Development of a high-throughput technique for screening Archaeal tetraether lipid cores and other alcohols in sediments and microbial cultures

Cezary Poplawski

PhD

University of York

Chemistry

June 2017

Abstract

A mild and effective protocol for the derivatisation of Archaeal tetraether lipid cores and other alcohols from cultured and sedimentary material has been developed using *N*-protected amino acids. The derivatives were prepared in excellent to near quantitative yields in approx. 2 h. The *N*-Boc and *N*-Fmoc derivatives exhibit very good chromatographic properties and considerable signal intensity improvement, in MS, of up to two orders of magnitude, relative to the native species. The fluorescence properties of the derivatives containing a strong chromophore (Fmoc), showed excellent detector response and allow very favourable detection limits to be achieved, comparable to those of MS detection.

Derivatisation of a total lipid extract from a soil using Fmoc-lysine(Boc) amino acid permitted the derivatives of sterols and alkanols to be chromatographed and detected by reversed phase LC-MS in under 20 min, which is significantly faster than the standard gas chromatography method. Due to the ability to selectively screen the mass spectral data for characteristic losses associated only with the derivatives, this approach enabled the identification a number of lipid components that were omitted during GC-MS analysis.

The novel APCI-LC-MS method was shown to be suited to application in rapid screening of GDGT tetraether lipids from Archaeal cultures and sediment extracts, with drastically reduced analytical run times and markedly improved separation of the cyclopentane ring-containing GDGT lipids. Although, the lack of resolution of crenarchaeol and its regioisomer precluded use of the TEX_{86} index to reconstruct the geological temperature, the calculation of TEX_{86}^{L} afforded an estimate of SST temperature that is very close to the value obtained from the native GDGTs using the normal phase based method with TEX_{86}.

List of Contents

List of Tables

List of figures

Acknowledgements

First and above all, I praise GOD ALMIGHTY, OUR FATHER, without whom nothing is possible and to whom I owe my very existence, for the strength, courage, good health and many blessings without which I could not have finished this research.

I would like to thank my wife, Justyna, for believing in me and for all your love, encouragement and continuing support you have given me ever since we met. I promise to reward you for all the sacrifices you have made and the lonely hours you have spent holding our family and house intact. I am very grateful for my children, Julia and Sebastian, for being my bright lights of happiness in the moments of despair and resignation during my research.

I would like to express my deepest gratitude to my research supervisor, Prof. Brendan Keely, for giving me the opportunity to do research, invaluable advice, critical comments, and particularly for your patience and guidance during the writing process and correction of this thesis.

I would like to thank my thesis advisor, Prof. Jane Thomas-Oates, for her patience, understanding, excellent comments and suggestions regarding my research, especially in its early stage, and for preparing me for my viva.

Much appreciation and thanks to Alice Bacci who helped me immensely during the development and optimization work on the derivatisation method and help with sample analyses.

Special thanks are in order to the staff of the Chemistry Graduate Office, Sharon Stewart, Helen Coombs, Rachel Crooks, Alice Duckett, for your invaluable support, care, empathy, and making sure, I submitted my thesis on time.

Heartfelt thank you to Phil Helliwell for letting me borrow laboratory equipment without which I would not have been able to continue my research.

Many to thanks to the members of the BJK group, Matt, Kim, Chris, Adam, Scott, Marina, Alex, Martina, for your friendship, company during many long hours in the lab and interesting and enlightening (at times) discussions.

Finally, I would like to thank mom for her love, caring, prayers and sacrifices for educating and preparing me for my future life. This thesis would not be possible without your help.

I would like to express my gratitude to The Department of Chemistry of the University of York for partial funding of my PhD studentship.

Author’s Declaration

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

Cezary Poplawski

Chapter 1 Introduction

1.1 Archaea

Phylogenetic analyses of 16S rRNA gene sequences of representative organisms including eukaryotes and prokaryotes led to discovery of a new taxonomic structure, resulting in the revision of the recognised categories by which life on Earth is classified. The studies recognised three separate domains: Archaea, Bacteria, and Eukarya (Woese and Fox, 1977). Notably, the new classification better reflects evolutionary relationships among organisms, thus leading to its widespread acceptance and adoption. The new domain, Archaea, is thought to have diverged from the ancestral line of eubacteria in the early life of the planet (Woese, Kandler and Wheelis, 1990). Some of the extant archaeal lineages are believed to closely resemble the ancient species due to the slow rate of evolutionary development (Woese, 1987). Within the domain, five phylogenetically dissimilar kingdoms have been recognised: *Euryarchaeota*, *Crenarchaeota*, *Thaumarchaeota*, *Korarchaeota* and *Nanoarchaeota* (**Figure 1-1**). While the *Crenarchaeota* include predominantly hyperthermophilic organism, the *Euryarchaeota* kingdom comprises physiologically diverse organisms represented by halophiles, thermophiles and methanogens. The two kingdoms represent the most abundant phyla occurring in the environment (Woese, Kandler and Wheelis, 1990; Spang *et al.*, 2010). Although they have been assigned on the basis of the specific archaeal 16S RNA sequences observed in the environment, their phylogenetic assignment has been debated (Brochier *et al.*, 2005). During the late 1990s, genomic studies and membrane lipid analysis of material from non-extreme marine and lacustrine environments resulted in identification of ubiquitous non-thermophilic groups of archaea, later assigned the separate phylum: Thaumarchaeota (DeLong *et al.*, 1998; Karner, DeLong and Karl, 2001; DeLong, 2006; Brochier-Armanet *et al.*, 2008).

The ubiquitous pelagic archaea are estimated to comprise approx. 20 % of the ocean picoplankton (Karner, DeLong and Karl, 2001). Further, several distinct clades of archaea have been recognized to play important roles in major cycles of elements. The ammonia – oxidising archaea (AOA) abundant in the marine (Francis *et al.*, 2005; Wuchter *et al.*, 2006; Walker *et al.*, 2010; Stahl and de la Torre, 2012), lacustrine (Caffrey *et al.*, 2007; Santoro *et al.*, 2008) and soil environments (Chen *et al.*, 2008; Tourna *et al.*, 2008; Zhang *et al.*, 2010) were found to be central to oxidation of ammonia (NH_3) to nitrite (NO_2^-) and are also involved in the fixation of nitrogen gas (N_2) (Cabello, Roldan and Moreno-Vivian, 2004; Dekas, Poretsky and Orphan, 2009) and denitrification processes (Cabello, Roldan and

Moreno-Vivian, 2004). Methanogenic archaea also dominate biogenic production of methane, a major greenhouse gas, the gas being the major product of their energy-conserving metabolism (Teske, Dhillon and Sogin, 2003; Feist *et al.*, 2006; Thauer *et al.*, 2008; Berg *et al.*, 2010; Conrad *et al.*, 2010). Methanotrophic archaea are capable of oxidation of methane as a source of energy, a process which often occurs in syntropy with sulfate reducing bacteria (Wakeham *et al.*, 2003; Dekas, Poretsky and Orphan, 2009; Mueller *et al.*, 2014). Archaea also play an important role in the cycling of sulfur: through a variety of processes involving both the production and oxidation of sulfidic compounds (Kletzin *et al.*, 2004; Kletzin, 2007; Liu, Beer and Whitman, 2012; Offre, Spang and Schleper, 2013; Fike, Bradley and Rose, 2015). Thus, archaea play key roles in the Earth's global biogeochemical cycles and greenhouse gas fixation.

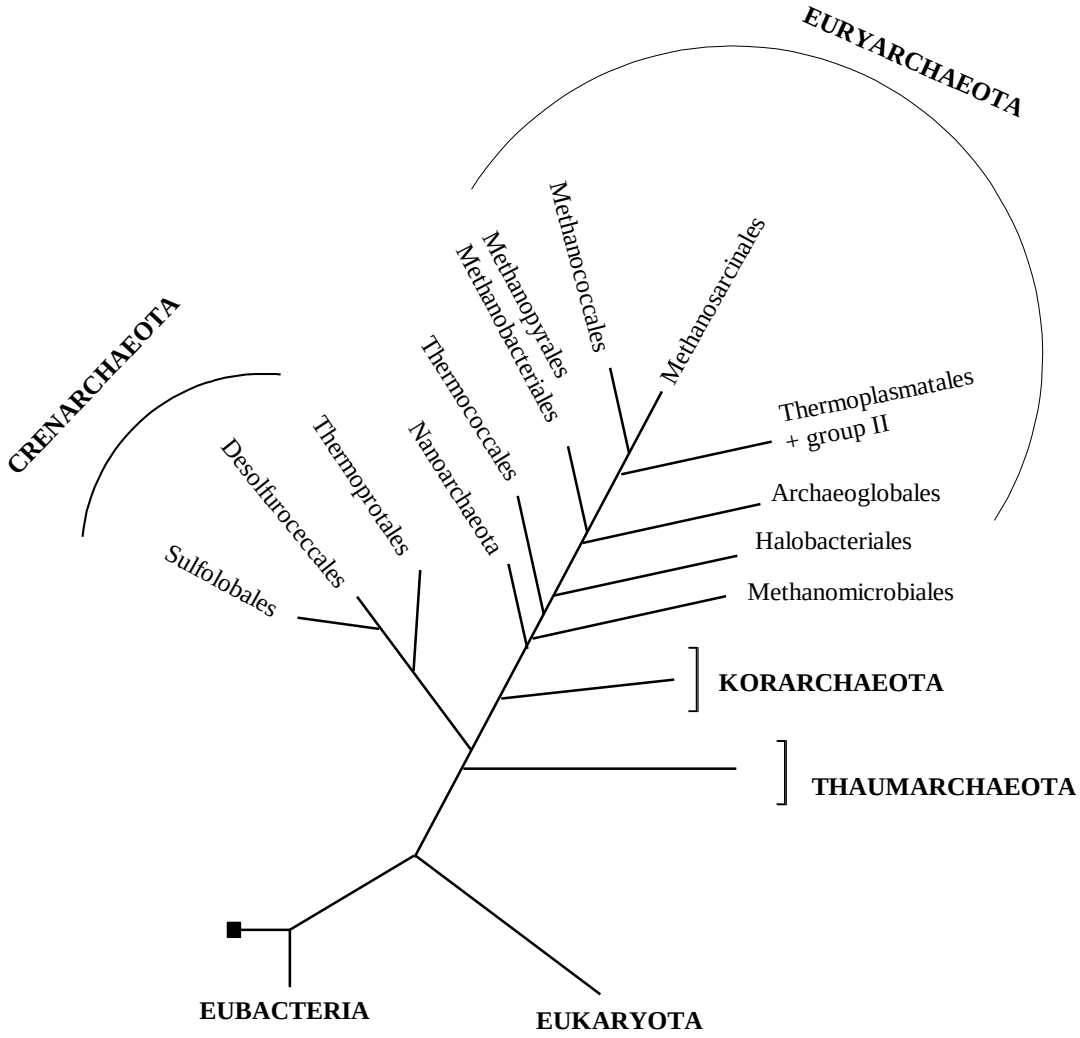


Figure 1-1 Schematic tree of life representing the three main domains: Eubacteria, Eukaryota and Archaea. The tree is further expanded to show the main phyla comprising the archaeal domain (reconstructed from Raymann *et al.*, 2015).

The ubiquity and diversity of archaea is not limited to marine environments, they have been found in other diverse settings such as peat bogs (Hopmans *et al.*, 2000; Kotsyurbenko *et al.*, 2004; Metje and Frenzel, 2005, 2007; Høj, Olsen and Torsvik, 2008; Schmidt, Hinrichs and Elvert, 2010; Puglisi *et al.*, 2014), soil (Leininger *et al.*, 2006; Adair and Schwartz, 2008; Chen *et al.*, 2008; Pearson *et al.*, 2008; Angel *et al.*, 2010; Pratscher, Dumont and Conrad, 2011; Bartossek *et al.*, 2012; Lim *et al.*, 2012; Daebeler, Gansen and Frenzel, 2013) and lacustrine environments (Karr *et al.*, 2006; Lehours *et al.*, 2007; Bomberg *et al.*, 2008; Schouten *et al.*, 2008; Callieri *et al.*, 2009; Jiang *et al.*, 2009; Lliros, 2010; Kan *et al.*, 2011, 2016; Hugoni *et al.*, 2013).

1.2 Archaeal lipids

The lipids that the members of the Archaea incorporate into their cell membranes are unique and intricate ether lipids. Typically, the lipid comprises an isoprenoid glycerol ether core (lipid core, LC) that is covalently capped with a polar head group (phosphate and or glycosyl (**Figure 1-2**). As such, the ether LC together with the polar head group comprise the membrane lipid: termed intact polar lipid (IPL). Notably the hydrophobic backbone of the LC is exclusively bound to the *sn*-2 and *sn*-3 position of the glycerol unit (Koga *et al.*, 1993). Distinctively, the lipids produced by Eukarya and Eubacteria are glycerol dialkyl esters where saturated or unsaturated fatty acids, predominantly straight chain components, are ester-bound to glycerol at positions *sn*-1 and *sn*-2. The exception to this phenomenon are some thermophilic bacteria capable to synthesize glycerol diethers and glycerol with ether and ester bound hydrocarbon chains within the same structure (Sinninghe Damsté *et al.*, 2014). Koga *et al.* (1963) were the first to isolate and identify the simplest ether lipid core, archaeol, a glycerol dialkyl diether (GDD-0) lipid core comprising of two phytanyl (3,7,11,15-tetramethylhexadecanyl, C_{20}) chains ether bound to a glycerol moiety and exhibiting the unique *sn*-2,3 stereochemistry. The most simple glycerol dialkyl glycerol tetraether lipid, caldarchaeol (GDGT-0) has two biphytanyl (C_{40}) chains with their termini ether bound to two glycerol units with the same unique *sn*-2,3 stereochemistry (**Figure 1-2**).

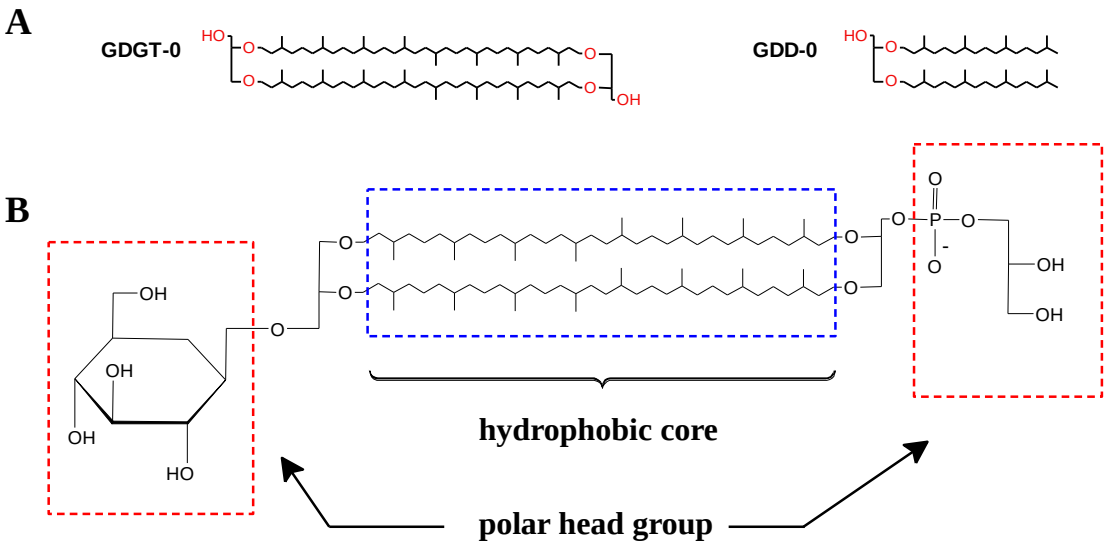


Figure 1-2 (A) Structures of common glycerol ethers: caldarchaeol (GDGT-0) and archaeol (GDD-0) and (B) an example of hypothetical monolayer-forming intact polar glycophospholipid with GDGT-0 as a hydrophobic lipid core. By analogy to the tetraether intact polar lipids, glycerol diether can also possess a polar head group.

The archaeal ether lipids are biosynthesized through a series of reaction steps, during which dimethylallyl diphosphate (DMAPP) is condensed with several isopentenyl diphosphate (IPP) units by IPP synthases during the propagation of the isoprenoidal chains (both C_{20} and C_{40}), which are subsequently bonded to the glycerol units (Villanueva, Sinninghe Damsté and Schouten, 2014). Finally, following saturation of the isoprenoid chains the attachment of the polar head groups takes place. Although, the formation of the tetraether lipid structures was speculated to occur through condensation of two diether glycerol units, the new biosynthetic route proposed by Villanueva seems more plausible as it is consistent with the analysed sequences of the enzymatic pathways and experimental evidence of the biosynthesis of the glycerol tetraether lipids.

The members of the Archaea inhabit a wide range of environments including opposite extremities of the environmental spectrum, whether the limiting factor is temperature, salinity, pH or other external environmental factors (Ciaramella, Pisani and Rossi, 2002; Cavicchioli, 2006; Chaban, Ng and Jarrell, 2006; Leigh *et al.*, 2011). Thus, in order to thrive in their surroundings, maintaining the membrane fluidity for optimal cell functioning is paramount. Hence, the membrane needs to be tight enough to permit proper functioning of the proton and ion pumps, yet it is also required to retain a necessary degree of fluidity to permit, for example, cell division (Elferink *et al.*, 1994; Van de Vossenberg, Driessen and

Konings, 1998). To counter potential environmental stress factors or to adapt to changes in their growth environment the archaea regulate the membrane parameters (e.g. fluidity and permeability) by modulating core lipid and polar head group compositions (Gliozzi, Relini and Chong, 2002; Chong *et al.*, 2012; Oger and Cario, 2013). Thus, it was observed that some thermophiles, methanogens and halophiles decrease the proportion of tetraether to macrocyclic diether and diether lipids in response to the increase in growth temperature (Sprott *et al.*, 1993; Lai, Springstead and Monbouquette, 2008). Moreover, some halophiles express exclusively diether lipids as their membrane constituents (Tornabene and Langworthy, 1979), while other methanogenic archaea incorporate unsaturation into their isoprenoidal chains as an adaptation to cold arctic conditions (Nichols *et al.*, 2004). Notably, another common adaptation mechanism among thermophilic archaea is the ability to modulate the degree of internal cyclization of the isoprenoidal chains in response to increase in the growth temperature. It has been demonstrated that with increasing growth temperature the number of cyclopentane rings incorporated into each of the biphytanyl chains increases (De Rosa, Esposito, *et al.*, 1980; De Rosa *et al.*, 1983; Gliozzi *et al.*, 1983; Uda *et al.*, 2001; Schouten *et al.*, 2003). Gabriel *et al.* (2000) proposed that the increased cyclization might help in maintaining the stability of the membrane during thermal expansion caused by the higher temperature, while Lai *et al.* (2008) suggested that thermally induced increase in cyclization might be linked to temperature-dependent genomic or enzymatic activities. However, the ability to regulate the temperature-induced degree of cyclization does not appear to be uniform among the archaea and other factors, such as pH, has considerable effect on the number of cyclopentane rings incorporated into GDGT structures. Indeed, as shown by Shimada *et al.* (2008) on cultivated thermo- and acidophilic *Thermoplasma acidophilum* the internal cyclization increased with increasing temperature as well as with decreasing pH. Notably, GDGT lipids with up to eight cyclopentane rings have been detected (De Rosa and Gambacorta, 1988). Further, a unique type of glycerol tetraether lipid containing four cyclopentane and one cyclohexane rings within the biphytane chain, crenarchaeol along with its regioisomer, has been identified exclusively in the phyla of Thaumarcheota (Sinninghe Damsté *et al.*, 2002; Schouten *et al.*, 2008; Pitcher *et al.*, 2010). The Thaumarcheota are non-thermophilic and occupy rather cold marine waters, thus the incorporation of the cyclohexane ring was suggested to provide optimal membrane lipid packing in order to sustain its life functions in lower temperatures (Sinninghe Damsté *et al.*, 2002).

Notably, the processes through which the archaeal species respond to given environmental stressors, or how they have adapted to the environment they occupy, are not addressed exhaustively in the discussion above. Broader discussion of this subject is widely available through a large body of published reports (Koga and Morii, 2007; Chong, 2010; Chong *et al.*, 2012; Schouten, Hopmans and Sinninghe Damsté, 2013). The structures of some of the archaeal glycerol ether commonly found in sedimentary, terrestrial, lacustrine, riverine and marine settings are illustrated in **Figure 1-3**.

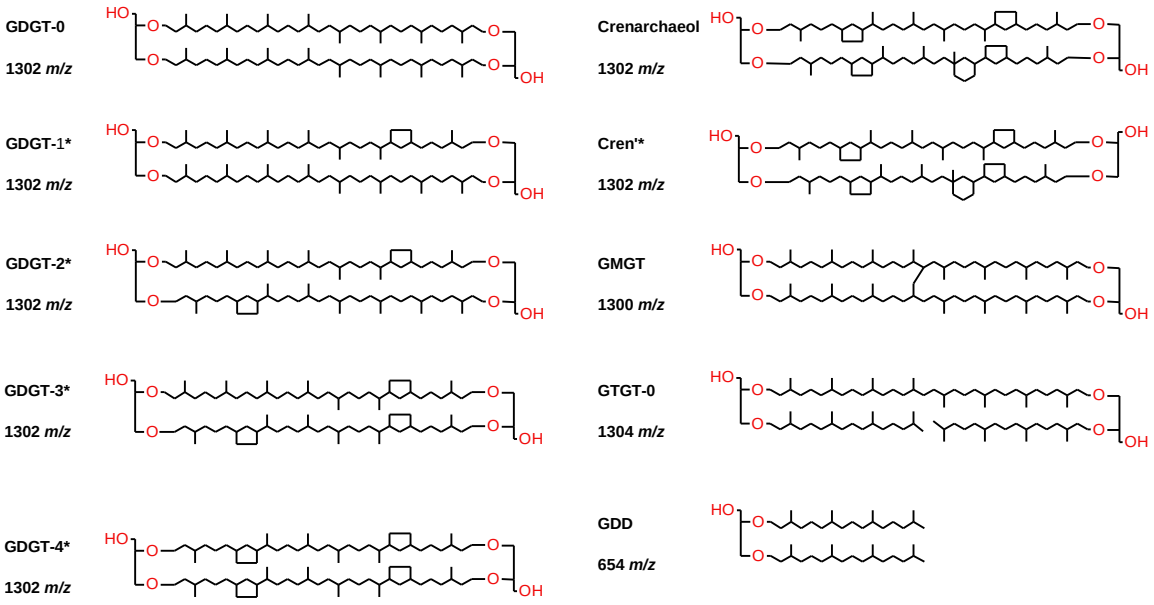


Figure 1-3 Structures of common archaeal lipid cores: diether archaeol (GDD) and tetraethers (GDGTs). The structures are annotated with their commonly accepted abbreviations along with their *m/z* values as observed during mass spectrometry detection. The arabic numerals next to the lipid name signifies the numbers of cyclopentane rings within the isoprenoidal core and an asterix indicates lipid cores used in palaeotemperature reconstructions.

1.3 The use of Archaeal lipids in palaeotemperature proxies

The use of biomarkers, remnant lipids specific to certain organisms found in sedimentary organic matter, has been developed and witnessed considerable expansion in the past three decades due to their contribution to the abilities to reconstruct palaeoenvironments over geological periods (Schouten, Hopmans and Sinninghe Damsté, 2013). This is in part due to

the different digenetic alterations these biomarkers undergo compared with the hard parts of fossils (e.g. calcareous foraminifera) and the absence of the skeletal forming organisms in certain environments (e.g. cold climates).

Archaea alter their cell membrane compositions in response to the temperature variations, producing lipids with different number of cyclopentane and cyclohexane rings which render the cell membrane more stable at higher temperatures (Gliozzi *et al.*, 1983). Schouten and co-workers (2002) found that the relative abundance of the GDGTs containing 1-3 cyclopentane rings and crenarchaeol regioisomer identified in sediments positively correlate to the increasing surface temperature of the water body at the time of formation of their deposition. This relationship was named the TEX_{86} (TetraEther indeX of tetraethers containing 86 carbons) and is expressed in Equation **1-1** (the structures of the lipid cores used in the TEX_{86} index are denoted with a star and shown in **Figure 1-3**).

$$TEX_{86} = \frac{[GDGT-2]+[GDGT-3]+[cren']}{[GDGT-1]+[GDGT-2]+[GDGT-3]+[cren']} \quad \textbf{(1-1)}$$

$$SST = (56.2 \times TEX_{86}) - 10.78 \quad (r^2 = 0.92, n = 43) \quad \textbf{(1-2)}$$

The index, originally calculated using 40 marine surface sediments, was subsequently correlated with the modern sea surface water temperature data and showed a strong linear correlation (Equation **1-2**). In a similar manner TEX_{86} values calculated using surface sediments from large lakes showed that the palaeotemperature reconstruction can be successfully applied to some continental lacustrine environments (Powers *et al.*, 2004, 2010). The isoprenoid GDGTs, including those used to calculate the TEX_{86} index, have also been found in soil and peat environments (Weijers *et al.*, 2006), thus contributions from such sources could potentially bias the TEX_{86} index, and in turn lead to erroneous determination of the palaeo surface temperatures in marine and lacustrine environments. To trace the origin of organic matter (OM) and account for the fluvial input of the terrestrial isoprenoid GDGTs into the aquatic environment Hopmans *et al.* (2004) introduced a proxy that uses the relative abundances of branched glycerol dialkyl glycerol tetraethers (brGDGTs, **Figure 1-4**), derived from anaerobic bacteria of terrestrial origin, versus the abundances of the crenarchaeol tetraether lipids that are dominant in aquatic settings.

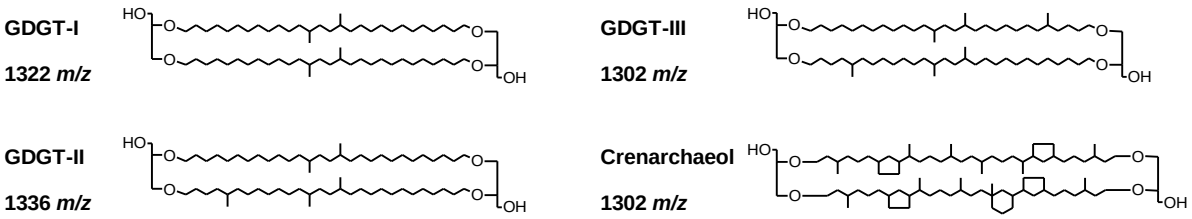


Figure 1-4 Structures of bacterial lipid cores: GDGT (I-III) and crenarchaeol used in the BIT index. The structures are annotated with their commonly accepted abbreviations along with *m/z* values of the protonated molecules as observed in mass spectrometry.

This BIT (branched versus isoprenoid tetraether) index (Equation **1-3**) is based on the observations that the brGDGTs are predominantly found in the terrestrial settings and are not likely to be transported with aeolian dust. Values of the index close to 1 indicate soil derived sources of the OM and values close to 0 suggest aquatic origins of the OM. In general, the TEX_{86} is considered to be applicable when reconstructing palaeotemperatures in settings where the BIT index is 0.1-0.3, suggesting little terrestrial OM input. Notably, it is crucial when reporting values of the BIT index to also report absolute GDGTs concentrations to avoid misinterpretation of the origin of the OM in cases where an increase in the abundance of crenarchaeol has also been noted (Castañeda *et al.*, 2010).

$$BIT = \frac{[GDGT-I]+GDGT-II]+[GDGT-III]}{\{Crenarchaeol]+[GDGT-I]+[GDGT-II]+[GDGT-III]} \qquad (\mathbf{1\text{-}3})$$

The experimental attempts to calibrate the TEX_{86} index, however, have not been successful mainly due to observations that pure cultures of Thaumarchaeota produce very little of the crenarchaeol regioisomer in mesocosm studies *versus* the ocean setting (Wuchter *et al.*, 2004) and the suggestion that the regioisomer may have a different provenance (Shah *et al.*, 2008). Thus, the index is reliant on the empirical core-top sediment calibration. In an attempt to increase the accuracy of the TEX_{86} index, it was recalibrated using an increased number of surface deposits (n = 287) from worldwide locations. This resulted in an improvement in the correlation to SST (r^2 = 0.935, n = 223), though a loss of the linear relationship was observed outside the temp range of 5 °C to 30 °C (Kim *et al.*, 2008). Subsequently, supported by the same data set, Liu *et al.* (2009) proposed that the SST may be inferred from the TEX_{86} data by applying a non-linear regression. A more extensive study, using a 426 global core-

top dataset Kim *et al.* (2010) proposed two modified indices, TEX_{86}^{L} (Equation **1-4**) and TEX_{86}^{H} (Equation **1-5**), where both indices are logarithmic formulae and are correlated to SST *via* linear regression Equations **1-6** and **1-7**).

$$TEX_{86}^{L} = \log\left(\frac{[GDGT-2]}{[GDGT-1]+[GDGT-2]+[GDGT-3]}\right) \quad (\mathbf{1\text{-}4})$$

$$TEX_{86}^{H} = \log\left(\frac{[GDGT-2]+[GDGT-3]+[cren']}{[GDGT-1]+[GDGT-2]+[GDGT-3]+[cren']}\right) \quad (\mathbf{1\text{-}5})$$

$$SST = (67.5 \times TEX_{86}^{L}) + 46.9 \quad (\mathbf{1\text{-}6})$$

$$SST = (68.4 \times TEX_{86}^{H}) + 38.6 \quad (\mathbf{1\text{-}7})$$

To date, the new indices offer very good correlations to the global SST dataset with the following qualifications: for palaeotemperature reconstructions where the SST temperatures fall below 15 °C the TEX_{86}^{L} should be used and TEX_{86}^{H} when the reconstructed SST is above 30 °C. In instances where the reconstructed temperatures may fall above and below 15 °C it is recommended that the TEX_{86}^{L} be used as it is more reliable in such situations (Kim *et al.*, 2010).

Although the TEX_{86} proxy was shown to be independent of ocean salinity and availability of nutrients, several reports have raised concerns about other factors that may bias the index values and lead to inaccuracy of the reconstructed temperatures. Some of the notable factors have been identified as: pH variations (Pearson *et al.*, 2008; Shah *et al.*, 2008), upwelling and lateral transport of lipids within the water column (Huguet *et al.*, 2007; Lee *et al.*, 2008), seasonal variations in productivity and/or transport of the lipids to ocean floor deposits (Wuchter *et al.*, 2006; Blaga *et al.*, 2009; Powers *et al.*, 2010), an unusual and localized crenarchaeotal ecology (Trommer *et al.*, 2009), persistent oxic degradation of the tetraether lipids (Huguet *et al.*, 2009), *Euryarchaeota* or benthic archaea contributions to the sedimentary lipid signatures (Sluijs *et al.*, 2006; Shah *et al.*, 2008; Blaga *et al.*, 2009; Powers *et al.*, 2010). Trommer *et al.* (2009) suggested, while interpreting considerable reconstructed

SST temperature bias between the northern and southern regions of the Red Sea, that in certain areas the use of a local calibration may be more appropriate, to better reflect contributions from the endemic archaeal populations. Lastly, factors associated with laboratory or analytical instrumentation imprecisions should be considered carefully when attempting geological temperature reconstruction using the TEX_{86} index.

1.4 Analysis of the archaeal lipids

1.4.1 Sample preparation: Lipids extraction and fractionation

In early research on archaeal lipids the extraction of the cellular material were normally performed using the Bligh and Dyer method (Bligh and Dyer, 1959), where the membrane lipids were isolated following a series of liquid-liquid extractions using varying ratios of methanol, chloroform and water. Due to the neutral conditions of the original extraction method, it was suspected that the conditions might not have been optimal for releasing all of the membrane cell lipids. Indeed, Nishihara *et al.* (1987) showed that by replacing the water with trichloroacetic acid in the extraction medium, the lipids recovery level increased six times. Thus, the modified method has found wide spread acceptance within scientific community and remains widely used at the present time. The extraction of the environmental samples such as sediments or water filtrates is usually achieved in two steps: first, the material is freeze dried and in the second step the lipids are extracted using organic solvents (methanol and dichloromethane) with the aid of Soxhlet or sonic irradiation (DeLong *et al.*, 1998). More recently, accelerated solvent extraction (ASE) with organic solvents has been developed for the isolation of organic material: the temperature and pressure of the extraction can be regulated with precision (Richter *et al.*, 1996; Schouten *et al.*, 2007). Notably, due to the high throughput and degree of automation offered by the ASE method it has become a standard method for isolation of the lipid material from sediments and soil matrixes. Moreover, in a comparison study of different extraction techniques of organic matter from sediments Schouten *et al.* (2007) demonstrated that there were no considerable differences between ASE, Soxhlet and ultrasonic methods in terms of the distribution of GDGTs.

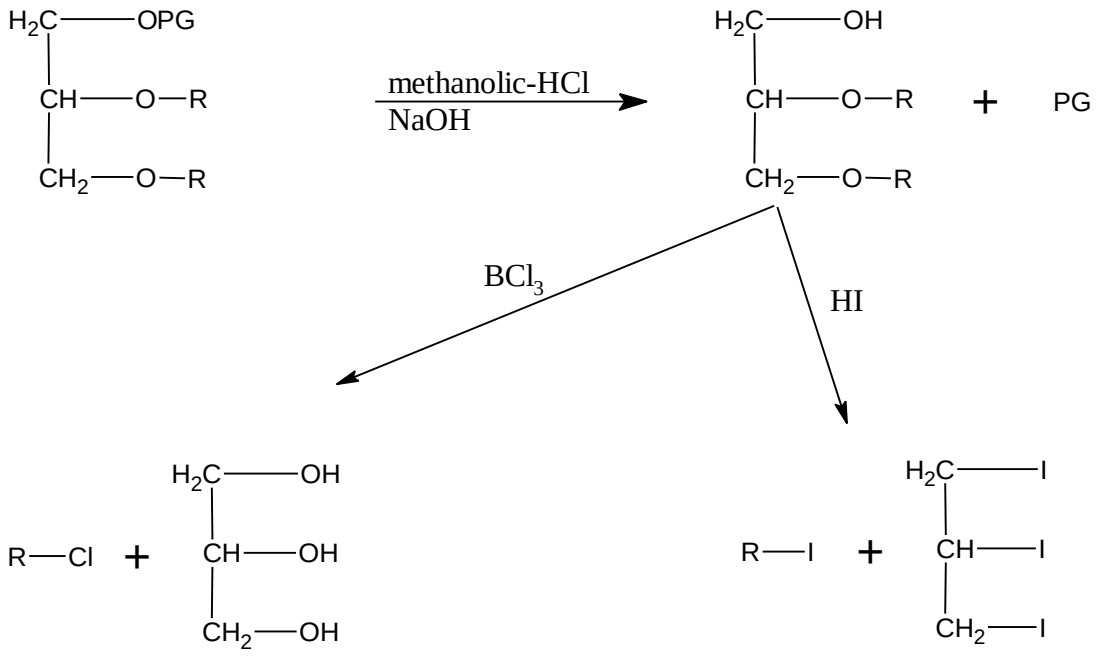
The organic material extracted from the matrix requires two main steps of preparation: purification to remove the interfering matrix, separation of the analytes by polarity to enable

enrichment of the fraction containing the target compounds. In the early days the separation of compounds by class was mainly performed by means of thin layer chromatography (TLC) (Mancuso, Nichols and White, 1986), which was later replaced by preparative column chromatography using alumina or silica gel, mainly due to the ease of the column preparation, high separation efficiency and column load and low price of the sorbent material (De Rosa *et al.*, 1983; Hopmans *et al.*, 2004; Schouten *et al.*, 2007). Alternatively, the total lipid extract, cellular material or isolated fractions can be saponified (hydrolysed) using methanolic solutions under acidic (HCl) or basic conditions (NaOH, KOH) in order to cleave the polar head groups and release of the lipid cores (GDGTs) (Kates, 1977; Trincone *et al.*, 1992). Although basic hydrolysis is often used it was argued that the acid hydrolysis releases the overwhelming majority of lipid cores (Nishihara and Koga, 1987; Trincone *et al.*, 1992) in some instances: to cleave aminophospho head groups acetolysis combined with strong acid hydrolysis was necessary in order to release the lipid cores (Nishihara and Koga, 1987).

Column chromatography protocols allow for the elimination, from the total lipid extract or hydrolysed fraction, of the first apolar fraction containing hydrocarbons and pigments (e.g. carotenoids) and collection of the archaeal GDGTs in a polar fraction (second), whilst retaining on column a large number of compounds of much higher polarity than the GDGT lipids (Joo, Shier and Kates, 1968; Tornabene and Langworthy, 1979; Sinninghe Damsté *et al.*, 2002; Pitcher *et al.*, 2009; Lengger *et al.*, 2012).

1.4.2 Lipid separation and identification

The initial methods of analysis of the archaeal lipids were termed indirect, owing to the structural components of the lipid being identified in a series of steps, subsequently allowing collation of the information gained in each step to reconstruct the structure of a given lipid compound (Kates, 1963, 1977). Thus, the components of the lipid extract are first saponified to allow cleavage and identification of the polar head group (De Rosa *et al.*, 1983) and the released lipid cores are subsequently separated and chemically degraded to enable characterisation by gas chromatography mass spectrometry (GC-MS) (De Rosa *et al.*, 1977; Hoefs *et al.*, 1997; Schouten *et al.*, 2000) (**Figure 1-5**).



R = isoprenoid hydrocarbon
PG = phospho or glycosidic polar head group

Figure 1-5 General scheme of chemical degradation of phospho or glycolipids of the dialkyl glycerol diether (GDD) moiety (modified from Kates, 1977)

During chemical degradation of the lipid cores the linkage between the glycerol and the hydrocarbon backbone is cleaved either by hydroiodic acid (HI) or by the treatment with trichloroborane (BCl_3), yielding glycerol and the constituent alkyl halide chains. The released alkyl halides are often treated with $LiAlH_4$ or $LiAlD_4$ in order to convert them into hydrocarbons, or deuterated hydrocarbons, which aides in their analysis (Chappe *et al.*, 1979; Michaelis and Albrecht, 1979; Chappe, Albrecht and Michaelis, 1982). The released isoprenoid hydrocarbons are separated by gas chromatography (GC) and identified using flame ionization (FID) or mass spectrometric (MS) detection. The main disadvantage of this approach is that the structural information of the individual GDGT lipids is irreversibly lost due to the destruction of the innate configuration of the lipid core. Thus, it is impossible to assign the distribution of the cyclic motifs between the particular lipids core, as well as within a given glycerol dialkyl glycerol tetraether molecule. Although this methodology allowed successful identification of a tremendous variety of archaeal lipid structures, it is not suitable for the TEX_{86} analysis.

Over the past three decades, a number of alternative methods have been developed with an emphasis on the analysis and detection of the archaeal lipids cores in their intact form with the intention of gleaning structural information for the individual analytes. Due to the high boiling points and relatively high polarity of the ether lipids, they are not readily amenable to gas chromatography due to strong retention on the column and considerable peak tailing. A high temperature gas chromatography analysis of the archaeal ether lipids as their trimethylsilyl (TMS) esters derivatives was demonstrated as a viable route of addressing the aforementioned problems, mainly by reducing the polarity and boiling point of the ether lipids by the TMS capping of the free hydroxyl groups. As such the ester lipids were eluted from the GC capillary column at elevated temperatures and detected either by FID or MS (Nichols *et al.,* 1993; Lim *et al.,* 2012). This approach, however, proved to be less than satisfactory, due mainly to the peak resolution being far below that expected from the capillary GC system and the difficulties of interfacing the HTGC with the MS detector (caused by the considerable temperature difference between the GC and MS). Additionally, the need for specialized equipment and expensive GC columns, have limited application of this approach.

The use of high-performance liquid chromatography, with a range of detection techniques, has also been explored. Some of the detection techniques include: post derivatisation ultraviolet detection (UV) (Martz, Sebacher and White, 1983), fluorescence (Bai and Zelles, 1997), refractive index (RI) detection (Mancuso, Nichols and White, 1986) and evaporative light scattering (ELSD) (Shimada *et al.*, 2002). Some of the reported limitations of these methods include: poor chromatographic separation, limited dynamic range and considerable baseline drift (RI); HPLC pump noise; low reproducibility and lack of structural information for individual GDGT analysed.

A milestone development in the direct analysis of the GDGT lipid cores was the use of a liquid chromatography coupled to a mass spectrometer *via* an atmospheric pressure chemical ionization (APCI) interface (Hopmans *et al.*, 2000). In the method developed, the separation of the individual lipid cores was achieved under normal phase conditions (NP) and the determination and unambiguous assignment of the analytes was performed according to their *m/z* values (mass-to-charge ratio of ionized molecules). Due to the "soft" ionisation in the APCI source the analytes are detected in the mass spectrum as protonated molecule ions $[M+H]^+$ with minor ions $[M-18+H]^+$ and $[M-74+H]^+$ being characteristic of loss of one molecule of water and the glycerol moiety, respectively. The invaluable works of Knappy *et al.* (2009; 2012) and Liu *et al.* (2012) provided tools for structural assignment of the

isoprenoidal backbone of the lipid cores, which led to the identification of plethora of new membrane lipid components of the diverse members of the Archaea. A wide variety of MS systems have been reported in the literature: single quadrupole (Hopmans *et al.*, 2000; Sinninghe Damsté *et al.*, 2002), triple quadrupole (Huguet, Fietz and Rosell-Melé, 2013; Inglis *et al.*, 2015), ion trap (Escala, Rosell-Melé and Masqué, 2007; Liu, Lipp and Hinrichs, 2011) and time-of-flight (TOF) (Liu *et al.*, 2012). While, the mass spectrometer is often operated in mass scanning range, potentially allowing detection of broad range lipid cores, selected ion monitoring (SIM) offers up to two orders of magnitude better detection limits owing to the shorter duty cycle of the MS detector (Schouten *et al.*, 2007).

Quantification of GDGT concentrations has been achieved by means of external calibration using a single isolated GDGT lipid core (Wuchter *et al.*, 2004) or by addition of a synthetic glycerol trialkyl tetraether lipid ($GTGT_{46}$) as an internal standard (Huguet *et al.*, 2006). Although, the calculation of the TEX_{86} index is performed on relative abundances of the lipid cores comprising the index and accurate quantification of the underlying components is not necessary for this purpose, the absolute concentrations are needed for calculating the BIT index and the assessment of the terrestrial input of organic matter into a lacustrine or marine setting (see discussion in Chapter 1.3).

It is also worth noting, that the analysis and detection of the archaeal polar lipids in their intact form (IPL), with the polar groups still attached to the GDGT lipid core, can be achieved by HPLC coupled to electrospray ionisation mass spectrometry (ESI-MS) (Sturt *et al.*, 2004; Pitcher *et al.*, 2009; Wörmer *et al.*, 2013; Zhu *et al.*, 2013). While the information derived from studies of the IPL distributions can be valuable in phylogenetic mapping in ecological environment studies (Sturt *et al.*, 2004), their application in palaeotemperature reconstruction has not been reported.

1.4.3 Derivatisation of the lipid cores

The approach of chemical modification (derivatisation) may be viewed reasonable when the desired effect falls within one of the following categories: first, to alter the physical or chemical behaviour of a particular functionality within a compound; second, to exert different or new properties that were not available in the native compound. Examples of the derivatisation of the archaeal lipid cores, which fall into the first category, are

trimethylsilylation of the terminal hydroxyl groups, which is commonly utilized in order to lower lipid polarity and thus affinity for the stationary in high temperature gas chromatography analysis. Another, good example is acetylation of hydroxyl groups for the assessment of the number of free hydroxyl functionalities, which led to identification of structural functionalities of novel lipids cores, *e.g.* glycerol dialkyl triol (*m/z* 1246) (Knappy and Keely, 2012; Knappy *et al.*, 2014). A second category comprises applications that transform the underlying lipid molecules into analytes readily amenable to alternative detection techniques, such as spectroscopy and fluorescence. Very few reports concern derivatisation of the archaeal lipid cores, detailing their chromatographic behaviour and detector response. Nonetheless, the following two examples are worthy of comment.

Derivatisation of the isolated lipid cores (archaeol and caldarchaeol) of methanogens *M. thermoautotrophicum* and *M. hungatei*, using 3,5-dinitrobenzoyl chloride (DNBC) in the presence of pyridine was reported by Demizu *et al.* (1992).

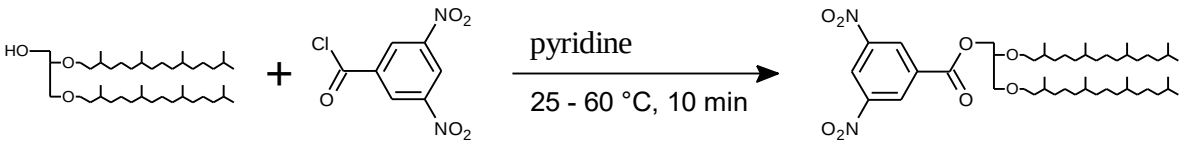


Figure 1-6 Scheme for derivatisation of the archaeol lipid core (GDD) using 3,5-dinitrobenzoyl chloride according to Demizu *et al.* (1992).

The modifications were carried out at moderate temperature (25 °C – 60 °C) and moderate excess of the derivatisation agent. The authors demonstrated that the derivatives of the main lipid cores (archaeol and caldarchaeol) were separated chromatographically on a C18 column under non-aqueous reversed phase chromatography (NARP, acetonitrile:isopropanol) and detected in UV (λ = 254 nm). The introduction of a chromophore allowed, with the use of an internal standard, estimation of the methanogenic cell material from environmental material with reasonable accuracy. Notably, however, the analytical separation developed failed to detect other known lipid components of the methanogenic archaea cell membranes, suggesting that a more comprehensive method is needed to address both the detection and resolution of the ether lipid components.

An alternative method of chemical modification to determine the archaeal ether linked glycerol lipids was reported three decades ago, by Ohtsubo *et al.* (1993), albeit the

derivatives were monitored under fluorescence detection. The authors demonstrated successful preparation of 9-antroyl esters of methanogenic archaeal lipid cores (**Figure 1-7**) and separation under normal phase chromatography conditions with detection by fluorescence.

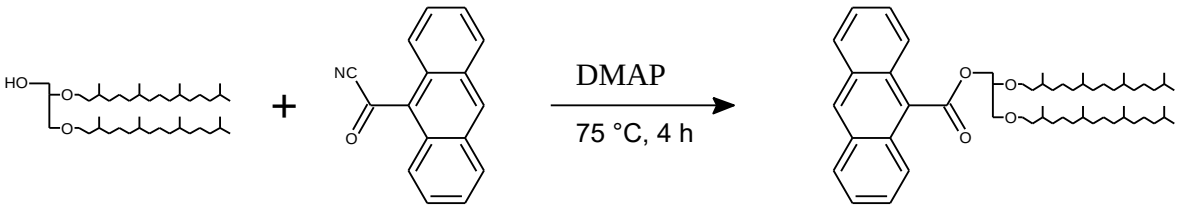


Figure 1-7 Scheme for the derivatisation of the archaeol lipid core (GDD) using 9-anthroyl nitrile according to Ohtsubo *et al.* (1993).

Although, the major lipid components were identified as the derivatives, the presence of many minor lipids in the chromatogram was not addressed. A subsequent report of derivatization of the ether lipids of six archaeal microorganisms and 2 environmental samples using 9-anthryoylnitrile, as reported by Bai *et al.* (1997), showed that the lipid cores can be effectively modified into structures readily amenable to analytical separation and fluorescence detection. This method does not allow unequivocal distinction of structures containing cyclopentane and cyclohexane motifs, which are used in the TEX_{86} palaeotemperature reconstruction.

1.4.4 Summary and aims

Thaumarchaeota, a phylum of Archaea, are ubiquitous microorganisms among the biosphere of the world's oceans. Their ability to adapt to changes in the growth environment is manifest in the alteration of the number of the cyclopentane and cyclohexane rings within the lipid core of the cell membrane lipids. Following the death of the organism the cleaved lipid cores are deposited on the ocean floor and, owing to their thermally and chemically stable glycerol – isoprenoid ether linkages, are well preserved over long geological timescales (GDGTs detected in a sample from Jurassic period, Kenig *et al.*, 1994). Thus, the deposited fossil lipids reflect the temperatures of the body of water at the time the lipids were formed by the

source organism, hence this correlation should allow reconstruction of past ocean's temperatures. Indeed, the empirically developed TEX_{86} index, using abundances of specific lipid core structure, has been calibrated to the available data of ocean temperatures and was shown to allow past sea surface temperature (SST) reconstruction with good accuracy.

At present, the analytical methodology employed in the separation and analysis of the lipid cores is overwhelmingly based on normal phase liquid chromatography (NP-LC) coupled with mass spectrometric (MS) detection. This approach, however, has been shown to have a number of limitations. The columns used to separate analytes during the analytical run require frequent back flushing due to adsorption of polar matrix components on the stationary phase, which in the long run may adversely affect the column performance. Although the lipid cores used in the TEX_{86} proxy elute relatively early in the analytical run, the nature of the underlying retention mechanisms on the stationary phase do not allow a rapid column flush of the column content in preparation for the next injection. The extended re-equilibration times between subsequent analytical runs, as well as system backpressure fluctuations, are commonly observed under NP-LC analysis. All these symptoms prevent this methodology from being used in a high throughput manner, effectively limiting the analytical capacity. Moreover, despite the use of new technology and small column packings, the current methods have not been shown to take advantage of the ultra-high pressure liquid chromatography systems, which have been demonstrated to offer improved analyte resolution and drastic reductions in analysis times.

Finally, many of the GDGTs in the TEX_{86} index occur in sediments in very low concentrations, requiring a considerable sample size (5-10 g) of the deposit to be used in order to record down core temperature variations, limiting the temporal resolution of the reconstruction.

Clearly, development of an alternative analytical approach that genuinely addresses the limitations of the present analytical methodology for the TEX_{86} index would benefit the broad scientific community involved in the reconstruction of ancient SST or study of climate change over geological timescales.

The main aim of the research work presented in this thesis was to develop a mild and efficient derivatisation protocol to permit high throughput analysis of GDGT lipid cores with reduced sample size requirements. The target for the approach was application in the characterisation

of GDGT lipid core distributions and for use in GDGT-based proxy measures such as the TEX_{86} index. The work described focuses on application of various amino acids to derivatise the GDGTs lipids under Steglich esterification conditions and the evaluation of the chromatographic, spectrometric and spectrofluorometric properties of the derivatives prepared in the study. Chapter 3 discusses the preparation of the *N*-Boc protected amino acid derivatives of 1-octadecanol (model lipid), method optimization as well as chromatographic and MS behaviour of the derivatives. Chapter 4 provides an assessment of the derivatisation of the model lipid using selected *N*-Fmoc protected amino acids, in addition to comprehensive evaluation of their chromatographic performance, MS behaviour and UV and fluorescence response. Chapter 5 focuses on the preparation of glycerol dialkyl diether lipid standard, preparation of its *N*-Fmoc-lysine(Boc) derivative as well as its chromatographic, UV and fluorescence and MS characterisation. Additionally, an *N*-Fmoc-lysine(Boc) derivative of cholesterol, and evaluation of its properties, are also discussed in this chapter. Finally, Chapter 6 investigates the application of the *N*-Fmoc-lysine(Boc) based derivatisation protocol to the lipid cores derived from *Sulfolobus acidocaldarius* (MR31), sediment (Oxford Clay S90-11) and a soil sample from an experimental piglet burial. Development of a novel reversed phase HPLC method for both the native lipids and their *N*-Fmoc-lysine(Boc) derivatives is also discussed.

Chapter 2 Experimental and analytical methods

2.1 General procedures

2.1.1 Solvents and reagents

All solvents used were of chromatographic grade (Fisher, Loughborough, UK), except methyl tert-butyl ether (MTBE, 99.8 %) was purchased from Scientific Laboratory Supplies (Nottingham, UK). The derivatisation reagents of certified purity were obtained from Sigma-Aldrich (Gillingham, UK): 1-octadecanol (99 %), DCC (99 %), DMAP (99 %), Boc amino acids (99 %), Fmoc-Phe (98 %), Fmoc-TRP (97 %), Fmoc-Gly (98 %), Fmoc-Pro (99 %), Fmoc-lysine(Boc) (99 %).

2.1.2 Glassware

To reduce risk of contamination, all the glassware used to carry chemical reactions were soaked in 1 % solution of Deacon-90 (Deacon, Hove, UK) overnight. This was followed by thorough rinsing with water, acetone and finally the glassware were pyro cleaned at 400 °C for at least 4h. In addition all HPLC mobile phase bottles, glass Pasteur pipettes, and HPLC vials were pyro cleaned in the same manner. The reagents were stored according to the manufacturer's recommendations.

2.1.3 Sample storage

The archaeal cellular material, soil and sediment samples was store at -20 °C prior to extraction. Extracted lipid material was stored at 4 °C for short term (up to 2 weeks) and at -20 °C for long term, unless stated otherwise.

2.2 Derivatisation methods

2.2.1 Derivatisation Method A

To a stirred solution of alcohol (1 eq.) in dry DCM, (5 mL) was added DMAP (0.1 eq.), *N*-Boc protected amino acid (1.1 eq.) and dicyclohexylcarbodiimide (EDC, 1.1 eq.). The reaction was stirred over ice for 30 min after which it was allowed to equilibrate to room temperature and reacted for a total of 3 h. Precipitated urea was filtered off and the filtrate evaporated *in vacuo*. The residue was reconstituted in DCM and washed twice with 0.5 M hydrochloric acid, saturated $NaHCO_3$, and then dried over $MgSO_4$. After filtration the dried solution was concentrated *in vacuo* and the product purified on silica gel column (0.55 cm × 5 cm) using DCM as eluent.

2.2.2 Derivatisation Method B

A solution of alcohol (5 × 10-3 mmol, 1 eq.), DMAP (0.5 eq.) and *N*-Boc protected amino acid (1.1 eq.) in dry DCM (5 mL) was cooled with stirring on an ice bath. EDC (1.1 eq.) was added in one portion and the solution was stirred at 0 °C for 2 h and at room temperature overnight. The solution was concentrated to dryness *in vacuo* and taken up in ethyl acetate (10 mL) and water (5 mL). The organic layer was separated, washed twice with 0.5 M hydrochloric acid, saturated $NaHCO_3$, water, and then dried over $MgSO_4$.The Solvent was removed *in vacuo* and the product was purified on silica gel column (0.55 cm × 5 cm) using DCM as eluent.

2.2.3 Derivatisation Method C

A glass screw top vial was charged with alcohol (1 eq.), DMAP (1 eq.), EDC (2 eq.), *N*-Boc protected amino acid (1.5 eq.), and DCM (2 mL). The vial was tightly capped, immediately placed in a laboratory ultrasonic cleaning bath and sonicated for 1 h. After that time the reaction mixture was passed through a short plug of silica gel and washed with DCM (2

column volumes), the product was eluted with hexane - ethyl acetate (1:1, v/v). The solvent was blown down under a stream of nitrogen gas leaving the pure product.

2.2.4 Derivatisation of lipid extracts form cellular material, sediments and soil samples

The lipid extracts derived from the archaeal cultures, sedimentary deposits and soil samples were derivatised according Method C, however the derivatisation reactants were used in large excess with regard to amount of the lipid extract. Assuming that the target lipid (e.g. cholesterol or caldarchaeol) comprised the whole amount of the extract the ratio of the reactants were DMAP/amino acid/EDC 10/50/75. Additionally, the sonication time was extended to 2 h.

2.2.5 Silylation of soil samples

Immediately prior to GC-MS analysis, the fraction iii of the soil extract (Chapter 2.4.2) was heated with *N,O*-bis(trimethylsilyl)trifluoroacetamide (BSTFA, 100 μL, containing 1 % of trimethylchlorosilane) and 5 drops of pyridine for 90 min at 60 °C before reducing to dryness under a gentle stream of nitrogen gas.

2.3 Preparation of lipid extracts

2.3.1 Archaeal cellular material

Sulfolobus acidocaldarius MR31 (Reilly, 2001) was grown aerobically at 75 °C by Daniela Barilla (University of York, UK) as reported previously (Knappy *et al.*, 2012). The cells were harvested, and subsequently stored at -20 °C until immediately prior to use.

2.3.2 Direct extraction or formation of ether lipid cores

The thawed cellular material (approx. 1g) was subjected to hydrolysis by refluxing in methanolic HCl (4.8 M, prepared from aqueous HCl and methanol) for 3 h to remove the polar head groups. The resulting solution was cooled, extracted with dichloromethane (DCM), dried by passage through a small column of anhydrous $MgSO_4$ and reduced *in vacuo*, yielding the total lipid extract.

2.3.3 Sediment samples

Approx. 20 g of the Peterborough member of the Oxford Clay collected from Stewartby (S90-11, Kenig *et al*., 1994), a sediment known to have appreciable GDGT contents was freeze dried by placing the unprocessed sample on a glass dish in a freeze dryer (Thermo Heto PowerDry PL3000). The sample was dried for 6 h at approximately 1 hPa achieving complete removal of ice. The dried material was then pulverized using a mortar and pestle and subsequently sieved through 1000 μm, 400 μm and 200 μm mesh. The sieve was agitated using sieve shaker (Endecottes Octagon Digital) for 5 min. The 200 μm fraction was collected for extraction and stored in a glass vial at -20 °C prior to use. The sediment sample was extracted using an ASE350 automated solvent extractor (Dionex, Sunnyvale, California, US) as reported previously (Schouten *et al*., 2007). Briefly, the soil was extracted three times with dichloromethane:methanol (9:1, v/v, 5 min, 100 °C and 1500 Psi), the solvent was removed *in vacuo* and the extract was fractionated using the procedure outlined in Chapter 2.4.1.

2.3.4 Soil samples

A 3-6 g sample of soil (C1, a non-gravefill control) from an experimental piglet burial (Hicks, 2016) was processed and extracted in the same manner as described in Chapter 2.3.3. The total lipid extract obtained was fractionated as outlined in Chapter 2.4.2

2.4 Fractionation of lipid cores

2.4.1 Cellular lipid extracts and sediments

Lipid core extracts from cultured materials were dissolved in minimal volumes of DCM and loaded onto glass pipette columns of activated alumina (5 mm × 50 mm) preconditioned with hexane:DCM (9:1). The apolar components were eluted with three column volumes of hexane:DCM (9:1) and the polar fraction containing liberated GDGT lipid cores were eluted with three column volumes of DCM:methanol (1:1). The solvent was removed under a gentle stream of nitrogen gas.

2.4.2 Soil samples

The extract was fractionated using a glass column (90 mm× 10 mm) packed with silica gel (750 mg). The extract was dissolved in dichloromethane and loaded onto the column. The column was eluted with three bed volumes of i) hexane, ii) hexane:toluene (1:1, v/v), iii) hexane:ethyl acetate (4:1, v/v), iv) dichloromethane:methanol (1:1, v/v) to produce four fractions containing: i) aliphatic hydrocarbons, ii) aromatic hydrocarbons, iii) alcohols and ester lipids and iv) acids, respectively. The solvent was removed using a rotary vacuum concentrator. Fraction iii was split in two, the first part was derivatised as outlined in Chapter 2.2.5 and subsequently screened for sterols and alkanols, whilst the second was derivatised as outlined in Chapter 2.2.3.

2.5 Mass spectrometry (LC-MS)

2.5.1 Mass spectrometer

HCTUltra ETD II ion trap mass spectrometer (Bruker, Daltonics, Coventry, UK) was used to record APCI positive mass spectra. The operational parameters were set at: nebulizer gas

(N_2) pressure 45 psi; drying gas (N_2) flow L min^{-1}; drying gas temperature 300 °C; vaporizer temperature 450 °C; capillary voltage -3500 V.

The analyte/sample solutions, which were directly infused into the ion trap, were delivered by syringe injection at 20 μL min^{-1} into a makeup solvent of methanol:dichloromethane (9:1, v/v) delivered at 0.3 mL min^{-1}, unless stated otherwise.

2.5.2 Tandem MS

MS/MS spectra were recorded using the HCT ion trap mass spectrometer during infusion and LC-MS analysis. The Auto MS^n feature, which automatically selects the base peak ion in the mass spectral scan for collision induced dissociation (CID) was used. The isolation width was set at 3 *m/z* window, the maximum accumulation time set to 40 ms and the fragmentation amplitude fixed at 1.2 V. In some instances, where indicated, the SmartFrag feature was used.

2.6 Sample analysis

2.6.1 Normal phase LC-MS

The purified lipid core extracts and *N*-Boc amino acids derivatives of 1-octadecanol were reconstituted in hexane:isopropanol (99:1, v/v) for normal phase LC-MS. The analysis was performed using Ultimate 3000 rapid separation liquid chromatograph (Dionex, Sunnyvale, California, US) coupled to the HCT ion trap mass spectrometer. Separation was achieved on Waters Spherisorb CN column (2.1 mm × 150 mm, 3.5 μm) maintained at 30 °C using a binary solvent system comprising of hexane:isopropanol delivered at 1 mL min^{-1}. The gradient program was consistent with the generally accepted method (Schouten *et al.*, 2007). Mainly, the analytes were eluted isocratically with 99 % A and 1 % B for 5 min, followed by a linear gradient to 2 % B in 45 min, where A) hexane and B) 2-propanol. The MS parameters were identical to those detailed in Chapter 2.5.1 and the scan range was set to *m/z* 950-1500 for the tetraether lipid cores and *m/z* 200-1000 for *N*-Boc amino acid derivatives of 1-octadecanol.

2.6.2 Reversed phase HPLC method for model compounds

The *N*-protected amino acid derivatives were analysed using an Ultimate 3000 (Dionex, Sunnyvale, California, US) rapid separations liquid chromatograph coupled to the HCTUltra ETD II ion trap mass spectrometer. The separation was achieved on a Dionex Acclaim RSLC 120 C18 column (2.1 mm × 150 mm, 2.2 μm). The column temperature and mobile phase composition are listed in **Table 2-1**. Method I was used to analyse the *N*-Boc derivatives of 1-octadecanol, method II was used to analyse *N*-Fmoc derivatives of 1-octadecanol and method III to analyse *N*-Fmoc-lysine(Boc) derivatives of 1,2-di-octadecyl-*rac*-glycerol (*r*-dOG) and cholesterol. The injection volume was set at 1 μL, unless stated otherwise.

Table 2-1 Reversed phase LC-MS methods gradient programs used in the analysis of the *N*-protected amino acid derivatives of the 1-octadecanol, cholesterol and *r*-dOG model lipids.

	Column temp (°C)	Flow rate mL min^{-1}	Time (min)	A	B	C	D
Method I	30	0.4	0 – 1.5	80	10	0	10
	30	0.4	1.5 - 7.0	67	10	0	23
	30	0.4	7.0 – 8.0	67	10	0	23
	30	0.4	8.0 – 9.0	80	10	0	10
	30	0.4	9.0 - 15	80	10	0	10
Method II	35	0.5	0 – 1.5	60	10	30	0
	35	0.5	1.5 – 10	55	5	35	0
	35	0.5	10 – 12	60	10	30	0
	35	0.5	12 – 17	60	10	30	0
Method III	40	0.5	0 – 1.5	60	10	30	0
	40	0.5	1.5 – 10	55	5	40	0
	40	0.5	10 – 12	60	10	30	0
	40	0.5	12 - 17	60	10	30	0

Note:
A – methanol
B – 0.5 % acetic acid in water
C – dichloromethane
D – ethyl acetate

2.6.3 Reversed phase HPLC method for the soil sample

The Fmoc-lysine(Boc) amino acid derivatives of cholesterol standard material as well as the soil sample extract were separated using Ultimate 3000 rapid separation liquid chromatograph (Dionex, Sunnyvale, California, US) coupled to the HCT ion trap mass spectrometer. Separation was achieved on Dionex Acclaim RSLC 120 C18 column (2.1 mm × 150 mm, 2.2 μm) maintained at 45 °C using a ternary solvent system, where A = methanol,

B = 0.5 % acetic acid in water, C = dichloromethane, delivered using a gradient program shown in **Table 2-2**, unless stated otherwise.

Table 2-2 Reversed phase LC-MS method IV gradient program, with percentage composition of solvents in the eluent

Time (min)	A	B	C
0	73	12	15
3	73	12	15
10	62	8	30
13	52	8	40
20	52	8	40
22	73	12	15

2.6.4 Reversed phase HPLC method for ether lipid cores and their derivatives

Analysis of the archaeal lipid core extracts and their Fmoc-lysine(Boc) amino acid derivatives were achieved using Ultimate 3000 rapid separation liquid chromatograph (Dionex, Sunnyvale, California, US) coupled to the HCT ion trap mass spectrometer. Separation was achieved on Phenomenex Kinetex PFP column (4.6 mm × 150 mm, 2.6 μm) maintained at 45 °C using eluent composed methanol/0.1 % acetic acid in water/methyl tert-butyl ether (40/20/40, v/v) delivered isocratically at 1 mL min^{-1}.

2.6.5 Gas chromatography - mass spectrometry

Analysis was performed using an Agilent 7860A gas chromatograph equipped with a 7683B series auto-sampler and coupled to a Waters GCT Premier time-of-flight mass spectrometer. A fused silica capillary column (Zebron, ZB-5, 30 m x 0.25 μm film thickness) was used

with helium as the carrier gas (1 mL m^{-1}) and the silylated total extract samples, dissolved in 100 μL of DCM, and 1 μL of the solution, were introduced *via* a split injection port (280 °C, split flow 1:5). The following temperature program was used: the oven was ramped from 70 °C to 130 °C at a rate 20 °C m^{-1}, and to 320 °C at a rate of 4 °C m^{-1} where it was held for 40 min. The MS transfer line was maintained at 300 °C and the electron ionisation energy was set at 70 eV. Spectra were produced using MassLynx v4.1 and identification of analytes was aided by comparison against the NIST 08 spectral library. Due to the complexity and the unpredictable content of the sample no internal standard was used.

2.7 Liquid chromatography with UV and fluorescence detection

2.7.1 HPLC-UV

The evaluation of the UV and fluorescence properties of the lipid derivatives with an active chromophore (Fmoc) was performed using an Agilent (Palo Alto, California, US) 1100-UV series chromatographic system. The separations were achieved using chromatographic conditions detailed in Chapter 2.6.2. The detector absorption wavelength was set at 263 nm.

2.7.2 HPLC-FLD

The appraisal of the fluorescent properties of lipid modified using amino acids possessing the *N*-Fmoc functionality were monitored using Jasco (Tokyo, Japan) FP-920 intelligent fluorescence detector (FLD) connected in series with the HPLC-UV system (Chapter 2.7.1), such that the UV detector effluent was directed into the inlet of the FLD detector. While the separation was controlled by the parameters of the LC methods outlined in Sction2.6.2, the FLD detector was settings were: absorption wavelength – 263 nm and emission – 309 nm.

Chapter 3 Preparation and characterization of *N*-Boc amino acids derivatives of 1-octadecanol

3.1 Introduction

Esterification is a fundamental chemical reaction both in nature and in synthetic chemical processes. The reaction, which effects the coupling of a carboxylic acid functionality with an alcohol, was first published by Fischer (1895).

In early experiments, the carboxylic acid was heated in the preferred alcohol in the presence of concentrated inorganic acid (one of H_2SO_4, H_3PO_4 or HCl) to yield an ester. This approach is limited in applicability to short-chain primary alcohols, giving poor results with secondary and sterically hindered tertiary alcohols due to their rapid decomposition. For example, *tert*-butyl alcohol dehydrates to form an inactive unsaturated by-product. Although, Lewis acids such as BF_3, $AlCl_3$, $InCl_3$ (Santos, Martinez and Mira, 1996; Mineno and Kansui, 2006; Jin *et al.*, 2014) have been used successfully under milder conditions, the need for rigorously dry conditions and the incompatibility with additional functional groups common in substrates limited the widespread application of Fischer esterification.

An alternative approach utilizes conversion of the acid to either an acyl chloride (Sonntag, 1953; Lillington, Trafford and Makin, 1981; Lepage and Roy, 1986; Villeneuve and Chan, 1997) or an anhydride (van Es and Stevens, 1965; Hassner, Krepski and Alexanian, 1978; Inanaga *et al.*, 1979), both of which are more reactive than the acid and undergo an addition reaction with alcohol in the presence of a base such as trimethylamine or pyridine. This method requires the separate preparation / availability of the acyl chloride or anhydride, adding to the steps in the process.

The importance of esterification has led to extensive experimental research over the past decades (Euranto, 1969; Iwasaki *et al.*, 2000; Nahmany and Melman, 2004; Saito *et al.*, 2005; Otera and Nishikido, 2010; Sirsam, Hansora and Usmani, 2016), resulting in the development of a direct reaction via acyl activation promoted esterification (Sheehan and Hess, 1955; Sheehan, Cruickshank and Boshart, 1961; Carraway and Koshland, 1972; Sunggak, Jae and Young, 1984; Saitoh, Shiina and Mukaiyama, 1998). Although, extensively used in peptide synthesis, the activation using carbodiimide had exhibited limited success in ester formation, due mainly to variable yields. Work by Neises and Steglich (1978), together with the discovery that addition of the basic catalyst 4-(dimethylamino)pyridine (DMAP) greatly increases the reaction yields, led to this route gaining considerable attention (DeTar and Silverstein, 1966; Hassner, Krepski and Alexanian, 1978). The ensuing efforts resulted in the discovery and evaluation of a plethora

of alternative carbodiimide activators and base catalysts (Sheehan and Hess, 1955; Höfle, Steglich and Vorbrüggen, 1978; Coste *et al.*, 1991; Han and Kim, 2004; Chan and Cox, 2007; Valeur and Bradley, 2009; Joullie and Lassen, 2010). Novel carbodiimides permit facile coupling reactions to be achieved, circumventing some of the most common limitations associated with other esterification methods including: decomposition of acid sensitive substrates or functional groups, low yields of sterically hindered reactants, the necessity for anhydrous conditions and the need for preparation of reactive intermediates, to mention a few.

The specific aspects of the Steglich esterification reaction: the carbodiimide activation mechanism, use of catalyst and the effects of reaction parameters on the yield of the product are discussed, in more detail below.

3.1.1 Steglich esterification reaction

Carbodiimides have been used extensively in peptide synthesis, the most prolific example being dicyclohexylcarbodiimide (DCC; Sheehan *et al*., 1955). Their function in the Steglich reaction is to activate the acyl group of the amino acid through formation of an *O*-acylisourea intermediate, which rapidly reacts with the amine group of another amino acid leading to formation of the amide product in good yield. It is also noteworthy that the only by-product in the reaction is urea, formed following attack of the amine on the *O*-acylisourea intermediate. Although this type of acyl activation has also been applied to ester synthesis, variable and poor yields have been reported (Sheikh *et al.*, 2010; Ter Horst *et al.*, 2010; Gilles *et al.*, 2015).

In 1978, Neises and Steglich reported on the esterification of carboxylic acids using DCC, noting tremendous increases both in reaction rates and in product yields when 4-dimethylaminopyridine was used as the base in the reaction. Following this discovery, other alkyl aminopyridines were synthesized and their properties as acylation catalyst evaluated by Höfle (1978). The intricacies of the reaction mechanism and the role of the acyl transfer catalyst have been studied in great detail (Boden and Keck, 1985; Xu *et al.*, 2005; Larionov, 2011).

The Steglich reaction proceeds under mild conditions in a low polarity aprotic solvent, such as dichloromethane, and may be carried out as a one-pot reaction at room temperature (**Figure 3-1**).

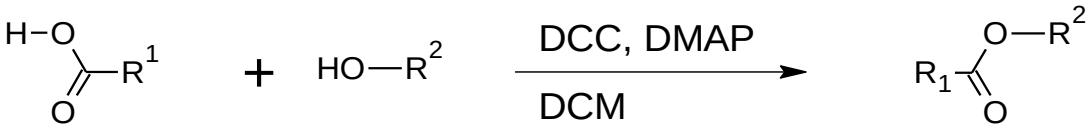


Figure 3-1 General scheme of the Steglich esterification reaction between carboxylic acid and alcohol using the DCC/DMAP coupling system.

The development of new types of carbodiimide activators and research into their modes of action has led to the availability of a vast array of coupling systems that can be "custom fit" to the synthetic purpose under consideration.

3.1.2 Coupling mechanism

Esterification using the Steglich catalyst proceeds as depicted in **Figure 3-2**. In the first step, the carbodiimide molecule is protonated following a proton transfer from the acid molecule. Then, addition of the carboxyl group to the carbodiimide function, most likely via an ion pair, forms an *O*-acylisourea, one of the most reactive acylating species (DeTar and Silverstein, 1966). This process occurs very rapidly. Unlike amines that react with the *O*-acylisourea readily, alcohols are weaker nucleophiles and the coupling is rather sluggish, often requiring hours or days to give product in good yields. In the second step, protonation of the *O*-acylisourea and simultaneous attack of DMAP on the carbonyl functionality leads to the formation of an acyl pyridinium intermediate. In the case of DCC, the urea by-product immediately precipitates. In the subsequent, and rate-determining, step the alcohol reacts with the acylated catalyst to form the ester and regenerate the (deactivated) catalyst.

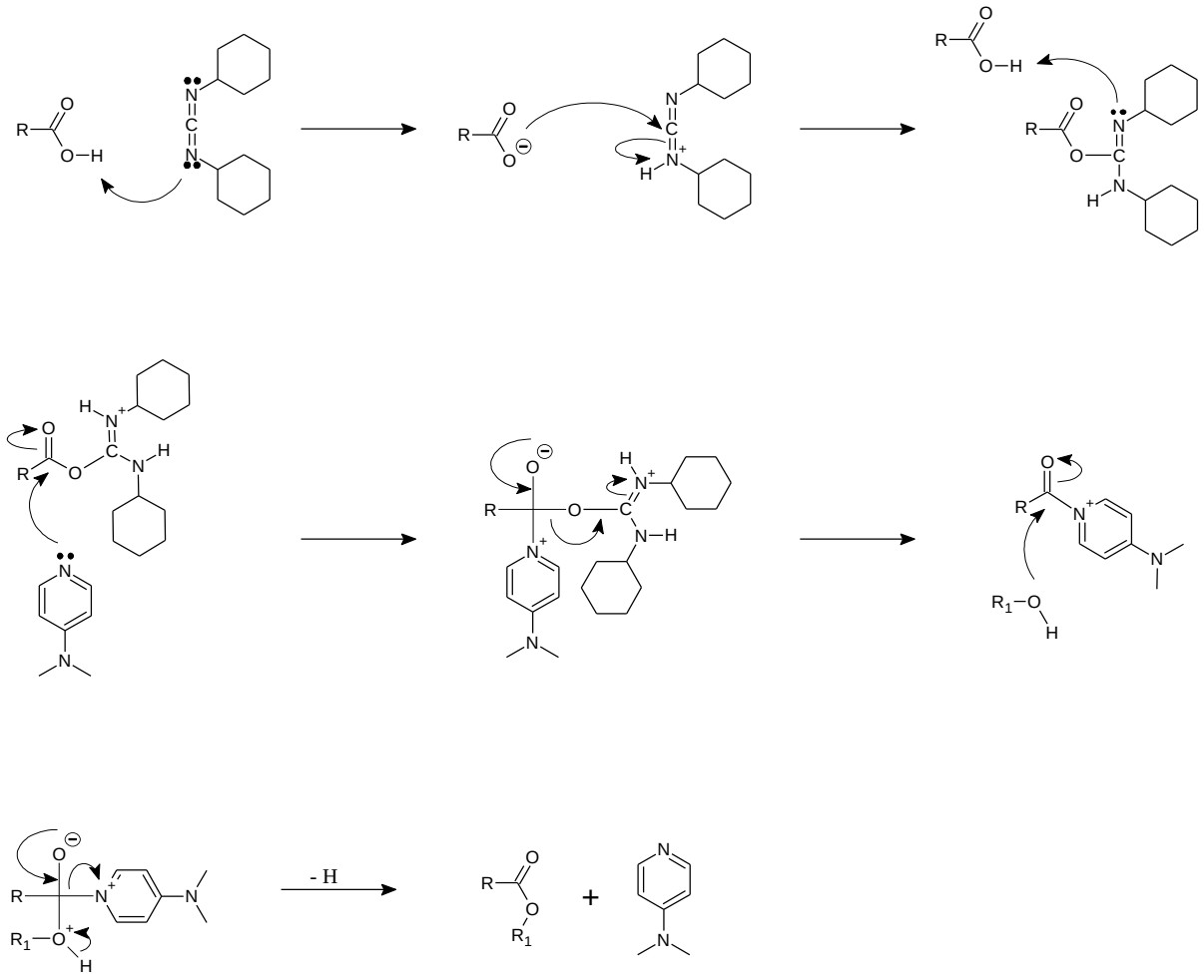
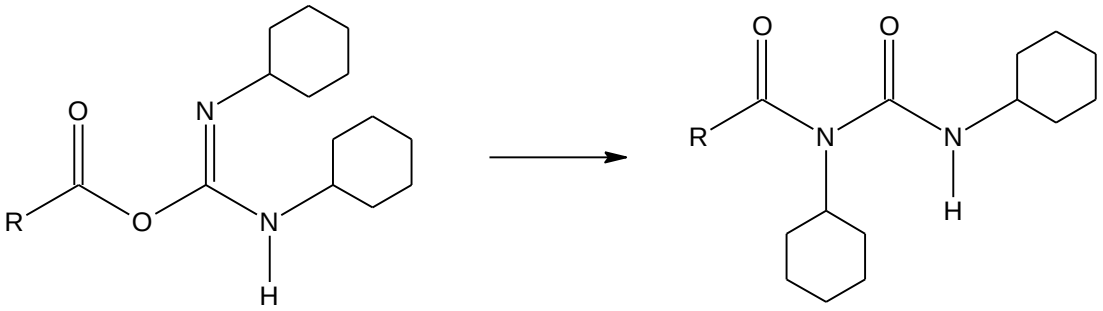


Figure 3-2 Mechanism of the Steglich esterification reaction using the DCC/DMAP coupling system (Neises and Steglich, 1985).

3.1.3 Factors affecting the reaction

It has been postulated that the Steglich reaction should be carried out in dichloromethane or an aprotic solvent of comparable polarity, such as diethyl ether or tetrahydrofuran (Neises and Steglich, 1985), solvents with higher polarity having been associated with lower reaction rates and lower product yields.

The highly reactive intermediate *O*-acylisourea may undergo a spontaneous competing reaction involving intramolecular acyl transfer to form *N*-acylisourea (**Figure 3-3)**. The unreactive *N*-acylisourea may significantly reduce product yield since its formation consumes the carboxylate without formation of the ester (Joullie and Lassen, 2010; Tsakos *et al.*, 2015).



***O*-acylisourea intermediate**

***N*-acylisourea intermediate**

Figure 3-3 General scheme of the collapse of the *O*-acylisourea, formed between DCC and carboxylic acid into *N*-acylisourea via intramolecular acyl transfer.

To prevent the formation of the unwanted *N*-acylisourea the reaction mixture is often initially cooled on ice for minutes up to hours, to slow its formation, before reaction is continued at room temperature. An alternative method suggests that an additive such as hydroxybenzotriazol (HOBt) may alleviate this problem (Koenig and Geiger, 1970). The HOBt readily protonates the *O*-acylisourea, aiding the formation of the urea by-product and the active intermediate and eliminating the possibility of the reaction being quenched at this point.

It has also been noted that excess of the carboxylic acid will lead to the formation of a symmetric acid anhydride, though this is not detrimental to the ester formation since anhydrides are also highly reactive species (DeTar and Silverstein, 1966; Hassner and Alexanian, 1978). Finally, it has been shown that the rate of the esterification reaction depends linearly on the concentration of DMAP (Xu *et al.*, 2005). Hence, increasing its ratio should result in higher reaction rates, allowing for faster ester formation and limiting the formation of the *N*-acylisourea.

3.1.4 Aims

There were three main objectives for this part of the study:

- to evaluate the applicability of the Steglich esterification method to the formation of *N*-Boc protected amino acids derivatives using model lipid: 1-octadecanol, and to develop and optimize the derivatization protocol so that it could be routinely applied in high throughput analysis of lipids extracted from sediment, microbial cultures and other samples settings;
- to investigate the chromatographic behaviour of the derivatives during the NP-HPLC mode that is regularly employed in GDGT analysis and explore the potential applicability of RP-HPLC to separation of the derivatives;
- to evaluate the mass spectrometric properties of the derivatives, including the fragmentation pathways during collision induced dissociation (CID) tandem mass spectrometry, to understand their potential benefits of structure elucidation of GDGT lipids.

3.2 Preparation of *N*-Boc amino acids derivatives of 1-octadecanol

3.2.1 Esterification using DCC activator

Selected *N*-protected amino acids were used as derivatisation agents in the evaluation of the applicability of the Steglich esterification to model lipid, the fatty alcohol, 1-octadecanol: *N*-Boc-phenylalanine (Boc-Phe), *N*-Boc-glycine (Boc-Gly), *N*-Boc-proline (Boc-Pro), and *N*-Boc-tryptophan (Boc-Trp). The selection of the amino acids was dictated mainly by their immediate availability in our laboratory though, as can be inferred from **Figure 3-4**, contrasting differences in the chemical structures of the amino acid side chains were also desired. It was envisioned that the size, rigidity, and presence of aromatic substituent might play an important role, potentially affecting the esterification reaction rates and the solubility, chromatographic and mass spectral properties of the ester derivatives.

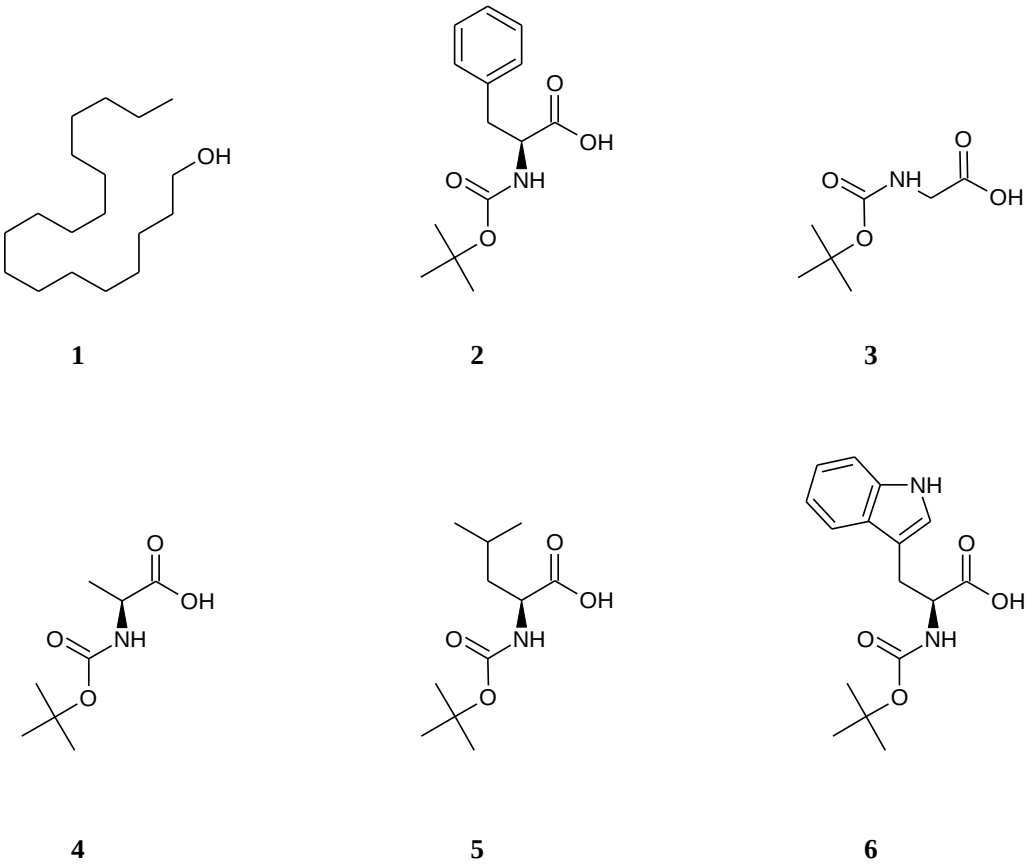


Figure 3-4. Structures of 1-octadecanol (**1**) and *N*-Boc protected amino acids: phenylalanine (**2**), glycine (**3**), alanine (**4**) leucine (**5**) and tryptophan (**6**).

N-Boc phenylalanine was selected as a derivatisation agent for the initial testing of the applicability of the original Steglich coupling of 1-octadecanol (Oct; Method A, for details see Chapter 2: Section 2.1.1).

The derivatisation (coupling) reaction was stirred over ice for 30 min after which it was allowed to equilibrate to room temperature and reacted for a total of 3 h. After reaction, the mixture was poured into vigorously stirred hexane to precipitate the urea by-product, followed by a normal aqueous work up and drying over magnesium sulfate. On evaporation of the solvent, the product was obtained as a white powder. The yield was calculated to be 76 % based on the weight of the product, which was satisfactory for demonstrating the utility of the method.

The purity of the ester products was evaluated by TLC and by APCI mass spectrometry (via direct infusion). Only two major ions at *m/z* 224.9 and *m/z* 418.3 were visible in the MS spectrum of the Oct-Boc-Phe product (**Figure 3-5**). The signal corresponding to the protonated molecule of the ester was not detected. Instead, a highly intense signal at *m/z* 418.3 was observed suggesting a loss of 100 Da from the protonated molecule. The formation of this ion is a characteristic for *N*-Boc amino acids and their esters and has been observed previously (Garner *et al.*, 1983). The mechanism of the formation of this ion is discussed in detail in subsequent sections.

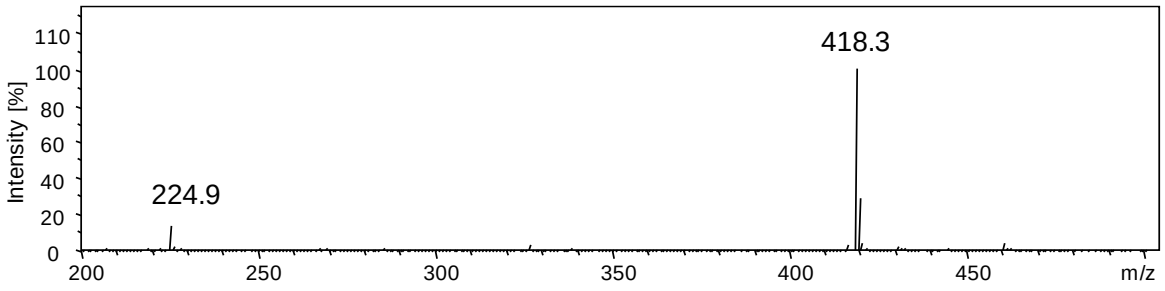


Figure 3-5 (+) APCI MS spectrum of Oct-Boc-Phe (**7**) ester obtained via direct infusion.

The signal at *m/z* 224.9 corresponds to the protonated molecule of dicyclohexylurea (DCU), a known by-product (contaminant) in the Steglich reaction, indicating that it was still present in the product, thus obscuring the real reaction yield. The removal of the DCU is commonly achieved by filtration of the precipitated urea, repeating the precipitation filtration cycles several times until no urea is present in the product.

Notably, during the course of the reaction, it was observed that the alcohol did not immediately dissolve in the solvent and it was present as solid floating on the surface for quite a long time after the cooling bath was removed. It was suspected that the low temperature applied at the start of the reaction lowered the already limited solubility of the alcohol, and that the reaction was not given enough time to reach the completion. Thus, by extending the reaction time at room temperature, more of the dissolved alcohol should be available to take part in the reaction which should result in higher yields. Although an increase in average crude product yield to 100 % was observed on extending the reaction time to 23 h (see **Table 3-1**), the presence of the urea by-product in the MS spectrum of the product suggested that the actual yield was lower.

Table 3-1 Yields of 1-octadecyl Boc-phenylalanine ester (**3-7**) formed in the Steglich coupling reaction using Method A with modifications.

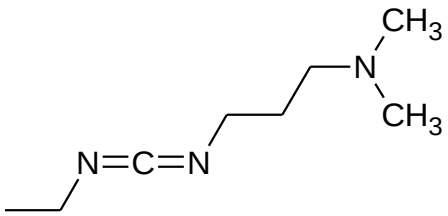
Total reaction time [h]	Time of cooling at 0 °C [h]	Crude yield [%] (*n*=3)
3	0.5	76 ± 12
23	3	100 ± 8
23	3	113[a] ± 9

[a] silica SPE cartridge was used to remove the urea by-product

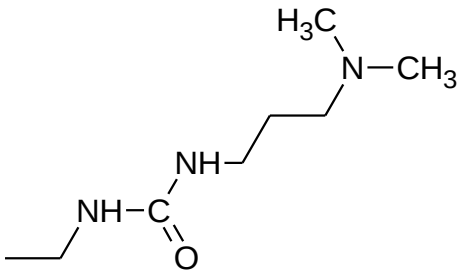
As an alternative, clean up of the crude product was attempted using silica solid phase extraction (SPE), with the expectation that the silica would adsorb the urea and allow for isolation of the product. This did not result in the desired outcome since the calculated average yield for the crude product was 113 % and the ion corresponding to urea by-product was again present in the mass spectrum of the product. It is possible that the complete removal of the undesired by-product could be achieved using multiple SPE cartridges to increase the overall retentive capacity of the sorbent, though this would result in an immediate and dramatic increase in the cost of carrying out the derivatisation protocol, potentially limiting its application due to financial constraints. Thus, the attention was shifted to an alternative approach.

3.2.2 Esterification using EDC activator

In the view of the hurdles in purification an alternative activating reagent, 1-ethyl-3[(3-dimethylamin)propyl]carbodiimide (EDC, **8**) was chosen as it offered significant potential advantage for reaction clean up, given that its urea by-product (EDU, **9**) is water soluble. The structures of EDC and its urea are shown in **Figure 3-6**. The presence of a tertiary amine in the structure of the EDU and its ability to protonate in a weakly acidic environment can be attributed to its partitioning in to the aqueous layer during liquid-liquid extraction (LLE) of the reaction mixture during the clean step of the esterification process (Joullie and Lassen, 2010; Tsakos *et al.*, 2015).



8



9

Figure 3-6 Structures of 1-ethyl-3-[(3-dimethylamino)propyl]carbodiimide (**8**) and its urea (**9**).

The 1-ethyl-3-[(3-dimethylamino)propyl] carbodiimide (EDC) has been shown to be successful in the activation and coupling of amino acids and short alcohols, such as *tert*-butanol, benzyl alcohol and methanol (Dhaon, Olsen and Ramasamy, 1982). The selected alcohols were used as protecting groups for the carboxyl groups of the amino acids intended for subsequent peptide synthesis. No comment on any attempts to apply this coupling method to higher or fatty alcohols could be found in the report.

Evaluation of the EDC activator was performed using Method B, which is taken from the method published by Dhaon *et al.* (1982) with details outlined in Chapter 2.2.2. The main changes from Method A were: the use of EDC activator instead of DCC and that the reaction

solution was cooled on an ice bath for 2 h prior to reaction being continued at room temperature overnight.

The yield of Oct-Boc-Phe (**7**) prepared using this method was 94 % (see **Table 3-2**). This was considered advantageous in terms of the conversion rate, and proved the usefulness of the EDC as carboxyl activator. Urea contamination, monitored by TLC, was not detected in the final product, suggesting that the aqueous work up achieved the sequestration of the urea by-product as intended. The product was identified by (+) APCI-MS (direct infusion) in which a highly intense signal at *m/z* 418.3, consistent with that expected for the Oct-Boc-Phe (**7**) derivative, was observed and no other ions were observed in the mass spectrum, indicating that the derivative was highly pure. The absence of mass ions pertaining to the coupler supports the decision to use an alternative acyl transfer agent and confirms that the increased polarity of the EDC and its urea should result in much more efficient isolation of the derivatives than when DCC was used as the coupler.

Following the successful preparation of Oct-Boc-Phe (**7**), reaction to prepare derivatives of 1-octadecanol using amino acids **3** through to **6** were carried out using Method B. The derivatives were obtained in very good to excellent yields (see **Table 3-2**)

Table 3-2 Yields of ester product of 1-octadecanol (**1**) derivatised using selected *N*-Boc amino acids according to Method B.

Derivatising Agent	**Crude yield [%], (*n*=3)**
Boc-Phe (**2**)	94 ± 4
Boc-Gly (**3**)	112 ± 6
Boc-Ala (**4**)	86 ± 4
Boc-Leu (**5**)	97 ± 5
Boc-Trp (**6**)	88 ± 6

High yields were achieved for the derivatives obtained using Boc-Ala and Boc-Trp amino acids and substantially better yields using Boc-Phe or Boc-Leu (**Table 3-2**). In the case where the derivatisation was performed using Boc-Gly, the recovered product yield was

calculated to be 112 %, suggesting that the product was impure. The source of the impurities was not investigated further. On repeating the derivatisation reaction using Boc-Gly the yield of the isolated ester was found to be 116 %, again indicating that the product was not pure. The purification of this derivative proved cumbersome. Despite changing the eluent composition and exploring application of solvents of differing polarities in column chromatography significant improvement in purification was not achieved, suggesting that the polarity of the Oct-Boc-Gly derivative was not very different to that of the impurities. Hence, the behaviour of product and impurities during silica gel column chromatography was very similar, resulting in marginal differences between their retention factors (R_f). Given that the application of eluents of different selectivity (ethyl acetate, dichloromethane) did not increase the difference between their R_f values, increasing the size of the column bed volume was considered as the only viable option. This would ultimately require longer elution times and volume of eluent in order to elute the product, rendering the method more time consuming. Alternatively, an acetylation of unreacted alcohol in the product material could be carried out to allow the amino acid derivative to be isolated more easily. Although these options could easily be incorporated into the derivatisation procedure should Boc-Gly be selected as the derivatising agent, in view of the much better performance of the other amino acids tested it was thought that any further optimization in the use of this particular amino acid should not be explored. It is reasonable, however, to assume that the esterification efficiency would be comparable to those observed for the other amino acids tested.

Comparing the pairs of amino acids that gave similar results in the coupling reaction using Method B, it is not immediately clear as to which factors led to higher yields. The structures of the Boc-Ala and Boc-Trp exhibit clear differences. The tryptophan's indole side chain is much larger than the simple methyl group found in alanine, hence, assuming that the steric effects were the rate limiting factor, it would be expected that the final product formation would be much faster for alanine than for tryptophan. This is probably not the case since the esters were produced with almost the same yield. Moreover, both of the amino acids are nonpolar in nature and dissolve well in the reaction solvent. Hence, limitations in solubility could not account for the observed differences in reactivity. Similarly, comparing leucine and phenylalanine, the similar reaction yields and higher yields than those for alanine and tryptophan could not be easily rationalised.

Taking into account that the reaction conditions were the same for all of the reactions (in terms of the solvent, temperature, ratios of the reagents, and the esterifying alcohol) the explanation possibly lies in the effects of the amino acid side chains on the dynamics and

structure of the acylpyridinium ion pair intermediate formed in the reaction. Studies of the mechanism of the DMAP catalysed acylation of alcohols using carbodiimide and acetic anhydride suggest that the nature of the counter ion in an ion pair with the acylpyridinium intermediate directly influences the rate of the acyl transfer and final product formation. Moreover, it has also been postulated that carboxylic acids in the reaction with carbodiimide can form the corresponding acid anhydrides. It could, therefore, be speculated that for the derivatisation reactions discussed, in presence of the EDC the amino acids form their corresponding acid anhydrides which subsequently undergo attack from the DMAP, leading to amino acid acylpyridinium ion pair intermediates and the amino acid carboxylate counter ion. Consequently, it may be suggested that any differences exhibited by the amino acids during this process will affect the reaction rates of the final product formation, leading to differences in product yield. Unfortunately, the reports on the DMAP catalysed acylation of alcohol did not investigate coupling examples such as those discussed in this section. Accordingly, the above reasoning cannot be verified. Exploration of the roles such effects may have on the kinetics of the reactions was not the subject of this research, thus they have not been investigated further.

It is worth noting that the amount of DMAP used in the reaction was increased to 0.5 eq., higher than the catalytic amount postulated by Steglich in his original report, and the reaction time was extended from 3 h to overnight. These changes, not surprisingly, led to higher product yields, in agreement with observations discussed previously and the expectation that by extending the reaction time and simultaneously increasing the ratio of the catalyst would have a cumulative effect and lead to increased product (derivative) yield.

3.2.3 Esterification using EDC activator under acoustic irradiation (sonochemical) conditions

Sonic irradiation has been shown to have positive impacts on chemical reactions including increasing reaction rates (Ando and Kimura, 1990; Srikrishna, Nagaraju and Sharma, 1992; Nakamura, Imanishi and Machii, 1994; El Fakih *et al.*, 1997; Lu, Cheng and Sheu, 1998; Cintas and Luche, 1999). Ultrasound induced chemical effects are attributed to extreme local conditions during cavitation and the collapse of cavitation bubbles. They are often linked with single electron transfer reactions involving formation of free radicals. However, it has been postulated that increased rates of reactions that follow ionic mechanisms are attributed

to mechanical effects, owing to enhancement in mass transfer (Mason, 1997; Naik and Doraiswamy, 1998; Luque de Castro, Priego-Capote and Peralbo-Molina, 2011). Fiamegos *et al.* (2004) demonstrated that a reaction that required stirring at ambient temperature for 80 min could be reduced to 10 min by sonication at 70 °C, while increasing product yield by 20-35 %. Although, the observed improvement was attributed to the transient hotspots of locally extreme conditions exerted by sonication the authors did not explain how the increased temperature contributed to the increased yields. Under sonic irradiation Kantharaju *et al.* (2006) obtained the methyl esters of a host of amino acids in almost quantitative yield just under 60 min *versus* 2-3 hours at reflux or 24 h under stirring. Jaita *et al.* (2015) demonstrated that methylation of *N*-protected Boc, and Fmoc glycine gave high yields (>80 %) in 20 min and 10 min, respectively, compared with 35 % under stirring.

The observation of relatively slow dissolution of 1-octadecanol under the conditions of Method A and the relatively long reaction times prompted evaluation of sonic irradiation in the derivatisation of the model alcohol using *N*-protected amino acids with the EDC/DMAP coupling system.

To test and optimize the derivatisation under acoustic irradiation conditions a trial reaction was performed using a modified Method B. Thus, the reactants, in unchanged ratios, were placed in a glass screw top vial (7 mL) containing dichloromethane (2 mL). The vial was tightly capped, immediately placed in a sonic cleaning bath and subjected to sonic irradiation. A number of parameters, such as bath water temperature, position and the level of immersion of the reaction vessel in the water bath, among others, have been reported to affect the ultrasonic intensity and the amount of the energy transferred to the reaction mixture (Nascentes *et al.*, 2001; VandenBurg and Price, 2012; Pchelintsev, Adams and Nelson, 2016). Trial sonic assisted derivatisation reactions of 1-octadecanol using Boc-Phe were carried out and the consumption of the alcohol was monitored via TLC. Several observations made over the course of the trial reactions reflect potential to optimize the reaction conditions. Firstly, the most optimal position for the glass beaker containing the reaction vial was in the centre of the sonic bath as visualized in **Figure 3-7**. Such positioning of the vial allowed the maximum disturbance of the reaction mixture by the sonic irradiation. Secondly, the time required for the alcohol to be fully consumed was 60 min. Thirdly, the highest reaction yields were observed when the amino acid:DMAP:EDC:alcohol ratio was 1.5:1:2:1. On the basis of the observations and empirical findings discussed above the derivatisation protocol, Method C, was formulated.

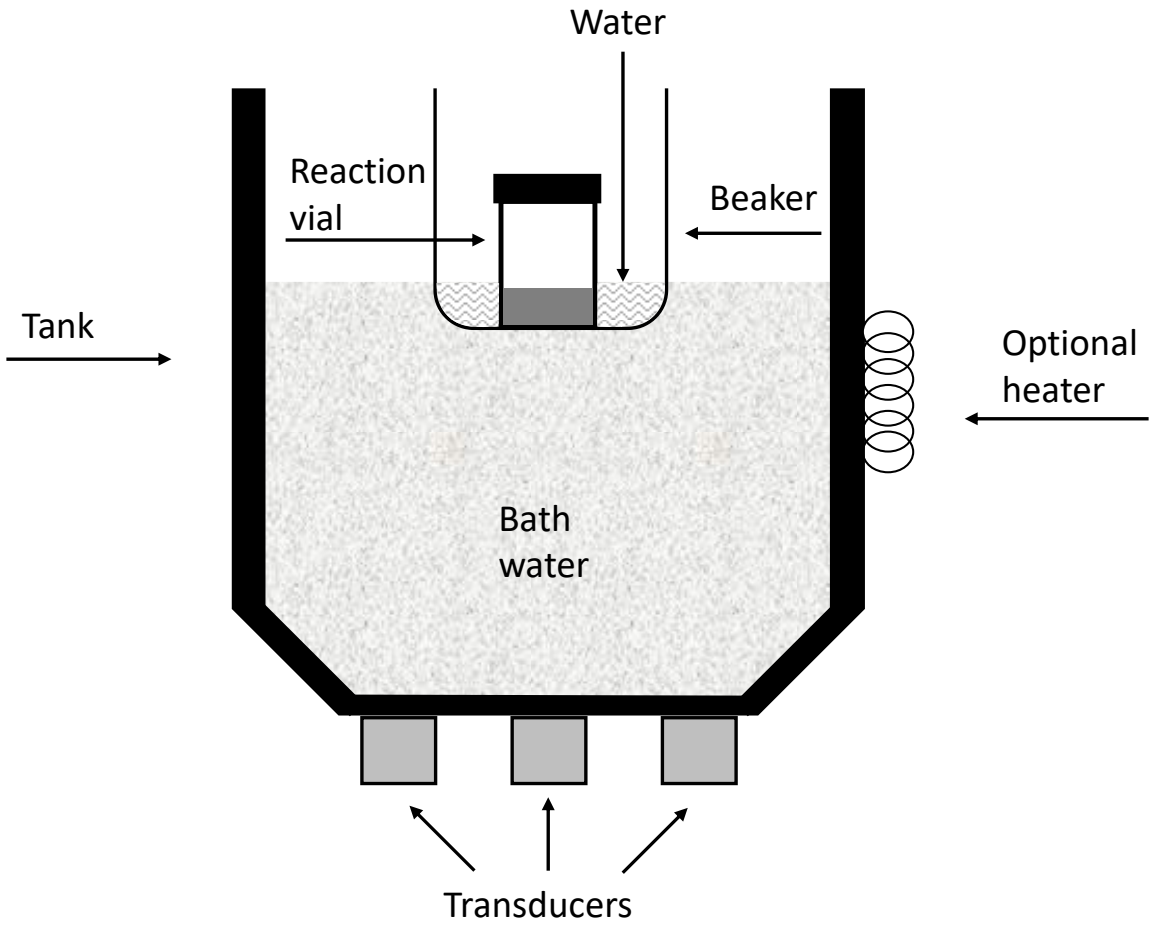


Figure 3-7 A scheme of a cross section of a sonication cleaning bath and the placement of a beaker containing the reaction vial during the sonic assisted derivatization reaction according to Method C.

During the derivatisation reactions in the sonic bath an increase of the bath water temperature was observed. The temperature measured using a laboratory thermometer showed an increase to 41 °C within the initial 30 min after which it remained essentially constant throughout the reaction time. Although some of the derivatisation reactions were carried out without allowing the bath water temperature to cool down the excellent yields obtained for those reactions suggest that this slight increase in water temperature did not affect the reaction condition. No evaluation of further increase in the bath water temperature on the reaction yields was performed. It is important to note that the elevated bath water temperature leads to increased evaporation rate, resulting in decrease in the bath water level. This must be monitored against the position of the reaction vial so that the reaction can be exposed to the maximum agitation forces as discussed above.

Having established the optimal reaction conditions with respect to the reaction vial position in the sonic bath, reactant ratio and reaction solvent volume, the subsequent derivatization reactions were carried out in triplicate on the same day in order to test repeatability of the protocol.

Under the conditions outlined in Method C, the derivatives of 1-octadecanol and selected *N*-Boc amino acids were obtained in excellent yields, in some cases nearly quantitative, as summarized in **Table 3-3**. The purity of the derivatives was determined using LC-MS.

Table 3-3 Results of derivatization of 1-octadecanol with selected *N*-Boc amino acids using EDC/DMAP coupling under sonic irradiation with accordance to Method C. The reactions were carried out in triplicates. Average yields and relative standard deviation (RSD) have are reported.

Derivative	**Yield [%]**				
	1	2	3	Average	RSD
Boc-Phe-Oct (7)	98.0	98.1	97.9	98.0	0.1
Boc-Gly-Oct (10)	95.5	95.2	95.4	95.4	0.2
Boc-Pro-Oct (11)	86.7	86.4	86.3	86.5	0.2
Boc-Trp-Oct (12)	92.9	92.4	92.6	92.6	0.3

The data in **Table 3-3** shows that excellent yields were obtained when Boc-Phe, Boc-Gly and Boc-Trp were used as derivatising agents. Only in the case of Boc-Pro was the average yield of the derivative low, at 86.5 %. It is also noteworthy that the tight RSD were achived for the triplicates for each of the derivatization reactions. The lower yield in case of Boc-Pro is most likely attributed to the proline ring-strain and the proximity of the *N*-Boc protecting group to the carboxyl group exerting additional steric overcrowding. Although no reason was offered, Li *et al*. (2008) reported only 76 % yield of methyl proline ester, compared with 96 % for glycine and 96 % for phenylalanine methyl esters for methylation reactions carried under the same conditions.

Interestingly the derivatization using Boc-Gly resulted in an excellent and highly reproducible yield. The reactions were monitored using TLC and no starting material

(alcohol) or other impurities were detected after 60 min. This contrasts with results obtained using Method B, where the derivatives exhibited excessive product yields for this amino acid.

The results show that derivatization reactions carried under the conditions of Method C lead to greatly improved yields and that sonication alleviated some of the issues encountered under stirred reaction conditions. The method is reliable; thus, it should find interest in applications where routine analysis of hydroxyl lipids as their derivatives is considered and where high precision and reproducibility are of importance.

Although, the results obtained for the proline derivatives are inferior to those for Phe, Gly and Trp, this does not preclude its use for the derivatization purposes: if it were selected a calculated derivatization efficiency, based on the initial amount of the native compound, would enable a correction factor to be incorporated for quantification.

The rise of the water bath temperature over the course of the reaction was, at first, considered as detrimental, due to possibility of the formation of the *N*-acylisourea, a deactivated by-product responsible for lowering of the yields of the product. This was thought to be especially true in the light of Steglich's report (Neises and Steglich, 1978), which advocated initial cooling of the reaction mixture to minimise this effect. On the other hand, some authors postulated that in the case of hindered reactants, heating the reaction mixture led to improvement in the effectiveness of the coupling. These two views on the reaction conditions appear to be in opposition to each other. Evaluation of the derivatization results given in **Table 3-3** indicates that such concerns are unfounded, in view of the excellent yields obtained for the majority of the coupling reactions. Further to that, it is reasonable to assume that the elevated reaction temperature improved the solubility of 1-octadecanol and, with acoustically enhanced mass transfer, resulted in greater yields and shorter reaction times, circumventing the low alcohol solubility observed for the stirred reactions. Notably, the similar reaction times and product yields obtained in reactions carried out with bath water temperature at ambient and ca. 40 °C suggest that the slightly elevated reaction temperature did not impact the reaction rates significantly.

It is reasonable to assume, that the acoustic irradiation enables the reacting species to reach appropriate energy levels and that the optimal orientation of the active intermediate, followed by the subsequent attack of the alcohol, was aided by an improved mass transfer during the formation and collapse of the cavitational bubble. These factors would have led to markedly increased reaction rates and the observed very high products yields. This is in

agreement with reports by Luche *et al.* (1990) who concluded that only mechanical effects of acoustic waves apply to reactions where the ionic intermediate is not generated via cavitation. Interestingly the effects are similar to hydrodynamic cavitation effected by high stirring rates (Mason, 1997).

3.3 Characterization of the derivatives of the model lipid compounds

3.3.1 Generalities

Lipid containing samples including the archaeal tetraether lipid cores are almost exclusively analysed by means of normal phase high performance liquid chromatography (NP-HPLC) interfaced with atmospheric pressure chemical ionization mass spectrometry (APCI-MS). This is mainly due to their highly non-polar character, pronounced retention under reversed-phase elution mode and ineffective ion formation under electrospray ionization mass spectrometry conditions (ESI-MS) (Hopmans *et al.*, 2000; Schouten *et al.*, 2007; Becker *et al.*, 2013). A recently reported reversed-phase (RP-HPLC) method of separation (Zhu *et al.*, 2013) resembles more the non-aqueous reversed-phase (NARP-HPLC) mode rather than classic RP-HPLC, mainly due to the absence of an aqueous component in the mobile phase. At present, there are no reports of a classical RP-HPLC method applied in the separation of tetraether lipid cores.

The oxycarbonyl moiety in the *N*-protecting groups of the amino acid derivatising agents used in this study follows a known fragmentation behaviour that has been studied previously for *N*-Boc (Garner *et al.*, 1983; Vairamani *et al.*, 1990). The ionization and fragmentation of this type of protected amino acid are well established, allowing the specific losses occurring on ionization or during tandem MS dissociation studies to be assigned, which can greatly assist in the reconstruction of the mass of the protonated molecule and in structural elucidation studies of the analyte under investigation.

The derivatised lipids do not have any functional group in their structure that can readily be protonated in solution, hence their ionization is not very efficient under electrospray ionization, and positive mode atmospheric pressure chemical ionization (+APCI) remains the only viable option. This technique is widely viewed as a soft ionization technique which, owing to limited fragmentation, often leads to the formation of a prominent protonated molecule $[M+H]^+$ or an adduct if a suitable salt is used, e.g. $[M+NH_4]^+$.

To date, there are no reports of implementation of a derivatisation method in the analysis of archaeal tetraether lipids in order to evaluate the potential benefits to their chromatographic behaviour and mass spectrometric response.

3.3.2 Detection of *N*-Boc amino acids derivatives of 1-octadecanol.

The mass spectra of the 1-octadecanol *N*-Boc amino acid modified derivatives were obtained during infusion of analyte solutions into the (+) APCI-MS *via* a syringe pump. The direct infusion of the analyte solution allowed the ionization mechanism to be established and the MS parameters to be tuned to optimize the ion formation process.

The ionization and ion formation of *N*-Boc protected amino acids has been characterized previously (Garner *et al.*, 1983; Du *et al.*, 2003; Zhu *et al.*, 2006; Reddy *et al.*, 2007; Zhang, Fang and Smagin, 2014). Although, the published reports do not pertain to hydroxyl lipids esters of *N*-protected amino acids, the ionization mechanisms discussed therein describe the observed ionization of the derivatives very well.

During the ionisation of the derivative molecule, the Boc group is cleaved off following a McLafferty type rearrangement, according to the mechanism shown in **Figure 3-8**, resulting in a fragment ion at 100 Da lower than that of the protonated molecule. This loss often results almost exclusively in a highly intense ion signal in MS. It is important to mention that, depending on the amino acid side chain and MS conditions, the ions resulting from a loss of 56 Da may also be present in the mass spectrum, however at a lower relative intensity.

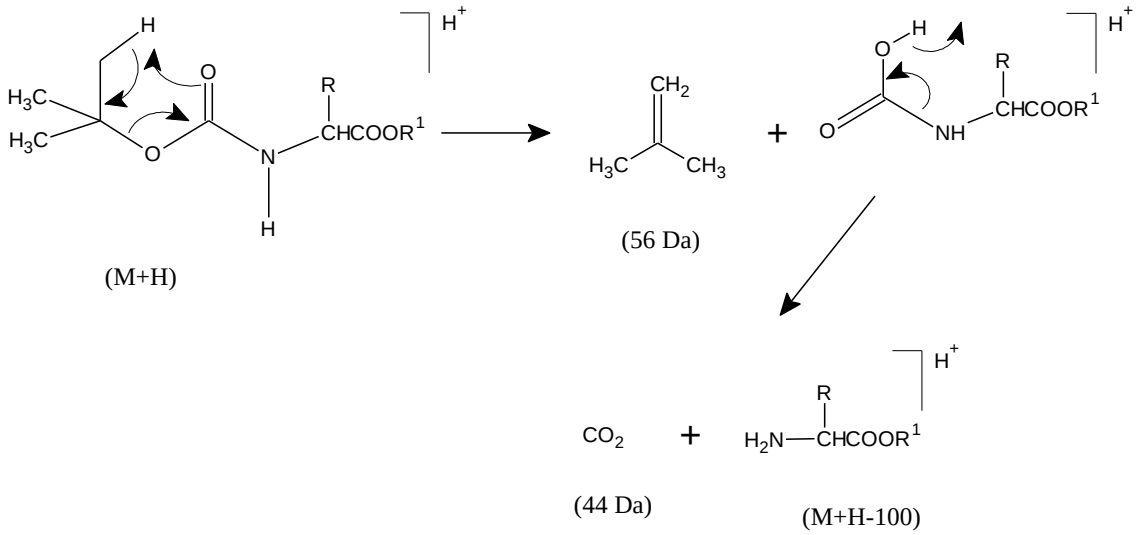


Figure 3-8 Mechanism of a McLafferty type rearrangement resulting in a loss of the protecting group in hydroxyl lipid esters of *N*-Boc amino acid. Note: R^1 represents lipid moiety, and R the amino acid side chain.

During the APCI ionization process no ions corresponding to the protonated molecule were observed in the mass spectrum, indicating that the cleavage of the Boc group occurs during the ionization step. In fact, the major ions observed for the *N*-Boc amino acid derivatives of 1-octadecanol were those of 100 Da lower than the value expected for the protonated molecule: *m/z* 328.3 for glycine, *m/z* 418.4 for phenylalanine, *m/z* 368.3 for proline and *m/z* 457.4 for tryptophan (see **Figure 3-9**A-D).

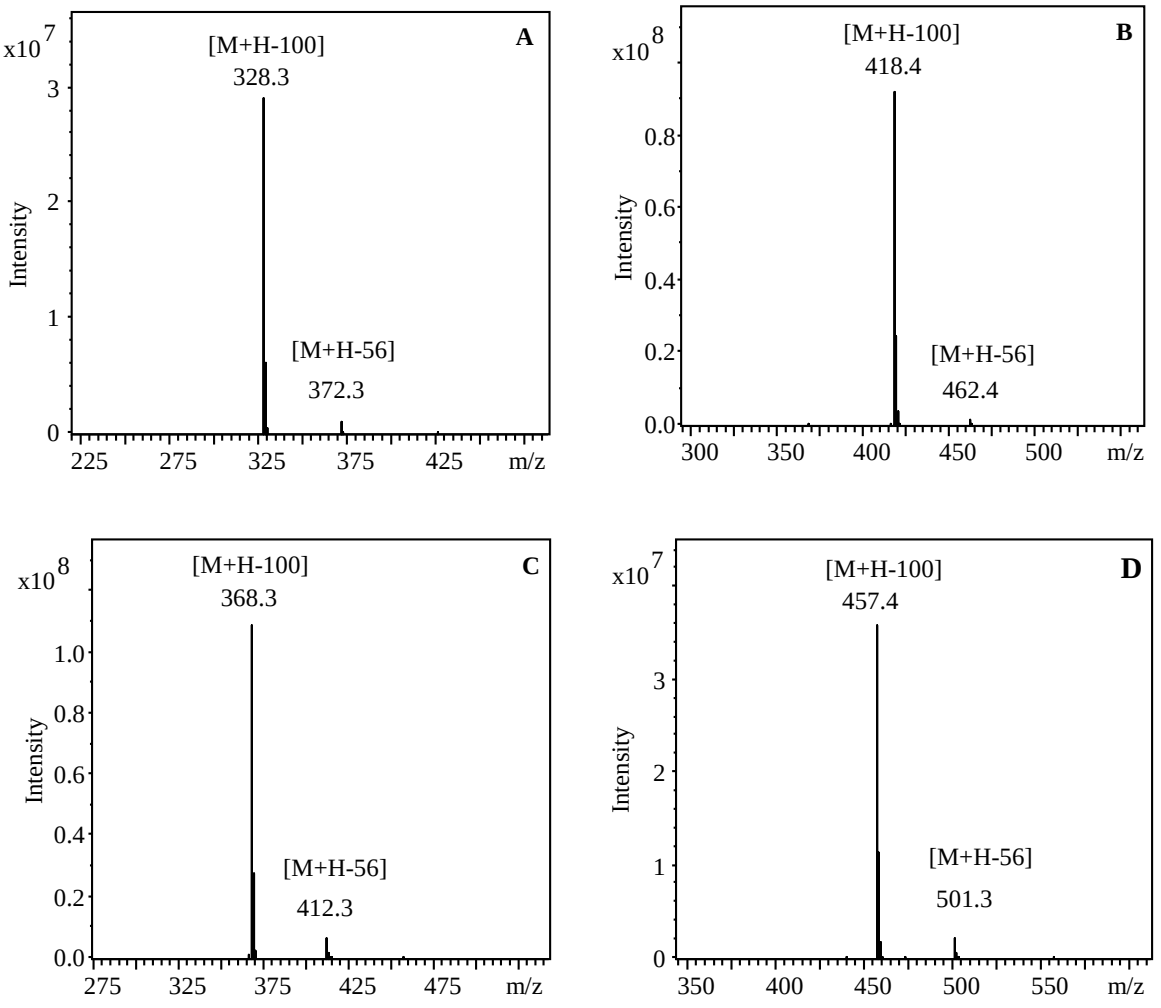


Figure 3-9 Mass spectra of the 1-octadecyl esters of N-Boc amino acids: A) glycine, B) phenylalanine, C) proline and D) tryptophan. The fragment of the spectrum shown represents the region where the most abundant ions were observed.

Interestingly, all derivatives have an even *m/z* value, with the exception of tryptophan, which has an odd value, conforming to the "nitrogen rule". According to this principle if the number of nitrogen atoms present in a molecule is odd, the molecular mass will be an odd number; if the number of nitrogen atoms present in the molecule is even (or zero) the molecular mass

of the molecule will be an even number. Hence, the protonated molecule of the former will be an even *m/z* ion and the protonated molecule of the latter will be an odd *m/z* ion. All of the derivatives tested contain only one nitrogen atom except tryptophan, which has two. This, indeed, is observed in mass spectra of the *N*-Boc amino acids derivatives.

The other, much less intense, ions formed during ionization of all of the derivatives result from a loss of isobutene (56 Da) from the protonated molecule. Judging by the low intensity of those ions it may be assumed that they are not very stable and fragment further, through expulsion of CO_2 (44 Da), leading to the fragment ion at 100 Da lower than the corresponding protonated molecule.

It is important to note that during the direct infusion (+) APCI-MS, the derivatives produced very similar mass spectra in terms of the number and type of the ions formed under both normal and reversed-phase chromatography solvents conditions. This in turn suggests that the ionization of the *N*-Boc amino acids derivatives is solvent independent (within the range of solvents tested) and may be characteristic of the mode of ionization. Notably no ions arising from the loss of the fatty alcohol or any ions related to it were observed in the mass spectra.

The MS behaviour of the *N*-Boc amino acid derivatives of the model lipid, mainly producing a highly intense and stable fragment ion accompanied by very little other fragmentation, allows the ion current to be narrowly focused and should result in appreciable response when analysing this type of derivative *via* APCI.

3.3.3 Normal phase high performance liquid chromatography mass spectrometry.

Due to the highly apolar character, similar to archaeal tetraether lipid cores, HPLC separation of the *N*-protected amino acid derivatives of 1-octadecanol was first performed under the normal phase APCI-MS conditions reported by Schouten *et al*. (2007). This method has become the standard analytical method for the GDGT lipid cores. The separation is achieved on a silica stationary phase chemically bonded with a propyl amino functionality, and the retention mechanism is based on adsorption of the analytes on the surface of the stationary phase. The mobile phase consists of a mixture of a non-polar solvent such as hexane or heptane (weak solvent) and more polar solvent such as ethyl acetate or isopropanol (strong solvent). The elution of the analytes is effected by increasing the eluotropic strength of the

mobile phase through an increase in the proportion of the strong solvent. In the applied method the mobile phase consisted of hexane - isopropanol (99:1, v/v).

A solution containing equimolar concentrations of each of the Boc-Gly, Boc-Phe and Boc-Trp derivatives of 1-octadecanol was prepared in hexane - isopropanol (99:1, v/v) and was analysed according to the standard method noted above (Chapter 2.6.1).

The derivatives were detected by recording the signal from the ion at *m/z* [M+H – 100]. The target ions for the derivatives tested were: *m/z* 328.3 for Boc-Gly-Oct (**10**), *m/z* 368.3 for Boc-Pro-Oct (**11**), *m/z* 418.4 for Boc-Phe-Oct (**7**), and *m/z* 457.4 for Boc-Trp-Oct (**12**). The mass chromatogram (**Figure 3-10**) shows that the derivatives can be successfully chromatographed in normal phase mode and that the peak geometry is very good and resembles that of a Gaussian distribution. Baseline separation of the derivatives was afforded for Boc-Phe and Boc-Trp derivatives whereas the Boc-Gly and Boc-Pro derivatives were only partially resolved. Although, it would be possible to resolve the pair in question by modifying the initial composition of the mobile phase this was not attempted because the separation was satisfactory to establish that these derivatives chromatograph well under the standard method routinely employed in analysis of tetraether lipid cores. It is noteworthy that the first three derivatives elute within the initial isocratic period of the gradient method Whereas the Boc-Trp derivative eluted much later.

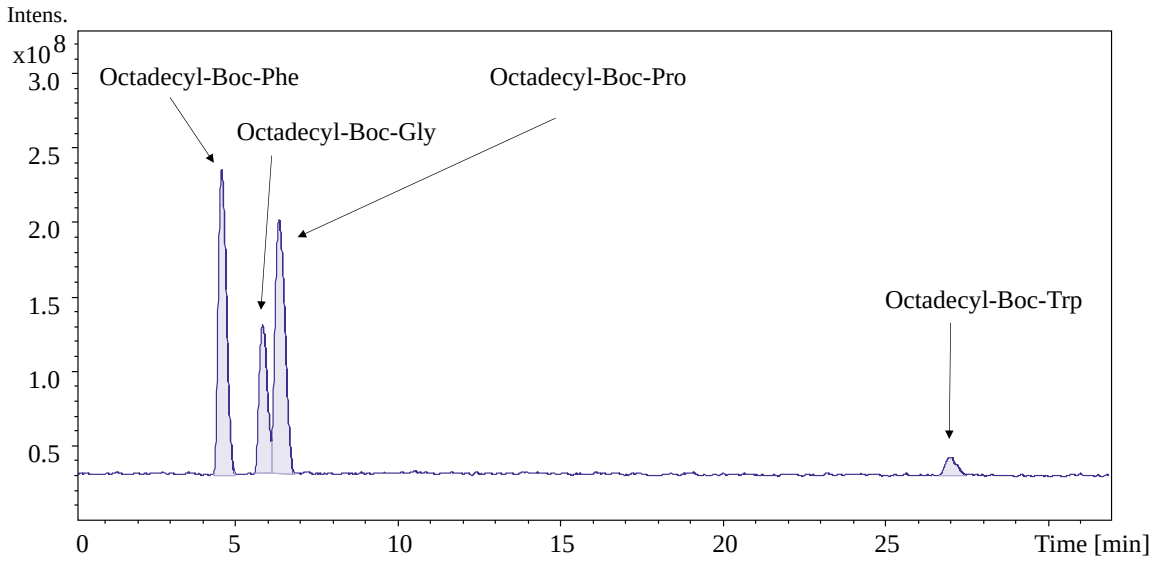


Figure 3-10 (+) APCI-MS base peak mass chromatogram of sample containing 1-octadecyl esters of Boc-Gly, Boc-Phe, Boc-Pro and Boc-Trp at 1 mM concentration. Separation was achieved using Waters Spherisorb CN column (2.1 mm × 150 mm, 3.5 μm) in normal phase mode.

The analysis of the Boc-protected amino acid derivatives allowed for a number of important observations to be made. First, the polarity of the derivatives analysed could be assessed by comparing their retention times to that of caldarchaeol (GDGT-0, approx. 12 min) under the same chromatographic conditions. The retention time of Boc-Phe-Oct is 4.6 min, Boc-Gly-Oct is 5.8 min, Boc-Pro-Oct is 6.3 min, and of Boc-Trp-Oct is 27.0 min. It is readily apparent that derivatives of Phe, Gly and Pro, which elute before GDGT-0, exhibit more apolar character and that only the Trp derivative is much more polar than the GDGT-0 tetraether lipid core. The short retention times suggest that there is little interaction between the derivatives and the stationary phase of the column and that increased adsorption is evident only for the Trp derivative. This is most likely explained by the differences in the side chain of the amino acids used in the derivatization.

The cyano group of the bonded phase exhibits dipolar properties due to the electronegativity difference between nitrogen and carbon, hence, it can act as a proton acceptor. Similarly, a dipole or induced dipole may exist in a compound having a partially substituted amino group such that the amino proton may carry a partial positive charge. Such analytes may interact with the cyano bonded stationary phase, resulting in their retention during chromatographic separation (see **Figure 3-11**). The degree of interaction depends on the accessibility of those hydrogens to interact with the stationary phase. All of the Boc-protected amino acids have their *N* terminus protected by the Boc group and experience steric hindrance with regard to its hydrogen atom. Thus, any potential interactions with the stationary phase may be limited. Comparing the side chains of the Boc-protected amino acid derivatising agents used it is readily apparent that only tryptophan has an additional hydrogen atom, located on the indole ring, capable of engaging in such interactions. This would imply that the retention time of the tryptophan derivative is longer mainly due to the stronger interaction than exists with the other derivatives. This is indeed in agreement with the observed elution order and much longer retention of the tryptophan derivative.

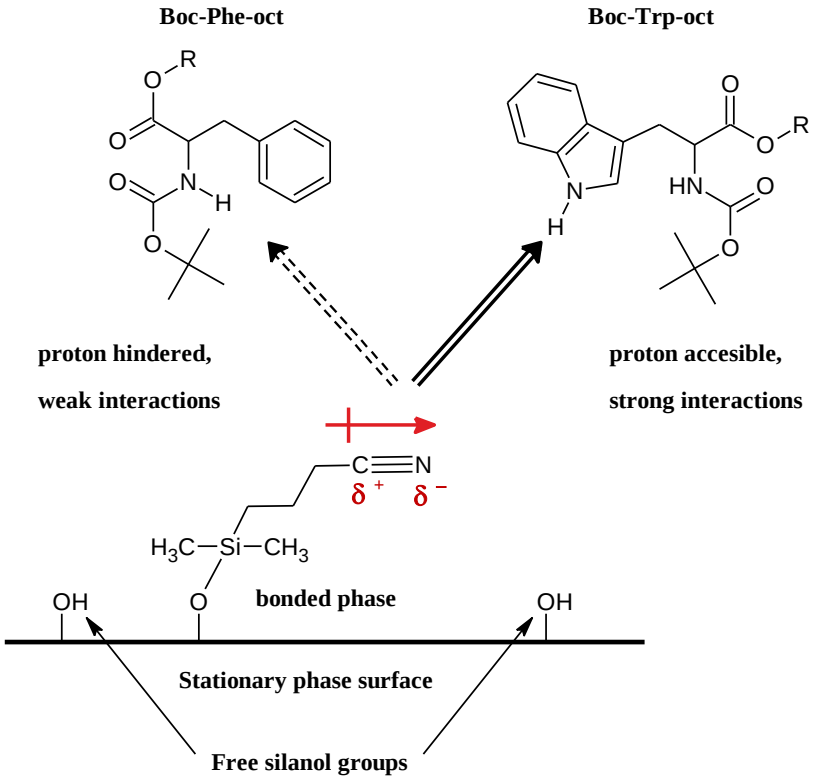


Figure 3-11 Scheme showing presence of a dipole on a cyano bonded column packing material and its potential interaction with amine protons of the Boc- protected amino acid derivatives (Phe and Trp). The R represents the 1-octadecyl group.

The second important observation is that differences in the peak areas of the derivatives are apparent in the mass chromatogram during normal phase chromatography. The peak areas of the derivatives prepared at the same concentration (1 mM) are given in **Table 3-4**. The evaluation of peak areas was performed using area percent relative to the derivative that produced the highest peak area value. This was done to allow easier comparison of the individual peak areas.

The derivatives were detected and identified based on the highly intense ion (target *m/z*) that is 100 Da lower than that of their corresponding protonated molecule. The ionized derivatives almost exclusively produced this ion, thus it was deemed fair to assume that, given the identical ionization conditions, these ions were formed in the same way. The largest peak areas were observed for the derivatives prepared using Boc-Pro and Boc-Phe. Somewhat surprisingly, the area percent of the peaks produced by the other two derivatives were much lower with Boc-Trp giving a very weak signal.

Table 3-4 Retention times and peak areas of Boc-protected amino acid derivatives of 1-octadecanol chromatographed using NP-HPLC-APCI-MS. Area % of each peak in the total chromatogram was also calculated.

Derivative	t_R [min]	Target *m/z*	Peak area	Area %
Boc-Phe-Oct (7)	4.6	418.4	3.58×10^9	99
Boc-Gly-Oct (10)	5.8	328.3	1.75×10^9	49
Boc-Pro-Oct (11)	6.3	368.3	3.60×10^9	100
Boc-Trp-Oct (12)	27.0	457.4	0.30×10^9	8

Note: concentration of each of the derivatives was approx. 1 mM, and 5mL of the solution was injected onto the column.

The differences in response may suggest that, although the ionization of the Boc-protected derivatives proceeds via the same mechanism the stability of the target ions used to detect these derivatives may not be the same, and that such potential instability may lead to partial in-source fragmentation with consequent reduction in signal and lower peak area for the Boc-Gly and Boc-Trp derivatives. Although the cause of such instability is not readily apparent, it may be related to differences in the mobile phase composition during the gradient elution, inherent structural composition/chemistry or lower solubility among the derivatives. The mobile phase composition at the retention time of Boc-Trp derivative contains a much higher proportion of isopropyl alcohol in hexane and is much more polar than that at the retention times of the other derivatives in the mixture. If, however, the increased polarity of the mobile phase, and mainly the higher content of isopropanol, was responsible for greater instability of the tryptophan derivative during ionisation, it could not explain the lower peak area of the glycine derivative, which elutes in the immediate proximity of both the proline and the phenylalanine derivatives. This may in turn suggest that the instability of the target ions of Gly and Trp derivatives is related to the amino acid structure and inability to retain the charge on the target ion formed during ionization, with subsequent fragmentation leading to formation of other secondary fragment ions that were not detected due to their low masses.

Alternatively, the reason for the smaller peak areas recorded for those two derivatives may be the result of their lower solubility in the sample solution (hexane - isopropanol, 9:1, v/v) compared with the proline and phenylalanine derivatives, ultimately resulting in fewer

analyte molecules being ionized and detected in MS. It is difficult to answer those questions definitively at this stage, and perhaps the comparison with the results obtained during RP-HPLC-APCI-MS can offer a means to explain the observed effects when normal phase conditions are applied.

The comparison of the signal intensity of the native lipid and its *N*-Boc protected amino acids derivatives, to assess a potential signal intensity improvement, proved unsuccessful mainly due to extremely poor ionization of the 1-octadecanol resulting in it not being detected in the APCI-MS mode. Nonetheless, it was demonstrated that the derivatization method is capable of circumventing this issue and extends the spectrum of analytical detection techniques for such classes of analytes beyond the usual practice of derivatization and subsequent gas chromatographic analysis.

3.3.4 Reversed phase high performance liquid chromatography mass spectrometry.

Evaluation of the chromatographic and mass spectrometric behaviour in reversed phase chromatography mass spectrometry (RP-HPLC-MS) was attempted to establish a separation of the derivatives of the model lipid compound. In particular, two factors were examined: comparison of the effect of the change in analyte polarity on retention on a standard C18 and the effect of the more polar mobile phase on ionization efficiency in MS, and ultimately on signal intensity. A reversed phase gradient method was developed in order to achieve chromatographic resolution of the *N*-Boc amino acid derivatives of 1-octadecanol.

Due to its high eluotropic strength the solvent used to solubilize samples for NP-HPLC analysis could not be used in RP-HPLC without risking the complete loss of retention of the analytes on the C18 column. Owing to the limited solubility of the derivatives in the solvent of first choice, pure methanol, combinations of methanol and other less polar solvents were investigated. Among the solvents evaluated were isopropanol, ethyl acetate, dichloromethane and chloroform. The solubilizing properties of the solvent compositions were tested by attempting to dissolve 1 mg of the derivatives in 1 mL of solvent. Due to its potential, as a protic solvent, to enhance the analyte protonation during the MS analysis the target was to find a combination in which methanol dominated the mix. Isopropanol was dismissed early as it failed to solubilize the derivatives even at a 1:1 ratio with methanol. By contrast, 20 % dichloromethane in methanol readily solubilized the derivatives, though its low boiling point might seem disadvantageous owing to the possibility that evaporation,

even at room temperature, could lead to changes in the solution concentration. Notably, this could cause a problem for repeat injections or when the sample is stored in solution for extended periods. Finally, it was determined that a combination of methanol and ethyl acetate or chloroform both offered very good solubilizing properties, the latter met the criteria at just 10 % ratio to methanol, while the former required 20 % - 25 % ratio. Hence, chloroform – methanol (1:9, v/v) solution was selected as the sample solvent.

Separation of the derivatives was initially tested using a mobile phase comprising methanol and a low proportion of water (5 %). The sample solvent proved to be too strong for this mobile phase composition, causing the analytes to elute prematurely as a single distorted peak. A change in the mobile phase composition to reflected the eluotropic strength of the sample solution was considered to circumvent this problem. Substituting part of the methanol component with isopropanol effected elution, though with very limited resolution and pronounced tailing of the peaks (results not shown). This was attributed to similarities in the properties of the two solvents not allowing adequate partitioning of the analytes between the mobile and stationary phases.

Ethyl acetate was next considered as the strong eluent of the mobile phase and methanol-water as the weak eluent: the hydrophobic analytes would be expected to partition into the stationary phase and only with increasing ethyl acetate would their partitioning between the mobile phase and the stationary phase occur, resulting in more discrete and orderly elution. Although this choice of solvents was not unprecedented, only a few reports of using such a mobile phase composition in RP chromatography could be found. The eluent composition is also compatible with the APCI mode of ionization.

The initial separation of a mixture of the derivatives of 1-octadecanol with Boc-Gly, Boc-Phe, Boc-Pro and Boc-Trp was tested on a Dionex Acclaim C18 column (5.0 mm × 2.1 mm, 2.2 μm) maintained at 30 °C and the mobile phase comprising methanol, ethyl acetate and water at proportions of 80/10/10, delivered at 0.4 mL min^{-1}.

Under isocratic conditions, the derivatives were resolved very well and showed only moderate peak tailing, though in a total elution time of over 30 min. This was deemed too long so a simple gradient program was devised to shorten the analysis time. The gradient program comprised a short isocratic period at initial conditions (MeOH/H_2O/ethyl acetate, 80/10/10) to compensate for any void volume effects, 5.5 min linear gradient to increase the ratio of ethyl acetate to 23 % while maintaining the aqueous proportion constant, followed by a short isocratic hold (1 min), and a return to the initial conditions (1 min). The system

was equilibrated for approx. 6 min before the next injection. Under such conditions, the backpressure remained below the limit of a conventional HPLC system. Notably, it was observed that addition of acetic acid to water (0.5 %, v/v) slightly improved signal intensity, hence the addition of acetic acid was incorporated in the final method.

The derivatives chromatographed well, with baseline or near baseline resolution between the analyte peaks (**Figure 3-12**). The derivatives show excellent peak geometries and very little tailing. It should be noted that although the achieved separation was deemed satisfactory, it should be possible to achieve full base line separation by modification of mobile phase gradient program, though with increased analysis time.

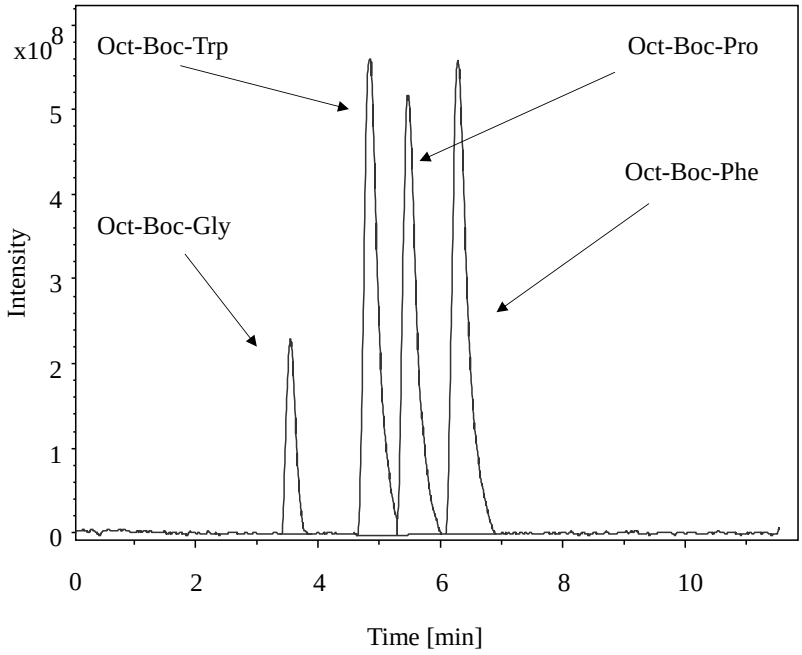


Figure 3-12 (+) APCI-MS base peak mass chromatogram of sample containing 1-octadecyl esters of Boc-Gly, Boc-Phe, Boc-Pro and Boc-Trp at 1 mM concentration. Separation was achieved using a Dionex Acclaim RSLC 120 C18 column (2.1 mm × 100 mm, 2.2 μm) under reversed phase mode.

The control of the retention time under RP was considered to be much better than under NP mode since the peaks eluted in a shorter time with well defined peak geometries and near baseline resolution. In contrast, the elution of the most retained derivative in NP mode, tryptophan, was effected by applying a gentle gradient program and required almost 30 min. The most important advantage of the RP method is that other highly polar potential constituents (e.g. impurities or components of real samples) are not retained in the RP mode and that the return and re-equilibration to the starting mobile phase composition can be

achieved by flushing the column with 10 column volumes. By contrast, in NP mode the column requires so called "back flushing" of the highly retained polar analytes/impurities at the end of each run, which is then followed by a lengthy re-equilibration step.

The peak areas corresponding to the target *m/z* of each of the individual derivatives were calculated (**Table 3-5**). The largest peak areas were generated by the Boc-Pro-Oct (**11**) and Boc-Phe-Oct (**7**) derivatives whereas the peak areas for Boc-Trp-Oct (**12**) and Boc-Gly-Oct (**10**) derivatives were around a factor of two less.

Table 3-5 Retention times and peak areas of *N*-Boc-protected amino acid derivatives of 1-octadecanol chromatographed using RP-HPLC-APCI-MS. Peak areas are calculated based on EIC for the target ions. Area % of each peak in the total chromatogram was also calculated.

Derivative	**t_R [min]**	**Target ion *m/z***	**Peak area**	**Relative Area (%)**
Boc-Gly-Oct (**10**)	7.2	418.4	2.5×10^9	25
Boc-Trp-Oct (**12**)	9.0	328.3	9.0×10^9	92
Boc-Pro–Oct (**11**)	9.8	368.3	8.4×10^9	86
Boc-Phe–Oct (**7**)	10.6	457.4	9.9×10^9	100

The recorded peak areas of the derivatives during the RP-LC-MS separation are comparable to those observed under normal phase conditions for Boc-Pro and Boc-Phe derivatives, whereas they differ considerably for Boc-Gly and Boc-Trp derivatives of 1-octadecanol. The largest difference in peak area percent between the two modes was observed for the Boc-Trp derivative: 8 % under NP and 92 % under RP chromatography (normalized values). The reasons behind such change is not known. It was previously suggested that lower solubility of the tryptophan and glycine derivatives could be responsible for fewer analyte molecules being available for ionization and causing lower signal intensity. If this were so, then it would be expected that the signal of the Boc-Gly derivative would be comparable to that of Boc-Phe and Boc-Pro., which was not observed. It is plausible, therefore that the observed increase in the relative peak area of the Boc-Trp derivative is associated with its ionization efficiency and charge stability, which may be solvent dependent. Similarly, the peak area for

the Boc-Gly derivative is approximately two times higher under NP than under RP conditions, which may be caused by cleavage of the ester bond resulting in a lower ion current focused into the target ion at *m/z* 328.3. This, however, was not investigated further.

Notably, although, the absolute value of the peak area recorded for each of the derivatives in normal phase was approximately twice that of peak area recorded under reversed phase the difference was not considered sufficiently beneficial to consider the NP chromatography mode to be superior to the RP mode.

3.3.5 Tandem mass spectrometric characterization of the *N*-Boc amino acid derivatives of 1-octadecanol

Ion trap tandem mass spectrometry results were obtained during direct infusion of the individual derivatives using APCI-MS and selection of the most intense ions for CID.

The $[M+H-100]^+$ ion of the Boc-Gly-Oct (**10**) derivative was isolated in the ion trap and subjected to resonance induced dissociation to yield the MS^2 spectrum. The ion current intensity loss, in the MS^2 spectrum of the ion at *m/z* 328.3 was around three orders of magnitude with no significant product ions being observed (**Figure 3-13**A–B). Dissociation of the derivative *via* ester bond cleavage would yield the protonated amino acid with the alkyl moiety of the alcohol being lost as a neutral species. The low *m/z* value of the potential product ion formed by this dissociation pathway would not have been captured in the MS^2 spectrum under the scan conditions employed, the mass of the ion being below the low mass cut-off (LMCO) of the ion trap. The LMCO is one of the drawbacks of using ion trap instruments, resulting in ions below a certain *m/z* value not being detected on account of their instability in the ion trap at the set q_z value typically used in the process. As such it is not known whether any of the precursor ion current produces ions with *m/z* values below the LMCO during CID. The cut-off value is usually at 1/4 of the *m/z* of the precursor. In this experiment the LMCO would be 82 Da, thus preventing the detection of the protonated molecule of glycine at *m/z* 76.3.

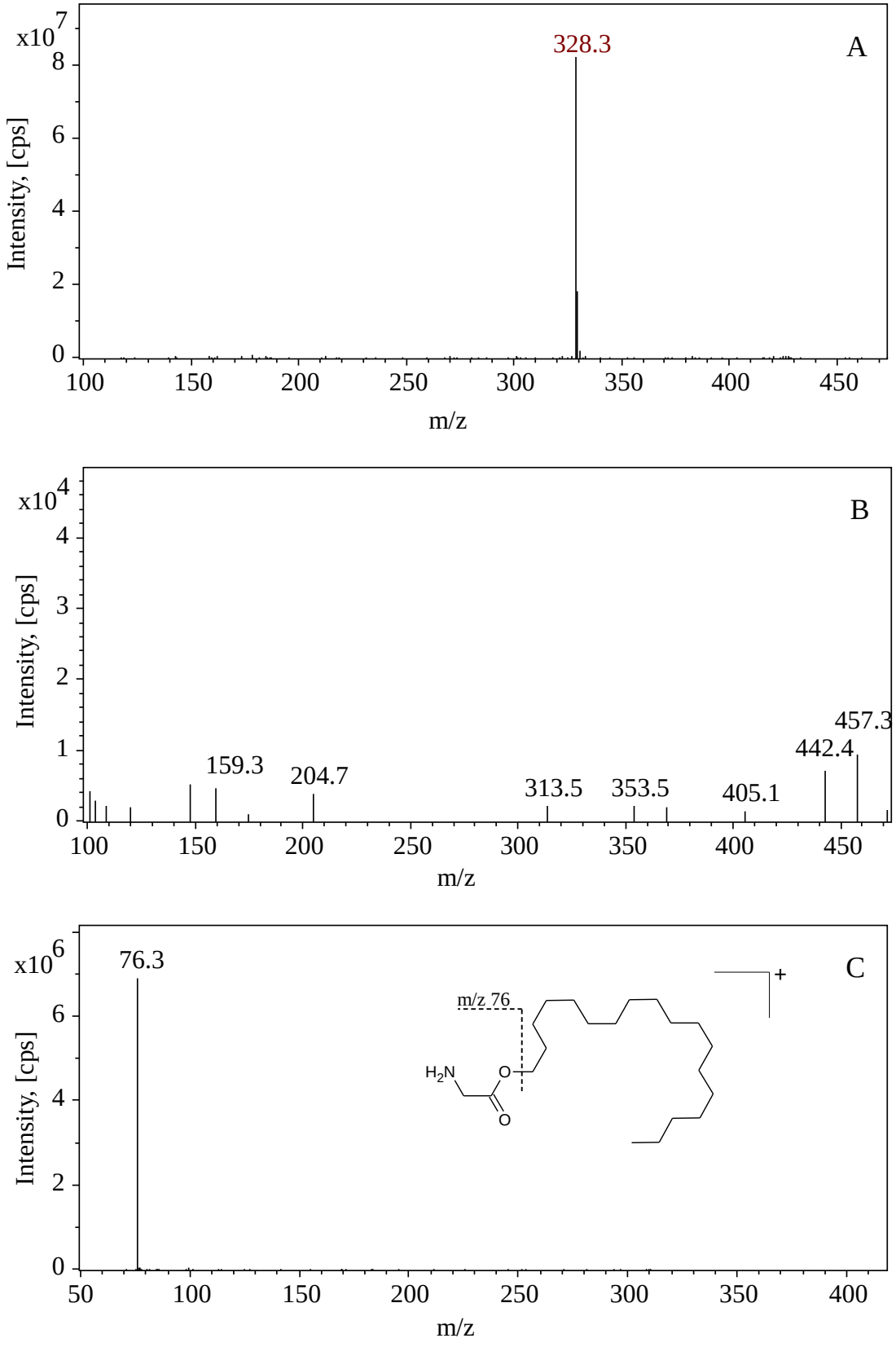


Figure 3-13 A) full mass spectrum B) MS^2 spectrum of the peak at *m/z* 328.3, C) PAN MS^2 spectrum of the peak at *m/z* 328.3 obtained during direct infusion of the Boc-Gly-Oct (**10**) derivative.

In order to circumvent this issue, the HCT ion trap (Bruker) is equipped with a feature called Panoramic (PAN) MS/MS, which can automatically set the *m/z* values of the precursor ion and of the smallest product ion to be captured, enabling CID to produce the most information rich spectrum. This allows for product ions as low as *m/z* 60 to be observed from precursors with *m/z* above 1000.

On activation of the PAN feature an intense product ion was detected at *m/z* 76.3 (**Figure 3-13**C), corresponding to the protonated molecule of glycine and confirming that the dissociation of the Boc-Gly-Oct (**10**) derivative proceeds *via* ester bond cleavage to yield the protonated amino acid, as had been suspected. The cleavage of the ester bond most likely proceeds mainly *via* a double hydrogen transfer (DHT). The hydrogen of the adjoining carbon migrating to the oxygen bridge followed by hydrogen transfer from a γ-carbon in a McLafferty style rearrangement. The intensity of this ion was only an order of magnitude lower than the precursor ion (*m/z* 328.3) selected for CID, suggesting that the efficiency of its formation is very high, thus allowing the ion current to be focused on this ion. This, however, indicates that the main dissociation pathway in tandem MS proceeds through the braking of the ester bond and that the charge is retained on the amino acid moiety. The alcohol (hydroxyl lipid) moiety is lost as a neutral fragment preventing any information about its structure to be gleaned from the tandem MS experiments. The MS^3 spectrum could not be recorded due to the product ions formed having a lower mass than was allowed to be observed even with use of the PAN-MS feature.

The ion at *m/z* 418.4 isolated during infusion of the Boc-Phe-Oct (**7**) derivative was subjected to resonance induced dissociation, producing an MS^2 spectrum that revealed only two prominent ions at *m/z* 166.3 and *m/z* 120.5 (**Figure 3-14**). The two product ions were assigned as the protonated molecule of phenylalanine and its immonium ion, respectively. Thus, the dissociation pathway is similar to that observed during tandem MS of Boc-Gly-Oct (**10**), indicating that the alcohol moiety was lost as a neutral molecule, hence no information about its structure could be inferred. The ester cleavage leading to the formation of the protonated amino acid molecule can be rationalised as following the same DHT mechanism observed for Boc-Gly-Oct (**10**). The immonium ion formation proceeds through the expulsion of CH_2O_2 (46 Da) as a neutral molecule, consistent with the mechanism described previously (Tsang and Harrison, 1976; Papayannopoulos, 1995; Ambihapathy *et al.*, 1997) to give the protonated immonium product ion (*m/z* 120).

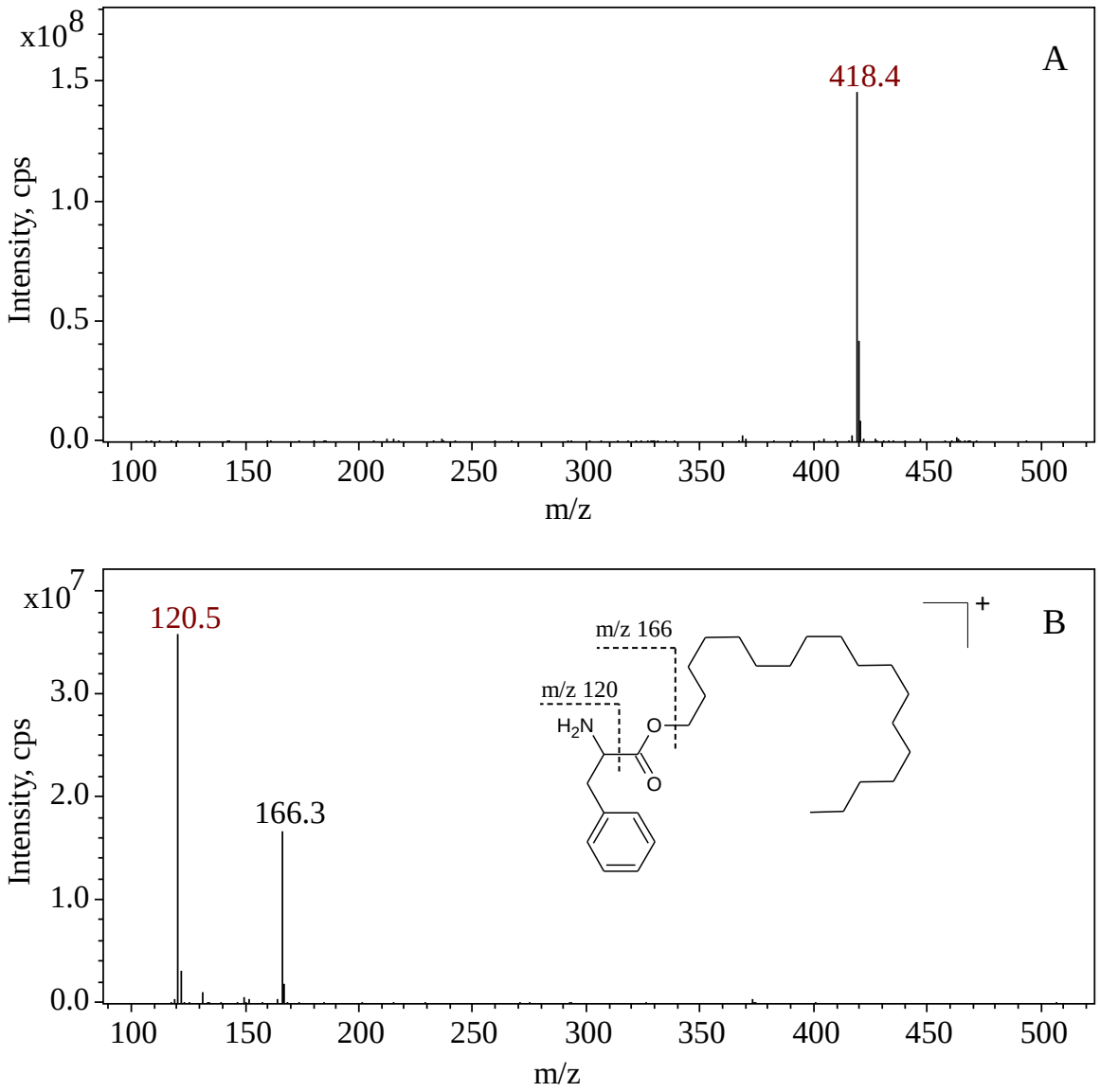


Figure 3-14 A) full mass spectrum B) MS^2 spectrum of the peak at m/z 418.4, obtained during direct infusion of the Boc-Phe-Oct (7) derivative.

The MS^2 spectrum of the base peak selected for dissociation during infusion of the of the Boc-Pro-Oct (**11**) derivative shows only one major ion, formed in the same manner as observed for the derivatives discussed above. The ester bond cleavage gave rise to a highly intense product ion at *m/z* 116.5 (**Figure 3-15**). Notably, however, subsequent CID did not yield any recognizable product ions, most likely due to proline dissociation to form small product ions that were not captured due to the LMCO. No product ions pertaining to the alcohol moiety were observed, prohibiting any structural information of the lipid to be learnt from the tandem MS studies.

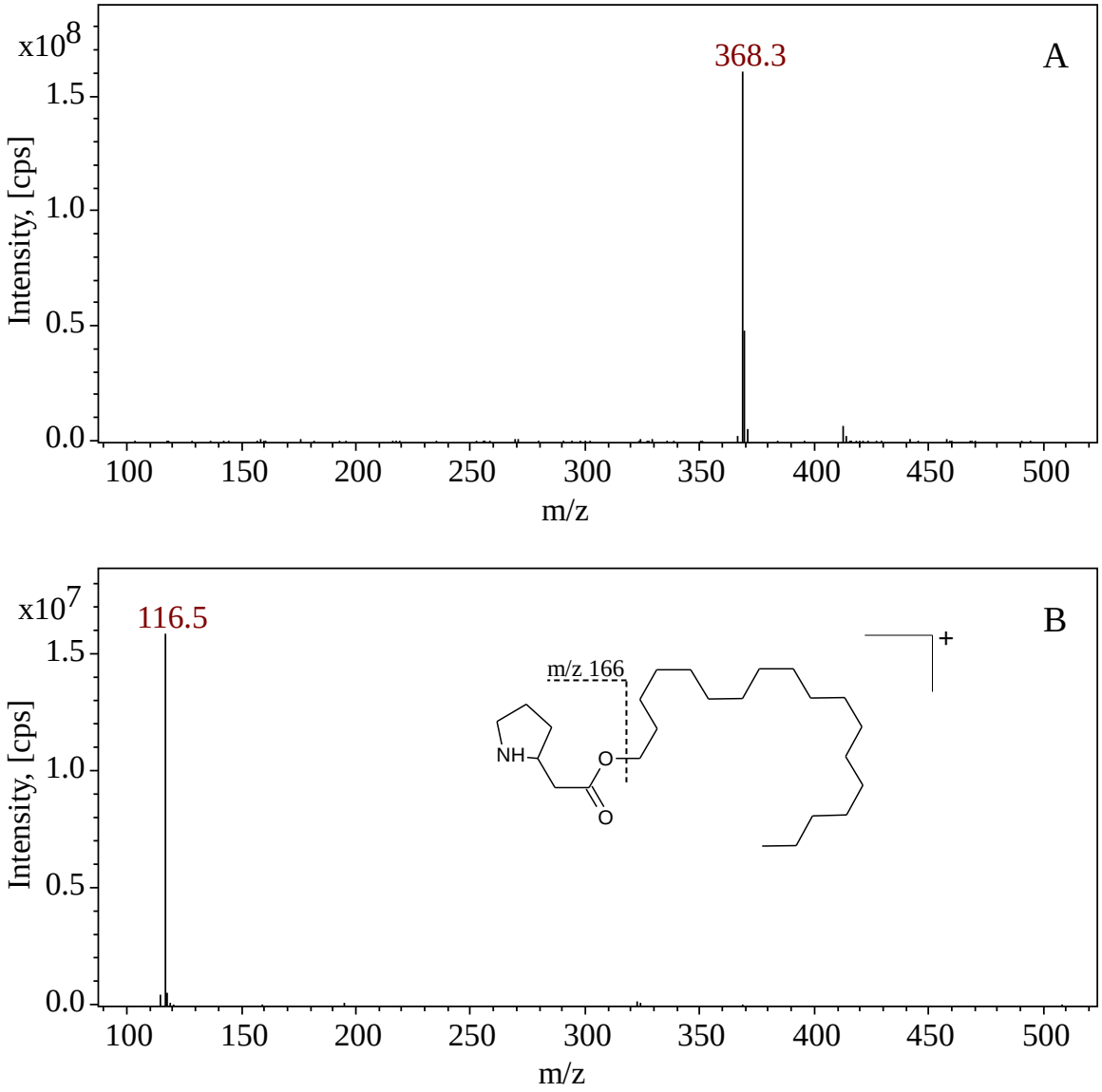


Figure 3-15 A) full mass spectrum B) MS^2 spectrum of the peak at *m/z* 368.3, obtained during direct infusion of the Boc-Pro-Oct (11) derivative.

Finally, to investigate the resonance induced dissociation of Boc-Trp-Oct (**12**) the base peak at *m/z* 457.7 in the full MS spectrum was selected (**Figure 3-16**A). This ion was formed *via* the same McLafferty type rearrangement as the base peak ions observed in the full MS spectra of all other *N*-Boc amino acid derivatives investigated, by a loss of the Boc moiety (-100 Da). The MS^2 spectrum shows an intense base peak at *m/z* 440.4 (**Figure 3-16**B). The formation of this ion results from the loss of the amine group (-17 Da), a characteristic feature of the tryptophan derivatives (Huang *et al.*, 1993). Notably, the subsequent dissociation of *m/z* 440.4 revealed a much richer MS^3 product ion spectrum indicating more complex and diverse dissociation behaviour (**Figure 3-16**C). The product ions observed in MS^3 are: *m/z* 412.3, *m/z* 380.4, *m/z* 190.2, *m/z* 144.3, *m/z* 130.5. The ion *m/z* 412.3 most

likely forms by the loss of ethane (28 Da) as a neutral fragment from the terminus of the alkyl moiety of the alcohol. Formation of *m/z* 190.2 represents a loss of 250 Da, attributed to simultaneous loss of ammonia and the alkyl group of the lipid moiety, both as neutral fragments. Although the loss of ammonia is not very common during the tandem MS of amino acids it has been observed in MS/MS spectra of tryptophan derivatives (Zhang, Fang and Smagin, 2014). Concurrent losses of the amine group, alkyl moiety of the alcohol, and CH_2O_2 (carboxyl), resulting in a total loss of 296 Da, gave rise to a prominent product ion at *m/z* 144.3. The product ion at *m/z* 130.5 represents a total loss of 310 Da from *m/z* 440.4 giving rise to protonated 3-methylidene-3*H*-indole. The formation of this ion was reported previously for *N*-ethoxycarbonyl ethyl esters of amino acids (Huang *et al.*, 1993; Zhang, Fang and Smagin, 2014).

The product ion at *m/z* 380.4 in the MS^3 spectrum of Boc-Trp-Oct (**12**) was not assigned.

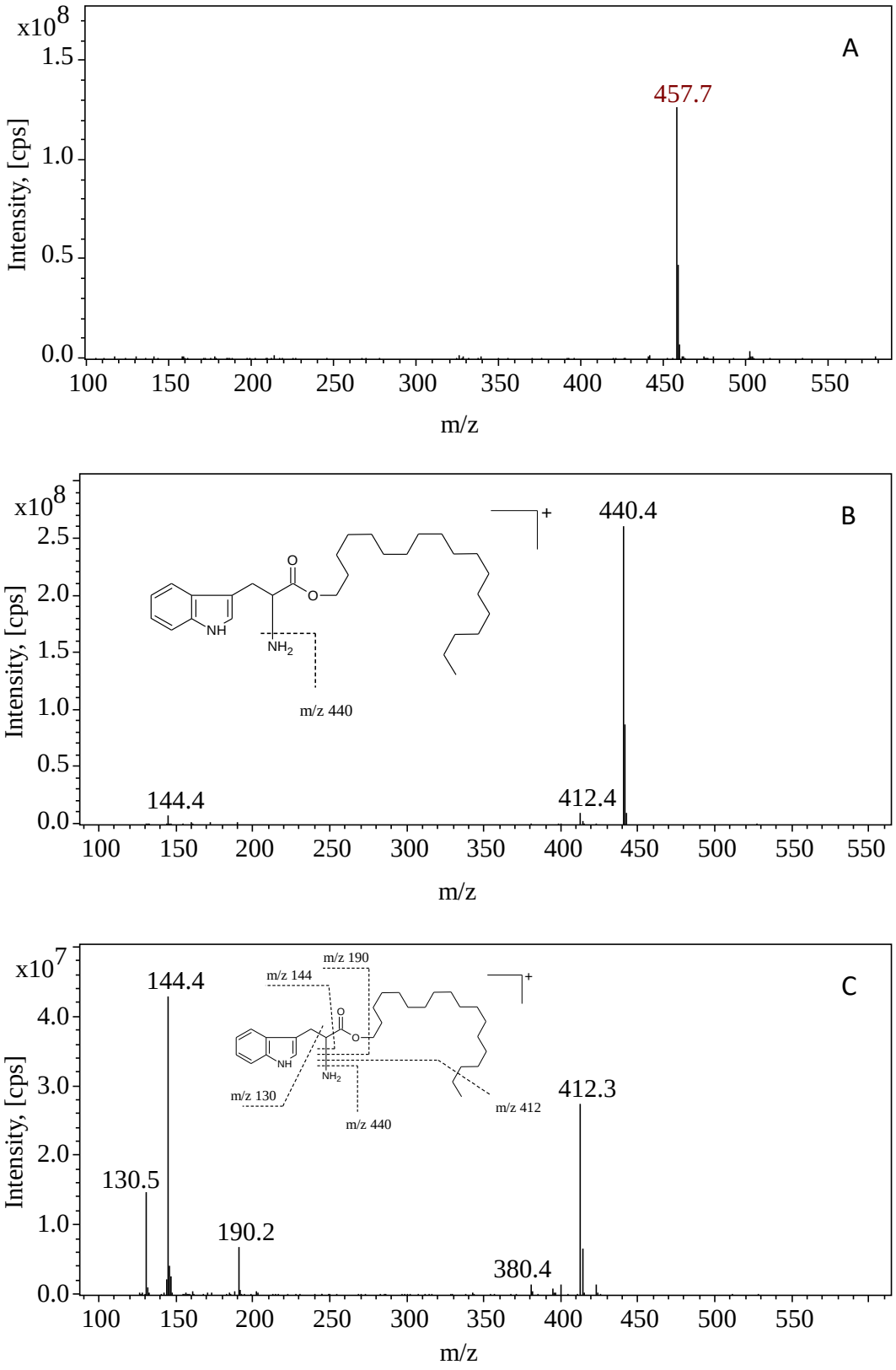


Figure 3-16 A) full mass spectrum B) MS^2 spectrum of the base peak at *m/z* 457.7, C) MS^3 spectrum of the base peak *m/z* 440.4 obtained during direct infusion of the Boc-Trp-Oct (**12**) derivative.

3.3.6 Conclusions

Application of the Steglich esterification reaction was successful in the preparation of *N*-Boc amino acids derivatives of 1-octadecanol, demonstrating the potential viability of this approach to chemical modification of hydroxylated lipids. The original reaction conditions were modified in order to improve the yield and minimize the reaction time. This was achieved by application of acoustic irradiation, by means of a standard sonic cleaning bath, and resulted in highly reproducible esterification reactions with excellent yields: 95 % for Boc-Gly, 98 % for Boc-Phe, 93 % for Boc-Trp, and moderately lower 86 % for Boc-Pro derivatives. Notably the use of the sonic bath allowed the reaction time to be reduced from several hours, including initial cooling on an ice bath, to just 2 h with no control of the temperature.

Further modifications and improvements, such as the change of the carbodiimide from DCC to EDC, enabled solubility of the carbodiimide in a broader range of solvents (including water) to be exploited and substitution of the commonly applied aqueous work-up with a direct silica gel column isolation of the derivative from the crude reaction, which is not only straightforward but also time saving.

The Boc-amino acid derivatives were chromatographed using HPLC-APCI-MS in both normal phase and reversed phase modes in order to compare their behaviour. Although separation of the derivatives was possible under NP chromatography mode, a mobile phase composition based on that in widespread use in the separation of archaeal tetraether lipid cores, was not deemed suitable. This was due to the poor resolution of the early and closely eluting derivatives of Boc-Phe, Boc-Gly, Boc-Pro (first 5 min) and the late eluting Boc-Trp with retention time of approx. 30 min.

An alternative RP method using methanol, ethyl acetate and aqueous acetic acid (0.5%, v/v) was developed and shown to offer improved, near baseline resolution of the derivatives, most notably within a short retention time window of less than 8 min. Importantly, the avoidance of the need to back flush the column to remove the polar analytes/impurities that accumulate on the normal phase column during routine analysis (Schouten *et al.*, 2000) is circumvented under the reversed phase mode due the polar components eluting early in the analytical run.

The mass spectrometric evaluation of the Boc-amino acid derivatives revealed stark disparities in the ionization potentials of the four derivatives analysed in normal phase

chromatography. Boc-Phe and Boc-Pro derivatives produced peaks of nearly the same area, Boc-Gly gave a peak area of about half and Boc-Trp gave a peak area of only 8 percent relative to Boc-Phe. By contrast, the derivatives analysed in reversed phase showed different peak area responses: Boc-Phe had the highest peak area with Boc-Trp at 92 % relative peak area with Boc-Pro and Boc-Gly, at 86 % and 25 % relative peak area, respectively.

Tandem mass spectrometry experiments of the derivatives revealed that CID proceeds *via* cleavage of the ester bond, giving rise to highly intense signals and, in each case, with the charge retained on the amino acid product ion. Notably, the alcohol moiety is lost as a neutral loss, preventing viable structural information about the (hydroxyl) lipid motif of the derivative being gained

Chapter 4 Preparation and characterization of *N*-Fmoc protected amino acid derivatives of 1-octadecanol

4.1 Introduction

The evaluation of the derivatisation of hydroxyl (alcohol) lipids using *N*-Boc protected amino acids in combination with acoustic irradiation shows that such modification can be carried out readily and quickly, as discussed in Chapter 3. Moreover, the high and consistent efficiency in the derivatisation suggests that the protecting group located on the *N*-terminus of the different amino acids examined does not influence the activation and acyl transfer processes though it may have impacts on their reaction rates and solubility in the reaction medium.

Many of the hydroxyl lipids (e.g. alkanols and sterols) that occur in sediments do not possess a chromophore and hence are not detectable under ultraviolet (UV) monitoring, rendering such mode of detection inapplicable. Although mass spectrometric detection is commonly used for identification of lipids, their quantification remains problematic owing to poor and variable ionization efficiencies. This is especially challenging when the ionization efficiency of an analyte is low, requiring a much larger sample size in order to enable detection. Similarly, when analytes in a sample examined by HPLC exhibit substantial differences in ionization efficiency the detection of the more weakly ionizing components can be obscured by saturation of the detector signal.

The fluorenylmethoxycarbonyl group (Fmoc), one of several *N*-protecting functionalities applicable to amino acids, confers excellent chromophore properties to the amino acid. A large number of amino acids protected with the Fmoc group are commercially available. The chemical hydrolysis of the Fmoc group is orthogonal to that of the Boc group, meaning that it is stable under acidic conditions whereas it can readily be removed using a base (e.g. pyrimidine). The chromophore of the Fmoc group exhibits strong UV absorption (λ_{max} 263 nm; ε 18950; Chang *et al.*, 2009), thus using *N*-Fmoc protected amino acids as derivatisation agents would allow detection of hydroxyl-containing lipids under UV or fluorescence detection modes as alternatives to MS detection. Thus, the development of a successful and high yielding derivatisation protocol will extend the range of available amino acid derivatisation agents beyond the *N*-Boc protected species discussed in Chapter 2.

The effective resolution of the *N*-Boc amino acid derivatives in reversed phase chromatography mode (discussed in Chapter 3) prompts the question of whether the more apolar Fmoc bearing derivatives could be resolved in a similar manner.

The *N*-Fmoc protected amino acids selected for evaluation in derivatisation of 1-octadecanol were: glycine (Fmoc-Gly, **13**), phenylalanine (Fmoc-Phe, **14)**, proline (Fmoc-Pro, **15)** and tryptophan (Fmoc-Trp, **16)**, so that the coupling efficiency and the properties of the derivatives they formed could be directly compared to *N*-Boc protected amino acids derivatives prepared previously.

4.2 Aims

Three main objectives were defined for the research discussed in this chapter:

- to evaluate the preparation and yields of the *N*-Fmoc amino acid derivatives of hydroxyl lipids using 1-octadecanol as a model lipid;
- to characterize the mass spectrometric, ultraviolet and chromatographic behaviours of the derivatives;
- to compare the properties of the orthogonally protected (Boc and Fmoc) amino acid derivatives of the model lipid.

4.3 Preparation of the derivatives

The derivatisation reactions between the model lipid and the selected *N*-Fmoc protected amino acids were carried out according to Method C (see Chapter 2.2.3). The ester derivatives were isolated, weighed and the calculated yields are reported in **Table 4-1**.

Table 4-1 Results of derivatisation of 1-octadecanol with selected *N*-Fmoc amino acids using EDC/DMAP coupling under sonic irradiation in accordance with Method C. The reactions were carried out in duplicate. Average yields and ranges are given.

	Yield [%]			
Derivative	1	2	Average	Range
Fmoc-Phe-Oct (18)	94.4	93.2	93.8	94.4 – 93.2
Fmoc-Gly-Oct (17)	96.1	95.6	95.9	96.1 – 95.6
Fmoc-Pro-Oct (19)	94.8	94.3	94.6	94.8 – 94.3
Fmoc-Trp-Oct (20)	96.6	96.8	96.7	96.6 – 96.8

The data in **Table 4-1** shows that excellent yields can be achieved when using the *N*-Fmoc amino acids as derivatisation agents under the conditions described in Method C. The reactions were carried out in duplicate and the very low RSD values for the replicates indicate high reproducibility of the derivatisation protocol.

The highest average yields of 96.7 % and 95.9 % were obtained for Fmoc-Trp (**16**) and Fmoc-Gly (**13**), respectively. The results for Fmoc-Pro (**15**) and Fmoc-Phe (**14**) derivatives were comparable, at 94.6 % and 93.8 %, respectively. Although the results for this pair are marginally lower than for the highest yielding Fmoc-amino acid pair, their yields in the sonochemical derivatisation protocol demonstrate the suitability of the method with these two amino acids and their excellent performance during the conversion. The narrow ranges of the yields of the esters of the *N*-Fmoc amino acids used suggest that the derivatisation procedure can routinely yield highly reproducible results.

Comparison of the average percentage yields of 1-octadecanol esters obtained using the orthogonal derivatisation with the *N*-Boc or the *N*-Fmoc protected amino acids shows that the yields for the latter are higher than the former, except for glycine. The phenylalanine

derivatisation of the model alcohol gave a higher yield for the *N*-Boc ester than for the *N*-Fmoc protected ester, 98 % *versus* approx. 94 %. By contrast, *N*-Fmoc protected esters of proline and tryptophan were obtained in higher yields than when the amino acids were used as *N*-Boc protected species: 94.6 % vs 86.5 % and 96.7 % vs 92.6 %, respectively. Finally, 1-octadecyl glycine esters were obtained in almost identical yields for both *N*-Fmoc and *N*-Boc amino acid: 95.9 % and 95.4 %, respectively.

It is somewhat surprising that the relatively small amino acid, glycine, performed essentially the same with regard to the *N*-protecting group despite the Fmoc group being much larger and having greater apolar character that the Boc group. Moreover, if the size of the amino acid side chain had any effect on reactivity it would be expected that glycine would outperform tryptophan or *vice versa*, depending on the nature of the influence of size. Additionally, the relatively low yield of the Boc-Pro ester was presumed to be attributed to the generation of strain due to the proline ring being in proximity to the Boc group, possibly hindering access to the carbonyl group during the acyl transfer step. Notably, however, the Fmoc-Pro yield (avg. 94.6 %) implies that this may not have been the case. In the case of phenylalanine, though, the Boc-Phe ester was formed in much higher yield, suggesting that Fmoc-Phe is less reactive than the Boc-Phe form.

The data suggest that the effectiveness of the *N*-protected amino acids as derivatisation agents in the reactions with the model lipid, 1-octadecanol, does not follow a particular trend and their reactivity appears to be intrinsically associated with the particular amino acid, thus requiring that their performance as such agents be evaluated on case-by-case basis.

4.4 Characterization of the *N*-Fmoc derivatives

4.4.1 Identification of the *N*-Fmoc amino acid derivatives of 1-octadecanol

The mass spectra of the 1-octadecanol esters of the *N*-Fmoc amino acids were acquired during analyte solution infusions into the (+) APCI-MS using a syringe pump. The MS instrument was tuned using solutions of individual derivatives.

The process of ion formation for *N*-Fmoc protected amino acids and their derivatives were studied previously (Du *et al.*, 2003; Zhu *et al.*, 2006), providing insight into their behaviour during mass spectrometric analysis.

The mass spectra of the *N*-Fmoc amino acid derivatives allowed identification and investigation of the ionization mechanism in a similar manner to the *N*-Boc derivatives. By contrast, with the *N*-Boc protected amino acid derivatives, the *N*-Fmoc derivatives showed two prominent ions.

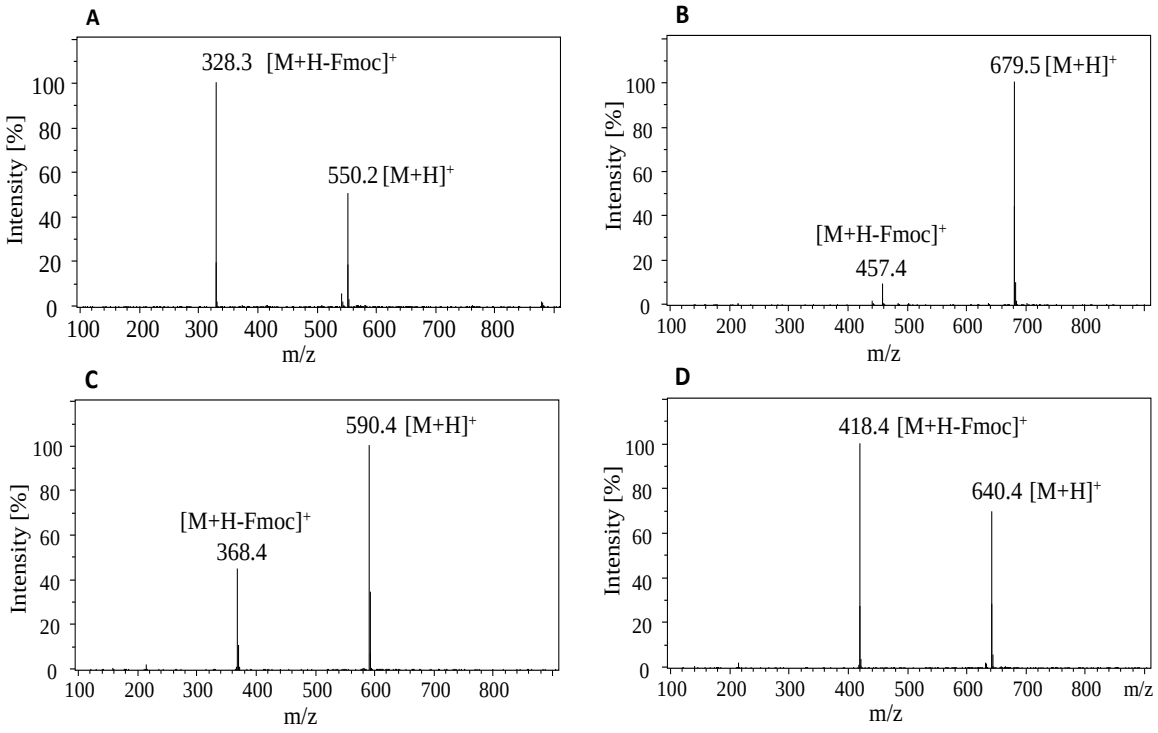


Figure 4-1 (+) APCI-MS mass spectra of the four 1-octadecyl ester derivatives: A) Fmoc-Gly, B) Fmoc-Trp, C) Fmoc-Pro, and D) Fmoc-Phe. All spectra were background corrected.

Of the four spectra examined (**Figure 4-1**A-D) the protonated molecule was the highest intensity signal in MS only for the Fmoc-Trp-Oct at *m/z* 679.5 (**Figure 4-1**B) and for Fmoc-Pro-Oct at *m/z* 590.4 (**Figure 4-1**C). The fragment ion resulting from loss of the Fmoc group (-222 Da) at *m/z* 368.4 is higher (approx. 50 %) for the proline derivative than for the tryptophan derivative (*m/z* 457.4; approx. 10 %), suggesting that the stability of the protonated molecule is much lower for the former than for the latter. For the Fmoc-Gly-Oct derivative (**Figure 4-1**A) the fragment ion resulting from the loss of the Fmoc group (*m/z* 328.3) is base peak as is the case for the Fmoc-Phe-Oct (**Figure 4-1**D) derivative, the ion being at *m/z* 418.4. The protonated molecule at *m/z* 640.4 is higher (approx. 70 %) for the former than for the latter, *m/z* 550.2 at approx. 50 %, indicating that the propensity for cleavage of the Fmoc group is more pronounced for the glycine derivative than for the phenylalanine derivative (**Table 4-2**).

Table 4-2 Ions observed in (+) APCI-MS during driect infusion of the Fmoc amino acid derivatives of 1-octadecanol. Relative abundance provided in parenthesis.

Derivative	**$[M+H]^+$**	**$[M+H-Fmoc]^+$**
Fmoc-Gly-Oct (**17**)	550.2 (52)	328.3 (100)
Fmoc-Trp-Oct (**20**)	679.5 (100)	457.4 (10)
Fmoc-Pro-Oct (**19**)	590.4 (100)	368.4 (50)
Fmoc-Phe-Oct (**18**)	640.4 (70)	418.4 (100)

Loss of the Fmoc group from the $[M+H]^+$ ion of the derivatives results from a McLafferty type rearrangement, where the γ-hydrogen migrates from the fluorenyl moiety to the carbonyl oxygen in the Fmoc group, followed by loss of dibenzofulvene (Fme, 178 Da) and subsequent loss of CO_2 (44 Da) (**Figure 4-2**). This interpretation is in agreement with the behaviour of ionized *N*-Fmoc amino acid derivatives reported previously by Ramesh *et al.* (2010, 2011).

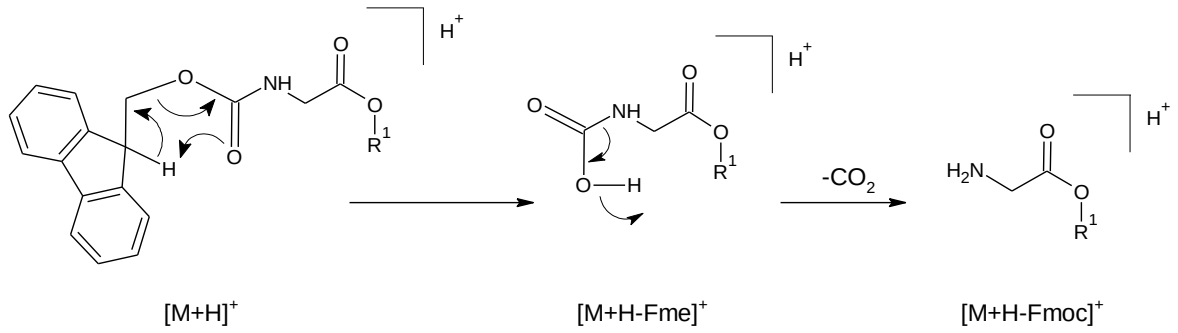


Figure 4-2 Mechanism of the McLafferty type rearrangement resulting in loss of the protecting group in hydroxyl lipid esters of *N*-Fmoc glycine amino acid. Note: R^1 represents the lipid moiety.

Notably, the absence of the ions corresponding to the loss of 178 Da in the mass spectra of all of the derivatives suggests that the rearrangement proceeds directly to produce the fragment ions resulting from the cleavage of the entire Fmoc group (-222 Da).

The cause for the differences in fragmentation is not immediately apparent, though it is likely to be dependent on the character of the amino acid side chain. Glycine, which gave the highest abundance of the fragment ion, lacks a side chain, having only a hydrogen atom at the α–position. On the other hand, tryptophan with its indole side chain, gave the lowest abundance of the fragment ion. Thus, it is possible, that the large indole structure will confer some degree of stabilization of the protonated derivative molecule, perhaps through the nitrogen atom located in the indole ring. The relatively high abundance of the fragment ion in the MS spectrum of the phenylalanine and proline derivatives indicates that, despite being larger than glycine, neither the aromatic (phenyl) nor the cyclic side chain substituent increases the stability of the protonated molecule. The true nature of this interaction is not known and was not investigated further.

Similarly, as was observed in the comparison of the MS behaviour of the *N*-Boc amino acid derivatives under the (+) APCI conditions employed, the *N*-Fmoc amino acid derivatives of the model lipid could not be compared directly with the native alcohol owing to it not ionising to any significant degree.

4.4.2 Chromatographic separation and detection of *N*-Fmoc amino acid derivatives of 1-octadecanol.

The successful separation of *N*-Boc protected amino acid derivatives of 1-octadecanol under RP chromatography prompted application of the same conditions to the *N*-Fmoc amino acid modified 1-octadecanol derivatives. The analysis of individual Fmoc derivatives was performed to enable direct comparison of the retention times of the two orthogonally protected derivative types and to establish the purity of each of the derivatives prepared.

The amino acid derivatives protected with the *N*-Fmoc group are much more apolar than their *N*-Boc alternatives, hence they are much more retained on the C18 column. The increase in the mobile phase eluotropic strength over the gradient program was not sufficient and ultimately failed to elute the derivatives within the elution window of the gradient programme. Although the most polar of the derivatives, Fmoc-Gly and Fmoc-Trp, showed acceptable peak shapes and were base line resolved, the Fmoc-Pro and Fmoc-Phe derivatives eluted during the return to starting conditions, causing significant peak shape distortion and tailing (results not included).

The eluotropic strength of the final gradient composition was increased to reduce the retention times of the derivatives and effect their faster elution. This was achieved by increasing the final content of ethyl acetate from 30 % to 35 % while maintaining the aqueous content unchanged. This change resulted in the desired effect and all of the derivatives were eluted in less than 15 min. The individual derivatives show excellent peak shapes and are well resolved under the chromatographic conditions applied (**Table 4-3**).

In order to enable direct comparison of the peak areas of the individual derivatives their solutions were prepared at a concentration of approx. 1 mM for each component and the mass spectrometer was tuned via infusion using those solutions to maximize their signal response. This was done to ensure that the ion current was focused on the target ions so that it would result in the highest signal intensity that could be recorded.

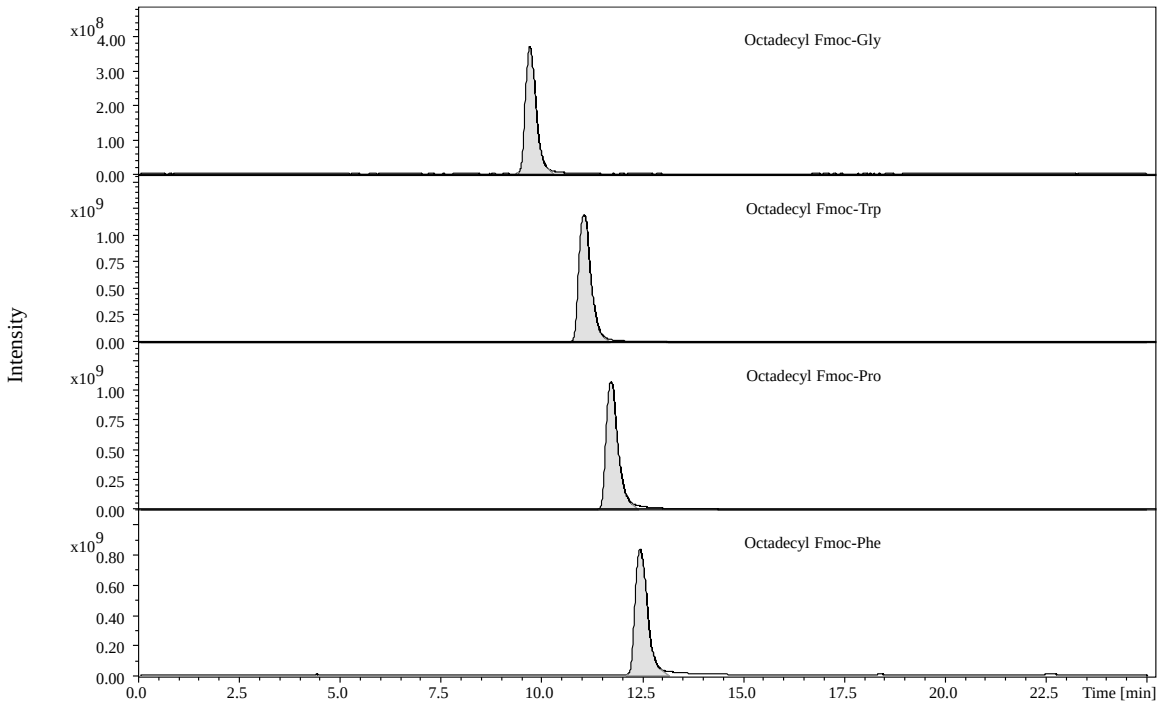


Figure 4-3 The base peak mass chromatograms of the individual *N*-Fmoc derivatives of 1-octadecanol from RP-HPLC-(+)APCI-MS analysis. The concentration of each of the derivatives is approx. 1 mM and the injection volume was 1 μL.

The retention times and measured peak areas of the derivatives are summarized in **Table 4-3** along with the target ion and relative peak area percent. The lipid derivatives elute in the same order as the Boc amino acid derivatives, suggesting that the mechanism governing their retention, analyte size and hydrophobicity, is the same for both types of derivatives. Evaluating the data in **Table 4-3** it is readily apparent that the largest peak was recorded for the Fmoc-Phe derivative. The relative peak areas of the other derivatives were calculated in relation to this value and are summarized in the last column (Area %) of the table. The second largest relative peak area, 57 %, was recorded for the Fmoc-Trp derivative, followed by 45 % for Fmoc-Pro, and 38 % for Fmoc-Gly derivatives. This indicates that the Fmoc-Phe derivative showed the highest ionization efficiency and/or best ion stability among the *N*-Fmoc derivatives tested. It is noteworthy that this derivative is less polar than the earlier eluting derivatives, suggesting that it is not the overall polarity that is responsible for the better ionization efficiency. Although the other derivatives showed peak areas of about half that of the Fmoc-Phe derivative the differences are small given the overwhelming improvement in response compared with the native alkanol. Hence, it should not be considered to be a determining factor during the selection of the derivatising agent.

Table 4-3 Retention times and peak areas of *N*-Fmoc-protected amino acid derivatives of 1-octadecanol chromatographed using RP-HPLC-APCI-MS. Peak areas calculated from the base peak in the range *m/z* 250 - 700. Area % of each peak in the total chromatogram is given.

Derivative	t_R [min]	Peak area	Area %
Fmoc-Gly-Oct (**17**)	9.7	5.0×10^9	38
Fmoc-Trp-Oct (**20**)	11.0	7.4×10^9	57
Fmoc-Pro-Oct (**19**)	11.7	5.9×10^9	45
Fmoc-Phe-Oct (**18**)	12.4	1.3×10^{10}	100

Evaluation of the response for mass spectrometric detection was performed by injecting approx. 10 pmol of each of the derivatives on column during a chromatographic run. All of the derivatives showed very good peaks that can be clearly distinguished rising above the baseline (**Figure 4-4**). The calculated signal-to-noise (*S/N*) values are: 131 for Fmoc-Gly, 136 for Fmoc-Trp, 196 for Fmoc-Pro and 189 for Fmoc-Phe. These values are significantly higher than the widely accepted *S/N* values for calculation of limit of detection (LOD) of 3. Thus, based on the 10 pmol of analyte injection on column and the measured *S/N* values the LOD for the individual derivatives was calculated as 0.2 pmol injected on column.

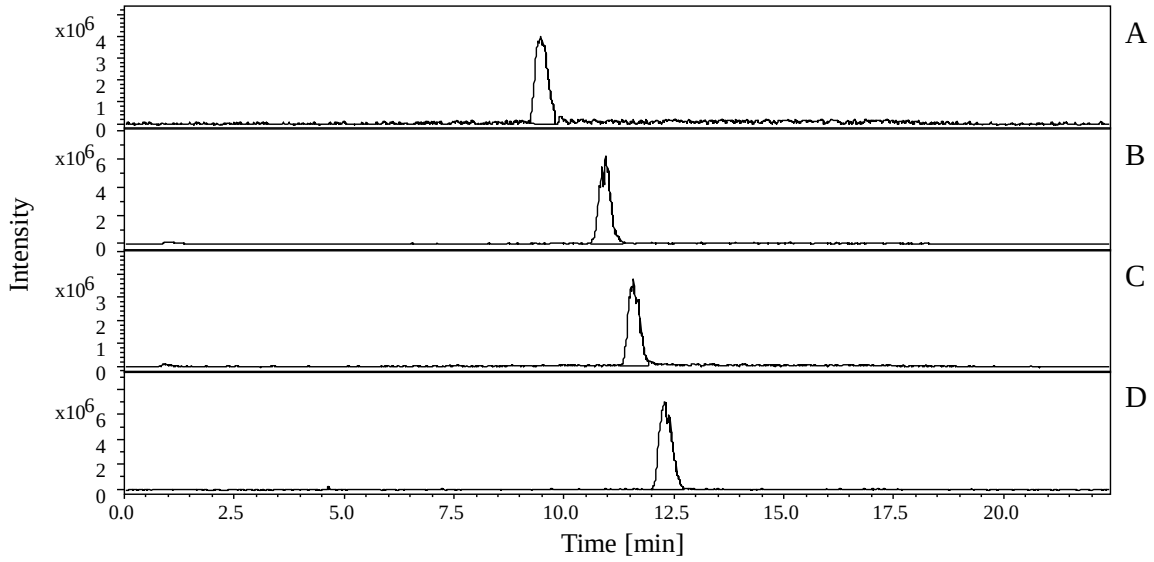


Figure 4-4 Base peak mass chromatograms of Fmoc amino acid modified 1-octadecanol derivatives: A) Fmoc-Gly-Oct, B) Fmoc-Trp-Oct, C) Fmoc-Pro-Oct and D) Fmoc-Phe-Oct.

The introduction of the fluorenylmethyloxycarbonyl (Fmoc) moiety into the derivatives permits their detection using ultraviolet (UV) or photodiode array (PDA) detectors in a wider range of solvents. The common settings of λ_{max} for detection that are reported in the literature are: 263 nm (Melucci *et al.*, 1999; Fabiani *et al.*, 2002; You *et al.*, 2010) and 265 nm (Chen *et al.*, 1999) for single wavelength and 263 nm (16 nm bandwidth) with a reference wavelength at 324 nm (8 nm bandwidth) (Henderson *et al.*, 2000).

Due to the configuration of the available analytical system the UV signal was evaluated using an HPLC-FLD-PDA instrument operated with the settings outlined by Henderson and co-workers (2000) for UV detection. Unlike the single wavelength UV detector, where only one wavelength can be monitored, the PDA allows the signal acquisition to be "bracketed" (bandwidth) around the target λ_{max} as well as signal correction *via* reference wavelength monitoring. This feature compensates for any undesired fluctuations in the acquired signal due, for example, to temperature or mobile phase composition. The fluorescence was monitored with excitation λ_{EX} = 263 and emission λ_{EM} = 309 nm, appropriate to the Fmoc group.

Unfortunately, due to the system developing an internal malfunction of the gradient proportioning system, the chromatography of the derivatives had to be performed under isocratic elution with the mobile phase composition fixed at A:B:C 55:10:35, where A = methanol, B = 0.5 % acetic acid in water, and C = ethyl acetate. Though, this mobile phase composition was similar to the composition at the time of elution of the derivatives, moderately increased peak width and tailing was observed. The elution order was the same as that discussed observed for HPLC-MS separations, with the retention times of the individual derivatives shifted towards lower retention times due to the isocratic elution where the mobile phase was less polar than during the initial stage of the gradient elution program. The UV and FLD chromatograms show the derivatives chromatographed well and have very good peak geometries with only minor tailing. Notably the peak area distribution among the derivatives is much more uniform than that observed in the MS chromatograms. To allow direct comparison of the signal response intensity the same concentration and injection volume as use during the HPLC-MS studies were employed in the HPLC-UV-FLD analysis.

Table 4-4 HPLC-UV-FLD results of the *N*-Fmoc amino acid modified 1-octadecanol derivatives during RP-HPLC separation.

Derivative	t_R [min]	Peak area UV [μAU]	Area %	Peak area FLD [mAU]	Area %
Fmoc-Gly-Oct (**17**)	6.2	8.3×10^5	89	3.1×10^7	92
Fmoc-Trp-Oct (**20**)	7.9	1.3×10^6	138	3.5×10^6	10
Fmoc-Pro-Oct (**19**)	8.7	9.1×10^5	98	3.3×10^7	99
Fmoc-Phe-Oct (**18**)	9.8	9.3×10^5	100	3.3×10^7	100

The highest peak area in UV detection was recorded for the Fmoc-Trp derivative, while lower values were observed for the Fmoc-Gly, Fmoc-Pro and Fmoc-Phe derivatives, with the peak area for the latter being slightly higher than the former two (**Table 4-4**).

The higher absorption for the tryptophan derivative may be explained by the additional absorption associated with the amino acid indole ring. Although the phenyl moiety in the phenylalanine derivative exhibits some chromophore properties its absorption coefficient (λ_{max} 258 nm, ε 188 cm^{-1} M^{-1}) is much smaller than for tryptophan ((λ_{max} 280 nm, ε 5600 cm^{-1} M^{-1}), thus it does not contribute as much to the overall absorption of the Fmoc-Phe derivative (Pace *et al.*, 1995). The two other derivatives, namely Fmoc-Pro and Fmoc-Gly, do not possess any additional chromophore in their structure beside the Fmoc group, which would suggest that any differences in their absorption can be attributed to their individual properties. These observations are in agreement with the literature where Fmoc-Trp had the highest coefficient of absorption and Fmoc-Phe and Fmoc-Trp, although lower than tryptophan, had similar values of the absorption coefficient as reported by Melucci and co-workers (1999) and Fabiani and co-workers (2002). Notably, by contrast with the absorption of the Fmoc-Gly derivative in this study, both authors reported absorption properties of the Fmoc-Gly to be slightly higher than that of Fmoc-Phe. The reason for such behaviour is unknown and may be associated with the difference in spectroscopic behaviour of this derivative under different detector parameters (detection wavelength, bandwidth and reference wavelength), which were different to that used in the literature reports.

The comparison of the peak area data from the fluorescence detector shows that the recorded peak area is virtually identical for all of the derivatives tested, with the exception of the tryptophan derivative for which the observed peak area was approx. one order of magnitude lower. This tendency of the fluorophore labelled tryptophan to show a markedly lower signal response in fluorescence detection is attributed to intramolecular quenching and has been reported previously (You *et al.*, 2007). The quenching of the emission energy of the Fmoc moiety proceeds *via* intramolecular resonance energy transfer to the indole ring of Trp. The requirement for such transfer to occur is that the donor (Fmoc) and the acceptor (Trp) need to be within a distance of less than 10 nm, and that the emission spectrum of the donor overlaps with the absorption spectrum of the acceptor. The maximum absorption for Trp, at 294 nm, indicates a partial overlap with the excitation wavelength employed (309 nm), supporting the suggestion that the indole moiety of tryptophan is responsible for the very low fluorescence signal intensity relative to the other Fmoc amino acid esters.

In order to assess the detection limits of both the ultraviolet and fluorescence detection techniques solutions of the four *N*-Fmoc amino acid derivatives of 1-octadecanol at a concentration of 10 μM were prepared and chromatographed as outlined above, keeping the injection volume fixed at 1 μL. This translates to 10 pmol of the analyte injected on column.

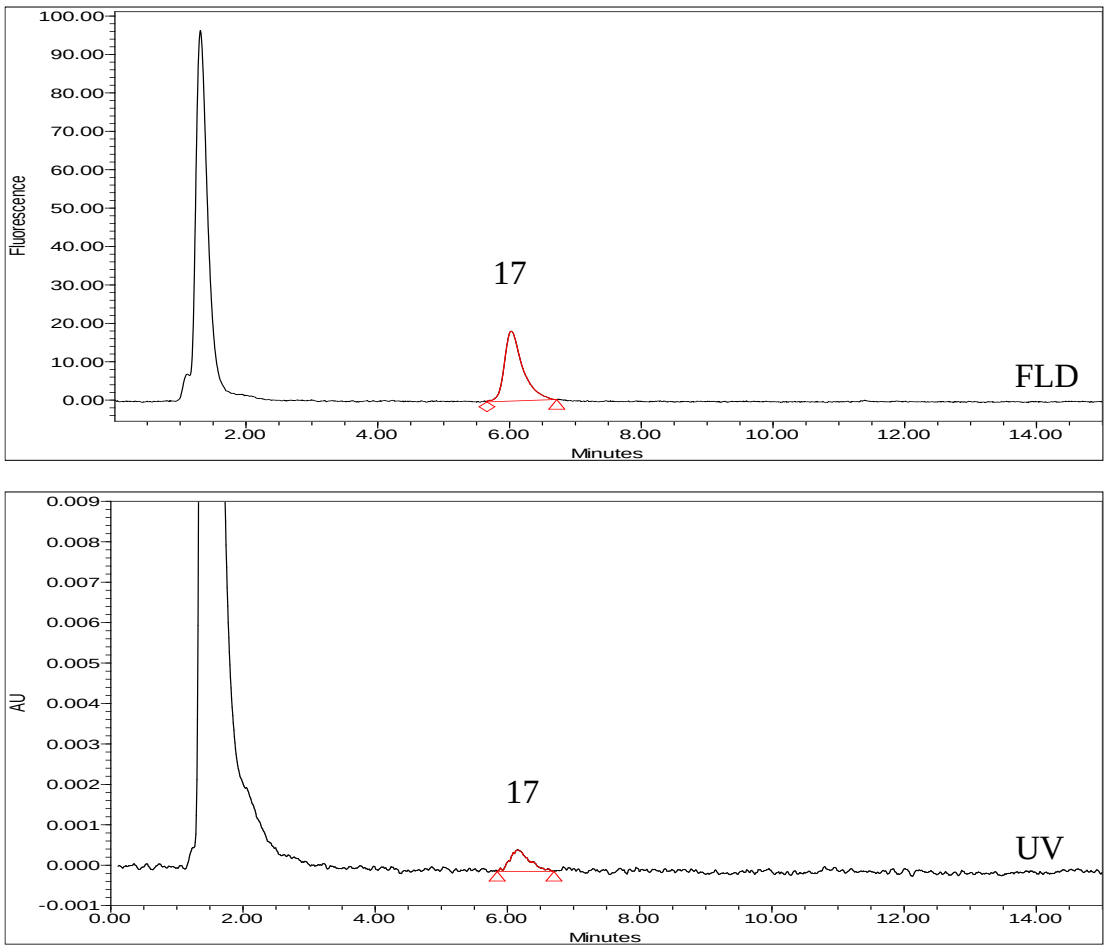


Figure 4-5 Chromatograms of Fmoc-Gly-Oct (**17**) derivative obtained during RP-HPLC analysis: A) fluorescence (λ_{EX} 263 nm, λ_{EM} 309 nm) B) UV detection (263 nm). Amount of the analyte injected on column was 10 pmol.

The analyte peak in the chromatogram in **Figure 4-5**A is clearly visible rising above the baseline making the integration process straightforward, whereas the peak in chromatogram B is very small with the baseline noise becoming a significant factor that may contribute to inaccurate peak integration, ultimately affecting peak area calculation. The calculated signal-to-noise (*S/N*) for the FLD detector was 125, thus, based on the requirement that LOD > 3 × *S/N*, the LOD for this analyte is calculated to be approx. 0.2 pmol. The *S/N* for the UV detector is 4 and as such, being only marginally lower than the LOD requirement, indicates that limit of detection for this technique is approx. 7 pmol.

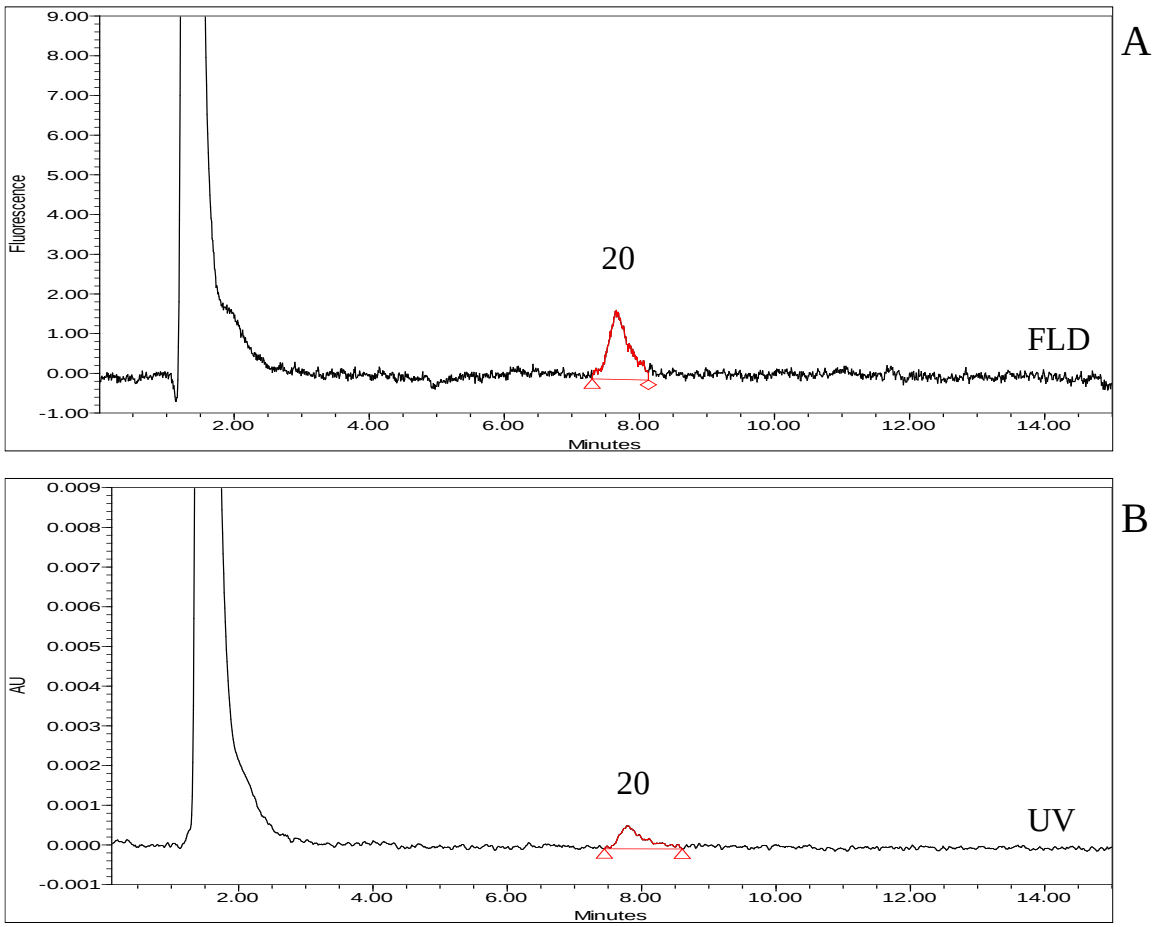


Figure 4-6 Chromatograms of Fmoc-Trp-Oct (**20**) derivative obtained during RP-HPLC analysis: A) fluorescence (λ_{EX} 263 nm, λ_{EM} 309 nm) B) UV detection (263 nm). Amount of the analyte injected on column was 10 pmol.

Due to the intramolecular quenching affecting the fluorescence properties of the tryptophan derivative of 1-octadecanol the signal intensity recorded in FLD is very weak (**Figure 4-6**A) yielding a small peak, which is only slightly higher than that observed in the UV trace **Figure 4-6**B. The lower peak height resulted in a higher LODs value of approx. 2 pmol for the FLD detection and approx. 6 pmol and for the UV mode.

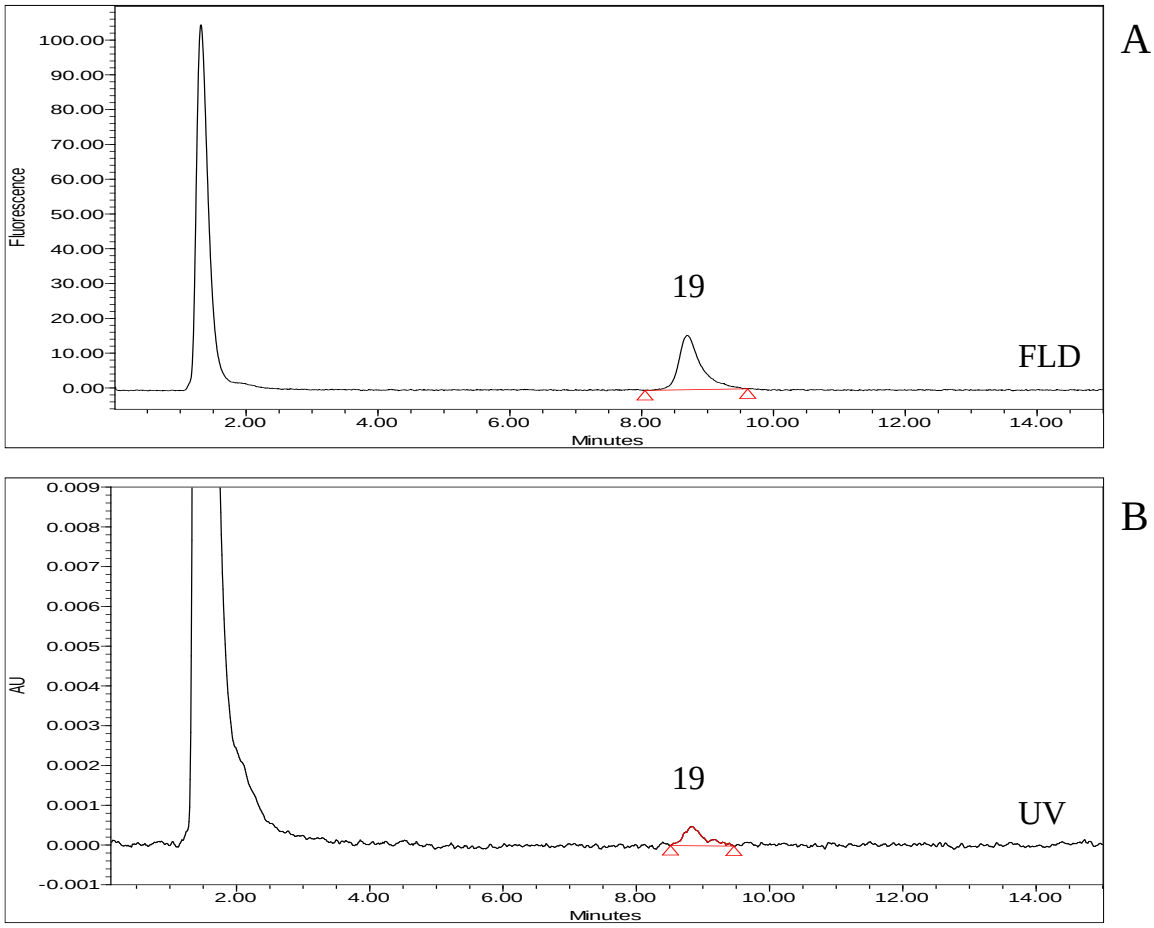


Figure 4-7 Chromatograms of Fmoc-Pro-Oct (**19**) derivative obtained during RP-HPLC analysis: A) fluorescence (λ_{EX} 263 nm, λ_{EM} 309 nm) B) UV detection (263 nm). Amount of the analyte injected on column was 10 pmol.

The UV and FLD traces of the Fmoc-Pro-Oct derivative resemble that of the Fmoc-Gly-Oct clearly showing a higher peak recorded with the use of fluorescence detector than the ultraviolet detector (**Figure 4-7**A-B), implying that the former technique has higher response and should provide a lower detection levels.

Indeed, analysis of the HPLC-UV and FLD data for the proline derivative, applying the same reasoning as for the previously discussed two derivatives allows the LOD to be calculated, yielding the following values: 8 pmol for the ultraviolet and approx. 0.2 pmol for the fluorescence detection.

Finally, the chromatograms recorded for the Fmoc-Phe-Oct derivative closely resemble those of glycine and proline derivatives with regards to the peak height, thus the expected LOD values should be also comparable (**Figure 4-8**A-B). The LOD values for the Fmoc-Phe-Oct derivative were established to be approx. 9 pmol for the UV detection and approx. 0.3 pmol for the FLD, in good agreement with values calculated for Fmoc-Gly and Fmoc-Pro derivatives.

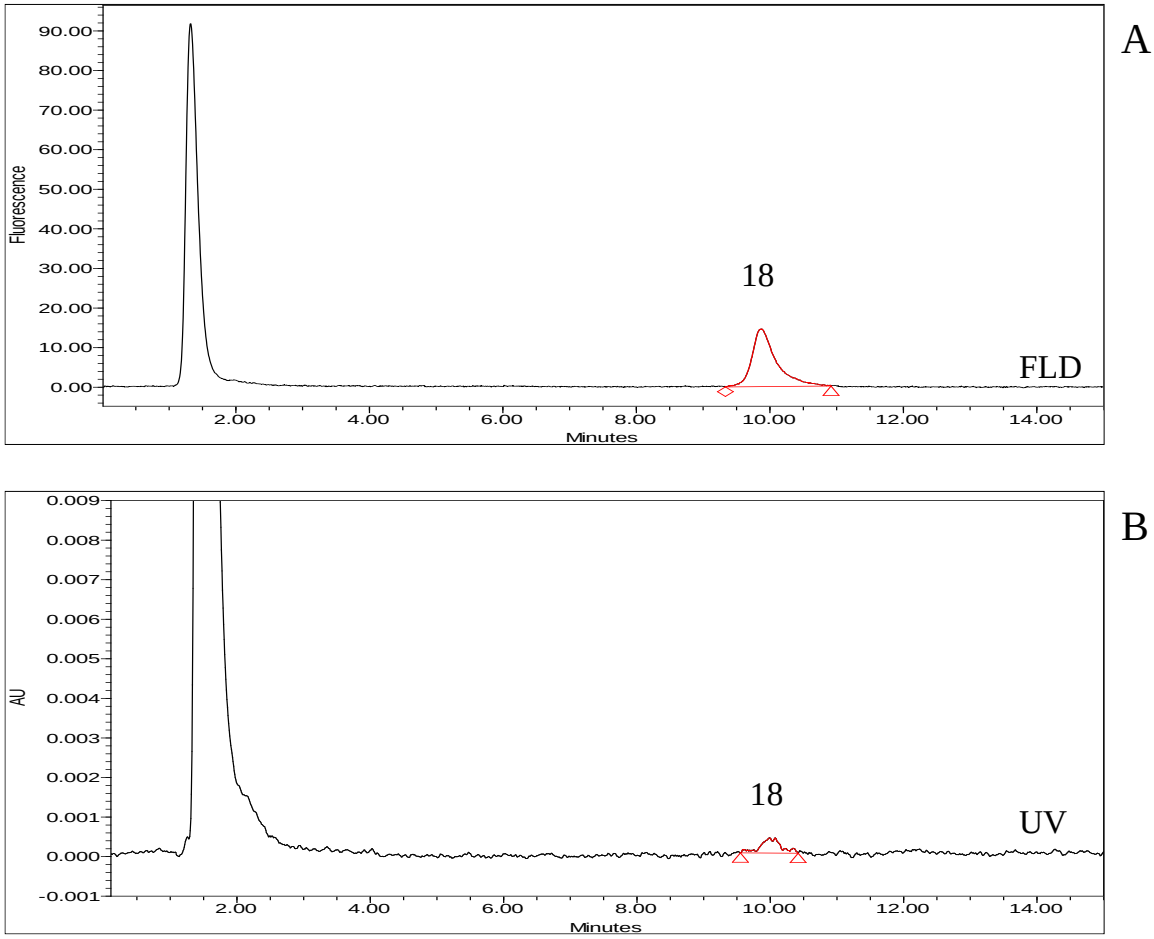


Figure 4-8 Chromatograms of Fmoc-Phe-Oct (**18**) derivative obtained during RP-HPLC analysis: A) fluorescence (λ_{EX} 263 nm, λ_{EM} 309 nm) B) UV detection (263 nm). Amount of the analyte injected on column was 10 pmol.

The calculated values of the peak area, signal-to-noise and the LOD for all four 1-octadecanol Fmoc amino acid modified derivatives are summarised in **Table 4-5**.

Table 4-5 HPLC-UV-FLD data of the 1-octadecanol N-Fmoc amino acid derivatives with calculated limits of detection (LOD) assuming that LOD > 3 x *S*/*N*. In all cases the amount injected on column was 10 pmol.

Derivative	**UV**			**FLD**		
	Peak area	*S/N*	LOD [pmol]	Peak area	*S/N*	LOD [pmol]
Fmoc-Gly-Oct (**17**)	1.04×10^4	4.0	7.5	3.54×10^5	125	0.2
Fmoc-Trp-Oct (**20**)	1.54×10^4	4.8	6.3	3.73×10^4	13	2.2
Fmoc-Pro-Oct (**19**)	1.03×10^4	3.7	8.1	3.69×10^5	120	0.2
Fmoc-Phe-Oct (**18**)	1.22×10^4	3.4	8.9	3.72×10^5	112	0.3

The data in **Table 4-5** indicates that the fluorescence response was uniform among all derivatives tested and the resulting limit of detection (LOD) is approx. 0.2 pmol, except for the Fmoc-Trp derivative that is 10 times worse due to intramolecular quenching. As such, the Fmoc-Trp derivative is not a good choice of derivatising agent for practical applications of the approach for quantification purposes. The behaviour of the derivatives during ultraviolet monitoring was broadly uniform, with LOD values ranging from 6.3 pmol (Fmoc-Trp) to approx. 9 pmol (Fmoc-Phe) suggesting that there are no obvious and immediate benefits that would influence the choice of the amino acid derivatisation agent when UV detection is considered.

4.4.3 Tandem mass spectrometric studies of *N*-Fmoc derivatives of 1-octadecanol

Investigation of the behaviour of the *N*-Fmoc amino acid modified 1-octadecanol derivatives was carried out using ion trap mass spectrometry, similar to the *N*-Boc amino acid derivatives as discussed in Chapter 3.3.5.

The protonated molecule was selected for CID to examine products of its dissociation. During the resonance induced dissociation of Fmoc-Gly-Oct (**17**) a prominent product ion was observed at *m/z* 328.3, confirming the direct formation of this ion from the protonated molecule **Figure 4-9**A. The two other minor ions in the MS^2 spectrum were not assigned. The base peak from the MS^2 spectrum was dissociated further and its product ions are shown in the MS^3 spectrum in **Figure 4-9**B. The product ion at *m/z* 282.4, 46 Da lower than the precursor ion, could potentially have formed *via* loss of a molecule of water from the amino acid functionality, accompanied by a loss of ethene from the alkyl chain. Additional product ions: *m/z* 182.3, 154.4, 127.4, 113.5, 99.7 observed in the MS^3 spectrum represent subsequent losses of 28 Da, and are characteristic of the dissociation of a linear alkyl hydrocarbon chain. Unfortunately, the marked decrease in the ion intensity suggests that CID beyond MS^2 may be of very limited use in structure elucidation studies of Fmoc-Gly derivatives. Most probable is the loss of the alkyl chain (de-esterification) and the glycine amino acid is not recorded due to LMCO.

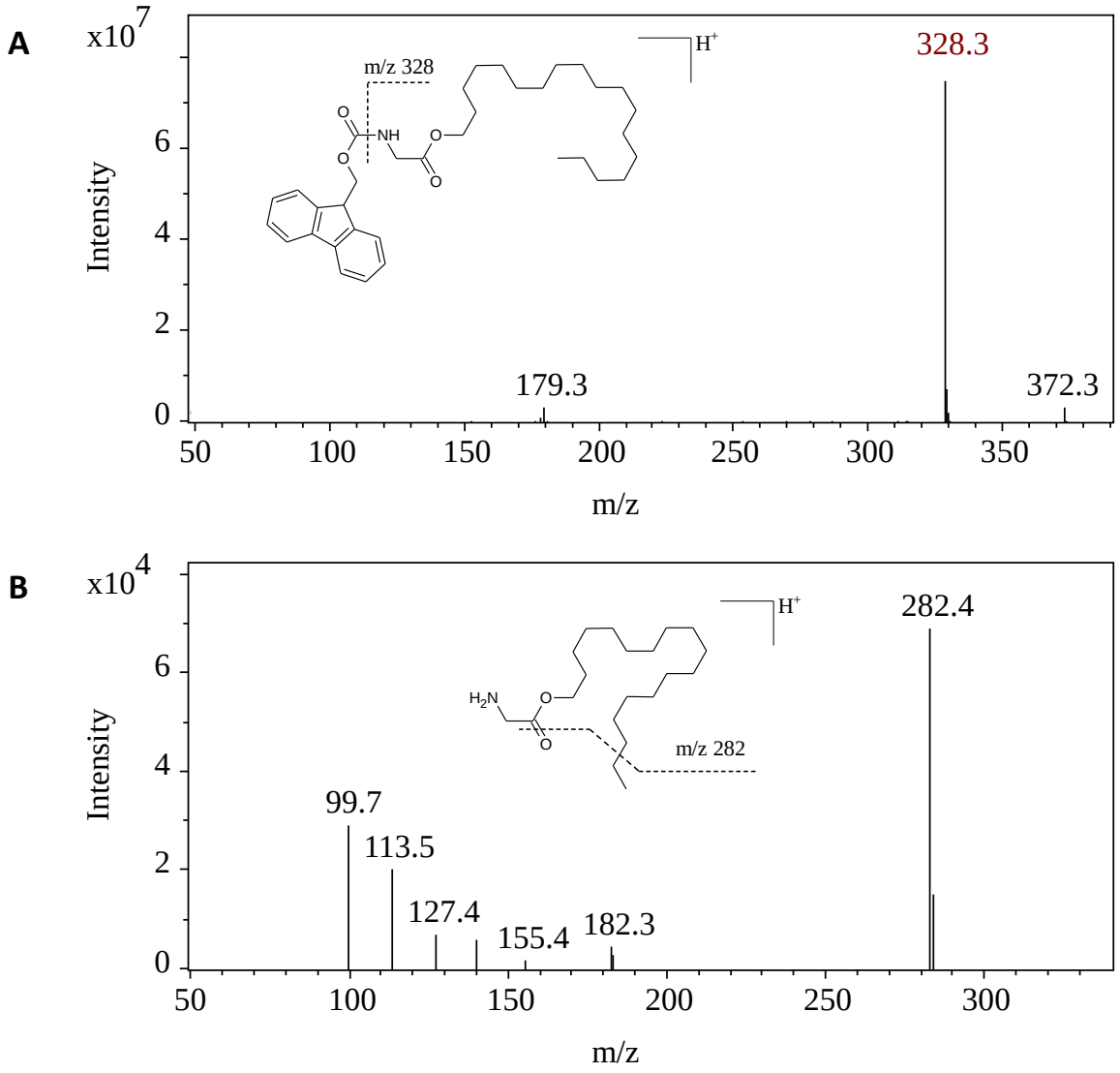


Figure 4-9 A) MS^2 spectrum of the peak at *m/z* 550.7, B) MS^3 spectrum of the peak at *m/z* 328.3 obtained during direct infusion of the Fmoc-Gly-Oct derivative.

The relatively high stability of the protonated molecule of the Fmoc-Trp-Oct derivative (*m/z* 679.5) is evident by it being base peak in the full MS spectrum **Figure 4-1**B. Two prominent product ions were observed in the MS^2 spectrum: *m/z* 457.4 and *m/z* 412.4 **(Figure 4-10**A**)**. Notably, these ions were also observed during the tandem MS studies of the Boc amino acid derivatives of 1-octadecanol as discussed in Chapter 3.3.5. They owe their formation to the loss of the Fmoc moiety and concomitant loss of Fmoc and ethene, respectively. Both ions we subjected to further dissociation to produce the MS^3 spectrum (**Figure 4-10**B-C). The main product ion resulting from dissociation of the base peak at *m/z* 457.4 was at *m/z* 440.4, indicating a loss of 17 Da. This ion is most likely formed by a loss of ammonia from the amine group of the MS^2 product ion. The presence of two minor product ions at *m/z* 412.5

and *m/z* 144.5 suggests that the stability of the main product ion is very high, resulting in the ion current being narrowly focused on this ion and highlighting its potential benefit as a diagnostic ion. The CID of the base peak ion at *m/z* 440.4 revealed a much richer MS^3 spectrum indicating, especially in the lower *m/z* end of the spectrum, that the derivative undergoes extensive further cleavages within the amino acid moiety of the derivative. The product ions at *m/z* 398.4 and *m/z* 380.4 were not assigned and are not believed to extend the understanding of the tandem MS dissociation of Fmoc-Trp-Oct derivative beyond that discussed above. No diagnostic product ions pertaining to specific dissociation of the lipid moiety of the derivative were detected, indicating that the major dissociation occurs within the amino acid part of the derivative, similarly to the Boc-Trp-Oct derivative discussed in the previous chapter.

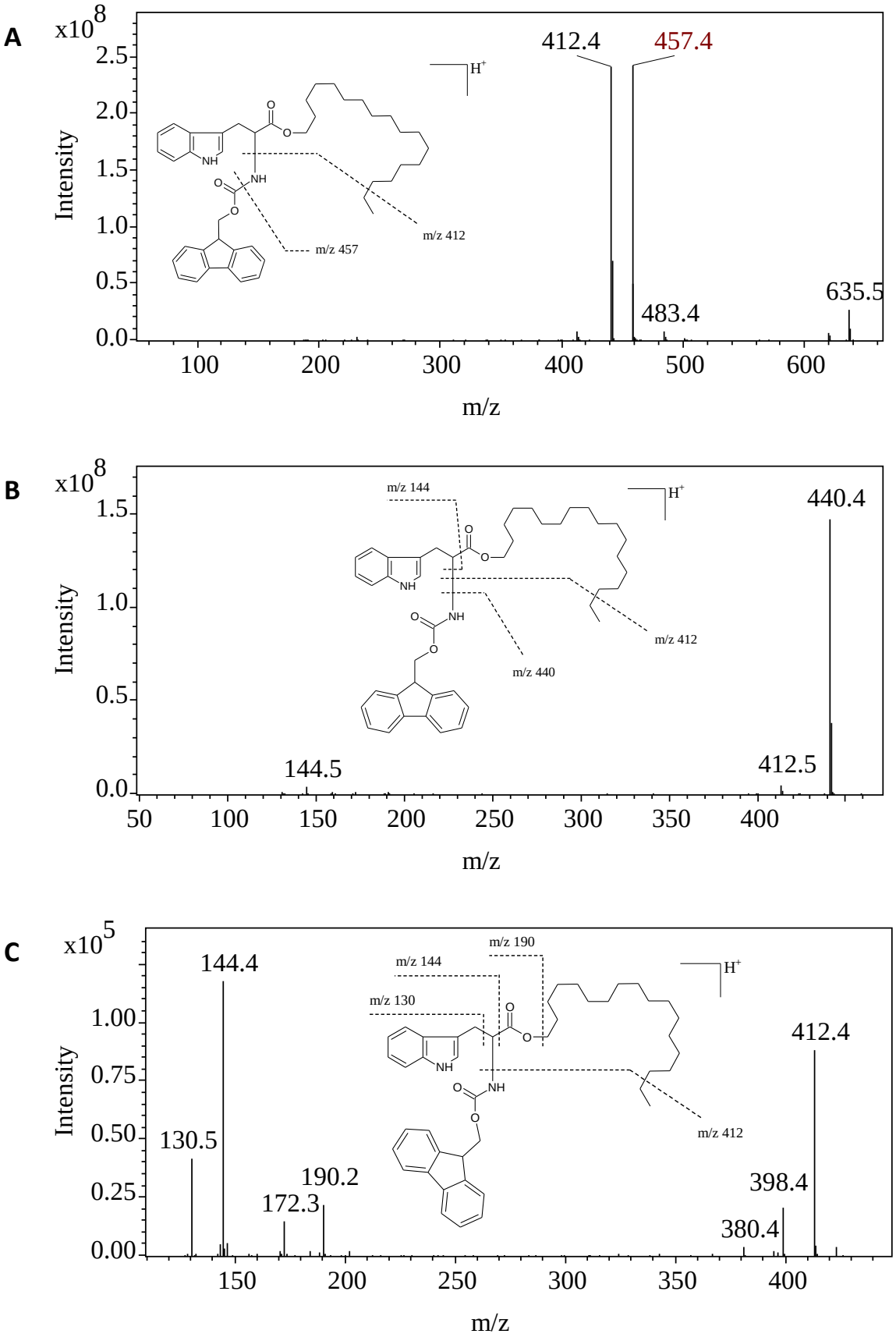


Figure 4-10 A) MS^2 spectrum of the peak at *m/z* 679.5, B) MS^3 spectrum of the peak at *m/z* 457.4, C) MS^3 spectrum of the peak at *m/z* 440.4 obtained during a direct infusion of the Fmoc-Trp-Oct derivative.

Analysis of the MS^2 spectrum produced by the resonance induced dissociation of the protonated molecule of the Fmoc-Pro-Oct derivative indicated that the cleavage of the Fmoc group leads to the formation of only one product ion at *m/z* 368.4 (**Figure 4-11**A-B). The dominance of this single ion suggests that the cleavage by which it is formed is energetically favoured and results in the formation of a stable product ion. Subsequent dissociation of the MS^2 base peak ion produced a rather simple MS^3 spectrum in which the prominent product ion at *m/z* 116.6 and a minor ion at *m/z* 322.4 were observed. The formation of the former is attributed to cleavage of the ester bond and the reconstitution of the amino acid molecule, with charge retention on the amino acid moiety. Interestingly, the same behaviour was observed during tandem MS studies of the Boc-Pro-Oct, suggesting that the dissociation of the derivative molecule is limited to the cleavage of the ester bond with formation of the amino acid, and as such does not offer any insight into the structure of lipid part of the derivative. The minor product ion at *m/z* 322.4 could not be assigned due to its very low intensity.

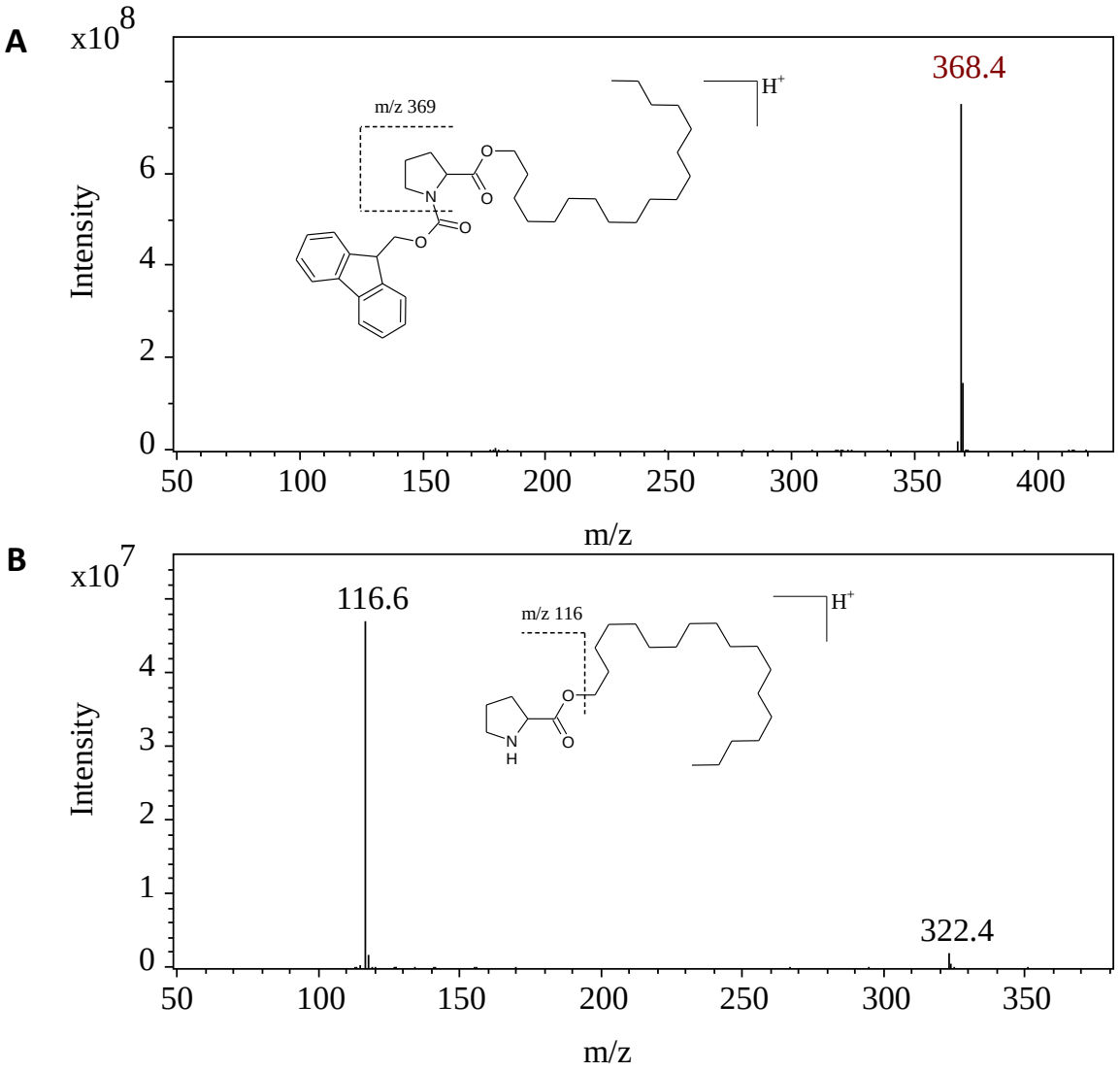


Figure 4-11 A) MS^2 spectrum of the peak at *m/z* 590.8, B) MS^3 spectrum of the peak at *m/z* 368.4, obtained during a direct infusion of the Fmoc-Pro-Oct derivative.

Finally, studies of the dissociation of the Fmoc-Phe-Oct derivative were performed by subjecting the protonated molecule of this derivative to multi stage tandem MS. The MS^2 spectrum of the protonated molecule of the derivative at *m/z* 640.8 revealed the ion at *m/z* 418.4 as the sole product ion (**Figure 4-12**A-B). This owes its origin to the loss of the Fmoc group from the protonated molecule of the derivative. Notably the same ion was observed in the mass spectrum of the Boc-Phe derivative of the model lipid during the tandem MS investigation. Further dissociation of this ion led to the formation of two prominent product ions at *m/z* 166.3 and at *m/z* 120.6. These ions were also been observed during the MS^2 CID dissociation of the Boc-Phe-Oct derivative. The former was formed as a result of a loss of the lipid moiety as a neutral molecule with regeneration of the phenylalanine amino acid.

The latter, was the product of a loss of the lipid alkyl moiety with simultaneous expulsion of CH_2O_2 in a similar manner as discussed in the case of Boc-Gly-Oct in Chapter 3.3.5. No diagnostic useful product ions pertaining to the lipid moiety were detected during the study of this derivative.

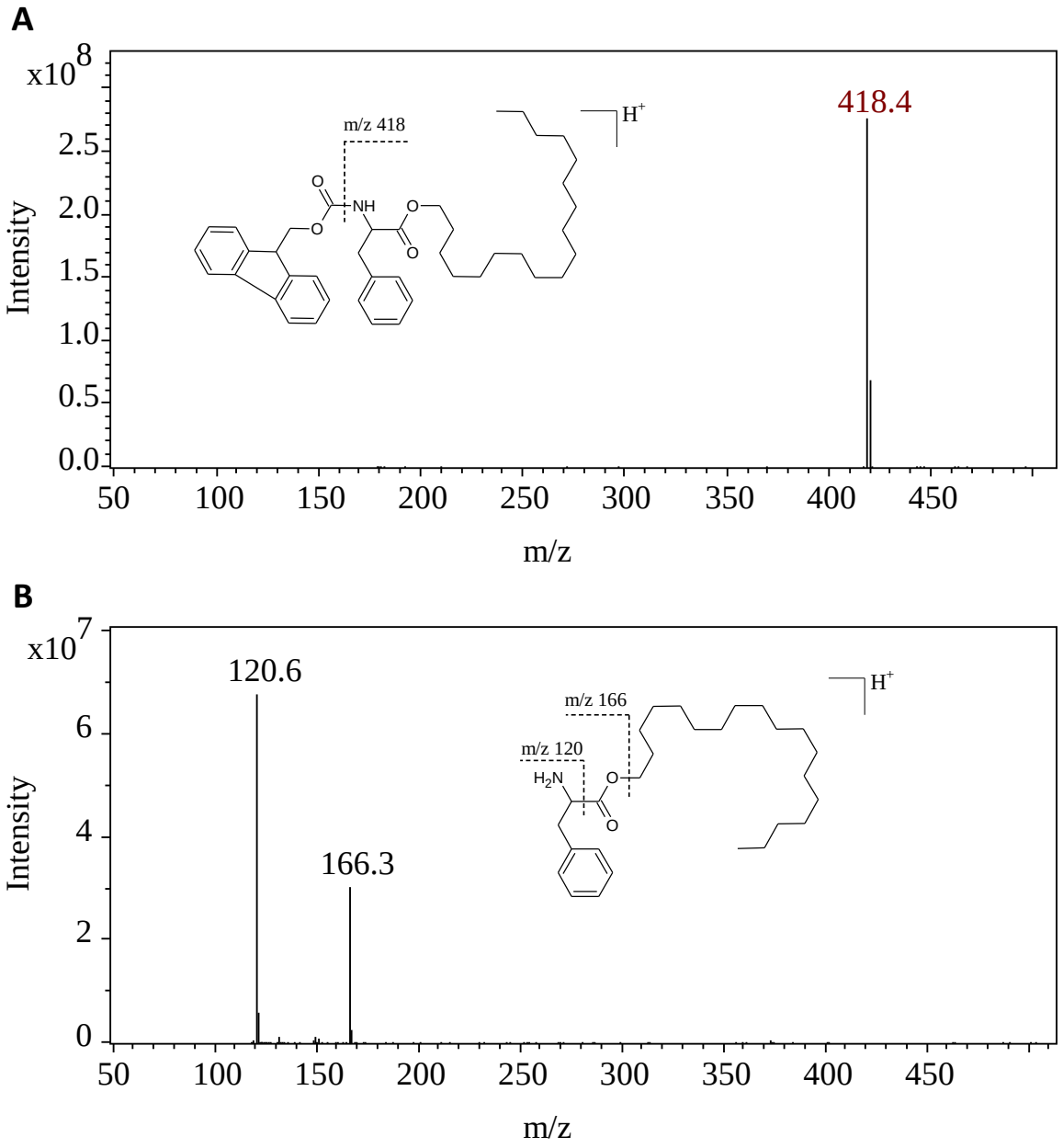


Figure 4-12 A) MS^2 spectrum of the peak at *m/z* 640.4, B) MS^3 spectrum of the peak at *m/z* 418.4, obtained during a direct infusion of the Fmoc-Phe-Oct derivative.

4.4.4 Conclusions

It has been demonstrated that 1-octadecanol can be modified using *N*-Fmoc protected amino acids in excellent yields (>94 %) under facile conditions assisted by acoustic irradiation. The isolation of the derivatives was straightforward and not labour intensive, which would permit the derivatisation to be utilized during daily operations in a high throughput manner.

The derivatives prepared were characterized using chromatography with mass spectrometry, ultraviolet and fluorescence detection. The analytes chromatographed well under the RP chromatography conditions, especially at higher analyte concentrations, and showed very good peak geometries with only slight tailing.

The signal intensity generated by the derivatives was very high, allowing the limit of detection to be calculated. The mass spectrometric limit of detection of approx. 0.2 pmol of analyte injected on column was uniform for all of the derivatives. The spectrophotometric LOD value varied between 6 pmol and 9 pmol for UV and the FLD was similar to that in MS except for the tryptophan derivative where it was 2.2 pmol. Notably, given that the LOD in fluorescence detection was comparable with the MS detection, it is apparent that the fluorescence methods can offer a truly alternative means of analysis. This in turn could enable many researchers to carry out lipid research using less costly equipment than any form of mass spectrometric equipment.

The MS behaviour of the derivatives was not uniform, and some derivatives showed significant fragment ions in full MS. The exception to that was the Fmoc-Trp derivative where most of the ion current was focused into the protonated molecule. During tandem MS studies the derivatives showed formation of high intensity product ions, associated predominantly with dissociation within the amino acid moiety, and no significant product ions pertaining to dissociation of the lipid moiety. Hence, no insight into the structural features of the lipid part of the derivative could be gained. This may be due to the nature of the model lipid, known for its poor ionization efficiency and lack of directed fragmentation under EI conditions. Nonetheless, the high intensity of the ions observed in full MS, as well as in the MS^2 scans, may be exploited in multiple reaction monitoring (MRM) scans following chromatographic resolution of the target analytes in complex samples of derivatised hydroxyl lipid constituents.

Chapter 5 Preparation and characterization of Fmoc-lysine(Boc) amino acid modified 1,2-di-*O*-octadecyl-*rac*-glycerol and cholesterol derivatives.

5.1 Introduction

The use of 1-octadecanol as a model compound in the development of a derivatisation protocol for lipids has been reported previously (Woo *et al.*, 2009). Although, it is a readily available and inexpensive test compound it may not adequately reflect the properties of lipid compounds for which this derivatisation method is intended. The fatty alcohol, having a long hydrocarbon chain with a hydroxyl group at its terminus, shares characteristics of the GDGT lipids including: non polar character, neutral charge and hydrophobicity. It is not known whether flexibility of the long hydrocarbon chain has any effect on the efficiency of ester formation under the derivatisation protocol or on the mass spectrometric behaviour of the *N*-protected amino acid esters of such a simple linear structure.

Significant differences exist between 1-octadecanol and lipids of other structural classes. For example, sterols may exhibit considerable structural rigidity and archaeal lipid cores possess multiple ether bound oxygen atoms in the immediate vicinity of the hydroxyl group that are involved in ester bond formation during the derivatisation. It is not known how those factors may affect the formation of the ester derivatives, their stability or behaviour during an analytical separation. It was, therefore, decided that the application of the derivatisation protocol established in the previous chapter should be evaluated directly on a model compound that better resembles the archaeal lipid cores both in structure and in chemical properties.

The use of a synthetic glycerol tetraether lipid as an internal standard was first introduced by Huguet *et al.*, (2006), for the purposes of monitoring the efficiency of the sample preparation process and to enable quantification of the tetraether lipids. The lipid labelled $GTGT_{46}$, denoting a glycerol trialkyl glycerol tetraether lipid containing a total of 46 carbon atoms, was employed (**Figure 5-1**).

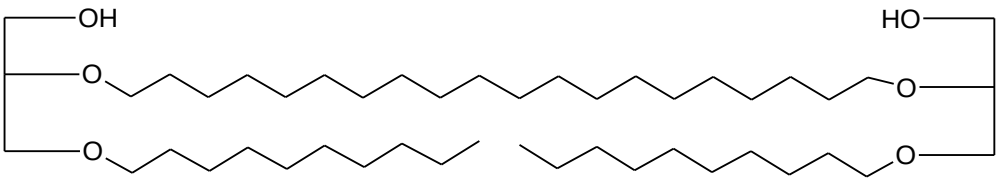


Figure 5-1 Structure of the $GTGT_{46}$ synthetic tetraether lipid standard material.

Although the $GTGT_{46}$ compound is available from commercial suppliers it is rather expensive. Thus, a synthesis of this type of compound was envisioned as a viable option to make a pure standard material for testing in the current study.

The *N*-Boc protected amino acid modified 1-octadecanol derivatives show excellent ionization properties during APCI while producing very little fragmentation, as discussed in Chapter 3.3.2. They do not, however, possess an active chromophore, hence their detection at low concentration is limited to mass spectrometry. While the *N*-Fmoc derivatives show excellent spectrometric and, except for Trp, fluorescent properties, they produce undesired fragment ions during the ionization process in MS. The presence of the fragment ions makes the derivatives less appealing owing to the potential complications they may cause in the analysis of complex mixtures in which components are not fully resolved (see Chapter 4).

It was decided to try to prepare a hybrid derivatising agent: a component offering the benefits of simple ionization in full MS *via* the presence of a Boc group, combined with the presence of a chromophore (Fmoc). Such a structure could alleviate the limitations of the *N*-Fmoc derivatives. Due to the capacity for orthogonal protection of the two amine groups, lysine offers the potential to meet the requirements. Hence, the Fmoc-lysine(Boc) amino acid (Fmoc-Lys-Boc) was selected for evaluation during derivatisation studies of cholesterol and the synthetic ether lipid.

5.2 Aims

The aims of the research discussed in this chapter were:

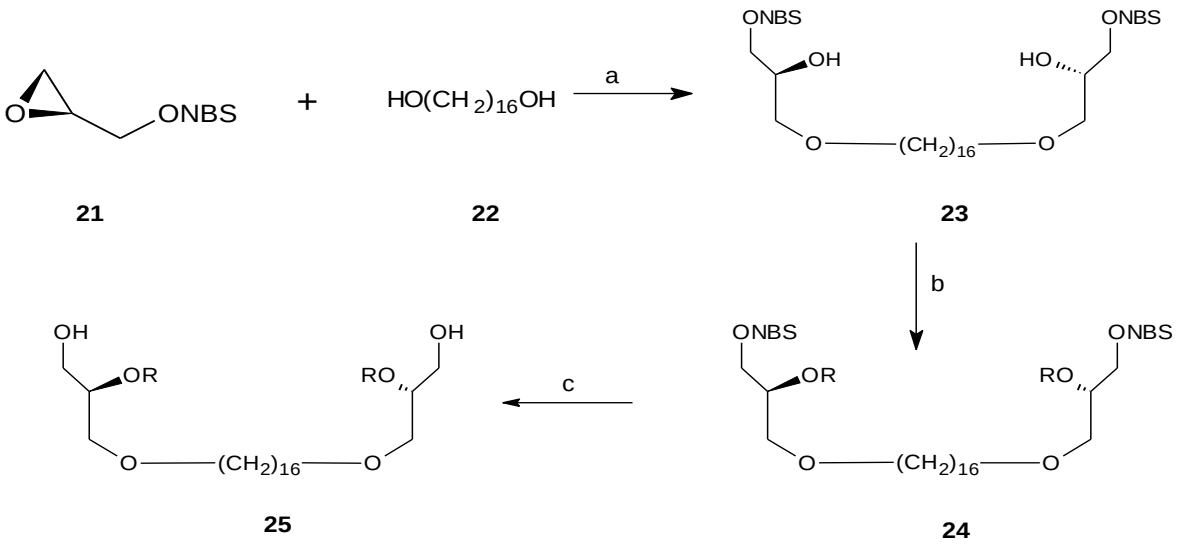
- to prepare a glycerol ether standard material that resembles the naturally occurring archaeal glycerol ether lipid archaeol;
- to evaluate the suitability of Fmoc-lysine(Boc) as a derivatising agent for modification of the synthetic glycerol ether and cholesterol;
- to investigate the chromatographic behaviour of the derivatives during RP-HPLC runs as well as their relative responses under APCI-MS, ultraviolet (UV) and fluorescent (FLD) detection modes.

5.3 Preparation of 1,2-di-*O*-octadecyl-*rac*-glycerol (*r*-dOG) lipid standard compound

Multiple synthetic routes have been reported thus far for the synthesis of glycerol tetraether lipids. These can be divided into two categories depending on whether: a) the chiral purity with regards to C2 position of glycerol is controlled by the use of stereo chemically pure starting material (Thompson *et al.*, 1994; Benvegnu, Brard and Plusquellec, 2004; Benvegnu *et al.*, 2005; Goodby *et al.*, 2007; Terme *et al.*, 2014) or b) racemic glycerol starting material leads to production of a racemic mixture of products (Paltauf and Spener, 1968; Raguse *et al*., 2000). From the plethora of reported synthetic protocols, attention was first given to considering those offering the least complexity, requiring relatively inexpensive reactants, least rigorous reaction conditions. Final selection was based on the strategy involving fewer steps and thus having the potential to give a satisfactory yield of the final product.

The synthesis of a chiral tetraether lipid, of the same stereochemistry as the naturally occurring tetraether lipids, was attempted following the method described by Thompson *et al.*, (1994). In this method a glycidol 3-nitrobenzenesulfonate (**21**) is reacted with 1,16-hexadecanediol (**22**) in presence of a catalytic amount of triflic acid under reflux in chloroform. According to the report, the triflic acid facilitates stereospecific ring opening of the glycidol ring and directs attack of the hydroxyl group at the less substituted position, C1, in the glycidol unit. The resulting diether diol (**23**) is reacted with an alkyl triflate to form a tetraether structure (**24**). Finally, the tetraether diol **25** is obtained by cleavage of the nitrobenzenesulfonate group with tetrabutylammonium hydroxide (TBAOH) in tetrahydrofuran (THF) (**Figure 5-2**).

The synthesis of a smaller glycerol trialkyl tetraether lipid, $GTGT_{38}$, was attempted mainly due to the lack of availability of α,ω-eicosanediol and the relatively low price of the shorter C_{16} terminal diol (**22**). Unfortunately, the first step where **21** and **22** were reacted could not be reproduced and the intermediate **23** was not produced. In the first attempt no product was recovered and the TLC indicated numerous side products with no compound corresponding to the R_f of the product reported by Thompson (1994). This was most likely caused by insufficiently dry reaction conditions: traces of moisture in the reaction could have quenched the triflic acid before it had a chance to facilitate the ring opening of **21**. The reaction was repeated using freshly distilled chloroform. This time the TLC monitoring indicated a compound with an R_f very close to that reported in the literature.



(a) CF_3SO_3H, $CHCl_3$, reflux
(b) $ROSO_2CF_3$, PS, CH_2Cl_2, reflux
(c) TBAOH, THF, rt
R = *n*-$(CH_2)_8OH$

Figure 5-2 Reaction scheme for the synthesis of a chiral tetraether diol lipid $GTGT_{38}$.

The presence of a number of closely eluting compounds, potential side products, was apparent. Side products may have been formed in the reaction of the highly reactive triflic acid with either of the reactants and, subsequently, further reaction with the products of the first reaction. Another possibility is that the acid may have reacted with the material of the seal used on the flask during the reaction as black spots were noted to have formed in the area of perforation where the acid was added using a syringe and needle. Inability to mitigate the issues rendered this method inapplicable in our laboratory setting.

After much consideration, it was decided that a simpler ether lipid may suffice as a standard testing material. Hence, 1,2 -di-*O*-octadecyl-*rac*-glycerol (*r*-dOG, **29**), similar to naturally occurring archaeol (2,3 -di-*O*-octadecyl-*sn*-glycerol), was selected. Unlike, the synthetic racemic material, the naturally occurring archaeol is a pure stereoisomer. The presence of two stereoisomeric structures in the synthetic *r*-dOG should not present any issues during reversed-phase chromatographic analysis: under such conditions, the stereoisomers are not discernible. The title compound (*r*-dOG, **29**) was synthesised in a typical Williamson etherification reaction under dry conditions (**Figure 5-3**). 1-*O*-benzyl-rac-glycerol (**26**) was directly alkylated with excess 1-bromohexadecane (**27**) in the presence of sodium hydride (NaH) in dry tetrahydrofuran (THF) under reflux conditions. Reaction was facilitated by a

catalytic amount of tetrabutylammonium iodide added to the reaction mixture to convert the bromide into iodide *in situ*, iodide being a better leaving group. The protected glycerol diether (**28**) was isolated on a silica gel column in 80 % yield as a slightly dark yellow oil. Subsequent debenzylation reaction using hydrogen gas (H_2) and palladium on carbon (Pd/C, 10 % load) as catalyst in ethyl acetate as solvent gave the diether glycerol (**29**). Following isolation on a silica gel column and solvent evaporation the product was obtained as a white powder in approx. 80 % yield. The melting point of the material was found to be 65 – 66.2 °C (lit. 62 – 63 °C).

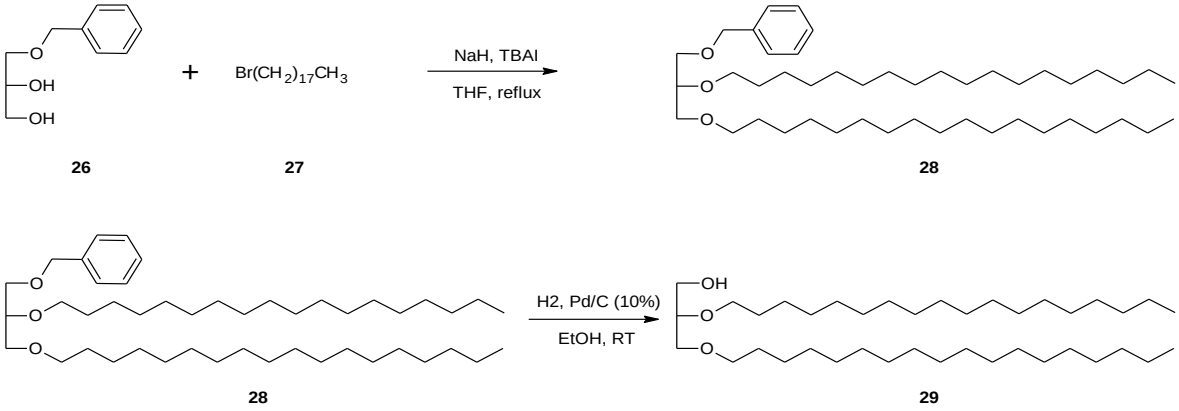


Figure 5-3 Synthetic route to 1,2-*O*-octadecyl-*rac*-glycerol (*r*-dOG, **29**) standard material.

The compound **29** was identified by the presence of the ion corresponding to its protonated molecule $[M+H]^+$ ion at *m/z* 597.5 in the APCI-MS and the product purity was assessed to be greater than 98 % (HPLC-APCI-MS) (**Figure 5-4**).

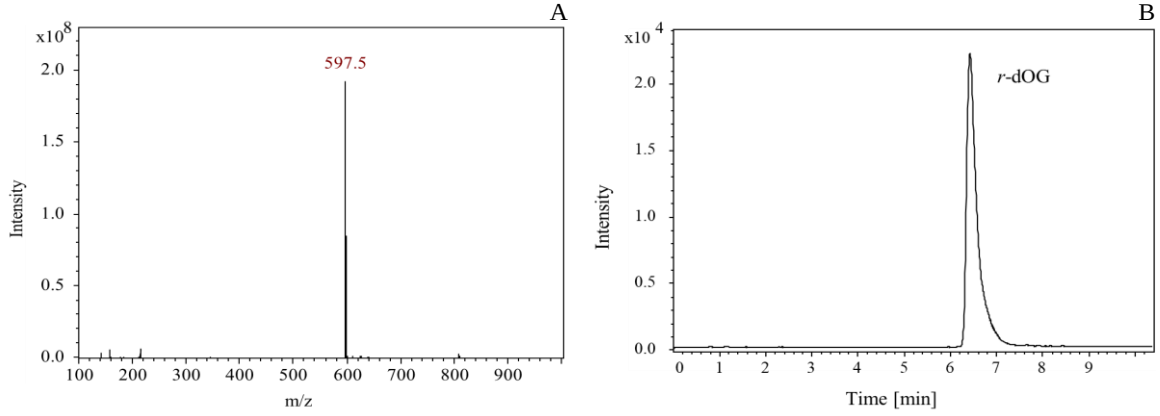


Figure 5-4 Synthetic glycerol diether **29** (*r*-dOG): A) full (+) APCI mass spectrum, B) HPLC-APCI-MS mass chromatogram.

5.4 Preparation of Fmoc-lysine(Boc) amino acid derivatives

Although the main target for the modification of the *r*-dOG was the Fmoc-Lys-Boc, derivative, the Boc-Phe and Fmoc-Phe amino acids were also prepared. The two orthogonally protected phenylalanine amino acids were selected because of the excellent performance they exhibited as derivatising agents for 1-octadecanol, and to allow direct comparison with the 1-octadecanol derivatives. The selection of orthogonally protected lysine (**Figure 5-5**) was dictated by the desire to exploit the benefits of the Boc group in the formation of a highly intense MS signal during the ionization process as well as being able to exploit the UV and, potentially, fluorescent activity of the Fmoc group during chromatographic monitoring and detection of the derivatised lipid molecules.

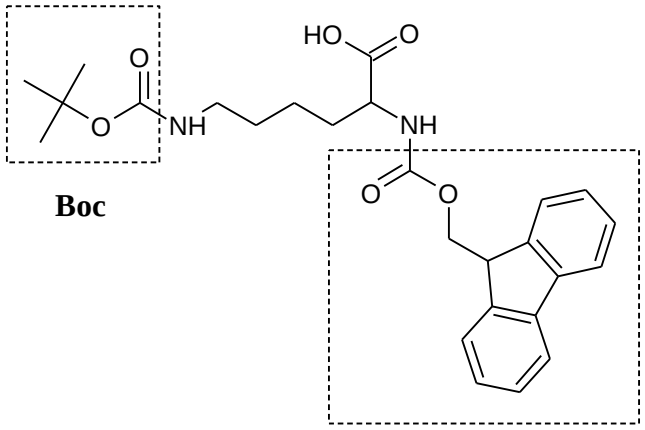


Figure 5-5 Structure of the orthogonally protected Fmoc-lysine(Boc) amino acid.

The derivatisation reactions were carried out according to Method C (Chapter 2.2.3) and were performed in triplicate. The calculated weight-based percentage yields of the isolated derivatives are summarized in **Table 5-1**, including the derivatives of 1-octadecanol using orthogonally protected phenylalanine amino acid.

Table 5-1 Results of derivatisation of 1-octadecanol and *r*-dOG model lipids with selected *N*-protected amino acids using in accordance to Method C. The reactions were carried out in triplicate. Average yields and relative standard deviation (RSD) are given.

Derivative	Yield [%]				
	1	2	3	Average	RSD
Boc-Phe-*r*-dOG	97.2	97.5	97.3	97.3	0.2
Fmoc-Phe-*r*-dOG	97.6	97.4	97.6	97.5	0.1
FLB-*r*-dOG	98.5	98.2	98.3	98.3	0.2
FLB-Octadecanol	97.8	97.9	98.2	98.0	0.2
Boc-Phe-Octadecanol*	98.0	98.1	97.9	98.0	0.1
Fmoc-Phe-Octadecanol *	94.4	93.2	N/A	93.8	N/A

Note: ***** denotes derivatives prepared in earlier studies using the same derivatisation protocol (Method C).

N/A the Fmoc-Phe-octadecanol derivatisation was performed in duplicate and as such the RSD was not calculated

The data shows that the derivatives of the archaeal lipid analog (*r*-dOG) can be prepared repeatedly in almost quantitative yields. The fractional differences between the replicates attest to the repeatability of the derivatisation protocol. The excellent yields of the alternatively protected derivatives of *r*-dOG prepared using phenylalanine suggest that the use of alternative amino acid modifying agents should lead to equally high yields, despite the nature of the protecting group. Moreover, comparing the yields of the derivatives of two different lipid molecules modified with Fmoc-Lys-Boc indicates only marginal differences in the recorded yields: 98.3 % for *r*-dOG *versus* 98.0 % for 1-octadecanol. Notably, the yields of the diether lipid derivatives of Fmoc-Lys-Boc are comparable to the yields of the derivatives obtained using Fmoc-Phe, further affirming its selection as the derivatisation agent of choice and emphasising the potential for it to be used in a wide range of applications with samples containing hydroxyl lipids.

The excellent results obtained during the preparation of derivatives of the two distinct model lipid classes (alkanol and glycerol dialkyl diether) prompted a question of whether this

derivatisation method can be applied to another class of lipid compounds, sterols, which occur abundantly in environmental samples. Should the application of this modification approach be successful it could offer real benefits to a broad spectrum of environmental scientists concerned with their monitoring and detection.

For the evaluation of the applicability of this methodology cholesterol, as an inexpensive and readily available sterol, was used. The preparation of its derivative using Fmoc-Lys-Boc under the conditions of Method C (see Chapter 2.2.3) was carried out in triplicate.

Table 5-2 Results of derivatisation of cholesterol with Fmoc-Lys-Boc amino acid using Method C. The reaction were carried out in triplicate. Average yield and relative standard deviation (RSD) are given.

Derivative	**Yield [%]**				
	1	2	3	Average	RSD
FLB-Cholesterol	97.8	98.2	98.0	98.0	0.2

The data in **Table 5-2** shows that excellent yields of the cholesterol derivative (average of 98 %, *n*=3) can be obtained using the Fmoc-Lys-Boc derivatising agent. The very small RSD value for the replicate yields indicates the high reproducibility of the reaction under conditions of Method C, ensuring almost quantitative conversions during routine modification. Interestingly, unlike the synthetic diether *r*-dOG, cholesterol is a secondary alcohol and, as such, it would be expected to react more slowly than a primary alcohol. This, however, did not impact on the yields for this derivatisation, indicating that the reaction conditions, and sonic irradiation in particular, provide suitable conditions to effect a near complete conversion of cholesterol.

The high product yields achieved in the derivatisation reactions of the three model compounds suggests that the conditions of the derivatisation protocol of Method C are suitable for a broad classes of hydroxyl lipids and lead to excellent conversion rates. This, in conjunction with a highly efficient acoustic agitation, resulted in the successful preparation of the derivatives of the three distinct lipid molecules modified using Fmoc-lysine(Boc).

5.4.1 Characterization of the derivatives of *r*-dOG and cholesterol

5.4.2 Detection of the Fmoc-lysine(Boc) derivatives of *r*-dOG and cholesterol

The full mass spectra of the Fmoc-Lys-Boc amino acid modified cholesterol and the diether lipid derivatives were produced during infusion of the analyte solution into the (+) APCI-MS *via* a syringe pump in the same manner as for the derivatives discussed earlier.

The Fmoc-Lys-Boc derivatives of glycerol dialkyl diether (*r*-dOG, **29**) and cholesterol exhibit a similar behaviour in the full MS to the *N*-Boc derivatives, mainly due to the presence of the Boc group in the amino acid moiety of the derivative. The intense ion, at 100 Da lower than the protonated molecule in the full MS scan of the Fmoc-Lys-Boc derivatives, is attributed to loss of the Boc group during ionization and, as such, is believed to be formed according to the mechanism of loss discussed in Chapter 3.3.2. Similarly, to the 1-octadecanol *N*-Boc amino acid derivatives, only a few fragment ions of very low intensity are observed in full MS, suggesting that the cleavage of the Boc group is the main fragmentation pathway. Notably, the presence of only a small number of low intensity fragment ions at *m/z* values below that of the base peak indicate its high stability (see **Figure 5-6**A-B) and the potential for it to be exploited during subsequent tandem MS dissociation. The mass losses from the protonated molecule of the derivatives are summarized in **Table 5-3**. The low intensity ion at *m/z* 1047.7 in the mass spectrum of the *r*-dOG-Fmoc-Lys-Boc derivative (**Figure 5-6**A) corresponds to the protonated molecule, and the fragment ion at *m/z* 991.6 to the ion formed by a loss of isobutylene (C_4H_8, 56 Da) as a neutral molecule. The characteristic loss the Boc group gave rise to the base peak at *m/z* 947.8, while the low abundance fragment ion at *m/z* 825.7 corresponds to a loss of 222 Da from the protonated molecule, which represents loss of the whole Fmoc group as a neutral molecule. The low relative intensity of the latter suggests that the Fmoc group remains largely intact during the MS ionization process. Importantly, no fragment ion resulting from ester bond cleavage of the derivative was detected.

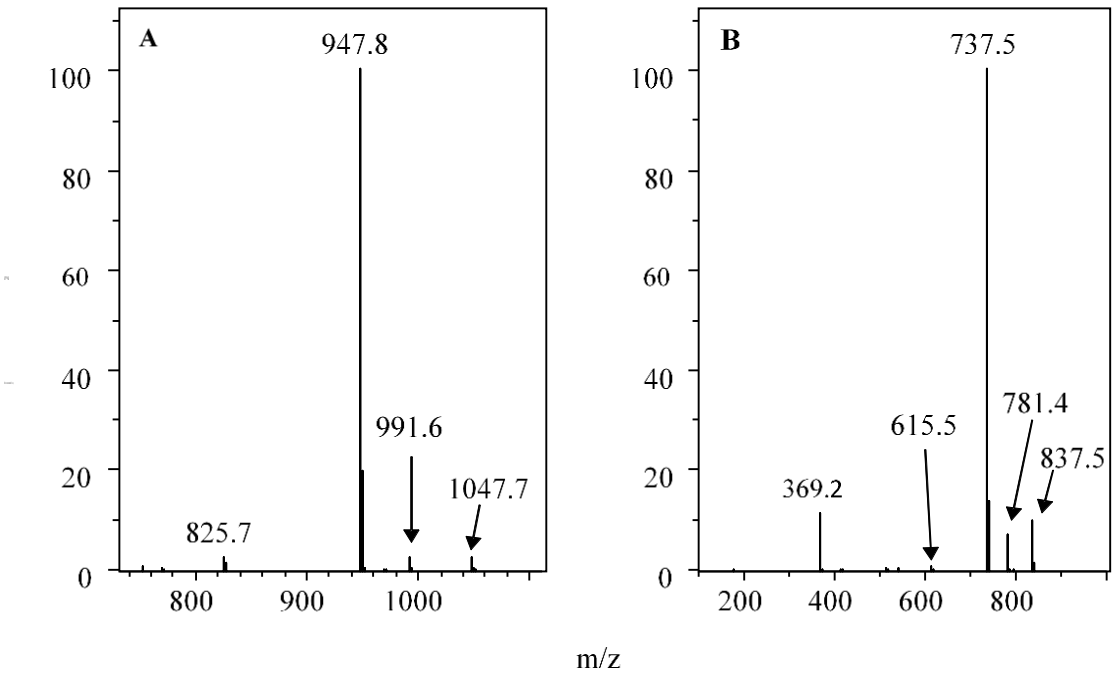


Figure 5-6 (+) APCI mass spectrum of the Fmoc-Lys-Boc derivatives of: A) synthetic glycerol diether (*r*-dOG) and B) cholesterol.

The ions recorded in the full mass spectrum of the cholesterol derivative reflect similar losses of neutral molecules to those detected for the synthetic glycerol diether, though with different relative intensities (**Figure 5-6**B).

The intensity of the protonated molecule at *m/z* 837.5 is *c.* 10 %, and the ion at *m/z* 781.4 (-56 Da, 8 %) corresponds to a neutral loss of isobutylene (**Table 5-3**). The loss of the Boc group again gave rise to the base peak at *m/z* 737.5 indicating this loss to be the most energetically favourable. The loss of the Fmoc group (-222 Da) resulted in a signal at *m/z* 615.5 of 1 % relative intensity. Finally, the loss of the entire of the amino acid moiety *via* ester bond cleavage, gave rise to a fragment ion at *m/z* 369.2 (*c.* 12 % relative intensity). Hence, though the loss of the Fmoc moiety is not favourable, the cleavage of the ester group is rather significant.

Table 5-3 Mass losses from the protonated molecule of the Fmoc-Lys(Boc) derivatives during direct infusion of the sample solution into (+) APCI-MS. Relative intensities of the ions are given in parentheses.

Mass Loss (Da)	FLB-*r*-dOG	FLB-Cholesterol
0	1047.7 (2.9)	837.5 (10.2)
-56 (C_4H_8)	991.6 (3.0)	781.4 (7.5)
-100 (Boc)	947.8 (100)	737.5 (100)
-222 (Fmoc)	825.7 (3.0)	615.5 (1.1)
-468 (Fmoc-Lys-Boc)	–	369.2 (11.5)

Strikingly, no ester bond cleavage was observed for the synthetic glycerol diether derivative, suggesting that the structure and therefore the chemistry of the hydroxyl lipid may play an important role in determining the type and the relative intensities of the fragment ions generated during ionization of their Fmoc-Lys-Boc derivatives during (+) APCI-MS. Although the nature of this difference has not been determined, it is suggested that the proximity to the ester bond of the double bond in cholesterol may be responsible for the enhanced loss in this case. This issue was not investigated further though it is noted that simple changes to the tune parameters of the MS instrument did not result in significant changes to the ion distributions in the mass spectrum of the cholesterol derivative.

5.4.3 HPLC-APCI-MS separation of Fmoc-Lys-Boc derivatives of *r*-dOG and cholesterol

The chromatographic separation of the synthetic diether lipid (*r*-dOG) and its Fmoc-Lys-Boc derivative was attempted on a C18 column under reversed phase chromatography mode using methanol:water:ethyl acetate (50:10:40, v/v) as mobile phase. As result of their greater apolar character than the *N*-Fmoc amino acid modified octadecanol derivatives ethyl acetate did not provide the desired chromatographic conditions. Two important observations were made when using ethyl acetate as the strong solvent: firstly, a high proportion of ethyl acetate (up to 65 %) had to be used to elute the analytes and secondly, the native and the derivative of the glycerol diether lipid both exhibited considerable peak tailing. The use of a higher proportion of ethyl acetate in the mobile phase restricted the proportion of water in the more polar methanolic component of the mobile phase to 15 %, due to solvent immiscibility. Additionally, an increased proportion of water led to increase in the system backpressure due to the increase of viscosity of the mobile phase. It is common practice to lower the system backpressure by increasing of the column temperature; this, however, resulted in deterioration of resolution between the native species and its derivative and led to considerable increases in peak tailing. Extensive peak tailing often reflects interactions between the analyte and active silanol sites of the stationary phase, indicating inefficient capping of those sites. Capping is a procedure for deactivation of the free silanol groups by silylation and is used during manufacturing of the stationary phase in order to minimize unwanted interactions. A common practical method of suppressing active sites is to use a diluted acid (i.e. formic or acetic) or ammonium salt of the acid in the mobile phase. Another cause of peak tailing could be inefficient analyte partitioning between the stationary and mobile phases, which can, in certain situations, be exacerbated by increased column temperatures. The lack of improvement in the peak tailing of the analytes when up to 0.5 % of an aqueous solution of acetic or formic acid was incorporated into the mobile phase rules out active silanol groups as the culprit, suggesting that the chromatographic conditions were not optimal for efficient analyte partitioning. It was suspected that the low proportion of the water and methanol mix to ethyl acetate did not offer a desired difference in polarity between the mobile phase components, reducing the speed of analyte partitioning. Additionally, the elevated column temperature may have increased the solubility of water in ethyl acetate, further diminishing the contrast of polarity between the aqueous methanol and ethyl acetate. Hence, the observed peak tailing might be attributed to ethyl acetate being an inadequately

strong solvent to chromatograph the native and the Fmoc-Lys-Boc derivative of the synthetic diether (*r*-dOG).

The excessive peak tailing impacted the signal–to-noise (*S/N*) ratio due to unnecessary analyte band broadening and consequent decrease in peak height, leading to poor limits of detection and quantification. This issue is often alleviated using a gradient elution program where, by gradually increasing the proportion of the strong eluent, the analyte may be focused to a narrow band, resulting in its detection as a narrow and sharp peak and thus giving higher *S/N* ratio. The higher the *S/N* ratio the lower the method detection limit (LOD) and *vice versa*. Unfortunately, application of gradient elution did not improve the analyte peak geometries, suggesting that an alternative solvent should be examined.

Dichloromethane was selected as the strong eluting solvent due to its excellent solubilizing properties and a lower polar character than ethyl acetate. It was thought that using a stronger non-polar solvent would permit much better control of the elution of the analytes due to a much greater difference in polarity between the aqueous-methanol and dichloromethane components of the mobile phase. A gradient elution program was developed (Chapter 2.6.2) which achieved good resolution between the native and modified *r*-dOG, while ensuring greater peak height due the analytes eluting in a much narrower band.

The synthetic diether eluted at 6.4 min and its derivative at 9.1 min, allowing for clear distinction between the two analytes as shown in **Figure 5-7**. The longer retention time of the derivative than the unmodified diether indicates the increase in the apolar character on derivatisation. Modest peak tailing is observed and is less pronounced for the derivative than for the native compound. It was considered satisfactory given the relatively low mobile phase flow rate of 0.5 mL min^{-1}. It is expected that the peak geometries would improve on adjusting the flow rate to a higher value to take full advantage of modern ultra-high pressure liquid chromatography (UHPLC) systems. In this work the back pressure limits of the conventional HPLC equipment used to study ultraviolet and fluorescence behaviour of the derivatives restricted the flow rate of the method to 0.5 mL min^{-1}.

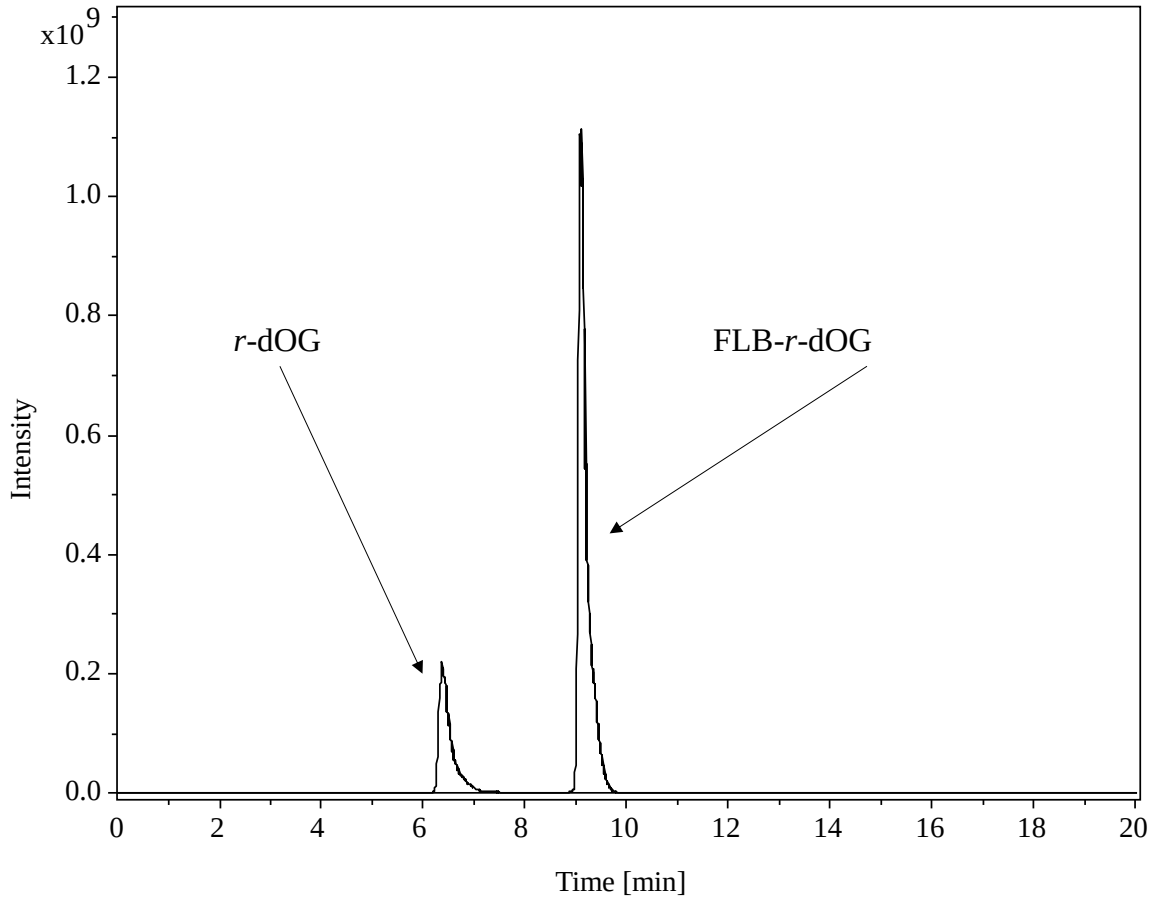


Figure 5-7 Extracted mass chromatograms for the synthetic diether (*m/z* 597.5, *r*-dOG) and its Fmoc-Lys-Boc derivative (*m/z* 947.8, FLB-*r*-dOG). The amount of injected on column was 100 nmoles.

Although the peak area of the derivative was only about four times higher than that of the native diether, the peak height increase of almost an order of magnitude for the former shows the benefits of the modification and indicates that the derivatisation might lead to improved detection limits (**Table 5-4)**. Comparing the peak geometries of the native and modified synthetic diether it is immediately apparent that the derivative peak is much narrower and sharper, only broadening towards the base of the peak. Such shape and the peak height should result in considerably better LOD values, and thus lead to detection of structures previously undetected or enable use of lower amounts of sample material.

Table 5-4 Retention times and peak areas of Fmoc-Lys-Boc derivatives of *r*-dOG and cholesterol obtained during RP-HPLC-APCI-MS analysis. The amount of injected on column was 100 nmoles.

Derivative	t_R [min]	Target ion *m/z*	Peak area
***r*-dOG**	6.4	597.5	3.5×10^9
FLB-*r*-dOG	9.1	947.8	1.3×10^{10}
Cholesterol	7.8	369.2	3.4×10^8
FLB-Cholesterol	10.6	737.5	2.2×10^9

Cholesterol is a more polar analyte than the *r*-dOG diether, mainly due to its more rigid and compact structure. Thus, its Fmoc-Lys-Boc derivative would be expected to be less polar than that of the diether, and as such it may not require the use of a very strong eluting solvent such DCM. Indeed, initial testing of ethyl acetate as the strong eluting solvent during chromatographic separation on a C18 column showed that it was capable of eluting both the native and the derivative of cholesterol, tolerating the presence of up to 15 % water in the mobile phase. It was found that holding the ratio of methanol:water:ethyl acetate at 73:15:12 (v/v) for 3 min was sufficient to allow the analytes to enter the column and circumvent any potential system void volume effects. This was followed by a fast gradient to methanol:water:ethyl acetate 58:12:40 (v/v) over 7 min to elute cholesterol and its derivative as separate peaks. The final composition was held for 2 min after which time the system was returned to the initial composition and allowed to equilibrate for 6 min (Chapter 2.6.3). The gradient elution program offered very good resolution between the native cholesterol and its derivative in less than 12 min at a flow rate of 0.5 mL min^{-1} (**Figure 5-8**). Notably, the peak geometries, similarly to the synthetic diether *r*-dOG were very good with only modest tailing.

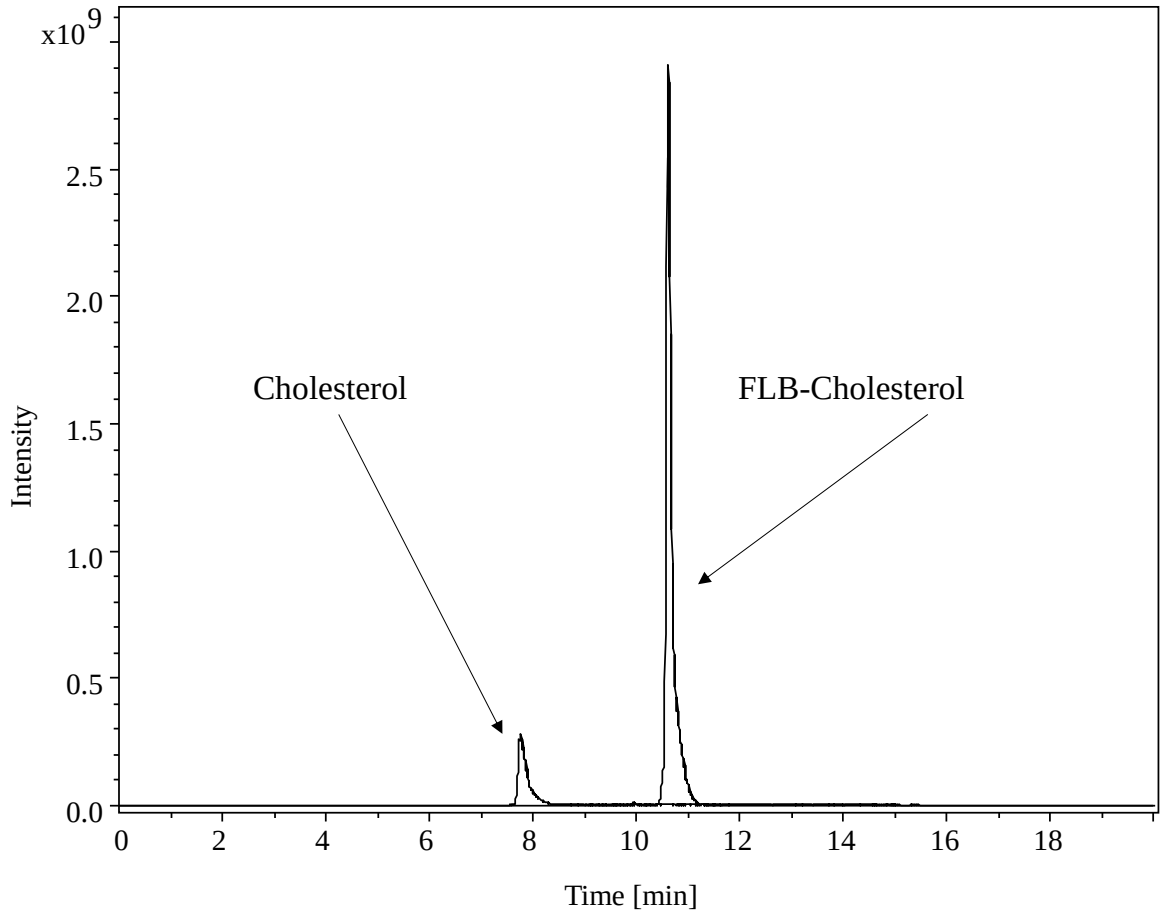


Figure 5-8 Extracted ion mass chromatograms of the cholesterol (*m/z* 369.2) and its Fmoc-Lys-Boc derivative (*m/z* 737.5). The amount of injected on column was 100 nmoles.

Similarly to the diether and its derivative, although the peak area was only slightly higher for the modified cholesterol *versus* the native compound, the peak area increased approx. 10 times as a result of the modification.

Changes in the initial mobile phase composition, to make it less polar, resulted in decreased resolution between the two peaks, though without a significant improvement in the peak shape of either of analyte. Although some improvement in the peak shape and thus peak height were observed at a higher flow rate (*i.e.* 0.7 mL min^{-1}) it caused the system backpressure to increase well beyond 400 bar.

Notably, however, as in the case of the synthetic diether *r*-dOG where the backpressure of the HPLC-UV-FLD used was limited of 400 bar, it is thought that the greater peak height that should be obtained under faster flow rates could improve the limit-of-detection and limit-of-quantification to levels close to those of the APCI-MS detector.

As expected, both derivatives produced highly intense ions at 100 units lower than the protonated molecule following characteristic loss of the Boc group as discussed previously (Chapter 3.3.2). The diether derivative (FLB-*r*-dOG) was detected as the protonated molecule at *m/z* 947.8 and cholesterol derivative (FLB-cholesterol) gave a base peak at *m/z* 369.2 due to a characteristic loss of one molecule of water from positions C3 in a manner commonly observed in APCI-MS (Raith *et al.*, 2005; Saldanha *et al.*, 2006; Cañabate-Díaz *et al.*, 2007).

The assessment of the dynamic signal response during HPLC-APCI-MS was performed by analysis of serial dilutions of a standard solution containing both the native analyte and its Fmoc-Lys-Boc derivative. The solution concentrations covered the range 1.0×10^{-4} M to 1.0×10^{-8} M. The injection volume of the analyte solution was fixed at 1 μL, representing a test range between 100 nmole to 10 pmole of the analyte injected on column. The evaluation of both the native and modified analyte was performed to rule out the potential bias of evaluating only the derivative and comparing the results to the literature values which might have caused biased or incorrect conclusions due to dependence of the performance of the ion trap mass spectrometer used in this study.

The mass chromatograms presented in **Figure 5-9**A show the FLB-*r*-dOG peaks recorded in the study with the peak for the injection of 10 pmol of the derivative being clearly represented. The calculated *S/N* ratio for this level is 95, suggesting that the limit of quantification (LOQ, 10 × *S/N*) lies an order of magnitude lower at *c.* 1 pmol injected on column and the limit-of-detection (LOD, 3 × *S/N*) at *c*. 0.3 pmol of the analyte on column.

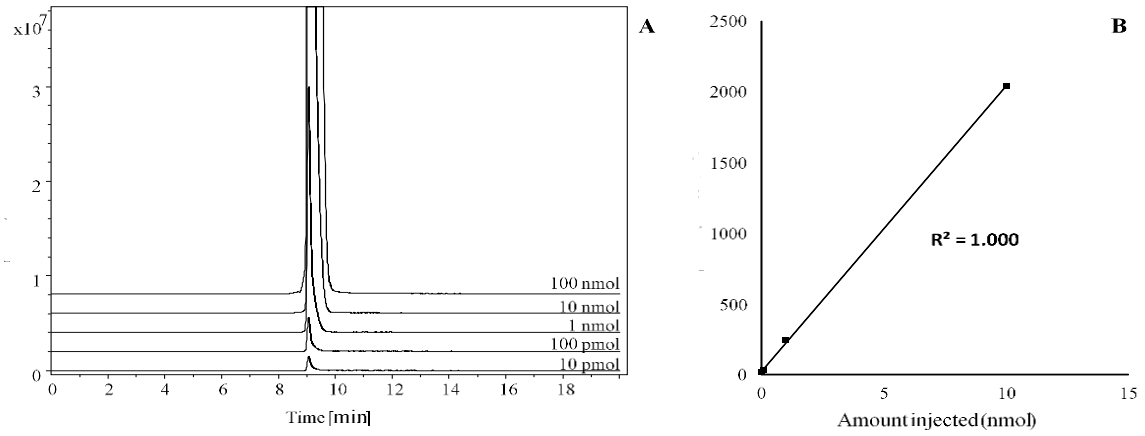


Figure 5-9 Evaluation of the dynamic response of FLB-*r*-dOG injected on column between 10 nmol and 10 pmol: A) extracted ion chromatogram of *m/z* 947.8, B) plot of the concentration *versus* peak area at *m/z* 947.8.

The signal response was found to be linear, with a coefficient of variation of 1.000 over four orders of magnitude (10 nmole to 10 pmole on column) as shown in **Figure 5-9**B. Although the tested range also included a level at 100 nmole it was found that the coefficient of variation was slightly lower, suggesting that the detector signal may be non-linear at the upper range and that polynomial fitting would be more appropriate over the extended concentration range.

The native synthetic diether was tested over the same concentration range. It was not detected at 10 pmol level and only a small peak was detected at 100 pmole injection on column (**Figure 5-10**A). The *S/N* for the latter was calculated to be approx. 5, thus suggesting that a LOD of *c*. 60 pmol and a LOQ of *c*. 0.2 nmole injected on column. The signal response was linear over three orders of magnitude (**Figure 5-10**B)

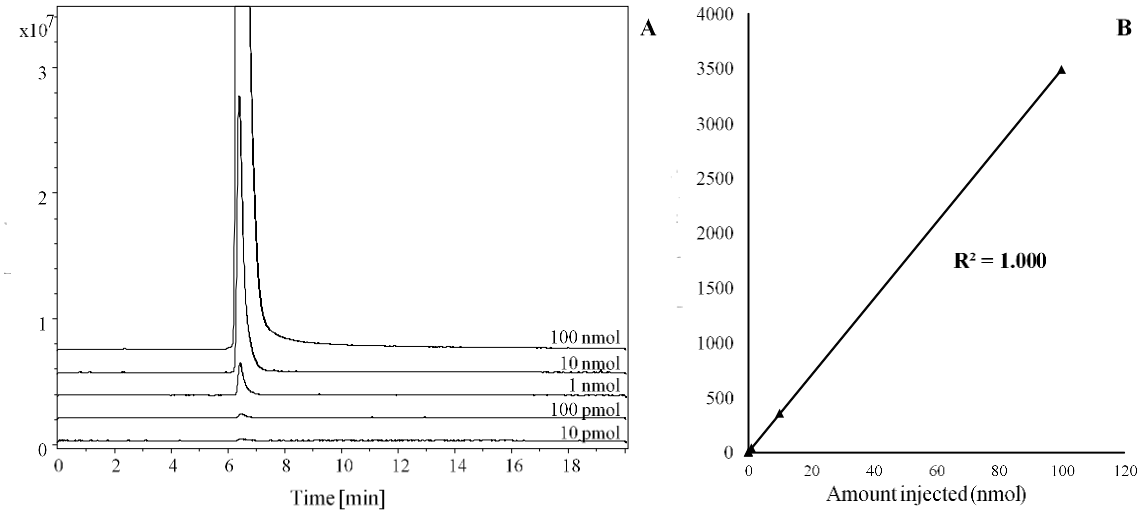


Figure 5-10 Evaluation of the dynamic response of *r*-dOG injected on column between 100 nmol and 10 pmol: A) extracted ion chromatogram of *m/z* 597.5, B) plot of the concentration *versus* peak area at *m/z* 597.5.

The APCI-MS LOD values of the native and the FLB derivative of the diether *r*-dOG indicate that the chemical modification resulted in signal intensity improvement of two orders of magnitude. This is a rather significant improvement, attesting to the benefits of the derivatisation and indicating that this might benefit the analysis of hydroxyl lipids by enabling the use of considerably less material and leading to detection of compounds that might have been missed due to the constraints of the LOD of the existing methods.

Evaluation of the HPLC-MS behaviour of cholesterol and its derivative was performed over the same range as the synthetic diether. The cholesterol derivative was not detected in the lowest level tested (**Figure 5-11**A) and the calculated *S/N* for the 0.1 nmol level was approx. 5.5. This suggests that the LOD is *c*. 60 pmol and the limit-of-quantification is *c*. 0.2 nmol. This is considerably higher than the values for the diether derivative and suggests that the response of the FLB derivative is dependent on the properties and behaviour of the particular class of hydroxyl lipid. The dynamic response was linear over four orders of magnitude (R^2 = 1.000 between 100 nmol and 0.1 nmol) as shown in **Figure 5-11**B.

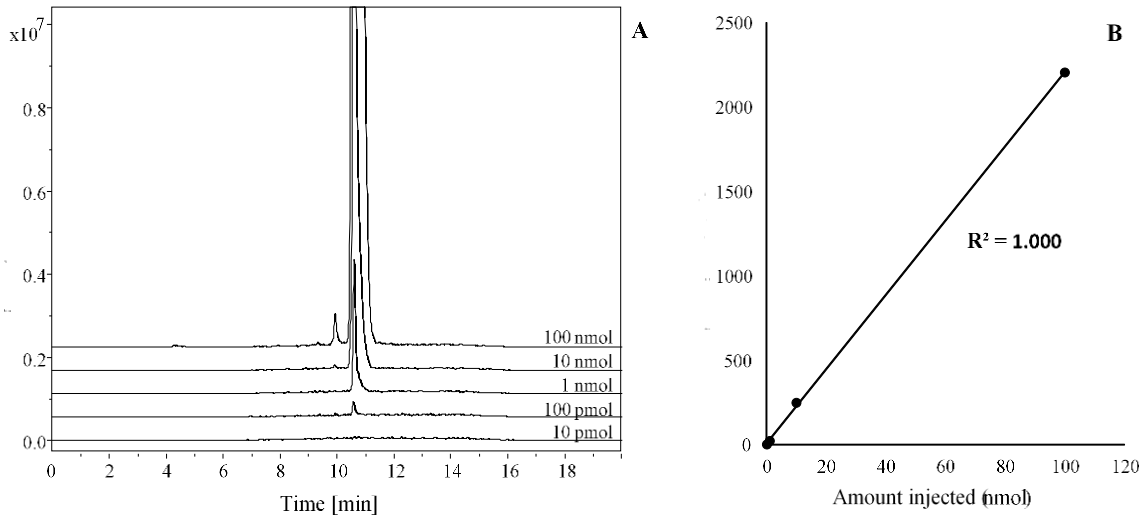


Figure 5-11 Evaluation of the dynamic response of FLB-cholesterol injected on column between 100 mmol and 10 pmol: A) extracted ion chromatogram of *m/z* 737.5, B) plot of the concentration *versus* peak area of the *m/z* 737.5.

Conversely, the native cholesterol was not detected below the 1 nmol level (see **Figure 5-12**A) and the calculated *S/N* was 6, which indicated that the LOD and LOQ are at approx. 0.5 nmol and 2 nmol on column, respectively. These values agree with the levels observed in APCI-MS detection (Raith *et al.*, 2005; Khajuria *et al.*, 2007). The signal response was linear between 1 pmol and 100 pmol on column (R^2 = 1.000, **Figure 5-12**B). Notably, the limits of detection of the FLB derivatives observed in HPLC-APCI-MS are comparable to those from GC-MS analysis (Son *et al.*, 2014; Chen *et al.*, 2015)

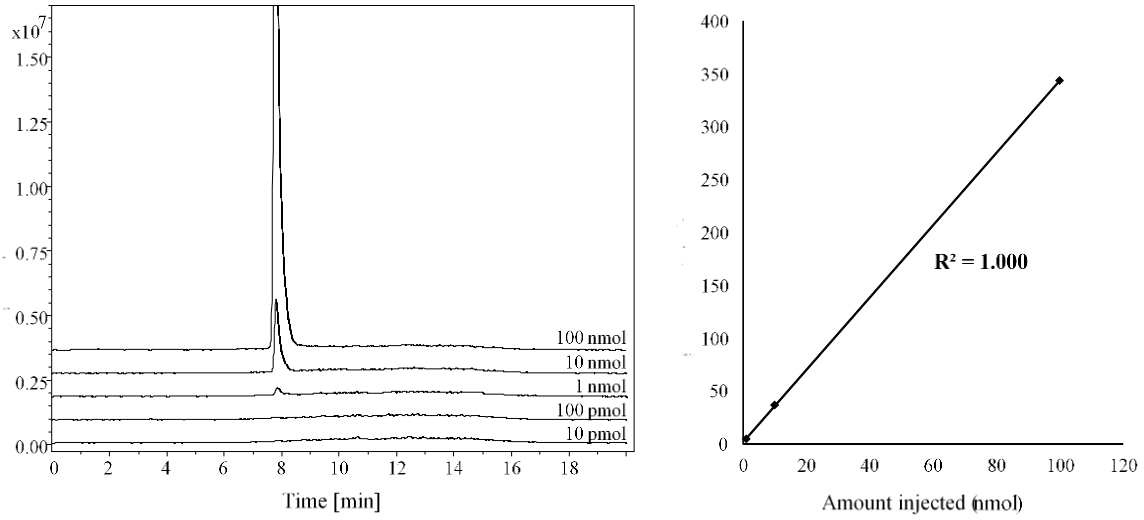


Figure 5-12 Evaluation of the dynamic response of cholesterol injected on column between 100 mmol and 10 pmol: A) extracted ion chromatogram of *m/z* 369.2, B) plot of the concentration *versus* peak area of the *m/z* 369.2.

The calculated LOD and LOQ of cholesterol and its derivative clearly shows that an improvement of one order of magnitude was accomplished as result of derivatisation. Notably the LOD for the diether lipid was found to be over two orders of magnitude lower (200 times) than for cholesterol, suggesting that the former might have higher ionization efficiency.

Notably, the response factor for the diether derivative is approx. 10 higher than that of the cholesterol derivative as is indicated by the much steeper slope of the linear fitting of the standard curve shown in **Figure 5-13**A.

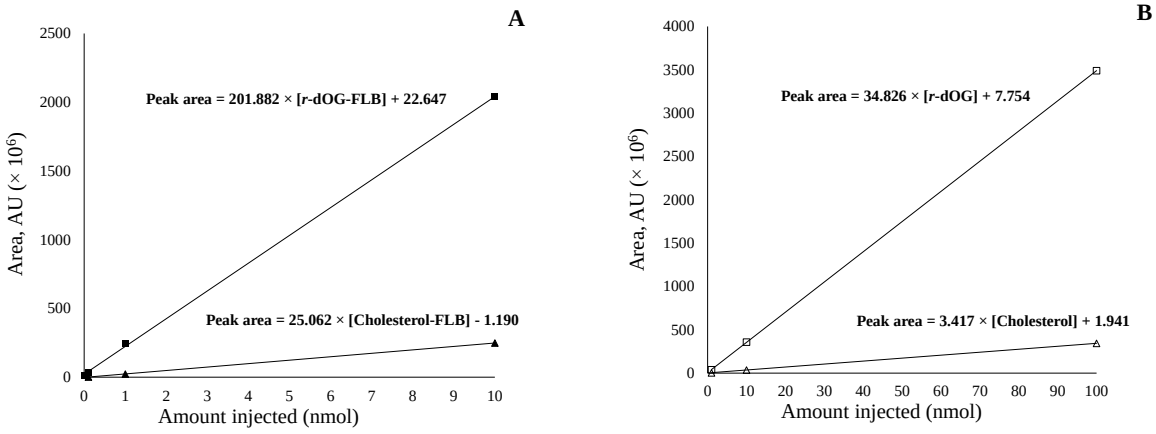


Figure 5-13 Standard curves for amount injected *versus* peak area: A) *r*-dOG and cholesterol modified with Fmoc-Lys-Boc derivatives and B) native lipids.

The same relation is true for the native diether and cholesterol: the response factor is around an order of magnitude higher for the former (**Figure 5-13**B). It is not apparent as to why the improvement in signal intensity was much higher for the derivative of the diether lipid than for cholesterol, , though given that the instrument was tuned for the individual compounds independently, this might suggest that the observed behaviour is related to an inherent property of the native lipid. In order to investigate the ionisation efficiencies of the FLB modified lipids more rigorously, testing of a broader range of lipid compounds is necessary.

Some consideration was given to the fact that the derivative of cholesterol undergoes a certain degree of fragmentation, as was indicated by the presence of fragment ions relating to the derivative and to complete loss of the derivatising group (FLB). Thus, extracted ion chromatograms which included the minor ions were constructed to assess whether this would improve the detection levels. Unfortunately, it was found that the improvement attributed to such operation was very modest and accounted to less than 10 %.

5.4.4 HPLC-FLD-UV characterisation of Fmoc-lysine(Boc) derivatives of *r*-dOG and cholesterol

Evaluation of the spectrophotometric and fluorescent properties of the derivatives was carried out using the same chromatographic conditions used for evaluation of their HPLC-MS properties (Chapter 2.7.1 and 2.7.2). Assessment of dynamic range and method sensitivity was performed using a series of dilutions of the derivatives of each of the lipids in a range of 100 nmol to 10 pmol injected on column. The HPLC detectors were connected in series such that the column effluent would first enter the UV detector before passing into the fluorescence detector, thus there is a slight retention time difference of the analytes as they were detected.

The lowest concentration for detection of the peak for the FLB-*r*-dOG derivative by UV was 1 nmol on column (**Figure 5-14**A) with *S/N* value of 14. This suggests that the LOD and LOQ values for this detection technique are *c*. 250 pmol and 750 pmol of analyte on column, respectively. The dynamic response range was found to be excellent over three orders of magnitude with an R^2 of 1.000 (**Figure 5-14**B).

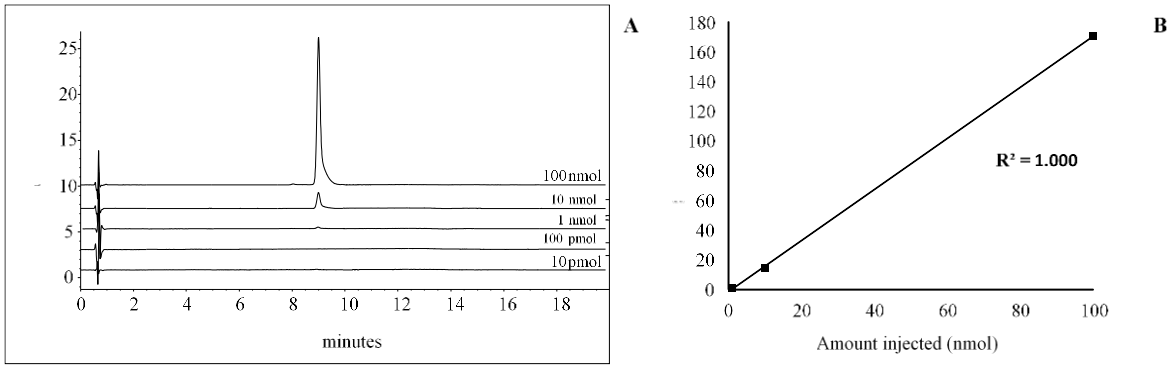


Figure 5-14 HPLC-UV chromatograms of FLB-*r*-dOG derivative (A), and plot of the concentration *versus* peak area (B).

Evaluation of the signal response in the fluorescence detector revealed that the diether derivative was not detected below 0.1 nmol on column (**Figure 5-15**A**)**. The signal-to-noise value for this level was 6, thus suggesting the LOD and LOQ values of *c*. 50 pmol and 140 pmol, respectively. These values clearly show that fluorescence detection is superior to the UV detection mode, giving improvements in LOD and LOQ of around 50 fold. The peak area response showed a very high correlation over the range 0.1 nmol to 100 nmol of analyte on column (R^2 = 1.000, **Figure 5-15**B).

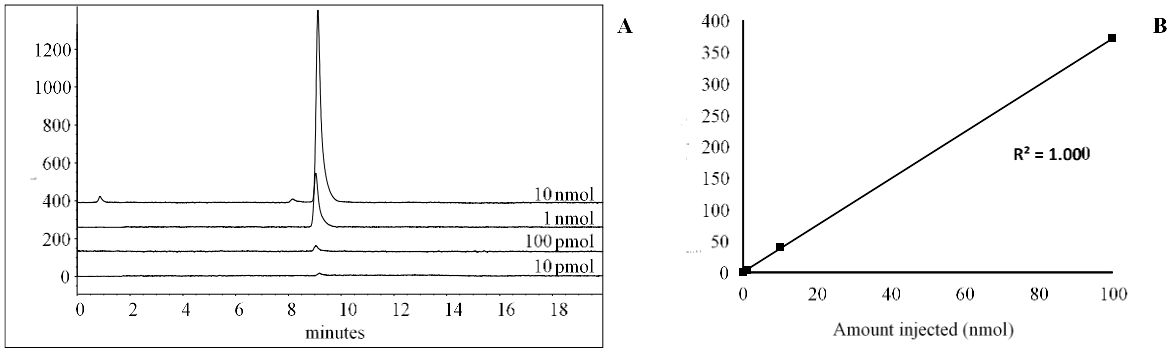


Figure 5-15 HPLC-FLD chromatograms of FLB-*r*-dOG derivative (A), and plot of the concentration *versus* peak area (B).

Although the chromatogram for the 10 pmol level shows a low level response it was found that the integrated peak area at this level was very unreliable and out of trend due to the baseline noise. Notably the calculated *S/N* value for this level would be below the critical value of *S/N*, and as such this level was rejected and considered as below the detection limit.

The UV detection performance of the FLB modified derivative of cholesterol was comparable to that of the diether, such that no peaks were detected for the two lowest levels tested, 10 pmol and 100 pmol of analyte on column (**Figure 5-16**A). The *S/N* value of the peak observed at 1 nmol level was 7, which suggests that the appropriate LOD and LOQ, are at *c.* 0.5 nmol and 1.5 nmol of analyte on column, respectively.

The calculated detection levels are only two times higher than the values calculated for the diether derivatives, suggesting that the UV detection response of both of the derivatised hydroxyl lipids are comparable and independent of the lipid molecule. This is an important observation as it indicates the possibility of using a derivative of a closely related structure as a reference standard when the native analyte is not available commercially or availability is limited by high cost.

The dynamic range for the FLB derivative of cholesterol was found to be linear over three orders of magnitude (between 1 nmol and 100 nmol on column) and showed excellent correlation between peak area and amount injected (R^2 = 1.000, **Figure 5-16**B).

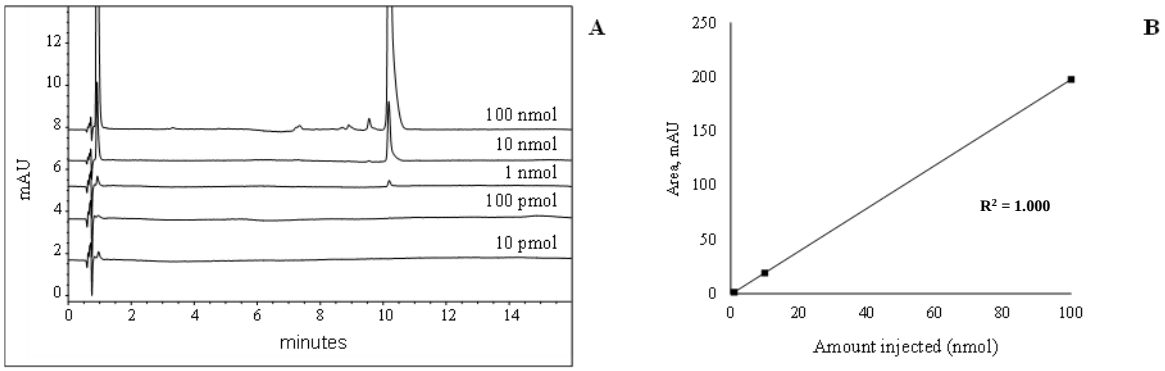


Figure 5-16 HPLC-UV chromatograms of FLB-cholesterol derivative (A) and plot of the concentration *versus* peak area (B).

Evaluation of the chromatograms recorded under fluorescence detector revealed that the analyte was not detected at less than 100 pmol injected on column (**Figure 5-17**A). The signal-to-noise (*S/N*) at this level was 4, which gives the LOD as 75 pmol on column. The calculated limit of quantification (10 × *S/N*) was found to be 250 pmol. The relation of the amount injected the peak area was linear over four orders of magnitude (0.1 nmol to 100 nmol on column) with a very high degree of correlation (R^2 = 1.000, **Figure 5-17**B).

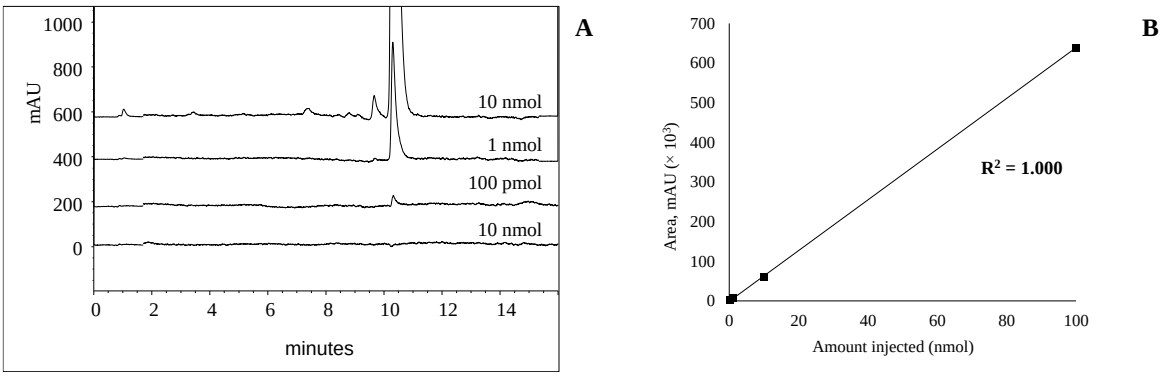


Figure 5-17 HPLC-FLD chromatograms of FLB-cholesterol derivative (A) and plot of the concentration *versus* peak area (B).

The investigation of the spectrophotometric and fluorescent properties of the diether and cholesterol modified using FLB revealed that their response in these detection modes is comparable to that observed for the selected *N*-Fmoc amino acid derivatives of 1-octadecanol described in Chapter 4.4.2. The LOD values for the FLB derivatives for UV detection were similar for the diether (*r*-dOG) and cholesterol at *c.* two times lower than for the derivatives of 1-octadecanol. The Fmoc-lysine(Boc) amino acid derivatives boasted modest improvement in the fluorescence detection of four and three fold for the diether and cholesterol, respectively, *versus* the derivatives of 1-octadecanol.

5.4.5 Tandem mass spectrometric characterization of the Fmoc-lysine(Boc) modified *r*-dOG and cholesterol derivatives

The dissociation pathways of the native *r*-dOG and cholesterol lipids as well as their FLB modified derivatives were studied by tandem mass spectrometry using an ion trap MS. Solutions of the individual analytes were infused into the HPLC column effluent *via* a syringe pump reflecting the mobile phase composition at the time of their respective elution. The ion trap MS was tuned to ensure the best conditions for the ionization of the analytes being studied. The tandem MS settings were set to adjust automatically to provide the best conditions for detection of the base peak selected for the subsequent dissociation.

The native diether lipid (*r*-dOG) was investigated first in order to understand its APCI-MS fragmentation behaviour and to be able to track its characteristic product ions during the tandem MS of the FLB derivative. Infusion of the solution of the native diether revealed a full MS spectrum where only one significant ion was observed, at *m/z* 597.5 (**Figure 5-18**). This ion corresponds to the protonated molecule of the diether lipid. The absence of any other significant fragment ions indicates high stability of the protonated molecule in the presence of the mobile phase.

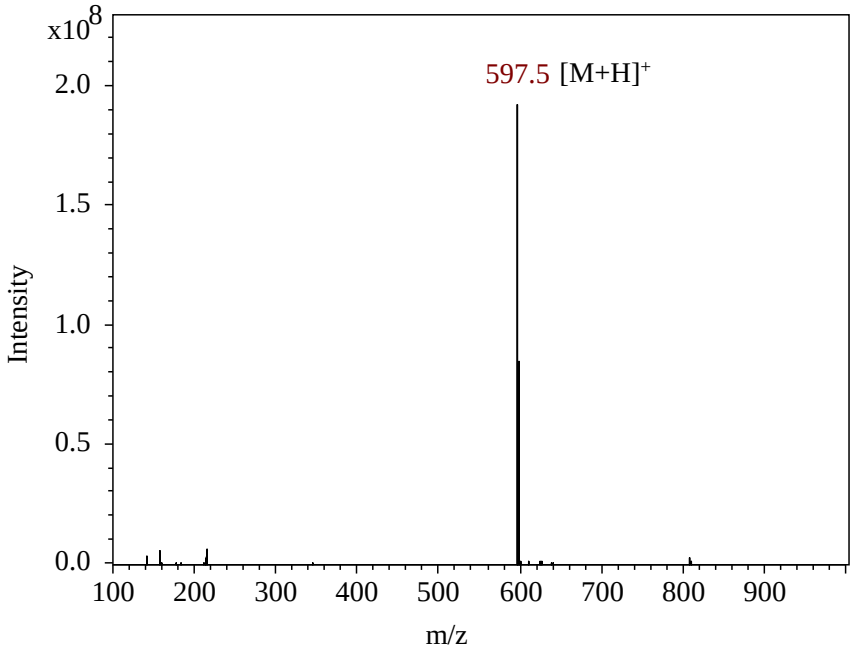


Figure 5-18 Full (+) APCI mass spectrum of the diether lipid (*r*-dOG).

Subsequent resonance induced dissociation of the base peak (protonated molecule) led to formation of a relatively high intensity product ion at *m/z* 345.2, which results from loss of one octadecyl chain from the protonated molecule (**Figure 5-19**A). The cleavage most likely occurs at the C2 position (*via* dissociation of bond labelled **c** in **Figure 5-20**) as suggested by similar observations during tandem MS studies of a similar structure: archaeol (Liu *et al.*, 2012). The relatively high intensity of this ion suggest that this type of bond dissociation is energetically preferred, giving a highly stable product ion.

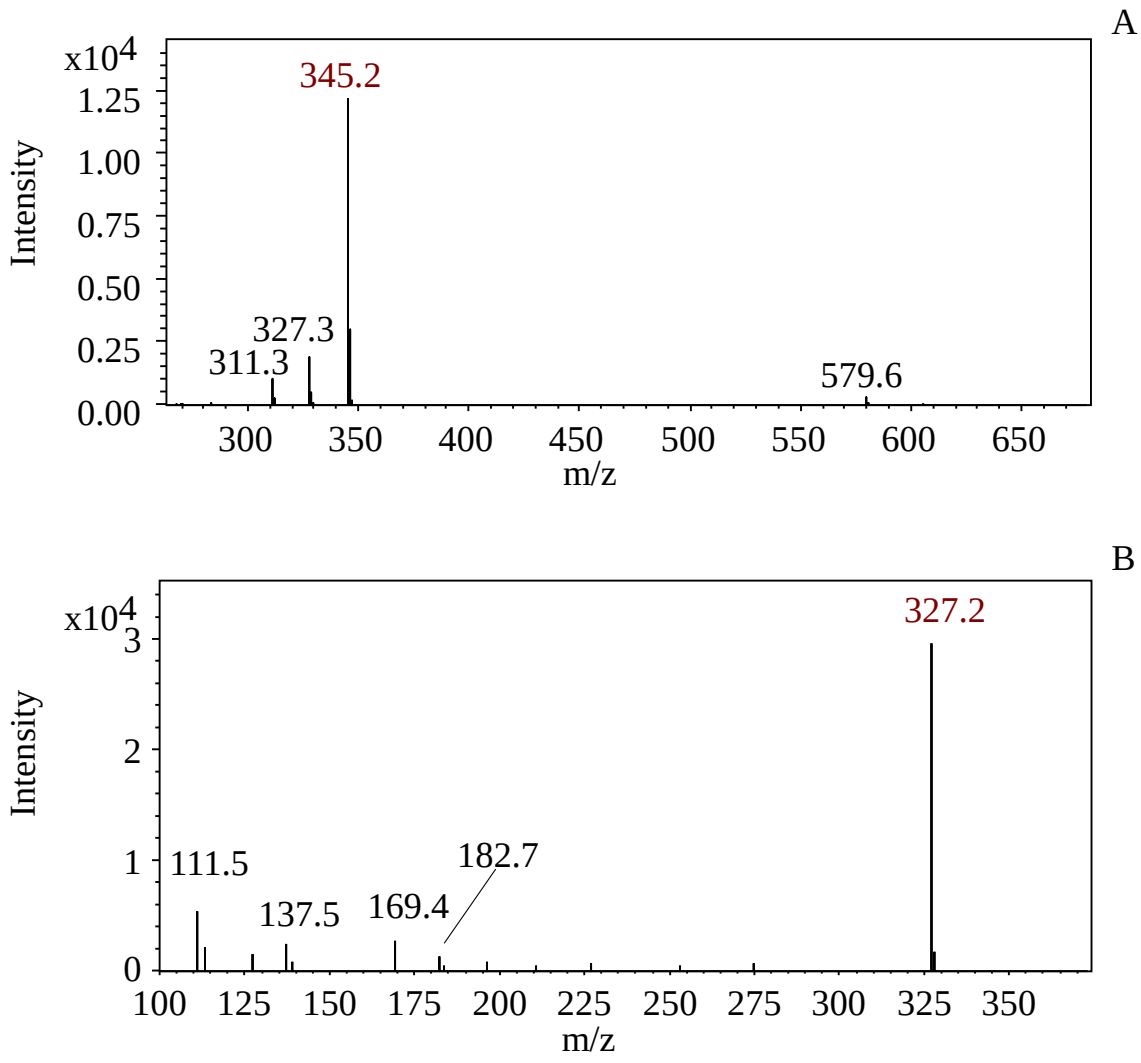


Figure 5-19 Tandem mass spectrum of the diether lipid *r*-dOG: A) MS^2 of the base peak at *m/z* 597.5, B) MS^3 of the base peak at *m/z* 345.2.

Lower intensity product ions were also observed in the MS^2 spectrum at *m/z* 579.6, *m/z* 327.3 and *m/z* 311.3. The first is 18 units lower than the protonated molecule, consistent with loss of one molecule of water from the terminal hydroxyl group (breaking of the bond labelled **a** in **Figure 5-20**). The product ion at *m/z* 327.3 is 270 units lower than the protonated molecule and results from loss of one molecule of 1-octadecanol (dissociation of bond **b**). Finally, the product ion at *m/z* 311.3, which is 286 units lower than the protonated molecule, is consistent with an aggregate loss of one 1-octadecanol moiety accompanied by loss of one molecule of water (breaking of bonds **a** and **b, Figure 5-20**). For the subsequent dissociation (MS^3) *m/z* 345.2 was selected (**Figure 5-19**B). The MS spectrum revealed an intense product ion at *m/z* 327.2, which suggest a loss of one molecule of water, being 18 Da lower than the precursor ion (dissociation of bonds **a** and **b**).

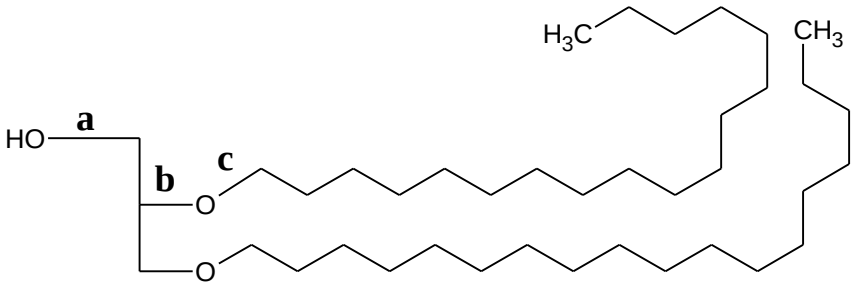


Figure 5-20 Structure of the diether lipid *r*-dOG with bonds that undergo dissociation during MS/MS studies indicated by the lower case letters (a-c).

Other product ions observed in the MS^3 spectrum, especially in the lower *m/z* region (between *m/z* 100 and *m/z* 200), are consistent with the dissociation pattern characteristic of hydrocarbons under electron ionisation conditions, i.e. *m/z* 182.7, *m/z* 169.4, *m/z* 137.5, and *m/z* 111.5. Further resonance dissociation studies of the ion at *m/z* 327.2 did not reveal any significant product ions that would be of practical use.

The full MS spectrum of the FLB-*r*-dOG derivative was dominated by a base peak ion at *m/z* 947.8, 100 Da lower than the protonated molecule, resulting from the characteristic loss of the Boc group. The formation of fragment ions revealed in the spectra of such derivatives was discussed in Chapter 5.4.2. The resonance induced dissociation of this ion gave rise to intense product ions in the MS^2 spectrum at *m/z* 903.8, *m/z* 751.7, *m/z* 708.7 (**Figure 5-21**A).

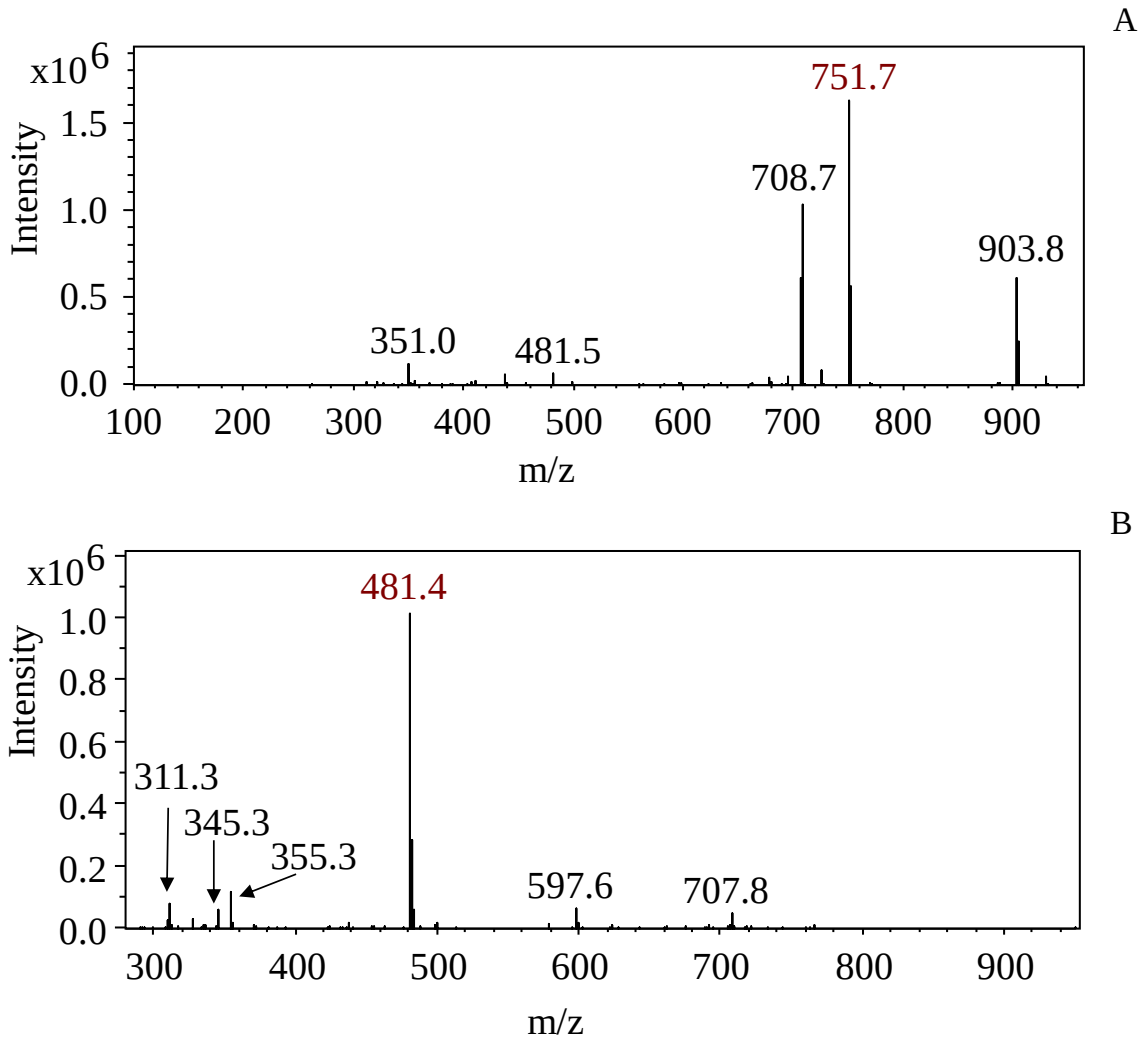


Figure 5-21 Tandem mass spectrum of the diether lipid *r*-dOG modified using FLB: A) MS^2 of the base peak at *m/z* 947.8, B) MS^3 of the base peak at *m/z* 751.7.

The most intense product ion at *m/z* 751.7 reflects a loss of 196 Da and most likely reflects a loss of the 9-fluorenemethanol *via* dissociation of bond **b** (**Figure 5-22**). This loss is sometimes observed during CID experiments of *N*-Fmoc protected amino acids as reported by Zhu *et al.* (2006). The second most intense product ion at *m/z* 708.7 is 239 Da lower than the precursor ion and results from simultaneous loss of the 9-fluorenemethanol and one molecule of NH_3 formally reflecting the dissociation of bonds **a** and **b** (**Figure 5-22**). Lastly, the product ion at *m/z* 903.8 is 44 Da lower than the precursor ion at *m/z* 947.8 was most likely formed by dissociation of bond **a** (**Figure 5-22**) resulting in a loss of a small neutral molecule comprising the amine group and two carbon atoms from the lysine alkyl side chain. Notably, subsequent dissociation of this ion gave rise to a single product ion of low intensity at *m/z* 708.7 (mass spectrum not included) discussed above. Considering the losses leading

to the formation of this ion, as discussed previously, it is likely that it is formed *via* a McLafferty rearrangement involving the carbonyl oxygen of the ester bond. The low abundance product ions at *m/z* 481.5 and *m/z* 351.0 were not assigned.

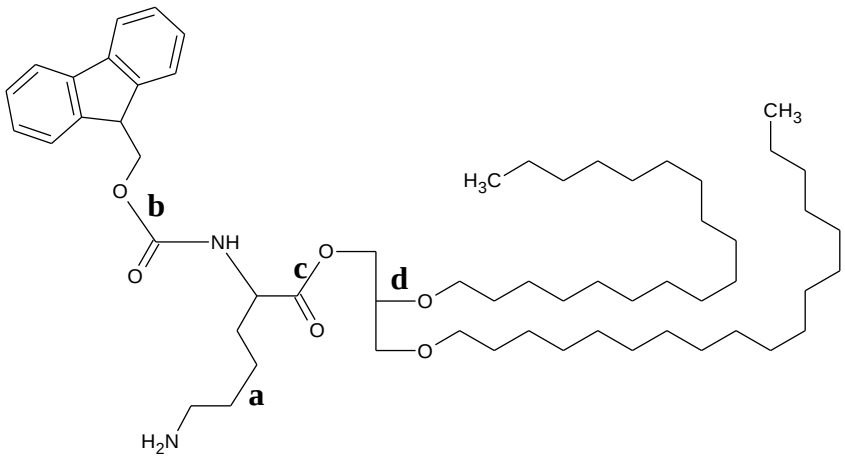


Figure 5-22 Structure of the diether lipid *r*-dOG modified using FLB with bonds undergoing dissociation indicated by the lower case letters (a-d).

Further resonance induced dissociation of the base peak ion at *m/z* 751.7, gave rise to an intense product ion at *m/z* 481.4 in the MS^3 spectrum (**Figure 5-21**B). This ion is 270 Da lower than the precursor ion and most likely represents loss of one molecule of 1-octadecanol as a result of dissociation of bond **d** (**Figure 5-22**). This type of loss was observed during the tandem MS studies of the native diether lipid discussed earlier in this subsection.

Other notable, although less intense, product ions in the MS^3 spectrum occur at *m/z* 707.8, *m/z* 597.6, *m/z* 355.3, *m/z* 345.3 and *m/z* 311.3. The first of those ions reflects loss of the Fmoc protecting group as has been discussed above and the second corresponds the protonated molecule of the native diether lipid being formed *via* cleavage of the ester bond and loss of the amino acid. The ions at *m/z* 345.3 and *m/z* 311.3 reflect losses observed in the tandem MS experiments of the native diether lipid, as discussed above. The low abundance ion at *m/z* 355.3 was not assigned. Further CID of the MS^3 base peak (*m/z* 481.4) did not reveal diagnostically useful product ions, thus suggesting that tandem MS studies beyond MS^3 reveal limited structural information.

Infusion of the solution of the FLB-cholesterol derivative into the HPLC effluent produced the full mass spectrum shown in **Figure 5-23**. The protonated molecule at *m/z* 837.4 was detected in very low abundance. The base peak ion at *m/z* 737.5 can be accounted for by loss of the butyloxycarbonyl moiety (Boc) from the protonated molecule.

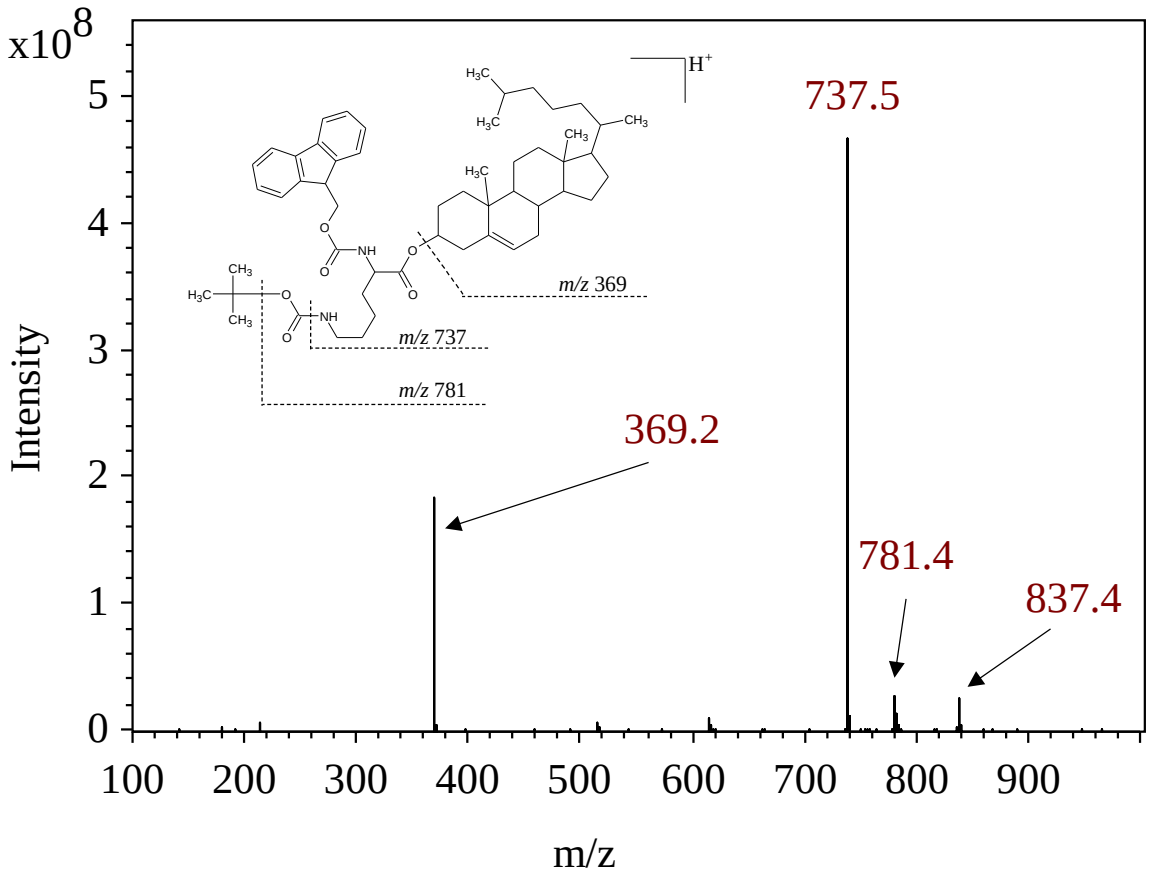


Figure 5-23 Full MS spectrum of cholesterol modified using FLB.

The low intensity fragment ion at *m/z* 781.4 is 56 Da lower than the protonated molecule and 44 units higher than the base peak ion at *m/z* 737.5. This fragment ion corresponds to loss of isobutylene as a neutral species from the protonated molecule, a loss that is often observed in *N*-Boc protected amino acids (see Chapter 3.3.2). The relatively intense (40 % relative abundance) fragment ion at *m/z* 369.2 is 468 Da lower than the protonated molecule, corresponding to the fragment ion observed in the mass spectrum of native cholesterol and reflecting cleavage of the ester bond and a loss of the amino acid moiety.

Resonance induced dissociation of the fragment ion at *m/z* 781.4 yielded an MS^2 spectrum showing an intense product ion at *m/z* 369.1 (**Figure 5-24**). Formation of this product ion, which is 412 units lower than the precursor ion, may be rationalised on the basis of the cleavage of the ester bond with a simultaneous loss of one molecule of water resulting in charge retention on the cholest-5-ene moiety. The ion at *m/z* 369.2 is present in full the MS spectra of both the cholesterol derivative (**Figure 5-23**) and the native cholesterol.

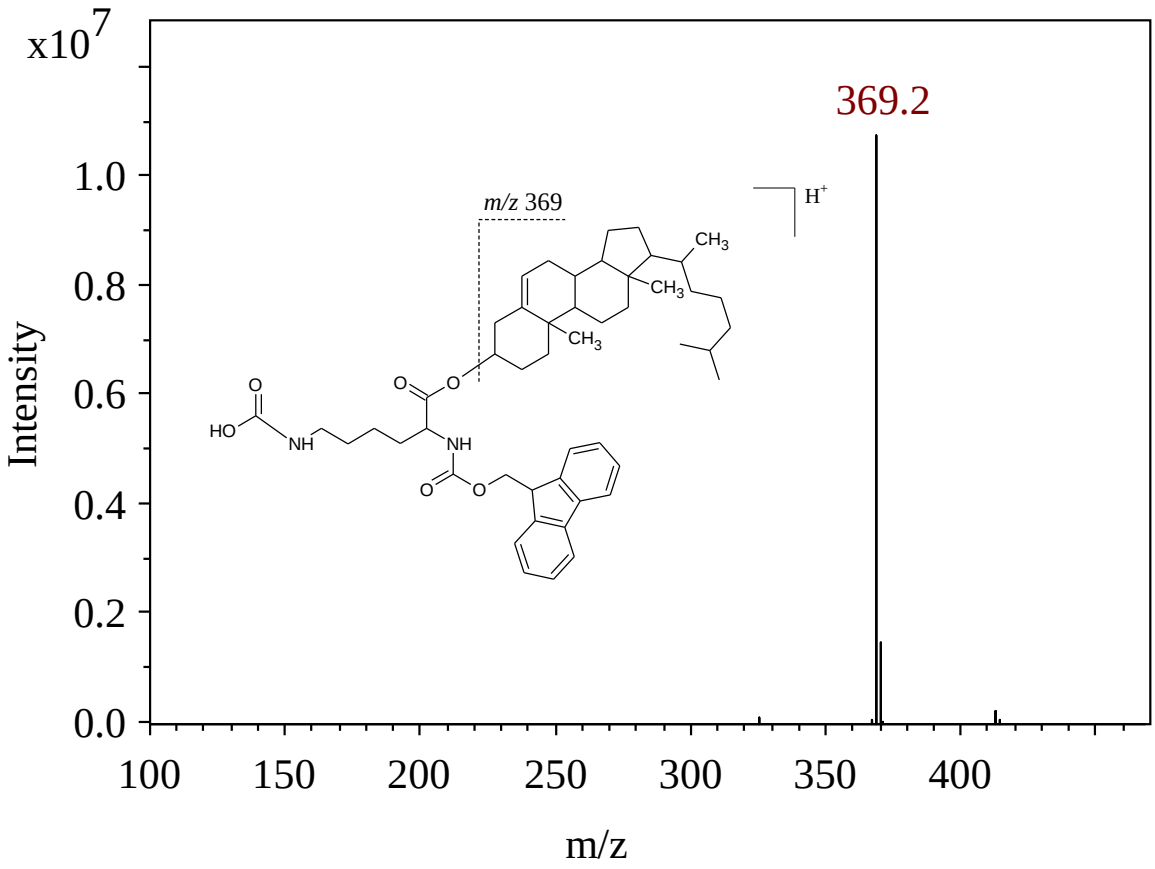


Figure 5-24 MS^2 spectrum of the fragment ion at *m/z* 781.7 of the cholesterol modified using FLB.

The base peak ion at *m/z* 737.5 observed in the full MS of cholesterol was also subjected to CID and gave rise to the product ion at *m/z* 369.2 (**Figure 5-25**). This product ion has the same *m/z* ratio as the ion observed in the full MS spectrum of the FLB-cholesterol and in the MS^2 spectrum of the ion at *m/z* 781, thus suggesting that the cleavage occurred at the ester bond with charge retention by the cholest-5-ene moiety. The absence of other product ions indicates that the charge is very well stabilised within the ion.

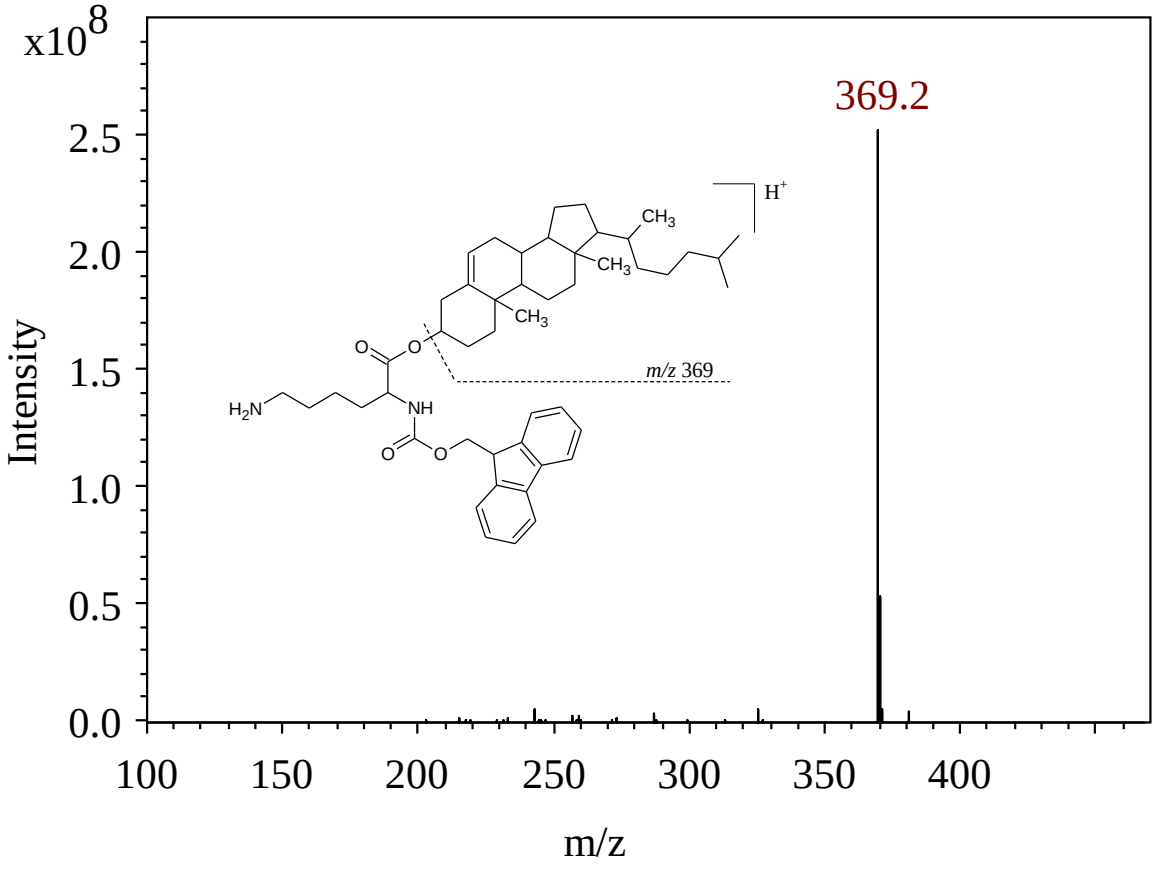


Figure 5-25 MS^2 spectrum of the fragment ion at *m/z* 737.5 of the cholesterol modified using FLB.

Subsequent resonance enhanced dissociation of the base peak at *m/z* 369.2 yielded the MS^3 spectrum shown in **Figure 5-26**. Only a few product ions were observed, at *m/z* 325, *m/z* 243, *m/z* 179, *m/z* 325 and *m/z* 130 and structures of these ions were not elucidated. The *m/z* ratio differences between the ions and the higher intensities of lower mass product ions suggest that the dissociation occurs *via* losses of large structural motifs and that the charge may be delocalized and stabilized differently from that which is commonly observed during CID studies of the native cholesterol.

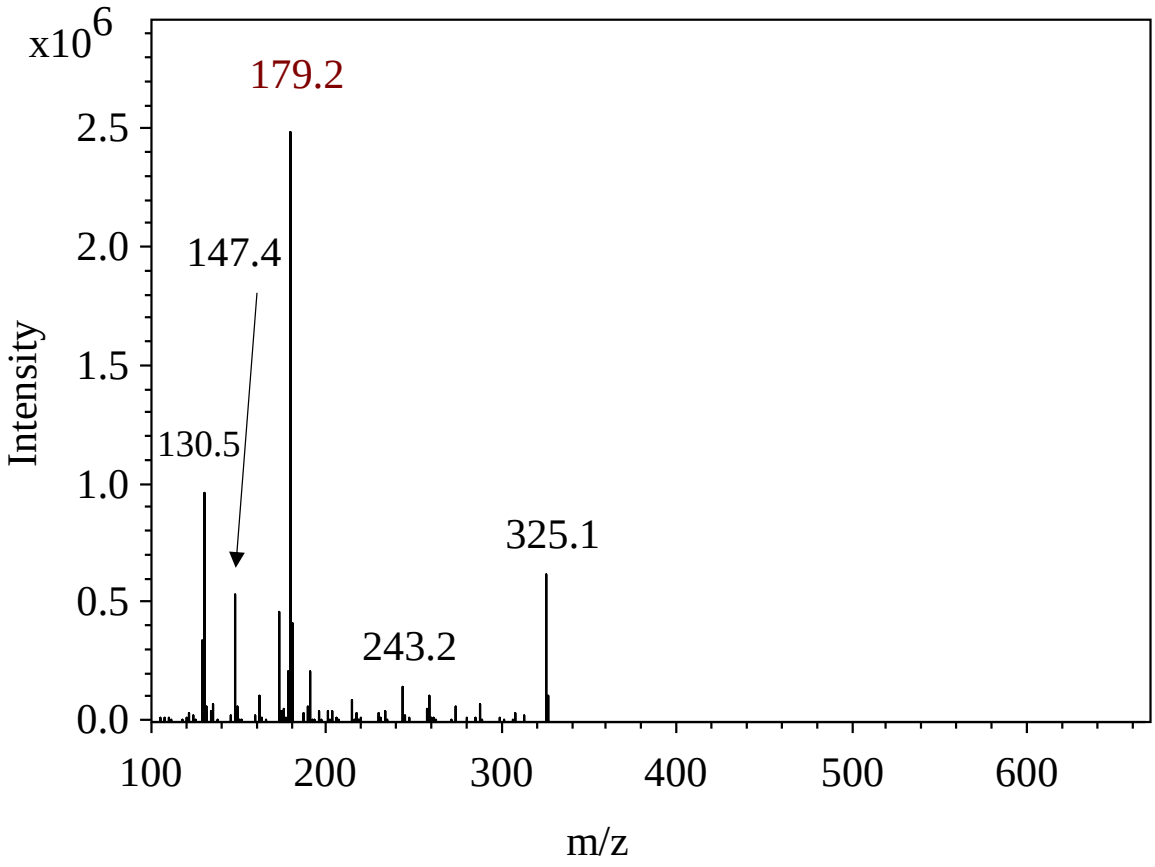


Figure 5-26 MS^3 spectrum of the fragment ion at *m/z* 369.2 of the cholesterol modified using FLB.

The stark differences in the dissociation pathways can be visualised by comparing the spectrum with the resonance induced dissociation spectrum of the base peak ion at *m/z* 369.2 formed during ionization of cholesterol by APCI-MS (see **Figure 5-27**). The presence of a multitude of product ions separated by 14 Da resembles the losses in hydrocarbon fragmentation under EI conditions, indicating that the dissociation progresses to the stage where the relative stabilities of the product ions formed are similar, with no discernible preference for any given pathway. Notably, the product ions at *m/z* 243.2 and *m/z* 147.4 can be seen the spectrum of both native and derivative of cholesterol, suggesting some degree of similarity in their provenance.

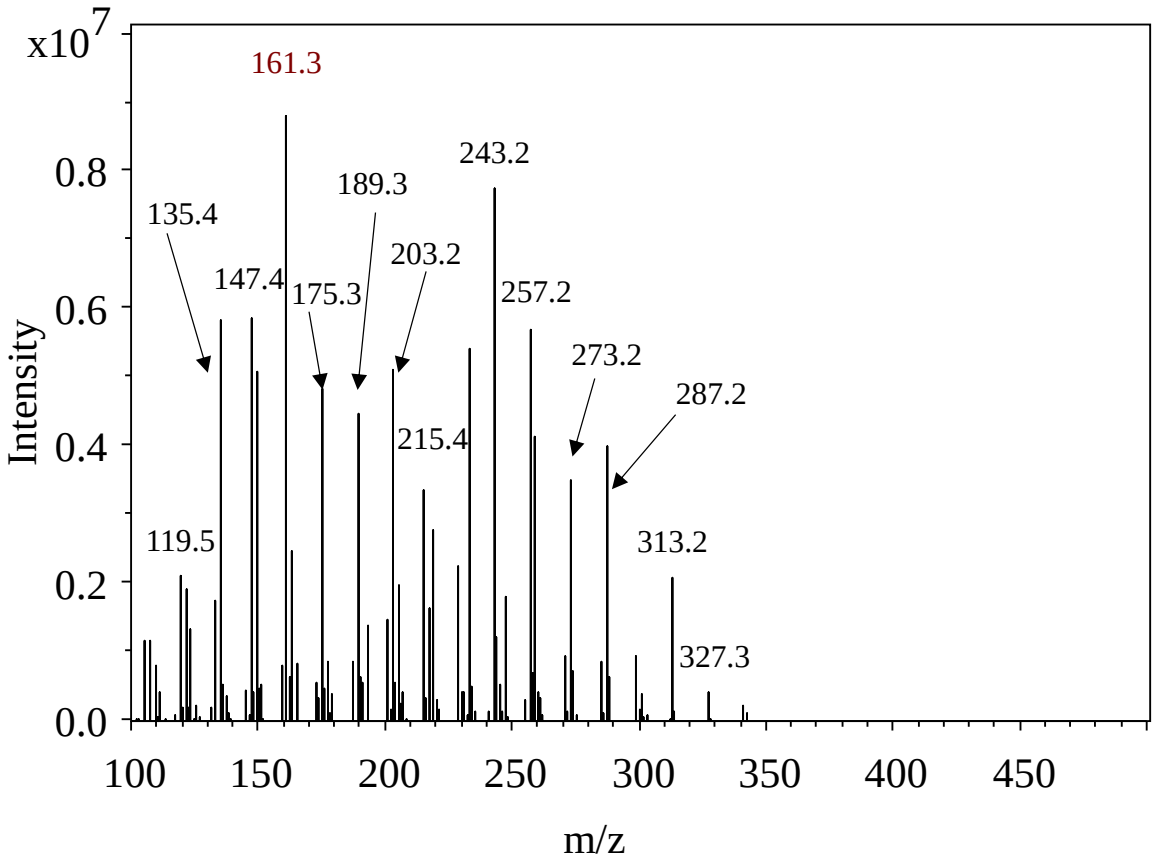


Figure 5-27 MS^2 spectrum of the fragment ion at *m/z* 369.2 of the cholesterol.

Although, the aspect of differences in the CID behaviour of the same product ion at *m/z* 369.2 of the native cholesterol and its FLB derivative is very intriguing and might have yielded novel valuable information, such investigation was outside the scope of this research and was not explored further.

5.4.6 Conclusions

The synthetic *rac*-1,2-di-octadecyl glycerol diether (*r*-dOG), resembling the naturally occurring archaeol glycerol diether, was synthesized in very good yield and its chromatographic and mass spectrometric behaviour were characterized. It was also demonstrated that derivatives of various lipids can be successfully prepared using Fmoc-lysine(Boc) amino acid in excellent yields. Notably, isolated yields above 98 % were achieved for both the FLB-*r*-dOG and FLB-cholesterol derivatives. The yields compare very well with the yields of *N*-Boc and *N*-Fmoc protected amino acids of octadecanol discussed in Chapter 3.2.3 and Chapter 4.3, which confirms the suitability of the derivatisation methodology for modifying various hydroxyl lipid classes.

The chromatographic performance of the derivatives was evaluated and it was shown that their chromatographic behaviour is superior to that of the native compounds. Additionally, introduction of an excellent chromophore (Fmoc group) allowed for the derivatives to be monitored under ultraviolet and fluorescence detection, greatly expanding the detection techniques for this type of lipid species. The derivatives each eluted in a sharp and narrow band, which results in increased limits-of-detection and limits-of-quantification by comparison with the native species. The limit of detection for HPLC-UV detection was found at 250 pmol and 500 pmol of the analyte on column for FLB-*r*-dOG and FLB-cholesterol, respectively, whereas for HPLC fluorescence detection the values were 50 pmol for the former and 75 pmol for the latter. The LOD for mass spectrometric detection of the FLB modified *r*-dOG and cholesterol were found to be 0.3 pmol and 60 pmol of the analyte on column, respectively. This represents a tremendous improvement in detection: two orders of magnitude. Notably, the detection of the FLB-*r*-dOG increased 200 times compared with the native diether, whereas the increase following derivatisation of cholesterol is around 10 times. The LOD values for MS and fluorescence detection are comparable for the cholesterol derivative, while the MS detection was far superior to the fluorescence (100 time increase) for the diether derivative.

The tandem MS behaviour of the FLB-cholesterol derivative reveals that it follows different dissociation pathways to the native cholesterol. Loss of the amino acid moiety with charge retention on the lipid moiety leads to a suite of pathways through which structure elucidation studies might benefit, with improvement to the ability to discern between closely related hydroxyl sterols.

Chapter 6 Preparation of derivatives of lipids from Archaeal extract and soil deposit

6.1 Introduction

At present the majority of analyses of sedimentary archaeal GDGTs for palaeothermometry via measurement of the TEX_{86} index are conducted using normal phase separation with mass spectrometric detection (NP-HPLC-APCI-MS), relying on the method developed by Hopmans *at al.* (2000) or its later modification by Schouten *et al.* (2007). The former method relies on the use of a standard internal diameter (ID) amino modified silica column (4.6 mm × 250 mm) with column packing beads of 5 μm, while the latter employs a smaller ID column (2.1 mm × 150 mm) with cyano modified silica packing of 3.5 μm bead size. The degree of separation achieved using both columns is comparable, the smaller ID column requires less solvent owing to the slower flow rate of the mobile phase. The separation of the analytes is governed by the differences in the extent of their adsorption onto the stationary phase (see discussion in Chapter 3) and has the ability to partially resolve the crenarchaeol and its regioisomer. The analytes are eluted using a mobile phase comprising hexane and isopropanol, whereby the proportion of isopropanol throughout the gradient program is relatively small (gradient from 1 % to 2 %). Although, with the use of modern HPLC pumps, such minimal changes in the mobile phase are not difficult to achieve during a gradient program, concerns have been raised relating to the system back pressure affecting the separation (Rattray *et al.*, 2013). Despite the palaeotemperature-relevant lipid cores eluting early in the chromatographic run (>15 min), the analytical run is lengthy. Thus, elution is continued for a total of approx. 50 min, followed by column equilibration for an additional 10 min. Due to the nature of the normal phase retention mechanism the more polar components of the sample matrix are retained on the column and have to be removed using a mobile phase of considerably higher polarity. This is achieved by flushing the column in the direction opposite to the normal flow through the column (back flushing), and is required approx. every 10 injections. The recurrent re-plumbing of the analytical column this requires not only adds to the analysis time, but also could affect column performance and lead to premature column failure.

Since the publication of the revised method for the determination of the TEX_{86} tetraether lipids (Schouten *et al.*, 2007) there have been numerous attempts to improve the resolution of co-eluting peaks. Those studies have been conducted in order to improve peak integration (and limit variations in temperature reconstruction) and to allow discovery of new structures that may have been obscured by poor peak resolution. The majority of the development work seems to have concentrated on the exploration of alternative column chemistries: amide and

silica (De Jonge *et al.*, 2013; Wörmer *et al.*, 2013), or use of reduced column packing particle size (< 2 μm) (Zech *et al.*, 2012; Becker *et al.*, 2013; Wörmer *et al.*, 2013). These attempts, however, resulted in only minor improvements in analyte resolution, partially due to the reliance on the same mobile phase solvents, mainly hexane and isopropanol. Other researchers tried using multiple columns connected in series to improve separation: two amide columns (Becker *et al.*, 2015), two silica columns (Hopmans, Schouten and Sinninghe Damsté, 2016) or even four silica columns (Yang *et al.*, 2015). Although, improved resolution of tetraether lipid cores was reported in all three cases, the markedly better results were observed using two silica columns, most likely due to a combination of two aspects: the size of the column packing (1.7 μm) and the packing type (solid core particles) (Hopmans, Schouten and Sinninghe Damsté, 2016).

Importantly, while the degree of chromatographic improvement is appreciable, the analytical protocols mentioned above have failed to eliminate the concerns relating to solvent system derived instability in backpressure and the necessity of repeatedly remove the highly polar matrix components retained on column. Moreover, the requirement of using multiple analytical columns is not only more expensive but also entails longer analytical runs. Alternatively, reversed phase separation methods have also been explored. Reversed phase chromatography (RP), due to its mechanism of analyte retention, offer a means of alleviating the inherent limitations of the normal phase separation. Notably, two reports provide thorough discussion of the literature relating to the use of C18 columns in the separation of the archaeal tetraether lipids. Wörmer *et al*. (2013) tested five C18 columns, both normal ID and particle size and UHPLC, in screening of standards as well as environmental samples. The method used a methanol - isopropanol mobile phase system (adjusted using HCO_2H and NH_3) in a gradient program. The authors noted an improved separation of the intact polar lipids and a considerable reduction of the background when using an UHPLC column (2.1 mm x 150 mm, 1.7 μm) compared with NP methods. Although some key parameters, such as the effect of column temperature, injection solvent, and buffers strength, were evaluated, the behaviour of the archaeal lipid cores was not discussed in detail. The same year Zhu *et al*. (2013) demonstrating considerable improvement in the separation and profiling of the bacterial and archaeal lipid core moieties using a novel C18 packing material. The analytes were separated in a non-aqueous reversed phase (NARP) mode using a combination of methanol and isopropanol (adjusted using HCO_2H and NH_3) in a gradient profile. Even though the separation was based on principles similar to those reported by Wörmer, the separation of analytes of broad range of polarities was accomplished due to a well-devised

gradient program. Although the authors claim comprehensive screening capabilities of the method; it appears to be geared more towards the intact polar lipids than the lipid cores, due mainly to the lipid cores of the TEX_{86} index eluting very late in the gradient program (between 55 – 65 min). Although, the separation of the index components that was achieved is good, the method failed to separate the crenarchaeol regioisomer, an important component of the TEX_{86} index. To circumvent the inability to assess the abundance of the regioisomer and still be able to infer the sea surface water temperature (SST) the TEX_{86}^{L} can be used (Kim *et al.*, 2010). Neither of the reversed phase methods discussed above suffers from the need of column back flushing, mainly because the highly polar sample matrix components elute early in the chromatographic run. Additionally, the column equilibration time between injections is much shorter than in NP, which is common for many RP methods.

The development of the reverse phase chromatographic methods for the separation of the archaeal tetraether lipids has proved to be cumbersome to many researchers. The difficulties pertain mainly to the extremely strong retention of the lipid cores on the stationary phase, and a loss of analyte resolution under higher eluotropic strength of the mobile phase. As such, the trade-off has been between shorter run time and acceptable analyte separation.

Distributions of the GDGTs in environmental samples vary substantially between the different sedimentary settings, with dominance of caldarchaeol and crenarchaeol often being very pronounced compared with other GDGT structures. This, in turn, necessitates injecting large volumes of sample material in order to detect the components that contribute in the TEX_{86} proxy. Elsewhere, especially in some ancient geological sediments, detection of the GDGTs may be impossible due to their very low abundance, rendering palaeotemperatures of from the distant past inaccessible.

Given the considerable potential benefits, including improved signal intensity, good chromatographic performance and excellent response in the fluorescence detection, of the *N*-protected amino acid derivatives evaluated in earlier chapters (see Chapter 3, 4 and 5) the preparation of the derivatives of tetraether lipid cores of cellular or environmental provenance seemed worthy.

6.2 Aims

The main objectives of the work discussed in this chapter were:

- to prepare the Fmoc-lysine(Boc) (FLB) derivatives of archaeal GDGT lipid cores sourced from microbial cellular material and extracts from sedimentary deposits and to evaluate the potential benefits in their analysis;
- to develop a novel RP separation method permitting profiling and characterisation of the lipid cores in their native as well as derivatised form and hence, to allow their characterization;
- to evaluate the extension of the application of the derivatization protocol to other lipid classes and to evaluate its potential benefits as an alternative to gas chromatographic (GC) profiling of sterols and alkanols in environmental and archaeological research.

6.3 Preparation and evaluation of the *N*-Fmoc lysine(Boc) derivatives of lipid cores derived from archaeal cellular material

6.3.1 Development of a novel reversed phase chromatography separation of GDGT lipid cores

The initial development work was performed on an Ace UltraCore SuperC18 column (3.0 mm × 150 mm, 2.5 μm), which is similar to that used by Zhu *et al.* (2013). However, this is a medium bore column, filled with a novel solid core particle packing material. The reasons for selecting the column were: first, the solid core technology often offers a better peak resolution due to the dynamics of analyte - stationary phase interactions, and secondly, the size of the column helps to maintain a low system backpressure and permits faster equilibration times. The column was tested under the conditions reported by Zhu *et al*. (2013) with ion trap (+)APCI-MS detection by analysing a GDGT lipid core fraction from aerobic thermoacidophilic crenarchaeol *Sulfolobus acidocaldarius* (for preparation see Chapter 2.3). The reason for using the lipid material from this microorganism is that its tetraether lipids distribution has been well characterised in the literature (De Rosa, Gambacorta, *et al.*, 1980; De Rosa and Gambacorta, 1988; Komatsu and Chong, 1998; Caforio and Driessen, 2017) (especially GDGT-0 through GDGT-5), which would help in identification and establishing the elution order. The resolution achieved was comparable with the published report (results not included). Given that the GDGTs required in the TEX_{86} index elute very late in the chromatographic run with the methanol – isopropanol mobile phase gradient elution it was rationalised that a solvent with greater eluotropic strength should effect elution with shorter retention times. During a preliminary work, ethyl acetate – methanol mobile phase was evaluated under the NARP conditions. The GDGT lipids eluted in a very narrow band and, despite attempts to effect separation under a gradient program, no reasonable resolution could be achieved (results not shown). Unfortunately, increasing the polarity of the mobile phase by part replacing methanol with water resulted in drastically increased retention and complete loss of resolution. For these reasons, the use of ethyl acetate as a component of the mobile phase was not evaluated further.

Building on the experience gained during development of the method to separate the synthetic glycerol diether (*r*-dOG) in its native form and as the FLB derivative, mainly using aqueous methanol (MeOH) and dichloromethane (DCM) (Chapter 5.4.3), it was considered that similar conditions should provide a reasonable starting point.

A portion of the *S. acidocaladrius* lipid core extract was eluted isocratically using MeOH/DCM/H_2O at a flow rate of 0.6 mL min^{-1}, while the column was maintained at 25 °C. The lipid cores (GDGT-0 to GDGT-5) were detected by MS using a scanning range of *m/z* 1290-1303 in order to aid identification and establish elution order. The base peak chromatogram revealed the target lipids to elute between 55 and 70 min, as shown **Figure 6-1**.

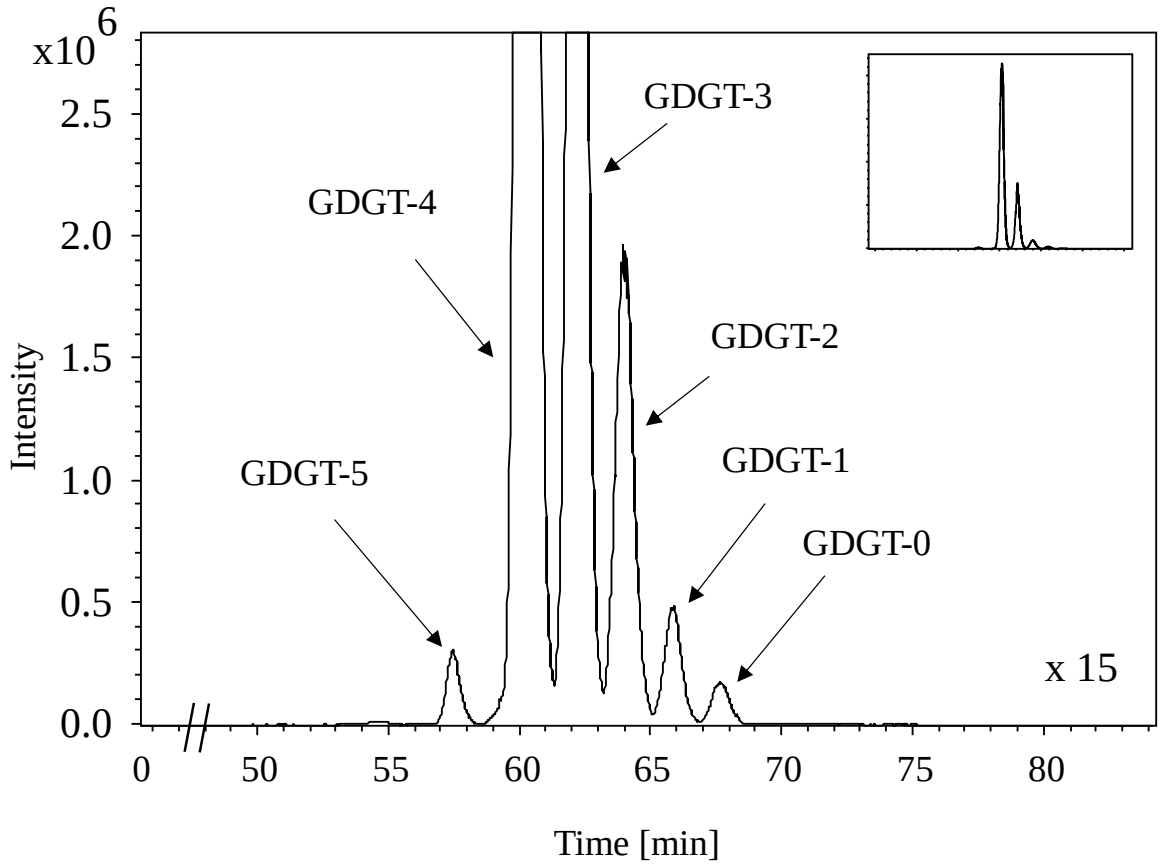


Figure 6-1 Base peak chromatogram of the polar fraction of *S. acidocaldarius* lipid extract obtained by RP-LC-APCI-MS. The inset shows the full-scale base peak chromatogram. Scan range *m/z* 1290-1303, column flow 0.6 mL min^{-1}.

The lipid cores eluted in the following order: the lipid with the greatest number of cyclopentane rings (GDGT-5) eluted first at 57.4 min, followed by the lipids with decreasing numbers of rings, with caldarchaeol (GDGT-0) eluting last at 67.7 min. The observed elution order is exactly the opposite of that observed when the lipids were chromatographed under the normal phase conditions **Figure 6-2**. Although, the GDGT lipid cores were eluted within a similar retention time window (approx. 10 min) the result in the RP mode is markedly

better, owing to the better analyte resolution and reduced background noise. This effect of the reduced background noise in RP is consistent with previous reports (Wörmer *et al.*, 2013; Zhu *et al.*, 2013).

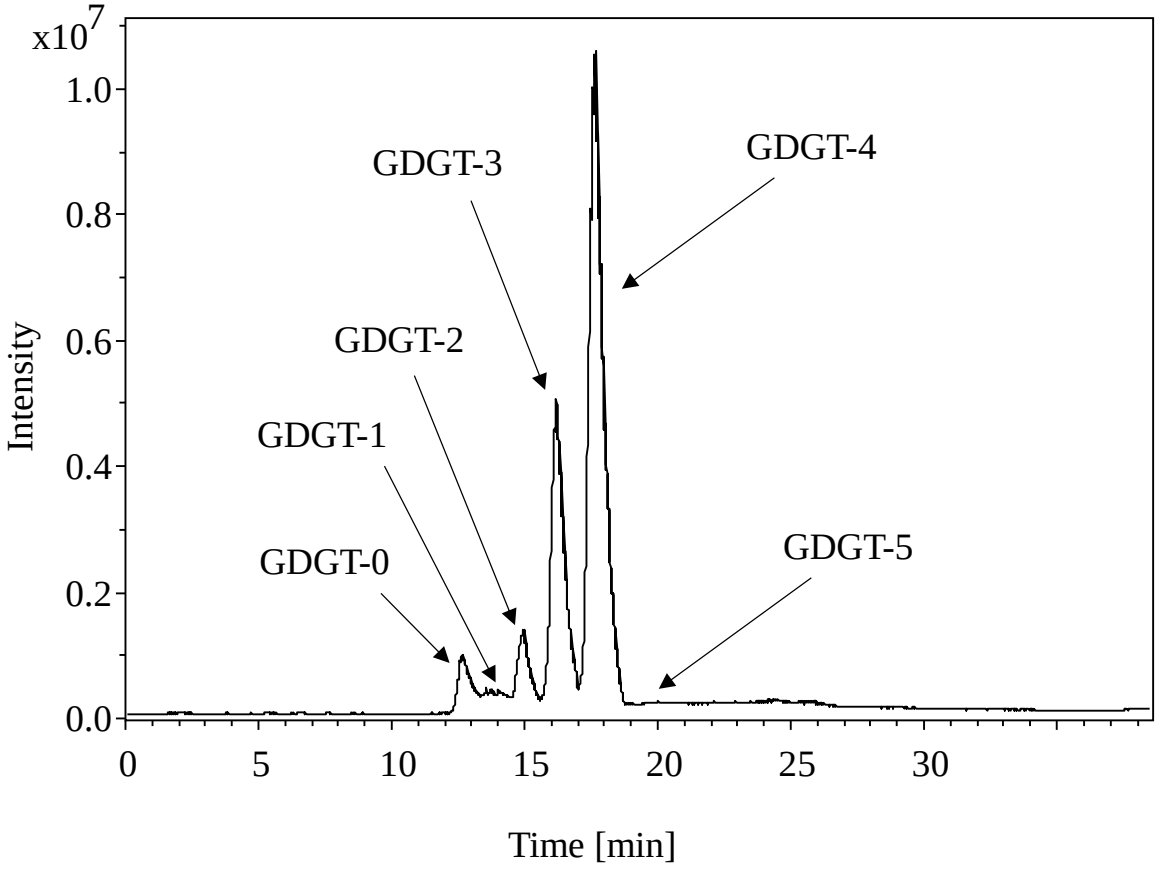


Figure 6-2 Base peak chromatogram of the polar fraction of *S. acidocaldarius* lipid extract obtained by NP-LC-APCI-MS. Scan range *m/z* 1290-1303, column flow 0.2 mL min^{-1}.

Significantly, the same amount the lipid extract was injected on column in both chromatographic modes. Whereas GDGT-5 was not separated under normal phase elution, it was well resolved from the later eluting GDGT-4 and displayed good peak shape in RP mode. While the overall signal intensity observed in both NP and RP modes was similar, the improved detection of GDGT-5 may be attributed to two factors: the analytes are better resolved in the RP mode and the reduction in background noise. Although, the GDGT-5 is not clearly resolved in NP mode due to the dominant GDGT-4 and the effect of peak tailing, its mass spectral information can be accessed *via* extracted ion chromatogram, thus allowing peak integration. The use of single ion monitoring (SIM) detection mode as opposed to mass scanning range mode can offer improved peak profiles of the analytes in the mass

chromatogram as reported previously (Wuchter *et al.*, 2004; Huguet *et al.*, 2006). This mode can be used with a single quadrupole MS systems whereas it cannot be achieved in exactly the same manner with ion trap MS, hence this feature was not utilised due to the use of ion trap MS system in this research work.

Although, under the initial RP-LC conditions the GDGTs were chromatographed with an appreciable degree of resolution, the retention times are too long for a high throughput method. An increase in the flow rate was considered in an attempt to reduce the retention times of the target GDGTs. The resulting system backpressure increase from doubling the flow rate to 1.2 mL min^{-1} was countered by increasing the column temperature to 45 °C and minimally increasing the content of water in the mobile phase. The lipid cores were eluted isocratically using mobile phase of composition MeOH/DCM/H_2O (49/35/16, v/v).

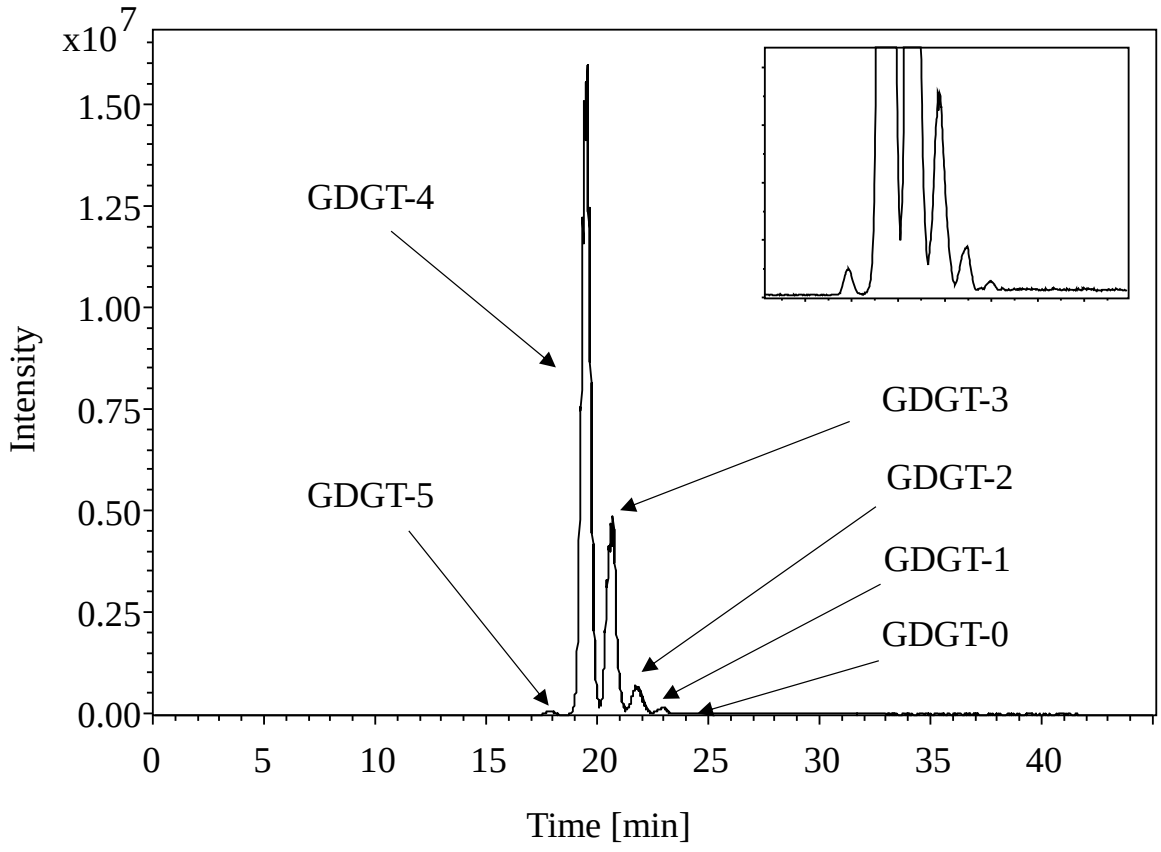


Figure 6-3 Base peak chromatogram of the polar fraction of *S. acidocaldarius* lipid extract obtained by RP-LC-APCI-MS. The inset shows resolution of the analytes near the baseline. Scan range *m/z* 1290-1303, column flow 1.2 mL min^{-1}.

The GDGTs were resolved reasonably well at the increased flow rate, eluting within a relatively narrow window and in less than 25 min as shown in **Figure 6-3**. However, the observed signal response slightly compared with the mass chromatogram obtained at a flow rate of 0.6 mL min^{-1} (**Figure 6-1**). This was most likely caused by the loss in the efficiency in ion transmission into the MS following the nebulisation and ionisation. Notably, at a flow rate of 1.5 mL min^{-1} almost complete loss of signal was observed. Given that the resolution was maintained at an increased flow rate, a further increase in the water content, and decrease in methanol, was explored in an attempt to further improve the resolution of the analytes. It was, however, quickly discovered that an increase of the water content above 15 %, while maintaining the DCM at 35 %, led to a loss of solvent miscibility, which affected both chromatography and detection.

The considerable retention of the lipid cores on the C18 column and the inability of alter the mobile phase components to effect better resolution whilst keeping the run time reasonably short, prompted consideration of an alternative stationary phase. A series of brief trials involving a number of standard columns having different column chemistries (C8, phenyl-hexyl, amide) showed that all of the columns tested were inferior in resolution of the GDGTs to the Ace UltraCore SuperC18 column tested thus far. Lastly, attention was turned to a column containing pentafluorophenyl propyl (PFP) stationary phase, reported to aid in separation of critical isomeric component pairs, often in high throughput mode (Richheimer, Tinnermeier and Timmons, 1992; van den Ouweland, Beijers and van Daal, 2011; Foo Wong *et al.*, 2014; Górnaś, Soliven and Segliņa, 2015). Moreover, replacing DCM with an organic solvent of a higher solubility in water was considered likely to circumvent the limitation on aqueous content in the mobile phase and ultimately lead to improved resolution of the GDGTs. Methyl tert-butyl ether (MTBE) was thus considered as a viable alternative to DCM. This solvent has a solubility in water of 4.3 (% w/w, 25 °C) compared with 1.3 (% w/w, 25 °C) for DCM (Smallwood, 1996). It was, therefore, envisaged that the PFP column, in conjunction with the MTBE based mobile phase, could offer unique separation conditions, potentially leading to improved resolution of GDGTs.

The separation of the polar fraction of the *S. acidocaldarius* lipid core extract was carried out in isocratic mode under the following conditions: mobile phase consisting of MeOH/MTBE/H_2O (50/30/20, v/v) was delivered at the flow rate of 1 mL min^{-1}, while the column was maintained at 45 °C. The GDGT lipids were resolved extraordinarily well, showing base line resolution and good peak geometries as shown in **Figure 6-4**. It is clear that the use of MTBE as a strong solvent played an immense part increasing the resolution of the lipid cores. The observed improvement in the separation of the lipid cores can most likely be attributed to the higher water solubility of the MTBE than DCM, and as such it allowed the aqueous component of the mobile phase to be raised significantly. Thus, the mobile phase compositions enabled the lipid cores to be eluted with an increased level of control by exploiting the subtle differences in their hydrophobic backbone. Additionally, the unique chemistry of the PFP column packing provided not only adequate retention of the target analytes, but also the ability to discriminate the GDGTs based on their polarity and differing solute - stationary phase interactions. As the lipids eluted considerably late in the chromatographic run, decreasing of the mobile phase polarity by increasing the proportion the MTBE was considered to shorten the retention times of the analytes.

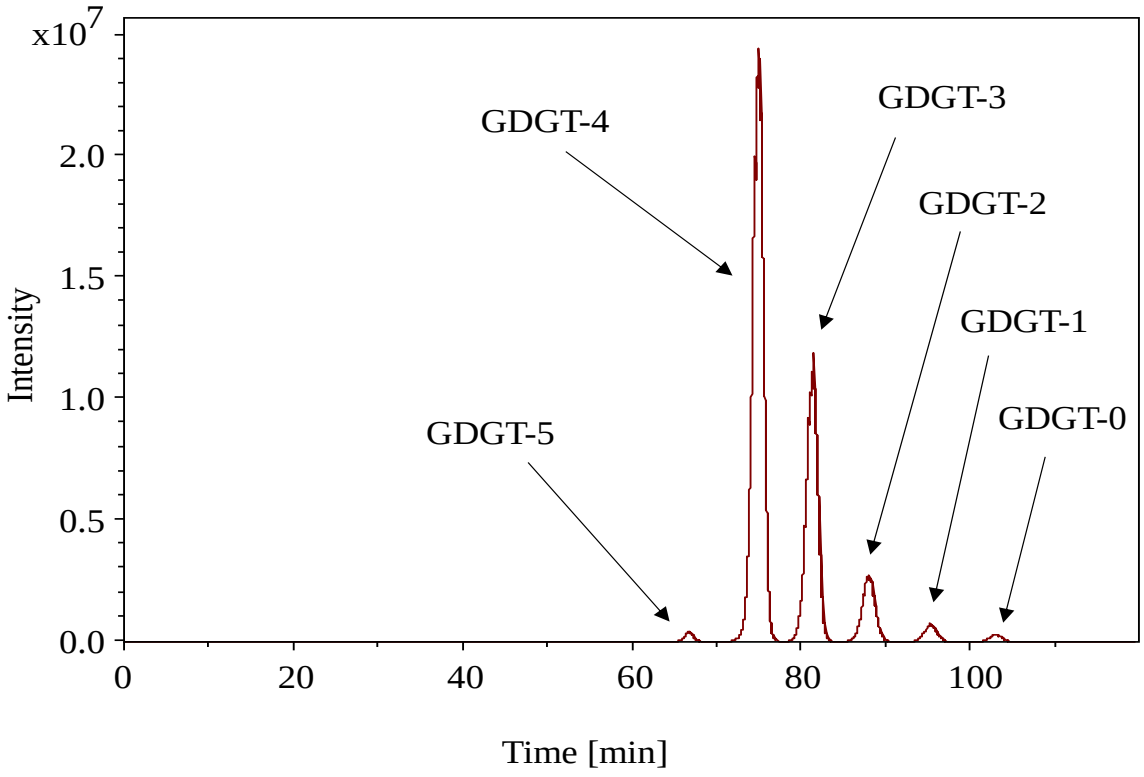


Figure 6-4 Base peak chromatogram of the polar fraction of *S. acidocaldarius* lipid extract obtained by RP-LC-APCI-MS. Scanning range *m/z* 1290-1303, column flow 1 mL min^{-1}.

While maintaining the flow and column temperature parameters and the ratio of water the same, the modification of the chromatographic conditions was focused on adjusting the methanol – MTBE ratio. After few trial runs, the appropriate mobile phase conditions were established to be MeOH/MTBE/H_2O (40/40/20, v/v). Under the new conditions, the lipid cores (GDGT-0 to GDGT-5) separated very well, achieving baseline resolution (**Figure 6-5**). Notably, the lipid cores eluted as very sharp peaks, considerably increasing the signal-to-noise ratio and allowing for the detection of GDGT-6 (inset in **Figure 6-5**), a component not previously detected using the DCM-based mobile phase. No considerable difference was observed in the distribution of the lipid cores due to the changes in the mobile phase composition.

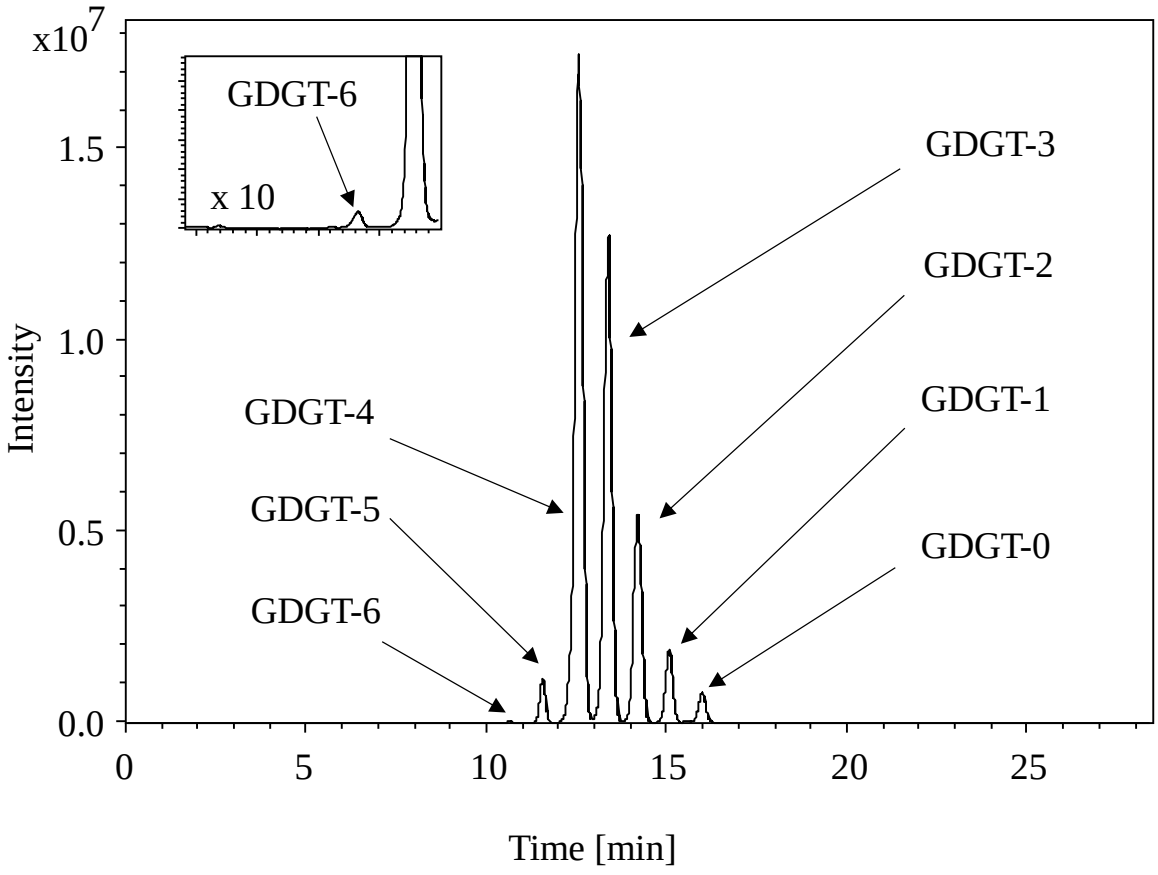


Figure 6-5 Base peak chromatogram of the polar fraction of *S. acidocaldarius* lipid extract obtained by RP-LC-APCI-MS. Scan range *m/z* 1290-1303, column flow 1 mL min^{-1}. The inset shows the degree of separation of the GDGT-6 near the baseline.

The lipid structures eluted within a very narrow retention time window (approx. 6 min) retaining the elution order established during previous RP-LC separations discussed in this section. Thus, the retention time increases with the decrease in the degree of the internal cyclisation of the isoprenoid core of the GDGTs, with caldarchaeol eluting last. The retention times and calculated column efficiency, *N*, for each of the target GDGT cores are summarised in the **Table 6-1**.

In practice, higher column efficiency correlated positively with column performance and the ability to maintain analytes movement through the column in a narrow band. This resulted in the analytes being detected as very sharp peaks.

Table 6-1 Retention times of the target lipid cores, their retention time, full width at half maximum ($w_{1/2}$) and column efficiency (*N*, Equation 6-1) of each of the GDGTs of interest.

Lipid core	*m/z* $[M+H]^+$	t_R [min]	$w_{1/2}$	*N*
GDGT-5	1292.5	11.6	0.2	18640
GDGT-4	1294.5	12.6	0.3	9770
GDGT-3	1296.5	13.4	0.2	24870
GDGT-2	1298.5	14.2	0.2	27930
GDGT-1	1300.5	15.1	0.2	31580
GDGT-0	1302.5	16.0	0.2	35460

$$N = 5.54 \times \left(\frac{t_R}{w_{1/2}}\right) \quad (6\text{-}1)$$

As the established ratio of methanol – water – MTBE mobile phase of the new method permitted the desired degree of analyte separation to be achieved, the impact of composition of the aqueous component was investigated. It was found that replacing the deionised water with a 10 mmol aqueous solution of ammonium acetate caused approx. 50 % decrease in the signal intensity of the lipids. On the other hand, replacing water with 0.1 % (v/v) aqueous solution of acetic or formic acid resulted in an increase in the detector response of two and four times, respectively. Thus, a 0.1 % solution of formic acid in water was incorporated into the final RP-LC method for the analysis of the archaeal lipid cores. Notably, attempts to

elute the lipid cores under a gradient elution program did not result in a significant improvement in resolution among the analytes most likely due to the distorted dynamics of analytes partitioning between the stationary and mobile phase. In summary, the novel method developed allows screening for the archaeal cyclopentane containing lipid cores to be achieved in under 16 min, making it the fastest RP-LC reported thus far, and meeting the analytical run time constrains of a high throughput method.

In order to enable evaluation of the separation of the Fmoc-lysine(Boc) (FLB) modified derivatives of the GDGTs (0 to 5 cyclopentane rings) a small portion of the *S. acidocaladrius* lipid cores was derivatised using a slightly modified Method C (see Chapter 2.2.3). The ratio of the amino acid/EDC/DMAP was increased to 50/75/10 equivalents with regard to the hydroxyl group of caldarchaeol, assuming that it comprised the whole of the lipid extract. This was done to account for the matrix components that could consume part of the derivatising reagents and lead to incomplete conversion of the GDGTs. The structure of a typical Fmoc-lysine(Boc) diester derivative of a GDGT lipid, here caldarchaeol (de-GDGT-0), is shown in **Figure 6-6**. The diester derivatives were detected as the fragment ion 100 Da lower than the protonated molecule. This ion was formed *via* a characteristic loss of the Boc group, a commonly observed feature in this type of derivatives (see Chapter 5.4.1), leading to the formation of a stable ion at *m/z* 2104.

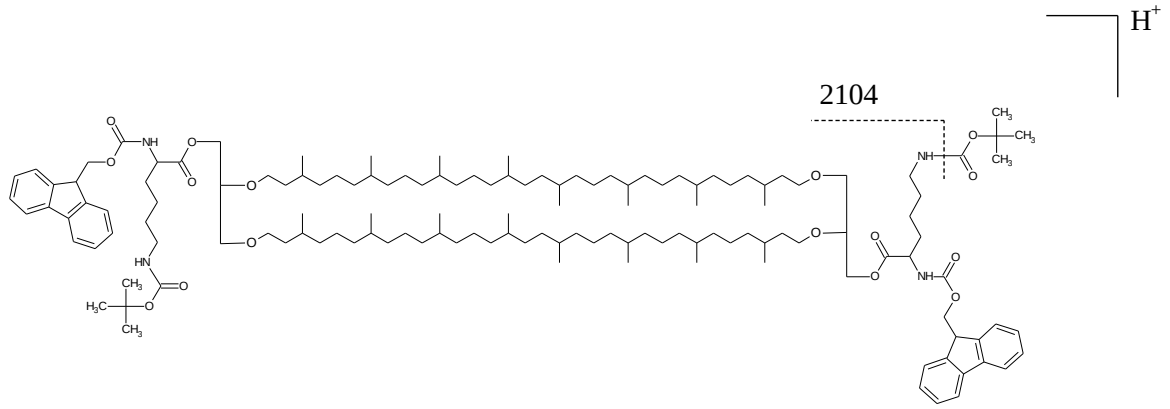


Figure 6-6 Structure of the caldarchaeol (GDGT-0) diester modified using Fmoc-lysine(Boc). Broken line indicates the bond breaking during the ionisation stage in MS. On ionisation the characteristic loss of one Boc group leads to the formation of the fragment ion *m/z* 2104.

The modified lipid cores were eluted using the method developed for the native lipid cores, to permit monitoring of any unreacted or partially derivatised lipid cores, which would elute before the diester derivatives. The derivatives separated well as evidenced in the mass chromatogram shown in **Figure 6-7**.

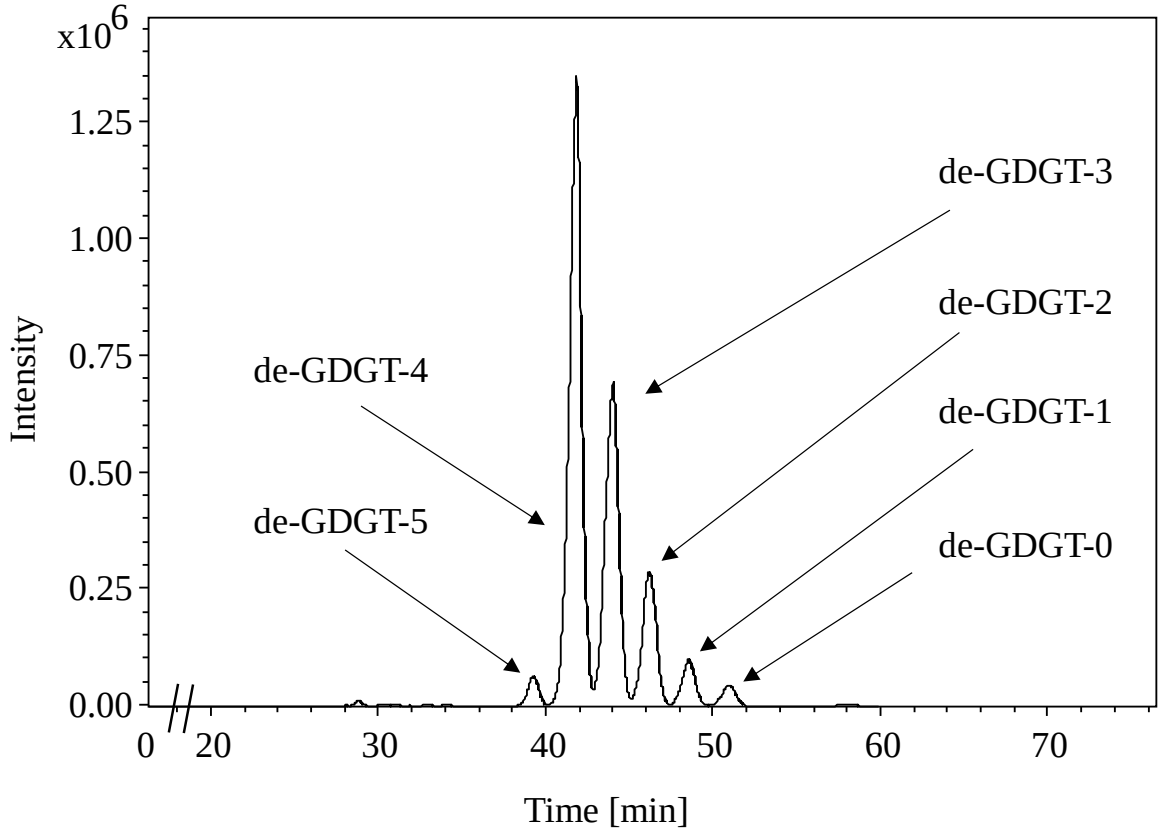


Figure 6-7 Base peak chromatogram of the FLB derivatives of lipid cores derived from *S. acidocaldarius* lipid extract obtained by RP-LC-APCI-MS. The prefix "de" denotes a diester derivative. Scan range *m/z* 2092-2105.

The derivatives followed the same elution order as the native lipid structures, such that the derivative of the lipid with the highest degree of cyclisation, diester de-GDGT-5 eluted first at 28.2 min and the caldarchaeol diester eluted last at 37.2 min. No residual levels of native GDGTs were detected within their normal elution time window (11-16 min; **Figure 6-5**) suggesting that complete conversion of the lipids was achieved or that the unreacted native lipids were well below the detection limit of the methods.

Notably, by lowering the aqueous 0.1 % formic acid proportion by 5 % while simultaneously increasing the proportion of methanol by the same degree (5 %) it was possible to reduce the

retention time by approx. 10 min as shown in **Figure 6-8**. The reduction in the retention times of the diester derivatives of GDGTs had insignificant effect on their resolution and no distinct change in their distribution was observed.

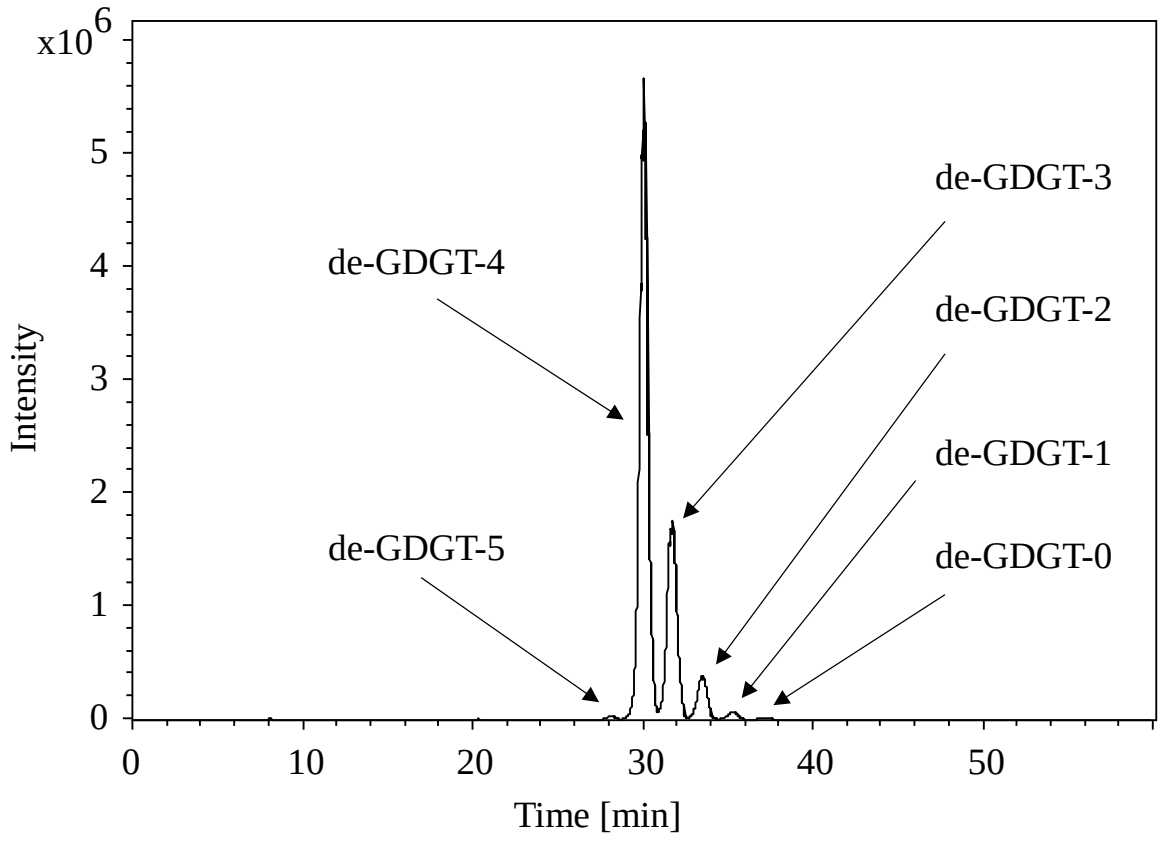


Figure 6-8 Base peak chromatogram of the FLB derivatives of lipid cores derived from *S. acidocaldarius* lipid extract obtained by RP-LC-APCI-MS. The prefix "de" denotes a diester derivative. Scan range *m/z* 2092-2105.

The adjustment of the mobile phase composition significantly reduced the analysis time and should find application in high throughput analysis applications for screening of the GDGT lipids containing cyclopentane rings in their isoprenoid backbone structure. In addition, the cyclopentane ring containing lipid core derivatives were monitored under fluorescence detection (FLD) in order to evaluate their detector response. Assuming that the fluorescence response of the derivatives is independent of the lipids core structure, comparison of their peak area ratios with the peak area ratios of the native lipids could shed light on the ionisation efficiencies of the latter during MS analysis. The HPLC-FLD chromatogram shown in **Figure 6-9** indicates that although the modified lipids were detected, the peaks are considerably broader than the peaks in the MS chromatogram. The cause of the peak

broadening is solely associated with the void volumes within the different chromatographic system that was used for fluorescence studies, as has been discussed in Chapter 5.4.4.

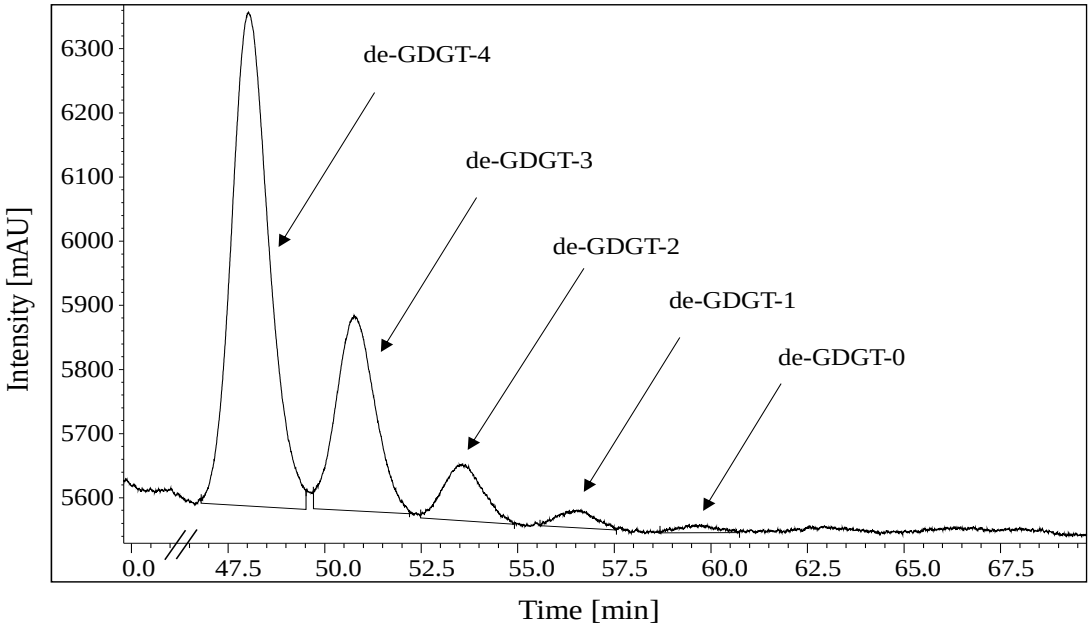


Figure 6-9 HPLC-FLD chromatogram of the FLB derivatives of lipid cores derived from *S. acidocaldarius* lipid extract. The prefix "de" denotes a diester derivative.

Comparison of the normalised MS responses of the archaeal tetraether lipids in their native form with the MS and FLD responses of the FLB derivatives (**Figure 6-10**) shows that the relative response for MS detection of the homologues is essentially the same. This suggest that the relative ionisation efficiencies of the native and the FLB derivatives are very similar for the respective GDGT homologues. By contrast, the responses for the GDGT (1-3) FLB derivatives in the FLD detector do not follow the same pattern as indicated by the lower relative levels of the de-GDGT-2 and de-GDGT-1 when normalised to de-GDGT-3. The normalised relative response of de-GDGT-2 to de-GDGT-3 in fluorescence is 30 % whereas it is 42 % for the derivative in MS and 44 % for the native lipid cores in MS. Similar differences are observed for the de-GDGT-1 derivative. The lower relative response of the lipid core derivatives in fluorescence compared with the MS detection could potentially be associated with reduction of fluorescence yield due to steric interference of the bulky Fmoc functionality, forcing the derivative molecule to take less energetically favoured conformation. Indeed, while investigating fluorescence properties of 1-dimethylaminonaphthalene-5-sulfonyl chloride (dansyl chloride) derivatives of a group of alkaloids, Nachtmann *et al.* (1975) reported a lower fluorescence intensity of adrenaline

triester compared with mono and diesters of ephedrine and cephaeline, respectively. The authors attributed this effect to a distortion out of the planar configuration caused by steric hindrance of the dansyl substituents. Similar structural differences resulting from conformational differences among GDGT lipids having different numbers of cyclopentane rings cannot be ruled out and should be considered in future studies.

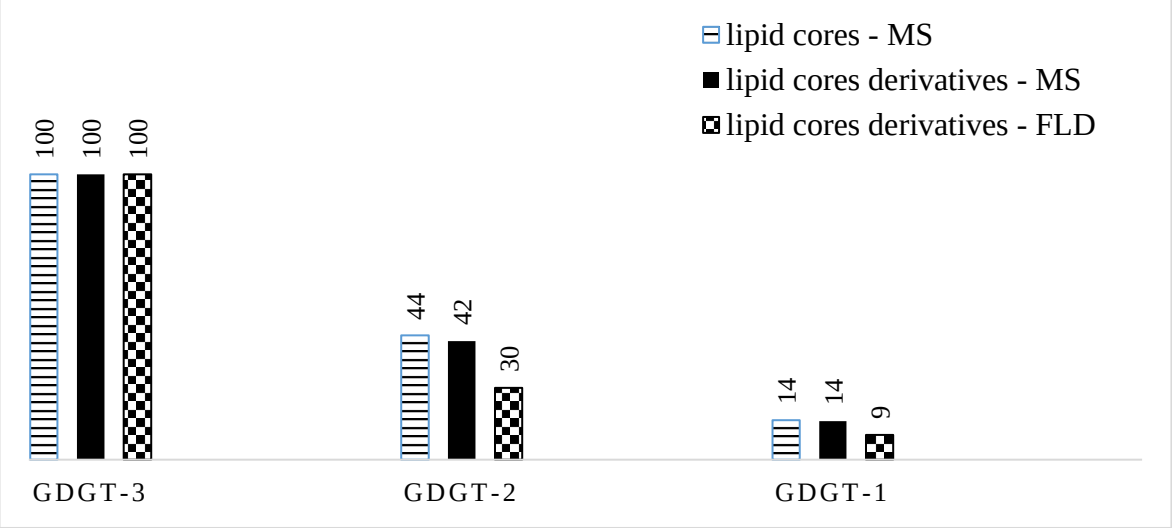


Figure 6-10 Comparison of the signal response (based on peak area) of the native and FLB-derivatives of the selected GDGTs. The detector response has been normalised in all series to the values for GDGT-3. Annotations: MS-mass spectrometry detector, FLD-fluorescence detector, derivative-FLB diester derivative of the lipid core.

The lower values observed for the FLD detector may also be attributable to the observed effect of peaks broadening, which in turn complicated integration. This may, however, be circumvented by analysing the FLB lipid derivatives on a modern UHPLC-FLD system equipped with an adequately sized flow cell to eliminate this effect. Overall, it is reasonable to conclude that the ionisation efficiency of the tetraether lipid cores, within the range of the structures evaluated, is very similar if not uniform. This is not to say that the same is true between other types of archaeal lipid core such as dialkyl, macrocyclic ether lipids, or structures containing additional hydroxyl groups or unsaturation within the hydrocarbon skeleton.

Despite the significantly lower the detection limits of the HPLC-FLD technique due to the observed peak broadening it was demonstrated that from the “proof-of-concept” point of view, the derivatisation enabled detection of the lipid cores as their (FLB) derivatives.

6.3.2 Derivatisation of the lipid extract from an ancient sediment

A polar fraction containing the isoprenoid tetraether lipid cores (GDGTs) from the Middle Jurassic period (ca. 160 Ma) shale of the Stewartby member of the Oxford clay formation, UK (sample S90-11, Kenig *et al.*, 1994) was prepared. The soil material was extracted using Dionex ASE according to a previously established method (Schouten *et al.*, 2007), followed by the hydrolysis to cleave the polar head groups, if present, and liberate the lipid cores as outlined by Knappy *et al*. (2012). Finally, the extract was fractionated into apolar and polar fractions, the latter containing the GDGT lipid cores. To a portion of the polar fraction a small amount of the synthetic glycerol diether (*r*-dOG, 6 μg) was added to allow monitoring of material losses during sample handling between subsequent analyses as well as the completeness of the derivatisation reaction. To maintain the fidelity of the comparative values from different analyses the same sample was used throughout this study. Thus, the same portion of the polar fraction was analysed under normal phase conditions (NP-LC-MS), reversed phase conditions (RP-LC-MS), and following derivatisation and analysis under reverse phase mode with MS and fluorescence detection. After each analysis, the sample solvent was blown down using a stream of nitrogen gas and then the sample was reconstituted in the appropriate solvent for the subsequent analysis. The sample injection was set to 5 μL to ensure comparison of equivalent amount of material injected on column. The derivatisation reaction of the lipid cores using *N*-Fmoc-lysine(Boc) was carried out according to Method C (see Chapter 2.2.3) using a large excess of DMAP/amino acid/EDC (10/50/75 equivalents), assuming that the caldarchaeol comprises the whole lipid extract, to ensure complete conversion of the GDGT lipid cores.

The analysis of the lipid cores under NP-LC-APCI-MS (see Chapter 2.6.1) revealed the distribution of the GDGTs, which is similar to that reported previously (Knappy *et al.*, 2012, 2015). The base peak chromatogram, as well as the extracted mass chromatograms of components of the TEX_{86} index, are shown in **Figure 6-11**. Although GDGT-0 was also detected, as shown in the base peak chromatogram (BPC), it is not shown in the extracted ion chromatogram (EIC) due to it not being used in the calculation of the TEX_{86} index.

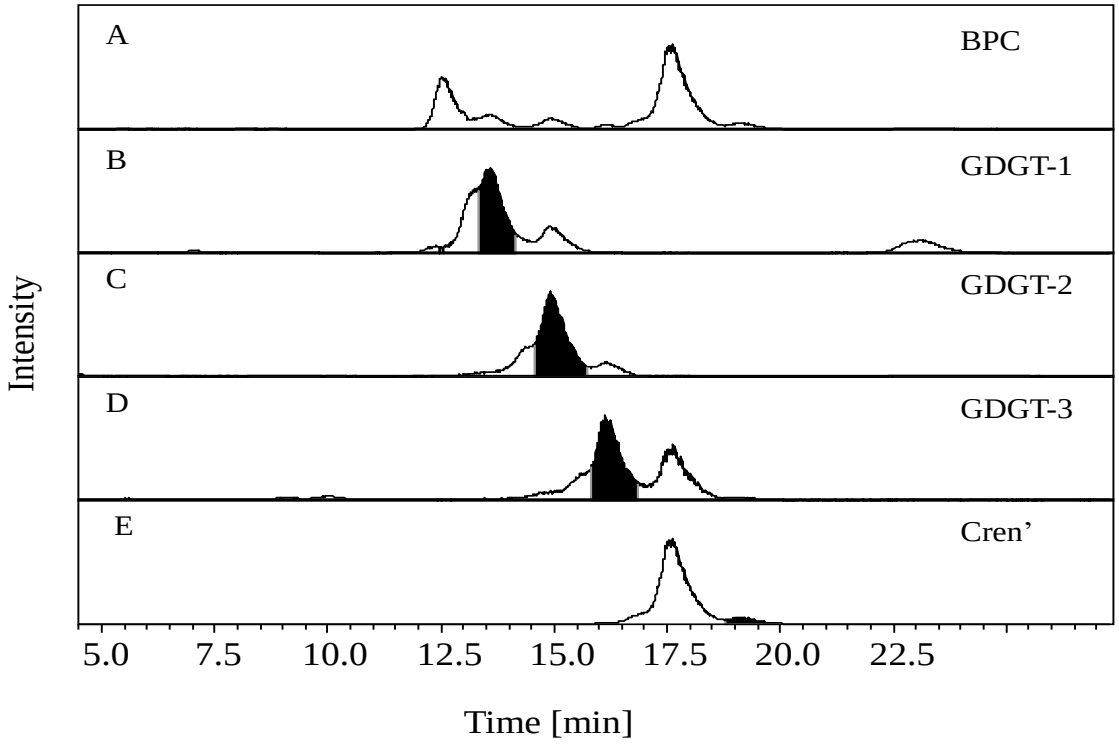


Figure 6-11 NP-LC-MS chromatograms of the polar fraction derived from the Oxford Clay OC-S90-11: A) base peak chromatogram (1290-1303), and extracted ion chromatograms corresponding to the protonated molecules of: B) GDGT-1, C) GDGT-2, D) GDGT-3, D) crenarchaeol regioisomer.

By contrast with the poor separation of the analytes under normal phase mode, the mass chromatogram (base peak) obtained from the analysis using the novel RP-LC-MS method (see Chapter 2.6.4) shows a mass chromatogram with baseline resolution of the GDGT lipid cores with differing numbers of cyclopentane rings **Figure 6-12**.

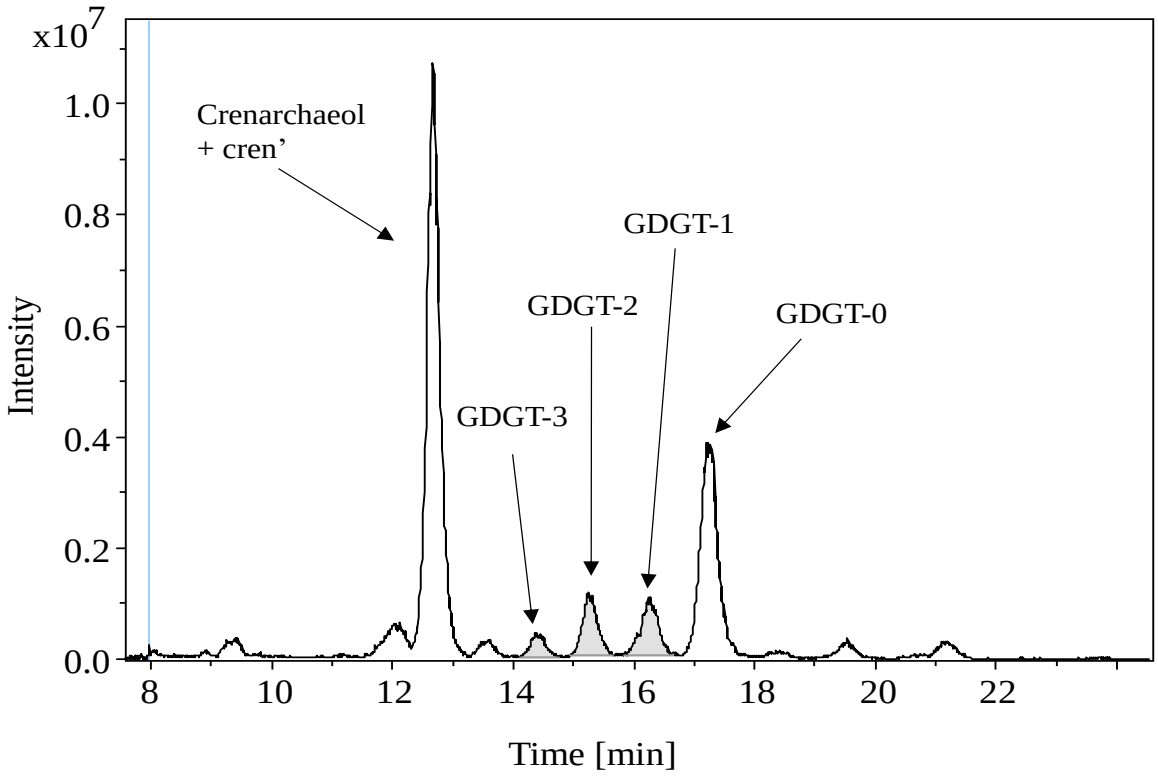


Figure 6-12 RP-LC-MS chromatograms of the polar fraction from the Oxford Clay OC-S90-11. Base peak chromatogram (*m/z* 1290-1303) showing major lipids, with integrated peaks for components included in the TEX_{86}^{L} index.

Notably, crenarchaeol regioisomer, a critical component of the TEX_{86} index, was not resolved under RP mode separation as it coelutes with crenarchaeol. This, however, can be alleviated by using the TEX_{86}^{L} version of the index for palaeotemperature reconstruction, as this index can be used to reconstruct geological temperatures where the anticipated temperature is below 15 °C, with reasonably good accuracy (Kim *et al.*, 2010). For sediments originating from warm or tropical regions this approach may not be effective, as the temperatures estimated using TEX_{86}^{L} would carry appreciable bias. Hence, the use of the RP mode as an analytical method can find its utility as means of rapid screening method for the assessment of presence of the tetraether lipid cores in a sample material before further sample processing and subsequent appraisal of the TEX_{86} index.

In an attempt to separate the crenarchaeol and its regioisomer in a fraction enriched in the regioisomer by preparative NP-HPLC, the polarity of the mobile phase was increased considerably, by replacing part of the MTBE with methanol. A partial, though not satisfying, separation of the two components was achieved in approximately 80 min. This suggests that full separation of crenarchaeol and its regioisomer would require extending the run time

considerably using this column. This was deemed impractical from the high-throughput perspective.

The analysis of the OC-S90-11 lipid extract derivatised using Fmoc-lysine(Boc) was carried under the same conditions as the analysis of the native lipids with both MS and fluorescence monitoring. Although the derivatives elute after 40 min, analysis under these conditions enables residual lipids cores to be accounted for in the same run. Unexpectedly, the mass chromatogram of the derivatised lipids revealed only a very weak signal corresponding to the most abundant lipid derivatives crenarchaeol (with cren') and caldarchaeol (GDGT-0)as shown in **Figure 6-13**.

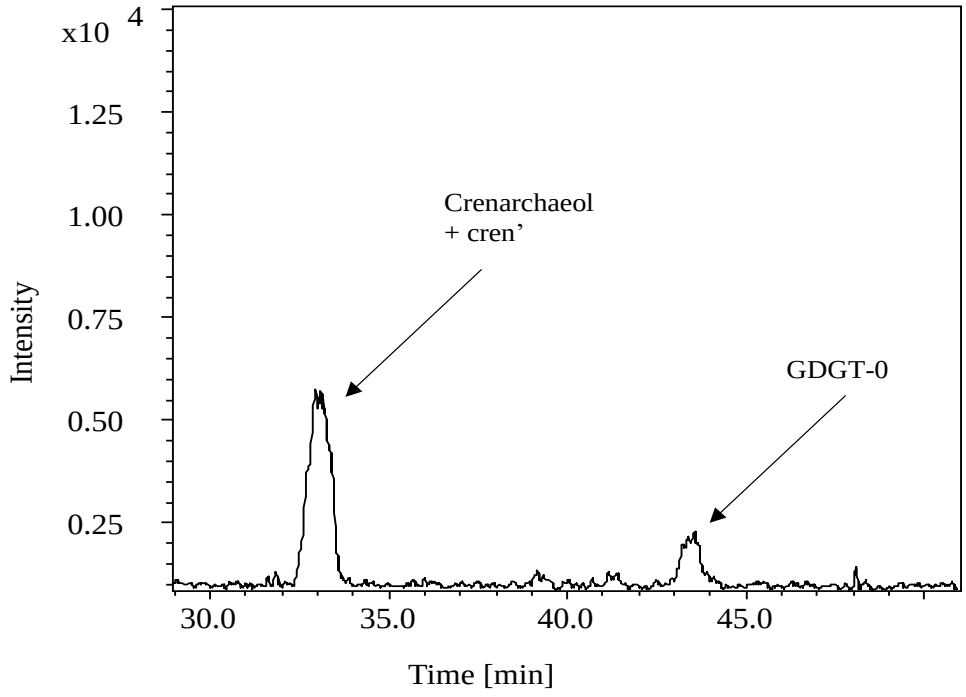


Figure 6-13 RP-LC-MS chromatograms of the derivatised polar fraction derived from the Oxford Clay OC-S90-11. Base peak chromatogram (*m/z* 2090-2105) showing the major GDGT lipid cores.

Comparing the signal response of the native and derivatised lipid cores, a dramatic drop in the signal intensity, of approx. two orders of magnitude, is clearly evident. This is in contrast with the observed signal intensity improvement observed after modifying the synthetic glycerol diether (*r*-dOG) with Fmoc-lysine(Boc), as discussed earlier (Chapter 5.4.3). Moreover, evaluation of the signal response of the *r*-dOG that was added to the lipid extract via the chromatograms obtained before and after derivatisation illustrates the enhancement

in the MS detection very clearly (**Figure 6-14**). The apparent peak area increase attributed to the derivatisation of the diether lipid comparing to the native lipid is approx. 90 times and is consistent with observations discussed earlier.

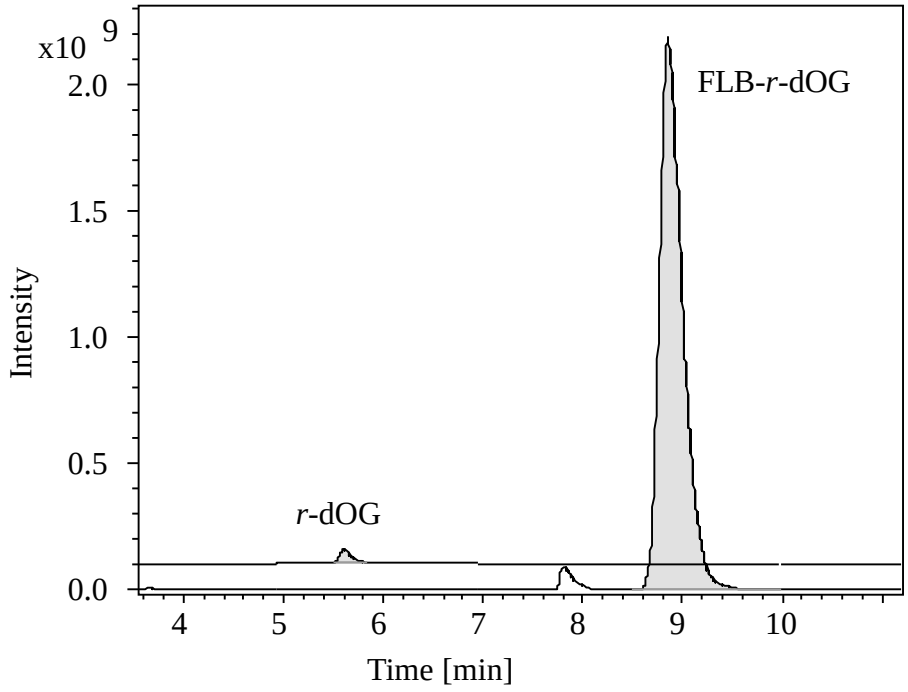


Figure 6-14 RP-LC-MS base peak chromatograms of the native diether *r*-dOG and its FLB derivative. The mass chromatograms were normalised. The peak eluting at approx. 7.8 min in the mass chromatogram of the derivative is a known impurity.

By analogy, it was expected that GDGTs modified using the same amino acid would exhibit a similar improvement in signal strength. To investigate potential losses of the material during sample transfer and silica gel column purification a separate portion of the polar fraction of the extract was first analysed using RP-LC-MS. It was subsequently loaded onto a silica gel column and eluted in the usual manner, as were the lipid derivatives and was then reconstituted and re-analysed. The calculated recovery was over 95 %, suggesting that losses of the material during sample handling were low. Furthermore, the derivatives used during the reversed phase method development had been prepared in earlier experiments from substantially smaller amounts of the lipid extract and the derivatives produced highly intense signals in the mass spectrometer. When the fresh lipid extracts from cultured archaea were later derivatised and analysed no enhancement in signal response the was observed following modification, in fact the signal intensity was lower, similar to the derivatised lipid extract of OC-S90-11. This suggested that a factor negatively affecting the response of the derivatives

must have occurred between the time of the derivatisation of the lipid extracts for method development studies and the evaluation of the effects of derivatisation on the sediment extract. Indeed, in that period the ion trap MS suffered simultaneous vacuum pump and electron multiplier failure. Although, the pump and electron multiplier were replaced and the detector subsequently recalibrated, the apparent instrument response did not improve with regard to the derivatised GDGT lipids. This may suggest that there are other underlying factors/conditions affecting the response for components having a relatively large mass (the derivatives have masses of approx. 2200 Da). This suggestion is put forward in the light of the markedly better performance of the synthetic diether derivative compared with its native form. Thus, it is likely that, following the failure of the instrument components, an unaccounted for factor that impacts on the analyte response at the high end of the mass range has resulted in the lower signal response for the derivatives.

Regrettably, the performance of the mass spectrometer could not be restored to its previous performance; hence, it was not possible to evaluate the full benefits of the signal intensity improvement in the derivatives of the GDGT lipids under optimal conditions.

6.3.3 Reconstruction of the SST using the Oxford Clay S90-11 sedimentary deposit

The sea surface water temperature (SST) values were reconstructed using the information obtained from both the normal phase and reversed phase analysis of the marine deposit extract OC-S90-11. The TEX_{86} index was calculated from the relative abundances of the lipid cores GDGT-1 to GDGT-3 and crenarchaeol regioisomer according to Kim *et al*. (2008). Due to co-elution of crenarchaeol and its regioisomer rendering their relative abundances inaccessible, in the case of the RP method, the TEX_{86}^{L} index was used instead. In addition, due to potentially high amount of organic pigments in the extract material, as suggested by the intense colouration of the lipid extract, a simple fractionation using a C18 solid phase extraction cartridge was performed. The isolated fraction containing the GDGTs was subsequently analysed using the reversed phase method. The relative abundances of the lipid cores are given in **Table 6-2**.

Table 6-2 Relative abundances of the GDGTs used to calculate TEX_{86} and TEX_{86}^{L} indexes. The asterisk denotes the lipid extract that was additionally fractionated using a C18 solid phase extraction cartridge.

Lipid core	*m/z* $[M+H]^+$	NP-LC-MS	RP-LC-MS	RP-LC-MS*
GDGT-1	1300.5	4.76×10^7	2.88×10^7	1.99×10^7
GDGT-2	1298.5	4.08×10^7	2.80×10^7	1.95×10^7
GDGT-3	1296.5	1.65×10^7	8.93×10^6	6.90×10^6
Cren'	1292.5	2.28×10^7	N/A	N/A

Note: N/A data not available

Notably, the ratio of the relative abundances of the lipid cores GDGT-1 to GDGT-3, analysed in RP mode, is similar for the sample before and after purification by solid phase extraction, suggesting that the removal of the potential matrix interference did not affect the GDGT distribution. Although the additional SPE clean-up step does not seem to be necessary, it may help to maintain the instrument free of build up of contamination.

Calculation of the BIT value is needed for the assessment of the origin of the organic matter in sediment samples, and ultimately for the appraisal of the terrestrial input, before the TEX_{86} index can be calculated and sea surface temperature SST reconstructed. Nevertheless, calculation of the TEX_{86} index values was performed to enable comparison of the index values obtained using the novel (RP) method with that obtained using the standard (NP) method. The TEX_{86} index values calculated using the three different index formula are given in **Table 6-3**. Using the data obtained from the normal phase method the TEX_{86} value was 0.63 and the TEX_{86}^{H} gave a value of -0.2. Notably, the TEX_{86}^{L} values using data form normal phase and reversed phase methods were -0.41 and -0.37, respectively.

Table 6-3 TEX_{86} index values calculated using different formula using data form, both normal and the reversed phase methods.

Method	TEX_{86}^{L}	TEX_{86}^{H}	TEX_{86}
NP-LC-MS	-0.41	-0.2	0.63
RP-LC-MS	-0.37	N/A	N/A

Note: N/A index value was not calculated

To illustrate more clearly how the index values generated by two different analytical methods affects the precision of reconstruction of the palaeotemperature the sea surface temperature (SST) was computed using different TEX_{86} index formula (**Table 6-4**). Notably, the SST reconstructed using the TEX_{86}^{L} index is 22.2 °C and 24.8 °C for normal and reversed phase methods, respectively. The bias of 2.6 °C is significant and suggests that by using the reversed phase method the reconstructed geological temperatures will be skewed towards higher temperatures, reflecting the rationale for the TEX_{86}^{L} in applications to reconstruct temperatures below 15 °C (Kim *et al.*, 2010). The SST reconstructed using the TEX_{86}^{H} (designated for use for temperatures above 15 °C) and the commonly used the TEX_{86} formula, using data from the normal phase method, are 24.7 °C and 24.5 °C, respectively. The close alignment of the two values reaffirms the use of TEX_{86}^{H}. Remarkably, despite using the TEX_{86}^{L} and data from the reversed phase method, the SST is in good agreement with both TEX_{86}^{H} and the TEX_{86} based on the data from the normal phase method: 24.8 °C *versus* 24.7 °C and 24.5 °C, respectively (**Table 6-4**).

Table 6-4 Sea surface water temperature (SST) reconstructed using, both normal and the reversed phase methods. The BIT index was not calculated for the reconstruction of SST.

	SST		
Method	TEX_{86}^{L}	TEX_{86}	TEX_{86}^{H}
NP-LC-MS	22.2	24.5	24.7
RP-LC-MS	24.8	N/A	N/A

Note: N/A index value was not calculated

The excellent correspondence of the TEX_{86}^{L} from RP-LC-MS and TEX_{86} from NC-LC-MS cannot be rationalised easily. It may, however, be a result of the improved resolution of the GDGT components of the TEX_{86}^{L} index by RP-LC and, as such, the proportions of the individual lipid cores being more reliable than those obtained from the normal phase method. Notably, this does not reconcile the fact that the TEX_{86} indices were developed using the normal phase method, and any potential correlation of the data for the two analytical methods was not anticipated. It is possible that, due to the similar ionisation potentials of the GDGTs in both methods and a higher precision in establishing of their abundance levels can permit the use of the reversed phase method-based TEX_{86}^{L} for the reconstruction of geological temperatures in settings with temperatures above 15 °C.

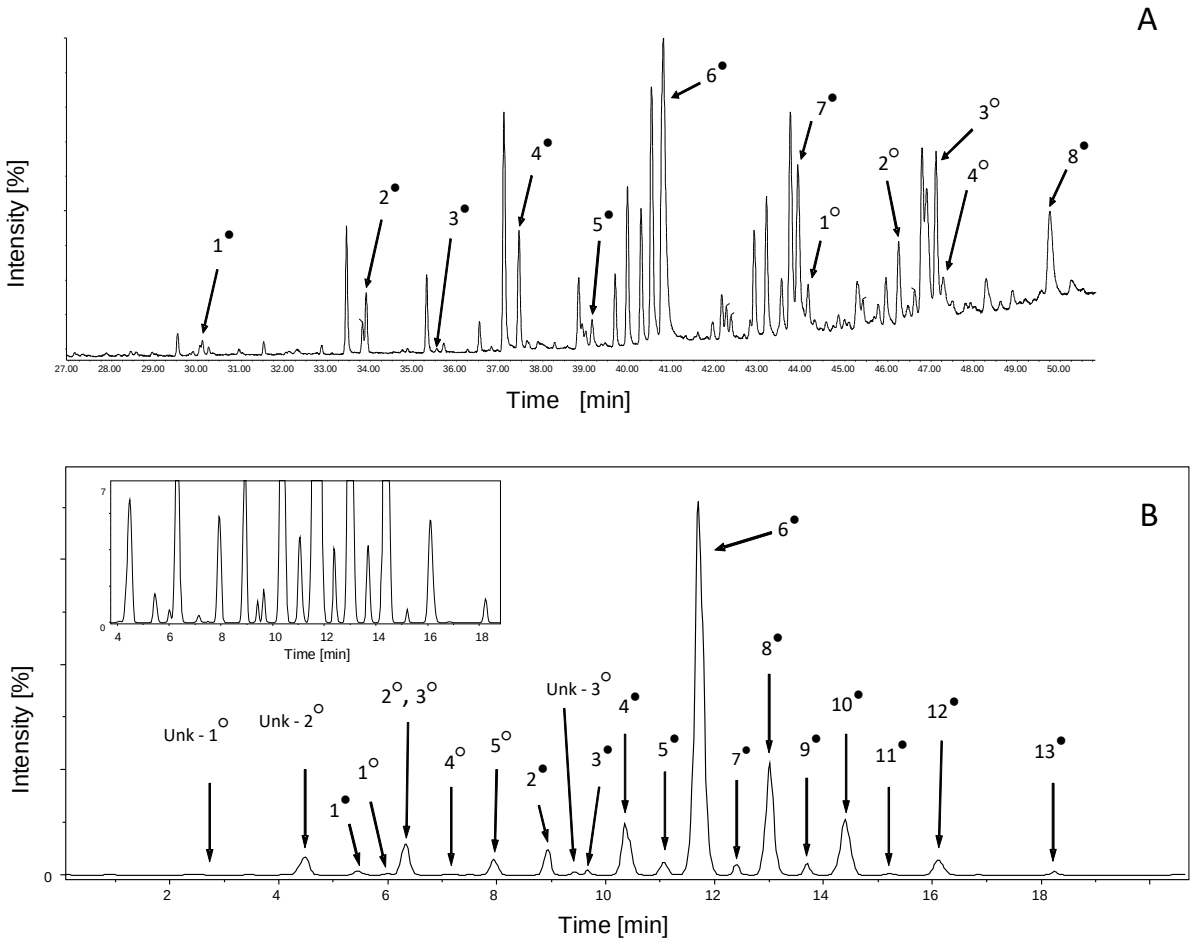
It must also be acknowledged that, despite the positive correlation between SST reconstructed using the reversed phase and the normal phase methods, the application of the new method was only tested on one sediment sample. Hence, in order to understand the relationship between the GDGT abundances obtained from the reversed phase method and the TEX_{86}^{L} and SST a much larger sediment sample population is required. It is envisaged that if no reasonable correlation can be established through, for example a correction factor, a new index can be constructed for the SST reconstruction of geological sedimentary settings. Finally, when the outlined goal is achieved the novel reversed phase method will allow the sample analysis to be reduced by three times.

6.3.4 Preparation and analysis of the Fmoc-lysine(Boc) derivatives of a grave soil extract lipids

The total lipid extract (TLE) from an experimental piglet burial soil from the InterArChive project was derivatised using Fmoc-lysine(Boc) amino acid according to Method C and analysed to evaluate the derivatisation approach as a potential alternative to the gas chromatography (GC) commonly employed in analysis of sterols and alkanols present in such sample material. The extract had previously been characterized by gas chromatography – mass spectrometry (GC-MS). The derivatised total lipid extract sample was chromatographed using the RP-LC-MS system according to Method V (see Chapter 2.6.4)

The chromatographic resolution achieved under the above gradient program was satisfactory and allowed detection of more of the components in TLE that had previously been detected using GC-MS. In addition, a host of components not observed during the original GC-MS analysis were identified. For comparison, the chromatograms obtained by both methods, GC-MS and HPLC-MS, are shown in **Figure 6-15**. The retention order is dramatically different: in GC-MS the *n*-alkanols and sterols elute towards the end of the chromatographic run (between 30 and 50 min) with the former eluting first (**Figure 6-15**A). By contrast, during HPLC-MS analysis, the sterol derivatives elute before the *n*-alkanol derivatives and the analytical run takes less than 20 min. Another important observation is that during GC separation all of the soil total lipid extract sample components, including hydrocarbons, derivatised acids, amides, will be detect even though they may not be the targeted analytes.

During the HPLC-MS separation, however, the transformation of the hydroxyl lipids into their derivatives allows their specific recognition using tandem MS and filtering the mass chromatogram for a specific neutral loss (NL). The mass chromatogram in **Figure 6-15**B has been reconstructed by filtering the mass signal to reflect the exact NL of 239.4 Da. By employing the MS to selectively extract signal from only the derivatised hydroxyl lipids their MS response is highlighted, and a much cleaner and more easily de-convoluted mass chromatogram was obtained. This approach allowed for a greater number of both alkanols, and sterols to be profiled (as the derivatives) than when the lipid extract was analysed using GC-MS (see legend in **Figure 6-15**). Although, the alkanols detected in the GC-MS ranged up to C30 it is expected that extending the run time and the temperature program would enable the detection of the longer chain alcohols. Notably, however, this would result in even longer analysis time and reduced overall sample throughput, rendering the derivatisation and HPLC-APCI-MS approach even more appealing.



Legend:

Black circle (●) *n*-alkanols: **1**$^{\bullet}$ - C_{20}, **2**$^{\bullet}$ - C_{22}, **3**$^{\bullet}$ - C_{23}, **4**$^{\bullet}$ - C_{24}, **5**$^{\bullet}$ - C_{25}, **6**$^{\bullet}$ - C_{26}, **7**$^{\bullet}$ - C_{28}, **8**$^{\bullet}$ - C_{30}, **9**$^{\bullet}$ - C_{18}, **10**$^{\bullet}$ - C_{27}, **11**$^{\bullet}$ - C_{29}, **12**$^{\bullet}$ - C_{31}, **13**$^{\bullet}$ - C_{32}, **14**$^{\bullet}$ - C_{34},

Open circle (○) sterols: **1**$^{\circ}$ - Δ^{5} – C_{27}, **2**$^{\circ}$ - $\Delta^{5,22}$ – C_{29}, **3**$^{\circ}$ - Δ^{5} – C_{29}, **4**$^{\circ}$ - C_{29}, **5**$^{\circ}$ - $\Delta^{5,7,22}$ – C_{28},

Figure 6-15 Separation of soil TLE sample from the experimental piglet burial from the InterArChive project. **A**) GC-MS chromatogram. The sample was methylated and silylated to derivatise carboxylic acid and hydroxyl functionalities. **B**) RP-LC-APCI-MS chromatogram. The mass chromatogram was reconstructed using a neutral loss of 239.4 Da to extract information of only the lipid derivatives. The inset shows excellent near the baseline separation of the derivatised components.

The highly focused reconstruction of the distributions of alkanol components is an important advantage of the derivatisation, due to the ability to target specifically the derivatives despite the presence of other sample matrix components, which may interfere (or co-elute) with the

target analytes during the commonly employed acquisition mode. Additionally, this approach can also be exploited using the multiple reaction monitoring (MRM) technique should only specific sample components be targeted in the analysis. This would greatly increase method response due to the reduction in the duty cycle of the MS.

The chromatographic separation method developed for the purpose of demonstrating the utility of this novel derivatisation approach could be improved to effect the baseline resolution of all of the derivatives. This in turn would allow detection of components not currently detected and which have been confirmed to be present in the sample by the GC-MS analysis.

6.3.5 Conclusions

A novel reversed phase chromatographic method for the screening of the archaeal tetraether lipids has been develop. The method uses a novel column packing material based of the pentafluorophenyl propyl chemistry. The method was demonstrated to perform exceptionally well under the mobile phase used in the separation. Notably, the use of the methyl-tert-butyl ether (MTBE) as the strong solvent in the aqueous-methanol mobile phase afforded base line separation, due to its excellent solubilising properties and relatively high water solubility. The method allows the profiling of cyclopentane containing GDGT lipid cores to be complete in just under 20 min, while achieving baseline separation for the components GDGT-0 to GDGT-6. Further, this method can also be used to monitor the GDGT lipid core derivatives (FLB) as they are eluted in approx. 50 min, while a modest alteration of the mobile phase composition reduces the run time to approx. 40 min.

The derivatisation of the archaeal lipid cores proceed smoothly using the Fmoc-lysine(Boc) amino acid (FLB). The presence of the strong chromophore (Fmoc) allowed the lipid cores to be monitored as derivatives in fluorescence detection. While the signal response in fluorescence was very good, the considerable peak broadening caused by the large detector cell led to a considerable decrease in the method detector response and introduced bias associated with the cumbersome nature of the detection system. The mass spectrometric performance of the FLB derivatives of the GDGT lipids was rather disappointing due to the signal intensity decrease, against the expected improvement observed in previous derivatisations of cholesterol and *r*-dOG diether. While the specific reason for the loss of the signal intensity of the lipid cores derivatives is not known, it appears to be associated with the recent failure of the instrument internal components and the inability to restore its performance to the level that had previously been achieved.

Despite the inability to separate the crenarchaeol and its regioisomer using the novel RP method, the sea surface water temperature (SST) was reconstructed based on the data obtained from the analysis of a Middle Jurassic period (160 Ma) sedimentary deposit using the TEX_{86}^{L} index. The reconstructed temperature of 24.8 °C was 2.6 °C higher compared with 22.2 °C reconstructed using the standard method (NP) and the TEX_{86}^{L} index. Strikingly, it correlated well with the SST of 24.5 °C reconstructed *via* the commonly used TEX_{86} base on the normal phase method. In view of this, it is suggested that further assessment of the

potential of the reversed phase method in aiding the transition from the commonly accepted normal phase mode to reversed phase mode is justified.

Finally, it was demonstrated that the derivatisation approach could be applied in the LC-MS analysis of a broad range of hydroxyl lipids (sterols and alkanols) offering and alternative to the lengthy GC-MS analysis that is commonly employed. It was demonstrated that while the lipid derivatives were very well separated in less than half the time (approx. 20 min) of the GC run, the derivatisation allowed for more lipids to be detected comparing to the GC based separation.

Chapter 7 Conclusions and future work

7.1 Overall Summary and Conclusions

The primary aim of the work presented in this thesis was to develop a derivatisation technique that would enable a high throughput technique screening for Archaeal tetraether lipid cores and other alcohols derived from diverse settings: microbial cultures, marine/lake sediments and soils. To satisfy these aims the work has concentrated of four key areas: (i) development of a facile and effective method for the derivatisation of hydroxyl lipids using amino acids; (ii) evaluation of the chromatographic properties of the lipid derivatives and their responses using different detection techniques; (iii) evaluation of the feasibility of derivatisation in the screening of tetraether lipids cores and other alcohols; (iv) development of a reversed phase separation method for high throughput profiling of the GDGT lipid cores used in calculation of the TEX_{86} index.

Derivatisation techniques currently in use rely on the application of lengthy, complicated protocols that use toxic, expensive reactants and offer inconsistent and moderate efficiency. This study sought to explore the application of the Steglich esterification reaction in preparation of *N*-protected amino acids of a range of alcohols. Improvement to the original reaction conditions, to enhance yield and minimise the reaction time, was achieved by carrying out the reaction under acoustic irradiation. The reaction carried out in a standard sonic cleaning batch resulted in highly reproducible and nearly quantitative yields. Notably, the use of the sonic bath led to reduction of the reaction time to 2 h (or less) compared with several hours under conventional, thermal, conditions. In addition, the product isolation was simplified using a direct silica gel column isolation that is both straightforward and rapid.

Evaluation of the MS behaviour of the *N*-protected amino acid derivatives revealed that they show an increased signal intensity compared with the native species. This was attributed to the presence of the amino acid moiety, which exhibits higher ionisation efficiency and charge stability than the native lipid molecules. Although tandem MS did not offer additional insight into the lipid structures, due to the charge often being retained by the amino acid moiety on cleavage of the ester bond, the Fmoc protected amino acid esters showed the formation of characteristic product ions, suggesting that an MRM scan may be used for highly specific lipid profiling.

The Fmoc protected amino acid derivatives of lipids and alcohols showed excellent signal response in UV and fluorescence, thus opening the way for this type of compound to be monitored using alternative detection modes to the commonly used mass spectrometric methods. Notably, the detection limits for fluorescence detection are comparable to those of MS. Overall, the derivatisation of the lipid models tested in this study resulted in significant increases in signal intensity, of up to two orders of magnitude, which could considerably improve trace level screening of hydroxyl lipids (e.g. cholesterol).

Development of the novel LC-MS method for profiling of the archaeal GDGT lipid cores (especially those containing cyclopentane rings) not only permitted the screening to be completed in under 20 minutes, but also resolved some of the underlying problems associated with existing methods. The considerably reduced analysis time and the ability to inject the total lipid extract dramatically reduce the time from sample preparation to analysis, demonstrating the method to be suitable for use in high throughput research work.

The derivatisation protocol was successfully extended to tetraether lipids from microbial cultures and ancient sediment material. Notably, however, similar enhancements to the MS detection as were observed for lipid models were not realised. This was most likely caused by factors associated with the recent failure of detector components of the MS instrument.

The analysis of the ancient sediment sample (160 Ma) using the reversed phase LC-MS method developed in the study allowed calculation of the TEX_{86}^{L} index values and the subsequent reconstruction of SST. The reconstructed geological temperature was in close agreement with the temperature reconstruction based on the commonly used normal phase TEX_{86} method.

7.2 Future work

The work presented here allowed preparation of novel types of derivatives, and revealed valuable information of their properties such as chromatographic behaviour, detector response and dissociation pathways during CID studies. This information will be invaluable when designing analytical protocols for many existing as well as novel chemical compounds that are difficult to separate from the sample matrix or present challenging chromatographic resolution from co-extracted analytes. In some instances, insufficiently sensitive or incompatible detection techniques may inhibit some analytes from being adequately profiled or quantified, thus requiring a reliance on measurement of relative abundances, which may be biased in a given detection technique due to the inherent properties of the analyte.

The work described in this thesis highlights a number of areas within the field of lipid analysis that merit additional research work.

- The evaluation of the derivatisation technique using a synthetic glycerol trialkyl glycerol tetraether lipid ($GDGT_{46}$) and subsequent characterization of the derivative would provide valuable insight into its behaviour under different detection modes, and help to establish its use as an internal standard for the absolute quantification of the ether lipids using fluorescence detection.

- The reversed phase LC-MS method that was developed for the profiling of the ether lipid cores used a normal bore column, which limits the available range of the mobile phase flow rates and requires rather large volumes of solvents. Reducing of the internal column diameter will resolve those issues and should result in improved resolution. In addition, alteration of the mobile phase composition in conjunction with the use of a narrow bore column may enable separation of the crenarchaeol and its regioisomer, which in turn would permit the use of the standard TEX_{86} index for palaeotemperature reconstruction.

- The successful preparation of the derivatives of cholesterol and other sterols suggests that, with improved chromatographic resolution, it may be possible to directly quantify the amount of cholesterol and its metabolites (coprostanol) instead of having to rely on their ratio for the evaluation of the anthropogenic contribution in environmental and archaeological research.

- The development of a mild and highly effective esterification protocol should also be evaluated as a means of preparing triacylglycerols (TAGs) of desired configurations, which can then be used as inexpensive internal standards in profiling of fats and oils.

Abbreviations

APCI	Atmospheric pressure chemical ionization
Boc	Tert-butyloxycarbonyl
BPC	Base peak chromatogram
CID	Collision induced dissociation
DCC	Dicyclohexylcarbodiimide
DCM	Dichloromethane
DHT	Double hydrogen transfer
DMAP	4-(Dimethylamino)pyridine
EDC	1-(3-(Dimethylamino)propyl)-3-ethyl-carbodiimide
EIC	Extracted ion chromatogram
ELSD	Evaporative light scattering
ESI	Electrospray ionization
FID	Flame ionization detector
FLB	Fmoc-lysine(Boc)
FLD	Fluorescence detector
Fmoc	Fluorenylmethyloxycarbonyl
GC – MS	Gas chromatography mass spectrometry
GC	Gas chromatography
GDD	Glycerol dialkyl diether
GDGT	Glycerol dialkyl glycerol tetraether
GTGT	Glycerol trialkyl glycerol tetraether
HPLC	High performance liquid chromatography
ID	Internal diameter
IPL	Intact polar lipid
LC	Liquid chromatography

LLE	Liquid-liquid extraction
LMCO	Low mass cut-off
LOD	Limit of detection
LOQ	Limit of quantification
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MS^n	Multistage tandem mass spectrometry
MTBE	tert-Butyl methyl ether
N	Column efficiency
n	Number of analytical replicates
NARP	Non-aqueous reversed phase
NL	Neutral loss
NP	Normal phase
PAN – MS	Panoramic MS/MS
PDA	Photodiode array
PFP	Pentafluorophenyl
Phytanyl	3,7,11,15-tetramethylhexadecanyl
q_z	Dimensionless stability parameter; axial solution of the Mathieu equation
r-dOG	1,2 -di-*O*-octadecyl-*rac*-glycerol
R_f	Retention factor
RP	Reversed phase
S/N	Signal – to - noise ratio
SIM	Single ion monitoring
SPE	Solid phase extraction
SST	Sea surface temperature
TEX_{86}	Tetraether index of tetraethers consisting of 86 carbon atoms

TLC	Thin layer chromatography
t_R	Retention time
UHPLC	Ultra high performance liquid chromatography
UV	Ultraviolet
$w_{1/2}$	Width at half height
λ_{EM}	Emission wavelength
λ_{EX}	Excitation wavelength

References

Adair, K. L. and Schwartz, E. (2008) ‘Evidence that ammonia-oxidizing archaea are more abundant than ammonia-oxidizing bacteria in semiarid soils of northern Arizona, USA’, *Microbial Ecology*, 56(3), pp. 420–426.

Ambihapathy, K. *et al.* (1997) ‘Pathways to immonium ions in the fragmentation of protonated peptides’, *Journal of Mass Spectrometry*. Chichester, Sussex, England: John Wiley & Sons, 32(2), pp. 209–215.

Ando, T. and Kimura, T. (1990) ‘Reactivity and selectivity in organic sonochemical reactions involving inorganic solids’, *Ultrasonics*, 28(5), pp. 326–332.

Angel, R. *et al.* (2010) ‘Biogeography of soil archaea and bacteria along a steep precipitation gradient’, *The ISME Journal*, 4(4), pp. 553–563.

Bai, Q. Y. and Zelles, L. (1997) ‘A Method for Determination of Archaeal Ether-Linked Glycerolipids By High Performance Liquid Chromatograpry With Fluorescence Detection As Their 94-Anthroyl Derivatives’, *Chemosphere*, 35(1/2), pp. 263–274.

Bartossek, R. *et al.* (2012) ‘Metagenomic analysis of ammonia-oxidizing archaea affiliated with the soil group’, *Frontiers in Microbiology*, 3(JUN).

Becker, K. W. *et al.* (2013) ‘An improved method for the analysis of archaeal and bacterial ether core lipids’, *Organic Geochemistry*. Elsevier Ltd, 61, pp. 34–44.

Becker, K. W. *et al.* (2015) ‘Rapid and simultaneous analysis of three molecular sea surface temperature proxies and application to sediments from the Sea of Marmara’, *Organic Geochemistry*, 85, pp. 42–53.

Benvegnu, T. *et al.* (2005) ‘Archaeosomes based on novel synthetic tetraether-type lipids for the development of oral delivery systems.’, *Chemical communications (Cambridge, England)*, (44), pp. 5536–5538.

Benvegnu, T., Brard, M. and Plusquellec, D. (2004) ‘Archaeabacteria bipolar lipid analogues: Structure, synthesis and lyotropic properties’, *Current Opinion in Colloid and Interface Science*, 8(6), pp. 469–479.

Berg, I. A. *et al.* (2010) ‘Autotrophic carbon fixation in archaea’, *Nature Reviews Microbiology*, 8(6), pp. 447–460.

Blaga, C. I. *et al.* (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), pp. 523–540.

Bligh, E. G. and Dyer, W. J. (1959) 'A rapid method of total lipid extraction and purification', *Canadian Journal of Biochemistry and Physiology*. NRC Research Press, 37(8), pp. 911–917.

Boden, E. P. and Keck, G. E. (1985) 'Proton-Transfer Steps in Steglich Esterification: A Very Practical New Method for Macrolactonization', *Journal of Organic Chemistry*. American Chemical Society, 50(50), pp. 2394–2395.

Bomberg, M. *et al.* (2008) 'Diversity and function of archaea in freshwater habitats', *Current Trends in Microbiology*, 4, pp. 61–89.

Brochier-Armanet, C. *et al.* (2008) 'Mesophilic crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota', *Nature Reviews Microbiology*, 6(3), pp. 245–252.

Brochier, C. *et al.* (2005) 'Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales?', *Genome Biology*, 6(5), p. R42.

Buckles, L. K. *et al.* (2014) 'Provenance of tetraether membrane lipids in a large temperate lake (Loch Lomond, UK): Implications for glycerol dialkyl glycerol tetraether (GDGT)-based palaeothermometry', *Biogeosciences*, 11(19), pp. 5539–5563.

Cabello, P., Roldan, M. D. and Moreno-Vivian, C. (2004) 'Nitrate reduction and the nitrogen cycle in archaea', *Microbiology*, 150(Pt 11), pp. 3527–3546.

Caffrey, J. M. *et al.* (2007) 'Ammonia oxidation and ammonia-oxidizing bacteria and archaea from estuaries with differing histories of hypoxia', *The ISME Journal*, 1(7), pp. 660–662.

Caforio, A. and Driessen, A. J. M. (2017) 'Archaeal phospholipids: Structural properties and biosynthesis', *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, pp. 1325–1339.

Callieri, C. *et al.* (2009) 'Bacteria, Archaea, and Crenarchaeota in the epilimnion and hypolimnion of a deep holo-oligomictic lake', *Applied and Environmental Microbiology*, 75(22), pp. 7298–7300.

Cañabate-Díaz, B. *et al.* (2007) 'Separation and determination of sterols in olive oil by HPLC-MS', *Food Chemistry*, 102(3), pp. 593–598.

Carraway, K. L. and Koshland, D. E. (1972) 'Carbodiimide Modification of Proteins', *Methods in Enzymology*. Academic Press, 25(C), pp. 616–623.

Castañeda, I. S. *et al.* (2010) 'Millennial-scale sea surface temperature changes in the eastern Mediterranean (Nile River Delta region) over the last 27,000 years', *Paleoceanography*, 25(1).

Cavicchioli, R. (2006) 'Cold-adapted archaea', *Nature Reviews Microbiology*, 4(5), pp. 331–343.

Chaban, B., Ng, S. Y. . and Jarrell, K. F. (2006) 'Archaeal habitats — from the extreme to the ordinary', *Canadian Journal of Microbiology*, 52(2), pp. 73–116.

Chan, L. C. and Cox, B. G. (2007) 'Kinetics of amide formation through carbodiimide/N-hydroxybenzotriazole (HOBt) couplings', *Journal of Organic Chemistry*, 72(23), pp. 8863–8869.

Chang, C.-D. *et al.* (2009) 'Preparation and Properties of Nα-9-Fluorenylmethyloxycarbonylamino Acids Bearing Tert.-Butyl Side Chain Protection', *International Journal of Peptide and Protein Research*, 15(1), pp. 59–66.

Chappe, B. *et al.* (1979) 'Fossil evidence for a novel series of archaebacterial lipids', *Naturwissenschaften*, 66(10), pp. 522–523.

Chappe, B., Albrecht, P. and Michaelis, W. (1982) 'Polar lipids of archaebacteria in sediments and petroleums', *Science*, 217(4554), pp. 65–66.

Chen, S. *et al.* (1999) 'Analysis of 9-Fluorenylmethyloxycarbonyl Derivatives of Catecholamines by High Performance Liquid Chromatography , Liquid Chromatography / Mass Spectrometry and Tandem Mass Spectrometry', *Rapid Communications in Mass Spectrometry*, 13, pp. 1869–1877.

Chen, X. P. *et al.* (2008) 'Ammonia-oxidizing archaea: Important players in paddy rhizosphere soil?', *Environmental Microbiology*, 10(8), pp. 1978–1987.

Chen, Y. Z. *et al.* (2015) 'Determination of cholesterol and four phytosterols in foods without derivatization by gas chromatography-tandem mass spectrometry', *Journal of Food and Drug Analysis*, 23(4), pp. 636–644.

Chong, P. L. G. (2010) ‘Archaebacterial bipolar tetraether lipids: Physico-chemical and membrane properties’, *Chemistry and Physics of Lipids*. Elsevier Ireland Ltd, 163(3), pp. 253–265.

Chong, P. L. G. *et al.* (2012) ‘On physical properties of tetraether lipid membranes: Effects of cyclopentane rings’, *Archaea*, pp. 2–9.

Ciaramella, M., Pisani, F. M. and Rossi, M. (2002) ‘Molecular biology of extremophiles: recent progress on the hyperthermophilic archaeon Sulfolobus’, *Antonie van Leeuwenhoek*, 81(1/4), pp. 85–97.

Cintas, P. and Luche, J.-L. (1999) ‘Green chemistry. The sonochemical approach’, *Green Chemistry*. Royal Society of Chemistry, 1(3), pp. 115–125.

Conrad, R. *et al.* (2010) ‘Methanogenic pathway, 13C isotope fractionation, and archaeal community composition in the sediment of two clear-water lakes of Amazonia’, *Limnology and Oceanography*, 55(2), pp. 689–702.

Coste, J. *et al.* (1991) ‘Oxybenzotriazole free peptide coupling reagents for N-methylated amino acids’, *Tetrahedron Letters*, 32(17), pp. 1967–1970.

Daebeler, A., Gansen, M. and Frenzel, P. (2013) ‘Methyl Fluoride Affects Methanogenesis Rather than Community Composition of Methanogenic Archaea in a Rice Field Soil’, *PLoS ONE*. Edited by J. Neufeld, 8(1), p. e53656.

Dekas, A. E., Poretsky, R. S. and Orphan, V. J. (2009) ‘Deep-Sea Archaea Fix and Share Nitrogen in Methane-Consuming Microbial Consortia’, *Science*, 326(5951), pp. 422–426.

DeLong, E. F. *et al.* (1998) ‘Dibiphytanyl ether lipids in nonthermophilic crenarchaeotes’, *Applied and Environmental Microbiology*, 64(3), pp. 1133–1138.

DeLong, E. F. (2006) ‘Archaeal mysteries of the deep revealed’, *Proceedings of the National Academy of Sciences*, 103(17), pp. 6417–6418.

Demizu, K. *et al.* (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), pp. 135–139.

DeTar, D. F. and Silverstein, R. (1966) ‘Reactions of Carbodiimides. I. The Mechanisms of the Reactions of Acetic Acid with Dicyclohexylcarbodiimide’, *Journal of the American Chemical Society*. ACS Publications, 88(5), pp. 1013–1019.

Dhaon, M. K., Olsen, R. K. and Ramasamy, K. (1982) ‘Esterification of N-Protected a-Amino Acids with Alcohol/Carbodiimide/4-(Dimethylamino)- pyridine. Racemization of Aspartic and Glutamic Acid Derivatives’, *The Journal of Organic Chemistry*, 47(5), pp. 1962–1965.

Du, J. *et al.* (2003) ‘Rearrangement mechanism of the sodium adducts of Fmoc protected amino acids’, *Chinese Science Bulletin*, 48(21), pp. 2317–2319.

Elferink, M. G. L. *et al.* (1994) ‘Stability and proton-permeability of liposomes composed of archaeal tetraether lipids’, *BBA - Biomembranes*, 1193(2), pp. 247–254.

van Es, A. and Stevens, W. (1965) ‘Mixed carboxylic acid anhydrides VI. Synthesis of aryl formates’, *Recueil des Travaux Chimiques des Pays-Bas*. WILEY-VCH Verlag, 84(10), pp. 1247–1252.

Escala, M., Rosell-Melé, A. and 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), pp. 161–164.

Euranto, E. K. (1969) ‘Esterification and ester hydrolysis’, in *Carboxylic Acids and Esters (1969)*. John Wiley & Sons, Ltd., pp. 505–588.

Fabiani, A. *et al.* (2002) ‘High-performance liquid chromatographic analysis of free amino acids in fruit juices using derivatization with 9-fluorenylmethyl-chloroformate.’, *Journal of chromatographic science*, 40(January), pp. 14–18.

El Fakih, H. *et al.* (1997) ‘Wittig-Horner Synthesis of Vinyl Sulfoxides from aryl alkyl or diaryl ketones under sonication’, *Journal für Praktische Chemie/Chemiker-Zeitung*. WILEY-VCH Verlag GmbH, 339(1), pp. 176–178.

Feist, A. M. *et al.* (2006) ‘Modeling methanogenesis with a genome-scale metabolic reconstruction of Methanosarcina barkeri’, *Molecular Systems Biology*, 2.

Fiamegos, Y. C., Nanos, C. G. and Stalikas, C. D. (2004) ‘Ultrasonic-assisted derivatization reaction of amino acids prior to their determination in urine by using single-drop microextraction in conjunction with gas chromatography’, *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, 813(1–2), pp. 89–94.

Fike, D. A., Bradley, A. S. and Rose, C. V. (2015) ‘Rethinking the Ancient Sulfur Cycle’, *Annual Review of Earth and Planetary Sciences*, 43(1), pp. 593–622.

Fischer, E. and Speier, A. (1895) ‘Darstellung der Ester’, *Berichte der deutschen chemischen Gesellschaft*, 28(3), pp. 3252–3258.

Foo Wong, Y. *et al.* (2014) ‘UPLC method for the determination of vitamin E homologues and derivatives in vegetable oils, margarines and supplement capsules using pentafluorophenyl column’, *Talanta*, 130, pp. 299–306.

Francis, C. A. *et al.* (2005) ‘Ubiquity and diversity of ammonia-oxidizing archaea in water columns and sediments of the ocean’, *Proceedings of the National Academy of Sciences*, 102(41), pp. 14683–14688.

Gabriel, J. L. and Chong, P. L. G. (2000) ‘Molecular modeling of archaebacterial bipolar tetraether lipid membranes’, *Chemistry and Physics of Lipids*, 105(2), pp. 193–200.

Garner, G. V. *et al.* (1983) ‘Fast atom bombardment mass spectrometry of butyloxycarbonyl protected (BOC) amino acids’, *Organic Mass Spectrometry*. John Wiley & Sons, Ltd., 18(11), pp. 486–488.

Gilles, V. *et al.* (2015) ‘A new, simple and efficient method of steglich esterification of juglone with long-chain fatty acids: Synthesis of a new class of non-polymeric wax deposition inhibitors for crude oil’, *Journal of the Brazilian Chemical Society*, 26(1), pp. 74–83.

Gliozzi, A. *et al.* (1983) ‘Effect of isoprenoid cyclization on the transition temperature of lipids in thermophilic archaebacteria’, *BBA - Biomembranes*, 735(2), pp. 234–242.

Gliozzi, A., Relini, A. and Chong, P. L. G. (2002) ‘Structure and permeability properties of biomimetic membranes of bolaform archaeal tetraether lipids’, *Journal of Membrane Science*, 206(1–2), pp. 131–147.

Goodby, J. W. *et al.* (2007) ‘Thermotropic liquid crystalline glycolipids’, *Chemical Society Reviews*, 36(12), pp. 1971–2032.

Górnaś, P., Soliven, A. and Segliņa, D. (2015) ‘Seed oils recovered from industrial fruit by-products are a rich source of tocopherols and tocotrienols: Rapid separation of α/β/γ/δ homologues by RP-HPLC/FLD’, *European Journal of Lipid Science and Technology*, 117(6), pp. 773–777.

Han, S. Y. and Kim, Y.-A. (2004) ‘Recent development of peptide coupling reagents in organic synthesis’, *Tetrahedron*, 60(11), pp. 2447–2467.

Hassner, A. and Alexanian, V. (1978) ‘Direct room temperature esterification of carboxylic acids’, *Tetrahedron Letters*, 19(46), pp. 4475–4478.

Hassner, A., Krepski, L. R. and Alexanian, V. (1978) ‘Aminopyridines as acylation catalysts for tertiary alcohols’, *Tetrahedron*, 34(14), pp. 2069–2076.

Henderson, J. W. *et al.* (2000) *Rapid, accurate, sensitive, and reproducible HPLC analysis of amino acids, Technical note*.

Hicks, S. A. (2016) *Fate and preservation of lipids in the soils of archaeological and experimental burials*. University of York.

Hoefs, M. J. L. *et al.* (1997) ‘Ether lipids of planktonic archaea in the marine water column?’, *Applied and Environmental Microbiology*, 63(8), pp. 3090–3095.

Höfle, G., Steglich, W. and Vorbrüggen, H. (1978) ‘4-Dialkylaminopyridines as Highly Active Acylation Catalysts.’, *Angewandte Chemie International Edition*, 17(8), pp. 569–583.

Høj, L., Olsen, R. A. and Torsvik, V. L. (2008) ‘Effects of temperature on the diversity and community structure of known methanogenic groups and other archaea in high Arctic peat’, *The ISME Journal*, 2(1), pp. 37–48.

Hopmans, E. C. *et al.* (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), pp. 585–589.

Hopmans, E. C. *et al.* (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), pp. 107–116.

Hopmans, E. C., Schouten, S. and Sinninghe Damsté, J. S. (2016) ‘The effect of improved chromatography on GDGT-based palaeoproxies’, *Organic Geochemistry*, 93, pp. 1–6.

Ter Horst, B. *et al.* (2010) ‘Asymmetric synthesis and structure elucidation of a glycerophospholipid from Mycobacterium tuberculosis’, *Journal of lipid research*, 51(5), pp. 1017–1022.

Huang, Z. H. *et al.* (1993) 'Characterization of N-ethoxycarbonyl ethyl esters of amino acids by mass spectrometry', *Journal of Chromatography A*. Elsevier, 635(2), pp. 271–281.

Hugoni, M. *et al.* (2013) 'Dynamics of ammonia-oxidizing Archaea and Bacteria in contrasted freshwater ecosystems', *Research in Microbiology*, 164(4), pp. 360–370.

Huguet, C. *et al.* (2006) 'An improved method to determine the absolute abundance of glycerol dibiphytanyl glycerol tetraether lipids', *Organic Geochemistry*, 37(9), pp. 1036–1041.

Huguet, C. *et al.* (2007) 'A study of the TEX 86 paleothermometer in the water column and sediments of the Santa Barbara Basin, California', *Paleoceanography*, 22, PA3203.

Huguet, C. *et al.* (2009) 'Effects of long term oxic degradation on the U37 K′, TEX86 and BIT organic proxies', *Organic Geochemistry*, 40(12), pp. 1188–1194.

Huguet, C., Fietz, S. and Rosell-Melé, A. (2013) 'Global distribution patterns of hydroxy glycerol dialkyl glycerol tetraethers', *Organic Geochemistry*, 57, pp. 107–118.

Inanaga, J. *et al.* (1979) 'A Rapid Esterification by Means of Mixed Anhydride and Its Application to Large-ring Lactonization', *Bulletin of the Chemical Society of Japan*, pp. 1989–1993.

Inglis, G. N. *et al.* (2015) 'Descent toward the Icehouse: Eocene sea surface cooling inferred from GDGT distributions', *Paleoceanography*. Wiley Online Library, 30(7), pp. 1000–1020.

Iwasaki, T. *et al.* (2000) 'Esterification', in *Kirk-Othmer Encyclopedia of Chemical Technology*. John Wiley & Sons, Inc., pp. 1–33.

Jaita, S., Phakhodee, W. and Pattarawarapan, M. (2015) 'Ultrasound-Assisted Methyl Esterification of Carboxylic Acids -Catalyzed by Polymer-Supported Triphenylphosphine', *Synlett*. 09.07.2015, 26(14), pp. 2006–2008.

Jiang, H. *et al.* (2009) 'Diversity and abundance of ammonia-oxidizing archaea and bacteria in Qinghai Lake, Northwestern China', *Geomicrobiology Journal*, 26(3), pp. 199–211.

Jin, B. *et al.* (2014) 'Lewis acid-catalyzed in situ transesterification/esterification of microalgae in supercritical ethanol', *Bioresource Technology*. Elsevier, 162, pp. 341–349.

De Jonge, C. *et al.* (2013) ‘Identification of novel penta- and hexamethylated branched glycerol dialkyl glycerol tetraethers in peat using HPLC-MS2, GC-MS and GC-SMB-MS’, *Organic Geochemistry*. Elsevier Ltd, 54, pp. 78–82.

Joo, C. N., Shier, T. and Kates, M. (1968) ‘Characterization and synthesis of mono- and diphytanyl ethers of glycerol.’, *Journal of lipid research*, 9(6), pp. 782–788.

Joullie, M. M. and Lassen, K. M. (2010) ‘Evolution of amide bond formation’, *Arkivoc*, 2010(8), pp. 189–250.

Kan, J. *et al.* (2011) ‘Archaea in Yellowstone Lake’, *The ISME Journal*, 5(11), pp. 1784–1795.

Kan, J. *et al.* (2016) ‘Geochemistry and mixing drive the spatial distribution of free-living archaea and bacteria in yellowstone lake’, *Frontiers in Microbiology*, 7(FEB).

Kantharaju, K. and Babu, V. V. S. (2006) ‘Ultrasound Accelerated Synthesis of Proteinogenic and α,α-Dialkylamino Acid Ester Salts.’, *ChemInform*, 37(50).

Karner, M. B., DeLong, E. F. and Karl, D. M. (2001) ‘Archaeal dominance in the mesopelagic zone of the Pacific Ocean.’, *Nature*, 409, pp. 507–510.

Karr, E. A. *et al.* (2006) ‘Biodiversity of methanogenic and other Archaea in the permanently frozen Lake Fryxell, Antarctica’, *Applied and Environmental Microbiology*, 72(2), pp. 1663–1666.

Kates, M. (1963) ‘Isolation and characterization of a diether analog of phosphatidyl glycerophosphate from Halobacteriurn cutirubrum’, *Biochmica Et Biophysica Acta*, 70, pp. 705–707.

Kates, M. (1977) ‘The phytanyl ether-linked polar lipids and isoprenoid neutral lipids of extremely halophilic bacteria’, *Progress in the Chemistry of Fats and Other Lipids*, pp. 301–342.

Kenig, F. *et al.* (1994) ‘Isotopic biogeochemistry of the Oxford Clay Formation (Jurassic), UK’, *Journal of the Geological Society*, 151(1), pp. 139–152.

Khajuria, R. K. *et al.* (2007) ‘Development of a Rapid Normal-Phase LC-Positive Ion APCI-MS Method for Simultaneous Detection and Quantitation of Cholesterol, Androst-4-ene-3, 17-dione, and Androsta-1, 4-diene-3, 17-dione’, *Journal of chromatographic science*. Preston Publication Inc, 45(8), p. 519.

Kim, J.-H. *et al.* (2008) 'Global sediment core-top calibration of the TEX86 paleothermometer in the ocean', *Geochimica et Cosmochimica Acta*, 72(4), pp. 1154–1173.

Kim, J.-H. *et al.* (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), pp. 4639–4654.

Kletzin, A. *et al.* (2004) 'Dissimilatory Oxidation and Reduction of Elemental Sulfur in Thermophilic Archaea', *Journal of Bioenergetics and Biomembranes*, pp. 77–91.

Kletzin, A. (2007) 'Metabolism of Inorganic Sulfur Compounds in Archaea', in *Archaea: Evolution, Physiology, and Molecular Biology*, pp. 261–274.

Knappy, C. S. *et al.* (2012) 'Structural complexity in isoprenoid glycerol dialkyl glycerol tetraether lipid cores of Sulfolobus and other archaea revealed by liquid chromatography-tandem mass spectrometry', *Chemistry and Physics of Lipids*. Elsevier Ireland Ltd, 165(6), pp. 648–655.

Knappy, C. S. *et al.* (2014) 'Identification of homoglycerol- and dihomoglycerol-containing isoprenoid tetraether lipid cores in aquatic sediments and a soil', *Organic Geochemistry*. Elsevier Ltd, 76, pp. 146–156.

Knappy, C. S. *et al.* (2015) 'Mono-, di- and trimethylated homologues of isoprenoid tetraether lipid cores in archaea and environmental samples: Mass spectrometric identification and significance', *Journal of Mass Spectrometry*, 50(12), pp. 1420–1432.

Knappy, C. S., Chong, J. P. J. and Keely, B. J. (2009) 'Rapid Discrimination of Archaeal Tetraether Lipid Cores by Liquid Chromatography-Tandem Mass Spectrometry', *Journal of the American Society for Mass Spectrometry*, 20(1), pp. 51–59.

Knappy, C. S. and Keely, B. J. (2012) 'Novel glycerol dialkanol triols in sediments: transformation products of glycerol dibiphytanyl glycerol tetraether lipids or biosynthetic intermediates?', *Chemical Communications*, 48(6), pp. 841–843.

Koenig, W. and Geiger, R. (1970) 'A new method for synthesis of peptides: activation of the carboxyl group with dicyclohexylcarbodiimide using 1-hydroxybenzotriazoles as additives', *Chemische Berichte*, 103(3), p. 788.

Koga, Y. *et al.* (1993) 'Ether polar lipids of methanogenic bacteria: structures, comparative aspects, and biosyntheses.', *Microbiological reviews*, 57(1), pp. 164–182.

Koga, Y. and Morii, H. (2007) ‘Biosynthesis of ether-type polar lipids in archaea and evolutionary considerations.’, *Microbiology and molecular biology reviews : MMBR*, 71(1), pp. 97–120.

Komatsu, H. and Chong, P. L. G. (1998) ‘Low permeability of liposomal membranes composed of bipolar tetraether lipids from thermoacidophilic archaebacterium Sulfolobus acidocaldarius’, *Biochemistry*, 37(1), pp. 107–115.

Kotsyurbenko, O. R. *et al.* (2004) ‘Acetoclastic and hydrogenotrophic methane production and methanogenic populations in an acidic West-Siberian peat bog’, *Environmental Microbiology*, 6(11), pp. 1159–1173.

Lai, D., Springstead, J. R. and Monbouquette, H. G. (2008) ‘Effect of growth temperature on ether lipid biochemistry in Archaeoglobus fulgidus’, *Extremophiles*, 12(2), pp. 271–278.

Larionov, E. (2011) *Activity and selectivity of DMAP derivatives in acylation reactions: experimental and theoretical studies, Synthesis*.

Lee, K. E. *et al.* (2008) ‘A study of the alkenone, TEX86, and planktonic foraminifera in the benguela upwelling system: Implications for past sea surface temperature estimates’, *Geochemistry, Geophysics, Geosystems*, 9(10).

Lehours, A. C. *et al.* (2007) ‘Phylogenetic diversity of archaea and bacteria in the anoxic zone of a meromictic lake (Lake Pavin, France)’, *Applied and Environmental Microbiology*, 73(6), pp. 2016–2019.

Leigh, J. A. *et al.* (2011) ‘Model organisms for genetics in the domain Archaea: methanogens, halophiles, Thermococcales and Sulfolobales.’, *FEMS microbiology reviews*, 35(4), pp. 577–608.

Leininger, S. *et al.* (2006) ‘Archaea predominate among ammonia-oxidizing prokaryotes in soils’, *Nature*, 442(7104), pp. 806–809.

Lengger, S. K. *et al.* (2012) ‘Comparison of extraction and work up techniques for analysis of core and intact polar tetraether lipids from sedimentary environments’, *Organic Geochemistry*. Elsevier Ltd, 47, pp. 34–40.

Lepage, G. and Roy, C. C. (1986) ‘Direct transesterification of all classes of lipids in a one-step reaction.’, *Journal of Lipid Research* , 27(1), pp. 114–120.

Li, J. and Sha, Y. (2008) ‘A Convenient Synthesis of Amino Acid Methyl Esters’, *Molecules*. Molecular Diversity Preservation International, 13(5), pp. 1111–1119.

Lillington, J. M., Trafford, D. J. H. and Makin, H. L. J. (1981) 'A rapid and simple method for the esterification of fatty acids and steroid carboxylic acids prior to gas-liquid chromatography', *Clinica Chimica Acta*. Elsevier Ltd, 111(1), pp. 91–98.

Lim, K. L. H. *et al.* (2012) 'Archaeol: An indicator of methanogenesis in water-saturated soils', *Archaea*, 2012.

Liu, X. L., Lipp, J. S. and Hinrichs, K. U. (2011) 'Distribution of intact and core GDGTs in marine sediments', *Organic Geochemistry*. Elsevier Ltd, 42(4), pp. 368–375.

Liu, X. L., Summons, R. E. and Hinrichs, K. U. (2012) 'Extending the known range of glycerol ether lipids in the environment: Structural assignments based on tandem mass spectral fragmentation patterns', *Rapid Communications in Mass Spectrometry*, 26(19), pp. 2295–2302.

Liu, Y., Beer, L. L. and Whitman, W. B. (2012) 'Sulfur metabolism in archaea reveals novel processes', *Environmental Microbiology*, 14(10), pp. 2632–2644.

Liu, Z. *et al.* (2009) 'Global Cooling During the Eocene-Oligocene Climate Transition', *Science*, 323(5918), pp. 1187–1190.

Lliros, M. (2010) 'Diversity, dynamics and activity of mesophilic Archaea in stratified freshwater lakes: Implications in biogeochemical cycles', *Institute of Aquatic Ecology*.

Lu, T.-J., Cheng, S.-M. and Sheu, L.-J. (1998) 'Ultrasound Accelerated Coupling Reaction of Grignard Reagents with 1,3-Dioxolanes of α,β-Unsaturated Aldehydes', *The Journal of Organic Chemistry*. American Chemical Society, 63(8), pp. 2738–2741.

Luche, J.-L. *et al.* (1990) 'Organic sonochenistry : A new interpretation and its consequences', *Tetrahedron Letters*, 31(29), pp. 4125–4128.

Luque de Castro, M. D., Priego-Capote, F. and Peralbo-Molina, A. (2011) 'The role of ultrasound in analytical derivatizations', *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, pp. 1189–1195.

Mancuso, C. A., Nichols, P. D. and White, D. C. (1986) 'A method for the separation and characterization of archaebacterial signature ether lipids.', *Journal of Lipid Research*, 27(1), pp. 49–56.

Martz, R. F., Sebacher, D. I. and White, D. C. (1983) 'Biomass measurement of methane forming bacteria in environmental samples', *Journal of Microbiological Methods*, 1(1), pp. 53–61.

Mason, T. J. (1997) ‘Ultrasound in synthetic organic chemistry’, *Chemical Society Reviews*. The Royal Society of Chemistry, 26(6), pp. 443–451.

Melucci, D. *et al.* (1999) ‘FMOC-Cl as derivatizing agent for the analysis of amino acids and dipeptides by the absolute analysis method’, *Chromatographia*, 49(5), pp. 317–320.

Metje, M. and Frenzel, P. (2005) ‘Effect of temperature on anaerobic ethanol oxidation and methanogenesis in acidic peat from a Northern Wetland’, *Applied and Environmental Microbiology*, 71(12), pp. 8191–8200.

Metje, M. and Frenzel, P. (2007) ‘Methanogenesis and methanogenic pathways in a peat from subarctic permafrost’, *Environmental Microbiology*, 9(4), pp. 954–964.

Michaelis, W. and Albrecht, P. (1979) ‘Molecular fossils of archaebacteria in kerogen’, *Naturwissenschaften*, 66(8), pp. 420–422.

Mineno, T. and Kansui, H. (2006) ‘High yielding methyl esterification catalyzed by indium(III) chloride.’, *Chemical & Pharmaceutical Bulletin*, 54(6), pp. 918–9.

Mueller, T. J. *et al.* (2014) ‘Methane oxidation by anaerobic archaea for conversion to liquid fuels’, *Journal of Industrial Microbiology and Biotechnology*, 42(3), pp. 391–401.

Nachtmann, F., Spitzy, H. and Frei, R. W. (1975) ‘Fluorescence derivatization for trace determination of some alkaloids and adrenaline’, *Analytica Chimica Acta*, 76(1), pp. 57–64.

Nahmany, M. and Melman, A. (2004) ‘Chemoselectivity in reactions of esterification.’, *Organic & biomolecular chemistry*, 2(11), pp. 1563–1572.

Naik, S. D. and Doraiswamy, L. K. (1998) ‘Phase transfer catalysis: Chemistry and engineering’, *AIChE Journal*, 44(3), pp. 612–646.

Nakamura, E., Imanishi, Y. and Machii, D. (1994) ‘Sonochemical Initiation of Radical Chain Reactions. Hydrostannation and Hydroxystannation of CC Multiple Bonds’, *The Journal of Organic Chemistry*. ACS Publications, 59(26), pp. 8178–8186.

Nascentes, C. C. *et al.* (2001) ‘Use of ultrasonic baths for analytical applications: a new approach for optimisation conditions’, *Journal of the Brazilian Chemical Society*, 12, pp. 57–63.

Neises, B. and Steglich, W. (1978) ‘Simple Method for the Esterification of Carboxylic Acids’, *Angewandte Chemie International Edition*, 17(7), pp. 522–524.

Neises, B. and Steglich, W. (1985) ‘Esterification of Carboxylic Acids with Dicyclohexylcarbodiimide/4-Dimethylaminopyridine: tert-Butyl Ethyl Fumarate’, *Organic Syntheses*. Wiley Online Library, p. 183.

Nichols, D. S. *et al.* (2004) ‘Cold adaptation in the Antarctic archaeon Methanococcoides burtonii involves membrane lipid unsaturation’, *Journal of Bacteriology*. American Society for Microbiology, 186(24), pp. 8508–8515.

Nichols, P. D. *et al.* (1993) ‘Analysis of archaeal phospholipid-derived di- and tetraether lipids by high temperature capillary gas chromatography’, *Journal of Microbiological Methods*, 18(1), pp. 1–9.

Nishihara, M. and Koga, Y. (1987) ‘Extraction and Composition of Polar Lipids from the Archaebacterium, Methanobacterium thermoautotrophicum: Effective Extraction of Tetraether Lipids by an Acidified Solvent1’, *The Journal of Biochemistry*, 101(4), pp. 997–1005.

Offre, P., Spang, A. and Schleper, C. (2013) ‘Archaea in Biogeochemical Cycles’, *Annual Review of Microbiology*, 67(1), pp. 437–457.

Oger, P. M. and Cario, A. (2013) ‘Adaptation of the membrane in Archaea’, *Biophysical Chemistry*. Elsevier B.V., 183, pp. 42–56.

Ohtsubo, S. *et al.* (1993) ‘A sensitive method for quantification of aceticlastic methanogens and estimation of total methanogenic cells in natural environments based on an analysis of ether-linked glycerolipids’, *FEMS Microbiol Ecol*, 12(1), pp. 39–50.

Otera, J. and Nishikido, J. (2010) *Esterification: Methods, Reactions, and Applications: Second Edition, Esterification: Methods, Reactions, and Applications: Second Edition*.

van den Ouweland, J. M. W., Beijers, A. M. and van Daal, H. (2011) ‘Fast Separation of 25-Hydroxyvitamin D3 from 3-Epi-25-Hydroxyvitamin D3 in Human Serum by Liquid Chromatography–Tandem Mass Spectrometry: Variable Prevalence of 3-Epi-25-Hydroxyvitamin D3 in Infants, Children, and Adults’, *Clinical Chemistry*, 57(11), p. 1618 LP-1619.

Pace, C. N. *et al.* (1995) ‘How to measure and predict the molar absorption coefficient of a protein.’, *Protein Science : A Publication of the Protein Society*. Cold Spring Harbor Laboratory Press, 4(11), pp. 2411–2423.

Paltauf, F. and Spener, F. (1968) ‘An improved synthesis of 1,2-dialkyl glycerol ethers and the synthesis of 14C-labelled trialkyl glycerol ethers’, *Chemistry and Physics of Lipids*. Elsevier, 2(2), pp. 168–172.

Papayannopoulos, I. A. (1995) ‘The interpretation of collision-induced dissociation tandem mass spectra of peptides’, *Mass Spectrometry Reviews*, 14(1), pp. 49–73.

Pchelintsev, N. A., Adams, P. D. and Nelson, D. M. (2016) ‘Critical Parameters for Efficient Sonication and Improved Chromatin Immunoprecipitation of High Molecular Weight Proteins’, *PLoS ONE*. Edited by M. Wu. San Francisco, CA USA: Public Library of Science, 11(1).

Pearson, A. *et al.* (2008) ‘Factors controlling the distribution of archaeal tetraethers in terrestrial hot springs’, *Applied and Environmental Microbiology*, 74(11), pp. 3523–3532.

Pitcher, A. *et al.* (2009) ‘Separation of core and intact polar archaeal tetraether lipids using silica columns: Insights into living and fossil biomass contributions’, *Organic Geochemistry*. Elsevier Ltd, 40(1), pp. 12–19.

Pitcher, A. *et al.* (2010) ‘Crenarchaeol dominates the membrane lipids of Candidatus Nitrososphaera gargensis, a thermophilic group I.1b Archaeon.’, *The ISME journal*, 4(4), pp. 542–52.

Powers, L. *et al.* (2004) ‘Crenarchaeotal membrane lipids in lake sediments: A new paleotemperature proxy continental paleoclimate reconstruction?’, *Geology*, 32(7), pp. 613–616.

Powers, L. *et al.* (2010) ‘Applicability and calibration of the TEX86 paleothermometer in lakes’, *Organic Geochemistry*, 41(4), pp. 404–413.

Pratscher, J., Dumont, M. G. and Conrad, R. (2011) ‘Ammonia oxidation coupled to CO2 fixation by archaea and bacteria in an agricultural soil’, *Proceedings of the National Academy of Sciences*, 108(10), pp. 4170–4175.

Puglisi, E. *et al.* (2014) ‘Changes in bacterial and archaeal community assemblages along an ombrotrophic peat bog profile’, *Biology and Fertility of Soils*, 50(5), pp. 815–826.

Raguse, B. *et al.* (2000) ‘The synthesis of archaebacterial lipid analogues’, *Tetrahedron Letters*, 41(16), pp. 2971–2974.

Raith, K. *et al.* (2005) ‘A new LC/APCI-MS method for the determination of cholesterol oxidation products in food’, *Journal of Chromatography A*, 1067(1–2), pp. 207–211.

Ramesh, M. *et al.* (2010) 'Characterization of Nα-Fmoc-protected ureidopeptides by electrospray ionization tandem mass spectrometry (ESI-MS/MS): differentiation of positional isomers.', *Journal of Mass Spectrometry*, 45(12), pp. 1461–72.

Ramesh, M. *et al.* (2011) 'Characterization of Nα-Fmoc-protected dipeptide isomers by electrospray ionization tandem mass spectrometry (ESI-MSn): effect of protecting group on fragmentation of dipeptides', *Rapid Communications in Mass Spectrometry*, 25(14), pp. 1949–1958.

Rattray, J., Muschitiello, F. and Smittenberg, R. (2013) *A comparison of HPLC MS methods for GDGT analysis ; should we make the switch ?*, Poster.

Raymann, K., Brochier-Armanet, C. and Gribaldo, S. (2015) 'The two-domain tree of life is linked to a new root for the Archaea', *Proceedings of the National Academy of Sciences*, 112(21), pp. 6670–6675.

Reddy, P. N. *et al.* (2007) 'Positive and Negative Ion Electrospray Tandem Mass Spectrometry (ESI MS/MS) of Boc-Protected Peptides Containing Repeats of L-Ala-γ4Caa/γ4Caa-L-Ala: Differentiation of Some Positional Isomeric Peptides', *Journal of the American Society for Mass Spectrometry*, 18(4), pp. 651–662.

Richheimer, S. L., Tinnermeier, D. M. and Timmons, D. W. (1992) 'High-performance liquid chromatographic assay of taxol', *Analytical Chemistry*. American Chemical Society, 64(20), pp. 2323–2326.

Richter, B. E. *et al.* (1996) 'Accelerated Solvent Extraction: A Technique for Sample Preparation', *Analytical Chemistry*. American Chemical Society, 68(6), pp. 1033–1039.

De Rosa, M. *et al.* (1977) 'Lipid structures in the Caldariella group of extreme thermoacidophile bacteria', *Journal of the Chemical Society, Chemical Communications*. The Royal Society of Chemistry, (15), pp. 514–515.

De Rosa, M., Gambacorta, A., *et al.* (1980) 'Complex lipids of Caldariella acidophila, a thermoacidophile archaebacterium', *Phytochemistry*, 19(5), pp. 821–825.

De Rosa, M., Esposito, E., *et al.* (1980) 'Effects of temperature on ether lipid composition of Caldariella acidophila', *Phytochemistry*, 19(5), pp. 827–831.

De Rosa, M. *et al.* (1983) 'Isoprenoid ethers; backbone of complex lipids of the archaebacterium Sulfolobus solfataricus', *Biochimica et Biophysica Acta (BBA)/Lipids and Lipid Metabolism*, 753(2), pp. 249–256.

De Rosa, M. and Gambacorta, A. (1988) ‘The lipids of archaebacteria’, *Progress in Lipid Research*, pp. 153–175.

Saito, Y. *et al.* (2005) ‘Simple and mild esterification of N-protected amino acids with nearly equimolar amounts of alcohols using 1-tert-butoxy-2-tert-butoxycarbonyl-1,2-dihydroisoquinoline’, *Tetrahedron Letters*, 46(8), pp. 1277–1279.

Saitoh, K., Shiina, I. and Mukaiyama, T. (1998) ‘O,O′-Di(2-pyridyl) Thiocarbonate as an Efficient Reagent for the Preparation of Carboxylic Esters from Highly Hindered Alcohols’, *Chemistry Letters*. The Chemical Society of Japan, 27(7), pp. 679–680.

Saldanha, T. *et al.* (2006) ‘HPLC Separation and Determination of 12 Cholesterol Oxidation Products in Fish: Comparative Study of RI, UV, and APCI-MS Detectors’, *Journal of Agricultural and Food Chemistry*. American Chemical Society, 54(12), pp. 4107–4113.

Santoro, A. E. *et al.* (2008) ‘Shifts in the relative abundance of ammonia-oxidizing bacteria and archaea across physicochemical gradients in a subterranean estuary’, *Environmental Microbiology*, 10(4), pp. 1068–1079.

Santos, A. M., Martinez, M. and Mira, J. A. (1996) ‘Comparison Study of Lewis Acid Type Catalysts on the Esterification of Octanoic Acid and n-Octyl Alcohol’, *Chemical engineering & technology*, 9, pp. 538–542.

Schmidt, F., Hinrichs, K. U. and Elvert, M. (2010) ‘Sources, transport, and partitioning of organic matter at a highly dynamic continental margin’, *Marine Chemistry*. Elsevier B.V., 118(1–2), pp. 37–55.

Schouten, S. *et al.* (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), pp. 14421–14426.

Schouten, S. *et al.* (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), pp. 265–274.

Schouten, S. *et al.* (2003) ‘Biogeochemical Evidence that Thermophilic Archaea Mediate the Anaerobic Oxidation of Methane Biogeochemical Evidence that Thermophilic Archaea Mediate the Anaerobic Oxidation of Methane’, *Applied and Environmental Microbiology*, 69(3), pp. 1680–1686.

Schouten, S. *et al.* (2007) 'Analytical Methodology for TEX 86 Paleothermometry by High-Performance Liquid Chromatography / Atmospheric Pressure Chemical Ionization-Mass Spectrometry', *Analytical chemistry*, 79(7), pp. 2940–2944.

Schouten, S. *et al.* (2008) 'An unusual isoprenoid tetraether lipid in marine and lacustrine sediments', *Organic Geochemistry*, Elsevier Ltd, 39(8), pp. 1033–1038.

Schouten, S., Hopmans, E. C. and Sinninghe Damsté, J. S. (2013) 'The organic geochemistry of glycerol dialkyl glycerol tetraether lipids: A review', *Organic Geochemistry*. Elsevier Ltd, 54, pp. 19–61.

Shah, S. R. *et al.* (2008) 'Origins of archaeal tetraether lipids in sediments: Insights from radiocarbon analysis', *Geochimica et Cosmochimica Acta*, 72(18), pp. 4577–4594.

Sheehan, J. C., Cruickshank, P. a. and Boshart, G. L. (1961) 'Convenient Synthesis of Water-Soluble Carbodiimides', *Journal of Organic Chemistry*, 26(7), pp. 2525–2528.

Sheehan, J. C. and Hess, G. P. (1955) 'A New Method of Forming Peptide Bonds', *Journal of the American Chemical Society*. American Chemical Society, 77(4), pp. 1067–1068.

Sheikh, M. C. *et al.* (2010) 'Mechanistic studies of DCC/HOBt-mediated reaction of 3-phenylpropionic acid with benzyl alcohol and studies on the reactivities of "active ester" and the related derivatives with nucleophiles', *Tetrahedron*. Elsevier Ltd, 66(36), pp. 7272–7278.

Shimada, H. *et al.* (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), pp. 556–563.

Shimada, H. *et al.* (2008) 'Effects of pH and temperature on the composition of polar lipids in *Thermoplasma acidophilum* HO-62', *Journal of Bacteriology*, 190(15), pp. 5404–5411.

Sinninghe Damsté, J. S. *et al.* (2002) 'Crenarchaeol: the characteristic core glycerol dibiphytanyl glycerol tetraether membrane lipid of cosmopolitan pelagic crenarchaeota', *The Journal of Lipid Research*, 43(10), pp. 1641–1651.

Sinninghe Damsté, J. S. *et al.* (2014) 'Ether- and ester-bound iso-diabolic acid and other lipids in members of Acidobacteria subdivision 4', *Applied and Environmental Microbiology*, 80(17), pp. 5207–5218.

Sirsam, R., Hansora, D. and Usmani, G. A. (2016) 'A Mini-Review on Solid Acid Catalysts for Esterification Reactions', *Journal of The Institution of Engineers (India): Series E*, 97(2), pp. 167–181.

Sluijs, A. *et al.* (2006) 'Subtropical Arctic Ocean temperatures during the Palaeocene/Eocene thermal maximum', *Nature*, 441(7093), pp. 610–613.

Smallwood, I. M. (1996) *Handbook of Organic Solvent Properties*. Oxford: Wiley.

Son, H.-H. *et al.* (2014) 'High-temperature GC-MS-based serum cholesterol signatures may reveal sex differences in vasospastic angina', *Journal of Lipid Research*, 55(1), pp. 155–162.

Sonntag, N. O. V (1953) 'The Reactions of Aliphatic Acid Chlorides.', *Chemical Reviews*. American Chemical Society, 52(2), pp. 237–416.

Spang, A. *et al.* (2010) 'Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum Thaumarchaeota', *Trends in Microbiology*, 18(8), pp. 331–340.

Sprott, G. D. *et al.* (1993) 'Hydroxydiether lipid structures in Methanosarcina spp. and Methanococcus voltae', *Applied and Environmental Microbiology*, 59(3), pp. 912–914.

Srikrishna, A., Nagaraju, S. and Sharma, G. V. R. (1992) 'Sonochemical Acceleration of Conversion of 2-Alkoxytetrahydrofurans to γ-Butyrolactones Synthesis of (±)-Quercus Lactone-A', *Synthetic Communications*. Taylor & Francis, 22(8), pp. 1127–1135.

Stahl, D. A. and de la Torre, J. R. (2012) 'Physiology and Diversity of Ammonia-Oxidizing Archaea', *Annual Review of Microbiology*, 66(1), pp. 83–101.

Sturt, H. F. *et al.* (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 : RCM*, 18(6), pp. 617–28.

Sunggak, K., Jae, I. L. and Young, K. K. (1984) 'Di-2-pyridyl carbonate: a new efficient coupling agent for the direct esterification of carboxylic acids', *Tetrahedron Letters*. Pergamon, 25(43), pp. 4943–4946.

Terme, N. *et al.* (2014) 'Modification of bipolar lipid conformation at the air/water interface by a single stereochemical variation', *Chemistry and Physics of Lipids*. Elsevier Ireland Ltd, 183, pp. 9–17.

Teske, A., Dhillon, A. and Sogin, M. L. (2003) 'Genomic markers of ancient anaerobic microbial pathways: Sulfate reduction, methanogenesis, and methane oxidation', in *Biological Bulletin*, pp. 186–191.

Thauer, R. K. *et al.* (2008) 'Methanogenic archaea: ecologically relevant differences in energy conservation', *Nature Reviews Microbiology*, 6(8), pp. 579–591.

Thompson, D. H. *et al.* (1994) 'Synthesis of Chiral Diether and Tetraether Phospholipids - Regiospecific Ring-Opening of Epoxy Alcohol Intermediates Derived From Asymmetric Epoxidation', *Journal of Organic Chemistry*, 59(15), pp. 2945–2955.

Tornabene, T. G. and Langworthy, T. A. (1979) 'Diphytanyl and dibiphytanyl glycerol ether lipids of methanogenic archaebacteria', *Science*, 203(4375), p. 51 LP-53.

Tourna, M. *et al.* (2008) 'Growth, activity and temperature responses of ammonia-oxidizing archaea and bacteria in soil microcosms', *Environmental Microbiology*, 10(5), pp. 1357–1364.

Trincone, A. *et al.* (1992) 'Distribution of Complex and Core Lipids within New Hyperthermophilic Members of the Archaea Domain', *Systematic and Applied Microbiology*, 15(1), pp. 11–17.

Trommer, G. *et al.* (2009) 'Distribution of Crenarchaeota tetraether membrane lipids in surface sediments from the Red Sea', *Organic Geochemistry*. Elsevier Ltd, 40(6), pp. 724–731.

Tsakos, M. *et al.* (2015) 'Ester coupling reactions – an enduring challenge in the chemical synthesis of bioactive natural products', *Natural Products Report*, pp. 605–632.

Tsang, C. W. and Harrison, A. G. (1976) 'Chemical ionization of amino acids', *Journal of the American Chemical Society*. American Chemical Society, 98(6), pp. 1301–1308.

Uda, I. *et al.* (2001) 'Variation in Molecular Species of Polar Lipids from Thermoplasma acidophilum Depends on Growth Temperature.', *Lipids*, 36(1), pp. 103–105.

Vairamani, M. *et al.* (1990) 'Mass spectra oft-butyloxycarbonyl (BOC)-protected peptides', *Organic Mass Spectrometry*, 25(2), pp. 97–100.

Valeur, E. and Bradley, M. (2009) 'Amide bond formation: beyond the myth of coupling reagents', *Chem Soc Rev*, 38(2), pp. 606–631.

VandenBurg, D. and Price, G. J. (2012) 'Ultrasound promoted Wurtz coupling of alkyl bromides and dibromides', *Ultrasonics Sonochemistry*, 19(1), pp. 5–8.

Villanueva, L., Sinninghe Damsté, J. S. and Schouten, S. (2014) ‘A re-evaluation of the archaeal membrane lipid biosynthetic pathway.’, *Nature reviews. Microbiology*. Nature Publishing Group, 12(6), pp. 438–48.

Villeneuve, G. B. and Chan, T. H. (1997) ‘A rapid, mild and acid-free procedure for the preparation of acyl chlorides including formyl chloride’, *Tetrahedron Letters*. Pergamon, 38(37), pp. 6489–6492.

Van de Vossenberg, J. L. C. M., Driessen, A. J. M. and Konings, W. N. (1998) ‘The essence of being extremophilic: The role of the unique archaeal membrane lipids’, *Extremophiles*, pp. 163–170.

Wakeham, S. G. *et al.* (2003) ‘Archaea mediate anaerobic oxidation of methane in deep euxinic waters of the Black Sea’, *Geochimica et Cosmochimica Acta*, 67(7), pp. 1359–1374.

Walker, C. B. *et al.* (2010) ‘Nitrosopumilus maritimus genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea’, *Proceedings of the National Academy of Sciences*, 107(19), pp. 8818–8823.

Weijers, J. W. H. *et al.* (2006) ‘Occurrence and distribution of tetraether membrane lipids in soils: Implications for the use of the TEX86 proxy and the BIT index’, *Organic Geochemistry*, 37(12), pp. 1680–1693.

Woese, C. R. (1987) ‘Bacterial evolution.’, *Microbiological Reviews*, 51(2), pp. 221–271.

Woese, C. R. and 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), pp. 5088–5090.

Woese, C. R., Kandler, O. and 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*, 87(12), pp. 4576–4579.

Woo, H. K. *et al.* (2009) ‘Phosphonium labeling for increasing metabolomic coverage of neutral lipids using electrospray ionization mass spectrometry’, *Rapid Communications in Mass Spectrometry*. NIH Public Access, 23(12), pp. 1849–1855.

Wörmer, L. *et al.* (2013) ‘Application of two new LC–ESI–MS methods for improved detection of intact polar lipids (IPLs) in environmental samples’, *Organic Geochemistry*, 59, pp. 10–21.

Wuchter, C. *et al.* (2004) ‘Temperature-dependent variation in the distribution of tetraether membrane lipids of marine Crenarchaeota : Implications for TEX 86 paleothermometry’, *Paleoceanography*, 19(4), pp. 1–10.

Wuchter, C. *et al.* (2006) ‘Archaeal nitrification in the ocean’, *Proceedings of the National Academy of Sciences*, 103(33), pp. 12317–12322.

Xu, S. *et al.* (2005) ‘The DMAP-catalyzed acetylation of alcohols - A mechanistic study (DMAP = 4-(dimethylamino)pyridine)’, *Chemistry - A European Journal*, 11(16), pp. 4751–4757.

Yang, H. *et al.* (2015) ‘The 6-methyl branched tetraethers significantly affect the performance of the methylation index (MBT′) in soils from an altitudinal transect at Mount Shennongjia’, *Organic Geochemistry*, 82, pp. 42–53.

You, J. *et al.* (2007) ‘Study of a new derivatizing reagent that improves the analysis of amino acids by HPLC with fluorescence detection: Application to hydrolyzed rape bee pollen’, *Analytical and Bioanalytical Chemistry*, 387(8), pp. 2705–2718.

You, J. *et al.* (2010) ‘An improved reagent for determination of aliphatic amines with fluorescence and online atmospheric chemical ionization-mass spectrometry identification’, *Analytica Chimica Acta*, 658(1), pp. 98–105.

Zech, R. *et al.* (2012) ‘Branched glycerol dialkyl glycerol tetraethers in Pleistocene loess-paleosol sequences: Three case studies’, *Organic Geochemistry*, 53, pp. 38–44.

Zhang, L.-M. *et al.* (2010) ‘Autotrophic ammonia oxidation by soil thaumarchaea’, *Proceedings of the National Academy of Sciences*, 107(40), pp. 17240–17245.

Zhang, M., Fang, C. and Smagin, G. (2014) ‘Derivatization for the simultaneous LC/MS quantification of multiple neurotransmitters in extracellular fluid from rat brain microdialysis’, *Journal of Pharmaceutical and Biomedical Analysis*, 100, pp. 357–364.

Zhu, C. *et al.* (2013) ‘Comprehensive glycerol ether lipid fingerprints through a novel reversed phase liquid chromatography-mass spectrometry protocol’, *Organic Geochemistry*. Elsevier Ltd, 65, pp. 53–62.

Zhu, Z.-T. *et al.* (2006) ‘A new fragmentation rearrangement of the N-terminal protected amino acids using ESI-MS/MS’, *Indian Journal of Biochemistry and Biophysics*, 43(6), pp. 372–376.