measures decreases as ring number increases. Thus, if GMGT lipids do indeed originate from GDGT lipids, the number of rings in the precursor GDGT must be a limiting factor on the degree of its conversion to its GMGT counterpart. Alternatively, it is conceivable that GMGT lipids **28**-**31** are formed by sequential incorporation of Cp rings into **27,** or an unsaturated precursor of this lipid, akin to the route proposed for the formation of GDGTs **14**-**17** from **2** (Weijers *et al.*, 2006a). The lipid distribution observed in *I.*

aggregans would also be consistent with this second scheme provided that the enzyme or enzyme suite involved in Cp ring-incorporation showed lower activity on **27** than on **2**. It is also possible that both proposed biosynthetic pathways or other routes may operate simultaneously to produce Cp ring-containing GMGTs. Irrespective of the route taken, the GMGT lipid profile of *I. aggregans* appears to be inconsistent with that of *A. boonei*, where no reduction in GMGT proportions for higher ring numbers was noted (Schouten *et al.*, 2008b). On the basis of the limited evidence obtained from these two organisms, it remains difficult to assign the production of GMGTs **28**-**31** to a particular biosynthetic route at the present time.

4.2.3. *Ignisphaera* sp. Tok37.S1, *Pyrobaculum* sp. AQ1.S2 and *Desulfurococcus* sp. Ket10.S1

4.2.3.1. Ether lipid profiles

The base peak chromatograms from LC-MS/MS analysis of lipid extracts from *Ignisphaera* sp. Tok37.S1, *Pyrobaculum* sp. AQ1.S2 and *Desulfurococcus* sp. Ket10.S1 strains are dominated by peaks relating to tetraether lipid cores, as illustrated in ion chromatograms for the mass spectral range *m/z* 1280-1320 (Fig. 4.7). GDGT lipids **2** and **14**-**16** (including **15***) were identified in the base peak chromatogram of each strain, with the structure of each lipid being confirmed by MS/MS. GDGT **17** was also identified in the base peak chromatograms for the *Ignisphaera* and *Pyrobaculum* species and could be identified in *Desulfurococcus* from a mass chromatogram for *m/z* 1293.5-1294.5. Similarly, GDGT lipids **14'**-**16'** were identified in mass chromatograms for each species, although no reliable MS/MS spectra were generated for any of these components. GDGT **17'** was identified in the extracts from *Ignisphaera* and *Pyrobaculum* species, again without support from MS/MS spectral data, but was not identified in the extract from the *Desulfurococcus* species. GTGT **3** was observed in ion chromatograms for *Pyrobaculum* sp. AQ1.S2 only, with the structure confirmed by MS/MS.

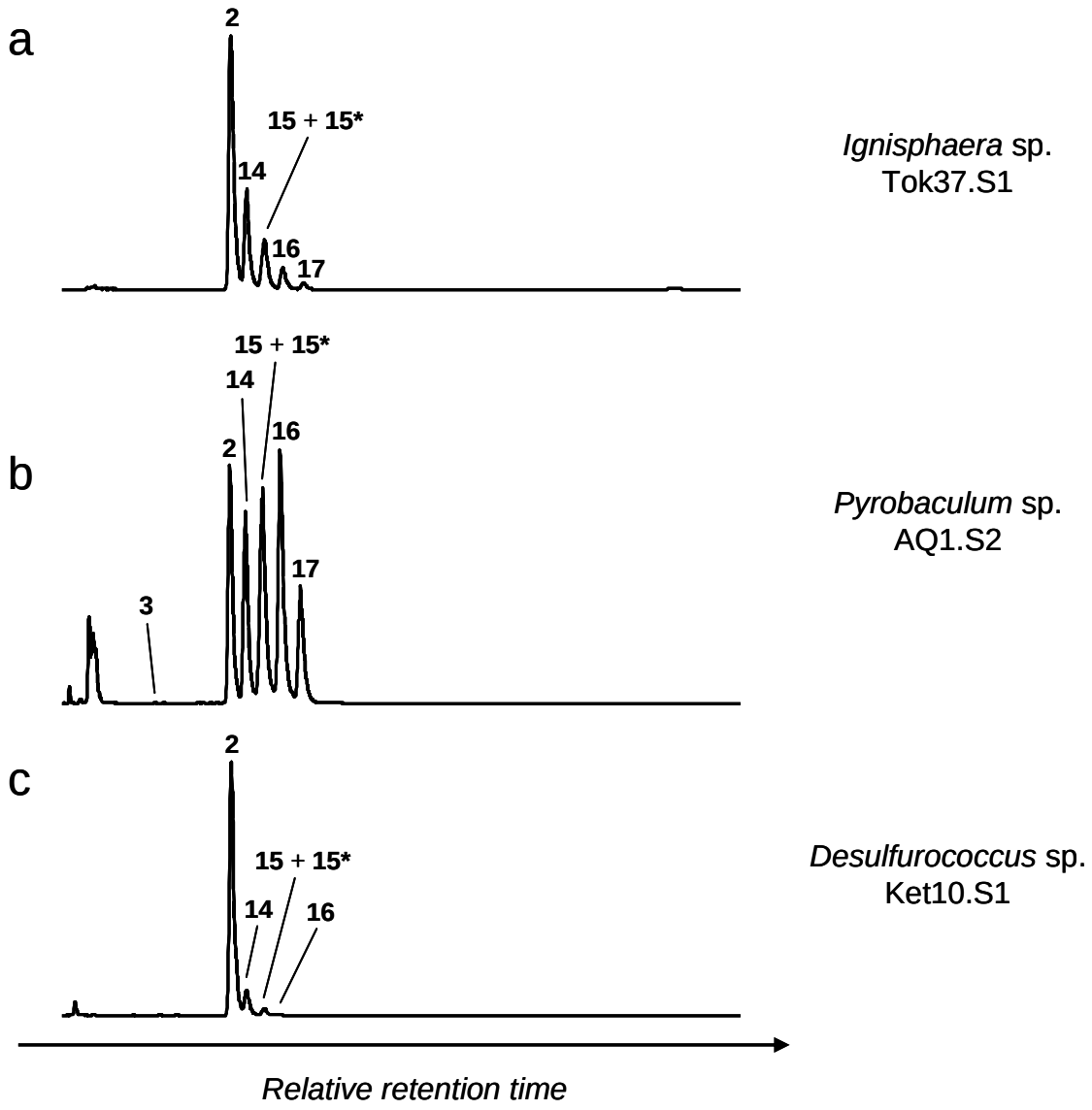


Fig. 4.7. HCT LC-MS ion chromatograms (*m/z* 1280-1320) for New Zealand hot spring isolates: a) *Ignisphaera* sp. Tok37.S1; b) *Pyrobaculum* sp. AQ1.S2; c) *Desulfurococcus* sp. Ket10.S1.

Ion chromatograms for the mass spectral ranges *m/z* 1307.5-1316.5 (Fig. 4.8b) and *m/z* 1321.5-1330.5 (Fig. 4.8c) extracted from the LC-MS/MS data for the *Pyrobaculum* species indicate a number of components which co-elute with GDGT lipids containing 1-4 Cp rings (shown in Fig. 4.8a) but which exhibit $[M+H]^+$ at +14 or +28 *m/z* units greater. This suggests that these components represent $GDGT_M$ lipids which contain the same number of cyclopentyl rings as the GDGT species

with which they co-elute (**14**-**17** and **14'**-**17'**), but which bear one (**82**-**83**, **83'**, **70**-**71** and **70'**-**71'**) or two (**84**-**85** and **85'**) additional carbon atoms.

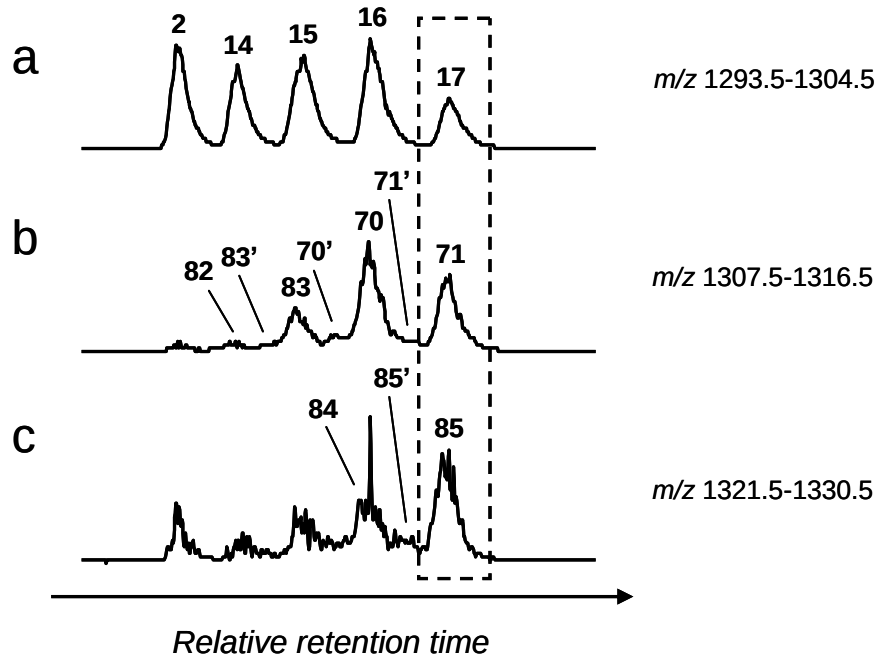


Fig. 4.8. HCT LC-MS ion chromatograms for the lipid extract from *Pyrobaculum* sp. AQ1.S2, highlighting: a) GDGT lipids; b) $GDGT_M$ lipids containing one additional carbon atom; c) $GDGT_M$ lipids potentially containing two additional carbon atoms. GDGT* and $GDGT_M$* lipids which co-elute with some of the components identified, are not labelled.

In order to verify the structures of the $GDGT_M$ lipids, the *Pyrobaculum* extract was first re-analysed using the scan range *m/z* 1307-1320, which facilitated selection of the structures containing one additional carbon atom for CID. Subsequently, the extract was analysed using the scan range *m/z* 1320-1330 to attempt to obtain CID of the structures containing two additional carbon atoms. Unfortunately, only four $GDGT_M$ lipids, **82**, **83**, **70** and **71**, and co-eluting $GDGT_M$* isomers **82***, **83*** and **70***, generated MS/MS spectra with sufficient signal strength for reliable structural interpretations to be made. The MS/MS spectrum of each $GDGT_M$ lipid contained product ions formed from loss of 0-3 molecules of water and/or C_3H_4O, indicative of two terminal, unmodified glycerol moieties within each structure. The MS/MS spectrum of **82**/**82*** (Fig. 4.9a) also contained product ions at *m/z* 755, 743, 741 and

739 formed *via* losses of, in turn, 558 Da ($[40_0]^0$), 570 Da ($[41_1]^0$), 572 Da ($[41_0]^0$) and 574 Da ($[40_1]^1$). These product ions, along with those in which additional functionality has been lost from the precursor, confirm a mixture of GDGT[40_1,41_0] (designated **82**) and GDGT[40_0,41_1] (designated **82***) isomers. The MS/MS spectrum of **83**/**83*** (Fig. 4.9b), on the other hand, shows product ions at *m/z* 755, 753, 743 and 741, corresponding to losses of 556 Da ($[40_1]^0$), 558 Da ($[40_0]^0$), 568 Da ($[41_2]^0$) and 570 Da ($[41_1]^0$) and suggestive of a mixture of GDGT[40_1,41_1] (designated **83**) and GDGT[40_0,41_2] (designated **83***) isomers. The MS/MS spectrum of **70**/**70*** (Fig. 4.9c) shows product ions at *m/z* 755, 753, 741 and 739, corresponsing to losses of 554 Da ($[40_2]^0$), 556 Da ($[40_1]^0$), 568 Da ($[41_2]^0$) and 570 Da ($[41_1]^0$), suggesting both GDGT[40_1,41_2] (designated **70**) and GDGT[40_2,41_1] (designated **70***) isomers are present in the extract. The presence of lipid **71** in the extract was confirmed following CID: the MS/MS spectrum obtained during LC-MS/MS analysis (Fig. 4.9d) contained product ions identical to those observed during CID of the lipid found in *S. acidocaldarius* (Section 3.2.4.5). Product ions observed in the spectrum of **71** include ions at *m/z* 753 and 739, formed *via* losses of 554 Da ($[40_2]^0$) and 568 Da ($[41_2]^0$), respectively, indicative of a GDGT[40_2,41_2] structure. Although the lipid structures were not confirmed by MS/MS analysis, $GDGT_M$ lipids **83'** and **70'-71'** are most likely chain methylated variants of GDGTs **15'-17'**, respectively. Similarly, $GDGT_M$ lipids **84**, **85** and **85'** are most likely to be higher homologues of **16**, **17** and **17'** in which either one chain has two additional methyl groups or both chains each have one additional methyl group. On the basis of the product ions of GDGTs **16** and **17** observed during LC-MS/MS analysis of the *Ignisphaera* sp. Tok37.S1 extract, this organism may synthesise trace levels of $GDGT_M$ lipids **70** and **71**, although further work is required to establish the presence of the components unequivocally. Similarly, the identification of trace levels of $GDGT_M$ lipids potentially in the *I. aggregans* AQ1.S1^T extract needs further verification.

LC-MS/MS analysis of each extract using an MS scan range of *m/z* 50-2000 provided no evidence for the expression of GDD **1** by any of the three isolates when grown at 90^{o}C.

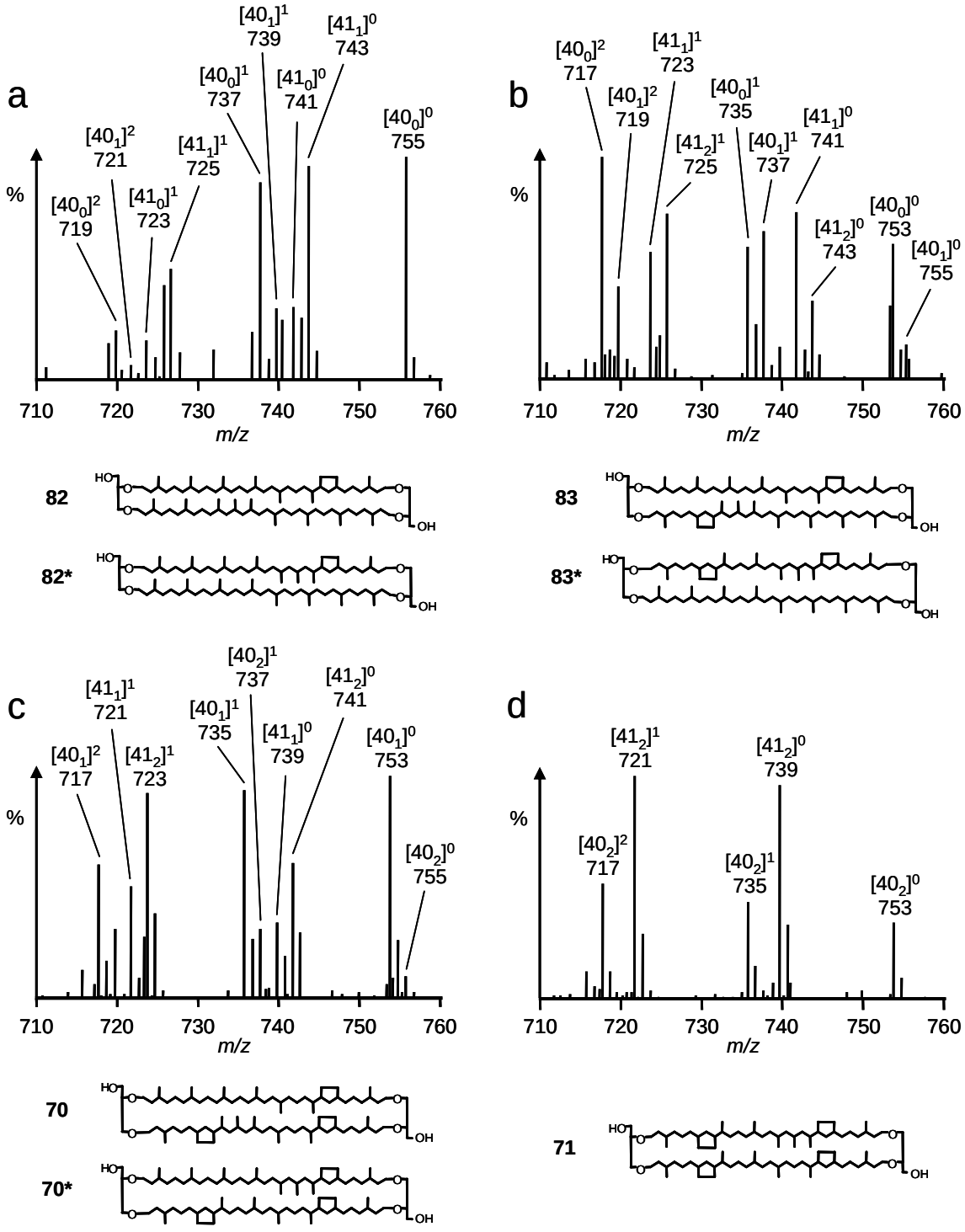


Fig. 4.9. Partial MS/MS spectra (HCT) of: a) $GDGT_M$ **82**/**82*** ($[M+H]^+$ *m/z* = 1314.4); b) $GDGT_M$ **83**/**83*** ($[M+H]^+$ *m/z* = 1312.4); c) $GDGT_M$ **70**/**70*** ($[M+H]^+$ *m/z* = 1310.4) and d) $GDGT_M$ **71** ($[M+H]^+$ *m/z* = 1308.4) from *Pyrobaculum* sp. AQ1.S2. $GDGT_M$ structures shown are representative structures; the additional methyl group in each is tentatively placed at the C-13 chain position in accordance with the known structure of $GDGT_M$ **26** (Galliker *et al.*, 1998).

4.2.3.2. Distributions and significance

As noted for *I. aggregans*, all three isolates exhibit both major and minor isomeric forms of GDGT lipids containing 0-4 Cp rings. The distribution of the major core lipids in *Pyrobaculum* sp. AQ1.S2 compares favourably with those of the three other species of the *Pyrobaculum* genus to have been profiled to date (Trincone *et al.*, 1992; Volkl *et al.*, 1993). By contrast, the ring-containing GDGT lipids identified in *Desulfurococcus* sp. Ket10.S1 were not observed in the only other member of the genus to have been profiled (*Desulfurococcus mobilis*, Lanzotti *et al.*, 1987), with **2** being the sole GDGT lipid identified. Provided that the taxonomic assignment of sp. Ket10.S1 is correct, its lipid profile indicates that members of the *Desulfurococcus* can express ring-containing tetraether lipids. Interestingly, lipids containing 5-8 Cp rings, as identified in species of *Sulfolobus* (Section 3.2.4), were not observed in any of the organisms nor in *I. aggregans* AQ1.S1^T, despite growth of each organism at temperatures which exceed the maximum viable growth temperature of the *Sulfolobus* species described. To date, expression of GDGT lipids containing more than four Cp rings in archaea has appeared to be restricted largely to the orders *Sulfolobales* and *Thermoplasmatales* (Section 1.4.1 and Schouten *et al.*, 2007c, Trincone *et al.*, 1989), although unique members of the *Thermoproteales* (*T. tenax*; Thurl and Schäfer, 1988) and of the Group I *Crenarchaeota* (*N. gargensis*; Pitcher *et al.*, 2010) appear to express lipids containing five Cp rings in addition to those with 0-4 Cp rings. The apparent specificity in the expression of GDGT core lipids containing more than four Cp rings to a restricted set of genera indicates that production of the more heavily cyclised core lipids may be a genetic, as opposed to adaptive, trait. As such, the absence of lipids containing five or more rings in the NZ hot spring isolates may reflect an inability to produce such structures, irrespective of the physiological conditions present during growth.

No GMGT lipids appear to be produced by *Pyrobaculum*, *Desulfurococcus* or by *Ignisphaera* Tok37.S1 when grown at 90^{o}C. This is particularly surprising in the latter case given the close phylogenetic relationship between Tok37.S1 and *I. aggregans* AQ1.S1^T, which does express a sizeable proportion of GMGT lipids. Given that 16S rRNA profiling of Tok37.S1 led to amplification of two distinct

RNA sequences (Niederberger *et al.*, 2006), there is a possibility that the culture is not axenic. If a contaminating organism is present and had out-competed Tok37.S1 during growth, the lipid profile taken from the cultured cells may be a distortion of that which would be obtained from a pure culture of Tok37.S1. Nevertheless, the fact that GMGT lipids appear to be expressed by *I. aggregans* AQ1.S1^{T} but not by other isolates from the hot springs may allow these lipids to be used as biomarkers for this species during exploratory profiling of New Zealand hot springs, complementing metagenomic typing made in the same environments. Further studies into the taxonomic distribution of GMGT lipids and into the physiological controls that regulate their expression are needed before this application will be fully practical.

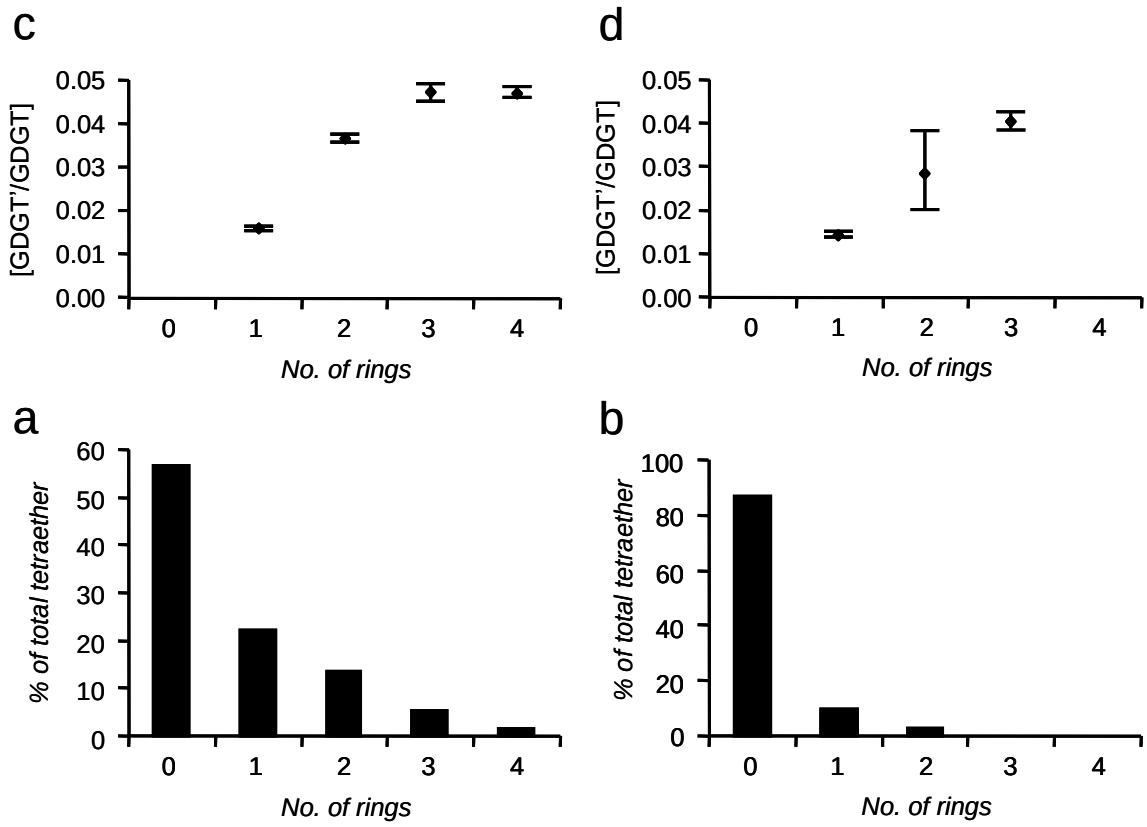


Fig. 4.10. Relative abundances of major GDGT isomers (GDGT) and minor GDGT isomers (GDGT') containing the same number of Cp rings, displayed as the proportion of each GDGT component as a percentage of the total tetraether lipid or as the ratio of GDGT' to GDGT in *Ignisphaera* sp. Tok37.S1 (a and c) and *Desulfurococcus* sp. Ket10.S1 (b and d). Error bars show ±1σ from the mean of analytical duplicates (HCT).

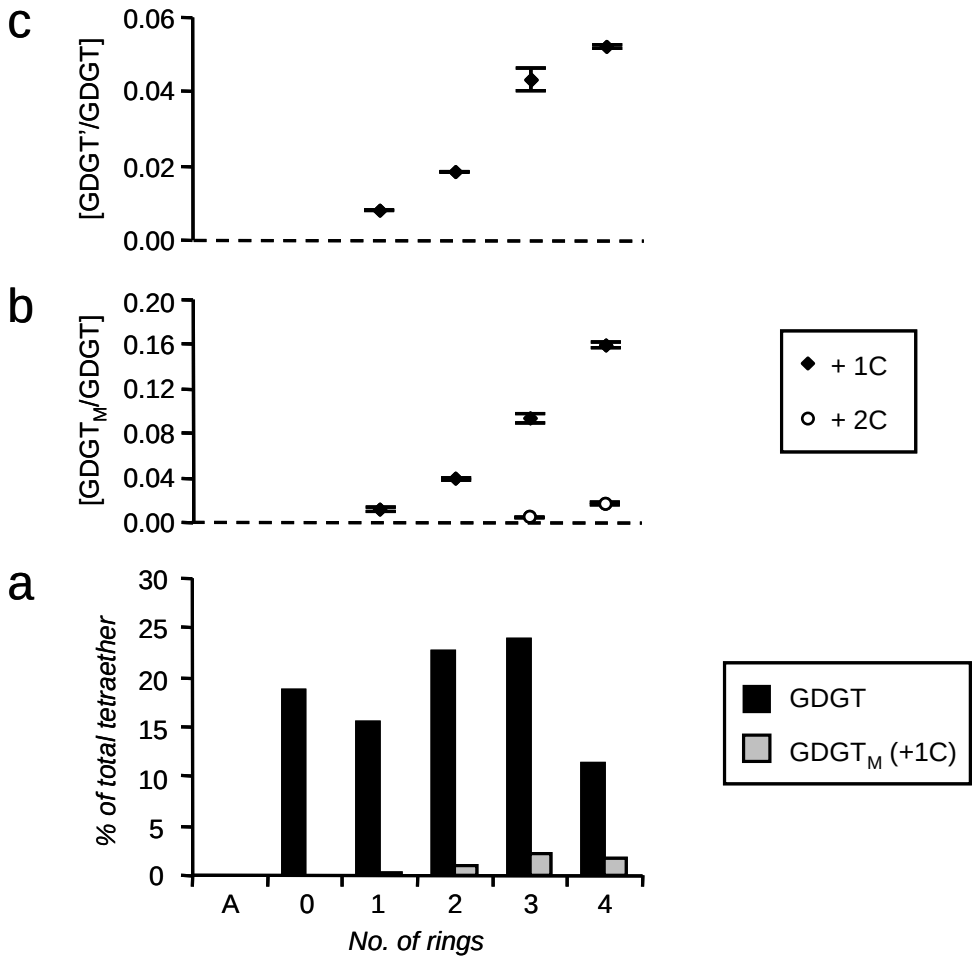


Fig. 4.11. Relative abundances of GTGT **3** (labelled A), major GDGT isomers (GDGT), minor GDGT isomers (GDGT') and methylated GDGTs ($GDGT_M$) containing the same number of Cp rings in *Pyrobaculum* sp. AQ1.S2: a) proportion of each GDGT or $GDGT_M$ (with one additional carbon atom) as a percentage of the total tetraether lipid; b) ratios of $GDGT_M$ containing one (+1C) or two (+2C) additional carbon atoms to GDGT; c) ratio of GDGT' to GDGT. Error bars show $\pm 1\sigma$ from the mean of analytical duplicates (HCT).

Ignisphaera sp. Tok37.S1 and *Desulfurococcus* sp. Ket10.S1 both express GDGT **2** as the major core lipid (Fig. 4.10) when grown at 90°C/pH 6, whereas the core lipid profile of *Pyrobaculum* sp. AQ1.S2 shows greater relative levels of expression of ring-containing GDGT lipids (Fig. 4.11). This is reflected in the degree of lipid cyclisation (calculated as the weighted average of lipids **2** and **14**-**17** only) in each

case, with the values for Tok37.S1 (0.72) and Ket10.S1 (0.17) both lower than that of AQ1.S2 (1.93). Calculations of the degree of lipid cyclisation that included minor GDGT lipids **14'-17'** gave similar, although marginally greater, values in each case.

The *Pyrobaculum* species was also observed to express $GDGT_M$ lipids containing up to four Cp rings, with the MS/MS spectra of some of the components confirming that the extra carbon in these structures was located on one or other isoprenoid chain. Components suggestive of $GDGT_M$ lipid structures containing three or four Cp rings and two additional carbon atoms were also observed, although the location of the extra carbon atoms was not confirmed. Each $GDGT_M$ component eluted marginally earlier than the homologous GDGT lipid during LC-MS/MS analysis, consistent with tetraethers methylated at the C-13 position of an isoprenoid chain which were identified during similar analysis of extracts from MTH(ΔH) and from species of *Sulfolobus,* described in Chapter 3. As the full structures of the lipids can not be confirmed by MS/MS analysis alone the assignment of the $GDGT_M$ lipids as C-13 methylated variants of GDGT lipids **14-17** and **14'-17'** is tentative. Irrespective of their precise structures, the identification of $GDGT_M$ lipids in a *Pyrobaculum* species extends the range of archaeal species which can utitilise tetraether lipid homologation during biosynthesis to the order *Thermoproteales*.

The most abundant $GDGT_M$ lipid in the *Pyrobaculum* extract, grown at 90^oC/pH 6 contains three Cp rings, the same number as in the most abundant GDGT lipid (Fig. 4.11a). By contrast, the most abundant of the $GDGT_M$ lipids containing two additional carbon atoms exhibited four Cp rings. The ratio of $GDGT_M$ to GDGT lipid for structures containing the same number of rings appears to increase with increasing ring number (Fig. 4.11b), with the same positive correlation to ring number observed for $GDGT_M$ lipids containing either one or two additional carbon atoms. The trends observed are, for the most part, consistent with those observed between GDGT and $GDGT_M$ lipids in *S. acidocaldarius* (Section 3.2.4.7), although *Pyrobaculu*m AQ1.S2 does not appear to express $GDGT_M$ **26**, which contains zero cyclopentyl rings and was found in elevated levels in *S. acidocaldarius*.

The ratio of GDGT' to major GDGT isomers containing the same number of Cp rings appears to show a general increase with increasing number of rings in each of the three isolates (Fig. 4.10c, 4.10d and 4.11c), although in *Ignisphaera* sp. Tok37.S1 the ratio is effectively the same for structures containing four rings and three rings. The lack of identification of a minor isomer of GDGT **17** is almost certainly due to this component being present in the extract at a concentration below the limit of the detection of the LC-MS/MS method used, especially considering the very low levels of **17** detected during the analyses. The trends observed are different to those noted for species of *Sulfolobus* (Section 3.2.4.7) and for *I. aggregans* AQ1.S1^{T} (Section 4.2.2.3), which show maximum GDGT':$GDGT_{maj}$ ratios for structures containing three Cp rings.

4.2.4. Other ether lipid components

In addition to the lipids described in Sections 4.2.1 and 4.2.2, a number of other components were identified in the lipid extracts from each organism. Although not all of the components could be characterised in each case, some of the components for which MS/MS spectra were obtained are discussed in further detail.

4.2.4.1. *Ignisphaera aggregans* AQ1.S1^{T}

In addition to GDGT and GMGT lipids, three series of minor components are also observed in the base peak chromatogram for *I. aggregans* (Fig. 4.12), eluting only after the IPA composition of the mobile phase had exceeded 2%. The late elution of all three families suggests that they represent structures containing functionality that makes them significantly more polar than both GDGT and GMGT lipids. The *m/z* of the protonated molecules of the lipids in the first series, **87**-**89**, are consistent with structures that are 48 Da greater than the corresponding lipids **15**-**17** (Table 4.4). Furthermore, the MS/MS spectra of **87**-**89** each show product ions formed from loss of the same biphytanyl chains as are found in their proposed GDGT counterpart, supporting structural similarities between the two sets of components. For example, the MS/MS spectrum of lipid **89** (Table 4.4) contains a product ion at *m/z* 787 formed from loss of 554 Da, indicating that it contains at least one biphytanyl chain

in which there are two incorporated Cp rings, a moiety also found in GDGT **17**. In addition, many of the product ions in the MS/MS spectrum of **89** are also shifted +48 *m/z* units relative to those of **17**. In spite of the mass spectral data presented, the nature of the structural difference between **87**-**89** and GDGTs **15**-**17** currently remains unclear, although it does appear that similar structural relatives of GDGTs **2** and **14** and of GMGTs **30**-**31** may also be synthesised by *I. aggregans* on account of the presence of other $[M+H]^+$ detected in the base peak chromatograms during LC-MS/MS. Unfortunately, these components were of insufficient signal strength to be selected for CID and, as such, can only be categorised tentatively. Components in the second and third series include lipid **58** and several ring-containing structural variants, lipids **75**-**77**, **86** and **90**-**94** (Table 4.4). Discussion of the nature of these compounds is provided in Section 5.2.3.2.

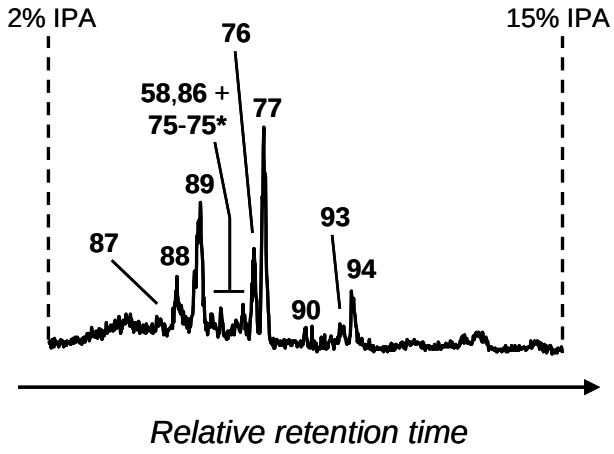


Fig. 4.12. Magnification of a region of the HCT LC-MS base peak chromatogram for a lipid extract from *I. aggregans* AQ1.S1^T relating to an eluent IPA proportion > 2%.

Other early eluting, minor components were observed in LC-MS/MS analyses using both conventional (*m/z* 900-1500) and extended (*m/z* 50-2000) scan ranges although, in the interests of brevity, these components are not discussed further.

4.2.4.2. *Ignisphaera* sp. Tok37.S1 and *Desulfurococcus* sp. Ket10.S1

Several late eluting components were also observed in the base peak chromatogram for the Tok37.S1 strain (Fig. 4.13), including lipid **58** and ring-containing structural variants **86** and **75**/**75***. An additional lipid, **95**, was also observed and appears, on the basis of the MS/MS spectrum (Table 4.4), to be a structural relative of lipids **87**-**89** containing no cyclopentyl rings. As such, this component probably represents a structurally modified variant of GDGT **2** in which a 48 Da structural appendage has been added. The base peak chromatogram of Ket10.S1 contained only one component which eluted after the eluent had reached a 2% IPA composition, with **95** determined as the only late-eluting lipid in the organism. In both organisms, very early eluting components were also observed in the base peak chromatograms generated from lipid extracts. Lipid **59** was identified in the base peak chromatogram for Tok37.S1 on the basis of the retention time, the *m/z* value of the $[M+H]^+$ and the MS/MS spectrum generated for the component during CID. Similarly, methyl etherified lipid **51**, probably an artifact generated during the extraction process, was also identified in Ket10.S1 on the basis of retention time, MS and MS/MS data. Other early eluting components proved more difficult to assign and are not discussed.

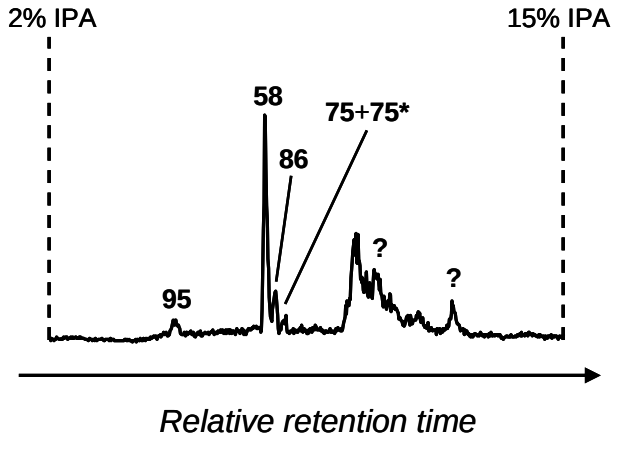


Fig. 4.13. Magnification of a region of the HCT LC-MS base chromatogram for a lipid extract from *Ignisphaera* sp. Tok37.S1 relating to an eluent IPA proportion > 2%.

Table 4.4. HCT MS (measured $[M+H]^+$ *m/z* values) and MS/MS (nominal *m/z* values of product ions) spectral data and assignment of minor, late eluting lipid components in species of *Ignisphaera* and *Desulfurococcus*, isolated from New Zealand hot springs.

Lipid	$[M+H]^+$	Product ions observed in MS/MS [a]	Assigned as:
95	1350.5	1331,1313,1293,1275,1273,1257; 791,773,757,755,717,699,681,679,663,643,**625**,623,607	**2** + 48 Da
87	1345.7	1327,1309,1289,1271,1269,1253; 789,771,755,753,715,697,679,641,**623**,621,605	**15** + 48 Da
88	1343.7	1325,1307,1287,1269,1267,1251; 789,787,771,769,755,753,751,715,713,697,695,679,677,641,639,**623**,621,605,603	**16** + 48 Da
89	1341.7	1323,1305,1285,1267,1265,1249; 787,769,753,751,713,695,677,639,**621**,619,603	**17** + 48 Da
58	1245.8	1227,1209; **669**,651,633,615,597,575,557	**2** - 56 Da
86	1243.8	1225,1207; 669,**667**,651,649,631,615,613,595,593,557	**14** - 56 Da
75 + **75***	1241.8	1223,1205; 669,**667**,665,651,649,647,633,631,629,615,613,611,595,593,575,573,571,569,557,555,553	**15/15*** - 56 Da
76	1239.8	1221,1203; 667,**665**,649,647,631,629,613,611,595,593,591,573,571,569,557,555,553	**16** - 56 Da
77	1237.8	1219,1201; **665**,647,629,611,593,591,573,571,569,555,553	**17** - 56 Da
90	1243.8	**1225**,1207,1189,1171,1169,1151,1149,1147,1133,1131,1129,1115	**27** - 56 Da
91	1241.7	1223,**1205**,1187,1169,1167,1149,1147,1145,1131,1129,1127,1113	**28** - 56 Da
92	1239.7	1221,**1203**,1185,1167,1165,1163,1147,1143,1129,1127,1125,1111	**29** - 56 Da
93	1237.7	1219,**1201**,1183,1165,1163,1145,1143,1141,1127,1125,1109	**30** - 56 Da
94	1235.7	1217,**1199**,1181,1163,1161,1143,1141,1139,1125,1123,1121,1107	**31** - 56 Da

[a] Product ions listed in bold face represent the base peak in MS/MS.

4.2.4.3. *Pyrobaculum* sp. AQ1.S2

The base peak chromatogram of the lipid extract from AQ1.S2 was devoid of peaks having MS/MS spectra consistent with tetraether lipid structure beyond 61.25 min, equivalent to the retention time window in which the polarity of the eluent exceeded a 2% IPA composition,. A number of very early eluting components (i.e. eluting significantly before GTGT **3**) were identified, including a series of components, **96-100**, which exhibit $[M+H]^+$ shifted -18 *m/z* units from those of GDGTs **2** and **14-17** (Fig. 4.14). The MS/MS spectrum of **96** ($[M+H]^+$ *m/z* = 1284.4) contains ions with

identical *m/z* to those observed in the MS^3 spectrum of **2**, obtained following selection of the precursor ion at nominal *m/z* 1283 during MS^n analyses of the lipids of MTH(ΔH) (Section 2.2.6). This indicates that **96** is most likely a dehydroxylated variant of GDGT **2** in which a hydroxyl group has been lost/eliminated. Lipids **97-100** most likely represent analogous dehydroxylated variants of ring-containing GDGTs **14-17**, with the observed MS/MS spectra of each component (not shown) being consistent with this assignment.

An additional series of early eluting components, **101**-**105**, with product ions exhibiting $[M+H]^+$ shifted +184 *m/z* units from those of GDGTs **2** and **14-17** were also identified in the base peak chromatogram of AQ1.S2 (Fig. 4.14). Each component shows product ions formed from neutral losses of 184 Da and 202 Da indicating that the lipids may be derivatives of **2** and **14-17** in which one of the terminal hydroxyls has been modified, with an alcoholic hydrogen being replaced by an apolar 185 Da group. Losses of 1-ene-biphytan-32-ol molecules containing 0-2 Cp rings from each structure during CID are also consistent with this assignment. Other early eluting components were also observed in the chromatogram for AQ1.S2 but proved difficult to assign on the basis of the MS/MS spectra obtained.

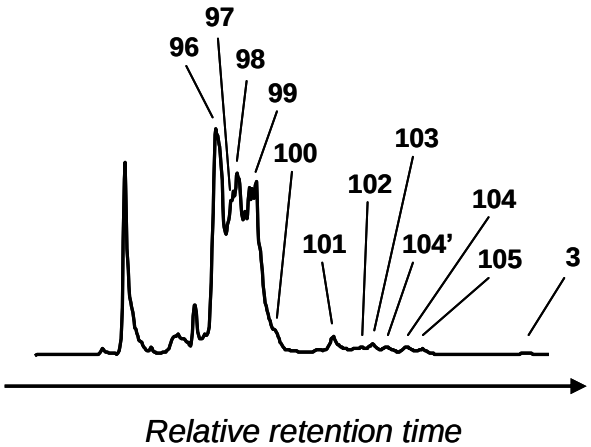


Fig. 4.14. Magnification of an early section of the HCT LC-MS base peak chromatogram for a lipid extract from *Pyrobaculum* sp. AQ1.S2.

4.2.4.4. Significance

Although the majority of the additional components described are found in levels significantly below those of the GDGT and GMGT lipids in the same extracts, some components are found to be significant contributors to the total lipid pool. Dehydroxylated lipids **96**-**100**, for example, are found in appreciable levels relative to those of GDGT lipids **2** and **14**-**17** in the extract from *Pyrobaculum* sp. AQ1.S2, suggesting that they may represent a substantial proportion of the membrane lipids of the organism. Unfortunately, the poor chromatographic resolution of **96**-**100** makes quantification of the individual components problematic. As such, conclusions as to whether **96**-**100** may be formed during biosynthesis by dehydroxylation of GDGT lipids containing the same number of Cp rings cannot currently be drawn.

Lipids **95** and **87**-**89** were identified in the extracts from the three members of the *Desulfurococcales* investigated but were not identified in the lipid extracts from cultures of *Pyrobaculum* sp. AQ1.S2 or in extracts from either MTH(ΔH) or three species of *Sulfolobus* (Chapter 3), each of which is categorised into a phylogenetic order distinct to the *Desulfurococcales*. Despite the fact that the levels of **95** and **87**-**89** expressed by the *Ignisphaera* and *Desulfurococcus* species were low, this highlights the potential of these components as unique marker compounds for the *Desulfurococcales*. Clearly, however, significant further screening of archaeal species lying both within and outside of the *Desulfurococcale*s is required before **95** and **87**-**89** can be used as order-specific biomarkers with any real confidence.

The MS spectra recorded during the column backflush procedure incorporated following LC-MS/MS of the extracts from each of the four organisms indicated that polar material with $[M+H]^+$ at $m/z > 1280$, possibly relating to tetraether lipid material, was retained on-column for the entirety of each analysis. This suggests that some highly polar components may also be synthesised by the hot spring isolates but have been missed on account of limitations in the LC-MS/MS employed during the analyses.

4.3. Conclusions

I. aggregans grown at 95^{o}C biosynthesises both diether and tetraether isoprenoid lipid cores for its cellular membrane, and contains one of the most complex distributions of the latter reported in cultured archaea to date. The source of such complexity arises from incorporation both of cyclopentane rings and isoprenoid chain cross-links in the tetraether lipid cores. It is presumed that both structural elements are used to regulate specific physicochemical properties of the cell membrane, particularly fluidity, at the high growth temperatures in which the organism thrives. No explanation is currently available to account for the inclusion of two mechanisms likely to influence membrane rigidity. The presence of GMGT lipids in *I. aggregans* is the first confirmation of this lipid core type in a *Crenarchaeote*, and indicates that these components can not be used as markers for *Euryarchaeota* in environmental samples. The monoalkyl lipids are almost certainly derived from caldarchaeol during biosynthesis, although the precise pathway to these compounds still needs to be established. In any case, *I. aggregans* represents an ideal organism for further study of isoprenoid lipid biosynthesis, particularly in terms of variations introduced into the membrane lipids in response to external stresses. This organism should also prove to be a key species in assessing the potential of GMGT lipids as biomarkers for thermophilic archaea in hot spring environments.

Three other isolates from New Zealand hot springs synthesise GDGT lipids but do not appear to express either GMGT or GDD lipids, at least not at the temperatures/pH values chosen for growth. The fact that each isolate synthesises the same ring-containing GDGT lipids highlights the lack of specificity among the distributions and the unsuitability of the compounds for use as biomarkers. Several other lipids that could potentially serve as more selective markers for particular orders, genera or species of archaea were identified in each extract. Further supporting studies are now required to investigate the scope of these compounds as biomarkers.

Methylated tetraether lipids were identified in a *Pyrobaculum* species, including components containing one or two extra carbon atoms by comparison with

previously reported ring-containing GDGT lipids. The identification of $GDGT_M$ lipids in this archaeon extends both the range of organisms in which these components have been identified and the degree of structural variation known for this lipid class. As *Pyrobaculum* sp. AQ1.S2 expresses relatively high levels of a variety of ring-containing $GDGT_M$ lipids, it may serve as an ideal source species in order to provide lipid material for more complete structural characterisation of the components by degradative analysis. Furthermore, the organism should prove to be an ideal candidate for physiological stressing to help understand the physicochemical factors which may account for $GDGT_M$ lipid formation.

Chapter 5

Lipid core profiles of aquatic sediments and a terrestrial soil

5.1. Introduction

5.1.1. Ether lipids in the environment

Lipid profiling of isolated microorganisms carries the distinct advantage that the analyses are of a controlled system. Cultures of individual organisms may be expected to produce a limited variety of lipids, leading to relatively simple profiles and a high probability that each component can be characterised. Although the lipids expressed and the relative expression of particular components is usually dependent on the growth conditions, these can be controlled, monitored and, if necessary, varied. Profiling of lipids in environmental samples is significantly more difficult; there are usually multiple sources of lipids, with the complexity of the structural variety present being dependent on the phylogenetic diversity of the fauna living in the environmental setting. More importantly, natural environments are dynamic systems in which there are a huge number of variables (e.g. temperature, pH, pressure, nutrient availability, light intensity etc.), changes in any of which can lead to alteration in lipid expression by the progenitor microorganisms. Furthermore, organisms of different species are free to interact with one-another, which can lead to blurring of genetic boundaries by gene transfer. The complex lipid profiles that are invariably found in natural environments as a result of these factors are extremely challenging targets for analysis. The challenge is even greater for geological samples, given that the possibility of making additional measurements to aid in identification has long since disappeared. In all cases, the inherent complexity can often be disentangled through the identification and interpretation of taxonomic marker compounds for specific phyla, genera or species.

As discussed in Chapter 1, isoprenoid ether lipids prove useful as biomarkers for the Archaea in the natural environment and have been used to this effect following their identification in a wide variety of the environments in which archaea can live. The limited number of compounds to have been identified, coupled with the broad expression of most of these throughout the domain (Section 1.4.1) makes more specific interpretation difficult. A greater array of structures is needed if more restricted chemotaxonomic assignments are to made possible. As alluded to in

Chapters 3 and 4, this can be achieved by more detailed profiling of the lipids from culture, particularly those of lesser prominence that are often overlooked. Alternatively, a bottom-up approach, in which the complex lipid extracts from environmental samples are screened for novel structures of potential chemotaxonomic value, may prove to be more effective. Approaches of this nature have the added benefit that they target lipids originating from organisms yet to be isolated in pure culture or which are impossible to culture using current microbiological techniques. These two possibilities encompass a significant proportion of archaea living in disparate environments throughout the world (e.g. Schleper *et al.*, 2005; Teske and Sørensen, 2008). In addition, ether lipids originating from the eubacteria are not neglected *via* this approach. Currently, even less is known about the origins of ether lipids derived from organisms categorised within this second domain.

5.1.2. Aims

The aims of the studies described in this chapter were to provide a comprehensive investigation of the ether lipid cores in both modern and ancient environmental settings and, where possible, to identify or partially characterise the structures using LC-MS/MS. The implications of previously reported and novel structures to ether lipid biosynthesis, ecological profiling and geochemical reconstruction were also considered.

5.2. Results and discussion

5.2.1. Samples

Samples from three distinct settings were selected for preliminary investigation into the diversity of ether lipid structures in environmental samples. A series of marls (e.g. IV-1.9, X-1, XI-1b etc.) from a Messinian evaporite sequence (Vena Del Gesso (VDG) basin, Northern Italy) and a Jurassic shale (S90-11) from the Stewartby section of the Oxford Clay formation, UK were selected as representative examples of ancient aquatic sediments. A soil taken from a domestic greenhouse in Norton, UK was also chosen as a representative terrestrial sample. Lipid cores were extracted *via* sonication of each sediment in MeOH/DCM, followed by centrifugation and filtration to remove particulate matter. Each extract was analysed using LC-MS/MS method B and a MS scan range of *m/z* 900-1500, unless otherwise stated.

5.2.2. Structural characterisation

5.2.2.1. Novel ether lipid structures

LC-MS base peak chromatograms for the lipid extracts from each environmental sample show a large number of peaks relating to ether lipids, some of which have been reported previously, others of which have not. While a significant proportion of the components have MS/MS spectra suggestive of ether lipid structures, not all provide spectra that are of particular use at the present time in the assignment of more specific structural features. Consequently, while the remainder of this chapter provides a preliminary investigation into the possible structural identities of some of the ether lipid components present in the complex lipid extracts from the environmental samples, it is important to stress that the report is not exhaustive. The extracts from the VDG sediments contain the greatest variety of novel structures and, for examples where the lipid origin is not explicitly stated, it can be assumed that the component was identified in the extract from at least one VDG horizon. Lipids of particular structural relevance that were absent in the VDG extracts but identified in

the extracts from the greenhouse soil or from the Oxford Clay sediment are also discussed.

5.2.2.2. Assignment and nomenclature

Tetraether lipids and neutral losses from their protonated molecules during CID are assigned as described in Chapters 2-4. The nomenclature used to describe both the structures and neutral losses is identical to that used in Chapters 3 and 4, with one notable exception. As tetraether lipids containing cyclohexyl rings have been reported previously in aquatic sediments (e.g. Damsté *et al.*, 2002b) and in soils (e.g. Weijers *et al.*, 2006b) it is essential that neutral loss of a chain containing a Ch ring is distinguished from loss of a chain which contains only Cp rings. As such, alkyl chains assigned to contain N carbon atoms and P Cp rings are denoted [N_P], as in Chapters 3 and 4. Alkyl chains assigned as containing N carbons, P Cp rings and H Ch rings, on the other hand, are denoted [N_{P+H}]. Structures described as being isoprenoid in nature are those which contain ether-bound alkyl chains which are based solely on isopranyl subunits. Structures discussed as being non-isoprenoid are those which contain ether-bound chains that cannot be based solely on isopranyl sub-units. Within this second category, structures described as being of "mixed" character are those thought to contain ether-bound chains which are not fully isoprenoid in nature, but which may contain a short polyisopranyl section.

5.2.3. Isoprenoid lipids

5.2.3.1. Known isoprenoid tetraether lipid components

Several lipids, exhibiting $[M+H]^+$ with *m/z* values ranging from 1292.3 to 1304.4, were detected in the base peak chromatograms of the lipid extracts from the environmental samples, eluting during the period of the elution program in which the eluent IPA concentration ranged from 1-2% (e.g. Fig. 5.1a-b). The relative retention times of the components, coupled to the *m/z* value of the $[M+H]^+$ observed in each case, suggest the presence of lipids **2, 14-16**, **24** and **24''** in each of the samples examined (Hopmans *et al.*, 2000; Schouten *et al.*, 2000). In addition, other lipids

(**14'**, **15'** and **16'**) sharing identical *m/z* values to components **14**, **15** and **16**, were also observed in individual ion chromatograms for the extracts, in each case eluting marginally earlier than their isobaric counterpart (Fig. 5.1c). An earlier eluting isomer of lipid **24**, component **24'**, was observed in the chromatograms for the aquatic sediments (Fig. 5.1c). Following online CID, the protonated molecules of **2**, **14-16**, **14'-15'**, **24**, **24'**, and **24''** all generated product ions in two regions of the MS/MS spectrum, with specific neutral losses of 1-3 water molecules and C_3H_4O observed in each case, confirming the components as GDGT lipids. Additional losses of $[40_0]^0$ (-558 Da; lost from **2**, **14'** and **14**) and/or $[40_1]^0$ (-556 Da; lost from **14'**, **14**, **15'**, **15** and **16**), $[40_2]^0$ (-554 Da; lost from **16**, **24'**, **24** and **24'**) and $[40_{2+1}]^0$ (-552 Da; lost from **24'**, **24** and **24''**) were also observed. In addition, low abundance product ions formed from losses of biphytadienes containing zero and two rings were observed in the MS/MS spectra of lipids **15** and **15'**, confirming the presence of trace amounts of co-eluting isomers **15*** and **15'***, respectively. In each of the lipids containing two structurally distinct isoprenoid chains, loss of the chain containing fewer rings was observed to be more favourable based on the relative abundances of the product ions in the MS/MS spectra. This effect was particularly prominent in the spectra for lipids **24'**, **24** ($[M+H]^+$ *m/z* = 1292.3) and **24''**, where product ions at *m/z* 739 (-552 Da), relating to loss of an isoprenoid chain containing one Ch and two Cp rings, were far less prominent than ions at *m/z* 737 (-554 Da), relating to loss of an isoprenoid containing two cyclopentyl rings only (e.g. for **24** in Fig. 5.2). Lipid **16'** was found in low abundance in both VDG and Oxford Clay sediments and, although selected for CID, gave weak, unreliable MS/MS spectra during analysis of the extracts from each.

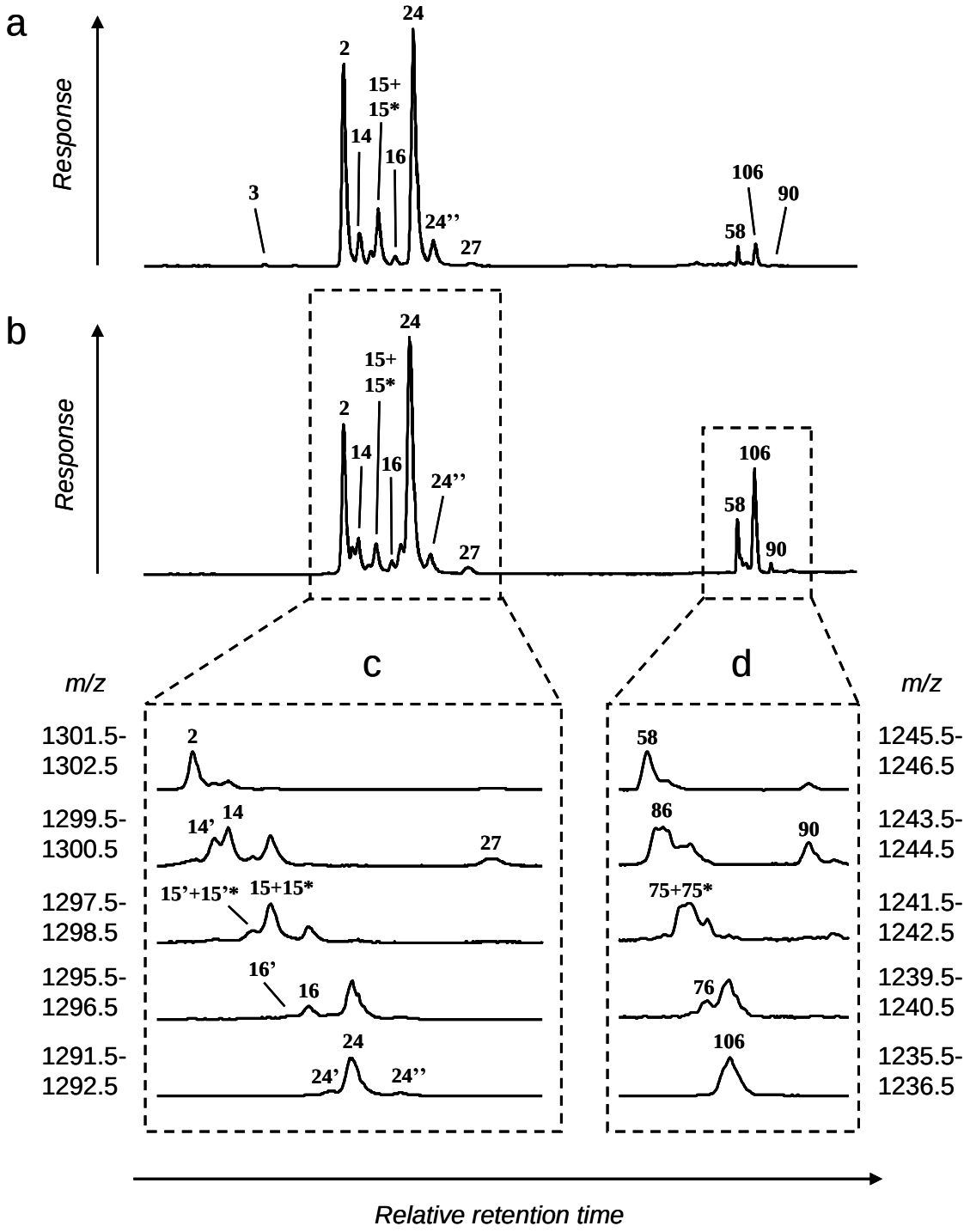


Fig. 5.1. HCT LC-MS ion chromatograms (*m/z* 1235.5-1304.5) for lipid extracts from: a) VDG horizon IV-1.9; b) Oxford Clay S90-11. c) and d) show individual ion chromatograms for several isoprenoid lipids found in the Oxford Clay sediment.

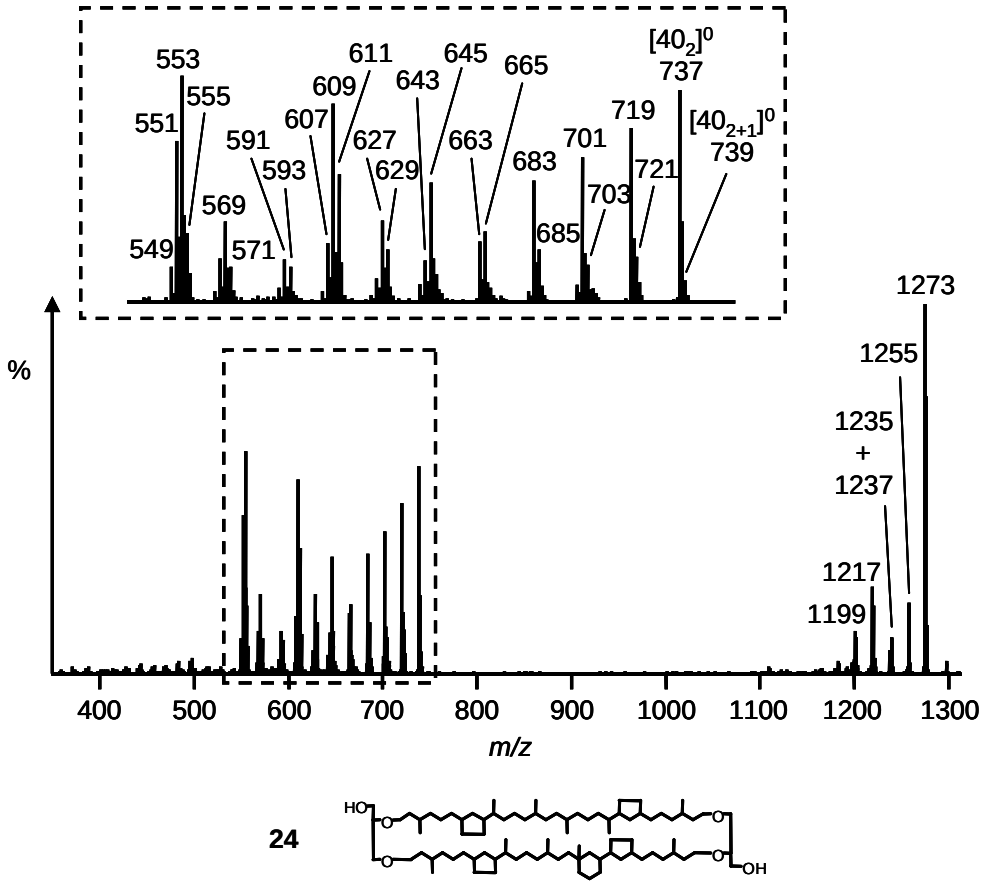


Fig. 5.2. MS/MS spectrum (HCT) of crenarchaeol, **24** ($[M+H]^+$ *m/z* = 1292.3). Inset shows expansion of a region of the spectrum.

The MS spectra obtained during elution of lipid **16** ($[M+H]^+$ *m/z* = 1296.4) in the analyses of the extracts from the aquatic sediments often showed a small additional peak for a component at *m/z* 1294.4 (e.g. VDG IV-1.9, Fig. 5.3a). Furthermore, individual ion chromatograms for *m/z* 1291.5-1292.5 and *m/z* 1293.5-1294.5 for the VDG extract show that the minor component partially co-elutes with **16** (Fig. 5.3b). It appears that the minor lipid is not discriminated from the $[M+H]^+$ of **16** by the 3 *m/z* unit window used during the precursor isolation procedure for CID and consequently, both **16** and the minor lipid are dissociated together, producing overlapping signals in MS/MS. This accounts for the presence of several product ions of low intensity in the averaged MS/MS spectrum obtained across the full peak width (i.e. across section τ in Fig. 5.3b). The ions present include *m/z* 1275 and *m/z* 737, which were not observed in the MS/MS spectrum of **16** obtained during LC-MS/MS of lipid extracts from *Sulfolobus* (see Table 3.4), suggesting these ions are

sourced from an alternative precursor. Confirmation that these two particular product ions originate from the minor lipid is provided by examination of the percentage abundances of each ion relative to product ions obtained from **16** in MS/MS spectral averages taken across sections of the chromatographic peak (i.e. across sections τ_1 and τ_2, Fig. 5.3b). The relative abundances of the two ions are greater during temporal section τ_1, in which the bulk of the minor lipid elutes, than in section τ_2 which represents a "cleaner" signal of **16**. The product ion at *m/z* 1275 relates to loss of water from the minor lipid. An ion at *m/z* 737 in the MS/MS spectra of **24** is assigned as a structure containing three remnant rings (i.e. two Cp and one Ch ring). Similarly, the 556 Da loss from the $[M+H]^+$ of the minor lipid, required to form a product ion at *m/z* 737, was also observed during CID of the $[M+H]^+$ of lipids **14**-**15**, where it is attributed to loss of $[40_1]^0$. As such, the minor component may represent a GDGT$[40_1,40_{2+1}]$ or GDGT$[40_1,40_3]$ structure, although further work is required to confirm this. A similar contamination of the product ion signal in the MS/MS spectrum of **24'** ($[M+H]^+$ *m/z* = 1292.4) was observed during analysis of the Oxford Clay sample, where minor ions at *m/z* 1271, 1253 and 1197 were detected, originating from a co-eluting lipid with $[M+H]^+$ at *m/z* 1290.4 (Fig. 5.3c).

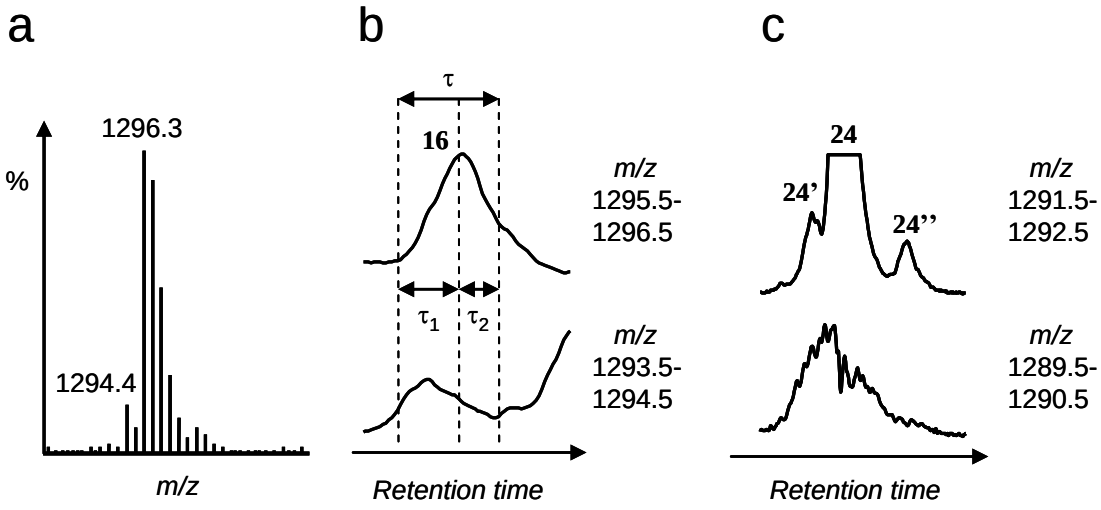


Fig. 5.3. a) Average mass spectrum obtained during elution of lipid **16** (i.e. during τ) in the LC-MS/MS (HCT) analysis of the lipid extract from the VDG IV-1.9 sediment; b) partial LC-MS ion chromatograms for lipid **16** and a co-eluting species in the same sample; c) partial LC-MS ion chromatograms (HCT) for lipids **24'** and **24** and a co-eluting species in the extract from the Oxford Clay S90-11 sediment.

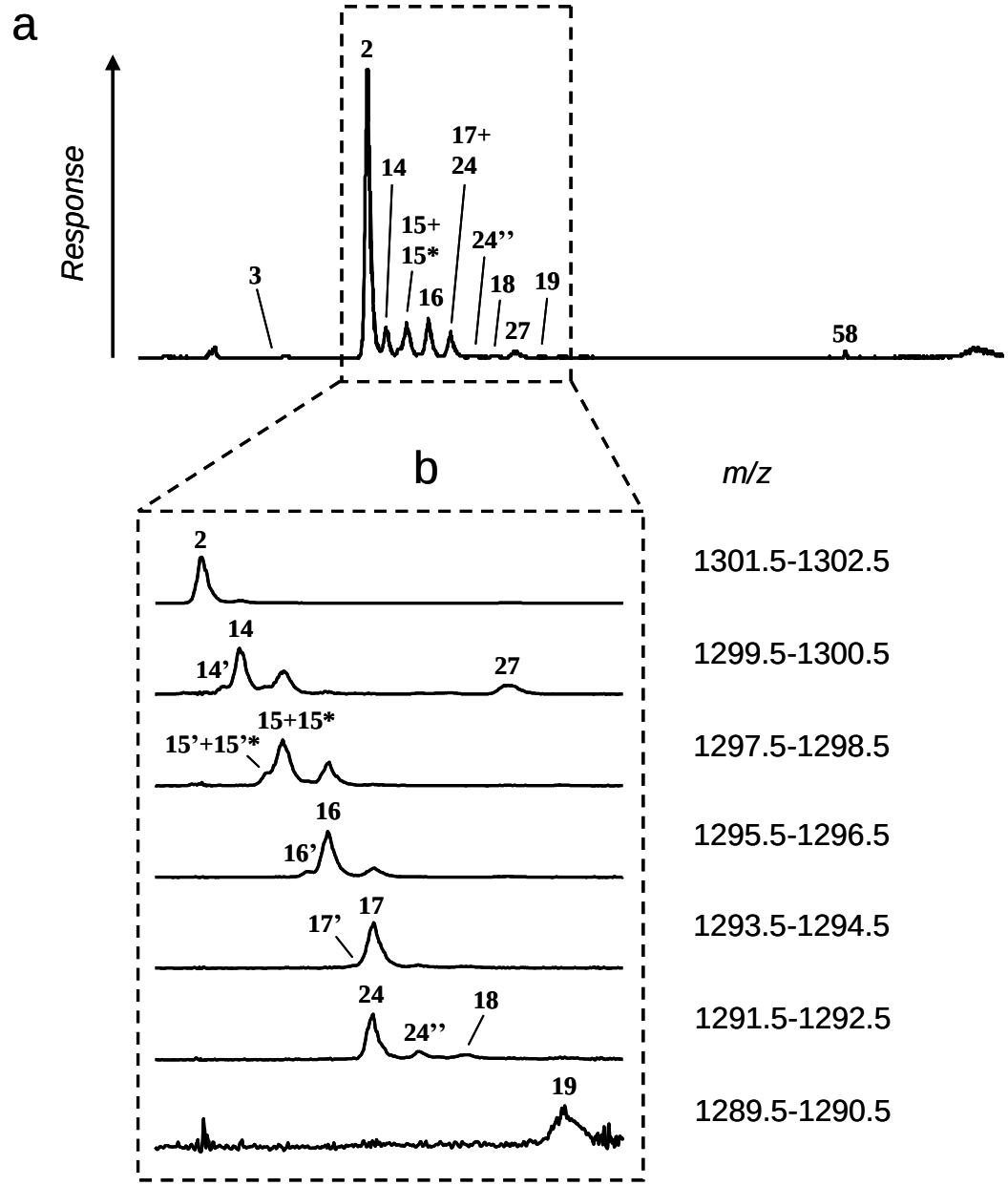


Fig. 5.4. HCT LC-MS ion chromatograms for the lipid extract from a greenhouse soil; a) base peak chromatogram (*m/z* 1235.5-1304.5); b) individual ion chromatograms highlighting isoprenoid lipids found in the sample.

Lipids **17**, **17'**, **18** and **19** were also identified either in the base peak chromatogram or in individual ion chromatograms for the lipid extract from the greenhouse soil (Fig. 5.4). The MS/MS spectra of **18** and **19** were too weak to be interpreted reliably, whereas the $[M+H]^+$ of **17'** was never the base peak ion in any MS scan and, consequently, was not selected for CID. Lipid **17** co-elutes with **24** under the chromatographic conditions employed during the LC-MS/MS studies (Fig. 5.4b), as was noted in previous LC-MS analyses of extracted archaeal core lipids in which identical or very similar chromatographic conditions were used (Weijers *et al.*, 2004; Zhang *et al.*, 2006; Schouten *et al.*, 2007c; de la Torre *et al.*, 2008; Pitcher *et al.*, 2009b). Although **17** is the more prominent lipid in the greenhouse soil, the MS/MS spectrum obtained from its $[M+H]^+$ at *m/z* 1293.7 contains product ions which appear to originate from both the intended precursor ion and the $[M+H]^+$ of co-eluting lipid **24**, exemplified by ions formed from loss of water from each precursor (*m/z* 1275 and 1273; Fig. 5.5a). A product ion at *m/z* 739 originates from the $[M+H]^+$ of both **17** and **24** *via* loss of $[40_2]^0$ and $[40_{2+1}]^0$, respectively, during CID, whereas an ion at *m/z* 737 originates exclusively from the $[M+H]^+$ of **24** *via* loss of $[40_2]^0$ (Fig. 5.5b-c).

Monoalkyl tetraether lipid **27** was identified in the base peak chromatograms for each extract on the basis of its retention time relative to **2**, the *m/z* value of the $[M+H]^+$ (*m/z* 1300.4) and from product ions observed in the MS/MS spectrum (Table 5.1), which compare favourably to those obtained for the component in extracts from thermophilic archaea (Sections 3.2.3 and 4.2.2.1). In extracts from the VDG horizons, an earlier eluting isobar of GMGT **27**, compound **27'**, was also observed (see Fig. 5.13). The $[M+H]^+$ of this lipid (*m/z* 1300.4) dissociates to give identical product ions to those observed from the $[M+H]^+$ of **27** during CID (Table 5.1). Trialkyl tetraether lipid **3** was also identified in the chromatograms for some of the aquatic sediments from comparison of LC-MS and MS/MS data to those of the authentic lipid in MTH(ΔH) (Section 3.2.3). Product ions observed to be formed from the $[M+H]^+$ of lipids **14**-**17** match those observed from these lipids in extracts from species of *Sulfolobus* (Section 3.2.4.5). Likewise, the MS/MS spectra of **14'**-**16'** contain the same product ions as those formed from components present in thermophilic and hyperthermophilic archaea (Sections 3.2.4.5 and 4.2.2.1). The

MS/MS data obtained for **24** and **24''** represent the first tandem mass spectral profiles of the lipids and support the assignments made *via* ^{1}H and ^{13}C NMR of the structures, and *via* electron ionisation mass spectra of their encapsulated isoprenoid hydrocarbons (Schouten *et al.*, 1998; Damsté *et al.*, 2002b).

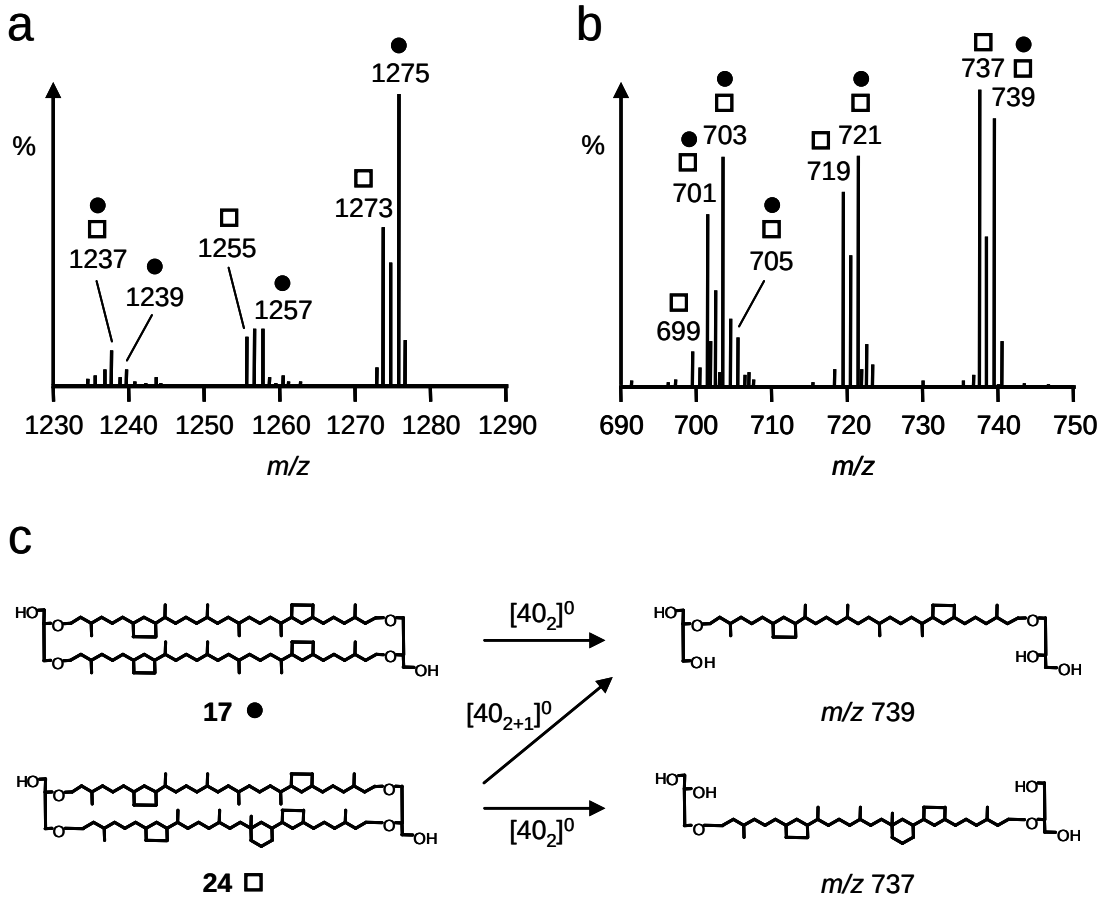


Fig. 5.5. a) and b) show regions of the MS/MS spectrum (HCT) obtained from co-eluting GDGTs **17** ($[M+H]^+$ *m/z* = 1293.7) and **24** ($[M+H]^+$ *m/z* = 1291.8) found in the greenhouse soil extract. Ions labelled with a filled circle originate from the $[M+H]^+$ of **17**, whereas those labelled with an open square originate from that of **24**; c) illustrates routes to product ions at *m/z* 739 and 737 from the two protonated precursors (ionising proton for each structure not shown).

Table 5.1. HCT MS data (measured $[M+H]^+$ *m/z* value) and MS/MS data (nominal *m/z* values of product ions) for GMGT and $GMGT_M$ lipid cores from VDG and Oxford Clay sediments and from the greenhouse soil.

		Isoprenoid		**Isoprenoid Homologue**		**Non-isoprenoid**				
Loss (Da)	Molecules Lost [a]	**27'**	**27**	**62**	**117**	**S2**	**S3**	**S4**	**S7**	**S8**
-0	$[M+H]^+$	1300.4	1300.4	1313.8	1327.8	1132.2	1118.2	1104.2	1034.1	1020.1
-18	[1,0,0]	**1281**[b]	**1281**	**1295**	**1309**	**1113**	**1099**	**1085**	**1015**	**1001**
-36	[2,0,0]	1263	1263	1277	1291	1095	1081	1067	997	983
-54	[3,0,0]	1245	1245	1259	1273	1077	1063	1049	979	965
-56	[0,1,0]	1243	1243	1257	1271	1075	1061	1047	977	963
-72	[4,0,0]	1227	1227	1241	1255	1059	1045	1031	961	947
-74	[1,1,0]	1225	1225	1239	1253	1057	1043	1029	959	945
-78	?	1221	1221	1235	1249	1053	1039	1025	955	941
-92	[2,1,0]	1207	1207	1221	1235	1039	1025	1011	941	927
-94	?	1205	1205	1219	1233	1037	1023	1009	939	925
-96	?	1203	1203	1217	1231	1035	1021	1007	937	923
-110	[3,1,0]	1189	1189	1203	1217	1021	1007	993	923	909
-112	[0,2,0]	1187	1187	1201	1215	1019	1005	991	921	907
-114	?	1185	1185	1199	1213	1017	1003	989		905
-128	[4,1,0]	1171	1171	1185	1199	1003	989	975	905	891
-130	[1,2,0]	1169	1169	1183	1197	1001	987	973	903	889
-146	?	1153	1153	1167	1181	985	971	957	887	873
-166	[3,2,0]	1133	1133	1147	1161	965	951	937	867	853
-168	?	1131	1131	1145	1159	963	949	935	865	851
-170	?	1129	1129	1143	1157	961	947	933	863	849
-184	[4,2,0]	1115	1115	1129	1143	947	933	919	849	835

[a] The number of simultaneous losses of each type of molecule during formation of the ion are listed as $[a,b,c_d]$, where a = water; b = C_3H_4O; c_d = biphytadiene containing "c" carbon atoms and "d" cyclopentyl rings (d = 0,1 or 2).

[b] Product ions listed in bold face represent the base peak in MS/MS.

5.2.3.2. Glycerol mono- and dialkanol triol lipids

In addition to the tetraether lipids, a series of novel, late eluting components, **58**, **86**, **75-76**, **106** and **90**, with $[M+H]^+$ ranging from *m/z* 1236.3 to 1246.3 was observed in the base peak chromatograms and in individual ion chromatograms for the extracts from aquatic sediments (Fig. 5.1). The retention times of the components indicate that each only elutes during LC-MS/MS analysis once the proportion of IPA in the eluent is substantially greater than 2% and consequently, that they are significantly more polar than isoprenoid GDGT lipids **2**, **14-16** and **24** and GMGT lipid **27**. Interestingly, each of the late eluting components (**58**, **86**, **75-76**, **106** and **90**) exhibits a $[M+H]^+$ at 56 *m/z* units less than one or more of the earlier eluting tetraether lipids (Fig. 5.1d). For example, GDGT **2** exhibits a $[M+H]^+$ at *m/z* 1302.4 and elutes a few minutes earlier than GDGT **24**, which exhibits a $[M+H]^+$ at *m/z* 1292.3. Novel lipid **58**, on the other hand, exhibits a $[M+H]^+$ at *m/z* 1246.3 and elutes before component **106**, which has a $[M+H]^+$ at *m/z* 1236.3. Similarly, lipid **90** (*m/z* 1244.3) elutes after lipids **106**, mirroring the later elution of GMGT **27** (*m/z* 1300.4) relative to GDGT lipid **24**.

The MS/MS spectra of lipids **58** (Fig. 5.6a) and **106** (Fig. 5.6b) show product ions in two regions of the spectrum and include those formed from loss of 1-2 molecules of water. By contrast with GDGT lipids, where the most favoured neutral loss in MS/MS is of a molecule of water, CID of the $[M+H]^+$ of **58** generates a base peak ion in MS/MS at *m/z* 669, formed *via* loss of 576 Da. Interestingly, a product ion with the same *m/z* ratio was also observed in the MS/MS spectrum of **2**, assigned as a protonated glycerol biphytanol monoether (see Section 2.2.5.2). Similarly, a 576 Da loss was also observed from the $[M+H]^+$ of **2** during CID and was attributed to loss of 1-ene-biphytan-32-ol (i.e. $[40_0]^1$; Section 2.2.5.2). Furthermore, the MS^3 spectrum generated from a precursor ion in MS^2 at *m/z* 1245, itself formed *via* loss of C_3H_4O from the $[M+H]^+$ of **2**, contains a product ion at *m/z* 669 generated *via* loss of 576 Da (Section 2.2.6). These observations strongly suggest that the $[M+H]^+$ of **58** has an identical structure to the product ion at nominal *m/z* 1245 in the MS/MS spectrum of **2**. Consequently, **58** represents a glycerol dialkanol triol (GDT) structure containing two hydroxyl-terminated biphytanyl chains, each ether linked *via* their

remaining carbon terminus to glycerol (representative structure shown in Fig. 5.7). The only structural difference between GDT **58** and GDGT **2** is the absence in the GDT of a C_3H_4O (56 Da) portion of one of the glycerol groups found in the GDGT. Assignment of **58** to a GDT structure also accounts for the greater preference for formation of the product ion at *m/z* 669 from the $[M+H]^+$: only a single elimination step is required for its formation whereas from the $[M+H]^+$ of **2**, at least two sequential eliminations are needed. Losses in MS/MS of 572 Da (*m/z* 663) and 570 Da (*m/z* 665) from the $[M+H]^+$ of **106** are attributed to $[40_2]^1$ and $[40_{2+1}]^1$ respectively, suggesting that **106** also has a GDT structure and is most likely a structural relative of GDGT **24,** and possibly **24'** and **24''** (representative structure shown in Fig. 5.7). The MS/MS spectra of **86** and **75-76** show losses of biphytanols containing 0-2 rings (Table 5.2), confirming that the components represent the GDT structural relatives of GDGTs **14-16** (representative structures shown in Fig. 5.7). A weak product ion at *m/z* 665 in the MS/MS spectrum of **75**, formed *via* loss of $[40_0]^1$ (-576 Da) from the $[M+H]^+$ (*m/z* = 1242.3) is suggestive of a co-eluting isomer, **75***, that contains zero rings on one isoprenoid chain and two rings on the other. This lipid is most likely structurally related to GDGT **15***, which contains the same combination of alkyl chains. It is notable that lipids **86**, **75-76** and **106** each elute as single peaks in the LC-MS chromatograms, as opposed to the evidence of isomerism observed for lipids **14-16** and **24**, which all have neighbouring peaks relating to at least one other isomeric form in the LC-MS ion chromatograms (Fig. 5.1c-d). Consequently, it is unclear at present if the peaks observed for compounds **58**, **86**, **76** and **106** represent individual components or mixtures of isobaric lipids. Similarly, the single peak observed for lipids **75** and **75*** could conceivably encompass other isobars of these two lipids. Regardless of the details, the presence of three primary hydroxyls in **58**, **86**, **75-76** and **106** would explain the substantial increase in the retention time of these components relative to **2**, **14-16** and **24**, which contain only two terminal hydroxyl groups.

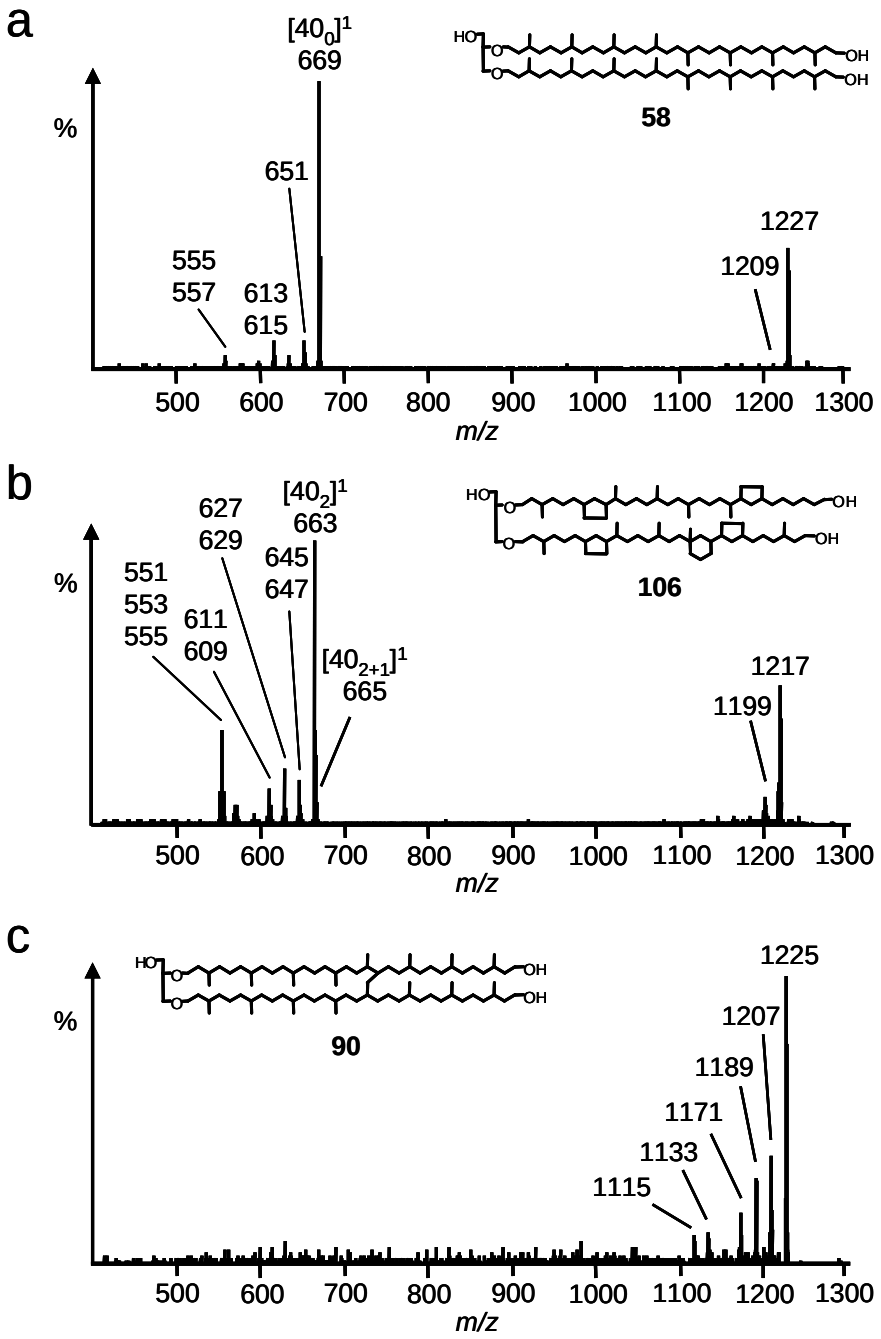


Fig. 5.6. MS/MS spectra (HCT) of: a) GDT **58** ($[M+H]^+$ *m/z* = 1246.3); b) GDT **106** ($[M+H]^+$ *m/z* = 1236.3); c) GMT **90** ($[M+H]^+$ *m/z* = 1244.3). Structures shown are representative structures; the location and/or orientation of the etherified glycerol group in each structure may differ from that shown. Likewise, the isoprenoid cross-link in GMT **90** may be placed differently.

The MS/MS spectrum of **90** (Fig. 5.6c and Table 5.2) shows product ions at *m/z* 1225, 1207, 1189 and 1171, formed from losses of 1-4 molecules of water, and ions of lower intensity at *m/z* 1133 (-110 Da) and 1115 (-128 Da), formed by loss of C_3H_4O and three or four molecules of water, respectively. No product ions consistent with loss of an isoprenoid hydrocarbon are observed, suggesting **90** most likely has a monoalkyl lipid structure. Each of the product ions formed from the $[M+H]^+$ of **90** is also observed in the MS/MS spectrum of GMGT **27**, where they were attributed to structures formed from loss of at least one molecule of C_3H_4O from the precursor (see Table 5.1). These observations, coupled with the *m/z* ratio observed for the $[M+H]^+$ of the lipid and the retention time of the component relative to **58** and the ring-containing GDT lipids lead to the assignment of **90** as a glycerol monoalkanol triol (GMT), a structural relative of GMGT **27** (representative structure shown in Fig. 5.7). Further support for this assignment is provided by consideration of the product ion intensity distributions observed in the MS/MS spectra of **90** and **27**. Although a product ion formed from loss of water is the base peak ion in the spectra of both lipids, it is more dominant in the spectrum of **90** than in that of **27**. The apparent preference for loss of water from **90** is consistent with the additional free hydroxyl in the structure compared to that of **27**.

GDTs **58**, **86** and **75-76** and GMT **90** were identified during LC-MS/MS analyses of various lipid extracts from species of *Methanothermobacter*, *Sulfolobus*, *Ignisphaera* and *Desulfurococcus* (see Sections 3.2.3.3, 3.2.4.5 and 4.2.4), where they exhibited MS/MS spectra containing identical product ions to those generated from the components in the environmental extracts. Not all of the components were observed in the extracts of each organism. Additional GDT lipids, **77-79**, **78'** and **79***, containing 4-6 Cp rings were also identified in two species of *Sulfolobus* (Section 3.2.4.5). Other GMT lipids (**91**-**94**), containing 1-4 Cp rings, were identified in *Ignisphaera aggregans* (Section 4.2.4.1). Representative structures for each GDT and GMT lipid are shown in Fig. 5.7.

Table 5.2. Product ions observed in the MS/MS spectra of GDT and GMT lipids identified in VDG and Oxford Clay sediments, a greenhouse soil and in thermophilic archaea belonging to the genera *Methanobacter*, *Sulfolobus*, *Ignisphaera* and *Desulfurococcus*.

			Isoprenoid													
			GDT									**GMT**				
			58	**86**	**75 75***	**76**	**77**	**78'**	**78**	**106**	**79 79***	**90**	**91**	**92**	**93**	**94**
Molecules lost	Loss (Da)	$[M+H]^+$ (nominal)	1245	1243	1241	1239	1237	1235	1235	1235	1233	1243	1241	1239	1237	1235
H_2O	-18		+[a]	+	+	+	+	+	+	+	+[b]	+	+	+	+	+
$2H_2O$	-36		+	+	+	+	+	+	+	+	+	+	+	+	+	+
$3H_2O$	-54											+	+	+	+	+
$4H_2O$	-72											+	+	+	+	+
$C_3H_4O + H_2O$	-74											+	+	+	+	+
$C_3H_4O + 2H_2O$	-92											+	+	+	+	+
?	-94											+	+	+	+	+
?	-96											+	+	+	+	+
$C_3H_4O + 3H_2O$	-110											+	+	+	+	+
?	-112											+	+	+	+	+
?	-114											+	+	+	+	+
$C_3H_4O + 4H_2O$	-128											+	+	+	+	+
$[30_0]^1$	-436															
$[31_0]^1$	-450															
$[32_0]^1$	-464															
$[33_0]^1$	-478															
$[34_0]^1$	-492															
$[35_0]^1$	-506															
$[36_0]^1$	-520															
$[37_0]^1$	-534															
$[40_0]^1$	-576		+	+	+											
$[40_1]^1$	-574			+	+	+										
$[40_2]^1$	-572				+	+	+	+	+	+	+					
$[40_3]^1$	-570							+	+		+					
$[40_{2+1}]^1$	-570									+						
$[40_4]^1$	-568										+					

[a] (+) = Product ion observed in the MS/MS spectrum.

[b] Product ions shaded grey (where shown) represent the ion observed as the base peak ion in MS/MS.

Non-isoprenoid																			
GDT														GMT					
K1	**K2**	**K3**	**L1**	**L2**	**L3 L3***	**L7**	**L9**	**L10**	**L11**	**L13 L13***	**L14**	**L15**		**T1**	**T2**	**T3**	**T4'**	**T4**
1161	1147	1133	1161	1147	1133	1077	1049	1035	1021	993	979	965		1075	1061	1047	963	963
+	+	+	+	+	+	+	+	+	+	+	+	+		+	+	+	+	+
+	+	+		+	+	+	+	+	+	+	+	+		+	+	+	+	+
														+	+	+	+	+
														+	+	+	+	+
														+	+	+	+	+
														+	+	+	+	+
														+	+		+	+
																		+
														+	+	+	+	+
														+	+	+	+	+
															+			+
														+	+	+	+	+
										+	+	+						
										+	+							
								+	+	+								
							+	+										
						+												
					+													
	+	+		+	+													
+	+		+	+	+													

Table 5.2. (cont.)

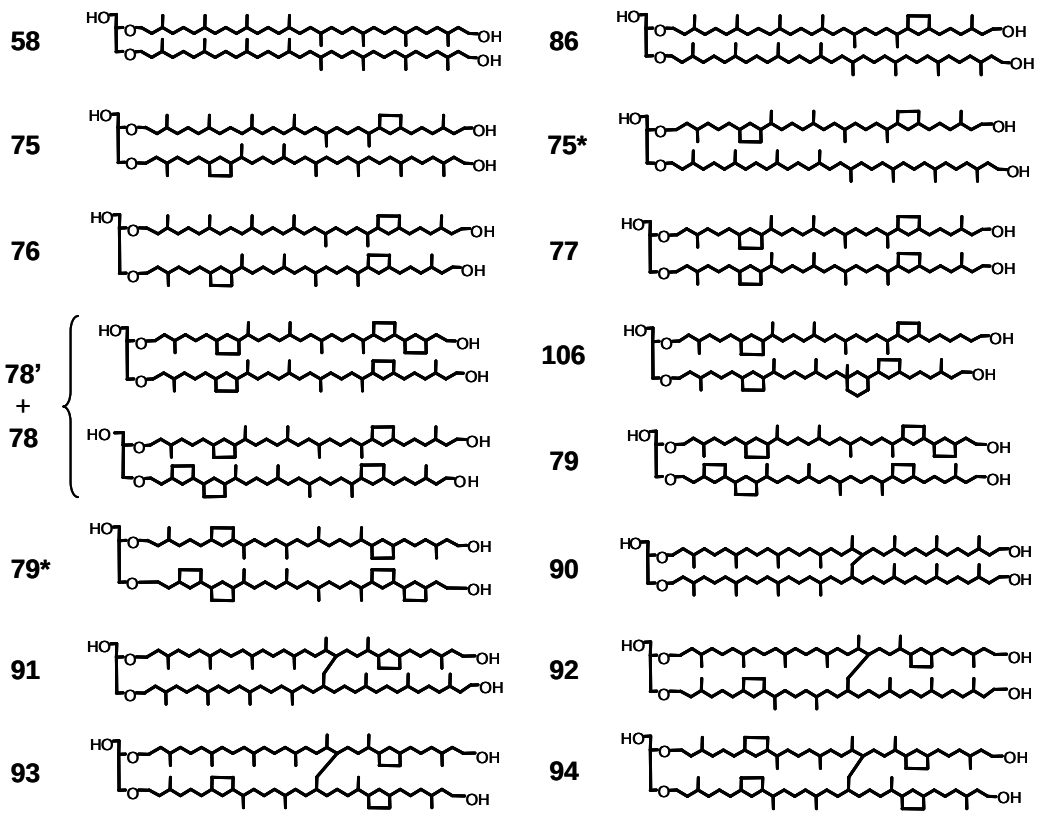


Fig. 5.7. Representative structures of GDT and GMT lipids found in lipid extracts from environmental samples and (hyper)thermophilic archaea. Note that the location and orientation of the etherified glycerol group and/or the position of the isoprenoid cross-link in each structure may differ from that shown.

In an attempt to verify the structural assignments of the triol (i.e. GDT and GMT) lipids, the lipid extract from the S90-11 Oxford Clay sample was treated with acetic anhydride in pyridine. It was envisaged that any structures in the extract which contained free hydroxyl groups would be acetylated under these conditions. Consequently, if **58**, **86**, **75-76**, **106** and **90** were triols, as suggested, they would be converted to triacetates following this procedure. Any structural modification could subsequently be monitored by examining the MS/MS spectra of the acetates formed. The acetylated extract was separated into three fractions on alumina, each of which was analysed separately *via* reversed-phase LC-MS/MS. Unfortunately, during these analyses it was clear that the acetylation reaction had not proceeded to completion, with far greater complexity being observed in the chromatograms for the acetylated fractions than could be accounted for by conversion of each component in the

original extract to a single acetylation product. For example, GDGTs **2** and **24** were present in monoacetylated and diacetylated forms amongst the reaction products. Unreacted **2** and **24** were also identified. The problem was further compounded by distribution of the acetates between the three column fractions, with some components being observed in the LC-MS chromatograms of two or more fractions. By analysing fraction 2 using a narrow mass spectral scan range, carefully chosen to encompass the *m/z* values expected for triacetylated variants of **58**, **86**, **75-76**, **106** and **90** (*m/z* 1361.5-1372.5), a series of components (**107-111**) with $[M+H]^+$ at 126 *m/z* units greater than those observed for **58**, **86**, **75-76** and **106** were revealed in the base peak chromatogram (Fig. 5.8a). The absence of a component similarly related to **90** may be due to the relatively low levels of the lipid in the original Oxford Clay extract. The MS/MS spectrum (Fig. 5.8b) of **107** ($[M+H]^+$ *m/z* = 1372.3) includes product ions at *m/z* 1329, 1311 and 1251, formed from losses of 42 (ketene), 60 (acetic acid) and 120 Da (two molecules of acetic acid), losses also observed during CID of the $[M+H]^+$ of **56**, a synthetic diacetate of GDGT **2** (Section 2.2.10.3). These product ions indicate that **107** contains at least two acetate groups. Product ions formed from losses of 636 (*m/z* 735), 696 (*m/z* 675) and 756 Da (*m/z* 615) are more informative and are attributed to loss of 1-acetoxybiphytan-32-ol with additional losses of one or two molecules of acetic acid. These observations agree favourably with the preferential loss from the $[M+H]^+$ of diacetylated tetraether **56** of a biphytanyl chain as an alcoholic as opposed to an unsaturated elimination product during CID (Section 2.2.10.3) and suggest that both of the biphytanyl chains in **107** are terminated with an acetate group. In addition to loss of the acetylated biphytanol chain, the subsequent loss of up to two further molecules of acetic acid from the $[M+H]^+$ of **107** show that this component contains a total of three acetate groups. The MS/MS spectrum (Fig. 5.8c) of **111** ($[M+H]^+$ *m/z* = 1362.2) shows product ions formed from losses of 630 Da (*m/z* 731) and 632 Da (*m/z* 729), relating to 1-acetoxybiphytan-32-ols containing three and two rings, respectively. As before, ions formed from losses of one (*m/z* 671 and 669) or two (*m/z* 611 and 609) additional molecules of acetic acid are also observed. Similarly, the $[M+H]^+$ of lipids **108-110** lose acetoxybiphytanols containing 0-2 rings and up to two further molecules of acetic acid during CID. In conclusion, the MS/MS data for **107-111** confirm that they represent the triacetylated derivatives of lipids **58**, **86**, **75-76** and **106**. By

inference, this would indicate that the latter components each contain three free hydroxyl groups, consistent with their proposed GDT structures.

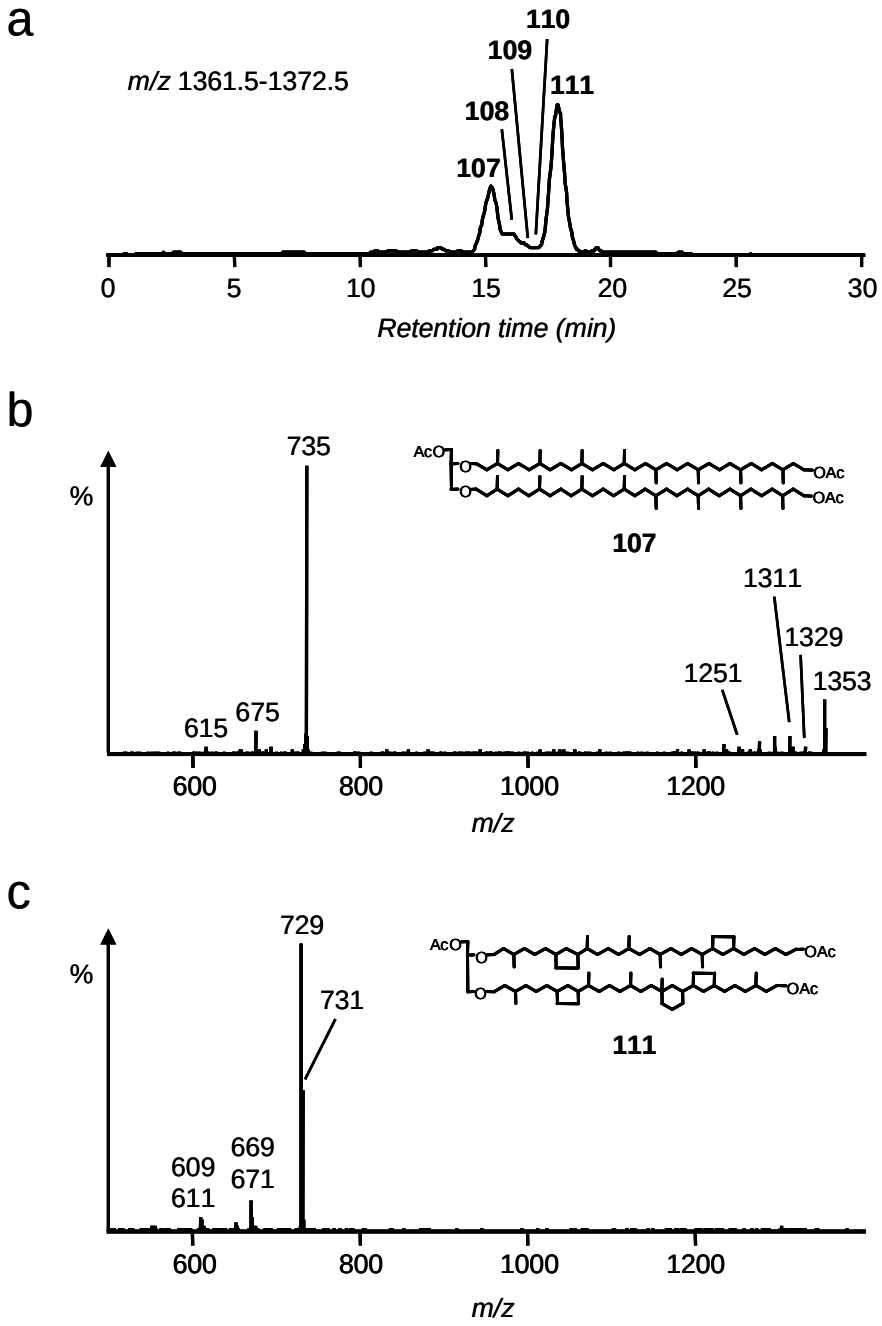


Fig. 5.8. a) Reversed-phase HCT LC-MS ion chromatogram for the acetylation products (fraction 2) formed from the Oxford Clay S90-11 extract; b) MS/MS spectrum (HCT) of **107** ($[M+H]^+$ *m/z* = 1372.3); c) MS/MS spectrum (HCT) of **111** ($[M+H]^+$ *m/z* = 1362.2).

5.2.3.3. Triether tetrol lipids

Two components, **112** and **112'**, each with $[M+H]^+$ at *m/z* 1320.4, were observed to elute at the very end of the base peak chromatogram of the S90-11 Oxford Clay extract using LC-MS/MS method B. Structure **112** had only partially eluted when the analysis was terminated (i.e. at 90 min). In order to facilitate complete elution of these components, the extract was re-analysed *via* LC-MS/MS using method C, in which the final isocratic eluent composition of 15% IPA was maintained for a further 15 min (total analysis time = 105 min; see Section 7.6.1). Using this program, **112** eluted fully, along with other components including **113** and **113'** which exhibit protonated molecules at *m/z* 1310.3 (Fig. 5.9). Reconstructed ion chromatograms for *m/z* 1317.5-1318.5 and 1315.5-1316.5 (not shown) indicate that components with *m/z* 1318.4 and 1316.4 partially co-elute with either **112**/**112'** or **113**/**113'**. Components **112** and **112'** have protonated molecules at 18 *m/z* units greater than that of GDGT **2**. Similarly, the protonated molecules of **113** and **113'** are 18 *m/z* units greater than those of GDGTs **24**, **24'** and **24''**. The differences in $[M+H]^+$ between **2** and **112**/**112'** suggests that the latter may formally represent structures in which a molecule of water has been added to **2**.

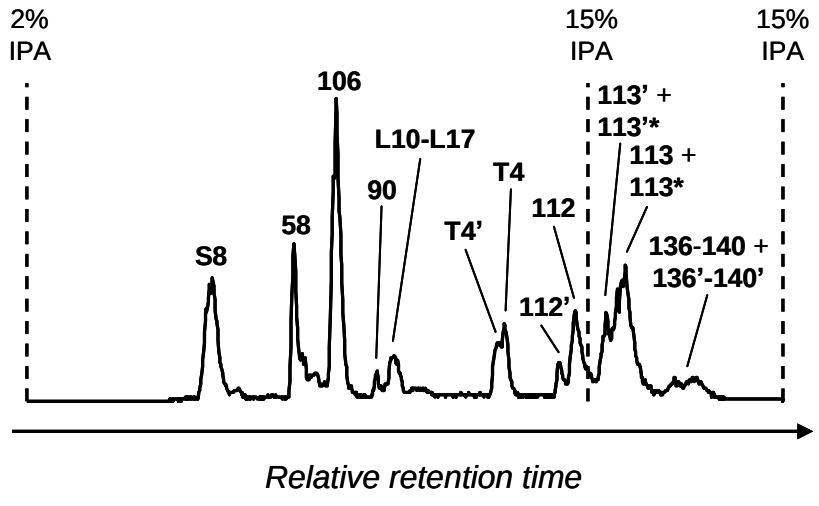


Fig. 5.9. Partial base peak chromatogram (*m/z* 900-1500; HCT) of the lipid extract from the Oxford Clay S90-11 sediment (recorded using LC-MS/MS method C), highlighting components eluting after the proportion of IPA in the mobile phase had exceeded 2% and, subsequently, been held isocratically at 15%.

The MS/MS spectra of **112** and **112'** (Table 5.3) each exhibit product ions formed from losses of 1-3 molecules of water (*m/z* 1301, 1283 and 1265), loss of C_3H_4O with simultaneous loss of one or two molecules of water (*m/z* 1245 and 1227) and loss of $[40_0]^1$ (-576 Da; *m/z* 743). Loss of $[40_0]^1$ was noted as the smallest isoprenoid-bearing molecule loss during CID of the $[M+H]^+$ of GDT **58**, which contains ether-bound biphytan-32-ol chains (Section 5.2.3.2). By association, this would indicate that **112** and **112'** may contain at least one similar moiety. As such, **112** and **112'** may be triether tetrol structures, related to **2** but in which the macrocycle has been opened by addition of water across a C-O ether bond (Fig. 5.10). As there are primary and secondary ether linkages in **2**, the opening of either would generate two distinct structures, perhaps explaining the presence of **112** and **112'** as an isomeric pair. On account of the reduced steric hindrance at the primary ether linkages, it may be expected that these are the most likely sites for oxidation/hydration, indicating that **112**, as the more prominent isomer in the chromatograms, might represent the primary ether cleavage product and **112'** the secondary ether cleavage product. These assignments would also explain the elution order of **112**, which is proposed to contain four primary hydroxyl groups, and **112'**, which is proposed to contain three primary hydroxyl groups and a less polar secondary hydroxyl group. The elution of **112** and **112'** later than both GDGT **2** and GDT **58**, which contain two and three free hydroxyls, respectively, is also consistent with the assigned structures of each.

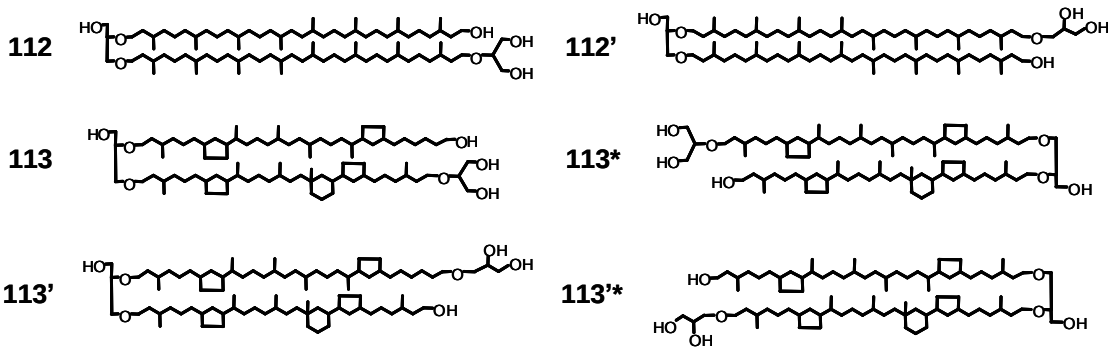


Fig. 5.10. Putative structures of isoprenoid glycerol triether tetrols identified in the Oxford Clay S90-11 sediment. Other isomeric forms may also be present.

The MS/MS spectra of **113** and **113'** (Table 5.3) also show ions formed from losses of 1-3 molecules of water (*m/z* 1291, 1273 and 1255), additional ions formed from concerted loss of C_3H_4O with 1-3 molecules of water (*m/z* 1235, 1217 and 1199) and ions consistent with losses of $[40_2]^1$ (-572 Da; *m/z* 737) and $[40_{2+1}]^1$ (-570 Da; *m/z* 739). Consequently, **113** and **113'** may represent structures in which an ether bond in **24**, **24'** or **24''** has been oxidised or hydrated. Since the two composite chains in each of **24**-**24''** contain differing numbers of rings, oxidation/hydration of either of the two primary ether linkages in these structures would lead to structurally distinct cleavage products (Fig. 5.10). As such, **113** may comprise a mixture of at least two types of isomer, one formed by cleavage of the primary ether linkage to a chain containing two Cp rings (designated **113**) and the other formed from cleavage of the primary ether linkage to a chain containing two Cp rings and one Ch ring (designated **113***). Similarly, oxidation/hydration of distinct secondary ether linkages would give two further isomeric structural types (Fig. 5.10); those having an etherified biphytan-32-ol chain containing two Cp rings and one Ch ring (designated **113'**) and those having an etherified biphytan-32-ol chain containing two Cp rings (designated **113'***). The components with $[M+H]^+$ at *m/z* 1318.4 and 1316.4 may represent triether tetrols related to GDGTs containing one or two Cp rings, respectively.

Table 5.3. HCT MS data (measured $[M+H]^+$ *m/z* values) and MS/MS data (nominal *m/z* values of product ions) for isoprenoid triether tetrols identified in an extract from the S90-11 Oxford Clay sediment.

Lipid	$[M+H]^+$	Product ions observed in MS/MS [a]	Assigned as:
112'	1320.4	**1301**,1283,1265,1245,1227; 743,725,707,689,669,651,633,615,557	**2** + 18 Da
112	1320.4	**1301**,1283,1265,1245,1227; 743,725,707,689,669,651,633,615,613,597,595,557	**2** + 18 Da
113'+113'*	1310.3	**1291**,1273,1255,1235,1217,1199; 739,737,721,719,703,701,685,683,665,663,647,645, 643,629,627,611,609,571,569,567,555,553,551	**24** + 18 Da
113+113*	1310.3	**1291**,1273,1255,1235,1217,1199; 739,737,721,719,701,685,683,665,663,647,645,643, 629,627,625,611,609,591,571,569,567,555,553,551	**24** + 18 Da

[a] Product ions listed in bold face represent the base peak in MS/MS.

A further complication could arise if **2** and **24**-**24''** found in the sediment are comprised of isomeric forms in which the glycerol moieties are arranged either *anti*

(e.g. **2a**) or *syn* (e.g. **2b**) to one-another. Cleavage at either the *sn*-2 or the *sn*-3 positions of the *syn* forms would generate distinct triether tetrol structures to those formed *via* cleavage of the *anti* forms at the same position. As such, **112**, **112'**, **113**/**113*** and **113'**/**113'*** may be accompanied by other co-eluting isomers, formed *via* cleavage of *syn* GDGT structures.

In spite of the preliminary evidence, all assignments of the triether tetrols are tentative and each component will require further analysis before more definitive assignments can be made.

5.2.3.4. Other potential isoprenoid lipids

A significant number of other components, potentially isoprenoid in nature, were observed in the ion chromatograms for the VDG sediment extracts, some found ubiquitously, others being specific to particular horizons. These include lipids with $[M+H]^+$ at nominal *m/z* values of 1417, 1401, 1399, 1349, 1347, 1343, 1333, 1329, 1317, 1315, 1299, 1297, 1295, 1290, 1285, 1275, 1271 and 1261. In many cases, MS/MS spectra were obtained, confirming an etherified glycerol functionality and/or the probable presence of isoprenoid alkyl chains within the structure. In some cases, a more specific structural assignment was possible, with monoacetylated caldarchaeol, monoformylated caldarchaeol, a caldarchaeol methyl ether and two isomers of monoacetylated crenarchaeol tentatively assigned in several horizons, although it is unclear as to whether these are authentic or artificial components. Despite the prominence of some of the other novel components in the LC-MS ion chromatograms, the MS/MS spectra and implications of these data are not further discussed in the interests of brevity.

5.2.3.5. Significance

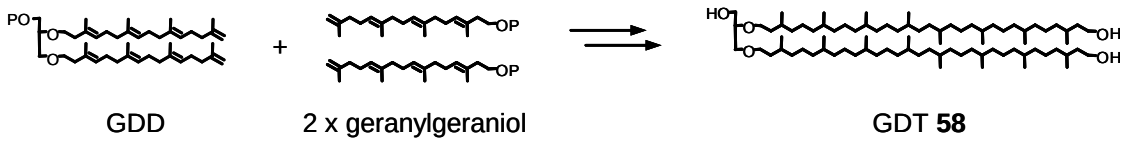
A series of components exhibiting mass spectra and tandem mass spectra consistent with isoprenoid tetraether lipid structures were identified in aquatic sediments from both depositional locations and in the greenhouse soil sample. The structures of lipids **2**, **3**, **14-16** have been fully characterised previously (Chapters 3 and 4 and

references therein) and, along with fully characterised lipid **24** and partially characterised lipid **24''** (Damsté *et al.*, 2002b), have been shown to be significant contributors to sedimentary organic matter in modern and ancient environments (e.g. Schouten *et al.*, 2000; Powers *et al.*, 2004; Kim *et al.*, 2008), including in VDG sediment (Hoefs *et al.*, 1997; Schouten *et al.*, 2000). Compounds **2**, **14**-**16** and **24** have also been identified previously in terrestrial soils (e.g. Weijers *et al.*, 2006b; 2007c). Although **17**-**19** have been identified in samples from a number of thermal environments (Zhang *et al.*, 2006; Pancost *et al.*, 2006; Schouten *et al.*, 2007c; Reigstad *et al.*, 2008; Pearson *et al.*, 2008), their identification in a soil from a lower temperature environment appears to be unprecedented. Lipid **27** has been identified in marine and lacustrine sediments (Schouten *et al.*, 2008a) and partially characterised on the basis of spectrometric and spectroscopic measurements (see Section 4.2.2.3). GMGT **27'** is almost certainly an isomer of GMGT **27**, potentially one in which the covalent bond tethering the isoprenoid chains is between a different pair of carbon atoms to those linked in the latter structure. Lipids **14'**, **15'**, **16'** and **24'** were identified in aquatic sediments from both depositional environments and have also been observed previously in other sediments (e.g. Kuypers *et al.*, 2002), where, on the basis of the *m/z* values of the $[M+H]^+$ and chromatographic behaviour, they were tentatively assigned as corresponding regioisomers of lipids **14**, **15**, **16** and **24**. The MS/MS spectra of **14'**, **15'** and **24'** show that they each lose isoprenoid chains containing the same numbers of rings as those lost from their structural counterpart (i.e. from **14**, **15** or **24**), strengthening this argument. Lipid **15***, identified as a trace component in each extract, has been reported in previous LC-MS analyses of the lipid extracts from two marine sediments (Schouten *et al.*, 2000; Pancost *et al.*, 2001b), where it was also observed to co-elute with **15**. In these two studies, isolation of GDGT lipids and their subsequent chemical degradation to their composite hydrocarbons was required in order for **15*** to be identified, a process too cumbersome to be applied routinely. This may explain the lack of other reports of this lipid in environmental samples, which have recently been profiled almost exclusively by LC-MS analysis. An additional isomer of **15**, **15'***, which co-elutes with **15'**, was also detected in a similar manner. As discussed in Section 4.2.2.3, further structural studies are necessary in order to assign all of the minor isomers conclusively. Non-isobaric lipids **17** and **24**, which co-elute during LC-MS/MS

analysis, appear to be indiscriminately selected for CID if both are present in the same extract. Attempts to assign tetraether lipid core structures *via* LC-MS/MS should be restricted to lipids which exhibit a "clean" peak (i.e. the peak does not overlay any co-eluting lipids with $[M+H]^+$ of similar *m/z*) in LC-MS (see Section 3.2.4.4). Nevertheless, the lack of precursor selectivity for **17** over **24** during CID, coupled to the unique origin of a product ion at *m/z* 737 in the MS/MS spectrum from the latter structure, may aid in confirmation of the presence of **24** in samples containing both lipids.

Late-eluting isoprenoid lipids **58**, **86**, **75-76**, **106** and **90** have been tentatively identified by LC-MS/MS, apparently representing the first such identifications of these components in environmental samples. The lack of previous reports of these structures is probably attributable to the retention of these components on-column for the duration of previous LC-MS analyses of environmental extracts, which typically employed a lower polarity in the final eluent (Hopmans *et al.*, 2000; Schouten *et al.*, 2007b; Escala *et al.*, 2007) than that used in the present study. The spectral data obtained for **58**, **86**, **75-76** and **106** suggest that they are GDT lipids: structures resembling GDGTs but with part of a terminal glycerol group absent. Similarly, **90** most likely represents a GMT structure, resembling GMGT **27** in an analogous fashion. Attempts to verify the structures of the GDT and GMT lipids *via* acetylation of the total polar lipid fraction from the Oxford Clay sediment were moderately successful, although acetylation of individual GDT lipids following isolation of each component from the total extracts should allow more definitive assignments to be made in future. In addition to the triols identified in the aquatic sediments, further structural variants, including GDT lipids containing 4-6 Cp rings and GMT lipids containing 1-4 Cp rings, were identified in lipid extracts from thermophilic archaea (Chapters 3 and 4). It is currently unclear as to which glycerol group is "missing" in the triols with respect to the corresponding GDGT/GMGT lipid. It is possible that each GDT/GMT lipid identified in the sediment, soil and culture extracts comprises a mixture of isomeric structures in which opposite glycerol termini are missing. If this is the case, these isomeric pairs are inseparable by LC-MS/MS.

The apparent structural similarity in the environmental samples between GDT/GMT lipids and GDGT/GMGT lipids containing the same number of rings is unlikely to be coincidental, although there are a number of possible ways in which the similarity could have arisen. The close structural relationship could be a result of biosynthetic steps that are conservative in the pathways to both tetraether and triol lipids. The biosynthesis of GDGT lipids is reasonably well understood (see Section 1.2.1 and Koga and Morii, 2007). To summarise, GDGT lipids are known to form following coupling of terminal carbon atoms found in two diether lipids. The diether precursors are formed at an earlier stage in the biosynthetic sequence, by attachment of two geranylgeraniol units to glycerol-1-phosphate (Koga and Morii, 2007). There is debate as to whether the isoprenoid chains in the diethers are reduced to saturated hydrocarbons before the coupling occurs to generate the completed tetraether (Nemoto *et al.*, 2003) or whether this reduction occurs afterwards (Eguchi *et al.*, 2000; 2003). Similarly, it is not known whether incorporation of Cp rings into the isoprenoid chains and formation of a cross-link between isoprenoid chains (as found in GMGT lipids) occur simultaneous to, or following completion of the coupling step (De Rosa *et al.*, 1980c; Weijers *et al.*, 2006a). Given that the carbon termini of the diether lipids and geranylgeraniol are the same, it is possible that small amounts of the latter may be incorporated into the coupling process. If two molecules of geranylgeraniol were co-condensed with one diether molecule, the requisite GDT-type structure would be established (Scheme 5.1). Further reduction and/or dephosphorylation could yield GDT **58**. Ring incorporation and/or isoprenoid cross-linking, either during or after coupling of geranylgeraniol units to the diether may, in turn, account for the formation of GDTs **86**, **75**-**79** and **106** and GMTs **90**-**94**.

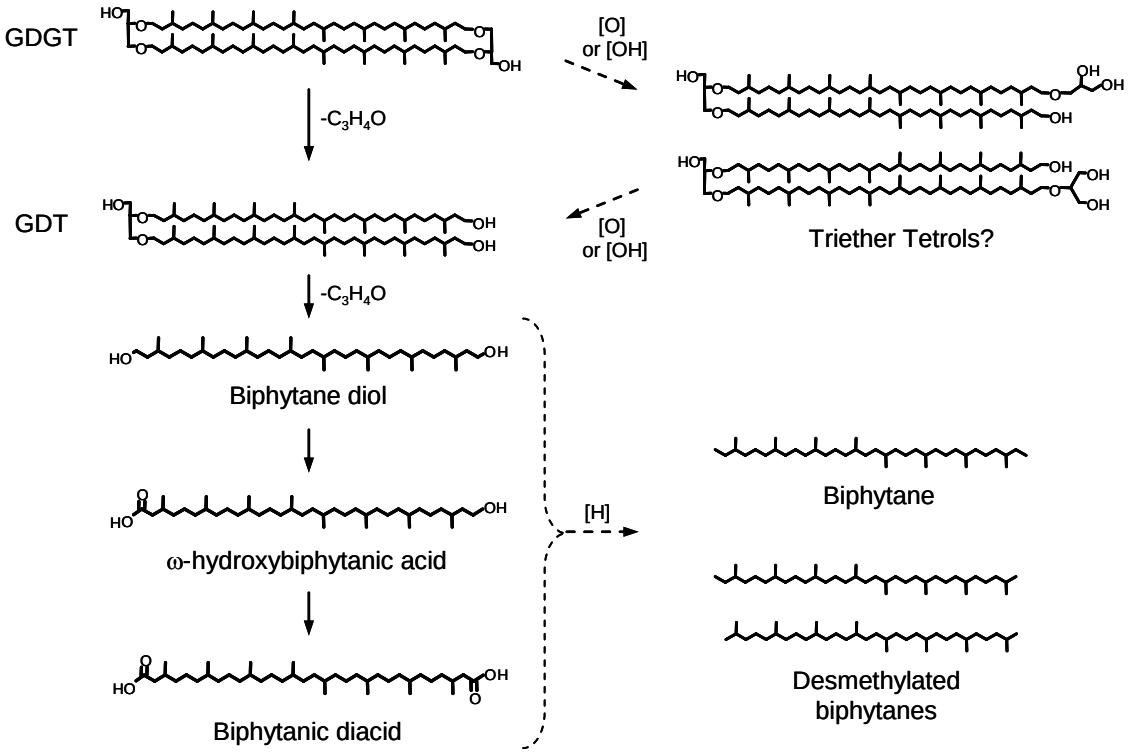


Scheme 5.1. Putative biosynthetic route to formation of GDT **58**. Although the GDD and geranylgeraniol precursors are shown to be both unsaturated and phosphorylated, the coupling could conceivably involve saturated or dephosphorylated molecules.

An alternative explanation for the presence of GDT and GMT lipids in the ancient sediments and in the soil sample is that they represent transformation products of tetraether lipids, formed *via* removal of a C_3H_4O unit from the end of a GDGT or GMGT structure, respectively (Scheme 5.2). This could potentially be effected by enzymatic modification of the tetraether lipids *in vivo,* used purposefully by the Archaea to modify the composition of the cellular membranes. Alternatively, the transformation could be extracellular, effected by ether-metabolising organisms, that have either grazed on the archaeal cells or actively/passively taken up liberated tetraether lipids following archaeal cell senescence. A substantial number of bacterial species are known to degrade xenobiotic ether-containing compounds (White *et al.*, 1996), although glycerol-bound ether lipids are not among the substrates currently known to be degraded by microorganisms. In the dogfish *Squalus acanthius,* glycerol alkyl monoethers were observed to be metabolically degraded *in vivo* by oxidative cleavage of the ether bond (Malins, 1968). Likewise, *in vitro* studies utilising enzymes from the livers of albino rats showed similar degradation of glycerol monoethers (Tietz *et al.*, 1964). Thus, enzymatic cleavage of glycerol ethers does occur in some higher organisms. In simulation studies to model the degradation of ether lipids in natural environments, an IPL based on a GDD **1** core was incubated under aerobic or anaerobic conditions in the presence of active sediments from a variety of environments (Harvey *et al.*, 1986). Whereas the lipid was rapidly (i.e. in a matter of days) converted to CO_2 in each of the standard experiments, degradation was completely inhibited if formaldehyde was added to the incubation flask or the flask was autoclaved prior to lipid addition. This led the authors to suggest that degradation of ether lipids in the natural environment is more likely a biotic as opposed to an abiotic process (Harvey *et al.*, 1986). Another relevant, compound-specific study has shown that GDGT lipid abundances reduce during passage through the digestional tracts of marine decapods, with higher levels found in the stomachs of the organisms than in their intestines (Huguet *et al.*, 2006a). Since no GDGT lipids were identified in either the livers or the gonads of the organisms, this drop is significant and might indicate a metabolic breakdown of the tetraethers. As no potential transformation products formed from the tetraether lipids were identified during the studies, it is impossible to assess whether GDT lipids may have accounted for some or all of the mass deficit.

GDT and GMT lipids in the aquatic sediments may, alternatively, have been generated by chemical modification of tetraether lipids whilst they reside in the water column or within the sediment itself. In particular, long-term oxic degradation has been implicated as a possible explanation for the reduction of GDGT concentrations in the oxidised sections compared to the anoxic sections of turbidites from the Madeira abyssal plain (Huguet *et al.*, 2008; 2009). In addition, sediment cores from two proximate locations (50 km apart) in the Arabian Sea show closer relative GDGT concentrations between horizons of the same age when those horizons were deposited under the same oxidative conditions (i.e. oxic, sub-oxic or anoxic) compared with horizons which were deposited under differing conditions at the two sites (Damsté *et al.*, 2002a; Schouten *et al.*, 2004). Another study, in which anoxic sediments from the Namibian upwelling region were relocated to various globally-distributed waters and stored for one year, showed that there was almost no difference in the concentrations of extractable GDGTs after this period, irrespective of the oxic status of the sites or whether indigenous microbes were allowed to admix with the samples throughout the year (Kim *et al.*, 2009c). This would indicate that oxic degradation of GDGT lipids is not a short-term process and implies that the triols in VDG and Oxford Clay sediments, if formed *via* this route, are unlikely to have been modified in the water column during early diagenesis. Prolonged oxic degradation could also account for the presence of GDT **58** in the soil extract (see Fig 5.4). The presence of GDT and GMT lipids in extracts from archaeal cellular material would appear to support a biogenic origin for these components. Given the harsher extraction procedures (i.e. hot acidic methanolysis) used to extract tetraether lipid cores from the culture samples and to degrade intact polar lipids in the soil in comparison to that used to obtain the lipids from the aquatic sediments (i.e. sonication in organic solvents at RT), the possibility that GDT and GMT lipids observed in the cultures and/or soil may be artifacts cannot be ruled out. Similarly, oxic degradation of lipid extracts from both archaeal cultures and sediments during their storage cannot be completely ruled out, although it is perhaps telling that the sample exhibiting the greatest relative ratios of GDT:GDGT and GMT:GMGT lipids, the Oxford Clay S90-11 shale, was that which had been stored at the lowest temperature.

Thermal degradation of GDGT lipids has been observed during hydrous pyrolysis of VDG sediment at temperatures of up to 330°C, in which both extractable residues and material assumed to be kerogen-bound were destroyed at 300°C (Schouten *et al.*, 2004). Similarly, flash and hydrous pyrolysis of GDD and GDGT lipids in other studies generated a series of alcoholic, carbonylic, acidic and/or aliphatic hydrocarbon degradation products formed *via* ether cleavage and additional decomposition of the alkyl chains within the structures (Rowland, 1990; Navale, 1992; Pease *et al.*, 1998). GDT and GMT lipids could potentially be formed *via* similar thermal cracking of the C-O ether bonds in GDGT and GMGT lipids, respectively. This is, however, a highly unlikely origin for the triols identified in the environmental samples investigated in the present study given the contemporary age of the greenhouse soil and the thermal immaturity of both the VDG (Damsté *et al.*, 1995) and Oxford Clay (Kenig *et al.*, 2004 and references therein) sediments.



Scheme 5.2. Putative dissociation pathway of GDGT lipids to their proposed transformation products, exemplified using structures which contain no Cp or Ch rings.

A series of components (**112-112'** and **113-113'**) identified in the Oxford Clay sediment extract exhibit MS spectra indicative of structures representing oxidised/hydrated variants of GDGT lipids **2** or **24-24''** found in the same sample. The MS/MS data obtained for **112-112'** and **113-113'** are also consistent with structures in which an additional molecule of H_2O has been added to the GDGT cores. Consequently, **112-112'** and **113-113'** were tentatively assigned as triether tetrol lipids, although further work is required to confirm both the suggested structure and the nature of the isomerism observed. Should the preliminary assignments prove to be correct, the tetrols would almost certainly represent intermediate structures in oxidative or hydrolytic breakdown of GDGT to GDT lipids whereby one, but not both, of the ether linkages joining the biphytanyl chains to an individual glycerol group has been cleaved (Scheme 5.2).

Biphytane diols and biphytanic diacids containing 0-3 rings (i.e. up to two Cp rings and one Ch ring) have been found, albeit not necessarily co-occurring, in a variety of environmental samples. These include ancient (Schouten *et al.*, 1998; Kuypers *et al.*, 2002; Pancost *et al.*, 2008) and recent (Hoefs *et al.*, 1997; Schouten *et al.*, 1998; Saito and Suzuki, 2010) sediments from marine environments, including mud domes (Pancost *et al.*, 2000; 2001b; Stadnitskaia *et al.*, 2003), carbonates (Kinnaman *et al.*, 2010), tubular concretions (De Boever *et al.*, 2009) and limestones deposited or produced in methane seep environments (Birgel *et al.*, 2008a; 2008b) and sediments from hydrothermally-influenced environments (Teske *et al.*, 2002; Schouten *et al.*, 2003a). On account of the presence of the isoprenoid chains in the structures, the diols and diacids have been suggested as possible transformation products of GDGT lipids (Schouten *et al.*, 1998; 2003a; Pancost *et al.*, 2006). This assertion seems to be supported by the identification of both biphytanic diacids and ω-hydroxybiphytanic acids containing 0-3 rings in silica sinters from New Zealand hot springs (Pancost *et al.*, 2006). In the sinter extracts, the distributions of both diacids and hydroxy-acids were similar with respect to the number of incorporated rings, both to one-another and to the distribution of similarly-cyclised biphytanes released following ether cleavage of GDGT lipids found in the same samples (Pancost *et al.*, 2006). Furthermore, $\delta^{13}C$ values of the biphytanic diacids in sediments from the Guaymas Basin approximately match those for biphytanes containing the same number of

rings released following ether cleavage of tetraether lipids found in the same samples (Schouten *et al.*, 2003a). Again, the distributions of the diacids were similar with respect to the number of rings to the liberated biphytanes, consistent with either derivation of the diacids from the GDGT precursors or, at the very least, an origin from the same set of source organisms. The presence of biphytane diols in surface sediments, on the other hand, has led to the alternative suggestion that the diols are more likely to be biogenic (Schouten *et al.*, 1998). Furthermore, the distributions of biphytane diols found in several surface and sub-surface sediments have been shown to be different to those of cyclised biphytanyl chains liberated following ether cleavage of intact polar tetraether lipids or tetraether lipid cores present in the same samples (Schouten *et al.*, 1998; Saito and Suzuki, 2010), suggesting that the biphytane diols are not derived directly from the tetraether lipids. In some sediments, the $\delta^{13}C$ values of the biphytanic acids were found to be different to those of the ether-cleaved biphytanes (Birgel *et al.*, 2008b), further suggesting that the diacids are unlikely to have been formed *via* transformation of the tetraether lipids in these samples. As such, the biphytanic acids may also be biogenic, perhaps being synthesised by unknown archaea (Schouten *et al.*, 1998; Birgel *et al.*, 2008b). Irrespective of whether biphytane diols, ω-OH biphytanic acids and biphytanic diacids are actively produced by the Archaea or are formed following the death of the progenitor archaeal cells, GDT lipids may represent intermediates in the degradative transformation of GDGT lipids to these compounds. For example, cleavage of one or both of the ether bonds to the remaining glycerol group in the GDT would yield one or two molecules of biphytane diol, respectively (Scheme 5.2). Subsequent mono- or bis-oxidation of the liberated biphytane diols could potentially yield the hydroxyacids and diacids, respectively. Further defunctionalisation of either biphytane diols and/or biphytanic acids during thermal maturation could result in formation of isoprenoid hydrocarbons or their desmethylated counterparts, as are often found in extremely antiquated sediments (e.g. from the Late Archaean; Ventura *et al.*, 2007; 2008) and crude oils (e.g. Moldowan and Seifert, 1979). Notably, the proposed pathways for degradation of GDGT lipids are in accordance with a similar pathway proposed for the degradation of GDD **1** to phytanol, phytanic acid and phytane in Dead Sea sediments (Nissenbaum *et al.*, 1972; Anderson *et al.*, 1977).

Degradation of GMGT lipids to GMT lipids, followed by further ether cleavage and oxidation may account for the isoprenoid tetraacids found in naphthenate deposits (Lutnaes *et al.*, 2006; 2007; Smith *et al.*, 2007). This would necessarily require the covalent link between isoprenoid chains in the GMGT lipids to be between carbons C-15 and C-15' (see Section 4.2.2.3 for further discussion). It is important to note that key intermediates in this process, namely C_{80} triether tetrols, tetrols and mono- or polyhydroxy acids, have yet to be identified in environmental samples. Furthermore, the only tetraacid for which potential GMGT and GMT precursors have been identified is a structure containing four rings (Lutnaes *et al.*, 2006; 2007; Smith *et al.*, 2007), which may be a product of degradation of **94**, itself formed from **31**.

GDGT lipids **14-16** and **24''** are used in the TEX_{86} index (Section 1.6.1), whereas lipid **24** is used in the BIT index (Section 1.4.2). The potential transformation of these components to GDT lipids **86**, **75-76** and **106** could have serious implications for the use of the two indices in palaeoreconstructions, particularly if some of the putative GDGT precursors are more readily transformed than others. Huguet *et al.* (2008) have shown that preferential degradation of **24** over non-isoprenoid GDGTs in oxidised sections of Madeira abyssal plain turbidites can lead to a substantial increase in the measured BIT value. Similarly, degradation of **14-16** and **24''** in the oxidised sections led to TEX_{86}-associated temperatures that differed by up to $6^{o}C$ from those determined from sediments in the unoxidised sections (Huguet *et al.*, 2009). This probably reflects preferential preservation of a matrix-protected subset of the these lipids derived from continental soil which carry a different temperature signal to those produced in the water column above the neighbouring west African shelf (Huguet *et al.*, 2009). Nevertheless, this example illustrates the inherent disruption that GDGT degradation can have on TEX_{86}–based reconstructions. Work is currently underway to investigate the relative distributions of GDT and GDGT lipids in the VDG and Oxford Clay sediments in order to investigate whether GDT formation may compromise calculated TEX_{86} or BIT values in these samples.

5.2.4. Isoprenoid lipid homologues

In addition to the isoprenoid lipid structures, a number of novel lipids which represent higher homologues of lipids **2**, **3**, **14**-**15** and **27** were also detected in the base peak chromatograms and in selected reconstructed ion chromatograms for several of the extracts. Some of the higher homologues were selected as precursors for CID in analysis of at least one sample, allowing structural information to be obtained. Others required re-analysis utilising a narrower, more appropriate *m/z* scan range of a sample in which the lipid in question was determined to be sufficiently abundant, thus allowing selection of the component for CID.

5.2.4.1. $GDGT_M$ lipids

During LC-MS/MS analysis of **2** in the soil extract, the base peak ion at *m/z* 1301.8 in the mass spectrum was accompanied by lower intensity peaks for components with $[M+H]^+$ at approximately 14 (*m/z* 1315.8; **26**), 28 (*m/z* 1329.8; **66**) and 42 (*m/z* 1241.7; **114**) *m/z* units higher. Additionally, appropriate reconstructed ion chromatograms showed that, although the four lipids partially co-elute, marginally earlier retention times are observed as *m/z* ratio increases (Fig. 5.11). The $[M+H]^+$ of **26** and **66** both accounted for the base peak ion in several scans in the LC-MS/MS analysis, allowing CID of each to be performed. The $[M+H]^+$ of **114**, on the other hand, was never the base peak and, consequently, an MS/MS spectrum was not recorded during the analysis. All product ions identified in the MS/MS spectrum of **26** matched those observed following CID of the $[M+H]^+$ of homocaldarchaeol, (Chapters 2 and 3), supporting the assignment of the lipid to this structure. Component **26** was also identified in ion chromatograms for the VDG extracts (see Fig. 5.13), although re-analysis of the XI-1b horizon extract by LC-MS/MS using a scan range of *m/z* 1310-1350 was required to obtain an MS/MS spectrum in this case**.** The MS/MS spectrum of **66** shows prominent product ions formed from losses of water, C_3H_4O, $[40_0]^0$ (*m/z* 771) and $[41_0]^0$ (*m/z* 757). A combination of $[40_0]$ and $[41_0]$ hydrocarbon chains in an individual GDGT lipid is inconsistent with the *m/z* ratio observed for the protonated molecule. Interestingly, a weak ion at *m/z* 743 and a more prominent ion at *m/z* 725 were also observed in the MS/MS spectrum and may

reflect neutral losses of $[42_0]^0$ and $[42_0]^1$, respectively. Consequently, **66** may comprise a mixture of two isomers, GDGT[41_0,41_0] (designated **66**) and GDGT[40_0,42_0] (designated **66***). A similar mixture of isomers was tentatively suggested for lipids with the same $[M+H]^+$ and relative retention times identified in a lipid extract from MTH(ΔH) (see Section 3.2.3.3). Further work will be required before the presence and/or structure of the putative isomers can be confirmed. A lipid consistent with **66** was also detected in ion chromatograms for the extracts from individual VDG horizons (see Fig. 5.13), although an attempt to verify the assignment *via* LC-MS/MS analysis of the XI-1b horizon extract using a scan range of *m/z* 1320-1350 yielded an insufficiently strong MS/MS spectrum for reliable interpretation to be made. Although no MS/MS data was obtained for lipid **114**, it is probable that this lipid represents a higher homologue of **2** containing two additional methyl groups on one biphytanyl chain and one additional methyl group on the other chain.

Lipids **82** (*m/z* 1313.7) and **115** (*m/z* 1327.8) were observed to co-elute with GDGT **14** in ion chromatograms for the soil extract (Fig. 5.11), with each exhibiting a retention time marginally less than that of **14**. Similarly, lipids **83** (*m/z* 1311.7) and **116** (*m/z* 1325.8) were observed to co-elute with GDGT **15** in the same chromatograms, again exhibiting retention times slightly less than that of **15**. Although no MS/MS spectra were obtained for these components in the soil extract, the retention times of **82** and **83** relative to those of **14** and **15** and the 14 *m/z* unit increase in the $[M+H]^+$ of each relative to their co-eluting GDGT counterpart are consistent with those observed for $GDGT_M$ lipids containing one or two Cp rings in a species of *Pyrobaculum* (Section 4.2.3.1). As such, **82** and **83** are tentatively assigned as the C-13 methylated homologues of **14** and **15**, respectively. Lipids **115** and **116** may be higher homologues of **14** and **15** which contain two more carbon atoms than the corresponding GDGT lipids.

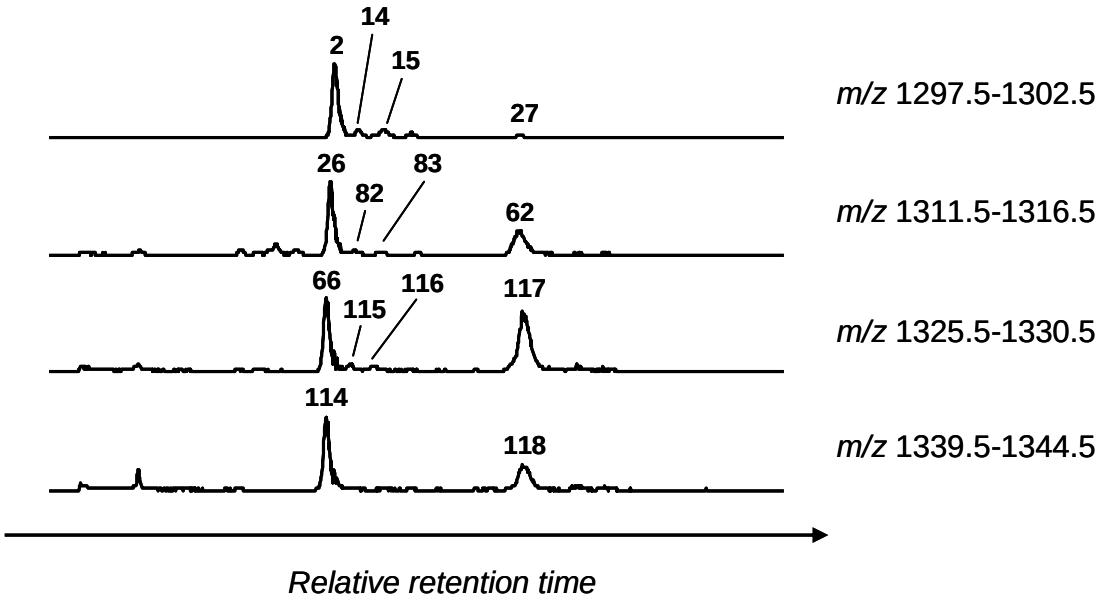


Fig. 5.11. Partial HCT LC-MS ion chromatograms for the lipid extract from a greenhouse soil, highlighting the presence of isoprenoid lipid homologues including $GDGT_M$ and $GMGT_M$ lipids.

5.2.4.2. $GMGT_M$ lipids

Ion chromatograms for the soil extract also indicated the presence of three higher mass species, **62** (*m/z* 1313.8), **117** (*m/z* 1327.8) and **118** (*m/z* 1341.8), which co-elute with **27** during LC-MS/MS (Fig. 5.11). Components **62** and **117** were also identified in ion chromatograms for extracts from the VDG horizons. The retention time of **62** is almost identical to that of **27**, whereas **117** and **118** have retention times which are slightly greater. Lipid **62** was selected for CID during LC-MS/MS of the extracts of the soil and the VDG IV-1b sediment and lipid **117** was selected for dissociation only during analysis of the soil extract. No MS/MS spectrum for lipid **118** was obtained during analysis of any of the extracts. The MS/MS spectrum (Table 5.1) and retention time of **62** are consistent with those observed during LC-MS/MS of H-shaped homocaldarchaeol in a lipid extract from MTH(ΔH) (see Section 3.2.3.3) allowing assignment to this structure. The MS/MS spectrum of **117** (Fig. 5.12 and Table 5.1) contains product ions each shifted by +28 *m/z* units from those observed in the spectrum for **27** and by +14 *m/z* units from those observed in the spectrum of **62**. Consequently, this lipid is likely to be a $GMGT_M$ lipid which

contains two carbons more than **27**, probably located on the C-13 chain positions of the isoprenoid hydrocarbon. The assignments made are supported by the structures of C_{81} and C_{82} Cp ring-containing naphthenic tetraacids with respect to their C_{80} analogues; rigorous structural characterisation showed that they are similarly methylated at C-13 chain positions (Lutnaes *et al.*, 2006; 2007; Smith *et al.*, 2007). Lipid **118** is probably a higher homologue of GMGT **27** which contains a C_{83} hydrocarbon.

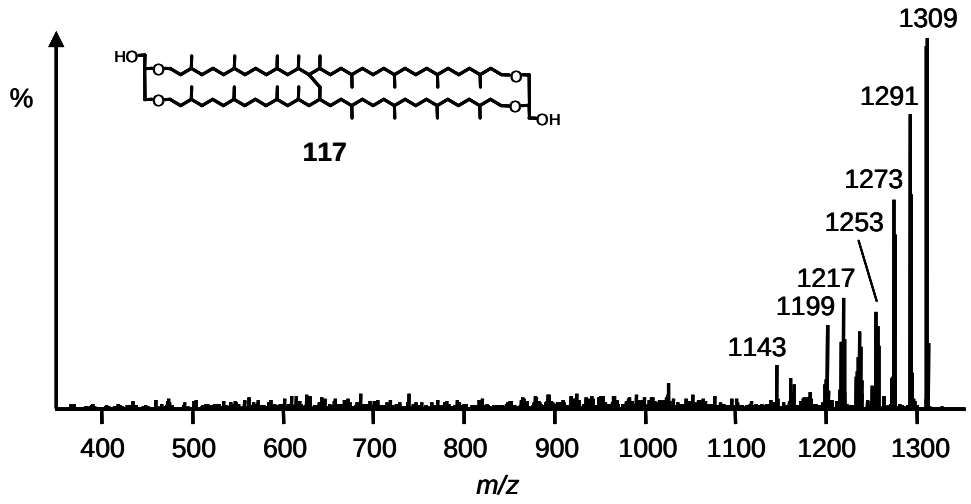


Fig. 5.12. MS/MS spectrum (HCT) of $GMGT_M$ **117** ($[M+H]^+$ *m/z* = 1327.8). As the position of the covalent link between the isoprenoid chains and placement of the additional methyl group are not known, the structure shown for **117** is based upon isoprenoid lipids containing C_{80} and C_{82} hydrocarbons reported previously (Morii *et al*., 1998; Lutnaes *et al*., 2007), but remains tentative nevertheless.

5.2.4.3. Absence of $GTGT_M$ and GDD_M lipids

No alkyl chain-derived homologues of GTGT **3** were identified in any of the extracts investigated. Analyses of the extracts from VDG horizons X-1 and VII-1b using a scan range of *m/z* 50-900 indicated the identity of an early eluting component as GDD **1** on the basis of its retention time (see Fig. 5.31), $[M+H]^+$ (*m/z* 653.6) and from a product ion at *m/z* 373 in the MS/MS spectrum (see Section 3.2.3.3). A corresponding chain methylated structural variant at nominal *m/z* 667 was not identified in either sediment.

5.2.4.4. Glycerol di- and trialkyl homoglycerol lipids

Additional lipid components exhibiting $[M+H]^+$ at *m/z* 1316.4 (**119**) and 1318.4 (**120**) were observed in the base peak chromatograms for several of the VDG horizon extracts (e.g. horizon XI-1b; Fig. 5.13). CID of the $[M+H]^+$ of **119** generated product ions in two regions of the MS/MS spectra, indicating a dialkyl-type tetraether structure (Fig. 5.14a). The $[M+H]^+$ of lipid **120**, on the other hand, produced product ions in three regions of the spectrum, confirming a trialkyl core-type structure (Fig. 5.14b). The MS/MS spectrum of **119** shows product ions formed from losses of water (-18 Da; *m/z* 1297), $C_3H_6O_2$ (-74 Da; *m/z* 1241) and $[40_0]^0$ (-558 Da; *m/z* 757), losses also observed for **2**, and abundant product ions at *m/z* 1245 and 1227, arising from unprecedented losses of 70 (C_4H_6O) and 88 Da (C_4H_6O + H_2O) from the precursor, respectively. Fragment ions formed from neutral losses of water (*m/z* 1297) or 88 Da (*m/z* 1227) from the $[M+H]^+$ are also observed in the MS spectrum of **119**. Interestingly, product ions formed from the 70 and 88 Da losses in MS/MS (highlighted in bold in Fig. 5.14a) have *m/z* values identical to those formed after loss of C_3H_4O and/or a molecule of water from the $[M+H]^+$ of **2**. This suggests not only that **2** and **119** represent an homologous pair of lipids, but that the additional carbon atom in **119** is located somewhere on one of the glycerol head groups. CID of the $[M+H]^+$ of **120** generated product ions formed *via* losses of an individual molecule of phytene (-280 Da; *m/z* 1037) or loss of this moiety with concerted loss of 1-2 molecules of water (-298 Da; *m/z* 1019 and -316 Da; *m/z* 1001), a 74 Da molecule (-354 Da; *m/z* 963) or an 88 Da molecule (-368 Da; *m/z* 949). Other product ions attributable to concerted losses of C_3H_4O, a 70 Da molecule, a second molecule of phytene or of biphytadiene from the $[M+H]^+$ were also observed. Notably, all product ions for which loss of 70 Da or 88 Da molecules are implicated (highlighted in bold in Fig. 5.14b) have identical *m/z* values to ions formed from the $[M+H]^+$ of **3** during CID. As such, **120** most likely represents a glycerol-modified homologue of **3**. To distinguish the lipids containing the unusual modified glycerol moiety, they are henceforth referred to as homoglycerol tetraether lipids (dialkyl = GDHT; trialkyl = GTHT).

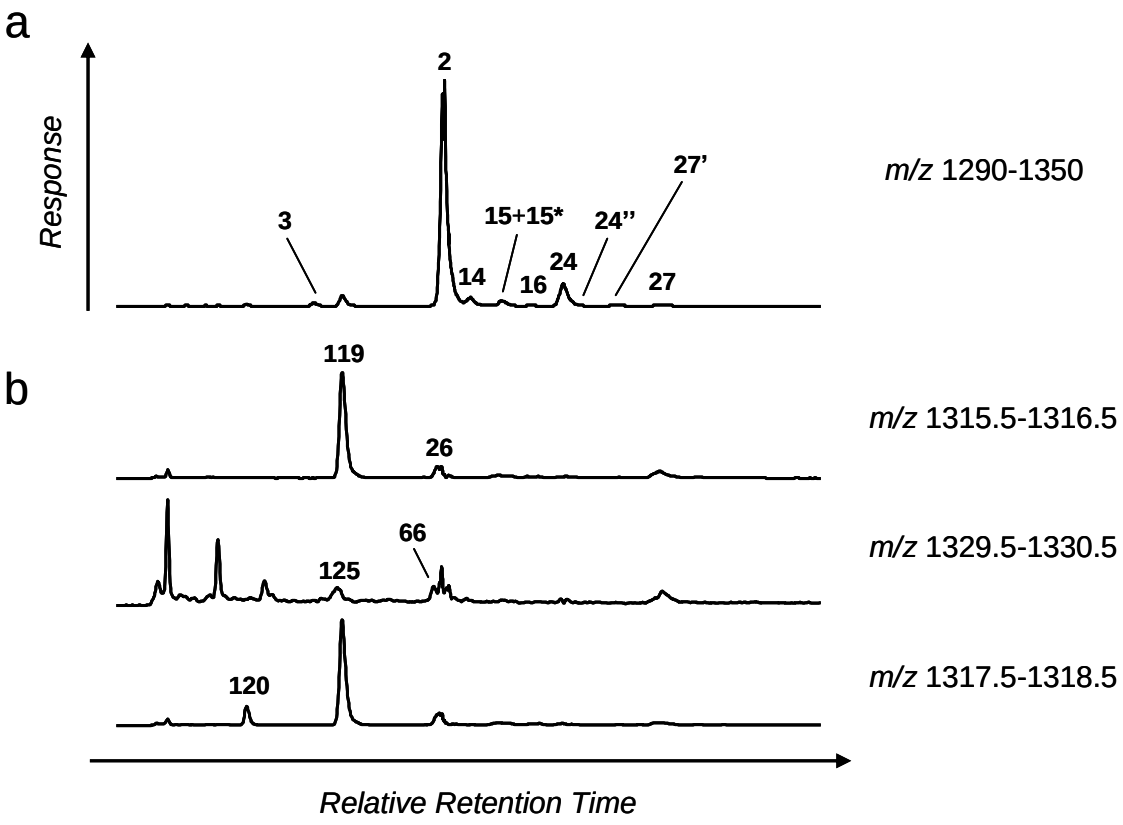


Fig. 5.13. Partial HCT LC-MS ion chromatograms for the lipid extract from VDG horizon VII-1b: a) ion chromatogram selecting for isoprenoid lipids and higher homologues; b) ion chromatograms selecting for specific isoprenoid higher homologues.

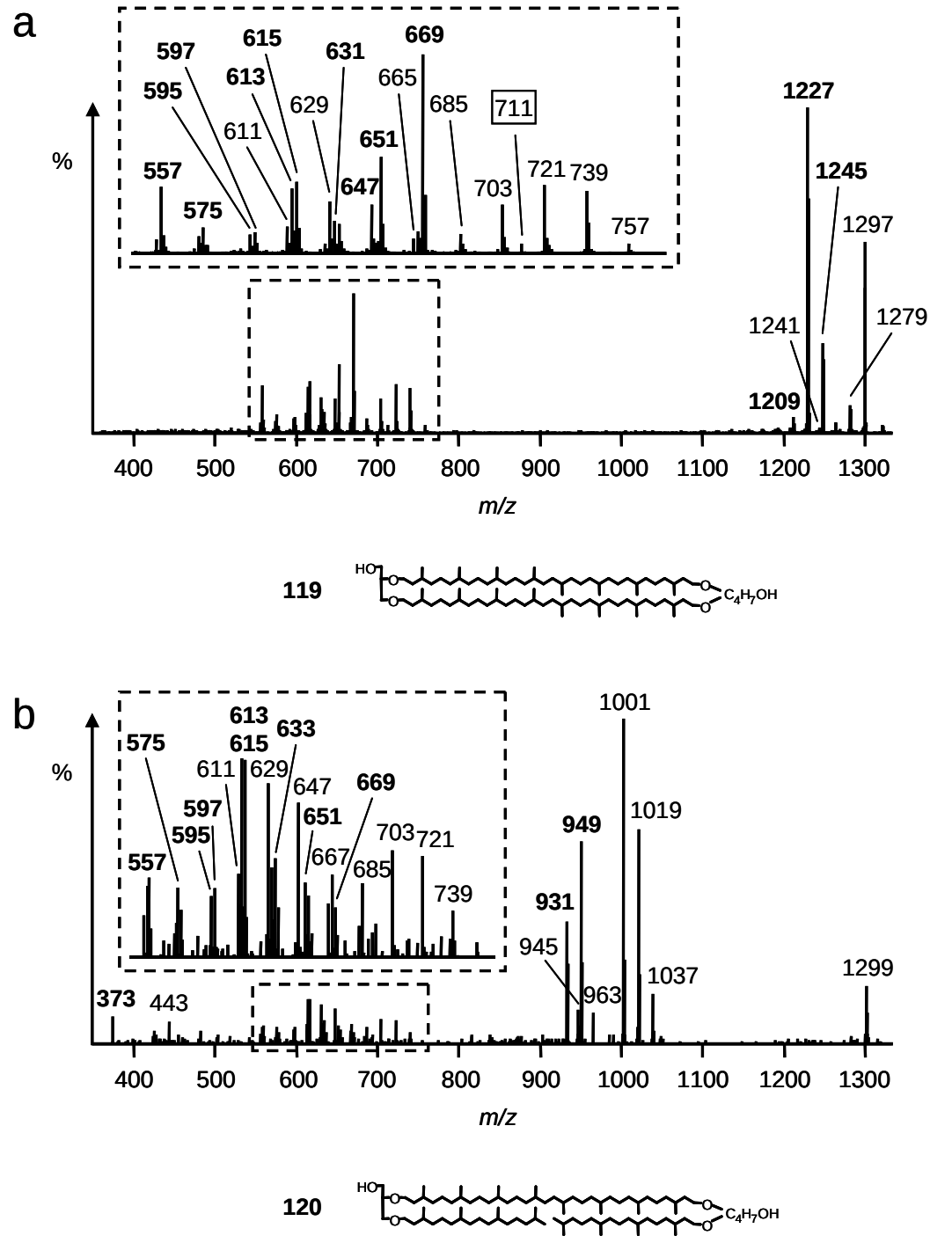


Fig. 5.14. MS/MS spectra (HCT) of: a) GDHT **119** ($[M+H]^+$ *m/z* = 1316.4); b) GTHT **120** ($[M+H]^+$ *m/z* = 1318.4). Insets show expansions of regions of the spectra. Ions in bold face were observed in the MS/MS spectra of GDGT **2** (for a) or GTGT **3** (for b), whereas other ions are shifted +14 *m/z* units from ions in these spectra. The ion at *m/z* 711 in a) cannot be assigned into either of these categories. The nature of the homoglycerol group shown in the representative structures is unknown.

It is important to note that another possible explanation for the 14 Da variation in mass between the GDGT/GTGT and GDHT/GTHT lipids could be addition of a carbonylic oxygen to a carbon contained within or adjacent to one of the glycerol termini in **2** and **3**. This possibility is unlikely, as such a modification would increase the hydrophilicity of the lipids and should lead to later elution times than for GDGT lipids, an effect observed in previous normal phase LC-MS analyses of ether and mixed ether/ester lipids in the eubacterial order *Thermotogales* (Damsté *et al.*, 2007). Since **119** and **120** elute significantly earlier than **2** and **3**, respectively, the modification distinguishing GDGT and GDHT lipids must increase the overall hydrophobicity of the molecules substantially. Addition of an extra carbon to the polar glycerol group, which is the functional group that undergoes the most significant interactions with the modified silanol groups of the normal phase column during the LC-MS separation, is consistent with this idea. Furthermore, the APCI MS spectra of mixed ether/ester lipids from the *Thermotogales* gave $[M+H-18]^+$ peaks as the base peak ions, whereas the base peak ion for tetraether lipids was the $[M+H]^+$ (Damsté *et al.*, 2007). These observations also support a lack of an ester functionality in **119**-**120**, which definitively exhibit a clear $[M+H]^+$ in MS. In order to provide further evidence that the 14 Da difference relates to an additional CH_2 group as opposed to an additional oxygen atom, **119** was isolated preparatively *via* sequential collections during several consecutive LC-MS/MS analyses of the VDG VII-1b sediment extract. The enriched lipid fraction was tested for purity *via* direct infusion into the LCQ mass spectrometer, whereby **119** (*m/z* 1316.1) was found to be the dominant tetraether core component (Fig. 5.15). Subsequent MALDI Qh-FT-ICR MS of the isolate gave a mass spectrum containing a variety of peaks, including a peak of low abundance at *m/z* 1338.3181 (Fig. 5.16) and peaks for three higher isotopomers of this species. A simulation of the monoisotopic peak expected for a species with ionic composition $C_{87}H_{174}O_6Na^+$ (predicted accurate *m/z* = 1338.3203) shows a close match to the ion at *m/z* 1338.3181, with an *m/z* error of only 1.6 ppm (i.e. dashed line in Fig. 5.16). Simulation of the peak expected for $C_{86}H_{170}O_7Na^+$ (predicted accurate *m/z* = 1328.2839), on the other hand, shows a poorer match, with an *m/z* error greater than 25 ppm (i.e. dotted line in Fig. 5.16). Similarly, the peaks predicted for higher isotopomers of $C_{87}H_{174}O_6Na^+$ show much closer matches to the experimental data measured for the higher isotopomers of the ion at *m/z* 1338.3181

than do the peaks predicted for the higher isotopomers of $C_{86}H_{170}O_7Na^+$. The FT-ICR MS signal strength for the ion at *m/z* 1338.3181 was low, making conclusive assignment of its elemental composition difficult. Nevertheless, it is encouraging that an ion consistent with a lipid component of molecular formula $C_{87}H_{174}O_6$ (i.e. a higher homologue of **2**) may be present in the fraction enriched with GDHT **119**, whereas an ion consistent with a lipid of formula $C_{86}H_{170}O_7$ (i.e. a derivative of **2** containing an additional carbonylic oxygen) could not be identified.

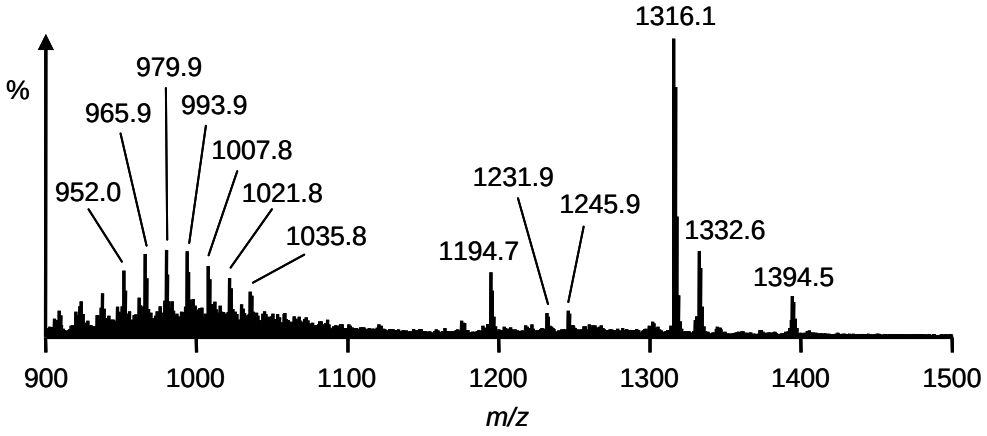


Fig. 5.15. Direct infusion MS (LCQ) spectrum of a fraction enriched in GDHT lipid **119** ($[M+H]^+$ *m/z* = 1316.1), prepared from the lipid core fraction extracted from VDG VII-1b sediment.

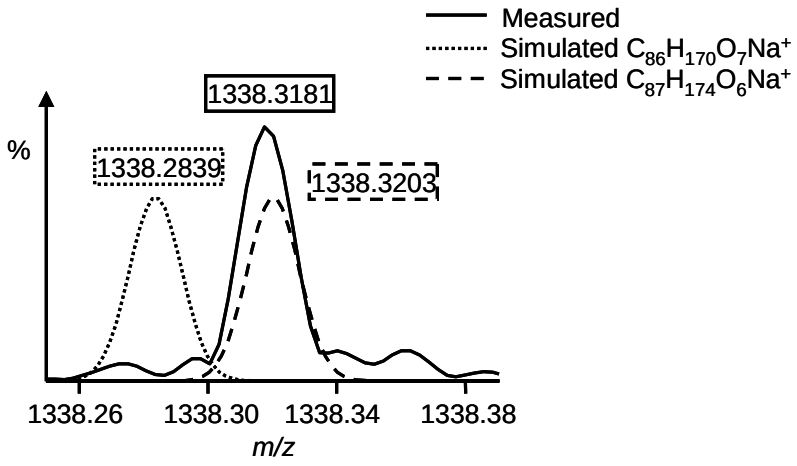


Fig. 5.16. Monoistopic peak of a component observed in the Qh-FT-ICR mass spectrum of an enriched fraction of GDHT **119**, prepared from the lipid core fraction extracted from VDG VII-1b sediment. Simulated peaks for $C_{86}H_{170}O_7Na^+$ and $C_{87}H_{174}O_6Na^+$ also shown.

From consideration of the retention time of **119** relative to **2,** O-methylation on glycerol can be ruled out as a possible source of the additional carbon in GDHT and GTHT lipids: lipid **51**, a glycerol *sn*-1 methyl ether of **2** observed in other LC-MS/MS analyses elutes significantly earlier than **119** (see Section 2.2.4). Furthermore, the methyl ether dissociates to lose methanol (-32 Da) during CID, a neutral loss that is not observed during dissociation of **119**. These observations imply that the unique head group in **119** and **120** contains an additional saturated C-C bond compared with glycerol. Unfortunately, the precise nature and position of the homologation in **119** and **120** cannot easily be ascertained by LC-MS/MS, although a product ion at *m/z* 711, observed in the MS/MS spectrum of **119** (Fig. 5.14a), may potentially hold the key to elucidation of the position of the extra carbon within the homoglycerol group. This ion cannot be attributed to loss of any combination of molecules of water, C_3H_4O and biphytadiene from the $[M+H]^+$ of **119**. Consequently, it must arise from an unusual dissociation which is not accessible to the $[M+H]^+$ of GDGT **2** and hence, be provided by dissociation of the only structural variation, the homoglycerol group, in **119**. Apart from the product ions discussed already, no other ions formed from small molecule losses (<70 Da) were observed in the MS/MS spectrum of **119**, indicating that this unusual homoglycerol scission does not, in itself, lead to any mass loss from the precursor, but merely provides opportunity for loss of a larger 604 Da fragment. Given that other product ions observed in the MS/MS spectrum indicate the presence of biphytanyl chains in the structure, the 604 Da molecule may be 1-ethoxy-biphyt-31-ene, lending support for a homoglycerol structure which is based on glycerol but with the additional carbon attached to an *sn*-3 carbon. Unfortunately, other structures based on a biphytanyl chain could also account for the 604 Da loss observed and, consequently, further studies will be required before the nature of the homoglycerol group can be fully determined.

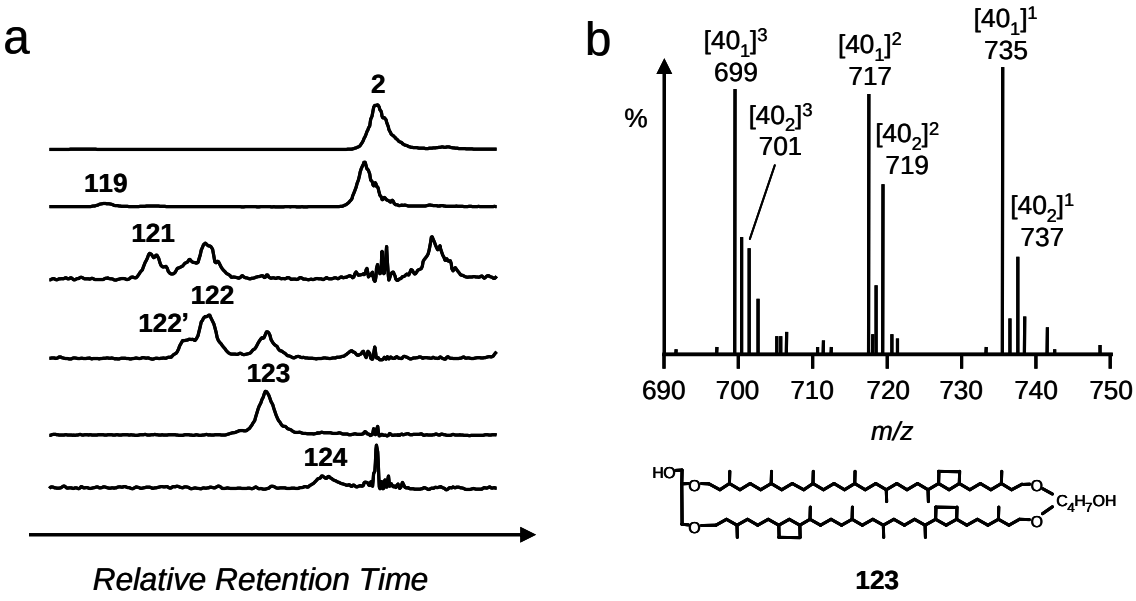


Fig. 5.17. a) Partial HCT LC-MS ion chromatograms for a lipid extract from a greenhouse soil, highlighting the presence of GDHT lipids. A chromatogram for GDGT **2** is provided for comparison; b) partial MS/MS spectrum (HCT) of GDHT **123** ($[M+H]^+$ *m/z* = 1309.8).

GDHT **119** (*m/z* 1315.8) was also detected in the greenhouse soil, along with a series of components, **121** (*m/z* 1313.8), **122'** (*m/z* 1311.7), **122** (*m/z* 1311.8), **123** (*m/z* 1309.8) and **124** (*m/z* 1307.7), which elute slightly later and exhibit $[M+H]^+$ at *m/z* values that are approximate multiples of 2 *m/z* units less than that of **119** (Fig. 5.17a). The MS/MS spectra of **121**-**123** and **122'** each show product ions formed from neutral losses of water and of 70 and 88 Da, confirming that each structure contains a homoglycerol group. The $[M+H]^+$ of lipid **124** was not selected as a precursor for CID during LC-MS/MS. The MS/MS spectrum of **123** (Fig. 5.17b) was of sufficient signal strength to allow the alkyl chains in the structure to be tentatively assigned from the product ions observed. Ions at *m/z* 737 and *m/z* 735 are most probably formed from loss of $[40_2]^1$ (-572 Da) and $[40_1]^1$ (-574 Da) respectively, indicating a GDHT[40_1,40_2] structure. Although the MS/MS spectra of **121**, **122** and **122'** were of insufficient signal strength for the alkyl chains in these structures to be similarly assigned, the differences in retention times and the values of the $[M+H]^+$ compared with those of **119** and **123** indicate that they most likely represent GDHT lipids containing one or two Cp rings. Similarly, **124** is probably a GDHT lipid

which contains four Cp rings. Lipids **119** and **121**-**122** were also identified in the Oxford Clay extract but, in this case, solely on the basis of retention time and the *m/z* value of the $[M+H]^+$ observed for each component.

5.2.4.5. Absence of homoglycerol diether and monoalkyl tetraether lipids

No components consistent with homoglycerol-containing variants of GMGT **27** were observed in any of the lipid extracts, either in the base peak chromatograms or in ion chromatograms selecting for *m/z* 1314 or 1328. Likewise, no homoglycerol-modified archaeol structures were observed in the analysis of the VII-1b extract using a scan range of *m/z* 50-900, despite this horizon showing the highest summed proportions of **119** and **120** of any sample.

5.2.4.6. Higher GDHT homologues

In some extracts, most notably in the VDG horizon III-1b extract, a lipid with a $[M+H]^+$ at *m/z* 1330.4 (**125**) was observed to elute immediately before **119** in the LC-MS base peak chromatograms (e.g. Fig. 5.18). Dissociation of the $[M+H]^+$ of **125** generated product ions in two regions of the MS/MS spectrum (Fig. 5.19), including ions formed from loss of 0-3 molecules of water (*m/z* 1311, 1293 and 1275), loss of 88 Da (*m/z* 1241) or from neutral loss of 84 Da with simultaneous losses of 0-2 molecules of water (*m/z* 1245, 1227 and 1209). Fragment ions formed from loss of water (*m/z* 1311) or 102 Da (*m/z* 1227) from the $[M+H]^+$ were also observed in the MS spectrum of **125**. Interestingly, the most favoured loss observed in both the MS and MS/MS spectra, that of 102 Da, generates a product with equivalent *m/z* value to that formed *via* loss of molecules of C_3H_4O and water from **2** and of C_4H_6O and water from **119**. Similarly, a product ion at *m/z* 669 in the spectrum of **125** could be attributable to neutral loss of 102 Da coupled with a simultaneous loss of 558 Da. This latter loss corresponds to expulsion of 1,31-biphytadiene, a molecule typically lost from the $[M+H]^+$ of **2** and **119** during CID. Other product ions generated from the $[M+H]^+$ of **125** are also consistent with loss of biphytadiene in conjunction with various small molecule losses (i.e. water, C_3H_4O, 84 Da fragment) discussed above and support a close structural relationship between

2, **119** and **125**. The 84 Da neutral loss is most likely elimination of C_5H_8O which, in turn, implies that **125** contains two carbon atoms more than **2**, both being located within the same modified glycerol group. As was observed following CID of the $[M+H]^+$ of **119**, a product ion at *m/z* 711 is present in the MS/MS spectrum of **125**. In the latter case, the ion was found in greater prominence than product ions of similar *m/z* value. The presence of this ion in the MS/MS spectra of both **119** and **125** suggests that they contain a common structural motif. This may indicate that **125** loses 1-propoxy-biphyt-31-ene (as opposed to 1-ethoxy-biphyt-31-ene proposed as the loss from **119**) during formation of the *m/z* 711 species. Accordingly, the tentative conclusion can be drawn that **125** represents a modification to structure **2** in which one of the *sn*-3 glycerol positions has been ethylated or dimethylated, an interpretation consistent with the elution order of **125**, **119** and **2** in LC-MS. Irrespective of the precise structure, **125** is designated as a GDH_2T lipid in order to clarify that one of the glycerol groups is homologated with two additional carbon atoms.

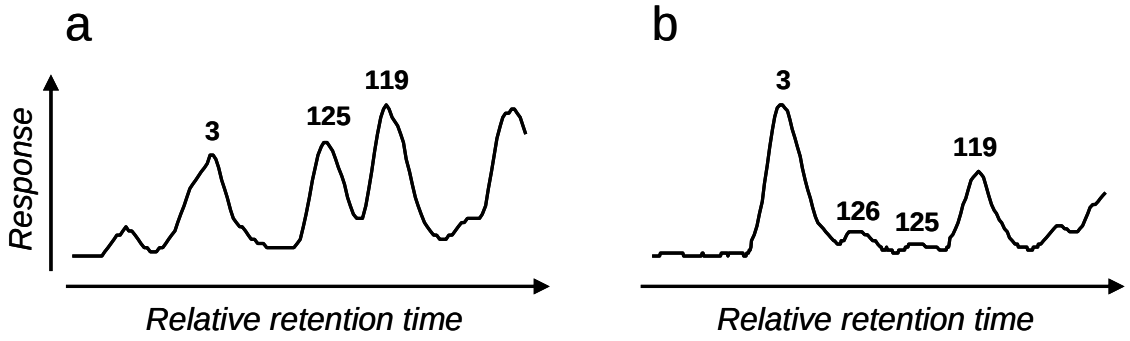


Fig. 5.18. Partial HCT LC-MS base peak chromatograms (*m/z* 900-1500) for the lipid extracts from: a) VDG III-1b; b) VDG IV-1.9 sediments, highlighting the elution of components **125** and **126**.

A second lipid with $[M+H]^+$ at *m/z* 1330.4 (**126**) elutes between **3** and **119** in the base peak chromatograms of some of the VDG lipid extracts, but is most obvious in the extract from horizon IV-1.9 (Fig. 5.18b). During analysis of this extract, a weak MS/MS spectrum was obtained for the lipid (not shown). The spectrum shows two regions of product ions, and includes ions generated from losses of 1-3 molecules of

water (*m/z* 1311, 1293 and 1275) and of a 70 Da molecule (C_4H_6O) plus 1-2 molecules of water (*m/z* 1241 and 1223). No ions consistent with loss of C_3H_4O or an 84 Da molecule (C_5H_8O) were observed, either with or without simultaneous water loss. Product ions formed from loss of neutral molecules that potentially contained one of the ether-bound alkyl chains were too weak to be assigned conclusively. The MS/MS data suggest that **126** does not contain an unmodified glycerol unit and, consequently, may have one extra carbon on each glycerol group than **2**. This assignment remains tentative and requires verification.

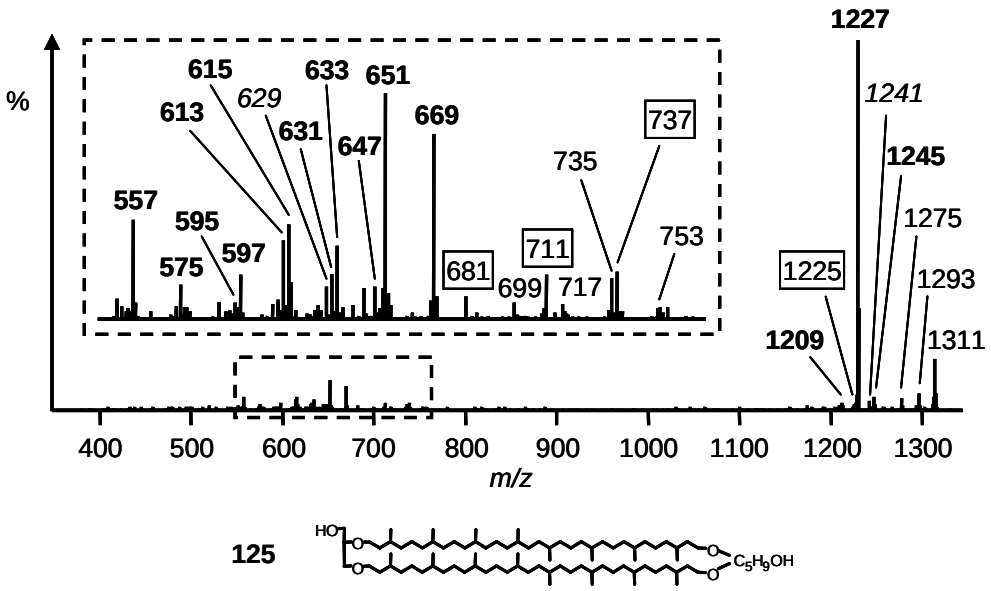


Fig. 5.19. MS/MS spectrum (HCT) of GDH_2T **125** ($[M+H]^+$ *m/z* = 1330.4) in the extract from VDG horizon III-1b. Inset shows expansion of a region of the spectrum. Ions in bold face were observed in the MS/MS spectrum of GDGT **2**. Ions in italic or in standard font are shifted +14 or +28 *m/z* units from ions in the spectrum of **2**, respectively. Ions enclosed in boxes cannot be assigned into any of these three categories.

5.2.4.7. Significance

The novel components of the extracts appear to relate to structures in which one, two or three extra carbon atoms have been added to previously reported tetraether lipids. Furthermore, the MS/MS spectra of some components indicate that these carbon atoms can be appendages either to the alkyl chains or glycerol moieties of known

isoprenoid tetraether structures. Methylation at a C-13 position of an isoprenoid chain at a late stage during biosynthesis was implicated as the source of the additional carbon in homocaldarchaeol (Section 3.2.3.6; Galliker *et al.*, 1998). As the position of methylation appears to remain saturated throughout biosynthesis of the lipid, it is apparent that highly unusual C-methyltransferase enzymes may be utilised by some archaea to invoke structural changes to the simple core tetraether lipids initially formed during biosynthesis. Lipids **26** and **62** in the VDG extracts show retention times and MS/MS data consistent with those observed during analyses of MTH(ΔH), suggesting a similar methylation at the C-13 chain position. It is noteworthy that other C_{41} hydrocarbons which appear to be homologous structural variants of biphytane have also been identified in samples from environmental settings. A 3,7,11,11,15,18,22,26,30-nonamethyldotriacontanyl chain (Fig. 5.20a) was liberated following HI-$LiAlH_4$ treatment of a polar lipid fraction isolated from an iron sulfide (Blumenberg *et al.*, 2007). Although a GMD containing this hydrocarbon was also identified and was probably the main source of the alkyl chain, an additional GDGT source could not be discounted. A 3,7,11,15,18,22,26,30-octamethyltritriacontanyl chain (Fig. 5.20b) has also been identified as a free hydrocarbon in Late Archaean sediments (Ventura *et al.*, 2007; 2008). In this case, it was suggested that the hydrocarbon was derived as a product of the thermal cracking of the C_{80} hydrocarbon found in GMGT **27** (Ventura *et al.*, 2008). Since substantial isomerisation of C_{40} biphytanes was noted in the same samples, it is also possible that the C_{41} chain found therein represents a rearrangement of an alternative chain containing the same number of carbon atoms. As such, the assignment of **26** and **62** from the environmental extracts as C-13 methylated structures will require further supporting evidence, perhaps *via* direct analysis of the ether-bound hydrocarbon chains found within these structures. Structures **66**, **114**, **117** and **118** are most likely further modifications to structures **2** and **27** in which a second and, subsequently, third methyl group has been transferred to an isoprenoid hydrocarbon. The elution of **117** and **118** later than **27** and **114** is noteworthy, as it implies that the di- and trimethylated structural variants of **27** are more hydrophilic than the monomethylated structure, unusual given that they contain more carbon atoms. Accordingly, the assignments remain tentative and further study into the structure of the lipids is

necessary. Similarly, structures **82-83** and **115-116**, suggested to be ring-containing $GDGT_M$ lipids, are tentatively assigned and will require further verification.

Isoprenoid chain methylation may also explain the presence of tetramethyl-, pentamethyl- and ethyltetramethylicosanes (Fig. 5.20c) in a Lower Albian black shale from France (Vink *et al.*, 1998). These hydrocarbons were believed to have originated from ether lipids of archaeal origin, although only the pentamethylicosane (PMI) chain is formed *via* isoprenoid biosynthetic pathways known to operate in these organisms. The other chains would require some form of further modification, such as methylation or demethylation of PMI (Vink *et al.*, 1998). This led the authors to conclude that the hydrocarbons may originate from an unsaturated tetramethylicosane-containing ether lipid intermediate, which is methylated to give PMI or ethyltetramethylicosane chain-containing lipid structures. Defunctionalisation of the homologous ether lipids formed during diagenesis would then provide the requisite hydrocarbon chains found in the shale. The tetraether lipid homologues identified here in the VDG sediments suggest that the chain methylations involved in modifications of the PMI chain need not, necessarily, have occurred at unsaturated carbons.

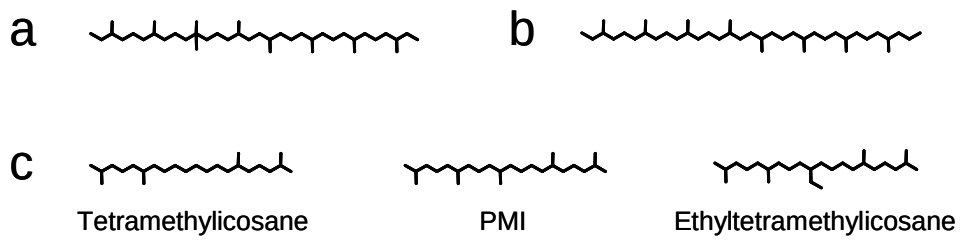


Fig. 5.20. a) and b) show the structures of methyl-branched C_{41} alkyl chains identified previously in aquatic sediments (Blumenberg *et al.*, 2007; Ventura *et al.*, 2007; 2008). c) shows the structures of polyalkylated icosanes identified in a Lower Albian shale (Vink *et al.*, 1998).

MS/MS evidence suggests that **119-123** are structures in which one of the terminal glycerol groups in **2**, **3** and **14-16** has been appended with an additional carbon. Similarly, the MS/MS spectrum of **125** suggests that a second carbon may have been

added to one of the glycerols in **2**. The nature of the homologation in the GDHT/GTHT lipids (i.e. methylene or methyl group) cannot be determined fully by MS/MS analysis and further studies will be required to verify the proposed structures. The modification of **2** and **3** to form **119**, **120** and **125** *via* formation of saturated C-C bonds following cell senescence in the aquatic and terrestrial environments is highly unlikely given the inherent energetic unfavourability of the process. Consequently, the homoglycerol groups must be incorporated during lipid biosynthesis, or otherwise generated during this process by modification of a glycerol group. Unusual lipids based on non-glyceride head groups containing extra methylene groups have been observed in some eukaryotes. These include structures containing 1,2,4-butanetriol moieties (Vaver *et al.*, 1964) which would be consistent with the MS/MS data observed for GDHT/GTHT lipids. Reports of these structures are, however, very rare and are unprecedented in archaea, suggesting that they are unlikely to be present among the archaeal lipids from the environmental samples. A more likely explanation for the unusual trihydric group in **119** and **120** would be an enzymatic C-methylation at one of the three carbons of a terminal glycerol group following the coupling of the two diether precursors to form **2** and **3**. A subsequent methylation of **119** at the same end of the molecule could generate **125**. Alternatively, **125** may be formed by a C-ethylation at one of the glycerols in **2**. Given that alkyl chain methylation at a saturated lipid carbon may be performed by some archaea, similar late-stage modifications of a glycerol group are also plausible. The absence of homoglycerol-modified diethers in the VDG horizon VII-1b extract suggests that the novel C_4 and C_5 trihydric head groups in **119**, **120** and **125** are not introduced into the lipids during the early stages of biosynthesis, although the possibility that their absence is the result of complete conversion of the homoglycerol diether lipids to **119**, **120** and **125** in all of the organisms that express these components cannot be ruled out. Tentative MS/MS assignments suggest that the alkylations which generate GDHT lipids may occur on a carbon at an *sn*-3 position of glycerol. Irrespective of their precise structure and mode of formation, the butan- and pentantriol groups found in GDHT lipids provide compelling evidence that precursors other than glycerol or calditol are utilised by archaea as scaffolds for membrane-spanning lipids.

Although the biogenic origin of the isoprenoid homologues is unclear, they probably originate from a variety of different archaea, especially considering the identification of $GDGT_M$ and GDHT lipids in samples from highly distinct environmental settings (i.e. sediments of different geological ages from two geographically distinct settings and in a modern day terrestrial soil). The identification of the lipid homologues in the soil indicates that chain methylation, and possibly glycerol methylation, can be utilised by mesophilic archaea as a process to modify tetraether structure. As monoalkyl tetraether lipids containing a homoglycerol group were not detected in the VDG lipid extracts, it is possible that homoglycerol-containing lipids and monoalkyl tetraether lipids are produced by distinct groups of organisms.

A novel tetraether lipid core with $[M+H]^+$ at *m/z* 1315.8 was recently identified in a diglycosylated polar lipid isolated from surface sediments in the Peru Margin (Lipp and Hinrichs, 2009). The authors assigned the lipid core as a GDGT containing one Cp ring and bearing a hydroxyl group at the C-3 position of a biphytanyl chain. $GDGT_M$ **26** and GDHT **119** in the aquatic sediments investigated in the present study both exhibit $[M+H]^+$ at the same nominal *m/z* value as the Peruvian lipid core. As such, it is possible that either structure could alternatively be the core of the diglycosidic lipid.

5.2.5. Non-isoprenoid tetraether lipids

5.2.5.1. Known non-isoprenoid tetraether lipid components

The LC-MS base peak chromatogram of the lipid extract from the greenhouse soil shows a series of abundant peaks relating to lipids which elute after isoprenoid GDGT lipids **24** and **27** but before the IPA composition of the eluent exceeds 2% (Fig. 5.21a). These include lipids **36** (*m/z* 1049.7), **35** (*m/z* 1035.7) and **34** (*m/z* 1021.7), which have been reported previously to be GDGT lipids containing the non-isoprenoid hydrocarbon chains 13,16-dimethyloctacosane ($[30_0]$) and/or 5,13,16-trimethyloctacosane ($[31_0]$) (Section 1.3.2; Damsté *et al.*, 2000).

The MS/MS spectrum of **35** (Fig. 5.22a) exhibits product ions formed *via* loss of 1-3 molecules of water (e.g. *m/z* 1017 and 999) and/or a molecule of C_3H_4O (e.g. *m/z* 961 and 943), consistent with the losses observed from isoprenoid GDGT lipids (Chapters 2 and 3). In addition, product ions at *m/z* 617 and *m/z* 603 relate to losses of $[30_0]^0$ (-418 Da) and $[31_0]^0$ (-432 Da) from the $[M+H]^+$ of **35**, respectively, consistent with the alkyl chains reported previously for the lipid. The MS/MS spectrum of **34** (Fig. 5.22b) exhibits a product ion at *m/z* 603, relating to loss of $[30_0]^0$ (-418 Da), again consistent with the reported structure of the lipid. The MS/MS spectrum of **36** (Fig. 5.22c), on the other hand, shows product ions at *m/z* 631, 617 and 603, representing losses of, in turn, $[30_0]^0$ (-418 Da), $[31_0]^0$ (-432 Da) and $[32_0]^0$ (-446 Da). The product ions observed suggest that **36** (GDGT[31_0,31_0]) may be accompanied by a second co-eluting isomer (**36***; GDGT[30_0,32_0]) which is also present in the soil extract.

Lipids **38** (*m/z* 1033.7), **41** (*m/z* 1031.7), **37** (*m/z* 1019.7) and **40** (*m/z* 1017.7) were identified in individual ion chromatograms for the soil extract (Fig. 5.21b), exhibiting retention times and $[M+H]^+$ consistent with GDGT structures containing 13,16-dimethyloctacosane and/or 5,13,16-trimethyloctacosane chains exhibiting 0-2 incorporated Cp rings (Section 1.3.2; Schouten *et al.*, 2000; Weijers *et al.*, 2006a). Cp ring-containing derivatives of **36** (i.e. **39** and **42**), also identified in previous analyses of lipid extracts from peats (Weijers *et al.*, 2006a), were not detected in the soil sample. The MS/MS spectrum of **37** (Fig. 5.22d) exhibits product ions at *m/z* 603 and *m/z* 601, formed from loss of $[30_1]^0$ (-416 Da) and $[30_0]^0$ (-418 Da), respectively. The MS/MS spectrum of **40** (Fig. 5.22e), on the other hand, exhibits a product ion at *m/z* 601 only, indicative of loss of $[30_1]^0$ from the precursor. Both spectra are consistent with the structures assigned previously for **37** and **40** (Weijers *et al.*, 2006a). Lipids **38** and **41** were concealed by co-elution with more abundant lipids **34** and **37**, respectively. As a consequence, MS/MS spectra were not obtained for these components.

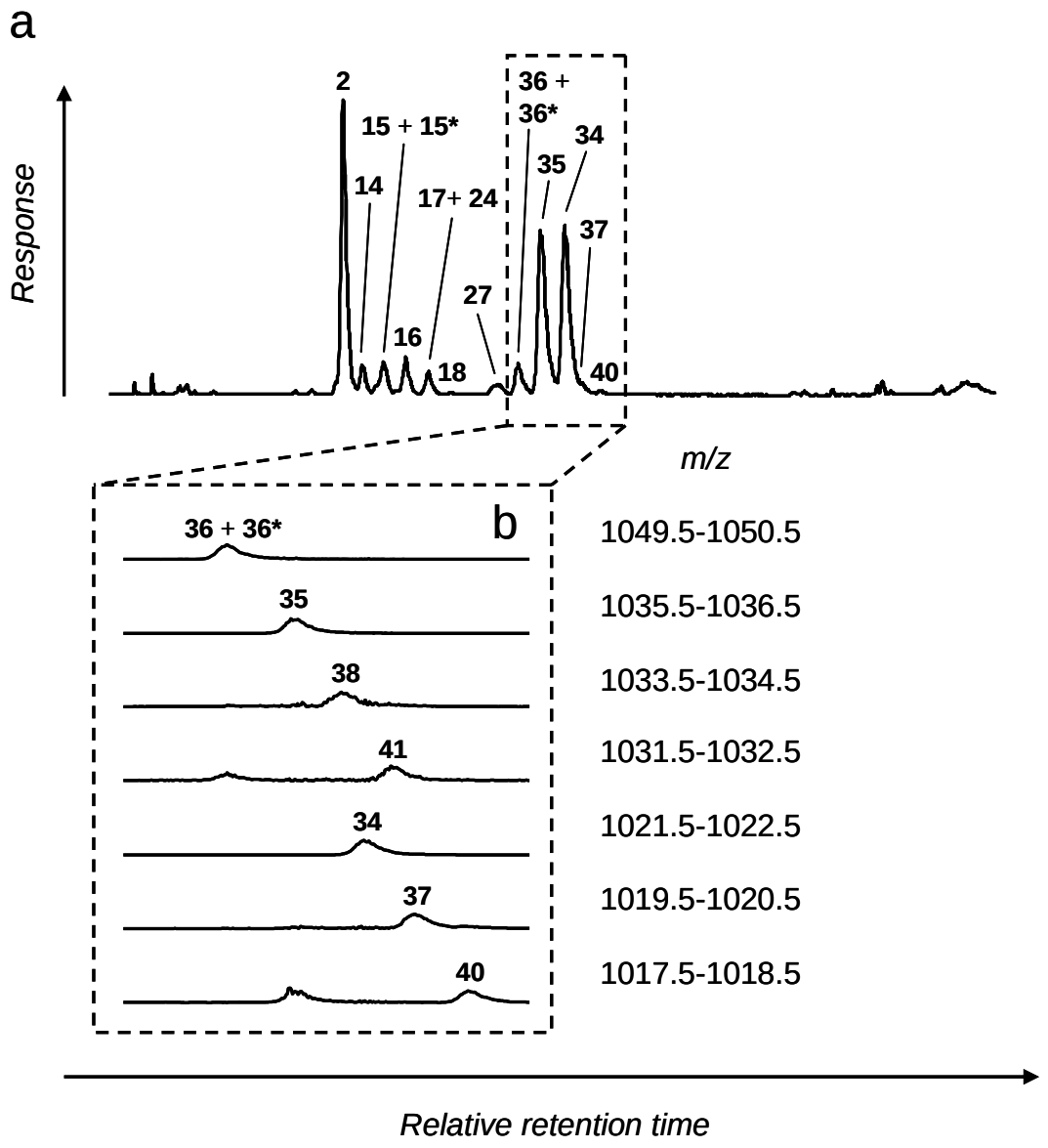


Fig. 5.21. HCT LC-MS ion chromatograms for the lipid extract from a greenhouse soil; a) base peak chromatogram (*m/z* 900-1500); b) individual ion chromatograms, highlighting the presence of GDGT lipids containing non-isoprenoid hydrocarbon chains.

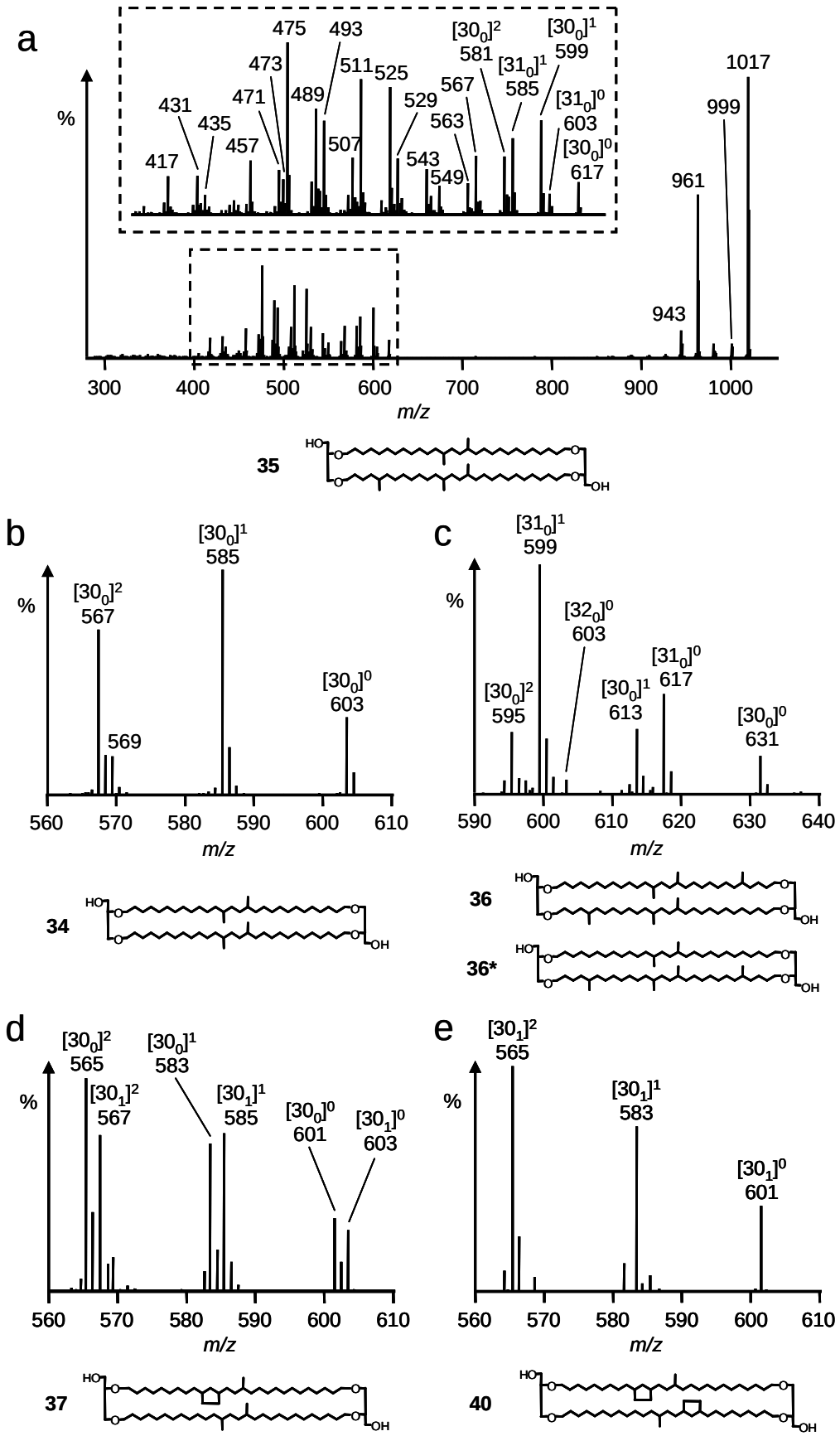

a
%
417
431
435
457
471
473
475
489
493
507
511
525
529
543
549
563
567
581
585
599
603
617
943
961
999
1017
300
400
500
600
700
800
900
1000
m/z
35
b
%
[30$_0$]^2 567
[30$_0$]^1 585
569
[30$_0$]^0 603
560
570
580
590
600
610
m/z
34
c
%
[31$_0$]^1 599
[32$_0$]^0 603
[30$_0$]^2 595
[30$_0$]^1 613
[31$_0$]^0 617
[30$_0$]^0 631
590
600
610
620
630
640
m/z
36
36*
d
%
[30$_0$]^2 565
[30$_1$]^2 567
[30$_0$]^1 583
[30$_1$]^1 585
[30$_0$]^0 601
[30$_1$]^0 603
560
570
580
590
600
610
m/z
37
e
%
[30$_1$]^2 565
[30$_1$]^1 583
[30$_1$]^0 601
560
570
580
590
600
610
m/z
40

Fig. 5.22. (opposite) MS/MS and partial MS/MS spectra (HCT) of GDGT lipids containing non-isoprenoid alkyl chains, identified in the lipid extract from a greenhouse soil: a) **35** ($[M+H]^+$ *m/z* = 1035.7). Inset shows expansion of a region of the spectrum; b) **34** ($[M+H]^+$ *m/z* = 1021.7); c) **36**/**36*** ($[M+H]^+$ *m/z* = 1049.7); d) **37** ($[M+H]^+$ *m/z* = 1019.7); e) **40** ($[M+H]^+$ *m/z* = 1017.7).

5.2.5.2. $C_nH_{2n}O_6$ & $C_nH_{2n+3}O_5$ homologous series of non-isoprenoid ether lipids

A series of components exhibiting $[M+H]^+$ values between nominal *m/z* 965 and 1273 were observed in the extracts from the VDG samples, most easily visualised in chromatograms for the VDG X-1 extract (Fig. 5.23a), where their abundance is high and co-occurring isoprenoid tetraether lipid complexity is low. Individual ion chromatograms indicate that many of these components lie within two homologous series, with molecular formulae of either $C_nH_{2n}O_6$ or $C_nH_{2n+3}O_5$ (Fig. 5.23b). The retention times of components from both series suggest that they may be further categorised into a series of eleven structurally distinct classes (Classes A-L). The majority of the components provided the base peak ion in an MS scan during analysis of at least one VDG horizon extract, allowing MS/MS spectra to be obtained. Other components, concealed under co-occuring, more abundant isoprenoid lipids were selected during LC-MS/MS analysis of the extract from the X-1 horizon using a truncated MS scan range of *m/z* 900-1270, chosen to exclude the majority of the isoprenoid lipid components. Some components were still not selected using the modified scan range and, consequently, are categorised solely on the basis of $[M+H]^+$ and relative retention time. The classification system is supported by the MS/MS data obtained, which show that components from each class, where selected for CID, undergo some common and some unique neutral losses, confirming that each represents a tetraether or modified tetraether structure containing non-isoprenoid alkyl chains. These characteristic losses are summarised in Tables 5.2 and 5.4 and are discussed individually.

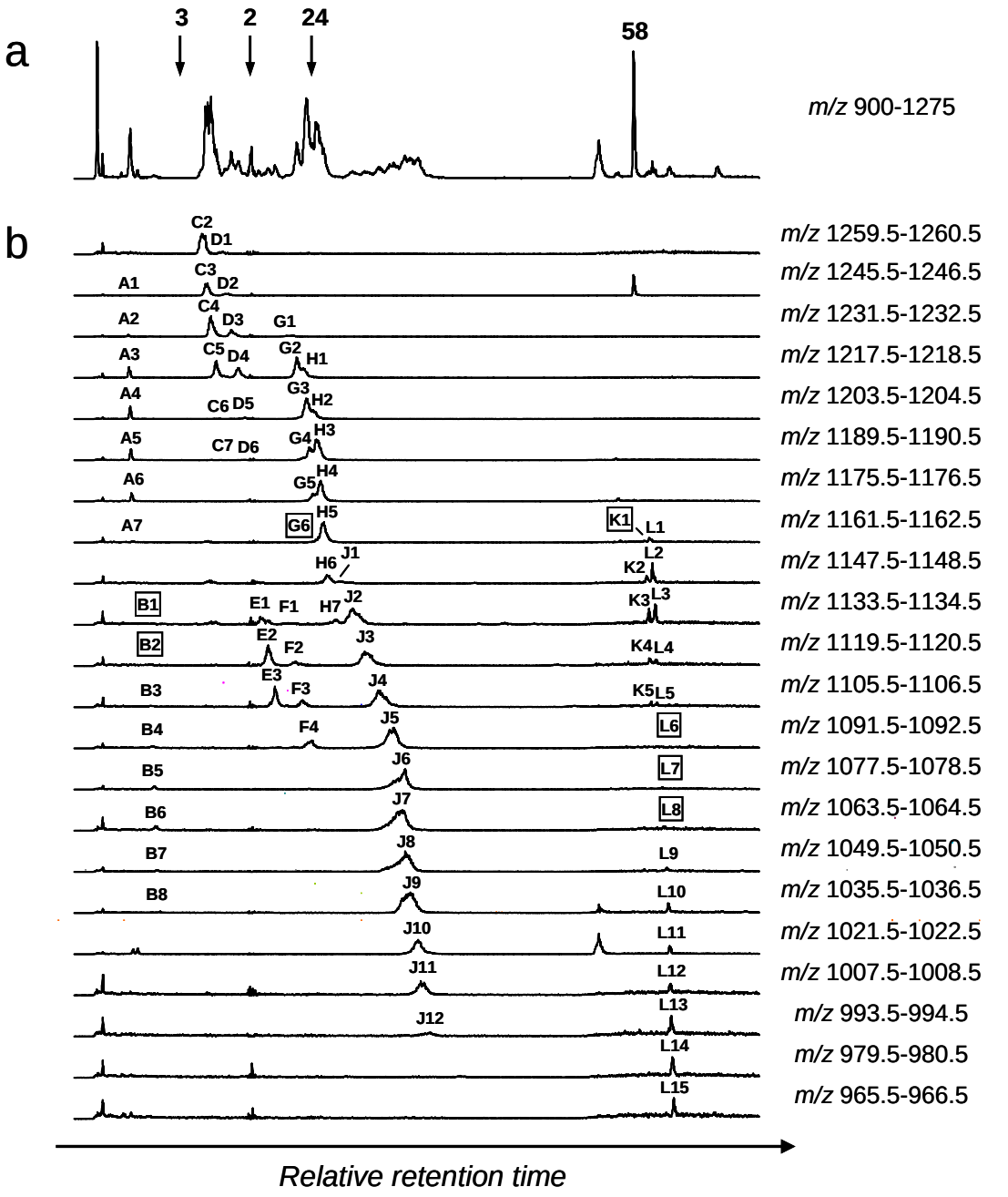


Fig. 5.23. HCT LC-MS ion chromatograms for a lipid extract from VDG horizon X-1, highlighting non-isoprenoid lipids with elemental composition $C_nH_{2n}O_6$ or $C_nH_{2n+3}O_5$: a) ion chromatogram (*m/z* 900-1275) selected to exclude the majority of isoprenoid lipid components. The elution times of GTGT **3**, GDGTs **2** and **24** and GDT **58** are included for comparison; b) individual ion chromatograms. Components labelled in a box were not observed in the chromatogram for the X-1 horizon, but were observed in those for other horizons from the VDG evaporitic sequence.

Table 5.4. Product ions[a] observed in the MS/MS spectra obtained for lipids in Classes A-J identified in VDG sediment.

			Class A (Methyl ether)					**Class C (GDH_2T)**				**Class D (GDHT)**				
			2	**3**	**4+4***	**5**	**6**	**2**	**3**	**4**	**5**	**2**	**3**	**4**	**5**	**6**
Molecules lost	Loss (Da)	$[M+H]^+$ (nominal)	1231	1217	1203	1199	1175	1259	1245	1231	1217	1245	1231	1217	1203	1189
H_2O	-18		+[a]	+	+	+	+	+	+	+	+	+	+	+	+	+
MeOH	-32		+	+	+	+	+									
$2H_2O$	-36							+	+	+	+	+	+	+	+	+
MeOH + H_2O	-50		+	+	+	+	+									
$3H_2O$	-54							+	+	+	+	+	+	+	+	
C_3H_4O	-56		+	+	+	+	+									
MeOH + $2H_2O$	-68			+	+	+										
C_4H_6O	-70											+	+	+		
$4H_2O$?	-72							+	+	+	+					
C_3H_4O + H_2O	-74		+	+	+	+	+					+	+	+	+	
C_5H_8O	-84							+	+	+	+					
C_3H_4O + MeOH	-88		+	+	+	+	+									
C_4H_6O + H_2O	-88											+	+	+	+	+
C_3H_4O + $2H_2O$	-92															
C_5H_8O + H_2O	-102							+[b]	+	+	+					
?	-102											+	+	+		
?	-104							+	+	+	+					
$C_3H_6O_2$ + MeOH	-106		+	+	+	+	+									
C_4H_6O + $2H_2O$	-106											+	+	+	+	
C_3H_4O + $3H_2O$?	-110							+	+	+	+					
C_5H_8O + $2H_2O$	-120							+	+	+	+					
$[30_0]^0$	-418															
$[31_0]^0$	-432															
$[32_0]^0$	-446															
$[33_0]^0$	-460															
$[34_0]^0$	-474															
$[35_0]^0$	-488				+	+	+									
$[36_0]^0$	-502			+	+	+				x	x			x	x	
$[37_0]^0$	-516		+	+	+			x	x	x		x	x	x		
$[38_0]^0$								x				x				

[a] (+) = Product ion observed in the MS/MS spectrum; (x) = Product ion not observed, but same alkyl chain composition inferred from other product ions in the MS/MS spectrum.

[b] Product ions shaded grey (where shown) represent the ion consistently observed as the base peak ion in MS/MS.

Class E (GDH_2T)			Class G (GDGT)			Class H (GDGT)							Class J (GDGT)								
1	**2**	**3**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**2**	**3**	**4+4***	**5**	**6**	**7**	**8+8***	**9**	**10**
1133	1119	1105	1217	1203	1189	1217	1203	1189	1175	1161	1147	1133	1133	1119	1105	1091	1077	1063	1049	1035	1021
+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	+	+	+	+	+	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+
	+	+	+	+	+	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+
			+	+	+	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+
	+	+																			
			+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	+	+																			
			+	+	+	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+
+	+	+																			
	+	+																			
	+	+																			
																			+	+	+
																		+	+	+	
	x	x													+	+	+	+	+		
	x													+	+	+					
											+		+	+	+						
					+				+	+	+										
				+				+	+	+											
			+	+	+	+	+														

Table 5.4. (cont.)

5.2.5.2.1. Class J – GDGT lipids containing C_{29} to C_{35} alkyl chains

Three components identified in the ion chromatograms for the VDG extracts, **J8-J10**, elute with retention times, and exhibit $[M+H]^+$ consistent with non-isoprenoid tetraether lipids **34-36** (Fig. 5.23b). The MS/MS spectrum of **J8** contains the same product ions as were observed during CID of the $[M+H]^+$ of **36** and co-eluting isomer **36*** (Fig. 5.22c) indicating that **J8** may comprise the same mixture of isomers (i.e. **J8** = **36**; **J8*** = **36***). By contrast, the MS/MS spectrum of **J9** (e.g. from the VDG IV-1.9 sample; Fig. 5.24a) is different to that obtained from the $[M+H]^+$ of **35** (greenhouse soil; Fig. 5.22a). The $[M+H]^+$ of **J9** (*m/z* 1036.1) dissociates to lose $[30_0]^0$ (-418 Da; *m/z* 617) and $[31_0]^0$ (-432 Da; *m/z* 603), losses also observed from the $[M+H]^+$ of **35**. In addition, several other ions were also observed in the MS/MS spectrum of **J9** which cannot have originated from the $[M+H]^+$ of **35** given its reported structure. As such, **J9** (GDGT[30_0,31_0]) appears to be accompanied by a co-eluting isomer of unknown structure. Similarly, the MS/MS spectrum of **J10** (e.g. from the Oxford Clay sample; Fig. 5.24b) contains product ions formed from losses of $[30_0]^0$ (-418 Da; *m/z* 603), $[30_0]^1$ (-436 Da; *m/z* 585) and $[30_0]^2$ (-454 Da; *m/z* 567) as were observed in the spectrum of **34** (Fig. 5.22b), but also contains other ions that were absent from the latter, presumably originating from a co-eluting isobar.

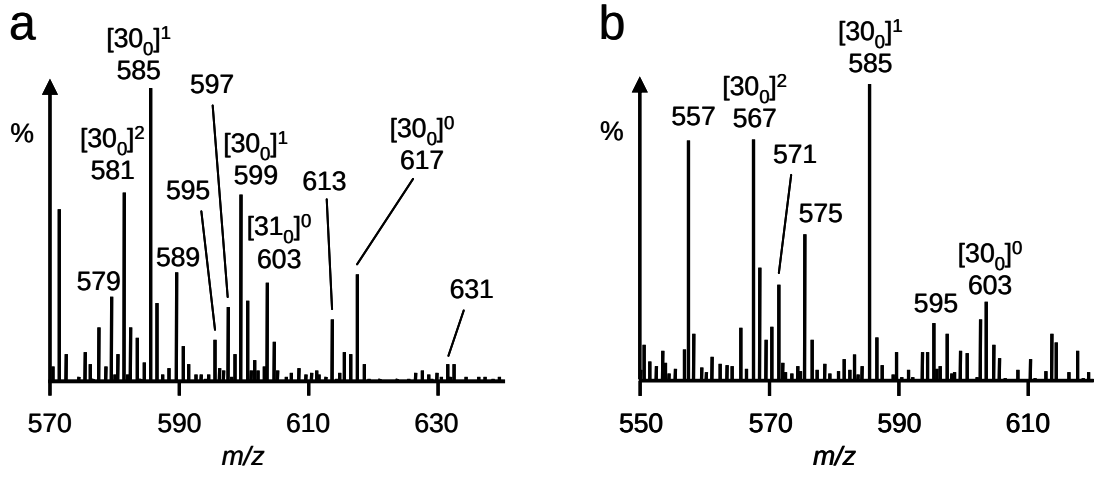


Fig. 5.24. Partial MS/MS spectra (HCT) of: a) **J9** (VDG IV-1.9; $[M+H]^+$ *m/z* = 1036.1); b) **J10** (Oxford Clay S90-11; $[M+H]^+$ *m/z* = 1022.1).

On the basis of the relative retention times of the components, **J8-J10** appear to lie within a continuous homologous series of lipids which includes components **J1-J7** and **J11-J12** (Fig. 5.23b). All other components within this series for which MS/MS spectra could be obtained generated product ions formed from loss of 1-3 molecules of water and/or a molecule of C_3H_4O, with particular alkyl chains also lost as dienes in each case (Table 5.4). The MS/MS spectral evidence indicates that **J2-J10** are GDGT lipids containing C_{30}-C_{34} alkyl chains. By proxy, **J1** presumably has a GDGT[34_0,35_0] structure. Likewise, **J11** and **J12** probably have GDGT[29_0,30_0] and GDGT[29_0,29_0] structures, respectively. The MS/MS spectrum obtained for **J4** ($[M+H]^+$ *m/z* = 1106.1) indicates losses of $[32_0]^0$, $[33_0]^0$ and $[34_0]^0$ during CID (Table 5.4), suggesting that the component has a GDGT[33_0,33_0] structure and has a co-eluting isomer (designated **J4***) with a GDGT[32_0,34_0] structure. An ion chromatogram which selects for **J4** and **J4*** in the extract from the VDG IV-1.9 horizon shows an earlier eluting species with the same nominal *m/z* (Fig. 5.25), possibly representing a third isomer, although no MS/MS spectrum was obtained for this component.

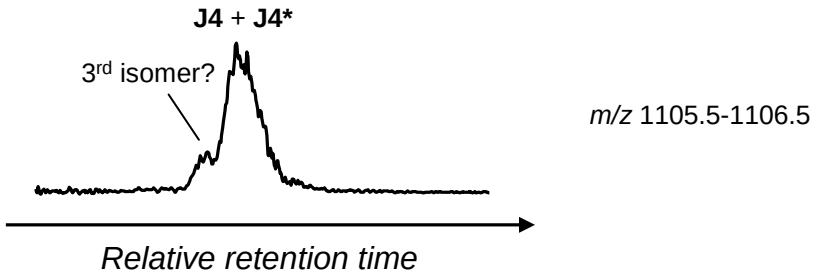


Fig. 5.25. Partial HCT ion chromatogram for a lipid extract from VDG horizon IV-1.9, indicating the presence of an isomer of **J4** and **J4*** ($[M+H]^+$ *m/z* = 1106.1).

5.2.5.2.2. Classes G and H – GDGT lipids containing C_{34} to C_{38} alkyl chains

Two other classes of components, Class G (**G1-G6**) and Class H (**H1-H7**), were identified in the ion chromatograms for the VDG extracts (Fig. 5.23b). The components for which MS/MS spectra were obtained show losses of 1-3 molecules of water and/or a molecule of C_3H_4O and also lose specific alkyl chains as dienes (Table 5.4). The sole exception, **H7**, gave a very weak MS/MS spectrum in which

product ions formed from loss of water with or without simultaneous loss of $C_3H_6O_2$ are the only ions which can be interpreted reliably. These data indicate that the Class G and H components are GDGT lipids containing C_{34}-C_{37} alkyl chains. By way of example, the MS/MS spectrum (Fig. 5.26) of **G3** ($[M+H]^+$ *m/z* = 1204.3) contains product ions at *m/z* 701 and *m/z* 687, formed from loss of $[36_0]^0$ (-502 Da) and $[37_0]^0$ (-516 Da), respectively. Consequently, **G3** is assigned to a GDGT[36_0,37_0] structure.

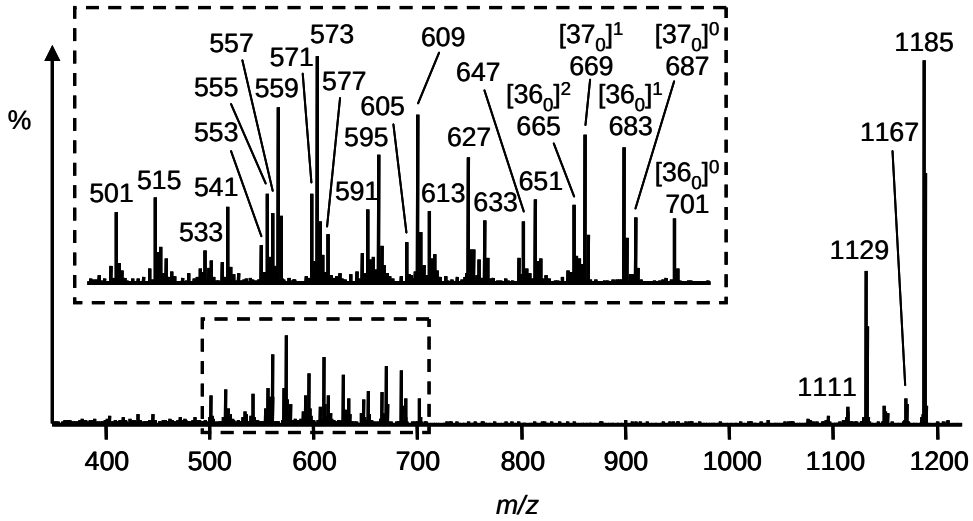


Fig. 5.26. MS/MS spectrum (HCT) of **G3** (VDG X-1; $[M+H]^+$ *m/z* = 1204.3). Inset shows expansion of a region of the spectrum.

The MS/MS spectrum (Fig. 5.27a) of component **G4** ($[M+H]^+$ *m/z* = 1190.2) shows product ions at *m/z* 701 and 673 formed from losses of $[35_0]^0$ (-488 Da) and $[37_0]^0$ (-516 Da) alkyl chains, respectively. This would indicate a GDGT[35_0,37_0] structure for the lipid, consistent with GDGT **50**, which has been identified previously in VDG sediment (Schouten *et al.*, 2000). The MS/MS spectrum (Fig. 5.27b) of isobaric lipid **H3** ($[M+H]^+$ *m/z* = 1190.3) shows a product ion at *m/z* 687 formed from loss of $[36_0]^0$ (-502 Da), indicating that it has a GDGT[36_0,36_0] structure. At first sight, product ions at *m/z* 687 and *m/z* 669 in the MS/MS spectrum of **G4** suggest loss of $[36_0]^0$ and $[36_0]^1$ (-520 Da) from the $[M+H]^+$ of the lipid. On the contrary, these ions are probably attributable to a contribution from ions formed from the $[M+H]^+$ of **H3**, on account of the lack of full chromatographic resolution of **H3** from **G4** during LC-MS/MS.

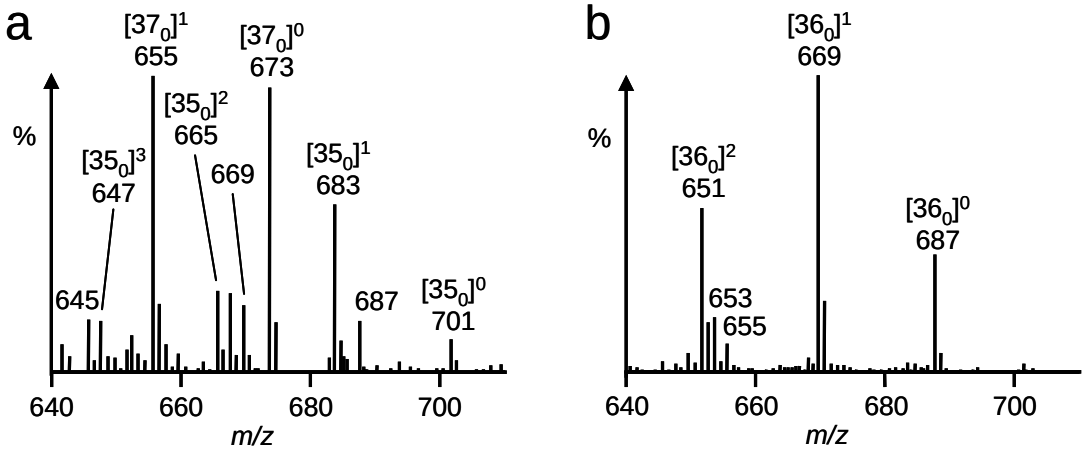


Fig. 5.27. Partial MS/MS spectra (HCT) of: a) **G4** (VDG IV-1.1; $[M+H]^+$ *m/z* = 1190.2); b) **H3** (VDG X-1; $[M+H]^+$ *m/z* = 1190.3).

Individual ion chromatograms (*m/z* 1189.5-1190.5) for some of the extracts from the VDG horizons indicate that, in addition to **G4** and **H3**, a third isobar eluting marginally before **G4** may also be present (e.g. VDG horizon IV-1.2; Fig. 5.28). Although no supporting MS/MS spectra was obtained for this lipid, it may represent an isomeric GDGT structure.

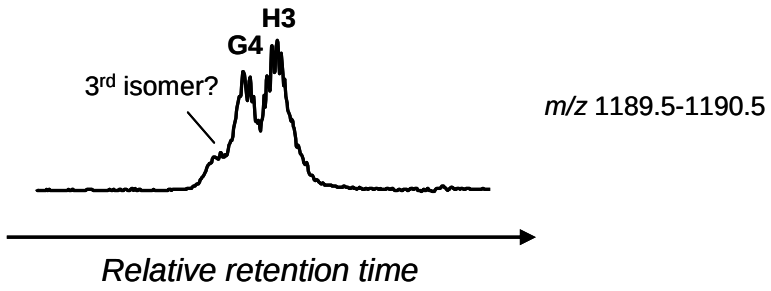


Fig. 5.28. Partial HCT ion chromatogram for a lipid extract from VDG horizon IV-1.2, indicating the presence of a possible additional isomer of **G4** and **H3** ($[M+H]^+$ *m/z* = 1190.2).

5.2.5.2.3. Classes A and B – GDGT methyl ethers

Class A and B components elute significantly before GTGT **3** during LC-MS/MS (Fig. 5.23b). Components within each class exhibit poor chromatographic resolution, although the two classes are well resolved from one-another. MS/MS spectra were

obtained for **A2-A6** but spectra obtained for Class B components were invariably too weak to be used for structural assignment. The MS/MS spectra of **A2-A6** (Table 5.4) contain product ions in two distinct spectral regions and include those formed from loss of methanol (-32 Da). These data suggest dialkyl-type tetraether core structures for the Class A components, with the terminal hydroxyl in one glycerol moiety modified as a methyl ether and the other glycerol moiety unchanged. Furthermore, the early elution times of these components relative to the native structure are consistent with those observed for **51**, an artificial methyl ether of **2** (Section 2.2.4).

Peaks **A2-A6** also show product ions in MS/MS formed from losses of C_{35}-C_{37} alkyl chains as dienes (Table 5.4). Interestingly, the alkyl chains in each structure are the same as those identified in Class G and H GDGT components which contain, in total, one carbon atom less. This suggests that the Class A components are probably methyl etherified derivatives of the Class G and H components. On the basis of the $[M+H]^+$ and relative retention times of Class B components in relation to Class J components, they most likely represent the methyl etherified derivatives of **J3-J10**.

5.2.5.2.4. Classes D and F – GDHT lipids

Components in Classes D and F exhibit retention times suggestive of an intermediate polarity to Class A-B and Class G-J components (Fig. 5.23b). Class F components were concealed *via* co-elution with members of Class G, resulting in an absence of accumulated MS/MS data for this compound class, even following the re-analysis of the VDG X-1 horizon extract using a restricted MS scan range. MS/MS spectra for all Class D components apart from **D1** were obtained (Table 5.4). CID of the $[M+H]^+$ of **D2-D6** resulted in product ions in two distinct regions of the MS/MS spectra, suggesting dialkyl-type core structures. Typically, the protonated molecules of Class D components expelled a 70 Da molecule, with additional losses of one (88 Da) or two (106 Da) molecules of water, amongst other molecules lost during CID. These losses are identical to those observed from the $[M+H]^+$ of GDHT **119** (Section 5.2.4.4), indicating that Class D components are GDHT lipids. Product ions formed from loss of 102 Da were also observed in the MS/MS spectra of some Class D lipids, a loss which was not observed from the $[M+H]^+$ of **119** during CID. Fragment

ions formed from the loss of water and 88 Da are also observed in the MS spectrum of each Class D GDHT lipid. The MS ion currents obtained for the $[M+H]^+$ of Class F components were too weak for similar fragment ions to be confirmed.

Product ions relating to neutral losses of $[N_0]^0$ and being consistent with the $[M+H]^+$ of each lipid were absent from the MS/MS spectra of Class D components, making assignment of the alkyl chains difficult. The lowest *m/z* value at which a prominent product ion is observed in the MS/MS spectrum of GDHT **119** is *m/z* 557 (Fig. 5.14a). This ion is also observed in the MS/MS spectrum of GDGT **2**, where it was assigned as $C_{40}H_{77}^+$ (Fig. 2.6 and Section 2.2.9.5). Considering that a $[40_0]$ chain was identified in GDHT **119** (Section 5.2.4.4), the ion at *m/z* 557 in the MS/MS spectrum of the isoprenoid GDHT lipid is, most likely, also $C_{40}H_{77}^+$. As this low *m/z* ion signifies the number of carbons in the constituent alkyl chains within GDHT **119**, the two prominent product ions at lowest *m/z* in the MS/MS spectra of **D2**-**D5** can allow tentative identification of the two alkyl chains within each Class D component. For example, the two lowest *m/z* ratios for prominent ions observed in the MS/MS spectrum of **D2** are at *m/z* 515 and 529; 42 and 28 *m/z* units less than *m/z* 557. This indicates that one alkyl chain in **D2** contains 37 carbon atoms, whilst the other contains 38 carbon atoms, indicating a possible GDHT$[37_0,38_0]$ structure. Similar assignment of **D3**-**D5** indicates that they represent GDHT lipids containing alkyl chains comprising 36 or 37 carbon atoms (Table 5.4). Interestingly, with the exception of **D1**, the $[M+H]^+$ of each Class D component is shifted +14 *m/z* units from components in Classes G and H which contain the same alkyl chains. It is possible, therefore, that **D2**-**D6** represent the GDHT analogues of the Class G or H GDGT components. Given the retention times of Class F components relative to Class J components, they may represent GDHT analogues of **J3**-**J6**.

5.2.5.2.5. Classes C and E – GDH$_2$T lipids

Components in Classes C and E elute marginally earlier during LC-MS/MS of the VDG extracts than isobaric components in Classes D and F, respectively (Fig. 5.23b). Component **C1** (nominal *m/z* 1273) was identified in some horizons but was absent in horizon X-1, explaining its absence in the chromatograms shown in Fig.

5.23b. Tandem MS spectra for **C2-C5** were obtained in conventional LC-MS/MS analyses. By contrast, peaks for Class E components lie under those of isoprenoid GDGT lipids and required re-analysis (MS scan range *m/z* 900-1270) of the X-1 horizon extract to obtain spectra for **E1-E3**. In nearly all cases, CID of the $[M+H]^+$ of Class C or E components gave rise to loss of an individual 84 Da molecule, with one (102 Da) or two (120 Da) additional molecules of water also being lost (Table 5.4). Fragment ions formed from loss of water or 102 Da are also observed in the MS spectrum of each component. Loss of 84 and 102 Da, characteristic losses observed during CID of the $[M+H]^+$ of GDH_2T lipid **125** (Section 5.2.4.6), indicate that Class C and E components are non-isoprenoid GDH_2T lipids.

As noted for Class D lipids, no product ions consistent with loss of $[N_0]^0$ were observed in the MS/MS spectra of Class C and E components. Given that the MS/MS spectrum of GDH_2T **125** (Fig. 5.19) contains a product ion at *m/z* 557, most likely corresponding to a $C_{40}H_{77}^+$ ion, the nature of the alkyl chains in **C2-C5** was inferred from the two lowest *m/z* values of prominent product ions in the MS/MS spectrum of each, as outlined in Section 5.2.5.2.4. For example, low *m/z* product ions in the spectrum of **C3** are observed at *m/z* 515 and *m/z* 533. The former is 42 *m/z* units less than *m/z* 557, suggestive of an ion containing 37 carbon atoms, and the latter is +18 *m/z* greater than the former, indicating that the ion formally contains an additional molecule of water. These ions suggest that **C3** contains two alkyl chains with 37 carbon atoms and, consequently, suggests that **C3** has a $GDH_2T[37_0,37_0]$ structure. Similarly, **C2** and **C4-C5** are assigned to structures with alkyl chains containing 36-38 carbon atoms, and **E1-E3** to structures with chains containing 32-33 carbons (Table 5.4).

Components within Class C for which alkyl chains could be assigned exhibit $[M+H]^+$ shifted by +28 *m/z* units from those of components in Classes G and H containing the same alkyl chains. This strongly suggests that the Class C components may be GDH_2T homologues of Class G or H GDGT lipids. The only possible exception is **C1**, for which a GDGT analogue was not identified in any of the VDG horizons examined. Class E components, likewise, are most probably GDH_2T homologues of Class J GDGT lipids.

5.2.5.2.6. Classes K and L – GDT lipids

A number of components of the homologous series elute in the ion chromatograms for the VDG extracts later than GDT **58** (Fig. 5.23b). Many of the late eluting compounds are chromatographically resolved into two apparent classes (Classes K and L). The MS/MS spectra of components in each class (Table 5.2) show product ions formed from losses of 1-2 molecules of water but do not show ions formed from loss of $[N_0]^0$. Interestingly, losses of $[N_0]^1$ were observed from the $[M+H]^+$ of each lipid for which a reliable MS/MS spectrum could be obtained, indicating that they can all lose a hydroxylated alkenyl chain. Furthermore, the number of carbon atoms in the chains lost from each Class K or L lipid were found to be the same as those lost as $[N_0]^0$ from lipids within Classes G-J, which exhibit $[M+H]^+$ values 56 *m/z* units greater. This relationship mirrors that observed between GDGT **2** and GDT **58**, which have $[M+H]^+$ which differ by 56 *m/z* units and lose biphytadiene and 1-ene-biphtyan-32-ol, respectively, as the smallest biphytanyl chain-bearing molecules during CID (see Sections 2.2.5.2 and 5.2.3.2). The relative retention times of Class K and L components and their Class G-J relatives also compare favourably with those of GDT **58** and GDGT **2**. Consequently, Class K and L lipids probably represent GDT counterparts of GDGT lipids in Classes G-J. Although nearly all spectra for Class K and L lipids are consistent with this assignment, the spectrum of **K3** shows product ions formed from loss of $[36_0]^1$. On the basis of MS/MS evidence, the C_{36} alkyl chain implicated by this loss is absent in **G4**, the most likely GDGT structural relative. Similarly, the MS/MS spectrum of **L3** contains product ions formed from loss of $[35_0]^1$ and $[37_0]^1$ during CID, both of which represent alkyl chains that are not lost from the $[M+H]^+$ of **H3**, its most likely GDGT relative. Until these anomalies can be better explained, the assignment of Class K and L components remains tentative.

5.2.5.3. $C_nH_{2n-2}O_6$ & $C_nH_{2n+1}O_5$ homologous series of non-isoprenoid ether lipids

Other components, with molecular formulae of $C_nH_{2n-2}O_6$ or $C_nH_{2n+1}O_5$ form third and fourth, albeit more limited, homologous series within the VDG sediment extracts. The components appear to be distributed temporally into three classes (Classes R-T) and are most easily visualised in ion chromatograms for the horizon III-1b extract (Fig. 5.29). MS/MS spectra for many of the components were obtained during LC-MS/MS analysis of the extract of at least one VDG horizon (Tables 5.1 and 5.2). The distinct spectral features of each class are discussed individually.

5.2.5.3.1. Class R – GDGT lipids containing one Cp ring

Ion chromatograms at 2 *m/z* units less than the $[M+H]^+$ of identified Class G-J components (Fig. 5.29b) revealed two components, **R1** and **R2**, in several of the VDG extracts. Lipid **R1** elutes immediately after **J9** and exhibits a $[M+H]^+$ at 2 *m/z* units less. Similarly, lipid **R2** elutes slightly later than **J10** and also shows a $[M+H]^+$ two *m/z* units lower than the Class J component. The retention times and $[M+H]^+$ of **R1** and **R2** are consistent with those of GDGTs **38** and **37**, respectively. Furthermore, the MS/MS spectrum of **R2** exhibits product ions formed from loss of $[30_0]^0$ and $[30_1]^0$, also consistent with that of **37** (Fig. 5.22d). Thus, Class R comprises a set of GDGT lipids containing one cyclopentyl ring, located on one of the alkyl chains.

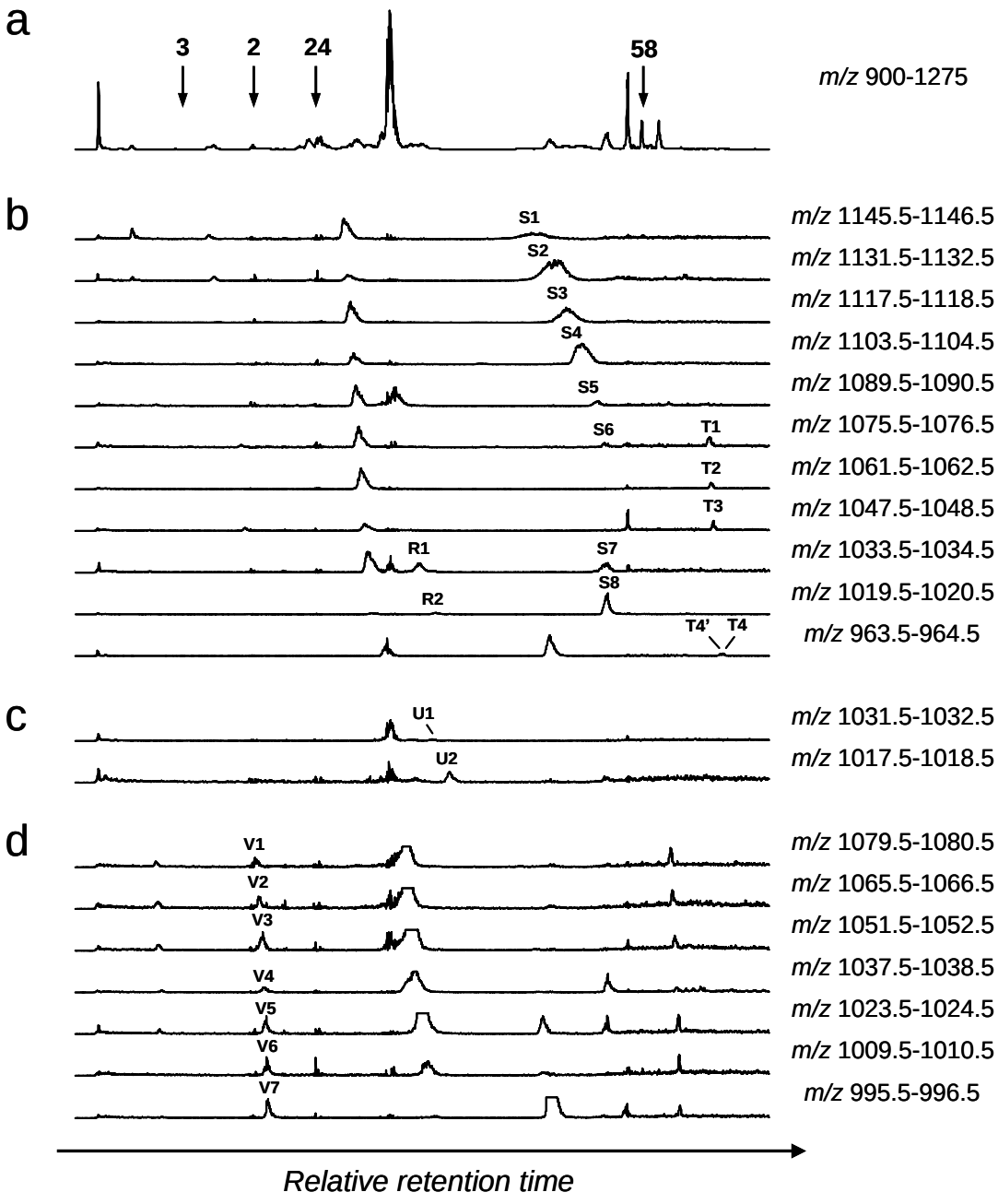


Fig. 5.29. HCT LC-MS ion chromatograms for a lipid extract from VDG horizon III-1b: a) ion chromatogram (*m/z* 900-1275) selected to exclude the majority of isoprenoid lipid components. The elution times of GTGT **3**, GDGTs **2** and **24** and GDT **58** are included for comparison; b-d) individual ion chromatograms highlighting the presence of non-isoprenoid lipid components with elemental composition $C_nH_{2n-2}O_6$ or $C_nH_{2n+1}O_5$ (b), $C_nH_{2n-4}O_6$ (c) or $C_nH_{2n+2}O_6$ (d).

5.2.5.3.2. Class S – GMGT lipids

Class S components begin to elute as the proportion of IPA in the eluent approaches 2%, with the majority eluting at higher elutropic compositions (Fig. 5.29b). Eight components (**S1**-**S8**) within this class were identified in the VDG extracts, with horizon III-1b extract showing the highest relative levels of the Class S components. As was observed for components in Class R, lipids **S1**-**S8** similarly exhibit $[M+H]^+$ at approximately 2 *m/z* units less than components of Class J, suggesting that they are also dehydrogenated counterparts. The MS/MS spectra of five Class S components (**S1**-**S4** and **S8**) were obtained during LC-MS/MS analysis of the VDG extracts, with the spectra for **S7**-**S8** also being obtained from the analysis of the greenhouse soil extract (Table 5.1). Each spectrum (e.g. **S8**; Fig. 5.30) shows product ions formed from losses of 0-4 molecules of water and/or 1-2 molecules of C_3H_4O and *via* other neutral losses, many of which cannot currently be attributed to specific molecules. Additional product ions formed from losses of 148, 150, 152, 154 and 186 Da were also observed following CID of each component, but are omitted from Table 5.1 in the interests of simplicity. No product ions consistent with losses of C_{30}-C_{38} alkene or diene molecules from the precursor were observed in any case. Such extensive loss of small neutral molecules from the precursors, coupled with the inability to lose an alkyl chain following CID is indicative of a GMGT core type (see Sections 2.2.4.2.4 and 4.2.2.1). Thus, Class S components are most likely structural variants of Class J lipids in which the two composite alkyl chains have been conjoined by a covalent bond, generating C_{60}-C_{61} (**S7**-**S8**) or C_{64}-C_{69} (**S1**-**S6**) hydrocarbons. Structural appendage of this nature would account for the requisite reduction in the *m/z* values of the $[M+H]^+$ and explain the observed increases in the retention times of Class S components compared with Class J components; both effects were noted previously for isoprenoid GMGT lipids and their GDGT counterparts (Chapters 2-4 and Schouten *et al.*, 2008a; 2008b). It is important to note that lipids **S1** and **S2** could also be related to Class H components, **H6** and **H7**, respectively, given the $[M+H]^+$ observed for the latter pair. In all VDG samples, the **S1**:**S2** ratio (calculated from LC-MS peak areas) more closely resembles the **J1**:**J2** ratio than the **H6**:**H7** ratio, suggesting that **S1** and **S2** are more likely derived from or related to Class J lipids. An additional component with $[M+H]^+$ at *m/z* 1047.6, which

elutes immediately before **S7**, was also observed in an ion chromatogram for the soil extract, hinting that a GMGT structural relative of GDGT **J8** may also be present in the sample. Unfortunately, no MS/MS spectrum was obtained for this component.

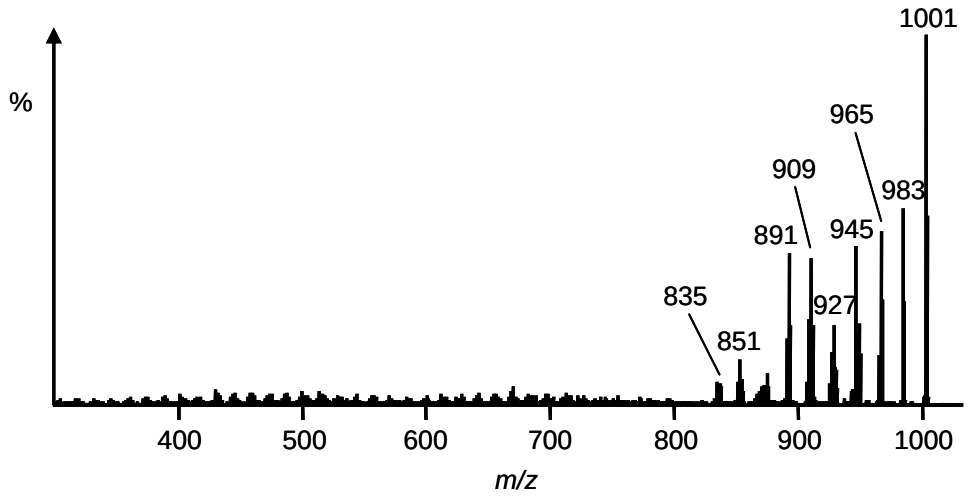


Fig. 5.30. MS/MS spectrum (HCT) of **S8** (VDG III-1b; $[M+H]^+$ *m/z* = 1020.1).

5.2.5.3.3. Class T – GMT lipids

Five very late eluting components, **T1**-**T4** and **T4'**, were identified in the ion chromatograms for some of the VDG extracts (Fig. 5.29b) and also in the S90-11 Oxford Clay extract (Fig. 5.9). MS/MS spectra obtained during analysis of the Oxford Clay extract indicates that the components typically lose up to four molecules of water and/or a molecule of C_3H_4O during CID of their $[M+H]^+$ (Table 5.2). Product ions formed from additional neutral losses of 94, 96, 112 and 114 Da were also observed for some of the Class T components. Not all of the neutral losses were observed in the MS/MS spectrum of each component, probably on account of the relatively low MS/MS ion currents obtained for each. Nevertheless, the neutral losses observed are consistent with those observed from the $[M+H]^+$ of GMT lipids **90**-**94** (see Sections 4.2.4.1 and 5.2.3.2). Furthermore, each Class T component has a $[M+H]^+$ at 56 *m/z* units less than a component in Class S, consistent with the difference in *m/z* values between GMGT lipids **27**-**31** and structurally related GMT lipids **90**-**94** (Chapter 4 and Section 5.2.3.2). The strong retention of Class T components, which elute after GMT **90**, is also consistent with the elution of non-

isoprenoid GMGT lipids (Class S) later than isoprenoid GMGT lipid **27**. Consequently, Class T components appear to represent GMT lipids comprising C_{66}-C_{68} (**T1**-**T3** inclusive) or C_{60} (**T4** and **T4'**) non-isoprenoid alkyl chains. The latter pair of lipids probably represent isomeric GMTs, each structurally related to GMGT **S8**, although the nature of the isomerism is unclear on the basis of the MS/MS evidence obtained.

5.2.5.4. $C_nH_{2n-4}O_6$ & $C_nH_{2n+2}O_6$ homologous series of non-isoprenoid ether lipids

Components belonging to a fifth and a sixth homologous series, with molecular formulae $C_nH_{2n-4}O_6$ and $C_nH_{2n+2}O_6$, respectively, were observed in the chromatograms for the VDG sediments (e.g. horizon III-1b; Fig. 5.29c-d) and categorised into two further classes (Class U and V). Unfortunately, no reliable MS/MS spectra were obtained for components in either class using either a conventional MS scan range (*m/z* 900-1500) or a truncated scan range (*m/z* 900-1200) specifically chosen to isolate Class V components for CID.

5.2.5.4.1. Class U – GDGT lipids containing two Cp rings

Ion chromatograms selecting for ions with $[M+H]^+$ at 4 *m/z* units less than those of **J9**-**J10** indicated the presence of **U1** and **U2** in the VDG extracts (Fig. 5.29c), at retention times and $[M+H]^+$ consistent with GDGT lipids **41** and **40**, respectively. Consequently, the two components are tentatively assigned to these structures.

5.2.5.4.2. Class V – GTGT lipids

Ion chromatograms for *m/z* values at two units greater than those of the $[M+H]^+$ of **J6**-**J12** exhibited a number of peaks relating to components **V1**-**V7**. The latter elute significantly earlier than the Class J components and were obscured beneath the peaks for isoprenoid lipids **2** and **14** in the extracts examined (Fig. 5.29d). Notably, each Class V component eluted later than GTGT **3**. The retention times of the Class V components relative to **3** and Class J components, coupled with the $[M+H]^+$ observed for each suggest that they may represent GTGT lipids, structurally related to non-isoprenoid GDGT lipids **J6**-**J12**, and are assigned accordingly.

5.2.5.5. Other non-isoprenoid lipids

Although there are a number of other components that may relate to tetraether lipids or close structural relatives containing non-isoprenoid alkyl chains in the ion chromatograms from VDG sediments, Oxford Clay sediments and from the greenhouse soil, the majority are not discussed in the interests of brevity. Three additional series of structurally relevant components identified in the samples, as exceptions, warrant further discussion.

5.2.5.5.1. Non-isoprenoid GDD lipids in VDG sediment

Extracts of two horizons from the VDG sequence, X-1 and VII-1b, were analysed *via* LC-MS/MS using a scan range of *m/z* 50-900 in order to identify potential GDD lipids. In addition to **1** (see Section 5.2.4.3), an homologous series ($C_nH_{2n+2}O_3$) of lipids with $[M+H]^+$ at nominal *m/z* 485-597 were also observed in each chromatogram, eluting a few minutes after the isoprenoid lipid (e.g. VDG X-1; Fig. 5.31). The components were most prominent relative to **1**, in the chromatogram for horizon X-1. Given the early elution of the components and the $[M+H]^+$ observed in each case, the components most likely represent GDD lipids containing a summed total of 28-36 carbons within the alkyl chains. Unfortunately, the MS/MS spectra of the components were invariably weak and no firm conclusions regarding the number of carbons on each specific alkyl chain within the structures could be drawn. The presence of GDD lipids in extracts from the other environmental samples was not investigated.

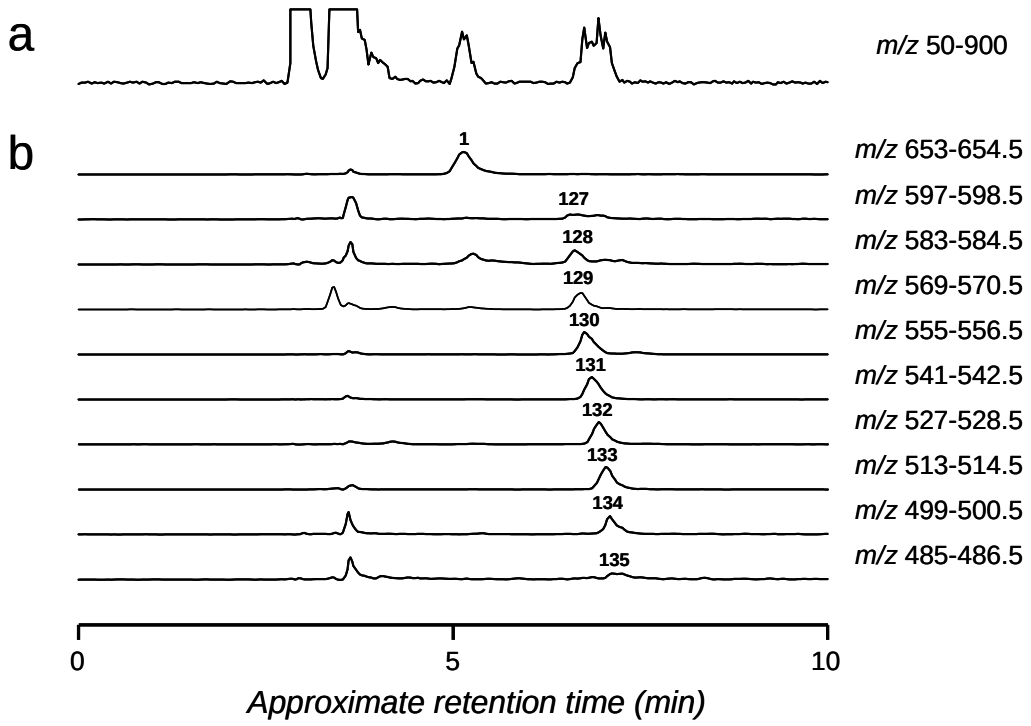


Fig. 5.31. Partial HCT LC-MS ion chromatograms for a lipid extract from VDG horizon X-1: a) base peak chromatogram; b) individual ion chromatograms, highlighting the presence of archaeol (**1**) and a series of non-isoprenoid GDD lipids (**127**-**135**).

5.2.5.5.2. Non-isoprenoid triether tetrols in Oxford Clay sediment

Components **136** (*m/z* 1096.1), **137** (*m/z* 1082.1), **138** (*m/z* 1068.1), **139** (*m/z* 1054.1) and **140** (*m/z* 1040.0) in the base peak chromatogram for the extract from S90-11 Oxford Clay sediment elute immediately after putative isoprenoid triether tetrols **113'** and **113** (Fig. 5.9). Each was also accompanied by an early eluting isomer, **136'**-**140'**. Lipids **136**-**140** and their isomers exhibit $[M+H]^+$ shifted by +18 *m/z* units from non-isoprenoid lipids **J6**-**J10**, also observed in ion chromatograms for the S90-11 sample. This mirrors the difference in *m/z* values observed between the $[M+H]^+$ of isoprenoid triether tetrols **112**/**112'** and **113**/**113'** and their proposed GDGT counterparts **2** and **24**/**24'**/**24''**. Weak MS/MS spectra were obtained for each component, in which each lipid is observed to lose 1-3 molecules of water and/or C_3H_4O during CID of the $[M+H]^+$. Unfortunately, MS/MS ion currents were insufficient for assignment of the alkyl chains to be made. On the basis of the

retention times and $[M+H]^+$ observed, **136-140** and **136'-140'** are tentatively assigned as non-isoprenoid triether tetrols, potentially derived from GDGT lipids **J6-J10**.

5.2.5.5.3. Polar non-isoprenoid lipids in greenhouse soil

Several prominent components were observed to elute after the IPA composition of the eluent had exceeded 2% during LC-MS/MS analysis of the lipid extract from the greenhouse soil (Fig. 5.32). These include components **141** (*m/z* 1035.7), **142** (*m/z* 1021.7), **143** (*m/z* 1081.7) and **144** (*m/z* 1067.7) which were not observed in the extracts from aquatic sediments. Although the nature of these components is unclear, the MS/MS spectra of each (Table 5.5) hints at an ether lipid structure based on a non-isoprenoid carbon backbone.

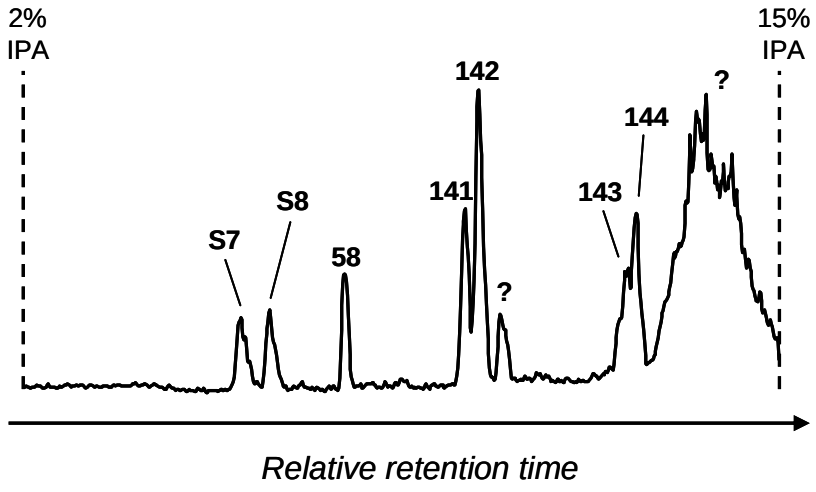


Fig. 5.32. Partial HCT base peak chromatogram (*m/z* 900-1500) of the lipid extract from a greenhouse soil, highlighting components eluting after the proportion of IPA in the mobile phase had exceeded 2%.

Table 5.5. HCT MS data (measured $[M+H]^+$ *m/z* values) and MS/MS data (nominal *m/z* values of product ions) of novel polar lipids in the lipid extract from the greenhouse soil.

Lipid	$[M+H]^+$	Product ions observed in MS/MS [a]
141	1035.7	1017,**999**,981,963,945,943,925,923,907,905,889; 563,549,545,531,527,525,513,511,509,507,505,493,491,489,475,473,471,457,449,431,417
142	1021.7	1003,**985**,967,949,931,929,911,909,893,891,875; 549,531,529,513,511,495,493,491,475,473,457,435,417
143	1081.7	1063,1049,**1031**,1013,1007,989,975,957; 617,603,599,595,585,581,539,525,521,507,489
144	1067.7	1049,1035,**1017**,999,993,975,961,943; 603,585,581,567,525,507,493,475,465,433

[a] Product ions listed in bold face represent the base peak in MS/MS.

5.2.5.6. GC and GC-MS of ether-cleaved hydrocarbons

In order to attempt to obtain more information regarding the structures of the non-isoprenoid components in the VDG sediments, the polar lipid extract from the VDG horizon in which these components were most prominent, VDG X-1, was subjected to ether cleavage with HI-$LiAlH_4$. The hydrocarbon fraction liberated was subsequently examined by GC and GC-MS. The chromatograms for each exhibit a number of peaks relating to hydrocarbons containing 14-40 carbon atoms (e.g. GC-MS chromatogram; Fig. 5.33). These include an extensive series of *n*-alkanes, iso- and anteiso-branched C_{15} and iso-branched C_{16} chains, 7-methylhexadecane (7Me-C_{16}), 1-cyclohexylundecane (ωCh-C_{17}), squalene, phytane (**I**) and acyclic biphytane (**II**). A $C_{14:1}$ chain containing a double bond or ring was also tentatively identified. An *n*-C_{29} chain was the only discernable chain in the chromatograms to contain between 29-38 carbon atoms.

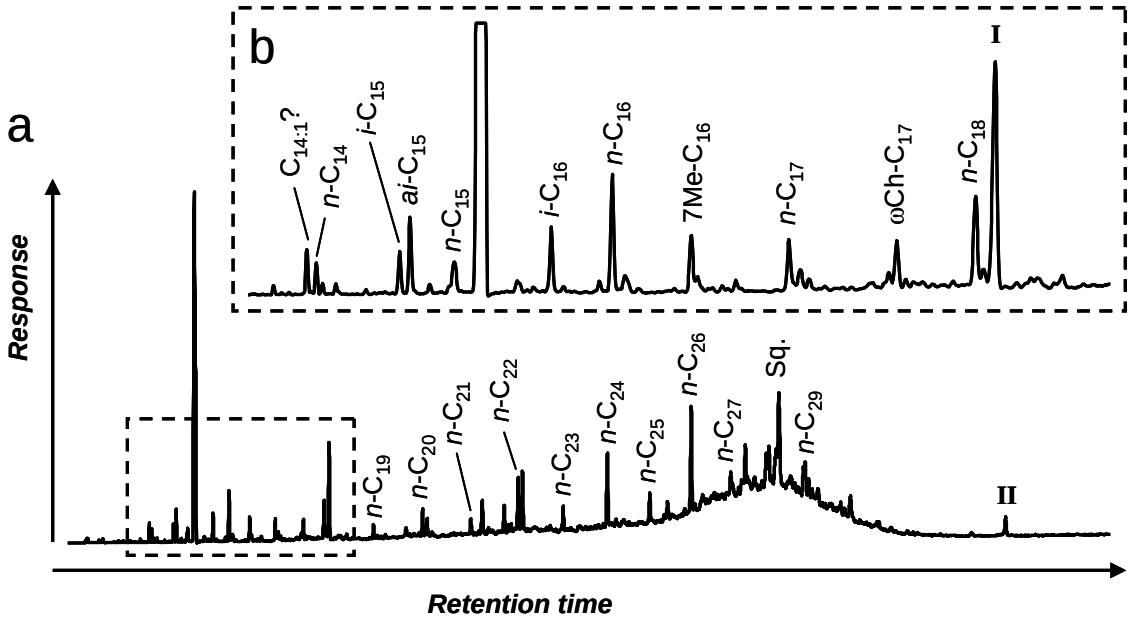


Fig. 5.33. a) GC-MS total ion chromatogram of the hydrocarbon fraction generated following treatment of a lipid core extract from VDG X-1 sediment with HI-$LiAlH_4$; b) shows an expansion of a region of the chromatogram. In both cases, only unambiguously assigned alkanes and alkenes are labelled. Sq. = squalene.

5.2.5.7. Significance

Non-isoprenoid GDGT lipids **34**-**38** and **40**-**41** were identified as components of all of the aquatic sediments, but were particularly prominent in the greenhouse soil extract. These observations are in accordance with previous studies in which **34**-**38** and **40**-**41** have been identified in marine (Schouten *et al.*, 2000; Hopmans *et al.*, 2004), lacustrine (Schouten *et al.*, 2000; Blaga *et al.*, 2009; 2010; Damsté *et al.*, 2009) and river (Herfort *et al.*, 2006b; Kim *et al.*, 2006; 2007) sediments but are typically found to be most prominent in samples from continental soils (Hopmans *et al.*, 2004; Weijers *et al.*, 2006b; 2007c) and peats (Schouten *et al.*, 2000; Weijers *et al.*, 2006a). Lipids **39** and **42** have also been identified in similar environments but are typically minor contributors to the total GDGT lipid (Weijers *et al.*, 2010), perhaps explaining their absence in any of the LC-MS/MS chromatograms generated in the present study. The MS/MS spectra of **34**-**35**, **37** and **40**, as isolated from the greenhouse soil, are consistent with their reported structures (Damsté *et al.*, 2000; Weijers *et al.*, 2006a). The MS/MS spectra obtained for **34**-**35** from the aquatic

sediments, on the other hand, show additional peaks which cannot be attributed to the proposed parent structures, possibly indicative of co-eluting isobars. It is important that these isobars are identified given that **34** and **35** are used in both BIT and MBT/CBT indices for palaeoenvironmental reconstruction (see Sections 1.4.2 and 1.6.2). If other isomers of **34** and **35** are present in aquatic sediments that are not indigenous components of soil, both BIT and MAAT reconstructions made *via* interpretations of the lipids in the sediment could be biased.

In addition to **36**, which has a GDGT[31_0,31_0] structure, a co-eluting isomer, **36***, with a GDGT[30_0,32_0] structure was identified in the MS/MS spectrum obtained from $[M+H]^+$ at *m/z* 1049.7 in the greenhouse soil. Similar observations were made during LC-MS/MS of the aquatic sediments (i.e. the MS/MS spectra of **J8** and **J8*** from the VDG samples). In the original characterisation of lipids **34**-**36**, the structures of GDGTs **34** and **35** were confirmed by ^{13}C and ^{1}H NMR and supported by GC-MS analyses of the liberated hydrocarbon chains, with methyl-branched C_{30} and C_{31} chains being identified (Damsté *et al.*, 2000). No hydrocarbon containing 32 carbon atoms was reported following GC-MS analysis of the ether cleavage products generated from the soil and peat lipid extracts initially used (Damsté *et al.*, 2000; Schouten *et al.*, 2000). As such, GDGT **36** was suggested to contain two C_{31} chains solely by association from the *m/z* value of its $[M+H]^+$ and LC-MS retention time relative to those of **34**-**35**. Recently, HI-$LiAlH_4$ degradation of a GDGT-enriched extract from a peat generated the C_{30} and C_{31} hydrocarbons assigned for GDGTs **34**-**36**, but also yielded a C_{32} chain, 5,13,16,24-tetramethyloctacosane (Fig. 5.34a; Weijers *et al.*, 2010). It was suggested that the latter component may originate from a GDGT lipid containing this chain and a 13,16-dimethyloctacosane as its ether-bound hydrocarbons, an assignment supported here by the identification of **36***, which is tentatively assigned the same structure (Fig. 5.34b). The presence of a second isomer may explain the broad and/or asymmetric peak shapes observed for **36** in previous LC-MS studies (Tierney *et al.*, 2009; Weijers *et al.*, 2010). The assignment of **36*** may also have implications for GDGT lipids **39** and **42**, originally assumed to contain two C_{31} chains; it is possible that these structures have isomeric forms which contain one C_{30} and one C_{32} chain.

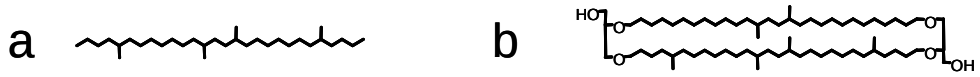


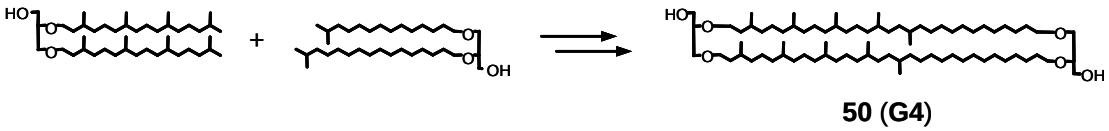
Fig. 5.34. a) The structure of a C_{32} methyl-branched hydrocarbon isolated from a GDGT-enriched fraction of a lipid extract from a peat (Weijers *et al.*, 2010); b) the tentative structure of GDGT **36***/**J8***.

A myriad of other non-isoprenoid lipids were detected *via* LC-MS analyses of the environmental extracts, particularly in the VDG extracts. These include a series of GDGT lipids, classified into five categories (Classes G-J, R and U) on the basis of the retention times and the molecular formulae inferred from the $[M+H]^+$ observed in each case. Many of the GDGT components were confirmed to contain C_{30}-C_{37} alkyl chains by LC-MS/MS, with others tentatively suggested to contain C_{29} or C_{38} chains on the basis of the elution order and their $[M+H]^+$ *m/z* values. Apart from **34**-**38** and **40**-**41**, the presence of which has already been addressed, only three other GDGT lipids containing non-isoprenoid alkyl chains have previously been identified; **44**, **45** and **50** (see Section 1.3.2). Components **44** (GDGT[30_0,32_0]) and **45** (GDGT[32_0,32_0]) have only been identified in the thermophilic eubacterium *Thermotoga maritima* (Damsté *et al.*, 2007). Structure **44** could, in principle, account for some or all of component **J8*** (along with/instead of **36***) observed in the chromatograms and **45** could similarly account for component **J6**. Given, however, that the ambient temperature at which the soil was collected and the probable temperature of northern Italy during the Messinian (Fauquette *et al.*, 2006) are both substantially lower than the viable growth range of *T. maritima* (55-90^{o}C; Huber *et al.*, 1986), an alternative, mesophilic source for **44** and **45** would be required. Lipid **50**, which has a GDGT[35_0,37_0] structure, has been observed previously in extracts from VDG sediment (Schouten *et al.*, 2000) and, consequently, accounts for component **G4**. This lipid is of particular significance on account of its mixed archaeal-eubacterial structural features (see Section 1.3.2 and discussion below).

It has been suggested that non-isoprenoid GDGT lipids are most probably formed from the coupling of two diether lipids in analogous fashion to the coupling of units of GDD **1** to make GDGT **2** in the Archaea (Schouten *et al.*, 2000; Weijers *et al.*,

2006a; Damsté *et al.*, 2007). If this were to be the case, the lipid cores of the GDD precursors might also be present in the VDG lipid extracts. By utilising LC-MS/MS with an extended scan range, ten GDD lipid cores were identified in the VDG horizon X-1 extract. These consisted of archaeol (**1**) and lipids **127**-**135**, which contain 4-12 fewer carbon atoms. The complete set of possible pair-wise combinations of the GDD components which were identified is sufficient to account for all Class G-J GDGT lipids present in the extract. Components **G1**-**G3** and **H1**-**H2** cannot be formed by combination of any two of **127**-**135**, implicating **1** as one of the GDD coupling partners in each case. Coupling of **1** with each of **128**-**135** would, in turn, account for all Class G-H components. Interestingly, the GDD series is truncated at **135**, which contains a total of 28 carbon atoms in the etherified chains. Coupling of this lipid to **1** would generate a GDGT lipid containing 68 chain carbon atoms in total, corresponding to the inferred number of chain carbon atoms at which Series H truncates (i.e. in **H7**). These observations provide preliminary evidence that each Class G and H lipid may be formed *via* the coupling of **1** with a non-isoprenoid GDD lipid and supports the structure of **50** shown in Fig. 1.9 with respect to the orientation of the alkyl chains. This would, by inference, imply that **50** is formed *via* coupling of archaeol to a GDD lipid containing *i*-C_{15} and *i*-C_{17} chains (Scheme 5.3). A GDD with the latter structure has yet to be conclusively identified in environmental samples or in eubacterial cultures (see Table 1.1). The $[M+H]^+$ of components **J3**-**J12** indicate that these lipids could not have originated from a coupling of **1** to any of **127**-**135**, but that they could have been formed by coupling of two of the non-isoprenoid GDD lipids. While **J1** and **J2** could be formed from **1** as one of the coupling partners, this is not a pre-requisite. As such, Class J components are tentatively suggested to form from coupling of two of **127**-**135**, in agreement with the route to **34** (i.e. **J8**) proposed previously (see Section 1.3.2; Weijers *et al.*, 2006a). Although it should be appreciated that the assignments made for Class G-J components are tentative, on the premise that the GDGT lipids can only be formed from GDD lipids identified in the fractions, Class G and H components would both exhibit some isoprenoid character, whereas Class J lipids would have none. This distinction may explain the greater chromatographic resolution of Class H from Class J components than of Class G from Class H components. The assignments, if correct, would also have important implications for

lipid biosynthesis given that the structures imply that lipids containing archaeol, the definitive archaeal lipid core, can be coupled to a number of different lipids which have an inferred eubacterial origin (Schouten *et al.*, 2000).



50 (**G4**)

Scheme 5.3. Putative biosynthetic route to formation of GDGT **50** (**G4**). Although the GDD precursors are shown to be saturated and dephosphorylated, the coupling could conceivably involve unsaturated and/or phosphorylated molecules.

In order to attempt to further elucidate the structures of the GDD and GDGT lipids, the ether-bound alkyl chains within the structures were liberated and the resultant hydrocarbons investigated *via* GC-MS analysis. Many of the smaller hydrocarbons (C_{14}-C_{20}) identified in the degradation mixture have been found previously as component chains in non-isoprenoid GDD and GMM lipids (see Table 1.1; Section 1.3.2). Of the previously identified GDD structures, those which contain only the C_{14}-C_{20} chains could account for all but one of the GDD lipids reported for the X-1 extract. The sole exception, GDD **128**, is inferred to contain a total of 35 carbon atoms in the two component ether-bound chains. This lipid may conceivably comprise a hitherto unreported combination of the liberated chains, such as a *n*-C_{18} chain plus a *n*-C_{17}/7-Me-C_{16} chain or a *n*-C_{19} chain plus a *i*-C_{16}/*n*-C_{16} chain. Given the previous identification of GDD **49** (see Section 1.3.2), which contains one isoprenoid and one *n*-alkyl chain, it is also possible that **128** could comprise a combination of a phytanyl plus an *n*-C_{15}/*i*-C_{15}/*ai*-C_{15} chain. Notably, an *i*-C_{17} chain, expected given the requirement for a GDD precursor containing this chain in the proposed route to **50** outlined in Scheme 5.3, was not identified among the degradates.

Apart from a general restriction to the hydrocarbons identified during GC-MS, no firm conclusions regarding the likely combination of these chains in GDDs **127**-**135** can be drawn. Firstly, the number of possible combinations which could account for **127**-**135** is too great to allow for specificity in assignment of any of the components.

Secondly, some of the hydrocarbons in the mixture may possibly have originated from degradation of GMM lipids as opposed to from **127**-**135**. No GDD lipids containing double bond equivalents were identified in the VDG X-1 lipid extract in the initial 10 min section of the LC-MS ion chromatogram, although it is possible that such components may elute later and have been overlooked. Nevertheless, it is distinctly possible that the ωCh-C_{17} chain and putative $C_{14:1}$ chains identified among the degradates may not originate from GDD lipids. GMM lipids containing either hydrocarbon have not been identified previously, although a fatty acid containing an ωCh-C_{17} chain has been identified in environmental settings associated with methanotrophic consortia (e.g. Blumenberg *et al.*, 2004; Krüger *et al.*, 2008). The $C_{14:1}$ chain could also potentially originate from a fatty acid (e.g. Blumenberg *et al.*, 2004). Similarly, an origin for other hydrocarbons identified in the GC-MS chromatograms from related fatty acid structures, as identified previously in methane seeps (e.g. Hinrichs *et al.*, 2000), cannot be discounted. Phytane (**I**; C_{20}) is one of the ether-bound components of **1** and **3**, both of which occur in the VDG X-1 sample. Although **I** in the mixture of degradates is probably sourced from these lipids, is not clear if other GDD lipids provide an additional source. Unfortunately, the ether cleavage of the polar extract yielded no discernable alkyl chains of length C_{29}-C_{38} apart from *n*-C_{29}, which is most likely generated *via* degradation of the corresponding *n*-alkanol or 1,*n*-alkanediol during the reaction (Weijers *et al.*, 2010). Despite the elevated relative proportion of non-isoprenoid GDGT components in the X-1 extract by comparison to extracts from the other marls, GDGT **2** is the dominant tetraether core, with this lipid consequently being the major source of **II** during ether cleavage. Given the small peak observed for **II** in the GC-MS ion chromatogram, it is likely that the C_{29}-C_{38} alkanes, if present, would be at concentrations too low to allow their identification. Weijers *et al.* (2010) have suggested that hydrocarbons cleaved from ether lipids are often swamped by those formed from alcohols also present in the polar lipid fractions unless these fractions are cleaned up *via* further preparatory steps. Given the presence of the extended series of *n*-alkanes identified in the mixture of degradates from the X-1 polar lipid extract, this seems to be the case in the present study. As such, the precise structures of the chains in the Class G-J components, apart from those in **34**-**36** (**J8**-**J10**) and **50** (**G4**), remain elusive.

Other compound classes (Classes A-F, K-L and S-T) assigned by LC-MS/MS for the VDG extracts appear, for the most part, to be structurally related to the non-isoprenoid GDGT lipids, often containing the same combinations of alkyl chains. Furthermore, most of the different classes (Classes A and B apart) mirror the differences in the isoprenoid lipid core types (as discussed in Sections 5.2.3 and 5.2.4), with non-isoprenoid GDHT, GDH_2T, GMGT, GDT and GMT lipids all having been identified. Non-isoprenoid GTGT lipids (Class V) were also tentatively identified in the VDG extracts. The complete set of lipid structures identified in the VDG sediments appear to indicate that both archaeol and non-isoprenoid GDD lipids are being forced through the same biosynthetic pathways, initially yielding GDGT lipids (probably *via* GTGT lipid intermediates) which are, perhaps, subsequently appended to the other structural classes. There are several potential scenarios that may account for these observations:

Eubacterial progenitor(s) – it possible that one or more mesophilic eubacterium may have inherited or acquired the ability to produce lipid features typical of the Archaea. Such behaviour has already been established for some eubacterial isolates, which produce GMM, GDD and GDGT lipids (see Section 1.3). It is interesting to note that many of species in which lipids of this nature have been identified belong to orders which are thought to be deep-rooted within the eubacteria (e.g. the *Thermotogales* or *Aquificales*; De Rosa *et al.*, 1988b; Huber *et al*. 1992; 1996; Jahnke *et al.*, 2001; Damsté *et al.*, 2007). It is tempting to suggest that this indicates an evolutionary conservation of archaeal-type biosynthetic features in these eubacteria, inherited from the last common ancestor shared with the Archaea. The possibility, however, of lateral transfer of the genes responsible for ether lipid formation between the archaeal and eubacterial domains must also be considered. The genome sequence for *T. maritima* (Nelson *et al.*, 1999), one of the deep-rooted organisms which produces ether lipids (De Rosa *et al.*, 1988b; Damsté *et al.*, 2007), contains a significant number of archaeal-like genes, believed to have been acquired by lateral gene transfer events between the two domains as opposed to an ancestral inheritance. Lateral gene transfer may also explain why species from other, less basally branched eubacterial phyla (e.g. the *Planctomycetes* or *Proteobacteria*; Damsté *et al.*, 2002c; 2005; Rütters *et al.*, 2001; Ring *et al.*, 2006) can also produce ether lipids.

Irrespective of how the archaeal-like genes were acquired, a mesophilic eubacterium possessing these genes could, in principle, be capable of producing the Class J GDGT lipids and structurally-related lipids in other classes which were identified in the VDG sediment.

Archaeal progenitors – it is also possible that one of more mesophilic archaeon may have acquired the ability to synthesise non-isoprenoid chains. The hyperthermophilic archaeon *Pyrococcus furiosus* is known to produce straight chain, branched and unsaturated C_{12}-C_{26} fatty acids and C_{16}-C_{22} α,ω-diacids containing mostly even but also sometimes odd numbers of carbon atoms (Carballeira *et al.*, 1997). The components were not identified in free acid form during the studies of the organism, indicating their esterification into complex lipids. In addition, *n*-hexadecanol and *n*-octadecanol have also been identified as free alcohols in the neutral lipid fraction of *P. furiosus* (Nishihara *et al.*, 2000) and are possibly derived from the corresponding fatty acids. Straight chain fatty acids and alkanes were similarly identified in *M. thermautotrophicus* (Tornabene *et al.*, 1978). Although the variety in the carbon chains observed in these organisms are not sufficient to account for the variety of structures found in the VDG sediment, it is apparent that some archaea may have the ability to produce chains based on acetate units in addition to those based on isoprene units.

Precursor uptake - if some species of archaea were capable of passive or active uptake of non-isoprenoid precursors such as GDD lipids **127**-**135** from the natural environment, these precursors may be scrambled into an early stage of isoprenoid lipid biosynthesis and carried right through the process to generate the other lipid classes. Encouragingly, this would explain the unusual condensation of archaeol with non-isoprenoid GDD lipids, the apparent route to Class G-H components. Unfortunately, it is harder to reconcile this suggestion with the apparent absence of GMGT structures relating to Class G-H components in the VDG sediments.

Thus, it remains to be seen which, if any, of these routes to the non-isoprenoid ether lipids may have been in operation in the VDG basin although each hypothesis should be testable in the future (see Chapter 6).

Class A and B components appear to be methyl-etherified derivatives of Class G-H and Class J components, respectively. Although **51**, the methyl-etherified derivative of **2** described for MTH(ΔH) is an artifact (Section 2.2.4.2.5), its formation was attributable to nucleophilic substitution of a terminal hydroxyl with MeOH during harsh acidic methanolytic extraction of *Methanothermobacter* cellular material. The milder conditions utilised during extraction of the VDG sediment are unlikely to invoke similar substitution of the terminal hydroxyls in the non-isoprenoid GDGT lipids. More importantly, although **51** was found as a trace component in some of the VDG horizons (Section 5.2.3.4), no methyl etherified derivatives of any of the other prominent isoprenoid lipids were identified. Consequently, Class A and B lipids are almost certainly non-artificial and represent genuine components found within the sediment.

Non-isoprenoid GDT and GMT lipids and the tentatively assigned triether tetrols present in the aquatic sediments may be derived from GDGT or GMGT lipids in Classes G-J *via* removal of a C_3H_4O portion of a terminal glycerol group as discussed for the isoprenoid lipids in Section 5.2.3.5. Interestingly, whereas isoprenoid GDT **58** was identified in the soil extract, non-isoprenoid GDT lipids were not present in this sample. This is at odds with the prominence of GDGTs **34-36** with respect to **2** in the base peak chromatogram of the extract. Clearly, the controls on isoprenoid and non-isoprenoid GDT formation in the soil were different.

5.3. Conclusions

In these preliminary studies, which examined the ether lipids present in ancient aquatic sediments and in a soil, a huge variety of components were unveiled, the majority of which are novel. It is probable that a significant number have been overlooked during LC-MS analyses of similar environmental samples attempted in the past, either because they could not be confirmed as ether lipids using a single stage of mass spectrometric analysis, or were not detected in analyses that employed single ion monitoring to screen for reported tetraether lipid cores only (see Schouten *et al.*, 2007b). While there has been a steady supply of novel structures reported from contemporary samples, the laborious and time-consuming nature of the purification

and degradative processes necessary to determine structure have no doubt proved a hindrance to new structural elucidations. Furthermore, novel components found in trace quantities could be difficult to characterise *via* this route without extraction of a substantial mass of sediment. The low limit of detection and ability to provide structural information during separation which are afforded from analysis by LC-MS/MS has allowed the suite of novel lipids to be reported for the first time using conservative amounts of sediment (less than 35 g).

Isoprenoid GDD, GTGT, GDGT and GMGT lipids, including those containing cyclopentyl and cyclohexyl rings, were identified in each of the environmental samples, generating MS/MS spectra consistent with their reported structures. This included three GDGT isomers containing two Cp rings, one of which would have been overlooked had LC-MS analysis (i.e. without online MS/MS) been employed. Additional isoprenoid structures to be identified were GDT lipids and a GMT lipid, assigned to be triolic structural variants of known tetraether lipids in which part of a terminal glycerol group is absent. Although several possible routes to the triols were considered, including biogenesis, intra- or extracellular enzymatic modification, oxic degradation and thermal degradation, their origin remains unclear. A series of components which represent potential intermediates in degradation of GDGT to GDT lipids, denoted glycerol triether tetrols, were tentatively identified and hint at a hydrolytic or oxidative cleavage of the tetraethers. Since the GDT and GMT lipids were found to be minor components in the lipid extracts from thermophilic archaea described in Chapters 2-4, a biosynthetic origin is also possible. Further work is required to confirm the structures of these lipids and to constrain their origins. To this end, studies using contemporary soils and sediments, in which the redox potential, temperature and microbiological diversity of the samples can all be directly monitored, should prove more enlightening. GDT lipids which are structurally related to ring-containing GDGT lipids used in both the TEX_{86} and BIT indices were identified. As such, if the triols do represent products of GDGT degradation, their identification could potentially have severe ramifications for the palaeoenvironmental reconstructions made using these indices previously.

Higher homologues of known isoprenoid GTGT, GDGT and GMGT lipids were identified in the extracts from the environmental samples, including structures appended with up to three additional carbon atoms, added to either the ether-bound chains or a glycerol moiety. It is difficult to assess the potential origin of such components at the current time, given that chain-methylation was noted for members of both the *Euryarchaeota* and *Crenarchaeota* in Chapters 3 and 4. Furthermore, additional work is required to provide more rigorous characterisation of the components. Nevertheless, the identification of these lipids could be of chemotaxonomic value in future palaeoecological reconstructions. Given that a potential link between temperature and chain methylation was noted in Chapter 3, the possibility that the tetraether homologues provide a record of *in situ* temperature should also be investigated.

The structural variety in non-isoprenoid GMM and GDD lipid cores identified in the past is extensive, whereas only twelve non-isoprenoid GDGT lipid cores have been noted previously (see Section 1.3.2). The LC-MS/MS analyses of the environmental samples described here have allowed over 100 non-isoprenoid tetraether lipids or related structures to be identified. Although the assignment of specific structural features remains tentative, this highlights the wealth of ether lipids of potential value as biomarkers which have yet to be characterised fully. Preliminary evidence based on the GDD lipids identified suggests that some GDGT components of the VDG sediment may have been formed *via* coupling of an archaeal GDD lipid with a non-isoprenoid GDD lipid, with precedent for the occurrence of the latter only existing in eubacteria. This identification implies a highly unusual biosynthetic pathway that may involve representatives from one or both of the domains. A more in depth profile of the lipids of the VDG, including a more thorough investigation of GDD lipids within the distinct sediment horizons may be necessary if these unusual lipids are to be understood.

LC-MS/MS analysis of the Oxford Clay sediment using HPLC method B was insufficient to elute all ether lipid components from the analytical column in 90 min. Although use of method C led to elution of additional components within 105 min, this method may still be inadequate; it is possible that other polar ether lipid cores of

environmental significance are still retained despite the increased polarity of the elution program employed. As such, the set of ether lipid structures reported should be considered to represent the minimum variety that may be found in such samples.

Chapter 6

Conclusions and future work

6.1. General conclusions

The primary aims of the work presented in this thesis were to develop LC-MS/MS methodology for identification of tetraether lipid cores and to investigate the feasibility of the technique for use in detailed lipid profiling. The results presented were segregated into three key areas; method development (Chapter 2), application of the method to lipids extracted from archaeal cellular material (Chapters 3 and 4) and subsequent application to those extracted from environmental samples (Chapter 5).

MS/MS analysis was incorporated into LC-MS methodology developed previously (Hopmans *et al.*, 2000) and was tested using lipids extracted from the archaeon *M. thermautotrophicus* (strain ΔH), which contains a relatively simple membrane lipid profile, in the place of an authentic isoprenoid tetraether lipid standard. Tetraether core lipids in the extract were dissociated effectively during CID, leading to the formation of several product ions for each lipid. Attempts to obtain reliable multistage tandem mass spectra of the lipid cores were, on the whole, unsuccessful. Even in cases where MS^3 spectra were obtained, little additional information regarding structure was gleaned over that inferred from the MS/MS spectra. This suggests that there is little potential value in attempting to extend LC-MS/MS to further stages of tandem mass spectrometric analysis in the future. MS/MS spectra were generated on both ion trap and Qh-FT-ICR instruments, suggesting that the technique, either used in a standalone fashion or coupled to HPLC, should be adaptable to other tandem instruments.

Product ions formed from losses of water and simultaneous loss of C_3H_4O (most likely representing a molecule of propenal) were amongst those observed in the MS/MS spectra. These ions are identical to fragment ions observed in conventional APCI MS spectra of the tetraether lipids (Hopmans *et al.*, 2000) which have been used as characteristic ions for lipids of this structural type. The extra ions observed in the MS/MS spectra prove to be of further characteristic value, particularly those formed from losses of an alkene or diene, which carry information regarding the number of carbon atoms and number of double bond equivalents in the ether-bound

alkyl chains. The general distribution of product ions in the MS/MS spectra is also characteristic for the type of lipid core dissociated. Consequently, LC-MS/MS analysis retains the advantageous low limits of detection and relatively high throughput afforded by LC-MS, but simultaneously provides specific information regarding structure that would previously have required isolation of each component and chemical degradation.

In-source CID MS and PAN MS/MS were both identified as alternative procedures for generating an extensive array of product ions from tetraether lipid core structures. Although these techniques allowed identification of additional ions to those observed in the conventional MS/MS experiments, the extra ions formed are currently of little characteristic value. Nevertheless, the identification of these low *m/z* ions indicates that both C-C bonds and C-O bonds can be broken during CID of the tetraether cores. The dissociation events which lead to both types of bond cleavages remain unclear, although it is apparent that distinct events can lead to product ions of the same *m/z*. In particular, some product ions may be formed by loss of an individual neutral molecule (e.g. 1-hydroxybiphyt-31-ene) or *via* simultaneous loss of two or more smaller molecules (1,31-biphytadiene and H_2O).

Application of LC-MS/MS to the lipid extracts obtained from thermophilic archaeal cellular material proved to be largely successful, with lipid cores identified previously being similarly identified using this method. Additionally, a number of novel archaeal isolates were also profiled for the first time using LC-MS/MS. The MS/MS spectra of the majority of the lipids are consistent with the structures reported for the components previously, confirming, in particular, that the technique can differentiate ether-bound chains containing differing numbers of Cp rings. The MS/MS spectra obtained for some lipids, on the other hand, suggest concealed lipid core isomerism. Although lipid isomerism in thermophilic archaea has been alluded to previously (e.g. Hopmans *et al.*, 2000; Pitcher *et al.*, 2010), the data obtained herein has both reinforced and extended these initial observations, with the nature of the structural isomerism for each component being constrained by the MS/MS spectra obtained. In addition, structures containing homologated isoprenoid chains (i.e. C_{41} or C_{81} chains) were identified for the first time. Preliminary data suggest that

these lipids are formed at a late stage during tetraether lipid biosynthesis, possibly by an unprecedented methylation at a saturated carbon atom. Although the functional role of the methylated tetraether lipid cores within the archaeal membrane remains enigmatic, results are presented that suggest expression of these lipids may be growth temperature dependent. The methylated components have been identified throughout the archaeal domain indicating that they represent a widely occurring modification to GDGT lipids. Although the $GDGT_M$ lipid cores may have limited use as phylum-specific biomarkers, they may be indicative of particular growth conditions. A number of other rarely reported and novel lipid cores were identified within the suite of organisms investigated, some of which were exclusive to a single organism. Encouragingly, these observations imply a potential role of the source-specific lipids as chemotaxonomic markers for the progenitor species.

Lipid extracts from environmental samples are inherently more complex than those from the microbes, with some samples containing over 100 ether lipid cores. These include an array of lipids inferred to be of archaeal origin and an even greater variety which may originate from eubacterial organisms. The variety identified far exceeds the ether lipid variety reported previously for organisms from either domain, indicating that only a narrow subset of the lipids of this nature produced in natural environments have been fully characterised to date. The lipids identified were classified on the basis of relative retention time, the $[M+H]^+$ observed in MS and by characteristic product ions identified in the MS/MS spectra. Components which may represent transformation products of tetraether lipids were identified among the environmental lipid set, although the precise route to their formation remains unclear. Nevertheless, their formation could have implications for the use of TEX_{86} and BIT indices to reconstruct palaeoenvironments from ancient sediments. Isoprenoid tetraether lipid isomerism and homologation were again noted in the environmental samples, indicating that mesophilic organisms can also vary their membrane lipid cores in this manner. Although the results presented in this thesis do not confirm the full structures of any of the new components, particular structural features of each lipid were determined from the LC-MS/MS data, allowing reasonable structures which are consistent with the data to be postulated. More

importantly, this initial screening highlights the need for further exploration of the lipid diversity in the natural environment.

To summarise, LC-MS/MS is a useful tool for characterisation of ether lipid cores, providing multi-dimensional data which can be used to gain insight into the fragmentation behaviour of known lipids or to contribute to the characterisation of novel lipid structures.

6.2. Conclusions regarding structural identification of ether lipid cores

Although full structural identification is not provided by LC-MS/MS, the studies described in this thesis, supported by the body of literature describing LC-MS profiles of the lipids from other archaea, eubacteria or environmental settings, provide a number of rules of thumb that can aid in identification of ether lipid structural features (classifications exemplified in Fig. 6.1).

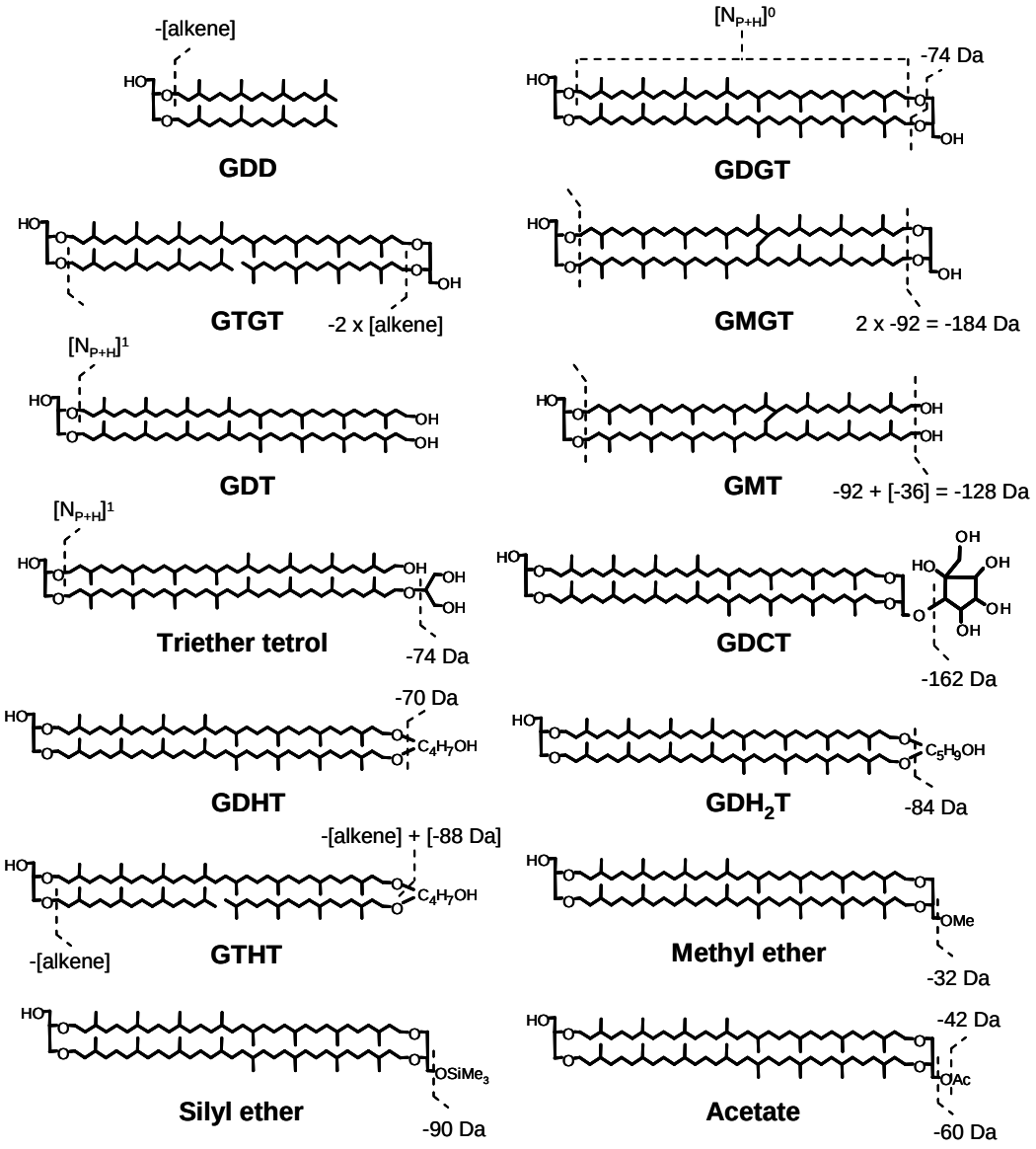


Fig. 6.1. Structural classifications of ether lipids discussed in the text (exemplified using structures which contain no rings) and their characteristic neutral losses.

6.2.1. Rules for relative retention time

1) The number of free hydroxyl groups in the structure is the dominant factor which influences retention time (Chapters 3 and 5). In general, the greater the number of free hydroxyl groups, the longer the retention; the retention time of polyols (GDCT) > tetrols (triether tetrols) > triols (GMT; GDT) > diols (GDGT; GTGT; GDHT; GDH_2T) > mono-ols (GDD), although other structural modifications can lead to some crossover of different lipid classes with respect to retention time.

2) Modifications to the polar termini in the lipids also lead to substantial changes in retention time. Appendages which increase the polarity of the moiety lead to later elution, while those that reduce the polarity lead to earlier elution. Adding a carbon atom (GDHT) or two carbon atoms (GDH_2T) to the propyl backbone of a glycerol group leads to earlier elution (Section 5.2.4), while adding an alkyl (i.e. Sections 2.2.4.2.5 and 5.2.5.2.3) or acyl (Section 3.2.3.3) group to the free hydroxyl in a glycerol leads to significantly earlier elution. Adding a carbonylic oxygen to convert an ether linkage into an ester linkage leads to later elution (Damsté *et al*., 2007).

3) The degree of conformational freedom of the etherified alkyl chains also appears to affect retention time; less constrained lipids elute earlier. Thus, GTGT lipids (not macrocyclic) elute before GDGT lipids (macrocyclic) which, in turn, elute before GMGT lipids (bridged macrocyclic) (Sections 2.2.4.1, 5.2.5.2 and 5.2.5.3) and GDT lipids (not macrocyclic) elute before GMT lipids (bridged) (Sections 5.2.3.2, 5.2.5.2 and 5.2.5.3).

4) The nature of the etherified alkyl chains also influences retention time although, as a systematic change (i.e. one ring or one carbon atom different), this is a lesser effect. Structures modified with more incorporated Cp rings elute later (Sections 3.2.4.2, 4.2.2.1, 5.2.3.1 and 5.2.5; Hopmans *et al*., 2000; Schouten *et al.*, 2000). Incorporation of a Ch ring has less of an effect on retention time (Section 5.2.3.1; Weijers *et al*., 2004; Pitcher *et al*., 2010). Structures containing more carbon atoms in the etherified chains, in general, elute earlier (Sections 3.2.3.3, 3.2.4.2, 4.2.3.1, 5.2.4 and 5.2.5).

6.2.2. Rules for APCI MS spectra

1) For di- and tetraether lipids, the $[M+H]^+$ is the base peak ion (Section 2.2.3.1 and Hopmans *et al*., 2000). For mixed ether-ester (i.e. triether-monoester, diether-diester or monoether-triester) lipids, the $[M+H-18]^+$ fragment ion is the base peak ion (Damsté *et al.*, 2007).

2) Tetraether lipids produce fragment ions at $[M+H-18]^+$ and $[M+H-74]^+$ in MS (Hopmans *et al*., 2000), although the relative prominence of these may, to some extent, be instrument dependent (Section 2.2.3.1).

3) GDHT lipids produce fragment ions at $[M+H-18]^+$ and $[M+H-88]^+$ in MS (Sections 5.2.4.4 and 5.2.5.2.4). GDH_2T lipids produce fragment ions at $[M+H-18]^+$ and $[M+H-102]^+$ in MS (Sections 5.2.4.6 and 5.2.5.2.5).

4) GDCT lipids may produce a number of fragment ions during MS, including $[M+H-162]^+$ (Section 3.2.4.3).

6.2.3. Rules for MS/MS spectra

1) During tandem MS, GMGT, GDGT and GTGT lipid cores produce product ions in one, two and three regions of the MS/MS spectra, respectively (Section 2.2.4.2). It is important to note that, while the number of regions is a characteristic spectral feature of each core type, the number of regions observed in the spectrum of an unknown lipid remains a necessary but not sufficient condition for assignment of a *particular* core type; the *m/z* values of all product ions observed must also be consistent with the proposed assignment. For example, while GTGT lipids categorically dissociate to give three major regions of product ions in the MS/MS spectra, identification of three regions of product ions in the MS/MS spectrum of an unknown does not necessarily confirm a GTGT structure for the lipid.

2) GDGT lipids can lose 1-3 molecules of water and/or 0-1 molecules of C_3H_4O from the glycerol termini without loss of an alkyl chain (Sections 2.2.4.2 and

2.2.5.2). GTGT lipids can lose 1-2 molecules of water without loss of an alkyl chain, depending on both the instrument used and lipid structure (Sections 2.2.4.2 and 2.2.5.3). GMGT lipids lose an extensive array of small molecules, including 1-4 molecules of water and/or 0-2 molecules of C_3H_4O from the glycerol termini (Sections 2.2.4.2 and 2.2.5.4). For dialkyl and trialkyl tetraether lipids, product ions formed from loss of mono-etherified chains as alkenes or di-etherified chains as dienes (i.e. $[N_{P+H}]^0$) are those most useful for assignment of the structures of the etherified chains. The two etherified alkyl chains in an individual GDGT lipid can eliminate as $[N_{P+H}]^0$ with differing favourabilities during CID (Section 3.2.4.4).

3) Characteristic neutral losses from other lipid classes include:

GDD	=	-[alkene] (Section 3.2.3.3)
GDT	=	-18 Da; -36 Da; $[N_{P+H}]^1$ as the smallest molecule to bear an ether-bound alkyl chain (Sections 5.2.3.2 and 5.2.5.2.6)
GMT	=	-72 Da; -128 Da as the heaviest favourable neutral loss (Sections 5.2.3.2 and 5.2.5.3.3)
Triether tetrol	=	-18 Da; -74 Da; $[N_{P+H}]^1$ as the smallest molecule to bear an ether-bound alkyl chain (Section 5.2.3.3)
GDHT	=	-70 Da; -88 Da (Sections 5.2.4.4 and 5.2.5.2.4)
GDH_2T	=	-84 Da; -102 Da; -104 Da (Sections 5.2.4.6 and 5.2.5.2.5)
GTHT	=	-[alkene]; -[alkene] + [-88 Da] (Section 5.2.4.4)
Methyl ethers	=	-32 Da; -50 Da (Sections 2.2.4.2.5 and 5.2.5.2.3)
Silyl ethers	=	-90 Da; -146 Da; $[N_{P+H}]^0$ as the smallest molecule to bear an ether-bound alkyl chain (Section 2.2.10.2)
Acetates	=	-42 Da; -60 Da; -116 Da; $[N_{P+H}]^1$ as the smallest molecule to bear an ether-bound alkyl chain (Sections 2.2.10.3 and 3.2.3.3)

The structural cleavages which result in these losses are highlighted for the ether lipids in Figure 6.1.

6.3. Future work

6.3.1. Relating to results presented in Chapter 2

One of the principle limitations of the LC-MS/MS methodology used is the lack of reproducibility between analyses with respect to component retention times (both absolute and relative). Although attempts were made to minimise this effect, variability between consecutive analyses was not fully eradicated. The use of the Dionex rapid separation 3000 liquid chromatograph led to less significant variation, suggesting that some of the variance relates to operational factors (e.g. the accuracy of solvent proportioning by the liquid chromatograph). On account of the differences in polarity between hexane and IPA, small inaccuracies in the proportioning of solvents could lead to substantial differences in retention time. One possible modification which may limit this problem would be the use of pre-mixed solvents (Peterse *et al.*, 2009a), consisting of the hexane:IPA proportions at the onset (99:1) and end (98:2) of HPLC method A. The differences in polarity between each of the pre-mixed solvents is significantly less than the difference between hexane and IPA, leading to a less significant influence of proportioning errors on retention time. Alternatively, a change in the solvents and/or column chemistry used during LC-MS/MS could help to reduce or eliminate retention time drift. In particular, use of a more polar column chemistry, such as a non-modified silica or a diol column may allow for use of a more polar eluent, overcoming the reliance on high hexane eluent compositions. Separation by reversed-phase HPLC may also prove to be a prudent choice for future development. In any case, more robust retention times should lead to improved confidence in characterisation of ether lipid components.

Other potential areas for improvement to the methodology lie in the precursor selectivity for CID and the robustness of MS/MS product ion currents in both conventional and PAN MS/MS. The 3 *m/z* unit window utilised during each procedure leads to complication of the MS/MS product ion distribution due to the simultaneous trapping and subsequent CID of higher isotopomers and species with similar *m/z* to the intended precursor. Reduction in the size of this window should simplify the resultant MS/MS spectra although, currently, any attempt to reduce the

isolation width leads to the complete lack of selection of any precursor ions. If this problem is to be overcome, an understanding is needed as to why the threshold isolation width required for these components is necessary. The variation in relative abundances of product ions in MS/MS and PAN MS/MS spectra of individual components in consecutive LC-MS/MS analyses reduces the potential for use of the MS/MS spectra in selective structural characterisation. It is unclear as to why this occurs and, consequently, is difficult to remedy. In the case of PAN MS/MS, the variation may be the result of inappropriate choices made during optimisation of the "Activation" and "Decay" parameters for CID, which were selected to provide maximum MS/MS ion current for ions with $m/z < 400$. An alternative optimisation which focuses on reproducibility may prove to be more useful in future. Should the reproducibility of the MS/MS ion currents for conventional CID be improved, quantification of co-eluting isobars which dissociate to yield different product ions may be possible from the ions unique to each precursor. Should the reproducibility of PAN MS/MS be improved, on the other hand, more specific information regarding the composition of the ether-bound alkyl chains may be obtainable *via* a statistical approach which accounts for all product ions with $m/z < 400$.

It will also be important to investigate the mechanisms in operation during CID of tetraether lipids in further detail in order to provide a better understanding of the formation of the product ions observed. Use of synthetic or semi-synthetic tetraether standards containing modified glycerol groups, incorporated ^{2}H, ^{13}C or ^{18}O isotopic labels and/or alkyl chains of various length and containing various degrees of methyl branching may allow some of the product ions formed to be characterised more fully. This, in turn, may facilitate more specific inference of structural features on the basis of the product ions observed in the MS/MS spectra of unknown lipid cores.

6.3.2. Relating to results presented in Chapters 3 and 4

Given that novel lipid cores were identified in MTH(ΔH) and species of *Sulfolobus* which were not identified during previous lipid profiles of these organisms, it is evident that some lipid structures of potential taxonomic value may have been overlooked in the past. In order, therefore, to assess the potential of novel structures

identified in the thesis for use as biomarkers, a much larger census of the Archaea is required. Although each novel lipid identified will require further structural analysis to ascertain the full structure, this does not preclude the use of the lipids as biomarkers. Some of the assignments made (e.g. of isomers to known Cp ring-containing tetraether lipid cores) could be consolidated by preparation of each individual lipid, followed by degradative analysis. Other components were identified at levels that may be too low for this approach to be effective. Consequently, further characterisation of such components may require development of more informative mass spectrometric methods with respect to structural elucidation. In any case, an appreciation as to whether further isomeric forms of some of the structures are also produced is required. For example, whether each organism produces distinct regioisomers with the glycerol groups aligned either *syn* or *anti* to one-another, or produces isomers in which the incorporated Cp rings are located in alternative positions along the isoprenoid chains is of particular importance.

Three prominent hydrocarbons were detected during GC and GC-MS analyses of the ether cleavage products obtained from the lipid extracts from MTH(ΔH) and were assigned as phytane, biphytane and 13-methylbiphytane. Whilst phytane was assigned using both the molecular ion and characteristic fragment ions observed in the EI mass spectrum, the equivalent spectra for biphytane and 13-methylbiphytane were each devoid of the molecular ion expected for the chain. Consequently, assignment of the two components was made by comparison of fragment ions in the spectra to those noted for the hydrocarbons previously. Given the atypical elution order of the two hydrocarbons during GC, determination of the molecular mass of each would allow for more definitive assignment. Chemical ionisation (CI) is a technique often used to produce ions that have less residual energy than those produced by EI (de Hoffmann and Stroobant, 2002). As such, CI often leads to a lesser degree of ion fragmentation and, consequently, affords easier recognition of the molecular species. Use of this ionisation technique during GC-MS analysis of the ether-cleaved hydrocarbons may, therefore, be more appropriate for determination of their molecular masses. Alternatively, it has been shown that methyl thioethers formed from biphytane diiodides, themselves liberated *via* treatment of tetraether lipid extracts with HI, produce more prominent molecular ions in EI mass spectra

than are observed in the spectra for structurally-related biphytanes, generated using the same ionisation technique (Schouten *et al.*, 1998). Consequently, treatment of the biphytane diiodides formed from the MTH(ΔH) lipid extracts with NaSMe as opposed to $LiAlH_4$ may provide derivatives of the hydrocarbons that prove to be of greater use in the determination of their molecular masses.

More needs to be learnt regarding the biosynthesis of tetraether lipids; in particular, how caldarchaeol is formed from lipids containing an archaeol core, how Cp rings and additional methyl groups are incorporated into the isoprenoid chains and how the isoprenoid chains may be conjoined during formation of GMGT lipids. As more and more full archaeal genome sequences become available, the potential for identifying and targeting the genes responsible for lipid biosynthesis is increasing. Lipid profiling of mutant archaeal strains in which genes have been knocked out may allow the genes responsible for the different biosynthetic modifications to be identified. MTH(ΔH), *S. acidocaldarius* and *I. aggregans* each have full or draft genomes available and, between them, express nearly all of the lipid core types identified. Consequently, these three organisms may be the most suitable archaea for generating knockout mutations in the interests of investigation of lipid biosynthesis in the future.

The physicochemical role fulfilled by incorporation of Cp rings has been widely documented and attributed to adaptation of the membrane in response to thermal or pH related stress. The roles fulfilled by $GDGT_M$, GMGT and $GMGT_M$ lipids, on the other hand, remain speculative, particularly given that some of these components also contain Cp rings. Future studies in which model organisms (e.g. *I. aggregans* for GMGT lipids, *Pyrobaculu*m sp. AQ1.S2 for $GDGT_M$ lipids and MTH(ΔH) for $GMGT_M$ lipids) are grown under controlled conditions of varying temperature, pH and nutrient availability may help to shed light on why the structural modifications to caldarchaeol are incorporated. Biological replication in such studies will also be a prerequisite if stress-induced variation in the lipid core production is to be extricated from any natural variability between batch cultures grown under identical conditions.

The need for greater sample throughput is also evident from the studies described in Chapters 3 and 4. LC-MS/MS analysis using method B takes 115 min per sample, leading to a current throughput of 10-12 samples per day. LC-MS methods proposed by Schouten *et al.* (2007b) and Escala *et al.* (2007; 2009) both have significantly higher throughput, but do not screen for more polar lipid cores. An LC-MS/MS method which provides more rapid profiling of the lipids but does not sacrifice the detection of the more polar components or the quality of the MS/MS spectra obtained should prove to be advantageous, particularly during profiling of large numbers of samples. A move to cyano column chemistry and a smaller column particle size than that of the amino column used in the present study would be expected to lead to earlier elution of the tetraether components with little to no sacrificial loss in resolution. Consequently, use of a column containing a sorbent of this nature may be a good starting point for future developments. The necessity for column backflushing between analyses is a feature of the current analytical protocol that prolongs each analysis and which could be targeted during further method development.

GDCT lipids were retained on the HPLC column for the duration of LC-MS/MS analyses of lipid extracts from species of *Sulfolobus* which were performed using HPLC method B, with these components only being removed during the column backflushing procedure. These examples indicate that, although LC-MS/MS method B does allow for identification of polar components missed in previous LC-MS analyses, it does not allow for identification of all tetraether components. Clearly, the highly polar components either need to be separated from other tetraether lipids by flash column chromatography prior to LC-MS/MS or, otherwise, the LC-MS/MS method must be altered to include an additional polar flush which facilitates their elution. Notably, and as a further incentive for method development, no instrumental method which can separate and detect GDCT lipid cores containing differing number of Cp rings has been developed to date. Given the recent recognition of GDCT cores in sediments deposited at moderate temperatures (Sturt *et al*., 2004; Biddle *et al*., 2006; Lipp *et al*., 2008; 2009) and peats (Liu *et al*., 2010), such a development may be of importance for higher resolution lipid profiling for samples of this nature.

6.3.3. Relating to results presented in Chapter 5

As noted for the thermophilic archaeal lipid extracts, LC-MS/MS method B does not allow for detection of all tetraether lipid cores in extracts from environmental samples; some of the triether tetrols were not eluted using this method. Although LC-MS/MS method C did allow for identification of additional components, there is no guarantee that other polar lipid cores do not remain fully retained on-column following use of this method. Again, this highlights the need for further LC-MS/MS method development.

It would be desirable to screen for the components identified in the VDG and Oxford Clay sediments in contemporary environments in which molecular genetic studies can be employed simultaneously. This may allow the origins of particular components to be better constrained to certain phylogenetic clades. Furthermore, the conditions conducive to formation of the lipids, such as temperature, pH and redox potential can be readily measured in contemporary environments.

GDGT lipids which appear to contain ether-bound hydrocarbons which may represent a hybrid between those typical of archaea and those typical of eubacteria are highly unusual and warrant further investigation. Three hypotheses which may explain the formation of these components were outlined in Section 5.2.5.7. Each could potentially be tested by determination of the glycerol stereochemistry in the lipids. If the lipids are sourced from an archaeon, they may be expected to have an *sn*-2,3 stereochemistry at each glycerol terminus. If they are sourced from a eubacterium, they may be expected to, instead, have *sn*-1,2 glycerol stereochemistry at each end. If, on the other hand, the formation of the GDGT lipids is *via* uptake of a eubacterial diether precursor by an archaeon, followed by coupling with a lipid containing archaeol as the lipid core, the two glycerol groups in the structure may be expected to have opposite stereochemistries.

Chapter 7

Experimental and analytical procedures

7.1. General procedures

7.1.1. Solvents and reagents

All solvents used were analytical grade or higher (Fisher, Loughborough, UK). HPLC grade hexane obtained from this supplier contained a number of contaminants which are amenable to APCI MS and can complicate mass spectra obtained. Prominent ions which relate to the contaminants are typically found at nominal *m/z* 1325, 663, 647 and 603 in the spectra. Where necessary, the contaminants can be removed from the hexane either by distillation (70-72^{o}C at 1 atm) or by passing the solvent through a column of activated alumina. All reagents used were obtained from Aldrich (Gillingham, UK) unless otherwise stated. Cotton wool used for filtrations was defatted *via* Soxhlet extraction with DCM for 72 h and dried prior to use.

7.1.2. Glassware

In order to reduce the risk of cross-contamination, all of the glassware used was rinsed with laboratory grade acetone followed by deionised water, before being soaked in a 1% solution of Decon-90 (Decon; Hove, UK) for 12-24 h. Glassware washed in this manner was subsequently subjected to extensive rinsing with water, deionised water and laboratory grade acetone and subjected to a final rinse with analytical grade acetone before use.

7.1.3. Sample storage

In general, and where possible, archaeal cellular material was stored at -20^{o}C prior to extraction. Harvested *S. acidocaldarius* cellular material was stored at room temperature as a pellet-supernatant pair for approximately 1 year. Cell pellets of the New Zealand hot spring isolates (Chapter 4) were thawed for a period of approximately 1 week during transport to the UK, but were immediately refrozen upon arrival. VDG sediments and the soil sample were stored at room temperature prior to extraction, whereas Oxford Clay sediment was stored at -20^{o}C. Extracted lipid core material (and derivatives and degradates thereof) were stored at +4^{o}C.

7.2. Preparation of lipid extracts from archaeal cellular material

7.2.1. Growth of organisms

MTH(ΔH) (DSM 1053) was grown at 45°C or 70°C in mineral salt media (at pH 6) by James Chong (University of York, UK) as described previously (Guy *et al*., 2004), with cells aerobically harvested and frozen as described above.

Sulfolobus acidocaldarius MR31 (Reilly *et al*., 2001) was grown aerobically at 75°C in Brock's medium (Brock *et al*., 1972), supplemented with 0.1% tryptone, 0.2% xylose and 20 μg/mL uracil, by Charlotte Nunn (University of Bath, UK). The pH was adjusted to 3.0-3.5 with sulfuric acid. The resultant broth was lyophilised and stored as described above. *Sulfolobus shibatae* (DSM 5389) and *Sulfolobus solfataricus* P2 (DSM 1617) were grown at approximately 75°C in basal salt media (Brock *et al*., 1972), adjusted to pH 3.5 and containing (per litre) 0.28 g KH_2PO_4, 1.30 g $(NH_4)_2SO_4$, 0.25 g $MgSO_4.7H_2O$, 0.25 g $CaCl_2.7H_2O$. The medium also contained a 1% trace element solution comprising: 20.0 mg $FeCl_3.6H_2O$, 4.5 mg $Na_2B_4O_7.10H_2O$, 1.8 mg $MnCl_2.4H_2O$, 0.22 mg $ZnSO_4.7H_2O$, 0.05 mg $CuCl_2.2H_2O$, 0.03 mg $Na_2MoO_4.4H_2O$, 0.03 mg $VOSO_4.2H_2O$ and 0.01 mg $CoSO_4.7H_2O$, 0.05% yeast extract and 0.2% glucose. Cultured cells were harvested and stored as described above.

The New Zealand (NZ) hot spring isolates are maintained in a collection at the Thermophile Research Unit (University of Waikato, Hamilton, NZ). Cells were inoculated into an appropriate growth medium, with all media at pH 6. *Ignisphaera aggregans* AQ1.S1^T was grown at 95°C, *Ignisphaera* sp. Tok37.S1 at 90°C and *Pyrobaculum* sp. AQ1.S2 at 90°C, by Hugh Morgan and Christine Tan (University og Waikato, NZ) as described previously (Niederberger *et al*., 2006; 2008). *Desulfurococcus* sp. Ket10.S1 was grown at 90°C as described previously (Patel *et al*., 1986). In each case, cells were aerobically harvested and stored as described above.

7.2.2. Extraction of IPLs

Cellular material (approx. 0.1-0.8 g) for the three species of *Sulfolobus* were thawed (where necessary) and extracted *via* a modified Bligh-Dyer technique (Nishihara and Koga, 1987a), whereby each was suspended in MeOH (20 mL) before chloroform (10 mL) and aqueous trichloroacetic acid (4 mL; 10% w:v) were added and the resultant mixture stirred for 2 h at room temperature. The solvent composition was then altered by addition of chloroform (10 mL) and water (10 mL) and the resultant biphasic mixture allowed to settle into layers over 1 h prior to isolation of the chloroform layer. The aqueous/methanol layer was further extracted with chloroform (2 × 20 mL) and the combined organic layers were reduced to dryness *in vacuo*.

7.2.3. Direct extraction or formation of ether lipid cores

Polar lipid extracts (from *Sulfolobus* species) or thawed cellular material (from MTH(ΔH) and the NZ isolates; approx. 0.1-1 g) were subjected to methanolysis by heating in methanolic HCl (4.8 M; prepared from aqueous HCl and MeOH) under reflux or in sealed vials in order to remove phosphatidyl and/or glycosyl polar head groups. In each case, the resulting solution was cooled and extracted with dichloromethane (DCM). Extracts were then combined, dried by passage through a small column of anhydrous $MgSO_4$ or Na_2SO_4 and reduced *in vacuo*. For acidic ethanolysis of MTH(ΔH) material grown at 70^{o}C (Section 2.2.4.2.5), ethanolic HCl (4.8 M) as opposed to methanolic HCl was used during extraction.

7.2.4. Flash column chromatography

The lipid core extracts from MTH(ΔH), *S. acidocaldarius* and the NZ isolates were dissolved in a minimal amount of DCM and loaded onto columns of activated alumina (5 mm x 50 mm) pre-washed with hexane:DCM (9:1[†]). Apolar components were removed *via* washing with three bed volumes of hexane:DCM (9:1). A polar fraction, containing ether lipid cores, was eluted in each case using three bed volumes of DCM:MeOH (1:1; Schouten *et al*., 2007b), with the exception of *S.*

[†] All solvent compositions are given as volume:volume ratios unless otherwise stated.

acidocaldarius, where a polar fraction was instead eluted using chloroform:MeOH (9:1; Nicolaus *et al*., 1990). Residual solvent in each case was removed under a stream of N_2. The core lipids of *S. shibatae* and *S. solfataricus* were not similarly purified.

7.2.5. Preparation of a sample for confirmation of GMGT 27 in *I. aggregans*

A small portion of the purified lipid cores extracted from *I. aggregans* AQ1.S1^T were dissolved in DCM, added to a portion of the cores extracted from MTH(ΔH) cellular material grown at 70^{o}C and the solvent subsequently removed under a stream of N_2. The resulting residue was used to confirm the presence of GMGT **27** in *I. aggregans*, as described in Section 4.2.2.1.

7.3. Preparation of lipid extracts from sediment and soil samples

7.3.1. Collection and preparation of ancient sediments and greenhouse soil

Approximately 3 g of each of a series of homogenised sediments from the Messinian VDG basin, Northern Italy were used in the studies. These included horizons II-T1, III-1b, IV-1b, IV-1.1, IV-1.2, IV-1.5, IV-1.7, IV-1.8, IV-1.9, IV-1.10, V-1b, VI-1b, VII-1b, VIII-1b, IX-1b, X-1 and XI-1b (as designated in Damsté *et al*., 1995; Keely *et al*., 1995). 20 g of a sample of Jurassic shale from the Stewartby member of the Oxford Clay formation, UK (sample S90-11; Kenig *et al*., 1994) was homogenised by gentle pulverisation prior to extraction. 32.6 g of a fine composted soil was collected (collection date February 2010) from a domestic greenhouse (Norton, North Yorkshire, UK). The measured air temperature inside the greenhouse during collection of the sample was approximately 5^{o}C.

7.3.2. Extraction of ether lipid cores from aquatic sediments

Sediments were extracted *via* sonication (10 min), centrifugation (10 min, 2000 rpm) and filtration of the resultant supernatant using a three solvent program comprising MeOH, followed by MeOH:DCM (1:1) and finally DCM (Hopmans *et al*., 2000).

The supernatants from each stage were combined and the solvent removed *in vacuo*. Core lipids were further purified by flash column chromatography as outlined for the lipid extracts from MTH(ΔH) in Section 7.2.4.

7.3.3. Extraction and degradation of polar lipids from soils

The greenhouse soil was sonicated (10 min) in 200 mL of MeOH:DCM:5% (w:v) aqueous trichloroacetic acid (2:1:0.8). Following centrifugation (10 min, 2000 rpm), the supernatant was collected, filtered and modified to a composition of MeOH:DCM:aqueous trichloroacetic acid (1:1:0.9) by addition of DCM and dionised H_2O. The modified supernatant was subjected to centrifugation (5 min, 2000 rpm) and the bottom (i.e. DCM) layer of the resultant biphasic mixture collected *via* pipette. The entire extraction procedure was repeated using the residual soil, with the DCM layers from both extractions collected, combined and reduced *in vacuo*. As the residue potentially contained both IPL and core lipid material, it was subjected to acidic methanolysis as outlined in Section 7.2.3. Liberated lipid cores were purified by flash column chromatography as outlined for the lipid extract from MTH(ΔH) in Section 7.2.4.

7.4. Derivatisation and degradation of ether lipid cores

7.4.1. Silylation

The silyl ethers described in Section 2.2.10.2 were formed from purified tetraether lipid cores extracted from MTH(ΔH) cellular material grown at 45^{o}C. The lipids were dissolved in 100 μL of BSTFA containing 1% TMS chloride, pyridine (3 drops) was added and the sample was heated at 60^{o}C for 1 h. Following the reaction, excess reagent was removed *in vacuo* and the derivatisation products reconstituted in hexane:IPA (99:1) prior to direct infusion MS.

7.4.2. Acetylation of lipid cores from MTH(ΔH)

The acetates described in Section 2.2.10.3 were also formed from purified tetraether lipid cores extracted from MTH(ΔH) cellular material grown at 45^{o}C using modifications to a previous method (Morii *et al*., 1998). The lipids were dissolved in 500 μL of acetic anhydride:pyridine (1:1) and the sample left at room temperature for 24 h. Following the reaction, excess solvent was removed *in vacuo* at 45^{o}C and the derivatisation products dissolved in hexane:IPA (99:1) prior to direct infusion MS.

The synthetic monoacetate described in Section 3.2.3.3 was formed from the purified tetraether lipid cores extracted from MTH(ΔH) cellular material using a method derived from that of Ogawa *et al.* (1998). The lipid material was dissolved in DCM and ~200 mg of silica added, with the solvent subsequently removed *in vacuo*. The adsorbed lipid preparation was suspended in 3 mL hexane under N_2, a drop of acetyl chloride was added and the suspension was heated at 60^{o}C for 2 h, with agitation every 30 min. The suspension was filtered and the filtrand washed with portions of DCM and MeOH, which were collected, combined with the initial filtrate and reduced to dryness. The residue was taken up in MeOH and the precipitate formed collected *via* filtration and washed with hexane, with the washings collected, reduced to dryness and dissolved in hexane:IPA (99:1) prior to direct infusion MS.

7.4.3. Acetylation/fractionation of lipid cores from the Oxford Clay S90-11 sample

Purified lipid cores extracted from the Oxford Clay S90-11 sample were acetylated (Section 5.2.3.2) *via* refluxing at 105^{o}C in 6 mL of acetic anhydride:pyridine (1:2) for 24 h. Excess solvent was removed *in vacuo*, after which the residue was reconstituted in the minimum amount of DCM and loaded onto a column of activated alumina (5 mm x 50 mm), pre-washed with hexane:DCM (9:1). Three fractions of varying polarity were eluted from the column; fraction 1 was eluted using three bed volumes of hexane:DCM (9:1), fraction 2 subsequently eluted using three bed volumes of DCM and, finally, fraction 3 eluted using three bed volumes of

DCM:MeOH (1:1). Residual solvent was removed under streams of N_2 and each fraction reconstituted in DCM:MeOH (1:1) prior to reversed-phase LC-MS/MS.

7.4.4. Formation of ether-cleaved hydrocarbons

Polar column fractions containing ether lipid cores extracted from MTH(ΔH) cells grown at 45°C and 70°C (Section 3.2.3.5) and from sediment of VDG horizon X-1 (Section 5.2.5.6) were heated for 4 h at 120-130°C in 2 mL of HI (57% aq.) to facilitate cleavage of the ether lipid cores to alkyl iodides (Koga and Morii, 2006). In each case, the reaction vessel was cooled and the solution extracted with 3 x 5 mL hexane. The organic layers were washed with saturated $NaHCO_3$ solution (10 mL) followed by 10% sodium thiosulfate solution (10 mL) and H_2O (10 mL), before being reduced *in vacuo*, flushed through small columns of $MgSO_4$ and reduced to dryness under a stream of N_2. The alkyl iodides generated from the MTH(ΔH) 70°C sample were subjected to an additional purification step, whereby they were dissolved in hexane:DCM (9:1), passed through a column of activated alumina (5 mm x 50 mm) using four bed volumes of the same solvent and dried under N_2.

Alkyl iodides were reduced to hydrocarbons by reaction with 2 mL of $LiAlH_4$ (2M solution in tetrahydrofuran) for 2-14 h at 70°C (Hoefs *et al*., 1997; DeLong *et al*., 1998). The solutions were cooled in ice and ethyl acetate (5 mL) was added cautiously to each, followed by addition of 5% NaCl solution (2 mL). The biphasic mixtures were agitated, allowed to settle and the ethyl acetate layers siphoned from the top. The remnant aqueous suspensions were each extracted with ethyl acetate (2 x 5 mL) and the organic layers from each round of extractions combined, reduced *in vacuo* and filtered through a column of $MgSO_4$ before being reduced to dryness. The residues from the MTH(ΔH) 45°C and the VDG X-1 sample were reconstituted in hexane:DCM (9:1), whereas that from the MTH(ΔH) 70°C sample was dissolved in hexane. Each was subsequently passed through a column of activated alumina (5 mm x 50 mm) using three bed volumes of the solvent used for dissolution and reduced to dryness under N_2.

An alternative method proved equally effective for the reduction of alkyl iodides generated from the MTH(ΔH) 45°C sample (Panganamala *et al*., 1971; Koga and Morii, 2006). The alkyl iodide preparation was dissolved in 2 mL acetic acid, 200 mg of zinc powder was added and the resulting suspension heated at 120°C for 2 h, following which the reaction vessel was cooled in ice before acetic acid (2 mL) and H_2O (5 mL) were added. The suspension was extracted with hexane (3 x 5 mL), with the organic layers subsequently washed with saturated $NaHCO_3$ solution (2 x 10 mL) and H_2O (5 mL), before being reduced *in vacuo*, filtered through a column of $MgSO_4$ and taken to dryness. The residue was further purified as described for the $LiAlH_4$ reduction procedure.

7.5. Mass spectrometric analyses

7.5.1. Ion trap MS

Direct infusion APCI positive ion mass spectra were recorded either on a Finnigan MAT LCQ (Thermo-Finnigan; San Jose, CA) or HCTultra ETD II (Bruker Daltonics; Coventry, UK) ion trap mass spectrometer, operated in both cases with scan range *m/z* 50-2000 unless otherwise stated. The LCQ operational parameters were set at: vaporiser temperature 450°C; sheath gas (N_2) pressure 30 (arb. units); auxiliary gas (N_2) pressure 0; capillary temperature 150°C; capillary voltage 47.8 kV; corona discharge current 5 μA; corona discharge voltage 5 kV. The HCT operational parameters were set at: nebuliser gas (N_2) pressure 15 psi; drying gas (N_2) flow 4 $Lmin^{-1}$; drying gas temperature 300°C; vaporiser temperature 450°C; capillary voltage -3500 V.

Solutions which were directly infused into the ion trap mass spectrometers were delivered by syringe injection at 20 $\mu Lmin^{-1}$ into a makeup flow of an additional solvent pumped at 0.3 $mLmin^{-1}$. Where hexane:IPA (99:1) was used for delivery, it was also used as the makeup flow. Where DCM or DCM:MeOH (1:1) was used for delivery, methanol was used as the makeup flow unless otherwise stated.

7.5.2. Instrument tuning

Ion optic parameters for each instrument were optimised for tetraether lipids *via* tuning on the $[M+H]^+$ of GDGT **2** (nominal *m/z* 1301), as found in a direct infusion of a portion of the lipid extracts from MTH(ΔH) grown at either 45°C or 70°C.

7.5.3. Ion trap MS^n

MS/MS and MS^n spectra, as described in Section 2.2.6 and Section 3.2.3.3 (in the latter case, for synthetic lipid **59** only), were recorded using the LCQ ion trap mass spectrometer during direct infusion. The isolation width was set to 3 *m/z* units in each case. The collision energy used to generate MS/MS spectra was 41%, whereas the collision energies used to generate MS^3 spectra, given explicitly in Fig. 2.7, differed depending on the selected precursor ion.

MS/MS spectra described in Sections 2.2.10 were recorded using the HCT ion trap mass spectrometer during direct infusion. Attempts to manually select precursors for MS/MS were unsuccessful using this instrument and, consequently, the spectra were recorded using the Auto MS^n feature, which automatically selects the base peak ion in a preliminary mass spectral scan for CID. To select for precursors that were not the base peak ion in the mass spectrum, the MS scan range was modified and set to encompass a narrow width of 5-20 *m/z* units, centred around the *m/z* value of the desired precursor. In each case, the isolation width was set to a 3 *m/z* unit window, with the maximum accumulation time set at 40 ms and fragmentation amplitude fixed at 2.0 V (SmartFrag not activated).

7.5.4. sCID MS

The sCID MS spectra described in Section 2.2.7 were recorded using the LCQ ion trap mass spectrometer during direct infusion. The source CID collision energies were set to the specific values described in Sections 2.2.7.2, 2.2.7.3 and 2..2.7.4. The MS/MS spectra recorded with sCID activated were obtained as described in Section 7.5.3.

7.5.5. Direct infusion PAN MS/MS

The PAN MS/MS spectrum described in Section 2.2.8.2 was recorded using the HCT ion trap mass spectrometer during direct infusion. Auto MS^n was used to acquire the spectrum, with PAN MS/MS activated and the PAN parameters set as: low CID 27%; min. frag. *m/z* 60; activation 100%; decay 50%; tickle 1000%.

7.5.6. ESI Qh-FT-ICR MS and MS/MS

High mass accuracy mass and tandem mass spectra of purified lipid cores extracted from MTH(ΔH) grown at 45^{o}C, as described in Section 2.2.9, were acquired using an 9.4 T Apex ultra Qh-FT-ICR hybrid instrument (Bruker Daltonics; Coventry, UK). The lipid cores were first dissolved in acetonitrile:IPA (80:20) before being further diluted by a factor of 25 into acetonitrile:IPA (80:20) containing 0.2% formic acid. The solution was directly infused into the instrument, with ionisation effected *via* an Apollo II dual source with ESI operated in positive ion mode. The ESI parameters were set at: nebuliser gas (N_2) flow 0.8 $Lmin^{-1}$; drying gas (N_2) flow 5.0 $Lmin^{-1}$; drying gas temperature 190^{o}C; capillary voltage 4.2 kV. The MS scan range was set to *m/z* 200-3000. During MS/MS, the precursor ion was selected for CID using a Q1 resolution of 20 *m/z* units. The hexapole collision energy was set to 32.5 eV, with Ar used as the collision gas.

7.5.7. MALDI Qh-FT-ICR MS

The high mass accuracy mass spectrum of a preparation enriched in GDHT **119**, as described in Section 5.2.4.4, was also acquired on the Apex ultra Qh-FT-ICR instrument, with ionisation effected by MALDI. The lipid preparation was dissolved in acetonitrile:IPA (80:20) and premixed with an α-cyano-4-hydroxycinnamic acid matrix. 1.5 μL of the resulting solution was subsequently spotted onto a MALDI plate and allowed to dry. The mass spectrum was acquired from 16 consecutive scans, each of 200 laser shots. The MS scan range was set to *m/z* 300-5000.

7.6. LC-MS/MS

7.6.1. Normal-phase LC-MS/MS

Purified lipid core extracts were reconstituted in hexane:isopropanol (99:1) *via* sonication for 2 min in preparation for LC-MS/MS analysis. LC-MS/MS was performed using either a Finnigan system comprising a Thermo Separations AS3000 autosampler, P4000 gradient pump and the LCQ ion trap mass spectrometer (Thermo-Finnigan; San Jose, CA, USA) or, otherwise, using an Ultimate 3000 rapid separation liquid chromatograph (Dionex; Sunnyvale, USA) coupled to the HCT ion trap mass spectrometer.

Separation was achieved on a Waters Spherisorb S5-NH_2 column (4.6 x 250 mm, 5 μm; Elstree, UK) using a binary solvent system, operated at a flow rate of 1 $mLmin^{-1}$ and comprising hexane:IPA as the eluent. Three different gradient programs were utilised, all representing minor adjustments to the program suggested in a previous literature method (Hopmans *et al*., 2000). Each program comprises a period of elution in the conventional flow direction for the column followed by a period of backflushing, in which the direction of eluent flow through the column is reversed and finally, a reconditioning of the column in the conventional direction with the initial eluent composition (Table 7.1). Method A was used during acquisition of data presented in Chapter 2, whereas method B was used to generate results presented in Chapters 3-5. Method C was used in one specific case, as described in Section 5.2.3.3. The injection volume in each case was 5 μL.

In all cases, ionisation was facilitated *via* an APCI source, operated in positive ion mode. For the LCQ instrument, the APCI parameters were identical to those given for direct infusion MS in Section 7.5.1. For the HCT instrument, all parameters were similarly identical to those given in Section 7.5.1 apart from the drying gas flow and temperature, which were increased to 8 $Lmin^{-1}$ and 325^{o}C, respectively, for LC-MS/MS analyses. Unless otherwise stated, the MS scan range was fixed at *m/z* 900-1500. MS/MS spectra were recorded automatically by selection and CID of the base peak ion in the preceding MS scan. MS/MS settings were identical to those

described in Section 7.5.3. The dwell time was set to 100 ms for the LCQ instrument. The maximum accumulation time for the HCT instrument was set to 40 ms.

Table 7.1. Normal-phase LC-MS/MS binary gradient programs, with the percentage compositions of IPA in the eluent shown.

	Time	Method		
	(min)	A	B	C
Elution	0	1	1	1
	5	1	1	1
	61.25	2	2	2
	90	-	15	15
	105	-	-	15
Backflush	10 min	10	25	-
	15 min	-	-	25
Recondition	15 min	1	1	1

7.6.2. Studies into normal-phase LC-MS/MS retention time drift

The studies investigating retention time drift during normal-phase LC-MS/MS analyses, as described in Section 2.2.4.1.3, each utilised lipid core material from MTH(ΔH) grown at 45^{o}C. Regulation of column temperature was achieved using a Shimadzu CTO-10Avp column oven (Shimadzu UK, Milton Keynes), set at 30^{o}C. The column with alternative cyano chemistry utilised was a Waters Spherisorb S3-CN column (4.6 x 150 mm, 3 μm; Elstree, UK). For the study in which part of the IPA in the eluent was replaced with DCM, an isocratic eluent composition of hexane:DCM:IPA (79.7:20:0.3) was used instead of gradient method A. Cholesterol used as a retention time marker was added to the lipid material as a solution in DCM, with the DCM removed under a stream of N_2 prior to analysis.

7.6.3. PAN LC-MS/MS

The PAN LC-MS/MS analyses described in Sections 2.2.8.3, 3.2.3.4, 3.2.4.6 and 4.2.2.2 were achieved *via* lipid separation and ionisation as discussed in Section 7.6.1, but with PAN MS/MS activated. PAN parameters were identical to those described in Section 7.5.5.

7.6.4. Reversed-phase LC-MS/MS

Reversed-phase LC-MS/MS analyses of the acetates of the Oxford Clay S90-11 lipid core extract, as described in Section 5.2.3.2, were performed using an Ultimate 3000 liquid chromatograph (Dionex; Sunnyvale, USA) coupled to the HCT ion trap mass spectrometer. Separation was achieved on a Waters XTerra RP_{18} column (4.6 x 100 mm, 3.5 μm; Elstree, UK) using an isocratic eluent flow comprising DCM:acetonitrile:MeOH (5:5:90) at a rate of 1 $mLmin^{-1}$. Ionisation was *via* APCI, with the APCI parameters identical to those described for normal-phase LC-MS/MS in Section 7.6.1. The MS scan range was set to *m/z* 50-2000 during preliminary screening of the samples, but was modified to *m/z* 1361.5-1372.5 during re-analysis of column fraction 2. MS/MS spectra were recorded using settings identical to those described in Section 7.5.3.

7.7. GC and GC-MS

7.7.1. Sample preparation

Both hexane and ethyl acetate proved to be effective as delivery solvents for GC-FID and GC-MS analyses of the hydrocarbon preparations.

7.7.2 GC-FID analysis of ether-cleaved hydrocarbons from MTH(ΔH)

GC-FID analyses of the hydrocarbons liberated from the lipid cores of MTH(ΔH) grown at 45^{o}C or 70^{o}C (Section 3.2.3.5) were performed using a Trace GC Ultra gas chromatograph (Thermo Scientific; Hemel Hempsted, UK) equipped with a DB-1

fused silica column (60 m, ID 0.32 mm, Film thickness 1.0 μm; J&W Scientific/Agilent Technologies, Stockport, UK). H_2 was used as the carrier gas. Samples were introduced by on-column injection of a plug of 1 μL of the delivery solution. The temperature program involved injection at 70^{o}C followed by an immediate ramp to 270^{o}C at 20^{o}C min^{-1}, a subsequent ramp to 320^{o}C at 1^{o}C min^{-1} and a 20 min isothermal period in which the temperature was held at 320^{o}C. The detector temperature was held at 350^{o}C throughout.

7.7.3. GC-FID analysis of ether-cleaved hydrocarbons from VDG horizon X-1

GC-FID analysis of the hydrocarbons liberated from the lipid cores of the VDG X-1 horizon (Section 5.2.5.6) was achieved using a Carlo Erba HRGC5300 Mega gas chromatograph (Carlo Erba Instruments, Milan, Italy) equipped with a HP-5 fused silica column (25 m, ID 0.32 mm, film thickness 1.05 μm; Hewlett Packard). H_2 was used as the carrier gas. Samples were, again, introduced by on-column injection of a plug of 1 μL of the delivery solution. The temperature program was identical to that used previously by Schouten *et al.* (2008b); the oven temperature at injection was 70^{o}C, which was increased to 130^{o}C at 20^{o}C min^{-1} and then to 320^{o}C at 4^{o}C min^{-1}, after which the temperature was held at 320^{o}C for 20 min. The detector temperature was held at 320^{o}C.

7.7.4. GC-MS analysis of ether-cleaved hydrocarbons

GC-MS analyses of all ether-cleaved hydrocarbons were performed using an Agilent 7890A gas chromatograph (Agilent Technologies, Stockport, UK) coupled to a Waters GCT Premier (Micromass UK LTD; Wythenshawe, UK) mass spectrometer. Separation was achieved on an Equity-5 fused silica column (30 m, ID 0.25 mm, Film thickness 0.25 μm; Supelco, Bellefonte) with He used as the carrier gas. Samples were introduced by split/splitless injection (split ratio 50:1) of a plug of 1 μL of the delivery solution. The column oven temperature program was based on that of Schouten *et al.* (2008b) with minor modifications; the initial oven temperature during injection was 50^{o}C, which was held for 1 min before an initial ramp to 130^{o}C at 20^{o}C min^{-1}, followed by a secondary ramp to 320^{o}C at 4^{o}C min^{-1},

after which the temperature was held at 320^{o}C for 20 min. The mass spectrometer scanned the range *m/z* 30-850 with the ionisation energy fixed at 70 eV.

7.8. Data processing and analytical considerations

7.8.1. LC-MS chromatograms and integration

In order to allow for the sub-unit resolution of the ion trap instruments employed, reconstructed ion chromatograms for LC-MS were generated by selecting all ions within ±0.5 *m/z* units of the target *m/z* value. Lipid relative abundances were calculated from individual ion chromatogram LC-MS peak areas, with ionisation efficiencies assumed to be constant for all of the lipids.

7.8.2. Analytical replication

Where analytical replicates were obtained, quantitative values are given as the mean of the repeat measurements. Errors quoted for such values represent one standard deviation (±σ) from the mean in each case. The number of replicates performed (n) is also provided in each case.

7.8.3. Mass and tandem mass spectra

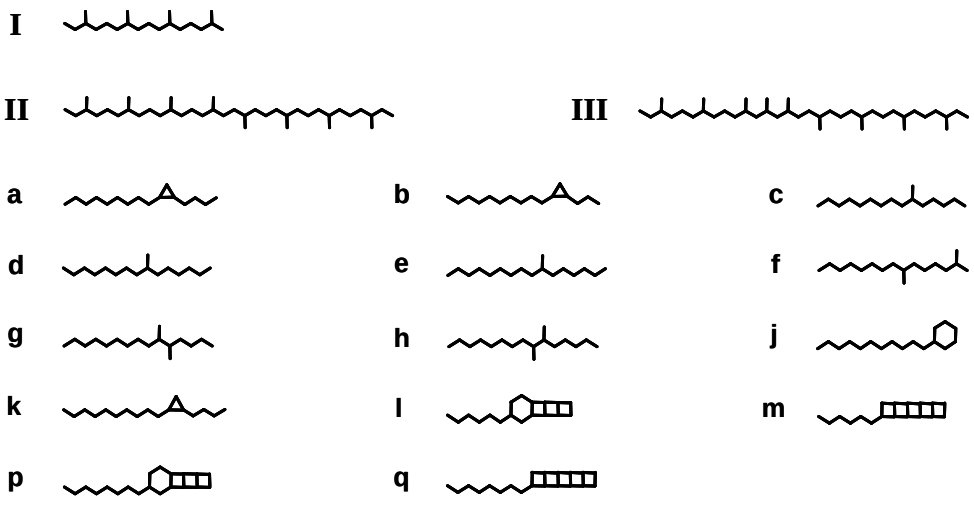
All ions identified in ion trap mass spectra are reported with their *m/z* value as measured on the ion trap instrument used, unless otherwise explicitly stated. Product ions in ion trap tandem mass spectra (i.e. MS/MS, MS^n, sCID MS, PAN MS/MS) are typically reported with their nominal *m/z* value, with the exception of those labelled in Fig. 2.4, where the average value measured for each during 12 replicate LC-MS analyses is shown. Ions in mass and tandem mass spectra recorded using the Qh-FT-ICR instrument are reported with their high mass accuracy *m/z* values.

Structural Appendix

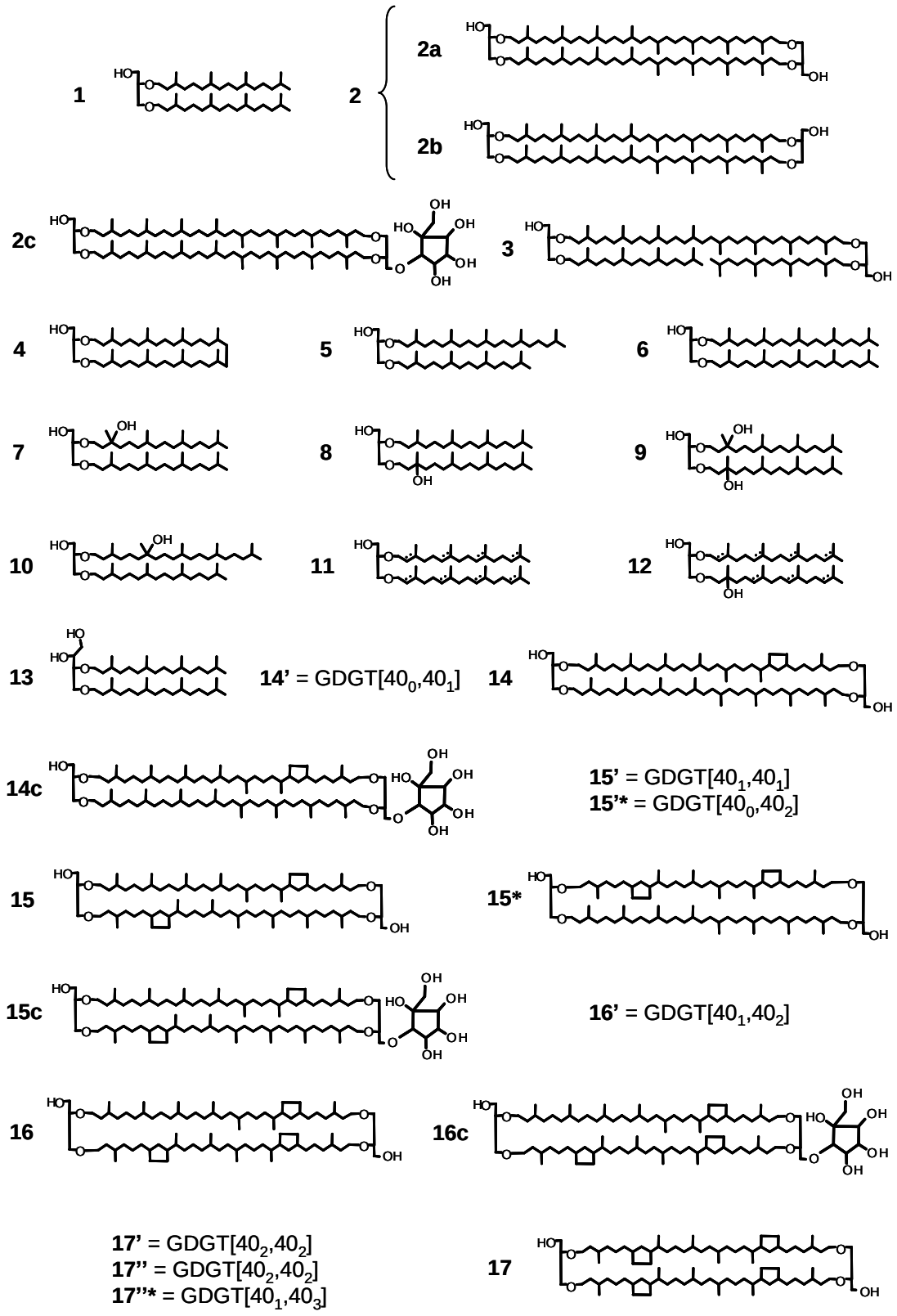
Structural representations

Provided in this appendix are representative structures for the ether lipid cores and hydrocarbons described in the preceding text. Although the structures of some of the lipids have been reported previously, often following rigorous structural analysis, many others are reported for the first time. For these latter lipid cores, the structures provided were reconstructed following consideration of both the LC-MS/MS data obtained and *a priori* knowledge of the limited structural variety within known archaeal lipids (e.g. the Cp rings are incorporated at specific positions). Nevertheless, the structures shown are provided for illustrative purposes only and should be treated as such. Most notably, the positions of cyclopentyl or cyclohexyl rings, methyl branching groups and/or hydroxyl groups in the etherified alkyl chains, the positions of covalent cross-links between etherified chains and the orientation of each chain with respect to the *sn*-2 and *sn*-3 carbinols of the two capping glycerol groups in each lipid may differ from those shown. Likewise, **2a**/**2b** and **26a**/**26b** aside, all tetraether structures are presented solely as isomers in which the glycerol groups are arranged *anti* to one-another. Isomers of each lipid in which the glycerol groups have a *syn* arrangement should also be considered as potential structural variants.

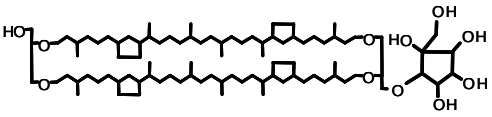
Hydrocarbons



Ether lipid cores

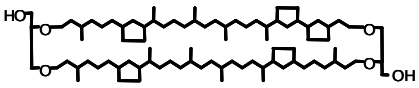


17c

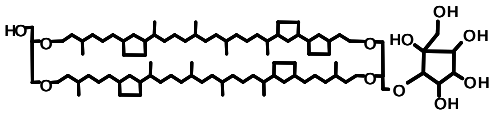


18’ = GDGT[40_2,40_3]

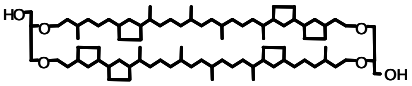
18



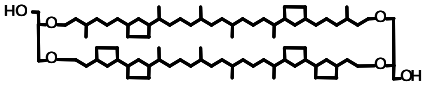
18c



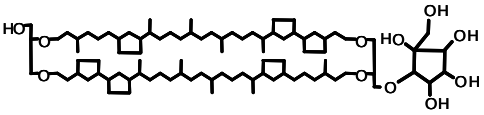
19



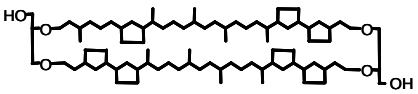
19*



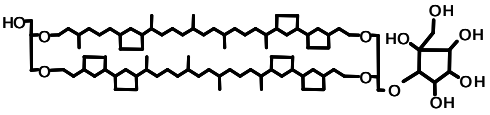
19c



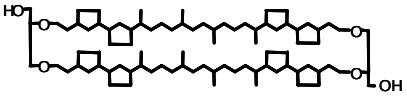
20



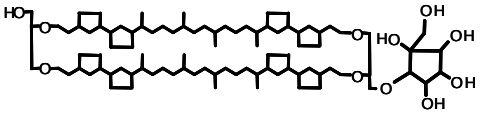
20c



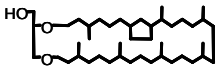
21



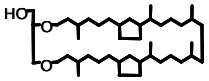
21c



22

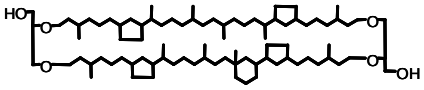


23

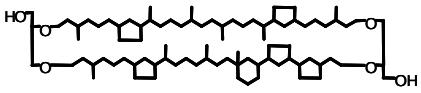


24’ = GDGT[40_2,40_{2+1}]
24’’ = GDGT[40_2,40_{2+1}]

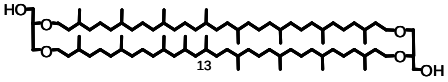
24



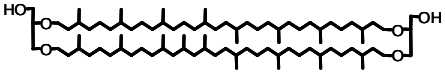
25



26a

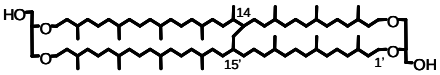


26b

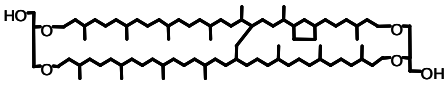


27’ = GMGT containing C_{80} chain

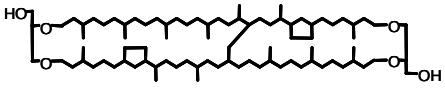
27



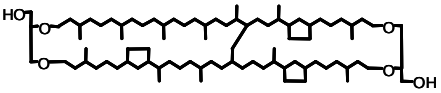
28



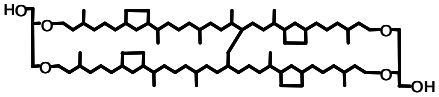
29



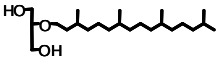
30



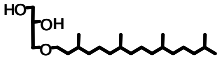
31



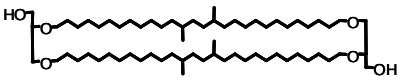
32



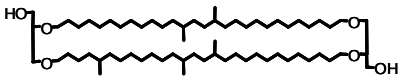
33



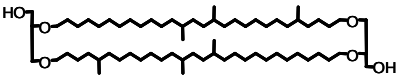
34 (**J10**)



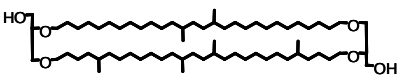
35 (**J9**)



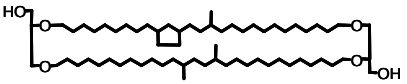
36 (**J8**)



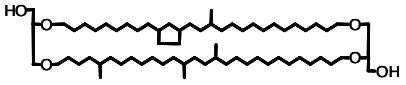
36* (**J8***)



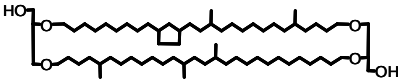
37 (**R2**)



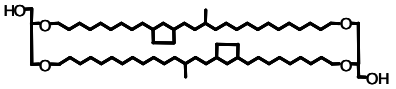
38 (**R1**)



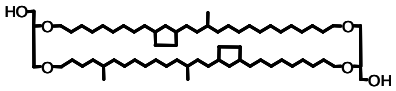
39



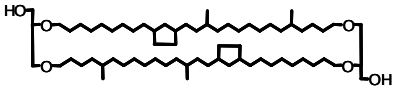
40 (**U2**)



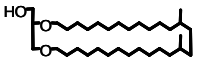
41 (**U1**)



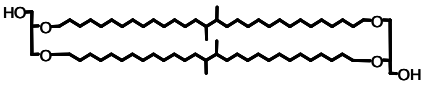
42



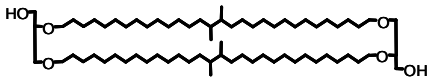
43



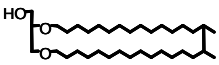
44



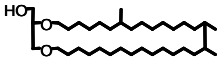
45



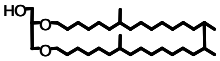
46



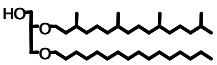
47



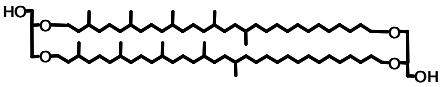
48



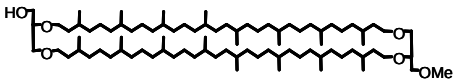
49



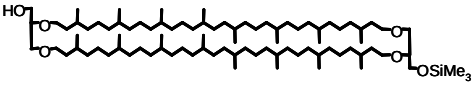
50 (**G4**)



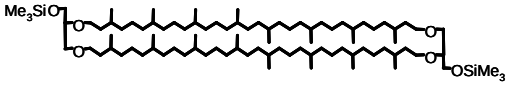
51



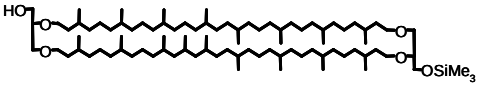
52



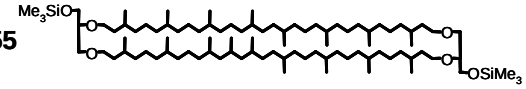
53



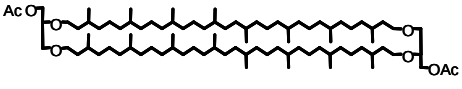
54



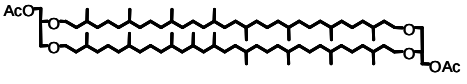
55



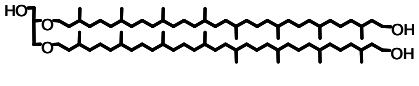
56



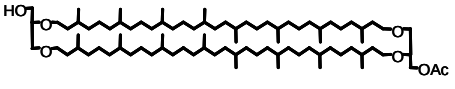
57



58



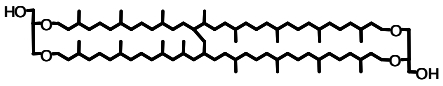
59



60 = Unknown dialkyl tetraether

61 = Unknown dialkyl tetraether

62

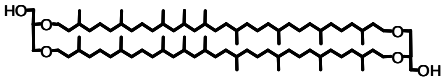


63 = Unknown dialkyl tetraether

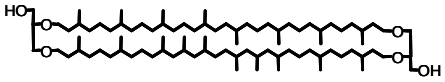
64 = Unknown dialkyl tetraether

65 = Unknown dialkyl tetraether

66

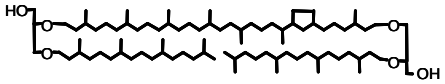


66*

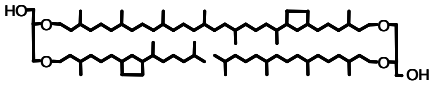


67 = Unknown dialkyl tetraether

68

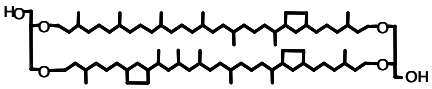


69

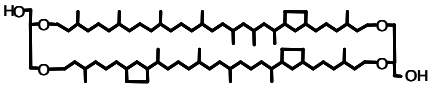


70' = $GDGT_M$ containing 3 Cp rings

70

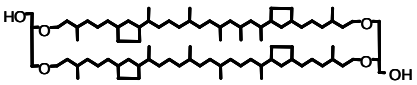


70*



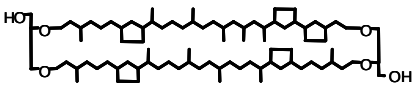
71' = $GDGT_M$ containing 4 Cp rings

71

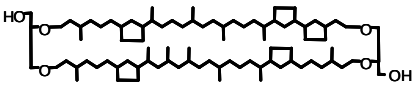


72' = $GDGT_M$ containing 5 Cp rings

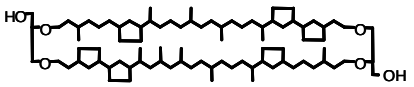
72



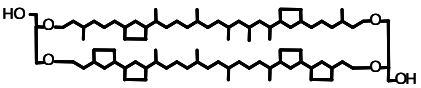
72*



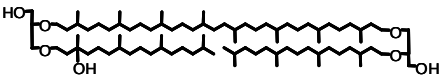
73



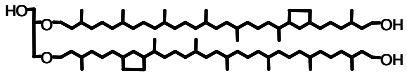
73*



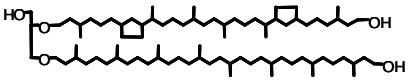
74



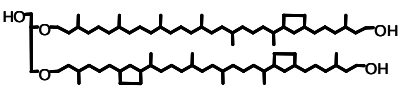
75



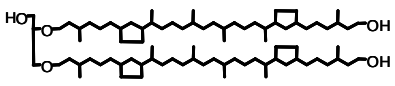
75*



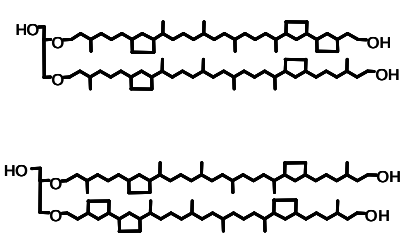
76



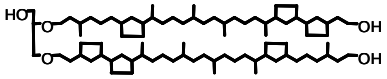
77



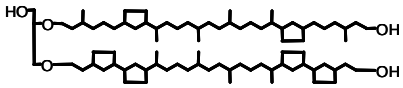
78' + **78**



79



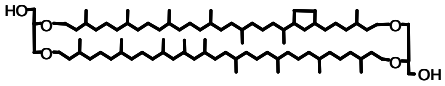
79*



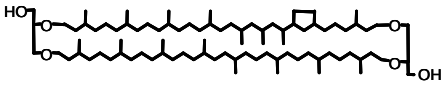
80 = Unknown GDGT

81 = Unknown dialkyl tetraether

82

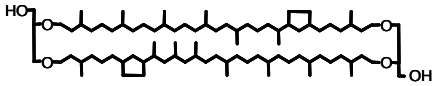


82*

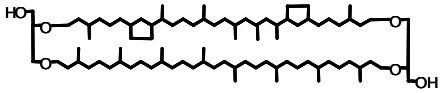


83' = $GDGT_M$ containing 2 Cp rings

83



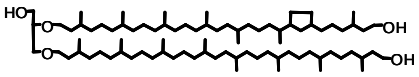
83*



84 = $GDGT_M$ containing 3 Cp rings

85 = $GDGT_M$ containing 4 Cp rings
85' = $GDGT_M$ containing 4 Cp rings

86

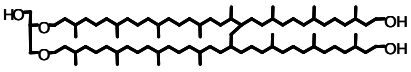


87 = Unknown dialkyl tetraether

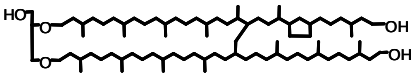
88 = Unknown dialkyl tetraether

89 = Unknown dialkyl tetraether

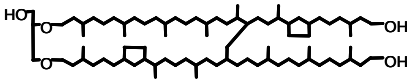
90



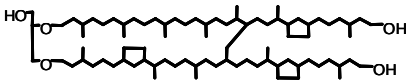
91



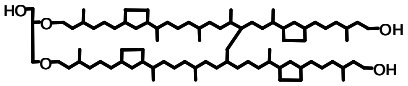
92



93

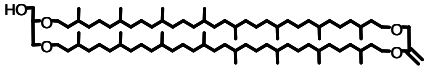


94

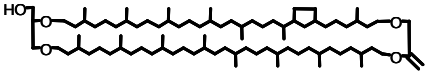


95 = Unknown dialkyl tetraether

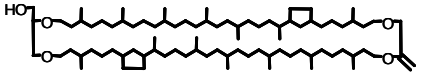
96



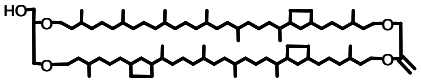
97



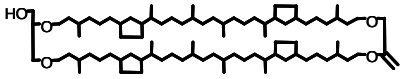
98



99



100



101 = Unknown dialkyl tetraether

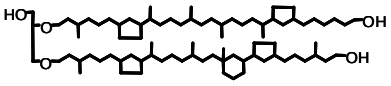
102 = Unknown dialkyl tetraether

103 = Unknown dialkyl tetraether

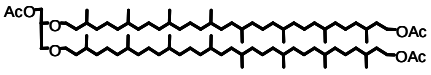
104' = Unknown dialkyl tetraether
104 = Unknown dialkyl tetraether

105 = Unknown dialkyl tetraether

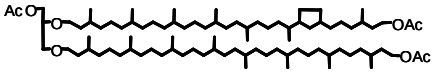
106



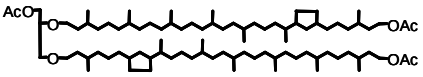
107



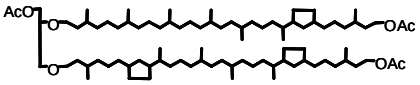
108



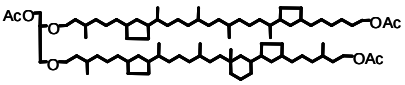
109



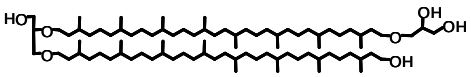
110



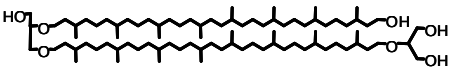
111



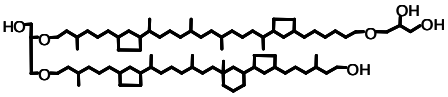
112'



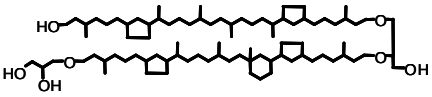
112



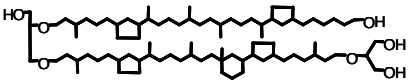
113'



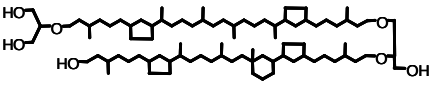
113'*



113



113*

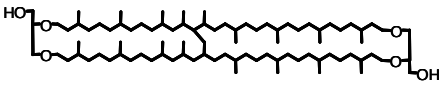


114 = GDGT[41_0,42_0]

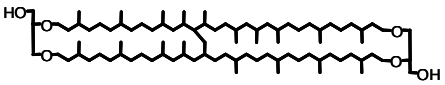
115 = $GDGT_M$ containing 1 Cp ring

116 = $GDGT_M$ containing 2 Cp rings

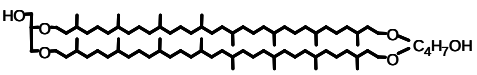
117



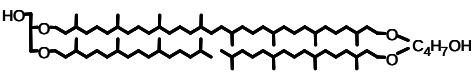
118



119

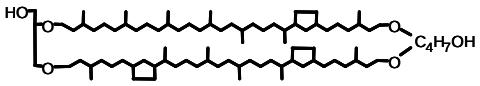


120

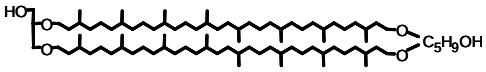


121 = GDHT containing 1 Cp ring

122' = GDHT containing 2 Cp rings
122 = GDHT containing 2 Cp rings

123 

124 = GDHT containing 4 Cp rings

125 

126 = Unknown GDHT

127 = GDD lipid; 36 chain carbon atoms

128 = GDD lipid; 35 chain carbon atoms

129 = GDD lipid; 34 chain carbon atoms

130 = GDD lipid; 33 chain carbon atoms

131 = GDD lipid; 32 chain carbon atoms

132 = GDD lipid; 31 chain carbon atoms

133 = GDD lipid; 30 chain carbon atoms

134 = GDD lipid; 29 chain carbon atoms

135 = GDD lipid; 28 chain carbon atoms

136 = Triether tetrol; 64 chain carbon atoms

137 = Triether tetrol; 63 chain carbon atoms

138 = Triether tetrol; 62 chain carbon atoms

139 = Triether tetrol; 61 chain carbon atoms

140 = Triether tetrol; 60 chain carbon atoms

141 = Unknown non-isoprenoid lipid

142 = Unknown non-isoprenoid lipid

143 = Unknown non-isoprenoid lipid

144 = Unknown non-isoprenoid lipid

A1 = GDGT[37_0,38_0] methyl ether?
A2 = GDGT[37_0,37_0] methyl ether
A3 = GDGT[36_0,37_0] methyl ether
A4 = GDGT[35_0,37_0] methyl ether
A4* = GDGT[36_0,36_0] methyl ether
A5 = GDGT[35_0,36_0] methyl ether
A6 = GDGT[35_0,35_0] methyl ether
A7 = GDGT[34_0,35_0] methyl ether?

B1 = GDGT[33_0,34_0] methyl ether?
B2 = GDGT[33_0,33_0] methyl ether?
B3 = GDGT[32_0,33_0] methyl ether?
B4 = GDGT[32_0,32_0] methyl ether?
B5 = GDGT[31_0,32_0] methyl ether?
B6 = GDGT[31_0,31_0] methyl ether?
B7 = GDGT[30_0,31_0] methyl ether?
B8 = GDGT[30_0,30_0] methyl ether?

C1 = $GDH_2T[38_0,38_0]$?
C2 = $GDH_2T[37_0,38_0]$
C3 = $GDH_2T[37_0,37_0]$
C4 = $GDH_2T[36_0,37_0]$
C5 = $GDH_2T[36_0,36_0]$
C6 = $GDH_2T[35_0,36_0]$?
C7 = $GDH_2T[35_0,35_0]$?

D1 = GDHT[38_0,38_0]?
D2 = GDHT[37_0,38_0]
D3 = GDHT[37_0,37_0]
D4 = GDHT[36_0,37_0]
D5 = GDHT[36_0,36_0]
D6 = GDHT[35_0,36_0]?

E1 = $GDH_2T[33_0,33_0]$?
E2 = $GDH_2T[32_0,33_0]$
E3 = $GDH_2T[32_0,32_0]$

F1 = GDHT[33_0,34_0]?
F2 = GDHT[33_0,33_0]?
F3 = GDHT[32_0,33_0]?
F4 = GDHT[32_0,32_0]?

G1 = GDGT[37_0,38_0]?
G2 = GDGT[37_0,37_0]
G3 = GDGT[36_0,37_0]
G4 = **50** = GDGT[35_0,37_0]
G5 = GDGT[35_0,36_0]?
G6 = GDGT[35_0,35_0]?

H1 = GDGT[37_0,37_0]
H2 = GDGT[36_0,37_0]
H3 = GDGT[36_0,36_0]
H4 = GDGT[35_0,36_0]
H5 = GDGT[35_0,35_0]
H6 = GDGT[34_0,35_0]
H7 = GDGT[34_0,34_0]?

J1 = GDGT[34_0,35_0]?
J2 = GDGT[34_0,34_0]
J3 = GDGT[33_0,34_0]
J4 = GDGT[33_0,33_0]
J4* = GDGT[32_0,34_0]
J5 = GDGT[32_0,33_0]
J6 = GDGT[32_0,32_0]
J7 = GDGT[31_0,32_0]
J8 = **36** = GDGT[31_0,31_0]
J8* = **36*** = GDGT[30_0,32_0]
J9 = **35** = GDGT[30_0,31_0]
J10 = **34** = GDGT[30_0,30_0]
J11 = GDGT[29_0,30_0]?
J12 = GDGT[29_0,29_0]?

K1 = GDT[37_0,37_0]
K2 = GDT[36_0,37_0]
K3 = GDT[36_0,36_0]
K4 = GDT[35_0,36_0]?
K5 = GDT[35_0,35_0]?

L1 = GDT[37_0,37_0]
L2 = GDT[36_0,37_0]
L3 = GDT[35_0,37_0]
L3* = GDT[36_0,36_0]
L4 = GDT[35_0,36_0]?
L5 = GDT[35_0,35_0]?
L6 = GDT[34_0,35_0]?
L7 = GDT[34_0,34_0]
L8 = GDT[33_0,34_0]?
L9 = GDT[33_0,33_0]
L10 = GDT[32_0,33_0]
L11 = GDT[32_0,32_0]
L12 = GDT[31_0,32_0]?
L13 = GDT[31_0,31_0]
L13* = GDT[30_0,32_0]
L14 = GDT[30_0,31_0]
L15 = GDT[30_0,30_0]

R1 = **38** = GDGT[30_1,31_0]
R2 = **37** = GDGT[30_0,30_1]

S1 = GMGT containing C_{69} chain?
S2 = GMGT containing C_{68} chain
S3 = GMGT containing C_{67} chain
S4 = GMGT containing C_{66} chain
S5 = GMGT containing C_{65} chain?
S6 = GMGT containing C_{64} chain?
S7 = GMGT containing C_{61} chain
S8 = GMGT containing C_{60} chain

T1 = GMT containing C_{68} chain
T2 = GMT containing C_{67} chain
T3 = GMT containing C_{66} chain
T4' = GMT containing C_{60} chain
T4 = GMT containing C_{60} chain

U1 = **41** = GDGT[30_1,31_1]
U2 = **40** = GDGT[30_1,30_1]

V1 = GTGT containing 64 chain C atoms?
V2 = GTGT containing 63 chain C atoms?
V3 = GTGT containing 62 chain C atoms?
V4 = GTGT containing 61 chain C atoms?
V5 = GTGT containing 60 chain C atoms?
V6 = GTGT containing 59 chain C atoms?
V7 = GTGT containing 58 chain C atoms?

References

Albers, S. V.; Szabó, Z.; Driessen, A. J. M.; (2006) Protein secretion in the Archaea: multiple paths towards a unique cell surface, *Nature Reviews Microbiology*, **4**, 7, 537-547.

Anderson, R.; Kates, M.; Baedecker, M. J.; Kaplan, I.; Ackman, R. G.; (1977) The stereoisomeric composition of phytanyl chains in lipids of Dead Sea sediments, *Geochimica et Cosmochimica Acta*, **41**, 1381-1390.

Bai, Q. Y.; Zelles, L.; (1997) A method for determination of archaeal ether-linked glycerolipids by high performance liquid chromatography with fluorescence detection as their 9-anthroyl derivatives, *Chemosphere*, **35**, 1-2, 263-274.

Balk, M.; Heilig, H. G. H. G.; van Eekert, M. H. A.; Stams, A. J. M.; Rijpstra, I. C.; Damsté, J. S. S.; de Vos, W. M.; Kengen, S. W. M.; (2009) Isolation and characterization of a new CO-utilizing strain, *Thermoanaerobacter thermohydrosulfuricus* subsp. *carboxydovorans*, isolated from a geothermal spring in Turkey, *Extremophiles*, **13**, 885-894.

Barns, S. M.; Delwiche, C. F.; Palmer, J. D.; Pace, N. R.; (1996) Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences, *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 17, 9188-9193.

Baudrand, M.; Grossi, V.; Pancost, R.; Aloisi, G.; (2010) Non-isoprenoid macrocyclic glycerol diethers associated with authigenic carbonates, *Organic Geochemistry*, in press.

Bauersachs, T.; Speelman, E. N.; Hopmans, E. C.; Reichart, G-J.; Schouten, S.; Damsté, J. S. S.; (2010) Fossilized glycolipids reveal past oceanic N_2 fixation by heterocystous cyanobacteria, *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 45, 19190-19194.

Bechtel, A.; Smittenberg, R. H.; Bernasconi, S. M.; Schubert, C. J.; (2010) Distribution of branched and isoprenoid tetraether lipids in an oligotrophic and a eutrophic Swiss lake: Insights into sources and GDGT-based proxies, *Organic Geochemistry*, **41**, 8, 822-832.

Beck, J. W.; Edwards, R. L.; Ito, E.; Taylor, F. W.; Recy, J.; Rougerie, F.; Joannot, P.; Henin, C.; (1992) Sea-Surface Temperature from Coral Skeletal Strontium/ Calcium Ratios, *Science*, **257**, 5070, 644-647.

Ben Ari, J.; Karni, M.; Apeloig, Y.; Mandelbaum, A.; (2003) The role of hydride migration in the mechanism of alcohol elimination from protonated ethers upon chemical ionization Experiment and theory, *International Journal of Mass Spectrometry*, **228**, 2-3, 297-306.

Ben Ari, J.; Navon, I.; Mandelbaum, A.; (2006) The effect of steric hindrance on the relative rates of anchimerically assisted alcohol eliminations from MH^+ ions of 2-substituted 1,4-dialkoxybutanes upon CI and CID - Experiment and theory, *International Journal of Mass Spectrometry*, **249**, 433-445.

Berkner, S.; Lipps, G.; (2008) Genetic tools for *Sulfolobus* spp.: vectors and first applications, *Archives of Microbiology*, **190**, 3, 217-230.

Biddle, J. F.; Lipp, J. S.; Lever, M. A.; Lloyd, K. G.; Sorensen, K. B.; Anderson, R.; Fredricks, H. F.; Elvert, M.; Kelly, T. J.; Schrag, D. P.; Sogin, M. L.; Brenchley, J. E.; Teske, A.; House, C. H.; Hinrichs, K-U.; (2006) Heterotrophic Archaea dominate sedimentary subsurface ecosystems off Peru, *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 10, 3846-3851.

Birgel, D.; Peckmann, J.; (2008a) Aerobic methanotrophy at ancient marine methane seeps: A synthesis, *Organic Geochemistry*, **39**, 12, 1659-1667.

Birgel, D.; Elvert, M.; Han, X. Q.; Peckmann, J.; (2008b) ^{13}C-depleted biphytanic diacids as tracers of past anaerobic oxidation of methane, *Organic Geochemistry*, **39**, 1, 152-156.

Blaga, C. I.; Reichart, G-J.; Heiri, O.; Damsté, J. S. S.; (2009) Tetraether membrane lipid distributions in water-column particulate matter and sediments: a study of 47 European lakes along a north-south transect, *Journal of Paleolimnology*, **41**, 3, 523-540.

Blaga, C. I.; Reichart, G-J.; Schouten, S.; Lotter, A. F.; Werne, J. P.; Kosten, S.; Mazzeo, N.; Lacerot, G.; Damsté, J. S. S.; (2010) Branched glycerol dialkyl glycerol tetraethers in lake sediments: can they be used as temperature and pH proxies? *Organic Geochemistry*, in press.

Blank, C. E.; (2009) Not so old Archaea - the antiquity of biogeochemical processes in the archaeal domain of life, *Geobiology*, **7**, 5, 495-514.

Blau, K.; Halket, J.; (1993) Handbook of derivatives for chromatography, 2nd edition, Wiley, Chichester.

Blumenberg, M.; Seifert, R.; Reitner, J.; Pape, T.; Michaelis, W.; (2004) Membrane lipid patterns typify distinct anaerobic methanotrophic consortia, *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 30, 11111-11116.

Blumenberg, M.; Seifert, R.; Petersen, S.; Michaelis, W.; (2007) Biosignatures present in a hydrothermal massive sulfide from the Mid-Atlantic Ridge, *Geobiology*, **5**, 4, 435-450.

Bode, M. L.; Buddoo, S. R.; Minnaar, S. H.; du Plessis, C. A.; (2008) Extraction, isolation and NMR data of the tetraether lipid calditoglycerocaldarchaeol (GDNT) from *Sulfolobus metallicus* harvested from a bioleaching reactor, *Chemistry and Physics of Lipids*, **154**, 2, 94-104.

Boetius, A.; Ravenschlag, K.; Schubert, C. J.; Rickert, D.; Widdel, F.; Gieseke, A.; Amann, R.; Jørgensen, B. B.; Witte, U.; Pfannkuche, O.; (2000) A marine microbial consortium apparently mediating anaerobic oxidation of methane, *Nature*, **407**, 6804, 623-626.

Bouloubassi, I.; Nabais, E.; Pancost, R. D.; Lorre, A.; Taphanel, M-H.; (2009) First biomarker evidence for methane oxidation at cold seeps in the Southeast Atlantic (REGAB pockmark), *Deep-Sea Research Part II - Topical Studies in Oceanography*, **56**, 23, 2239-2247.

Boumann, H. A.; Hopmans, E. C.; van de Leemput, I.; Op den Camp, H. J. M.; van de Vossenberg, J.; Strous, M.; Jetten, M. S. M.; Sinninghe Damsté, J. S.; Schouten, S.; (2006) Ladderane phospholipids in anammox bacteria comprise phosphocholine and phosphoethanolamine headgroups, *FEMS Microbiology Letters*, **258**, 2, 297-304.

Bradley, A. S.; Fredricks, H.; Hinrichs, K-U.; Summons, R. E.; (2009a) Structural diversity of diether lipids in carbonate chimneys at the Lost City Hydrothermal Field, *Organic Geochemistry*, **40**, 12, 1169-1178.

Bradley, A. S.; Hayes, J. M.; Summons, R. E.; (2009b) Extraordinary ^{13}C enrichment of diether lipids at the Lost City Hydrothermal Field indicates a carbon-limited ecosystem, *Geochimica Et Cosmochimica Acta*, **73**, 1, 102-118.

Brassell, S. C.; Eglinton, G.; Marlowe, I. T.; Pflaumann, U.; Sarnthein, M.; (1986) Molecular Stratigraphy: a New Tool for Climatic Assessment, *Nature*, **320**, 6058, 129-133.

Brochier, C.; Gribaldo, S.; Zivanovic, Y.; Confalonieri, F.; Forterre, P.; (2005) Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales?, *Genome Biology*, **6**, R42, 1-10.

Brochier-Armanet, C.; Boussau, B.; Gribaldo, S.; Forterre, P.; (2008) Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota, *Nature Reviews Microbiology*, **6**, 3, 245-52.

Brock, T. D.; Brock, M. L.; (1966) Temperature optima for algal development in Yellowstone and Iceland hot springs, *Nature, London*, **209**, 733-734.

Brock, T. D.; Brock, K. M.; Belly, R. T.; Weiss, R. L.; (1972) *Sulfolobus*: A New Genus of Sulfur-Oxidizing Bacteria Living at Low pH and High-Temperature, *Archiv Fur Mikrobiologie*, **84**, 1, 54-68.

Bruins, A. P.; (1998) Mechanistic aspects of electrospray ionization, *Journal of Chromatography A*, **794**, 1-2, 345-357.

Burhan, R. Y. P.; Trendel, J. M.; Adam, P.; Wehrung, P.; Albrecht, P.; Nissenbaum, A.; (2002) Fossil bacterial ecosystem at methane seeps: Origin of organic matter from Be'eri sulfur deposit, Israel, *Geochimica Et Cosmochimica Acta*, **66**, 23, 4085-4101.

Byrdwell, W. C.; (2001) Atmospheric pressure chemical ionization mass spectrometry for analysis of lipids, *Lipids*, **36**, 4, 327-346.

Caillon, E.; Lubochinsky, B.; Rigomier, D.; (1983) Occurrence of Dialkyl Ether Phospholipids in *Stigmatella aurantiaca* DW4, *Journal of Bacteriology*, **153**, 3, 1348-1351.

Carballeira, N. M.; Reyes, M.; Sostre, A.; Huang, H.; Verhagen, M. F. J. M.; Adams, M. W. W.; (1997) Unusual fatty acid compositions of the hyperthermophilic archaeon *Pyrococcus furiosus* and the bacterium *Thermotoga maritima*, *Journal of Bacteriology*, **179**, 8, 2766-2768.

Chang, E. L.; (1992) US Patent, no. 5,098,588.

Chen, L.; Brügger, K.; Skovgaard, M.; Redder, P.; She, Q.; Torarinsson, E.; Greve, B.; Awayez, M.; Zibat, A.; Klenk, H-P.; Garrett, R. A.; (2005) The genome of *Sulfolobus acidocaldarius*, a model organism of the *Crenarchaeota*, *Journal of Bacteriology*, **187**, 14, 4992-4999.

Chevalier, N.; Bouloubassi, I.; Stadnitskaia, A.; Taphanel, M-H.; Lorre, A.; Damsté, J. S.; Pierre, C.; (2010) Distributions and carbon isotopic compositions of lipid biomarkers in authigenic carbonate crusts from the Nordic margin (Norwegian Sea), *Organic Geochemistry*, **41**, 9, 885-890.

Clarke, N. G.; Hazlewood, G. P.; Dawson, R. M. C.; (1980) Structure of diabolic acid-containing phospholipids isolated from *Butyrivibrio* sp., *Biochemical Journal*, **191**, 2, 561-569.

Cole, R. B.; (2000) Some tenets pertaining to electrospray ionization mass spectrometry, *Journal of Mass Spectrometry*, **35**, 7, 763-772.

Comita, P. B.; Gagosian, R. B.; (1983) Membrane Lipid from Deep-Sea Hydrothermal Vent Methanogen: a New Macrocyclic Glycerol Diether, *Science*, **222**, 4630, 1329-1331.

Comita, P. B.; Gagosian, R. B.; Pang, H.; Costello, C. E.; (1984) Structural Elucidation of a Unique Macrocyclic Membrane Lipid from a New, Extremely Thermophilic, Deep-Sea Hydrothermal Vent Archaebacterium, *Methanococcus jannaschii*, *Journal of Biological Chemistry*, **259**, 24, 5234-5241.

Cunningham, C.; Glish, G. L.; Burinsky, D. J.; (2006) High amplitude short time excitation: A method to form and detect low mass product ions in a quadrupole ion trap mass spectrometer, *Journal of the American Society for Mass Spectrometry*, **17**, 1, 81-84.

Dale, J. A.; Mosher, H. S.; (1973) Nuclear magnetic resonance enantiomer regents. Configurational correlations *via* nuclear magnetic resonance chemical shifts of diastereomeric mandelate, O-methylmandelate, and α-methoxy-α-trifluoromethylphenylacetate (MTPA) esters, *Journal of the American Chemical Society*, **95**, 2, 512-519.

Damsté, J. S. S.; Frewin, N. L.; Kenig, F.; De Leeuw, J. W.; (1995) Molecular Indicators for Paleoenvironmental Change in a Messinian Evaporitic Sequence

(Vena-Del-Gesso, Italy). 1: Variations in Extractable Organic Matter of ten Cyclically Deposited Marl Beds, *Organic Geochemistry*, **23**, 6, 471-483.

Damsté, J. S. S.; Hopmans, E. C.; Pancost, R. D.; Schouten, S.; Geenevasen, J. A. J.; (2000) Newly discovered non-isoprenoid glycerol dialkyl glycerol tetraether lipids in sediments, *Chemical Communications*, 17, 1683-1684.

Damsté, J. S. S.; Rijpstra, W. I. C.; Hopmans, E. C.; Prahl, F. G.; Wakeham, S. G.; Schouten, S.; (2002a) Distribution of membrane lipids of planktonic *Crenarchaeota* in the Arabian sea, *Applied and Environmental Microbiology*, **68**, 6, 2997-3002.

Damsté, J. S. S.; Schouten, S.; Hopmans, E. C.; van Duin, A. C. T.; Geenevasen, J. A. J.; (2002b) Crenarchaeol: the characteristic core glycerol dibiphytanyl glycerol tetraether membrane lipid of cosmopolitan pelagic crenarchaeota, *Journal of Lipid Research*, **43**, 10, 1641-1651.

Damsté, J. S. S.; Strous, M.; Rijpstra, W. I. C.; Hopmans, E. C.; Geenevasen, J. A. J.; van Duin, A. C. T.; van Niftrik, L. A.; Jetten, M. S. M.; (2002c) Linearly concatenated cyclobutane lipids form a dense bacterial membrane, *Nature*, **419**, 6908, 708-712.

Damsté, J. S. S.; Rijpstra, W. I. C.; Strous, M.; Jetten, M. S. M.; David, O. R. P.; Geenevasen, J. A. J.; van Maarseveen, J. H.; (2004) A mixed ladderane/*n*-alkyl glycerol diether membrane lipid in an anaerobic ammonium-oxidizing bacterium, *Chemical Communications (Cambridge)*, 22, 2590-2591.

Damsté, J. S. S.; Rijpstra, W. I. C.; Geenevasen, J. A. J.; Strous, M.; Jetten, M. S. M.; (2005) Structural identification of ladderane and other membrane lipids of planctomycetes capable of anaerobic ammonium oxidation (anammox), *FEBS Journal*, **272**, 16, 4270-4283.

Damsté, J. S. S.; Rijpstra, W. I. C.; Hopmans, E. C.; Schouten, S.; Balk, M.; Stams, A. J. M.; (2007) Structural characterization of diabolic acid-based tetraester, tetraether and mixed ether/ester, membrane-spanning lipids of bacteria from the order *Thermotogales*, *Archives of Microbiology*, **188**, 6, 629-641.

Damsté, J. S. S.; Ossebaar, J.; Schouten, S.; Verschuren, D.; (2008) Altitudinal shifts in the branched tetraether lipid distribution in soil from Mt. Kilimanjaro (Tanzania): Implications for the MBT/CBT continental palaeothermometer, *Organic Geochemistry*, **39**, 8, 1072-1076.

Damsté, J. S. S.; Ossebaar, J.; Abbas, B.; Schouten, S.; Verschuren, D.; (2009) Fluxes and distribution of tetraether lipids in an equatorial African lake: Constraints on the application of the TEX_{86} palaeothermometer and BIT index in lacustrine settings, *Geochimica Et Cosmochimica Acta*, **73**, 14, 4232-4249.

Dawson, P. H.; (1986) Quadrupole Mass Analyzers: Performance, Design and Some Recent Applications, *Mass Spectrometry Reviews*, **5**, 1, 1-37.

De Boever, E.; Birgel, D.; Thiel, V.; Muchez, P.; Peckmann, J.; Dimitrov, L.; Swennen, R.; (2009) The formation of giant tubular concretions triggered by anaerobic oxidation of methane as revealed by archaeal molecular fossils (Lower Eocene, Varna, Bulgaria), *Palaeogeography Palaeoclimatology Palaeoecology*, **280**, 1-2, 23-36.

de Hoffmann, E.; Stroobant, V.; (2002) Mass Spectrometry: Principles and Applications, 2nd edition. Wiley, Chichester.

de la Torre, J. R.; Walker, C. B.; Ingalls, A. E.; Könneke, M.; Stahl, D. A.; (2008) Cultivation of a thermophilic ammonia oxidizing archaeon synthesizing crenarchaeol, *Environmental Microbiology*, **10**, 3, 810-818.

De Rosa, M.; Gambacorta, A.; Bu'lock, J. D.; (1975) Extremely Thermophilic Acidophilic Bacteria Convergent with *Sulfolobus acidocaldarius*, *Journal of General Microbiology*, **86**, 156-164.

De Rosa, M.; De Rosa, S.; Gambacorta, A.; Bu'lock, J. D.; (1977a) Lipid Structures in *Caldariella* Group of Extreme Thermoacidophile Bacteria, *Journal of the Chemical Society-Chemical Communications*, 15, 514-515.

De Rosa, M.; De Rosa, S.; Gambacorta, A.; Minale, L.; Bu'lock, J. D.; (1977b) Chemical Structure of Ether Lipids of Thermophilic Acidophilic Bacteria of the *Caldariella* Group, *Phytochemistry*, **16**, 12, 1961-1965.

De Rosa, M.; De Rosa, S.; Gambacorta, A.; (1977c) ^{13}C-NMR assignments and biosynthetic data for the ether lipids of *Caldariella*, *Phytochemistry*, **16**, 12, 1909-1912.

De Rosa, M.; De Rosa, S.; Gambacorta, A.; Bu'lock, J. D.; (1980a) Structure of Calditol, a New Branched-Chain Nonitol, and of the Derived Tetraether Lipids in Thermoacidophile Archaebacteria of the *Caldariella* Group, *Phytochemistry*, **19**, 2, 249-254.

De Rosa, M.; Esposito, E.; Gambacorta, A.; Nicolaus, B.; Bu'lock, J. D.; (1980b) Effects of Temperature on Ether Lipid Composition of *Caldariella acidophila*, *Phytochemistry*, **19**, 5, 827-831.

De Rosa, M.; Gambacorta, A.; Nicolaus, B.; Sodano, S.; Bu'lock, J. D.; (1980c) Structural Regularities in Tetraether Lipids of *Caldariella* and Their Biosynthetic and Phyletic Implications, *Phytochemistry*, **19**, 5, 833-836.

De Rosa, M.; Gambacorta, A.; Nicolaus, B.; Ross, H. N. M.; Grant, W. D.; Bu'Lock, J. D.; (1982) An Asymmetric Archaebacterial Diether Lipid from Alkaliphilic Halophiles, *Journal of General Microbiology*, **128**, 2, 343-348.

De Rosa, M.; Gambacorta, A.; Nicolaus, B.; Grant, W. D.; (1983a) A C_{25},C_{25} Diether Core Lipid from Archaebacterial Haloalkaliphiles, *Journal of General Microbiology*, **129**, 8, 2333-2338.

De Rosa, M.; Gambacorta, A.; Nicolaus, B.; Chappe, B.; Albrecht, P.; (1983b) Isoprenoid Ethers: Backbone of Complex Lipids of the Archaebacterium *Sulfolobus solfataricus*, *Biochimica Et Biophysica Acta*, **753**, 2, 249-256.

De Rosa, M.; Gambacorta, A.; Lanzotti, V.; Trincone, A.; Harris, J. E.; Grant, W. D.; (1986) A Range of Ether Core Lipids from the Methanogenic Archaebacterium *Methanosarcina barkeri*, *Biochimica et Biophysica Acta*, **875**, 3, 487-492.

De Rosa, M.; Gambacorta, A.; (1988a) The Lipids of Archaebacteria, *Progress in Lipid Research*, **27**, 3, 153-175.

De Rosa, M.; Gambacorta, A.; Huber, R.; Lanzotti, V.; Nicolaus, B.; Stetter, K. O.; Trincone, A.; (1988b) A New 15,16-Dimethyl-30-Glyceroxytriacontanoic Acid from Lipids of *Thermotoga maritima*, *Journal of the Chemical Society Chemical Communications*, 19, 1300-1301.

de Souza, L. M.; Müller-Santos, M.; Iacomini, M.; Gorin, P. A. J.; Sassaki, G. L.; (2009) Positive- and negative-tandem mass spectrometric fingerprints of lipids from the halophilic Archaea *Haloarcula marismortui*, *Journal of Lipid Research*, **50**, 7, 1363-1373.

DeLong, E. F.; (1992) Archaea in coastal marine environments, *Proceedings of the National Academy of Sciences of the United States of America*, **89**, 12, 5685-5689.

DeLong, E. F.; King, L. L.; Massana, R.; Cittone, H.; Murray, A.; Schleper, C.; Wakeham, S. G.; (1998) Dibiphytanyl ether lipids in nonthermophilic Crenarchaeotes, *Applied and Environmental Microbiology*, **64**, 3, 1133-1138.

DeLong, E. F.; Taylor, L. T.; Marsh, T. L.; Preston, C. M.; (1999) Visualization and enumeration of marine planktonic archaea and bacteria by using polyribonucleotide probes and fluorescent in situ hybridization, *Applied and Environmental Microbiology*, **65**, 12, 5554-5563.

DeLong, E. F.; (2006) Archaeal mysteries of the deep revealed, *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 17, 6417-6418.

DeLong, E. F.; (2007) Microbial Domains in the Ocean: A Lesson from the Archaea, *Oceanography*, **20**, 2, 124-129.

DeLuca, S. J.; Voorhees, K. J.; Langworthy, T. A.; Holzer, G.; (1986) Capillary Supercritical Fluid Chromatography of Archaebacterial Glycerol Tetraether Lipids, *Journal of High Resolution Chromatography & Chromatography Communications*, **9**, 3, 182-185.

Demizu, K.; Ohtsubo, S.; Kohno, S.; Miura, I.; Nishihara, M.; Koga, Y.; (1992) Quantitative Determination of Methanogenic Cells Based on Analysis of Ether-Linked Glycerolipids by High-Performance Liquid-Chromatography, *Journal of Fermentation and Bioengineering*, **73**, 2, 135-139.

Ding, X.; Yang, W-J.; Min, H.; Peng, X-T.; Zhou, H-Y.; Lu, Z-M.; (2010) Isolation and characterization of a new strain of *Methanothermobacter marburgensis* DX01 from hot springs in China, *Anaerobe*, **16**, 1, 54-59.

Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B.; (1968) Molecular Beams of Macroions, *Journal of Chemical Physics*, **49**, 5, 2240-2249.

Donders, T. H.; Weijers, J. W. H.; Munsterman, D. K.; Kloosterboer-van Hoeve, M. L.; Buckles, L. K.; Pancost, R. D.; Schouten, S.; Damsté, J. S. S.; Brinkhuis, H.; (2009) Strong climate coupling of terrestrial and marine environments in the Miocene of northwest Europe, *Earth and Planetary Science Letters*, **281**, 3-4, 215-225.

Druckman Y.; Weissbrod T.; Aharon P.; (1994) Evidence for methane and hydrogen sulfide venting imprinted on a Quaternary eolianite from southern Israel. *Geo-Marine Letters*, **14**, 170–176.

Dumitrescu, M.; Brassell, S. C.; Schouten, S.; Hopmans, E. C.; Damsté, J. S. S.; (2006) Instability in tropical Pacific sea-surface temperatures during the early Aptian, *Geology*, **34**, 10, 833-836.

Egorova, K.; Antranikian, G.; (2005) Industrial relevance of thermophilic Archaea, *Current Opinion in Microbiology*, **8**, 6, 649-655.

Eguchi, T.; Takyo, H.; Morita, M.; Kakinuma, K.; Koga, Y.; (2000) Unusual double-bond migration as a plausible key reaction in the biosynthesis of the isoprenoidal membrane lipids of methanogenic archaea, *Chemical Communications*, 16, 1545-1546.

Eguchi, T.; Nishimura, Y.; Kakinuma, K.; (2003) Importance of the isopropylidene terminal of geranylgeranyl group for the formation of tetraether lipid in methanogenic archaea, *Tetrahedron Letters*, **44**, 16, 3275-3279.

Ekiel, I.; Sprott, G. D.; (1992) Identification of degradation artefacts formed upon treatment of hydroxydiether lipids from methanogens with methanolic HCl, *Canadian Journal of Microbiology*, **38**, 764–768.

Elderfield, H.; Ganssen, G.; (2000) Past temperature and $\delta^{18}O$ of surface ocean waters inferred from foraminiferal Mg/Ca ratios, *Nature*, **405**, 6785, 442-445.

Elkins, J. G.; Podar, M.; Graham, D. E.; Makarova, K. S.; Wolf, Y.; Randau, L.; Hedlund, B. P.; Brochier-Armanet, C.; Kunin, V.; Anderson, I.; Lapidus, A.; Goltsman, E.; Barry, K.; Koonin, E. V.; Hugenholtz, P.; Kyrpides, N.; Wanner, G.; Richardson, P.; Keller, M.; Stetter, K. O.; (2008) A korarchaeal genome reveals insights into the evolution of the Archaea, *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 23, 8102-8107.

Ellen, A. F.; Albers, S-V.; Huibers, W.; Pitcher, A.; Hobel, C. F. V.; Schwarz, H.; Folea, M.; Schouten, S.; Boekema, E. J.; Poolman, B.; Driessen, A. J. M.; (2009)

Proteomic analysis of secreted membrane vesicles of archaeal *Sulfolobus* species reveals the presence of endosome sorting complex components, *Extremophiles*, **13**, 1, 67-79.

Epstein S.; Mayeda, T. K.; (1953) Variations of the $^{18}O/^{16}O$ ratio in natural waters, *Geochimica et Cosmochimica Acta*, **4**, 213-224.

Erez, J.; Luz, B.; (1983) Experimental paleotemperature equation for planktonic foraminifera, *Geochimica et Cosmochimica Acta*, **47**, 6, 1025-1031.

Escala, M.; Rosell-Melé, A.; Masqué, P.; (2007) Rapid screening of glycerol dialkyl glycerol tetraethers in continental Eurasia samples using HPLC/APCI-ion trap mass spectrometry, *Organic Geochemistry*, **38**, 1, 161-164.

Escala, M.; Fietz, S.; Rueda, G.; Rosell-Mele, A.; (2009) Analytical Considerations for the Use of the Paleothermometer Tetraether $Index_{86}$ and the Branched vs Isoprenoid Tetraether Index Regarding the Choice of Cleanup and Instrumental Conditions, *Analytical Chemistry*, **81**, 7, 2701-2707.

Evans, C. S.; Startin, J. R.; Goodall, D. M.; Keely, B. J.; (2000) Optimisation of ion trap parameters for the quantification of chlormequat by liquid chromatography/mass spectrometry and the application in the analysis of pear extracts, *Rapid Communications in Mass Spectrometry*, **14**, 2, 112-117.

Fauquette, S.; Suc, J-P.; Bertini, A.; Popescu, S-M.; Warny, S.; Taoufiq, N. B.; Villa, M-J. P.; Chikhi, H.; Feddi, N.; Subally, D.; Clauzon, G.; Ferrier, J.; (2006) How much did climate force the Messinian salinity crisis? Quantified climatic conditions from pollen records in the Mediterranean region, *Palaeogeography Palaeoclimatology Palaeoecology*, **238**, 1-4, 281-301.

Fenn, J. B.; (2002) Electrospray ionization mass spectrometry: How it all began, *Journal of Biomolecular Techniques*, **13**, 3, 101-118.

Ferrante, G.; Ekiel, I.; Patel, G. B.; Sprott, G. D.; (1988) A Novel Core Lipid Isolated from the Aceticlastic Methanogen, *Methanothrix concilii* GP6, *Biochimica Et Biophysica Acta*, **963**, 2, 173-182.

Ferrante, G.; Brisson, J-R.; Patel, G. B.; Ekiel, I.; Sprott, G. D.; (1989) Structures of Minor Ether Lipids Isolated from the Aceticlastic Methanogen, *Methanothrix concilii* GP6, *Journal of Lipid Research*, **30**, 10, 1601-1609.

Forster, A.; Schouten, S.; Baas, M.; Damsté, J. S. S.; (2007) Mid-Cretaceous (Albian–Santonian) sea surface temperature record of the tropical Atlantic Ocean, *Geology*, **35**, 10, 919-922.

Fox, G. E.; Stackebrandt, E.; Hespell, R. B.; Gibson, J.; Maniloff, J.; Dyer, T. A.; Wolfe, R. S.; Balch, W. E.; Tanner, R. S.; Magrum, L. J.; Zablen, L. B.; Blakemore, R.; Gupta, R.; Bonen, L.; Lewis, B. J.; Stahl, D. A.; Luehrsen, K. R.; Chen, K. N.; Woese, C. R.; (1980) The phylogeny of prokaryotes, *Science*, **209**, 4455, 457-463.

Francis, C. A.; Beman, J. M.; Kuypers, M. M. M.; (2007) New processes and players in the nitrogen cycle: the microbial ecology of anaerobic and archaeal ammonia oxidation, *ISME Journal*, **1**, 1, 19-27.

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

Fries, E.; Püttmann, W.; (2002) Analysis of the antioxidant butylated hydroxytoluene (BHT) in water by means of solid phase extraction combined with GC/MS, *Water Research*, **36**, 9, 2319-2327.

Fuhrman, J. A.; McCallum, K.; Davis, A. A.; (1992) Novel major archaebacterial group from marine plankton, *Nature*, **356**, 6365, 148-149.

Gabriel, J. L.; Chong, P. L. G; (2000) Molecular modeling of archaebacterial bipolar tetraether lipid membranes, *Chemistry and Physics of Lipids*, **105**, 193-200.

Galliker, P.; Gräther, O.; Rümmler, M.; Fitz, W.; Arigoni, D.; (1998) New Structural and Biosynthetic Aspects of the Unusual Core Lipids from Archaebacteria. In *Vitamin B12 and B12-Proteins*, Kräutler, B., Arigoni, D., Golding, B. T., Eds.; Wiley VCH, Hoboken, p. 447-458.

Gattinger, A.; Günther, A.; Schloter, M.; Munich, J. C.; (2003) Characterisation of *Archaea* in soils by polar lipid analysis, *Acta Biotechnologica*, **23**, 1, 21-28.

Gibson, J. A. E.; Miller, M. R.; Davies, N. W.; Neill, G. P.; Nichols, D. S.; Volkman, J. K.; (2005) Unsaturated diether lipids in the psychrotrophic archaeon *Halorubrum lacusprofundi*, *Systematic and Applied Microbiology*, **28**, 1, 19-26.

Gliozzi, A.; Relini, A.; Chong, P. L-G.; (2002) Structure and permeability properties of biomimetic membranes of bolaform archaeal tetraether lipids, *Journal of Membrane Science*, **206**, 1-2, 131-147.

Gonthier, I.; Rager, M-N.; Metzger, P.; Guezennec, J.; Largeau, C.; (2001) A di-O-dihydrogeranylgeranyl glycerol from Thermococcus S 557, a novel ether lipid, and likely intermediate in the biosynthesis of diethers in Archaea, *Tetrahedron Letters*, **42**, 15, 2795-2797.

Good, A.; Durden, D. A.; Kebarle, P.; (1970) Ion-Molecule Reactions in Pure Nitrogen and Nitrogen Containing Traces of Water at Total Pressures 0.5-4 Torr - Kinetics of Clustering Reactions Forming $H^+(H_2O)_n$, *Journal of Chemical Physics*, **52**, 1, 212-221.

Grant, W. D.; Pinch, G.; Harris, J. E.; De Rosa, M.; Gambacorta, A.; (1985) Polar Lipids in Methanogen Taxonomy, *Journal of General Microbiology*, **131**, 3277-3286.

Grant, W. D.; Ross, H. N. M.; (1986) The ecology and taxonomy of halobacteria, *FEMS Microbiology Letters*, **39**, 1-2, 9-15.

Gräther, O.; Arigoni, D.; (1995) Detection of Regioisomeric Macrocyclic Tetraethers in the Lipids of *Methanobacterium thermoautotrophicum* and Other Archaeal Organisms, *Journal of the Chemical Society-Chemical Communications*, 4, 405-406.

Grogan, D. W.; (1989) Phenotypic Characterization of the Archaebacterial Genus *Sulfolobus*: Comparison of Five Wild-Type Strains, *Journal of Bacteriology*, **171**, 12, 6710-6719.

Grogan, D.; Palm, P.; Zillig, W.; (1990) Isolate B12, Which Harbors a Virus-Like Element, Represents a New Species of the Archaebacterial Genus *Sulfolobus*, *Sulfolobus shibatae*, sp. nov., *Archives of Microbiology*, **154**, 6, 594-599.

Guilhaus, M.; Selby, D.; Mlynski, V.; (2000) Orthogonal acceleration time-of-flight mass spectrometry, *Mass Spectrometry Reviews*, **19**, 2, 65-107.

Gulik, A.; Luzzati, V.; De Rosa, M.; Gambacorta, A.; (1988) Tetraether Lipid Components from a Thermoacidophilic Archaebacterium – Chemical Structure and Physical Polymorphism, *Journal of Molecular Biology*, **201**, 2, 429-435.

Guy, C. P.; Majernik, A. I.; Chong, J. P. J.; Bolt, E. L.; (2004) A novel nuclease-ATPase (Nar71) from archaea is part of a proposed thermophilic DNA repair system, *Nucleic Acids Research*, **32**, 21, 6176-6186.

Hafenbradl, D.; Keller, M.; Thiericke, R.; Stetter, K. O.; (1993) A novel unsaturated archaeal ether core lipid from the hyperthermophile *Methanopyrus kandleri*, *Systematic and Applied Microbiology*, **16**, 165–169.

Harrison, A. G.; (1999) Energy-resolved mass spectrometry: a comparison of quadrupole cell and cone-voltage collision-induced dissociation, *Rapid Communications in Mass Spectrometry*, **13**, 16, 1663-1670.

Harvey, H. R.; Fallon, R. D.; Patton, J. S.; (1986) The effect of organic matter and oxygen on the degradation of bacterial membrane lipids in marine sediments, *Geochimica et Cosmochimica Acta*, **50**, 795-804.

Hedrick, D. B.; Guckert, J. B.; White, D. C.; (1991) Archaebacterial Ether Lipid Diversity Analyzed by Supercritical Fluid Chromatography: Integration with a Bacterial Lipid Protocol, *Journal of Lipid Research*, **32**, 4, 659-666.

Herfort, L.; Schouten, S.; Boon, J. P.; Damsté, J. S. S.; (2006a) Application of the TEX_{86} temperature proxy to the southern North Sea, *Organic Geochemistry*, **37**, 12, 1715-1726.

Herfort, L.; Schouten, S.; Boon, J. P.; Woltering, M.; Baas, M.; Weijers, J. W. H.; Damsté, J. S. S.; (2006b) Characterization of transport and deposition of terrestrial organic matter in the southern North Sea using the BIT index, *Limnology and Oceanography*, **51**, 5, 2196-2205.

Herndl, G. J.; Reinthaler, T.; Teira, E.; van Aken, H.; Veth, C.; Pernthaler, A.; Pernthaler, J.; (2005) Contribution of *Archaea* to total prokaryotic production in the deep Atlantic Ocean, *Applied and Environmental Microbiology*, **71**, 5, 2303-9.

Hinrichs, K-U.; Hayes, J. M.; Sylva, S. P.; Brewer, P. G.; DeLong, E. F.; (1999) Methane-consuming archaebacteria in marine sediments, *Nature*, **398**, 6730, 802-805.

Hinrichs, K-U.; Summons, R. E.; Orphan, V.; Sylva, S. P.; Hayes, J. M.; (2000) Molecular and isotopic analysis of anaerobic methane-oxidizing communities in marine sediments, *Organic Geochemistry*, **31**, 12, 1685-1701.

Hoefs, M. J. L.; Schouten, S.; De Leeuw, J. W.; King, L. L.; Wakeham, S. G.; Damsté, J. S. S.; (1997) Ether lipids of planktonic archaea in the marine water column, *Applied and Environmental Microbiology*, **63**, 8, 3090-3095.

Hoehler, T. M.; Alperin, M. J.; Albert, D. B.; Martens, C. S.; (1994) Field and laboratory studies of methane oxidation in an anoxic marine sediment: Evidence for a methanogen-sulfate reducer consortium. *Global Biogeochemical Cycles*, **8**, 451-463.

Hofmann, P.; Stüsser, I.; Wagner, T.; Schouten, S.; Damsté, J. S. S.; (2008) Climate-ocean coupling off North-West Africa during the Lower Albian: The Oceanic Anoxic Event 1b, *Palaeogeography Palaeoclimatology Palaeoecology*, **262**, 3-4, 157-165.

Hopmans, E. C.; Schouten, S.; Pancost, R. D.; van der Meer, M. T. J.; Damsté, J. S. S.; (2000) Analysis of intact tetraether lipids in archaeal cell material and sediments by high performance liquid chromatography/atmospheric pressure chemical ionization mass spectrometry, *Rapid Communications in Mass Spectrometry*, **14**, 7, 585-589.

Hopmans, E. C.; Weijers, J. W. H.; Schefuß, E.; Herfort, L.; Sinninghe Damsté, J. S.; Schouten, S.; (2004) A novel proxy for terrestrial organic matter in sediments based on branched and isoprenoid tetraether lipids, *Earth and Planetary Science Letters*, **224**, 1-2, 107-116.

Horning, E. C.; Horning, M. G.; Carroll, D. I.; Dzidic, I.; Stillwell, R. N.; (1973) New Picogram Detection System Based on a Mass Spectrometer with an External Ionization Source at Atmospheric Pressure, *Analytical Chemistry*, **45**, 6, 936-943.

Huber, R.; Langworthy, T. A.; König, H.; Thomm, M.; Woese, C. R.; Sleytr, U. B.; Stetter, K. O.; (1986) *Thermotoga maritima* sp. nov. Represents a New Genus of Unique Extremely Thermophilic Eubacteria Growing up to 90^{o}C, *Archives of Microbiology*, **144**, 4, 324-333.

Huber, R.; Wilharm, T.; Huber, D.; Trincone, A.; Burggraf, S.; König, H.; Rachel, R.; Rockinger, I.; Fricke, H.; Stetter, K. O.; (1992) *Aquifex pyrophilus* gen. nov. sp. nov., represents a novel group of marine hyperthermophilic hydrogen-oxidizing bacteria, *Systematic and Applied Microbiology*, **15**, 340-351.

Huber, R.; Rossnagel, P.; Woese, C. R.; Rachel, R.; Langworthy, T. A.; Stetter, K. O.; (1996) Formation of ammonium from nitrate during chemolithoautotrophic growth of the extremely thermophilic bacterium *Ammonifex degensii* gen. nov. sp. nov., *Systematic and Applied Microbiology*, **19**, 40-49.

Huber, H.; Hohn, M. J.; Rachel, R.; Fuchs, T.; Wimmer, V. C.; Stetter, K. O.; (2002) A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont, *Nature*, **417**, 6884, 63-67.

Huguet, C.; Cartes, J. E.; Damsté, J. S. S.; Schouten, S.; (2006a) Marine crenarchaeotal membrane lipids in decapods: Implications for the TEX_{86} paleothermometer, *Geochemistry Geophysics Geosystems*, **7**, Q11010.

Huguet, C.; Hopmans, E. C.; Febo-Ayala, W.; Thompson, D. H.; Damsté, J. S. S.; Schouten, S.; (2006b) An improved method to determine the absolute abundance of glycerol dibiphytanyl glycerol tetraether lipids, *Organic Geochemistry*, **37**, 9, 1036-1041.

Huguet, C.; Schimmelmann, A.; Thunell, R.; Lourens, L. J.; Damsté, J. S. S.; Schouten, S.; (2007) A study of the TEX_{86} paleothermometer in the water column and sediments of the Santa Barbara Basin, California, *Paleoceanography*, **22**, 3, PA3203.

Huguet, C.; de Lange, G. J.; Gustafsson, Ö.; Middelburg, J. J.; Damsté, J. S. S.; Schouten, S.; (2008) Selective preservation of soil organic matter in oxidized marine sediments (Madeira Abyssal Plain), *Geochimica Et Cosmochimica Acta*, **72**, 24, 6061-6068.

Huguet, C.; Kim, J-H.; de Lange, G. J.; Damsté, J. S. S.; Schouten, S.; (2009) Effects of long term oxic degradation on the $U^{K'}_{37}$, TEX_{86} and BIT organic proxies, *Organic Geochemistry*, **40**, 12, 1188-1194.

Huguet, C.; Martens-Habbena, W.; Urakawa, H.; Stahl, D. A.; Ingalls, A. E.; (2010) Comparison of extraction methods for quantitative analysis of core and intact polar glycerol dialkyl glycerol tetraethers (GDGTs) in environmental samples, *Limnology and Oceanography-Methods*, **8**, 127-145.

Iribarne, J. V.; Thomson, B. A.; (1976) On the Evaporation of Small Ions from Charged Droplets, *Journal of Chemical Physics*, **64**, 6, 2287-2294.

Ishii, K.; Muβmann, M.; MacGregor, B. J.; Amann, R.; (2004) An improved fluorescence in situ hybridization protocol for the identification of bacteria and archaea in marine sediments, *FEMS Microbiology Ecology*, **50**, 3, 203-212.

Izard, C.; Libermann, C.; (1978) Acrolein, *Mutation Research*, **47**, 2, 115-138.

Jaeschke, A.; Rooks, C.; Trimmer, M.; Nicholls, J. C.; Hopmans, E. C.; Schouten, S.; Damsté, J. S. S.; (2009) Comparison of ladderane phospholipid and core lipids as indicators for anaerobic ammonium oxidation (anammox) in marine sediments, *Geochimica Et Cosmochimica Acta*, **73**, 7, 2077-2088.

Jaffrés, J. B. D.; Shields, G. A.; Wallmann, K.; (2007) The oxygen isotope evolution of seawater: A critical review of a long-standing controversy and an improved geological water cycle model for the past 3.4 billion years, *Earth-Science Reviews*, **83**, 1-2, 83-122.

Jagersma, G. C.; Meulepas, R. J. W.; Heikamp-de Jong, I.; Gieteling, J.; Klimiuk, A.; Schouten, S.; Damsté, J. S. S.; Lens, P. N. L.; Stams, A. J. M.; (2009) Microbial diversity and community structure of a highly active anaerobic methane-oxidizing sulfate-reducing enrichment, *Environmental Microbiology*, **11**, 12, 3223-32.

Jahnke, L. L.; Eder, W.; Huber, R.; Hope, J. M.; Hinrichs, K-U.; Hayes, J. M.; Des Marais, D. J.; Cady, S. L.; Summons, R. E.; (2001) Signature lipids and stable carbon isotope analyses of octopus spring hyperthermophilic communities compared with those of *Aquificales* representatives, *Applied and Environmental Microbiology*, **67**, 11, 5179-5189.

James, A. T.; Martin, A. J. P.; (1952) Gas-Liquid Partition Chromatography - a Technique for the Analysis of Volatile Materials, *Analyst*, **77**, 921, 915-932.

Jenkyns, H. C.; Forster, A.; Schouten, S.; Damsté, J. S. S.; (2004) High temperatures in the Late Cretaceous Arctic Ocean, *Nature*, **432**, 7019, 888-892.

Jones, W. J.; Leigh, J. A.; Mayer, F.; Woese, C. R.; Wolfe, R. S.; (1983) *Methanococcus jannaschii* sp. nov., an Extremely Thermophilic Methanogen from a Submarine Hydrothermal Vent, *Archives of Microbiology*, **136**, 4, 254-261.

Jonscher, K. R.; Yates, J. R.; (1997) The quadrupole ion trap mass spectrometer - A small solution to a big challenge, *Analytical Biochemistry*, **244**, 1, 1-15.

Joo, C. N.; Shier, T.; Kates, M.; (1968) Characterization and Synthesis of Mono- and Diphytanyl Ethers of Glycerol, *Journal of Lipid Research*, **9**, 6, 782-788.

Jung, S.; Zeikus, J. G.; Hollingsworth, R. I.; (1994) A New Family of Very Long-Chain α,ω-Dicarboxylic Acids Is a Major Structural Fatty Acyl Component of the Membrane Lipids of *Thermoanaerobacter ethanolicus* 39E, *Journal of Lipid Research*, **35**, 6, 1057-1065.

Kaneda, T.; (1991) Iso- and anteiso-fatty acids in bacteria: biosynthesis, function, and taxonomic significance, *Microbiological Reviews*, **55**, 2, 288-302.

Karner, M. B.; DeLong, E. F.; Karl, D. M.; (2001) Archaeal dominance in the mesopelagic zone of the Pacific Ocean, *Nature*, **409**, 6819, 507-510.

Karpas, Z.; Eiceman, G. A.; Ewing, R. G.; Harden, C. S.; (1994) Collision Induced Dissociation Studies of Protonated Alcohol and Alcohol-Water Clusters by Atmospheric Pressure Ionization Tandem Mass-Spectrometry. Part 2. Ethanol, Propanol and Butanol, *International Journal of Mass Spectrometry and Ion Processes*, **133**, 1, 47-58.

Kates, M.; Palameta, B.; Yengoyan, L. S.; (1965) Aliphatic Diether Analogs of Glyceride-Derived Lipids. II. Synthesis of Naturally Occurring L-2,3,-Di-*O*-3',7',11',15'-Tetramethylhexadecyl Glycerol and Its D Isomer, *Biochemistry*, **4**, 8, 1595-1599.

Kates, M.; Joo, C. N.; Palameta, B.; Shier, T.; (1967) Absolute Stereochemical Configuration of Phytanyl (Dihydrophytyl) Groups in Lipids of *Halobacterium cutirubrum*, *Biochemistry*, **6**, 11, 3329-3338.

Kates, M.; Ether Lipids in Extremely Halophilic Bacteria. In *Ether Lipids: Chemistry and Biology*, Snyder. F., Ed.; Academic Press: New York, (1972); p. 351.

Kaur, G.; Mountain, B. W.; Pancost, R. D.; (2008) Microbial membrane lipids in active and inactive sinters from Champagne Pool, New Zealand: Elucidating past geothermal chemistry and microbiology, *Organic Geochemistry*, **39**, 8, 1024-1028.

Keely, B. J.; Blake, S. R.; Schaeffer, P.; Maxwell, J. R.; (1995) Distributions of Pigments in the Organic Matter of Marls from the Vena Del Gesso Evaporitic Sequence, *Organic Geochemistry*, **23**, 6, 527-539.

Kelley, D. S.; Karson, J. A.; Früh-Green, G. L.; Yoerger, D. R.; Shank, T. M.; Butterfield, D. A.; Hayes, J. M.; Schrenk, M. O.; Olson, E. J.; Proskurowski, G.; Jakuba, M.; Bradley, A.; Larson, B.; Ludwig, K.; Glickson, D.; Buckman, K.; Bradley, A. S.; Brazelton, W. J.; Roe, K.; Elend, M. J.; Delacour, A.; Bernasconi, S. M.; Lilley, M. D.; Baross, J. A.; Summons, R. E.; Sylva, S. P.; (2005) A serpentinite-hosted ecosystem: The Lost City hydrothermal field, *Science*, **307**, 5714, 1428-1434.

Kenig, F.; Hayes, J. M.; Popp, B. N.; Summons, R. E.; (1994) Isotopic biogeochemistry of the Oxford Clay formation (Jurassic), UK, *Journal of the Geological Society*, **151**, 139-152.

Kenig, F.; Hudson, J. D.; Damsté, J. S. S.; Popp, B. N.; (2004) Intermittent euxinia: Reconciliation of a Jurassic black shale with its biofacies, *Geology*, **32**, 421-424.

Kim, J-H.; Schouten, S.; Buscail, R.; Ludwig, W.; Bonnin, J.; Damsté, J. S. S.; Bourrin, F.; (2006) Origin and distribution of terrestrial organic matter in the NW Mediterranean (Gulf of Lions): Exploring the newly developed BIT index, *Geochemistry Geophysics Geosystems*, **7**, Q11017.

Kim, J-H.; Ludwig, W.; Schouten, S.; Kerhervé, P.; Herfort, L.; Bonnin, J.; Damsté, J. S. S.; (2007) Impact of flood events on the transport of terrestrial organic matter to the ocean: A study of the Têt River (SW France) using the BIT index, *Organic Geochemistry*, **38**, 10, 1593-1606.

Kim, J-H.; Schouten, S.; Hopmans, E. C.; Donner, B.; Damsté, J. S. S.; (2008) Global sediment core-top calibration of the TEX_{86} paleothermometer in the ocean, *Geochimica Et Cosmochimica Acta*, **72**, 4, 1154-1173.

Kim, J-H.; Buscail, R.; Bourrin, F.; Palanques, A.; Damsté, J. S. S.; Bonnin, J.; Schouten, S.; (2009a) Transport and depositional process of soil organic matter during wet and dry storms on the Têt inner shelf (NW Mediterranean), *Palaeogeography Palaeoclimatology Palaeoecology*, **273**, 3-4, 228-238.

Kim, J-H.; Crosta, X.; Michel, E.; Schouten, S.; Duprat, J.; Damsté, J. S. S.; (2009b) Impact of lateral transport on organic proxies in the Southern Ocean, *Quaternary Research*, **71**, 2, 246-250.

Kim, J-H.; Huguet, C.; Zonneveld, K. A. F.; Versteegh, G. J. M.; Roeder, W.; Damsté, J. S. S.; Schouten, S.; (2009c) An experimental field study to test the stability of lipids used for the TEX_{86} and $U^{K'}_{37}$ palaeothermometers, *Geochimica Et Cosmochimica Acta*, **73**, 10, 2888-2898.

Kim, J-H.; van der Meer, J.; Schouten, S.; Helmke, P.; Willmott, V.; Sangiorgi, F.; Koç, N.; Hopmans, E. C.; Damsté, J. S. S.; (2010) New indices and calibrations derived from the distribution of crenarchaeal isoprenoid tetraether lipids: Implications for past sea surface temperature reconstructions, *Geochimica Et Cosmochimica Acta*, **74**, 16, 4639-4654.

Kinnaman, F. S.; Kimball, J. B.; Busso, L.; Birgel, D.; Ding, H.; Hinrichs, K-U.; Valentine, D. L.; (2010) Gas flux and carbonate occurrence at a shallow seep of thermogenic natural gas, *Geo-Marine Letters*, **30**, 3-4, 355-365.

Koga, Y.; Akagawa-Matsushita, M.; Ohga, M.; Nishihara, M.; (1993a) Taxonomic Significance of the Distribution of Component Parts of Polar Ether Lipids in Methanogens, *Systematic and Applied Microbiology*, **16**, 3, 342-351.

Koga, Y.; Nishihara, M.; Morii, H.; Akagawa-Matsushita, M.; (1993b) Ether polar lipids of methanogenic bacteria: structures, comparative aspects, and biosyntheses, *Microbiological Reviews*, **57**, 1, 164-182.

Koga, Y.; Morii, H.; Akagawa-Matsushita, M.; Ohga, M.; (1998) Correlation of polar lipid composition with 16S rRNA phylogeny in methanogens. Further analysis of lipid component parts, *Bioscience Biotechnology and Biochemistry*, **62**, 2, 230-236.

Koga, Y.; Morii, H.; (2005) Recent advances in structural research on ether lipids from archaea including comparative and physiological aspects, *Bioscience Biotechnology and Biochemistry*, **69**, 11, 2019-2034.

Koga, Y.; Morii, H.; (2006) Special methods for the analysis of ether lipid structure and metabolism in archaea, *Analytical Biochemistry*, **348**, 1, 1-14.

Koga, Y.; Morii, H.; (2007) Biosynthesis of ether-type polar lipids in archaea and evolutionary considerations, *Microbiology and Molecular Biology Reviews*, **71**, 1, 97-120.

Kon, T.; Nemoto, N.; Oshima, T.; Yamagishi, A.; (2002) Effects of a squalene epoxidase inhibitor, terbinafine, on ether lipid biosyntheses in a thermoacidophilic archaeon, *Thermoplasma acidophilum*, *Journal of Bacteriology*, **184**, 5, 1395-1401.

Könneke, M.; Bernhard, A. E.; de la Torre, J. R.; Walker, C. B.; Waterbury, J. B.; Stahl, D. A.; (2005) Isolation of an autotrophic ammonia-oxidizing marine archaeon, *Nature*, **437**, 7058, 543-546.

Kramer, J. K. G.; Sauer, F. D.; (1991) Changes in the Diether-to-Tetraether-Lipid Ratio During Cell Growth in *Methanobacterium thermoautotrophicum*, *FEMS Microbiology Letters*, **83**, 1, 45-50.

Krishnan, L.; Sprott, G. D.; (2008) Archaeosome adjuvants: Immunological capabilities and mechanism(s) of action, *Vaccine*, **26**, 17, 2043-2055.

Krüger, M.; Blumenberg, M.; Kasten, S.; Wieland, A.; Känel, L.; Klock, J-H.; Michaelis, W.; Seifert, R.; (2008) A novel, multi-layered methanotrophic microbial mat system growing on the sediment of the Black Sea, *Environmental Microbiology*, **10**, 8, 1934-1947.

Kuypers, M. M. M.; Blokker, P.; Hopmans, E. C.; Kinkel, H.; Pancost, R. D.; Schouten, S.; Damsté, J. S. S.; (2002) Archaeal remains dominate marine organic matter from the early Albian oceanic anoxic event 1b, *Palaeogeography Palaeoclimatology Palaeoecology*, **185**, 1-2, 211-234.

Kuypers, M. M. M.; Sliekers, A. O.; Lavik, G.; Schmid, M.; Jørgensen, B. B.; Kuenen, J. G.; Damsté, J. S. S.; Strous, M.; Jetten, M. S. M.; (2003) Anaerobic ammonium oxidation by anammox bacteria in the Black Sea, *Nature*, **422**, 6932, 608-611.

Lai, D.; Springstead, J. R.; Monbouquette, H. G.; (2008) Effect of growth temperature on ether lipid biochemistry in *Archaeoglobus fulgidus*, *Extremophiles*, **12**, 2, 271-278.

Langworthy, T. A.; (1977) Comparative Lipid Composition of Heterotrophically and Autotrophically Grown *Sulfolobus acidocaldarius*, *Journal of Bacteriology*, **130**, 3, 1326-1332.

Langworthy, T. A.; Tornabene, T. G.; Holzer, G.; (1982) Lipids of Archaebacteria, *Zentralblatt Fur Bakteriologie Mikrobiologie Und Hygiene I Abteilung Originale C-Allgemeine Angewandte Und Okologische Mikrobiologie*, **3**, 2, 228-244.

Langworthy, T. A.; Holzer, G.; Zeikus, J. G.; Tornabene, T. G.; (1983) Iso- and Anteiso-Branched Glycerol Diethers of the Thermophilic Anaerobe *Thermodesulfotobacterium commune*, *Systematic and Applied Microbiology*, **4**, 1, 1-17.

Lanzotti, V.; De Rosa, M.; Trincone, A.; Basso, A. L.; Gambacorta, A.; Zillig, W.; (1987) Complex Lipids from *Desulfurococcus mobilis*, a Sulfur-Reducing Archaebacterium, *Biochimica Et Biophysica Acta*, **922**, 2, 95-102.

Lee, S.; Kang, S.; Kim, J. N.; Jung, S.; (2002) Structural analyses of the novel phosphoglycolipids containing the unusual very long bifunctional acyl chain, α,ω-13,16-dimethyloctacosanedioate in *Thermoanaerobacter ethanolicus*, *Bulletin of the Korean Chemical Society*, **23**, 12, 1778-1784.

Lee, K. E.; Kim, J-H.; Wilke, I.; Helmke, P.; Schouten, S.; (2008) A study of the alkenone, TEX_{86}, and planktonic foraminifera in the Benguela Upwelling System: Implications for past sea surface temperature estimates, *Geochemistry Geophysics Geosystems*, **9**, Q10019.

Leininger, S.; Urich, T.; Schloter, M.; Schwark, L.; Qi, J.; Nicol, G. W.; Prosser, J. I.; Schuster, S. C.; Schleper, C.; (2006) Archaea predominate among ammonia-oxidizing prokaryotes in soils, *Nature*, **442**, 7104, 806-809.

Levsky, J. M.; Singer, R. H.; (2003) Fluorescence in situ hybridization: past, present and future, *Journal of Cell Science*, **116**, 14, 2833-2838.

Liefkens, W.; Boon, J. J.; De Leeuw, J. W.; (1979) On the Occurrence of Alkyl- and Alk-1-Enyl-Diacylglycerides in the Lugworm Arenicola Marina, *Netherlands Journal of Sea Research*, **13**, 3-4, 479-486.

Lin, H-Y.; Rockwood, A.; Munson, M. S. B.; Ridge, D. P.; (1993) Proton Affinity and Collision-Induced Decomposition of Ethoxylated Alcohols: Effects of Intramolecular Hydrogen-Bonding on Polymer Ion Collision-Induced Decomposition, *Analytical Chemistry*, **65**, 15, 2119-2124.

Lipp, J. S.; Morono, Y.; Inagaki, F.; Hinrichs, K-U.; (2008) Significant contribution of Archaea to extant biomass in marine subsurface sediments, *Nature*, **454**, 7207, 991-994.

Lipp, J. S.; Hinrichs, K-U.; (2009) Structural diversity and fate of intact polar lipids in marine sediments, *Geochimica Et Cosmochimica Acta*, **73**, 22, 6816-6833.

Liu, Z.; Pagani, M.; Zinniker, D.; DeConto, R.; Huber, M.; Brinkhuis, H.; Shah, S. R.; Leckie, R. M.; Pearson, A.; (2009) Global Cooling During the Eocene-Oligocene Climate Transition, *Science*, **323**, 5918, 1187-1190.

Liu, X-L.; Leider, A.; Gillespie, A.; Gröger, J.; Versteegh, G. J. M.; Hinrichs, K-U.; (2010) Identification of polar lipid precursors of the ubiquitous branched GDGT orphan lipids in a peat bog in Northern Germany, *Organic Geochemistry*, **41**, 7, 653-660.

Lo, S-L.; Montague, C. E.; Chang, E. L.; (1989) Purification of Glycerol Dialkyl Nonitol Tetraether from *Sulfolobus acidocaldarius*, *Journal of Lipid Research*, **30**, 6, 944-949.

Louris, J. N.; Brodbelt-Lustig, J. S.; Cooks, R. G.; Glish, G. L.; Van Berkel, G. J.; McLuckey, S. A.; (1990) Ion Isolation and Sequential Stages of Mass Spectrometry in a Quadrupole Ion Trap Mass-Spectrometer, *International Journal of Mass Spectrometry and Ion Processes*, **96**, 2, 117-137.

Lubeck, M.; Brekenfeld, A.; Gebhardt, C.; Baessmann, C.; (2007) Novel CID fragmentation techniques in ion traps for improved peptide characterization, *Application Note*, Bruker Daltonics.

Lutnaes, B. F.; Brandal, Ø.; Sjöblom, J.; Krane, J.; (2006) Archaeal C_{80} isoprenoid tetraacids responsible for naphthenate deposition in crude oil processing, *Organic & Biomolecular Chemistry*, **4**, 4, 616-620.

Lutnaes, B. F.; Krane, J.; Smith, B. E.; Rowland, S. J.; (2007) Structure elucidation of C_{80}, C_{81} and C_{82} isoprenoid tetraacids responsible for naphthenate deposition in crude oil production, *Organic & Biomolecular Chemistry*, **5**, 12, 1873-1877.

Macalady, J. L.; Vestling, M. M.; Baumler, D.; Boekelheide, N.; Kaspar, C. W.; Banfield, J. F.; (2004) Tetraether-linked membrane monolayers in *Ferroplasma* spp: a key to survival in acid, *Extremophiles*, **8**, 5, 411-419.

Malins, D. C.; (1968) Metabolism of Glycerol Ether-Containing Lipids in Dogfish (*Squalus acanthias*), *Journal of Lipid Research*, **9**, 6, 687-692.

Mancuso, C. A.; Nichols, P. D.; White, D. C.; (1986) A Method for the Separation and Characterization of Archaebacterial Signature Ether Lipids, *Journal of Lipid Research*, **27**, 1, 49-56.

March, R. E.; (1997) An introduction to quadrupole ion trap mass spectrometry, *Journal of Mass Spectrometry*, **32**, 4, 351-369.

March, R. E.; (2000) Quadrupole ion trap mass spectrometer, in *Encyclopedia of Analytical Chemistry*, Meyers, R. A., Ed.; Wiley, Chichester; pp. 11848-11872.

Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S.; (1998) Fourier transform ion cyclotron resonance mass spectrometry: A primer, *Mass Spectrometry Reviews*, **17**, 1, 1-35.

Marshall, A. G.; Hendrickson, C. L.; (2002) Fourier transform ion cyclotron resonance detection: principles and experimental configurations, *International Journal of Mass Spectrometry*, **215**, 59-75.

Matsuno, Y.; Sugai, A.; Higashibata, H.; Fukuda, W.; Ueda, K.; Uda, I.; Sato, I.; Itoh, T.; Imanaka, T.; Fujiwara, S.; (2009) Effect of Growth Temperature and Growth Phase on the Lipid Composition of the Archaeal Membrane from *Thermococcus kodakaraensis*, *Bioscience Biotechnology and Biochemistry*, **73**, 1, 104-108.

McEwen, C. N.; Larsen, B. S.; (2009) Ionization Mechanisms Related to Negative Ion APPI, APCI, and DART, *Journal of the American Society for Mass Spectrometry*, **20**, 8, 1518-1521.

McLafferty, F. W.; Tureĉek, F.; (1993) Interpretation of mass spectra, 4th edition, University Science Books, Sausalito.

Meany, D. L.; Xie, H.; Thompson, L. V.; Arriaga, E. A.; Griffin, T. J.; (2007) Identification of carbonylated proteins from enriched rat skeletal muscle mitochondria using affinity chromatography-stable isotope labelling and tandem mass spectrometry, *Proteomics*, **7**, 1150-1163.

Ménot, G.; Bard, E.; Rostek, F.; Weijers, J. W. H.; Hopmans, E. C.; Schouten, S.; Damsté, J. S. S.; (2006) Early reactivation of European rivers during the last deglaciation, *Science*, **313**, 5793, 1623-1625.

Menzel, D.; Hopmans, E. C.; Schouten, S.; Damsté, J. S. S.; (2006) Membrane tetraether lipids of planktonic Crenarchaeota in Pliocene sapropels of the eastern Mediterranean Sea, *Palaeogeography Palaeoclimatology Palaeoecology*, **239**, 1-2, 1-15.

Meot-Ner, M.; (1983) The Ionic Hydrogen Bond. 2. Intramolecular and Partial Bonds. Protonation of Polyethers, Crown Ethers, and Diketones, *Journal of the American Chemical Society*, **105**, 15, 4906-4911.

Middleditch, B. S.; (1989) Analytical artefacts: GC, MS, HPLC, TLC, and PC (Journal of Chromatography Library, Volume 44), Elsevier Science Publishers, Amsterdam.

Moldowan, J. M.; Seifert, W. K.; (1979) Head-to-head linked isoprenoid hydrocarbons in petroleum, *Science*, **204**, 169-171.

Moraes, L. A. B.; Eberlin, M. N.; (2001) Ketalization of gaseous acylium ions, *Journal of the American Society for Mass Spectrometry*, **12**, 2, 150-162.

Morgan, H. W.; (2009) Personal communication.

Morii, H.; Koga, Y.; (1993) Tetraether type polar lipids increase after logarithmic growth phase of *Methanobacterium thermoautotrophicum* in compensation for the decrease of diether lipids, *FEMS Microbiology Letters*, **109**, 2-3, 283-287.

Morii, H.; Eguchi, T.; Nishihara, M.; Kakinuma, K.; König, H.; Koga, Y.; (1998) A novel ether core lipid with H-shaped C_{80}-isoprenoid hydrocarbon chain from the hyperthermophilic methanogen *Methanothermus fervidus*, *Biochimica Et Biophysica Acta-Lipids and Lipid Metabolism*, **1390**, 3, 339-345.

Morii, H.; Yagi, H.; Akutsu, H.; Nomura, N.; Sako, Y.; Koga, Y.; (1999) A novel phosphoglycolipid archaetidyl(glucosyl)inositol with two sesterterpanyl chains from the aerobic hyperthermophilic archaeon *Aeropyrum pernix* K1, *Biochimica Et Biophysica Acta-Molecular and Cell Biology of Lipids*, **1436**, 3, 426-436.

Morlender-Vais, N.; Mandelbaum, A.; (1997a) Proton transfer between functional groups in MH^+ ions of stereoisomeric diethers on chemical ionization and collision-induced dissociation conditions, *International Journal of Mass Spectrometry*, **167**, 13-20.

Morlender-Vais, N.; Mandelbaum, A.; (1997b) The role of hydrogen migration in the mechanism of alcohol elimination from MH^+ ions of ethers upon chemical ionization, *Journal of Mass Spectrometry*, **32**, 10, 1124-1132.

Morth, S.; Tindall, B. J.; (1985) Evidence That Changes in the Growth Conditions Affect the Relative Distribution of Diether Lipids in Haloalkaliphilic Archaebacteria, *FEMS Microbiology Letters*, **29**, 3, 285-288.

Murae, T.; Muraoka, R.; Kitajima, F.; (2001) Examination of intact ether lipid components of thermoacidophile *Sulfolobus* sp. by ESI/TOF mass spectrometry, *Journal of Mass Spectrometric Society of Japan*, **49**, 5, 195-200.

Murae, T.; Takamatsu, Y.; Muraoka, R.; Endoh, S.; Yamauchi, N.; (2002) Facile distinction of neutral and acidic tetraether lipids in archaea membrane by halogen atom adduct ions in electrospray ionization mass spectrometry, *Journal of Mass Spectrometry*, **37**, 2, 209-215.

Navale, V.; (1992) A Study of Chemical Transformation of Glycerol Ether Lipids to Hydrocarbons by Flash and Hydrous Pyrolysis, *Journal of Analytical and Applied Pyrolysis*, **23**, 2, 121-133.

Nelson, K. E.; Clayton, R. A.; Gill, S. R.; Gwinn, M. L.; Dodson, R. J.; Haft, D. H.; Hickey, E. K.; Peterson, J. D.; Nelson, W. C.; Ketchum, K. A.; McDonald, L.; Utterback, T. R.; Malek, J. A.; Linher, K. D.; Garrett, M. M.; Stewart, A. M.; Cotton, M. D.; Pratt, M. S.; Phillips, C. A.; Richardson, D.; Heidelberg, J.; Sutton, G. G.; Fleischmann, R. D.; Eisen, J. A.; White, O.; Salzberg, S. L.; Smith, H. O.; Venter, J. C.; Fraser, C. M.; (1999) Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*, *Nature*, **399**, 6734, 323-329.

Nemoto, N.; Shida, Y.; Shimada, H.; Oshima, T.; Yamagishi, A.; (2003) Characterization of the precursor of tetraether lipid biosynthesis in the thermoacidophilic archaeon *Thermoplasma acidophilum*, *Extremophiles*, **7**, 3, 235-243.

Nes, W. D.; (2000) Sterol methyl transferase: enzymology and inhibition, *Biochimica Et Biophysica Acta-Molecular and Cell Biology of Lipids*, **1529**, 1-3, 63-88.

Nes, W. D.; Song, Z.; Dennis, A. L.; Zhou, W.; Nam, J.; Miller, M. B.; (2003) Biosynthesis of phytosterols - Kinetic mechanism for the enzymatic C-methylation of sterols, *Journal of Biological Chemistry*, **278**, 36, 34505-34516.

Nichols, P. D.; Franzmann, P. D.; (1992) Unsaturated Diether Phospholipids in the Antarctic Methanogen *Methanococcoides burtonii*, *FEMS Microbiology Letters*, **98**, 1-3, 205-208.

Nichols, P. D.; Shaw, P. M.; Mancuso, C. A.; Franzmann, P. D.; (1993) Analysis of Archaeal Phospholipid-Derived Diether and Tetraether Lipids by High Temperature Capillary Gas Chromatography, *Journal of Microbiological Methods*, **18**, 1, 1-9.

Nichols, D. S.; Miller, M. R.; Davies, N. W.; Goodchild, A.; Raftery, M.; Cavicchioli, R.; (2004) Cold adaptation in the antarctic archaeon *Methanococcoides burtonii* involves membrane lipid unsaturation, *Journal of Bacteriology*, **186**, 24, 8508-8515.

Nicolaus, B.; Trincone, A.; Esposito, E.; Vaccaro, M. R.; Gambacorta, A.; De Rosa, M.; (1990) Calditol Tetraether Lipids of the Archaebacterium *Sulfolobus solfataricus* – Biosynthetic Studies, *Biochemical Journal*, **266**, 3, 785-791.

Niederberger, T. D.; Götz, D. K.; McDonald, I. R.; Ronimus, R. S.; Morgan, H. W.; (2006) *Ignisphaera aggregans* gen. nov., sp nov., a novel hyperthermorphilic crenarchaeote isolated from hot springs in Rotorua and Tokaanu, New Zealand, *International Journal of Systematic and Evolutionary Microbiology*, **56**, 965-971.

Niederberger, T. D.; Ronimus, R. S.; Morgan, H. W.; (2008) The microbial ecology of a high-temperature near-neutral spring situated in Rotorua, New Zealand, *Microbiological Research*, **163**, 5, 594-603.

Niemann, H.; Elvert, M.; (2008) Diagnostic lipid biomarker and stable carbon isotope signatures of microbial communities mediating the anaerobic oxidation of methane with sulphate, *Organic Geochemistry*, **39**, 12, 1668-1677.

Nikitas, P.; Pappa-Louisi, A.; Agrafiotou, P.; (2004) New insights on the retention mechanism of non-polar solutes in reversed-phase liquid chromatographic columns, *Journal of Chromatography A*, **1034**, 1-2, 41-54.

Nishihara, M.; Koga, Y.; (1987a) Extraction and Composition of Polar Lipids from the Archaebacterium, *Methanobacterium thermoautotrophicum*: Effective Extraction of Tetraether Lipids by an Acidified Solvent, *Journal of Biochemistry*, **101**, 4, 997-1005.

Nishihara, M.; Morii, H.; Koga, Y.; (1987b) Structure Determination of a Quartet of Novel Tetraether Lipids from *Methanobacterium thermoautotrophicum*, *Journal of Biochemistry*, **101**, 4, 1007-1015.

Nishihara, M.; Morii, H.; Koga, Y.; (1989) Heptads of Polar Ether Lipids of an Archaebacterium, *Methanobacterium thermoautotrophicum*: Structure and Biosynthetic Relationship, *Biochemistry*, **28**, 1, 95-102.

Nishihara, M.; Koga, Y.; (1990) Natural occurrence of archaetidic acid and caldarchaetidic acid (di- and tetraether analogues of phosphatidic acid) in the archaebacterium *Methanobacterium thennautotrophicum*, *Biochemistry and Cell Biology*, **68**, 91-95.

Nishihara, M.; Koga, Y.; (1991) Hydroxyarchaetidylserine and hydroxyarchaetidyl-myo-inositol in *Methanosarcina barkeri*: polar lipids with a new ether core portion, *Biochimica et Biophysica Acta*, **1082**, 2, 211-217.

Nishihara, M.; Nagahama, S.; Ohga, M.; Koga, Y.; (2000) Straight-chain fatty alcohols in the hyperthermophilic archaeon *Pyrococcus furiosus*, *Extremophiles*, **4**, 5, 275-277.

Nishihara, M.; Morii, H.; Matsuno, K.; Ohga, M.; Stetter, K. O.; Koga, Y.; (2002) Structural analysis by reductive cleavage with $LiAlH_4$ of an allyl ether choline-phospholipid, archaetidylcholine, from the hyperthermophilic methanoarchaeon *Methanopyrus kandleri*, *Archaea*, **1**, 2, 123-131.

Nissenbaum A.; Kaplan I. R.; (1966) Origin of the Be'eri (Israel) sulfur deposit. *Chemical Geology*, **1**, 295–316.

Nissenbaum, A.; Baedecker, M . J.; Kaplan, I. R.; (1972) Organic geochemistry of Dead Sea sediments, *Geochimica et Cosmochimica Acta*, **36**, 709-727.

Oba, M.; Sakata, S.; Tsunogai, U.; (2006) Polar and neutral isopranyl glycerol ether lipids as biomarkers of archaea in near-surface sediments from the Nankai Trough, *Organic Geochemistry*, **37**, 12, 1643-1654.

Ogawa, H.; Amano, M.; Chihara, T.; (1998) Facile and highly selective monoacylation of symmetric diols adsorbed on silica gel with acetyl chloride, *Chemical Communications*, 4, 495-496.

Oppermann, B. I.; Michaelis, W.; Blumenberg, M.; Frerichs, J.; Schulz, H. M.; Schippers, A.; Beaubien, S. E.; Krüger, M.; (2010) Soil microbial community changes as a result of long-term exposure to a natural CO_2 vent, *Geochimica et Cosmochimica Acta*, **74**, 9, 2697-2716.

Orphan, V. J.; Hinrichs, K-U.; Ussler, W.; Paull, C. K.; Taylor, L. T.; Sylva, S. P.; Hayes, J. M.; Delong, E. F.; (2001) Comparative analysis of methane-oxidizing archaea and sulfate-reducing bacteria in anoxic marine sediments, *Applied and Environmental Microbiology*, **67**, 4, 1922-1934.

Orphan, V. J.; House, C. H.; Hinrichs, K-U.; McKeegan, K. D.; DeLong, E. F.; (2002) Multiple archaeal groups mediate methane oxidation in anoxic cold seep sediments, *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 11, 7663-7668.

Pancost, R. D.; Damsté, J. S. S.; de Lint, S.; van der Maarel, M. J. E. C.; Gottschal, J. C.; (2000) Biomarker evidence for widespread anaerobic methane oxidation in Mediterranean sediments by a consortium of methanogenic archaea and bacteria, *Applied and Environmental Microbiology*, **66**, 3, 1126-1132.

Pancost, R. D.; Bouloubassi, I.; Aloisi, G.; Damsté, J. S. S.; (2001a) Three series of non-isoprenoidal dialkyl glycerol diethers in cold-seep carbonate crusts, *Organic Geochemistry*, **32**, 5, 695-707.

Pancost, R. D.; Hopmans, E. C.; Damsté, J. S. S.; The MEDINAUT Shipboard Scientific Party; (2001b) Archaeal lipids in Mediterranean cold seeps: Molecular

proxies for anaerobic methane oxidation, *Geochimica Et Cosmochimica Acta*, **65**, 10, 1611-1627.

Pancost, R. D.; Pressley, S.; Coleman, J. M.; Benning, L. G.; Mountain, B. W.; (2005) Lipid biomolecules in silica sinters: indicators of microbial biodiversity, *Environmental Microbiology*, **7**, 1, 66-77.

Pancost, R. D.; Pressley, S.; Coleman, J. M.; Talbot, H. M.; Kelly, S. P.; Farrimond, P.; Schouten, S.; Benning, L.; Mountain, B. W.; (2006) Composition and implications of diverse lipids in New Zealand Geothermal sinters, *Geobiology*, **4**, 2, 71-92.

Pancost, R. D.; Coleman, J. M.; Love, G. D.; Chatzi, A.; Bouloubassi, I.; Snape, C. E.; (2008) Kerogen-bound glycerol dialkyl tetraether lipids released by hydropyrolysis of marine sediments: A bias against incorporation of sedimentary organisms?, *Organic Geochemistry*, **39**, 9, 1359-1371.

Panganamala, R. V.; Sievert, C. F.; Cornwell, D. G.; (1971) Quantitative Estimation and Identification of O-Alkyl Glycerols and Alkyl Iodides and Their Hydrocarbon Derivatives, *Chemistry and Physics of Lipids*, **7**, 4, 336-&.

Pape, T.; Blumenberg, M.; Seifert, R.; Egorov, V. N.; Gulin, S. B.; Michaelis, W.; (2005) Lipid geochemistry of methane-seep-related Black Sea carbonates, *Palaeogeography Palaeoclimatology Palaeoecology*, **227**, 1-3, 31-47.

Paściak, M.; Holst, O.; Lindner, B.; Mordarska, H.; Gamian, A.; (2003) Novel bacterial polar lipids containing ether-linked alkyl chains, the structures and biological properties of the four major glycolipids from *Propionibacterium propionicum* PCM 2431 (ATCC 14157^{T}), *Journal of Biological Chemistry*, **278**, 6, 3948-3956.

Patel, B. K. C.; Jasperse-Herst, P. M.; Morgan, H. W.; Daniel, R. M.; (1986) Isolation of Anaerobic, Extremely Thermophilic, Sulphur Metabolising Archaebacteria from New Zealand Hot Springs, *New Zealand Journal of Marine and Freshwater Research*, **20**, 3, 439-445.

Patwardhan, A. P.; Thompson, D. H.; (1999) Efficient synthesis of 40- and 48-membered tetraether macrocyclic bisphosphocholines, *Org. Lett*. **1**, 2, 241-243.

Paul, W.; Steinwedel, H. S.; (1960) US Patent, no. 2,939,952.

Pearson, A.; McNichol, A. P.; Benitez-Nelson, B. C.; Hayes, J. M.; Eglinton, T. I.; (2001) Origins of lipid biomarkers in Santa Monica Basin surface sediment: A case study using compound-specific $\Delta^{14}C$ analysis, *Geochimica Et Cosmochimica Acta*, **65**, 18, 3123-3137.

Pearson, A.; Huang, Z.; Ingalls, A. E.; Romanek, C. S.; Wiegel, J.; Freeman, K. H.; Smittenberg, R. H.; Zhang, C. L.; (2004) Nonmarine crenarchaeol in Nevada hot springs, *Applied and Environmental Microbiology*, **70**, 9, 5229-5237.

Pearson, A.; Pi, Y.; Zhao, W.; Li, WJ.; Li, Y.; Inskeep, W.; Perevalova, A.; Romanek, C.; Li, S.; Zhang, C. L.; (2008) Factors controlling the distribution of archaeal tetraethers in terrestrial hot springs, *Applied and Environmental Microbiology*, **74**, 11, 3523-3532.

Pease, T. K.; Van Vleet, E. S.; Barre, J. S.; Dickins, H. D.; (1998) Simulated degradation of glyceryl ethers by hydrous and flash pyrolysis, *Organic Geochemistry*, **29**, 4, 979-988.

Pernthaler, A.; Pernthaler, J.; Amann, R.; (2002) Fluorescence in situ hybridization and catalyzed reporter deposition for the identification of marine bacteria, *Applied and Environmental Microbiology*, **68**, 6, 3094-3101.

Peterse, F.; Kim, J-H.; Schouten, S.; Kristensen, D. K.; Koç, N.; Damsté, J. S. S.; (2009a) Constraints on the application of the MBT/CBT palaeothermometer at high latitude environments (Svalbard, Norway), *Organic Geochemistry*, **40**, 6, 692-699.

Peterse, F.; Schouten, S.; van der Meer, J.; van der Meer, M. T. J.; Damsté, J. S. S.; (2009b) Distribution of branched tetraether lipids in geothermally heated soils: Implications for the MBT/CBT temperature proxy, *Organic Geochemistry*, **40**, 2, 201-205.

Peterse, F.; van der Meer, M. T. J.; Schouten, S.; Jia, G.; Ossebaar, J.; Blokker, J.; Damsté, J. S. S.; (2009c) Assessment of soil *n*-alkane δD and branched tetraether membrane lipid distributions as tools for paleoelevation reconstruction, *Biogeosciences*, **6**, 12, 2799-2807.

Peterse, F.; Nicol, G. W.; Schouten, S.; Damsté, J. S. S.; (2010) Influence of soil pH on the abundance and distribution of core and intact polar lipid-derived branched GDGTs in soil, *Organic Geochemistry*, **41**, 10, 1171-1175.

Pitcher, A.; Hopmans, E. C.; Schouten, S.; Damsté, J. S. S.; (2009a) Separation of core and intact polar archaeal tetraether lipids using silica columns: Insights into living and fossil biomass contributions, *Organic Geochemistry*, **40**, 1, 12-19.

Pitcher, A.; Schouten, S.; Damsté, J. S. S.; (2009b) In Situ Production of Crenarchaeol in Two California Hot Springs, *Applied and Environmental Microbiology*, **75**, 13, 4443-4451.

Pitcher, A.; Rychlik, N.; Hopmans, E. C.; Spieck, E.; Rijpstra, W. I. C.; Ossebaar, J.; Schouten, S.; Wagner, M.; Sinninghe Damsté, J. S.; (2010) Crenarchaeol dominates the membrane lipids of Candidatus *Nitrososphaera gargensis*, a thermophilic Group I.1b Archaeon, *ISME J*, **4**, 4, 542-552.

Powers, L. A.; Werne, J. P.; Johnson, T. C.; Hopmans, E. C.; Damsté, J. S. S.; Schouten, S.; (2004) Crenarchaeotal membrane lipids in lake sediments: A new paleotemperature proxy for continental paleoclimate reconstruction?, *Geology*, **32**, 7, 613-616.

Powers, L.; Werne, J. P.; Vanderwoude, A. J.; Damsté, J. S. S.; Hopmans, E. C.; Schouten, S.; (2010) Applicability and calibration of the TEX_{86} paleothermometer in lakes, *Organic Geochemistry*, **41**, 4, 404-413.

Prahl, F. G.; Muehlhausen, L. A.; Zahnle, D. L.; (1988) Further evaluation of long-chain alkenones as indicators of paleoceanographic conditions, *Geochimica et Cosmochimica Acta*, **52**, 9, 2303-2310.

Qiu, D-F.; Games, M. P. L.; Xiao, X-Y.; Games, D. E.; Walton, T. J.; (1998) Application of high-performance liquid chromatography/electrospray mass spectrometry for the characterization of membrane lipids in the haloalkaliphilic archaebacterium *Natronobacterium magadii*, *Rapid Communications in Mass Spectrometry*, **12**, 14, 939-946.

Qiu, D-F.; Xiao, X-Y.; Walton, T. J.; Games, M. P. L.; Games, D. E.; (1999) High-performance liquid chromatography/atmospheric pressure chemical ionization mass spectrometry of phospholipids in *Natronobacterium magadii*, *European Mass Spectrometry*, **5**, 3, 151-156.

Qiu, D-F.; Games, M. P. L.; Xiao, X-Y.; Games, D. E.; Walton, T. J.; (2000) Characterisation of membrane phospholipids and glycolipids from a halophilic archaebacterium by high-performance liquid chromatography/electrospray mass spectrometry, *Rapid Communications in Mass Spectrometry*, **14**, 17, 1586-1591.

Rafferty, J. L.; Zhang, L.; Siepmann, J. I.; Schure, M. R.; (2007) Retention mechanism in reversed-phase liquid chromatography: A molecular perspective, *Analytical Chemistry*, **79**, 17, 6551-6558.

Rattray, J. E.; van de Vossenberg, J.; Hopmans, E. C.; Kartal, B.; van Niftrik, L.; Rijpstra, W. I. C.; Strous, M.; Jetten, M. S. M.; Schouten, S.; Damsté, J. S. S.; (2008) Ladderane lipid distribution in four genera of anammox bacteria, *Archives of Microbiology*, **190**, 1, 51-66.

Reeburgh, W. S.; (2007) Oceanic methane biogeochemistry, *Chemical Reviews*, **107**, 2, 486-513.

Reigstad, L. J.; Richter, A.; Daims, H.; Urich, T.; Schwark, L.; Schleper, C.; (2008) Nitrification in terrestrial hot springs of Iceland and Kamchatka, *FEMS Microbiology Ecology*, **64**, 2, 167-174.

Reilly, M. S.; Grogan, D. W.; (2001) Characterization of intragenic recombination in a hyperthermophilic archaeon via conjugational DNA exchange, *Journal of Bacteriology*, **183**, 9, 2943-2946.

Reysenbach, A-L.; Liu, Y.; Banta, A. B.; Beveridge, T. J.; Kirshtein, J. D.; Schouten, S.; Tivey, M. K.; Von Damm, K. L.; Voytek, M. A.; (2006) A ubiquitous thermoacidophilic archaeon from deep-sea hydrothermal vents, *Nature*, **442**, 7101, 444-447.

Ring, M. W.; Schwär, G.; Thiel, V.; Dickschat, J. S.; Kroppenstedt, R. M.; Schulz, S.; Bode, H. B.; (2006) Novel iso-branched ether lipids as specific markers of developmental sporulation in the myxobacterium *Myxococcus xanthus*, *Journal of Biological Chemistry*, **281**, 48, 36691-36700.

Romano, I.; Poli, A.; Finore, I.; Huertas, F. J.; Gambacorta, A.; Pelliccione, S.; Nicolaus, G.; Lama, L.; Nicolaus, B.; (2007) *Haloterrigena hispanica* sp. nov., an extremely halophilic archaeon from Fuente de Piedra, southern Spain, *International Journal of Systematic and Evolutionary Microbiology*, **57**, 1499-1503.

Rossel, P. E.; Lipp, J. S.; Fredricks, H. F.; Arnds, J.; Boetius, A.; Elvert, M.; Hinrichs, K-U.; (2008) Intact polar lipids of anaerobic methanotrophic archaea and associated bacteria, *Organic Geochemistry*, **39**, 8, 992-999.

Rowland, S. J.; (1990) Production of acyclic isoprenoid hydrocarbons by laboratory maturation of methanogenic bacteria, *Organic Geochemistry*, **15**, 9-16.

Rueda, G.; Rosell-Melé, A.; Escala, M.; Gyllencreutz, R.; Backman, J.; (2009) Comparison of instrumental and GDGT-based estimates of sea surface and air temperatures from the Skagerrak, *Organic Geochemistry*, **40**, 2, 287-291.

Ruepp, A.; Graml, W.; Santos-Martinez, M-L.; Koretke, K. K.; Volker, C.; Mewes, H. W.; Frishman, D.; Stocker, S.; Lupas, A. N.; Baumeister, W.; (2000) The genome sequence of the thermoacidophilic scavenger *Thermoplasma acidophilum*, **407**, 508-513.

Rütters, H.; Sass, H.; Cypionka, H.; Rullkötter, J.; (2001) Monoalkylether phospholipids in the sulfate-reducing bacteria *Desulfosarcina variabilis* and *Desulforhabdus amnigenus*, *Archives of Microbiology*, **176**, 6, 435-442.

Saiki, R. K.; Gelfand, D. H.; Stoffel, S.; Scharf, S. J.; Higuchi, R.; Horn, G. T.; Mullis, K. B.; Erlich, H. A.; (1988) Primer-Directed Enzymatic Amplification of DNA with a Thermostable DNA Polymerase, *Science*, **239**, 4839, 487-491.

Saito, H.; Suzuki, N.; (2010) Distribution of acyclic and cyclic biphytanediols in recent marine sediments from IODP Site C0001, Nankai Trough, *Organic Geochemistry*, **41**, 9, 1001-1004.

Sako, Y.; Nomura, N.; Uchida, A.; Ishida, Y.; Morii, H.; Koga, Y.; Hoaki, T.; Maruyama, T.; (1996) *Aeropyrum pernix* gen. nov., sp. nov., a novel aerobic hyperthermophilic archaeon growing at temperatures up to 100^{o}C, *International Journal of Systematic Bacteriology*, **46**, 4, 1070-1077.

Schleper, C.; Jurgens, G.; Jonuscheit, M.; (2005) Genomic studies of uncultivated archaea, *Nature Reviews Microbiology*, **3**, 6, 479-488.

Schouten, S.; Hoefs, M. J. L.; Koopmans, M. P.; Bosch, H-J.; Damsté, J. S. S.; (1998) Structural characterization, occurrence and fate of archaeal ether-bound acyclic and cyclic biphytanes and corresponding diols in sediments, *Organic Geochemistry*, **29**, 5-7, 1305-1319.

Schouten, S.; Hopmans, E. C.; Pancost, R. D.; Damsté, J. S. S.; (2000) Widespread occurrence of structurally diverse tetraether membrane lipids: Evidence for the ubiquitous presence of low-temperature relatives of hyperthermophiles, *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 26, 14421-14426.

Schouten, S.; Hopmans, E. C.; Schefuβ, E.; Damsté, J. S. S.; (2002) Distributional variations in marine crenarchaeotal membrane lipids: a new tool for reconstructing ancient sea water temperatures?, *Earth and Planetary Science Letters*, **204**, 1-2, 265-274.

Schouten, S.; Wakeham, S. G.; Hopmans, E. C.; Damsté, J. S. S.; (2003a) Biogeochemical evidence that thermophilic archaea mediate the anaerobic oxidation of methane, *Applied and Environmental Microbiology*, **69**, 3, 1680-1686.

Schouten, S.; Hopmans, E. C.; Forster, A.; van Breugel, Y.; Kuypers, M. M. M.; Damsté, J. S. S.; (2003b) Extremely high sea-surface temperatures at low latitudes during the middle Cretaceous as revealed by archaeal membrane lipids, *Geology*, **31**, 12, 1069-1072.

Schouten, S.; Hopmans, E. C.; Damsté, J. S. S.; (2004) The effect of maturity and depositional redox conditions on archaeal tetraether lipid palaeothermometry, *Organic Geochemistry*, **35**, 5, 567-571.

Schouten, S.; Forster, A.; Panoto, F. E.; Damsté, J. S. S.; (2007a) Towards calibration of the TEX_{86} palaeothermometer for tropical sea surface temperatures in ancient greenhouse worlds, *Organic Geochemistry*, **38**, 9, 1537-1546.

Schouten, S.; Huguet, C.; Hopmans, E. C.; Kienhuis, M. V. M.; Damsté, J. S. S.; (2007b) Analytical methodology for TEX_{86} paleothermometry by high-performance liquid chromatography/atmospheric pressure chemical ionization-mass spectrometry, *Analytical Chemistry*, **79**, 7, 2940-2944.

Schouten, S.; van der Meer, M. T. J.; Hopmans, E. C.; Rijpstra, W. I. C.; Reysenbach, A-L.; Ward, D. M.; Damsté, J. S. S.; (2007c) Archaeal and bacterial glycerol dialkyl glycerol tetraether lipids in hot springs of Yellowstone National Park, *Applied and Environmental Microbiology*, **73**, 19, 6181-6191.

Schouten, S.; Baas, M.; Hopmans, E. C.; Damsté, J. S. S.; (2008a) An unusual isoprenoid tetraether lipid in marine and lacustrine sediments, *Organic Geochemistry*, **39**, 8, 1033-1038.

Schouten, S.; Baas, M.; Hopmans, E. C.; Reysenbach, A-L.; Damsté, J. S. S.; (2008b) Tetraether membrane lipids of *Candidatus* "*Aciduliprofundum boonei*", a cultivated obligate thermoacidophilic euryarchaeote from deep-sea hydrothermal vents, *Extremophiles*, **12**, 1, 119-124.

Schouten, S.; Hopmans, E. C.; Baas, M.; Boumann, H.; Standfest, S.; Konneke, M.; Stahl, D. A.; Damsté, J. S. S.; (2008c) Intact membrane lipids of "*Candidatus* Nitrosopumilus maritimus," a cultivated representative of the cosmopolitan

mesophilic group I crenarchaeota, *Applied and Environmental Microbiology*, **74**, 8, 2433-2440.

Schouten, S.; Eldrett, J.; Greenwood, D. R.; Harding, I.; Baas, M.; Damsté, J. S. S.; (2008d) Onset of long-term cooling of Greenland near the Eocene-Oligocene boundary as revealed by branched tetraether lipids, *Geology*, **36**, 2, 147-150.

Schouten, S.; Hopmans, E. C.; van der Meer, J.; Mets, A.; Bard, E.; Bianchi, T. S.; Diefendorf, A.; Escala, M.; Freeman, K. H.; Furukawa, Y.; Huguet, C.; Ingalls, A.; Ménot-Combes, G.; Nederbragt, A. J.; Oba, M.; Pearson, A.; Pearson, E. J.; Rosell-Melé, A.; Schaeffer, P.; Shah, S. R.; Shanahan, T. M.; Smith, R. W.; Smittenberg, R.; Talbot, H. M.; Uchida, M.; Van Mooy, B. A. S.; Yamamoto, M.; Zhang, Z.; Damsté, J. S. S.; (2009) An interlaboratory study of TEX_{86} and BIT analysis using high-performance liquid chromatography-mass spectrometry, *Geochemistry Geophysics Geosystems*, **10**, Q03012.

Schouten, S.; Middelburg, J. J.; Hopmans, E. C.; Damsté, J. S. S.; (2010) Fossilization and degradation of intact polar lipids in deep subsurface sediments: A theoretical approach, *Geochimica Et Cosmochimica Acta*, **74**, 13, 3806-3814.

Schubert, C. J.; Coolen, M. J. L.; Neretin, L. N.; Schippers, A.; Abbas, B.; Durisch-Kaiser, E.; Wehrli, B.; Hopmans, E. C.; Damsté, J. S. S.; Wakeham, S.; Kuypers, M. M. M.; (2006) Aerobic and anaerobic methanotrophs in the Black Sea water column, *Environmental Microbiology*, **8**, 10, 1844-1856.

Sehgal, S. N.; Kates, M.; Gibbons, N. E.; (1962) Lipids of *Halobacterium cutirubrum*, *Canadian Jouranl of Physiology and Pharmacology*, **40**, 1, 69-81.

Shah, S. R.; Mollenhauer, G.; Ohkouchi, N.; Eglinton, T. I.; Pearson, A.; (2008) Origins of archaeal tetraether lipids in sediments: Insights from radiocarbon analysis, *Geochimica Et Cosmochimica Acta*, **72**, 18, 4577-4594.

Sharma, R. B.; Blades, A. T.; Kebarle, P.; (1984) Protonation of Polyethers, Glymes and Crown Ethers, in the Gas Phase, *Journal of the American Chemical Society*, **106**, 3, 510-516.

She, Q.; Singh, R. K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M. J.; Chan-Weiher, C. C-Y.; Clausen, I. G.; Curtis, B. A.; De Moors, A.; Erauso, G.; Fletcher, C.; Gordon, P. M. K.; Heikamp-de Jong, I.; Jeffries, A. C.; Kozera, C. J.; Medina, N.; Peng, X.; Thi-Ngoc, H. P.; Redder, P.; Schenk, M. E.; Theriault, C.; Tolstrup, N.; Charlebois, R. L.; Doolittle, W. F.; Duguet, M.; Gaasterland, T.; Garrett, R. A.; Ragan, M. A.; Sensen, C. W.; Van der Oost, J.; (2001) The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2, *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 14, 7835-7840.

Shimada, H.; Nemoto, N.; Shida, Y.; Oshima, T.; Yamagishi, A.; (2002) Complete polar lipid composition of *Thermoplasma acidophilum* HO-62 determined by high-performance liquid chromatography with evaporative light-scattering detection, *Journal of Bacteriology*, **184**, 2, 556-563.

Shimada, H.; Nemoto, N.; Shida, Y.; Oshima, T.; Yamagishi, A.; (2008) Effects of pH and temperature on the composition of polar lipids in *Thermoplasma acidophilum* HO-62, *Journal of Bacteriology*, **190**, 15, 5404-5411.

Shvily, R.; Müller, T.; Apeloig, Y.; Mandelbaum, A.; (1997) Direct evidence for anchimeric assistance in alcohol elimination from gas-phase MH^+ ions of 1,4-dialkoxycyclohexanes under chemical ionisation. Experiment and theory, *Journal of the Chemical Society-Perkin Transactions 2*, 6, 1221-1234.

Sigsby, M. L.; Day, R. J.; Cooks, R. G.; (1979a) Fragmentation of Even Electron Ions: Protonated Amines and Esters, *Organic Mass Spectrometry*, **14**, 10, 556-561.

Sigsby, M. L.; Day, R. J.; Cooks, R. G.; (1979b) Fragmentation of Even Electron Ions. Protonated Ketones and Ethers, *Organic Mass Spectrometry*, **14**, 5, 273-280.

Skoog, D. A.; Holler, F. J.; Neiman, T. A.; (1998) Principles of Instrumental Analysis, 5th edition. Saunders College Publishing, London.

Sluijs, A.; Schouten, S.; Pagani, M.; Woltering, M.; Brinkhuis, H.; Damsté, J. S. S.; Dickens, G. R.; Huber, M.; Reichart, G-J.; Stein, R.; Matthiessen, J.; Lourens, L. J.; Pedentchouk, N.; Backman, J.; Moran, K.; Expedition 302 Scientists; (2006) Subtropical arctic ocean temperatures during the Palaeocene/Eocene thermal maximum, *Nature*, **441**, 7093, 610-613.

Smith, D. R.; Doucette-Stamm, L. A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Bashirzadeh, R.; Blakely, D.; Cook, R.; Gilbert, K.; Harrison, D.; Hoang, L.; Keagle, P.; Lumm, W.; Pothier, B.; Qiu, D..; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; Caruso, A.; Bush, D.; Safer, H.; Patwell, D.; Prabhakar, S.; McDougall, S.; Shimer, G.; Goyal, A.; Pietrokovski, S.; Church, G. M.; Daniels, C. J.; Mao, J-I.; Rice, P.; Nölling, J.; Reeve, J. N.; (1997) Complete genome sequence of *Methanobacterium thermoautotrophicum* ΔH: Functional analysis and comparative genomics, *Journal of Bacteriology*, **179**, 22, 7135-7155.

Smith, B. E.; Sutton, P. A.; Lewis, C. A.; Dunsmore, B.; Fowler, G.; Krane, J.; Lutnaes, B. F.; Brandal, Ø.; Sjöblom, J.; Rowland, S. J.; (2007) Analysis of 'ARN' naphthenic acids by high temperature gas chromatoagraphy and high performance liquid chromatography, *Journal of Separation Science*, **30**, 3, 375-380.

Smith, R. W.; Bianchi, T. S.; Savage, C.; (2010) Comparison of lignin phenols and branched/isoprenoid tetraethers (BIT index) as indices of terrestrial organic matter in Doubtful Sound, Fiordland, New Zealand, *Organic Geochemistry*, **41**, 3, 281-290.

Smittenberg, R. H.; Sachs, J. P.; (2007) Purification of dinosterol for hydrogen isotopic analysis using high-performance liquid chromatography-mass spectrometry, *Journal of Chromatography A*, **1169**, 1-2, 70-76.

Snyder, L. R.; Poppe, H.; (1980) Mechanism of Solute Retention in Liquid-Solid Chromatography and the Role of the Mobile Phase in Affecting Separation - Competition Versus "Sorption", *Journal of Chromatography*, **184**, 4, 363-413.

Snyder, L. R.; Schunk, T. C.; (1982) Retention Mechanism and the Role of the Mobile Phase in Normal-Phase Separation on Amino-Bonded-Phase Columns, *Analytical Chemistry*, **54**, 11, 1764-1772.

Spang, A.; Hatzenpichler, R.; Brochier-Armanet, C.; Rattei, T.; Tischler, P.; Spieck, E.; Streit, W.; Stahl, D. A.; Wagner, M.; Schleper, C.; (2010) Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum Thaumarchaeota, *Trends in Microbiology*, **18**, 8, 331-340.

Sprott, G. D.; Ekiel, I.; Dicaire, C.; (1990) Novel, Acid-Labile, Hydroxydiether Lipid Cores in Methanogenic Bacteria, *Journal of Biological Chemistry*, **265**, 23, 13735-13740.

Sprott, G. D.; Meloche, M.; Richards, J. C.; (1991) Proportions of Diether, Macrocyclic Diether, and Tetraether Lipids in *Methanococcus jannaschii* Grown at Different Temperatures, *Journal of Bacteriology*, **173**, 12, 3907-3910.

Sprott, G. D.; Dicaire, C. J.; Choquet, C. G.; Patel, G. B.; Ekiel, I.; (1993) Hydroxydiether Lipid Structures in *Methanosarcina* Spp. and *Methanococcus voltae*, *Applied and Environmental Microbiology*, **59**, 3, 912-914.

Sprott, G. D.; Agnew, B. J.; Patel, G. B.; (1997) Structural features of ether lipids in the archaeobacterial thermophiles *Pyrococcus furiosus*, *Methanopyrus kandleri*, *Methanothermus fervidus*, and *Sulfolobus acidocaldarius*, *Canadian Journal of Microbiology*, **43**, 5, 467-476.

Sprott, G. D.; Brisson, J-R.; Dicaire, C. J.; Pelletier, A. K.; Deschatelets, L. A.; Krishnan, L.; Patel, G. B.; (1999) A structural comparison of the total polar lipids from the human archaea *Methanobrevibacter smithii* and *Methanosphaera stadtmanae* and its relevance to the adjuvant activities of their liposomes, *Biochimica Et Biophysica Acta-Molecular and Cell Biology of Lipids*, **1440**, 2-3, 275-288.

Stadnitskaia, A.; Baas, M.; Ivanov, M. K.; Van Weering, T. C. E.; Damsté, J. S. S.; (2003) Novel archaeal macrocyclic diether core membrane lipids in a methane-derived carbonate crust from a mud volcano in the Sorokin Trough, NE Black Sea, *Archaea*, **1**, 3, 165-173.

Stadnitskaia, A.; Muyzer, G.; Abbas, B.; Coolen, M. J. L.; Hopmans, E. C.; Baas, M.; van Weering, T. C. E.; Ivanov, M. K.; Poludetkina, E.; Damsté, J. S. S.; (2005) Biomarker and 16S rDNA evidence for anaerobic oxidation of methane and related carbonate precipitation in deep-sea mud volcanoes of the Sorokin Trough, Black Sea, *Marine Geology*, **217**, 1-2, 67-96.

Stadnitskaia, A.; Bouloubassi, I.; Elvert, M.; Hinrichs, K-U.; Damsté, J. S. S.; (2008) Extended hydroxyarchaeol, a novel lipid biomarker for anaerobic methanotrophy in cold seepage habitats, *Organic Geochemistry*, **39**, 8, 1007-1014.

Stafford, G. C.; Kelley, P. E.; Syka, J. E. P.; Reynolds, W. E.; Todd, J. F. J.; (1984) Recent Improvements in and Analytical Applications of Advanced Ion Trap

Technology, *International Journal of Mass Spectrometry and Ion Processes*, **60**, SEP, 85-98.

Stetter, K. O.; Thömm, M.; Winter, J.; Wildgruber, G.; Huber, H.; Zillig, W.; Janecovic, D.; König, H.; Palm, P.; Wunderl, S.; (1981) *Methanothermus fervidus*, sp. nov., a Novel Extremely Thermophilic Methanogen Isolated from an Icelandic Hot-Spring, *Zentralblatt Fur Bakteriologie Mikrobiologie Und Hygiene I Abteilung Originale C-Allgemeine Angewandte Und Okologische Mikrobiologie*, **2**, 2, 166-178.

Stevens, J. F.; Maier, C. S.; (2008) Acrolein: Sources, metabolism, and biomolecular interactions relevant to human health and disease, *Molecular Nutrition & Food Research*, **52**, 1, 7-25.

Sturt, H. F.; Summons, R. E.; Smith, K.; Elvert, M.; Hinrichs, K-U.; (2004) Intact polar membrane lipids in prokaryotes and sediments deciphered by high-performance liquid chromatography/electrospray ionization multistage mass spectrometry - new biomarkers for biogeochemistry and microbial ecology, *Rapid Communications in Mass Spectrometry*, **18**, 6, 617-628.

Sugai, A.; Uda, I.; Itoh, Y. H.; Itoh, T.; (2004) The Core Lipid Composition of the 17 Strains of Hyperthermophilic Archaea, *Thermococcales*, *Journal of Oleo Science*, **53**, 1, 41-44.

Swain, M.; Brisson, J-R.; Sprott, G. D.; Cooper, F. P.; Patel, G. B.; (1997) Identification of β-L-gulose as the sugar moiety of the main polar lipid *Thermoplasma acidophilum*, *Biochimica Et Biophysica Acta-Lipids and Lipid Metabolism*, **1345**, 1, 56-64.

Tachibana, A.; Yano, Y.; Otani, S.; Nomura, N.; Sako, Y.; Taniguchi, M.; (2000) Novel prenyltransferase gene encoding farnesylgeranyl diphosphate synthase from a hyperthermophilic archaeon, *Aeropyrum pernix* - Molecular evolution with alteration in product specificity, *European Journal of Biochemistry*, **267**, 2, 321-328.

Teira, E.; Reinthaler, T.; Pernthaler, A.; Pernthaler, J.; Herndl, G. J.; (2004) Combining catalyzed reporter deposition-fluorescence in situ hybridization and microautoradiography to detect substrate utilization by bacteria and archaea in the deep ocean, *Applied and Environmental Microbiology*, **70**, 7, 4411-4414.

Teske, A.; Hinrichs, K-U.; Edgcomb, V.; de Vera Gomez, A.; Kysela, D.; Sylva, S. P.; Sogin, M. L.; Jannasch, H. W.; (2002) Microbial diversity of hydrothermal sediments in the Guaymas Basin: Evidence for anaerobic methanotrophic communities, *Applied and Environmental Microbiology*, **68**, 4, 1994-2007.

Teske, A.; Sørensen, K. B.; (2008) Uncultured archaea in deep marine subsurface sediments: have we caught them all? *ISME J.*, **2**, 3-18.

Thiel, V.; Peckmann, J.; Seifert, R.; Wehrung, P.; Reitner, J.; Michaelis, W.; (1999) Highly isotopically depleted isoprenoids: Molecular markers for ancient methane venting, *Geochimica Et Cosmochimica Acta*, **63**, 23-24, 3959-3966.

Thiel, V.; Heim, C.; Arp, G.; Hahmann, U.; Sjövall, P.; Lausmaa, J.; (2007a) Biomarkers at the microscopic range: ToF-SIMS molecular imaging of Archaea-derived lipids in a microbial mat, *Geobiology*, **5**, 4, 413-421.

Thiel, V.; Toporski, J.; Schumann, G.; Sjövall, P.; Lausmaa, J.; (2007b) Analysis of archaeal core ether lipids using Time of Flight-Secondary Ion Mass Spectrometry (ToF-SIMS): Exploring a new prospect for the study of biomarkers in geobiology, *Geobiology*, **5**, 1, 75-83.

Thurl, S.; Schäfer, W.; (1988) Lipids from the Sulphur-Dependent Archaebacterium *Thermoproteus tenax*, *Biochimica Et Biophysica Acta*, **961**, 2, 253-261.

Tierney, J. E.; Russell, J. M.; (2009) Distributions of branched GDGTs in a tropical lake system: Implications for lacustrine application of the MBT/CBT paleoproxy, *Organic Geochemistry*, **40**, 9, 1032-1036.

Tierney, J. E.; Russell, J. M.; Eggermont, H.; Hopmans, E. C.; Verschuren, D.; Damsté, J. S. S.; (2010) Environmental controls on branched tetraether lipid distributions in tropical East African lake sediments, *Geochimica Et Cosmochimica Acta*, **74**, 17, 4902-4918.

Tietz, A.; Lindberg, M.; Kennedy, E. P.; (1964) A New Pteridine-Requiring Enzyme System for Oxidation of Glyceryl Ethers, *Journal of Biological Chemistry*, **239**, 12, 4081-4090.

Tornabene, T. G.; Wolfe, R. S.; Balch, W. E.; Holzer, G.; Fox, G. E.; Oro, J.; (1978) Phytanyl-Glycerol Ethers and Squalenes in Archaebacterium *Methanobacterium thermoautotrophicum*, *Journal of Molecular Evolution*, **11**, 3, 259-266.

Trincone, A.; De Rosa, M.; Gambacorta, A.; Lanzotti, V.; Nicolaus, B.; Harris, J. E.; Grant, W. D.; (1988) A Simple Chromatographic Procedure for the Detection of Cyclized Archaebacterial Glycerol-Bisdiphytanyl-Glycerol Tetraether Core Lipids, *Journal of General Microbiology*, **134**, 3159-3163.

Trincone, A.; Lanzotti, V.; Nicolaus, B.; Zillig, W.; De Rosa, M.; Gambacorta, A.; (1989) Comparative Lipid Composition of Aerobically and Anaerobically Grown *Desulfurolobus ambivalens*, an Autotrophic Thermophilic Archaeobacterium, *Journal of General Microbiology*, **135**, 2751-2757.

Trincone, A.; Nicolaus, B.; Palmieri, G.; De Rosa, M.; Huber, R.; Huber, G.; Stetter, K. O.; Gambacorta, A.; (1992) Distribution of Complex and Core Lipids within New Hyperthermophilic Members of the *Archaea* Domain, *Systematic and Applied Microbiology*, **15**, 1, 11-17.

Trommer, G.; Siccha, M.; van der Meer, M. T. J.; Schouten, S.; Damsté, J. S. S.; Schulz, H.; Hemleben, C.; Kucera, M.; (2009) Distribution of Crenarchaeota tetraether membrane lipids in surface sediments from the Red Sea, *Organic Geochemistry*, **40**, 6, 724-731.

Turich, C.; Freeman, K. H.; Bruns, M. A.; Conte, M.; Jones, A. D.; Wakeham, S. G.; (2007) Lipids of marine Archaea: Patterns and provenance in the water-column and sediments, *Geochimica Et Cosmochimica Acta*, **71**, 13, 3272-3291.

Uda, I.; Sugai, A.; Itoh, Y. H.; Itoh, T.; (2000) Characterization of caldarchaetidylglycerol analogs, dialkyl-type and trialkyl-type, from *Thermoplasma acidophilum*, *Lipids*, **35**, 10, 1155-1157.

Uda, I.; Sugai, A.; Itoh, Y. H.; Itoh, T.; (2001) Variation in molecular species of polar lipids from *Thermoplasma acidophilum* depends on growth temperature, *Lipids*, **36**, 1, 103-105.

Uda, I.; Sugai, A.; Itoh, Y. H.; Itoh, T.; (2004) Variation in Molecular Species of Core Lipids from the Order *Thermoplasmales* Strains Depends on the Growth Temperature, *Journal of Oleo Science*, **53**, 8, 399-404.

Upasani, V. N.; Desai, S. G.; Moldoveanu, N.; Kates, M.; (1994) Lipids of Extremely Halophilic Archaeobacteria from Saline Environments in India: a Novel Glycolipid in *Natronobacterium* Strains, *Microbiology-Uk*, **140**, 1959-1966.

Urey, H. C.; (1947) The thermodynamic properties of isotopic substances, *Journal of the Chemical Society*, 562-581.

Vailaya, A.; Horváth, C.; (1998) Retention in reversed-phase chromatography: partition or adsorption?, *Journal of Chromatography A*, **829**, 1-2, 1-27.

van Dongen, B. E.; Roberts, A. P.; Schouten, S.; Jiang, W-T.; Florindo, F.; Pancost, R. D.; (2007) Formation of iron sulfide nodules during anaerobic oxidation of methane, *Geochimica Et Cosmochimica Acta*, **71**, 21, 5155-5167.

van Dongen, B. E.; Semiletov, I.; Weijers, J. W. H.; Gustafsson, Ö. R.; (2008) Contrasting lipid biomarker composition of terrestrial organic matter exported from across the Eurasian Arctic by the five great Russian Arctic rivers, *Global Biogeochemical Cycles*, **22**, 1, GB1011.

Vaver, V. A.; Prokazova, N. V.; Bergel'son, L. D.; (1964) New types of neutral lipids, *Russian Chemical Bulletin*, **13**, 2, 392.

Ventura, G. T.; Kenig, F.; Reddy, C. M.; Schieber, J.; Frysinger, G. S.; Nelson, R. K.; Dinel, E.; Gaines, R. B.; Schaeffer, P.; (2007) Molecular evidence of Late Archean archaea and the presence of a subsurface hydrothermal biosphere, *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 36, 14260-14265.

Ventura, G. T.; Kenig, F.; Reddy, C. M.; Frysinger, G. S.; Nelson, R. K.; Van Mooy, B.; Gaines, R. B.; (2008) Analysis of unresolved complex mixtures of hydrocarbons extracted from Late Archean sediments by comprehensive two-dimensional gas chromatography (GCxGC), *Organic Geochemistry*, **39**, 7, 846-867.

Vink, A.; Schouten, S.; Sephton, S.; Damsté, J. S. S.; (1998) A newly discovered norisoprenoid, 2,6,15,19-tetramethylicosane, in Cretaceous black shales, *Geochimica Et Cosmochimica Acta*, **62**, 6, 965-970.

Völkl, P.; Huber, R.; Drobner, E.; Rachel, R.; Burggraf, S.; Trincone, A.; Stetter, K. O.; (1993) *Pyrobaculum aerophilum* sp. nov., a Novel Nitrate-Reducing Hyperthermophilic Archaeum, *Applied and Environmental Microbiology*, **59**, 9, 2918-2926.

Voyksner, R. D.; Lee, H.; (1999) Improvements in LC/electrospray ion trap mass spectrometry performance using an off-axis nebulizer, *Analytical Chemistry*, **71**, 7, 1441-1447.

Wagner, F.; Rottem, S.; Held, H-D.; Uhlig, S.; Zähringer, U.; (2000) Ether lipids in the cell membrane of *Mycoplasma fermentans*, *European Journal of Biochemistry*, **267**, 20, 6276-6286.

Wakeham, S. G.; Amann, R.; Freeman, K. H.; Hopmans, E. C.; Jørgensen, B. B.; Putnam, I. F.; Schouten, S.; Damsté, J. S. S.; Talbot, H. M.; Woebken, D.; (2007) Microbial ecology of the stratified water column of the Black Sea as revealed by a comprehensive biomarker study, *Organic Geochemistry*, **38**, 12, 2070-2097.

Walsh, E. M.; Ingalls, A. E.; Keil, R. G.; (2008) Sources and transport of terrestrial organic matter in Vancouver Island fjords and the Vancouver–Washington Margin: A multiproxy approach using $\delta^{13}C_{org}$, lignin phenols, and the ether lipid BIT index, *Limnology and Oceanography*, **53**, 3, 1054-1063.

Wang, Q-F.; Li, W.; Yang, H.; Liu, Y-L.; Cao, H-H.; Dornmayr-Pfaffenhuemer, M.; Stan-Lotter, H.; Guo, G-Q.; (2007) *Halococcus qingdaonensis* sp. nov., a halophilic archaeon isolated from a crude sea-salt sample, *International Journal of Systematic and Evolutionary Microbiology*, **57**, 600-604.

Ward, D. M.; Brassell, S. C.; Eglinton, G.; (1985) Archaebacterial Lipids in Hot-Spring Microbial Mats, *Nature*, **318**, 6047, 656-659.

Wasada, H.; Tsutsui, Y.; Yamabe, S.; (1996) Ab initio study of proton affinities of three crown ethers, *Journal of Physical Chemistry*, **100**, 18, 7367-7371.

Wasserfallen, A.; Nölling, J.; Pfister, P.; Reeve, J.; de Macario, E. C.; (2000) Phylogenetic analysis of 18 thermophilic *Methanobacterium* isolates supports the proposals to create a new genus, *Methanothermobacter* gen. nov., and to reclassify several isolates in three species, *Methanothermobacter thermautotrophicus* comb. nov., *Methanothermobacter wolfeii* comb. nov., and *Methanothermobacter marburgensis* sp. nov., *International Journal of Systematic and Evolutionary Microbiology*, **50**, 43-53.

Weijers, J. W. H.; Schouten, S.; van der Linden, M.; van Geel, B.; Damsté, J. S. S.; (2004) Water table related variations in the abundance of intact archaeal membrane lipids in a Swedish peat bog, *FEMS Microbiology Letters*, **239**, 1, 51-56.

Weijers, J. W. H.; Schouten, S.; Hopmans, E. C.; Geenevasen, J. A. J.; David, O. R. P.; Coleman, J. M.; Pancost, R. D.; Damsté, J. S. S.; (2006a) Membrane lipids of mesophilic anaerobic bacteria thriving in peats have typical archaeal traits, *Environmental Microbiology*, **8**, 4, 648-57.

Weijers, J. W. H.; Schouten, S.; Spaargaren, O. C.; Damsté, J. S. S.; (2006b) Occurrence and distribution of tetraether membrane lipids in soils: Implications for the use of the TEX_{86} proxy and the BIT index, *Organic Geochemistry*, **37**, 12, 1680-1693.

Weijers, J. W. H.; Schefuβ, E.; Schouten, S.; Damsté, J. S. S.; (2007a) Coupled thermal and hydrological evolution of tropical Africa over the last deglaciation, *Science*, **315**, 5819, 1701-1704.

Weijers, J. W. H.; Schouten, S.; Sluijs, A.; Brinkhuis, H.; Damsté, J. S. S.; (2007b) Warm arctic continents during the Palaeocene-Eocene thermal maximum, *Earth and Planetary Science Letters*, **261**, 1-2, 230-238.

Weijers, J. W. H.; Schouten, S.; van den Donker, J. C.; Hopmans, E. C.; Damsté, J. S. S.; (2007c) Environmental controls on bacterial tetraether membrane lipid distribution in soils, *Geochimica Et Cosmochimica Acta*, **71**, 3, 703-713.

Weijers, J. W. H.; Panoto, E.; van Bleijswijk, J.; Schouten, S.; Rijpstra, W. I. C.; Balk, M.; Stams, A. J. M.; Damsté, J. S. S.; (2009) Constraints on the Biological Source(s) of the Orphan Branched Tetraether Membrane Lipids, *Geomicrobiology Journal*, **26**, 6, 402-414.

Weijers, J. W. H.; Wiesenberg, G. L. B.; Bol, R.; Hopmans, E. C.; Pancost, R. D.; (2010) Carbon isotopic composition of branched tetraether membrane lipids in soils suggest a rapid turnover and a heterotrophic life style of their source organism(s), *Biogeosciences Discussions*, **7**, 3691-3734.

Welsch, T.; Frank, H.; Vigh, G.; (1990) Silanol Effects in Reversed-Phase Liquid-Chromatography, *Journal of Chromatography*, **506**, 97-108.

White, G. F.; Russell, N. J.; Tidswell, E. C.; (1996) Bacterial scission of ether bonds, *Microbiological Reviews*, **60**, 1, 216-232.

Wilson, M. A.; (2004) PhD Thesis, University of York, UK.

Woese, C. R.; Fox, G. E.; (1977) Phylogenetic Structure of the Prokaryotic Domain: the Primary Kingdoms, *Proceedings of the National Academy of Sciences of the United States of America*, **74**, 11, 5088-5090.

Woese, C. R.; (1987) Bacterial evolution, *Microbiological Reviews*, **51**, 2, 221-271.

Woese, C. R.; Kandler, O.; Wheelis, M. L.; (1990) Towards a Natural System of Organisms: Proposal for the Domains Archaea, Bacteria, and Eucarya, *Proceedings of the National Academy of Sciences of the United States of America*, **87**, 12, 4576-4579.

Wuchter, C.; Schouten, S.; Boschker, H. T. S.; Damsté, J. S. S.; (2003) Bicarbonate uptake by marine Crenarchaeota, *FEMS Microbiology Letters*, **219**, 2, 203-207.

Wuchter, C.; Schouten, S.; Coolen, M. J. L.; Damsté, J. S. S.; (2004) Temperature-dependent variation in the distribution of tetraether membrane lipids of marine *Crenarchaeota*: Implications for TEX_{86} paleothermometry, *Paleoceanography*, **19**, 4, PA4028.

Wuchter, C.; Schouten, S.; Wakeham, S. G.; Damsté, J. S. S.; (2005) Temporal and spatial variation in tetraether membrane lipids of marine *Crenarchaeota* in particulate organic matter: Implications for TEX_{86} paleothermometry, *Paleoceanography*, **20**, 3, PA3013.

Wuchter, C.; Schouten, S.; Wakeham, S. G.; Damsté, J. S. S.; (2006) Archaeal tetraether membrane lipid fluxes in the northeastern Pacific and the Arabian Sea: Implications for TEX_{86} paleothermometry, *Paleoceanography*, **21**, 4, PA4208.

Yamamoto, M.; Okino, T.; Sugisaki, S.; Sakamoto, T.; (2008) Late Pleistocene changes in terrestrial biomarkers in sediments from the central Arctic Ocean, *Organic Geochemistry*, **39**, 6, 754-763.

Yang, Y-H.; Lee, K.; Jang, K-S.; Kim, Y-G.; Park, S-H.; Lee, C-S.; Kim, B-G.; (2009) Low mass cutoff evasion with q_z value optimization in ion trap, *Analytical Biochemistry*, **387**, 1, 133-135.

Zeikus, J. G.; Wolfe, R. S.; (1972) *Methanobacterium thermoautotrophicus* sp. n., an Anaerobic, Autotrophic, Extreme Thermophile, *Journal of Bacteriology*, **109**, 2, 707-713.

Zeikus, J. G.; Wolfe, R. S.; (1973) Fine Structure of *Methanobacterium thermoautotrophicum*: Effect of Growth Temperature on Morphology and Ultrastructure, *Journal of Bacteriology*, **113**, 1, 461-467.

Zeikus, J. G.; (1977) Biology of Methanogenic Bacteria, *Bacteriological Reviews*, **41**, 2, 514-541.

Zenobi, R.; Knochenmuss, R.; (1998) Ion formation in MALDI mass spectrometry, *Mass Spectrometry Reviews*, **17**, 5, 337-366.

Zhang, C. L.; Pearson, A.; Li, Y-L.; Mills, G.; Wiegel, J.; (2006) Thermophilic temperature optimum for crenarchaeol synthesis and its implication for archaeal evolution, *Applied and Environmental Microbiology*, **72**, 6, 4419-4422.

Zillig, W.; Stetter, K. O.; Wunderl, S.; Schulz, W.; Priess, H.; Scholz, I.; (1980) The *Sulfolobus caldariella* Group: Taxonomy on the Basis of the Structure of DNA-Dependent RNA Polymerases, *Archives of Microbiology*, **125**, 3, 259-269.

Zillig, W.; (1993) Confusion in the Assignments of *Sulfolobus* Sequences to *Sulfolobus* Species, *Nucleic Acids Research*, **21**, 22, 5273-5273.

Zink, K-G.; Vandergoes, M. J.; Mangelsdorf, K.; Dieffenbacher-Krall, A. C.; Schwark, L.; (2010) Application of bacterial glycerol dialkyl glycerol tetraethers (GDGTs) to develop modern and past temperature estimates from New Zealand lakes, *Organic Geochemistry*, **41**, 9, 1060-1066.