Improvement and evaluation of FMOC-Lys-Boc derivatisation for the detection of steroidal and tetraether lipid core biomarkers in soils and sediments

NATTA WIRIYAKUN

PhD

University of York

Chemistry

September 2021

Abstract

Steroids and glycerol dialkyl glycerol tetraether (GDGT) core lipids (biological markers containing at least one alcohol group) are widely used in the interpretation of contemporary environments and in archaeological and palaeoenvironmental studies. The conventional methods for their analysis are gas chromatography-mass spectrometry (GC-MS) of steroidal alcohols as the silylether derivatives and high performance liquid chromatography-mass spectrometry (HPLC-MS) for native tetraether lipids. Due to their low concentrations in soils and sediments, quantification of these biomarkers by the conventional analytical methods may be limited by the amount of materials available for extraction. An initial method for Fmoc-lysine-Boc (FLB) derivatisation of alcohols using multiple detectors (mass spectrometry; MS, ultraviolet-visible; UV-vis spectrophotometry, and fluorescence; FL) has been re-examined and the analytical capabilities improved and validated. Refinements to all steps of the method for preparation and analysis of FLB derivatives of C_{27} steroidal alcohols (cholesterol, coprostanol, epicoprostanol, and 5α-cholestanol) and GDGTs have significantly extended the applicability of the method. FLB derivatives give improved response (by approximately 200-1,000 times) over conventional methods, suggesting the method to be suitable in situations where analyte concentrations are at trace levels and/or where improved stratigraphic resolution is required. The refined methods for measurement of C_{27} steroidal alcohols and GDGTs performed well for all aspects of quantification (repeatability, reproducibility, recovery). Measurement of C_{27} steroidal lipids as FLB derivatives in extracts from archaeological soils and estuary sediments gave comparable values with those obtained by the conventional method. Detection of isoprenoid GDGTs (*i*-GDGTs) and branched GDGTs (*br*-GDGTs) as their FLB derivatives in sediment samples from a round-robin study (Schouten *et al.* 2013) gave comparable values for the tetraether index of tetraethers consisting of 86 carbons (TEX_{86}) and branched and isoprenoid tetraether (BIT) indices with those obtained in the study. The better repeatability of analysis of FLB-GDGTs by FL detector (<5.9 %RSD, n=3) than by MS detection (<12.5 %RSD) demonstrates the suitability of this approach for measurement of TEX_{86} and BIT indices.

List of Contents

List of Tables

List of Figures

Acknowledgements

I would like to express my deepest gratitude to my research supervisor, Prof. Brendan Keely, for giving me the opportunity of being polished to become a better researcher through invaluable advice, critical comments, and emotionally support throughout my study. I also would like to express my appreciation for your particular patience and guidance during the writing process and correction of this thesis. This research was conducted in the Department of Chemistry, University of York and received funding from a Royal Thai Government scholarship.

I would like to express my sincere thanks to Ed Bergström for technical support with LC-MS instrument, and Karl Heaton for acquiring the GC-MS data. Daniela Barilla and her team are thanked for providing Archaeal cell culture material. Martina Conti is thanked for being a mentor and training me to work in laboratory properly, while Cezary Poplawski and Alice Bacci are thanked for partial development the FLB derivatisation method and demonstration of proof of the concept from successful screening of alcohol biomarker compounds in archaeological samples.

Author's Declaration

I declare that this thesis is a presentation of original work, and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Natta Wiriyakun

Chapter 1

Introduction

A variety of biological markers containing one of more alcohol groups are recognised and are widely used in the interpretation of palaeoenvironments and in environmental analysis (Bull *et al.* 1999; Stott *et al.* 1999; Schouten *et al* 2013). These biomarkers have conventionally been analysed by gas chromatography-mass spectrometry (GC-MS) approaches in which their alcohols are typically derivatised as the silylethers (Jeng *et al.* 1997; Sutton and Rowland 2012; Lengger *et al.* 2018). Quantification of these components can be compromised in situations where the availability of soils and sediments is limited due to low amounts of analyte (Sistiaga *et al.* 2014; White *et al.* 2018; von der Lühe *et al.* 2020). The work described in this thesis focuses on improving and evaluating a novel derivative for the analysis of alcohols by high performance liquid chromatography (HPLC), with a view to improving the analytical capabilities for the detection of alcohols. The method has been applied to two key groups of biomarkers which have alcohol groups: steroids (see Section 1.1) and tetraether lipids (see Section 1.2), owing to their importance and utility in studies of contemporary environments and in archaeological and palaeoenvironmental studies (Bull *et al.* 1999; Schouten *et al* 2013; Prost *et al.* 2017).

1.1 Steroidal lipids – cholesterol and its transformation products

The human gut microbiome produces a unique profile of steroidal transformation products which have been exploited as biomarkers in fields including geochemistry, medicine, forensic science, archaeology, and environmental analysis (von der Lühe *et al.* 2013, 2018; Prost *et al.* 2017; Schött *et al.* 2018). Faecal profiles of sterols and metabolic derivatives, which are formed by microbial-mediated reactions in the gut, are determined by the nature of the foodstuffs recently consumed by an individual and are affected by human health conditions such as obesity, dysfunctional regulation of the immune system, homoeostasis of blood lipids, food intervention and antibiotic treatment (Clavel *et al.* 2017; Schött *et al.* 2018; Kenny *et al.* 2020). For instance, steroidal profiles in faeces have been highlighted as one of the indicators to diagnose the status of atherosclerosis which can be used to evaluate the risk of cardiovascular disease (CVD) (Nie *et al.* 2019; Vors *et al.* 2020). Faecal steroidal profiles have also been commonly used in environmental settings to monitor levels of faecal contamination in water sources such as river, lagoons and estuaries or in aquatic sediments, and to monitor the effects of sewage treatments implemented to address acute or chronic levels of contamination (Matić *et al.* 2014; Battistel *et al.* 2015; Matić *et al.* 2016; de Melo *et al.* 2019). Recently, faecal steroidal biomarkers in water have also been utilised for quality control in *Prochilodus lineatus* fisheries and an oyster farm impacted by effluents discharged from urban waters or estuaries (Florini *et al.* 2020; Speranza *et al.* 2020). Faecal steroidal profiles in rivers in the proximity of an area of intense agricultural activity were combined with nitrogen isotope ($\delta^{15}N$) measurements to distinguish between livestock faeces and chemical fertilisers as sources of nitrate pollution (Nakagawa *et al.* 2019). Faecal steroidal profiles in soils and sediments have also been used as signatures for past human and mammalian activities in areas lacking visible evidence (Bethell *et al.* 1994; Evershed *et al.* 1997; Sistiaga *et al.* 2014; Harrault *et al.* 2019). Profiles of faecal contamination in archaeological contexts have revealed the dynamics of human populations impacted by climate change or natural catastrophes, demonstrating the potential to elucidate how ancient humans adapted and responded to palaeoenvironmental change (White *et al.* 2018, 2019; McWethy *et al.* 2020; Keenan *et al.* 2021).

Faecal steroidal biomarkers were measured as supporting evidence in archaeological and forensic sites where human occupation or use was known, either from visible evidence or from records (White *et al*. 2018; von der Lühe *et al.* 2018, 2020; Schroeter *et al.* 2020). In several of those cases, the coprostanol index values measured did not reach the threshold value for human contamination, were close to the critical value (von der Lühe *et al.* 2018; Schroeter *et al.* 2020) or were much lower (White *et al.* 2018; von der Lühe *et al.* 2020). Failure to detect faecal steroidal signatures in these reports was accounted for either by loss of steroidal lipids in soils due to flooding or erosion, or to the limited availability of materials. Hence, recognition of faecal contamination in archaeological and environmental settings from faecal steroid profiles may sometimes be compromised by steroid concentrations being close to the limit of detection of the conventional method, suggesting the need for improved analytical capabilities.

1.1.1 Transformation pathways of cholesterol and cholestanols in human intestine

Cholesterol is necessary for the correct functioning of the human body: it is a crucial component of all cell membranes and a precursor for biosynthesis of other compounds such as other steroids, bile acids and vitamin D (Song *et al.* 2014; Maekawa *et al.* 2015; Faridi *et al.* 2017; Li *et al.* 2019). About 20-25 % of cholesterol in the intestine of humans comes from diet, the rest being produced by organs such as the liver, adrenal glands and the gonads (Norum *et al.* 1983; Bays *et al.* 2008). Cholesterol delivered to the intestine is metabolised by two main pathways: a serial transportation to serum by Niemann–Pick C1-like 1 protein (NPC1L1) located on the brush border membrane of intestinal epithelial plasma cells (Davis *et al.* 2004; Garcia-Calvo *et al.* 2005; Lammert *et al.* 2005), and conversion of cholesterol to coprostanol by resident gut microbiota (Kenny *et al.* 2020). High circulating cholesterol in the bloodstream is strongly related to the development and progression of CVD which causes approximately 30% of global deaths (Clavel *et al.* 2017; Kenny *et al*. 2020; Vors *et al.* 2020). Unabsorbed cholesterol in the intestine is converted to coprostanol and excreted in faeces (Björkhem *et al.* 1971; Eyssen *et al.* 1973; Eyssen *et al.* 1973; Sadzikowski *et al.* 1977; Veiga *et al.* 2005). Recognition that the conversion of

cholesterol to coprostanol in the human intestine balances serum cholesterol levels shed light on a possible clinical strategy to decrease a risk of CVD by altering the gut microbiota combined with controlling the levels of metabolic lipid absorption between the intestine and serum (Nie *et al.* 2019; Vors *et al.* 2020).

An understanding of the microbial metabolic pathways for cholesterol-to-coprostanol transformation also suggested the potential development of a framework for recognising signatures of faecal profiles in soils and sediments to using them as invisible markers for human faecal contamination in cases where coprostanol is the dominant C_{27} steroidal transformation product. The transformation of cholesterol to coprostanol occurs either by direct (*I*) or by indirect (*II*) pathways (Figure 1.1) (Björkhem *et al.* 1971; Wakehman 1989, Freier *et al.* 1994; Li *et al.* 1995). The type of transformation mechanism depends on the medium to which cholesterol is exposed. Cholesterol reductase directly reduces the double-bond of C-5 of cholesterol to form 5β(H)-cholestanol, otherwise termed coprostanol (Figure 1.1, *I*). By contrast, some microbiota in the gastrointestinal tracts of mammals oxidise the C-3 hydroxyl group of cholesterol to form 5-cholesten-3-one (Figure 1.1, *II-1*) which is subsequently isomerised to 4-cholesten-3-one (Figure 1.1, *II-2*). Subsequent reduction of the C-4 double bond forms coprostanone (Figure 1.1, *II-3*), which is further reduced to coprostanol (Figure 1.1, *II-4*). Most cholesterol in the mammal intestine is converted to coprostanol *via* the indirect pathway (*II*) by members of the gut microbiota such as *Eubacterium coprostanoligenes* (*E. coprostanoligenes*) (Eyssen *et al.* 1973; Sadzikowski *et al.* 1977; Brinkley *et al.* 1982; Freier *et al.* 1994). Coprostanol excreted in faeces constitutes more than 50% of the total faecal sterols in rats and humans, reflecting its potential as a faecal biomarker of mammalian input (Bhattacharyya 1986; Veiga *et al.* 2005).

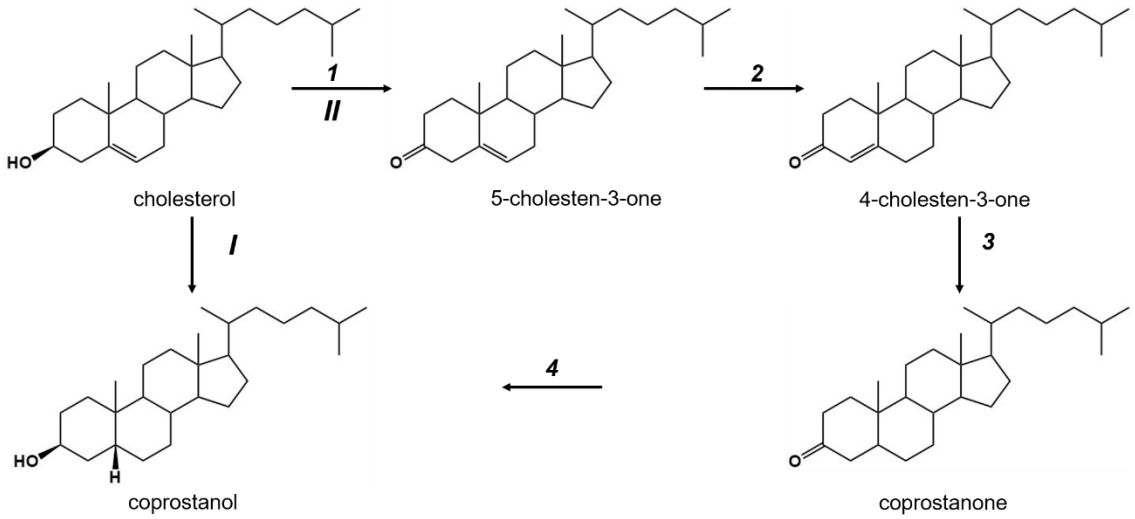


Figure 1.1. Proposed reaction pathways for conversion of cholesterol to coprostanol (Björkhem *et al.* 1971). *I* = direct pathway; *II* = indirect pathway; 1 = oxidation, 2 = isomerisation, 3 = reduction, 4 = reduction.

1.1.2 Signatures of cholesterol and cholestanols in soils and sediments

Deposition of faecal organic matter to soils and sediments can lead to the preservation of the steroidal lipid profiles. Cholesterol and coprostanol can both be preserved or undergo further transformation (Figure 1.2).

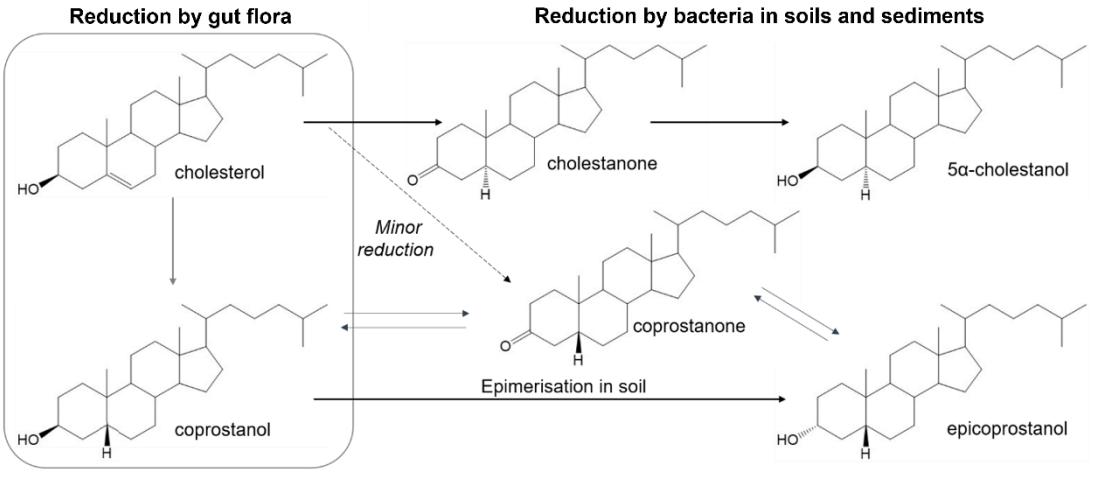


Figure 1.2. Schematic of the transformation pathways for reduction of cholesterol to 5α-stanols and 5β-stanols in the mammalian gut and the environment (soils and sediment) (Prost *et al*. 2017).

Differentiating steroidal transformation products formed by different groups of microorganisms has great potential for indicating human activity and occupation. Depending on the nature of the microbial flora to which it is exposed, the dominant animal sterol (the $C_{27:1}$ component cholesterol) is reduced to form one of two epimers: 5β(H) cholestanol, otherwise termed coprostanol, or 5α(H)-cholestanol (Björkhem *et al.* 1971; Eyssen *et al.* 1973; Veiga *et al.* 2005). Thus, microbes in the gastrointestinal tracts of mammals generate coprostanol, whereas the natural microbiota in sediments produce the 5α-epimer (Björkhem *et al.* 1971; Eyssen *et al.* 1973; Bull *et al.* 2002). The 5β-epimer further undergoes inversion of the 3β-hydroxyl group to form the 3α-epimer, termed epicoprostanol in soils and sediments (Bull *et al.* 2002; Prost *et al.* 2017). Understanding the transformation pathways of cholesterol has enabled cholesterol and its stanols to be used as faecal biomarkers in a broad range of applications.

The extremely slow degradation of sterols and stanols in soils and sediments has enabled their preservation as potential biomarkers in archaeological contexts of greater than ten thousand years (Sistiaga *et al.* 2014; McWethy *et al.* 2020; Schroeter *et al.* 2020). The good preservation of steroidal lipids in soils and sediments is aided by a low leaching rate of the compounds (due to their low water solubility), along with their slow degradation under anaerobic conditions (Walker *et al.* 1982; Birk *et al.* 2011; Lloyd *et al.* 2012). The uniqueness of the profiles of cholesterol transformation products in soils and sediments makes them promising indicators of faecal input. Thus, these markers can reflect human occupation, including in situations where visible evidence (such as human remains, grave goods, burial structures and past land use) are lacking (Evershed *et al.* 1997; Bull *et al.* 1999; Evershed 2008; Guillemot *et al.* 2017; Pickering *et al.* 2018; Schroeter *et al.* 2020).

Identifying human activity and occupation based on the signatures of cholesterol and cholestanols in soils and sediments presents challenges owing to their occurrence along with various other steroidal lipids (derived from plants, fungi, other animals and their faecal remains) in the environment (Prost *et al.* 2017; von der Lühe *et al.* 2018; Harrault *et al.* 2019). Depending on the sources of faecal inputs, the dominant plant sterols (the $C_{29:2}$ component stigmasterol and the $C_{29:1}$

component β-sitosterol) are reduced to form one of two epimers, 5α(H) stigmastanol or 5β(H)-stigmastanol, *via* similar pathways to cholesterol (Figure 1.2). Differentiating steroidal transformation products produced from different foodstuffs and animal species has allowed their unique profiles to be used for determining human or animal activity and occupation (Prost *et al.* 2017, 2018; Harrault *et al.* 2019). The dominance and persistence of the C_{27} 5β-stanols (coprostanol and epicoprostanol) in soils and sediments reflect their input in faeces of omnivores and carnivores, while high concentrations of the C_{29} 5β-stanols (5β-stigmastanol and 5β-epistigmastanol) reflect input from herbivore faeces (Derrien *et al.* 2011; Prost *et al.* 2017, 2018). The dominance and persistence of the C_{27} steroidal lipids (cholesterol and its transformation products) in the soils beneath buried porcine and human remains enabled recognition of inputs from fluid generated upon decomposition of the bodies as alternative biomarkers in forensic science (von der Lühe *et al.* 2013, 2018; Pickering *et al.* 2018). The persistence of cholesterol and coprostanol in sewage outflow to rivers has also been used to trace pollution plumes and estimate impacts of human occupation on the environment (Grimalt *et al.* 1990; Stefens *et al.* 2007; Bujagić *et al.* 2016; Tort *et al.* 2017; de Melo *et al.* 2019).

High concentrations of 5β-stanols in soils and sediments have been established as reflecting the presence of human faecal matter. For example, a ratio of coprostanol: 5β-stigmastanol over 1.5: 1 correlated well with soils of various ages obtained from locations expected to contain human faecal matter: a 17th century latrine; a medieval garderobe and a suspected Roman cesspit (Bethell *et al.* 1994). Exploring the transformation of cholesterol and coprostanol in soils and sediments led to the establishment of reliable criteria for assessment of human faecal contamination in environmental samples or archaeological contexts: a ratio of coprostanol to total cholesterol transformation products greater than 0.7 (Bull *et al.* 1999; Pickering *et al.* 2018). Coprostanol signatures from faecal matter have been used to identify human occupation and to explore meal reconstructions of different species (Homo sapiens and Neanderthals) or recent meals consumed by humans (Sistiaga *et al.* 2014; Prost *et al.* 2017) and primates (Sistiaga *et al.* 2015).

1.1.3 Quantification of cholesterol and cholestanols

Analysis of cholesterol and the full suite of transformation products that commonly occur in the complex distributions in soils and faecal materials requires an analytical technique providing both satisfactory separation efficiency and favourable detection limits. Gas chromatography-mass spectrometry (GC-MS) is the conventional method for determination of steroidal lipids as their trimethylsilyl (TMS) ethers (Bethell *et al.* 1994; Battistel *et al.* 2015; Prost *et al.* 2018), the TMS derivatives having greater ionisation efficiency and better chromatographic peak shapes than the native steroids. Gas-chromatography (GC) provides chromatographic retention times for each species and mass spectrometry enables their identification.

Quantification of cholesterol and cholestanols in archaeological samples is often limited by their extremely low concentrations. Due to these low concentrations, many studies of archaeological samples have used strategies for enhancing GC signal response, including increasing injection volumes (Sistiaga *et al.* 2014), or use of splitless injection (Lin *et al.* 1978; Bethell *et al.* 1994; Evershed *et al.* 1997; Hansen *et al.* 2011). Although these approaches increase response, they may reduce resolution of chromatographic separation due to co-elution of steroidal lipids with other components from sample overloading (Jayasinghe *et al.* 1998; Liu *et al.* 2013; Chen *et al.* 2015). The signal responses of sterols and stanols have also been deconvoluted for quantification using either selective ion monitoring (SIM) (Bethell *et al.* 1994; Evershed *et al.* 1997; Sistiaga *et al.* 2014; Prost *et al.* 2018) or multiple reaction monitoring (MRM) (Sistiaga *et al.* 2014) in order to obtain only targeted components and filter out signals generated from other components in the sample matrix.

Despite the recognition of human faecal contamination based on the coprostanol ratio commonly giving good agreement with the samples obtained from human faeces (Cuevas-Tena *et al.* 2017; Prost *et al.* 2017) and environmental settings (Pratt *et al.* 2008; Bujagić *et al.* 2016; de Melo *et al.* 2019), its implementation to signal human occupation has failed in archaeological contexts from some well-known areas of ancient cultural occupation. These include Paisley cave (Sistiaga *et al.* 2014), Cahokia (White *et al.* 2018), Lake Chatyr Kol along the ancient silk road (Schroeter

et al. 2020) and the Maya lowlands (Keenan *et al.* 2021), possibly implying the need for an alternative analytical approach to provide more reliable steroidal analysis.

1.1.3.1 Detection of native steroids

Whereas the gas chromatography-tandem mass spectrometry analysis of native steroidal compounds (i.e. without derivatisation) has been reported for general food samples at part-per-million (ppm) levels (Chen *et al.* 2015), direct liquid chromatography-mass spectrometry (HPLC-MS) analysis in extracts from aquatic sediments has been reported at ppm to part per billion (ppb) detection limits (Table 1.1) (Matić *et al.* 2014; Bataglion *et al.* 2015, 2016; Bujagić *et al.* 2016; Jauković *et al.* 2017). Notably, however, co-elution of various steroids and their transformation products requires the use of tandem mass spectrometry (MS/MS) to deconvolute the profiles, and the methods gave variations of up to 20% relative standard deviation (%RSD) (Matić *et al.* 2014; Bujagić *et al.* 2016; Jauković *et al.* 2017). These methods reflect limited chromatographic resolution and insecurity in quantification. The LC-MS signal response for native steroidal alcohols was also affected by dehydration of the components on ionisation, suggesting the use of a derivatising agent would improve stability and ionisation efficiency (Higashi and Shimada 2004; Kirk 2007).

Similarly, high performance liquid chromatography (LC) with ultraviolet-visible spectrophotometry (UV-Vis) detection has been reported for direct detection of native coprostanol and cholesterol in wastewater (Moliner-Martinez *et al.* 2010). Although the method allowed detection of the native cholesterol and coprostanol down to ppb levels without sample preparation, it required large sample volume injection (200 μL) using a modified sample injection loop made from GC capillary tubing attached directly to the LC-UV. The approach was limited by the low molar absorption coefficients of the UV absorption bands of steroidal lipids (Moliner-Martinez *et al*. 2010). The limitation in UV detection of steroidal compounds as the native forms is also evident from the three-times worse detection limit (LOD) values than those obtained using

a corona charged aerosol (CAD) detector (Fibigr *et al.* 2017). While direct analysis of steroidal lipids in human hepatocyte cells has been demonstrated using UHPLC-triple quadrupole MS combined with an increased injection volume (5 μL), the method depended on tandem MS to refine the steroid signals and required two different LC separation conditions to distinguish cholesterol from other sterols, reflecting inefficient chromatographic resolution (Skubic *et al.* 2020). The low concentration levels of native steroidal lipids typical in environmental samples led to those studies using strategies to enhance MS or UV response, including collecting larger samples (up to 20 g of sediments) and increasing injection volumes (10-20 μL). Such requirements may limit the potential for application to archaeological samples where the availability of materials is often extremely limited.

Table 1.1. Summary of literature review of the reported methods for native steroidal lipids detection.

Analyte	Real sample, g	Analysis method (injection volume)	Analytical figures					Note	Reference
			Linearity	LOD	LOQ	RSD /%	Recovery / %		
C_{27} – C_{29} sterols and their 5β-stanols, triterpenols	Sediments from Antarctica, Continental Shelf, São Sebastião Channel, and Santos Estuary, 20 g	UHPLC-MS/MS (10 µL)	20-1000 ng/ml	1.2-3 ng/mL, ppb level	3.9-9.9 ng/mL, ppb level	0.3-10.2	83-107	Used microwave-assisted for lipids extraction. Required deconvolution to pull out the peaks	Bataglion *et al.* (2015)
C_{27} – C_{29} sterols and their 5β-stanols, triterpenols	Sediments from estuarine- lagoonal system, 20 g	UHPLC-MS/MS (10 µL)	20-1000 ng/ml	1.2-3 ng/mL, ppb level	3.9-9.9 ng/mL, ppb level	0.3-10.2	83-107	Same method as (Bataglion *et al.* 2015)	Bataglion *et al.* (2016)
C_{27} and C_{29} sterols and their stanols, hormones	River sediments, 2 g	HPLC-MS/MS (10 µL)	N.D.	0.8–18 ng/g, ppb level	2.5-60 ng/g, ppb level	up to 20	73–118	>24 hours for lipid extraction Required deconvolution to pull out the peaks	Matić *et al.* (2014)
C_{27} sterols and their stanols, hormones	River sediments in Serbia, 2 g	HPLC-MS/MS (10 µL)	N.D.	0.8–18 ng/g, ppb level	2.5-60 ng/g, ppb level	up to 20	73–118	Same as (Matić et al. 2014)	Bujagić *et al.* (2016)

Analyte	Real sample, g	Analysis method (injection volume)	Analytical figures					Note	Reference
			Linearity	LOD	LOQ	RSD /%	Recovery / %		
C_{27} and C_{29} sterols and their stanols, hormones	Surface water and wastewater, 100 mL	HPLC-MS/MS (10 µL)	1–500 µg/L	2.8-10 ng/L, ppt level	9.3-50 ng/L, ppt level	up to 20	70–114	Quantified by 6-points standard addition (required sample at least for 6 portions)	Jauković *et al.* (2017)
Faecal sterols (cholesterol, coprostanol)	Wastewater, N.D.	HPLC-UV (up to 200 µL for real sample)	5–125 ug/L for cholesterol 40–5000 ug/L for coprostanol	1.3-10 ug/L, ppb level	4-30 ug/L, ppb level	up to 6	86–107	Used TRB-5 capillary GC column as an injection loop to increase sample volume.	Moliner-Martinez *et al.* (2010)
Phytosterols (e.g. ergosterol, brassicasterol, campesterol, fucosterol, β-sitosterol, stigmastanol)	Food supplements (capsules, <1.2 g and tablets, <20 g)	UHPLC-tandem UV/ corona charged aerosol (2 µL)	0.3 – 1.2 µg/mL	0.4 – 0.6 µg/mL, ppm level	N.D.	1.0 – 6.7	95.4 – 103.4	Used a core–shell phenyl-hexyl column providing a separation within 8.5 min. Available used for UV or corona-charged aerosol detectors.	Fibigr *et al.* (2017)

Analyte	Real sample, g	Analysis method (injection volume)	Analytical figures					Note	Reference
			Linearity	LOD	LOQ	RSD /%	Recovery / %		
Cholesterol and its late part of synthesis	Cultured human hepatocytes, 10^6 cells	UHPLC-MS/MS/MS (5 μL)	N.D.	23.4 – 56.3 ng/mL, ppb level	78.1 – 187.7 ng/mL, ppb level	3.5-9.4	N.D.	Used MRM mode to selectively detect the target components. Required a different mobile phase system to detect cholesterol separately	Skubic *et al.* (2020)
Cholesterol Brassicasterol Stigmasterol β-sitosterol Campesterol	Edible oils, 1 g Other foods (chicken eggs, milk powder, beverages, dietary supplement foods), 5 g	GC-MS/MS (1 μL)	2 – 100 mg/L	2 mg/kg, ppm level	N.D.	2-12	91-100	Determine cholesterol and four phytosterols as their native forms within 20 minutes	Chen *et al.* (2015)

1.1.3.2 Chemical modification

Approaches to enhance the response for steroids and their transformation products by using derivatisation agents have been reported previously (Table 1.2) (Jayasinghe *et al.* 1998; Liu *et al.* 2013; Resende *et al.* 2014; Ito *et al.* 2017, 2018; Sun *et al.* 2017; Schött *et al*. 2018; Nakakuni *et al.* 2019; Wang *et al.* 2019; Nzekoue *et al*. 2020; Yu *et al.* 2020). The ether and silylether steroid derivatives, produced by reaction with tetramethylammonium hydroxide (TMAH) (Nakakuni *et al.* 2019) and flophemesyl chloride (Jayasinghe *et al.* 1998), respectively, have enabled GC-MS detection of steroidal compounds at part per million (ppm) levels (0.1 μg mL^{-1}). Large volume injection (LVI) of up to 20 μL has been used and combined with multiple reaction monitoring (MRM) to refine signal responses of human faecal steroids as the TMS derivatives prepared using *N*-methyl-*N*-trimethylsilyl-trifluoracetamide (MSTFA) (Kunz and Matysik 2019).

Derivatising agents can increase the ionisation efficiency of steroids and improve their LC separation (Sun *et al.* 2017; Schött *et al.* 2018; Wang *et al.* 2019). A *N*,*N*-dimethyl glycine (DMG) derivative has been reported for quantification of steroidal lipids in clinical human faeces by isotope dilution liquid chromatography-mass spectrometry (LC-MS) (Schött *et al.* 2018). While it enables quantification of sterols and 5α- and 5β-stanols, problems with coelution required use of two separate HPLC columns. Cholesterol and phytosterols in food supplements have also been quantified as thiol derivatives formed using *N*-(4-(carbazole-9-yl)-phenyl)-*N*-maleimide labelled with ethylenedithiol (NCPM-SH), signal response being enhanced up to 560-fold (Yu *et al.* 2020). The method requires *in situ*-synthesis of NCPM and the separation condition allows the determination of C_{27} – C_{29} sterols, offering almost baseline separation within a 5 min-interval using MS detection in MRM mode. Phytosterols have also been reported to form ester derivatives of epigallocatechin (EGC) (Wang *et al.* 2019) or 4′-carboxy-substituted rosamine (CSR) (Zhao *et al*. 2016; Sun *et al.* 2017). Additional steps are required for preparation of both derivatising agents and the separation conditions were reported only for separation of sterol derivatives. A *p*-toluenesulfonyl isocyanate (PTSI) ester derivative of lupeol enabled quantification of the component in rat plasma that reflects the potential to use for pharmacokinetic study (Wang *et al.*

2021). Although the method detected lupeol within 2 min, precision was up to 15% RSD, reflecting the need for further improvement for faecal steroid profiling and quantification.

Ultraviolet-visible spectrophotometry (UV-Vis) and fluorescence spectrophotometry (FL) are favourable detection techniques for absolute quantification, and both have lower cost than GC-MS and LC-MS (Nelson 2011; Ito *et al.* 2017; Nzekoue *et al.* 2020). The use of UV-Vis or FL detectors for determination of sterols and stanols has been reported to give detection limits in the same range as MS (ppm to ppb levels) (Liu *et al.* 2013; Resende *et al.* 2014; Ito *et al.* 2017; Nzekoue *et al.* 2020). Strategies to increase UV detection by derivatisation include use of benzyl chloride as derivatising agent for quantification of phytosterols (Liu *et al.* 2013) and faecal sterols in sediments (Resende *et al.* 2014). While the methods demonstrated applicability for determination of the steroidal lipids, co-elution of certain phytosterols (Liu *et al.* 2013) and a shift in chromatographic retention time of the derivatives of steroids between the standard mixture and real samples (Resende *et al.* 2014) occurred. Both limitations can be attributed to the large injection volume (20 μL) reducing resolution due to increased longitudinal diffusion. Dansyl chloride was used to quantify phytosterols and their stanols in food supplements (Nzekoue *et al.* 2020) and spent coffee grounds (Nzekoue *et al.* 2020) using HPLC separation with UV and MS detectors. The method provided greater security of qualification by MS, though some phytosterols co-eluted due to large injection volume (10-20 μL), reflecting the need to refine conditions for extension of the method to separate the 5α- and 5β-phytostanols. The ester derivatives of cholesterol and phytosterols with either naproxen acyl chloride (Lin *et al.* 2007) or Bodipy FL (Nelson 2011) were reported for foods and biological samples. The baseline obtained with naproxen acyl chloride was not flat, and column flushing with ethanol was required for Bodipy FL after each chromatographic run to remove fluorescent residues limiting practical application. Anthroyl cyanide (ACN) derivatives enabled use of FL detection of phytosterols (Ito *et al.* 2017, 2018). Though reliable quantification of phytosterols in plant foods was evident, the LC chromatograms show a partial co-elution of stigmasterol and campesterol, possibly resulting from the sample injection volume of 10 μL.

The reported approaches for determination of steroidal lipids with derivatisation typically require the use of sample injection volumes of 5-10 μL to obtain signal responses of the components, implying a need to explore other approaches to improve responses in combination and achieve baseline separation of the analytes. The attempts to improve steroidal lipid analysis via derivatisation reflect the need for an analytical method having greater security in detection and quantification. To the best of the authors knowledge, there is no report of a method offering adequate chromatographic separation and reliable quantification of steroidal lipids that could potentially be used as an alternative to GC-MS.

Table 1.2. Summary of literature review of the reported derivatising agents for GC and HPLC detections of steroidal lipids.

Analyte	Derivatising agent	Real sample, g	Analysis method (injection volume)	Analytical figures					Note	Reference
				Linearity	LOD	LOQ	RSD / %	Recovery / %		
The C_{27} - C_{29} sterols	Tetramethylammonium Hydroxide (TMAH)	Sediment from Japan tideland, N.D.	GC-MS (N.D.)	N.D.	N.D.	N.D.	N.D.	N.D.	Available analysis for both ester and hydroxyl compounds to form their ether derivatives as the TMAH gave hydrolysis and methylation simultaneously.	Nakakuni *et al.* (2019)
The C_{27} and C_{29} sterols and their stanols	Pentafluorophenyldimethylsilyl (flophemesyl chloride)	Raw sewage river water, 20 mL to 2 L	GC-ECD GC-MS (1 µL)	0.5-50 µg/ml	0.1–0.6 ug/ml, ppm level	N.D.	up to 3.6	80–110	Gave the LOD less than one order magnitude of the TMS derivatives by GC-FID	Jayasinghe *et al.* (1998)
The C_{27} - C_{29} sterols	*N*-methyl-*N*-trimethylsilyl-trifluoracetamide (MSTFA)	Human faeces, 2 g	GC-triple quadrupole-MS (20 µL)	0.5-50 µg/ml	0.09–12.9 µg/mg, ppm level	0.5 - 142 µg/mg, ppm level	up to 15	85–115	Used MRM to refine the MSTFA derivatives selectively	Kunz and Matysik (2019)

Analyte	Derivatising agent	Real sample, g	Analysis method (injection volume)	Analytical figures					Note	Reference
				Linearity	LOD	LOQ	RSD / %	Recovery / %		
The C_{27} - C_{29} sterols	*N*,*N*-dimethylglycine (DMG)	Human faeces, 2 g	UHPLC-MS/HRMS (5 µL)	0.0019-137.68 nmol/mg dw	0.003−0.09 nmol/mg dw, ppm level	0.026−0.301 nmol/mg dw, ppm level	up to 16	88-111	Required two separation systems for sterols and stanols detection separately. Peaks overlapping required deconvoluted mode to pull out the interested peaks.	Schött *et al.* (2018)
The C_{27} - C_{29} sterols	*N*-(4-(carbazole-9-yl)-phenyl)-*N*-maleimide labelled derivative of ethylenedithiol (NCPM-d_0-SH)	Vegetable oil (soybean oil and peanut oil), 50 mg	UHPLC-tandem MS (N.D.)	2.5-1,000 µg/kg	0.15-0.40 mg/kg, ppb level	0.50-1.30 mg/kg, ppb level	up to 5	>90	In-house synthesised NCPM required. The detection sensitivities of thiol-derivatives containing drugs improved by 53 to 560-fold.	Yu *et al.* (2020)
Phytosterols cholesterol	Epigallocatechin (EGC)	N.D.	HPLC-MS (N.D.)	N.D.	N.D.	N.D.	N.D.	N.D.	Required an in-house preparation of EGC. Demonstration of Cholesterol-Redcing Activity and Inhibition of lipid oxidation in O/W emulsion of β-carotene.	Wang *et al.* (2019)

Analyte	Derivatising agent	Real sample, g	Analysis method (injection volume)	Analytical figures					Note	Reference
				Linearity	LOD	LOQ	RSD / %	Recovery / %		
Phytosterols	4′-carboxy-substituted rosamine (CSR)	Functional foods and medicinal herbs, 0.5 g for solids or 5 mL for liquids	UHPLC-MS/MS (N.D.)	0.10–100 ng/mL	5–1.5 pg/µL, ppb level	30–100 pg/µL, ppb level	2.8–8.3.	88.4–109.6	Required an in-house preparation of CSR.	Sun *et al.* (2017)
Lupeol	*p*-toluenesulfonyl isocyanate	Rat plasma, 100 µL	UHPLC-ESI-MS/MS (5 µL)	2.5-250 ng/mL	1 ng/mL, ppb level	10 ng/mL, ppb level	±15.	88.7–109.8%	Application to pharmacokinetic study of lupeol in rat plasma after oral administration	Wang *et al.* (2021)
Phytosterols	Benzoyl chloride	N.D.	HPLC-UV (20 µL)	0.2–5.0 M	89 ng	N.D.	N.D.	>95	Detected at λ 254 nm	Liu *et al.* (2013)
Faecal sterols	Benzoyl chloride	River sediments, 30 g	HPLC-UV (20 µL)	0.2-1.0 mg/ml	1.90 - 4.17 mg/L, ppm level	6.3-13.9 mg/L, ppm level	0.73-14	65-89	Used an ultra-sonication for lipids extraction	Resende *et al.* (2014)

Analyte	Derivatising agent	Real sample, g	Analysis method (injection volume)	Analytical figures					Note	Reference
				Linearity	LOD	LOQ	RSD / %	Recovery / %		
The C_{27} to C_{29} sterols and plant stanols	Dansyl chloride	Food supplement pills, 0.5 g	HPLC-DAD HPLC-MS (10 µL)	0.5-100 µg/ml	2.3 ng/mL, ppb level	7.7 ng/mL, ppb level	1.2–2.7	87-91	Increased the UV signal response of the derivatives for 23-fold for plant sterols and a 400-fold increment for plant stanols.	Nzekoue *et al.* (2020)
β-sitosterol campesterol stigmasterol cycolartenol	Dansyl chloride	Spent coffee grounds (SCG), 1 g	HPLC-DAD (20 µL)	0.5-100 µg/ml	9 - 15 ng/mL, ppb level	29 - 50 ng/mL, ppb level	0.3–6.2	N.D.	Phytosterol profiling to distinguish an origin of coffee beans	Nzekoue *et al.* (2020)
Cholesterol β-sitosterol Stigmastanol stigmasterol	Naproxen acyl chloride	Cow milk, soy milk, saliva, urine, 100 – 500 µL	HPLC-FL (10 µL)	0.1 – 2.0 µM	25 nM, ppm level	7.7 ng/mL, ppm level	1.8 – 5.3	99 - 104	Observed non-baseline background	Lin *et al.* (2007)
Cholesterol Byostatin	Bodipy FL	N.D.	HPLC-FL (20 µL)	10 fmol – 1 pmol	10 fmol	N.D.	up to 15	N.D.	Needed flushing column with ethanol for residue removal	Nelson (2011)

Analyte	Derivatising agent	Real sample, g	Analysis method (injection volume)	Analytical figures					Note	Reference
				Linearity	LOD	LOQ	RSD / %	Recovery / %		
Cholesterol and phytosterols	1-anthroyl cyanide (ACN)	Land plant and marine algae (1 g)	HPLC-FL (10 µL)	1.95-500 µg/ml	0.25-0.40 µg/mL, ppm level	0.83-1.35 µg/mL, ppm level	<2	85- 94	Partial co-elution of stigmasterol and campesterol.	Ito *et al.* (2017)
Cholesterol and phytosterols	1-anthroyl cyanide (ACN)	Algae from Japanese coast (1 g)	HPLC-FL (10 µL)	1.95-500 µg/ml	0.25-0.40 µg/mL, ppm level	0.83-1.35 µg/mL, ppm level	<2	85- 94	Same method as M. Ito *et. al.* (2017)	Ito *et al.* (2018)

1.2 Archaeal lipids related to palaeoenvironmental analysis

The Archaea produce unique membrane lipids containing glycerol dialkyl glycerol tetraether (GDGT) lipid cores. The GDGTs have been exploited in various proxies in the field of paleoenvironmental analysis: the well-known tetraether index of lipids with 86 carbons (TEX_{86}) for reconstruction of past sea surface water temperature (Huguet *et al.* 2007; Schouten *et al.* 2007; Escala *et al.* 2009; Powers *et al.* 2010; Schouten *et al.* 2013) being the most widely used. The unique profiles of Archaeal membrane lipids are affected by the *in*-situ conditions (e.g., temperature and pH) at the time cell growth occurs (Wuchter *et al.* 2004; Jain *et al.* 2014). For example, the GDGT lipids of Archaea found in marine water columns (Schouten *et al.* 2002; Wuchter *et al.* 2004) give different structural profiles to those obtained from volcanic and hydro-thermal vents (Brock *et al.* 1972; Jones *et al.* 1983). The signatures of GDGTs found in non-extreme environments (e.g., ocean, lakes, sediments, and soils) were also recognised to represent prolific Archaeal clades and to show potential as indicators in the global cycling of carbon and nitrogen (Jarrell *et al.* 2011; Schouten *et al.* 2013). Thus, GDGT biomarkers are widely used as proxies for environmental conditions based on differences in relative abundances of structurally distinct components within the distributions extracted from sediments, see for example (Stetter 2006; Berg *et al.* 2010).

1.2.1 Diversity of Archaeal GDGT membrane lipids and their transformation pathways

The GDGT membrane lipids of Archaea uniquely function as a monolayer and exhibit a structural diversity that has been exploited in the fields of microbiology and geochemistry (Jarrell *et al.* 2011; Schouten *et al.* 2013; Jain *et al.* 2014). Thus, Archaeal membrane lipids are membrane-spanning and so differ from the bilayers of bacterial and eukaryote membranes. The Archaeal membrane lipids comprise isoprene-based alkyl chains that are ether-linked to the glycerol moiety at the position of *sn*-glycerol-1-phosphate (G1P) (Figure 1.3b), while bacterial and eukaryotic membrane lipids comprise fatty acid chains ester linked to a single glycerol group at the *sn*-

glycerol-3-phosphate (G3P) position (Figure 1.3a). The GDGTs, which occur as ether-linked phospholipids in the Archaeal membrane, have been recognised to distinguish Archaea from other domains of life, the Eukarya, and the Bacteria (Woese and Fox 1977; Zillig 1991; Embley *et al.* 1992; Burggraf *et al.* 1994; Hugenholtz *et al.* 1998; Woese 2004). The presence of the ether bond in the Archaeal membrane provides greater resistance to degradation in extreme temperatures and pH than the ester linkage in membranes of Bacteria and Eukarya, reflected in the unique ability of Archaea to survive in extreme conditions (Van de Vossenberg *et al.* 1998). Archaeal membrane lipids show a distinct difference in structural diversity from bacterial and eukaryotic membranes, the difference in their membranes reflects phylogenetic differences apparent from the tree of life (Figure 1.4).

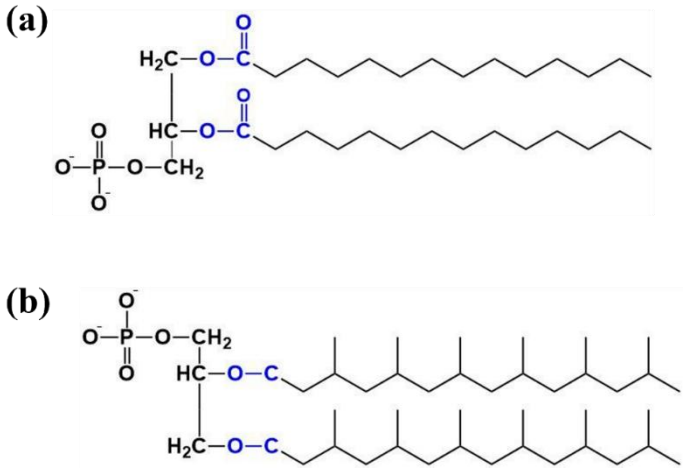


Figure 1.3. Chemical structures of membrane phospholipids (a) Bacteria, (b) in certain Archaea (for example halophiles). Blue label: The dominant difference in lipids is an ester-linked glycerol of bacteria, while the ether-linked glycerol appeared for Archaea.

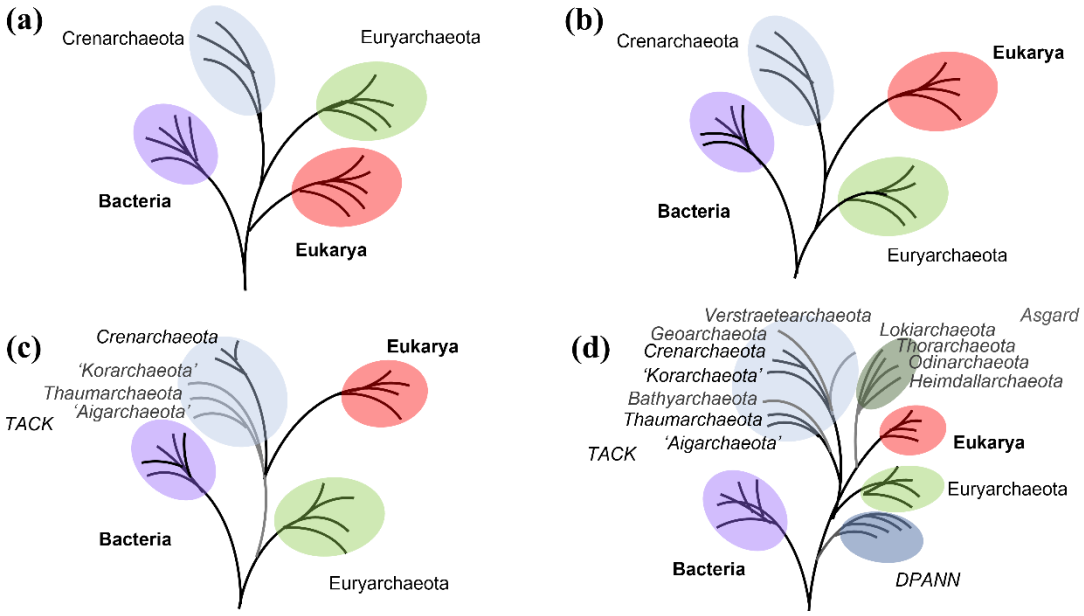


Figure 1.4. A schematic representation of an evolution of the tree of life showing the relationships between Eukaryotes and Archaea over the past 40 years. (a) Archaea and Eukarya each represent a monophyletic group and share a unique common ancestor to the exclusion of Bacteria. (b) *Crenarchaeota* show the closer relationship with eukaryotes than *Euryarchaeota*, (c) Novel *Thaumarchaeota*, *Aigarchaeota*, *Crenarchaeota* and *Korarchaeota* (*TACK*) superphylum reported, (d) Addinal two novel superphyla, *Asgard* and *DPANN* are represented (Eme *et al.* 2017).

The revelation of Archaeal diversity based on 16S rRNA gene sequence-based approaches has advanced understanding of the origins of Archaea, their evolution, metabolic capabilities, and ecological impact. This has led to several changes to our understanding of the tree of life over the past 40 years (Woese 2004; Villanueva *et al.* 2014). At first it was thought that Archaea and Eukarya shared a monophyletic group originating from the same ancestor and separated from Bacteria (Figure 1.4a). This resulted in the classification of Archaea into just two phyla, the *Euryarchaeota* and the *Crenarchaeota* (Figure 1.4a) (Buckles *et al.* 2013; Villanueva *et al.* 2014; Eme *et al.* 2017). Uncovering Archaeal diversity through advanced DNA sequencing and genomic technologies revealed that phylum *Crenarchaeota* has a closer relationship to eukaryotes than *Euryarchaeota* (Figure 1.4b). The *Crenarchaeota* was further shown to have evolved into

various branches, resulting in a reorganisation of the phylogeny into *Thaumarchaeota*, *Aigarchaeota*, *Crenarchaeota* and *Korarchaeota* (*TACK*) superphyla (Figure 1.4c). Recently, the *Asgard* superphylum has been included in the tree of life, suggesting that eukaryotes originated from within the Asgard archaea or that they represented a sister group to them (Figure 1.4d). The *Diapherotrites*, *Parvarchaeota*, *Aenigmarchaeota*, *Nanohaloarchaeota*, and *Nanoarchaeota* (*DPANN*) superphylum has also been introduced into tree of life, while the TACK superphylum has revealed several branches. The updating of the tree of life has reflected greater diversity among the Archaea with new phyla being defined (Figure 1.4), showing that there is much more to be understood regarding the process of eukaryogenesis. The gap in the evolutionary relationship between the three domains of life and the origin of eukaryotes, implies that further developments in genetic methodology are required in order to offer better confidence in resolution and to allow the genomes of uncultivated Archaea to be explored.

Studies of cell cultures and environmental gene sequencing reveal that the different environmental conditions under which Archaea grow provide the unique profiles of Archaeal lipids (Table 1.3). The diversity of Archaeal membrane lipids (Table 1.3) has enabled further classification of ether lipids into two major structures: archaeol (in which one glycerol moiety is linked by two phytanyl (C_{20}) chains at the *sn*-2,3 positions of glycerol) and glycerol dialkyl glycerol tetraethers (GDGT) in which two glycerol groups are etherified with two biphytanes (C_{40}) chains (Figure 1.5). Variation in GDGT structures includes differences in the number of cyclopentane rings, with up to 8 moieties in the structure described as 'GDGT-x', where x is the number of cyclopentane rings within the structure.

The dominant isoprenoid ether lipid core is the C_{86} component comprising a macrocyclic diglycerol lipid with etherified C_{40} biphytanes (BPs) (Figure 1.6). The C_{86} isoprenoid chains of GDGTs (*i*-GDGTs) also have cycloalkyl rings, double bonds, methyl groups and/or hydroxyl groups in their structures. For instance, up to four cyclopentyl rings can be present in each of the BP chains (BP_0 to BP_4; Figure 1.6), leading to the GDGT structures with up to eight rings in total

(GDGT-1_a to -7_a; Figure 1.6). Furthermore, other forms of glycerol tetraethers are also produced: glycerol monoalkyl glycerol tetraethers (GMGTs), in which the biphytanes (BP) are conjoined via a covalent cross link (e.g. GMGT-0_a in Figure 1.6), and glycerol trialkyl glycerol tetraethers (GTGTs), where the macrocycle is incomplete (e.g. GTGT-0_a in Figure 1.6).

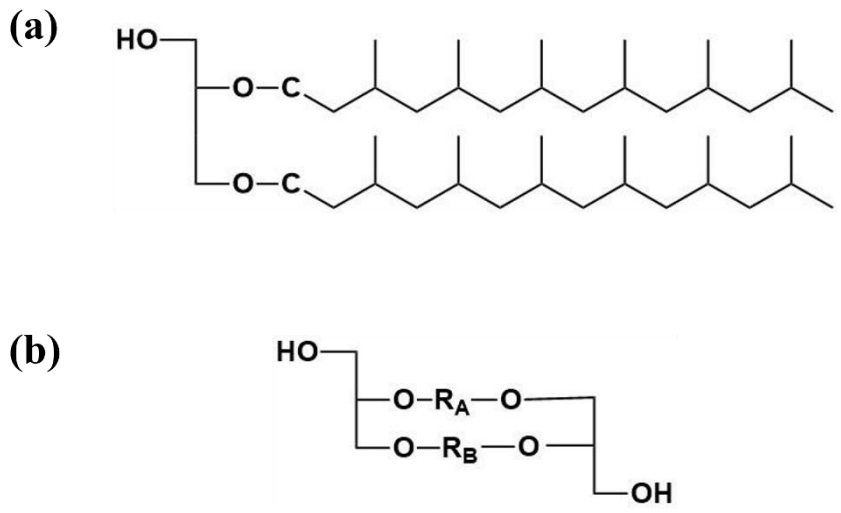


Figure 1.5. Chemical structures of (a) Archaeol, (b) Glycerol dibiphytanyl glycerol tetraether (GDGT). R_A, R_B = the C_{40} biphytanes (BPs) arranged in different configurations.

Table 1.3. Distribution of Archaeal membrane lipids in different orders of the *Euryarchaeota*, *Crenarchaeota* and *Thaumarchaeota* (Villanueva *et al.* 2014).

Archaea	Temperature*	pH‡	Metabolism	Archaeol	Extended archaeol	GDGT-0	GDGT-1-4	GDGT-5-8	Crenarchaeol
Euryarchaeota									
Halobacteriales	Mesophile	Neutrophile or alkalophile	Heterotrophy	✓	✓				
Methanosarcinales	Mesophile	Neutrophile	Methanogenesis	✓					
Methanopyrales	Hyperthermophile	Neutrophile	Methanogenesis	✓					
Methanococcales	Mesophile or Thermophile	Neutrophile or alkalophile	Methanogenesis	✓					
Thermococcales	Thermophile or Hyperthermophile	Neutrophile	Sulphur dependent	✓		✓			
Methanobacteriales	Mesophile or Thermophile	Neutrophile	Methanogenesis	✓		✓			
Archaeoglobales	Mesophile or Thermophile	Alkalophile	Sulphur dependent			✓			
Methanomicrobiales	Mesophile	Neutrophile	Methanogenesis			✓			
Thermoplasmatales§	Mesophile or Thermophile	Acidophile	Sulphur dependent			✓	✓	✓	

Archaea	Temperature*	pH‡	Metabolism	Archaeol	Extended archaeol	GDGT-0	GDGT-1-4	GDGT-5-8	Crenarchaeol
Crenarchaeota									
Thermoproteales	Thermophile or Hyperthermophile	Neutrophile or Acidophile	Sulphur dependent			✓	✓	✓	
Sulfolobales	Thermophile or Hyperthermophile	Acidophile	Sulphur dependent			✓	✓	✓	
Acidilobales	Hyperthermophile	Acidophile	Organotroph			✓	✓	✓	
Desulfurococcales	Hyperthermophile	Neutrophile	Sulphur dependent			✓	✓		
Thaumarchaeota									
Cenarchaeale	Mesophile	Neutrophile	Autotroph; ammonia oxidizer			✓	✓		✓
Nitrosopumilales	Mesophile	Neutrophile	Autotroph‖; ammonia oxidizer			✓	✓		✓
Nitrososphaerales	Mesophile or Thermophile	Neutrophile	Autotroph; ammonia oxidizer			✓	✓		✓

*Hyperthermophiles have an optimal growth temperature of at least 80 °C; mesophiles grow at a moderate temperature, typically between 20 °C and 45 °C.
‡Neutrophiles grow at almost neutral pH (pH 5−8); alkalophiles grow at alkaline pH (pH >8); acidophiles grow at acidic pH (pH <5).
§Deep hydrothermal vent Euryarchaeota 2 (DHVE-2) cluster (*Candidatus* Aciduliprofundum boonei), closely related to the order Thermoplasmatales, synthesize GDGT-0 and GDGT-1 to GDGT-4.
‖Autotroph, but some reports have suggested possible mixotrophy.

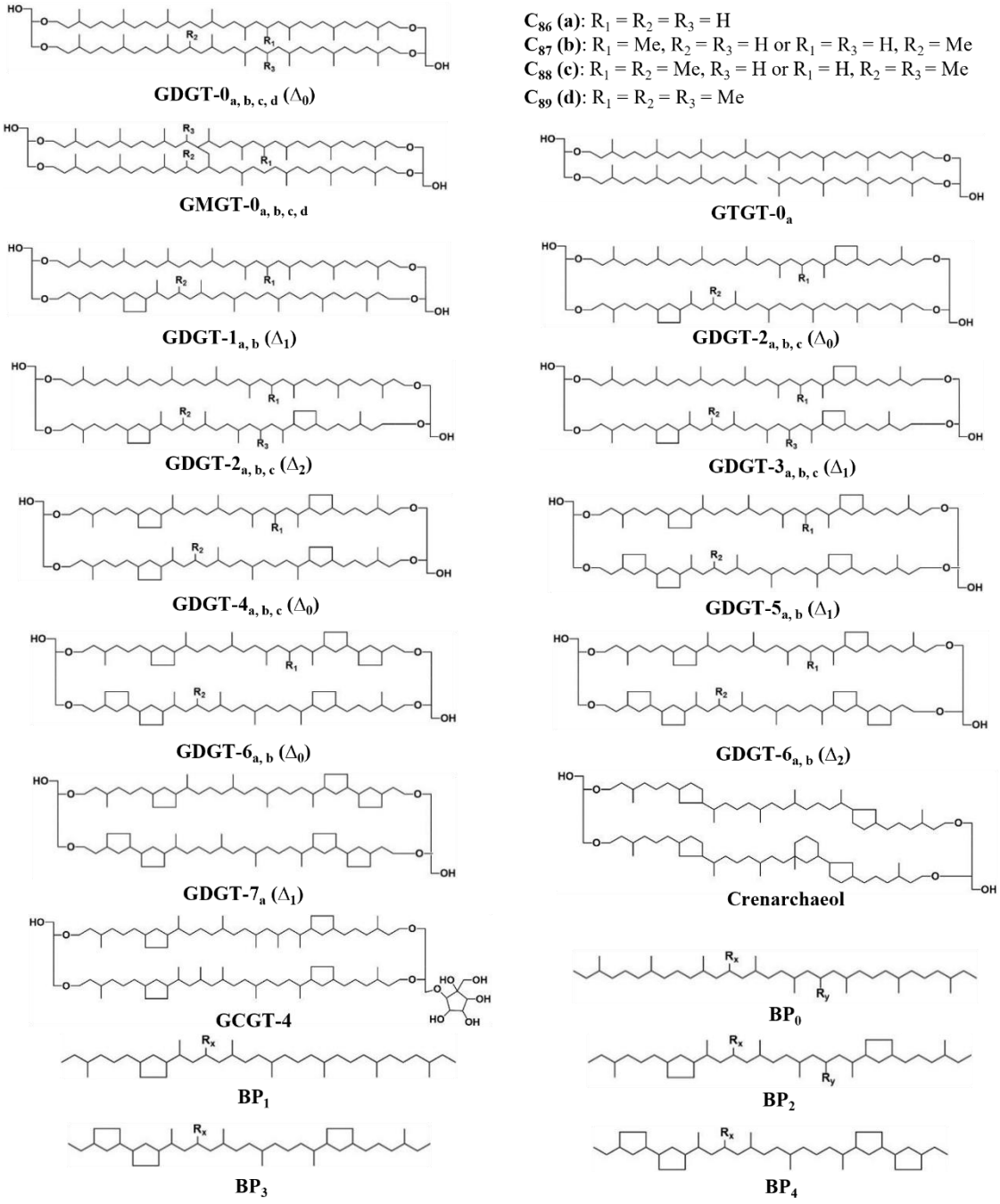


Figure 1.6. Archaeal tetraether lipid cores and hydrocarbons discussed in the text reviewed by Knappy *et al.* 2012. Δ_n = Ring distribution, where n is the numerical difference in the number of rings between etherified chains. C_{40} (BP): $R_x = R_y = H$.

The unique structure of Archaeal membrane lipids provides chemical stability that enables Archaea to survive in extreme environmental conditions such as high temperatures, high salinity or extreme pH (Thompson *et al.* 1992; Kates 1993). The isoprene-based alkyl chains of Archaea (the building blocks of the apolar region) exhibit diversity in the number of cyclopentane rings in the GDGTs related to growth conditions (Jensen *et al.* 2015; Quehenberger *et al.* 2020). The structural variations in the alkyl chains (see structures in 1.6) has been suggested to control membrane permeability (Van Der Meer *et al.* 2007; Pearson *et al.* 2008; Powers *et al.* 2010; Schouten *et al.* 2013). GDGT lipids of mesophilic or neutrophilic Archaea are also ubiquitous in non-extreme environments such as the oceans (Kuypers *et al.* 2001; Pitcher *et al.* 2011; Pearson and Ingalls 2013; Zeng *et al.* 2019). Different classes of the *Euryarchaeota, Crenarchaeota* and *Thaumarchaeota* have unique GDGT profiles that may relate, directly or indirectly, to their roles in different global biogeochemical cycles; for example, *Euryarchaeota* participate in methane cycling while *Thaumarchaeota* participate in nitrogen cycling (Schouten *et al.* 2002; Liu *et al.* 2009; Pitcher *et al.* 2011; Schouten *et al.* 2013; Villanueva *et al.* 2014). Thus, the *i*-GDGT profiles of Archaea in the lake waters show that *Thaumarchaeota* function above the oxycline/nitrocline in Lake Challa in East Africa (Buckles *et al.* 2013).

The living Archaea have membranes formed from intact polar lipids (IPL) comprising various isoprenoid ether lipid cores functionalised with sugar and/or phosphate head groups (Sturt *et al.* 2004; Lipp and Hinrichs 2009). After Archaea die, cellular remains deposited in soils or the underlying sediments of aquatic environments are transformed. The majority of the polar head groups of Archaea are lost by hydrolysis within days to several tens of millennia depending on depositional conditions, leaving the GDGT lipid cores (*i*-GDGTs) (Xie *et al.* 2013; Liu *et al.* 2016). The lipid cores can be preserved over geological timescales, or they can be further transformed to by-products including carboxyl derivatives, alcohols, and hydrocarbons (Michael *et al.* 1979; Schouten *et al.* 1998, 2003). Thus, *i*-GDGTs signatures from different depths in marine sediments have been widely used to reconstruct the profiles of past surface seawater temperatures (SST) (Schouten *et al.* 2002; Kim *et al.* 2008; Kim *et al.* 2015).

1.2.2 Significance of Archaeal lipids in soils and sediments related to past surface sea water reconstruction

Crenarchaeota and *Thaumarchaeota* represent approximately 20% of the picoplankton in the ocean, occurring in the photic and aphotic zones of the water column (Karner *et al.* 2001; Hoshino and Inagaki 2019). The GDGT membrane lipids of these Archaeal phyla were detected in relatively high amounts in recent and ancient marine sediments, and their potential to act as indicators of sea water temperature was demonstrated (Schouten *et al.* 2002; Wuchter *et al.* 2004; Villanueva *et al.* 2014). Studies of Archaeal lipids isolated from marine waters having different surface water temperatures revealed that the GDGT profiles contained greater proportions of the structures containing more cyclopentane moieties at higher temperatures. For instance, *Crenarchaeota* living at temperatures >60°C in geothermally heated soils or water and in sulphur-rich hot springs or hydrothermal vents contain a wide range of GDGT membrane lipids with from 0 up to 8 cyclopentane rings (Schouten *et al.* 2007; Pearson *et al.* 2008; Villanueva *et al.* 2014). The *Crenarchaeota* from temperatures higher than 80ºC revealed lower proportions of the higher numbers of cyclopentane rings than those obtained from lower temperature, although with some exceptions observed at particular locations (Van Der Meer, *et al.* 2007; Pearson *et al.* 2008; Schouten *et al.* 2013). Variation in the proportions of different GDGT lipid structures in *Sulfolobales* was influenced by a change in environmental temperature or pH. Nutrient limitation led to a shift in their specific growth rates, reflected through a change in the ratio of diether to tetraether lipids while maintaining the average cyclisation number of tetraether lipids (Jensen *et al.* 2015; Quehenberger *et al.* 2020). By contrast, *Thaumarchaeota* mostly grow at moderate temperatures (20-45°C) and produce GDGT lipids having a maximum of four cyclopentane rings and including the specific GDGT lipid crenarchaeol (Figure 1.6). Crenarchaeol possesses an extra cyclohexane ring in its structure (in addition to four cyclopentane rings) and has been used as the biomarker for *Thaumarchaeota*. *Thaumarchaeota* have been found in marine sediments, and their GDGT lipids have been measured to reflect temperature in surface waters (Sinninghe Damsté *et al.* 2002; Villanueva *et al.* 2014).

1.2.2.1 TEX_{86} index

Isoprenoid GDGT lipids produced by marine Archaea have been utilised as biomarkers for SST assessment based on the relative abundances of alkyl hydrocarbon skeletons with different numbers of cyclopentane rings, commonly 0 - 3 rings (GDGT-0–GDGT-3; see structures in Figure 1.6). In addition to producing crenarchaeol (Figure 1.6), *Thaumarchaeota* also produce its stereoisomer designated crenarchaeol'. Temperature dependent variations in the distributions of *i*-GDGTs led to the establishment of a proxy for sea surface water temperature named the 'tetraether index of tetraethers consisting of 86 carbons (TEX_{86})' index: the ratio of a ratio of GDGT lipid cores consisting of cyclopentane rings from 2-4 rings in the structures with crenarchaeol to total GDGT lipid cores consisting of cyclopentane rings (see the equation 1.1) (Schouten *et al.* 2002; Wuchter *et al.* 2004).

$$TEX_{86} = \frac{([GDGT-2]+[GDGT-3]+[crenarchaeol\prime]}{[GDGT-1]+[GDGT-2]+[GDGT-3]+[crenarchaeol\prime]} \quad (1.1)$$

where crenarchaeol' = crenarchaeol regioisomer, which was later recognised to be a *cis* stereoisomer of the cyclohexane moiety (Sinninghe Damsté *et al.* 2018).

Screening GDGT profiles in surface sea sediments (0-2 cm depth) from different worldwide locations allowed their TEX_{86} values to be correlated to the recorded annual mean sea surface temperature (SST) with the linear regression (r^2) reaching 0.92. This relationship reflects the potential to assess SST values from unknown sediments using equation 1.2 (Kim *et al.* 2008).

$$T = -10.78 + 56.2\,TEX_{86} \quad (r^2 = 0.935), n = 223 \quad (1.2)$$

where T = annual mean sea surface temperature (SST, °C)

The TEX_{86} index has been proposed as another tool for the reconstruction of past temperature in ancient sediments alongside other common proxies based on plankton foraminifera or haptophyte

algae, $\delta^{18}O$ and U^{k}_{37} respectively, as discussed below (Erez and Luz 1983; Brassell *et al.* 1986). Measurement of the calcite skeletons of planktonic foraminifer *Globigerinoides sacculifer*, deposited under isotopic equilibrium in waters having a wide range of temperatures (14 to 30°C), led the establishment of proxies for reconstruction of sea water temperature based on the oxygen isotope ratio ($\delta^{18}O$) and the ratio of magnesium over calcium (Mg/Ca) (Erez and Luz 1983). Subsequently, an evaluation from recent and ancient sediments of the variation in the degrees of unsaturation in long-chain unsaturated ketones produced by haptophyte algae gave a good correlation with the variations in the $\delta^{18}O$ signal for the calcareous skeletons of certain planktonic foraminifera, reflecting the potential to establish another proxy for assessment of SST, termed the $U^{K'}_{37}$ ratio (Brassell *et al.* 1986). Necessarily, the SST reconstructed from planktonic foraminifera cannot be determined from sediments deposited below the carbonate compensation depth owing to the dissolution of carbonate skeletons (Spero *et al.* 1997; Barker *et al.* 2005; Sadekov *et al.* 2008).

The analysis of alkenones to assess the $U^{K'}_{37}$ ratio has been reported in ancient sediments up to *ca*. 55 million years old, possibly reflecting a limited evolution of the alkenone unsaturation for SST reconstruction by this index (Schneider *et al.* 2001; Brassell *et al.* 2014; Robinson *et al.* 2017). The period over which reconstruction of past sea water temperature is possible using the TEX_{86} index extends to at least 190 million years (Kuypers *et al.* 2001; O'Brien *et al.* 2017; Robinson *et al.* 2017), reflecting the great advantage of this proxy compared with other proxies for paleoenvironmental temperature reconstruction.

A large variation, up to ± 4ºC in some marine sediments, was evident in past SST estimated using equation 1.2, reflecting a gap in the relationship of *i*-GDGTs to crenarchaeol stereoisomer that required further examination (Schouten *et al.* 2007; Kim *et al.* 2008; Steinig *et al.* 2020). Several attempts to refine TEX_{86} index and SST calibrations have been reported, revealing that the uncertainty derives from the amount of variation in the crenarchaeol regioisomer (Kim *et al.* 2008, 2010; Liu *et al.* 2009; Tierney and Tingley 2014). Distributional variations in the crenarchaeol regioisomer have been attributed to variations in *Crenarchaeota* production, with factors such as

water column depth, seasonality, and different oceanic settings thought to be influential. An attempt to refine the TEX_{86} index and narrow the SST bias has been reported by screening of GDGT profiles from extended surface sea sediments (426 samples), resulted in modified TEX_{86} indices that operate over two different temperature ranges: below 15°C; TEX_{86}^{L} (see the equation 1.3) and above 15°C; TEX_{86}^{H} (see the equation 1.4)(Kim *et al.* 2010).

$$TEX_{86}^{L} = log\left(\frac{([GDGT-2]}{[GDGT-1]+[GDGT-2]+[GDGT-3]]}\right) \quad (1.3)$$

$$TEX_{86}^{H} = log\left(\frac{([GDGT-2]+[GDGT-3]+[crenarchaeol\prime]}{[GDGT-1]+[GDGT-2]+[GDGT-3]+[crenarchaeol\prime]}\right) \quad (1.4)$$

$$SST = 67.5 \times TEX_{86}^{L} + 46.9 \quad (r^2 = 0.86, n = 396, p < 0.0001) \quad (1.5)$$

$$SST = 68.4 \times TEX_{86}^{H} + 38.6 \quad (r^2 = 0.87, n = 255, p < 0.0001) \quad (1.6)$$

Though the SST estimates based on TEX_{86}^{L} and TEX_{86}^{H} gave good correlations to the global SST dataset (see the calculated equations 1.5 and 1.6), standard errors of ±4°C were still obtained in some samples while TEX_{86}^{H} values from the Red Sea had to be omitted (Kim *et al.* 2008). The unusually high bias of SST values in surface sediments obtained from the Red Sea was later shown to result from a difference in the amounts of *Crenarchaeota* between the northern and southern regions, the latter being strongly influenced by inflowing water from the open ocean (Trommer *et al.* 2009). Reported calibrations for TEX_{86} that have coefficient of determination (r^2) values less than 0.9 have been questioned as models for the relationship between TEX_{86} and SST, possibly leading to underestimation of temperatures at high TEX_{86}^{H} values (Tierney and Tingley 2014). A data processing approach, utilising Bayesian regression to adjust TEX_{86} values, gave estimated SST values that matched the recorded data, suggesting that the previously determined TEX_{86}^{H} values were underestimated due to impacts of other environmental factors (salinity, nitrate, seasonality, depth). Furthermore, a study using the Deep-time Model Intercomparison Project (DeepMIP) (Hollis *et al.* 2019), a machine learning model developed to

evaluate the accuracy and a precision of all temperature proxies, recommended not using the TEX^{H}_{86} calibration for SST determination due to the known bias from exponential regression dilution. The numerous revaluations of TEX_{86} values could imply that the correlation issue between TEX_{86} and SST could be influenced by the limits to which the analytical method can estimate the relative proportions of individual *i*-GDGTs. In addition, the natural structural complexity of *i*-GDGT distributions in sediments also imposes a limitation. This suggests that further improvements in the analytical methodology are required.

1.2.2.2 BIT index

Another class of GDGT lipid commonly found in aquatic sediments comprises glycerol dialkyl ether linked with branched alkyl hydrocarbon skeletons, termed 'branched GDGTs' (*br*-GDGTs) (see the structures in Figure 1.7). Branched GDGTs are produced by bacteria that inhabit the soils in terrestrial environments such as lakes, continental lacustrine, and peats (Weijers *et al.* 2006). Profiles of *br*-GDGT profiles can be recognised in freshwater and marine aquatic environments as a result of the transport of terrestrial sediment, e.g in land run-off and rivers (Hopmans *et al.* 2004; Weijers *et al.* 2006, Schouten *et al.* 2009). An abundance of *br*-GDGTs in aquatic settings indicates significant fluvial- or runoff-derived input of soil organic matter. Like isoprenoid GDGTs, branched GDGTs can be preserved in sediments over archaeological timescales, reflecting the potential to reconstruct past activities within the hydrological cycle from terrestrial inputs in geological age sediments (Hopmans *et al.* 2004; Yang *et al.* 2014).

Variations in terrestrial organic matter input lead to changes in the amounts of *br*-GDGTs and crenarchaeol introduced into marine sediments. A proxy, established to evaluate the contribution of these components, termed the branched and isoprenoid tetraether (BIT) index, establishes the proportions of *br*-GDGTs to crenarchaeol (see the equation 1.7) (Hopmans *et al.* 2004). BIT index values close to 1 indicates high terrestrial inputs while values close to 0 represent exclusively to aquatic inputs. Typically the BIT index is used either in multiproxy studies focusing on the abundance and distribution of terrestrial organic matter in continental margins

(Hopmans *et al.* 2004; Escala *et al.* 2009; Peterse *et al.* 2009; Peterse *et al.* 2015) or in climate studies to assess extraneous inputs of crenarchaeol' and hence the reliability of TEX_{86} values (Weijers *et al.* 2006).

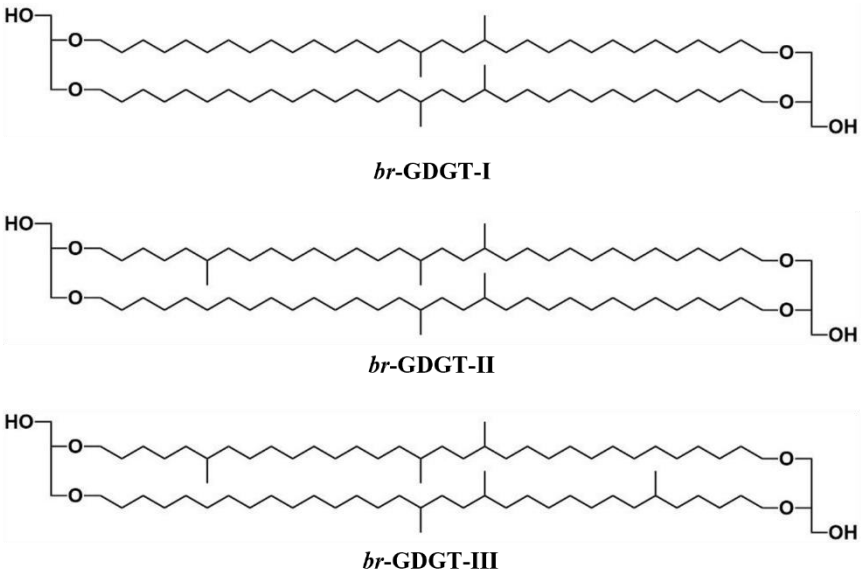


Figure 1.7. Branched glycerol dialkyl glycerol tetraethers lipid (*br*-GDGT) cores found in terrestrial sediments, lacustrine, soils and peats (Hopmans *et al.* 2004).

$$BIT = \frac{([br-GDGT-I]+[br-GDGT-II]+[br-GDGT-III]}{[br-GDGT-I]+[br-GDGT-II]+[br-GDGT-III]+[crenarchaeol]} \quad (1.7)$$

Thus, the BIT index has been used to indicate if organic matter (OM) input in depositional environments is from terrestrial or aquatic settings. Analysis of terrestrial soil and peats gave predominant signals of *br*-GDGT-1-3, resulting in BIT values close to 1, whereas the GDGT lipids in surface water sediments from a river basin were dominated by crenarchaeol, giving BIT values close to 0 (Hopmans *et al.* 2004; Weijers *et al.* 2006). The TEX_{86} values should be considered overestimates when the BIT values are >0.3 due to the influence of terrestrial input of soil-derived isoprenoid GDGTs, strongly suggesting that both indices should be measured (Weijers *et al.* 2006). Another issue of concern relates to the relative amounts of *i*-GDGTs *vs.* *br*-GDGTs in different terrestrial regions. Notably, there is no universal cut-off for BIT values

as BIT does not relate to terrestrial organic matter fluxes in a straightforward manner (Weijers *et al.* 2006; Schouten *et al.* 2013). Furthermore, *br*-GDGTs can also be produced *in-situ* in marine environments (Peterse *et al.* 2009). Hence the BIT index may not provide a meaningful cutoff value to evaluate the reliability of TEX_{86} values (Schouten *et al.* 2013). The BIT index in coastal or open marine sediment is commonly very low, reflecting extremely low levels of *br*-GDGTs. Under such conditions, the measurement of *br*-GDGTs carries large errors and a refinement of the analytical process used to quantify GDGT lipids is warranted.

Limitations of past temperature reconstruction was evaluated by comparing TEX_{86} and BIT values in sediments and their fractionated GDGT extracts from different locations (Salt Pond, Carolina Margin, Arabian Sea) analysed by 35 different laboratories in a round-robin study (Schouten *et al.* 2009, 2013). The comparison reflects the effects of different sediment extraction techniques and different analytical instrumentation. The round-robin study showed that SST assessments from TEX_{86} varied over the range 1.3 - 3.0°C, which is similar to the extent of variation from other palaeotemperature proxies. Notably, the BIT values obtained from the same samples were in range of 0.013 to 0.042, indicating good reproducibility among the laboratories involved in the analysis. Comparison of the measured TEX_{86} values of each sample showing the most similar trends in their values, either lower or higher than the mean, reflects the consistency within individual laboratories with no effect from different extraction methods. Comparing the MS signal response of GDGT lipids from each sample detected using different types of mass spectrometer reveals lower signal response from ion trap mass spectrometers than those obtained from single quadrupole mass spectrometers at a statistically significant level (Student's *t* test, $p < 0.05$), reflecting a need for standard GDGT mixtures in order to correct the MS signal response of each instrument. The study illustrated that the current challenges in SST assessment via the TEX_{86} and BIT indices relate to the reliability of conventional method for GDGTs analysis in resolving structural complexity among GDGTs, establishment of calibration models and interlaboratory applicability.

1.2.3 Analysis of Archaeal GDGT lipids

The analytical procedure for measurement of GDGT lipids obtained either from living Archaea or extracted from cellular remains in sediments requires several elements: a purified concentrate containing the components, a satisfactory separation to resolve the structurally complex components and a favourable detection method that enables quantification of trace amounts. The extraction of GDGT lipid cores utilises extraction techniques such as Bligh & Dyer, ultrasonication, microwave, Soxhlet extraction or accelerated solvent extraction (ASE) with dichloromethane and methanol to produce an extractant described as a polar lipid (PL) extract (Pitcher *et al.* 2009; Lengger *et al.* 2012; Schouten *et al.* 2013). Comparison of TEX_{86} values determined from sediment extracts gave SST variations in range of -0.8 to 0.9°C regardless of extraction method, indicating that the results are not significantly impacted by the method of extraction (Pitcher *et al.* 2009; Schouten *et al.* 2013). The PL extract could contain a broad range of lipid classes including isoprenoid GDGT lipid cores with or without the phosphate and/or sugar polar head groups (Figure 1.8). These are cleaved by acidic methanolysis to obtain the GDGT lipids cores which are isolated and purified using activated silica or alumina column chromatography with increasingly polar solvents, to produce a liberated GDGT lipid core fraction (Knappy *et al*. 2009; 2012).

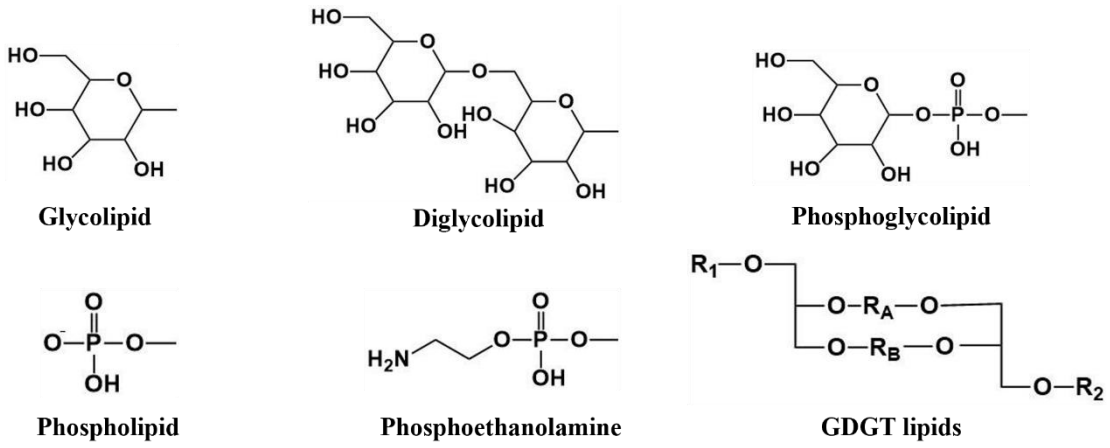


Figure 1.8. Examples of head groups commonly consisted of Archaeal tetraether lipid cores (Pitcher *et al.* 2009; Law and Zhang 2019). R_1, R_2 = head groups or H. R_A, R_B = different forms of C_{40} biphytanyl.

Conventional gas chromatography-mass spectrometry (GC-MS) methods are not suitable for GDGT analysis; high temperature GC-MS has been achieved but it is not routinely available (Sutton and Rowland 2012; Lengger *et al.* 2018). As an alternative to GC-MS analysis of intact GDGT lipid cores, chemical degradation approaches have been used, with characterisation of the constituent alkyl chains. The GDGT lipid cores are commonly analysed in detail by liquid or gas chromatographic-based separations with mass spectrometric detection. The peaks in the mass chromatograms obtained are identified as GDGT lipid cores of a series of components exhibiting protonated molecules with mass to charge ratios in the range *m/z* 1286-1302. The relative amount of GDGT lipid cores can be estimated using a synthesised C_{46} GTGT as an internal standard (Figure 1.9a) or by comparing the measured peak areas to a standard curve of a GDGT-0 standard (Wuchter *et al.* 2004; Huguet *et al.* 2006).

1.2.3.1 Gas chromatographic-based separation

Lennger *et al.* noted lower GDGT responses of 70 eV EI-GC-MS than FID (Lengger *et al.* 2018) which they attributed to low ionisation efficiency. Furthermore, due to the relatively high molecular masses (1000 to >1300 Da) of the GDGT lipid cores, high temperature GC was required. An approach of chemical degradation of the ether bonds to obtain hydrocarbon chains, e.g., biphytanes, has been used to obtain sufficient volatility for analysis (De Rosa *et al.* 1988; Schouten *et al.* 1998). The ether bonds of the isoprenoid GDGT lipid cores can be cleaved using hydroiodic acid (HI) or boron tribromide (BBr_3) to obtain the alkyl halides (Schouten *et al.* 1998; Weijers *et al.* 2010; Liu *et al.* 2018). The isolated alkyl halides are further hydrodehalogenated by lithium aluminium hydride ($LiAlH_4$) or lithium triethyl-borohydride (superhydride) to give the biphytane or branched hydrocarbons (Liu *et al.* 2018; Liu *et al.* 2016; Schouten *et al.* 1998; Weijers *et al.* 2010). This approach to detect GDGT lipids in peats gave insufficient yield to recover hydrocarbon GDGT skeletons as their signatures are easily swamped by other hydrocarbons formed by the acid-catalysed dehydration of alcohols (Weijers *et al.* 2010). Furthermore, the use of $LiAlH_4$ for reduction of GDGT skeletons can produce lithium hydroxide solids, which increase the complexity of the subsequent purification step. An alternative approach

for hydrogenation employed Adams' catalyst (platinum dioxide, PtO_2) and offers greater simplicity in formation of the products (Kaneko *et al.* 2011; Law and Zhang 2019). Another GC-based approach to detect the GDGT hydrocarbon used soft ionisation, specifically supersonic molecular beam (SMB) ionisation operated at a low ionization energy (~20 eV) (De Jonge *et al.* 2013; Sinninghe Damsté *et al.* 2018). The method improved the stability of the molecular ions which were unstable under EI at 70 eV, revealing novel discovery of hexamethylated *br*-GDGTs in Siberian peat (De Jonge *et al.* 2013) and identification of the *cis* stereoisomer of crenarchaeol' (Sinninghe Damsté *et al.* 2018). The SMB approach gives the potential for more complete consideration of EI fragmentation with lower levels of background ions (Seemann *et al.* 2015; Tsizin *et al.* 2017). Notably, however, approaches involving ether cleavage give inconsistent reaction yields and cannot provide the information about the combinations of the biphytanes in GDGT lipid cores and so are not applicable in determination of TEX_{86} or BIT proxies (Schouten *et al.* 2013; Lengger *et al.* 2018; Law and Zhang 2019).

An alternative approach to GC-based GDGT analysis uses high temperature gas chromatography with flame ionisation detection (HTGC-FID) operated at 400°C to detect the GDGTs in their TMS ether forms (Sutton and Rowland 2012; Lengger *et al.* 2018). The method also provides the ability to couple with time-of-flight mass spectrometry for structural identification and shows the potential for quantification of both isoprenoid GDGTs and branched GDGTs using external calibration of the response relative to an internal standard of authenticated GDGT (Lengger *et al.* 2018). The HTGC-MS approach is also applicable to the determination of the δ^2H values of GDGT lipids, indicating the potential to reconstruct environmental conditions (Lengger *et al.* 2021). The applicability of HTGC-MS is currently challenged by incomplete resolution of individual GDGT species and low relative abundances of the molecular ion peaks of GDGT lipids compared to background ions in the spectra (Lengger *et al* 2018; Law and Kai 2019).

1.2.3.2 Liquid chromatographic-based separation

The GDGT lipid cores can be detected in their native form by UHPLC-MS using an atmospheric pressure chemical ionization (APCI) or electrospray ionization (ESI) as detectors. The analysis of GDGT lipids is commonly performed by normal-phase liquid chromatography (NPLC) coupled to APCI source (Hopmans *et al.* 2000; Schouten *et al.* 2007). Conventionally, silica columns modified with amino- or cyano-groups have been used to separate GDGT lipid cores based on differences in polar interactions and in steric configurations. The typical APCI mass spectra of GDGT lipids show dominant protonated molecular ions $[M+H]^+$ with minor intensity ions corresponding to $[M-18+H]^+$ and $[M-74+H]^+$, resulting from losses of water and glycerol, respectively. The fragment ions have been used as characteristic profiles in the identification of archaeal ether lipids (Schouten *et al.* 1998). Structural complexity in GDGT lipid cores arises due to isomeric structures (De Rosa *et al.* 1980, 1983). The GDGT lipid cores containing even numbers of cyclopentane rings were attributed structures in which the two etherified chains are identical whereas those containing an odd number of rings were ascribed inequivalent isoprenoid chains containing numbers of rings which differ by one. The original assumption regarding the difference between chains of either 0 or 1 cyclopentane rings is incorrect and differences of more than one are evident from tandem mass spectrometry analysis (Knappy *et al.* 2012). The isomeric forms of GDGT lipid cores also have different distributions of isoprenoid chains containing numbers of cyclopentane rings by two, for instance GDGT-2 having one isoprenoid chain without cp ring and another chain having 2 cp rings, GDGT-4 having one cp ring and three cp rings on its each etherified chain (Figure 1.6). Isomeric GDGT lipids having different distributions of cp rings between isoprenoid chains, e.g. isomers of GDGT-4, coelute as the resulting of different in polarity being negligible. A silica column coated with diol groups was used to screen the structural diversity of intact polar core lipids by LC-ESI-ion trap MS (Sturt *et al.* 2004; Yoshinaga *et al.* 2011). This method has limited capability to distinguish different lipid cores due to being insufficiently apolar to retain apolar components in the column.

Approaches using low polarity stationary phase columns such as hydrophilic interaction chromatography (HILIC) and reverse-phase liquid chromatography (RPLC) columns enabled a

refinement of separation efficiency and discovery of new Archaeal ether lipids. Comparison of GDGT separations using RP and HILIC columns reveals that HILIC columns offer shorter retention times and greater resolution than the NP column functionalised with diols (Wörmer *et al.* 2013). This column also enables the use of water and water-miscible organic solvents that are more amenable to ESI. Notably, the use of RP columns has enabled separation of IPL and GDGT lipid cores in a single analysis with improved chromatographic separation of the GDGT lipid cores than with the HILIC columns (Wörmer *et al.* 2013). The elution order under RP conditions depends on the molecular dimensions and the number of unsaturated or cyclic moieties in the structure (Zhu *et al.* 2013; Rattray and Smittenberg 2020). The RP separation results in sharper peaks, higher signal-to-noise ratio and greater response, suggesting the potential to use in exploring structural complexity within GDGTs. Unfortunately, the use of RP columns for separation of GDGT lipid cores has not been widely considered, due either to the longer analysis time than for NP (Wörmer *et al.* 2013; Zhu *et al.* 2013) or partial co-elution between isoprenoid GDGT-4 lipid species (Rattray and Smittenberg 2020).

The GDGTs are commonly detected in UHPLC-MS using selected ion monitoring (SIM) mode (Schouten *et al.* 2007) or selected/multiple reaction monitoring (SRM/MRM) mode (Pitcher *et al.* 2011; Chen *et al.* 2016). The amount of GDGTs can be assessed based on the signal response of the protonated molecules of GDGTs relative to a C_{46} GTGT (Figure 1.9a) as an internal standard, due to a limited availability of the commercial GDGT standards for external calibration (Huguet *et al.* 2006). The MS signal responses obtained from the GDGTs having different structures, (e.g., *i*-GDGTs and *br*-GDGTs) results in some insecurity in quantification due to uncertainties regarding the relative ionisation efficiencies and the instrumental tuning (Schouten *et al.* 2009, 2013). Many approaches have been reported to refine GDGT analysis, including use of UHPLC-MS to improve chromatographic behaviour (Hopmans *et al.* 2016), time-of-flight secondary ion MS (Sjövall *et al.* 2008) to improve MS detection efficiency, laser desorption coupled to Fourier-transform ion cyclotron resonance MS (Wörmer *et al.* 2014, 2019), and hybrid ion trap-orbitrap MS (Jensen *et al.* 2015). GDGT lipid cores and their IPL derivatives have been detected in living Archaeal membranes by matrix-assisted laser desorption/ionization coupled to

time-of-flight mass spectrometry (MALDI-TOF/MS) (Macalady *et al.* 2004; Angelini *et al.* 2010; Lobasso *et al.* 2012). These approaches provide greater security in identification but still suffer limitations in quantification owing to the lack of authentic standards.

Discrepancies in MS response caused by differences in instrument tuning can be avoided using spectrophotometric-based detection methods such as ultraviolet-visible (UV-vis) and fluorescence (FL), both of which are favourable techniques for quantification and offer the ability for absolute quantification of analytes (Ohtsubo *et al*. 1993; Ito *et al.* 2017; Sun *et al.* 2017). The signal response of the molecules obtained from the UV-vis or FL are dependent on the value of absorption coefficient (ε) or fluorescence quantum yield (Φ_F) of the active molecule, potentially offering an absolute measurement of the quantity of analytes. A few reports considered derivatisation with an 9-anthroyl group to detect Archaeal ether core lipids by LC-FL (Ramasha *et al.* 1989; Ohtsubo *et al.* 1993; Bai and Zelles 1997). The method enables detection of Archaeal membrane lipids with the use of 1,2-di-*O*-hexadecyl-*rac*-glycerol (dihexadecylglycerol) as internal standard (Bai and Zelles 1997). Though the method shows potential for use in investigations of methanogenic ecosystems in the ppm levels (2 ng of tetraether lipids and 0.4 mg/kg for environmental samples), profiling of GDGT lipids is limited by co-elution of GDGT peaks. Such a spectrophotometric approach would enable trace analysis while providing security in quantification without a requirement for an internal standard for each GDGT structural type.

A synthesized C_{46} GDGT has been used as an internal standard for measurement of GDGTs by MS detection. It is, however, expensive, and not widely available commercially (Huguet *et al.* 2006). The availability of an internal standard that can be used as a retention time marker would enable measurements to be performed without recourse to an MS detector, thereby increasing capacity and opportunity for new researchers and laboratories to perform GDGTs analysis. Previous work at York examined the use of fluorenylmethoxycarbonyl-lysine-*tert*-butoxycarbonyl (FMOC-Lysine-Boc, FLB, Figure 1.9c) derivatisation of GDGT alcohol groups (Poplawski 2017). Separation of lipid profiles employed a pentafluorophenyl (PFP) column and analysis using three different detectors (MS, UV, FL) (Poplawski 2017). The *N*-Fmoc protected

amino acids of GDGT lipid cores isolated from *Sulfolobus acidocaldarius* MR31 (*S. acidocaldarius* MR31) along with a laboratory-synthesised 1,2-di-*O*-octadecyl-*rac*-glycerol (r-dOG, Figure 1.9b) internal standard were prepared and an acceptable RP-HPLC separation developed on a PFP column. The results demonstrate the potential for development of an analytical method to measure long-chain alcohols (octadecanol and GDGTs) with MS, UV-vis and FL detectors while offering security in quantification and an opportunity for application by researchers lacking MS facilities.

As part of his study to improve the detection and analysis of the GDGT lipid cores of Archaea by FLB derivatisation, Poplawski examined Fmoc-amino acid-Boc as a versatile derivatising agent allowing for UV, FL and MS detection (Poplawski 2017). The Fmoc group exhibits an excellent chromophore (UV absorption: λ_{max} 263 nm; ε 18950; Chang *et al.* 2009), allowing detection of hydroxyl-containing lipids under UV or fluorescence detection as alternatives to MS detection (Melucci *et al.* 1999). The Boc group is widely used for *N*-protection of amino acids, and the derivatives are easy deprotected, allowing response to MS detection (Ho and Ome 1987). Poplawski examined the suitability of *N*-Fmoc protected amino acids as derivatisation agents for GDGT lipid cores by using 1-octadecanol as model compound due to a limited availability of GDGT materials (Poplawski 2017). Either Fmoc or Boc groups linked with different amino acids, glycine (Gly), phenylalanine (Phe), proline (Pro), tryptophan (Tyr), and lysine (Lys) were evaluated for their signal responses to UV, FL and MS detection. All derivatising agents tested were used successfully with octadecanol, obtaining overall product yields by weight of more than 90%, indicating near complete derivatisation. The FLB derivative was selected to use for further assessment of its suitability to detect GDGT lipid cores as it contains two basic nitrogen groups (Fmoc and Boc), that could potentially enhance the MS signal response by facilitating protonation. Unfortunately, time restrictions limited the analysis to a preliminary screening of lipids in a geological sample where signal suppression of GDGTs in the derivative form was identified as a significant obstacle (Poplawski 2017).

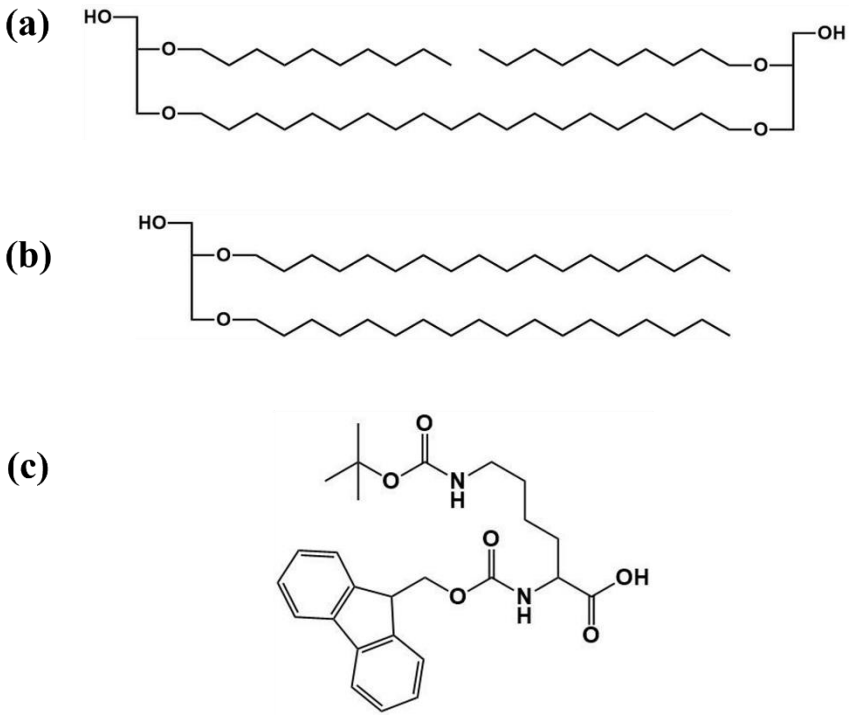


Figure 1.9. Chemical structures of (a) C_{46} GTGT, (b) 1,2-di-*O*-octadecyl-*rac*-glycerol (r-dOG), (c) Fmoc-Lysine-Boc (FLB).

1.3 Aims and scope of this study

Steroidal and GDGT lipids have been widely used as biological biomarkers for paleoenvironmental and environmental interpretations and have in common the presence of at least one of alcohol group. This feature enables the use of conventional GC-MS methods for steroidal and GDGT lipids analysis following derivatisation of the alcohol groups to silyl ethers. Quantification of these components by the conventional GC-MS approaches can be compromised in situations of limited availability of materials either due to low amounts of analyte or limited volatility in the case of GDGTs. The work described in this thesis focuses on improving and evaluating a novel derivative for the analysis of alcohols by HPLC with the view to enhancing the analytical capabilities for detecting alcohols. The method has been applied to two key groups of biomarkers having alcohol groups, steroids and tetraether lipids, owing to their importance and utility in studies of contemporary environments and in archaeological and paleoenvironmental studies.

This thesis presents a systematic study of hydroxyl lipids derivatised with FLB by Steglich esterification, carried out to assess the suitability of the approach for the analysis of steroidal and tetraether alcohols in sediments. The approach has the potential to enhance the detection of trace level steroids in fields such as geochemistry, medicine, archaeology, and environmental analysis. Chapter 2 describes experimental details, while Chapter 3 presents the development and evaluation of an analytical method to assess FMOC-Lysine-Boc (FLB) as derivatising agent for measurement of steroidal alcohols in soils and sediments and to demonstrate applicability of the method to archaeological samples. The refinement of the analytical steps of derivatisation, sample clean-up and LC separation has been examined systematically. Evaluation of an internal standard for use in detection and quantification of cholesterol and its related stanols (coprostanol and 5α-cholestanol) has been carried out. The full range of analytical figures: repeatability, reproducibility, limit of detection, limit of quantification, linearity and recovery have been reported. Chapter 4 presents the application of the FLB derivatisation approach to measure cholesterol and its related stanols in a wide range of archaeological and environmental samples for the purpose of detecting human occupation and activity. The coprostanol ratio of the samples was re-evaluated by UHPLC-MS, UHPLC-UV and HPLC-FL based on the relative ratios of the signals from the different detectors. Chapter 5 presents a method development of GDGTs analysis by FLB to measure proxies based on the Archaeal lipids. The refinement of the analytical steps of sample preparation, derivatisation, sample clean-up and MS condition have been examined systematically. Applicability of the method to determine and quantify GDGT lipid profiles in FLB form has been tested and compared with analysis of the native form for the purpose of evaluating signal enhancement. The TEX_{86} and BIT indices of geological sediments from an international round-robin study have been re-assessed by UHPLC-MS, UHPLC-UV and HPLC-FL for the purpose of demonstrating a suitability of the method. Chapter 6 summarises points of this work and proposes ideas to extend the method for further study.

Chapter 2

Experimental

2.1 Chemical and materials

All solvents used (methanol, ethyl acetate, hexane, dichloromethane, HPLC water, chloroform, methyl tertbutyl ether) were of analytical grade or higher (Fisher, Loughborough, UK). All chemicals used were of analytical grade: acetic acid, cholesterol, 5α-cholestanol, trichloroacetic acid, hydrochloric acid (Fisher, Loughborough, UK), coprostanol (Insight Biotechnology, Wembley, UK), and heptadecanol (Sigma Aldrich, Gillingham, UK). Derivatisation reagents of certified purity were obtained from Sigma Aldrich (Gillingham, UK): 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride, EDC (99%), 4-(dimethylamino) pyridine, DMAP (99%), Fmoc-lysine-Boc (99%). All glassware was cleaned by soaking in Decon-90 (1% *v/v*, Deacon, Hove, UK) at least 24 hours prior to rinsing with deionised water and acetone, respectively, and dried. All glassware was then baked (450 °C, 6 h) prior to use using a Pyroclean Trace oven (Barnstead International, USA) to remove organic contaminants.

2.2 Preparation of lipid extracts from soils, sediments and microbial cell pastes

2.2.1 Steroidal lipids extraction from soils and sediments

Soil samples were obtained in 2010 and 2011 during excavations of burials from Hungate, York (sample prefixes SK51350 and 51151). The samples were prepared following the procedures of the InterArChive project (Pickering *et al.* 2018). Briefly, soil samples were freeze-dried to remove moisture before sieving to obtain the fine material (< 200 μm) for analysis (approximately 3-5 g). Extractable organic matter was recovered by accelerated solvent extraction (ASE) (ASE 350, Dionex, Hemel Hempstead, UK) using a mixture of dichloromethane (DCM) and methanol (9:1, *v/v*) for 5 min at 100°C, 1500 psi, 3 cycles. After extraction, the solvent was removed using a rotary vacuum concentrator (RVC 2–25, Christ, Osterode am Harz, DE). The extracts were fractionated using homemade chromatography columns (5 cm × 0.55 mm *i.d.*) of silica gel 60 (40-45 mg) eluted sequentially using three column volumes each of: (1) hexane, (2) hexane:toluene (1:1. *v/v*), (3) hexane:ethyl acetate (4:1, *v/v*), (4) DCM:methanol (1:1, *v/v*), to generate hydrocarbon (HC), aromatic hydrocarbon (ArHC), medium polar (MP) and highly polar

(HP) fractions, in turn. The fractionated samples were dried (RVC) and stored at -20°C until used. The MP fractions were used for steroidal analysis in this work.

2.2.2 Extraction of Archaeal tetraether lipids extraction from cell pastes

Archaea species of *Sulfolobus* were grown aerobically in basal salt medium (Brock *et al.* 1972) at approximately 75°C, in pH range 3.0-3.5 adjusted using sulfuric acid (H_2SO_4) and were provided by Professor Daneila Barillà (Department of Biology, University of York). Cells (approximately 0.1-0.8 g) of *Sulfolobus solfataricus* (*S. solfataricus*) and *Sulfolobus acidocalarius* MR31 (*S. acidocalarius* MR31) were thawed at room temperature prior to extraction via a modified Bligh-Dryer technique to recover Archaeal polar lipids (C. S. Knappy et al. 2012b). The cells were suspended in methanol (20 mL), chloroform (10 mL) and aqueous trichloroacetic acid (4 mL; 10% *w/v*). The mixture was stirred at room temperature for 2 hours. To the mixture was then added chloroform (10 mL) and water (10 mL) prior to standing for 1 hour to enable separation of the two phases; the chloroform layer was isolated. The aqueous/methanol layer was further extracted with chloroform (2 × 20 mL), the organic layers combined and reduced in *vacuo* to give the lipid core extract.

Lipid core extracts (from *S. solfataricus* or *S. acidocalarius* MR31 species) or thawed cellular material (all other organisms; approximately 0.1-1 g) were subjected to acidic methanolysis by refluxing in methanolic HCl (4.8 M; prepared from aqueous HCl and MeOH) at 100°C in order to remove modified phosphate and/or glycosyl polar head groups. Lipid cores were partitioned into dichloromethane (DCM), passed through a plug of anhydrous magnesium sulfate ($MgSO_4$) to remove moisture, and reduced in *vacuo*.

The lipid core extracts were dissolved in a minimal amount of DCM and loaded onto columns of activated alumina (5 mm × 50 mm) pre-washed with hexane: DCM (9:1, *v/v*). Apolar components were removed via washing with three bed volumes of hexane: DCM (9:1, *v/v*) and were discarded. Unless stated otherwise, the polar fraction containing ether lipid cores was eluted using three bed

volumes of hexane:ethyl acetate (1:1, *v/v*) before being dried under a gentle stream of nitrogen to give the polar lipid fraction.

2.3 Derivatisation of lipid extracts

2.3.1 Derivatisation of steroidal compounds using FMOC-Lysine-Boc

Steroidal compounds (cholesterol, coprostanol and 5α-cholestanol) were derivatised by esterification with FMOC-Lys-Boc (FLB). Cholesterol and its cholestanol derivatives were prepared according to the method developed by Poplawski (Poplawski 2017). To the sterol (*c*. 5 mg) was added DMAP (1 eq.), EDC (2 eq.), FMOC-Lys-Boc (1.5 eq.) and DCM (2 mL) and the solution was sonicated (60 min), unless stated otherwise. For the archaeological and environmental soil samples, the ratio of the reactants (DMAP/FLB/EDC) was 12/25/50 and the sonication time was extended to 120 min. The reaction mixture was passed through a short plug of silica gel in a glass column and washed with DCM (4 column volumes) to remove unreacted reagents. The product was eluted with hexane-ethyl acetate (3:1 *v/v*) twice with two column volumes to give fractions 1 and 2, respectively. Fraction 1 was discarded. Fraction 2 was dried under a gentle stream of nitrogen before storing dry. Fraction 2 was reconstituted in DCM (1 mL) with sonication for 2 min prior to HPLC analysis.

2.3.2 Derivatisation of steroidal compounds using BSTFA

The medium polar fraction from soil extracts was reacted with 100 µL of *N,O*-bis (trimethylsilyl) trifluoroacetamide (BSTFA, containing 1% trimethylchlorosilane) and pyridine (5 drops) at 70°C for 90 min. The reaction mixture was dried under a gentle stream of nitrogen to give the trimethylsilyl ether (TMS) lipid fraction. The silylated ether lipid sample was stored at 4°C and dissolved in DCM (100 µL) for GC analysis.

2.3.3 Derivatisation of tetraether lipids from lipid core extracts using FMOC-amino acid-Boc

The lipid core extracts of GDGTs from *S. solfataricus* and *S. acidocaldarius* cultures were derivatised by esterification with FMOC-Lys-Boc according to the method developed by Poplawski (Poplawski 2017). The ratio of the reactants (DMAP/FLB/EDC) was 10/50/75 in DCM (2 mL), unless stated otherwise. The sonication time was extended to 120 min. The reaction mixture was passed through a short plug of silica gel in a glass column (0.5 cm × 5 cm) and washed with DCM (3 column volumes) to remove unreacted reactants. The product was eluted with hexane:ethyl acetate (1:1, *v/v*) before being dried under a gentle stream of nitrogen. The resulting sample was stored dry and diluted in DCM (100 µL) for UHPLC analysis.

2.4 Structural characterisation of derivative compounds by mass spectrometry

An HCTUltra ETD II ion trap mass spectrometer (Bruker, Daltonics, Coventry, UK) was used in positive mode with either an ESI or an APCI ion source. For APCI, the operational parameters were set as follows: nebulizer gas (N_2) pressure 45 psi; drying gas (N_2) flow 3 L min^{-1}; drying gas 300°C; vaporizer temperature 450°C; capillary voltage -3700 V. The sample solutions were dissolved in DCM (1 mL) before directly injecting (1 µL) into the ion trap by syringe at 0.4 mL min^{-1}, unless stated otherwise. For ESI, the operating parameters were set as follows: nebulizer pressure 20 psi, dry gas flow 5 L min^{-1}, drying temperature 300ºC, unless stated otherwise. Mass spectra were acquired over the range 100–1000 *m/z*, unless otherwise stated. Tandem MS (MS^n) spectra were recorded using an HCTUltra ETD II ion trap mass spectrometer. 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 to 3 *m/z*, the maximum accumulation time set to 40 ms and the fragmentation amplitude fixed at 1.2 V. In some instances, the smartFrag feature was used.

2.5 Analysis of steriodal lipids

2.5.1. Ultra high performance liquid chromatography with mass spectrometry (UHPLC-MS) and/or diode array detector (UHPLC-DAD)

The steroid Fmoc-lysine-Boc derivatives of cholesterol, coprostanol and 5α-cholestanol were analysed using an Ultimate 3000 rapid separation chromatograph (Dionex, Sunnyvale, California, US) coupled with the HCTUltra ETD II ion trap spectrometer (Bruker, Daltonics, Coventry, UK). The sample volume of 1 μL was injected. Separation was carried out on an Acquity BEH C_{18} column, 2.1 mm × 150 mm, 1.7 μm) maintained at 45°C. The ternary solvent system comprising A = methanol, B = 0.5% acetic acid in water, C = dichloromethane was delivered with a constant flow rate of 0.3 mL min^{-1}. The gradient changed from 68:11:21 (*v/v*) A:B:C to 52:8:40 (*v/v*) over 7 min followed by change to 52:8:40 (*v/v*) of A:B:C over 5 min, the final composition being held for 8 min. The solvent was returned to the initial composition and held for 5 min for column re-equilibrium. The MS detection was operated as stated in Section 2.4. The diode array detector (DAD) (Dionex, Sunnyvale, California, US) was monitored over the range of 190-400 nm (2 nm resolution) and recorded at peak width response time >0.05 min (1 s). All analyses were performed in triplicated (n = 3), unless stated otherwise.

2.5.2 High performance liquid chromatography with fluorescence (HPLC-FL)

The steroid FLB derivatives (chol-FLB, cop-FLB and 5α-chol-FLB) were analysed using an Agilent 1100-UV series chromatograph (Palo Alto, California, US) with a Jasco FP-920 fluorescence detector operated with excitation wavelength (λ_{ex}) 263 nm and emission wavelength (λ_{em}) 309 nm. Owing to the pressure limit of the fluorescence detector (<400 bars), the separation was carried out using Symmetry C_{18} column (2.1 mm × 150 mm, 3.5 μm) maintained at 45°C using the same mobile phase composition and gradient programme as section 2.5.1 using a 1-μL injected sample volume. All analyses were performed in triplicate (n = 3), unless stated otherwise.

2.5.3 Separation of the steroidal lipids by gas chromatography with mass spectrometry (GC-MS)

The trimethylsilyl ether (TMS) lipids from soil extracts were analysed using a 7860A gas chromatograph (Agilent, Wokingham, UK) equipped with a 7683B Series auto sampler coupled to a GCT Premier Micromass time of flight mass spectrometer (Waters, Elstree, UK). Separation was achieved on a fused silica capillary column (Zebron, ZB-5, 30 m × 0.25 mm i.d., 0.25 μm film thickness, Phenomenex, Macclesfield, UK). The sample volume of 1 μL was injected onto the column using a split/splitless injector operated in split mode (280°C, split flow 1:5). The oven temperature was programmed as follows: initial temperature 70ºC ramping to 130ºC at 20ºC/min followed by ramp to 320ºC at 4ºC/min; hold for 40 min. Mass spectra were acquired over the range 50–750 *m*/z. Helium was used as a carrier gas at flow rate of 1 mL min^{-1}. A time of flight (TOF) mass detector (Agilent, Workingham, UK) was used with the GC/MSD interface set at 300ºC and MS source at 180 ºC. The ionization energy was 70 eV. Mass spectra were acquired over the range 50–750 *m/z* units. Instrument control, data acquisition and processing was by MassLynx V4.1 (Waters, Elstree, UK). Components were identified from their retention times and comparison of their mass spectra with the NIST database.

2.6 Tetraether lipids from GDGT extracts

2.6.1 Ultra high performance liquid chromatography with mass spectrometry (UHPLC-MS) and diode array detector (UHPLC-DAD)

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. The sample volume of 1 μL was injected. Separation was achieved on a Phenomenex Kinetex PFP column (4.6 mm × 150 mm, 2.6 μm) maintained at 45°C and eluting with methanol/0.1 % acetic acid in water/methyl tertbutyl ether (40/20/40, *v/v*) delivered isocratically at 0.9 mL min^{-1}. For MS detection, an HCTUltra ETD II ion trap mass spectrometer (Bruker, Daltonics, Coventry, UK) was used with an APCI or

ESI positive ion source as described in Section 2.4.2, excepted the vaporizer temperature of APCI-MS was operated at 500°C.

2.6.2 Separation of the FLB derivatives of GDGTs by high performance liquid chromatography with diode array detector (DAD) and fluorescence (HPLC-UV and HPLC-FL)

The FLB derivatives of tetraether lipids from either Archaeal cell or sediment extracts, were analysed using an HPLC-FL as described in Section 2.5.2. The sample volume of 1 μL was injected. The separation was carried out using Phenomenex Luna PFP (2) column (2.0 mm × 150 mm, 3.0 μm) maintained at 45°C. The binary solvent system comprising of methanol/0.1 % acetic acid in water/methyl tertbutyl ether was delivered with a constant flow rate of 0.4 mL min^{-1}. The gradient changed from 46.4:13.6:40 (*v/v*) to 41.0:13.6:45.4 (*v/v*) over 15 min followed by change to 40.15:13.6:40.8 (*v/v*) over 3 min and maintained for 5 min. The final composition was reached to 40.2:13.6:41.08 (*v/v*) and being held for 5 min. The solvent was returned to the initial composition and held for 20 min for column re-equilibrium, unless stated otherwise. The fluorescence detector was operated as stated in Section 2.5.2. The diode array detector (DAD) (Shimadzu UV-1800 spectrophotometer) was monitored at wavelength 263 nm (2 nm resolution) and recorded at peak width response time >0.05 min (1 s). All analyses were performed in triplicate (n = 3), unless stated otherwise.

Chapter 3

Method development for the analysis of cholesterol and related stanols as their Fmoc-Lys-Boc derivatives

3.1 Introduction

Gas chromatography-mass spectrometry (GC-MS) is the conventional method for detecting cholesterol and its related stanols (coprostanol, epicoprostanol, 5α-cholestanol) as their trimethyl silyl ether (TMS) derivatives (Bethell *et al.* 1994; Battistel *et al.* 2015; Prost *et al.* 2018). The limit of quantification of these components by GC-MS presents a challenge in situations where the availability of materials is limited either from the naturally of low concentrations or small sample sizes. Detection of steroidal lipids in soils and sediments for archaeological and environmental applications has required use of a wide range of injection volumes (1-20 μL, spitless mode; Table 1.1). Gas-chromatography (GC) provides chromatographic retention times for each species and mass spectrometry (MS) enables their identification. Where large injection volumes are used baseline separation of closely eluting peaks such as the stanols may not be resolved among the structurally complex soils or sediment extracts. Hence, an alternative approach to improve responses for these components would be beneficial.

FLB is a promising derivatising agent for lipid analysis (both GDGTs and sterols) owing to the ease and efficiency of the reaction and the response of the derivatives to fluorescence (FL), ultra-violet (UV) and mass spectrometry (MS) detection. The chemical structure of FLB suggests that the Boc group offers the potential to enhance MS signal response (Du *et al.* 2003; Zhu *et al*. 2006), while the Fmoc group provides the properties of being chromophore and fluorophore, enabling both UV and FL detection (Eissler *et al*. 2017). Thus, the use of FLB as derivatising agent could offer the potential to screen for the presence of stanols related to cholesterol by MS, while the quantification could be performed by FL detector due to the selectivity of the derivatisation to compounds having hydroxyl groups (e.g. steroidal lipids and alkanols). Thus, FL chromatograms have the potential to offer greater security in quantification.

As part of a study to improve the detection and analysis of the glycerol dialkyl glycerol tetraether (GDGT) lipid cores of Archaea, Poplawski (2017) demonstrated that FMOC-Lysine-Boc (FLB see Figure 3.1a for chemical reaction) could be used to detect cholesterol in the soil around the remains of a piglet in the soil of an experimental burial site from the InterArChive project. The

method required a lower injection volume (1 μL) than many similar studies reported in the literature (Table 1.1 and 1.2). The use of FLB derivatives enables analysis by UHPLC-MS and may also offer improved detection limits than the conventional TMS ethers that are analysed by GC-MS. Unfortunately, time restrictions limited the analysis to a preliminary screening of lipids in the sample. Hence, the reliability of the method to identify and quantify lipids was not evaluated.

Poplawski demonstrated a successful derivatisation of cholesterol as the FLB derivative (chol-FLB) with an average yield by weight of 98 ± 0.2 % (n =3) (Poplawski 2017). The MS spectrum of chol-FLB gave the protonated molecule at *m/z* 837.5 (10.2% relative abundance, %RA) with fragment ions at *m/z* 781.5 (7.5 %RA), *m/z* 737.5 (100 %RA), *m/z* 615.5 (1.1 %RA) and *m/z* 369.2 (11.5 %RA). The fragment ions result from the losses of the *tert*-butyl (C_4H_9) group (−56 Da), the butyloxycarbonyl (Boc) group (-100 Da), fluorenylmethyloxycarbonyl (FMOC) group (-222 Da) and the cholesterol residue after ester-bond cleavage (−468.5 Da), in turn, confirming the formation of the FLB derivative of cholesterol. These characteristic mass losses can be used to indicate the formation of FLB derivative with other alcohols. Despite chol-FLB being prepared in excellent yield, the HPLC-FL chromatograms revealed moderate amounts of residues, reflecting a need for a refinement of the derivatisation process. Furthermore, the UHPLC-MS chromatograms of steroidal lipids in archaeological soil did not distinguish FLB derivatives of 5β-stanols and 5α-stanol related to cholesterol, indicating the need for an improved LC separation.

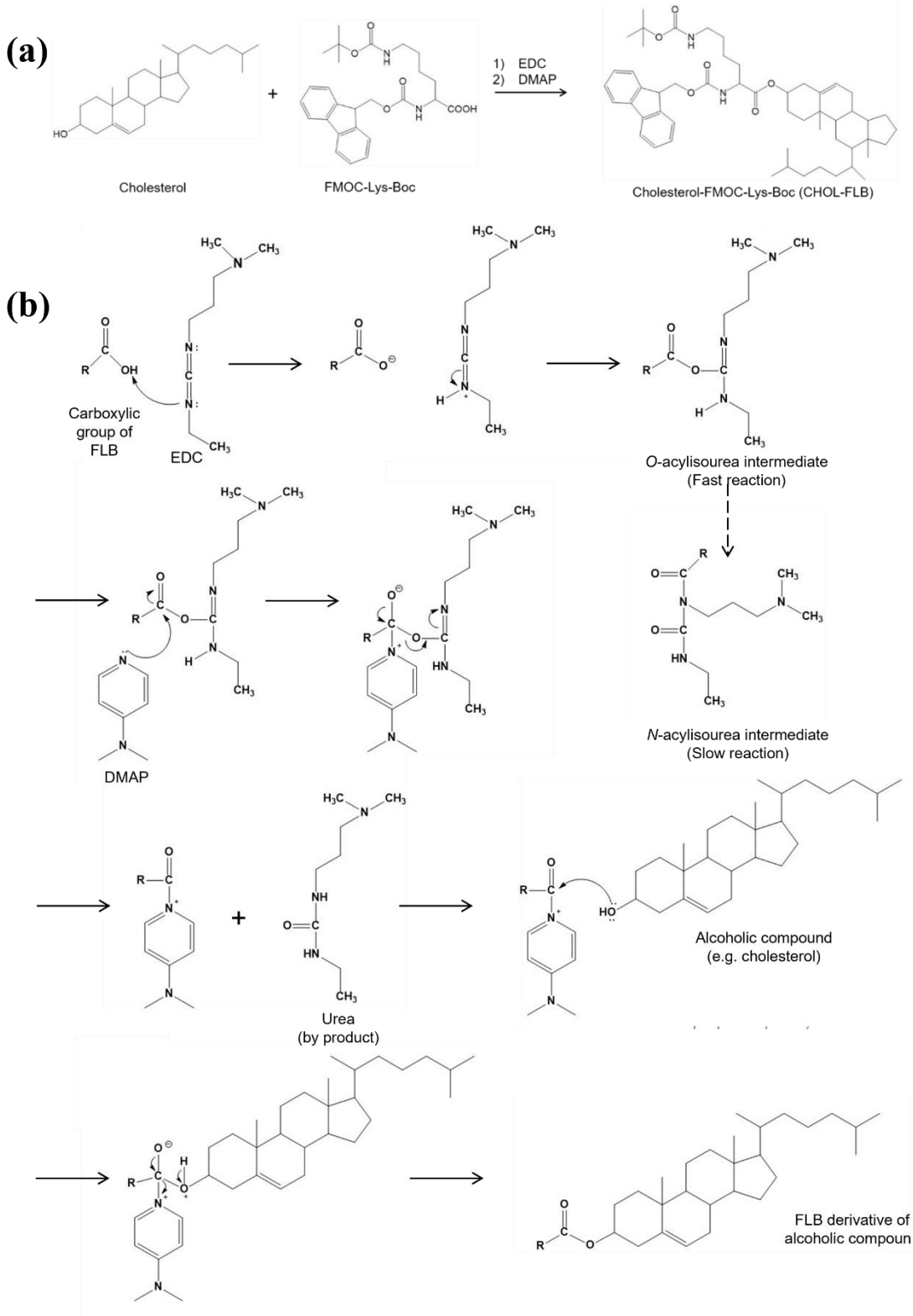


Figure 3.1. Schemes of (a) Derivatisation reaction of cholesterol by FMOC-Lys-Boc developed by Poplawski (2017), (b) Mechanism of FLB derivatisation.

The work described in this chapter involves the systematic evaluation of FLB derivatives of cholesterol and its transformation products (coprostanol and 5α-cholestanol) as an alternative approach for measuring steroidal alcohols in soils and sediments. The FLB derivatives were prepared by Steglich esterification as proposed by Poplawski 2017. The responses for steroidal FLB derivatives were obtained using different detectors (MS, UV and FL) and compared with the conventional GC-MS analysis of TMS ethers to assess the potential of improving detection limit. Refinements to the whole analytical process (derivatisation process, sample preparation and separation steps) were explored and a full-range method validation (working range, limit of detection, limit of quantification, repeatability, reproducibility, recovery) was assessed to evaluate the modified analytical method.

3.2 Aims

The overall aim of the work described in this chapter was to generate improved signal response for steroidal lipids by the following specific strategies:

1. To explore the potential to improve the completeness of the derivatisation process of using a standard of cholesterol (see Section 3.3.1.1).
2. To explore the MS^n dissociation pathway of the FLB derivatives of cholesterol and two of its reduction products (coprostanol and 5α-cholestanol) and their impacts on detection (see Section 3.3.1.2).
3. To determine if modifications could be made to the sample preparation and separation steps in order to improve and streamline the analytical process of cholesterol and related stanols (see Section 3.3.2).

3.3 Rationale, results and discussion

3.3.1 Formation of FMOC-Lys-Boc (FLB) steroids

3.3.1.1 Optimisation of the conditions for the formation of FMOC-Lys-Boc (FLB) derivatives of steroids

The FLB derivatives of steroidal lipids were prepared using Steglich esterification and the conditions for derivatisation with FLB were evaluated using cholesterol as a model. Conditions used for derivatisation of cholesterol were those of Poplawski 2017: reagent proportion was 1.5: 1: 2 of FLB: DMAP: EDC to 1 mole of cholesterol with sonication at room temperature for an hour to give the crude reaction mixture. After one hour of reaction time there was no longer any observable fluorescence response by thin layer chromatography (TLC) at the R_f value for cholesterol (data not shown), reflecting completion of the reaction. The residual reagents were removed from the crude reaction mixture using column chromatography on silica gel, based on the different polar interactions of the FLB derivatives and the derivatisation agent residues. Poplawski used a mixture of hexane:ethyl acetate (1:1, %*v/v*) to remove residue to give the isolated chol-FLB (Poplawski 2017). The isolated chol-FLB was weighed to estimate the product yield from derivatisation. Repeated preparation of chol-FLB gave an average yield by weight of 98 ± 0.2 % (n =3). However, incomplete removal of residue was evident from analysis of the derivative by HPLC using a reversed phase C_{18} column with ultraviolet (UV) and fluorescence (FL) detection (Poplawski 2017). The HPLC-FL chromatograms of chol-FLB prepared by the condition that Poplawski used gave two moderate peaks at retention times 1 and 2 min, one major peak at retention time 15.7 min and relatively small peaks between 3 and 14 min (Figure 3.2a). The earliest eluting peak was identified as FLB residue, and the major peak was identified as chol-FLB. The HPLC-FL chromatograms of chol-FLB reflect that the reaction yield estimated by weight is not accurate and the derivatisation conditions used resulted in incomplete derivatisation.

The conditions for derivatisation of steroidal lipids were further explored in this work to optimise the formation of chol-FLB and maximise signal response of steroidal compounds as the FLB

derivatives. The purification step was performed using a less polar mixture of hexane: ethyl acetate (3:1, %*v/v*) as eluent to increase retention of reagent residues on the silica gel. The isolated chol-FLB was analysed by HPLC-FL to obtain the peak area of the product. The HPLC-FL chromatograms show chol-FLB eluting at a retention time of 15.7 min, confirming formation of one major product. A small residue remained after purification, eluting in the retention time window 0-3 min (Figure 3.2b).

The dominant peak of chol-FLB obtained from the HPLC-FL chromatograms reflects that the isolated chol-FLB was produced as a single pure product. Hence, completeness of the derivatisation reaction will be reflected when the maximum peak area of the desired product is obtained, and the percentage yield can be estimated by either the fluorescence signal or weight. To ensure completion of the reaction the esterification was carried out using various molar ratios of FLB to cholesterol (Table 3.1), chol-FLB was isolated by column chromatography and analysed by HPLC-FL (Figure 3.2b). Product yields by mass and the FL peak areas responses obtained from the HPLC-FL separation are given in Figure 3.2c.

Table 3.1. Mole ratios of cholesterol and derivatising agents (FLB, DMAP, EDC).

Mole ratio of cholesterol: FLB: DMAP: EDC	**cholesterol (moles)**	**FLB (moles)**	**DMAP (moles)**	**EDC (moles)**
1: 0.5: 0.7: 1.5	6.0×10^{-6}	3.4×10^{-6}	4.1×10^{-6}	8.9×10^{-6}
1: 1.1: 1.0: 1.8	6.0×10^{-6}	6.2×10^{-6}	6.6×10^{-6}	1.1×10^{-5}
1: 1.8: 1.1: 2.3	2.1×10^{-5}	3.8×10^{-5}	2.4×10^{-5}	4.9×10^{-5}
1: 2.6: 1.6: 3.4	1.5×10^{-5}	3.8×10^{-5}	2.4×10^{-5}	4.9×10^{-5}
1: 5.6: 3: 7	6.7×10^{-6}	3.8×10^{-5}	2.3×10^{-5}	4.6×10^{-5}
1: 10: 2.4: 21	5.2×10^{-6}	5.3×10^{-5}	1.2×10^{-5}	1.1×10^{-4}
1: 15: 10: 21	6.0×10^{-6}	9.2×10^{-5}	6.5×10^{-5}	1.2×10^{-4}
1: 20: 5: 25	7.8×10^{-6}	1.5×10^{-4}	3.9×10^{-5}	2.0×10^{-4}
1: 24: 9.7: 43	7.3×10^{-6}	1.8×10^{-4}	7.0×10^{-5}	3.1×10^{-4}

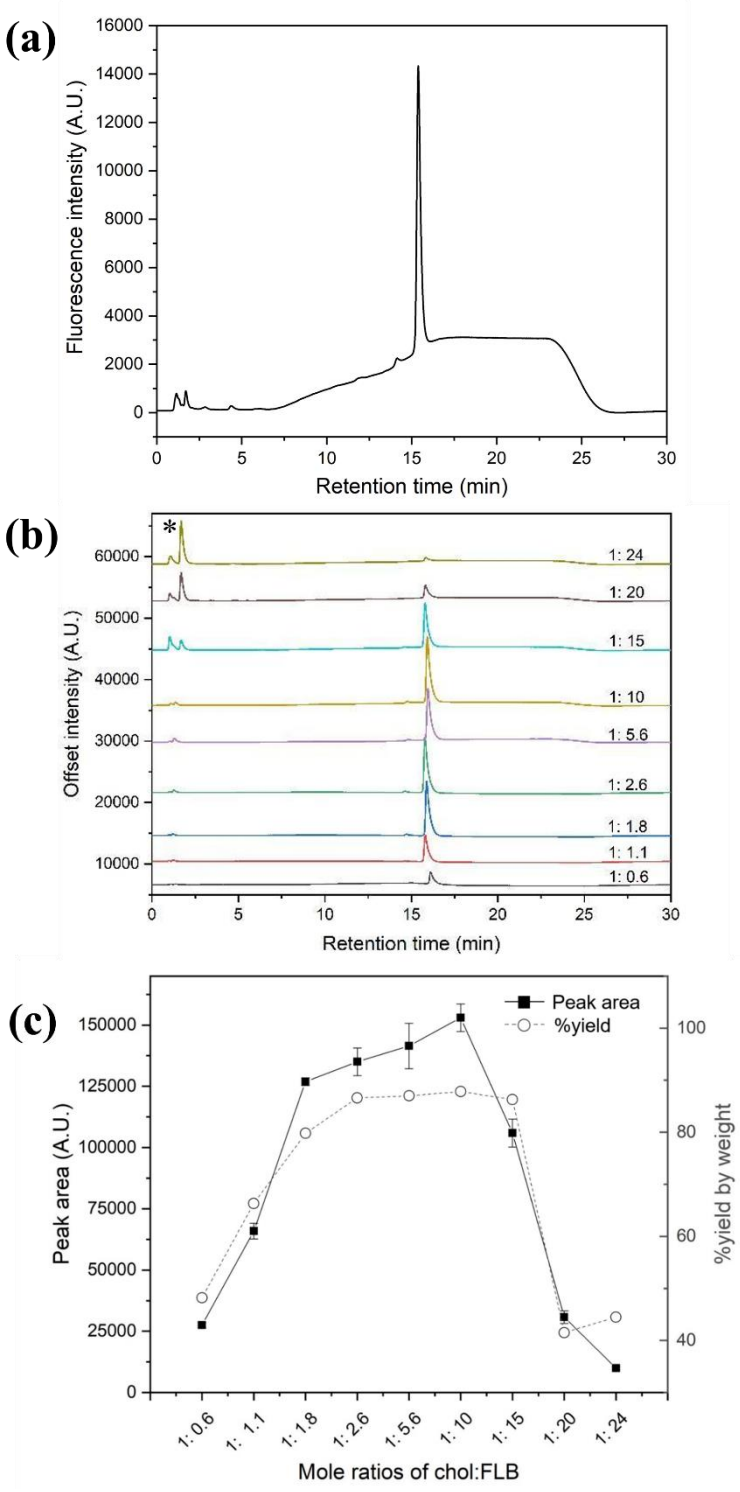


Figure 3.2. (a) HPLC chromatograms of cholesterol-FLB (chol-FLB) prepared using Poplawski's condition, (b) HPLC chromatograms of chol-FLB, prepared using various ratios of derivatising reagents (mole ratio as cholesterol: FLB), recorded using fluorescence detection (asterisk peak = FLB residue), (c) fluorescence peak area (n = 3) and corresponding percentages by weight of chol-FLB prepared using various molar ratios of reactants. In all cases, chol-FLB was normalised to the molar concentration of cholesterol used in the reaction.

The initial reaction conditions gave incomplete derivatisation of cholesterol indicated by the low yield of chol-FLB (48%) and low fluorescence peak area. The yield and fluorescence peak area increased with increasing proportion of FLB, maximising when the amount of the FLB reached between 2.6 and 10 molar equivalents of cholesterol, the latter reflecting that the derivatisation reached completion. A decrease in the signal response of the FLB derivative and corresponding increase in the response of the residue remaining after purification was observed as the amount of derivatising agent was raised further. This reflects a formation of *O*-acylisourea formed between EDC and carboxylic acid transformed into *N*-acylisourea via intramolecular acyl transfer (Joullie and Lassen, 2010; Tsakos *et al*. 2015). The *N*-acylisourea intermediate could not continue esterification, resulting a decrease in yield of chol-FLB. Hence, for maximum signal response of the steroidal derivative, with minimal FLB residue, the suggested ratio of reagents is 10: 2.5: 20 (FLB: DMAP: EDC to 1 mole of steroid). Repeated preparation of the FLB derivative of cholesterol gave a mean yield by weight of 89 ± 4 % (n =8), reflecting good reproducibility in the yield of chol-FLB under these conditions. The suggested ratio of the derivatising reagents was used to prepare the FLB derivatives of coprostanol and 5α-cholestanol in the same manner.

3.3.1.2 Structural characterisation of cholesterol and cholestanols by mass spectrometry

3.3.1.2.1 FLB derivative of cholesterol (chol-FLB)

The chol-FLB derivative was characterised by APCI mass spectrometry. An aliquot of chol-FLB (1 µL) was directly infused into an ion trap mass spectrometer in order to record its full MS and tandem MS^n spectra. The full mass spectrum of chol-FLB (Figure 3.3) compares well with the result obtained by Poplawski (2017) and confirms the formation of the desired product (structure in inset image of Figure 3.3): the protonated molecule occurs at *m/z* 837.5 (57% RA) and a prominent ion at *m/z* 737.5 (45% RA) corresponds to the loss of the butyloxycarbonyl (Boc) group (-100 Da), as is commonly observed for FLB derivatives (Poplawski 2017). The chol-FLB derivative also forms an ion at *m/z* 781.4 (15% RA) corresponding to neutral loss of the *tert*-butyl (C_4H_9) group (−56 Da) from the protonated molecule, while the ion at *m/z* 369.2 (100% RA)

corresponds to the cholesterol residue after ester-bond cleavage (−468.5 Da). The ion at *m/z* 615.5 (1% RA) corresponds to chol-FLB after loss of the fluorenylmethyloxycarbonyl (FMOC) group (-222.0 Da). The MS spectrum of chol-FLB shows two low intensity ions that were not elucidated by Poplawski (2017): the ion at *m/z* 179.0 (3% RA) corresponds to a cleavage of the 9-fluorenylmethyl group from chol-FLB; the minor ion at *m/z* 854.5 (3% RA) corresponds to a water adduct of chol-FLB with water [chol-FLB+H_2O]$^+$, possibly reflecting incomplete declustering in the ionisation source (*cf* Kolakowski *et al.* 2004; Kostiainen and Kauppila 2009)

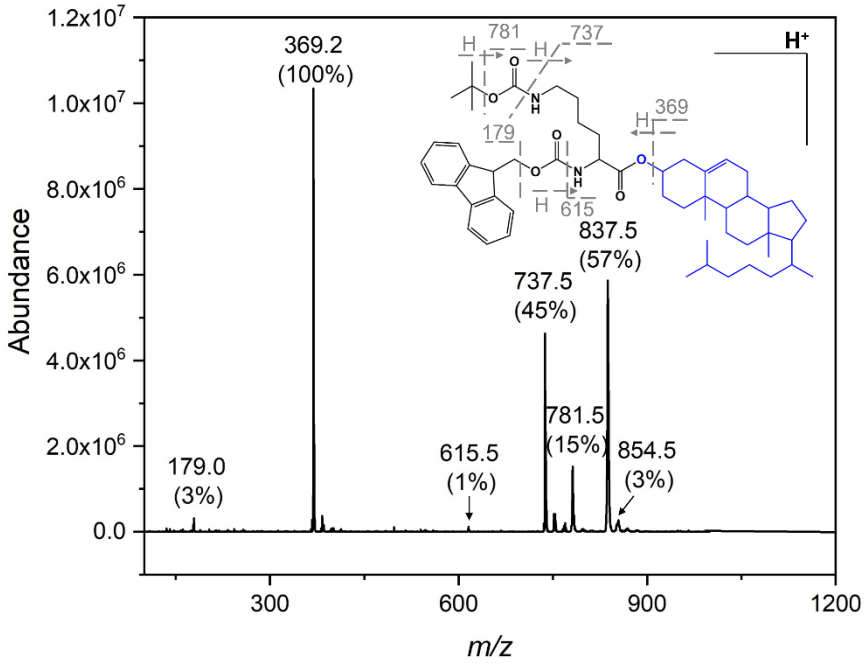


Figure 3.3. Infusion APCI MS spectrum of cholesterol derivatised with FLB (chol-FLB). Concentration of chol-FLB was 12 pmol infused with 0.4 mL min^{-1} and recorded the MS spectrum for 0.6 min. Inset image of chemical structure of chol-FLB indicating the fragmentation.

Notably, the relative abundances of the fragment ions are different than reported by Poplawski (2017). Poplawski (2017) reported a relative abundance of 10.2% for the protonated molecule and abundances for the fragment ions as: *m/z* 737.5, 100%; *m/z* 369.2, 11.5%; *m/z* 781.5, 7.5%. The difference in ion yields in the spectrum most likely reflects differences in the MS experimental setup, possibly causing a difference in ionisation efficiency. The MS spectrum of

chol-FLB recorded by Poplawski was from a direct infusion of chol-FLB in a makeup solvent of methanol: DCM (9:1, *v/v*), while this study used a mixture of methanol: ethyl acetate: 0.1% acetic acid (62: 30: 8, %*v/v*) as a makeup solvent instead in order to generate closely comparable MS spectra to those obtained during UHPLC-MS analysis.

Rationalisation of the dissociation pathways accounting for the fragment ions of chol-FLB confirms that the correct structure was produced. Prominent ions in the full mass spectrum were selected to further investigate their dissociation by tandem mass spectrometry (Figures 3.4). The product ion spectra were generated from five chol-FLB precursor ions as follows: (1) *m/z* 837.5, (2) *m/z* 781.5, (3) *m/z* 737.5, (4) *m/z* 615.5, (5) *m/z* 369.2.

All tandem mass spectra obtained from chol-FLB confirm the formation of the desired product. The most abundant product ions in the three tandem mass spectra shown in Figure 3.4a-c result from cleavages of the amide and ester-bonds of the Boc and FLB moieties (inset image Figure 3.3). The losses of these groups confirm the chol-FLB structure and matched the losses observed in the full mass spectrum of chol-FLB (Figure 3.3).

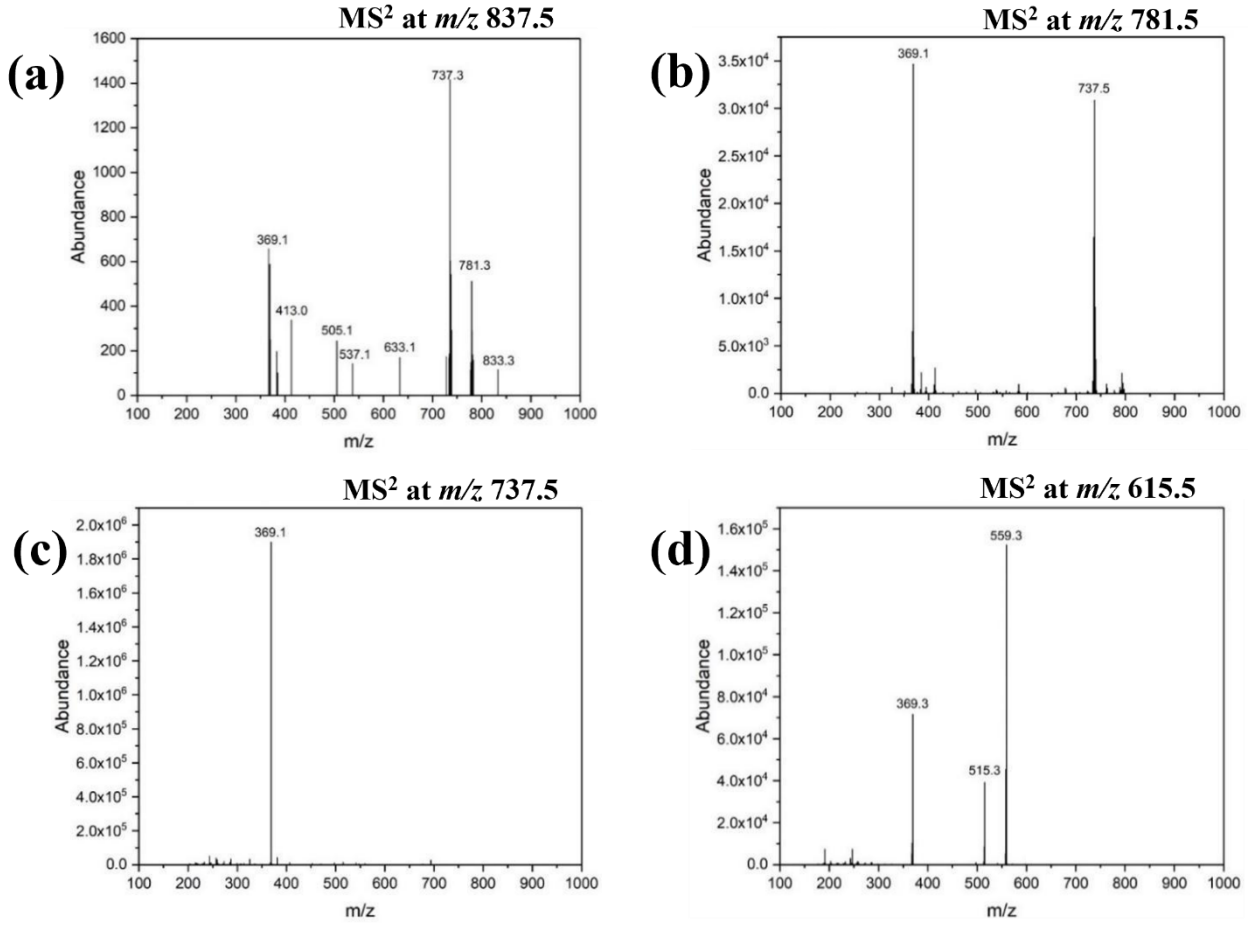


Figure 3.4. APCI MS^2 spectra of chol-FLB from mass-to-charge ratios (a) *m/z* 837.5, (b) *m/z* 781.5, (c) *m/z* 737.5, (d) *m/z* 615.5. Concentration of chol-FLB was 12 pmol infused at 0.4 mL min^{-1} and record the MS spectrum for 0.6 min.

The MS^2 spectrum from precursor ion *m/z* 369.2 shows prominent ions at *m/z* 179 and *m/z* 325 (Figure 3.5a). These ions were assigned by comparison with the MS^3 spectrum of chol-FLB (Figure 3.5b) and the mass spectrum of native cholesterol (Figure 3.6). The MS^3 spectrum of chol-FLB precursor ion *m/z* 737.5 shows that the product ion at *m/z* 369 is formed from the chol-FMOC-Lys derivative after the loss of the Boc group. The prominent product ions occur at *m/z* 179, *m/z* 325, *m/z* 173, *m/z* 130, *m/z* 147, *m/z* 191 and *m/z* 243. These ions mostly match those observed in the MS^2 spectrum of chol-FLB from m/z 369.2 (Figure 3.5a) and match the result obtained by Poplawski (Poplawski 2017). The MS^2 spectrum from precursor ion *m/z* 615.5 (formed by loss of the FMOC group) shows prominent ion at *m/z* 559 and moderate ions at *m/z* 515 and *m/z* 369, corresponding to cleavages of *tert*-butyl (C_4H_9) group (-56 Da), Boc group (-100 Da) and Boc-lysine chain (-246 Da) (Figure 3.4d).

The mass spectrum of native cholesterol contained only one prominent ion at *m/z* 369.3 (Figure 3.6a). The ion arises from the facile loss of water (−18 Da) from the protonated molecule, as has been observed previously for cholesterol (Poplawski 2017). As expected, the MS^2 spectrum differs from that of chol-FLB but is very similar to the MS^2 spectrum of cholesterol reported by Poplawski (2017). Notably, though some ions in the MS^3 spectrum of chol-FLB from *m/z* 737.5 and from *m/z* 369.2 match ions in the MS^2 spectrum of cholesterol (*m/z* 203, *m/z* 215, *m/z* 243 and *m/z* 259) their low abundance suggests that loss to form cholesterol contributes only weakly to the spectrum.

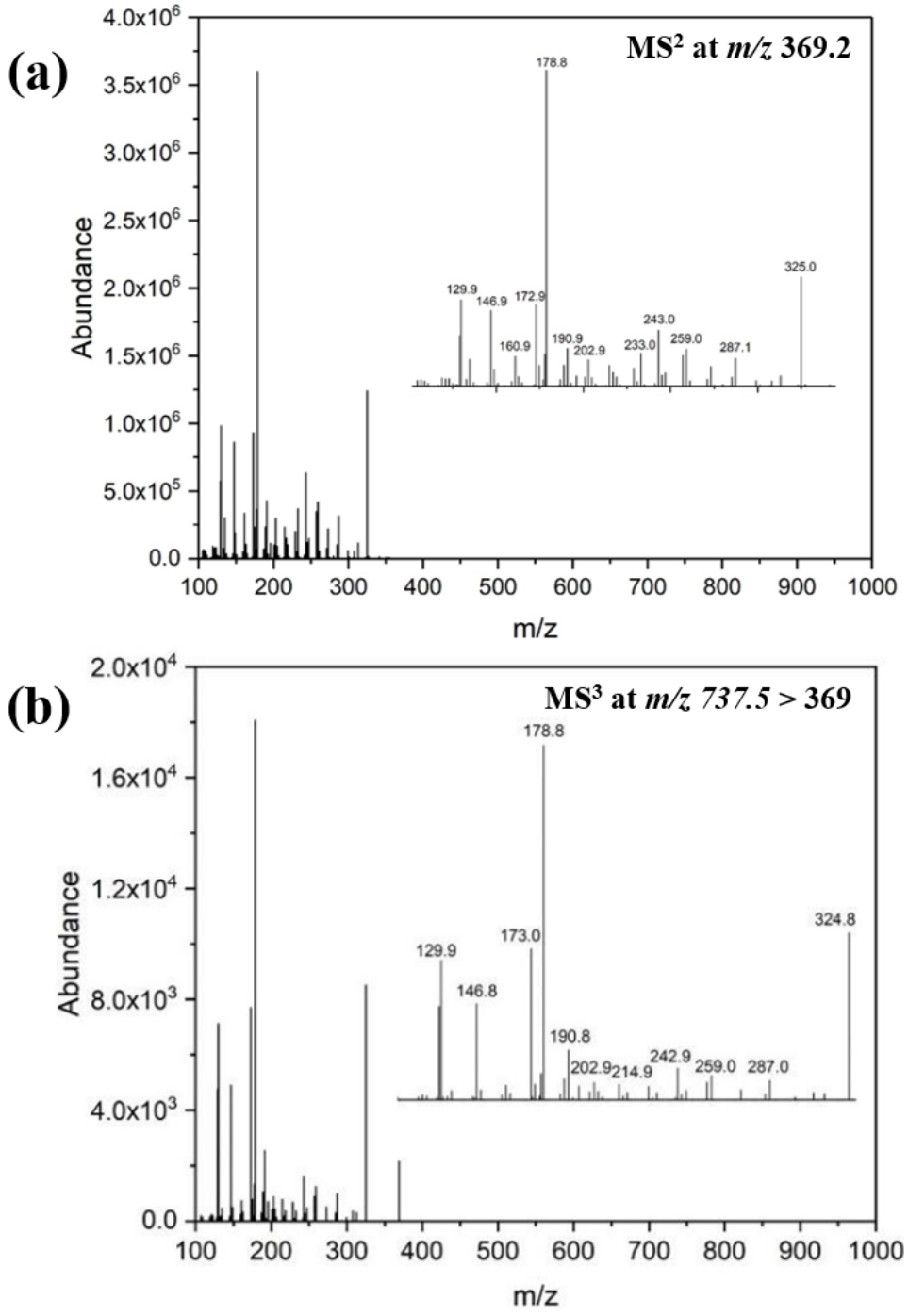


Figure 3.5. Multistage mass spectra of chol-FLB (a) MS^2 spectrum from *m/z* 369.2. (b) MS^3 spectrum from *m/z* 737.5 and *m/z* 369.2. Expanded spectra over the range *m/z* 100-330 are inset. Concentration of chol-FLB was 12 pmol infused with 0.4 mL min^{-1} and record the MS spectrum for 0.6 min.

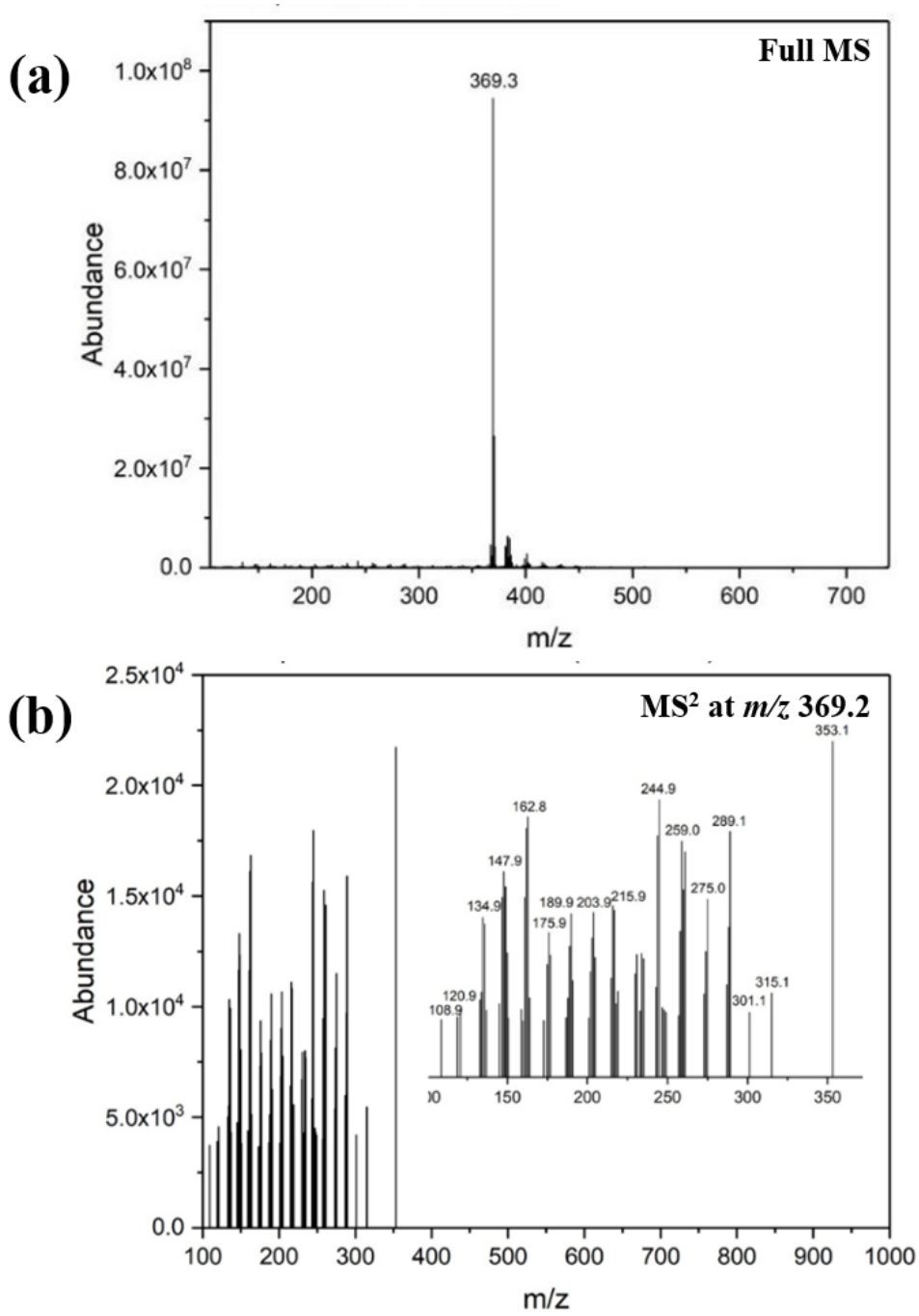


Figure 3.6. Mass spectra of native cholesterol (a) Full mass spectrum. (b) MS^2 spectrum from *m/z* 369.3 with expansion over the range *m/z* 100-350 (inset image). Concentration of cholesterol was 1 nmol infused at 0.4 mL min^{-1} and record the MS spectrum for 0.6 min.

Poplawski (2017) noted the difficulty in interpreting the product ions formed from the chol-FLB and cholesterol precursor ion *m/z* 369, though further investigation of this aspect was beyond the scope of his research. In an attempt to rationalise the dissociation of the chol-FLB precursor ion

at *m/z* 369 the spectrum was compared with the full mass spectrum of the derivatising agent, native FMOC-lysine-Boc (FLB). The full and tandem mass spectra of FLB reveal an interesting finding when compared with the MS^n spectrum from the chol-FLB precursor ion at *m/z* 369. Only a single ion at *m/z* 369.0 was observed in the full spectrum (Figure 3.7a), attributable to a fragment ion formed by facile loss of the Boc group (-100 Da) of FLB (468 Da). The MS^2 spectrum from *m/z* 369.0 shows prominent peaks at *m/z* 130, *m/z* 147, *m/z* 173, *m/z* 179 and *m/z* 191 (Figure 3.7b), possibly indicating loss of the FMOC group linked with lysine at different positions. These ions perfectly match those ions in the MS^2 spectrum of the chol-FLB ion at *m/z* 369 (Figure 3.5a) that do not match with the MS^2 spectrum of native cholesterol (Figure 3.6b). Native cholesterol and FLB both fragments to give an ion at *m/z* 369, hence the presence of this ion can be due either to the presence of FLB or of cholesterol.

All the full MS and tandem MS spectra of chol-FLB confirm the presence of both cholesterol and FMOC-Lysine-Boc in the structure, ensuring a successful preparation of chol-FLB.

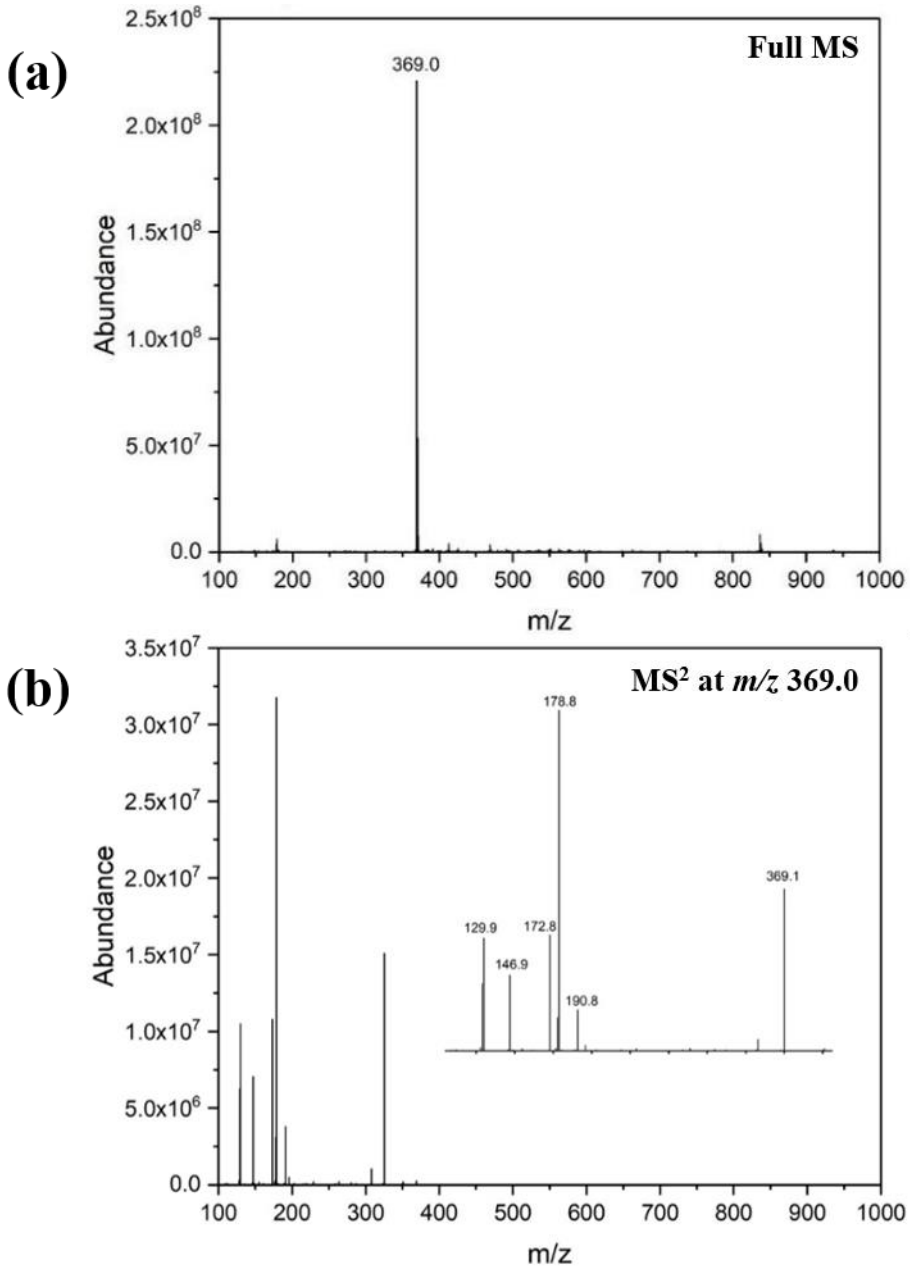


Figure 3.7. Mass spectra of native FMOC-Lys-Boc (FLB). (a) Full mass spectrum. (b) MS^2 spectrum from *m/z* 369.2 with expansion over the range *m/z* 100-350 (inset image). The concentration of FLB was 2 pmol infused at 0.4 mL min^{-1} and record the MS spectrum for 0.6 min.

3.3.1.2.2 FLB derivative of coprostanol (cop-FLB)

The FLB derivative of coprostanol (cop-FLB) was prepared *via* esterification in the same manner as chol-FLB (see Section 3.3.1.1) and the structure characterised by mass spectrometry. The mass

spectrum (Figure 3.8) shows a protonated molecule of moderate intensity at *m/z* 839.7 and a base peak fragment ion at *m/z* 739.7 corresponding to loss of the butyloxycarbonyl (Boc) group (-100 Da). The cop-FLB derivative also forms the minor intensity ion at *m/z* 783.6 corresponding to neutral loss of the tert-butyl (C_4H_9) group (−56 Da) from the protonated molecule, while a moderate intensity ion at *m/z* 369.1 corresponds to the FLB residue after the Boc-bond cleavage (−100 Da). The full mass spectrum of cop-FLB also shows two minor intensity ions: at *m/z* 371.3, corresponding to the coprostanol residue after ester-bond cleavage (−468.5 Da), and at *m/z* 383.3 (-456 Da) presently not assigned. The mass spectrum of cop-FLB shows characteristic fragmentation that compares favourably with that observed for the cholesterol counterpart and is consistent with the structure of cop-FLB (inset image in Figure 3.8), confirming the formation of the desired product.

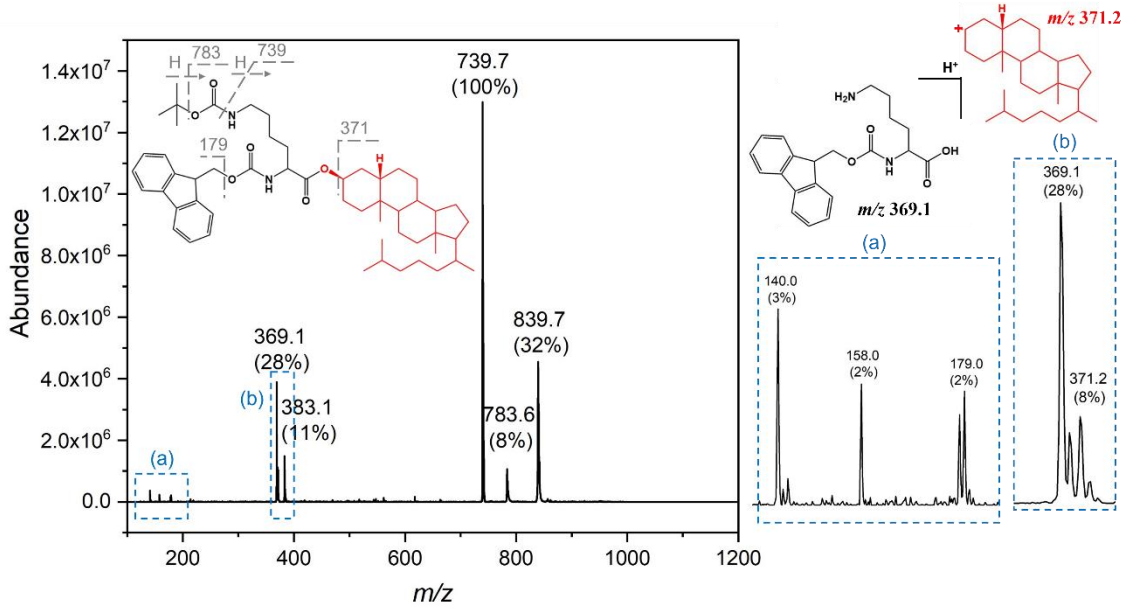


Figure 3.8. APCI MS spectra of coprostanol derivatised with FLB (cop-FLB) over different ranges of mass-to-charge ratio *m/z* 100-1000 with expansion over the range *m/z* 350-400. Concentration of cop-FLB was 12 pmol. Inset image of chemical structure of cop-FLB showing the fragmentation. Expanded full MS spectra of area marked (a) *m/z* 100-190 and (b) *m/z* 360-375 are shown.

The dissociation pathways for cop-FLB were also characterised by tandem-MS (MS^2) (Figure 3.9). The MS^2 spectrum of the protonated ion at *m/z* 839.6 contains a minor intensity product ion at *m/z* 783 and base peak at *m/z* 739, arising from losses of *tert*-butyl group, the entire Boc group, respectively (Figure 3.9a). The MS^2 spectrum of the fragment ion at *m/z* 783.6 also contains an ion at *m/z* 739 as base peak, corresponding to the loss of Boc group from the FLB derivative, along with a minor intensity ion at *m/z* 369 (Figure 3.9b). The MS^2 spectrum of the fragment ion at *m/z* 739.6 produced a base peak ion at *m/z* 369, corresponding to coprostanol loss, and a minor intensity ion at *m/z* 325 resulting from loss of coprostanol + CO_2 via acyl ester cleavage (-415 Da) (Figure 3.9c). The MS^2 spectrum of the fragment ion at *m/z* 369.1 (Figure 3.9e) produced ions at *m/z* 325, *m/z* 191, *m/z* 179, *m/z* 147 and *m/z* 130, matching the fragmentation of the derivatising group as previously shown in the MS^n spectra of chol-FLB and native FLB (Section 3.3.1.2.1). The MS^2 spectrum of the fragment ion at *m/z* 383.1 produced ions at *m/z* 339, *m/z* 187, *m/z* 179, and *m/z* 144, matching the fragmentation of the FLB at different positions (Figure 3.9d).

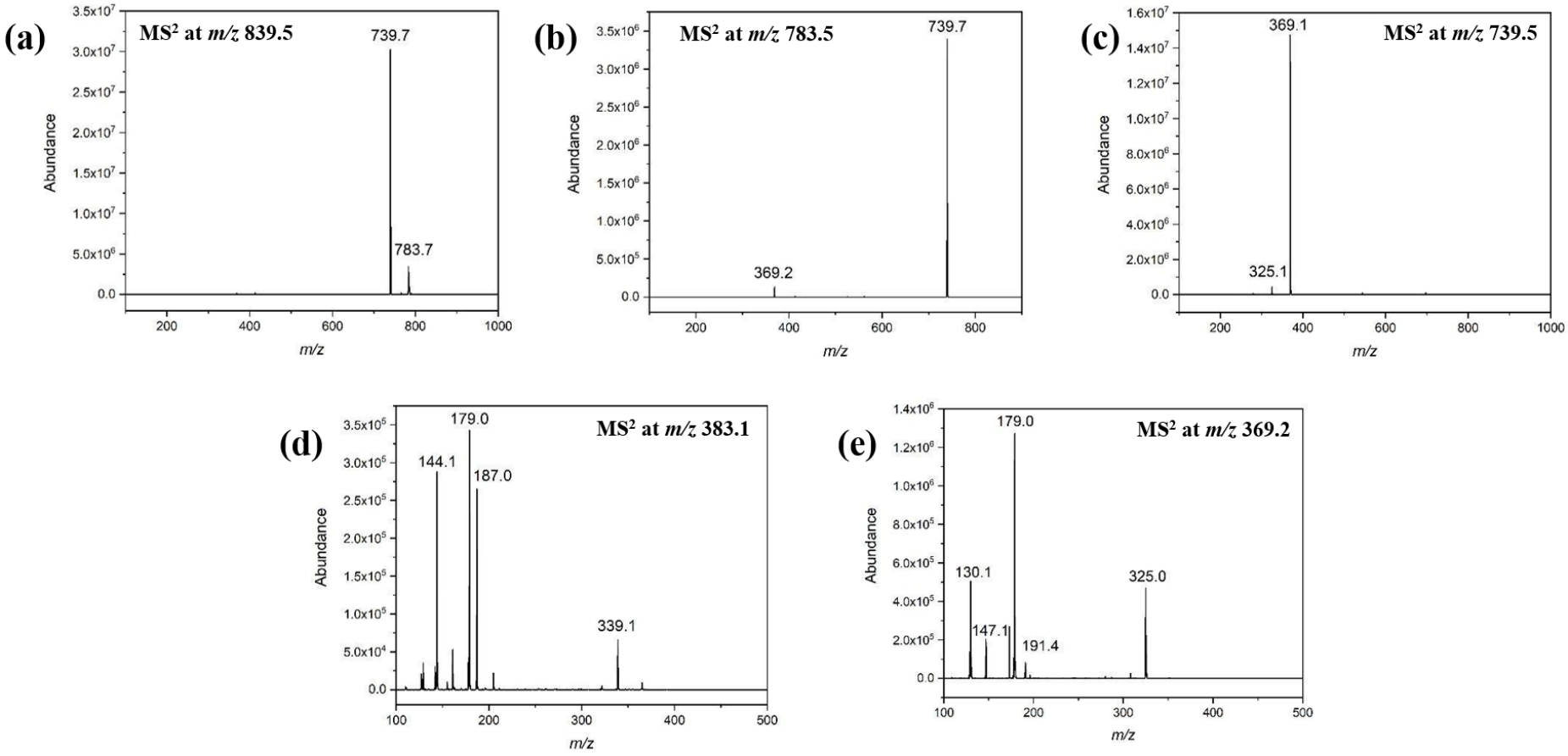


Figure 3.9. APCI MS^2 spectra of cop-FLB from mass-to-charge ratios (a) *m/z* 839.5, (b) *m/z* 783.5, (c) *m/z* 739.5, (d) *m/z* 383.1, (e) *m/z* 369.5. Concentration of cop-FLB was 12 pmol.

3.3.1.2.3 FLB derivative of 5α-cholestanol (5α-chol-FLB)

The mass spectrum of the FLB derivative of 5α-cholestanol (5α-chol-FLB) exhibits a protonated molecule of moderate intensity at *m/z* 839.7, a base peak fragment ion at *m/z* 739.7 and a moderate intensity ion at *m/z* 369.1 (Figure 3.10). The characteristic losses to form the fragment ions confirm the formation of the desired product (inset image in Figure 3.10). The mass spectral fragmentation of 5α-chol-FLB is very similar to that of cop-FLB, though the latter gave rise to a greater number of minor ions, possibly related to the difference in stereochemistry. The dissociation of 5α-chol-FLB was further characterised by tandem-MS (MS^2) (Figure 3.11). The tandem MS spectra were very similar to those of the 5β epimer (coprostanol) FLB derivative, confirming the same origins of product ions in both the 5α and 5β forms.

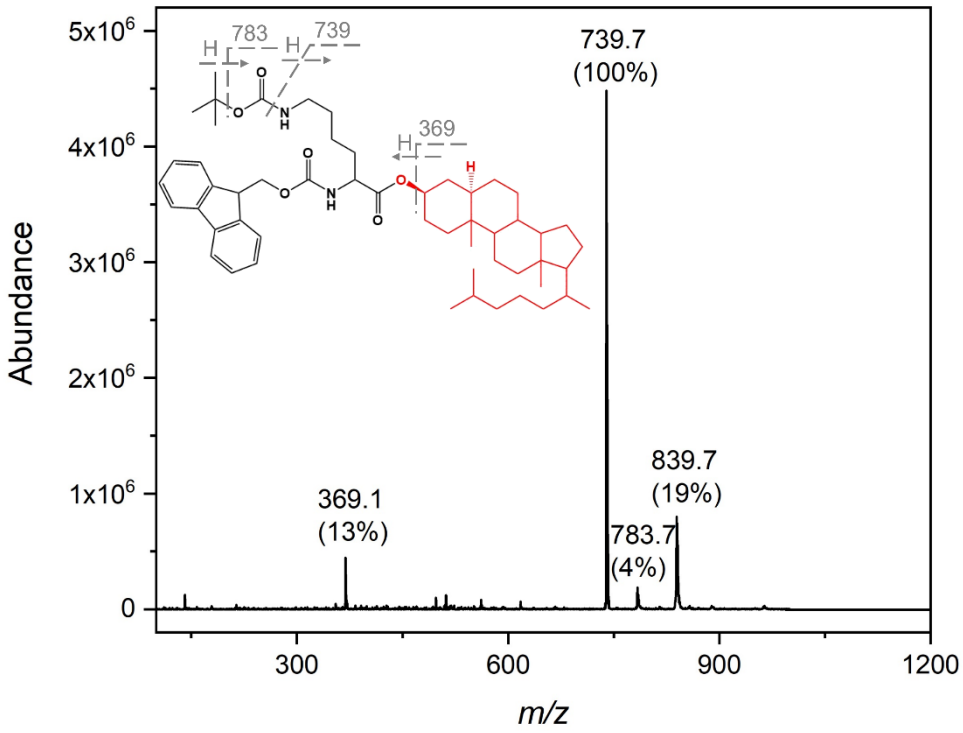


Figure 3.10. APCI MS spectrum of 5α-cholestanol-FLB. Concentration of 5α-chol-FLB was 13 pmol. Chemical structure of 5α-chol-FLB showing fragmentation is inset.

The tandem-MS (MS^2) spectra of 5α-chol-FLB reveal the same dissociation pathways as cop-FLB owing to it being the epimer form, confirming the formation of the correct structure (Figure

3.11). The MS^2 spectrum of the protonated ion at *m/z* 839.6 produces a minor intensity product ion at *m/z* 783, a base peak product ion at *m/z* 739, corresponding to losses of the *tert*-butyl group and the entire Boc group, respectively (Figure 3.11a). The MS^2 spectrum of the fragment ion at *m/z* 783.6 produce a base peak produced ion at *m/z* 739, arising from the loss of Boc group from the FLB derivative (Figure 3.11b). The MS^2 spectrum of the fragment ion at *m/z* 739.6 produced a base peak ion at *m/z* 369 and a minor intensity ion at *m/z* 325, corresponding to the losses of 5α-cholestanol at the different positions on its structure (Figure 3.11c). The MS^2 spectrum of the fragment ion at *m/z* 369.1 produced ions at *m/z* 191, *m/z* 179, *m/z* 147 and *m/z* 130, matching the cleavage the FMOC at different positions, while the ions at *m/z* 351, *m/z* 325 correspond to the losses from different positions of the ester group of FLB: water (-18 Da) and carbon dioxide (-44 Da), respectively (Figure 3.11d).

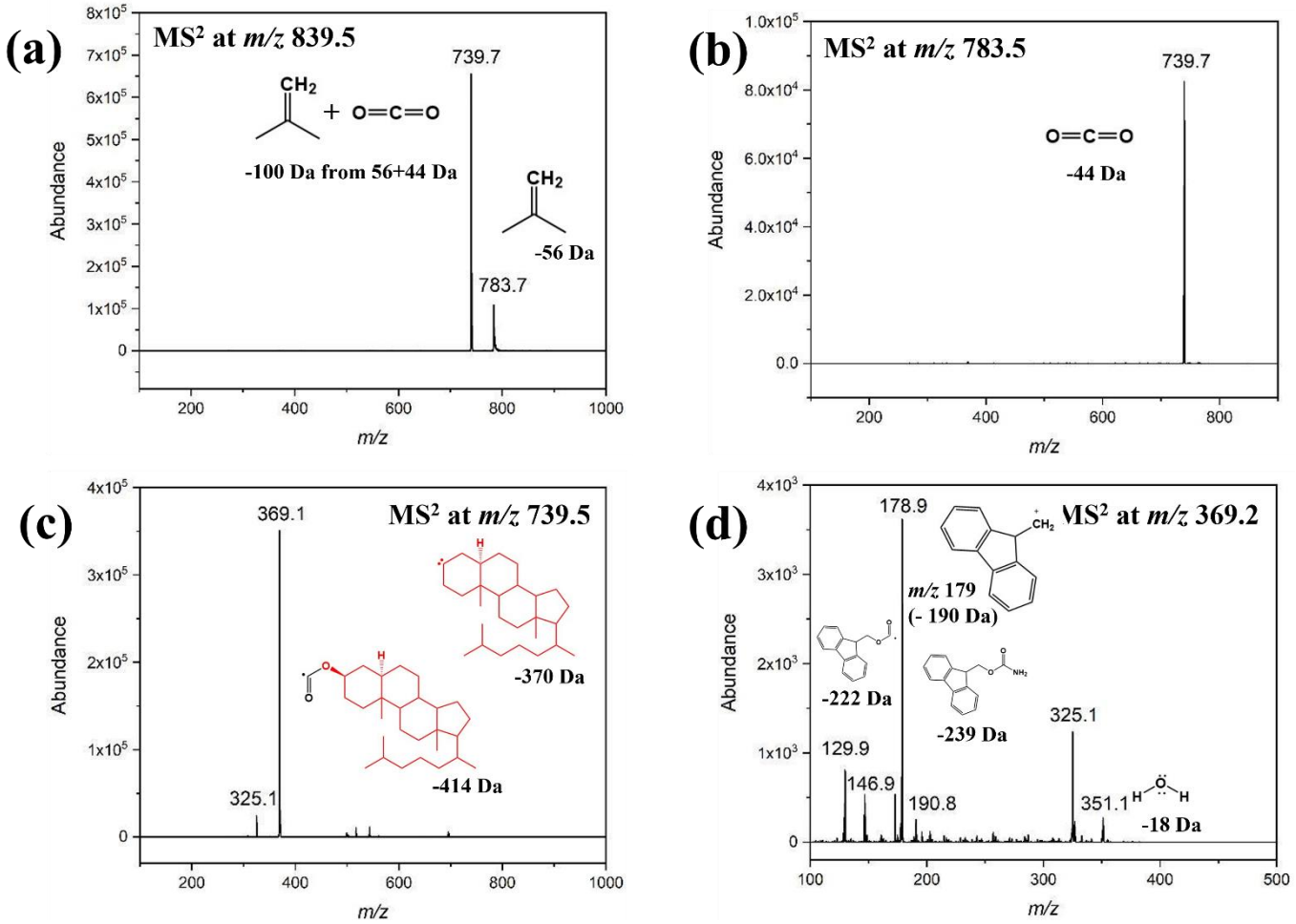


Figure 3.11. APCI MS^2 spectra of 5α-chol-FLB from mass-to-charge ratios (a) *m/z* 839.5, (b) *m/z* 783.5, (c) *m/z* 739.5, (d) *m/z* 369.5. The concentration of 5α-chol-FLB was 13 pmol. Chemical structures of cleavages of molecules from 5α-chol-FLB are inset.

3.3.2 Development of the analytical method for separation of FLB derivatives of C_{27} steroidal lipids

3.3.2.1 Separation of FLB derivatives of C_{27} steroidal lipids (cholesterol, coprostanol and 5α-cholestanol) by (U)HPLC equipped with various detectors

The FLB derivatives of C_{27} steroidal lipids (cholesterol, coprostanol and 5α-cholestanol) exhibit MS ions that relate to the presence of the Boc and FMOC groups. At equal concentration, the MS signal response of chol-FLB was higher than its native form by about 10 times (Poplawski 2017), indicating the potential to use FLB for enhancing MS response of cholesterol and thus offering a better limit of detection. The additional ultra-violet and fluorescence detection capabilities made possible by the presence of the Fmoc group provides the opportunity to detect FLB derivatives by MS, UV and FL detectors. Thus, the detection of steroidal FLB derivatives by MS, UV and FL detectors offers a greater security of quantification. In the case that a UHPLC-MS-UV-FL system is available, the steroids can be separated and identified by MS detector while quantification can rely on the UV or FL detectors, the UV/FL selectivity providing specificity in the chromatograms. This section reports the development of a separation method for C_{27} steroidal lipids (cholesterol, coprostanol, 5α-cholestanol) by UHPLC equipped with UV-MS detectors, while, owing to the lack of UHPLC-FL facilities, an HPLC-FL system was used to demonstrate fluorescence detection of these.

3.3.2.1.1 MS detection

Poplawski reported a UHPLC separation for chol-FLB with MS and UV detection, though other C_{27} steroidal lipids (coprostanol and 5α-cholestanol) were not included (Poplawski 2017). A mixture of the FLB derivatives of cholesterol and its reduced counterparts (coprostanol and 5α-cholestanol) was analysed by UHPLC-MS (Figure 3.13) in the same manner as Poplawski (2017). Each component was examined separately by LC-MS analysis to determine retention time and peak area in order to confirm identification in a mixture of the components (Figure 3.12). The steroidal derivatives eluted at slightly different retention times; 18.3 min for cop-FLB, 18.4 min

for chol-FLB and 19.2 min for 5α-chol-FLB. Assuming that a mixture of these sterols would exhibit the same elution behaviour, the base peak chromatograms would show partial co-elution of cop-FLB and chol-FLB in the retention time range 18.2 – 18.5 min, with complete separation of 5α-chol-FLB. The UHPLC-MS chromatogram of a mixture of the coprostanol, cholesterol and 5α-cholestanol FLB derivatives showed only two peaks, evidencing that the method that Poplawski used is not capable of resolving all three C_{27} steroidal lipids (Figure 3.13).

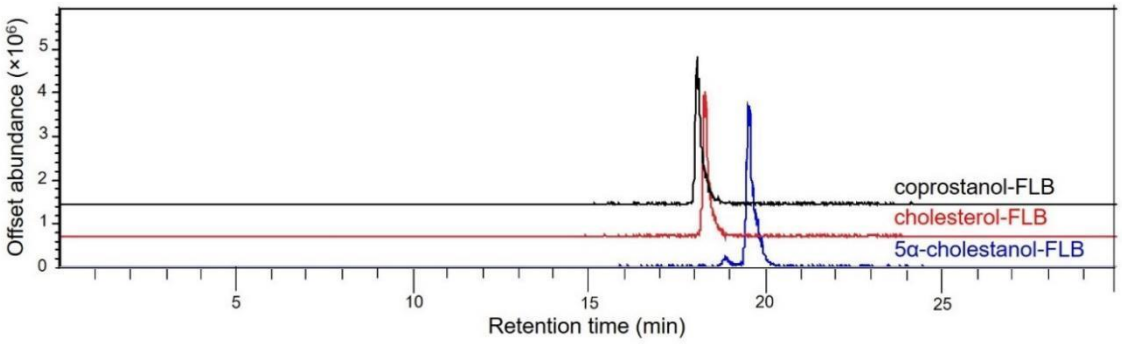


Figure 3.12. APCI base peak chromatograms of steroidal-FLB derivatives (chol-FLB, cop-FLB and 5α-chol-FLB) obtained during separate analyses. Concentrations were 9 pmol (chol-FLB), 12 pmol (cop-FLB) and 13 pmol (5α-chol-FLB). The separation condition is shown in Section A.1.1.

The base peak chromatogram of the mixture showed the presence of two prominent peaks at retention times of 14.6 min and 15.1 min (Figure 3.13a). Notably, the earlier elution of the steroidal-FLB derivatives than in the previous analysis of the individual standards is attributed to the built up of back pressure in the HPLC system. The pressure increased in the HPLC system not only effected the earlier elution of steroidal mixture (Figure 3.13) compared to those from individual standards (Figure 3.12), but also resulted in differences in peak areas of both chromatograms. Assuming that the sterols would exhibit the same elution behaviour, the order of elution -would be cop-FLB (peak 1), chol-FLB (peak 2) and 5α-chol-FLB (peak 3). The peak at 15.1 min (peak 3), which should be 5α-chol-FLB, has a peak area 60% greater than of that of the individual standard (Figure 3.13b). The peak at 14.6 min (peak 1+2), which is expected to be

a co-elution of cop-FLB with chol-FLB, has 50% less peak area than the sum of the standards during their separate analysis (Figure 3.13b). The results possibly reflect differences in ionization efficiency.

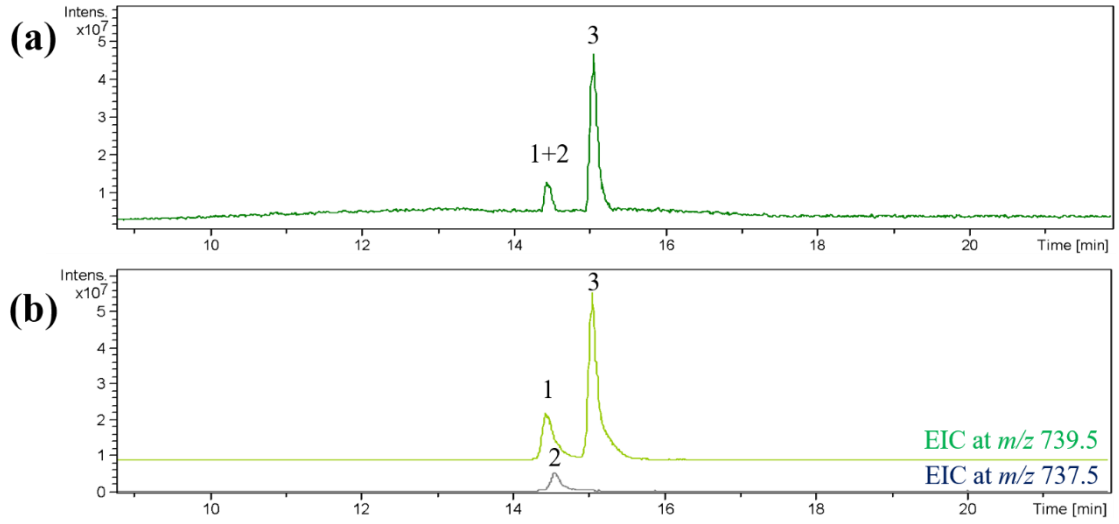


Figure 3.13. Reversed phase LC-MS chromatograms of steroid mixture (chol-FLB, cop-FLB and 5α-chol-FLB) (a) base peak chromatogram; peak 1; 14.6 min and peak 2; 15.1 min, (b) extracted ion chromatograms obtained at *m/z* 737.5 and *m/z* 739.5. The separation condition is shown in Section A.1.1. Concentrations of the individual components of the mixture on column were chol-FLB (9 pmol), cop-FLB (12 pmol) and 5α-chol-FLB (13 pmol).

Extracted ion chromatograms of the mixture at *m/z* 737.7 and *m/z* 739.7 revealed co-elution of chol-FLB and one of the stanol-FLB derivatives at 14.6 min (Figure 3.13b), the stanol being assigned as cop-FLB on the basis of the earlier retention time (Figure 3.12). The mass spectrum of each peak confirmed the co-eluting species in peaks 1+2. The mass spectrum (Figure 3.14a-b) indicates co-elution of coprostanol with cholesterol, evidenced by the presence of fragment ions at *m/z* 615.5, *m/z* 617.7, *m/z* 737.7, *m/z* 739.7, *m/z* 781.5, *m/z* 783.6, *m/z* 837.7 and *m/z* 839.7 which matched the ions obtained from the FLB derivatives of cholesterol and its stanols. The mass spectra of cop-FLB and chol-FLB were isolated by integration with a peak width of 0.1 min

(Figure 3.14c-d). The mass spectrum obtained from the peak at retention time of 15.1 min matches that of 5α-chol-FLB (Figure 3.14e). There are significant differences in the mass spectra of the steroidal mixture on column to those obtained previously from direct infusion (Figure 3.3, 3.8, 3.10), the relative abundances (%RA) of the ions do not match, for example the presence of the moderate intensity ions at *m/z* 615.5 and *m/z* 617.5 correspond to the loss of the fluorenylmethyloxycarbonyl (FMOC) group (-222 Da) were 15-60 %RA for these components as a mixture while the signal response of each individual component was about 1 %RA. The differences suggest that is not possible to use %RA of the ions obtained via infusion for the purpose of identifying from the analytes on column, this could possibly be due to changes in ionisation efficiency arising from the mobile phase gradient during separation.

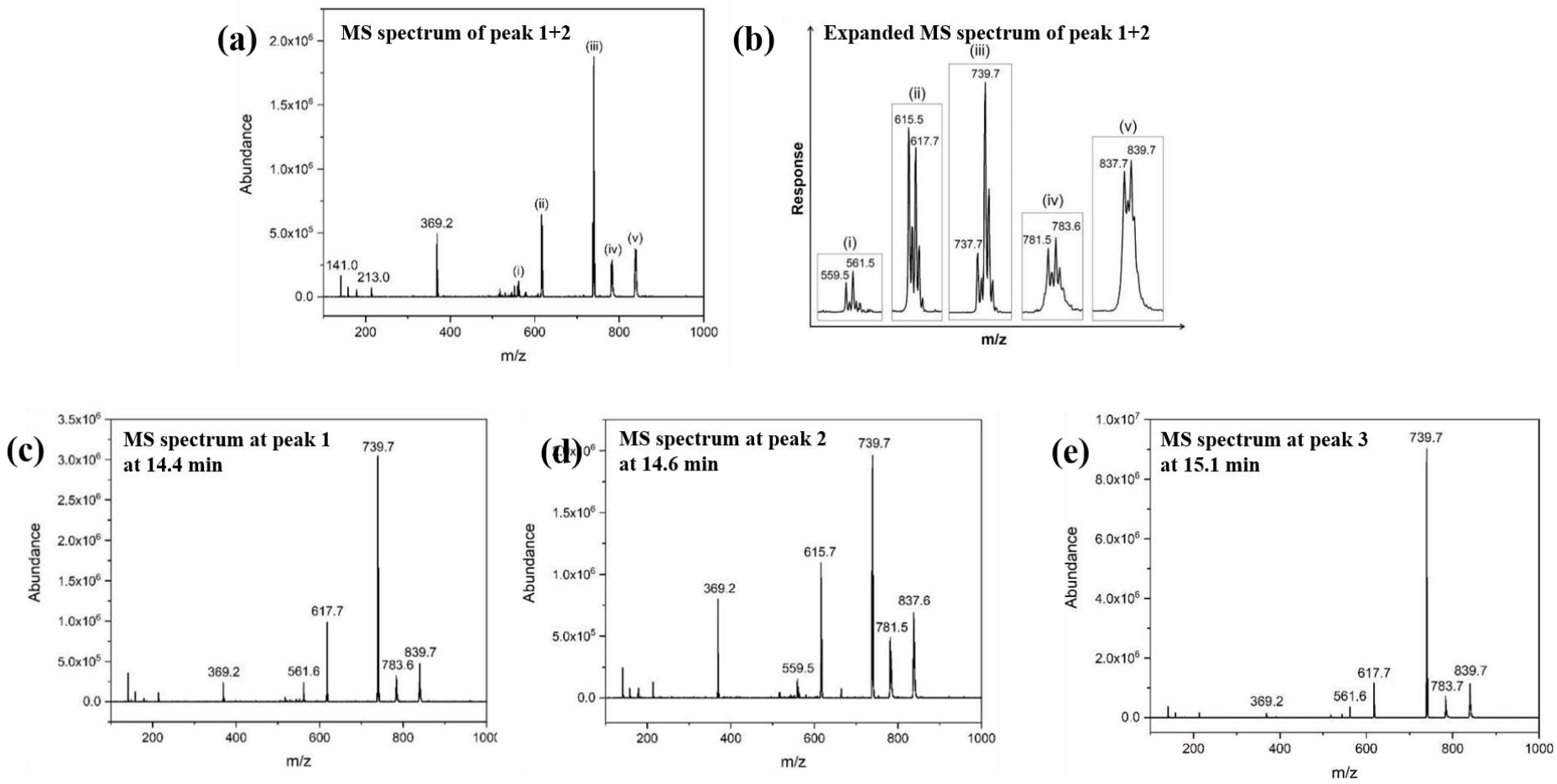


Figure 3.14. APCI MS spectra of steroidal FLB mixture (a) at 14.6 min; (b) expanded MS spectrum at 14.6 min showing ion clusters labelled i - v. (c) 14.4 min; (d) 14.6 min; (e) 15.1 min. Steroidal lipids mixture comprised chol-FLB (9 pmol), cop-FLB (12 pmol) and 5α-chol-FLB (13 pmol).

3.3.2.1.2 UV and FL detection

A mixture of FLB derivatives of cholesterol, 5α-cholestanol and 5β-cholestanol (coprostanol) was detected by UV-vis and FL during LC analysis (Figure 3.15). In order to generate comparable results, the LC conditions used were the same as those used for MS detection (Figure 3.12-3.13). The components were assigned based on the LC-MS analysis, the co-elution of cop-FLB and chol-FLB being expected. Partial separation was observed in the UV-vis chromatogram and complete co-elution in that obtained with the FL detector. Separation of 5α-chol-FLB was observed using both detectors.

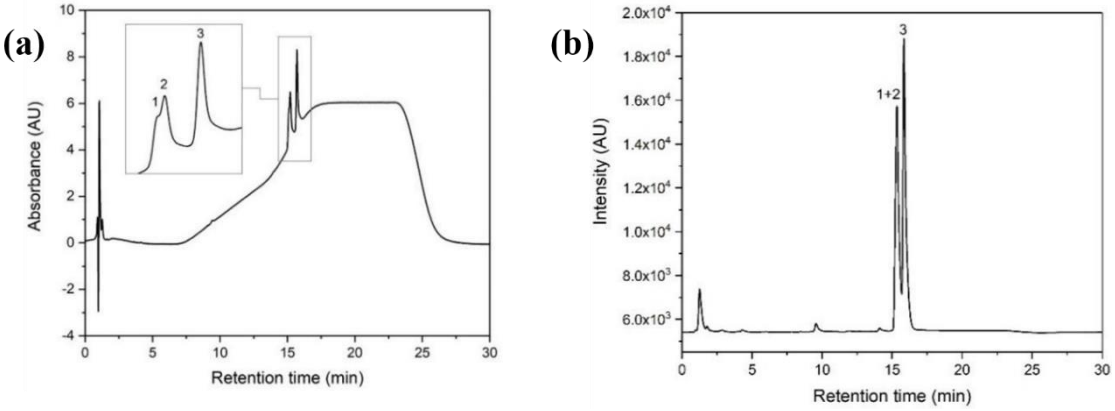


Figure 3.15. Chromatograms of C_{27} steroidal lipids mixture recorded using different detectors (a) ultra-violet, (b) fluorescence. Assigned peaks: 1 = coprostanol-FLB, 2 = cholesterol-FLB and 3 = 5α-cholestanol-FLB. The separation condition is shown in Section A.1.1. Steroidal lipids mixture was chol-FLB (9 pmol), cop-FLB (12 pmol) and 5α-chol-FLB (13 pmol).

The signal response in the chromatograms was highest for the FLB derivative of 5α-cholestanol. Assuming a signal response of one for 5α-chol-FLB, the relative responses of the signals of cop-FLB and 5α-chol-FLB were expressed as ratios (Table 3.2). The ratio of chol-FLB: cop-FLB: 5α-chol-FLB obtained can possibly be used as a basis for quantification of sterols when only partial separation occurs. The differences in signal of the steroidal mixture with the various detectors possibly reflects differences in physical properties of cholesterol and cholestanol. For example, cop-FLB has a better ionisation efficiency than chol-FLB, resulting a higher ratio of

cop-FLB: chol-FLB for MS detection. However, the cop-FLB appears to have a lower absorption coefficient than chol-FLB, reflecting the lower ratio of cop-FLB: chol-FLB for UV detection. The steroidal FLB derivatives have the greatest signal response with the FL detector, reflecting its potential for use in quantification even at very low analyte concentrations.

Table 3.2. Ratio of each peak of sterol-FLB responding to different detectors (n = 3).

Sterol-FLB	Ratio of peak area for detector selected		
	MS[1]	**UV[2]**	**FL[1]**
cop-FLB	0.21 ± 0.01 (5.6%)	0.46 ± 0.02 (4.3%)	0.90 ± 0.01 (0.1%)
chol-FLB	0.11 ± 0.01 (6.8%)	0.62 ± 0.01 (1.5%)	
5α-chol-FLB	1.00 ± 0.00 (N.D.)	1.00 ± 0.00 (N.D.)	1.00 ± 0.00 (N.D.)

[1]peak area, [2]peak height, N.D. (not determined).

3.3.2.2 Preliminary screening of FLB derivatives of C_{27} steroidal lipids (cholesterol, coprostanol and 5α-cholestanol) in a Roman grave soil from Hungate, York

Assessment of Fmoc-Lys-Boc (FLB) as a derivatising reagents for measurement of C_{27} steroidal lipids (cholesterol, coprostanol and 5α-cholestanol) in archaeological samples was performed by UHPLC-MS. Cholesterol and cholestanols were determined in a Roman age grave soil sample from Hungate, York (skeleton number SK51350), the sample was obtained from grave fills underneath the pelvis base (Pickering *et al.* 2018). The FLB derivatives of samples were prepared using derivatising agents in ratio of 50:75:10 of FLB:EDC:DMAP to 1 mole (estimated) of soil extract as recommended by Poplawski (2017). The products obtained from soil extracts after FLB derivatisation and clean-up by hexane: ethyl acetate (1:1, *v/v*) were separated by UHPLC with MS to obtain the base peak chromatogram of the FLB derivatives (Figure 3.16). The FLB derivatives of cholesterol, coprostanol and 5α-cholestanol were identified based on the fragmentation patterns from both full MS and MS^2 spectra obtained from Section 3.3.1.2.

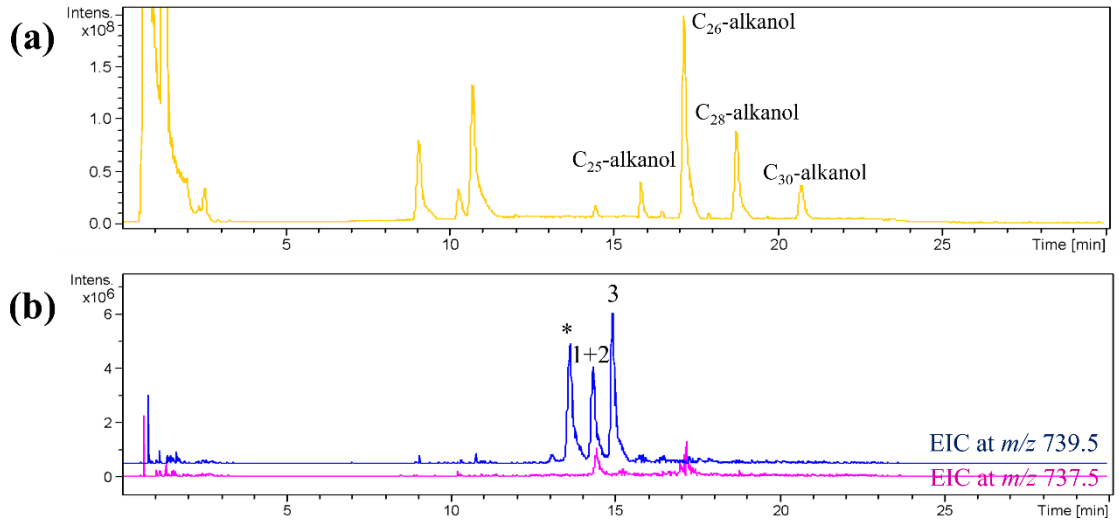


Figure 3.16. APCI MS chromatograms of the medium polar fraction of the grave soil extract from SK51350 (pelvis) (a) base peak chromatogram, (b) extracted ion chromatograms obtained from *m/z* 737.5 and *m/z* 739.5. Assigned peaks: 1 = cop-FLB, 2 = chol-FLB, 3 = 5α-chol-FLB, * = epicoprostanol-FLB (epicop-FLB). The separation condition is shown in Section A.1.1.

Successful detection of cholesterol, coprostanol and 5α-cholestanol at retention times of 14.3, 14.5 and 15.1 min was evident from the extracted ion chromatograms (*m/z* 737.7, *m/z* 739.7) (Figure 3.16). The results show a close match to the standard mixture (Figure 3.14). A prominent peak eluting at retention time 13.5 min possibly reflects the presence of another transformation product of C_{27}-sterol, epicoprostanol (asterisk peak in Figure 3.16b). The presence of cholesterol and cholestanols in the soil extract reflects contributions from human activity based on the samples being grave fills, confirming the potential to use these samples for evaluation of the suitability of FLB as derivatising agent for the detection of steroidal alcohols in soils and sediments. However, the APCI MS chromatogram of the soil extract shows prominent peaks eluting in the first 5-min (Figure 3.17, black trace), reflecting the presence of a large amount of residue from the soil extract. The signal responses of residues from the derivatisation process were still observed from the chromatograms of the following blank injections (data not shown); reflecting carry-over of residues and retention of material in the system that could lead to column blockage and shortening column lifetime.

3.3.2.3 Modification of the analytical method for FLB derivatives in sediment extracts

The chromatograms obtained both from standard sterols and from soil extracts separated by the conditions that Poplawski reported (Poplawski 2017) show incomplete resolution of cop-FLB and chol-FLB, showing limitation of the HPLC separation. Thus, the analytical method for the FLB derivatives of cholesterol and its reduction products was examined to evaluate if improvement in separation could be made by the following strategies: (1) altering the sample clean-up step for sediment extracts (see Section 3.3.2.3.1), (2) use of an alternative HPLC column (see Section 3.3.2.3.2), (3) gradient programming of the HPLC mobile phase (see Section 3.3.2.3.3).

3.3.2.3.1 Altering the sample clean-up step for sediment extracts

The FMOC-Lys-Boc (FLB) was first demonstrated as a derivatising agent for screening of cholesterol in grave soils (Poplawski, 2017). In the original method, the products from FLB derivatisation were collected from a crude reaction mixture using column chromatography and a mixture of hexane: ethyl acetate (1:1, %*v/v*) as eluent. In this study, the FLB derivatives of soil extracts (number 51150 from Hungate, York) were prepared in the same manner as the original method. The chromatograms showed dominant signal responses of residues from the derivatisation process in addition to the FLB derivatives, indicating that the sample preparation steps were not optimal. Considering that polar components elute prior to apolar components on a reversed phase column, the residues are considerably more polar than the C_{27} steroidal lipids. It was assumed that a decrease in polarity of the eluent used in the clean-up step by column chromatography would improve the isolation of the steroidal lipid products from a crude reaction mixture after the FLB derivatisation reaction, minimising residues of left-over reagent. The sample clean-up step for the FLB derivatives of soil extracts was examined. An organic extract from the soil of a Roman cesspit (number 51151 from Hungate, York) was derivatised with FLB and then divided into two equal portions. One of the portions was cleaned-up using eluent (hexane-ethyl acetate, 1:1 *v/v*; solvent polarity (*P'*) of 2.25; *P'* based on Burdick & Jackson Solvent Guide (Snyder *et al.* 1997)) and the other using a less polar eluent (hexane-ethyl acetate,

3:1 *v/ v; P'* of 1.18). The isolated FLB derivatised extracts were separated by (U)HPLC with FL, UV and MS detection to obtain signal responses of the product (Figure 3.17).

The modified clean-up method using hexane: ethyl acetate, 3:1 *v/v* showed almost complete removal of the early eluting peaks, the small amounts of residue only being evident in the FL chromatogram (Figure 3.17a). The response for the steroidal FLB derivatives on all three detectors (FL, UV and MS) was similar to that observed in the chromatograms from the original method, reflecting no loss of C_{27}-sterols. Almost complete removal of residue while maintaining signal response for the analytes clearly indicates the suitability of the modified sample clean-up step. Hence, for maximum signal response of the steroidal derivative with minimal FLB residue, the suggested eluent of residue removal from all derivatised reaction products is hexane: ethyl acetate, 3:1 *v/v*.

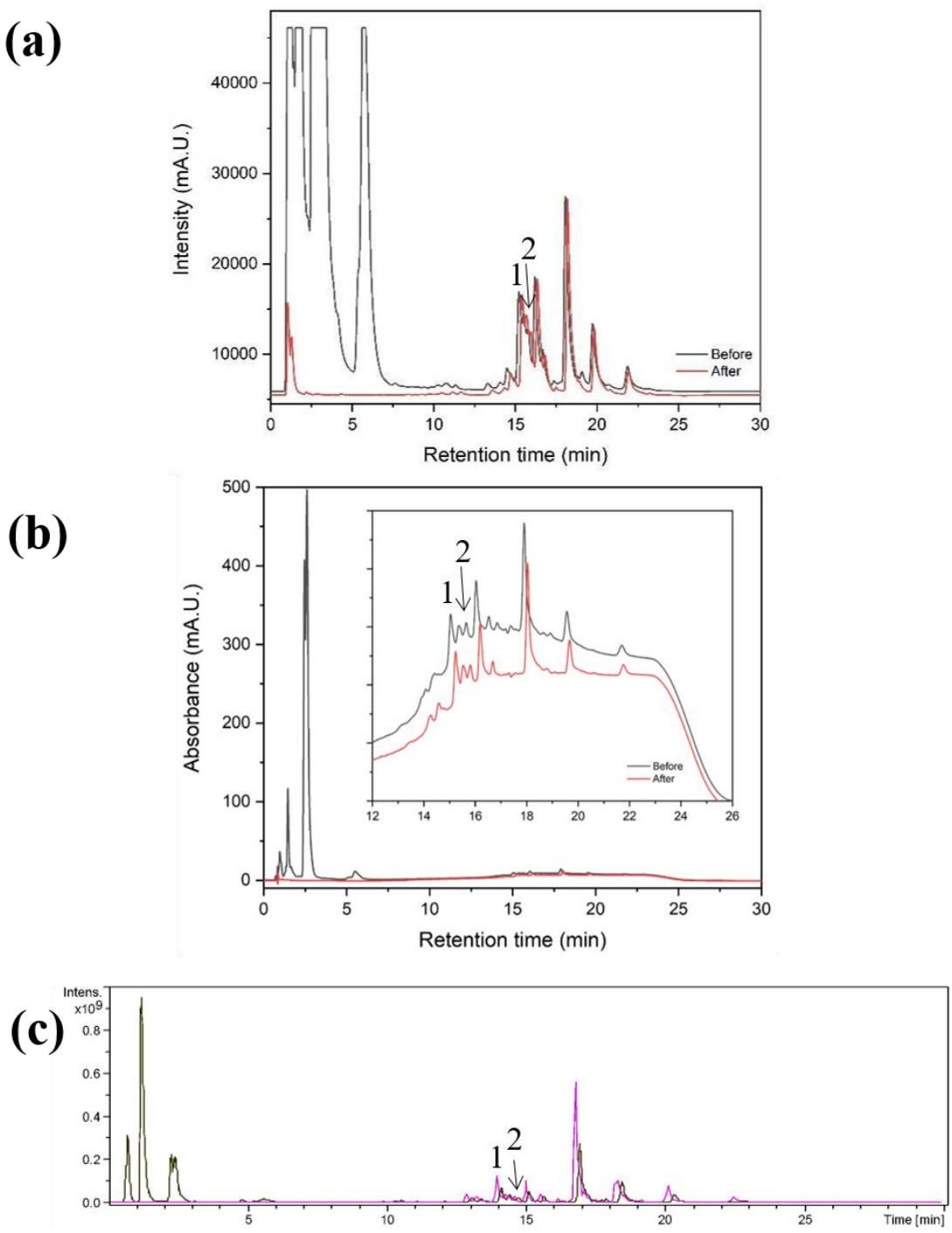


Figure 3.17. Chromatograms of the FLB derivatives of the medium polar fraction of the soil extracts (Hungate 51151, York) cleaned-up by different eluents recording using different detectors: (a) fluorescence, (b) ultra-violet, (c) APCI MS. Before (black trace) eluted using hexane: ethyl acetate (1:1 *v/v*), After (coloured trace) eluted using hexane: ethyl acetate (3:1 *v/v*). Assigned peaks: 1 = coprostanol, 2 = 5α-cholestanol. The separation condition is shown in Section A.1.1.

3.3.2.3.2 Evaluation of various HPLC columns

Improvement to the (U)HPLC resolution of the FLB derivatives of cholesterol and coprostanol of the method developed by Poplawski (2017) was attempted by using various HPLC columns (Figure 3.18). Separation of the mixture of FLB derivatives of cholesterol, coprostanol and 5α-cholestanol was evaluated by comparing peak shapes of chol-FLB and cop-FLB obtained from each column. The LC conditions used (see Section A.1.1) were the same as those used for the Dionex acclaim C_{18} UHPLC column in order to generate comparable results, while the UV and FL detectors were used because of the greater limitation of these detectors for deconvolution of the co-elution components.

The separation of the mixture of chol-FLB, cop-FLB and 5α-chol-FLB using the Dionex acclaim C_{18} HPLC column, designated as column A (Figure 3.18, black trace) only resolves two peaks at retention times between 15-17 min as discussed previously (see Figure 3.15).

Three different HPLC columns (Columns B-C, Figure 3.18) were examined using the same mobile phase composition to provide comparability. The HPLC chromatograms of the FLB-steroidal mixture obtained from using the Acquity BEH C_{18} column, 'column B', and from the Luna C_{18} column, 'column C' gave two dominant peaks, evidencing coelution of cop-FLB and chol-FLB and complete separation of 5α-chol-FLB. The HPLC chromatograms obtained using the C_{18} Waters Symmetry column, 'column D', gave two coeluting peaks at 15.1 min and 15.8 min (Figure 3.18, green line). Partial resolution of chol-FLB and cop-FLB was evident in the earlier eluting peak, reflecting the potential to distinguish the FLB forms of cholesterol and coprostanol using this column. Partial separation was later confirmed from the elution times of each individual sterol (data not shown). Thus, column D (Waters Symmetry) was selected for further attempts to improve resolution of chol-FLB and cop-FLB by variation of the mobile phase.

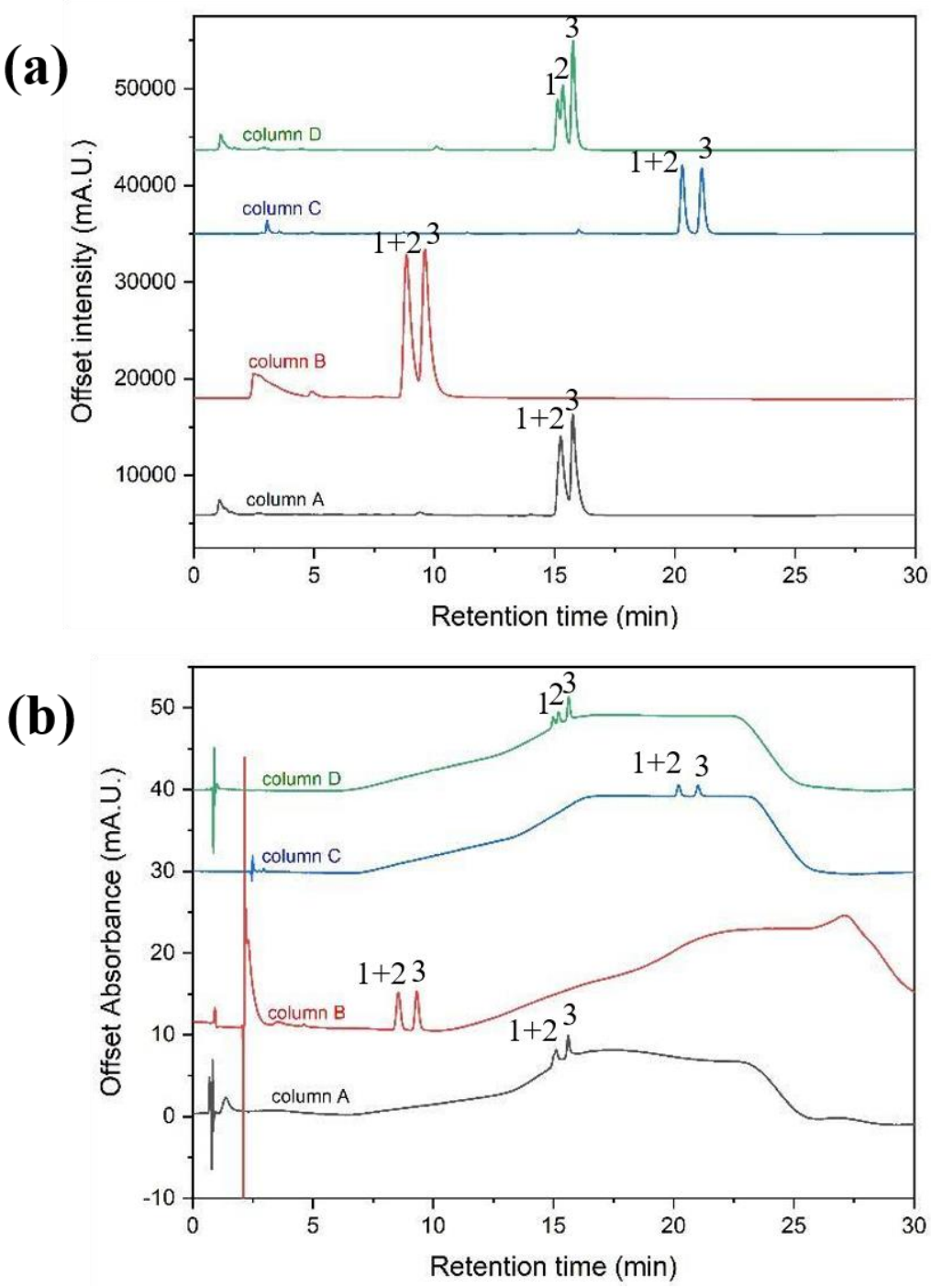


Figure 3.18. Chromatograms of the FLB derivatives of steroidal mixture recording using different detectors: (a) fluorescence and (b) ultra-violet. Column A = Dionex Acclaim C_{18} (2.2 µm, 2.1 mm×100 mm), column B = Acquity BEH C_{18} (1.7 µm, 2.1 mm×150 mm), column C = Luna C_{18} (3.0 µm, 4.6 mm×150 mm), column D = Symmetry C_{18} (3.5 µm, 2.1 mm×150 mm). Sterol mixture was chol-FLB (9 pmol), cop-FLB (12 pmol) and 5α-chol-FLB (13 pmol). Assigned peaks: 1 = coprostanol, 2 = cholesterol, 3 = 5α-cholestanol. The separation conditions are given in Section A.1.1.

3.3.2.3.3 Modification of the HPLC mobile phase gradient

The resolution of chol-FLB and cop-FLB was determined on decreasing the polarity of the mobile phase composition to allow the two components to elute over a narrower polarity range. Over the retention time window where the cop-FLB and chol-FLB eluted (15-17 min), the mobile phase composition of the original gradient (Condition A; Figure 3.19, black trace) was methanol/ ethyl acetate/ 0.1% *v/v* acetic acid of 52/40/8. The possibility of increasing the separation by decreasing the polarity of mobile phase gradient was examined (see the details in Table A.1-A.4 in Appendix A). Thus, the original (see Table A.1) and three more apolar mobile phase gradients were compared (see Table A.2-A.4). The HPLC chromatograms were recorded using both FL and UV detectors (Figure 3.19). The chromatograms obtained show all three peaks of the steroid FLB derivatives, identification of each peak being based on retention time. The peak widths were evaluated to enable the calculation of resolution (R_s) values when using each gradient programme.

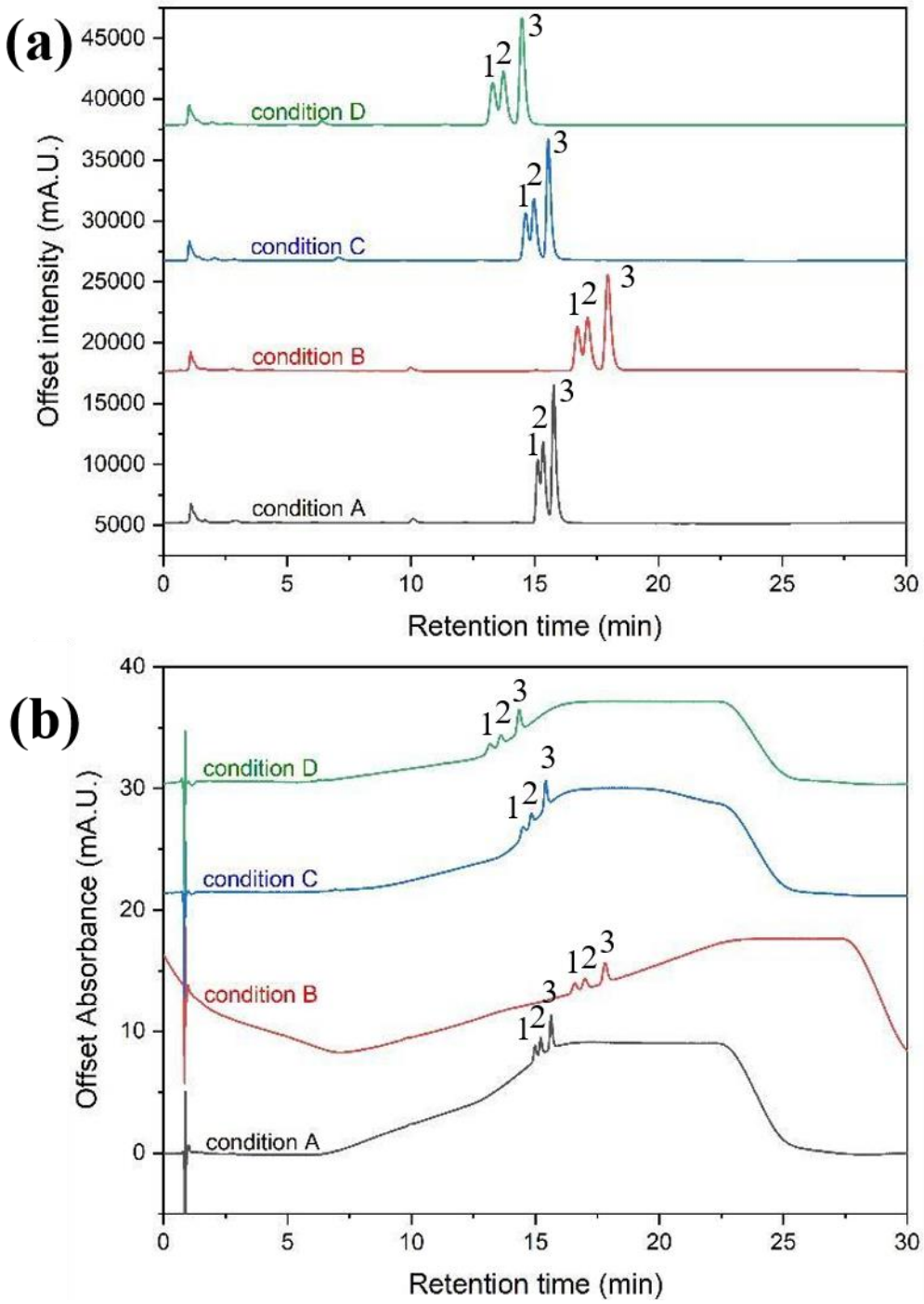


Figure 3.19. Chromatograms of the FLB derivatives of steroidal mixture recording using different detectors: (a) fluorescence and (b) ultra-violet. HPLC separations were operated by Symmetry C_{18} column (3.5 µm, 2.1 mm×150 mm) using various mobile phase components; condition A (Table A.1), condition B (Table A.2), condition C (Table A.3), condition D (Table A.4). Sterol mixture was chol-FLB (9 pmol), cop-FLB (12 pmol) and 5α-chol-FLB (13 pmol). Assigned peaks: 1 = coprostanol, 2 = cholesterol, 3 = 5α-cholestanol. The separation condition is shown in Section A.1.2.

The FL, UV and MS chromatograms show separation of all three components with UV and MS showing close to baseline resolution (Figure 3.19-3.20). The resolution of chol-FLB from each of the other two compounds in the mixture was determined (Table 3.3). Based on a resolution value above 1.5, 5α-chol-FLB was clearly distinguished using HPLC conditions B, C & D. Partial co-elution was observed for chol-FLB and cop-FLB when using HPLC conditions B, C & D for FL and UV detectors. The lower degree of resolution of chol-FLB and cop-FLB observed with the FL detector can be attributed to the larger flow cell volume of the detector. Thus, resolution was greatest using HPLC condition D. The separation condition D with MS detection (Figure 3.20) gave resolution values greater than 1.5 for cop-FLB: chol-FLB and chol-FLB: 5α-chol-FLB, which reflects its suitability for use in the determination of the coprostanol ratio in archaeological samples. Hence, condition D was selected for further use as the mobile phase gradient programming for separation of C_{27} steroid FLB derivatives.

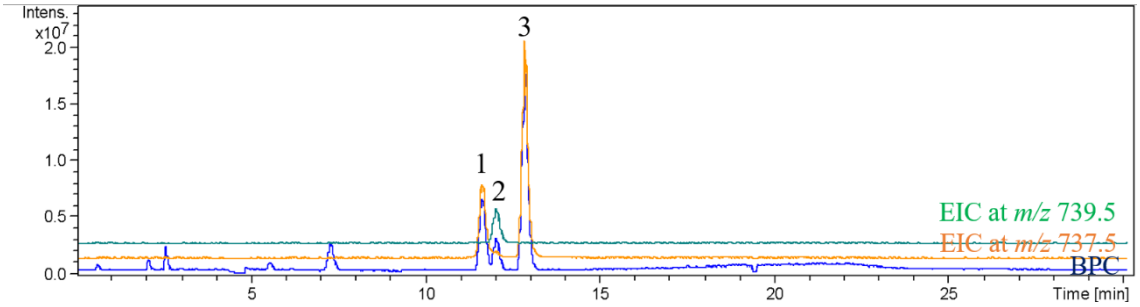


Figure 3.20. APCI MS chromatograms of the FLB derivatives of steroidal mixture recorded as base peak chromatogram and extracted ion chromatograms obtained from *m/z* 737.5, *m/z* 739.5. HPLC separations were on a Symmetry C_{18} column (3.5 µm, 2.1 mm×150 mm) using condition D (Table A.4). Steroidal lipids mixture was chol-FLB (9 pmol), cop-FLB (12 pmol) and 5α-chol-FLB (13 pmol). Assigned peaks: 1 = coprostanol, 2 = cholesterol, 3 = 5α-cholestanol.

Table 3.3. Calculated resolution of cholesterol and related stanols obtained from the chromatograms separated using HPLC conditions A-D with various detectors.

Separation condition	Resolution* of peaks (cop-FLB *vs.* chol-FLB)			Resolution* of peaks (chol-FLB *vs.* 5α-chol-FLB)		
	FL	UV	MS	FL	UV	MS
A	0.7	1.0	N.D.	1.4	2.0	N.D.
B	0.9	1.2	N.D.	1.9	2.4	N.D.
C	0.9	1.2	N.D.	1.6	2.0	N.D.
D	1.0	1.2	1.5	1.8	2.3	1.9

*Resolution (R_s) was evaluated from an equation: $R_s = [1.18\times(t_2-t_1)]/(W_{0.5,1}+W_{0.5,2})$ (Snyder *et al.* 1997)

N.D. = Non-determined, FL = fluorescence, UV = ultra-violet, MS = mass spectrometry

As detailed in Table 3.3 for condition D, a R_s value of 1.5 was observed for the separation of chol-FLB and cop-FLB by HPLC-MS, though partial overlap of the peaks at baseline was still apparent. The potential for improvement using UHPLC was investigated in an attempt to achieve baseline resolution of cholesterol and coprostanol. A mixture of the FLB derivatives of cholesterol, coprostanol and 5α-cholestanol was separated using an Acquity BEH C_{18} column (1.7 µm, 2.1 mm × 150 mm) with a UHPLC method modified from the solvent system by substituting ethyl acetate with dichloromethane (see Section 2.5.1.1 for final separation condition) and the MS chromatogram was recorded (Figure 3.21). The MS chromatogram shows baseline separation of all three components and gave R_s values for each adjacent component pair in the mixture above 2.3, clearly reflecting its potential for use in the determination of the coprostanol ratio in archaeological samples.

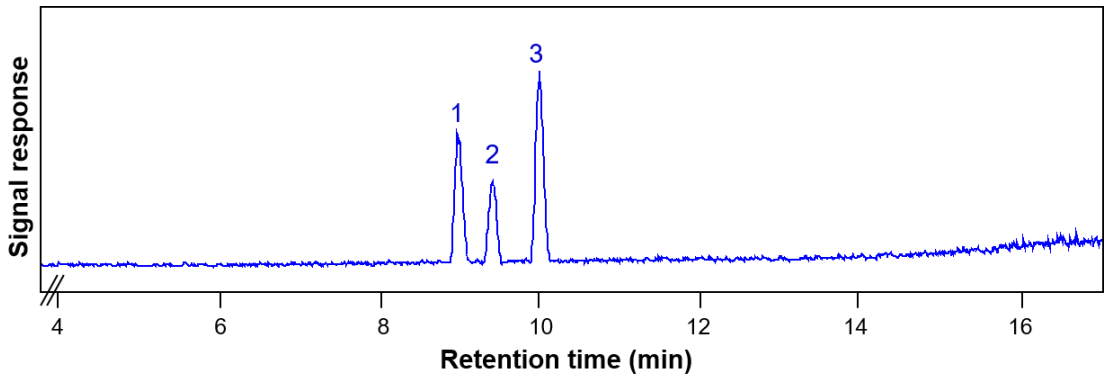


Figure 3.21. Partial APCI-MS based peak chromatogram of the FLB derivatives of steroidal mixture. Sterol mixture was chol-FLB (9 pmol), cop-FLB (12 pmol) and 5α-chol-FLB (13 pmol). Assigned peaks: 1 = coprostanol, 2 = cholesterol, 3 = 5α-cholestanol. *Resolution (*R*s) was evaluated using the equation: $R_s = [1.18 \times (t_2\text{-}t_1)]/(W_{0.5,1}+W_{0.5,2})$ (Snyder *et al.* 1997).

3.3.2.3.4 Assessment of internal standard for cholesterol and its reduction products (coprostanol and 5α-cholestanol)

An internal standard can compensate for changes in sample size or response due to instrument variations and can be used to correct for samples losses during preparation steps. An internal standard for cholesterol and its reduction products was considered to obtain a greater security of quantification. The selection of an appropriate internal standard for cholesterol and its stanols was based on the chromatographic behaviours observed in the soil extracts, where steroids in the range $C_{27} - C_{29}$ and *n*-alkanols in the range $C_{25} - C_{30}$ were present. The *n*-alkanol, 1-heptadecanol, was selected to assess its suitability as an internal standard due to it having a long alkyl chain and a hydroxyl group and being readily available. The FLB derivative of 1-heptadecanol (1-hep-FLB) was prepared in the same manner as the cholesterol-FLB (Figure 3.22).

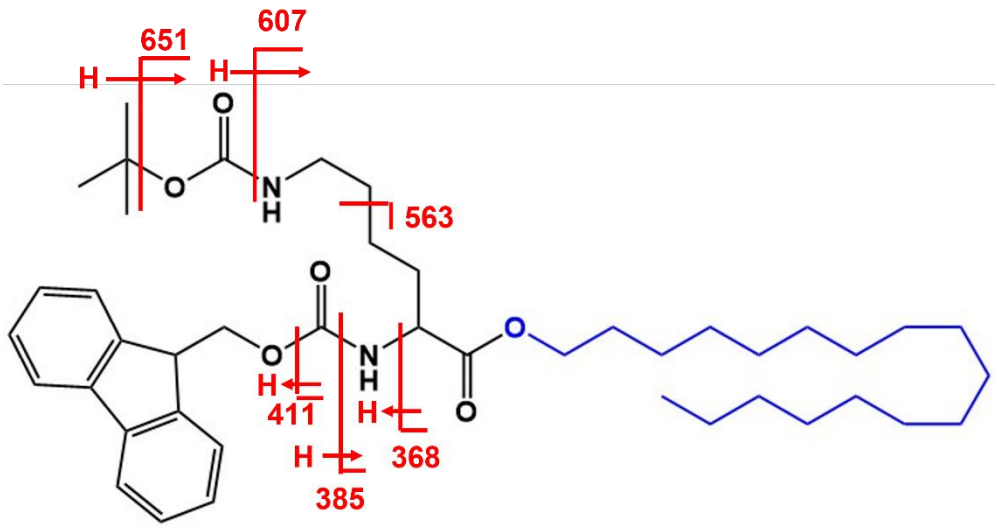


Figure 3.22. Chemical structure of 1-hep-FLB showing origins of fragment ions in the APCI MS spectrum.

Isolated 1-hep-FLB was analysed by UHPLC-MS to determine its chromatographic behaviour and confirm the structure of the product (Figure 3.23). The UHPLC-MS chromatogram shows 1-hep-FLB eluting at a retention time of 7 min, confirming the formation of one major pure product (Figure 3.23a). The product was produced in high percentage yield by weight of the isolated 1-hep-FLB (93 ± 2 %, n = 5).

The MS behaviour of 1-hep-FLB was examined by MS and MS^2 (Figure 3.23b-e). The full MS spectrum of 1-hep-FLB shows a protonated molecule at *m/z* 707.5 and a prominent ion at *m/z* 607.5, corresponding to the loss of Boc group (-100 Da). All MS^2 spectra of 1-hep-FLB show product ions at *m/z* 651, *m/z* 607, *m/z* 563, *m/z* 411, *m/z* 385 and *m/z* 368 corresponding to cleavages of Boc and Fmoc groups at various positions as shown in the Figure 3.22, confirming the formation of the desired product.

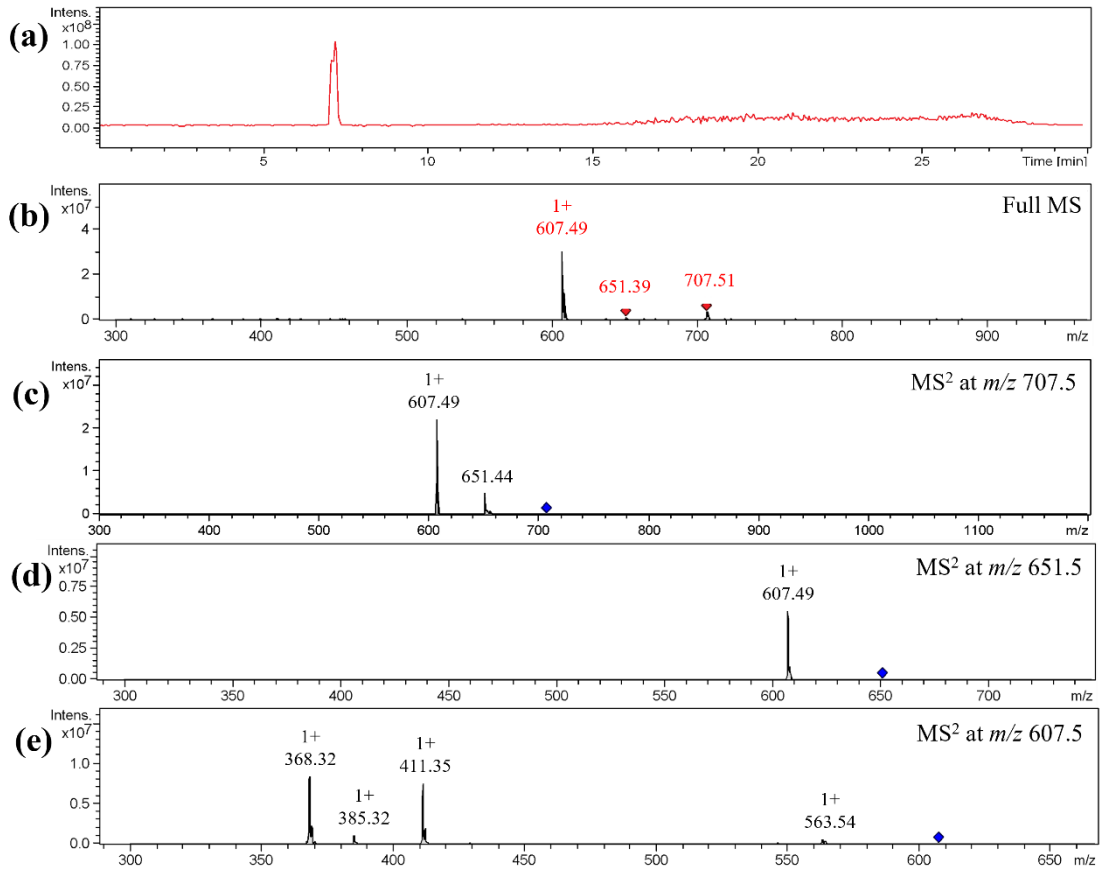


Figure 3.23. (a) APCI base peak chromatogram of 1-heptadecanol derivatised with FLB (1-hep-FLB), APCI MS and MS^2 spectra of 1-hep-FLB (b) Full MS and MS^2 spectra from precursors (c) *m/z* 707.5, (d) *m/z* 651.5 and (e) *m/z* 607.5. Concentration of 1-hep-FLB was 12 pmol injected on column.

Isolated 1-hep FLB was analysed by (U)HPLC with FL, UV and MS detectors to compare signal responses. The (U)HPLC chromatograms obtained from all three detectors (FL, UV and MS) show the isolated 1-hep-FLB eluting at a retention time of 7 min, confirming a formation of one major product with no residue (Figure 3.24).

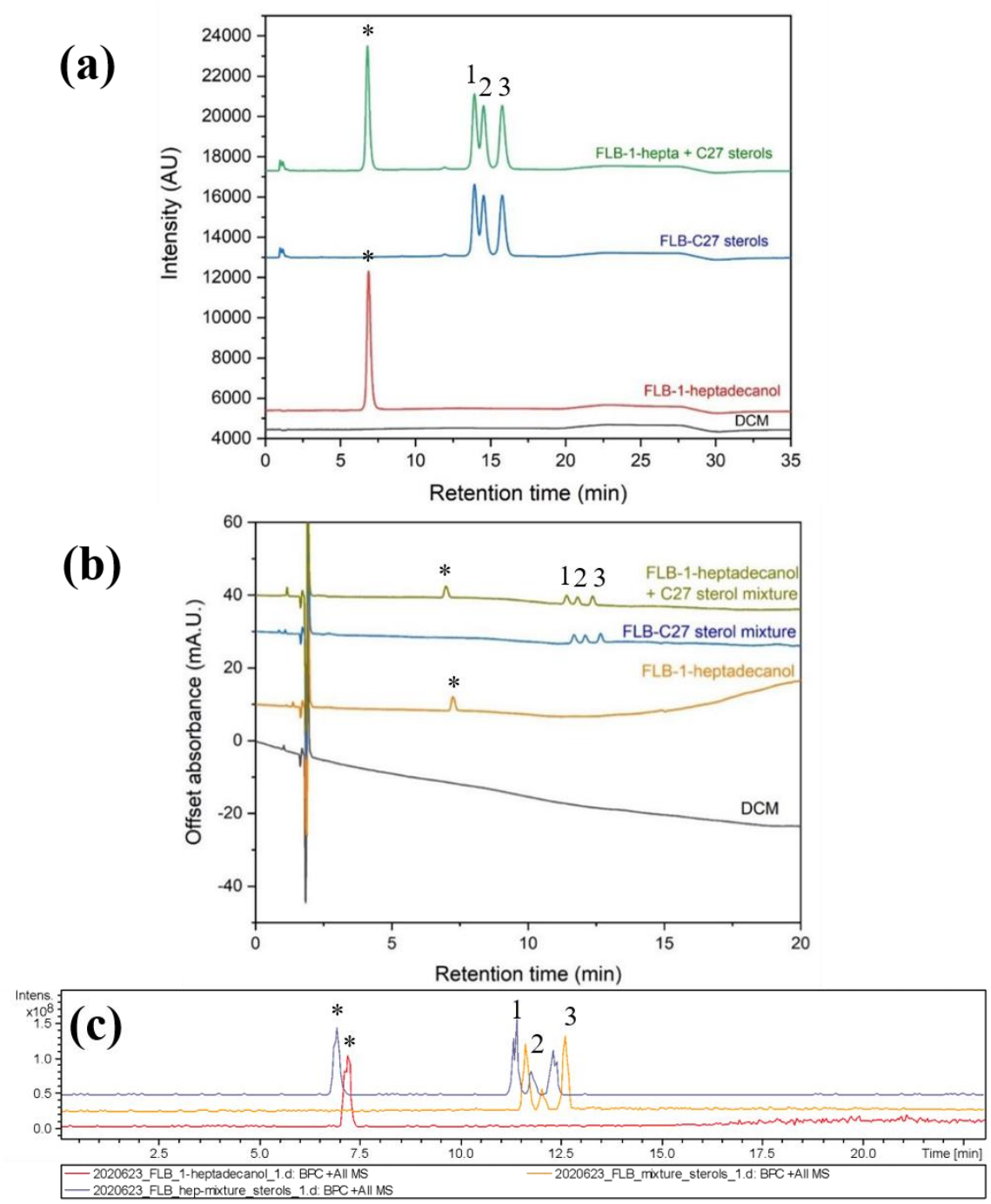


Figure 3.24. (U)HPLC chromatograms of 1-heptadecanol derivatised with FLB (1-hep-FLB), the FLB derivatives of steroidal mixture with and without 1-hep-FLB corresponding to various detectors: (a) fluorescence, (b) ultra-violet, (c) APCI MS. Concentration of 1-hep-FLB was 16 µM (12 pmol injected on column, asterisk peak). Steroidal mixture was cop-FLB (9 pmol, peak 1), chol-FLB (9 pmol, peak 2) and 5α-chol-FLB (9 pmol, peak 3).

The retention time and response of 1-hep-FLB were compared with those for the FLB derivatives of cholesterol, coprostanol and 5α-cholestanol. Chromatographic separations show that 1-hep-FLB (peak *) is separated well from the C_{27} steroidal-FLB derivatives. LC-MS chromatograms of the soil extracts examined show no peaks eluting at the retention time of hep-FLB (data not shown), reflecting the potential to use 1-hep-FLB as an internal standard.

3.3.2.3.5 Analytical figures of the developed method

The reliability in identification and quantification of C_{27} steroidal lipids as the FLB derivatives was assessed by an investigation of a full-range analytical figures including working range, limit of detection, limit of quantification, repeatability, reproducibility, and recovery (*cf* McNaught and Wilkinson 1997).

3.3.2.3.5.1 Working range, LOD and LOQ

The working range for measurement of FLB derivatives of the three C_{27} steroidal lipids with MS, UV and FL detection was assessed from peak area measurements of chol-FLB, cop-FLB and 5α-chol-FLB during (U)HPLC separation of a series of concentrations of a mixture prepared by serial dilution (Figure 3.25-2.27). The serial dilutions were performed to examine the limit of detection based on $3 \times S/N$, and the linearity in the response (Miller and Miller 2010; Christian *et al.* 2014). The derivatives show excellent responses for both MS and UV detection with limit of detection (LOD) at the picomolar level (0.2 pmol and 0.8 pmol, respectively). The LODs obtaining from the method are lower than the values reported by Poplawski (75 pmol and 500 pmol for FL and UV detection, respectively) (Poplawski 2017). The limits of quantification (LOQ) were in range of 0.8-2.2 pmol. Sensitivity and dynamic range were evaluated based on peak area for both MS and UV detection methods (Figure 3.25b and 3.26b). The relationship of the steroidal-FLB peak area with molar concentration gave the linear responses for MS detection (r^2 >0.99, Synder *et al.* 1997; FDA guidance 2020) over two different concentration ranges: (1) 3.5 – 70 pmol for chol-FLB, and (2) 3.5-35 pmol for cop-FLB and 5α-chol-FLB (Table 3.4). While the peak area of

these three steroidal-FLB gave the linear responses for UV detection (r^2 >0.99) over the measured concentration range of 3.5 – 70 pmol (Table 3.4).

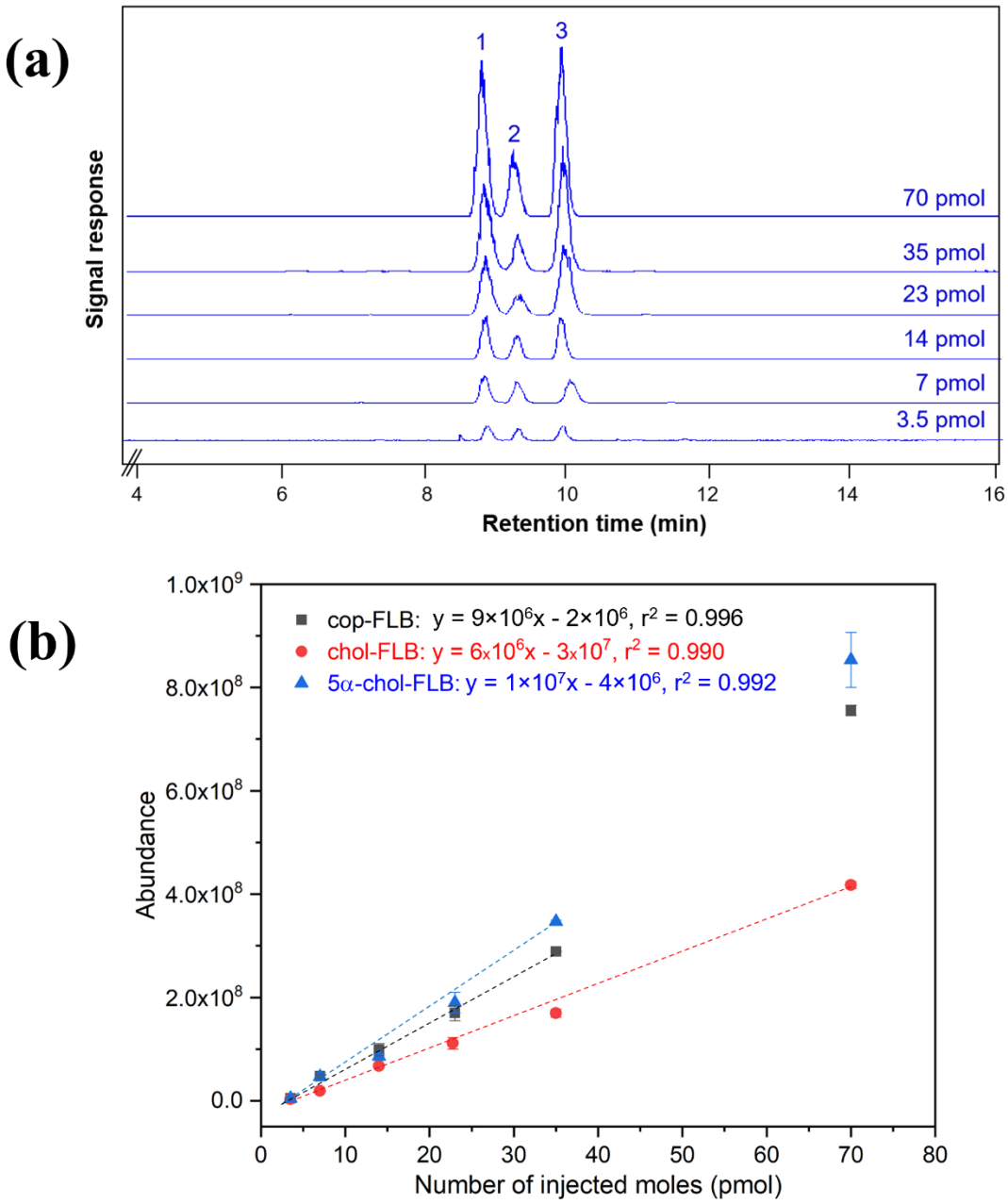


Figure 3.25. (a) Partial APCI MS chromatograms of the FLB derivatives of a mixture comprising cop-FLB (peak 1), chol-FLB (peak 2) and 5α-chol-FLB (peak 3) analysed at various concentration levels. UHPLC separations employed an Acquity BEH C_{18} (1.7 µm, 2.1 mm×150 mm) at a flow rate of 0.3 mL min^{-1}. (b) Calibration curve of MS response versus concentration (3.5-70 pmol).

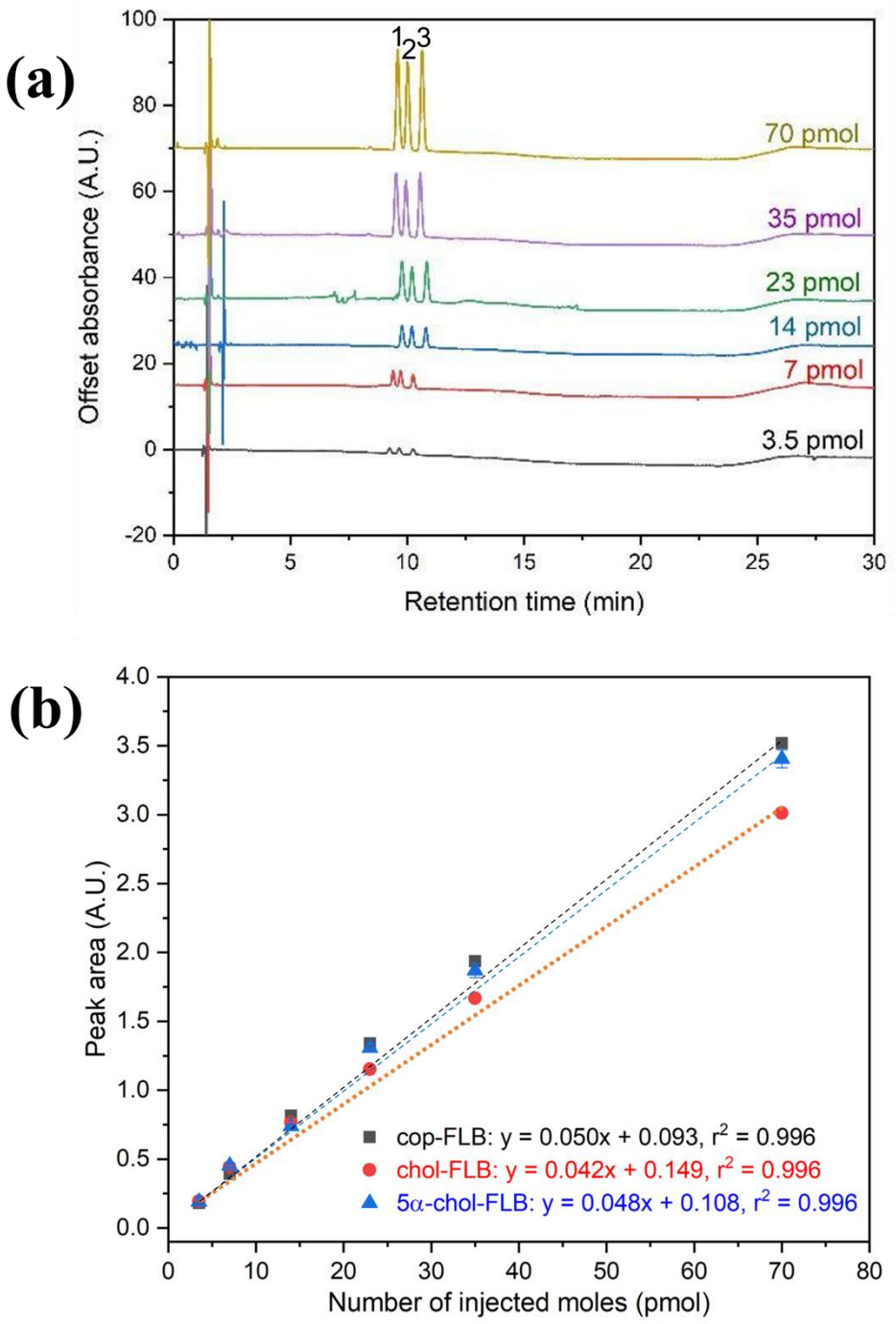


Figure 3.26. (a) UHPLC UV-DAD chromatograms of a mixture comprising cop-FLB (peak 1), chol-FLB (peak 2) and 5α-chol-FLB (peak 3) analysed at various concentration levels. UHPLC separations employed an Acquity BEH C_{18} (1.7 µm, 2.1 mm×150 mm) at a flow rate of 0.3 mL min^{-1}. (b) Calibration curve for UV response versus concentration (3.5-70 pmol).

For fluorescence detection, the steroidal derivatives were separated using HPLC due to the pressure limit of the detector (< 400 bar). The HPLC chromatographic behaviour of the derivatives confirmed reliable retention time stability (<2% RSD, *n* = 3 for each concentration) though a small degree of peak tailing is evident (Figure 3.27a). The FL detector gave a linear response for all the measured concentrations in the range of 35 fmol – 7 pmol, r^2 value of all above 0.99 (Figure 3.27b and Table 3.5). The LOD and LOQ obtained from fluorescence detection were 0.25 fmol and 0.93 fmol, respectively.

The results indicate that steroidal-FLB compounds give a higher response by fluorescence detection than by ultraviolet or MS detection, by approximately 1,000 times. The higher fluorescence signal response for the steroidal-FLB derivatives is very similar to that observed by Poplawski, who reported that the linear ranges of chol-FLB using both methods were 100 pmol – 10 nmol and 1 – 100 nmol for FL and UV detection, respectively (Poplawski 2017). The significantly greater response by fluorescence provides compelling evidence that FLB is a promising derivatising agent for trace analysis of cholesterol in environmental samples at the ppb level, which is to the same concentration ranges than these steroidal compounds reported in the literature (Tables 1.1 and 1.2).

The LODs and LOQs obtained using MS and UV detection to monitor the FLB derivatives shows that the C_{27} steroidal lipids can be detected at the ppm level, equating to the concentration ranges of these steroidal compounds reported in the literature (Tables 1.1 and 1.2). For instance, the reported LOD ranges of C_{27} steroidal lipids were 3−90 pmol for human faeces and 2 pmol – 47 nmol for river sediments, while the reported LOQ ranges of the same types of samples were 26−301 pmol and 6 pmol -15 nmol, respectively. The better LOD and LOQ values obtained from the FLB derivatives by MS, UV and FL demonstrate improved capability over the conventional analysis. Thus, the potential exists to use the approach in a wide range of applications including analysis of archaeological, environmental, or medical samples. In situations where the approach and the conventional GC-MS based approach may be limited by the amounts of the C_{27} steroidal

lipids available (low concentrations or small sample size) the use of FL detection offers the potential to detect components down to femtomole (fmol) levels.

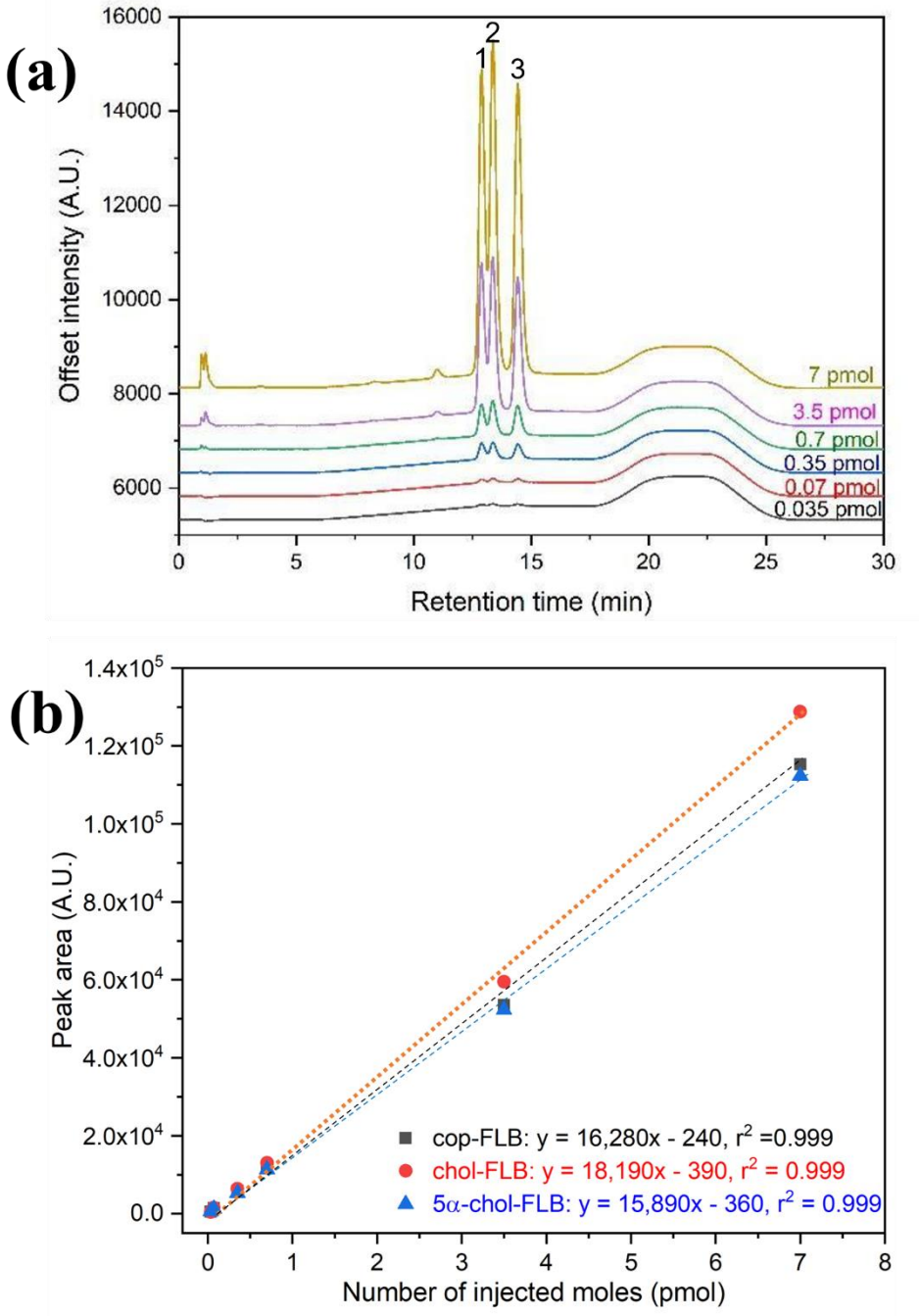


Figure 3.27. (a) HPLC chromatograms of the FLB derivatives of a mixture comprising cop-FLB (peak 1), chol-FLB (peak 2) and 5α-chol-FLB (peak 3) analysed at various concentration levels. HPLC separations employed a Symmetry C_{18} column (3.5 µm, 2.1 mm×150 mm) at a flow rate of 0.5 mL min^{-1}. (b) Calibration curve of FL response versus concentration (35 fmol -7 pmol).

Table 3.4. Analytical characteristic of the LC methods corresponding with different detectors for FLB derivatives of each C_{27} steroidal lipids.

Compound	t_r (min)	Calibration curve	Correlation factor (r^2)	Calibration range (nmol), 1 µL inj.	LOD (mol)	LOQ (mol)
UHPLC-MS						
Coprostanol	9.5	$y = 9\times10^6x - 2\times10^6$	0.996	3.5 – 35	0.35 pmol	1.2 pmol
Cholesterol	9.9	$y = 6\times10^6x - 3\times10^7$	0.990	3.5 – 70	0.55 pmol	1.8 pmol
5α-cholestanol	10.5	$y = 1\times10^7x - 4\times10^6$	0.992	3.5 – 35	0.24 pmol	0.8 pmol
UHPLC-UV-DAD						
Coprostanol	9.3	$y = 0.050x + 0.093$	0.996	3.5 – 70	0.76 pmol	2.16 pmol
Cholesterol	9.7	$y = 0.042x + 0.149$	0.996	3.5 – 70	0.66 pmol	1.88 pmol
5α-cholestanol	10.2	$y = 0.048x + 0.108$	0.996	3.5 – 70	0.57 pmol	1.63 pmol
HPLC-Fluorescence						
Coprostanol	13.3	$y = 16{,}280x - 240$	0.999	0.035 – 7	0.28 fmol	0.93 fmol
Cholesterol	13.8	$y = 18{,}190x - 390$	0.999	0.035 – 7	0.26 fmol	0.88 fmol
5α-cholestanol	15.0	$y = 15{,}890x - 360$	0.999	0.035 – 7	0.25 fmol	0.83 fmol

3.3.2.3.5.2 Repeatability and reproducibility (retention times & peak area)

The reliability of the UHPLC and HPLC chromatographic behaviour of the C_{27} steroidal FLB derivatives (chol-FLB, cop-FLB and 5α-chol-FLB) was examined by replicate separation of the mixture of FLB derivatives with MS and UV detection and by HPLC with FL detection. The retention times and measured peak areas of each C_{27} steroidal FLB derivative were determined both within one day and between days (n = 3 for each condition; Table 3.5).

Table 3.5. Repeatability and reproducibility relative standard deviations (%RSD) for retention times and peak areas of FLB derivatives for the three C_{27} steroidal lipids n = 3 for each condition.

Component	%RSD Retention time (%)		%RSD Peak area (%)	
	Repeatability	**Reproducibility**	**Repeatability**	**Reproducibility**
UHPLC-MS				
Coprostanol	1.9 – 4.8	1.4 – 3.2	0.3 – 11.7	1.9 – 4.0
Cholesterol	2.1 – 4.6	2.1 – 3.2	1.5 – 12.6	4.4 – 12.7
5α-cholestanol	1.7 – 4.4	1.4 – 3.1	0.9 – 10.1	1.6 – 10.3
UHPLC-UV-DAD				
Coprostanol	1.7 – 3.6	1.5 – 3.3	0.9 – 1.1	4.5 – 6.9
Cholesterol	1.7 – 4.3	1.3 – 3.9	0.5 – 2.1	0.1 – 4.9
5α-cholestanol	1.8 – 4.0	2.1 – 4.0	0.6 – 3.2	4.6 – 8.0
HPLC-Fluorescence				
Coprostanol	0.04 – 0.06	0.5 – 2.1	0.6 – 5.0	2.6 – 10.6
Cholesterol	0.05 – 0.11	0.8 – 2.1	1.6 – 11.0	2.5 – 10.4
5α-cholestanol	0.04 – 0.12	0.4 – 2.2	1.2 – 5.0	1.4 – 5.7

The reliability of retention times and measured peak areas from analyses performed on the same day is expressed as repeatability, and those measured on different conditions as reproducibility (*cf* McNaught and Wilkinson 1997). In this work, the time interval for measuring reproducibility was three weeks for performing the experiment. Good reproducibility in retention times (<2% RSD) was demonstrated despite a small degree of peak tailing (Figure 3.25). For MS and UV detections, the peak areas of the steroidal-FLB derivatives gave acceptable RSD of in the range of 0.3-13% for repeatability and reproducibility (*cf* McNaught and Wilkinson 1997). While HPLC-FL showed comparable values to UHPLC-MS, these values could be expected to improve

for UHPLC. The variation of %RSD obtained using peak area was up to 12.6%, equating to the %RSD ranges of these steroidal compounds reported in the literature which was up to 20% (Tables 1.1 and 1.2). Thus, the UHPLC method provides an opportunity to determine low levels of steroidal alcohols with good reproducibility, giving confidence in quantification.

3.3.2.3.5.3 Recovery

An assessment of the recovery of C_{27} steroidal lipids as their FLB derivatives was carried out using archaeological soil extracts obtained from five different depths in the same Roman cesspit core. The mixture of standard C_{27} steroidal lipids (cholesterol, coprostanol and 5α-cholestanol) were spiked into each cesspit soil at three different concentrations: low, medium, and high (4 µM, 35 µM, 50 µM). The internal standard, 1-heptadecanol (5 µg), was also added into all samples. The spiked cesspit soils were prepared and derivatised in the same manner as the soils and sediments described previously (see the procedure in Section 2.2.1 and 2.3.1). The resulting spiked FLB-soil extracts were separated by (U)HPLC methods with all three detectors (see the procedure in Section 2.5.1) to obtain chromatograms and peak areas of chol-FLB, cop-FLB and 5α-chol-FLB. Recovery of the C_{27} steroidal lipids was measured by comparing the peak areas of FLB-C_{27} steroidal lipids from spiked soil extracts with peak areas of FLB-C_{27} steroidal lipids from unspiked soil extracts (Figure 3.28-3.30) (Snyder *et al.* 1997). The LC separations for each condition were replicated (n = 6) for all five cesspit soils. The average recovery for FLB derivatives of each C_{27} sterol (Figure 3.28-3.30) show the acceptable range of 81 to 118% (*cf* McNaught and Wilkinson 1997; Synder *et al.* 1997), indicating that the method enables reliable recovery and quantification of coprostanol, cholesterol and 5α-cholestanol with all three detectors (MS, UV and FL).

The stability of the FLB derivatives with cholesterol and cholestanols was assessed to determine if extracts could be stored. A mixture of chol-FLB, cop-FLB and 5α-chol-FLB was repeatedly analysed by LC-FL to obtain peak area of their compounds every two days. The FL response of the FLB derivatives of the mixture were constant for 14 days with the %RSDs less than 5%. Re-

analysis of the extract after a further 2 days showed a slight decrease in response indicating that, the steroidal lipids in the FLB form should be analysed within two weeks.

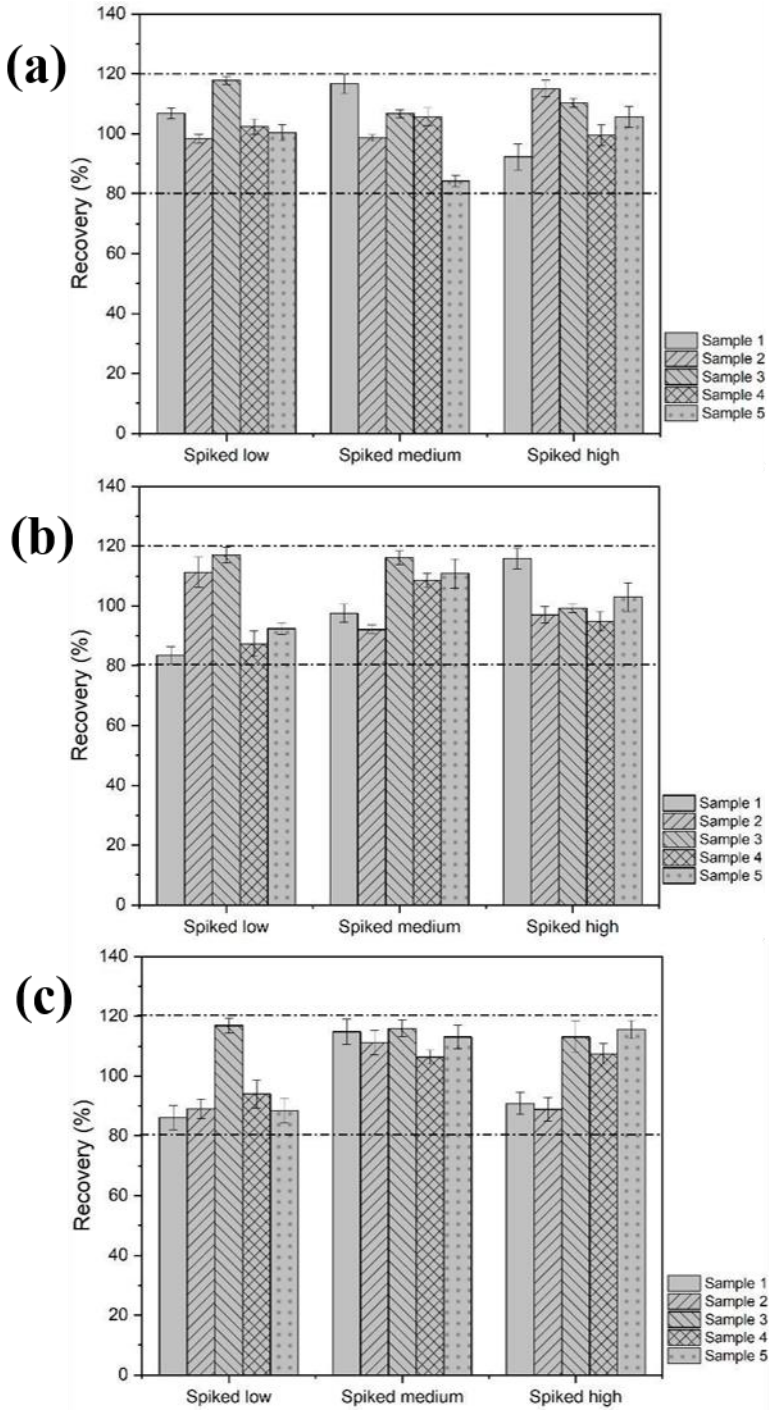


Figure 3.28. Percentage recovery measurements for cholesterol and its cholestanols detected by UHPLC-MS (a) coprostanol, (b) cholesterol, (c) 5α-cholestanol from five cesspit soils (n = 6). Dotted line = acceptable range of the recovery (80-120%).

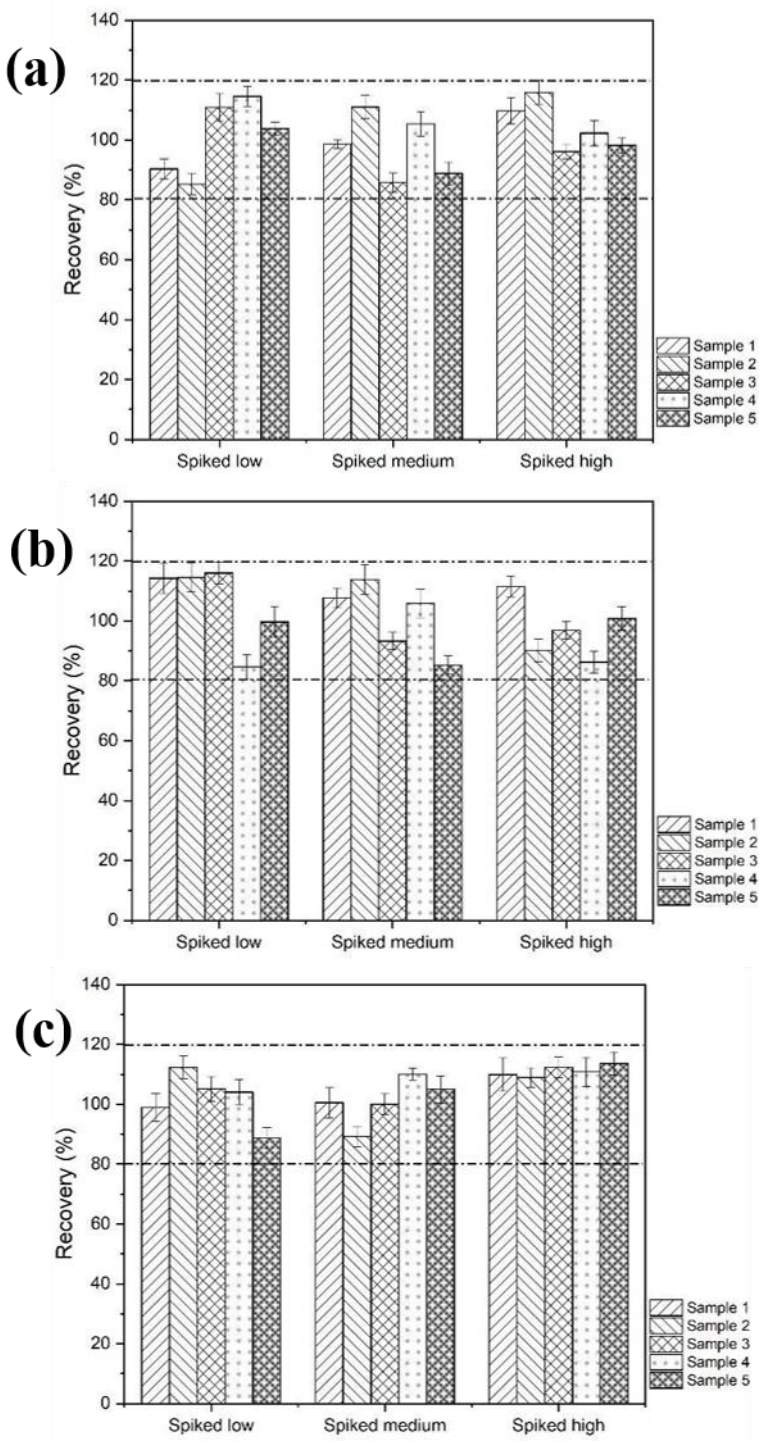


Figure 3.29. Percentage recovery measurements for cholesterol and its cholestanols detected by UHPLC-UV (a) coprostanol, (b) cholesterol, (c) 5α-cholestanol from five cesspit soils (n = 6). Dotted line = acceptable range of the recovery (80-120%).

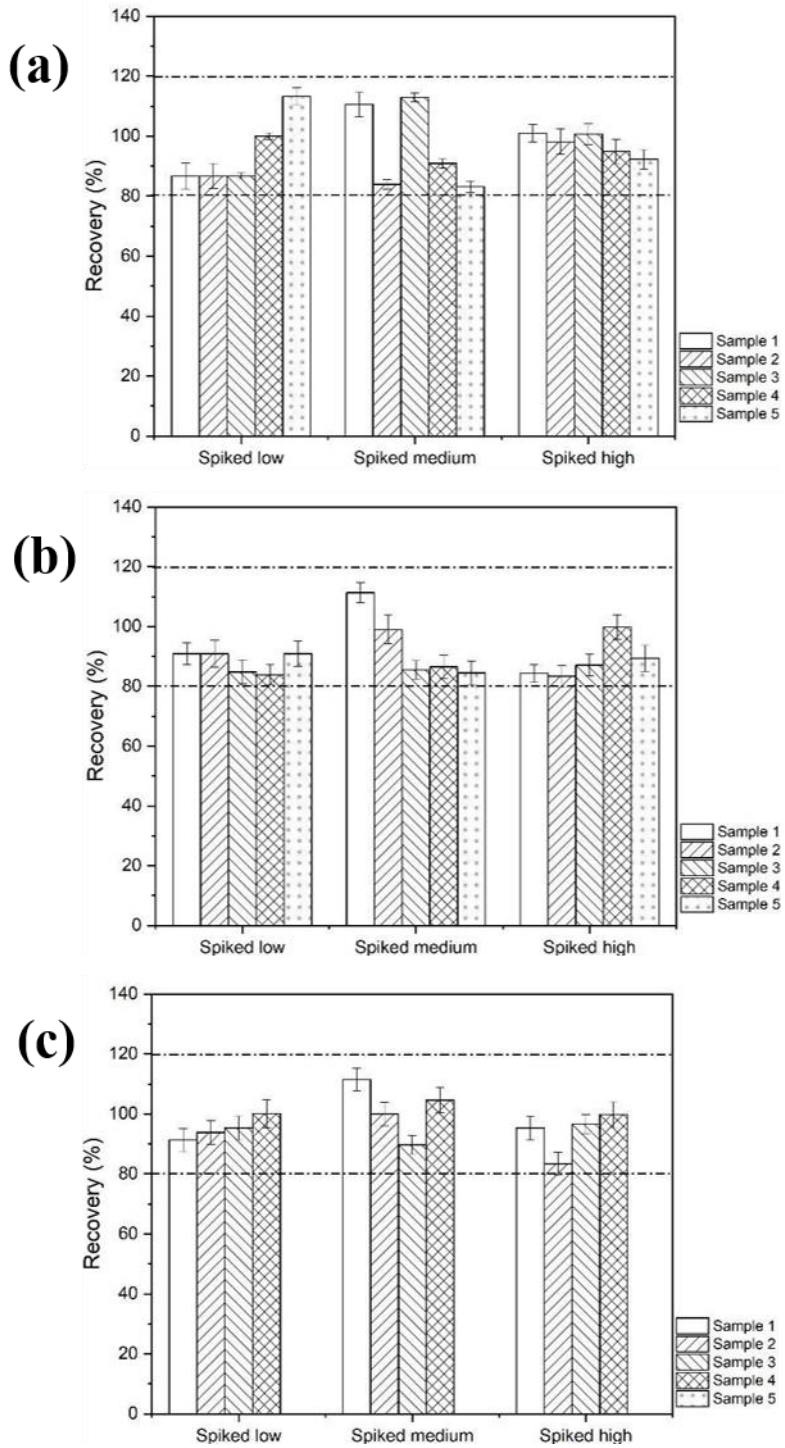


Figure 3.30. Percentage recovery measurements for cholesterol and its cholestanols detected by HPLC-FL (a) coprostanol, (b) cholesterol, (c) 5α-cholestanol from five cesspit soils (n = 6). Dotted line = acceptable range of the recovery (80-120%).

3.4 Conclusion

The method using Fmoc-Lys-Boc (FLB) as a derivatising agent for the measurement of cholesterol and its reduction products (coprostanol and 5α-cholestanol) to detect human activity in archaeological contexts was evaluated systematically.

The FLB derivatives of cholesterol and cholestanols were successfully prepared by Steglich esterification; the completeness of the derivatisation process for standard cholesterol was improved by increasing the ratio of derivatising agent to steroid relative to the proportions suggested by Poplawski (2017). Structural characterisation of FLB derivatives of C_{27} steroidal lipids (cholesterol, coprostanol and 5α-cholestanol) by mass spectrometry confirms the formation of the desired products. A study of the dissociation pathways of FLB derivatives of cholesterol and its reduction products by multistage mass spectrometry reveals characteristic fragmentation of FLB derivatives that can inform the identification of unknown alcoholic components.

The preliminary screening of cholesterol and its transformation products by (U)HPLC showed that the method previously proposed by Poplawski (2017) needed to be modified to remove residues from the derivatisation process and to achieve baseline resolution of cholesterol and coprostanol. The analytical method was modified successfully, showing an improvement in signal responses of cholesterol, coprostanol and 5α-cholestanol. The residues from the derivatisation process were mostly removed using a modified sample clean-up with a more apolar eluent. The co-elution of cholesterol and coprostanol in soil extracts was resolved by modification of the (U)HPLC conditions, resulting in baseline separation of these two analytes. A full-range of analytical characteristics (repeatability and reproducibility for intraday and interday, working range, limit of detection, limit of quantification and recovery) of the refined method demonstrate that the method provides reliable analysis of coprostanol, cholesterol and 5α-cholestanol with all three detectors (MS, UV and FL). The lower LODs and LOQs in standard solutions obtained from FL detection than the more widely used detectors (MS, UV) offers the greater potential to detect C_{27} steroidal lipids at trace levels, reflecting the potential to use for wide range applications for archaeological, environmental, or medical samples. The FLB derivatives

with cholesterol and cholestanols gave the maximum signal responses at least two weeks after preparation, reflecting the ability to use this approach as a routine analysis with a security for quantification. Hence, the refinement of an approach for measurement of cholesterol, coprostanol and 5α-cholestanol as the FLB derivatives is suitable for identifying and quantifying steroidal compounds in soil extracts.

Chapter 4

Application of the measurement of cholesterol and its stanol derivatives

4.1 Introduction

The study of human occupation and activity in the archaeological records typically focuses on the recording, recovery and analysis of visible artefacts, human remains or living spaces and/or structures (O'Connor *et al.* 2011; Pickering *et al.* 2018). While some organic materials from human tissues or faeces can be well-preserved over long timescales, their visual recognition is rare and generally limited to processes such as mummification and anoxic depositional environments where preservation is exceptional (Evershed *et al.* 1988; Gülacara *et al.* 1990; Mayer *et al.* 1997; Prost *et al.* 2017; Pickering *et al.* 2018). Organic materials that enter burial environments have the potential to yield biological marker compounds which, though invisible to traditional archaeological methods, have potential as chemical signatures of human occupation (Bujagić *et al.* 2016; von der Lühe *et al.* 2018; Pickering *et al.* 2018).

The analysis of organic matter has been exploited as a tool in several fields including geochemistry, medicine, forensic science, archaeology, and environmental analysis (von der Lühe *et al.* 2013; Prost *et al.* 2017; Schött *et al.* 2018; Harrault *et al.* 2019). Various types of organic material such as bone collagen have been detected in archaeological and human remains (Evershed 2008) and various other proteins and lipid signatures from body tissues have been recognised (Evershed *et al.* 1988; Gülacara *et al.* 1990; Mayer *et al.* 1997; O'Connor *et al.* 2011). Bone collagen can survive over archaeological timescales and is routinely isolated for stable isotope and radiocarbon analysis (Corr *et al.* 2005, 2008). Likewise, the characterised organic residues from the wrappings of Egyptian mummies have provided information on the preparations used by ancient embalmers (Buckley *et al.* 2001). Those examples show that organic components can relate to a variety of sources such as body tissues, faeces, stomach, and gut contents, and from formulations used in preparing the body for burial. Cholesterol and coprostanol in soils and sediments have revealed information on the decomposition of human and porcine remains, and the distribution patterns of coprostanol used to identify human occupation and activity in areas that lack visible features (von der Lühe *et al.* 2013, 2018; Sistiaga *et al.* 2014). Such organic signatures have the potential to inform the archaeological, forensic science or environmental

interpretations. Notably, assessment of coprostanol and cholesterol to identify human activity or occupation can only be evaluated by comparing the concentrations of these components with other related transformation products in soil samples, and with those of the surrounding soils as reference background signatures.

Cholesterol and C_{29} steroidal lipids occur commonly as background signatures in soils and sediments, relating to the origins of organic matter from both plants and animals. Cholesterol reflects as an input from the degradation of all eukaryotic cells, animal tissues, plants, root exudates and several fungi species, while specific C_{29} steroidal lipids (β-sitosterol, 5β-stigmastanol and 5α-stigmastanol) evidence organic matter derived from plant-based sources or herbivore activity (Prost *et al.* 2017, 2018). Incorporation of cholesterol from human/animal bones and tissues into soils can significantly increase the amount of cholesterol in soils. The cholesterol concentration found in soils underneath decaying pig bodies was three times higher than those obtained from surrounding soils (von der Lühe *et al.* 2013), while the distribution of cholesterol between 8 and 12 cm depth in the soil could be detected for over a year after human cadaver removal (von der Lühe *et al.* 2018). Thus, the higher concentration of cholesterol in the area of a suspected burial than in the surrounding soils can provide evidence of *in situ* body decomposition (von der Lühe *et al.* 2018; Pickering *et al.* 2018).

In order to allow steroidal lipid signatures to be meaningfully employed, it is necessary to gain insights into the nature and extent of their alteration in the environment. As such, the potential of coprostanol and its related stanols as indicators to assess faecal inputs in soils and sediments has been evaluated extensively (Walker *et al.* 1982; Grimalt *et al.* 1990; Bethell *et al.* 1994; Evershed *et al.* 1997; Bull *et al.* 1999; Stott *et al.* 1999). Degradation of cholesterol and coprostanol from mammal faeces after deposition can be influenced by the environment. The steroids can be further reduced to the 5α-epimer or transformed to the 3α-epimer (Björkhem *et al.* 1971; Eyssen *et al.* 1973; Bull *et al.* 2002). Cholesterol is reduced to 5α(H)-cholestan-3β-ol (hereafter, 5α-cholestanol) by soil microbial flora, while coprostanol is transformed to 5β(H)-cholestan-3α-ol (epicoprostanol) *via* epimerisation (Bull *et al.* 2002; Prost *et al.* 2017).

Cholesterol and its related stanols can be preserved in soils or sediments over archaeological timescales, the signatures having the potential to reveal aspects of human occupation and/or activity in the area. Recognition of the relationships between the steroid structures led to the establishment of criteria for recognition of human faecal contamination, the coprostanol ratio (Evershed *et al.* 1997; Bull *et al.* 1999). Values of the ratio of coprostanol to total cholesterol transformation products greater than 0.7 are indicative of human faeces (Evershed *et al.* 1997; Bull *et al.* 1999). Coprostanol signatures from faecal matter have also been used to reconstruct meals of different species (Homo sapiens and Neanderthals) or recent meals consumed by humans (Prost *et al.* 2017; Sistiaga *et al.* 2014) and primates (Sistiaga *et al.* 2015). Likewise, high concentrations of faecal sterols in water and soils enable the monitoring of sewage pollution (Stefens *et al.* 2007; Birk *et al.* 2011; Bujagić *et al.* 2016). By contrast, the potential for quantification of cholesterol-related steroids in human occupation soils has received little attention.

The conventional analytical process for determining cholesterol and its related stanols in soils and sediments comprises three main steps; (1) extraction of organic matter from raw materials, (2) lipid fraction based on compound class, (3) derivatisation and detection of the silyl ethers by gas chromatography-mass spectrometry (GC-MS) (Bethell *et al.* 1994; Battistel *et al.* 2015; Prost *et al.* 2018). The soils or sediments are freeze-dried to remove moisture, before grinding to particle sizes of less than 200 μm diameter. The organic matter can be extracted from the dried and ground samples by various techniques such as sonication (Isobe *et al.* 2002; Birk *et al.* 2011; von der Lühe *et al.* 2018), Soxhlet extraction (Grimalt *et al.* 1990; Tort *et al.* 2017) or accelerated solvent extraction (ASE) (Prost *et al.* 2017; Pickering *et al.* 2018). The resulting organic matter total extracts comprise various compound classes including hydrocarbon, aromatic hydrocarbon, alcohol, and carboxylic acid components. The organic matter total extracts are fractionated by silica column chromatography and collected steroidal components recovered within the alcohol or lipid fraction. The alcohols are converted to their trimethyl silyl ether (TMS) derivatives (see mechanism of reaction in Figure 4.1) in order to improve volatility and ionisation efficiency for

GC-MS analysis. Gas-chromatography (GC) provides chromatographic retention times for each species and mass spectrometry enables their identification.

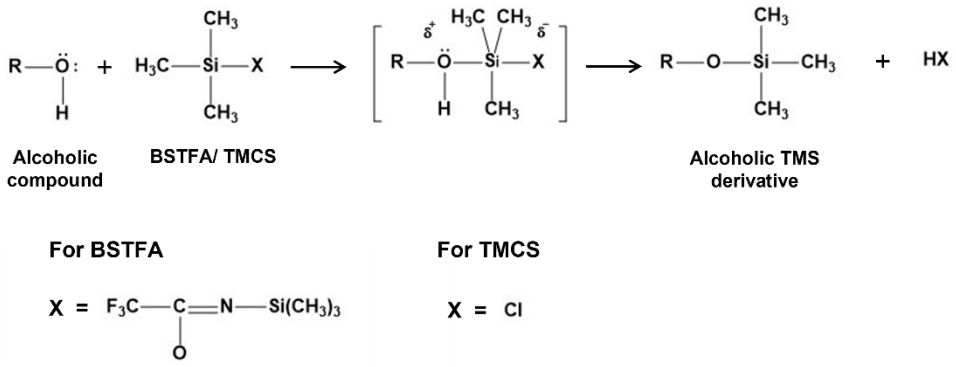


Figure 4.1. Scheme of silylation reaction of alcoholic compound derivatised with trimethylsilyl (TMS) reagent. The TMS reagent comprises BSTFA (*N*,*O*-bis(trimethylsilyl) trifluoroacetamide) and TMCS (trimethylchlorosilane).

Despite the detection of cholesterol and its related stanols by GC-MS being the conventional method, the quality of the chromatograms obtained has received little attention. With the separation of compounds by gas chromatography (GC) relying mainly on the boiling point of each compound, high numbers of lipids can lead to GC chromatograms exhibiting low separation efficiency (see Figure 4.2). This example chromatogram shows that the relative signal responses of cholesterol and its related stanols are much lower than those of the co-occurring *n*-alkanols.

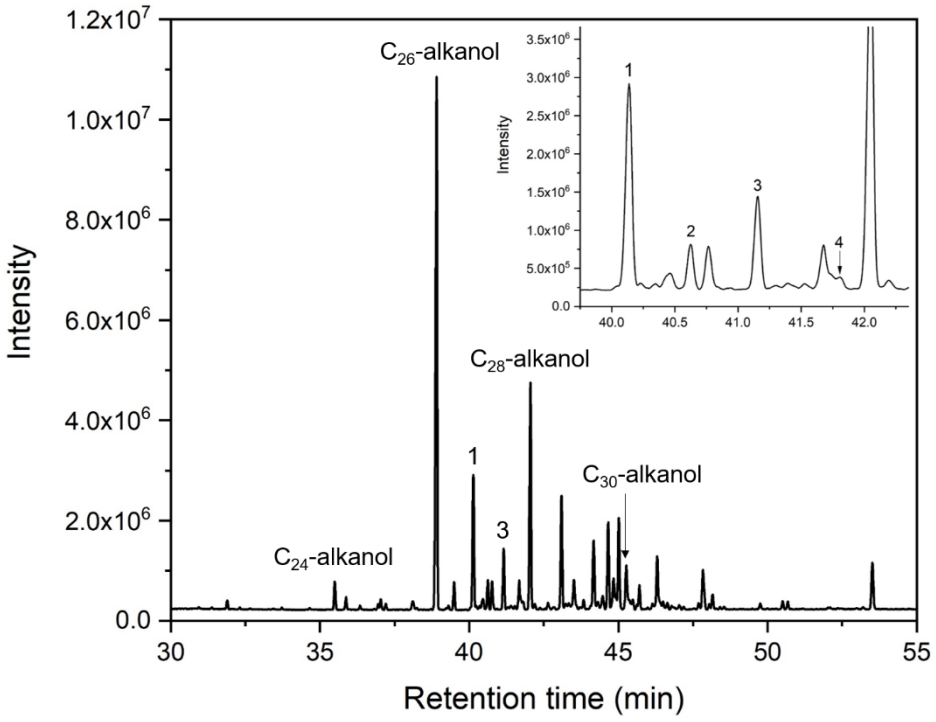


Figure 4.2. Partial GC-MS chromatogram of the medium polar fraction of the soil extract from cesspit core number 51150 at depth 29-30 cm, from Hungate, York (0.8 mg) derivatised with TMS. Assigned peaks: 1 = coprostanol, 2 = epicoprostanol, 3 = cholestan-3-one, 4 = 5α-cholestanol.

As seen from Figure 4.2, the chromatogram obtained from the soil extract shows co-elutions between some components due to its complexity. Likewise, in cases where samples are of limited availability (e.g., archaeological samples), some components may give peaks exhibiting low signal-to-noise ratio (Jayasinghe *et al.* 1998; Kunz and Matysik 2019; Nakakuni *et al.* 2019) possibly leading to misinterpretation. These obstacles could be overcome by development of an analytical technique providing both satisfactory separation efficiency and favourable detection limits for the determination of cholesterol and the full suite of transformation products obtained from the complex distributions that commonly occur in samples such as soils and faecal material.

Various derivatisation agents have been used to modify steroids and their transformation products, some with the aim of enhancing their MS signal response paralleled with improving separation efficiency (see Section 1.1.3.2). The use of UV-Vis or FL detectors has been reported

for determination of sterols and stanols, giving detection limits in the same range as MS detection (ppm to ppb levels) (see Section 1.1.3.2). Unfortunately, the reported methods have not been applied to the determination of cholesterol and related stanols in soils and sediments and give similar detection limits to GC-MS (ppm level).

Another derivatising agent, FMOC-Lys-Boc (FLB), has been reported to enhance the signal response of lipids and to be amenable to detection by MS, UV-vis and FL detectors (Poplawski 2017). The use of Fmoc-Lys-Boc (FLB) as a derivatising agent for the measurement of cholesterol and its reduced products (coprostanol and 5α-cholestanol) was evaluated systematically by LC separation combined with MS, UV, and FL detections (see Chapter 3). The initial studies suggest that the derivatives may offer greater security and improve detection limits compared with the conventional trimethylsilyl ethers analysed by GC-MS. The refinement of the approach for measurement of steroidal FLB derivatives indicates their potential for identifying and quantifying steroidal compounds in soil extracts for various applications in archaeology, environmental and forensic science.

4.2 Aims

The overall aim of the work described in this chapter was to demonstrate the extended applicability of the application of FLB derivatives of cholesterol and its reduction products (coprostanol and 5α-cholestanol) to indicate past human activity from signatures in different types of soils and sediments. Specific objectives were as follows:

1. To explore the potential to improve the completeness of the derivatisation process when applied to soils/sediments (see Section 4.3.2).

2. To apply the method (see the detail in Section 2.5.1) for assessment of human faecal contamination in archaeological samples from a Roman-Scandinavian cesspit core (Hungate, York) and to Roman grave soils collected from specific anatomical locations beneath human skeletons (Hungate, York) (see Section 4.3.3.1).

3. To test the application of the method to sediments obtained from the Humber estuary, Hull (see Section 4.3.3.2).

4.3 Results and discussion

4.3.1 UHPLC-MS Identification of FLB derivatives of components in soil extracts

The refined UHPLC-MS analysis method (see Section 2.5.1) was applied to the FLB derivatives of cholesterol, coprostanol and 5α-cholestanol in the extract of a soil sample from an Anglo-Scandinavian cesspit from Hungate, York (sample number 51151, for sample details see Section 4.3.3.1). The most abundant peak in the base peak chromatogram of the FLB derivatives (peak *, Figure 4.3) was assigned as 1-hexacosanol from the mass spectrum: protonated molecule at *m/z* 833.6 with the base peak at *m/z* 733.6 and a minor peak at *m/z* 777.6. The fragmentation is characteristic of FLB derivatives (Chapter 3) allowing a confirmation of the mass of the alcohol, 383.6 g/mol, matching with 1-hexacosanol. The peak intensity is consistent with the GC-MS analysis of the silylated total extract. The efficiency of the MS response was evaluated using 1-hexacosanol. The MS response of the FLB derivative in the soil extracts indicates an enhancement in the signal of around about 560 times compared with the native alcohol. The result confirms that FLB derivatisation significantly enhances the signals for alcohols, indicating its suitability for determination of trace amounts of cholesterol and cholestanol.

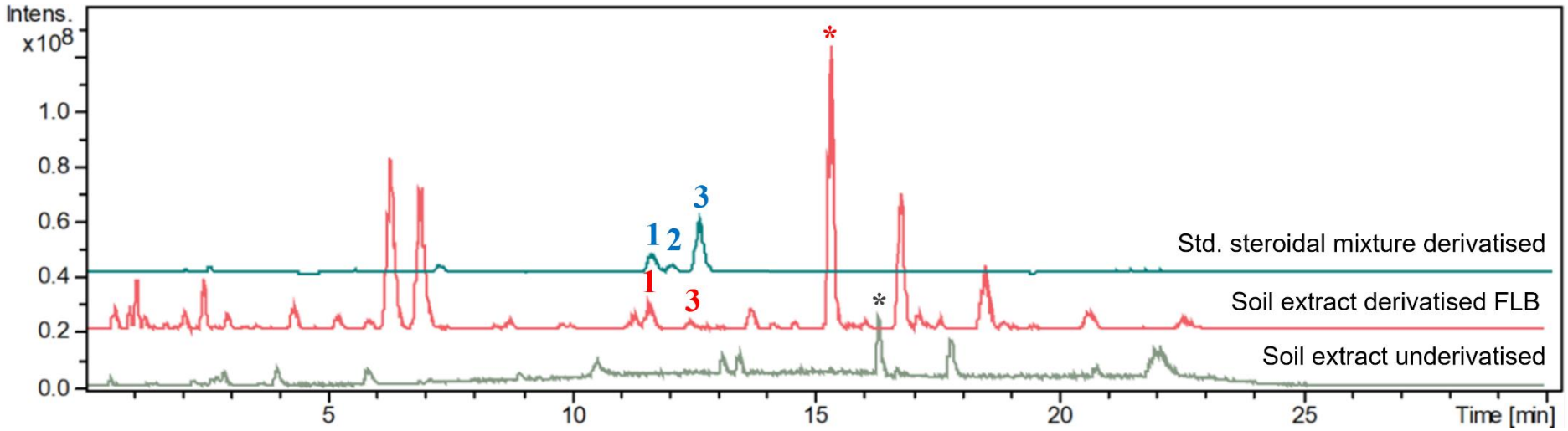


Figure 4.3. APCI MS base peak chromatograms of an extract of medium polar fraction of the soil extract from an Anglo-Scandinavian cesspit, Hungate, York detected as the native form (underivatised), and as the FLB derivatives. The chromatogram of the FLB derivatives of a mixture of C_{27} steroidal standards is included for comparison (1 = coprostanol, 2 = cholesterol, 3 = 5α-cholestanol, asterisk = hexacosanol). LC separations were carried out on a Symmetry C_{18} column (3.5 µm × 2.1 mm × 150 mm). The ratio of derivatising FLB reagent was 40 moles per mole of soil extract.

4.3.2 Evaluation of the amount of derivatising agent required for extracts from archaeological samples

Given that a variety of alcoholic compounds occur in archeological samples, the ratio of the derivatising reagents was re-examined to ensure complete derivatisation of steroidal compounds in soil sample extracts. Soil from an Anglo-Scandinavian cesspit core at depth 30-31 cm (Hungate, York) was used as the test material (for sample details see Section 4.3.3.1). Organic extracts obtained by accelerated solvent extraction (ASE) using DCM: methanol (9:1, *v/v*) were fractionated by column chromatography using hexane: ethyl acetate (4:1, *v/v*) (see procedure in Section 2.2.1). The medium polar fraction obtained from the soil extract was silylated and analysed by GC-MS to screen for the presence of trimethylsilylether (TMS) steroidal derivatives (see procedure in Section 2.3.2). The main peaks in the GC-MS chromatogram were identified by comparison with spectra in the NIST database (Figure B.1, Table B.1), the main components being identified as long-chain alcohols and steroidal compounds. Mass spectral identifications were confirmed using extracted ion chromatograms (EIC) to check elution orders. The presence of C_{27} steroidal lipids including 5β-cholestan-3β-ol (coprostanol), 5β-cholestan-3α-ol (epicoprostanol), 5α-cholestan-3β-ol (5α-cholestanol) reflect animal contributions to the soils, while C_{29} steroidal lipids such as β-sitosterol and stigmastanol reflect plant contributions to the soil examined (Bull *et al.* 1999; Prost *et al.* 2017). In the majority of soils, sterols are accompanied by alkanols in the range C_{26} – C_{30}, the latter confirming terrestrial higher plant contributions in the soil extracts (Simpson *et al.* 1998; Bull *et al.* 1999). Successful GC-MS detection of cholestanols in the cesspit sample confirms the potential to use these samples for evaluation of an applicability of FLB as a derivatising agent. The GC-MS analysis shows the presence of various alcohols in the sample, with alkanols being in significantly higher concentrations than the steroids. The high levels of *n*-alkanols confirm the need to re-evaluate the ratio of derivatising reagents required for use with soil extracts to ensure complete derivatisation of steroidal compounds and maximum signal response.

The soil extract from the same cesspit sample was subjected to FLB derivatisation following the procedure outlined in Chapter 2 (see Sections 2.2.1 and 2.3.1). The derivatisation was carried out using various molar ratios of the derivatisation reagents (Table 4.1). The GC-MS chromatogram of the soil extract from the cesspit core (Hungate, York) revealed hexacosanol (C_{26}-alkanol) to be the most abundant organic compound in the sample, hence the mole of soil extract was roughly estimated using the molecular weight of hexacosanol (M_w = 383 g/mol) and described as 'the molar fraction of C_{26} alkanol'. The products from FLB derivatisation were separated by (U)HPLC using multiple detectors (FL, UV and MS) to obtain signal responses from each detector (Figure 4.4, 4.5, 4.7 and 4.8).

Table 4.1. Mole ratios of soil extracts and derivatising agents (FLB, DMAP, EDC).

Mole ratio of Soil extract: FLB: DMAP: EDC	Soil extract (moles)	FLB (moles)	DMAP (moles)	EDC (moles)
Control	1.03×10^{-6}	-	-	-
1: 3: 4: 9	1.03×10^{-6}	3.2×10^{-6}	4.1×10^{-6}	9.4×10^{-6}
1: 6: 4: 9	1.03×10^{-6}	6.2×10^{-6}	4.1×10^{-6}	8.9×10^{-6}
1: 10: 8: 16	7.7×10^{-7}	8.1×10^{-6}	6.6×10^{-6}	1.2×10^{-5}
1: 16: 8: 16	7.7×10^{-7}	1.2×10^{-5}	6.6×10^{-6}	1.2×10^{-5}
1: 25: 12: 50	7.7×10^{-7}	1.9×10^{-5}	9.8×10^{-6}	3.8×10^{-5}
1: 32: 16: 64	7.7×10^{-7}	2.5×10^{-5}	1.2×10^{-5}	5.0×10^{-5}
1: 40: 8: 74	1.03×10^{-6}	4.0×10^{-5}	8.2×10^{-6}	7.6×10^{-5}

4.3.2.1 HPLC-FL and UV detections

The HPLC-FL chromatograms of the soil extract show cop-FLB eluting at a retention time of 15.2 min (peak 1, Figure 4.4a), confirming formation of the FLB derivative. The chromatograms also show a residue from the derivatisation eluting in the retention time window 0-7 min (Figure 4.4a), indicating an excess of the reagent. The peak areas of the FLB derivatives obtained from the (U)HPLC separation of the soil extracts (Figure 4.6a) reveal coprostanol as the dominant

steriodal component while the most abundant alcohol in the soil extract is hexacosanol eluting at 18.2 min. The presence of the two types of hydroxyl components (steroidal alcohols and long-chain alkanols) indicates that FLB derivatisation could offer more selective detection of a range of alcoholic compounds, providing highly selective chromatograms and excellent signal response. Structurally different alcohols may give different yields in the reaction to produce FLB derivatives. Hence, the FL and UV signal responses of coprostanol and hexacosanol were reported by peak area obtained from the HPLC separations to find the optimum derivatising condition (Figure 4.4 and 4.5).

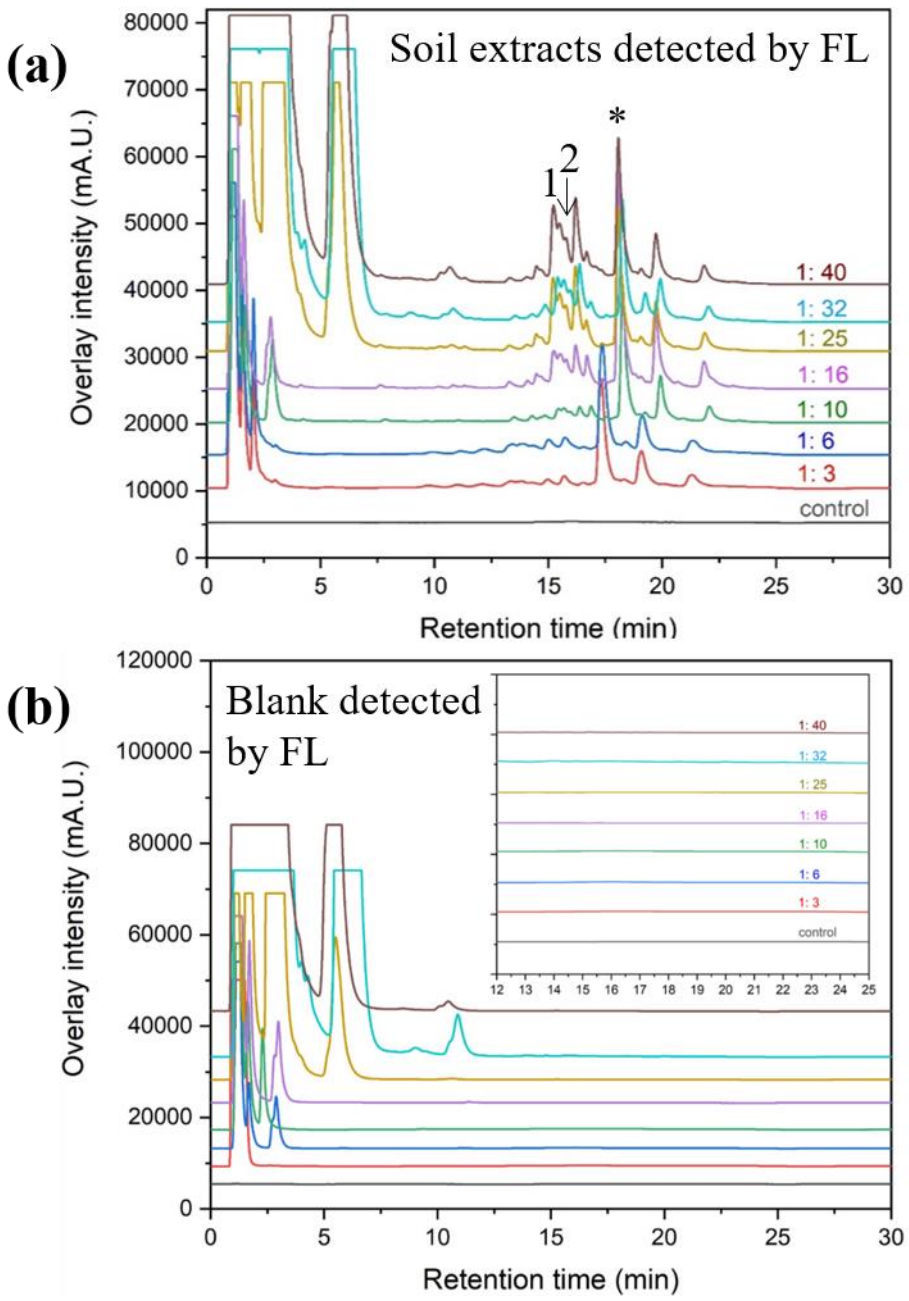


Figure 4.4. Chromatograms of the FLB derivatives of of the medium polar fraction of the soil extract and blank samples prepared using various ratios of derivatising agents (mole ratio as soil extract: FLB) recorded using fluoresence detector: detection of (a) soil extract and (b) blank samples (*n* =3). Assigned peaks: 1 = coprostanol (15.1 min) , 2 = 5α-cholestanol (15.7 min), asterisk (*) = hexacosanol (18.2 min). LC separations were carried out on a Symmetry C_{18} column (3.5 µm, 2.1 mm × 150 mm) (see Section 2.5.2).

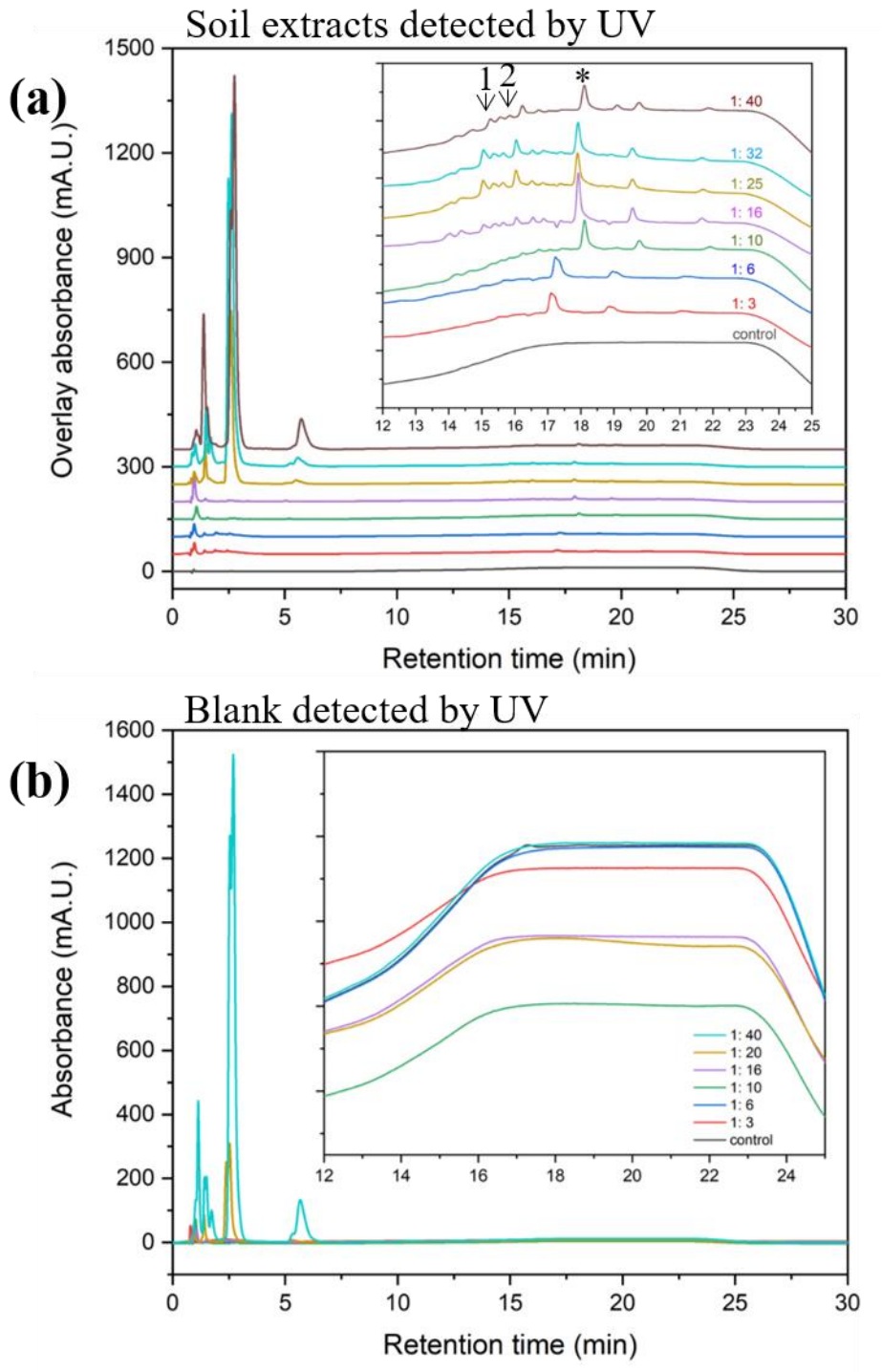


Figure 4.5. Chromatograms of the FLB derivatives of of the medium polar fraction of the soil extract and blank samples prepared from various derivatising ratios (mole ratio as soil extract: FLB) recording using ultra-violet detector: detection of (a) soil extracts and (b) blank samples (*n* =3). Assigned peaks: 1 = coprostanol (15.0 min) , 2 = 5α-cholestanol (15.6 min), asterisk (*) = hexacosanol (18.2 min). LC separations were carried out on a Symmetry C_{18} column (3.5 µm, 2.1 mm × 150 mm) (see Section 2.5.2).

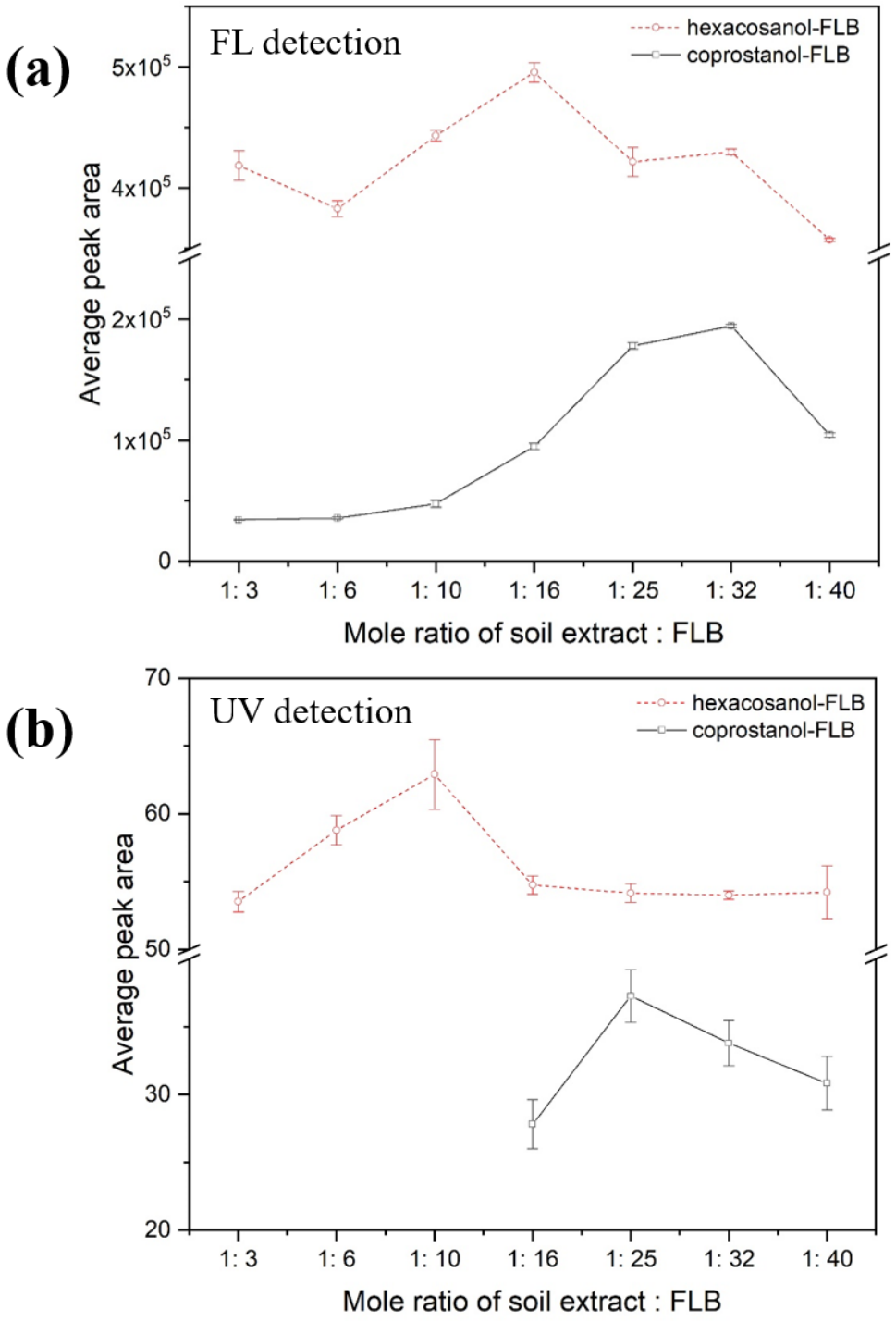


Figure 4.6. Peak areas of coprostanol-FLB and hexacosanol-FLB detected in soil extract prepared using various molar ratios of reactants: (a) fluorescence detection and (b) ultraviolet-visible detection. All soil extracts were prepared from the same sample weight before derivatisation (*n* =3).

Incomplete derivatisation of analytes under the initial reaction conditions is evident from an increase in the overall signal response of the soil extracts obtained from FL and UV detectors when the amount of derivatising agent was increased (Figure 4.4a, 4.5a). Based on the previous results where a mixture of steroid standards eluted in the retention time range 14-15 min, signal response in the chromatograms for steroidal lipids in their FLB forms were highest when the mole ratio of FLB was at least 25 times that of the C_{26} *n*-alkanol. However, a decrease in response of FLB derivatives was observed when the mole ratio reached 40 times. The results indicate that the amount of FLB is a crucial factor in the response for steriodal compounds in soil extracts. By contrast, the amount of derivatising agent appears to have little effect on the response of compounds eluting in the retention time window 18-20 min (Figure 4.4a, 4.5a), possibly suggesting a difference in reactivity of these components, which are predominantly *n*-alkanols.

Peak profiles in the first ten minutes of the chromatograms of the blank samples (Figure 4.4b, 4.5b) matched the chromatograms of soil extracts prepared using the same derivatisation ratios (Figure 4.4a, 4.5a). This reflects a presence of residue from the derivatisation process after sample clean-up. The excess is undesirable as it could lead to saturation of active sites in the column, blockage and thus limit separation efficiency. The sample clean-up step was later modified to address this issue (see Section 3.3.2.3.1).

4.3.2.2 MS detection

To achieve comparability, the HPLC MS signal response for of the medium polar fraction of the soil extract derivatised with the different amounts of FLB reagents (Figure 4.8) was monitored under the same conditions used for FL and UV detection (Figure 4.4, 4.5). Sample blanks, prepared for each different ratio of reagents, were also analysed by HPLC-MS (Figure 4.7).

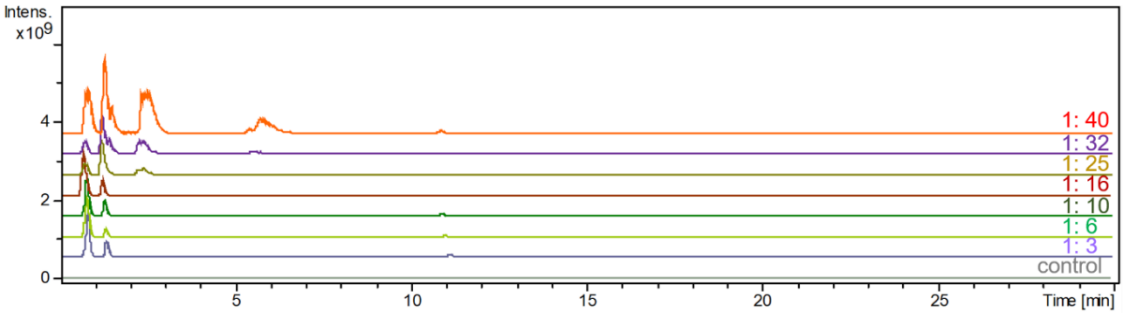


Figure 4.7. APCI MS base peak chromatograms of the FLB derivatives of blank samples prepared from various ratios of derivatising reagents (mole ratio as soil extract: FLB). LC separations were carried out on a Symmetry C_{18} column (3.5 µm, 2.1 mm × 150 mm) (see Section 2.5.2).

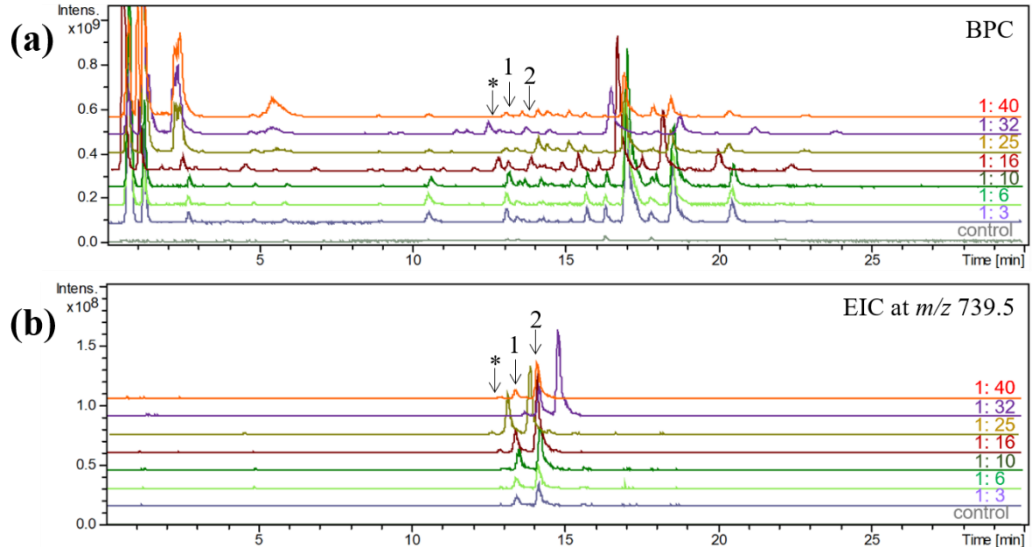


Figure 4.8. APCI MS of the FLB derivatives of the medium polar fraction of the soil extract prepared from various ratios of reagent to extract (mole ratio as soil extract: FLB), (a) base peak chromatograms, (b) extracted ion chromatograms obtained at *m/z* 739.5. Assigned peaks: 1 = coprostanol, 2 = 5α-cholestanol, * = epicoprostanol. LC separations were carried out on a Symmetry C_{18} column (3.5 µm, 2.1 mm × 150 mm) (see Section 2.5.2).

Screening of the C_{27} stanols (cop-FLB and 5α-chol-FLB) was carried out by generating extracted ion chromatograms at *m/z* 739.5 (Figure 4.8). Mass spectra for each peak in the EIC were compared with those obtained from the corresponding standard FLB derivatives to confirm the presence of the analytes (data not shown). The prominent peaks obtained from EIC at *m/z* 739.5 (Figure 4.8b) clearly show the presence of C_{27} stanols in FLB form in the soil extract, reflecting changes in the signal response when changing the proportion of derivatising agent to soil extract. The signal response was greater when the amount of FLB was increased, reaching a plateau at 25 times, before slightly dropped when the ratio of soil extract: FLB was 1:40, as previously observed from HPLC-UV and HPLC-FL. A decrease in response for cop-FLB with further increase in FLB matched that observed for chol-FLB (see Figure 3.2b and 3.2c), most likely reflecting the formation of an unwanted *N*-acylisourea intermediate leading to quenching of the reaction and reduction in yield of ester (Joullie and Lassen, 2010; Tsakos *et al.* 2015).

In order to examine signatures of human activity recorded by the sample, peaks of cop-FLB and 5α-chol-FLB were identified by matching the retention times in the soil extract with those obtained from the FLB derivatives of standard sterols (see Chapter 3, Section 3.3.2.5.1): elution order coprostanol, cholesterol and 5α-cholestanol. The mass spectra of each component perfectly matched those obtained from the FLB derivatives of standard C_{27} steroidal lipids (see Figure 3.22). The coprostanol ratio was evaluated by determining peak areas of coprostanol, epicoprostanol and 5α-cholestanol to indicate feacal human contamination, while the peak area of cholesterol was also determined to identify the possibility of sources due to decompostion of body tissues. The ion chromatograms at *m/z* 739.5, the base peak in cholestanol-FLB dissocation pathways, were extracted to confirm the presence of cholestanols (Figure 4.8b). The prominent signal for cop-FLB (Figure 4.8b, peak 1), is consistent with the soil extract used in this demonstration being from a cesspit core that has been reported to contain abudant coprostanol (*cf* Prost *et al.* 2017). A small amount of 5α-chol-FLB was also detected (Figure 4.8b, peak 2), meaning that the coprostanol ratio can be estimated. The peak marked * (Figure 4.8b) was identified as epicoprostanol based on its MS spectrum exactly matching that of coprostanol and its response relative to coprostanol being similar to that obtained from GC-MS (Figure 4.1). The

amount of each of the other C_{27} steriodal lipids was measured based on their UHPLC peak area using external standard calibration (see Section 3.3.2.3.5.1) with epicoprostanol being estimated using the calibration curve of coprostanol owing to the similarity of their structures. The coprostanol ratio (Figure 4.9) was estimated from the FLB derivatives for each of the different ratios of derivatising agents (Table 4.1). All of the different ratios or reagents gave coprostanol ratio values above 0.7, matching the criteria to indicate human activity (Bull *et al.* 1999). Small increases in the value of the coprostanol ratio of the soil extract were evident as the amount of FLB for derivatisation was increased, with a plateau for values of the FLB ratio of 1: 25 and above, reflecting there is some difference in reactivity between components used in the ratio until completion of FLB derivatisation. In order to obtain the maximum signal response for soil samples with minimal FLB residue, the suggested ratio of derivatising agents is 25: 12: 50 (FLB: DMAP: EDC to 1 mole of sample). The suggested ratio of the derivatising reagents was further used to prepare the FLB derivatives of sterols in various other archaeological contexts (Section. 4.3.3.1).

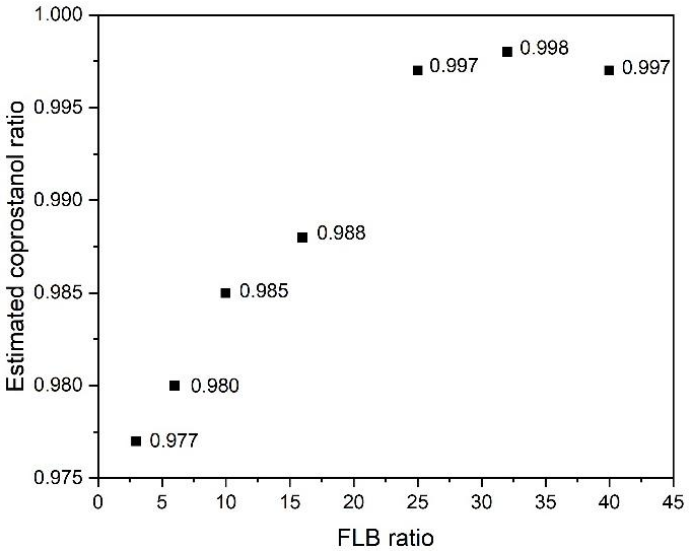


Figure 4.9. Estimated coprostanol ratio of the soil extract from cess pit core (Hungate 51151, York) derivatised with FLB in different proportions to the soil extract.

4.3.3 Quantification of cholesterol and cholestanols in archaeological contexts (compared with GC-MS result)

4.3.3.1 Archaeological samples

Determination of cholesterol and cholestanols in archaeological soil samples from Hungate, York was carried out using the refined separation condition (Section 2.5.1). Samples from cesspit 1 were obtained from fourteen levels and samples from a Roman grave (context number C51364) and a closely associated cesspit, cesspit 2, comprised grave fill at different anatomical locations around the human remains and a sample from within the cesspit. The FLB derivatised soil extracts were separated by UHPLC with MS detection to obtain the base peak chromatogram of the FLB derivatives (Figure 4.11). Analysis of trimethylsilylether (TMS) derivatised extracts by GC-MS was also carried out to identify the steroidal components in their TMS form. The total ion chromatograms of the TMS derivatised extracts show complex distributions with many closely eluting components within the region where the C_{27} steroidal components elute (40-42 min, Figures 4.2 and B.1). Successful detection of cholesterol and cholestanols in the grave soil and cess pit samples confirms the potential to use these samples for evaluating the FLB derivatising approach for steriodal alcohols analysis.

4.3.3.1.1 Cesspit 1: Anglo-Scandinavian cesspit (Hungate, York)

A core from an archaeological cesspit, interpreted as Anglo-Scandinavian in age, was collected from Hungate, York. The core represents a depth profile of around 40 cm and shows three distinct zones distinguished by colour (Figure 4.10). Zone 1, 0-9 cm depth, is grey-brown soil with light brown and grey patches at 4.5 to 6 cm and at 7-8 cm, respectively. Zone 2, 9 – 29 cm depth, is dark brown soil along with slanted laminations and stones in the range of 20-28 cm. Zone 3, 29-40 cm depth, is very dark brown soil with a high-water content. Two yellow brown patches with sandy material occur between 33-34 cm and 36-40 cm.

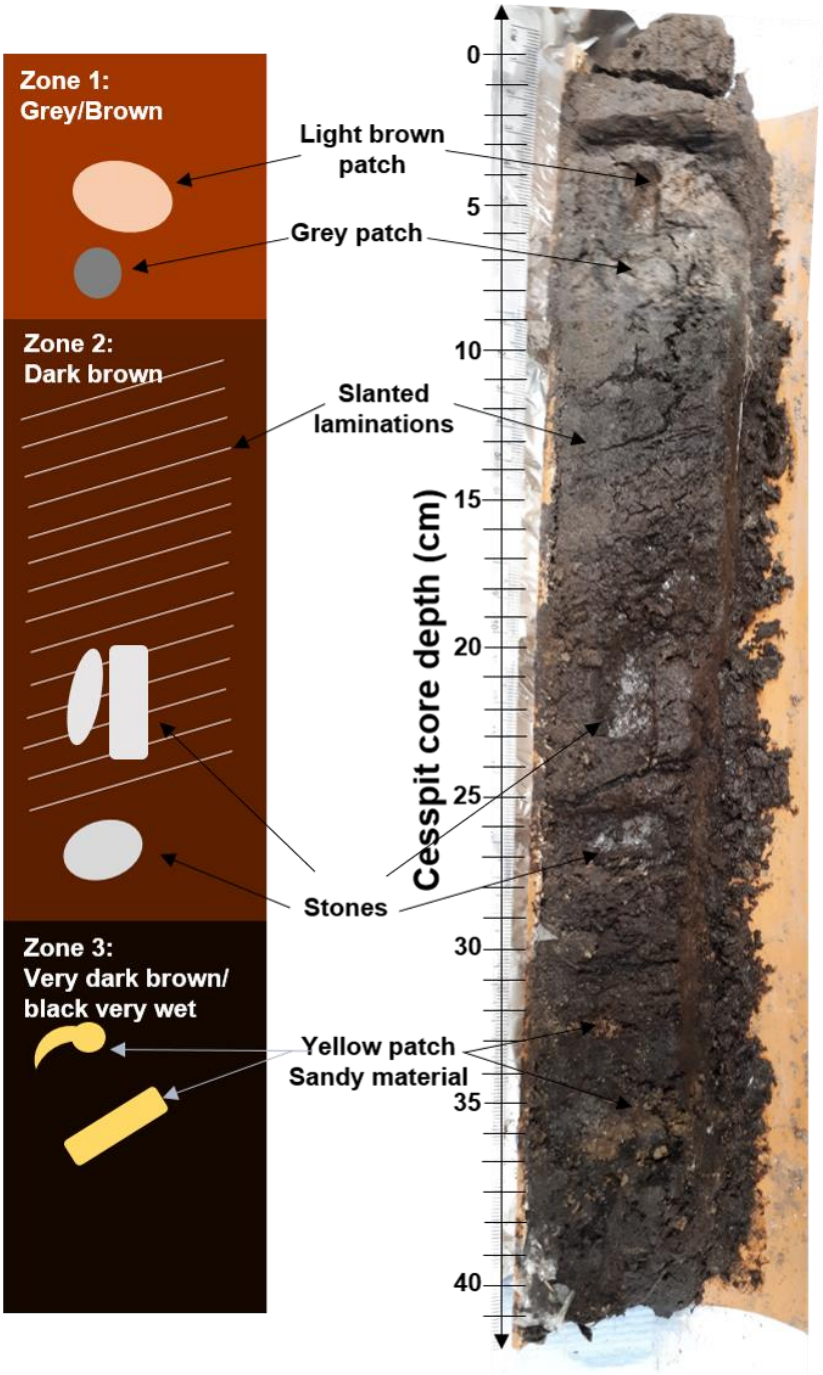


Figure 4.10. Diagrammatic representation of the nature of the soil in the Anglo-Scandinavian cesspit core collected from Hungate, York together with a photograph showing the three zones.

4.3.3.1.1.1 Signatures of cholesterol and related stanols to assess signatures of human activity

Determination of cholesterol and related stanols as the FLB derivatives (see Section 2.5.1) was carried out on all fourteen samples from the Anglo-Scandinavian cesspit core. The samples were from depths (1) 0-1 cm, (2) 4-5 cm, (3) 8-9 cm, (4) 9-10 cm, (5) 13-14 cm, (6) 15-16 cm, (7) 16-17 cm, (8) 24-25 cm, (9) 25-26 cm, (10) 27-28 cm, (11) 28-29 cm, (12) 29-30 cm, (13) 30-31 cm and (14) 39-40 cm. The FLB derivatised soil extracts were separated by UHPLC with MS detection to obtain the base peak chromatogram of the FLB derivatives (Figure 4.11).

Successful detection of coprostanol, cholesterol, and 5α-cholestanol at retention times of 9.4, 9.9 and 10.4 min in the soil extracts was evident from extracted ion chromatograms (*m/z* 737.5 and *m/z* 739.9) (Figure 4.11) and mass spectra (Figure B.2). The retention times show close matches to the corresponding standards analysed as a mixture (Figure 3.25). The mass spectra obtained from the samples were cross-checked with those obtained from the standard mixture, the good matches confirming the presence of cholesterol and cholestanols (data not shown). Epicoprostanol was identified as giving a prominent peak at retention time 8.7 min, base peak ion at *m/z* 739.9 and the same fragmentation pattern as that obtained from coprostanol. Its abundance relative to coprostanol was similar to that observed by GC-MS (Figures 4.2 and B.1 , Table B.1).

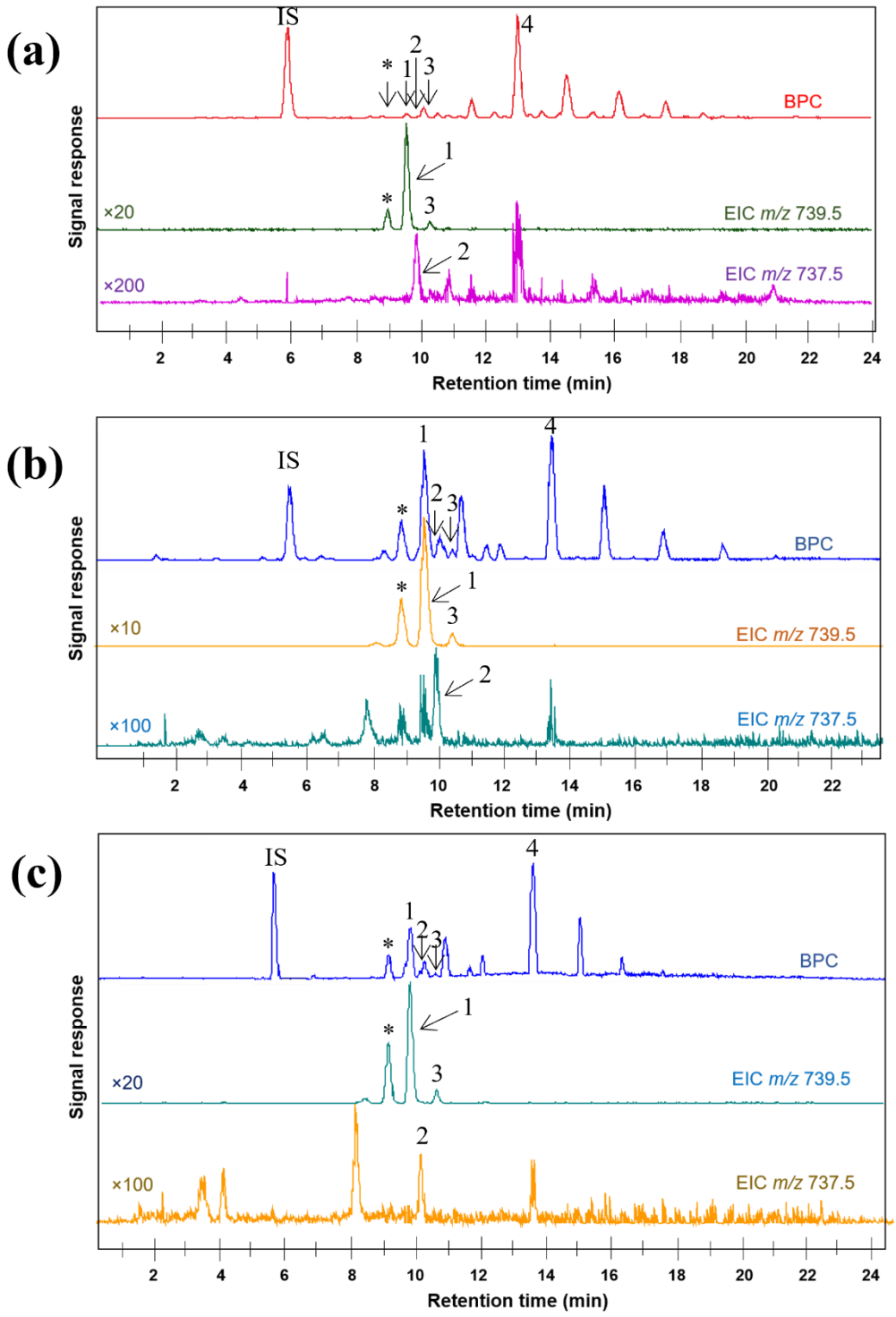


Figure 4.11. APCI MS chromatograms of the FLB derivatives of the medium polar fraction of the soil extracts from different depths of cesspit: (a) 0-1 cm, (b) 28-29 cm, (c) 30-31 cm. Base peak chromatogram and extracted ion chromatograms obtained from *m/z* 737.5, *m/z* 739.5. Assigned peaks: * = epicoprostanol, 1 = coprostanol, 2 = cholesterol, 3 = 5α-cholestanol, 4 = 1-hexacosanol, IS = 1-heptadecanol as an internal standard.

The presence of cholesterol and cholestanols in all of the soil extracts reflects contributions from human activity, consistent with the samples being from a cesspit (Figure 4.12 and Table 4.2). In the majority of samples, coprostanol is the major component, accompanied by smaller amounts of other forms of cholestanols (epicoprostanol, 5α-cholestanol) and very low levels of cholesterol. The dominance of coprostanol reflects inputs of faecal matter over all of the time period represented. The profiles of the C_{27} steriodal lipids reveal a significant increase of cholestanols (coprostanol, epicoprostanol and 5α-cholestanol) and a minor change in cholesterol in the lower parts of the core (Figure 4.12). The amount of coprostanol at the top layers of the core (0-1 cm and 4-5 cm) was in the range of nanograms per gram of dried soil, rising to microgram per gram levels at the boundary with zone 2 of the core (8-9 cm). Overall, the amount of coprostanol in the layers deeper than 8-9 cm increased throughout the core with fluctuations in some layers (16-17 cm and 25-26 cm). The general increase in the concentrations of the steroidal components with depth could be mainly attributable either to leaching of fluids during the decay of the contents of the pit or to changes in its use. Given the irregularity in the changes in concentration with depth and the differences between the sterols and stanols, the latter appears more likely. This interpretation is also supported by total organic carbon contents of the soils, which show similar depth profiles to the steroidal components (Stubbs, M., unpublished MChem project report, University of York). Thus, the coprostanol profile possibly reflects a change in inputs of faecal matter over the period the cesspit was in use with the main activity being associated with the levels below 24 cm depth.

Table 4.2. FLB derivatives of cholesterol and its reduced products detected from soil extracts obtained from the suspected Anglo-Scandinavian cesspit (Hungate, York), *n* =3.

Cesspit core depth (cm)	Cholesterol and cholestanols ($\times 10^{-9}$ gram per gram of dried soil, ng/g dw)			
	coprostanol	epicoprostanol	5α-cholestanol	cholesterol
Zone 1: Light brown				
0-1	28 ± 1	2.45 ± 0.03	N.D.	4.5 ± 0.2
4-5	240 ± 8	30 ± 0.4	N.D.	40 ± 0.2
8-9	2,660 ± 65	180 ± 7	85 ± 3	74 ± 2
Zone 2: Dark brown				
9-10	2,150 ± 90	160 ± 10	87 ± 0.4	19 ± 0.1
13-14	3,160 ± 60	1,460 ± 20	430 ± 20	20 ± 0.3
15-16	5,990 ± 230	1,820 ± 60	140 ± 3	43 ± 1
16-17	3,150 ± 30	750 ± 8	220 ± 7	170 ± 5
24-25	26,900 ± 1,210	6,770 ± 290	1,630 ± 80	100 ± 4
25-26	2,960 ± 90	290 ± 10	300 ± 10	54 ± 2
27-28	26,100 ± 1,100	7,370 ± 240	320 ± 10	130 ± 4
28-29	20,600 ± 550	9,300 ± 430	1,120± 20	270 ± 10
Zone 3: Very dark brown				
29-30	25,900 ± 800	5,510 ± 250	1,400 ± 60	110 ± 5
30-31	47,700 ± 2,040	8,210 ± 140	3,210 ± 160	120 ± 6
39-40	26,600 ± 1,130	6,740 ± 180	770 ± 30	N.D.

N.D. = Non-determined

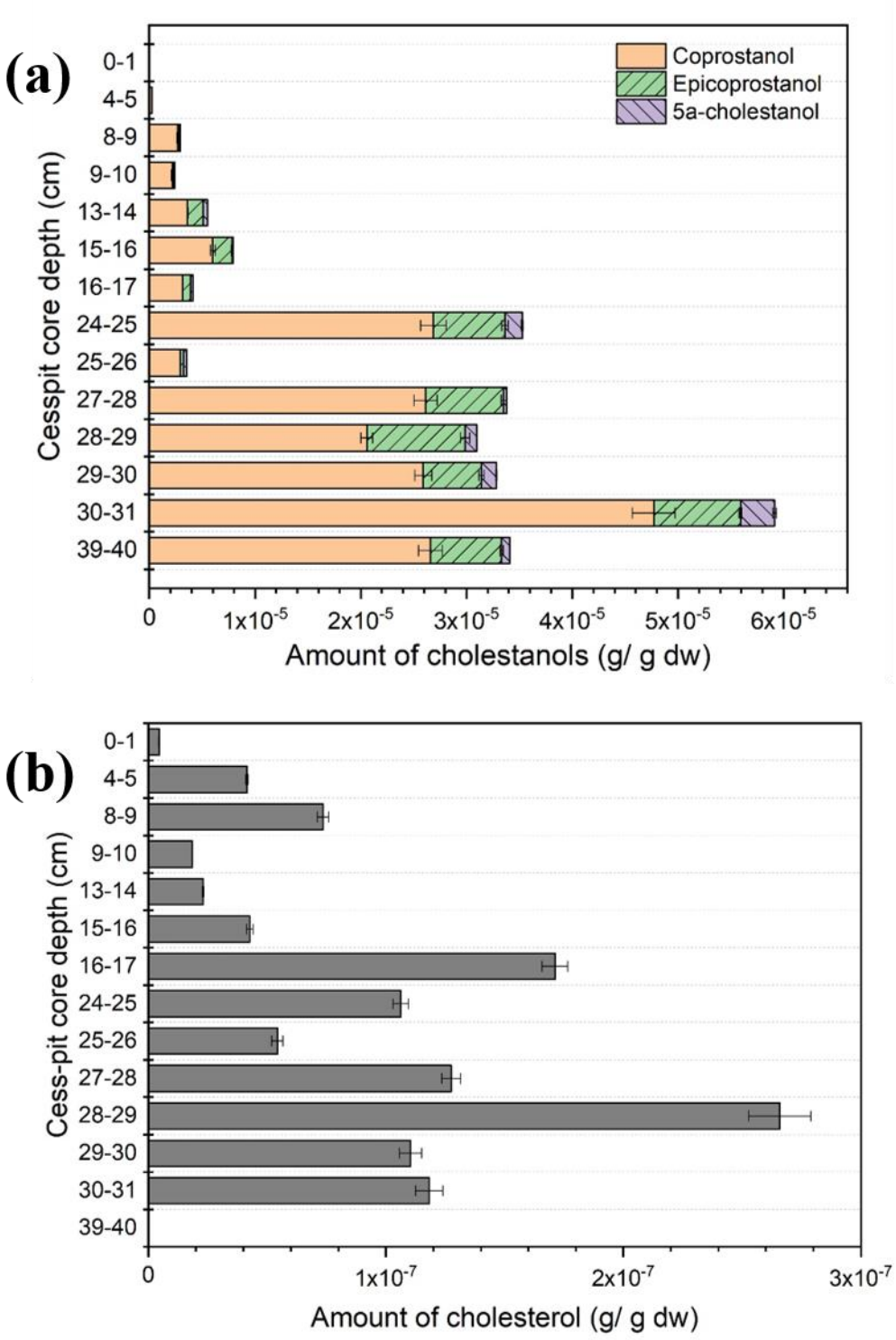


Figure 4.12. The measured amounts of C_{27} steroidal lipids detected from soil extracts obtained from different depths of the suspected Anglo-Scandinavian cesspit (Hungate, York) (a) C_{27} stanols (coprostanol, epicoprostanol and 5α-cholestanol), (b) cholesterol, n =3.

Based on the successful UHPLC-MS identification of cholesterol and its transformation products and the deconvolution of signals using extracted ion chromatograms to identify the presence of cholestanols and cholesterol, the ratio of 5β/(5β+5α) cholestanols was estimated for the different levels. The fourteen different depths all gave coprostanol ratios >0.9, indicating feacal contamination (Figure 4.12a). The conventional method for detection of steriodal lipids by GC-MS was run for all same samples to enable comparison of the results with FLB derivatisation. In order to identify all cholestanol peaks in this study, double the amount of the medium polar fraction of the same soil extract was required for GC-MS analysis compared with that used for FLB derivatisation,. This reflects that the steroidal lipid concentrations were close to the detection limit of the conventional method; it also demonstrates that FLB derivatisation can refine the signals of these biomarkers in soils owing to the smaller sample requirements. The values matched those obtained from the same soil samples by GC-MS of the TMS derivatives, reflecting good correspondence between the FLB derivatives and the conventional TMS derivatives (Figure 4.13a). Coprostanol can be produced by all mammals; higher values of the coprostanol ratio (>0.9) have previously been reported from the faeces of humans, pigs and dogs (Prost *et al.* 2017; Harrault *et al.* 2019). In particular, the coprostanol ratios from the Anglo-Scandinavian cesspit reveal differences in the organic residues recovered throughout the full depth sequence. Coprostanol was the dominant component among the C_{27} steroidal lipids (coprostanol ratio >0.9), confirming that the site was used as a cesspit in which faecal matter of higher mammals such as human and pigs was deposited (Prost *et al.* 2017; Harrault *et al.* 2019). Notably, a decrease in the concentrations of C_{27} steroidal lipids in the upper layers reflects a change in the use of the site over time. The amount of coprostanol was in range of ppm level (2.1 – 47.7 μg/g) in the deeper levels and ppb (ng/g) levels with an absence of 5α-cholestanol at the upper layers of the core. These differences reflect the deposition of much lower concentrations of faecal matter in the cesspit during the later phases of its use. Notably, however, the similarities in the coprostanol ratio at all depths suggests the nature of the material depositied in the cesspit was similar over the period of its use.

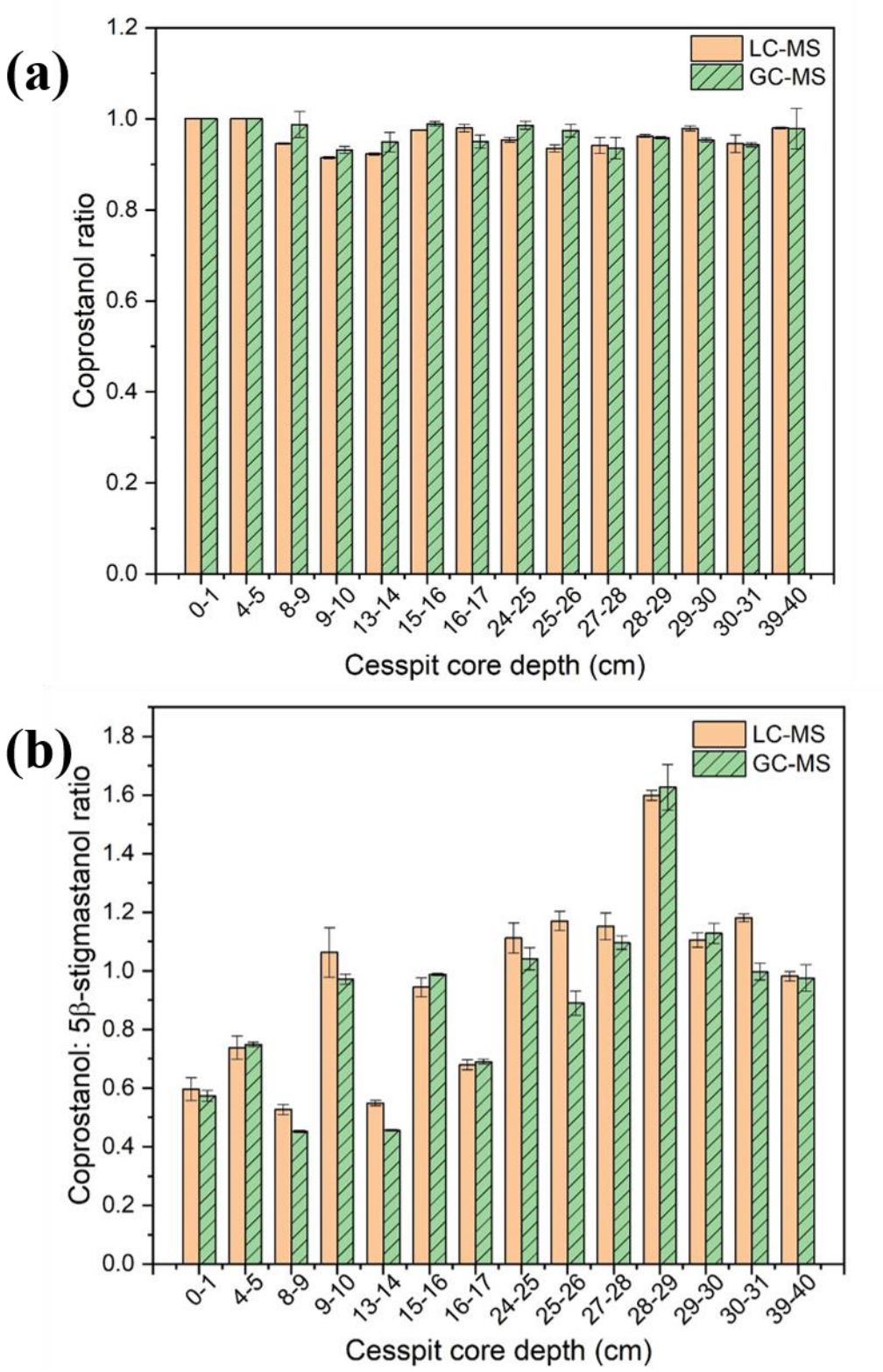


Figure 4.13. The measured C_{27} steriodal lipids and 5β-stigmastanol from soil extracts obtained from cesspit core at different depths (a) coprostanol ratio, (b) coprostanol: 5β-stigmastanol ratio, *n*=3.

Cholesterol is the predominant sterol in most animal tissue, while β-sitosterol is a typical sterol of plants (Birk *et al.* 2011; Prost *et al.* 2017). Conversion of cholesterol and β-sitosterol by mammalian gut microflora produces coprostanol and 5β-stigmastanol, respectively. The uniqueness of 5β-stanols produced in the mammalian gut has enabled their use to reflect the nature of foodstuffs ingested from their profiles in faeces or in soils, sometimes enabling distinction of different mammalian feeding behaviours (herbivore, omnivore, and carnivore) (Sistiaga *et al.* 2014; Guillemot *et al.* 2017; Prost *et al.* 2017; Harrault *et al.* 2019). For example, faeces of herbivores or vegetarian humans contain higher concentrations of 5β-stigmastanol than coprostanol while non-vegetarian human faeces contain a dominant amount of coprostanol (Evershed *et al.* 1997; Sistiaga *et al.* 2015; Guillemot *et al.* 2017; Prost *et al.* 2018). This led to the establishment of another criteria for use in archaeological sites to indicate human occupation: a ratio of coprostanol to 5β-stigmastanol over 1.5 (Bethell *et al.* 1994). The β-sitosterol and its related stanols were identified in the same manner as the C_{27} steroidal lipids (Figure 4.16 and Figure B.6). The ratio of coprostanol to 5β-stigmastanol of all fourteen cesspit core samples was assessed in order to evaluate further the nature of inputs in the suspected cesspit core (Figure 4.13b). A high coprostanol: β-stigmastanol ratio (>1.5) was obtained from only one soil extract (28-29 cm) from the cesspit core, confirming that the dominant input was from omnivores with only one level that may potentially have been used as a cesspit for human faecal material. Low coprostanol: β-stigmastanol ratios (<1.5) were obtained from the rest of the soil layers from the cesspit core, reflecting that the dominant input was from herbivores and carnivores. Thus, the results imply that the cesspit was used for deposition of faecal matter from various origins (human, higher mammals and other animals). The coprostanol: β-stigmastanol ratios determined by LC-MS all matched the values obtained from the same soil samples by GC-MS of the TMS derivatives.

Comparison of the signal responses for cholesterol and its transformation products obtained from the soil extracts by UHPLC-MS and GC-MS shows that the FLB derivatives provide a substantial improvement in response compared with the conventional TMS derivatives (Figures 4.2, 4.11 and B.1). In the same soil extract, the FLB derivatisation generates a chromatogram with lower noise

than those obtained from TMS derivatising agent. Furthermore, the approach of FLB derivatisation was performed using around half of the amount of the soil extract than for silylation. The cholestanol-FLB derivatives gave a calculated limit of detection, based on $3 \times S/N$ (Miller and Miller 2010; Christian *et al.* 2014), of 0.05 pmol on column compared with 1 pmol on column for the TMS derivatives. Relative standard deviation (%RSD, n=3) for cholestanols-FLB detections were 5.6% lower than those from TMS derivatives (12.3%), reflecting greater precision in quantification. Thus, the method provides an opportunity to overcome several limitations of organic matter detection in archeological samples with the greater reproducibility offering a greater security to interpretations. In cases where archeaological materials contain steroidal compounds close to the detection limit of MS, the use of the FLB derivatising agent can enable identification and quantification of steriodal alcohols.

4.3.3.1.1.2 Comparison of the concentration estimates of C_{27} steroids obtained by three different detectors (MS, UV-vis, FL)

The FMOC group of the FLB derivatives acts both as a chromophore and a fluorophore, enabling detection in HPLC by UV-vis and fluorescence detection. Determination of cholesterol and cholestanols from the fourteen cesspit core samples enabled successful detection of epicoprostanol, coprostanol, cholesterol, and 5α-cholestanol at retention times of 8.5, 9.3, 9.9 and 10.4 min in the soil extracts from the UHPLC-UV (Figure 4.14a) and at retention times of 11.1, 13.2, 13.9 and 14.7 min in the HPLC-FL chromatograms, the latter being run using HPLC owing to the lack of availability of a FL detector compatible with UHPLC (Figure 4.14b). The amounts of cholesterol and related stanols were evaluated in the same manner with those obtained from MS detection (Table 4.3 and Table B.2-B.5).

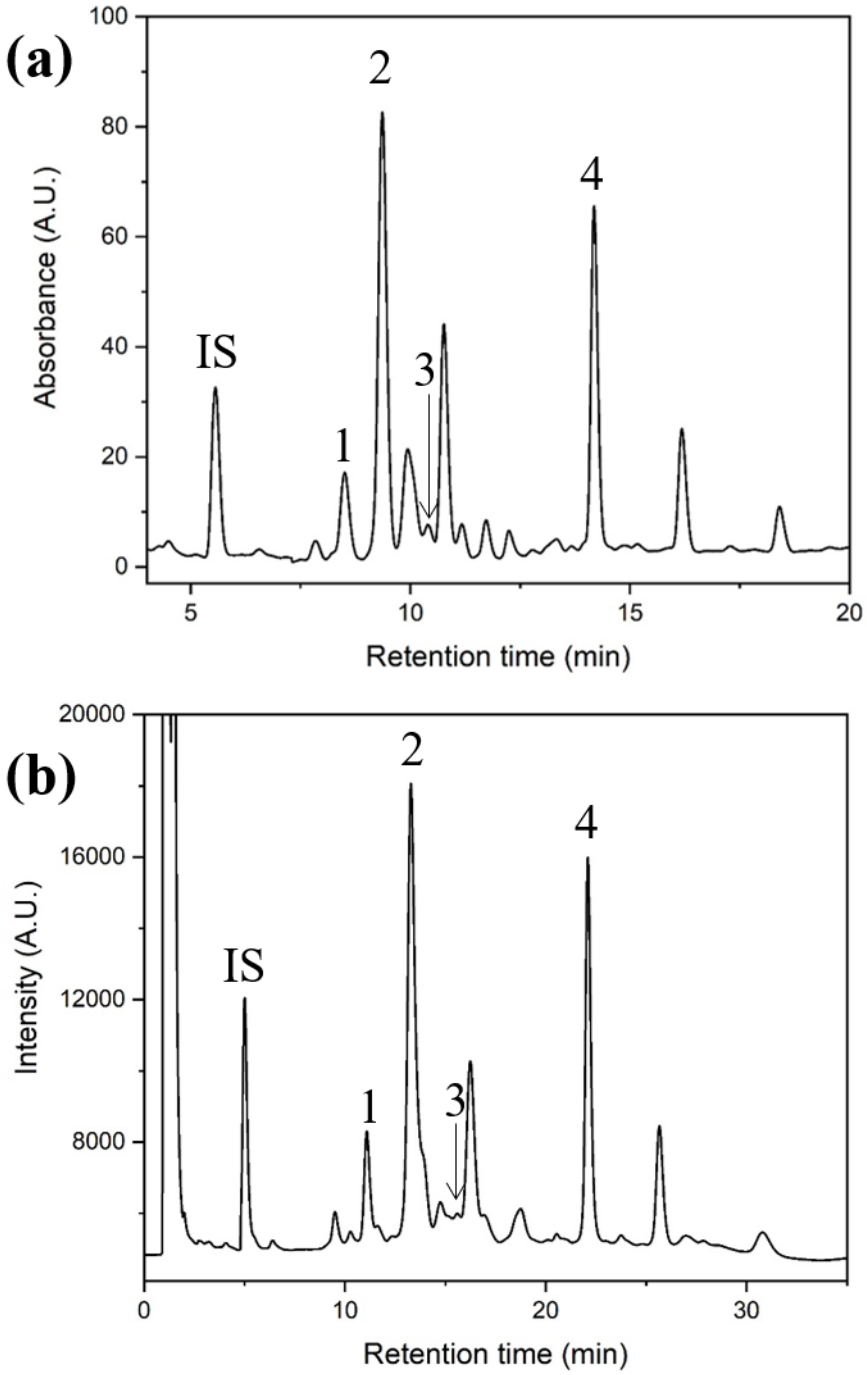


Figure 4.14. Partial chromatograms of the FLB derivatives the medium polar fraction of the soil extract obtained from cesspit core depth at 28-29 cm recording using different detectors: (a) UV and (b) FL (*n* =3). Assigned peaks: 1 = epicoprostanol (11.2 min for FL), 2 = coprostanol (9.3 min for UV, 13.7 min for FL) , 3 = 5α-cholestanol (10.4 min for UV, epicoprostanol (11.2 min for UV, 15.1 min for FL), 4 = 1-hexacosanol (11.2 min for FL). IS: 1-heptadecanol (5.5 min for UV, 5 min for FL).

Comparison of signal responses by UHPLC-MS, UHPLC-UV and HPLC-FL reveals that quantification of cholesterol and related stanols by UV and FL detectors provides comparable estimations of the amounts of the components compared with conventional MS detection for cholesterol and its transform products (Figure 4.15, Table B.2-B.5). In the same soil extract, the amounts of C_{27} steroidal lipids estimated by UV detection gave similar results to those obtained from the MS detector (Figure 4.11 and Figure 4.14a). By contrast, the concentrations obtained using the FL detector were slightly lower, most likely a result of the chromatographic system operating at lower flow rate and not completely resolving some peaks, e.g. coprostanol and cholesterol, to baseline (Figure 4.14b). Based on the results described in Chapter 3, Section 3.3.2.3.5.1, it is clear that FLB derivatives of steriodal components detected by fluorescence offer the better signal response than that obtained from detection by UV and MS detection by about 1,000 times, reflecting the potential to use FL detection as an alternative approach to detect trace amounts of the components. Thus, the method provides an opportunity to determine steroidal lipids in cases where archaeological materials contain steroidal compounds close to detection limit of MS, and quantification by FL detector can provide a reliable alternative.

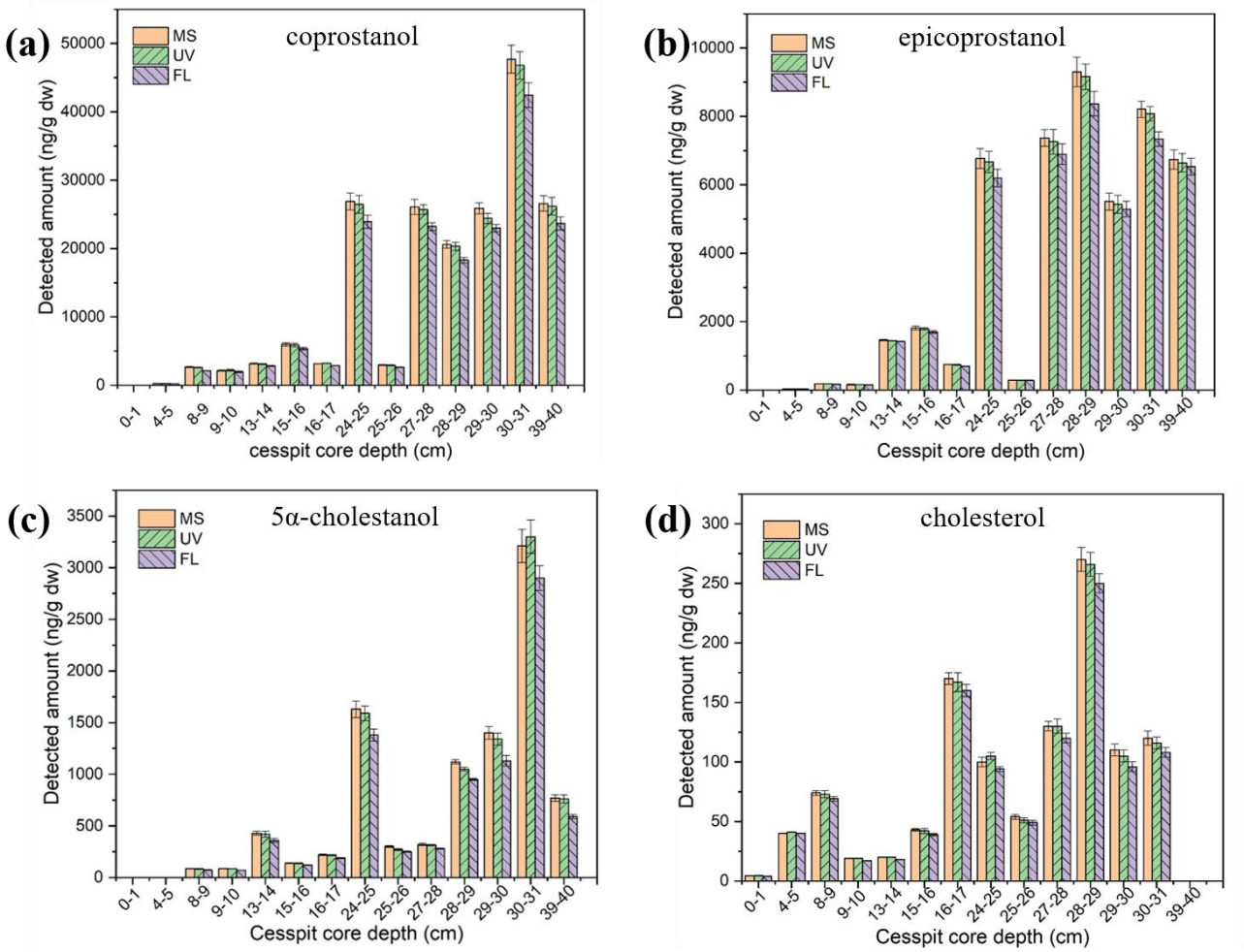


Figure 4.15. The measured C_{27} steriodal lipids from soil extracts obtained from cesspit core at different depths: (a) coprostanol, (b) epicoprostanol, (c) 5α-cholestanol, (d) cholesterol. The signal responses of the FLB derivatives of each sample were detected by three different detectors (MS, UV, FL), *n*=3.

4.3.3.1.1.3 Identification of other alcoholic components

The mass spectral fragmentation of the FLB derivatives can reveal the presence of other alcohols based on three observations: (1) a prominent fragment ion arising from the specific neutral loss from the molecular ion of the entire Boc group (-100 Da), (2) a minor losses arising from cleavage of the tert-butyl group (-56 Da), (3) a fragment ion corresponding to FLB (*m/z* 369.1). This information was used to identify and quantify other hydroxyl compounds in the medium polar fraction of the soil extracts from cesspit core as FLB derivatives without the need for standards of those compounds. Thus, identifications were based on full APCI MS and MS^2 spectra obtained from each peak in the LC-MS chromatograms (see the MS spectra in Figure B.3-B.6). A total of 33 components were identified with a further 5 unidentified alcohols being observed (Figure 4.16, Table 4.3). The massses of the underivatised compounds were cross-checked with the identifications obtained from GC-MS (Table B.1). For example, identification of 5β-stigmastanol was carried out based on the fragmentation pathways for FLB derivatives and confirmed from the GC-MS intepretation.

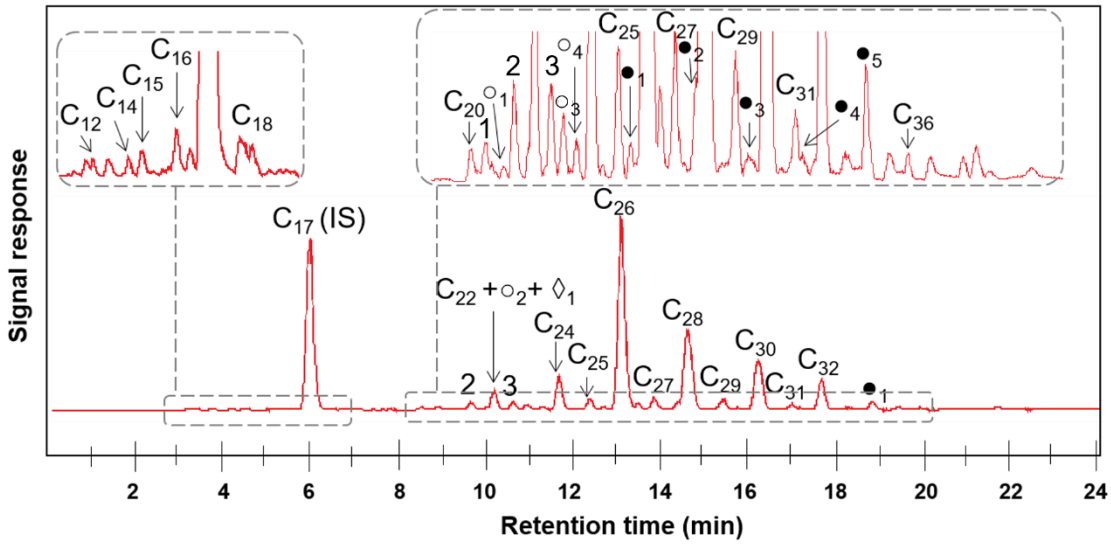


Figure 4.16. APCI MS chromatograms of the FLB derivative of the medium polar fraction of the soil extract obtained from cesspit core depth at 0-1 cm. Assigned peaks: 1 = epicoprostanol, 2 = coprostanol, 3 = 5α-cholestanol, C_n = *n*-alkanols, $○_1$ = β-tocopherol, $○_2$ = γ –tocopherol, $○_3$ = 5β-stigmastan-3β-ol, $○_4$ = 5α-stigmastan-3β-ol, $◊_1$ = β-amyrin, $●_n$ = unknown FLB derivatives.

The LC-MS spectra of the FLB derivatised the medium polar fraction of the soil extracts show *n*-alkanols, in the range C_{18} – C_{30}, as the most abundant compounds. The highest response was from hexacosanol with even-numbered alcohols in higher relative abundance than the odd numbered components, reflecting inputs from plant sources (Bull *et al.* 1999). An attempt to identify peaks having low signal response in the MS chromatograms revealed the presence of the FLB derivatives of β-sitosterol with its reduced counterparts, 5β-stigmastanol and 5α-stigmastanol, matching with those from TMS derivatives. FLB-derivatives of C_{30} sterols, triterpanols, tocopherols and amyrins were also identified. Considering the MS chromatograms of the same soil extracts obtained from LC and GC separations, GC-MS chromatograms comprise peaks for TMS ethers of hydroxyl components together with any other compounds that can be ionised. By contrast, the LC-MS chromatograms of the FLB derivatives give responses only for hydroxyl components, reflecting the potential for selective detection of alcohols including steroidal lipids and eliminating the impact of overlapping components as occurs in GC-MS chromatograms. Notably, the other *n*-alkanols, in the range C_{12} – C_{16} and C_{31},C_{32} and C_{36}, were identified by LC-MS, while these components were not detected by GC-MS, evidencing the use of FLB as a derivatising agent to both enhance signal response and to recognise the presnce of undiscovered hydroxyl components.

Based on the hydroxyl compounds identified from the LC-MS spectra, the peaks observed in the chromatograms from UV and FL detection were labelled to demonstrate the extended applicability of the FLB derivatives (Figure 4.14). The components other than the C_{27} steriodal lipids were not quantified from FL and UV chromatograms, as this was beyond the scope of the project. Overall, the results indicate the potential for the use of this derivatising agent as alternative approach to analyse a wide range of alcohols in the organic matter of soils and sediments by MS, UV and FL detection.

Table 4.3. List of FLB derivative of alcoholic compounds identified from the suspected Anglo-Scandinavian cesspit (Hungate, York).

Compund identified, labelled peak	RT (min) observed by LC-MS	Molecular mass of component in different forms (g/mol)		Detected by GC-MS
		FLB form	Native form	
C_{27} steroiodal lipids				
Epicoprostanol, 1	9.0	839.7	388.7	✓
Coprostanol, 2	9.3	839.7	388.7	✓
Cholesterol	9.9	837.5	386.7	✓
5α-cholestanol, 3	10.4	839.7	388.7	✓
C_{29} steroiodal lipids				
5β-Stigmastan-3β-ol, $\circ_3$	11.0	867.7	416.7	✓
β-sitosterol, 4	11.3	865.7	414.7	✓
5α-Stigmastan-3β-ol, $\circ_4$	11.5	867.7	416.7	✓
C_{30} steroiodal lipids				
3β-lanost-9(11), 24-dien-3-ol	11.2	877.5	426	✓
Lanost-8-en-3-ol (3β)	12.7	879.5	428	✓
Triterpene				
β-Amyrin, $\lozenge_1$	10.2	877.6	427.1	✓
α-Amyrin	N.D.	877.6	427.1	✓
Lupeol	N.D.	877.5	426.7	✓
Alkanols				
1-dodecanol, C_{12}	3.2	637.5	186.9	N.D.
1-tetradecanol, C_{14}	3.9	665.6	214.6	N.D.
1-pentadecanol, C_{15}	4.2	679.6	228.6	N.D.
1-hexadecanol, C_{16}	4.7	693.6	243.0	N.D.
1-Heptadecanol, C_{17} (IS)	5.9	707.6	257.0	✓
1-octadecanol, C_{18}	7	721.6	270.6	N.D
1-icosanol, C_{20}	8.3	749.6	298.6	N.D.
1-docosanol, C_{22}	10.2	777.7	326.6	✓
1-tetracosanol, C_{24}	11.7	805.7	354.7	✓
1-pentacosanol, C_{25}	12.5	819.8	368.7	✓
1-hexacosanol, C_{26}	13.2	833.5	382.6	✓

Compund identified, labelled peak	RT (min) observed by LC-MS	Molecular mass of component in different forms (g/mol)		Detected by GC-MS
		FLB form	Native form	
1-heptacosanol, C_{27}	13.8	847.8	396.7	✓
1-octacosanol, C_{28}	14.7	861.8	410.7	✓
1-nonacosanol, C_{29}	15.5	875.8	424.8	✓
1-triacosanol, C_{30}	16.3	889.9	438.8	✓
1-hentriacontanol, C_{31}	17.1	903.8	451.2	N.D.
1-dotriacontanol, C_{32}	17.7	917.9	467.3	N.D.
1-hexatriacontanol, C_{36}	19.6	974.9	523.0	N.D.
Tocotrienols				
γ-tocopherol, $\circ_2$	10.2	876.5	416.5	✓
α-tocopherol	N.D.	881.2	430.7	✓
β-tocopherol, $\circ_1$	8.5	867.5	416.5	✓
Unknown FLB compounds				
Unknown 1, $\bullet_1$	12.8	907.8	456.46	N.D.
Unknown 2, $\bullet_2$	14.5	963.9	512.84	N.D.
Unknown 3, $\bullet_3$	15.7	991.9	540.46	N.D.
Unknown 4, $\bullet_4$	17.3	1020.9	570.38	N.D.
Unknown 5, $\bullet_5$	18.9	1104.1	N.D.	N.D.

4.3.3.1.2 Cesspit 2: Grave soil samples (Hungate, York)

The Roman age grave (C51364, 1st – 4th CE) was sampled for the InterArChive project in 2010 and 2011 during excavation of the Hungate site (Pickering *et al.* 2018). The grave showed poor preservation of the skeletal remains both in terms of completeness and in the physical condition of the bones. Thus, the Roman burial had lost much structural integrity and was incomplete, particularly regarding the upper left side of the remains (Figure 4.17). The samples were obtained from grave fill at three different anatomical locations around the human remains: (1) material excavated from a nearby cesspit, (2) abdominal region, (3) left pelvis. The products of soil extracts obtained from derivatisation with FLB were separated by LC with MS detection to obtain the base peak chromatogram of the FLB derivatives (Figure 4.18).

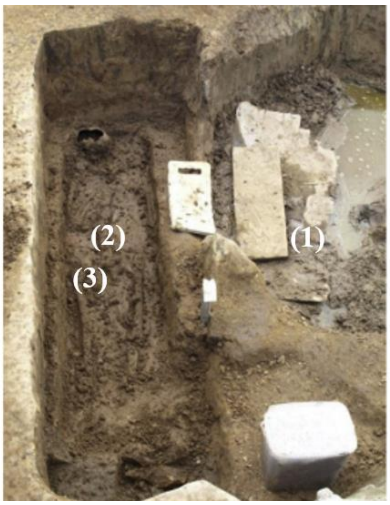


Figure 4.17. Roman age grave (C51364, 1st – 4th C CE) containing the skeletal remains of an adult (Roman age~3rd century CE) and a nearby Anglo-Scandinavian age cesspit (Context 2652) (Pickering *et al.* 2018). Three positions of the collected samples in this work: (1) material excavated from cess pit, (2) abdominal region, (3) left pelvis.

Successful detection of epicoprostanol, coprostanol, cholesterol and 5α-cholestanol at retention times of 7.9, 9.2, 9.6 and 10.2 min was evident from the extracted ion chromatograms at *m/z* 737.7 and at *m/z* 739.7 and mass spectra of the soil extracts (Figure 4.18, Table 4.4 and 4.5). The presence of cholesterol and cholestanols in all the soil extracts reflects contributions from the human remains within the grave fills. The amounts of C_{27} steroidal lipids were measured in the same manner as cesspit core from Hungate, York (See Section 4.3.3.1.1). Coprostanol is the major component, accompanied by smaller amounts of the epicoprostanol and 5α-cholestanol with a trace amount of cholesterol. The concentration of coprostanol in soils excavated from the cesspit was in range of micrograms per gram of dried soil while the amounts of epicoprostanol and 5α-cholestanol were in range of nanograms per gram of dried soil, reflecting deposition of faecal matter within the cesspit context (Table 4.4). The amounts of all cholestanols (coprostanol, epicoprostanol, 5α-cholestanol) were in the range of micrograms per gram of dried soil underneath abdominal region, while the amounts of these components were in range of nanograms per gram of dried soil underneath the left pelvis, reflecting a difference in inputs of faecal matter at different the anatomical positions within the grave.

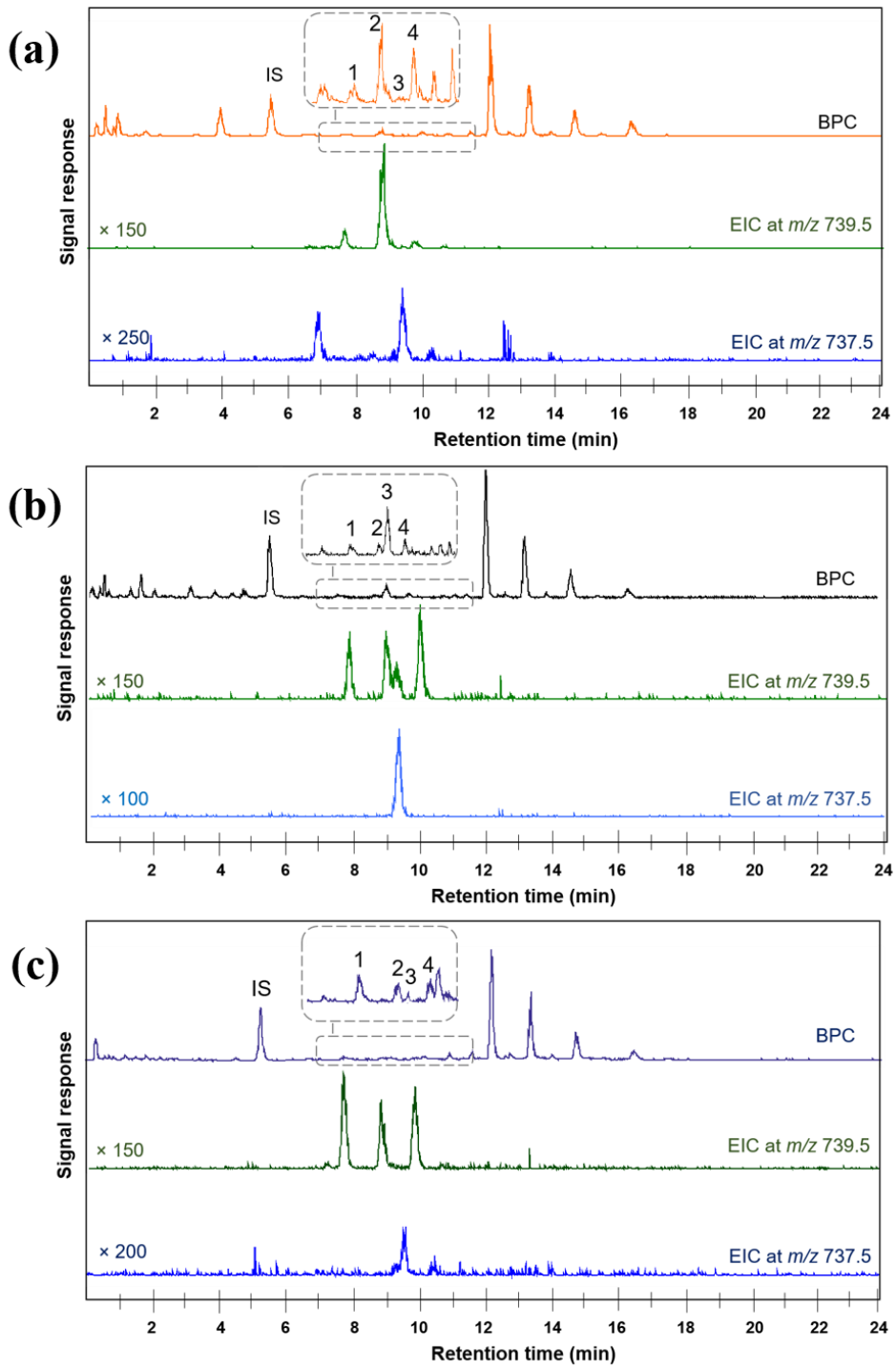


Figure 4.18. APCI MS chromatograms of the FLB derivatives of the medium polar frations of soil extracts from Hungate C51364 and the associated cesspit (a) material excavated from cesspit, (b) abdominal region, (c) left pelvis. Base peak chromatograms (BPC) and extracted ion chromatograms (EIC) obtained from *m/z* 737.5, *m/z* 739.5. Assigned peaks: 1 = epicoprostanol, 2 = coprosanol, 3 = cholesterol, 4 = 5α-cholestanol, IS = 1-heptadecanol as an internal standard.

Table 4.4. The measured cholestanols detected from soil extracts obtained from Roman age grave soils (Hungate C51364, York), *n* =3.

Soil extract from different positions	Detected cholestanols ($\times 10^{-9}$ gram per gram of dried soil, ng/g dw)			Coprostanol ratio (5β/5β+5α)		
	coprostanol	epicoprostanol	5α-cholestanol	FLB	TMS	(Pickering *et al.* 2018)
Excavated material	220 ± 10	2 ± 0.1	0.7 ± 0.04	0.99	0.90	> 0.9
Abdominal region	140 ± 6	140 ± 5	120 ± 6	0.70	0.72	< 0.75
Left pelvis	2 ± 0.1	2.3 ± 0.1	1.6 ± 0.08	0.73	0.72	< 0.75

Table 4.5. The measured cholesterol detected from soil extracts obtained from Roman age grave soils (Hungate C51364, York), *n* =3

Soil extract from different positions	Detected cholesterol ($\times 10^{-9}$ gram per gram of dried soil, ng/g dw)	
	FLB	TMS
Excavated material	1.3 ± 0.6	0.9 ± 0.1
Abdominal region	590 ± 30	560 ± 10
Left pelvis	0.1 ± 0.004	0.1 ± 0.005

Based on the successful UHPLC-MS identification of cholesterol and its transformation products and the deconvolution of signals using extracted ion chromatograms to identify the presence of cholestanols and cholesterol, the ratio of 5β/(5β+5α) cholestanols was estimated from the extracts of the different soil samples. The high ratios (>0.7) revealed in all soil extracts match the values obtained from the same soil samples (Hungate C51364) by GC-MS of the TMS derivatives and the estimated values that were reported by Pickering *et. al.* (Pickering *et al.* 2018) (Table 4.4). For the soil extracts obtained from the Roman age grave soil, the high index values (>0.9) from the excavated material evidences the migration of faecal materials from the cesspit into the grave,

suggested to be responsible for the decomposition of the left-hand side of skeleton (Pickering *et al.* 2018). By contrast, the lower values (>0.7) obtained from the soils collected from underneath abdominal region and left pelvis of human remains inside the grave is consistent with an input from a less strongly reducing, though still faecal source. The coprostanol ratios obtained from the UHPLC-MS are very similar to the GC-derived values, reflecting the suitability of FLB as alternative derivatising agent to evaluate human-influenced impacts in soils and sediments.

Estimation of cholesterol in all of the soil extracts was carried out in the same manner as for the cesspit core from Hungate, York (See Section 4.3.3.1.1.1) (Table 4.5). The trace amounts of cholesterol per gram of soil from the cesspit and from underneath the left pelvis was in the range of nanograms per gram of dried soil, significantly less than the concentrations of the cholestanols. Cholesterol concentrations in all soil extracts were not significantly different, possibly reflecting reduction of cholesterol to cholestanone (shown as a conspicuous peak in the GC-MS chromatograms) by soil bacteria. Remarkably, the concentration of cholesterol in soil underneath the abdominal region of the human remains was much larger than for the cholestanols, in range of micrograms per gram of dried soil. Detection of significant levels of cholesterol in an archaeological burial site possibly indicates an origin from the human remains where the bones have decomposed, releasing cholesterol into the soils (Pickering *et al.* 2018). Hence, the results demonstrate the potential for the use of this derivatising agent as an alternative approach to analyse the organic matter of archaeological soils and sediments.

4.3.3.2 Environmental samples from Humber estuary, Hull

Determination of cholesterol and cholestanols using the same approach as Section 4.3.3.1 was carried out on samples collected from different locations in the Humber estuary Humberside, UK (Figure 4.19). Two types of sediments were obtained from Humber estuary, Hull: a sediment core (*) and surface sediments (1, 2). The sediment core samples were from up-river from the Humber bridge at different depths (*): (1) 2-3 cm, (2) 4-5 cm, (3) 6-7 cm, (4) 8-9 cm, (5) 10-11 cm, (6) 12-13 cm, (7) 13-14 cm, (8) 14-15 cm, while two surface sediments collected near

Grimsby, closer to the Humber estuary: (1) surface 5 cm, (2) surface 2 cm. All selected samples were screened for steroidal compounds by GC-MS (data not shown) and the FLB derivatisation approach was applied with separation by UHPLC with MS detection to obtain the base peak chromatogram of the FLB derivatives (Figure 4.20).

Figure 4.19. Map of the Humber estuary showing the location of sample sites. Assigned locations of collected samples: * = Sediment core at up-river from Humber bridge site 1 (2-15 cm), (1) = surface sediment 5 cm, (2) = surface sediment 2 cm. The sediments were collected in 2000 by Mark C. Evans. The Humber estuary with a catchment area of 25,000 km^2, receives flows from upland rivers including the Derwent and Ouse and the more industrialised rivers Aire and Trent (adapted from Google Maps).

Successful detection of coprostanol, cholesterol, and 5α-cholestanol at retention times of 8.5, 9.0 and 9.5 min in the soil extracts was evident from extracted ion chromatograms (*m/z* 737.5 and *m/z* 739.9) and mass spectra (Figure 4.20). The retention times show close matches to the standard mixture (Figure 3.25). The mass spectra obtained from the samples were cross-checked with those obtained from the standard mixture, confirming the presence of cholesterol and

cholestanols. Epicoprostanol was identified as giving a prominent peak at retention time 8.0 min, base peak ion at *m/z* 739.9 and the same fragmentation pattern as that obtained from coprostanol.

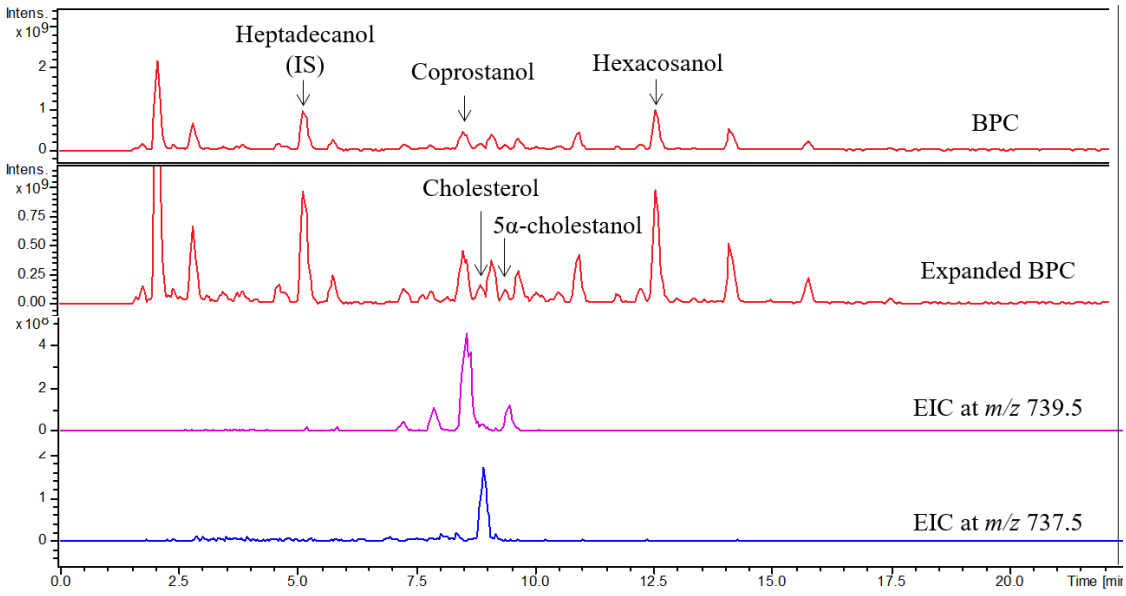


Figure 4.20. APCI MS chromatograms of the FLB derivatives of the medium polar fraction of sediment extract from Humber estuary site 2 (surface 2 cm). Base peak chromatograms (BPC) and extracted ion chromatograms (EIC) obtained from *m/z* 737.5, *m/z* 739.5.

The presence of cholesterol and cholestanols, in all of the soil extracts, reflect contributions from human activity, consistent with the samples being from riverine/estuarine settings influenced by major sites of humans occupation (Figure 4.21-4.22 and Table 4.6). In the majority of samples, coprostanol is the major component, accompanied by smaller amounts of other forms of cholestanols (epicoprostanol, 5α-cholestanol) and very low levels of cholesterol. The dominance of coprostanol in all depths of the sediment core and in the surface sediments near Grimsby reflects input of faecal matter and its transportation *via* the river flow. The profiles of the C_{27} steriodal lipids reveal changes of cholestanols (coprostanol, epicoprostanol and 5α-cholestanol) and cholesterol in the site over time (Figure 4.21-4.22, Table 4.6). All three stanols were detected from all depths of river sediment core, except for 2-3 cm depth (Figure 4.21). The changes in the

profile and concentrations within the core possibly reflects changes in the profile of steroidal lipids transportedation of these biomarker along the river. Alternatively, the decrease in concentrations could reflect increasing extents of degradation over time (Figure 4.21), though with a small increase at depth 12-13 cm. No significant trends could be recognised in the relative concentrations of cholestanols and cholesterol, possibly reflecting a variation in the compositions of the source material carried in the river over time.

Table 4.6. FLB derivatives of cholesterol and its reduced products detected from sediment extracts obtained from Humber estuary site from Hull, n =3.

Sediment from different depths (cm)	**Detected cholesterol and cholestanols ($\times 10^{-9}$ gram per gram of dried soil, ng/g dw)**			
	coprostanol	**epicoprostanol**	**5α-cholestanol**	**cholesterol**
Sediment core from up river				
2-3	330 ± 10	N.D.	N.D.	N.D.
4-5	650 ± 20	110 ± 5	210 ± 9	190 ± 20
6-7	510 ± 20	85 ± 1	200 ± 8	170 ± 8
8-9	420 ± 20	100 ± 4	80 ± 2	60 ± 3
10-11	440 ± 10	70 ± 1	20 ± 1	50 ± 3
12-13	390 ± 30	110 ± 4	180 ± 10	100 ± 8
13-14	370 ± 30	60 ± 5	45 ± 2	30 ± 3
14-15	330 ± 20	60 ± 1	30 ± 1	40 ± 2
Surface sediments from up river bridge				
Site 1	350 ± 10	90 ± 2	50 ± 2	50 ± 2
Site 2	460 ± 10	90 ± 4	80 ± 6	160 ± 6

N.D. = Non-determined

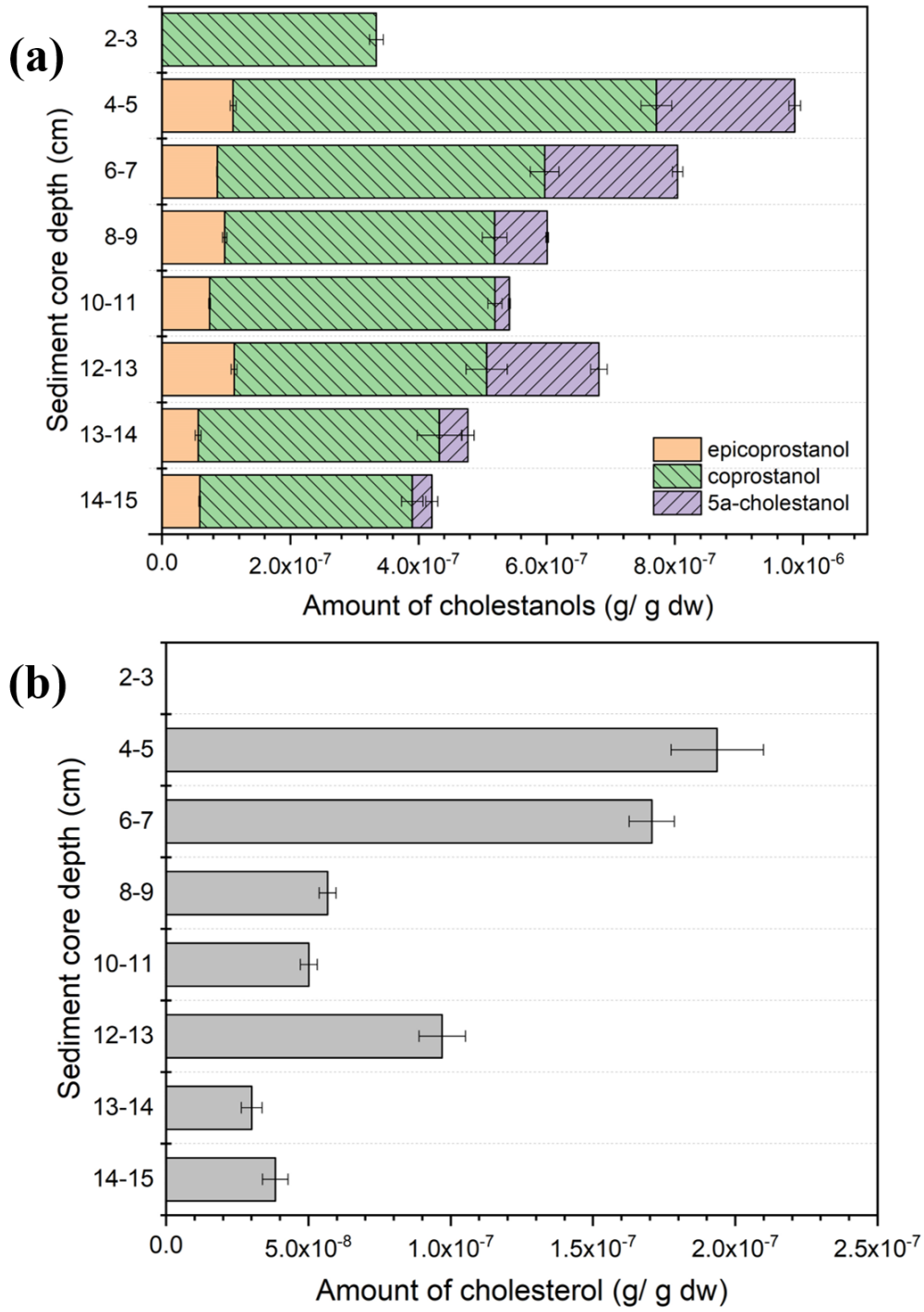


Figure 4.21. The measured amounts of C_{27} steroidal lipids detected from sediment extracts obtained from different depths of Humber estuary site from Hull (a) C_{27} stanols (coprostanol, epicoprostanol and 5α-cholestanol), (b) cholesterol, *n* =3.

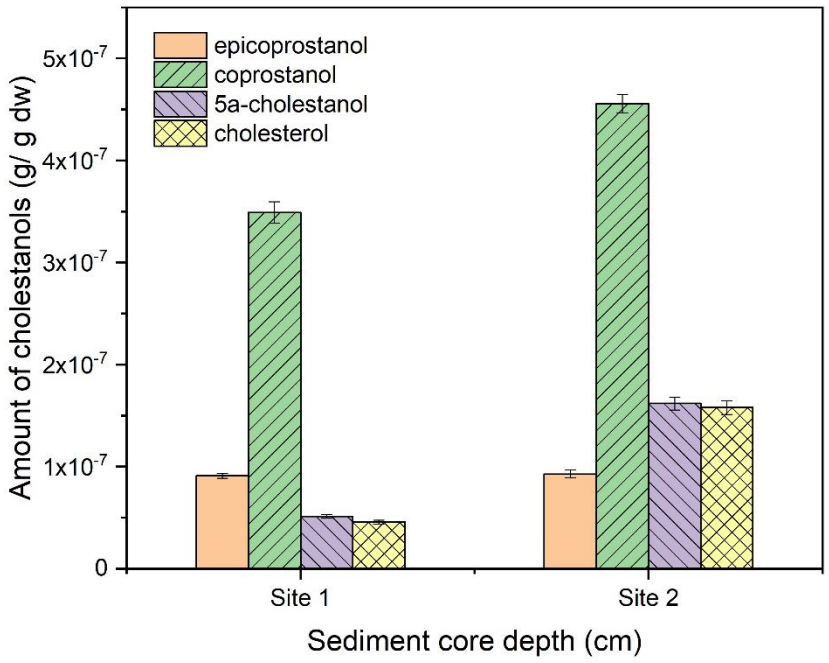


Figure 4.22. The measured amounts of C_{27} steroidal lipids (coprostanol, epicoprostanol, 5α-cholestanol and cholesterol) detected from sediment extracts obtained from surface sediment at different locations of Humber estuary site from Hull, *n* =3.

Based on the successful UHPLC-MS identification of cholesterol and its transformation products and the deconvolution of signals using extract ion chromatograms, the ratio of 5β/(5β+5α) cholestanols was estimated for the different levels. The seven different depths all contained coprostanol ratio >0.7, indicating faecal contamination (Table 4.7). Notably, 5α-cholestanol could not be detected at the top layer of sediment core. The FLB derivatisation was performed using the same the amounts of sediment extract as used for silylation. The coprostanol ratios obtained by LC-MS (>0.7) match the values obtained from GC-MS of the TMS derivatives for the corresponding samples where detected. Notably, the coprostanol ratio value in most of sediment core samples was not evaluated by GC-MS due to the inability to determine 5α-cholestanol, possibly reflecting that the amount of 5α-cholestanol was close to detection limit of GC-MS, while the same amount of 5α-cholestanol could be determined using an approach of FLB derivatisation. Notably, the changes in the concentrations of C_{27} steroidal lipids with depth in the

sediment core most likely reflect changes in human activity in the site over time. The amount of coprostanol was in range of ppb level (20 – 650 ng/g). The similar amounts of C_{27} steroidal lipids from surface sediments from site 1 and 2 possibly evidences the nature of the material transported along the river.

Table 4.7. The measured C_{27} steriodal lipids and 5β-stigmastanol from sediment extracts obtained from Humber estuary site from Hull, *n* =3.

Sediment from different depths (cm)	**Coprostanol ratio**		**Ratio of coprostanol: 5β-stigmastanol**	
	LC-MS	**GC-MS**	**LC-MS**	**GC-MS**
Sediment core from up river				
2-3	N.D.	N.D.	1.22 ± 0.04	0.74± 0.02
4-5	0.77 ± 0.02	N.D.	1.18 ± 0.04	N.D.
6-7	0.75 ± 0.02	N.D.	1.08 ± 0.05	0.97± 0.04
8-9	0.87 ± 0.01	N.D.	1.45 ± 0.01	0.85± 0.04
10-11	0.94 ± 0.02	N.D.	1.50 ± 0.05	N.D.
12-13	0.74 ± 0.01	N.D.	0.95 ± 0.04	N.D.
13-14	0.93 ± 0.01	N.D.	1.48 ± 0.05	1.54 ± 0.40
14-15	0.88 ± 0.01	0.88 ± 0.02	1.52 ± 0.02	1.60 ± 0.40
Surface sediments from up river bridge				
Site 1	0.85 ± 0.01	0.87 ± 0.04	2.07 ± 0.05	1.48 ± 0.40
Site 2	0.87 ± 0.01	0.92 ± 0.01	2.30 ± 0.05	1.54 ± 0.50

N.D. = Non-determined

The ratio of coprostanol to 5β-stigmastanol of all sediment samples was assessed in order to evaluate further the nature of the inputs in the site (Table 4.7). A high coprostanol: β-stigmastanol ratio (>1.5) was obtained from only four sediment core extract (8-9 cm, 10-11 cm, 13-14 cm, 14-15cm), confirming that dominant inputs were from both omnivore and herbivore with four levels that may potential represent human faecal material. The surface sediments from both site (site 1 and 2) gave the high coprostanol: β-stigmastanol ratios (>1.5), confirming the dominant inputs were from human faecal material consisent with the fact that the areas were influenced by human activity.

Determination of cholesterol and cholestanols from the eight sediment core samples and two surface sediments enabled successful detection of epicoprostanol, coprostanol, cholesterol, and 5α-cholestanol at retention times of 7.5, 8.3, 8.9 and 9.4 min in the extracts from the UV (Figure 4.23a) and at retention times of 12.1, 14.2, 14.9 and 15.7 min in the FL chromatograms which were run using HPLC owing to the lack of availability of a FL detector compatible with UHPLC (Figure 4.23b). The amounts of cholesterol and its related stanols were evaluated in the same manner with those obtained from MS detection (Figure 4.24).

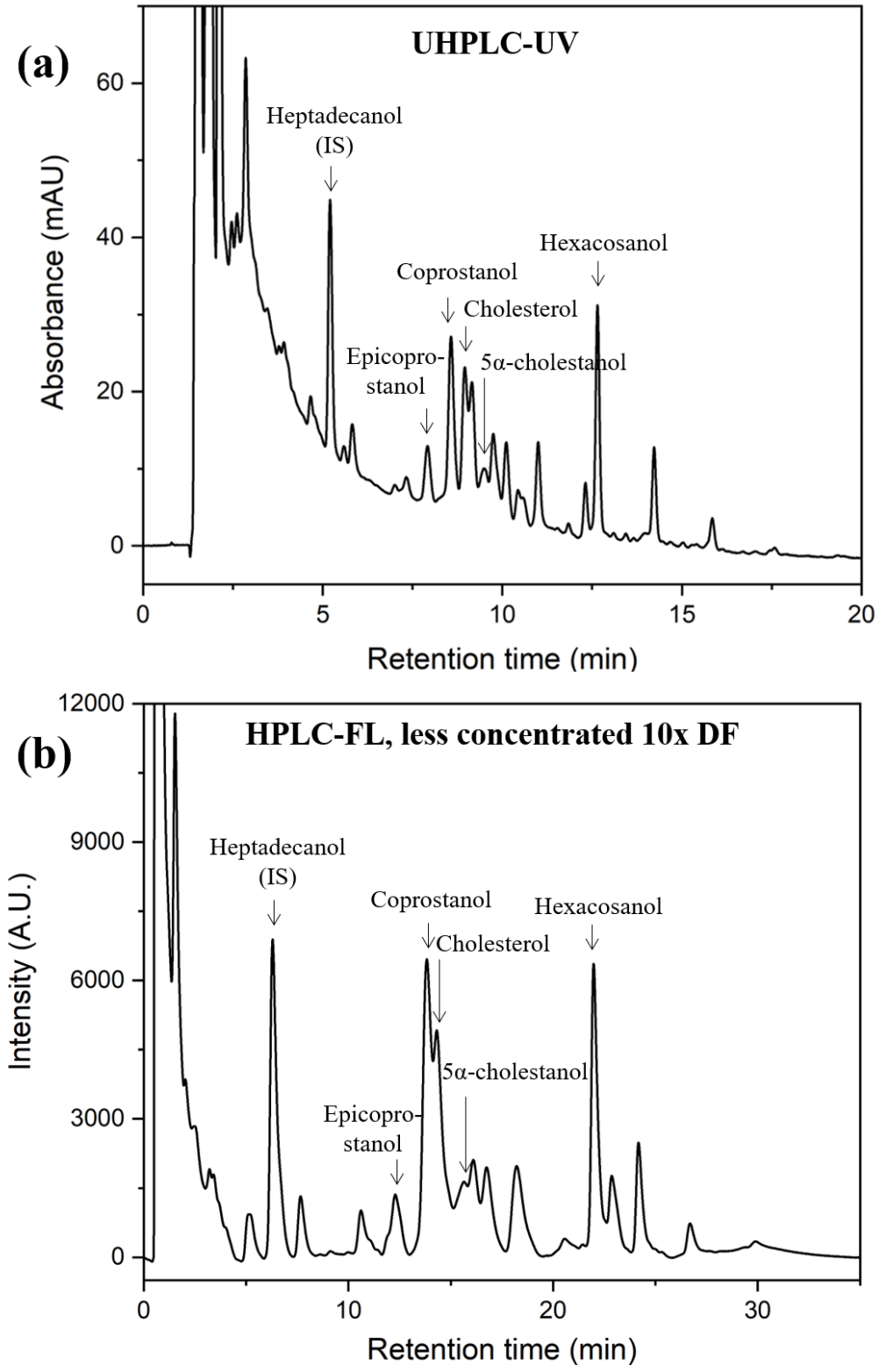


Figure 4.23. Partial chromatograms of the FLB derivatives of the medium polar fraction of the sediment extract from Humber estuary site 2 (surface 2 cm) recording using different detectors: (a) UV and (b) FL (*n* =3).

Comparison of signal responses by UHPLC-MS, UHPLC-UV and HPLC-FL reveals that quantification of cholesterol and its related stanols by UV and FL detectors provides comparable estimations of the amounts of the components compared with conventional MS detection for cholesterol and its transformation products (Figure 4.24). In the same soil extract, the amounts of C_{27} steroidal lipids estimated by UV detection gave similar results to those obtained from MS detection (Figure 4.20 and Figure 4.23a). By contrast, the concentrations obtained using the FL detector were slightly lower, most likely a result of the chromatographic system operating at lower flow rate and not completely resolving some peaks, e.g. coprostanol and cholesterol, to baseline (Figure 4.23b). In the same sediment extract, the FLB derivatives were detected by FL using less sample than was required for MS and UV: 10 times dilution. The lower sample requirement reflects the better signal response of fluorescence detection than UV and MS detection. Thus, the method demonstrates an opportunity to use FLB for a detection of organic matter in environmental samples and the greater signal response that it offers will give greater security to the interpretations. In cases where environmental materials contain steroidal compounds close to detection limit of GC-MS, the use of the FLB derivatising agent can provide a reliable alternative for quantification by either MS or FL detectors.

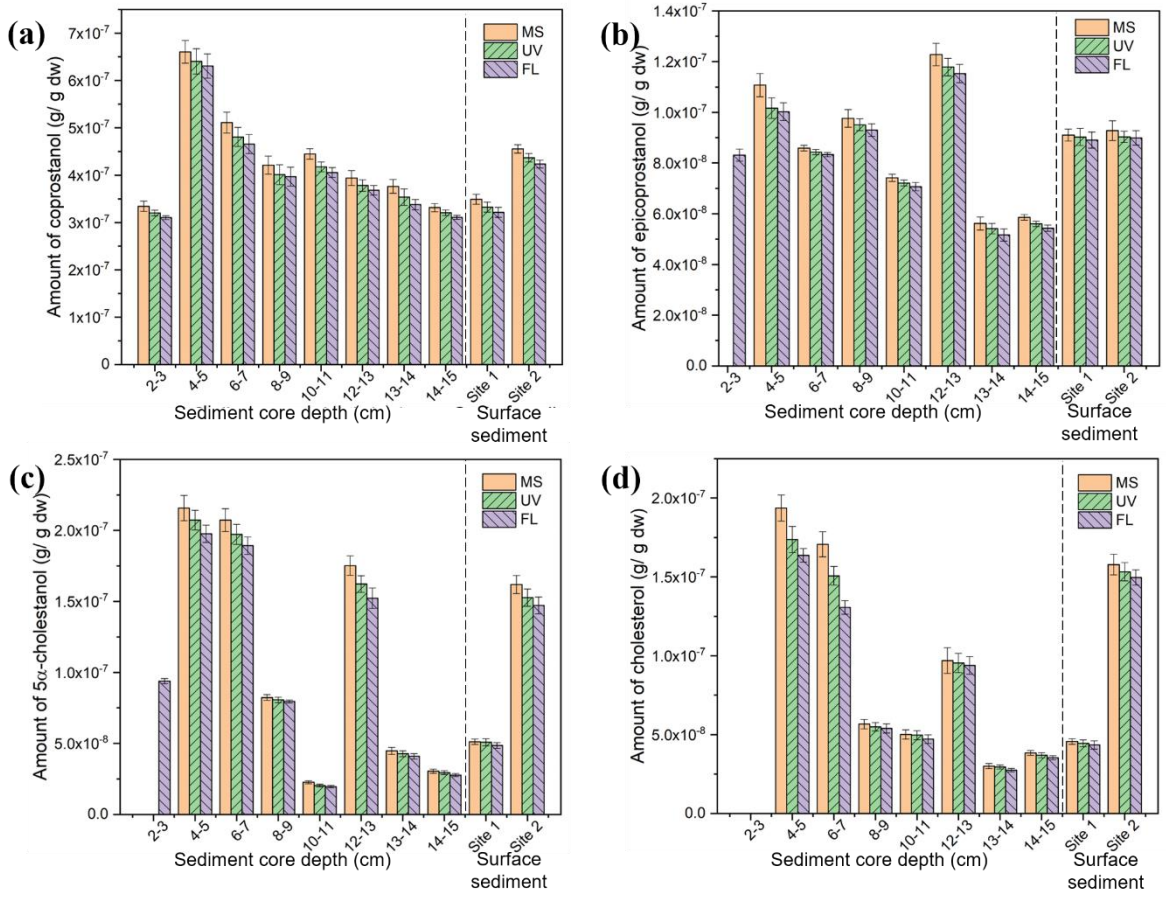


Figure 4.24. The measured C_{27} steriodal lipids from soil extracts obtained from sediment extracts obtained from different depths of Humber estuary site from Hull and surface sediments: (a) coprostanol, (b) epicoprostanol, (c) 5α-cholestanol, (d) cholesterol. The signal responses of the FLB derivatives of each sample were detected by three different detectors (MS, UV, FL), *n*=3.

4.4 Conclusion

The use of Fmoc-Lys-Boc (FLB) as a derivatising agent for the measurement of cholesterol and its reduced transformation products (coprostanol and 5α-cholestanol) to detect human activity was demonstrated in archaeological and environmental contexts.

Completeness of the derivatisation process for cholesterol and cholestanols in archaeological soil was improved by increasing the ratio of derivatising agent to sample solution relative to the proportions suggested for a cholesterol standard solution (Chapter 3) and Poplawski (2017). The optimised ratio of derivatising agents is 25: 12: 50 (FLB: DMAP: EDC to 1 mole of sample), and the refined method to detect FLB derivatives is described in Section 2.5.1 for UHPLC, and in Section 2.5.2 for HPLC.

An approach for FLB derivatisation of cholesterol and cholestanols was applied to a range of archeological samples (different fourteen depths from cesspit core and grave soil underneath human remains) and environmental samples (one sediment core and two surface sediments both from an estuary). The values of coprostanol ratios obtained are very similar to those obtained from GC-MS. Comparison of signal responses for steroidal components in archaeological soils and environmental sediment extracts analysed by HPLC-MS and GC-MS reveals that use of FLB as derivatising agent provides a substantial improvement in response compared with a conventional TMS derivative for cholesterol and its transform products. In the environmental samples, 5α-cholestanol was not detected in most of sediment core samples by GC-MS, while it could be determined using an approach of FLB derivatisation, evidencing the potential greater security for quantification by FLB derivatives.

Thus, the FLB approach provides an opportunity to overcome several of the limitations of organic matter detection in both archaeological and environmental samples and the greater repeatability that it will offer greater security to the interpretations. In cases where the materials have contain alcohol compounds close to detection limit of MS, the use of the FLB derivatising agent can provide an opportunity for quantification by FL detector.

Chapter 5

Method development for analysis of GDGTs as their Fmoc-Lys-Boc derivatives

5.1 Introduction

The use of glycerol dialkyl glycerol tetraethers (GDGTs) as biological biomarkers for palaeoenvironmental studies often entails analysis of large sample sets. The conventional method employs normal phase (NP), either cyano or amide columns, to separate *i*-GDGTs and APCI MS for identification and quantification (Hopmans *et al.* 2000; Schouten *et al.* 2007; Hopmans *et al.* 2016). The latest method for GDGT separation uses two UHPLC silica columns (BEH HILIC columns, 2.1 × 150 mm, 1.7 µm; Waters) operated at maximum back pressure of 230 bar (Hopmans *et al.* 2016). Although this NP column system achieves separation of the GDGT lipid core structures within 30 min, the total time for each chromatographic run is approximately 90 min and a 20 min column re-equilibrium time is required for every run in order to regenerate the stationary phase. Furthermore, column back-flushing is required every 20 runs to clean the columns (Hopmans *et al.* 2016). The long chromatographic run time reflects (1) the variations in GDGT lipid classes present in environmental samples, (2) the incomplete elution of highly polar components of the sample matrix. The changes in the mobile phase gradient (Hopmans *et al.* 2016) can be explained by considering *i*-GDGT analysis. The initial mobile phase composition of hexane: isopropanol (82:18, *v/v*) equates to a solvent polarity (*P'*) of 0.78. This composition is held for 25 min, followed by a slight increase in polarity to *P'* = 1.43 (hexane: isopropanol; 65: 35, *v/v*) at 50 min, followed by a further increase in polarity to *P'* = 3.90 (hexane: isopropanol; 0: 100, *v/v*). The mobile phase gradient programme enables *i*-GDGTs lipid core separation over a narrow polarity window (less than 0.8%). The matrix components, having higher polarity, require greater polarity (*P'* > 3.9: up to 5 times more than *i*-GDGTs) to be eluted, leading to the requirement the additional 20 min re-equilibration of the column. The incomplete elution of highly polar components of the sample matrix leads to their retention in the column, reducing column efficiency and shorten column lifetime. This is minimised by back-flushing the column every 20 runs. This practical issue limits opportunities for automation of the analytical method in order to handle large environmental sample sets. Alternative analytical methods for separating *i*-GDGT lipid cores and effecting elution of highly polar components have been developed based on reversed phase (RP) columns (Wörmer *et al.* 2013; Zhu *et al.* 2013; Liu *et al.* 2019; Rattray

and Smittenberg 2020). Those methods allow distinction of *i*-GDGT-0-3, crenarchaeol and its regioisomer with partial co-elution, eliminating the need for column regeneration between runs. The relative proportions of *i*-GDGTs contributing to the TEX_{86} index were found to be comparable with those obtained from NP columns, reflecting the suitability of RP separations for use in *i*-GDGT analysis (Rattray and Smittenberg 2020). Notably, however, the estimated TEX_{86} index determined by these RP methods gave a broader range of variation in surface water temperature (up to 3°C) than the NP method, reflecting a further need of analytical method development.

Mass spectrometry (MS) is the conventional detector for GDGT analysis as the native species (Schouten *et al.* 2007; Law and Zhang 2019). Variations in quantification of these components by MS presents a challenge in situations where the MS signal response differs. This may be due either to different GDGT structures (e.g. *i*-GDGTs and *br*-GDGTs) having different ionisation efficiencies or to use of different types of MS detector (Schouten *et al.* 2009, 2013; Lengger *et al.* 2018). Discrepancies in MS response caused by differences in instrument tuning can be avoided by spectrophotometric-based detection methods such as ultraviolet-visible (UV-vis) and fluorescence (FL), both of which are favourable techniques for quantification and offer the ability for absolute quantification of analytes. A few studies have considered derivatisation of archaeal ether lipid cores with an 9-anthroyl group for detection by LC-FL (Ramasha *et al.* 1989; Ohtsubo *et al.* 1993; Bai and Zelles 1997). Although this method showed potential for use in investigations of methanogenic ecosystems, the characterisation of the lipid profiles is limited by co-elution of GDGT lipid derivatives.

The analysis of Fmoc-lysine-Boc (FLB) derivatives of glycerol dialkyl glycerol tetraether (GDGT) was first demonstrated by Poplawski (2017). Two different core-shell RP columns were evaluated, an Ace UltraCore Super C_{18} column (3.0 mm × 150 mm, 2.5 μm) and a Kinetex® pentafluorophenyl (PFP) column (4.6 mm × 150 mm, 2.6 μm). These were used for the separation of GDGT lipid cores isolated from *Sulfolobus acidocaldarius* MR31 (*S. acidocaldarius* MR31) in the native form and as the FLB derivatives. The ACE column gave incomplete separation of *i*-GDGTs 0-5 within

25 min. The PFP column gave improved separation of *i*-GDGTs 0-5 within 20 min, generating a chromatographic peak for each GDGT and exhibiting high column efficiency (plate number, *N*, range 9,770-35,460). The PFP column was selected for use in the analysis of the FLB derivatives of *i*-GDGT-0-5 and baseline separation was achieved within 52 min. The relative MS responses of *i*-GDGT-1-3 FLB derivatives gave the same relative proportions as those obtained from the native species, indicating the suitability of FLB derivatisation to determine GDGT distributions. Surprisingly, however, the MS peak heights of the FLB derivatives were lower than those obtained from their native forms by about 10-1,000 times, flagging a concern regarding ionisation efficiency. This low response would limit GDGT lipid quantification by MS. By contrast, more reliable quantification of GDGT lipids could be achieved by UV and FL detection due to the efficiencies of the underlying spectroscopic processes along with the use of a laboratory-synthesised 1,2-di-*O*-octadecyl-*rac*-glycerol (r-dOG) as an internal standard. The HPLC separation of FLB derivatives of *i*-GDGT-1-3 was also achieved within 60 min, though with peak broadening observed due to the lower pressure (<400 bars). Although the peak areas of the derivatives measured with FL detection were in the same relative proportions as the native species, the low separation efficiency reflects the need for an improved analytical method.

This chapter presents a systematic study to investigate the signal suppression of GDGTs in the FLB form and to determine if the method could be improved to make it applicable for routine analysis. The FLB derivatisation process proposed by Poplawski was repeated to generate a comparable result (see Section 5.3.1). The tandem MS dissociation pathways of GDGT lipids in both native and FLB forms were examined to develop an understanding of their dissociation and potential for structural characterisation (see Section 5.3.2.4). A complete analytical procedure, from sample preparation to separation method and detection was systematically refined (See Section 5.3.2). The applicability of the refined analytical method for measurement of *i*-GDGTs and *br*-GDGTs was demonstrated using the sediments from an international round-robin study (Schouten *et al.* 2013) (see Section 5.3.3). The GDGTs obtained from FLB derivatisation were compared with those obtained from the conventional method to assess the potential for GDGT lipid quantification at the levels they occur in typical sediments.

5.2 Aims

The overall aim of the work described in this chapter was to conduct a systematic evaluation of the applicability of FLB as a derivatising agent to determine and quantify GDGT lipid profiles for use in the TEX_{86} and BIT indices. Specific objectives were as follows:

1. To explore the potential for improvement in the completeness of the derivatisation process when applied to soils/sediments (see Section 5.3.1).

2. To refine the analytical method for measurement of GDGT lipids using three different detectors (MS, UV and FL) (see Section 5.3.2).

3. To demonstrate the suitability of the method for assessment of TEX_{86} and BIT indices by application to geological sediments from an international round-robin study (see Section 5.3.3).

5.3. Rationale, results and discussion

5.3.1 Assessment of the preparation of FLB derivatives of glycerol dialkyl glycerol tetraethers (FLB-GDGTs) using Poplawski's approach

5.3.1.1 Native GDGTs from S. solfataricus

The native GDGT lipid cores were extracted from a cell culture of *S. solfataricus* (see growth condition in Section 2.2.2) by a modified Bligh-Dyer technique and acidic methanolysis (Nishihara and Koga, 1987). The native lipid extract and the extract derivatised with FLB were separated by UHPLC-MS (Figure 5.1) using the separation conditions of Poplawski (Poplawski 2017). The mass spectrum of each peak was compared with the identification based on retention time order reported in the literature (Table 5.1).

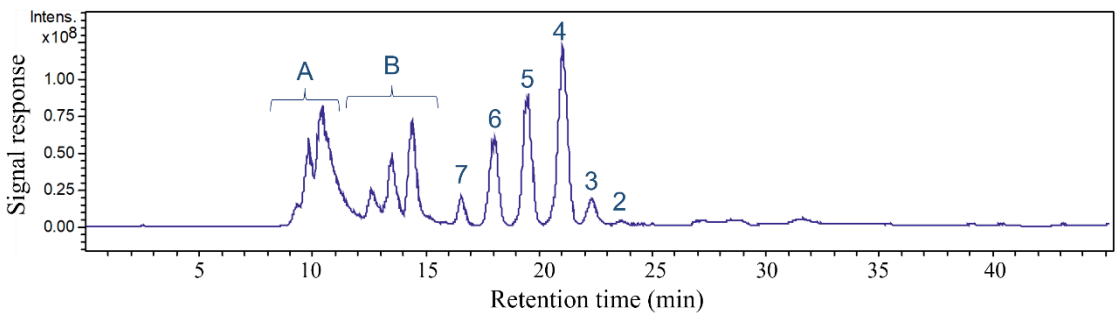


Figure 5.1. APCI UHPLC-MS based peak chromatograms of the lipid extract from *S. solfataricus* in native form: before derivatisation. Assigned peak group A = glycerol dialkyl calditol tetraether (GDCT) lipid cores, group B = unassigned, 2-7 = *i*-GDGT-2-7. The components were prepared and analysed according to Poplawski (2017).

The base peak chromatogram of the native lipid extract from *S. solfataricus* (Figure 5.1) exhibits peaks within three main groups (A, B and 2-7). The mass spectra of the earliest eluting group (group A) gave protonated molecules between *m/z* 1452.6-1458.4 in the mass spectra (Figure 5.2, Figure C.1), confirming the presence of C_{92} GDGTs having calditol etherified to the GDGT core (*cf* Koga and Morii, 2005). The term GDCT was suggested for these components (Knappy *et al.* 2015). The calditol etherified lipids are acid tolerant to methanolysis and are commonly found in the lipid extracts prepared from *Sulfolobus* species (Knappy *et al*. 2009; Zeng *et al.* 2018). The

second group (group B), eluting later, gave protonated molecules in the range *m/z* 1496.4-14969.4 and *m/z* 1510.4-1511.4 (Figure 5.2, Figure C.2), presently not assigned. The later eluting group (2-7) gave protonated molecules in the range *m/z* 1288.4-1299.4 (Figure 5.3-5.4), consistent with *i*-GDGT-2-7. The full MS and MS^2 spectra of peaks 2-7 were explored to identify the structural complexity in *i*-GDGTs and the MS^2 spectra were compared with those of known lipids present in extracts from a previous report to confirm identifications (Knappy *et al.* 2012, 2015). For example, the full MS spectrum of the protonated molecule of *i*-GDGT 4 (peak 4 in Figure 5.1) gave the base peak ion at *m/z* 1294.4 and a fragment ion at *m/z* 1276.4 corresponding to the loss of water (Figure 5.4a). The tandem MS^2 spectra of the ions at *m/z* 1294.4 and *m/z* 1276.4 gave similar fragment ions, reflecting common sub-structures in each ion (Figure 5.4b-c). The product ions at *m/z* 1276.4, *m/z* 1258.4, *m/z* 1250.4, *m/z* 1220.4, *m/z* 1202.4, arise from cleavages of one molecule of water (-18 Da), two molecules of water (-36 Da), propenal (C_3H_4O, -57 Da), hydroxyacetone ($C_3H_6O_2$, -74 Da), glycerol ($C_3H_8O_3$, -92 Da) at different positions of glycerol, respectively.

Structural variations in the biphytanyl hydrocarbon chains were identified from the MS^2 spectra of the ions between *m/z* 500 and *m/z* 750 (Figure 5.5). For example, the MS^2 spectra of GDGT-4 (peak 4 in Figure 5.1) gave product ions in the range *m/z* 737 to *m/z* 741, corresponding to the losses of biphytadienes with different numbers of cyclopentanyl rings (BP_1-BP_3) (Figure 5.5, Table 5.1). The isomeric forms of *i*-GDGT-4 eluting in the retention time window 20-22 min (Figure 5.1, peak 4) were recognised from the MS^2 spectra of the precursor ion at *m/z* 1294.4 (*i*-GDGT-4) integrated based on a 1-min peak width interval (Figure 5.5). Thus, the MS^2 spectrum of the *i*-GDGT-4 peak 4_a (precursor ion at *m/z* 1294.4, Figure 5.5.c) gave a unique product ion at *m/z* 739.7 (-554.6 Da; $C_{40}H_{74}$) indicating the Δ_0 structure whereas the MS^2 spectrum of *i*-GDGT-4 peak 4_a'' (precursor ion at *m/z* 1294.5, Figure 5.5.b) exhibited product ions at *m/z* 737.7 (-556.7 Da; $C_{40}H_{76}$), at *m/z* 739.7 (-554.6 Da; $C_{40}H_{74}$) and at *m/z* 741.7 (-554.7 Da; $C_{40}H_{72}$) corresponding to cleavages of esterified chain containing one (BP_1), two (BP_2) and three (BP_3) cyclopentane rings, respectively, consistent with co-eluting Δ_0 and Δ_2 isomeric forms of *i*-GDGT-4. Thus, the identification of various etherified chains of *i*-GDGT-4 reflects the overlapping of isomeric forms

within one chromatographic peak (Knappy *et al.* 2012). Notably, the product ion profiles only discriminate among the isomeric forms and do not relate directly to abundance. Structural complexity within other chromatographic peaks exhibiting base peak ions in range *m/z* 1288.4-1303.4 was recognised in the similar manner to *i*-GDGT-4 (Table 5.1). Successful characterisation of *i*-GDGTs 0-7 isolated from membrane lipids of *S. solfataricus* indicated the suitability of this extract for use as a standard GDGT mixture in the modification of the analytical method. Notably, the results demonstrate that the Archaeal lipid extraction technique used does not completely removed intact polar head groups (IPLs) (Sturt *et al.* 2004; Knappy *et al.* 2012).

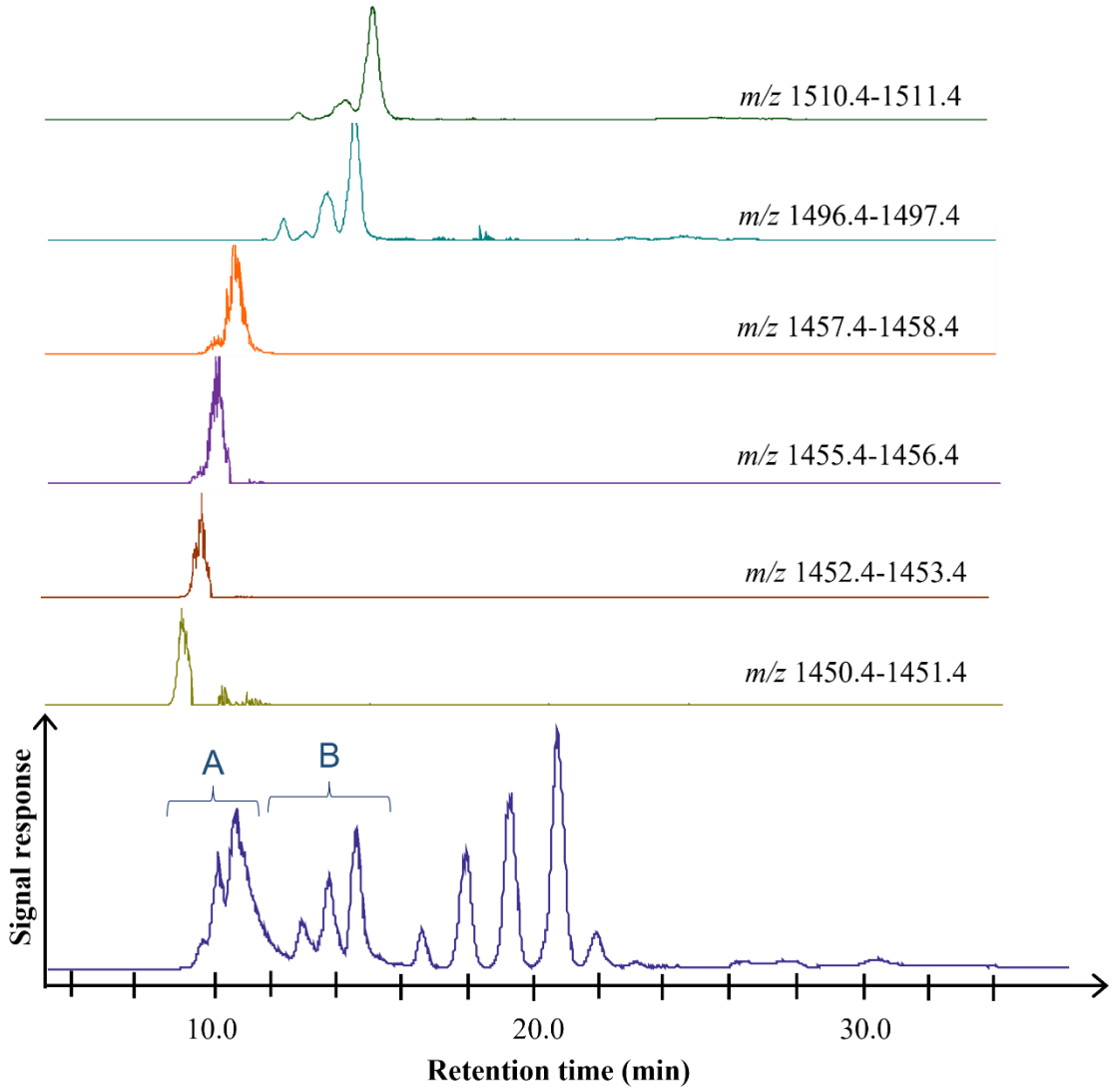


Figure 5.2. APCI UHPLC-MS based peak chromatograms of the lipid extract from *S. solfataricus* and extracted ion chromatograms in range of *m/z* 1450-1511 showing perfect alignment of retention times with peaks in groups A and B.

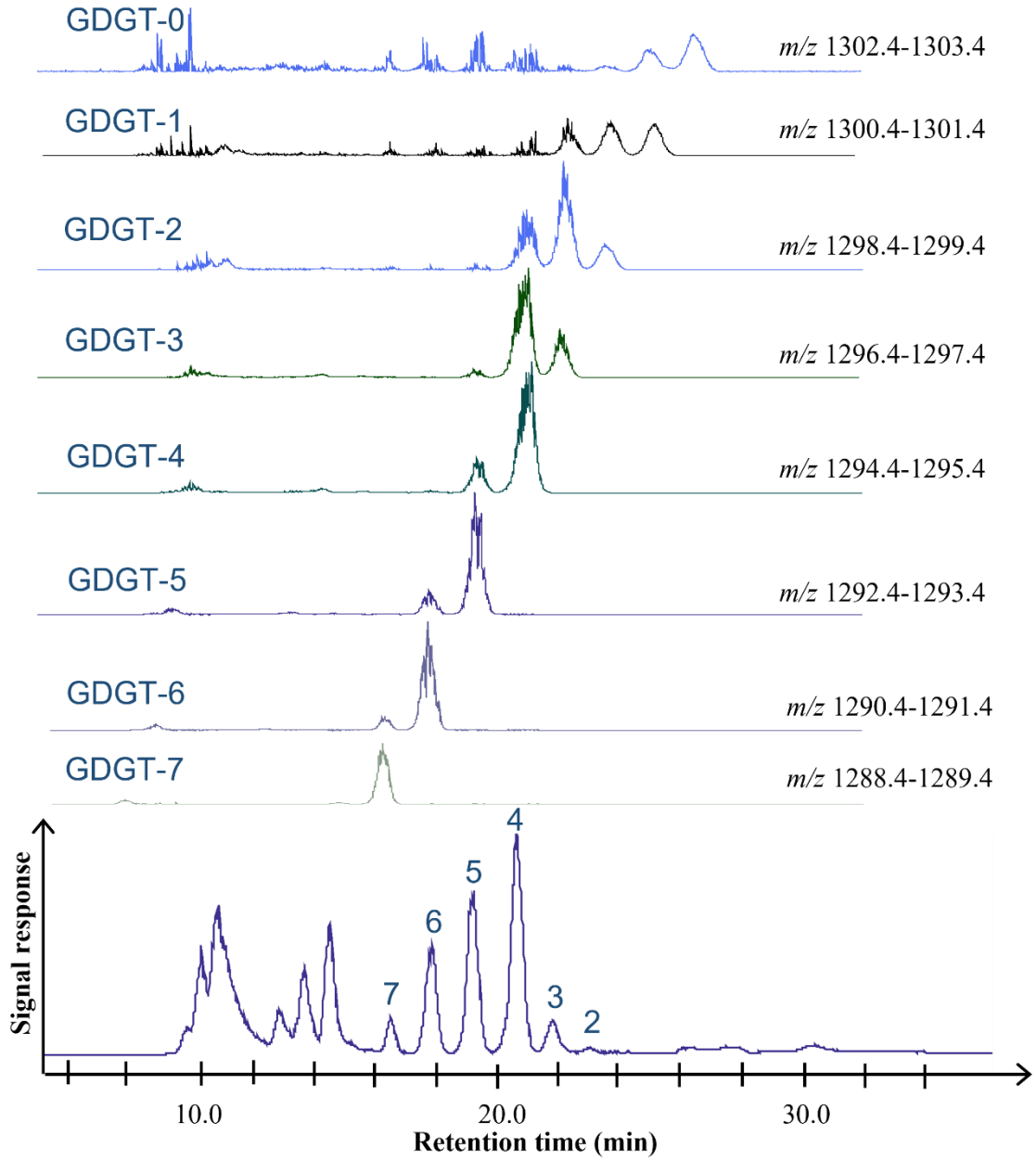


Figure 5.3. APCI UHPLC-MS based peak chromatograms of the lipid extract from *S. solfataricus* and extracted ion chromatograms in the range *m/z* 1288-1302 showing perfect alignment in retention times with peaks of *i*-GDGT-0-7.

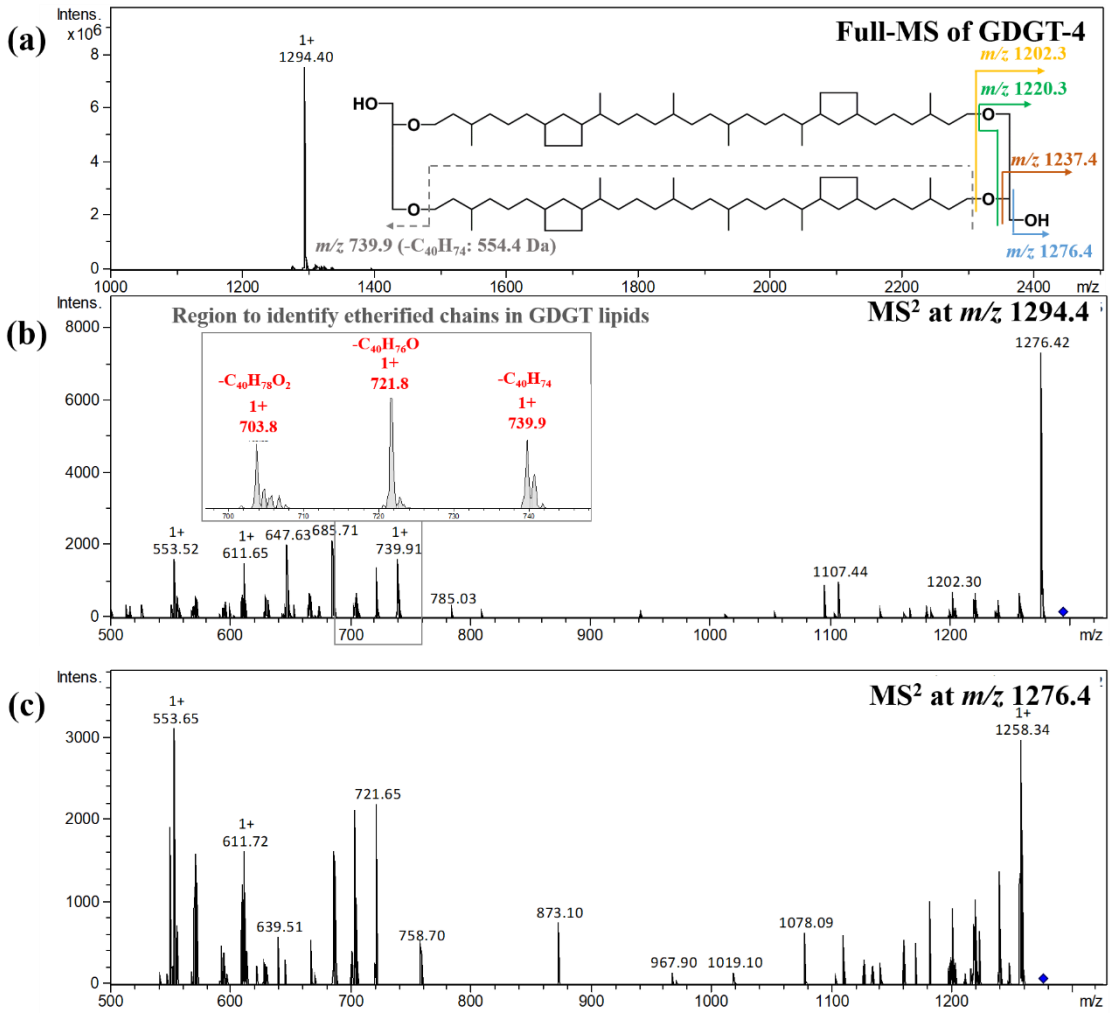


Figure 5.4. APCI MS spectra of native glycerol dialkyl glycerol tetraether-4 (*i*-GDGT-4) obtained from the lipid extract from *S. solfataricus* recorded in different modes: (a) full MS spectrum, (b) and (c) MS^2 spectra from precursor ions at *m/z* 1294.4 and *m/z* 1276.4, respectively. The inset image shows the structure of *i*-GDGT-4 with fragmentation indicated.

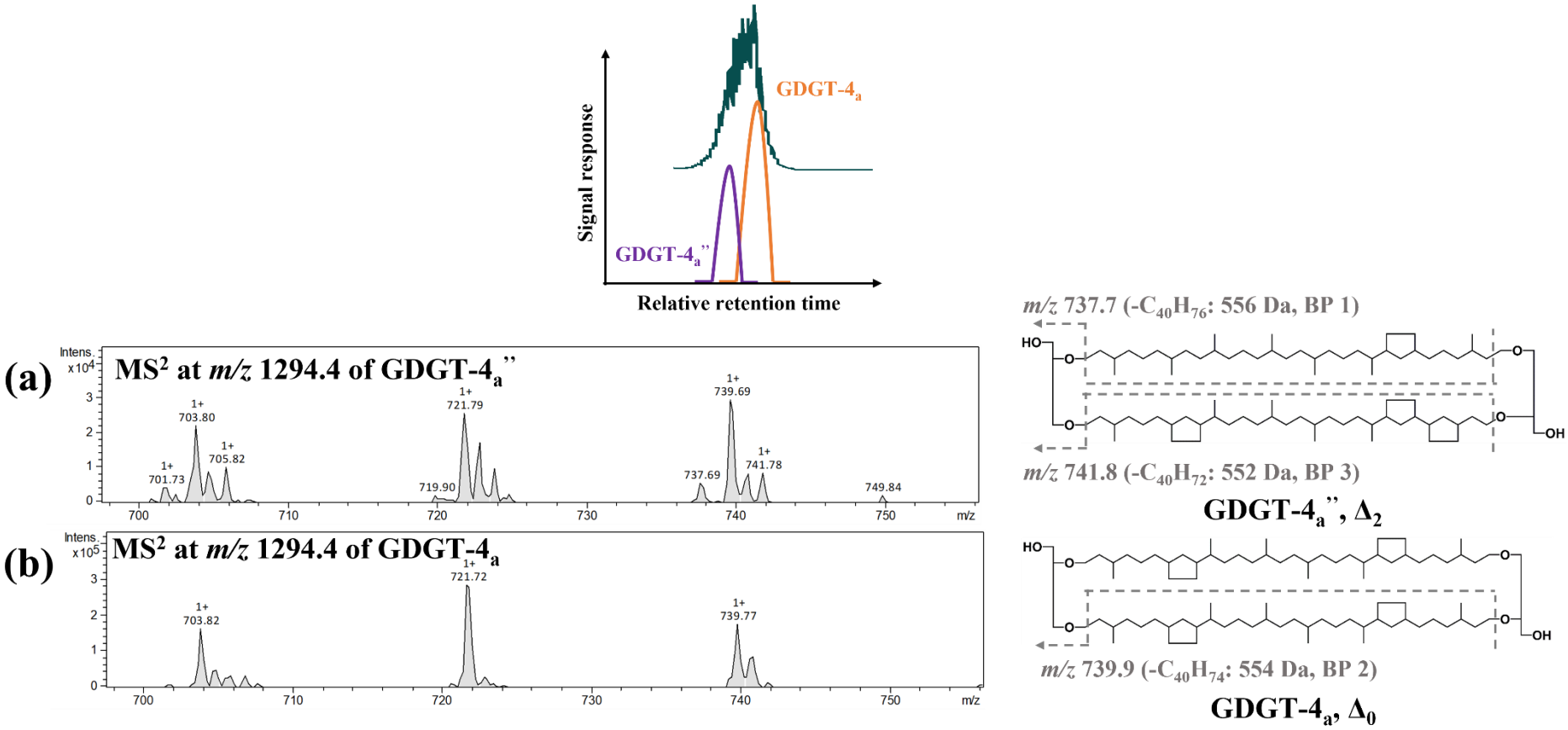
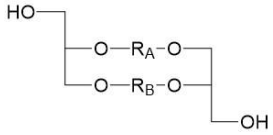


Figure 5.5. MS^2 spectra of *S. solfataricus* native glycerol dialkyl glycerol tetraether-4 (*i*-GDGT-4) from precursor *m/z* 1294.4, shown over the range *m/z* 700- 750. The spectra were obtained at retention times (a) 20.5 min, (b) 21.5 min. The inset shows the structure of *i*-GDGT-4 with fragmentation giving loss of the of biphytane chains illustrated.

Table 5.1. Mass spectral data and identification of etherified chains of native glycerol dialkyl glycerol tetraether lipid cores obtained from the *S. solfataricus* extracts.

i-GDGTs = 

Lipid	$[M+H]^+$ (*m/z*)	Δ_n	Characteristic product ions (*m/z*)	Etherified chains		Presence in *S. solfataricus* extracts	
				R_A	R_B	This work	Refs.
0_a	1302.3	Δ_0	745.9, 725.7	BP_0	BP_0	✓	✓
$1_a'$	1300.3	Δ_1	743.6, <u>741.9</u>	BP_0	BP_1	✓	✓
2_a	1298.3	Δ_0	741.9	BP_1	BP_1	✓	✓
$2_a''$	1298.3	Δ_2	743.4, <u>739.8</u>	BP_0	BP_2	✓	✓
$3_a'$	1296.4	Δ_1	741.6, <u>739.7</u>	BP_1	BP_2	✓	✓
$4_a''$	1294.4	Δ_2	<u>739.7</u>, 737.7, 741.8	BP_1	BP_3	✓	✓
4_a	1294.4	Δ_0	739.9	BP_2	BP_2	✓	✓
$5_a'$	1292.4	Δ_1	739.6, <u>737.9</u>	BP_2	BP_3	✓	✓
6_a	1290.4	Δ_0	737.8	BP_3	BP_3	✓	✓
$6_a''$	1290.4	Δ_2	739.5, <u>735.7</u>	BP_2	BP_4	✓	✓
$7_a'$	1288.4	Δ_1	737.8, <u>735.8</u>	BP_3	BP_4	✓	✓

Base peak ions were underlined., Refs = (Knappy *et al.* 2015; Knappy *et al.* 2012).

5.3.1.2 FLB derivatives of GDGTs extracted from S. solfataricus

The *S. solfataricus* polar lipid fraction was derivatised with FLB using the method of Poplawski (2017) and the crude reaction product was cleaned up by column chromatography using an eluent of hexane: ethyl acetate (1:1, *v/v*) to remove residues from the reaction. The resulting FLB-GDGT fraction was analysed by UHPLC-MS using the same APCI MS condition (Section 5.3.1.1) used for native GDGTs to enable evaluation of the completeness of the derivatisation process (Figure 5.6).

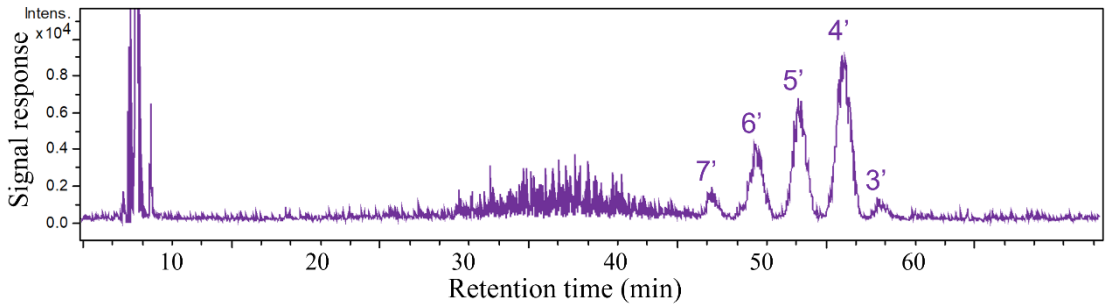


Figure 5.6. APCI UHPLC-MS based peak chromatograms of the lipid extract from *S. solfataricus* as FLB derivatives: after derivatisation with FLB reagent in ratio of 1: 50, Assigned peak 3'-7' = FLB-*i*-GDGT-3-7. The components were prepared and analysed according to Poplawski (Poplawski 2017).

A poor MS signal was obtained for FLB derivatised GDGTs (FLB-GDGTs) using the reagents in the proportions specified by Poplawski (FLB: DMAP: EDC = 50: 10: 75; Poplawski 2017). The base peak chromatogram of the FLB-GDGTs (Figure 5.6) exhibits a group of peaks (3'-7') at later retention times than the native GDGTs, reflecting the more apolar character of the analytes following derivatisation. The mass spectra of peaks (3'-7') exhibit characteristic fragmentation of FLB derivatives and confirm they are derivatives of GDGTs. For example, the mass spectrum of peak 4' (Figure 5.7b) has the base peak at *m/z* 1294.3 corresponding to the prominent ion of the native form of GDGT-4 (peak 4 in Figure 5.1, 5.7a) arising from loss of two molecules of FLB (-900 Da). The mass spectrum also contains a moderate intensity ion at *m/z* 2095.7 consistent with loss of the butyloxycarbonyl (Boc) group (-100 Da) from the protonated molecule as is commonly observed for FLB derivatives (Poplawski 2017). Thus, peak 4 is assigned as FLB-GDGT 4 esterified with two FLB groups. The mass spectrum of FLB-GDGT-4 also show three minor intensity ions: at *m/z* 1423.4, at *m/z* 1751.6, at *m/z* 2197.6 presently not assigned. Likewise, the mass spectra of the other peaks labelled 3'-7' also show characteristic fragmentation of fully FLB derivatised GDGTs confirming the peaks as di-esterified FLB-GDGTs 3-7. Notably, however, as noted by Poplawski (2017), the MS signal response for the FLB-GDGTs (Figure 5.6) is much lower than that obtained from the native form (Figure 5.1) by

approximately 1,000 times, reflecting either incomplete recovery or signal suppression. The low response negates the benefits of the FLB derivatives over the native GDGTs and would preclude their use in routine analysis. Reasons for the low response were considered to include the presence of hydroxyl groups from other alcohols (for example GDCTs) in the GDGT extracts which could also be derivatised with FLB (Figure 5.1, Figure 5.2), possibly causing a low yield of FLB-GDGTs from over-consumption of the derivatising agent. This could be tested by using a modified sample clean-up procedure designed to remove GDCTs.

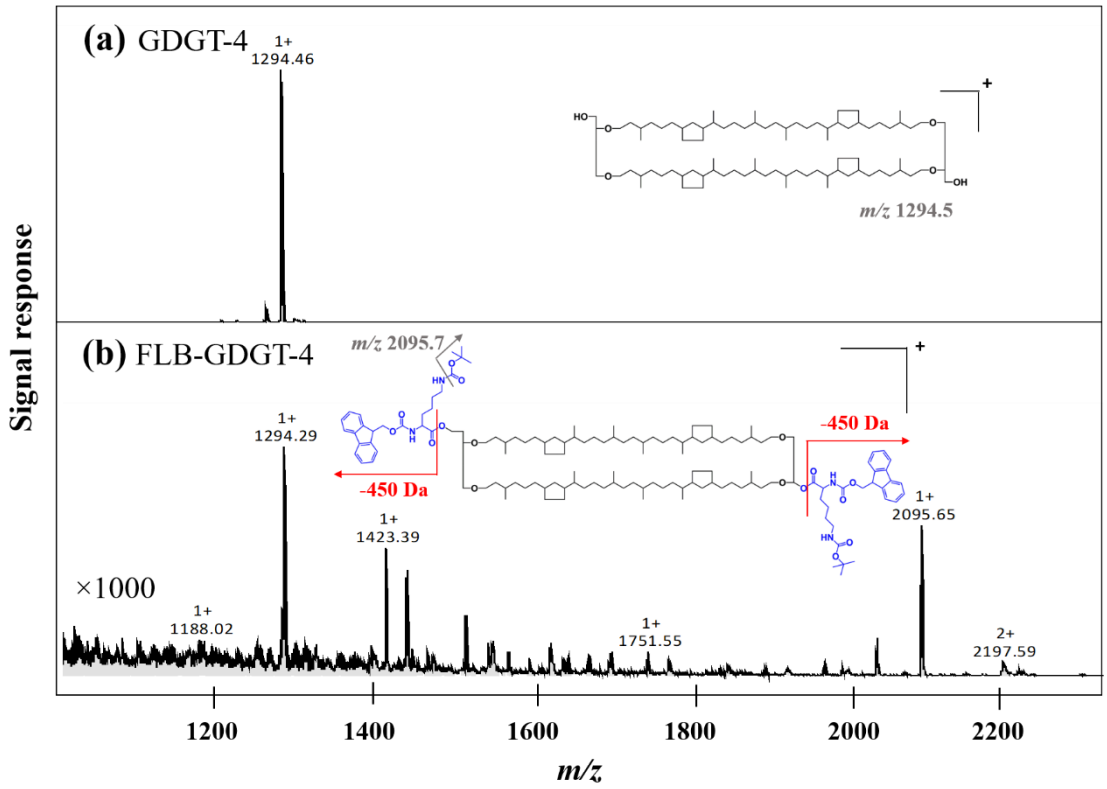


Figure 5.7. APCI MS spectra of (a) GDGT 4, (b) FLB-GDGT 4 of lipid extract from *S. solfataricus*, *m/z* 1000-2300. The MS spectra were acquired using the same MS conditions as employed by Poplawski (2017).

5.3.2 Modification of the analytical method for GDGT analysis as Fmoc-Lys-Boc derivatives

5.3.2.1 Evaluation of the amount of derivatising agent required for dihydroxyl compounds

As Poplawski (2017) was unable to demonstrate the effective use of FLB as a derivatising agent for GDGT analysis due to the weak MS response for the derivatives a systematic examination of the stages in the process was carried out to identify where signal was lost. Due to limited availability of GDGT lipid core material, 1,12-dodecanediol was used as a model compound in this study, like GDGTs, it is a dihydroxylated lipid. The amount of derivatising agent to 1,12-dodecanediol was prepared in various ratios (Table 5.2). The residual reagents were removed from the crude reaction mixture using column chromatography and the FLB derivative of 1,12-dodecanediol (1,12-FLB diester) was eluted using a mixture of hexane: ethyl acetate (1:1, *v/v*). The isolated 1,12-FLB diester was analysed by HPLC using fluorescence (FL) detection to determine signal response of the product (Figure 5.8a).

Table 5.2 Mole ratios of 1,12-dodecanediol and derivatising agents (FLB, DMAP, EDC).

Mole ratio of 1,12-dodecanediol: FLB: DMAP: EDC	**1,12-dodecanediol (moles)**	**FLB (moles)**	**DMAP (moles)**	**EDC (moles)**
1: 1.5: 1.1: 2	1.5×10^{-5}	2.2×10^{-5}	1.6×10^{-5}	3.0×10^{-5}
1: 1.7: 1.3: 2.5	2.4×10^{-5}	4.0×10^{-5}	3.4×10^{-5}	6.1×10^{-5}
1: 2.5: 2.6: 5	1.2×10^{-5}	3.0×10^{-5}	3.1×10^{-5}	5.9×10^{-5}
1: 3: 2: 4	1.5×10^{-5}	4.5×10^{-5}	3.1×10^{-5}	6.1×10^{-5}
1: 4: 2.6: 5	1.5×10^{-5}	6.4×10^{-5}	3.9×10^{-5}	7.8×10^{-5}
1: 10: 3: 9	1.6×10^{-5}	1.6×10^{-4}	4.9×10^{-5}	1.5×10^{-4}
1: 16: 10: 20	1.2×10^{-5}	2.0×10^{-4}	1.3×10^{-4}	2.5×10^{-4}
1: 20: 10: 40	1.5×10^{-5}	3.2×10^{-4}	1.5×10^{-4}	6.3×10^{-4}

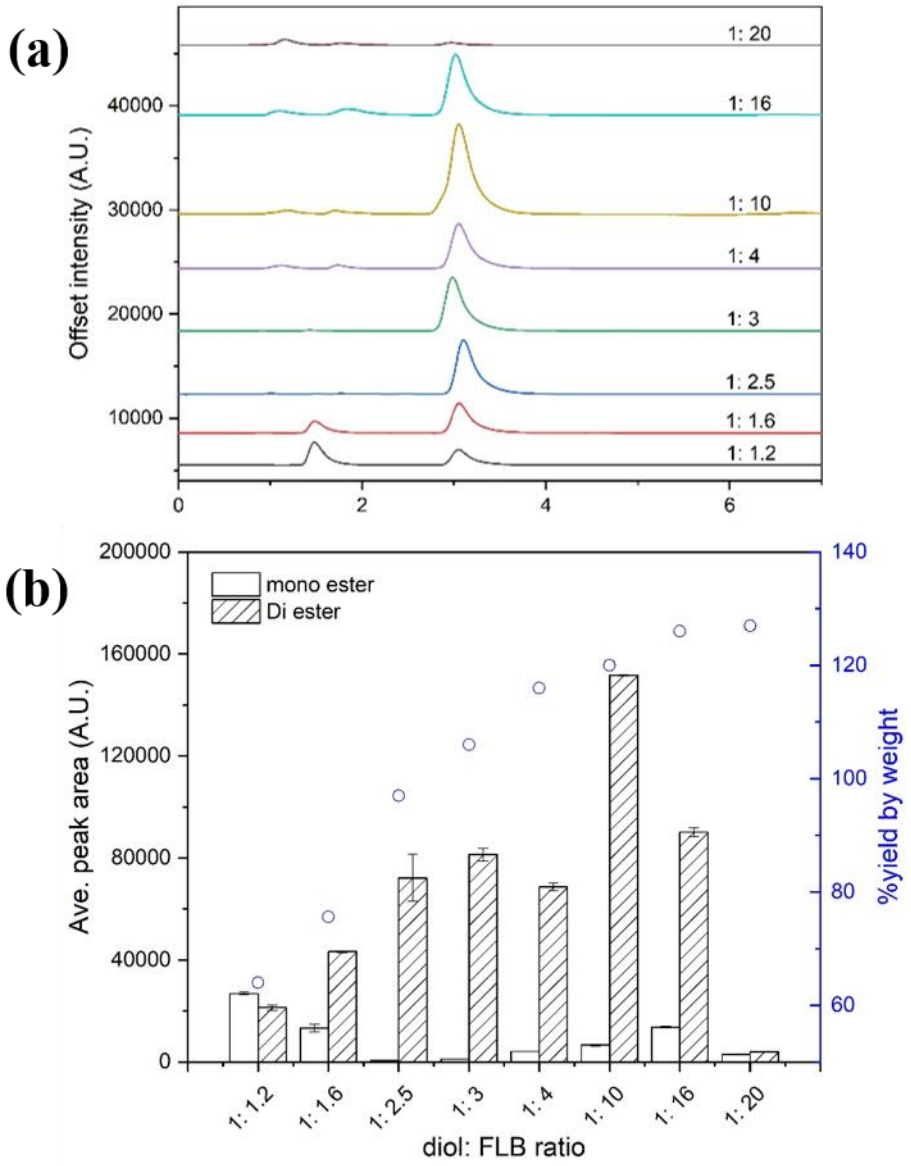


Figure 5.8. (a) chromatograms of the 1,12-dodecanediol-FLB derivative (1,12-FLB diester) prepared using various ratios of reagents (mole ratio as 1,12-dodecanediol: FLB) recording using fluorescence detection, (b) Peak area of fluorescent signals of the diol-FLB corresponding to the percentages by weight (n =3). All prepared diol-FLBs were normalised to the same concentration of initial 1,12-dodecanediol before HPLC analysis. Separation was carried out using a Dionex Acclaim RSLC120 C_{18} column (2.1 mm × 150 mm, 2.2 µm) maintained at 45°C using eluent comprising methanol/0.5 % acetic acid in water/ethyl acetate (73/12/15, *v/v*) delivered isocratically at 0.4 mL min^{-1}. The FL detection was performed using a Jasco FP-920 fluorescence detector with excitation wavelength (λ_{ex}) 263 nm and emission wavelength (λ_{em}) 309 nm.

The HPLC-FL chromatograms show a dominant peak eluting at a retention time of 3 min confirming a successful esterification of 1,12-dodecanediol with FLB and a minor peak at retention time of 1.5 min (Figure 5.8a). Based on the chemical structure of 1,12-dodecanediol, polarity of the C_{18} column and capability of fluorescence response, the major peak obtained from the chromatogram is suggested to be the FLB di-ester form of 1,12-dodecanediol (1,12-FLB diester). The minor peak in the chromatogram is suggested to be the mono-ester FLB form of 1,12-dodecanediol (1,12-FLB monoester). The yields of the 1,12 FLB mono- and di-ester were determined from the overall product yield and proportionality of peak areas from the HPLC-FL separation (Figure 5.8b).

The incomplete derivatisation of 1,12-dodecanediol under the initial reaction conditions is evident from an increase in the overall signal responses of the 1,12-FLB diester when the amount of derivatising agent was increased, indicating an increased percentage yield of the product (Figure 5.8b). The FL signal response of the 1,12-FLB diester was highest and the percentage of the product reached 100% when the amount of FLB to diol was 10 times, reflecting complete derivatisation. A significant decrease in signal response of FLB derivatives was observed when the amount of derivatising agent was raised to above 16 times, possibly reflecting formation of unwanted *N*-acylisourea intermediate and quenching of the esterification reaction yield. In order to obtain the maximum signal response for diol compounds, the suggested ratio of derivatising agents is 10: 3: 10 (FLB: DMAP: EDC to 1 mole of diol).

5.3.2.2 Modification of sample clean-up step for GDGTs extract

In the method the derivatised lipid extract is cleaned up using column chromatography to remove residual intact polar lipids. To test the efficiency of the procedure, the GDGT extract from a cell culture of *S. solfataricus* was divided into two equal portions. One of the portions was cleaned-up using the conventional eluent (DCM:MeOH, 1:1 *v/v*) (Knappy *et al.* 2012) and the other was cleaned-up using sequential elution starting from a low polarity hexane:ethyl acetate (2:1, %*v/v*) mixture through increasingly polar mixtures and finally eluting with DCM:MeOH (1:1, %*v/v*).

The isolated GDGT extracts were separated by UHPLC-MS to obtain the peak areas of the resulting lipid cores from each eluent. Characterisation of the resulting lipid cores was achieved by mass spectrometry (MS and MS^2).

The base peak chromatograms of the native GDGT extracts collected using each eluent all exhibited six peaks eluting in the retention time range 15-30 min; those peaks gave protonated molecules in the range *m/z* 1292.4-1302.4 (Figure 5.9a). The MS^2 spectrum of all of the peaks gave perfect matches to the GDGT lipids containing 0-4 rings identified previously in the extracts (Knappy *et al.* 2012, 2015). The dominant peaks in the UHPLC-MS chromatograms reflect the native GDGTs isolated by each eluent, the extent of their recovery being reflected in their peak areas (Figure 5.9b).

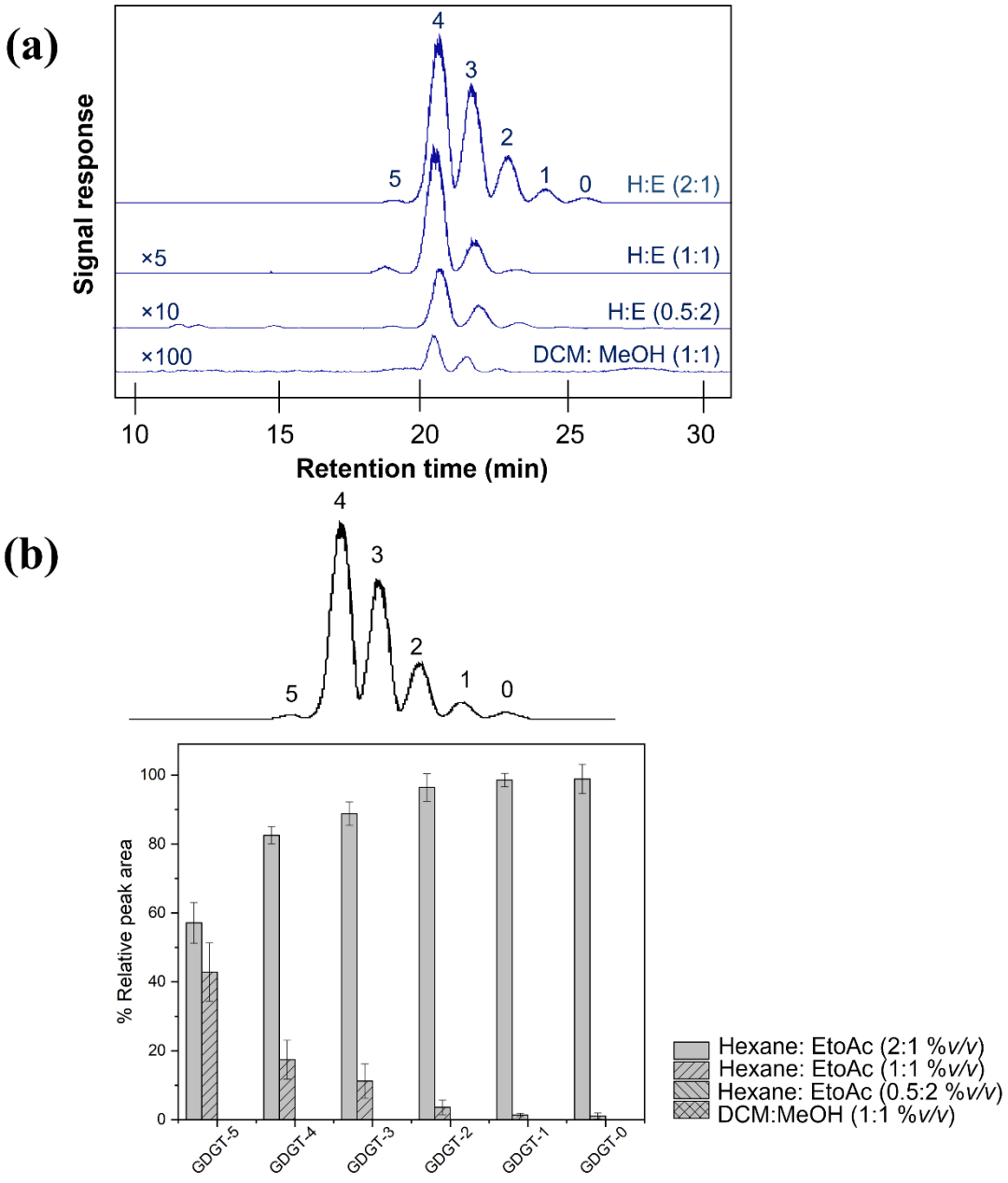


Figure 5.9. (a) APCI UHPLC-MS base peak chromatograms of the native GDGTs extracted from *S. solfataricus*, *m/z* 900-1500 by sequential extraction with the solvent combinations listed. (b) Peak area of each assigned peak when using different eluents. Assigned peaks: 1 = *i*-GDGT 5 (*m/z* 1292.4), 2 = *i*-GDGT 4 (*m/z* 1294.4), 3 = *i*-GDGT 3 (*m/z* 1296.4), 4 = *i*-GDGT 2 (*m/z* 1298.4), 5 = *i*-GDGT 1 (*m/z* 1300.4), 6 = *i*-GDGT 0 (*m/z* 1302.4). The GDGTs extract was derivatised with the ratio of FLB: DMAP: EDC (10: 5: 20). Eluents: H:E = hexane: ethyl acetate, DCM: MeOH = dichloromethane: methanol.

The initial eluent, hexane:ethyl acetate (2:1, *v/v*), gave almost complete isolation of the GDGTs (Figure 5.9b) as indicated by the percentage peak area (60-100%) with an absence of peaks in the retention time window 0-15 min. The peak areas of the GDGTs sequentially isolated using hexane: ethyl acetate (1:1, *v/v*) were lower for GDGTs 0-3 (*m/z* 1302.4, *m/z* 1300.4, *m/z* 1298.4, *m/z* 1296.4), less than 15%, while those for GDGTs 4-5 (*m/z* 1294.4, *m/z* 1292.4) were between 20-45%, reflecting a significant recovery level. No significant peak areas were evident for extracts eluted with higher polarity solvent mixtures (hexane:ethyl acetate, 0.5:2 *v/v* and dichloromethane:methanol, 1:1 *v/v*); the peak areas of less than 0.9% reflect almost complete isolation of GDGTs achieved by the more apolar eluent combinations.

Separation of the native GDGTs by UHPLC-MS show almost complete removal of the early eluting peaks when using the mixtures of hexane-ethyl acetate (1:1 and 2:1, *v/v*) as eluent: no intact polar lipids were observed. Only small amounts of residue were evident in the MS chromatogram from the subsequent fraction obtained using the more polar eluent combination: DCM: MeOH (1:1 %*v/v*), confirming successful recovery of GDGTs from the cell extracts. Following removal of the GDCT residues from the extract, the MS response was similar to that observed in the chromatogram from the original method, reflecting no significant loss of GDGTs. The almost complete removal of residue while maintaining signal response for the GDGTs clearly indicates the suitability of the modified sample clean-up step (Figure 5.9b). The result suggests that GDGT lipids can be collected selectively from the Archaeal extracts by eluting with a mixture of hexane: ethyl acetate (1:1, *v/v*) instead of the conventional solvent (DCM: MeOH, 1:1 *v/v*).

5.3.2.3 Modification of the lipid extraction process for Archaeal cultures

The previous section shows that GDGT lipids were collected selectively from the Archaeal extracts of *S. solfataricus* P2 (DSM 1617) by eluting with a mixture of hexane:ethyl acetate (1:1, *v/v*) instead of the conventional solvent (DCM:MeOH, 1:1 *v/v*). Due to the limited availability of cell material of *S. solfataricus* P2 (DSM 1617), another *Sulfolobus* species, *Sulfolobus acidocaldarius* MR31 (*S. acidocaldarius* MR31) was used for further experiments. The sample

clean-up step was re-assessed in the same manner as that used for *S. solfataricus* P2 (DSM 1617). Lipids were extracted from *Sulfolobus acidocaldarius* MR31 (*S. acidocaldarius* MR31) by cell lysis in a mixture of chloroform, methanol and trichloroacetic acid prior to cleavage of head groups by methanolysis using methanol-HCl at 100°C (Knappy *et al.* 2012). The crude lipid extract was cleaned up using column chromatography to remove residual intact polar lipids. The GDGT lipids were collected selectively from the Archaeal extracts by eluting with a mixture of hexane: ethyl acetate (1:1, *v/v*).

The crude GDGT extract was analysed by UHPLC-MS to obtain the absolute signal response. The base peak chromatogram shows a series of small peaks eluting before 18 min, indicating low levels of intact polar lipids (IPLs) remaining in the extract (Figure 5.10a). The peaks eluting over the retention time range 18-25 min gave protonated molecules in the range *m/z* 1290.5-1302.5, corresponding to GDGTs lipid cores GDGT-0-6. Essentially complete removal of IPLs was achieved using a chromatographic clean-up on activated alumina eluted with hexane:ethyl acetate (1:1, *v/v*) (Figure 5.10a). The base peak chromatograms of the GDGT extract both before and after clean-up gave a dominant peak at *m/z* 1294.5, corresponding to *i*-GDGT-4. The base peak chromatogram also gave peaks at *m/z* 1290.5, *m/z* 1292.5, *m/z* 1296.5, *m/z* 1298.5, *m/z* 1300.5, *m/z* 1302.5, reflecting the presence of *i*-GDGTs 0-6 in the extract. The MS chromatogram of the purified GDGT extract was essentially identical to that of the crude extract except for the absence of IPLs or other peaks eluting earlier than the GDGTs, indicating that the purified product is suitable for studying the derivatisation step.

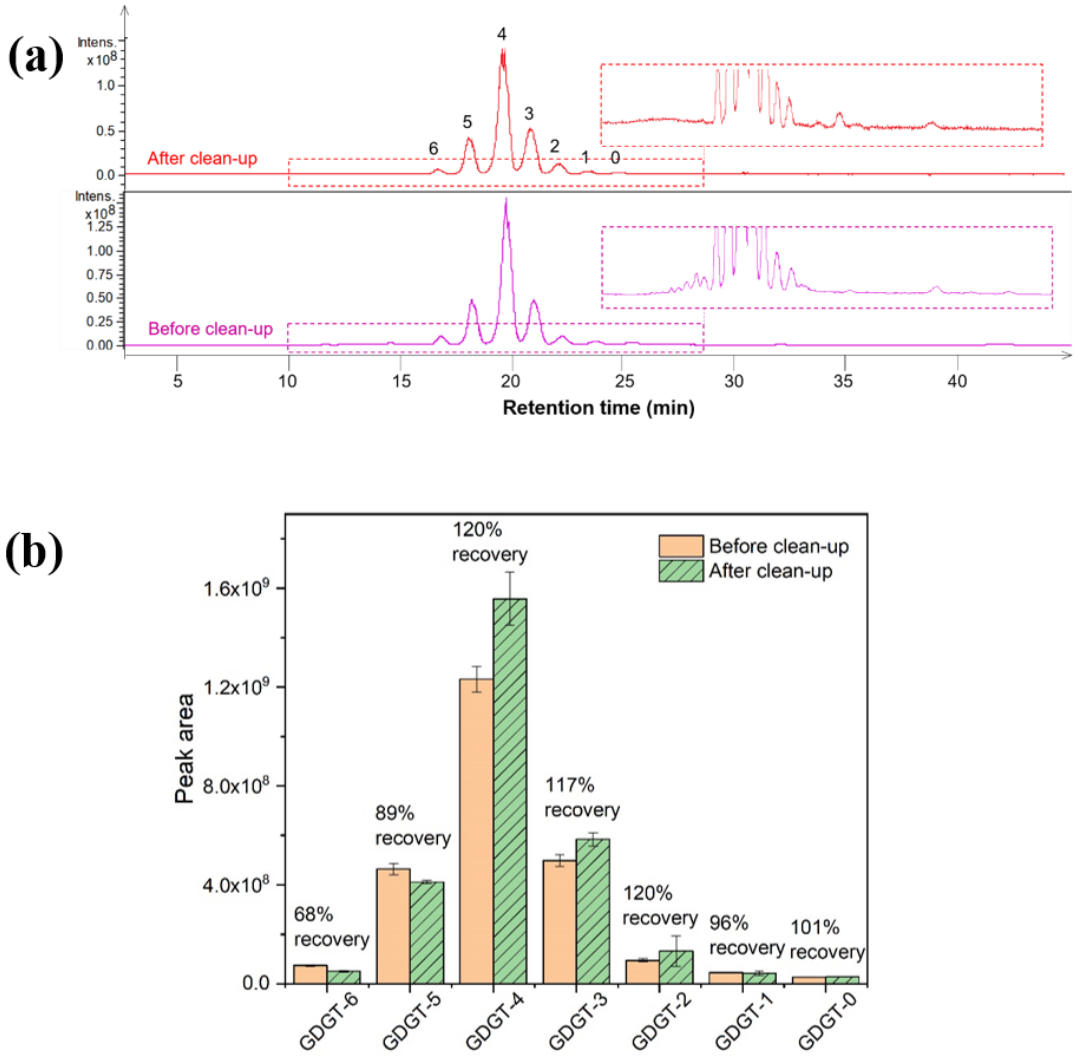


Figure 5.10. (a) APCI UHPLC-MS base peak chromatogram of the native GDGTs extract from *S. acidocaldarius* MR31 before and after clean-up, *m/z* 900-1500. (b) UHPLC-MS peak areas obtained from individual native GDGT peaks before and after clean-up process. The GDGT extract was collected using hexane: ethyl acetate (1:1, *v/v*), *n* = 3. Peak 0: GDGT-0, *m/z* 1302, peak 1 = GDGT-1, *m/z* 1300, peak 2 = GDGT-2, *m/z* 1298, peak 3 = GDGT-3, *m/z* 1296, peak 4 = GDGT-4, *m/z* 1294, peak 5 = GDGT-5, *m/z* 1292, peak 6 = GDGT-6, *m/z* 1290.

UHPLC-MS peak area measurements were used to determine the efficiency of the recovery of individual native GDGTs before and after clean-up. Individual GDGTs in the purified extract had recoveries between 67% and 120% recovery compared with the values obtained before clean-up (Figure 5.10b), reflecting essentially complete recovery of the native GDGTs in the extract.

The result confirms that GDGT lipids can be purified from the Archaeal extracts with complete removal of IPLs by eluting with a mixture of hexane:ethyl acetate (1:1, *v/v*).

The purified GDGT lipids extracted from *S. acidocaldarius* MR31 were derivatised with FLB using the ratio of 10: 5: 10 (FLB: DMAP: EDC to 1 mole of GDGTs). The FLB-GDGT extract was isolated and analysed in the same manner as for *S. sulfolitaricus*. The UHPLC base peak chromatograms of the FLB-GDGTs gave three groups of peaks (Figure 5.11a). The first eluted over the same retention time range as the native GDGTs with the protonated molecules in the range *m/z* 1290.5-1298.5, corresponding to *i*-GDGTs 2-6. The absence of GDGTs 0 and 1 is attributed to the low signal response for these compounds in the chromatogram. The other two groups of peaks were identified based on the structures of GDGTs, the estimated protonated molecules of GDGTs in mono- and di-ester forms and the polarity characteristic of the C_{18}-PFP column. The second group of peaks, eluting over the range 30-40 min, gave protonated molecules in the range *m/z* 1740.5-1748.5, corresponding to the mono-ester form of GDGTs 2-6 (mono-FLB-GDGTs). The third group of peaks, eluting over the range 50-70 min, gave protonated molecules in the range *m/z* 2191.5-2199.5, corresponding to the di-esters form of GDGTs 2-6 (di-FLB-GDGTs). The two groups of peaks of the mono- and di-ester FLB-GDGTs obtained from the UHPLC-MS chromatograms show that the preparation process produced pure GDGT-FLB derivatives in two forms. The different extents of esterification indicate incomplete derivatisation.

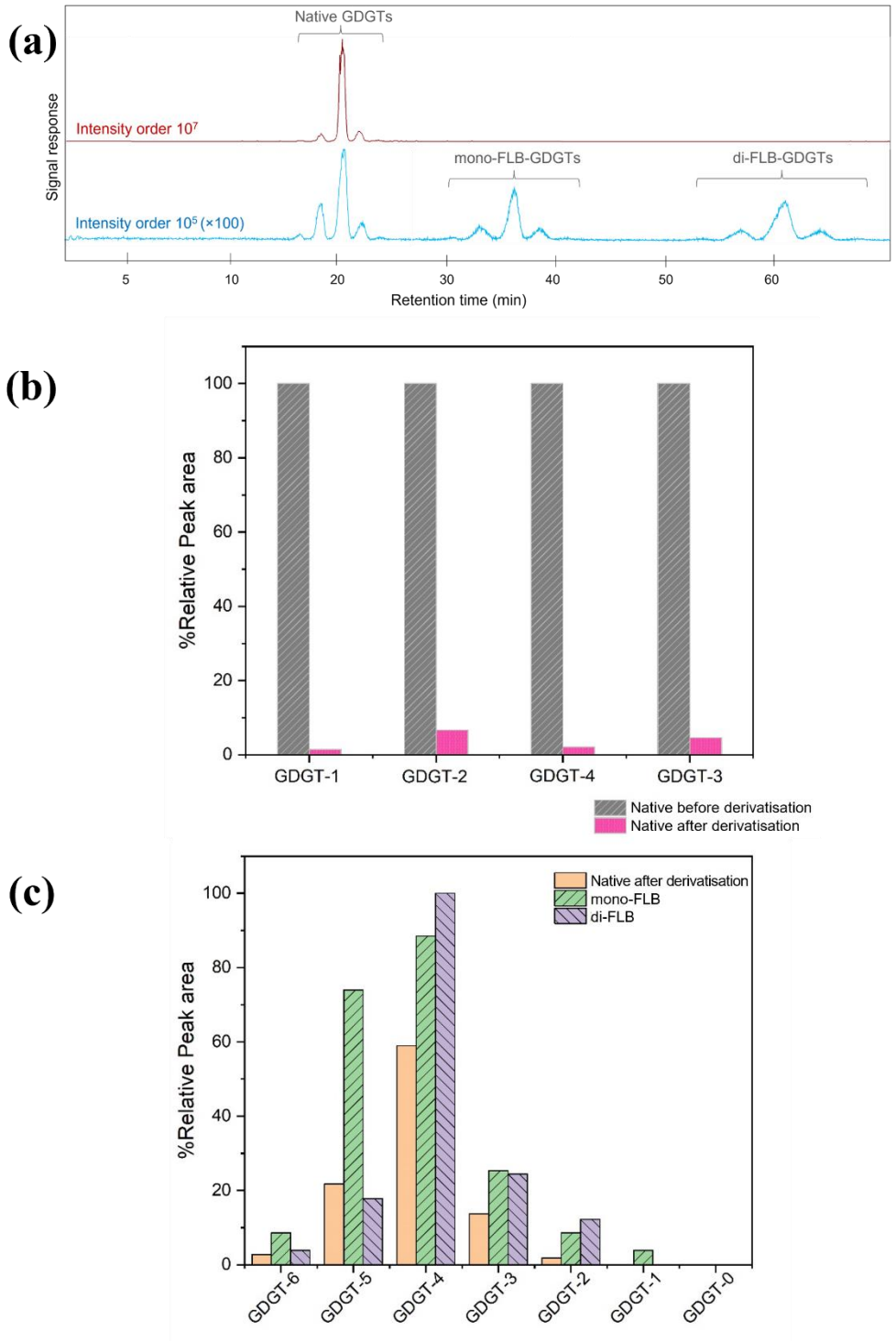


Figure 5.11. (a) APCI base peak chromatograms of native GDGTs from *S. acidocaldarius* MR31 (upper trace) and FLB derivatives (lower trace) obtained using the same molar concentration of extract, (b) Peak area of MS signals of native GDGTs before derivatisation and native GDGTs after derivatisation, (c) Peak area of MS signals of the FLB-GDGTs in different species. The FLB-GDGTs was prepared using derivatising agents of FLB: DMAP: EDC as mole ratio of 10: 5: 10 to 1 mole of GDGTs).

The efficiency of the derivatisation was estimated by comparing MS peak areas of the native GDGTs before and after derivatisation (Figure 5.11b). The proportion of peak area of native form of GDGT-2-6 remaining in the BPC of FLB derivatives shows incomplete conversion to the esters. Notably, the response for the BPC of the FLB-GDGTs is about 100 times lower than that of the native GDGTs (Figure 5.11a). The peak areas of the native mono-FLB and di-FLB GDGTs were expressed as relative percentage areas to reflect the MS response for each species (Figure 5.11c). The most abundant components are di-FLB and the native GDGTs are least abundant. Given that the native GDGTs represent less than 10% of the extract that was derivatised, the signal intensity of the mono-FLB and di-FLB are surprisingly low when compared to the loss of MS signal of the native form after derivatisation (see Section 5.3.2.6).

5.3.2.4 Formation of Fmoc-Lys-Boc (FLB) GDGTs

5.3.2.4.1 Full MS structural characterisation of GDGTs

The mass spectrum of each peak in the chromatogram of the purified FLB-GDGTs (Figure 5.11a, blue trace) was analysed to identify the products (Figure 5.12, Table 5.3). Lipid group A of (Figure 5.11a) were confirmed as the native GDGTs (Figure 5.1 and 5.12a). The mass spectra obtained from lipids group B (Figure 5.12b, Table 5.3) show a series of protonated molecules in the range of *m/z* 1743.4-1751.4 and base peak fragment ions in the range *m/z* 1643.5-1651.5 corresponding to the loss of Boc group (-100 Da), which compares favourably with the FLB derivatives (Poplawski 2017). Group B also form moderate intensity ions: in the range *m/z* 1292.4-1302.4, corresponding to loss of the FLB residue (−450 Da), confirming the presence GDGTs as the derivative form. Two other moderate intensity ions in the mass spectra of group in the range *m/z* 1447.5-1455.5 represent losses of the Boc and alkyl fluorenyl groups (-296 Da), and in the range *m/z* 1478.9-1473.4 (-271 Da), presently not assigned. Group B FLB-GDGTs also show minor intensity ions in the range *m/z* 1674.5-1670.5, corresponding to neutral loss from a free hydroxyl group (-$C_3H_6O_2$; −74 Da) from the protonated molecule. Thus, the mass spectra of group B FLB-GDGTs show fragmentation characteristic both of FLB derivatives and of the

native GDGTs (Figure 5.13a), confirming the formation of GDGTs in the form of mono-FLB ester (mono-FLB-GDGT).

The mass spectra of the FLB derivatives of group C GDGTs (Figure 5.12c, Table 5.3) exhibit base peak ions in the range *m/z* 2093.7-2098.8, corresponding to loss of the Boc group (-100 Da). The group C FLB derivatives in Table 5.3 show a series of fragment ions in the range of *m/z* 1293.5-1302.4, indicating the presence of *i*-GDGT 0-4 in their structures (-900 Da) (see fragmentation origins of the structure in Figure 5.13b). The mass spectra of group C FLB-GDGTs also show other moderate intensity ions: in the range *m/z* 1446.5-1448.5 (-746 Da), in the range *m/z* 1474.5-1476.6 (-720 Da) and in the range *m/z* 1628.6-1631.5 (-566 Da), all presently not assigned. The characteristic losses to form the fragment ions confirm the formation of the desired product in the form of diester (di-FLB-GDGTs).

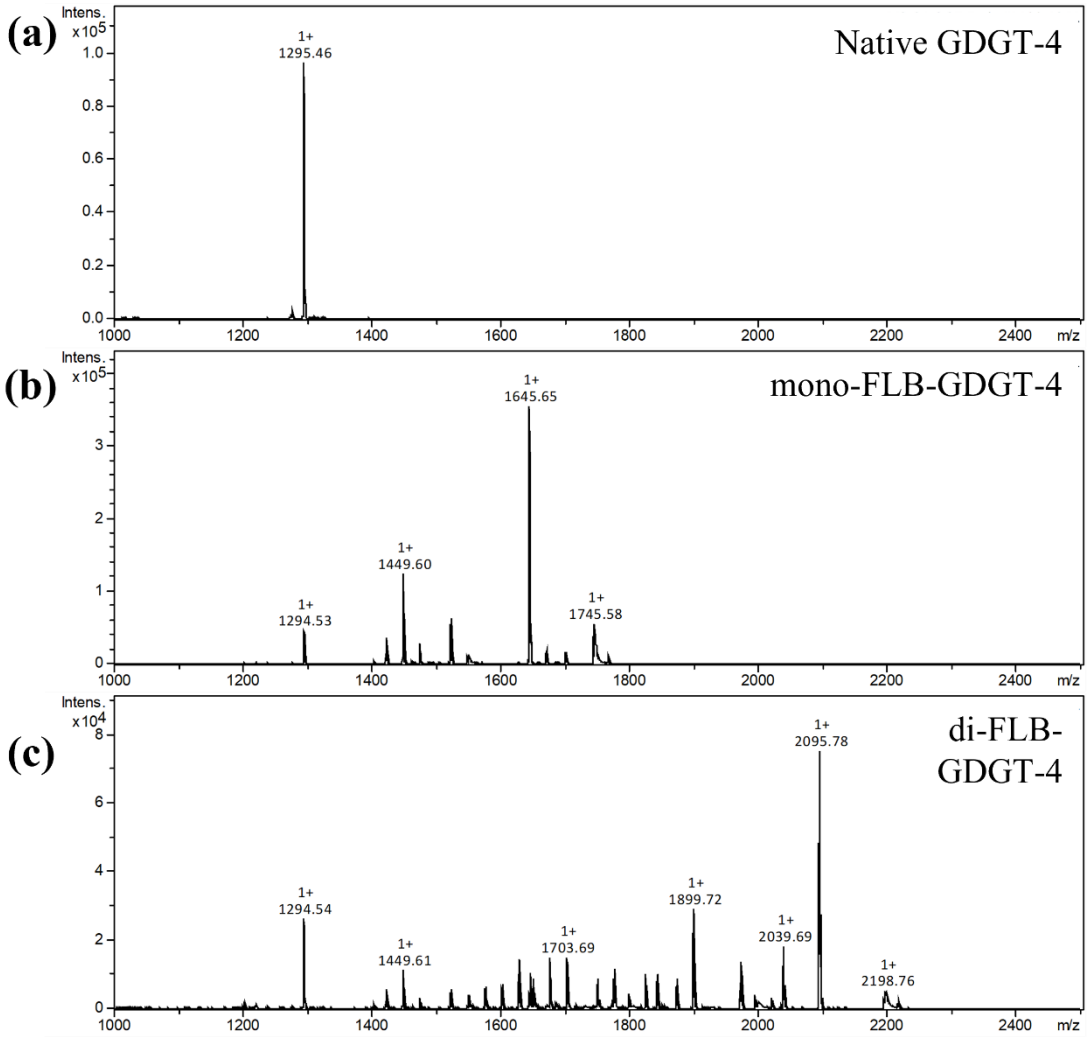


Figure 5.12. Full APCI MS spectra of GDGT-4 of lipid extract from *S. solfataricus* in different forms (a) native form: GDGT 4, (b) mono-FLB-GDGT-4, (c) di-FLB-GDGT-4, acquired using an ion trap spectrometer operated over the range of *m/z* 1000-2500.

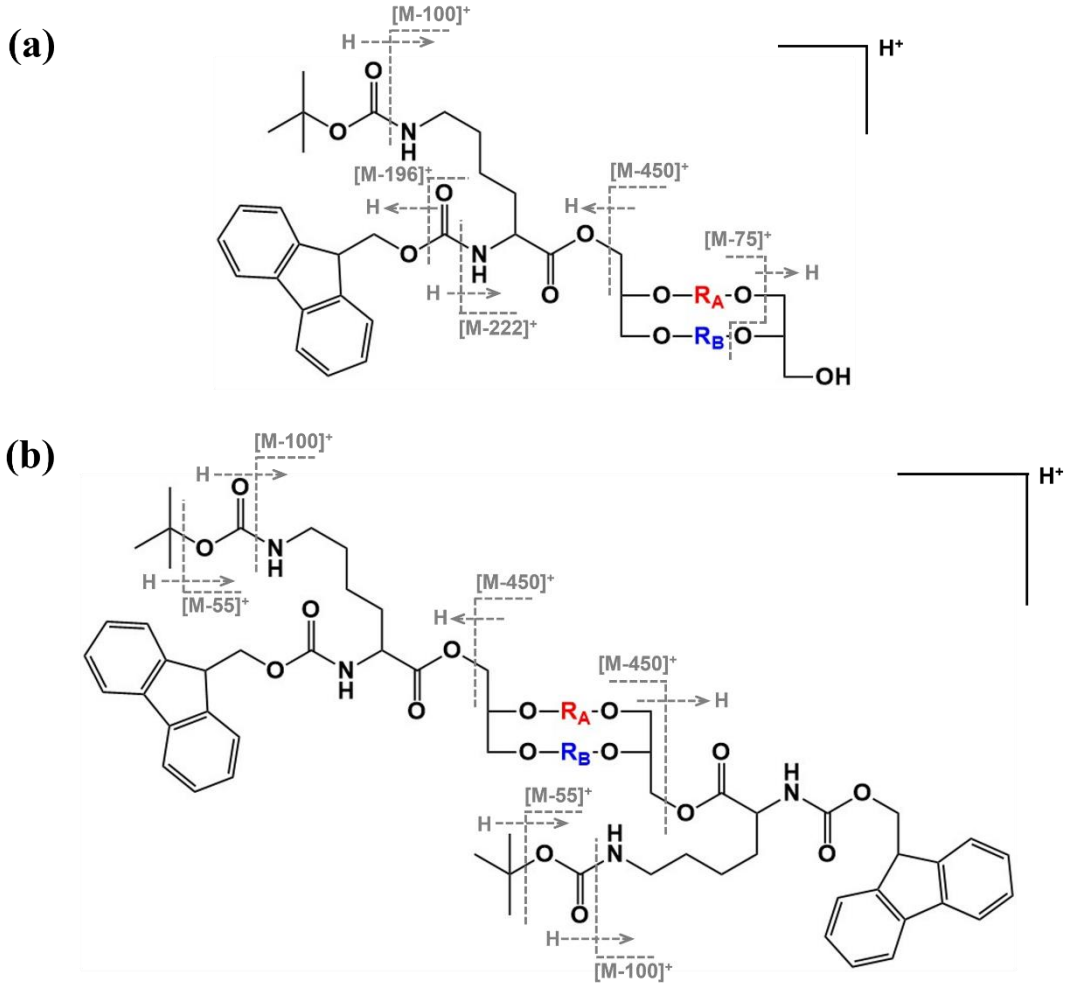


Figure 5.13. Origins of fragment ions in the APCI MS spectrum of GDGTs derivatives in the form of (a) mono-FLB-GDGT, (b) di-FLB-GDGTs. R_A, R_B = different etherified chains of diglycerol tetraether lipid cores.

5.3.2.4.2 MS^2 dissociation pathways of FLB derivatised GDGTs

5.3.2.4.2.1 Mono-FLB-GDGTs

The dissociation pathways for mono-FLB-GDGT were characterised by tandem MS (MS^2) (Table 5.3, Figure 5.14). The MS^2 spectra were scanned in auto mode to obtain both full MS and MS^2 spectra of the components in the same run. For example, the full MS spectrum of mono-FLB-GDGT-4 (Figure 5.14a) contained the same ions as observed in Figure 5.12b, the difference in relative intensities most likely reflect variations in the energetics of ionisation between runs. The MS^2 spectrum of the precursor ion at *m/z* 1645.6 (Figure 5.14b), representing the mono-FLB-

GDGT-4 without Boc group, produced the base peak product ion at *m/z* 1448.6 and a moderate intensity ion at *m/z* 1405.6, arising from the losses of the alkyl fluorenyl group (-196 Da) and the entire Fmoc group (-222 Da), respectively. The MS^2 spectrum of the precursor ion at *m/z* 1448.6 (Figure 5.14c), representing the mono-FLB-GDGT-4 without Boc and alkyl fluorenyl groups, gave the base peak product ion at *m/z* 1431.5 and a moderate intensity ion at *m/z* 1294.5, reflecting the cleavage of ammonia from the lysine chain (-17 Da) and the presence of *i*-GDGT-4, respectively. The precursor ion at *m/z* 1294.5 (Figure 5.14d) produced the similar MS^2 spectrum to those obtained from the MS^2 spectrum of native *i*-GDGT-4 (Figure 5.4b), confirming the presence of *i*-GDGT-4 as the FLB derivative. The MS^2 spectra of all of the other protonated molecules in the range *m/z* 1751.6-1743.6 exhibited the same manner of MS^2 dissociation as those obtained from mono-FLB-GDGT-4 giving rise to ions differing according to their number of cyclopentyl rings (Table 5.3), confirming the formation of mono-ester form of *i*-GDGT-1-5. Assignment of the MS^2 spectra of the mono-FLB-GDGT reflects a good match to the fragmentation behaviour of the derivatising group observed for the FLB derivatives of sterols (Chapter 3).

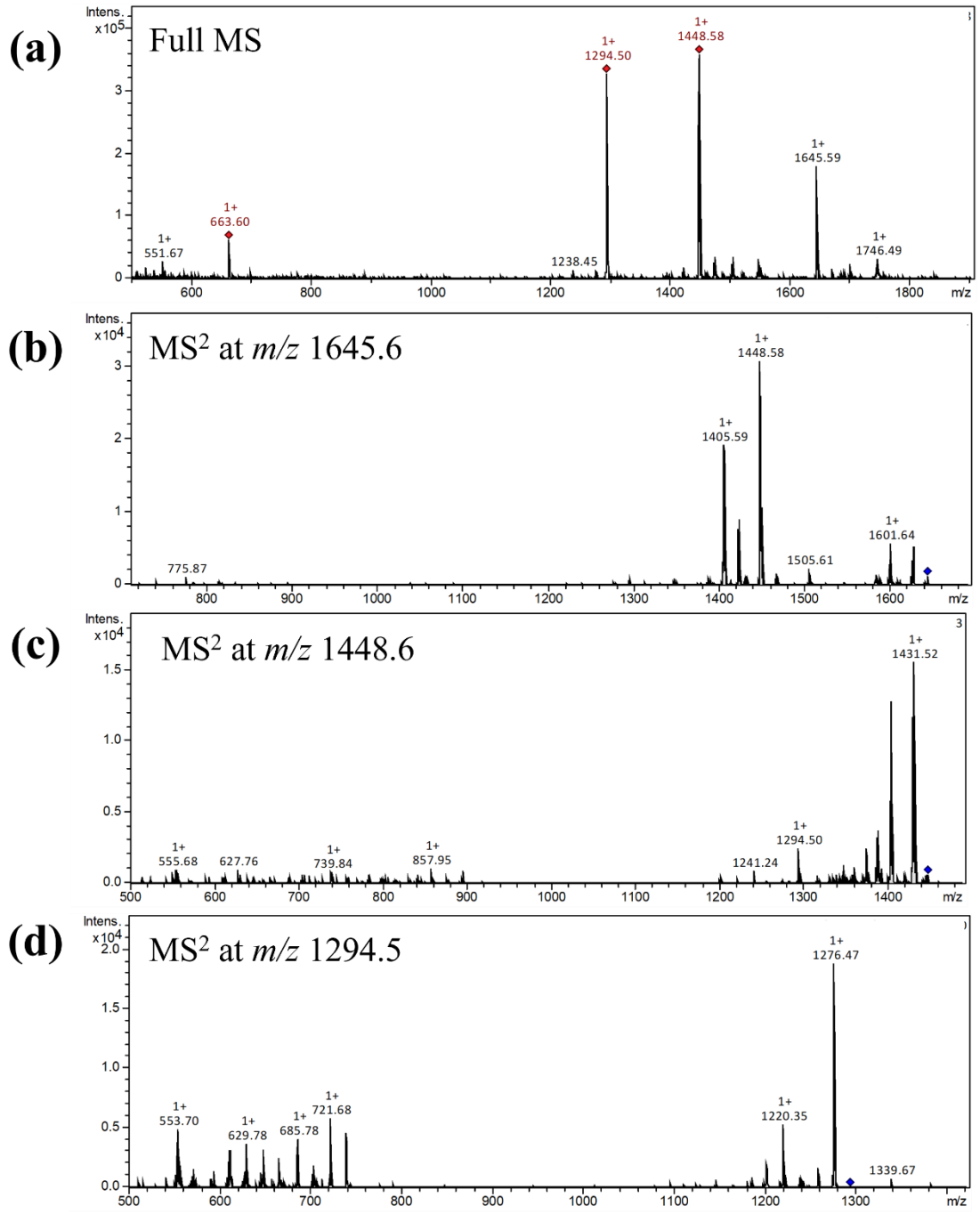


Figure 5.14. APCI MS and MS^2 spectra of mono-FLB GDGT-4 isolated from *S. solfataricus* (mono-FLB-GDGT-4) (a) Full MS and MS^2 spectra from precursors: (b) *m/z* 1645.6, (c) *m/z* 1488.6 and (d) *m/z* 1294.5. All spectra were recorded over the range of *m/z* 500-2500.

Table 5.3. Mass spectral data and identifications of FLBs derivatised with diglycerol tetraether lipid cores obtained from the *S. solfataricus* extracts.

Lipid	$[M+H]^+$ (*m/z*)	Characteristic fragment ions (*m/z*)	Precursor ion (*m/z*)	Characteristic product ions (*m/z*)
GDGTs				
1_a	1300.5		N.D.	N.D.
2_a	1298.4		1298.4	724.0, 721.7
3_a	1296.5		1296.5	739.8
4_a	1294.5		1294.5	741.8, 739.7
5_a	1292.5		N.D.	N.D.
mono-FLB-GDGTs				
1_a	1751.6	1651.7, 1454.6, 1301.5	N.D.	N.D.
2_a	1749.5	1674.51, 1648.6, 1553.5, 1478.85, 1452.6, 1298.5	1648.7	1452.5, 1409.6
			1298.5	1280.4, 1220.5, 721.8, 555.8
3_a	1747.5	1672.5, 1647.6,1477.4, 1450.6, 1296.5	1646.5	1450.5, 1407.5
			1296.5	1278.4, 1220.5, 739.8, 721.8, 553.8
4_a	1745.5	1670.5, 1645.5, 1474.5, 1448.6, 1294.4	1645.5	1448.6, 1405.6
			1294.5	1276.4, 1220.5, 739.8, 721.5, 553.8
5_a	1743.6	1643.6, 1473.4, 1292.4	N.D.	N.D.
di-FLB-GDGTs				
2_a	2199.8	2098.8, 1298.5	2098.8	2043.7, 1999.7, 1649.6
3_a	2197.8	2096.8, 1631.5, 1476.6, 1450.6, 1296.5	2096.8	2041.7, 1997.7, 1801.6, 1647.2
			1296.5	1278.4, 1220.5, 739.8, 721.8, 553.8
4_a	2195.8	2094.7, 1628.6, 1474.6, 1448.6, 1294.4	2094.7	2039.7, 1995.7, 1799.7, 1645.6
			1294.5	1276.4, 1220.5, 739.8, 721.5, 553.8
5_a	2193.8	2093.7, 1292.4	N.D.	N.D.

GDGT = glycerol dialkyl glycerol tetraether, N.D. = Not determined, base peak ions were underlined.

5.3.2.4.2.2 Di-FLB-GDGTs

The dissociation pathways of the di-FLB-GDGTs were characterised by tandem MS (MS^2) (Table 5.3, Figure 5.15). For example, the full mass spectrum of di-FLB-GDGT-4 (Figure 5.15a) gave the matched fragment ions with those obtained from Figure 5.12c: at *m/z* 2094.8 and at *m/z* 1448.6, corresponding to the loss of one Boc group (-100 Da) and losses of one FLB, Boc and alkyl fluorenyl groups (-746 Da), respectively. Significant differences in relative abundance may be attributable to the low concentrations in the incompletely derivatisation fraction. The MS^2 spectrum of the fragment ion at *m/z* 2094.8 (Figure 5.15b), representing the di-FLB-GDGT-4 with loss of one Boc group, gave base peak product ion at *m/z* 1995.7 corresponding to the loss of another Boc group. The fragment ion at *m/z* 2094.8 also produced three product ions of minor intensity: *m/z* 2039.7 arising from the loss of *tert*-butyl group (-55 Da), *m/z* 1645.5 corresponding to entire group of another FLB (-450 Da), and *m/z* 1799.7 showing simultaneously cleavages of one Boc and the alkyl fluorenyl groups (-296 Da). The MS^2 spectrum of precursor ion *m/z* 1448.6 (Figure 5.15c), representing the di-FLB-GDGT-4 following loss of an entire FLB group, Boc and the alkyl fluorenyl group, produced the base peak product ion at *m/z* 1404.5 (-44 Da) possibly arising from the loss of either a carboxyl group by rearrangement or a C_2H_6N radical from the lysine chain. High resolution MS would be required to distinguish the possibility (Glish and Vacher 2003; Perry *et al.* 2008; Zubarev and Makarov 2013). A moderate intensity ion at *m/z* 1294.5 formally corresponding to the *i*-GDGT-4 protonated molecule. The product ions from *m/z* 1294.5 (Figure 5.15d) matched those observed in the MS^2 spectra from *m/z* 1294.5 for mono-FLB-GDGT-4 (Figure 5.14d) and native *i*-GDGT-4 (Figure 5.4b), confirming the GDGT-4 lipid core and presence of two FLB groups in the structure. The MS^2 spectra of the other precursor ions in the range *m/z* 2098.8-2094.8 exhibited similar MS^2 dissociation as those obtained from di-FLB-GDGT-4 (Table 5.3), confirming the presence of di-ester from of GDGT giving rise to ions differing according to their degrees of cyclopentanyl rings.

All full MS and MS^2 spectra of GDGTs obtained after derivatised with FLB in the ratio of 10: 3: 10 (FLB: DMAP: EDC to 1 mole of diol compound) have revealed the presence of GDGTs in three different forms (native, mono-ester and di-ester forms), reflecting incomplete derivatisation.

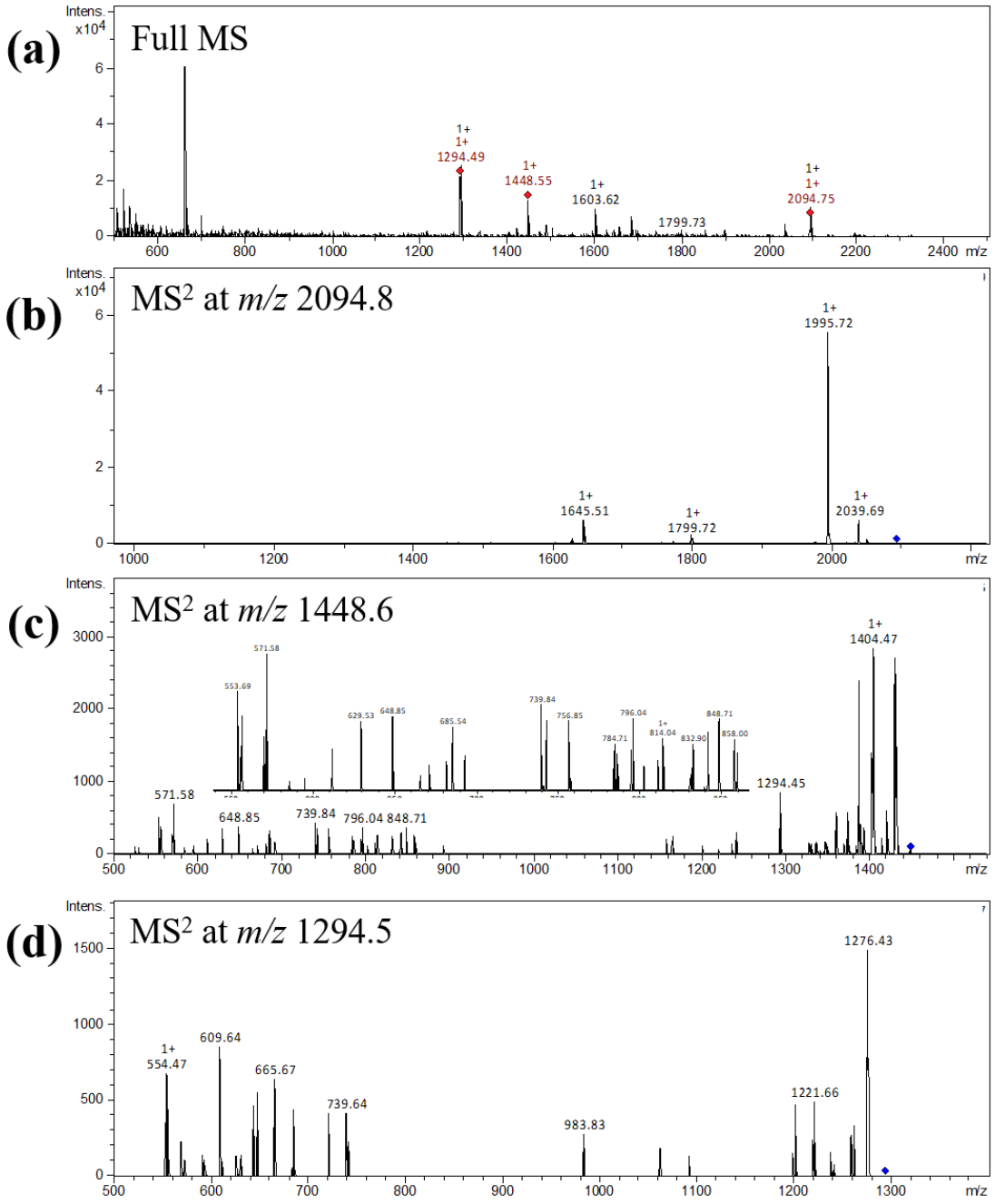


Figure 5.15. APCI MS and MS^2 spectra of di-FLB-GDGT-4 isolated from *S. solfataricus* (a) Full MS and MS^2 spectra from precursors: (b) *m/z* 2094.8, (c) *m/z* 1448.6 and (d) *m/z* 1294.5. All spectra were recorded over the range *m/z* 500-2500.

5.3.2.5 Evaluation of the amount of derivatising agent required for GDGTs extracts

The FLB derivatisation of GDGT extracts was re-evaluated in order to obtain the maximum yield of the di-derivatised GDGTs. The proportions of the derivatising agents to GDGT extract were

prepared in various mole ratios (Table 5.4). The products from each reaction were separated by UHPLC with MS detection (Figure 5.16). In order to reflect the effectiveness of the derivatisation process, the measured peak area of the MS response of each product was compared with that obtained from the GDGT extract before derivatisation (GDGT: FLB: DMAP: EDC ratio of 1:0: 0: 0 in Table 5.4).

Table 5.4. Mole ratios of GDGT extract and derivatising agents (FLB, DMAP, EDC).

Mole ratio of GDGTs: FLB: DMAP: EDC	**GDGTs* (moles)**	**FLB (moles)**	**DMAP (moles)**	**EDC (moles)**
1: 0: 0: 0	4.34×10^{-6}	-	-	-
1: 1.5: 0.4: 2	4.34×10^{-6}	5.62×10^{-6}	1.73×10^{-6}	8.27×10^{-6}
1: 10: 5: 19	3.56×10^{-7}	3.54×10^{-6}	1.72×10^{-6}	6.68×10^{-6}
1: 20: 17: 27	9.60×10^{-7}	1.92×10^{-5}	1.50×10^{-5}	2.57×10^{-5}
1: 23: 19: 24	4.34×10^{-6}	1.02×10^{-4}	8.19×10^{-5}	1.04×10^{-4}
1: 35: 28: 36	2.91×10^{-6}	1.02×10^{-4}	8.19×10^{-5}	1.04×10^{-4}
1: 70: 56: 72	1.46×10^{-6}	1.02×10^{-4}	8.19×10^{-5}	1.04×10^{-4}

**The moles of GDGTs extract was estimated based on the weight of the obtained samples and divided by the molecular weight of i-GDGT-4 (1294.5 g/mol) as the most dominant component in the soil extract.*

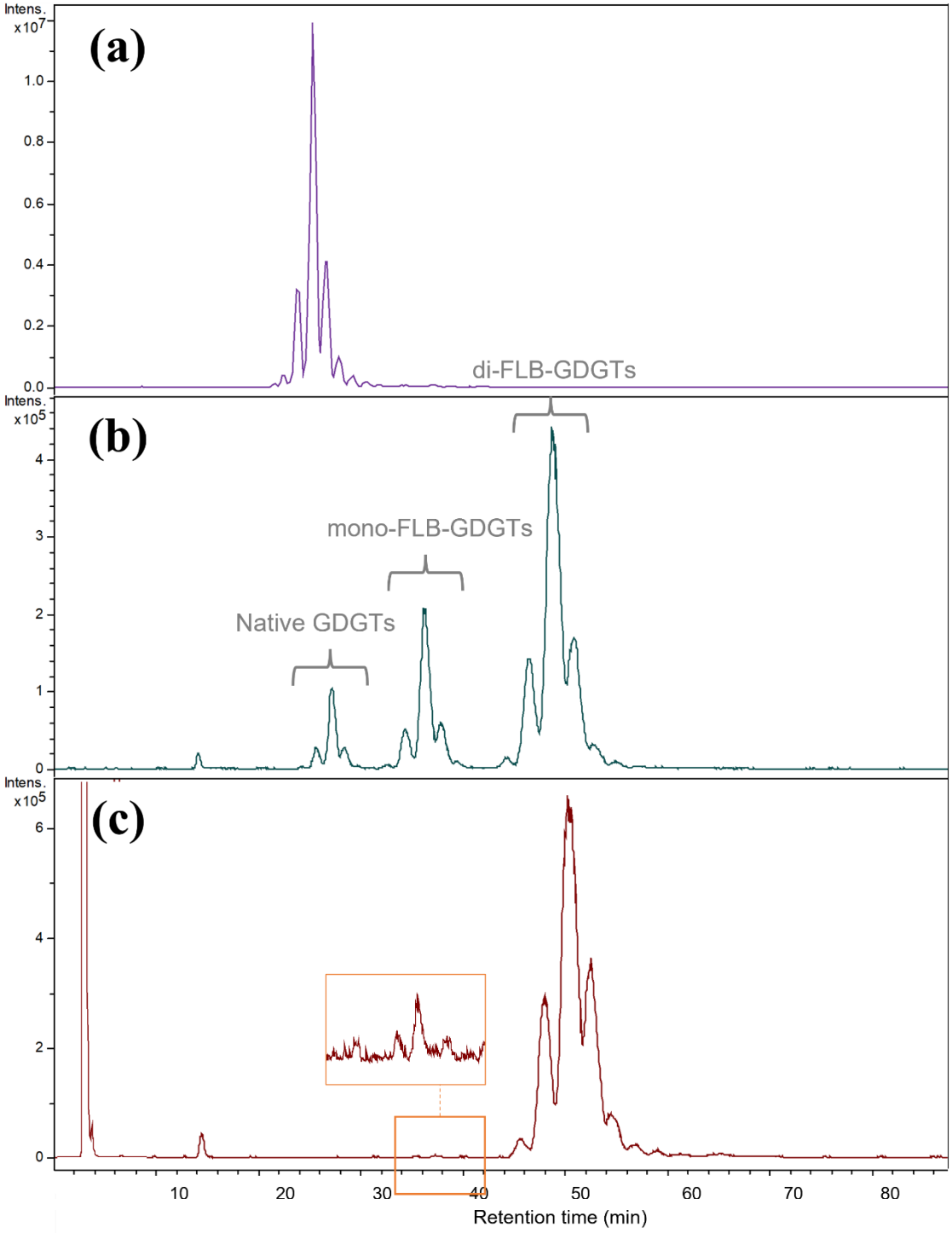


Figure 5.16. Representative APCI base peak chromatograms of GDGTs: (a) native GDGTs (b) and (c) FLB-GDGTs prepared using two different mole ratios of derivatising agents (FLB: DMAP: EDC) to 1 mole of GDGTs, 20: 17: 27 and 23: 19: 24, respectively. In all cases, FLB-GDGTs was normalised to the molar concentration of GDGTs used in the reaction.

Derivatisation of GDGTs is indicated by an obvious decrease in the MS signal response of the native GDGTs with increasing proportions of derivatising agent (Figure 5.17a). The incomplete derivatisation of the GDGT extract under the initial reaction conditions is also evident from the increase in the overall signal responses of mono-FLB-GDGTs and di-FLB-GDGTs when the proportions of derivatising agent was increased (Figure 5.17b). An increase in the amount of derivatising agents gave greater MS response of the mono-FLB-GDGTs, the highest value being when the amount of FLB reached 20 mole equivalents of GDGTs, the yield decreasing significantly at above 23 mole equivalents (Figure 5.17b). Similarly, the MS response of di-FLB-GDGTs was slightly increased and highest when the amount of FLB reached 23 mole equivalents of GDGTs (Figure 5.17c). The amount of FLB was raised further up to 35 mole equivalents and gave a decrease in response of the di-FLB-GDGTs (Figure 5.17c) with a corresponding increase in the response for reagent residues (data not shown) accompanying (and causing) MS signal suppression. The ratio of derivatising agents of 23: 19: 24 (FLB: DMAP: EDC to 1 mole of GDGTs extract) gave predominantly the di-ester form of GDGTs, reflected in the high response of di-FLB-GDGTs with minimum response for mono-FLB-GDGTs and native species.

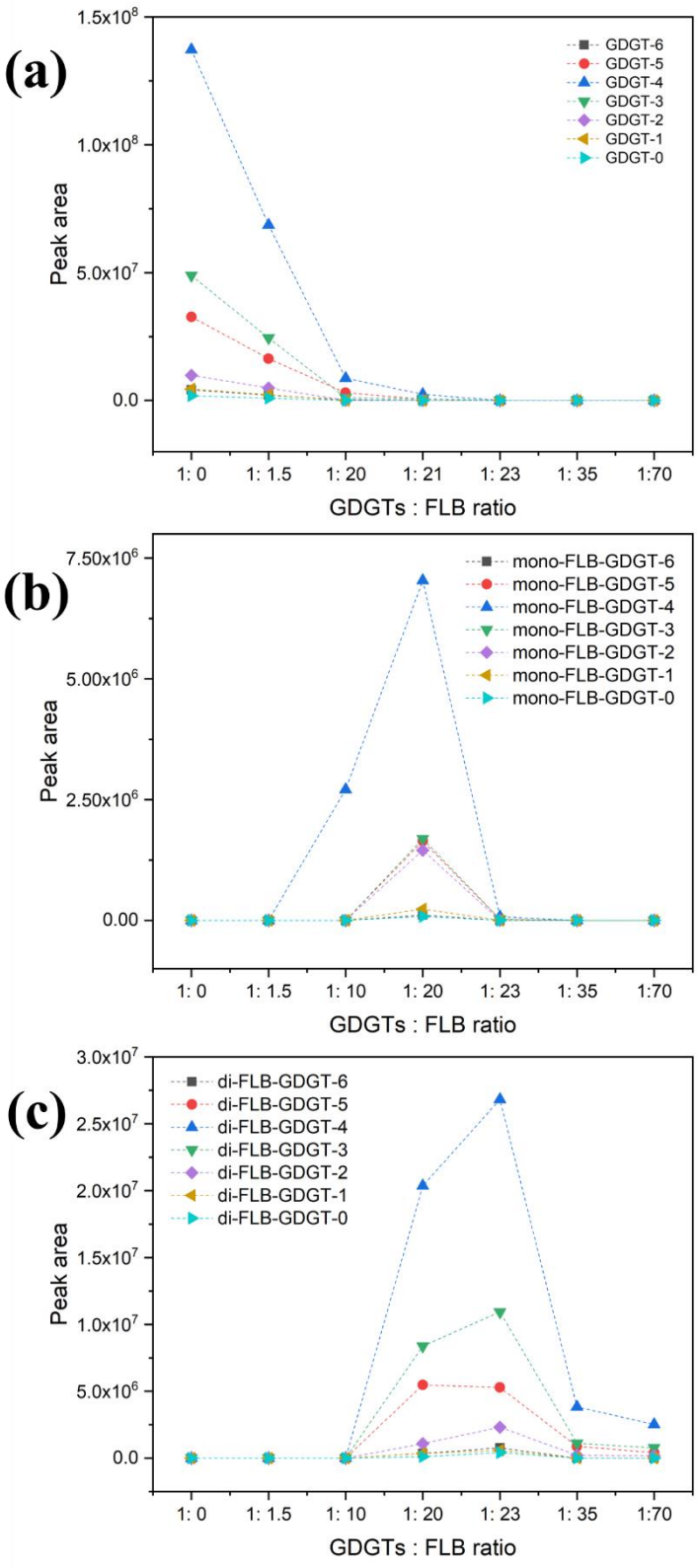


Figure 5.17. MS peak areas of the esterified FLB-GDGTs using various ratios of reagents (mole ratio as GDGTs: FLB) corresponding as the different GDGT lipid species (a) native, (b) mono-FLB-GDGTs, (c) di-FLB-GDGTs. In all cases, FLB-GDGTs was normalised to the molar concentration of GDGTs used in the reaction.

The MS response of the GDGT derivatives was compared with the response for the native GDGTs by peak area measurement before and after derivatisation (Figure 5.18). The GDGT lipids were prepared at the same molar concentration of extract to obtain a comparable result. The MS signal response of FLB-GDGTs was in the range 16-23% compared to that obtained from the native GDGTs, representing a significant decrease in response rather than the expected increase.

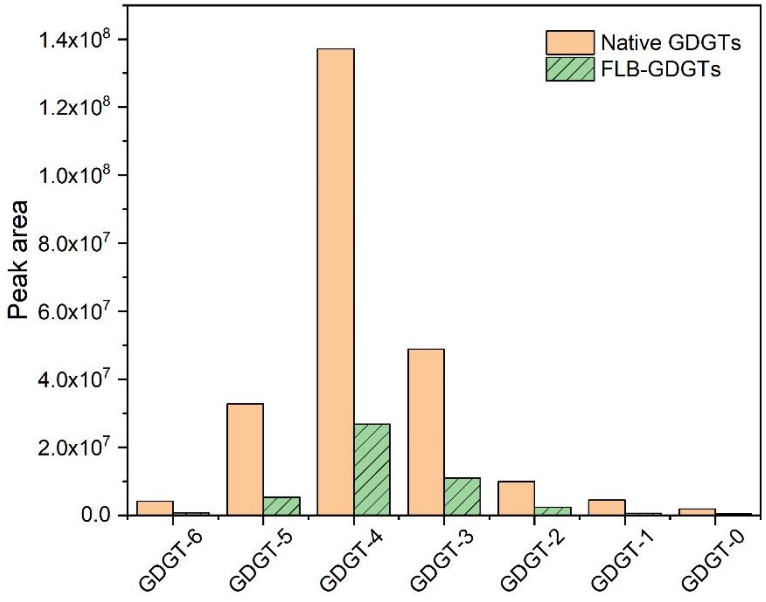


Figure 5.18. Peak area of MS signals of GDGTs before and after derivatisation obtained using the same molar concentration of extract. The FLB-GDGTs was prepared using derivatising agents of FLB: DMAP: EDC in mole ratios of 23: 19: 24 to 1 mole of GDGTs.

5.3.2.6 Examination of the MS responses of FLB-GDGTs

Structural analysis of large compounds (MW >1000 Da) such as GDGTs by MS can be challenging due to incomplete ionisation, potentially leading to suppression of the response of GDGTs and unreliable quantification (Lengger *et al.* 2018). The response of FLB-GDGTs being around 10 times less than that of the native GDGTs (Figure 5.18) may in part be associated with their high molecular weight (>2000 Da). An attempt to increase the signal for the FLB-GDGTs

was explored by examination of the ionisation conditions including APCI vaporiser temperature (Section 5.3.2.6.1) and the use of electrospray ionisation (ESI) (Section 5.3.2.6.2).

5.3.2.6.1 Effect of vaporiser temperatures on signal response of di-ester-GDGTs

The effect of changing the APCI MS vaporiser temperature was examined using the FLB derivatised GDGT extract from *S. acidocaldarius* MR31. The temperature was varied either side of the normal operating conditions of 450°C under which conditions di-FLB-GDGT-4 ($[M\text{-}Boc]^+$ at *m/z* 2095.5) was the main component in the full MS spectra (Figure 5.19c). In this study, the FLB-GDGTs (1 µL) were detected following direct infusion into the APCI source in the same solvent and at the flow rate used for UHPLC analysis. The vaporiser temperature was set to temperatures of 400, 425, 450, 460, 470, 480, 490 and 500°C, while the other ionisation conditions (dry gas temperature, nebuliser temperature) were held constant to obtain the comparable results. The MS spectra were recorded over the range *m/z* 1900-2500 in order to focus only on the signal response of di-FLB-GDGT species (Figure 5.19). At all vaporiser temperatures the FLB-GDGTs show the dominant peak at *m/z* 2095, corresponding to the base peak of di-FLB-GDGT-4 arising from loss of the Boc group (-100 Da). The signal response of FLB-GDGTs corresponding to the different vaporisation temperatures was assessed by comparing intensities of peaks at *m/z* 1973, *m/z* 2039, *m/z* 2095, *m/z* 2117 and *m/z* 2217 (Figure 5.20). These peaks relate to fragmentation of the FLB-GDGTs at different positions in their structures. The dominant peaks at *m/z* 1973 and *m/z* 2095 correspond to cleavages of Fmoc group (-222 Da) and Boc group (-100 Da) from di-FLB-GDGT-4, respectively, while two sodium adduct peaks arise from di-FLB-GDGT-4 without Boc group (*m/z* 2117) and di-FLB-GDGT-4 (*m/z* 2217), respectively. The response of the peak at *m/z* 2095 obtained from the current condition at 450°C increases substantially with increasing vaporiser temperature (Figure 5.20). A greater ion intensity for the FLB-GDGTs was obtained when the vaporisation temperature was increased above 480°C, reflecting better ionisation of the derivatives. An accompanying increase occurred in the intensities of the fragment ions with that arising from loss of Fmoc (*m/z* 1973)

becoming similar in intensity to *m/z* 2095. This maybe either a consequence of better ion transmission from the source or it may reflect more energy being imparted during ionisation, resulting in the ion being more vibrationally excited (Cooks *et al.* 1991; March 1997). The vaporisation temperature of 500°C was therefore selected as this is the maximum limit of the APCI MS detector.

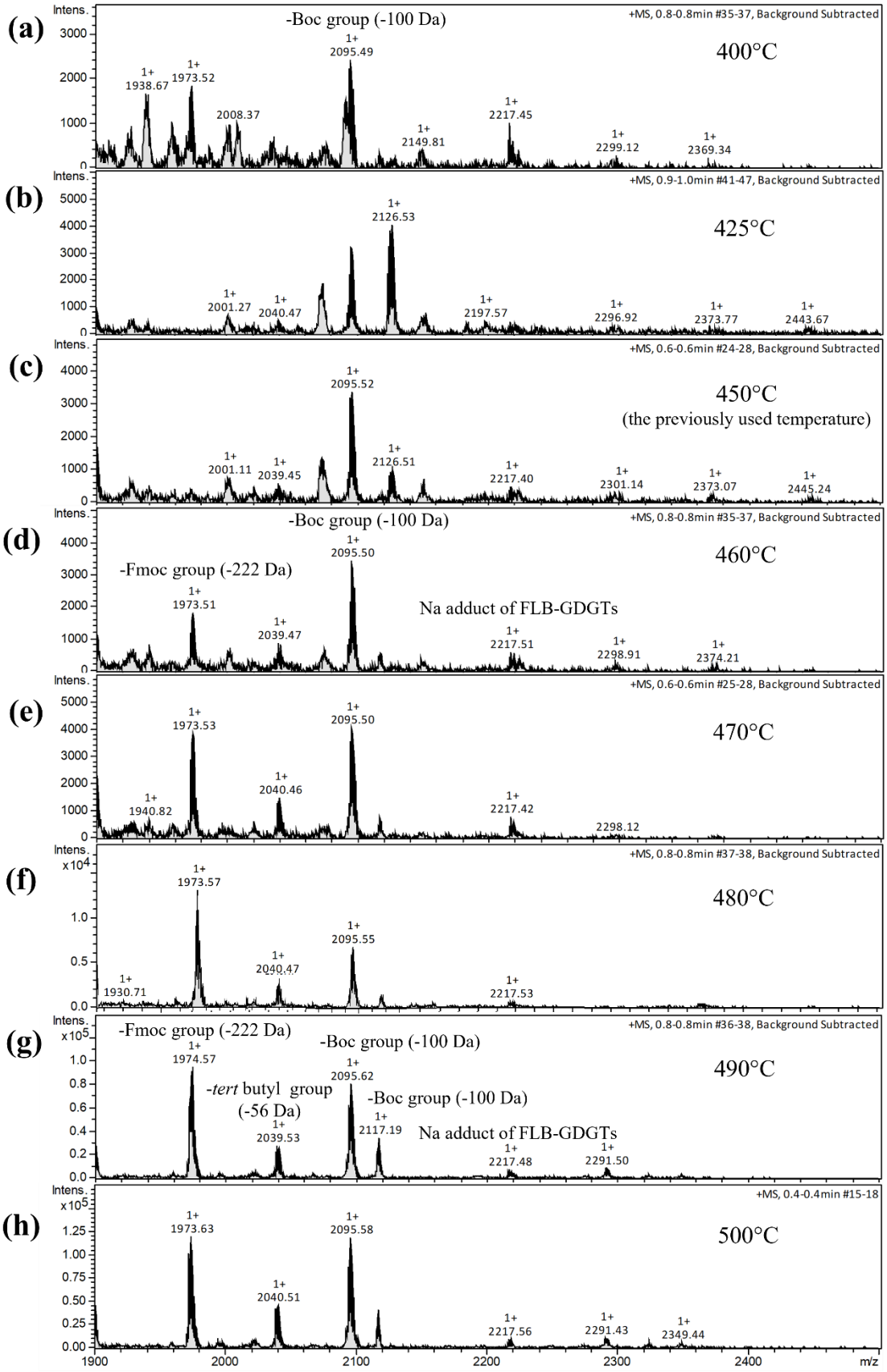


Figure 5.19. APCI-MS spectra of FLB derivatives of GDGT extract from *S. acidocaldarius* MR31 detected at various vaporisation temperatures (400-500°C).

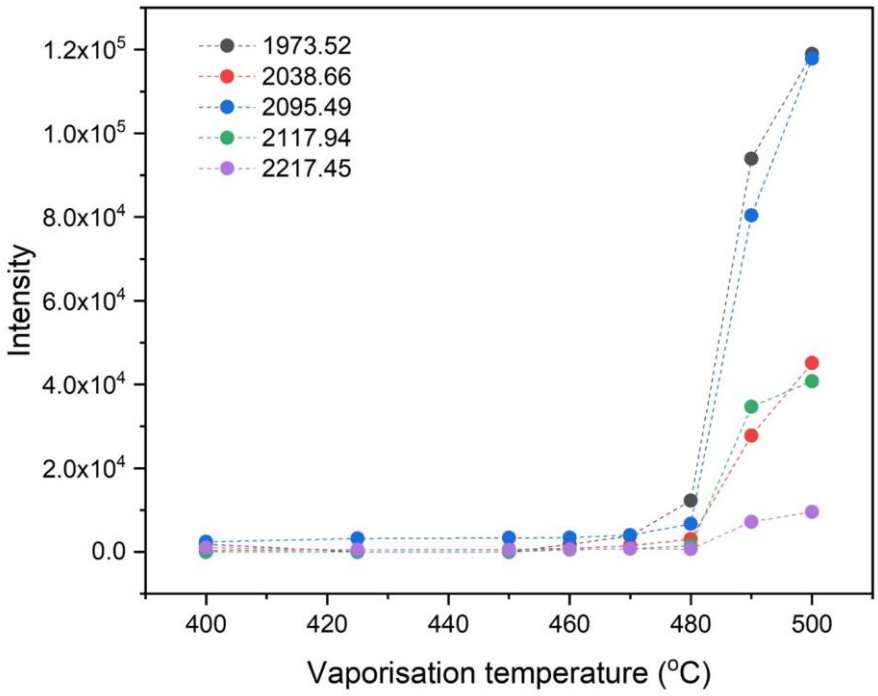


Figure 5.20. Peak area of FLB GDGT derivatives of an extract from *S. acidocaldarius* MR31 detected at different *m/z* values and ionised using different vaporisation temperatures (400-500°C).

To enable comparison of the behaviour of the native GDGTs and the esters, chromatographic separation of the incompletely derivatised extract from *S. acidocaldarius* was carried out with the vaporisation temperature set to 500°C (Figure 5.21b). Compared with the vaporisation temperature of 450°C (Figure 5.21a), the peak areas of the di-FLB-GDGTs were greater whereas those of the mono-FLB-GDGTs and the native species were lower (Figure 5.22) reflecting the improved ion transmission for the larger molecule, di-FLB-GDGT.

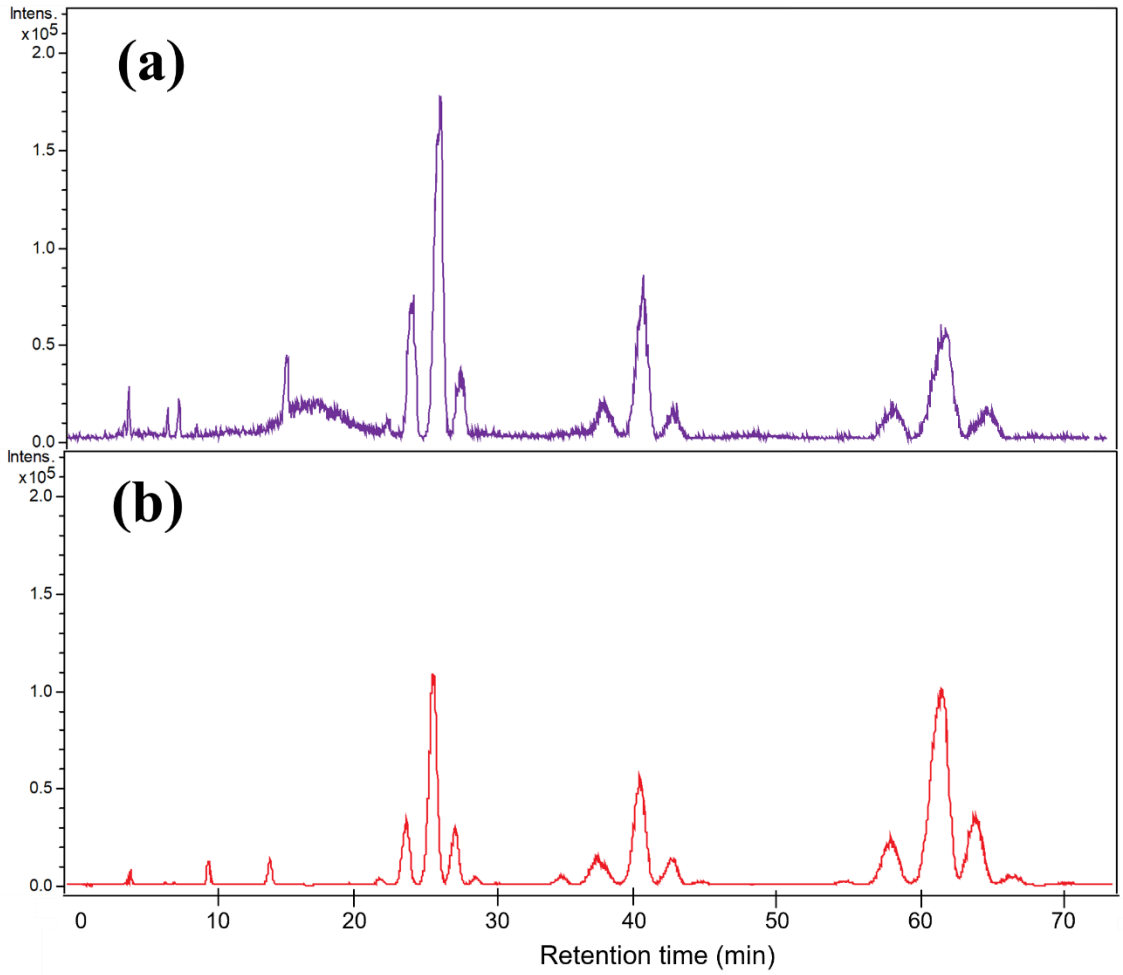


Figure 5.21. APCI base peak chromatograms of FLB-GDGTs from *S. acidocaldarius* MR31 obtained using the different vaporisation temperatures (a) 450°C, (b) 500°C. The FLB-GDGTs was prepared using derivatising agents of FLB: DMAP: EDC in mole ratios 20: 17: 27 to 1 mole of GDGTs. In all cases, the FLB-GDGT was normalised to the same molar concentration of extract used in reaction.

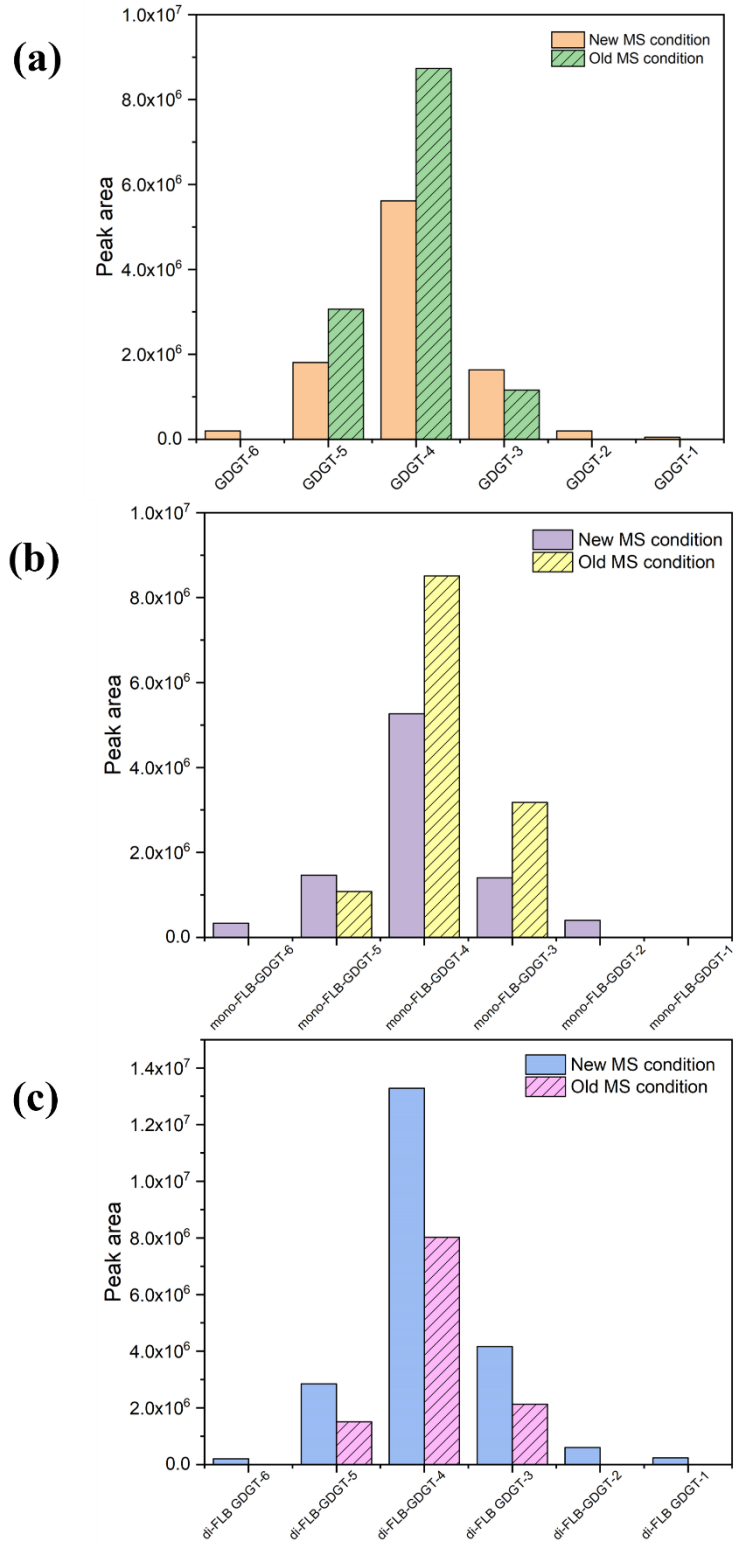


Figure 5.22. Peak area of MS signals of FLB-GDGTs species from *S. acidocaldarius* MR31 in different species (a) native form, (b) mono-FLB-GDGTs, (c) di-FLB-GDGTs. The FLB-GDGTs was prepared using derivatising agents of FLB: DMAP: EDC as mole ratio of 20: 17: 27 to 1 mole of GDGTs.

5.3.2.6.2 Detection of FLB-GDGTs by electron spray ionisation (ESI)

The use of electrospray ionisation (ESI) to detect FLB-GDGTs was evaluated. The incompletely derivatised GDGT extract from *S. acidocaldarius* MR31, previously analysed by UHPLC-APCI MS (Figure 5.21b), gave an ESI base peak chromatogram in which the most abundant components are di-FLB-GDGTs (Figure 5.23). The response of FLB-GDGTs by ESI was evaluated by comparing the summed peak areas of di-FLB-GDGTs with those obtained from APCI (Figure 5.24). The peak areas of di-FLB-GDGTs species from ESI were greater than those obtained from APCI by approximately 1,000 times with signal to noise ratio (*S*/*N*) about 7 times higher, indicating better ionisation efficiency for the FLB derivatives. The detection of FLB-GDGTs by ESI was therefore selected as an alternative detection method for subsequent analyses of GDGTs (see Section 2.6.1).

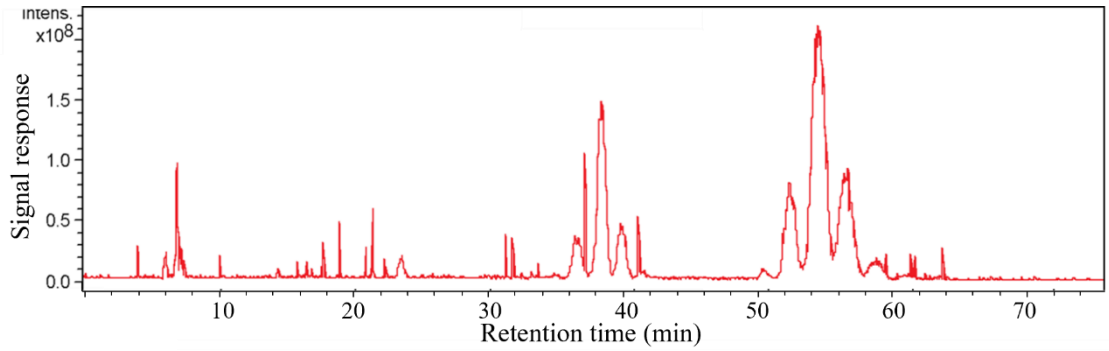


Figure 5.23. ESI base peak chromatograms of FLB-GDGTs from *S. acidocaldarius* MR31. The FLB-GDGTs were prepared using derivatising agents of FLB: DMAP: EDC in mole ratios 20: 17: 27 to 1 mole of GDGTs.

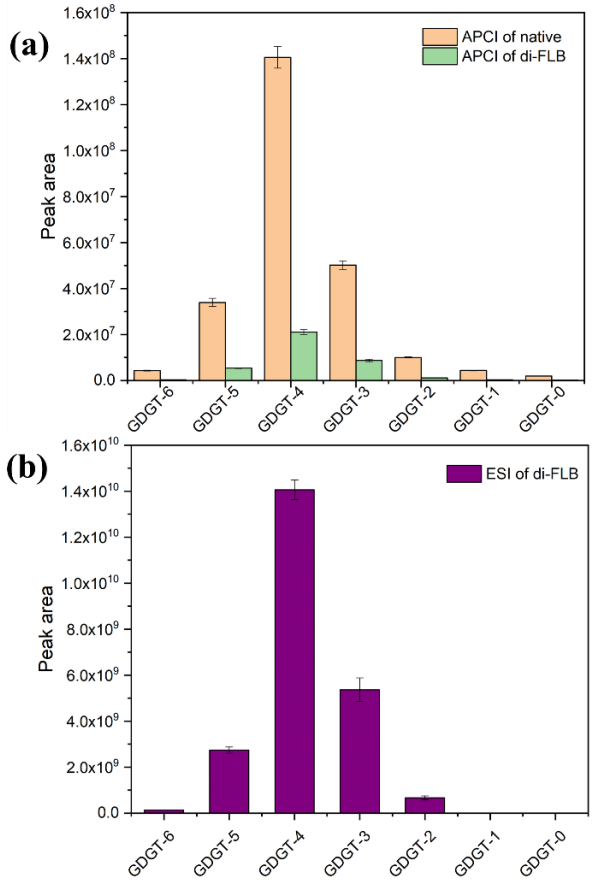


Figure 5.24. Peak areas of MS signals of FLB-GDGTs species from *S. acidocaldarius* MR31 ionised using (a) APCI, (b) ESI. The FLB-GDGTs were prepared using derivatising agents of FLB: DMAP: EDC in mole ratios 23: 19: 24 to 1 mole of GDGTs.

5.3.2.7 Examination of fluorescent signal response of FLB-GDGTs

The FLB GDGT derivatives were separated by LC-FL in order to test the ionisation efficiency of the different MS sources. The incompletely derivatised FLB-GDGT extract from *S. acidocaldarius* MR31 showed differences in the proportions of mono-FLB-GDGTs and di-FLB-GDGTs, possibly caused by inefficient ionisation. The same sample were re-analysed by UHPLC-APCI MS (Figure 5.25a) and HPLC-FL (Figure 5.25b). The separation condition of HPLC-FL was the same as used in UHPLC-MS, except that the flow rate had to be reduced to 0.2 mL min^{-1} due to the pressure limit of the instrument (<400 bars). The HPLC-FL chromatogram provides comparable results owing to the use of the same stationary phase and mobile phase composition. The chromatographic resolution of FLB-GDGTs by HPLC-FL,

however, was reduced owing to the slower flow rate (Figure 5.25b). The HPLC-FL gave two groups of peaks, the first eluting within the retention time window of 75-110 min and the second eluting between 150-210 min. The peaks were identified by comparing the response of the same sample analysed by UHPLC-MS, the first group being mono-FLB GDGTs and the second group being di-FLB-GDGTs while the native GDGTs were not detected due to the lack of a fluorescence-active group in their structures. The elution times of FLB-GDGTs by HPLC-FL (200 min) was considerably longer than for UHPLC-MS (60 min) and did not achieve baseline separation as the pressure limit requires operation at a much lower flow rate (*c.* 4 times lower). The UHPLC-APCI MS and HPLC-FL chromatograms showed that same trend in relative peak areas between mono-FLB-GDGT and di-FLB-GDGTs, reflecting that the current MS condition can be used to detect the FLB derivatives by either APCI or ESI but that ESI will enable more reliable quantification.

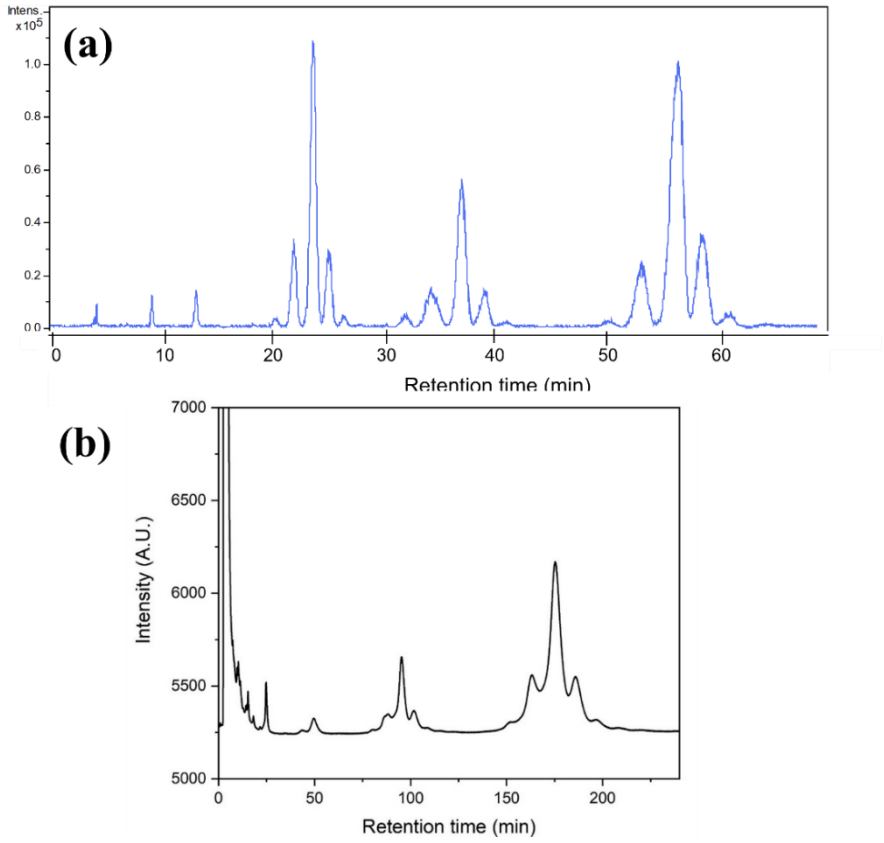


Figure 5.25. Chromatograms of FLB-GDGT extract from *S. acidocaldarius* MR31 obtained using (a) UHPLC-APCI MS operated at 500°C, (b) HPLC-FL. The FLB-GDGTs was prepared using derivatising agents of FLB: DMAP: EDC in mole ratios 20: 17: 27 to 1 mole of GDGTs.

5.3.2.8 Assessment of 1,2-di-O-octadecyl-rac-glycerol (r-dOG) as an internal standard for measurement of GDGTs

Use of an appropriate internal standard can correct for changes in sample size or concentration, samples losses during preparation steps and instrumental variation (Synder *et al.* 1997, Huguet *et al.* 2006). An internal standard for GDGTs, a synthetic diether lipid 1,2-di-*O*-octadecyl-*rac*-glycerol (r-dOG) was prepared by Poplawski (2017). The structural similarity of r-dOG to naturally occurring archaeol (2,3-di-*O*-octadecyl-*sn*-glycerol) leads to similar chromatographic behaviour, and the diol character to similar behaviour during derivatisation (Poplawski 2017). The internal standard r-dOG was analysed by UHPLC-MS to determine its chromatographic and MS behaviour under the conditions developed in the current work (see Figure C.3). The UHPLC-MS chromatogram of the standard exhibited a single peak at a retention time of 11 min, the MS spectrum of which gave a protonated molecule at *m/z* 597.8, confirming the purity of the r-dOG standard. The r-dOG was esterified with FLB in the same manner as the GDGTs, the resulting product being FLB-r-dOG (see the structure in Figure 5.26f).

The isolated FLB-r-dOG was analysed by UHPLC-MS to determine its chromatographic and MS behaviour under the conditions used in this work (Figure 5.26). The UHPLC-MS chromatogram shows FLB-r-dOG eluting at a retention time of 15 min, confirming a formation of a single pure product (Figure 5.26a). The percentage of yield by weight of the isolated FLB-r-dOG was 94 ± 1% (n = 5).

The full MS and MS^2 spectra of FLB-r-dOG were examined following ionisation by APCI (Figure 5.26b-e) and by ESI MS (Figure 5.20b-d). The full APCI MS spectrum of FLB-r-dOG shows the protonated molecule at *m/z* 1047.9 with a prominent ion at *m/z* 947.9, corresponding to the loss of a Boc group (-100 Da). The spectrum also contains a medium intensity fragment ion at *m/z* 751.8, arising from loss of both 9-fluorenemethanol (-196 Da) and Boc (-100 Da) groups from the protonated molecule of FLB-r-dOG. The MS^2 spectra of FLB-r-dOG show product ions at *m/z* 903, *m/z* 751 (Figure 5.26d) and at *m/z* 707 (Figure 5.26e), corresponding to cleavages of Boc and Fmoc groups at the various positions as shown for FLB-r-dOG in (Figure 5.26f), confirming

the presence of FLB-r-dOG. The product ion at *m/z* 597 from the precursor ion at *m/z* 751.8 (Figure 5.26e) corresponds to the loss of the rest of the FLB moiety to give an ion corresponding to native r-dOG.

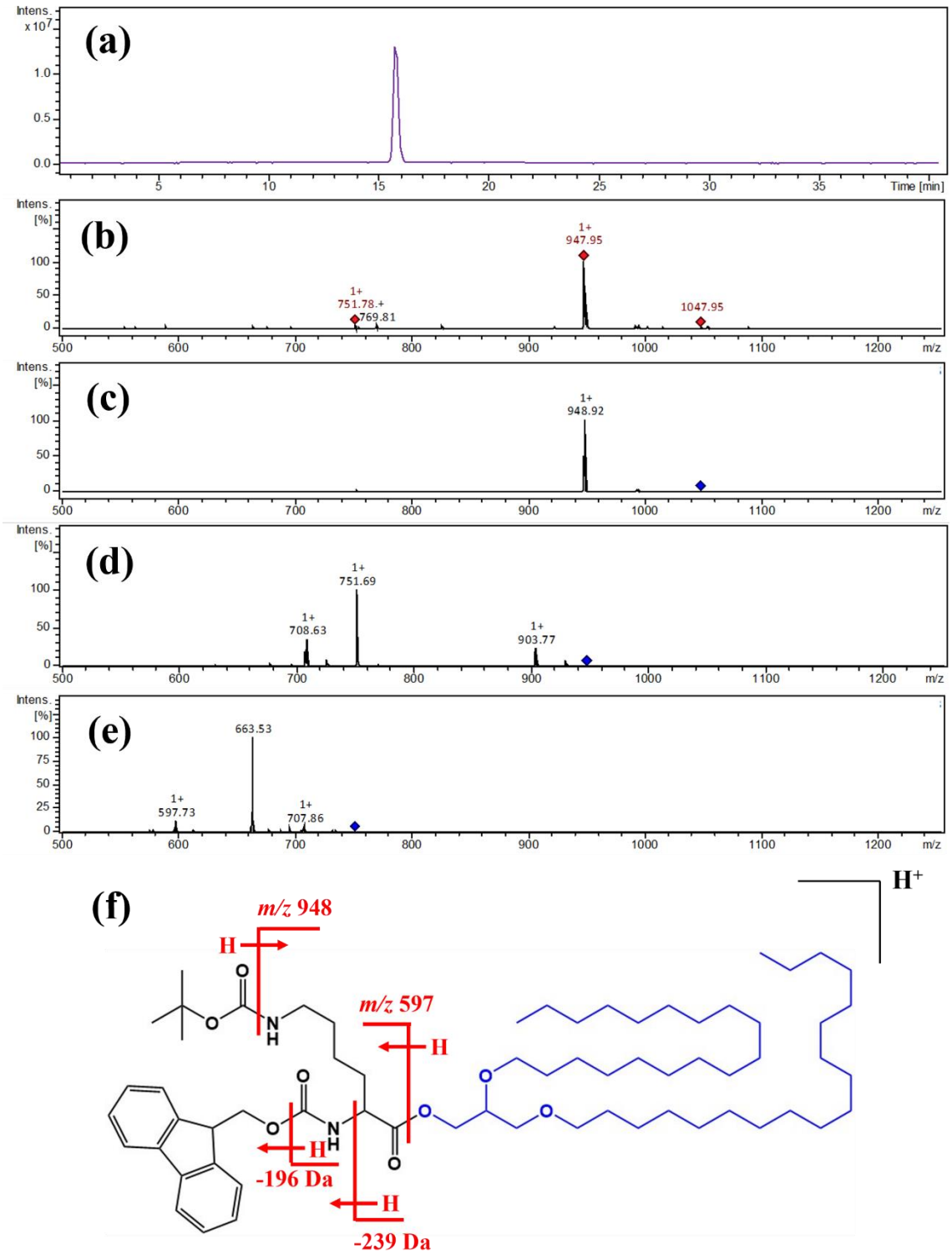


Figure 5.26. (a) APCI base peak chromatogram of 1,2-di-*O*-octadecyl-*rac*-glycerol derivatised with FLB (FLB-r-dOG), (b) Full APCI MS spectrum (c) MS^2 spectrum from precursor *m/z* 1047.9, (d) MS^2 spectrum from precursor *m/z* 947.9, (e) MS^2 spectrum from precursor *m/z* 751.8 and (f) chemical structure of FLB-r-dOG showing origins of fragment ions in the APCI MS spectrum. Concentration of FLB-r-dOG was 20 pmol injected on column.

The UHPLC-ESI MS chromatogram also shows a single peak of FLB-r-dOG eluting at a retention time of 15 min (Figure 5.27a). Notably, the UHPLC-ESI chromatogram exhibits spikes as background, possibly resulting from instability in the electrospray flow (Loo 2000; Page *et al.* 2007). The full ESI MS spectrum of FLB-r-dOG shows an adduct ion at *m/z* 1069.6 corresponding to the formation of sodium adduct (M+23 Da) with a minor intensity ion at *m/z* 947.5 arising from the loss of Boc group (-100 Da) from the protonated molecule (Figure 5.27b). The ESI MS^2 spectrum of FLB-r-dOG from precursor ion at *m/z* 1069.6 (Figure 5.27c) gave product ions at *m/z* 1013 and *m/z* 969 corresponding to cleavages of Boc group (-100 Da) and *tert*-butyl group (-56 Da), respectively. The precursor ion at *m/z* 947.5 (Figure 5.27d) gave product ions at *m/z* 751 and *m/z* 708, corresponding to losses of a 9-fluorenemethanol group (-196 Da) and an Fmoc amide protecting group (-239 Da), respectively. The same precursor ion also generated a product ion at *m/z* 903 (-44 Da), possibly arising from losses of either C_2H_6N from lysine chain as a radical or CO_2 at Fmoc chain. The dissociations observed from FLB-r-dOG by ESI MS matched those obtained from APCI MS.

The isolated FLB-r-dOG was analysed by (U)HPLC with FL, UV and MS detectors to compare signal responses. The (U)HPLC chromatograms obtained from three detectors (FL, UV and MS) with isocratic elution show the isolated FLB-r-dOG eluting at a retention time of 15 min for UHPLC condition (Figure 5.28a) and 10 min for HPLC run (Figure 5.28b-c), confirming a formation of one major product with no residue. Considering the longer elution times of FLB-GDGTs by HPLC-FL than for UHPLC-MS (Figure 5.25d), the HPLC separation was adjusted to gradient mode to obtaining faster separation (see the separation condition in Section 2.6.2, except that a lower flow rate of 0.2 mL min^{-1} was used).

The retention time and signal response of FLB-r-dOG was compared with those of the FLB-GDGT derivatives in the extracts. Chromatographic separations show FLB-r-dOG (peak indicated with *) separated well from the GDGTs (Figure 5.28). The (U)HPLC chromatograms of the FLB-GDGT extract shows that no peaks eluted at the retention time of FLB-r-dOG, reflecting the potential to use FLB-r-dOG as an internal standard.

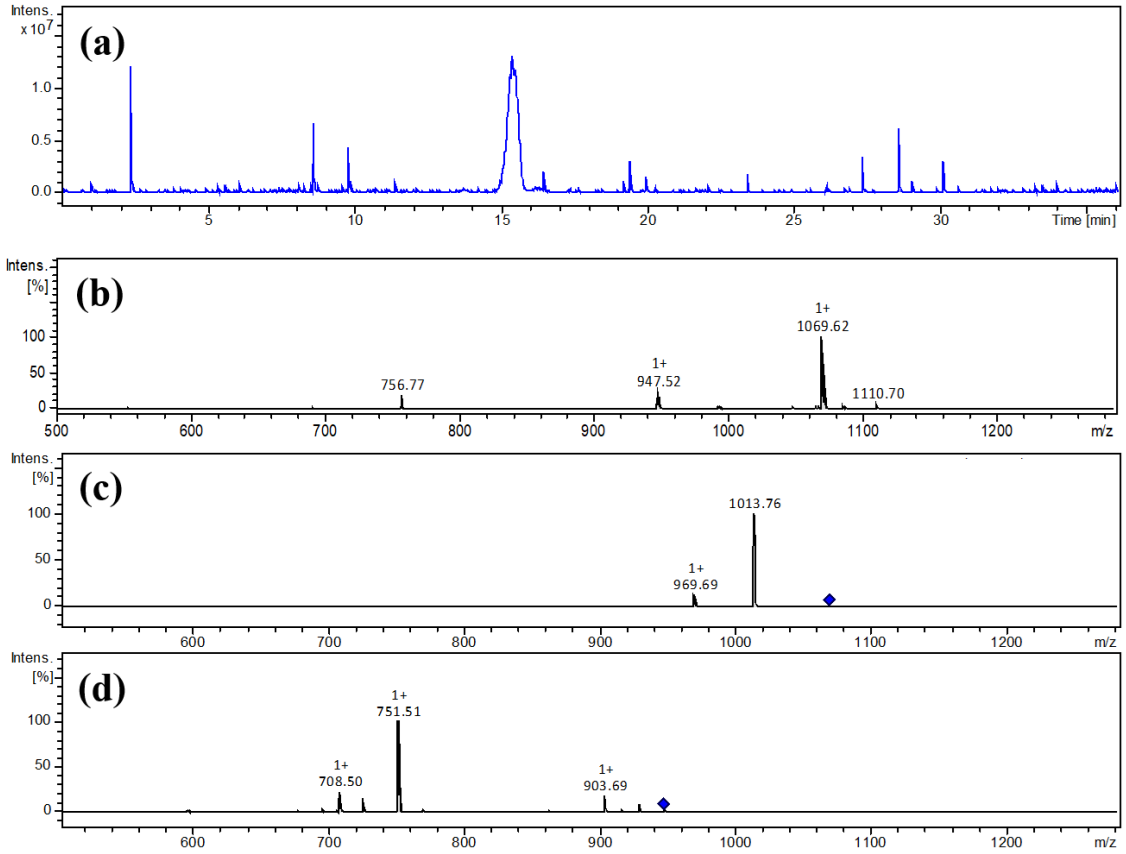


Figure 5.27. (a) ESI base peak chromatogram of 1,2-di-*O*-octadecyl-*rac*-glycerol derivatised with FLB (FLB-r-dOG), ESI MS and MS^2 spectra of FLB-r-dOG, (b) Full MS and MS^2 spectra from precursors (c) *m/z* 1069.6 and (d) *m/z* 947.5. Concentration of FLB-r-DOG was 20 pmol injected on column.

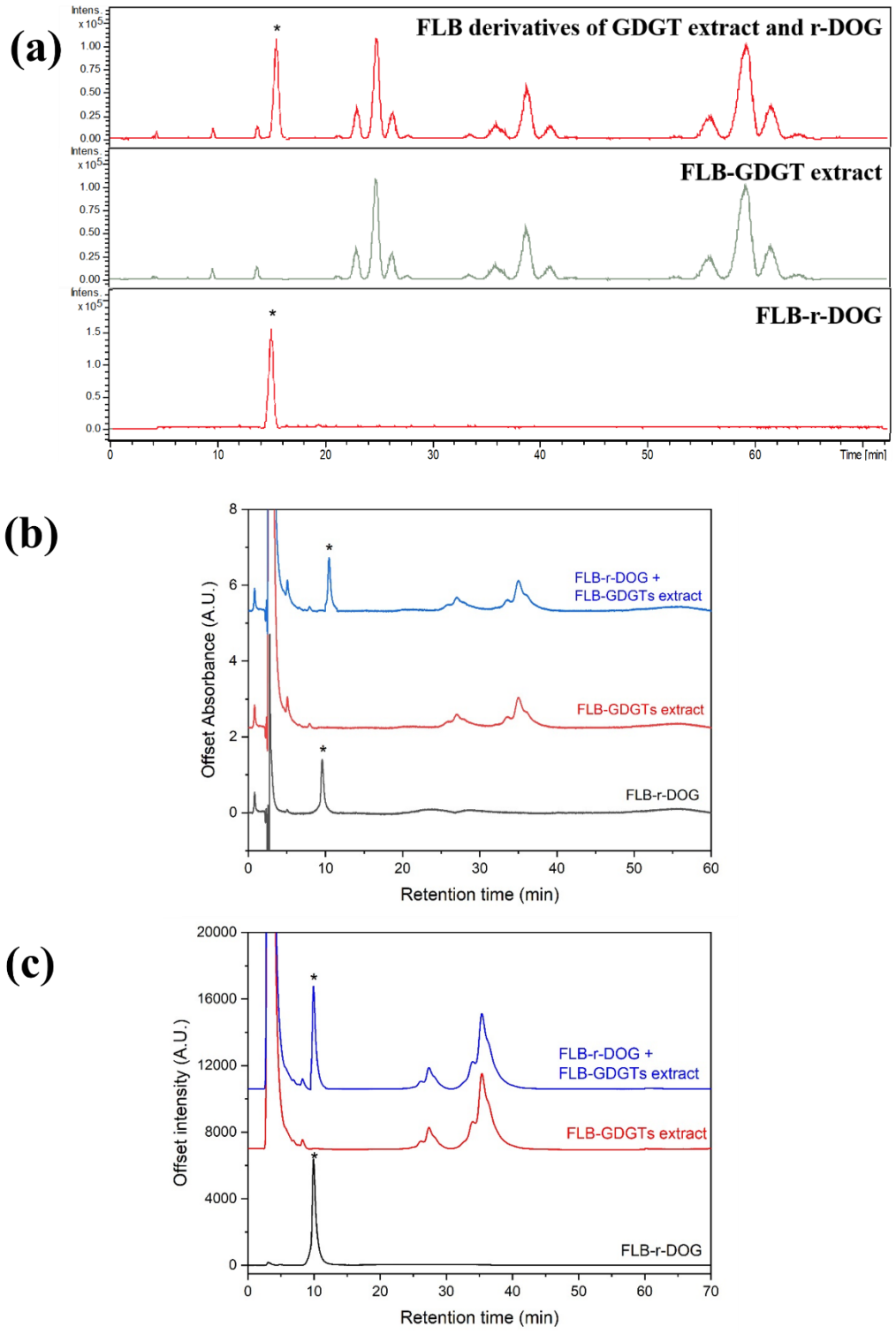


Figure 5.28. (U)HPLC chromatograms of 1,2-di-*O*-octadecyl-*rac*-glycerol derivatised with FLB (FLB-r-dOG) and of GDGT extracts with and without FLB-r-dOG monitored using different detectors: (a) ESI MS, (b) ultraviolet, (c) fluorescence. Concentration of FLB-r-dOG was 12 pmol injected on column. GDGTs were extracted from *S. acidocaldarius* MR31.

Successful measurement of FLB-GDGTs by HPLC-FL offers the potential to quantify GDGTs. In the absence of instrumentation capable of operating at >400 bar and consequent overlapping of the GDGT peaks, application of the HPLC-FL approach is limited to screening to assess completeness of derivatisation. A further refinement of the HPLC-FL separation was achieved using a newly purchased Luna PFP (2) column (Figure 5.29). Using the same sample and mobile phase composition as in Figure 5.28c (see the method details in Section 2.6.2) but with the flow rate increased from 0.2 to 0.4 mL min^{-1} to achieve the same back pressure (~280 bars), the resolution of the GDGT derivatives was improved (Figure 5.29). Identification of components based on chromatographic retention times and comparing relative peak areas with UHPLC-MS (Figure 5.28a) enabled distinction of mono- and di-ester FLB-GDGTs 0-6. The FLB-GDGTs separated by this HPLC condition gave reproducible retention times (<0.1%RSD, *n*=3) and FL signal responses (<4%RSD for peak area, <2%RSD for peak height, *n*=3), reflecting the potential to use this HPLC-FL method to obtain the reliable results for qualitative and quantitative analysis of GDGTs.

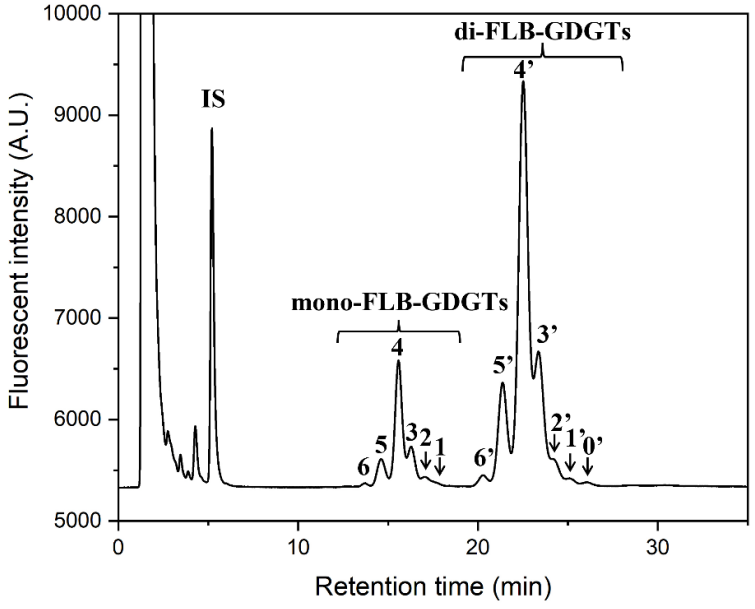


Figure 5.29. HPLC chromatogram of the FLB derivatives of GDGT extracts with FLB-r-dOG corresponding to FL detector. The FLB-GDGTs extracted from *S. acidocaldarius* MR31 were prepared using derivatising agents of FLB: DMAP: EDC as mole ratio of 20: 19: 27 to 1 mole of GDGTs (*n*=3). The separation condition is described in Section 2.6.2

5.3.2.9 Assessment of accuracy of the refined analytical method for measuring GDGTs by HPLC-FL

The esterification of GDGTs with FLB was systematically developed by modification of the whole analytical method from sample clean-up steps before and after derivatisation, the amounts of derivatising agent to obtain a complete derivatisation, and the MS and separation conditions (see Sections 5.3.2.1-5.3.2.7). The refined method (see method details in Section 2.6.2 but operated at flow rate 0.2 mL/min) offers enhanced MS signal responses of FLB derivatives compared to their native forms, with the potential to detect the components using MS, UV, and FL detection. Given that signal suppression observed in the MS detection of GDGTs can result in poor reproducibility between runs (Schouten *et al.* 2009; Lengger *et al.* 2018), spectroscopic detection methods, e.g. ultra-violet and fluoresence, could offer the potential to measure absolute concentrations of the components and thus improve the security of quantification. Reliable security in qualification and quantification of GDGTs lipids by the refined FLB method was assessed by measuring FL signal responses of FLB derivatives to evaluate their molar concentrations. Only three different concentrations of GDGT extracts were used in this study due to limited availability of the GDGT lipid core material. The molar concentrations of GDGTs extracted from *S. acidocaldarius* MR31 was estimated based on the most abundant GDGT species, *i*-GDGT-4 (1294 g/mol). An aliquot of r-dOG (0.15 μmoles) was spiked into each concentration of GDGT extract (0.15, 0.30, 0.44 μmoles) as an internal standard prior to derivatisation with FLB. The FLB-GDGTs were prepared using the refined method (see detail in Sections 2.2.2, 2.3.3 and 2.6.2). The isolated FLB-GDGTs were analysed by HPLC-FL with the chromatographic condition adjusted to shorten the analysis time, except that the lower flow rate of 0.2 mL min^{-1} was used (Figure 5.30a). The FLB-r-dOG peak eluted at a retention time of 10 min while the FLB-GDGTs eluted as a group of peaks within the retention time window 30-40 min. The FLB-GDGTs were identified by both comparing the retention times of HPLC-FL chromatograms of FLB-GDGTs extracts appeared in mono- and di-ester forms (Figure 5.28c), and UHPLC-MS (Figure 5.16) via the protonated molecules in range *m/z* 2192-2200, reflecting complete derivatisation to the di-ester forms of *i*-GDGTs 1-5 in all three different concentrations.

The concentrations of GDGTs measured using FLB-r-dOG as an internal standard (Figure 5.30b) were in good agreement with those obtained from the native GDGTs before derivatisation which was estimated by weighing (r^2 >0.99), reflecting no significant loss of GDGTs throughout the refined method.

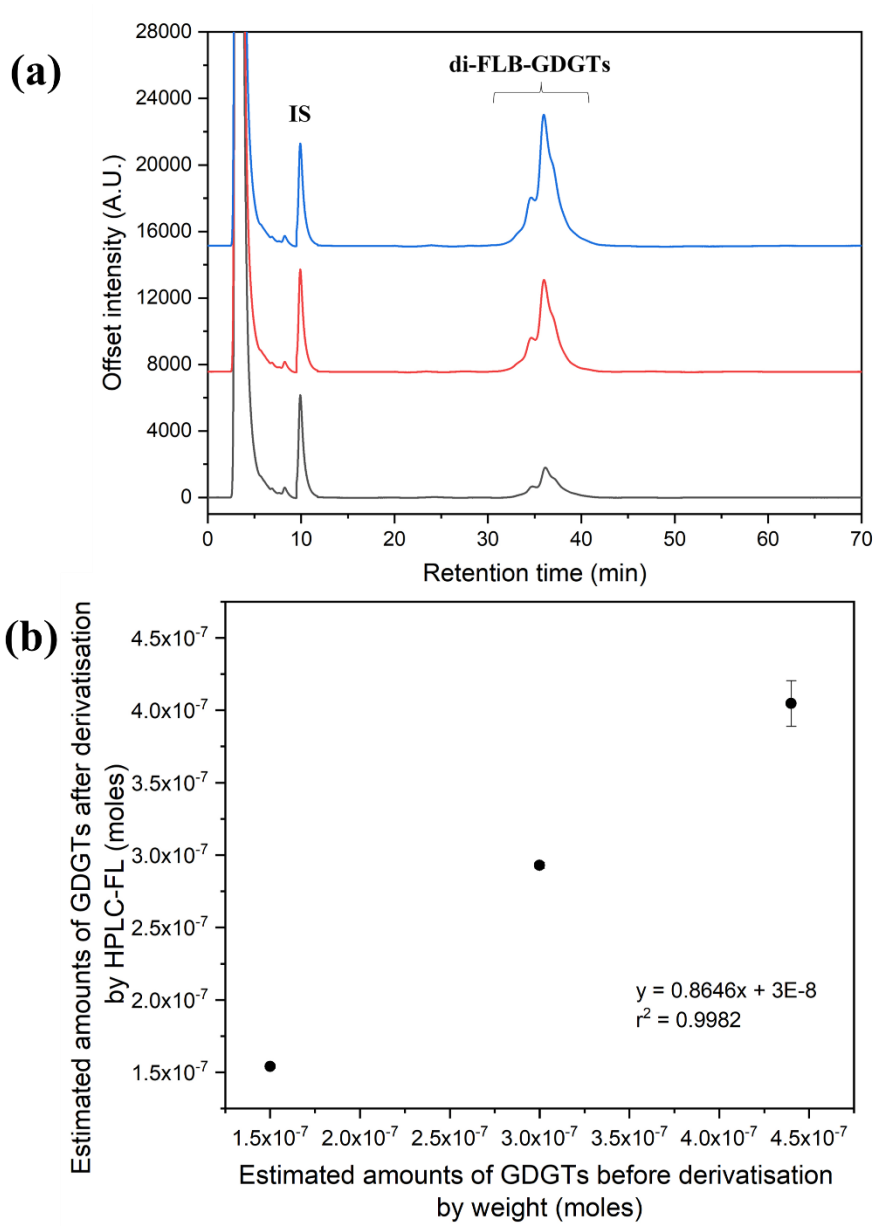


Figure 5.30. (a) HPLC-FL chromatograms of the FLB derivatives of GDGT extracts with FLB-r-DOG at three different concentrations. The FLB-GDGTs extracted from *S. acidocaldarius* MR31 were prepared using derivatising agents of FLB: DMAP: EDC as mole ratio of 23: 19: 24 to 1 mole of GDGTs, (b) The concentrations of FLB-GDGTs estimated from HPLC-FL compared to the amounts of native GDGTs before derivatisation (*n*=3).

5.3.3 Demonstration of the use of FLB to measure glycerol dialkyl glycerol tetraethers (GDGTs) in sediments

Determination of GDGTs using the modified analytical approach (Section 5.3.2) was carried out using the sediments and GDGTs extracts from a round-robin study (Schouten *et al.* 2013). The samples were three different homogenised sediments (labelled as sediment A, B, C), one sediment extract (labelled as extract F), and two mixtures of isolated branched GDGTs and crenarchaeol (labelled as extracts D and F) as designated in the round-robin study project (Schouten *et al.* 2013). Briefly, sediment A was collected from Salt Pond, Falmouth, Massachusetts, USA (41°32'N, 70°37'W; water depth 3 m). Sediment B was from a 46 kg box core from the Carolina Margin (35°50'N, 74°50'W; water depth *ca.* 600 m). Sediment C was derived from the upper part of a piston core (TY92-310G; 16°03'N, 52°71'E; 880 m water depth; 0 to 42 cm depth) taken in the Arabian Sea. Fractions D and E were mixtures of three isolated GDGT standards: crenarchaeol, *br*-GDGT-I, and *br*-GDGT-II the *br*-GDGTs having been isolated from a large extract of sediment derived from a piston core taken in the Drammensfjord, Norway (D2-H; 59°40.11'N, 10°23.76'E; water depth 113 m; sediment depth 746– 797 cm), while crenarchaeol was isolated from the remainder of Sediment C. Fraction F was a polar fraction isolated from Sediment C in a concentration of 2 mg mL^{-1}. The samples were extracted and polar fractions containing ether lipid cores were isolated using the same approach employed for the previous measurements made in the round-robin study project. The polar fractions of all samples were further fractionated to isolate only GDGTs using the modified clean-up process (see Section 5.3.2.2). All isolates were also screened for native GDGT components by UHPLC-APCI MS, and the extracted ion chromatograms used to identify the presence of *i*-GDGTs and *br*-GDGTs and to confirm the match to the previous data for each sample prior to derivatisation with FLB. The isolated GDGTs were derivatised by FLB (see Section 5.3.2.3) and separated by (U)HPLC using three different detectors (MS, UV, FL). The *i*-GDGTs and *br*-GDGTs in both native and FLB forms were detected in all samples were further used to estimate TEX_{86} (see Section 5.3.3.1) and BIT (see Section 5.3.3.2) indices.

5.3.3.1 Isoprenoid GDGTs (i-GDGTs)

5.3.3.1.1 Isoprenoid GDGTs profiles: native and FLB

Determination of FLB isoprenoid GDGTs (*i*-GDGTs) (see Section 2.6) was carried out on four samples from the round-robin study project. The sediments (A, B, C) and extract F were selected as they were previously reported for calculation of the TEX_{86} index (Schouten *et al.* 2013).

5.3.3.1.1.1 Native isoprenoid GDGTs

The GDGT extract of each sample was separated by UHPLC with MS detection to obtain the base peak chromatogram of the native GDGTs (Figure 5.31). In all of the samples *i*-GDGTs 0-6 were evident in the retention time range of 20-30 min from extracted ion chromatograms (*m/z* 1292.5 - *m/z* 1302.5) and mass spectra (illustrated for Sediment C in Figure 5.32). The retention times show close matches to the GDGT extracts from *S. acidocaldarius* and *S. solfataricus* cell cultures (Figure 5.3). The mass spectra were cross-checked with those obtained from the GDGT extracts, confirming the presence of isoprenoid GDGTs containing 0-4 cyclopentane rings (Figure 5.32). Notably, this represent a peak eluted before crenarchaeol peak (Figure 5.31, red trace: *S/N* =12), it exhibits the protonated molecule at *m/z* 1290.4 and the MS^2 spectrum (see Figure C.4, Table C.1) shows that the ion dissociates in the same manner as GDGT 0-4. As the presence of *i*-GDGT-6 has not been reported in marine sediment (Table 1.3, Schouten *et al.* 2013), the MS^2 spectrum of the precursor ion at *m/z* 1290.3 was characterised (see Figure C.4, Table C.1). The base peak of the product ion at *m/z* 737.8, corresponding to the loss of a BP_3 (-552 Da). The same precursor ion also produced the moderate intensity of product ions at *m/z* 1272.2, *m/z* 1199.3, and *m/z* 645.5, corresponding to the loss of one group of water (-18 Da), one glycerol group (-92 Da), and the loss of BP_3 connected to one glycerol group connected the BP_3, -644.5 Da), respectively. These product ions exhibit the characteristic peaks of *i*-GDGT containing one biphytane having 3 cyclopentane rings in its structure, potentially being GDGT-6. Given that the MS response of the putative GDGT-6 peak was more than 100 times lower than the other *i*- GDGTs, e.g. *S/N* of crenarchaeol peak was 965. It is possible that others have not noted this minor component in

marine sediments as extracts are typically analysed using targeted SIM approaches in the narrow mass range of *m/z* 900-1500 which most likely omit its protonated molecule. While the full scan range used in this work was *m/z* 500-2000. Therefore, the use of high-resolution MS techniques, such as orbitap or FT-ICR, to re-characterise the peak is suggested to clarify this peak.

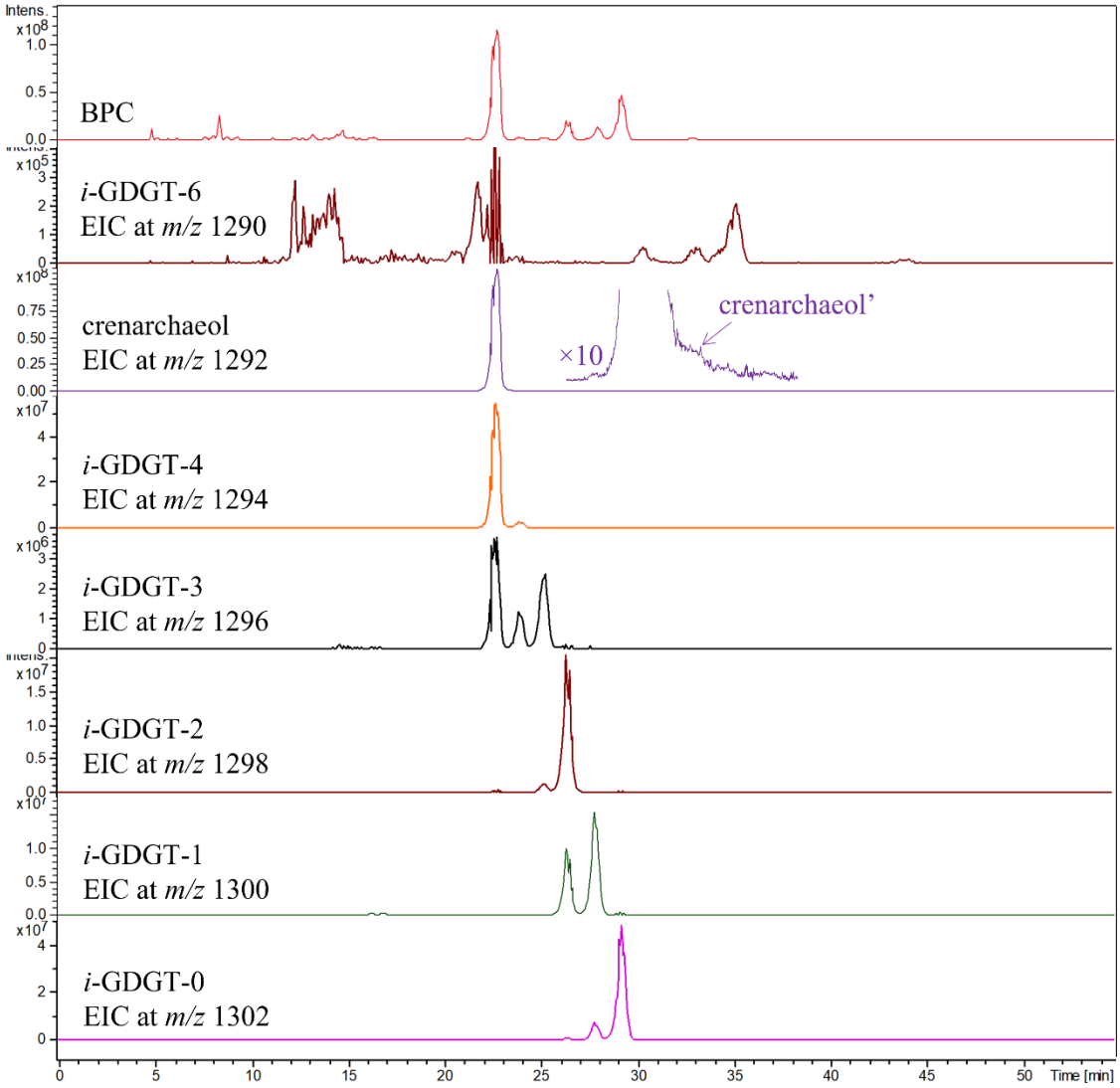


Figure 5.31. APCI base peak chromatograms and extracted ion chromatograms in range of *m/z* 1290-1302 for native GDGTs isolated from sediment C.

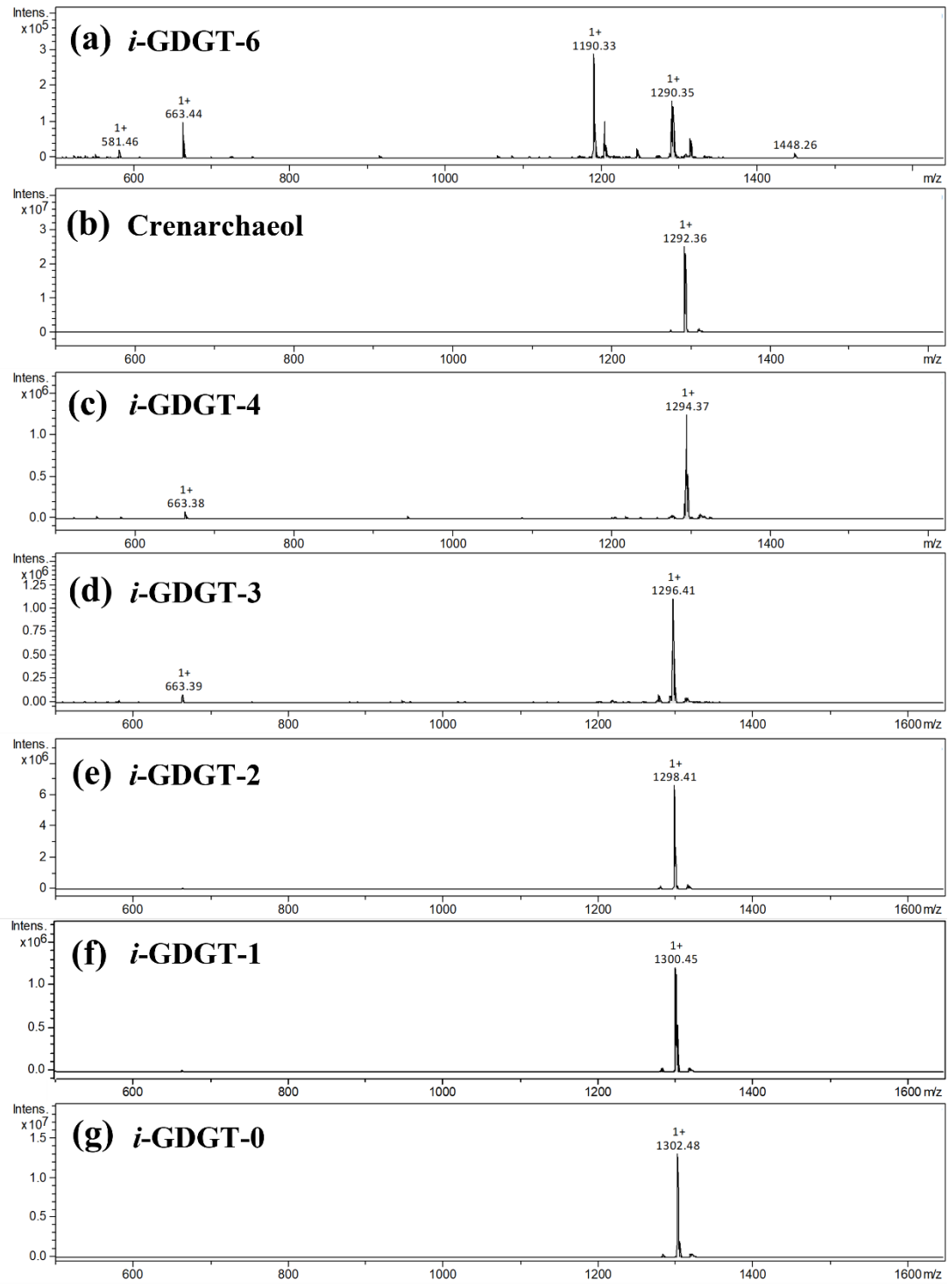


Figure 5.32. Full APCI MS spectra obtained from chromatograms of the native GDGT extract from sediment C (retention time window of 20.0-28.6 min). Each peak appeared as base peak in the range *m/z* 1290-1302.

The mass spectrum of peak at retention time of 22.6 min (Figure 5.31) exhibited the base peak at *m/z* 1292.4, consistent with the presence of crenarchaeol and crenarchaeol sterioisomer (crenarchaeol'). These two components were later assigned by comparing the relative responses of components with those of the same samples analysed in the previous studies (Schouten *et al.* 2009, 2013). The greater signal response was identified as crenarchaeol, while the minor intensity of the tiny shoulder peak was recognised as crenarchaeol'.

5.3.3.1.1.2 FLB derivative of isoprenoid GDGTs

The purified GDGT extracts obtained from each sample were derivatised by FLB (see Section 2.3.3) and analysed by the refined (U)HPLC methods (see Section 2.6). Successful detection of FLB derivatives of *i*-GDGTs 0-6 at retention time range of 48-62 min in all selected samples was evident from the extracted ion chromatograms (*m/z* 2213.5 - *m/z* 2225.5) and mass spectra (illustrated for Sediment C in Figure 5.33-5.34). The retention times show close matches to the FLB-GDGTs obtained from *S. acidocaldarius* and *S. solfataricus* cell cultures (Figure 5.33). The mass spectra obtained from the samples were cross-checked with those obtained from the FLB-GDGT extracts, matching the presence of isoprenoid GDGTs in the FLB form (Figure 5.34). The GDGT profile in each sample was identified in the same manner as the native GDGTs, revealing the presence of *i*-GDGTs 0-6 in the FLB form. The FLB derivatives of crenarchaeol sterioisomer (crenarchaeol') and crenarchaeol were also identified by comparing the relative intensities of the base peak ions at *m/z* 2215.5 at the retention times of 48 min, confirming the presence of crenarchaeol and its sterioisomer (crenarchaeol'), respectively.

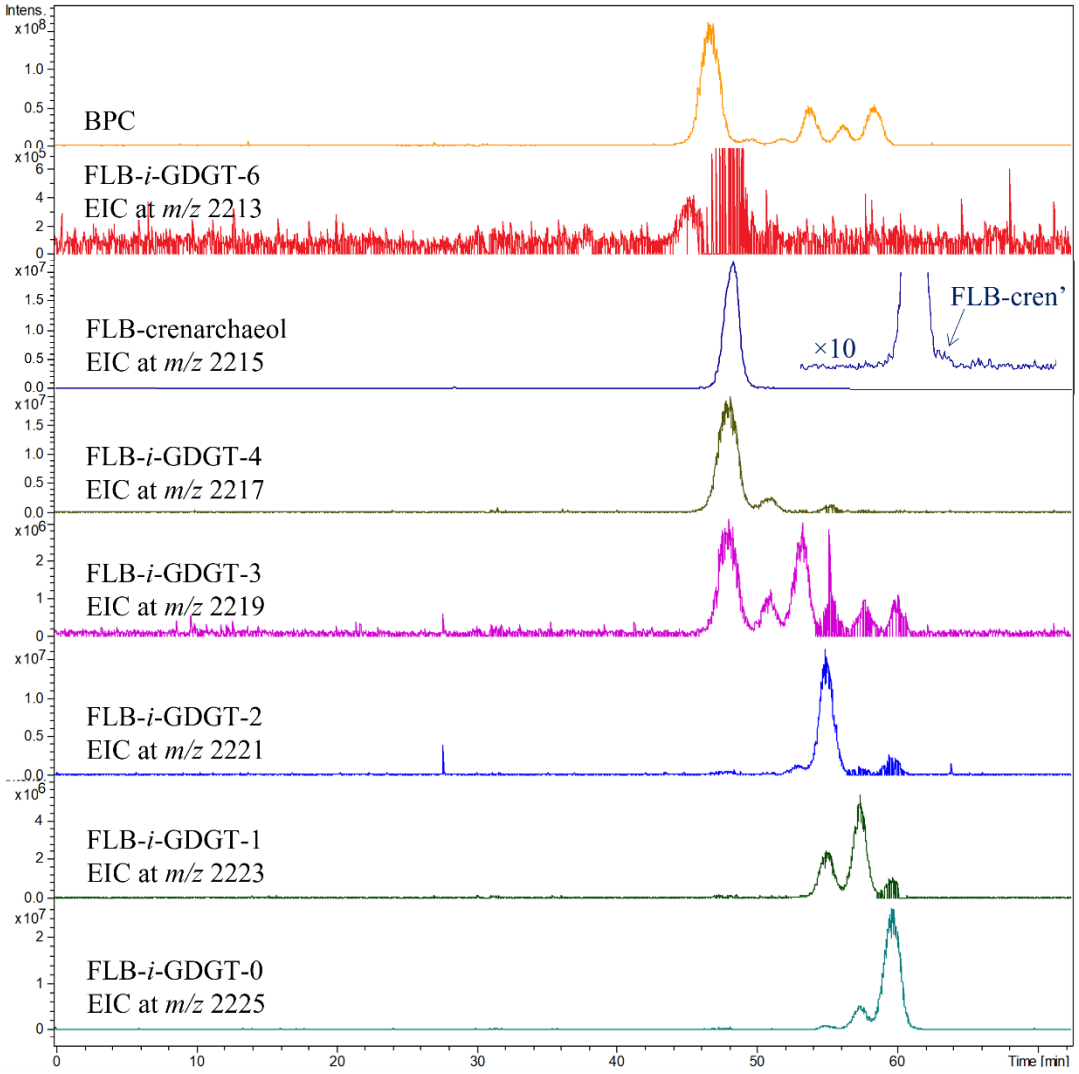


Figure 5.33. ESI base peak and extracted ion chromatograms for FLB-GDGTs from sediment C over the range *m/z* 2213-2225.

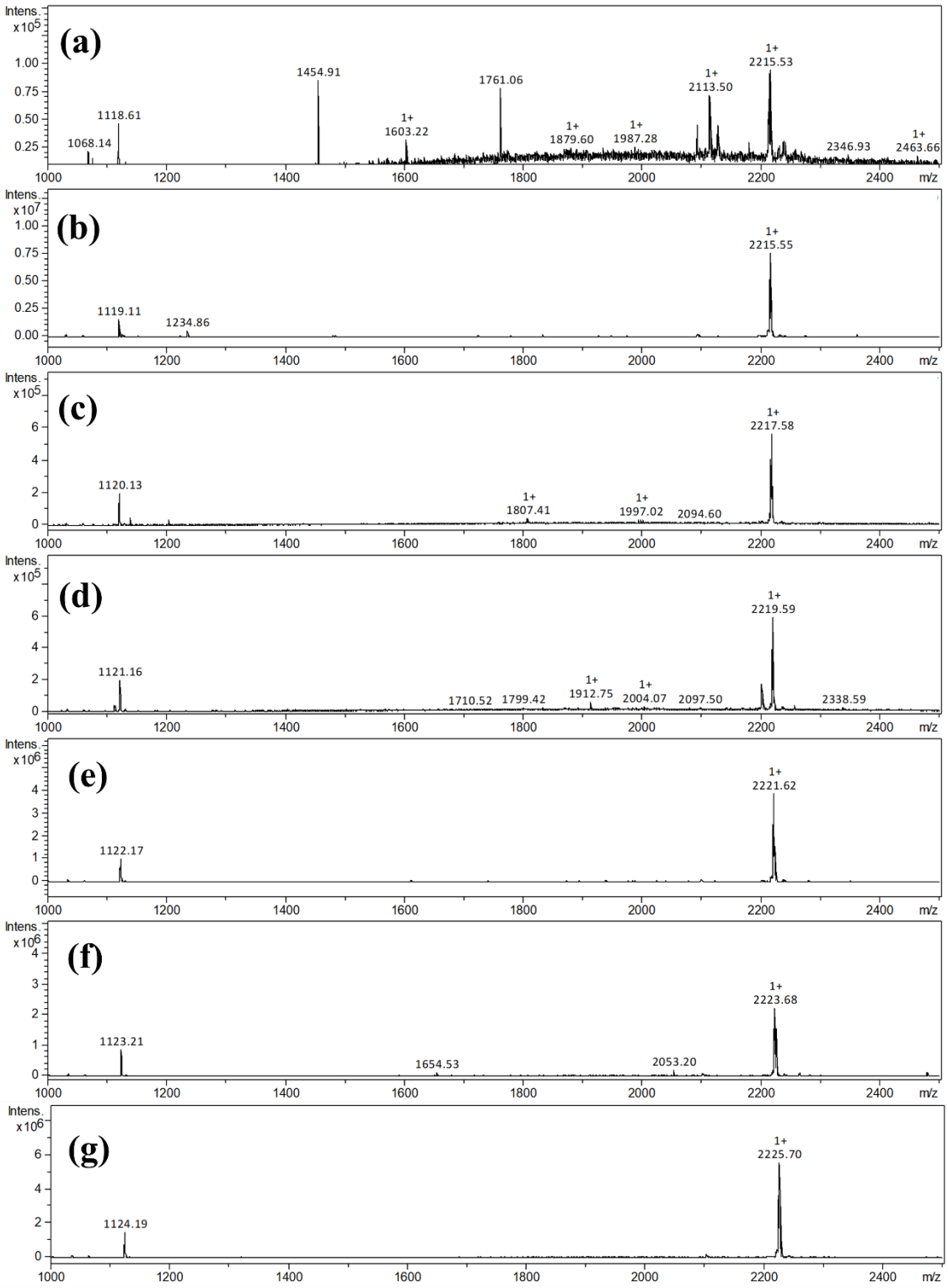


Figure 5.34. Full ESI MS spectra obtained from chromatograms of FLB-GDGTs from sediment C over the retention time window 48-62 min. Each peak appeared as base peak in the range *m/z* 2213-2225.

Sediments A (Salt Pond), B (Carolina Margin) and C (Arabian Sea) show the presence of *i*-GDGTs in the sediments of a lake, a lacustrine pelagic environment and a marine enviornment in turn (Schouten *et al.* 2013). In the majority of samples, crenarchaeol and *i*-GDGT-0 are the two major components, accompanied by smaller amounts of other forms of *i*-GDGTs (*i*-GDGT-1 to 4, *i*-GDGT-6 and crenarchaeol'). The dominance of crenarchaeol reflects input of *Thaumarchaeota* in all of the samples analysed. Higher amounts of crenarchaeol and *i*-GDGTs 0-3 were observed in the samples obtained from the open marine sediments (sample C, extract F) compared with those obtained from a terrestrial sample (sediment A) and the continential margin (sediment B) (Schouten *et al.* 2013). The Archaeal GDGT profiles of the samples were used to estimate as the TEX_{86} index proxy for sea surface water temperature.

5.3.3.1.2 Estimation of TEX_{86} index from FLB-isoprenoid GDGTs

The FMOC group of the FLB derivatives acts both a chromophore and a fluorophore, enabling detection in (U)HPLC by both UV-vis and fluorescence detection. Determination of *i*-GDGTs from the four extracts was based on successful detection of crenarchaeol, *i*-GDGT-2, *i*-GDGT-1 and *i*-GDGT-0 at retention times of 51, 60, 63 and 66 min in the samples from the UV (Figure 5.35a) and at retention times of 21.5, 24.3, 25.3 and 26.2 min in the FL chromatograms (Figure 5.35b) which were run using HPLC owing to the lack of availability of a FL detector compatible with UHPLC. The amounts of isoprenoid GDGTs were evaluated by fluorescence detection (see Table 5.5). Comparison of signal responses of FLB-GDGTs by UHPLC-MS, UHPLC-UV and HPLC-FL was made. For FL detection the extract was diluted two-fold compared with those used in UV and MS detection to avoid peak overloading of crenarchaeol. The necessity for dilution reflects the greatly improved response and the potential to quantify Archaeal lipids when concentrations are below MS detection limit .

Table 5.5. FLB-isoprenoid GDGT derivatives detected in sediment extracts of samples from the round robin study, *n* =3.

Sample	Isoprenoid GDGTs ($\times 10^{-3}$ gram per gram of polar fraction extract of sample soil, mg g^{-1})				
	crenarchaeol'	***i*-GDGT-3**	***i*-GDGT-2**	***i*-GDGT-1**	***i*-GDGT-0**
Sediment A	0.16 ± 0.01	N.D.	0.23 ± 0.01	0.29 ± 0.01	1.45 ± 0.10
Sediment B	0.16 ± 0.01	N.D.	0.21 ± 0.01	0.53 ± 0.02	3.65 ± 0.10
Sediment C	1.88 ± 0.04	1.72 ± 0.04	13.1 ± 0.16	11.2 ± 0.12	40 ± 2
2Extract F	1.71 ± 0.02	1.58 ± 0.04	11.5 ± 0.20	9.44 ± 0.40	11.4 ± 0.40

The extracted ion chromatogram of each sample was evaluated and the proportions of (*i*-GDGT-2 + *i*-GDGT-3 + crenarchaeol')/ (*i*-GDGT-1 + *i*-GDGT-2 + *i*-GDGT-3 + crenarchaeol'), expressed as TEX_{86} values (Table 5.6, Figure 5.36). The amounts of *i*-GDGT 2 -3 and crenarchaeol in the Salt Pond and Carolina Margin samples were lower than those obtained from the Arabian Sea sample, reflecting the different archaeal input and resulting in lower TEX_{86} values (Table 5.6, Figure 5.36). Notably, a large variation in TEX_{86} values was reported in the round-robin study and was attributed to the use of different spectrometers, the results indicate that ion trap spectrometers generally gave slightly lower TEX_{86} values than quadrupole instruments (Schouten *et al.* 2013). The ion trap values differ consistently from the values obtained using linear quadrupole instuments. This may be due partly to inherent differences in *S/N* between ion trap and linear quadrupole instrument (Cooks *et al.* 1991, March 1997). The TEX_{86} values obtained from this work, for both native and FLB forms, were comparable with the values reported previously using the ion trap spectrometer. The TEX_{86} values obtained from the sediment A were in the same range with those reported in the literature, while the TEX_{86} values from the sediment B and C with the extract F were slightly lower (Table 5.6, Figure 5.36). The slightly lower TEX_{86} values give rise to lower values of the estimated temperature than those obtained from the literature (Schouten *et al.* 2013). This could be attributed either to degradation of the polar lipid extracts of each sample over time, or low *S/N* ratio of crenarchaeol' generated by the ion trap. The TEX_{86} estimates obtained from this work were translated to temperature

(Table 5.7); the repeatability in temperature was less than 1°C which is comparable with other values reported (Schouten *et al.* 2013). However, the reproducibility of the translated temperature obtained from interlaboratory study was up to 3°C (Schouten *et al.* 2013). This reflects that the *S/N* ratio of each *i*-GDGT species could vary with different mass spectrometers, resulting a variation of the proportions of the relevant *i*-GDGTs. The use of FLB as derivatising agent provides an opportunity to overcome low MS response for GDGTs in sediment samples by FL detection in cases where sediments contain GDGT compounds close to the detection limit of MS. By contrast, fluorescence detection enabled measurement of concentrations in all of the samplesand gave ratio values very similar to those obtained by MS of the native GDGTs, offering a reliable alternative for quantification. The close agreement between the MS and the FL results strongly suggests that the ion trap MS results are more reliable than those obtained from the linear ion traps from which the reference data originates.

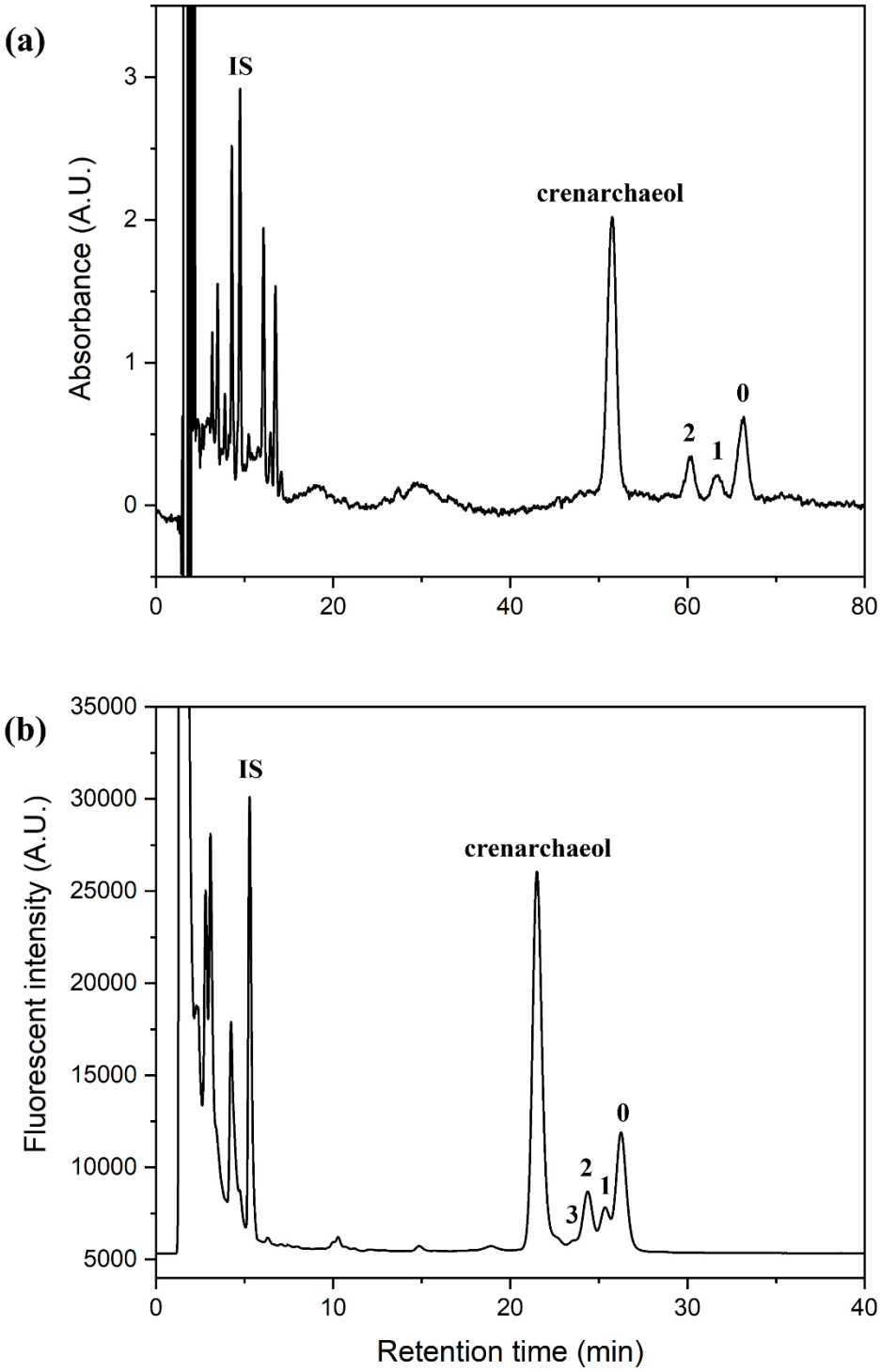


Figure 5.35. Partial chromatograms of the FLB derivatives of sediment extract from sediment C recording using different detectors: (a) UV and (b) FL (*n* =3). Assigned peaks: crenarchaeol (51 min for UV, 21.5 min for FL), 3 = *i*-GDGT-3 (23.5 min for FL), 2 = *i*-GDGT-2 (60 min for UV, 24.3 min for FL) , 1 = *i*-GDGT-1 (63 min for UV, 25.3 min for FL), 0 = *i*-GDGT-0 (66 min for UV, 26.2 min for FL). IS: r-dOG (13.5 min for UV, 5.2 min for FL). The concentration of FLB derivatives detected by FL was half that used for MS and UV detection.

Table 5.6. TEX_{86} index values estimated from the FLB isoprenoid GDGT derivatives of sediment extracts obtained from four of the round robin study samples, *n* =3.

Sample	Estimated TEX_{86}				
	MS	**UV**	**FL**	**MS***	**Reference**
Sediment A	0.46 ± 0.02 (4.3 %RSD)	N.D.[1]	0.57 ± 0.01 (1.8 %RSD)	0.44 ± 0.02 (4.5 %RSD)	0.41 – 0.65
Sediment B	0.33 ± 0.01 (3.0 %RSD)	N.D.[2]	0.41 ± 0.02 (4.9 %RSD)	0.41 ± 0.02 (4.9 %RSD)	0.47 – 0.65
Sediment C	0.60 ± 0.01 (1.7 %RSD)	0.56 ± 0.01 (1.8 %RSD)	0.60 ± 0.02 (3.3 %RSD)	0.59 ± 0.01 (1.7 %RSD)	0.63 – 0.74
Extract F	0.59 ± 0.03 (5.1 %RSD)	0.54 ± 0.01 (1.9 %RSD)	0.61 ± 0.02 (3.3 %RSD)	0.58 ± 0.02 (3.4 %RSD)	0.64 – 0.74

N.D.[1] = non-determined as crenarchaeol' could not identified., N.D.[2] = non-determined as crenarchaeol' and *i*-GDGT 0-4 were not detected., MS* = evaluated values from native GDGTs, Reference was from (Schouten *et al*. 2013).

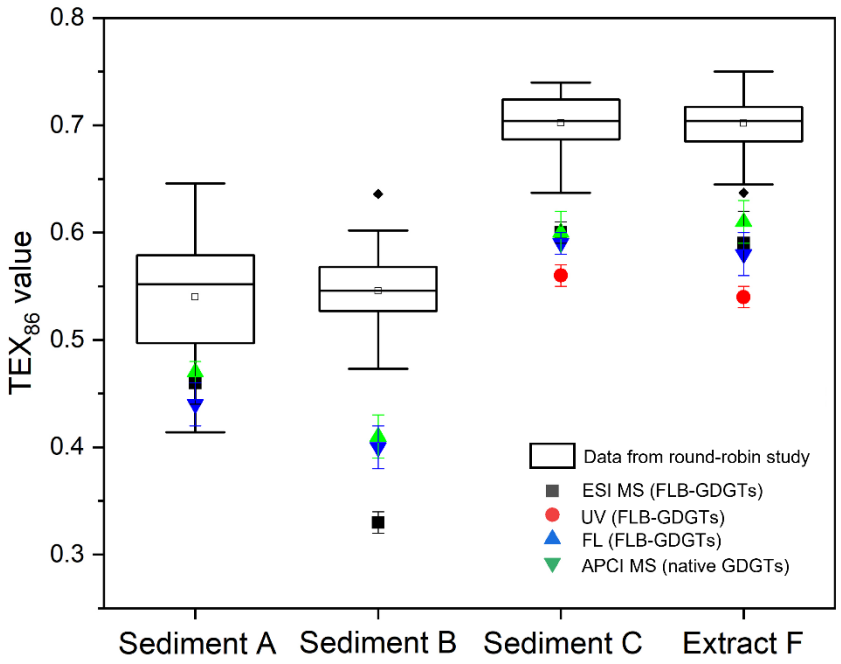


Figure 5.36. Box plots showing TEX_{86} values estimated from the FLB isoprenoid GDGT derivatives of sediment extracts obtained from four of the round robin study samples, *n* =3. The different symbols represent the data acquired using different detectors. Reference ranges are shown as box plots reffering to the TEX_{86} values of the same samples reported in the round rodin study (Schouten *et. al* 2013). The horizontal line within the box represents the median sample value. The square symbol (▫) indicates the mean sample value and the diamond symbol (♦) represents the outliers. Box indicates lower 25% and upper 75% percentiles and the whisker bars represents lower 10% and upper 90% percentiles.

Table 5.7. Calculated surface sea temperature (°C) from TEX_{86} index values estimated from the FLB isoprenoid GDGT derivatives of sediment extracts obtained from four of the round robin study samples, *n* =3.

Sample	Calculated surface sea temperature (°C)				
	MS	**UV**	**FL**	**MS***	**Reference**
Sediment A	15 ± 0.7 (4.3 %RSD)	N.D.[1]	21 ± 0.4 (1.8 %RSD)	14 ± 0.6 (4.5 %RSD)	12.3 – 25.8 (± 13.5°C)
Sediment B	8 ± 0.2 (3.0 %RSD)	N.D.[2]	12 ± 0.6 (4.9 %RSD)	12 ± 0.6 (4.9 %RSD)	15.6 – 25.8 (± 10.1°C)
Sediment C	23 ± 0.4 (1.7 %RSD)	21 ± 0.4 (1.8 %RSD)	23 ± 0.8 (3.3 %RSD)	22 ± 0.4 (1.7 %RSD)	24.6 – 30.8 (± 6.2°C)
Extract F	22 ± 1.1 (5.1 %RSD)	20 ± 0.4 (1.9 %RSD)	24 ± 0.8 (3.3 %RSD)	22 ± 0.8 (3.4 %RSD)	25.2– 30.8 (± 5.6°C)

5.3.3.2 Branched GDGTs (br-GDGTs)

5.3.3.2.1 Branched GDGTs profiles: native and FLB

Distributions of GDGTs with non-isoprenoidal carbon skeletons are also commonly found in soils and in marine sediments where they reflect the transportation of organic matter from terrestrial soils and peats into the oceans (Weijers *et al.* 2006). The amounts of these bacterial-derived GDGTs relative to crenarachaeol is expressed by the branched isoprenoid tetraether (BIT) index, which is used in the validation of TEX_{86} measurements. Determination of branched GDGTs (*br*-GDGTs) as FLB derivatives (see Section 2.6) was carried out on six samples from the round-robin study (Schouten *et al.* 2013).

5.3.3.2.1.1 Native branched GDGTs

The samples were sediments (A, B, C) and extracts (D, E, F); the isolated GDGT extract of each sample was separated by UHPLC with MS detection (see the method details in Section 2.6.1) to obtain the base peak chromatogram of the native *br*-GDGTs. The mixture of three isolated GDGT standards: crenarchaeol, *br*-GDGT-I, *br*-GDGT-II (labelled as extract E) gave three peaks in the

base peak chromatogram (Figure 5.37a). The peak at retention time 19 min gave an ion at *m/z* 1292.4, confirming the presence of crenarchaeol. Successful detection of *br*-GDGT-I and *br*-GDGT-II at retention times of 9.5 and 10.5 min in the extracts was evident from extracted ion chromatograms (*m/z* 1022.5 and *m/z* 1036.5) and their mass spectra (Figure 5.37b-c). The method used in this work has not been able to distinguish the 5- and 6- methyl *br*-GDGTs, suggesting that further modification of the separation method is required. The mass spectra obtained from the samples match those reported in the literature (Hopmans *et al.* 2004; Schouten *et al.* 2009; Yang *et al.* 2014), confirming the presence of branched GDGTs and crenarchaeol. The *br*-GDGTs and crenarchaeol in other samples were identified in the same manner as for extract E.

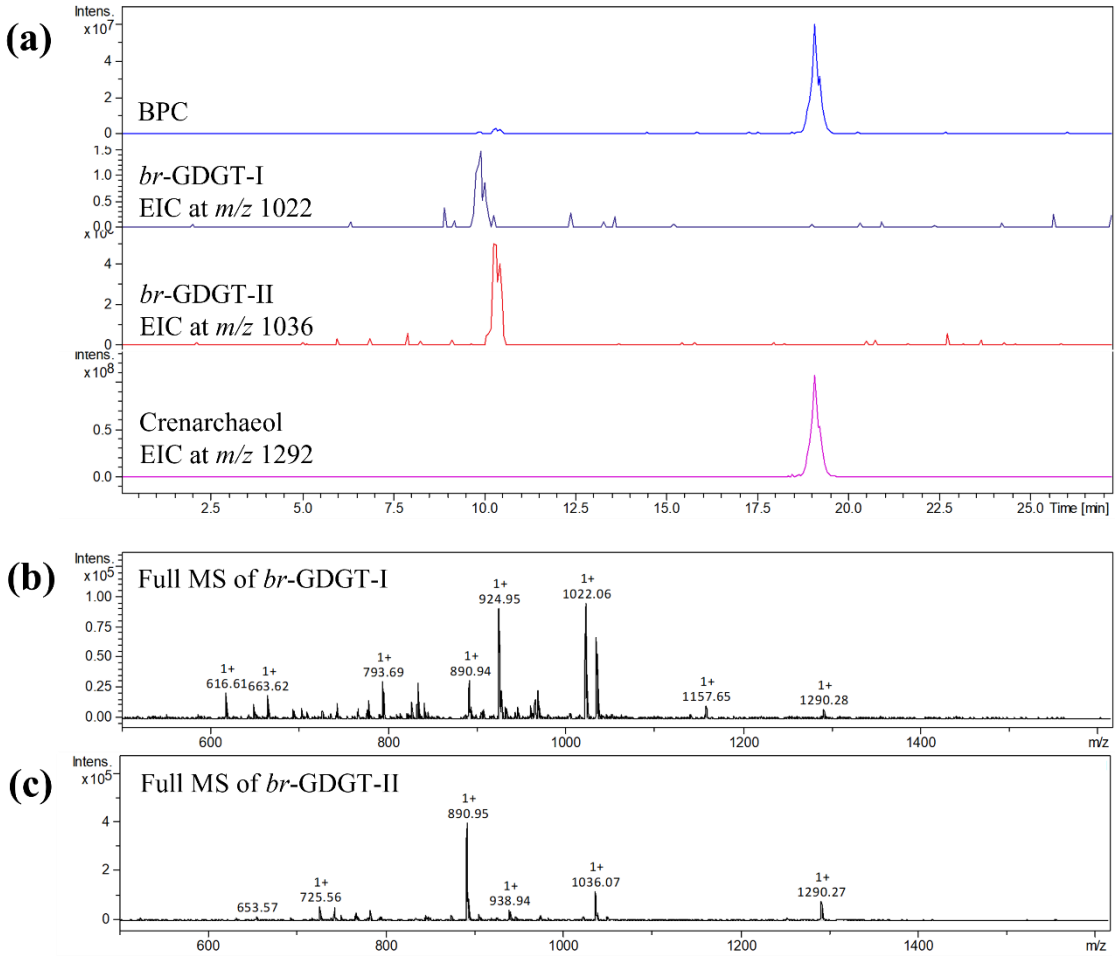


Figure 5.37. (a) APCI base peak and extracted ion chromatograms of native GDGTs from extract E recorded over range *m/z* 1022-1036 and at *m/z* 1292, (b) Full MS spectrum of *br*-GDGT-I (*m/z* 1022), (c) Full MS spectrum of *br*-GDGT-II (*m/z* 1036).

5.3.3.2.1.2 FLB derivative of branched GDGTs

FLB derivatives of *br*-GDGTs 1-2 were identified in extract E at retention times of 29 and 30 min (Figure 5.37a) using extracted ion chromatograms (*m/z* 1945.1 and *m/z* 1959.1) corresponding to the presence of sodium adducts of *br*-GDGT-I and *br*-GDGT-II, respectively. Each chromatographic peak was characterised from their full MS and MS^2 spectra (Figure 5.37b-d). For example, the UHPLC-MS peak at 29 min gave a base peak ion at *m/z* 1945.1 corresponding to the presence of the sodium adduct of FLB-*br*-GDGT-I (Figure 5.37b). The product ions *m/z* 1845 and *m/z* 1745 obtained from the MS^2 precusor ion at *m/z* 1945.1 correspond to loss of one and of both Boc groups (-100 Da and -200 Da, respectively), confirming the di-ester form of GDGT-I (Figure 5.37c). The base peak ion at *m/z* 1959.1 (Figure 5.37d) obtained from the UHPLC-ESI MS chromatogram at retention time 30 min exhibited the same losses as observed for FLB-*br*-GDGT-I (Figure 5.37e) with ions shifted by 14 *m/z* consitent with a homologue having one additional carbon, assigned as FLB-*br*-GDGT-II. The FLB-*br*-GDGT profiles in the other samples were identified in the same manner as for extract E, revealing the presence of FLB-*br*-GDGTs I-III. Crenarchaeol-FLB was identified in all selected samples by the mass spectra giving base peak at *m/z* 2215.5 at the retention time of 57 min. Thus, the BIT index was determined for all of the sediment extracts analysed.

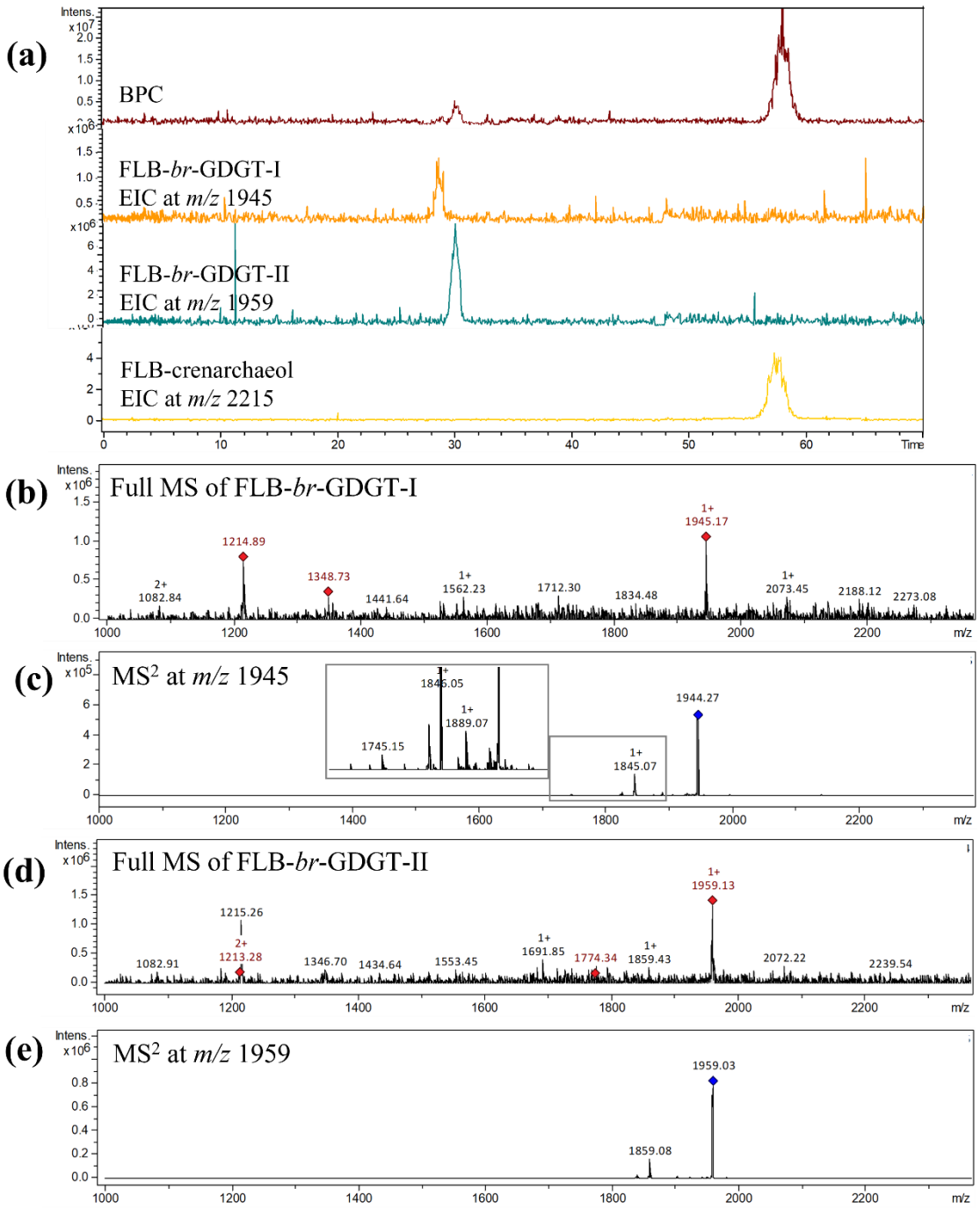


Figure 5.38. (a) ESI base peak and extracted ion chromatograms of FLB derivatives from extract E recorded over the range *m/z* 1945-2215. The MS spectra of FLB-*br*-GDGT-I showing (b) full MS and (c) MS^2 spectra from precursor ion at *m/z* 1945. The MS spectra of FLB-*br*-GDGT-II showing (d) its full MS and (e) MS^2 spectrum of *br*-GDGT-II from precursor ion at *m/z* 1959.

5.3.3.2.2 Estimation of BIT index from FLB-branched GDGTs

Branched-GDGTs were detected in the FL HPLC chromatograms of all six samples at retention times of 10.3 (*br*-GDGT-I) and 10.8 (*br*-GDGT-II) along with crenarchaeol (t_R 21.9 min) (Figure 5.39b) whereas only crenarchaeol was detected at a retention time of 51 min in the UV UHPLC chromatogram (Figure 5.39a). The amounts of the GDGTs were evaluated using r-dOG as an internal standard (Table 5.8). The profiles reveal a significant difference in the relative amounts of *br*-GDGTs to crenarcheol: the Salt Pond sample had approximately 20 times more than the open marine system. The BIT index values for each sample were determined for each of the different detectors (Table 5.9, Figure 5.40).

Table 5.8. The amounts of FLB derivatives of branched GDGTs detected in sediment extracts obtained from the different samples from round-robin study, n =3.

Sample	**Branched GDGTs and crenarchaeol ($\times 10^{-3}$ gram per gram of polar fraction extract of sample soil, mg g^{-1})**			
	***br*-GDGT-I**	***br*-GDGT-II**	***br*-GDGT-III**	**crenarchaeol**
Sediment A	0.36 ± 0.02	0.76 ± 0.05	0.93 ± 0.05	0.38 ± 0.03
Sediment B	0.12 ± 0.01	0.21 ± 0.01	0.37 ± 0.01	180 ± 12
Sediment C	0.08 ± 0.005	0.01 ± 0.0005	0.19 ± 0.006	8 ± 0.5
Extract D	1.13 ± 0.02	1.21 ± 0.010	N.D.	4.4 ± 0.2
Extract E	5.7 ± 0.1	15.3 ± 0.4	N.D.	110 ± 5
Extract F	0.12 ± 0.01	0.24 ± 0.01	0.95 ± 0.04	25 ± 1

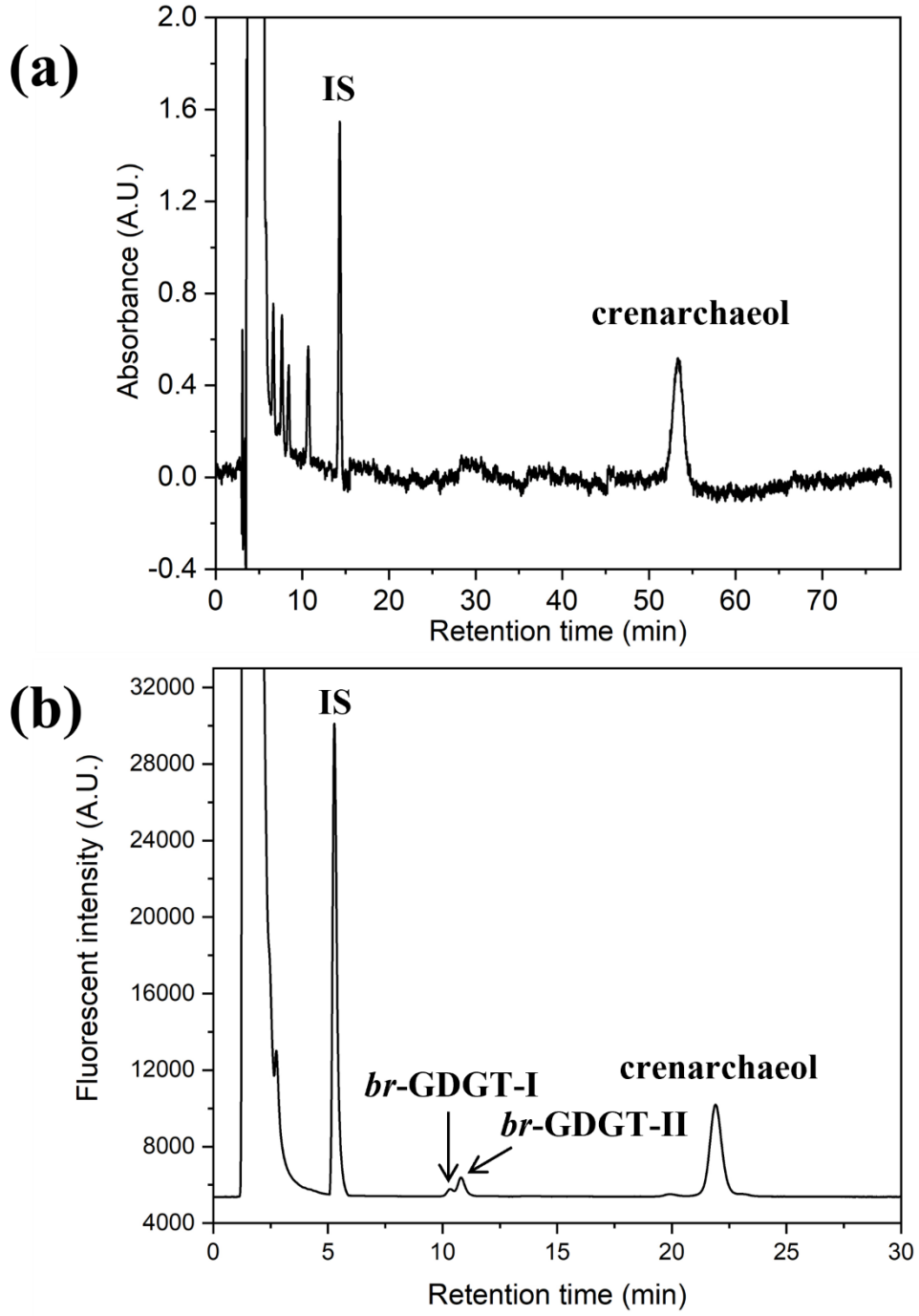


Figure 5.39. Partial chromatograms of the FLB derivatives of the extract obtained from extract E detected using: (a) UV and (b) FL (*n* =3). Assigned peaks:, 1 = *br*-GDGT-I (11 min for FL), 2 = *br*-GDGT-II (13 min for FL) , crenarchaeol (51 min for UV, 21.5 min for FL), IS: r-dOG (14 min for UV, 5.2 min for FL).

Table 5.9. BIT values estimated from FLB derivatives of branched GDGTs detected from sedimental extracts obtained from the different samples from round robin study project, *n* =3.

Sample	Estimated BIT value				
	ESI MS	**UV**	**FL**	**APCI MS***	**Reference**
Sediment A	0.842 ± 0.021 (2.4 %RSD)	N.D.	0.840 ± 0.041 (4.8 %RSD)	0.870 ± 0.040 (4.6 %RSD)	0.897 – 0.990
Sediment B	0.017 ± 0.002 (11.8 %RSD)	N.D.	0.0362 ± 0.0012 (3.3 %RSD)	N.D.	0.021 – 0.285
Sediment C	0.034 ± 0.003 (8.8 %RSD)	N.D.	0.035 ± 0.001 (2.9 %RSD)	0.037 ± 0.001 (2.7 %RSD)	0.008 – 0.247
Extract D	0.363 ± 0.007 (1.9 %RSD)	N.D.	0.340 ± 0.020 (5.9 %RSD)	N.D.	0.390 – 0.830
Extract E	0.160 ± 0.020 (12.5 %RSD)	N.D.	0.183 ± 0.012 (5.6 %RSD)	0.142 ± 0.011 (7.1 %RSD)	0.136 – 0.706
Extract F	0.011 ± 0.005 (4.5 %RSD)	N.D.	0.045 ± 0.002 (4.4 %RSD)	0.013 ± 0.001 (7.7 %RSD)	0.010 – 0.296

N.D. = non-determined *br*-GDGT-1-3 were not detected. MS* = evaluated values from native GDGTs. Reference was from (Schouten *et al.* 2013).

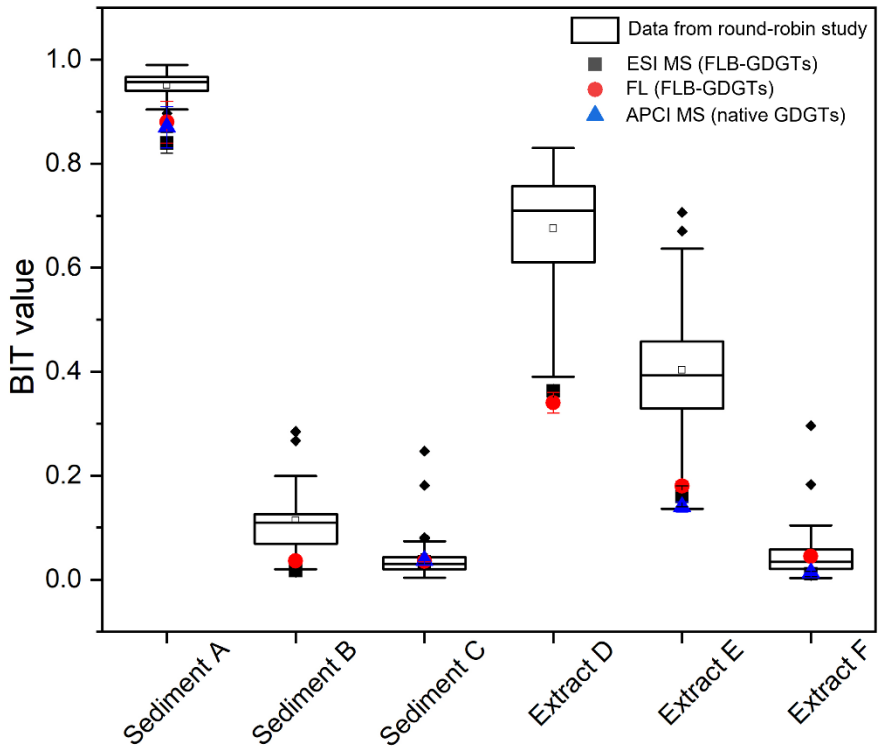


Figure 5.40. Box plots showing BIT values estimated from FLB derivatives of branched GDGTs detected from sediment extracts obtained from samples from the round robin study project, *n* =3. The different symbols represent the data acquired using different detectors. Reference ranges are shown as box plots reffering to the BITvalues of the same samples reported in the round rodin study (Schouten *et. al* 2013). The horizontal line within the box represents the median sample value. The square symbol (▫) indicates the mean sample value and the diamond symbol (♦) represents the outliers. Box indicates lower 25% and upper 75% percentiles and the whisker bars represents lower 10% and upper 90% percentiles.

The BIT values obtained from this work for both native GDGTs and their FLB derivatives were comparable with the reported values (Table 5.9, Figure 5.40), consistent with the use of an ion trap spectrometer which was found previously to give slightly lower BIT values than quadrupole instruments (Schouten *et al.* 2013) The BIT values of sediment B and the extracts (E, F) were in the same range with those reported in the literature, while the values from sediments A and C and extract D were slightly lower (Table 5.9, Figure 5.40). The slightly lower BIT values could be

attributable to a difference in the *S/N* ratio of each *br*-GDGT and crenarcheol generated by different mass spectrometers. Notably, the extracts (D and E) were GDGT mixtures prepared by weight, giving the known molar ratios of *br*-GDGT-I, *br*-GDGT-II and creanarcheol as BIT values of 0.440 and 0.158, respectively. The BIT values obtained from the literature generally significantly overestimated the stated values for the BIT index, except for the values from two of the ion trap instruments (Schouten *et al.* 2013). The BIT values obtained from this work were closer to the stated BITvalues of the standard mixtures, showing the potential for obtaining accurate molar-based quantification.

The relative standard deviations (%RSD) of BIT values from different MS detectors reported in the literature reveal that the repeatability was up to 11% while the reproducibility was up to 66% (Schouten *et al.* 2013). The %RSD of BIT values evaluated from FLB-GDGTs by FL detection (<5.9%) were lower than those from MS detection (12.5%), showing the potential for improvement in quantification.

Considering the reported values of both TEX_{86} and BIT indices (Schouten *et al.* 2013), the ion trap values differ consistently from the values obtained using linear quadrupole instuments and that may be due to higher *S/N* for the former (Cooks *et al.* 1991; March 1997). The TEX_{86} and BIT values estimated from the FLB derivatives by ion trap MS were comparable with the values from FL, seemingly reflecting greater accuracy and precision of the ion trap compared with linear quadrupole instruments. Thus, the use of FLB as derivatising agent provides an opportunity to assess these GDGT-based proxies and the greater reproducibility that it offers will give greater security to the interpretations. In cases where sediment extracts contain *br*-GDGT compounds close to detection limit of MS, the use of the FLB derivatising agent can provide a reliable alternative for quantification by FL detector.

5.4 Conclusion

The use of Fmoc-lysine-Boc (FLB) as a derivatising agent for measurement of glycerol dialkyl glycerol tetraether (GDGT) lipids to reconstruct their relative populations in geological sediments was evaluated systematically. The GDGT lipids were successfully isolated from the polar fractions obtained from Archaeal cell cultures (*S. solfataricus* and *Sulfolobus acidocaldarius* MR31). The FLB derivatives of GDGT lipids from Archaeal cell cultures were prepared by Steglich esterification with improved yields by decreasing the ratio of derivatising agent to lipid extract from the proportions suggested by Poplawski (2017). Structural characterisation of FLB-GDGT derivatives (FLB-GDGT 0-6) by mass spectrometry confirms the formation of the desired products. A study of the MS^2 dissociation pathways of FLB derivatives of GDGT lipids reveals characteristic losses for FLB derivatives that can inform the identification of GDGTs with two FLB molecules. Comparison of UHPLC-APCI MS responses of GDGTs in their native and FLB forms showed that the method previously proposed by Poplawski (2017) needed to be modified to achieve a maximum response for the FLB derivatives. The analytical method was refined successfully, however the response for APCI of the FLB derivatives was still lower relative to the native GDGTs. The MS signals of FLB derivatives of GDGTs detected by ESI MS gave higher signal response than those obtained from APCI MS approximately 1,000 times for FLB and 100 times for the native GDGTs. The FLB-GDGT derivatives gave about 200 times improvement in response by fluorescence detection than by MS, reflecting the suitability of this approach and the resulting greater security in quantification. HPLC-FL analysis of GDGT components in sediment extracts from a round-robin study reveals that of the FLB derivatives give a substantial improvement in response compared with a conventional MS detector for isopreniod GDGTs and branched GDGTs extracted from sediments. In the same sediment extract, derivatisation by FLB at half the concentration (Figure 5.35b) provides a chromatogram with a considerably lower level of noise than those obtained with APCI MS detection (Figure 5.34). Evaluation of the TEX_{86} and BIT indices based on determination of FLB-GDGT derivatives gave values comparable with those obtained from the native form using the same mass spectrometer. Determination of FLB-

GDGTs gives a significant improvement in relative standard deviations (%RSD, *n*=3) for detection by FL (5.9% lower than those from MS detector: 12.5%), reflecting greater security in quantification.

The derivatisation approach provides an opportunity to overcome several of the limitations in the detection of Archaeal lipids in geological sediments, and the greater repeatability that it offers will give greater security to the interpretations. The further refinement in signal response of FLB-GDGTs from FL detection could be explored using a UHPLC compatible detector. Once the maximum signal response is obtained, the reliability of the response for FLB-GDGTs detection by MS, UV and FL could be compared directly. In cases where sediments contain GDGT compounds close to the detection limit of MS, the use of the FLB derivatising agent can provide a reliable alternative for quantification by FL detector.

Chapter 6

Conclusions and Future work

6.1 Overall conclusions

The approach using Fmoc-Lys-Boc (FLB) as a derivatising agent for measurement of steroidal lipids (cholesterol, coprostanol and 5α-cholestanol) and glycerol dialkyl glycerol tetraether (GDGT) lipids was evaluated systematically, showing the potential to use the method to detect human activity in archaeological and environmental contexts, or to measure Archaeal tetraether lipids in geological sediments, including for reconstruction of past surface water temperature.

Successful formation of cholesterol and cholestanols esterified with FLB using the derivatisation condition proposed by Poplawski (2017) was confirmed by structural characterisation of FLB derivatives of C_{27} steroidal lipids (cholesterol, coprostanol and 5α-cholestanol) by mass spectrometry. The preliminary screening of cholesterol and its transformation products using the UHPLC conditions proposed by Poplawski (2017) revealed some unwanted residues from the derivatisation process and the co-elution of cholesterol and coprostanol peaks, reflecting the need to modify the analytical method.

Generation of an improved signal response for the FLB derivatives of cholesterol and its related stanols was obtained by modification of the method from the sample preparation step to the separation method. Completeness of the derivatisation process for cholesterol was improved by increasing the ratio of derivatising agent to standard solution relative to the proportions suggested by Poplawski (2017). Repeated preparation of the FLB derivative of cholesterol gave a mean yield by weight of 89 ± 4% (n =8), reflecting good yield and reproducibility under these conditions. The residues from the derivatisation process were mostly removed using a modified sample clean-up step with a more apolar eluent relative to that suggested by Poplawski (2017). Baseline resolution of cholesterol and coprostanol in the UHPLC separation was achieved by changes to the stationary phase and the mobile phase gradient composition. The modified analytical method showed an improvement in the signal responses for cholesterol, coprostanol and 5α-cholestanol recorded using mass spectrometry and ultra-violet/visible spectrophotometry. Measurement of cholesterol, coprostanol and 5α-cholestanol by UHPLC with MS and UV

detection was evaluated and the capability to determine these components by FL detector was assessed by HPLC separation owing to the lack of UHPLC-FL facilities. The FLB derivatives of cholesterol and cholestanols gave maximum signal response even two weeks after preparation, reflecting the suitability of approach for routine analysis and reliable quantification. The FLB derivative of 1-heptadecanol was shown to be suitable for use as an internal standard, the desirable product being formed in high percentage yield by weight (93 ± 2%, n = 5). A full-range of analytical characteristics (repeatability and reproducibility, working range, limit of detection, limit of quantification and recovery) of the modified method was evaluated, demonstrating reliable analysis of coprostanol, cholesterol and 5α-cholestanol with all three detectors (MS, UV and FL). The more favourable LODs and LOQs of the FLB derivatives obtained from FL detection than with more widely used MS and UV detectors offer greater potential to detect C_{27} steroidal lipids at trace levels, extending the potential range of applications for archaeological, environmental, and potentially medical samples.

The approach for FLB derivatisation of cholesterol and cholestanols was applied to a range of archeological samples from Hungate, York (fourteen different depths from a cesspit core and grave soil underneath human remains) and environmental samples (one sediment core and two surface sediments both from the Humber estuary, Hull). The values for the coprostanol ratio (>0.7) confirms signatures of human faecal contamination at all of the sites. The coprostanol ratios measured from the FLB derivatives were similar to those obtained from GC-MS. Comparison of the HPLC-MS and GC-MS signal responses for steroidal components in archaeological soils and environmental sediment extracts demonstrates that the use of FLB as a derivatising agent provides a substantial improvement in response compared with the conventional TMS derivatives for cholesterol and its transform products. In the same soil extract, the derivatisation of sample by FLB required around half of the amount of the soil extract than for silylation. Furthermore, the chromatograms exhibited lower noise levels than those obtained with the TMS derivatising agent. The more favourable limit of detection of the FLB derivatives analysed by LC-MS (0.05 pmol on column calculated based on 3 × *S/N* ratio) compared to the

conventional trimethylsilyl ethers that are analysed by GC-MS (1 pmol on column, 3 × *S/N*) offers greater security and improved detection limits. In situations where the conventional GC-MS based approach may be limited by the amounts of the C_{27} steroidal lipids available (low concentrations or small sample size), the use of FLB offers the potential to detect components down to picomole level by either MS or UV detection or down to femtomole (fmol) level for FL detection; the presence of the FMOC provides the ability to quantify the FLB derivatives at trace level.

An exploration of the dissociation pathways of the FLB derivatives of cholesterol and two of its reduction products (coprostanol and 5α-cholestanol) by full MS and multistage tandem (MS^n) mass spectrometry revealed three characteristic fragmentations of the derivatives: (1) a prominent fragment ion arising from the specific neutral loss from the molecular ion of the entire Boc group (-100 Da), (2) a minor loss arising from cleavage of the tert-butyl group (-56 Da), (3) a fragment ion corresponding to FLB (*m/z* 369.1). These fragmentation patterns were used to detect the presence of other alcohols as their FLB derivatives without the need for access to standard compounds, revealing signatures of *n*-alkanols, triterpene, tocotrienols and C_{29} steroidal lipids in soil samples. In the same soil extracts, successful identification of a wider range of carbon numbers (C_{12}-C_{36}) were detected for *n*-alkanols as the FLB derivatives compared to those obtained from TMS derivatives (C_{22}-C_{30}), demonstrating the potential for the approach to reveal unknown compounds. The signatures of coprostanol relative to C_{29} stanol reveal information of human faecal inputs related to ingestion of particular types of foodstuff (herbivores and omivores) recorded within each layer of the cesspit, offering greater potential to extend the understanding of past human behaviour.

The FLB derivatives of GDGT lipids isolated from archaeal cell cultures (*S. solfataricus* and *Sulfolobus acidocalarius* MR31) were prepared using the derivatisation conditions proposed by Poplawski (2017) and formation of the desired products was confirmed by mass spectral matching to the full MS spectra previously reported by Poplawski (2017). The lower signal response for FLB derivatives of GDGTs compared with that of the native forms was investigated to provide

an explanation for the poor signal responses of FLB derivatives and to develop an approach to enhance their response.

The poor signal responses of FLB-GDGTs derivatives obtained using the MS conditions employed by Poplawski (2017) was attributed in part to the presence of hydroxyl groups from GDCTs in the extracts, reducing the yield of FLB-GDGTs. A modified sample clean-up procedure was developed by which the GDCTs in the polar fraction of Archaeal cell cultures extracts were mostly removed using a more apolar eluent relative to that used by Poplawski (2017). Mass spectral dissociation pathways of FLB derivatives of GDGTs were further elucidated by full MS and MS^2, revealing characteristic fragmentation of FLB derivatives of GDGTs with one and two FLB molecules, enabling confirmation of the incompleteness of the derivatisation process. The completeness of the derivatisation process for GDGT lipids from Archaeal cell cultures was improved by decreasing the ratio of derivatising agent to lipid solution relative to the proportions suggested by Poplawski (2017).

Comparison of the UHPLC-APCI MS signal responses of GDGTs in their native and FLB forms showed that the MS conditions used by Poplawski (2017) leads to incomplete ionisation, increased response for FLB-GDGTs being obtained by raising the APCI vaporisation temperature from 450ºC to 500ºC. The maximum APCI vaporisation temperature (500ºC) still gave lower signal responses (*c*. × 10) for FLB-GDGTs compared with native GDGTs, reflecting incomplete ionisation/ion transmission by APCI. Improved ionisation efficiency for FLB-GDGTs was successful achieved using ESI MS, providing greater response (*c*. × 100) than obtained from the native forms in APCI mode. Thus, the ionisation of FLB-GDGTs by ESI is recommended for FLB-GDGT analysis.

Assessment of 1,2-di-*O*-octadecyl-*rac*-glycerol (r-dOG) as an internal standard for GDGT quantification demonstrated that the pure FLB product was generated in high percentage yield by weight (94 ± 1%, $n = 5$), indicating the suitability of its use to quantify GDGTs.

Determination of FLB derivatives of GDGTs by MS and UV detection was evaluated using reversed phase UHPLC separation on a C_{18} PFP column and the ability to determine the derivatives by FL detection was assessed by HPLC separation (due to the lack of UHPLC-FL facilities). The UHPLC separations of GDGTs were achieved within 30 min for the native forms and within 60 min for FLB derivatives with no build-up of back-pressure occurring between runs. Thus, the method appears to be more robust than the conventional normal phase separation which normally exhibits build-up of back-pressure within 10 runs. The detection of FLB-GDGTs by FL detector required half the concentration required for ESI MS detection. Comparison of signal responses for FLB-GDGTs with the three detectors reveals FL detection to be about 200 times better than MS detection, indicating that the approach will give greater security for quantification and is suitable for trace analysis. In situations where the use of FLB offers the potential to identify GDGT components by MS detection, the absolute amounts of different GDGT structures can be determined directly using FL detection.

The use of FLB as a derivatising agent for the measurement of GDGTs in sediment extracts from a round-robin study revealed the ability to detect both isopreniod GDGTs (*i*-GDGTs) and branched GDGTs (*br*-GDGTs) in the samples. Evaluation of the TEX_{86} and BIT indices based on FLB derivatives gave estimated comparable values with those obtained from the native form, however these values were still close to the lower extreme of those reported in the round-robin study. Given the closeness of the MS measured values to those measured by FL detection, the lower values can be attributed to the better *S*/*N* ratio obtained with the ion trap compared with linear quadrupole instruments (Cooks *et al.* 1991; March 1997). The better repeatability for FLB-GDGTs determined using FL detection (<5.8 %RSD, n=3) than with conventional MS detection (12.5%RSD, n=3), indicates the potential for greater security in quantification by fluorescence detection. Thus, the use of FLB to measure GDGTs by FL detection provides an opportunity to overcome several of the limitations in the routine assessment of Archaeal lipids in geological sediments, and the greater reproducibility that it gives will provide greater security to the interpretations arising from estimation of the commonly determined GDGT indices. In cases

where geological materials contain GDGT lipids close to detection limit of MS, the use of the FLB derivatising agent can enable quantification by FL detector.

6.2 Future work

The approach using FLB as a derivatising agent to determine steroidal lipids (cholesterol, coprostanol and 5α-cholestanol) developed in this work has been shown to be suitable for application to archaeological and environmental samples. Further refinement of the approach could be considered in order to extend its potential to routine analysis and to a wider range of applications, the following five aspects being areas for consideration:

1. To alleviate the need for fractionation prior to derivatisation, the potential for derivatisation of total extracts from soils and sediments could be examined and compared with GC-MS analysis of TMS derivatives of C_{27} steroidal lipids (cholesterol, coprostanol and 5α-cholestanol) obtained from total extracts.

2. The completeness of the derivatisation reaction could be assessed by monitoring for the presence of excess reagent. Hence development of a rapid method to detect excess reagent from its fluorescence response would be advantageous.

3. Signatures of cholesterol and related sterols and stanols could be determined in a wider range of samples, including forensic and medical samples, to extend the applications of the approach to other environmental challenges and to examine crime scenes and metabolic activity. For example, application of the FLB derivatisation approach to the analysis of sterols in crime scene material such as the soils underneath human cadavers or to obtain total sterol profiles of algae, could be explored. In addition, FLB derivatives could provide significant potential benefit in the detection of hydroxy cholesterol which is involved in metabolic signalling of p450 enzyme activity (Griffiths *et al.* 2017; Zu *et al.* 2020).

4. The improved capability for trace analysis could be evaluated by UHPLC separation equipped with MS and FL detection to measure FLB derivatives of cholesterol, coprostanol and 5α-cholestanol in situations where the conventional GC-MS based approach is limited by the amounts of the C_{27} steroidal lipids available (low concentrations, below 30 ng/g of dry soil, or small sample size less than 5 g).

5. The amounts of cholesterol and its related stanols determined by FL detection were slightly different than those determined by MS and UV detection. Hence, examination of whether UHPLC-FL chromatograms would determine if comparable results to MS and UV detection would be obtained under conditions with comparable resolution. Partial co-elution of cholesterol and γ-tocopherol was noted in the UHPLC-MS chromatograms from some of the Humber estuary sediment extracts, suggesting a UHPLC method could also be further modified to achieve baseline separation between these two components.

The modified approach for the use of FLB as a derivatising agent to obtain better signal response for GDGTs revealed that the measured amounts of GDGTs gave slightly lower estimates for the TEX_{86} and BIT indices than those reported for sediment extracts in the round-robin study. Thus, suggestions that further evaluation of the suitability of the approach for GDGT analysis could include the following four aspects:

1. The MS conditions for detection of FLB derivatives of GDGTs could be further explored to test if the ionisation efficiency of the various GDGT structures is maximal. Strategies to improve ionisation efficiency could include changes to the ESI conditions, the use of other types of mass spectrometers including orbitraps and FT-ICR, and changes in mobile phase composition to increase polarity.

2. Assuming that different GDGT isomers co-elute, further modification of the UHPLC separation could be explored to attempt separation of isomeric structures. This would require changes in mobile phase composition or use of different reversed-phase stationary phases. Alternatively, a different separation mechanism such as capillary

electrophoresis equipped with either MS or fluorescence detection could be explored given the potential for better resolution within an electro-osmotic induced flow (Mischak *et al.* 2009; Stolz *et al.* 2019).

3. The supression of the APCI MS response of the FLB-GDGTs (M_w > 2000 Da) compared with the native structures (M_w >1000 Da) may indicate an inverse relationship between ionisation efficiency and mass. Smaller fluorophore containing derivatisation agents could be explored to evaluate if they would be more suitabile derivatising agents for GDGTs.

4. The reversed phase UHPLC separation developed in the study could be applied to examination of microbial extracts with the purpose of identify novel components and unusual GDGT distributions (see the method detail in Section 2.6.2). For example, in Figures 5.1 and 5.2, the group peaks labelled B and peaks eluted in the retention time window 27-40 min are presently unassigned and could be targeted for identification.

In conclusion, the work has demonstrated significant advances in the use of FLB for detection of biological markers in soils and sediments which have one of more alcohol groups: steroidal lipids and tetraether lipid cores. The detection capabilities of the improved FLB approach for the analysis of these biomarkers with variety of detectors (MS, UV, and FL) indicates a significant advantage of the use of these derivatives. Quantification of these components by FLB allows the improved capability for trace analysis in situation where the conventional approaches are unable to give reliable results due to low amounts of analyte close to the detection limit of the instruments. The improved capability of measuring steroidal and tetraether lipid cores as FLB derivatives offers the potential for application in studies of palaeoenvironments and in environmental analysis.

Appendix A

Method development for the analysis of cholesterol and related stanols as their Fmoc-Lys-Boc derivatives

A.1 Separation conditions for the FLB derivatives of coprostanol and 5α-cholestanol

A.1.1 Separation conditions of the FLB derivatives of cholesterol and its reduced products proposed by Poplawski (2017)

The cholesterol-Fmoc-lysine-Boc derivative (chol-FLB), coprostanol-Fmoc-lysine-Boc derivative (cop-FLB) and 5α-cholestanol-Fmoc-lysine-Boc derivative (5α-chol-FLB) were analysed using an Ultimate 3000 rapid separation chromatograph (Dionex, Sunnyvale, California, US) coupled with the HCTUltra ETD II ion trap spectrometer operated as described in the section on mass spectrometry. Separation was carried out by Dionex Acclaim RSLC 120Å C_{18} 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 = ethyl acetate; delivered using the gradient program in Table A.1, unless stated otherwise.

Table A.1. Gradient programming of reversed phase LC-MS for separation of cholesterol-FLB derivatives (Poplawski 2017).

Time (min)	Percentage of composition in the solvent eluent (%)		
	Methanol	**0.5% Acetic acid**	**Ethyl acetate**
0	73	12	15
3	73	12	15
10	62	8	30
13	52	8	40
20	52	8	40
22	73	12	15

A.1.2 Separation conditions of the FLB derivatives of cholesterol and its reduced products by different HPLC mobile phase gradients

The chol-FLB, cop-FLB and 5α-chol-FLB, were analysed using an Agilent 1100-UV series chromatograph (Palo Alto, California, US) with the UV detector set at 263 nm and a Jasco FP-920 fluorescence detector operated with excitation wavelength (λ_{ex}) 263 nm and emission wavelength (λ_{em}) 309 nm. The diode array detector (DAD) was monitored over the range 190-400 nm (2 nm resolution) and recorded at peak width response time > 0.05 min (1 s). Separation was carried out using a Symmetry C_{18} column (3.5 µm, 2.1 mm×150 mm) maintained at 45°C using a ternary solvent system, where A = methanol, B = 0.5% acetic acid in water, C = ethyl acetate; delivered using gradient program (Table A.1) with a constant flow rate of 0.5 mL min^{-1}, unless stated otherwise. Each sample of sterol derivatives was analysed in triplicate (n = 3).

Table A.2. Gradient programming of reversed phase HPLC separation stated as condition B.

Time (min)	Percentage of composition in the solvent eluent (%)		
	Methanol	**0.5% Acetic acid**	**Ethyl acetate**
0	73	12	15
3	73	12	15
7	67.7	11	21.3
12	62	10	30
20	52	8	40
25	52	8	40
27	73	12	15
30	73	12	15

Table A.3. Gradient programming of reversed phase HPLC separation stated as condition C.

Time (min)	Percentage of composition in the solvent eluent (%)		
	Methanol	**0.5% Acetic acid**	**Ethyl acetate**
0	68.8	11.2	20
3	68.8	11.2	20
7	67.7	11	21.3
12	62	10	30
20	52	8	40
25	52	8	40
27	73	12	15
30	73	12	15

Table A.4. Gradient programming of reversed phase HPLC separation stated as condition D.

Time (min)	Percentage of composition in the solvent eluent (%)		
	Methanol	**0.5% Acetic acid**	**Ethyl acetate**
0	67.7	11	21.3
3	67.7	11	21.3
7	62	10	30
12	52	8	40
20	52	8	40
25	67.7	11	21.3
27	67.7	11	21.3
30	67.7	11	21.3

Appendix B

Application of the measurement of cholesterol and its stanols derivatives

B.1 Screening of steroidal compounds in cesspit core from Hungate, York by conventional GC-MS

Screening of steriodal compounds by GC-MS was performed using four soil samples obtained from different locations within an Anglo-Scandivanian cesspit and adjacent Roman grave (Hungate, York). Analysis of trimethylsilylether (TMS) derivatised organic soil extracts was carried out by GC-MS to obtain chromatograms of the steroidal components in their TMS form. The total ion chromatograms of the derivatised extract show complex distributions with many closely eluting components within the region where the C_{27} steroidal components elute (40-42 min, see some examples of chromatograms in Figure B.1).

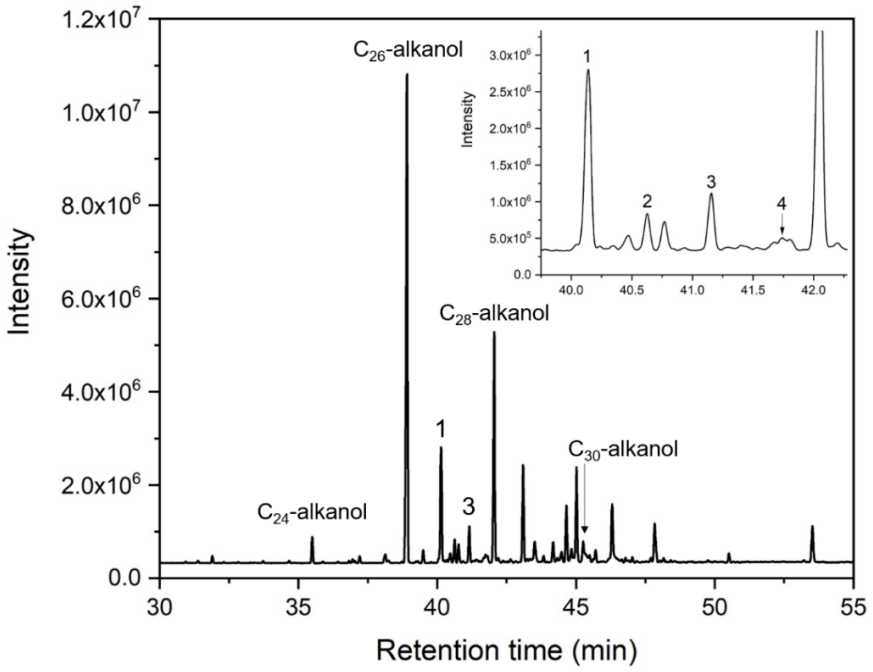


Figure B.1. Partial GC-MS chromatogram of the medium polar fraction of the soil extract from cesspit core at depth 30-31 cm, Hungate 51150 (0.6 mg) derivatised with TMS. Assigned peaks: 1 = coprosanol, 2 = epicoprostanol, 3 = cholestan-3-one, 4 = 5α-cholestanol.

Table B.1. List of organic compounds identified in the soil extracts from Anglo-Scandinavian cesspit (Hungate, York).

Compound identified	RT (min)	Cesspit core depth						
		0-1 cm	4-5 cm	8-9 cm	9-10 cm	13-14 cm	15-16 cm	16-17 cm
Sterols								
Coprostanol	40.14	✓	✓	✓	✓	✓	✓	✓
Epicoprostanol	40.63	-	✓	✓	✓	✓	✓	✓
5α-cholestanol	41.74	✓	✓	✓	✓	✓	✓	✓
Cholestan-3-one	41.15	✓	-	✓	✓	-	-	✓
Stigmastanol, 2-ethylbutarate	43.09	✓	✓	✓	✓	✓	✓	✓
Stigmastanol	43.51	✓	✓	✓	✓	✓	✓	-
Stigmasterol	43.73	-	-	-	-	-	-	✓
24-Ethyl-δ(22)-coprostenol	43.88	-	-	-	-	-	-	✓
Lanost-8-en-3-ol (3β)	43.83	-	-	-	-	✓	-	-
Stigmastane, 3-oxo	44.17	-	-	-	✓	-	-	✓
β-sitosterol	44.48	-	-	-	✓	-	-	✓

Compound identified	RT (min)	Cesspit core depth						
		0-1 cm	4-5 cm	8-9 cm	9-10 cm	13-14 cm	15-16 cm	16-17 cm
3β-lanost-9(11), 24-dien-3-ol	45.49	-	-	-	✓	-	-	-
Triterpene								
β-Amyrin	44.65	-	✓	✓	✓	✓	✓	✓
α-Amyrin	45.26	-	✓	✓	✓	✓	✓	✓
Lupeol	45.71	-	-	-	-	✓	-	-
Alkanols								
1-heptadecanol (internal standard)	21.92	✓	✓	✓	✓	✓	✓	✓
1-docosanol	31.89	✓	✓	✓	✓	✓	✓	✓
1-tetracosanol	35.50	✓	✓	✓	✓	✓	✓	✓
1-pentacosanol	37.21	-	✓	✓	✓	✓	✓	✓
1-hexacosanol	38.91	✓	✓	✓	✓	✓	✓	✓
1-heptacosanol	40.47	✓	✓	✓	✓	✓	✓	✓
1-octacosanol	42.05	✓	✓	✓	✓	✓	✓	✓

Compound identified	RT (min)	Cesspit core depth						
		0-1 cm	4-5 cm	8-9 cm	9-10 cm	13-14 cm	15-16 cm	16-17 cm
1-triacosanol	45.02	-	✓	✓	✓	✓	✓	✓
Others								
Neophytadiene	16.61	-	-	-	✓	-	✓	✓
Dibutyl phthalate	19.26	-	-	-	✓	-	-	✓
Plamitic acid	21.43	-	-	-	✓	-	-	-
Stearic acid	25.74	-	-	-	✓	-	-	-
Unknown	27.54	-	-	-	✓	-	-	-
1,2-benzenedicarboxylic acid bis(2-methylpropyl) ester	31.39	✓	-	-	-	-	-	✓
13-docosenamide	35.87	-	-	-	-	-	-	✓
octadecanamide	36.34	-	-	-	-	-	-	✓
Hexacosene	38.09	-	-	-	-	-	-	✓
γ-tocopherol	39.49	✓	✓	✓	✓	✓	✓	-
α-tocopherol	41.68	-	✓	✓	-	✓	✓	-

Compound identified	RT (min)	Cesspit core depth						
		0-1 cm	4-5 cm	8-9 cm	9-10 cm	13-14 cm	15-16 cm	16-17 cm
β-tocopherol	43.48	-	-	-	-	✓	-	-
bisphenol A monomethyl ester	40.77	-	✓	-	✓	-	-	✓
Cholet-5-ene, 3β-cholro	46.30	-	-	-	✓	-	-	-
Unknown	47.83	-	-	-	✓	✓	✓	✓
Tris(2,4-di-tert-butylphenyl)phosphate	48.15	-	-	-	-	-	-	✓
4,6-dibenzothiophene diyl bis(diphenyl)	50.51	-	-	✓	✓	✓	-	-
Irganox 1076	50.68	-	✓	-	-	✓	✓	✓
Unknown	53.52	-	-	-	✓	-	-	-

✓ means the component was detected, while – means the component was non-detected.

Table B.1 (cont). List of organic compounds identified in the soil extracts from Anglo-Saxon cesspit (Hungate, York).

Compound identified	RT (min)	Cesspit-core depth						
		24-25 cm	25-26 cm	27-28 cm	28-29 cm	29-30 cm	30-31 cm	39-40 cm
Sterols								
Coprostanol	40.14	✓	✓	✓	✓	✓	✓	✓
Epicoprostanol	40.63	✓	✓	✓	✓	✓	✓	-
5α-cholestanol	41.74	✓	✓	✓	✓	✓	✓	✓
Cholestan-3-one	41.15	✓	✓	✓	✓	✓	✓	✓
Stigmastanol, 2-ethylbutarate	43.09	✓	✓	✓	✓	✓	✓	✓
Stigmastanol	43.51	✓	✓	✓	✓	✓	✓	✓
Lanost-8-en-3-ol (3β)	43.83	✓	-	✓	✓	-	✓	-
Stigmastane, 3-oxo	44.17	✓	✓	✓	✓	✓	✓	-
β-sitosterol	44.48	-	-	✓	-	✓	✓	✓
Stigmasterol	43.73	-	-	-	-	-	-	-
24-Ethyl-δ(22)-coprostenol	43.88	-	-	-	-	-	-	-

Compound identified	RT (min)	Cesspit-core depth						
		24-25 cm	25-26 cm	27-28 cm	28-29 cm	29-30 cm	30-31 cm	39-40 cm
3β-lanost-9(11), 24-dien-3-ol	45.49	-	-	-	-	✓	-	-
Triterpene								
β-Amyrin	44.65	✓	✓	✓	✓	✓	✓	✓
α-Amyrin	45.26	✓	✓	✓	✓	✓	✓	✓
Lupeol	45.71	-	-	-	-	-	-	✓
Alkanols								
1-heptadecanol (an internal standard)	21.92	✓	✓	✓	✓	✓	✓	✓
1-docosanol	31.89	✓	✓	✓	✓	✓	✓	✓
1-tetracosanol	35.50	✓	✓	✓	✓	✓	✓	✓
1-pentacosanol	37.21	-	✓	✓	✓	✓	✓	✓
1-hexacosanol	38.91	✓	✓	✓	✓	✓	✓	✓
1-heptacosanol	40.47	✓	✓	✓	✓	✓	✓	✓
1-octacosanol	42.05	✓	✓	✓	✓	✓	✓	✓

Compound identified	RT (min)	Cesspit-core depth						
		24-25 cm	25-26 cm	27-28 cm	28-29 cm	29-30 cm	30-31 cm	39-40 cm
1-triacosanol	45.02	✓	✓	✓	✓	✓	✓	✓
Others								
Neophytadiene	16.61	✓	-	✓	-	-	✓	✓
Dibutyl phthalate	19.26	-	✓	✓	-	✓	✓	-
Plamitic acid	21.43	-	-	✓	-	✓	✓	-
Stearic acid	25.74	-	-	✓	✓	✓	✓	-
Unknown	27.54	-	✓	✓	-	-	✓	-
1,2-benzenedicarboxylic acid bis(2-methylpropyl) ester	31.39	-	-	-	-	-	-	-
13-docosenamide	35.87	-	-	✓	-	✓	-	-
octadecanamide	36.34	-	-	✓	-	✓	-	-
Lignoceric acid	37.21	-	-	-	-	-	-	-
Hexacosene	38.09	-	-	-	-	-	-	-
γ-tocopherol	39.49	✓	✓	✓	✓	✓	✓	✓

Compound identified	RT (min)	Cesspit-core depth						
		24-25 cm	25-26 cm	27-28 cm	28-29 cm	29-30 cm	30-31 cm	39-40 cm
α-tocopherol	41.68	✓	✓	✓	✓	✓	✓	✓
β-tocopherol	43.74	✓	✓	✓	✓	✓	✓	✓
bisphenol A monomethyl ester	40.77	-	-	✓	-	✓	-	-
Cholet-5-ene, 3β-cholro	46.30	-	-	✓	✓	✓	✓	-
Unknown	47.83	-	-	✓	✓	✓	✓	-
Tris(2,4-di-tert-butylphenyl)phosphate	48.15	✓	-	-	-	-	-	-
4,6-dibenzothiophene diyl bis(diphenyl)	50.51	-	-	✓	✓	✓	✓	✓
Irganox 1076	50.68	✓	✓	-	-	-	-	✓
Unknown	53.52	-	✓	-	✓	-	✓	✓

✓ means the component was detected, while – means the component was non-detected.

B.2 Determinatination of steroidal compounds in cesspit core from Hungate, York by FLB derivatives

Detection of steriodal compounds by UHPLC-MS was performed using soil samples obtained from an Anglo-Scandivinian cesspit core (Hungate, York) (See Section 4.3.3.1.1.1).

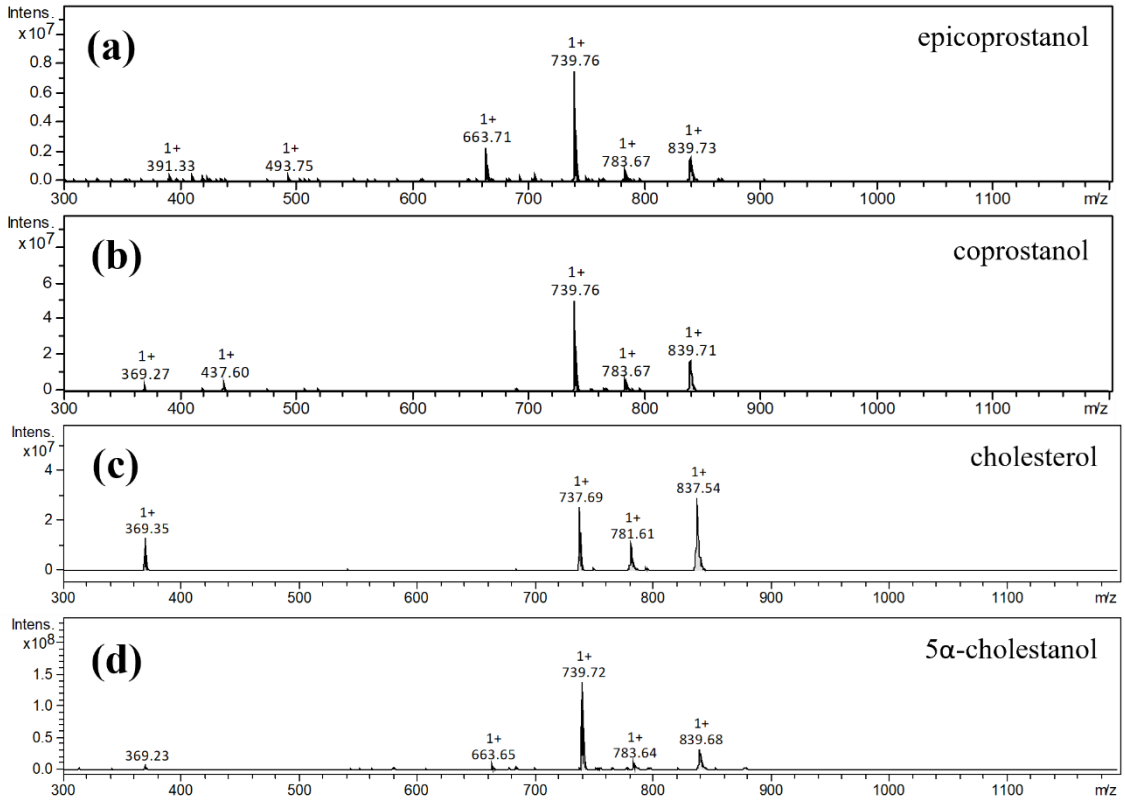


Figure B.2. APCI MS spectra of the identified FLB derivatives with C_{27} steriodal lipids found in cesspit core (Hungate, York): (a) epicoprostanol (9.0 min), (b) coprostanol (9.3 min), (c) cholesterol (9.9 min), (d) 5α-cholestanol (10.4 min).

Table B.2. The evaluted amount of FLB derivative of coprostanol (cop-FLB) obtained from soil extracts of the suspected Anglo-Scandinavian cesspit (Hungate, York) recording by three detectors (MS, UV and FL) , *n* =3.

Cesspit core from different depths (cm)	**Detected coprostanol (×10^{-9} gram per gram of dried soil, ng/g dw)**		
	UHPLC-MS	**UHPLC-UV**	**HPLC-FL**
Zone 1: Light brown			
0-1	28 ± 1	26 ± 1	25 ± 1
4-5	240 ± 8	230 ± 10	200 ± 10
8-9	2,660 ± 65	2,610 ± 70	2,130 ± 30
Zone 2: Dark brown			
9-10	2,150 ± 90	2,180 ± 100	1,950 ± 80
13-14	3,160 ± 60	3,110 ± 70	2,830 ± 60
15-16	5,990 ± 230	5,890 ± 260	5,330 ± 190
16-17	3,150 ± 30	3,200 ± 40	2,850 ± 30
24-25	26,900 ± 1,210	26,480 ± 1,310	23,940 ± 960
25-26	2,960 ± 90	2,900 ± 90	2,630 ± 70
27-28	26,100 ± 1,100	25,700 ± 700	23,230 ± 600
28-29	20,600 ± 550	20,340 ± 580	18,300 ± 430
Zone 3: Very dark brown			
29-30	25,900 ± 800	24,400 ± 730	23,000 ± 540
30-31	47,700 ± 2,040	46,800 ± 2,000	42,450 ± 1,810
39-40	26,600 ± 1,130	26,200 ± 1,300	23,670 ± 950

Table B.3. The evaluted amount of FLB derivative of epicoprostanol (epicop-FLB) obtained from soil extracts of the suspected Anglo-Scandinavian cesspit (Hungate, York) recording by three detectors (MS, UV and FL) , *n* =3.

Cesspit core from different depths (cm)	**Detected epicoprostanol (×10^{-9} gram per gram of dried soil, ng/g dw)**		
	UHPLC-MS	**UHPLC-UV**	**HPLC-FL**
Zone 1: Light brown			
0-1	2.45 ± 0.03	2.41 ± 0.05	2.34 ± 0.05
4-5	30 ± 0.4	30 ± 0.3	28 ± 0.2
8-9	180 ± 7	177 ± 5	170 ± 4
Zone 2: Dark brown			
9-10	160 ± 10	160 ± 7	156 ± 6
13-14	1,460 ± 20	1,440 ± 13	1420 ± 10
15-16	1,820 ± 60	1,790 ± 40	1,700 ± 40
16-17	750 ± 8	740 ± 12	700 ± 10
24-25	6,770 ± 290	6,670 ± 310	6,200 ± 260
25-26	290 ± 10	285 ± 12	280 ± 10
27-28	7,370 ± 240	7,260 ± 360	6,900 ± 300
28-29	9,300 ± 430	9,160 ± 370	8,370 ± 360
Zone 3: Very dark brown			
29-30	5,510 ± 250	5,430 ± 270	5,290 ± 230
30-31	8,210 ± 140	8,080 ± 210	7,340 ± 210
39-40	6,740 ± 180	6,640 ± 270	6,530 ± 250

Table B.4. The evaluted amount of FLB derivative of 5α-cholestanol (5α-cholFLB) obtained from soil extracts of the suspected Anglo-Scandinavian cesspit (Hungate, York) recording by three detectors (MS, UV and FL) , *n* =3.

Cesspit core from different depths (cm)	**Detected 5α-cholestanol (×10^{-9} gram per gram of dried soil, ng/g dw)**		
	UHPLC-MS	**UHPLC-UV**	**HPLC-FL**
Zone 1: Light brown			
0-1	N.D.	N.D.	N.D.
4-5	N.D.	N.D.	0.076 ± 0.003
8-9	85 ± 3	84 ± 4	72 ± 3
Zone 2: Dark brown			
9-10	87 ± 0.4	85 ± 1	72 ± 0.3
13-14	430 ± 20	420 ± 30	360 ± 20
15-16	140 ± 3	138 ± 5	120 ± 4
16-17	220 ± 7	217 ± 5	190 ± 5
24-25	1,630 ± 80	1,590 ± 70	1,380 ± 60
25-26	300 ± 10	270 ± 10	250 ± 7
27-28	320 ± 10	315 ± 7	280 ± 5
28-29	1,120± 20	1,050 ± 18	950 ± 20
Zone 3: Very dark brown			
29-30	1,400 ± 60	1,340 ± 60	1,130 ± 50
30-31	3,210 ± 160	3,300 ± 160	2,900 ± 120
39-40	770 ± 30	760 ± 40	590 ± 20

N.D. = Non-determined

Table B.5. The evaluted amount of FLB derivative of cholesterol (chol-FLB) obtained from soil extracts of the suspected Anglo-Scandinavian cesspit (Hungate, York) recording by three detectors (MS, UV and FL) , *n* =3.

Cesspit core from different depths (cm)	Detected cholesterol (×10^{-9} gram per gram of dried soil, ng/g dw)		
	UHPLC-MS	**UHPLC-UV**	**HPLC-FL**
Zone 1: Light brown			
0-1	4.5 ± 0.2	4.6 ± 0.3	4 ± 0.2
4-5	40 ± 0.2	41 ± 0.3	40 ± 0.2
8-9	74 ± 2	73 ± 3	69 ± 2
Zone 2: Dark brown			
9-10	19 ± 0.1	19 ± 0.1	17 ± 0.1
13-14	20 ± 0.3	20 ± 0.2	18 ± 0.1
15-16	43 ± 1	42 ± 2	39 ± 1
16-17	170 ± 5	167 ± 8	160 ± 5
24-25	100 ± 4	105 ± 3	94 ± 2
25-26	54 ± 2	51 ± 2	49 ± 2
27-28	130 ± 4	130 ± 6	120 ± 4
28-29	270 ± 10	266 ± 10	250 ± 8
Zone 3: Very dark brown			
29-30	110 ± 5	105 ± 5	96 ± 4
30-31	120 ± 6	116 ± 5	108 ± 4
39-40	N.D.	N.D.	N.D.

N.D. = Non-determine

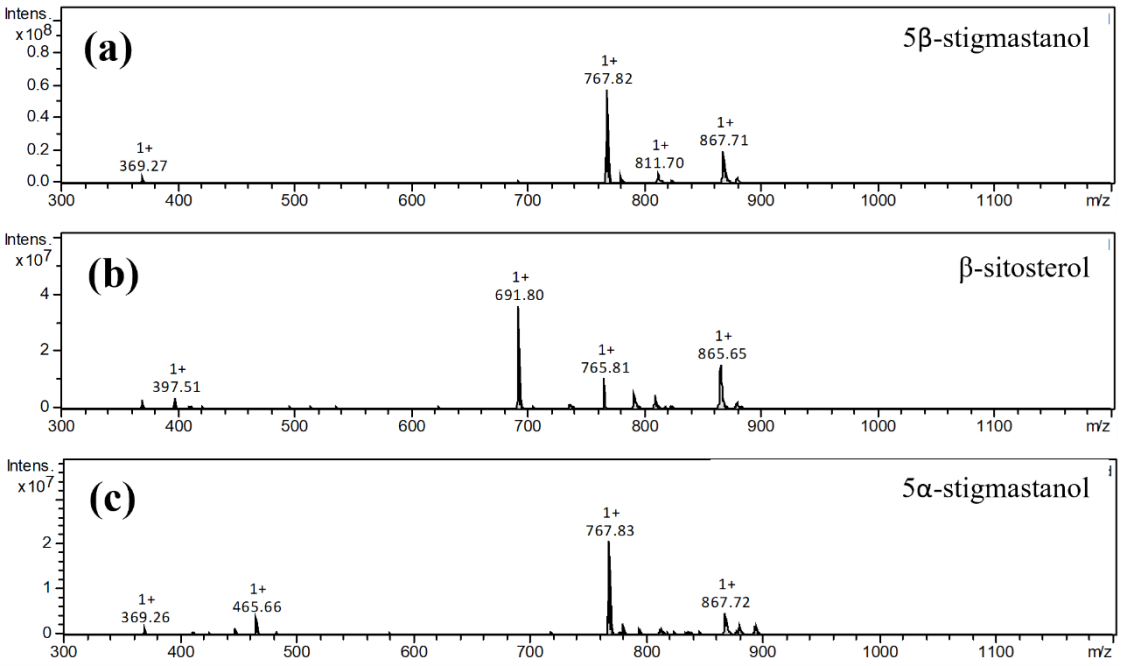


Figure B.3. APCI MS spectra of the identified FLB derivatives with C_{29} steriodal lipids found in cesspit core (Hungate, York): (a) 5β-stigmastanol (11.0 min), (b) β-sitosterol (11.3 min), (c) 5α-stigmastanol (11.5 min).

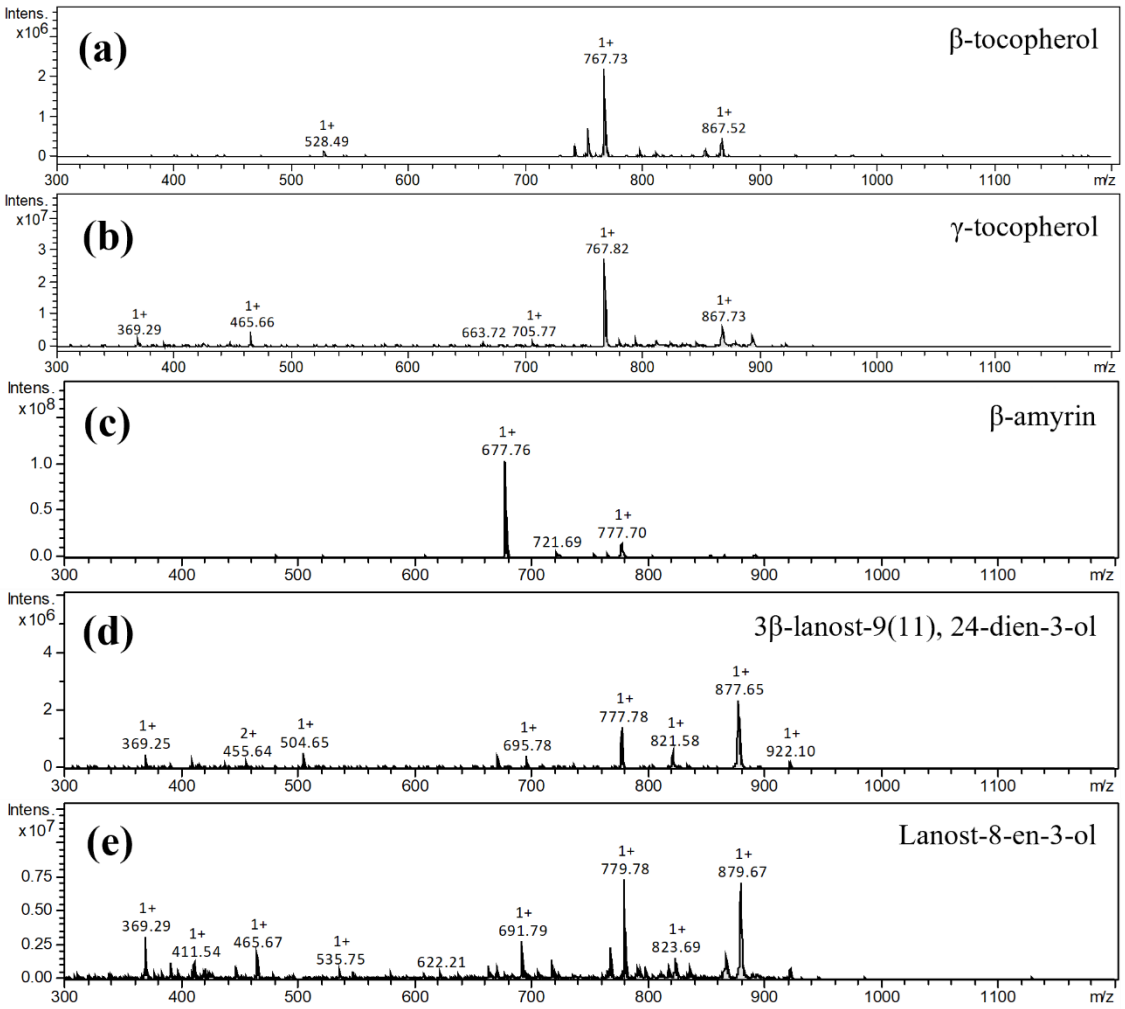


Figure B.4. APCI MS spectra of the identified FLB derivatives with C_{30} steroidal lipids, tocotrienols and tripenes found in cesspit core (Hungate, York): (a) β-tocopherol (10.6 min), (b) γ-tocopherol (11.5 min), (c) β-amyrin (10.2 min), (d) 3β-lanost-9(11), 24-dien-3-ol (11.2 min), (e) lanost-8-en-3-ol (12.7 min).

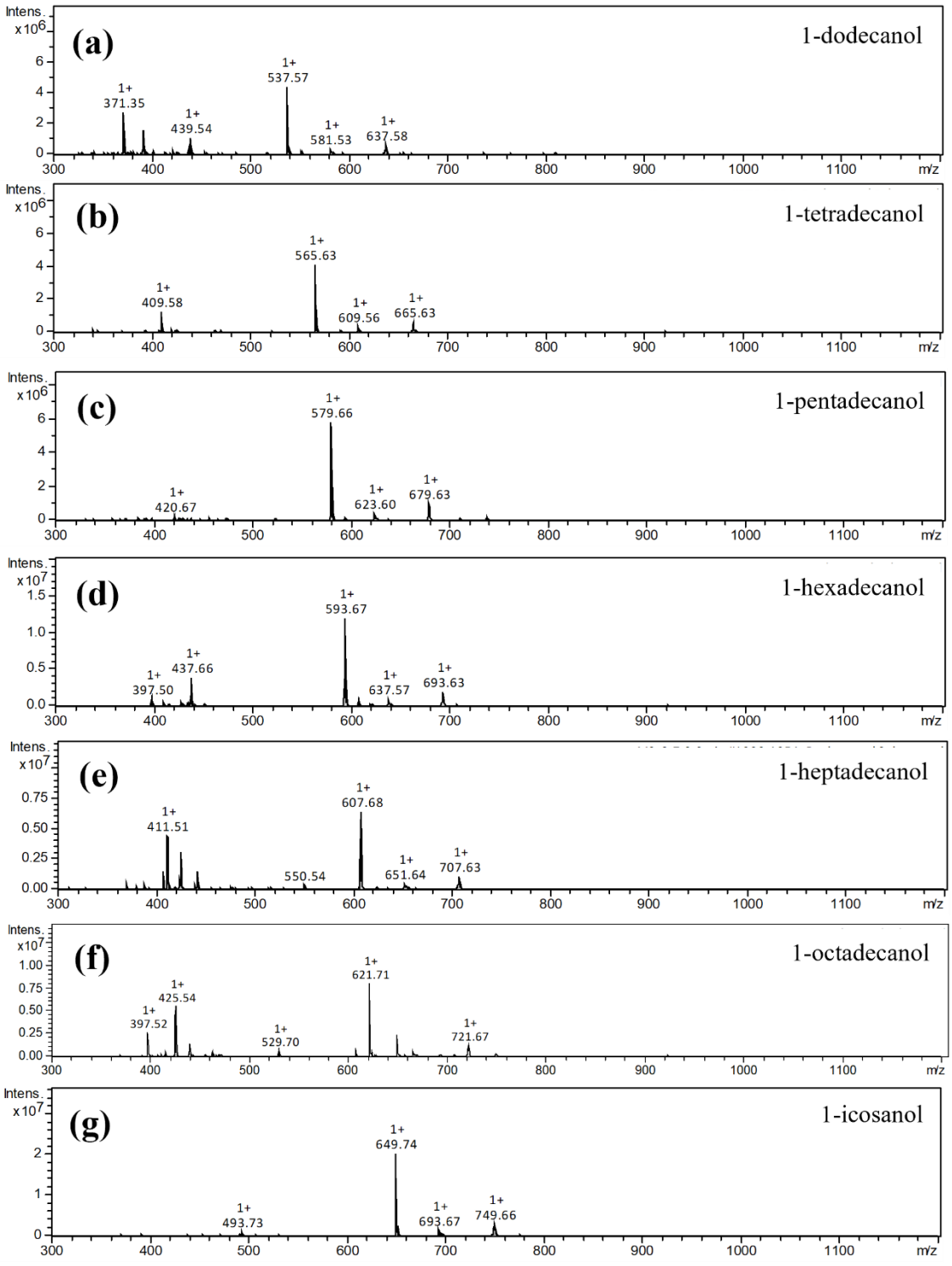


Figure B.5. APCI MS spectra of the identified FLB derivatives with *n*-alkanols found in cesspit core (Hungate, York): (a-r) the identified *n*-alkanols were in range of C_{12}-C_{36}.

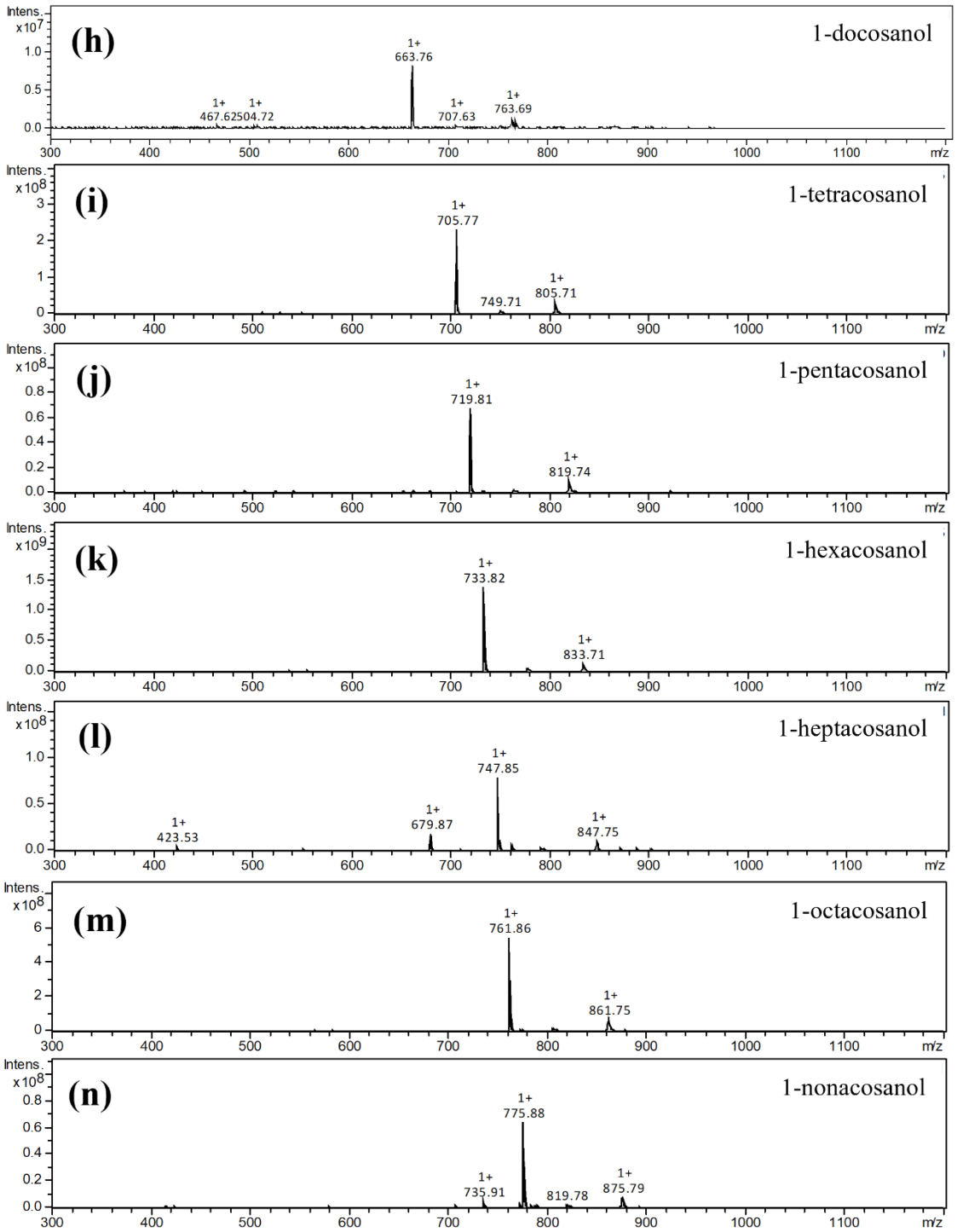


Figure B.5 (cont.). APCI-MS spectra of the identified FLB derivatives with *n*-alkanols in cess pit found in cesspit core (Hungate, York): (a-r) the identified *n*-alkanols were in range of C_{12}-C_{36}.

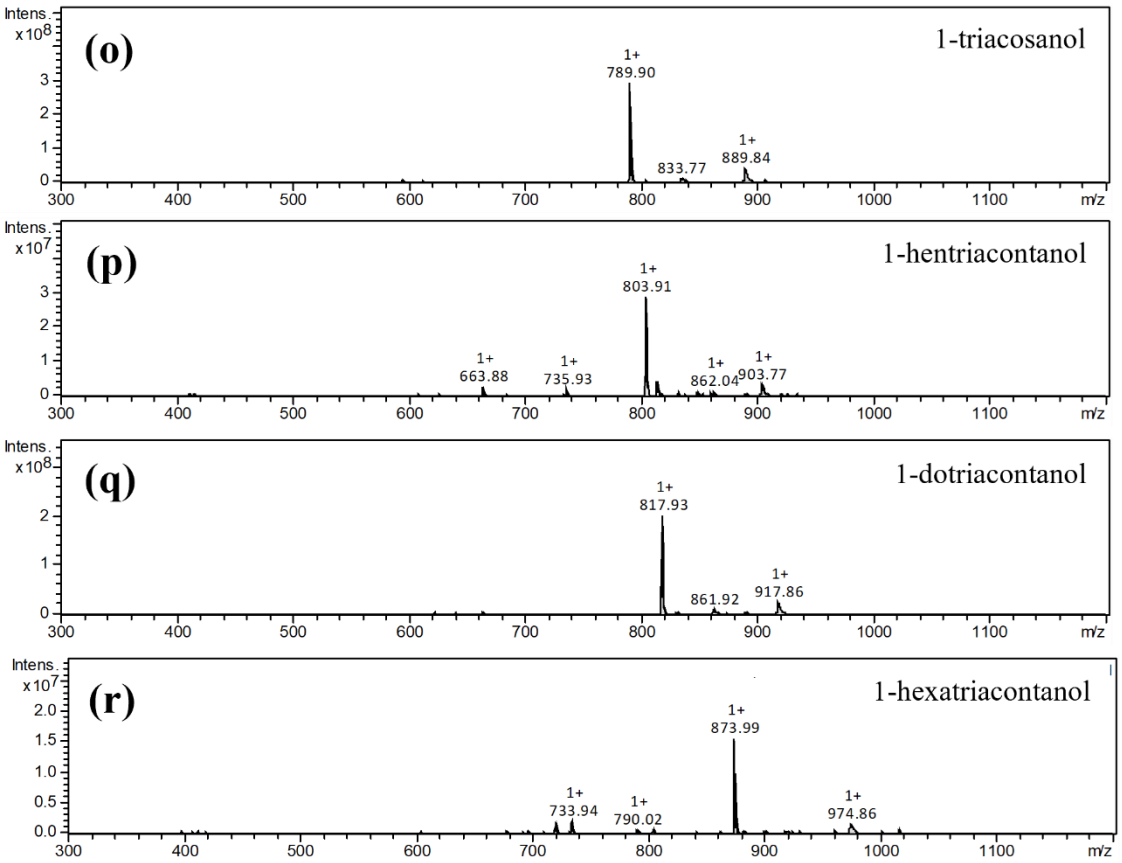


Figure B.5 (cont.). APCI-MS spectra of the identified FLB derivatives with n-alkanols found in cesspit core (Hungate, York): (a-r) the identified *n*-alkanols were in range of C_{12}-C_{36}.

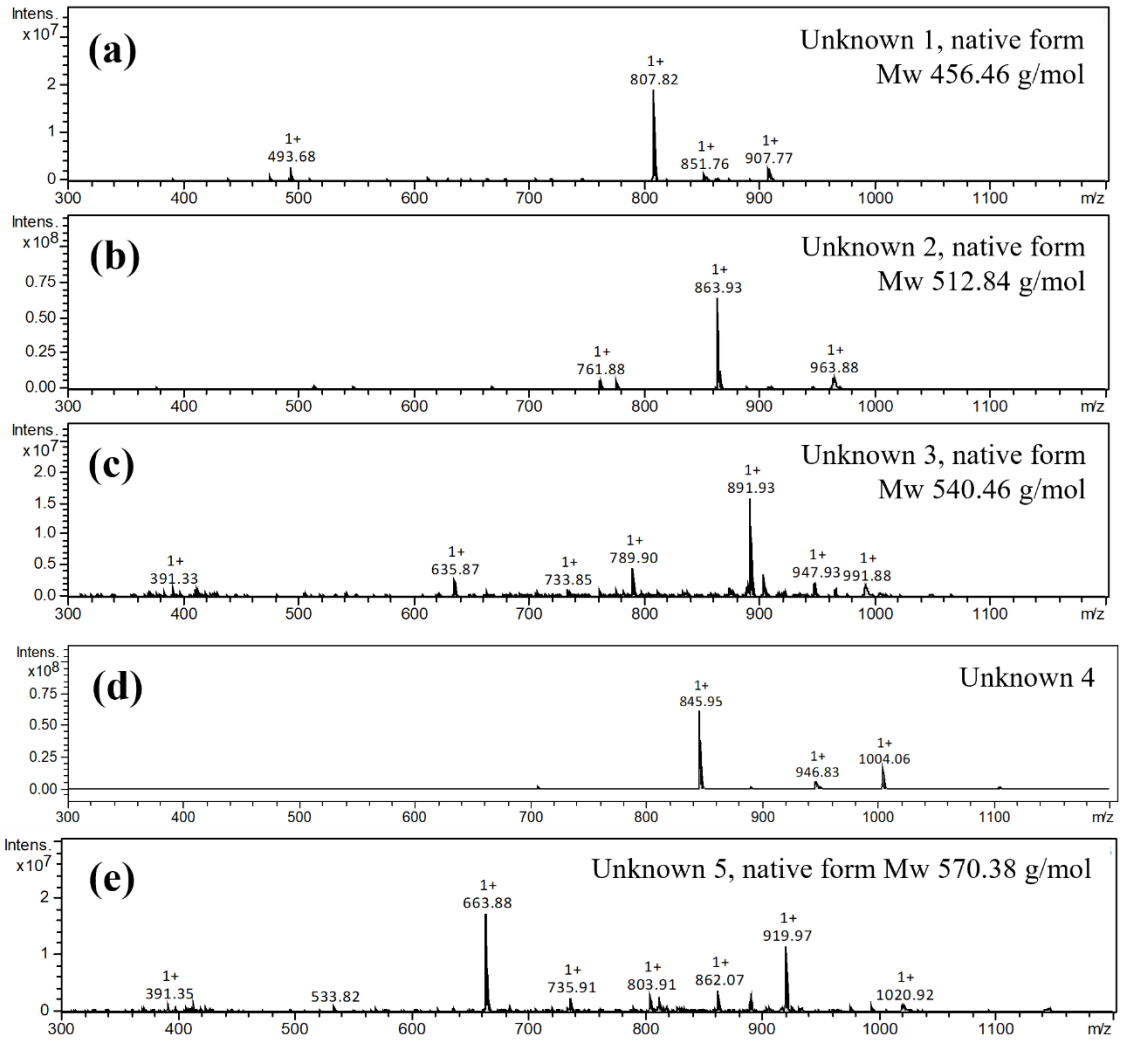


Figure B.6. APCI MS spectra of the identified FLB derivatives with unknown components found in cesspit core (Hungate, York): (a) unknown 1 (12.8 min), (b) unknown 2 (14.5 min), (c) unknown 3 (15.7 min), (d) unknown 4 (17.3 min), (e) unknown 5 (18.9 min).

Appendix C

Method development for analysis of GDGTs as their Fmoc-Lys-Boc derivatives

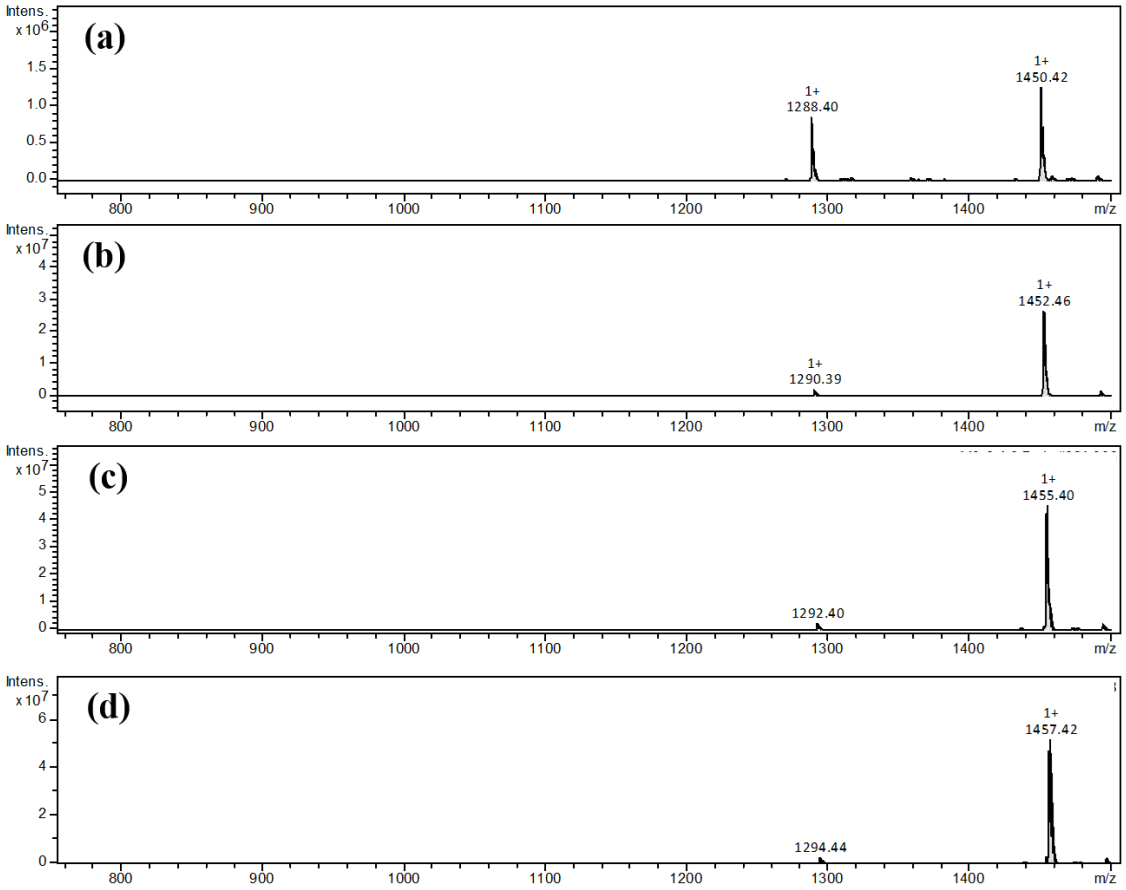


Figure C.1. APCI MS spectra of glycerol dialkyl calditol tetraethers (GDCTs) obtaining from peak group A of the lipid extract from *S. solfataricus* in retention time window 12-15 min (Figure 5.2).

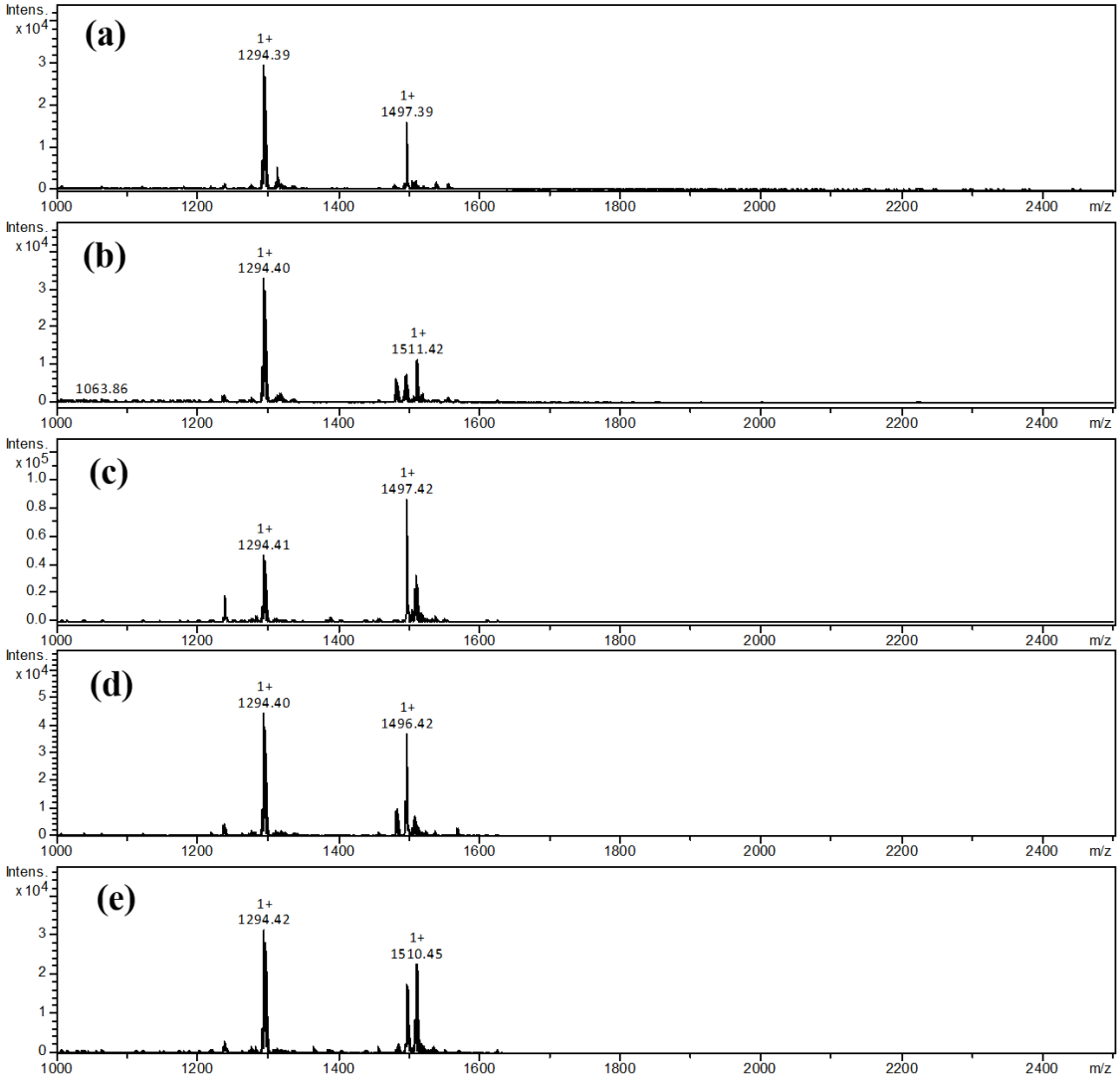


Figure C.2. APCI MS spectra of peak group B of the lipid extract from *S. solfataricus* in retention time window 15-20 min (Figure 5.2).

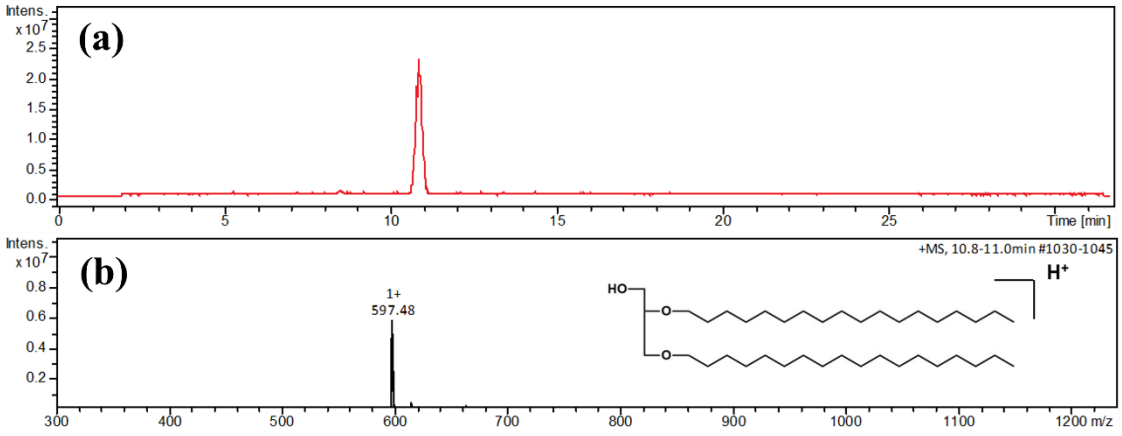


Figure C.3. (a) APCI base peak chromatogram of 1,2-di-*O*-octadecyl-*rac*-glycerol (r-dOG), (b) Full MS spectrum of r-dOG inset with its chemical structure. Concentration of r-dOG was 20 pmol injected on column.

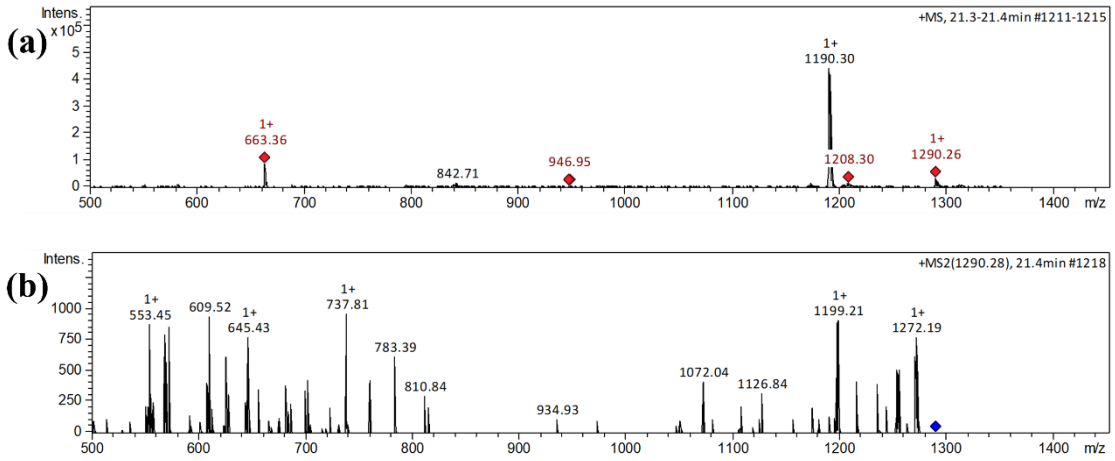


Figure C.4. APCI MS spectra of the suspected *i*-GDGT-6 (a) Full MS spectrum of, (b) MS^2 spectrum of the protonated molecule at *m/z* 1290.3 referring to the peak (Figure 5.31, red trace).

Table C.1. Assignment of the product ions produced from the precursor ion at *m/z* 1290.3 (Figure C.4).

Mass	Molecule	Abundance [%]	MS loss from *m/z* 1290	Assignment
737.63	**$[M+H]^+$**	**100**	**552.37**	**Neutral loss of BP_3, $C_{40}H_{72}$**
1272.26	$[M+H]^+$	59.25	17.74	Loss of water
553.43	**$[M+H]^+$**	**76.17**	**736.57**	**BP_3, $C_{40}H_{73}^+$**
1199.23	$[M+H]^+$	76.31	92.77	Loss of one glycerol group
645.54	$[M+H]^+$	66.48	644.46	Neutral loss of BP_3, $C_{40}H_{72}$ and glycerol group
643.54	$[M+H]^+$	51.26	646.46	Neutral loss of BP_2, $C_{40}H_{72}$ and glycerol group
719.53	$[M+H]^+$	46.98	570.47	Neutral loss of BP_3 chain attached with O
1254.28	$[M+H]^+$	38.68	35.72	$2H_2O$
555.02	$[M+H]^+$	9.82	734.98	BP_2, $C_{40}H_{75}^+$

Abbreviations

AA	Acetic acid
ASE	Accelerated solvent extraction
ACN	Anthroyl cyanide
APCI	Atomic pressure chemical ionisation
BBr_3	Boron tribromide
BIT	Branched and isoprenoid tetraether
Boc	Butyloxycarbonyl
BPC	Base peak chromatogram
br-GDGT	Branched glycerol dialkyl glycerol tetraether
CE	Common Era
Chol-FLB	Cholesterol-Fmoc-Lysine-Boc
Cop-FLB	coprostanol-Fmoc-Lysine-Boc
cm	Centimetre
CSR	4′-carboxy-substituted rosamine
Da	Dalton
DCM	Dichloromethane
DMG	*N*,*N*-dimethyl glycine
DMAP	4-(dimethylamino) pyridine
EDC	1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride
EGC	Epigallocatechin
EIC	Extracted ion chromatograms
ESI	Electron spray ionisation
EtOAC	Ethyl acetate
FMOC	Fluorenylmethyloxycarbonyl

FL	Fluorescence
FLB	Fmoc-lysine-Boc
GC	Gas chromatography
GC-MS	Gas chromatography-mass spectrometry
GDCT	Glycerol dialkyl calditol tetraethers
GDGT	Glycerol dialkyl glycerol tetraether
GMGT	Glycerol monoalkyl glycerol tetraether
GTGT	Glycerol trialkyl glycerol tetraether
HI	Hydroiodic acid
HILIC	Hydrophilic interaction chromatography
HPLC-FL	High performance liquid chromatography-fluorescence
HPLC-UV	High performance liquid chromatography-ultraviolet
HTGC-FID	High temperature gas chromatography with flame ionisation detection
i.d.	Internal diameter
i-GDGT	Isoprenoid glycerol dialkyl glycerol tetraether
IPL	intact polar lipid
IS	Internal standard
$LiAlH_4$	Lithium aluminium hydride
LOD	Limit of detection or detection limit
LOQ	Limit of quantification
LVI	Large volume injection
MeOH	Methanol
mm	Millimetre
mL	Millilitre
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MS^n	Multistage tandem mass spectrometry

m/z	Mass-to-charge ratio
MTBE	Methyl tertbutyl ether
n	Number of analytical replicates
NCPM-SH	*N*-(4-(carbazole-9-yl)-phenyl)-*N*-maleimide labelled with ethylenedithiol
NP	Normal phase
NPLC	Normal-phase liquid chromatography
PFP	pentafluorophenyl
ppm	part per million
ppb	part per billion
P’	polarity
LC	liquid chromatography
r^2	linear regression
r-dOG	1,2-di-*O*-octadecyl-rac-glycerol
R_f	Retention factor
R_s	Resolution
RP	Reverse-phase
RPLC	Reverse-phase liquid chromatography
%RA	Percentage of relative abundance
%RSD	Percentage of relative standard deviation
SIM	Single ion monitoring
S/N	Signal-to-noise ratio
SRM/MRM	selected/multiple reaction monitoring
SST	Sea surface water temperature
TEX_{86}	Tetraether index of tetraethers consisting of 86 carbons
TEX_{86}^{L}	Tetraether index of tetraethers consisting of 86 carbons at temperature lower than 15°C

TEX_{86}^{H}	Tetraether index of tetraethers consisting of 86 carbons at temperature higher than 15°C
TMAH	Tetramethylammonium hydroxide
TMS	Trimethyl silyl ether
TLC	Thin layer chromatography
v/v	Volume per volume
UHPLC-MS	Ultra high performance liquid chromatography-mass spectrometry
UV	Ultra-violet
UV-vis	Ultraviolet-visible
1-hep-FLB	1-heptadecanol-Fmoc-Lysine-Boc
5α-chol-FLB	5α-cholestanol-Fmoc-Lysine-Boc
μm	Micrometre or micron
μL	Microlitre

References

Angelini, R.; Babudri, F.; Lobasso, S.; Corcelli, A. 2010. "MALDI-TOF/MS Analysis of Archaebacterial Lipids in Lyophilized Membranes Dry-mixed with 9-aminoacridine." *Journal of Lipid Research* 51(9): 2818-2825.

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): 263-274.

Barker, S.; Cacho, I.; Benway, H.; Tachikawa, K. 2005. "Planktonic Foraminiferal Mg/Ca as a Proxy for Past Oceanic Temperatures: A Methodological Overview and Data Compilation for the Last Glacial Maximum." *Quaternary Science Reviews* 24(7-9 SPEC. ISS.): 821–834.

Bataglion, G. A.; Meurer, E.; de Albergaria-Barbosa, A. C. R.; Bícego, M. C.; Weber, R. R.; Eberlin, M. N. 2015. "Determination of Geochemically Important Sterols and Triterpenols in Sediments Using Ultrahigh-Performance Liquid Chromatography Tandem Mass Spectrometry (UHPLC−MS/MS)." *Analytical Chemistry* 87(15): 7771-7778.

Bataglion, G. A.; Henrique, H.; Koolen, F.; Weber, R. R.; Eberlin, M. N. 2016. "Quantification of Sterol and Triterpenol Biomarkers in Sediments of the Cananéia-Iguape Estuarine-Lagoonal System (Brazil) by UHPLC-MS / MS." *International Journal of Analytical Chemistry* 8361375: 1-8.

Battistel, D.; Piazza, R.; Argiriadis, E.; Marchiori, E.; Radaelli, M.; Barbante, C. 2015. "GC-MS Method for Determining Faecal Sterols as Biomarkers of Human and Pastoral Animal Presence in Freshwater Sediments." *Analytical bioanalyitcal chemistry* 407: 8505–8514.

Bays, H. E.; Neff, D.; Tomassini, J. E.; Tershakovec, A. M. 2008. "Ezetimibe: Cholesterol Lowering and Beyond." *Expert Review of Cardiovascular Therapy* 6(4): 447–470.

Berg, I. A.; Kockelkorn, D.; Ramos-Vera, W. H.; Say, R. F.; Zarzycki, J.; Hügler, M.; Alber, B. E.; Fuchs, G. 2010. "Autotrophic Carbon Fixation in Archaea." *Nature Reviews Microbiology* 8(6): 447-460.

Bethell, P. H.; Goad, L. J.; Evershed, R. P. 1994. "The Study of Molecular Markers of Human Activity: The Use of Coprostanol in the Soil as an Indicator of Human Faecal Material." *Journal of Archaeological Science* 21: 619–632.

Bhattacharyya, A. K. 1982. "Differences in Uptake and Esterification of Saturated Analogues of Cholesterol by Rat Small Intestine." *American Journal of Physiology - Gastrointestinal and Liver Physiology* 251(4): 495-500.

Birk, J. J.; Teixeira, W. G.; Neves, E. G.; Glaser, B. 2011. "Faeces Deposition on Amazonian Anthrosols as Assessed from 5β-Stanols." *Journal of Archaeological Science* 38(6): 1209–1220.

Björkhem, I.; Gustafsson, J. 1971. "Mechanism of Microbial Transformation of Cholesterol into Coprostanol." *European Journal of Biochemistry* 21(3): 428–432.

Brassell, S. C.; Eglinton, G.; Marlowe, I. T.; Pflaumann, U.; Sarnthein, M. 1986. "Molecular Stratigraphy: A New Tool for Climatic Assessment." *Nature* 320(6058): 129-133.

Brassell, S. C. 2014. "Climatic Influences on the Paleogene Evolution of Alkenones." *Paleoceanography* 29(3): 255–272.

Brinkley, A. W.; Gottesman, A. R.; Mott, G. E. 1982. "Isolation and Characterization of New Strains of Cholesterol-reducing Bacteria from Baboons." *Applied and Environmental Microbiology* 43(1): 86-89.

Brock, T. D.; Brock, K. M.; Belly, R. T.; Weiss, R. L. 1972. "Sulfolobus: A New Genus of Sulfur-oxidizing Bacteria Living at Low pH and High Temperature." *Archiv für Mikrobiologie* 84(1): 54-68.

Buckles, L. K.; Villanueva, L.; Weijers, J. W. H.; Verschuren, D.; Sinninghe Damsté, J. S. 2013. "Linking Isoprenoidal GDGT Membrane Lipid Distributions with Gene Abundances of Ammonia-Oxidizing Thaumarchaeota and Uncultured Crenarchaeotal Groups in the Water Column of a Tropical Lake (Lake Challa, East Africa)." *Environmental Microbiology* 15(9): 2445–2462.

Buckley, S. A.; R. P. Evershed. 2001. "Organic Chemistry of Embalming Agents in Pharaonic and Graeco-Roman Mummies." *Nature* 413(6858): 837–841.

Bull, I. D., Simpson, I. A.; Van Bergen, P. F.; Evershed, R. P. 1999. "Muck 'n' Molecules: Organic Geochemical Methods for Detecting Ancient Manuring." *Antiquity* 73(279): 86–96.

Bull, I. D.; Lockheart, M. J.; Elhmmali, M. M.; Roberts, D. J.; Evershed, R. P. 2002. "The Origin of Faeces by Means of Biomarker Detection." *Environment International* 27(8): 647-654.

Burggraf, S.; Mayer, T.; Amann, R.; Schadhauser, S.; Woese, C. R.; Stetter, K. O. 1994. "Identifying Members of the Domain Archaea with rRNA-targeted Oligonucleotide Probes." *Applied and Environmental Microbiology* 60(9): 3112-3119.

Chen, Y. Z.; Kao, S. Y.; Jian, H. C.; Yu, Y. M.; Li, J. Y.; Wang, W. H.; Tsai, C. W. 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): 636–644.

Chen, Y.; Zhang, C.; Jia, C.; Zheng, F.; Zhu, C. 2016. "Tracking the Signals of Living Archaea: A Multiple Reaction Monitoring (MRM) Method for Detection of Trace Amounts of Intact Polar Lipids from the Natural Environment." *Organic Geochemistry* 97: 1-4.

Christian, G. D.; Dasgupta, P. K.; Schug, K. A. 2014. "*Chapter 3 statistics and data handling in analytical chemistry*." In: Gayle, A. and Bull, K. (eds) *Analytical chemistry*, 7th edition.

Wiley, Hoboken: 105–106.

Clavel, T.; Gomes-Neto, J. C.; Lagkouvardos, I.; Ramer-Tait, A.E. 2017. "Deciphering Interactions between the Gut Microbiota and the Immune System via Microbial Cultivation and Minimal Microbiomes." *Immunological Reviews* 279(1): 8–22.

Cooks, R. G.; Glish, G. L.; Mcluckey, S. A.; Kaiser, R. E. 1991. "Ion Trap Mass Spectrometry." *Chemical Enginerring News* 69(12): 26-41.

Corr, L. T.; Richards, M. P.; Jim, S.; Ambrose, S. H.; Mackie, A.; Beattie, O.; Evershed, R. P. 2008. "Probing Dietary Change of the Kwäday Dän Ts'ìnchi Individual, an Ancient Glacier Body from British Columbia: I. Complementary Use of Marine Lipid Biomarker and Carbon Isotope Signatures as Novel Indicators of a Marine Diet." *Journal of Archaeological Science* 35(8): 2102–2110.

Corr, L. T.; Sealy, J. C.; Horton, M. C. Evershed, R. P. 2005. "A Novel Marine Dietary Indicator Utilising Compound-Specific Bone Collagen Amino Acid Δ13C Values of Ancient Humans." *Journal of Archaeological Science* 32(3): 321–330.

Cuevas-Tena, M.; Alegría, A.; Lagarda, M. J. 2017. "Determination of Fecal Sterols Following a Diet with and without Plant Sterols." *Lipids* 52(10): 871–884.

Davis, H. R.; Zhu, L.; Hoos, L. M.; Tetzloff, G.; Maguire, M.; Liu, J.; Yao, X. Iyer, S. P. N.; Lam, M. Lund, E. G.; Detmers, P. A.; Graziano, M. P.; Altmann, S. W. 2004. "Niemann-Pick C1 like 1 (NPC1L1) Is the Intestinal Phytosterol and Cholesterol Transporter and a Key Modulator of Whole-Body Cholesterol Homeostasis." *Journal of Biological Chemistry* 279(32): 33586–33592.

Derrien, M.; Jardé, E.; Gruau, G.; Pierson-Wickmann, A. C. 2011. "Extreme Variability of Steroid Profiles in Cow Feces and Pig Slurries at the Regional Scale: Implications for the Use of Steroids to Specify Fecal Pollution Sources in Waters." *Journal of Agricultural and*

Food Chemistry 59(13): 7294–7302.

Jintang, D.; Li, Y.; Zhu, Z.; Chen, Y.; Zhao, Y. 2003. "Rearrangement Mechanism of the Sodium Adducts of Fmoc Protected Amino Acids." *Chinese Science Bulletin* 48(21): 2317–2319.

Eissle, S.; Kley, M.; Bächle, D.; Loidl, G.; Meier, T.; Samson D.; 2017. "Substitution Determination of Fmoc-Substituted Resins at Different Wavelengths." *Journal of Peptide Science* 23(10): 757–762.

Embley, T. M.; Finlay, B. J.; Thomas, R. H.; Dyal, P. L. 1992. "The Use of rRNA Sequences and Fluorescent Probes to Investigate the Phylogenetic Positions of the Anaerobic Ciliate *Metopus palaeformis* and its Archaeobacterial Endosymbiont." *Journal of General Microbiology* 138(7): 1479-1487.

Eme, L.; Spang, A.; Lombard, J.; Stairs, C.W.; Ettema, T. J. G. 2017. "Archaea and the Origin of Eukaryotes." *Nature Reviews Microbiology* 15(12): 711–723.

Erez, J. and Luz, B. 1983. "Experimental Paleotemperature Equation for Planktonic Foraminifera." *Geochimica et Cosmochimica Acta* 47(6): 1025-1031.

Escala, M.; Fietz, S.; Rueda, G.; Rosell-Melé, A. 2009. "Analytical Considerations for the Use of the Paleothermometer Tetraether Index86 and the Branched vs Isoprenoid Tetraether Index Regarding the Choice of Cleanup and Instrumental Conditions." *Analytical Chemistry* 81(7): 2701–2707.

Evershed, R. P. 2008. "Organic Residue Analysis in Archaeology: The Archaeological Biomarker Revolution." *Archaeometry* 50(6): 895–924.

Evershed, R. P. and Connolly, R. C. 1988. "Lipid Preservation in Lindow Man." *Naturwissenschaften* 75(3): 143–145.

Evershed, R P; Bethell, P. H.; Reynolds, P. J.; Walsh, N. J. 1997. "Manuring : Analysis of

Modern Experimental Material And." *Journal of Archaeological Science*: 485–495.

Eyssen, H. J.; Parmentier, G. G.; Compernolle, F. C.; De Pauw, G.; Piessens-Denef, M. 1973. "Biohydrogenation of Sterols by Eubacterium ATCC21,408-Nova Species." *European Journal of Biochemistry* 36: 411–421.

Faridi, K. F.; Lupton, J. R.; Martin, S. S.; Banach, M.; Quispe, R.; Kulkarni, K.; Jones, S. R.; Michos, E. D. 2017. "Vitamin D deficiency and non-lipid biomarkers of cardiovascular risk." *Archives of Medical Science* 13(4): 732-737.

Fibigr, J.; Šatínský, D.; Solich, P. 2017. "A UHPLC Method for the Rapid Separation and Quantification of Phytosterols Using Tandem UV/Charged Aerosol Detection – A Comparison of Both Detection Techniques." *Journal of Pharmaceutical and Biomedical Analysis* 140: 274–280.

Florini, S.; Shahsavari, E.; Aburto-Medina, A.; Khudur, L. S.; Mudge, S. M.; Smith, D. J.; Ball, A. S. 2020. "Are Sterols Useful for the Identification of Sources of Faecal Contamination in Shellfish? A Case Study." *Water (Switzerland)* 12(11): 1–14.

Freier, T. A.; Beitz, D. C.; Li, L.; Hartman, P. A. 1994. "Characterization of Eubacterium coprostanoligenes sp. nov., a Cholesterol-Reducing Anaerobe." *International Journal of Systematic Bacteriology* 44(1): 137-142.

Garcia-Calvo, M.; Lisnock, J. M.; Bull, H. G.; Hawes, B. E.; Burnett, D. A.; Braun, M. P.; Crona, J. H.; Davis, H. R.; Dean, D. C.; Detmers, P. A. *et al.* .2005. "The Target of Ezetimibe Is Niemann-Pick C1-Like 1 (NPC1L1)." *Proceedings of the National Academy of Sciences of the United States of America* 102(23): 8132–8137.

Glish, G. L. and Vacher, R. W. 2003. "The Basic of Mass Spectrometry in the Twenty-first Century." *Nature review*s 2: 140-150.

Griffiths, W. J.; Abdel-Khalik, J.; Yutuc, E.; Morgan, A. H.; Gilmore, I.; Hearn, T.; Wang, Y.

2017. “Cholesterolomics: An Update.” *Analytical Biochemistry* 524: 56-67.

Grimalt, J. O.; Fernández, P.; Bayona, J. M.; Albaigés. J. 1990. “Assessment of Fecal Sterols and Ketones as Indicators of Urban Sewage Inputs to Coastal Waters.” *Environmental Science and Technology* 24(3): 357–363.Guillemot, T.; Jacob, J.; Le Milbeau, C. 2017. “Fecal Biomarker Imprints as Indicators of Past Human Land Uses : Source Distinction and Preservation Potential in Archaeological and Natural Archives.” 81: 79–89.

Gülacara, F. O.; Susini, A.; Klohn, M. 1990. “Preservation and Post-Mortem Transformations of Lipids in Samples from a 4000-Year-Old Nubian Mummy.” *Journal of Archaeological Science* 17(6): 691–705.

Hansen, M.; Krogh, K. A.; Halling-Sørensen, B.; Björklund, E. 2011. “Determination of Ten Steroid Hormones in Animal Waste Manure and Agricultural Soil using Inverse and Integrated Clean-up Pressurized Liquid Extraction and Gas Chromatography-tandem Mass Spectrometry.” *Analytical Methods* 5(3): 1087-1095.

Harrault, L.; Milek, K.; Jardé, E.; Jeanneau, L.; Derrien, M.; Anderson, D. 2019. “Faecal Biomarkers Can Distinguish Specific Mammalian Species in Modern and Past Environments.” *PLoS ONE* 14(2): 1–26.

Higashi, T.; Shimada, K. 2004. “Derivatization of Neutral Steroids to Enhance Their Detection Characteristics in Liquid Chromatography-Mass Spectrometry.” *Analytical and Bioanalytical Chemistry* 378(4): 875–882.

Hollis, C. J.; Dunkley Jones, T.; Anagnostou, E.; Bijl, P. K.; Cramwinckel, M. J.; Cui, Y. 2019. “The DeepMIP Contribution to PMIP4: Methodologies for Selection, Compilation and Analysis of Latest Paleocene and Early Eocene Climate Proxy Data, Incorporating Version 0.1 of the DeepMIP Database.” *Geoscientific Model Development* 12(7): 3149-3206.

Hopmans, E. E.; Schouten, S.; Pancost, R. D.; van der Meer, M. T.; Sinninghe Damsté, J. S. 2000.

"Analysis of Intact Tetraether Lipids in Archaeal Cell Material and Sediments by High Performance Liquid Chromatography/Atmospheric Pressure Chemical Ionization Mass Spectrometry." *Rapid Commun Mass Spectrom* 14(7): 585–589.

Hopmans, E. C.; Weijers, J. W. H.; Schefuß, E.; Herfort, L.; Sinninghe Damsté, J. S.; Schouten, S. 2004. "A Novel Proxy for Terrestrial Organic Matter in Sediments Based on Branched and Isoprenoid Tetraether Lipids." *Earth and Planetary Science Letters* 224(1–2): 107–116.

Hopmans, E. C.; Schouten, S.; Sinninghe Damsté, J. S. 2016. "The Effect of Improved Chromatography on GDGT-Based Palaeoproxies." *Organic Geochemistry* 93: 1–6.

Hoshino, T. and Inagaki, F. 2019. "Abundance and Distribution of Archaea in the Subseafloor Sedimentary Biosphere." *ISME Journal* 13(1): 227–231.

Hugenholtz, P.; Goebel, B. M.; Pace, N. R. 1998. "Impact of Culture-independent Studies on the Emerging Phylogenetic View of Bacterial Diversity." *Journal of Bacteriology* 180(18): 4765-4774.

Huguet, C.; Hopmans, E. C.; Febo-Ayala, W.; Thompson, D. H.; Sinninghe Damsté, J. S.; Schouten, S. 2006. "An Improved Method to Determine the Absolute Abundance of Glycerol Dibiphytanyl Glycerol Tetraether Lipids." *Organic Geochemistry* 37(9): 1036-1041.

Isobe, K. O.; Tarao, M.; Zakaria, M. P.; Chiem, H. N.; Minh, L. Y.; Takada, H. 2002. "Quantitative Application of Fecal Sterols Using Gas Chromatography - Mass Spectrometry to Investigate Fecal Pollution in Tropical Waters: Western Malaysia and Mekong Delta, Vietnam." *Environmental Science and Technology* 36(21): 4497–4507.

Ito, M.; Ishimaru, M.; Shibata, T.; Hatate H.; Tanaka, R. 2017. "High-Performance Liquid Chromatography with Fluorescence Detection for Simultaneous Analysis of Phytosterols (Stigmasterol, β-Sitosterol, Campesterol, Ergosterol, and Fucosterol) and Cholesterol in

Plant Foods." *Food Analytical Methods* 10(8): 2692–2699.

McNaught, A. D. and Wilkinson, A. 1997. "IUPAC. Compendium of Chemical Terminology." 2nd ed. (the "Gold Book"). *Blackwell Scientific Publications*, Oxford. Online version (2019) created by Chalk, S. J.. ISBN 0-9678550-9-8. https://doi.org/10.1351/goldbook.

Ito, M.; Koba, K.; Hikihara, R.; Ishimaru, M.; Shibata, T.; Hatate, H.; Tanaka, R. 2018. "Analysis of Functional Components and Radical Scavenging Activity of 21 Algae Species Collected from the Japanese Coast." *Food Chemistry* 255: 147-156.

Jain, S.; Caforio, A.; Driessen, A. J. M. 2014. "Biosynthesis of Archaeal Membrane Ether Lipids." *Frontiers in Microbiology* 5(NOV): 1–16.

Jarrell, K. F.; Walters, A. D.; Bochiwal, C.; Borgia, J. M.; Dickinson, T.; Chong, J. P. J. 2011. "Major Players on the Microbial Stage: Why Archaea Are Important." *Microbiology* 157(4): 919–936.

Jauković, Z. D.; Svetlana, D. G.; Ivana, V.; Matić, B.; Laušević, M. D. 2017. "Determination of Sterols and Steroid Hormones in Surface Water and Wastewater Using Liquid Chromatography-Atmospheric Pressure Chemical Ionization-Mass Spectrometry." *Microchemical Journal* 135: 39–47.

Jayasinghe. L.Y.; Marriott, P.J.; Carpenter, P.D.; Nichols, P.D. 1998. "Application of Pentafluorophenyldimethylsilyl Derivatization for Gas Chromatography–electron-Capture Detection of Supercritically Extracted Sterols." *Journal of Chromatography A* 809: 109–120.

Kirk, J. M. 2007. "Mass Spectrometric Analysis of Steroid Hormones for Application in Analysis of Bovine Urine." University of York.

Koga, Y. and Morii, H. 2005. "Recent Advances in Structural Research on Ether Lipids from Archaea Including Comparative and Physiological Aspects." *Biosci. Biotechnol. Biochem*

69(11): 2019-2034.

Jeng, W. L.; Huh, C. A.; Chen, C. L. 1997. “Alkanol and Sterol Degradation in a Sediment core from the Continental Slope Off Southwestern Taiwan.” *Chemosphere* 35(11): 2515-2523.

Jensen, S. M.; Neesgaard, V. L.; Skjoldbjerg, S. L. N.; Brandl, M.; Ejsing, C. S.; Treusch, A. H. 2015. “The Effects of Temperature and Growth Phase on the Lipidomes of Sulfolobus Islandicus and Sulfolobus Tokodaii.” *Life* 5(3): 1539–1566.

Jones, W. J.; Leigh, J. A.; Mayer, F.; Woese, C. R.; Wolfe, R. S. 1983. “*Methanococcus Jannaschii* sp. nov., an Extremely Thermophilic Methanogen from a Submarine Hydrothermal Vent.” *Archives of Microbiology* 136(4): 254-261.

De Jonge, C.; Hopmans, E. C.; Stadnitskaia, A.; Rijpstra, W.; Irene C.; Hofland, R.; Tegelaar, E.; Sinninghe Damsté, J. S. 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* 54: 78–82.

Kaneko, M.; Kitajima, F.; Naraoka, H. 2011. “Stable Hydrogen Isotope Measurement of Archaeal Ether-bound Hydrocarbons.” *Organic Geochemistry* 42(2): 166-172.

Karner, M. B.; Delong, E. F.; Karl, D. M. 2001. “Archaeal Dominance in the Mesopelagic Zone of the Pacific Ocean.” *Nature* 409(6819): 507–510.

Kates, M. 1993. “Biology of Halophilic Bacteria, Part II - Membrane Lipids of Extreme Halophiles: Biosynthesis, Function and Evolutionary Significance.” *Experientia* 49(12): 1027–1036.

Keenan, K.; Imfeld, A.; Johnston, K.; Breckenridge, A.; Gélinas, Y.; Douglas, P. M. J. 2021. “Molecular Evidence for Human Population Change Associated with Climate Events in the Maya Lowlands.” *Quaternary Science Reviews* 258: 106904.

Kenny, D. J.; Plichta, D. R.; Shungin, D.; Koppel, N.; Hall, A. B.; Fu, B.; Vasan, R. S.; Shaw, S. Y.; Vlamakis, H.; Balskus, E. P.; Xavier, R. J. 2020. "Cholesterol Metabolism by Uncultured Human Gut Bacteria Influences Host Cholesterol Level." *Cell Host and Microbe* 28(2): 245-257.

Kim, J.-H.; Schouten, S.; Hopmans, E. C.; Donner, B.; Sinninghe Damsté, J. S. 2008. "Global Sediment Core-Top Calibration of the TEX_{86} Paleothermometer in the Ocean." *Geochimica et Cosmochimica Acta* 72(4): 1154–1173.

Kim, J. H.; van der Meer, J.; Schouten, S.; Helmke, P.; Willmott, V.; Sangiorgi, F.; Koç, N.; Hopmans, E. C.; Sinninghe Damsté, J. S. 2010. "New Indices and Calibrations Derived from the Distribution of Crenarchaeal Isoprenoid Tetraether Lipids: Implications for Past Sea Surface Temperature Reconstructions." *Geochimica et Cosmochimica Acta* 74(16): 4639-4654.

Kim, R. A.; Lee, K. E.; Bae, S. W. 2015. "Sea Surface Temperature Proxies (Alkenones, Foraminiferal Mg/Ca, and Planktonic Foraminiferal Assemblage) and Their Implications in the Okinawa Trough." *Progress in Earth and Planetary Science* 2(1): 1–16.

Knapp, D. R. 1979. "Handbook of Analytical Derivatization Reactions." John Wiley & Sons, New York.

Knappy, C. S.; Chong, J. P. J.; 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): 51-59.

Knappy, C. S.; Barillà, D.; de Blaquiere, J. P. A.; Morgan, H. W.; Nunn, C. E. M.; Suleman, M.; Tan, C. H. W.; Keely, B. J. 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* 165(6): 648–655.

Knappy, C.; Barillà, D.; Chong, J.; Hodgson, D.; Morgan, H.; Suleman, M.; Tan, C.; Yao, P.; Keely, B. 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): 1420–1432.

Kolakowski, B. M.; Grossert, J. S.; Ramaley, L. 2004. “Studies on the Positive-Ion Mass Spectra from Atmospheric Pressure Chemical Ionization of Gases and Solvents Used in Liquid Chromatography and Direct Liquid Injection.” *Journal of the American Society for Mass Spectrometry* 15(3): 311-324.

Kostiainen, R. and Kauppila, T. J. 2009. “Effect of Eluent on the Ionization Process in Liquid Chromatography–mass spectrometry.” *Journal of Chromatography A* 1216(4): 685-699.Kunz, S. and Matysik, S. 2019. “A Comprehensive Method to Determine Sterol Species in Human Faeces by GC-Triple Quadrupole MS.” *Journal of Steroid Biochemistry and Molecular Biology* 190(January): 99–103.

Kuypers, M. M.; Blokker, P.; Erbacher, J.; Kinkel, H.; Pancost, R. P.; Schouten, S.; Sinninghe Damste, J. S. 2001. “Massive Expansion of Marine Archaea during a Mid-Cretaceous Oceanic Anoxic Event.” *Science* 293(5527): 92–94.

Lammert, F. and Wang, D. Q. H. 2005. “New Insights Into the Genetic Regulation of Intestinal Cholesterol Absorption.” *Gastroenterology* 129(2): 718–734.

Law, K. P. and Zhang, C. L. 2019. “Current Progress and Future Trends in Mass Spectrometry-Based Archaeal Lipidomics.” *Organic Geochemistry* 134: 45–61.

Lenggera, S. K.; Sutton, P. A.; Rowland, S. J.; Hurley, S. J.; Pearson, A.; Naafs, B. D. A.; Dang, X.; Inglis, G. N.; Pancost, R. D. 2018. “Archaeal and Bacterial Glycerol Dialkyl Glycerol Tetraether (GDGT) Lipids in Environmental Samples by High Temperature-Gas Chromatography with Flame Ionisation and Time-of-Flight Mass Spectrometry Detection.” *Organic Geochemistry* 121: 10–21.

Lengger, S. K. Weber, Y.; Taylor, K. W. R.; Kopf, S. H.; Berstan, R.; Bull, I. D.; Mayser, J.-P.; Leavitt, W. D.; Blewett, J.; Pearson, A.; Pancost, R. D. 2021. "Determination of the δ 2 H Values of High Molecular Weight Lipids by High-temperature Gas Chromatography Coupled to Isotope Ratio Mass Spectrometry." *Rapid Communications in Mass Spectrometry* 35(4): 1-10.

Li, L.; Freier, T. A.; Hartman, P. A.; Young, J. W.; Beitz, D. C. 1995. "A Resting-cell Assay for Cholesterol Reductase Activity in Eubacterium Coprostanoligenes ATCC 51222." *Applied Microbiology and Biotechnology* 43(5): 887-892.

Li, L. H.; Dutkiewicz, E. P.; Huang, Y. C.; Zhou, H. B.; Hsu, C. C. 2019. "Analytical methods for Cholesterol Quantification." *Journal of Food and Drug Analysis* 27(2): 375-386.

Lin, D. S.; Connor, W. E.; Napton, L. K.; Heizer, R. F. 1978. "The Steroids of 2000-year-old Human Coprolites." *Journal of Lipid Research* 19(2): 215-221.

Lin, Y. T.; Wu, S. S.; Wu, H. L. 2007. "Highly Sensitive Analysis of Cholesterol and Sitosterol in Foods and Human Biosamples by Liquid Chromatography with Fluorescence Detection." *Journal of Chromatography A* 1156(1-2 SPEC. ISS.): 280–287.

Lipp, J. S. and Hinrichs, K. U. 2009. "Structural Diversity and Fate of Intact Polar Lipids in Marine Sediments." *Geochimica et Cosmochimica Acta* 73(22): 6816-6833.

Liu, S. and Ruan, H. 2013. "A Highly Sensitive Quantification of Phytosterols through an Inexpensive Derivatization." *Chemistry and Physics of Lipids* 166(1) 18-25.

Liu, X. L.; Birgel, D.; Elling, F. J.; Sutton, P. A.; Lipp, J. S.; Zhu, R.; Zhang, C.; Könneke, M.; Peckmann, J.; Rowland, S. J.; Summons, R. E.; Hinrichs, K. U. 2016. "From Ether to Acid: A Plausible Degradation Pathway of Glycerol Dialkyl Glycerol Tetraethers." *Geochimica et Cosmochimica Acta* 183: 138-152.

Liu, X. L.; Lipp, J. S.; Birgel, D.; Summons, R. E.; Hinrichs, K. U. 2018. "Predominance of

Parallel Glycerol Arrangement in Archaeal Tetraethers from Marine Sediments: Structural Features Revealed from Degradation Products." *Organic Geochemistry* 115: 12-23.

Liu, X. L.; Russell, D. A.; Bonfio, C.; Summons, R. E. 2019. "Glycerol Configurations of Environmental GDGTs Investigated Using a Selective Sn2 Ether Cleavage Protocol." *Organic Geochemistry* 128: 57–62.

Liu, Z.; Pagani, M.; Zinniker, D.; DeConto, R.; Huber, M.; Brinkhuis, H.; Shah, S. R.; Leckie, M.; Pearson A. 2009. "Global Cooling During the Eocene-Oligocene Climate Transition." *Science (New York, N.Y.)* 323(February): 1187–1190.

Lloyd, C. E. M.; Michaelides, K.; Chadwick, D. R.; Dungait, J. A. J.; Evershed, R. P. 2012. "Tracing the Flow-driven Vertical Transport of Livestock-derived Organic Matter Through Soil using Biomarkers." *Organic Geochemistry* 43: 56-66.

Lobasso, S.; Lopalco, P.; Angelini, R.; Vitale, R.; Huber, H.; Müller, V.; Corcelli, A. 2012. "Coupled TLC and MALDI-TOF/MS Analyses of the Lipid Extract of the Hyperthermophilic Archaeon *Pyrococcus furiosus*." Archaea 2012(957852): 1-10.

Loo, J. A. 2000. "Electrospray Ionization Mass Spectrometry: a Technology for Studying Noncovalent Macromolecular Complexes." *International Journal of Mass Spectrometry* 200(1): 175-186.

von der Lühe, B.; Dawson, L. A.; Mayes, R. W.; Forbes, S. L.; Fiedler, S. 2013. "Investigation of Sterols as Potential Biomarkers for the Detection of Pig (*S. Domesticus*) Decomposition Fluid in Soils." *Forensic Science International* 230(1–3): 68–73.

von der Lühe, B.; Birk, J. J.; Dawson, L.; Mayes, R. W.; Fiedler, S. 2018. "Steroid Fingerprints: Efficient Biomarkers of Human Decomposition Fluids in Soil." *Organic Geochemistry* 124: 228–237.

von der Lühe, B.; Prost, K.; Birk, J. J.; Fiedler, S. 2020. "Steroids Aid in Human Decomposition

Fluid Identification in Soils of Temporary Mass Graves from World War II." *Journal of Archaeological Science: Reports* 32(July): 102431.

Maekawa, M. and Fairn, G. D. 2015. "Complementary probes reveal that phosphatidylserine is required for the proper transbilayer distribution of cholesterol." *Journal of Cell Science* 128(7): 1422-1433.

Macalady, J. L.; Vestling, M. M.; Baumler, D.; Boekelheide, N.; Kaspar, C. W.; Banfield, J. F. 2004. "Tetraether-linked Membrane Monolayers in *Ferroplasma* spp: A Key to Survival in Acid." *Extremophiles* 8(5): 411-419.

March, R. E. 1997. "An Introduction to Quadrupole Ion Trap Mass Spectrometry." *Journal of Mass Spectrometry* 32: 351-369.

Matić Bujagić, I.; Grujić, S.; Jauković, Z.; Laušević, M. 2016. "Sterol Ratios as a Tool for Sewage Pollution Assessment of River Sediments in Serbia." *Environmental Pollution* 213: 76–83.

Matić, I.; Grujić, S.; Jauković, Z.; Laušević, M. 2014. "Trace Analysis of Selected Hormones and Sterols in River Sediments by Liquid Chromatography-Atmospheric Pressure Chemical Ionization – Tandem Mass Spectrometry Sevi C." 1364: 117–127.

Mayer, B. X.; Reiter, C.; Bereuter, T. L. 1997. "Investigation of the Triacylglycerol Composition of Iceman's Mummified Tissue by High-Temperature Gas Chromatography." *Journal of Chromatography B: Biomedical Applications* 692(1): 1–6.

McNaught, A. D. and Wilkinson A. 1997. "IUPAC. Compendium of Chemical Terminology (the "Gold Book")." 2nd ed., *Blackwell Scientific Publications*, Oxford.

McWethy, D. B.; Alt, M.; Argiriadis, E.; Battistel, D.; Everett, R.; Pederson, G. T. 2020. "Millennial-Scale Climate and Human Drivers of Environmental Change and Fire Activity in a Dry, Mixed-Conifer Forest of Northwestern Montana." *Frontiers in Forests and Global Change* 3(April): 1–16.

de Melo, M. G.; da Silva, B. A.; Costa, G. S.; da Silva Neto, J. C. A.; Soares, P. K.; Val, A. L.; Chaar, Ja. S.; Koolen, H. H. F. Bataglion, G. A. 2019. "Sewage Contamination of Amazon Streams Crossing Manaus (Brazil) by Sterol Biomarkers." *Environmental Pollution* 244: 818–826.

Michael Moldowan, J. and Seifert, W. K. 1979. "Head-to-head Linked Isoprenoid Hydrocarbons in Petroleum." *Science* 204(4389): 169-171.

Miller, J. N.; Miller, J. C. 2010. "Calibration Methods: Regression and Correlation." In *Statistics and Chemometrics for Analytical Chemistry*, 6th edition, Pearson, Prentice Hall: 124–126.

Mischak, H.; Coon, J. J.; Novak, J.; Weissinger, E. M.; Schanstra, J.; Dominiczak, A. F. 2009. "Capillary Electrophoresis–mass Spectrometry as a Powerful Tool in Biomarker Discovery and Clinical Diagnosis: an Update of Recent Developments." *Mass Spectrometry Reviews* 28(5): 703-724.

Moliner-Martinez, Y.; Herráez-Hernández, R.; Molins-Legua, C.; Campins-Falcó, P. 2010. "Improving Analysis of Apolar Organic Compounds by the Use of a Capillary Titania-based Column: Application to the Direct Determination of Faecal Sterols Cholesterol and Coprostanol in Wastewater Samples." *Journal of Chromatography A* 1217(28): 4682-4687.

Nakagawa, K.; Amano, H.; Berndtsson, R.; Takao, Y.; Hosono, T. 2019. "Use of Sterols to Monitor Surface Water Quality Change and Nitrate Pollution Source." *Ecological Indicators* 107(June): 105534.

Nakakuni, M.; Yamasaki, Y.; Yoshitake, N.; Takehara, K.; Yamamoto, S. 2019. "Methyl Ether-Derivatized Sterols and Coprostanol Produced *via* Thermochemolysis Using Tetramethylammonium Hydroxide (TMAH)." *Molecules* 24(4040):1-19.

Nelson, T. J. 2011. "Fluorescent High-Performance Liquid Chromatography Assay for Lipophilic Alcohols." *Analytical Biochemistry* 419(1): 40–45.

Nie, J., L.; Zhang, G. Z.; Du, X. 2019. “Quercetin Reduces Atherosclerotic Lesions by Altering the Gut Microbiota and Reducing Atherogenic Lipid Metabolites.” *Journal of Applied Microbiology* 127(6): 1824–1834.

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 Solvent.” *Journal of Biochemistry* 101(4): 997-1005.

Norum, K. R.; Berg, T.; Helgerud, P.; Drevon, C. A. 1983. “Transport of Cholesterol.” *Physiological Reviews* 63(4): 1343–1419.

Nzekoue, F. K.; Khamitova, G.; Angeloni, S.; Sempere, A. N.; Tao, J.; Maggi, F.; Xiao, J.; Sagratini, G.; Vittori, S.; Caprioli, G. 2020. “Spent Coffee Grounds: A Potential Commercial Source of Phytosterols.” *Food Chemistry* 325(April) 126836.

Nzekoue, F. K.; Caprioli, G.; Ricciutelli, M.; Cortese, M.; Alesi, A.; Vittori, S.; Sagratini, G. 2020. “Development of an Innovative Phytosterol Derivatization Method to Improve the HPLC-DAD Analysis and the ESI-MS Detection of Plant Sterols/Stanols.” *Food Research International* 131(January): 108998.

O'Brien, C. L.; Robinson, S. A.; Pancost, R. D.; Sinninghe Damsté, J. S.; Schouten, S. *et al.* 2017. “Cretaceous Sea-Surface Temperature Evolution: Constraints from TEX_{86} and Planktonic Foraminiferal Oxygen Isotopes.” *Earth-Science Reviews* 172(July): 224–247.

O'Connor, S.; Ali, E.; Al-Sabah, S.; Anwar, D.; Bergström, Ed *et al.* 2011. “Exceptional Preservation of a Prehistoric Human Brain from Heslington, Yorkshire, UK.” *Journal of Archaeological Science* 38(7): 1641–1654.

Ohtsubo, S.; Kanno, M.; Miyahara, H.; Kohno, S.; Koga, Y.; Miura, I. 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*

Microbiology Ecology 12(1): 39-50.

Page, J. S.; Kelly, R. T.; Tang, K.; Smith, R. D. 2007. "Ionization and Transmission Efficiency in an Electrospray Ionization–Mass Spectrometry Interface." *Journal of American Society for Mass Spectrometry* 18: 1582-1590.

Pearson, A.; Pi, Y.; Zhao, W.; Li, W. J.; Li, Y.; Inskeep, W.; Perevalova, A.; Romanek, C.; Li, S.; Zhang, C. L. 2008. "Factors Controlling the Distribution of Archaeal Tetraethers in Terrestrial Hot Springs." *Applied and Environmental Microbiology* 74(11): 3523–3532.

Pearson, A.; Ingalls, A. E. 2013. "Assessing the Use of Archaeal Lipids as Marine Environmental Proxies." *Annual Review of Earth and Planetary Sciences* 41(1): 359–384.

Perry, R. H.; Cooks, R. G.; Noll, R. J. 2008. "Orbitrap Mass Spectrometry: Instrumentation, Ion Motion and Applications." *Mass spectrometry reviews* 27(6); 661-699.

Peterse, F.; Moy, C. M.; Eglinton, T. I. 2015. "A Laboratory Experiment on the Behaviour of Soil-Derived Core and Intact Polar GDGTs in Aquatic Environments." *Biogeosciences* 12(4): 933–943.

Peterse, F.; Kim, J. H.; Schouten, S.; Kristensen, D. K.; Koç, N.; Sinninghe Damsté, J. S. 2009. "Constraints on the Application of the MBT/CBT Palaeothermometer at High Latitude Environments (Svalbard, Norway)." *Organic Geochemistry* 40(6): 692–699.

Pickering, M. D.; Ghislandi, S.; Usai, M. R.; Wilson, C.; Connelly, P.; Brothwell, D. R.; Keely, B. J. 2018. "Signatures of Degraded Body Tissues and Environmental Conditions in Grave Soils from a Roman and an Anglo-Scandinavian Age Burial from Hungate, York." *Journal of Archaeological Science* 99(September 2016): 87–98.

Pitcher, A.; Hopmans, E. C.; Schouten, S.; Sinninghe Damsté, J. S. 2009. "Separation of Core and Intact Polar Archaeal Tetraether Lipids using Silica Columns: Insights into Living and Fossil Biomass Contributions." *Organic Geochemistry* 40(1): 12-19.

Pitcher, A.; Villanueva, L.; Hopmans, E. C.; Schouten, S.; Reichart, G. J.; Sinninghe Damsté, J. S. 2011. “Niche Segregation of Ammonia-Oxidizing Archaea and Anammox Bacteria in the Arabian Sea Oxygen Minimum Zone.” *ISME Journal* 5(12): 1896–1904.

Poplawski, Cezary. 2017. “Development of a High-Throughput Technique for Screening Archaeal Tetraether Lipid Cores and Other Alcohols in Sediments and Microbial Cultures.” University of York.

Powers, L.; Werne, J. P.; Vanderwoude, A. J.; Sinninghe Damsté, J. S.; Hopmans, E. C.; Schouten, S. 2010. “Applicability and Calibration of the TEX_{86} Paleothermometer in Lakes.” *Organic Geochemistry* 41(4): 404–413.

Pratt, C.; Warnken, J.; Leeming, R.; Arthur, M. J.; Grice, D. I. 2008. “Degradation and Responses of Coprostanol and Selected Sterol Biomarkers in Sediments to a Simulated Major Sewage Pollution Event: A Microcosm Experiment under Sub-Tropical Estuarine Conditions.” *Organic Geochemistry* 39(3): 353–369.

Prost, K.; Birk, J. J.; Lehndorff, E.; Gerlach, R.; Amelung, W. 2017. “Steroid Biomarkers Revisited - Improved Source Identification of Faecal Remains in Archaeological Soil Material.” *PLoS ONE* 12(1): 1–30.

Prost, K.; Bradel, P. L.; Lehndorff, E.; Amelung, W. 2018. “Steroid Dissipation and Formation in the Course of Farmyard Manure Composting.” *Organic Geochemistry* 118: 47–57.

Quehenberger, J.; Pittenauer, E.; Allmaier, G.; Spadiut, O. 2020. “The Influence of the Specific Growth Rate on the Lipid Composition of Sulfolobus Acidocaldarius.” *Extremophiles* 24(3): 413–420.

Ramasha, C. S.; Pickett, W. C.; Murthy, D. V. K.; 1989. “Sensitive Method for Analysis of Phospholipid Subclasses and Molecular Species as 1-anthoryl Derivative of their Diglycerides.” *Journal of Chromatography* 491: 37-48.

Rattray, J. E., and Smittenberg, R. H. 2020. “Separation of Branched and Isoprenoid Glycerol Dialkyl Glycerol Tetraether (GDGT) Isomers in Peat Soils and Marine Sediments Using Reverse Phase Chromatography.” *Frontiers in Marine Science* 7(September) 1-17.

Resende, M. F.; Santos, M. D. R.; Matos, R. C.; Matos, M. A. C. 2014. “The Analysis of Faecal Sterols in Sediment Samples by HPLC-UV using Ultrasound-assisted Treatment.” *Analytical Methods* 6(24): 9581-9587.

Robinson, S. A.; Ruhl, M.; Astley, D. L.; Naafs, B. D. A.; Farnsworth, A. J.; Bown, P. R.; Jenkyns, H. C.; Lunt, D. J.; O'Brien, C.; Pancost, R. D.; Markwick, P. J. 2017. “Mesozoic Climates and Oceans – a Tribute to Hugh Jenkyns and Helmut Weissert.” *Sedimentology* 64(1): 1-15.

Robinson, S. A.; Ruhl, M.; Astley, D. L.; Naafs, B. D. A.; Farnsworth, A. J.; Bown, P. R.; Jenkyns, H. C.; Lunt, D. J.; O'Brien, C.; Pancost, R. D.; Markwick, P. J. 2017. “Early Jurassic North Atlantic Sea-Surface Temperatures from TEX_{86} Palaeothermometry.” *Sedimentology* 64(1): 215–230.De Rosa, M.; Gambacorta, A.; Nicolaus, B.; Sodano, S.; Bu'Lock, J. D. 1980. “Structural Regularities in Tetraether Lipids of Caldariella and their Biosynthetic and Phyletic Implications.” *Phytochemistry* 19(5): 833-836.

De Rosa, M. and Gambacorta, A. 1988. “The Lipids of Archaebacteria.” *Progress in Lipid Research* 27(3): 153-175.

Sadekov, A.; Stephen M. E.; De Deckker, P.; Kroon, D. 2008. “Uncertainties in Seawater Thermometry Deriving from Intratest and Intertest Mg/Ca Variability in Globigerinoides Ruber.” *Paleoceanography* 23(1): 1–12.

Sadzikowski, M. R.; Sperry, J. F.; Wilkins, T. D. 1977. “Cholesterol-Reducing Bacterium from Human Feces.” *Applied and Environmental Microbiology* 34(4): 355–362.

Schneider, R. 2001. “Alkenone Temperature and Carbon Isotope Records: Temporal Resolution,

Offsets, and Regionality." *Geochemistry, Geophysics, Geosystems* 2(1): 2000GC000060.

Schött, H. F.; Krautbauer, S.; Höring, M.; Liebisch, G. Matysik, S. 2018. "A Validated, Fast Method for Quantification of Sterols and Gut Microbiome Derived 5α/β-Stanols in Human Feces by Isotope Dilution LC-High-Resolution MS." *Analytical Chemistry* 90(14): 8487–8494.

Schouten, S.; Hoefs, M. J. L.; Koopmans, M. P.; Bosch, H. J.; Sinninghe Damsté, J. S. 1998. "Structural Characterization, Occurrence and Fate of Archaeal Ether-bound Acyclic and Cyclic Biphytanes and Corresponding Diols in Sediments." *Organic Geochemistry* 29(5-7): 1305-1319.

Schouten, S.; Wakeham, S. G.; Hopmans, E. C.; Sinninghe Damsté, J. S. 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): 1680-1686.

Schouten, S.; Forster, A.; Panoto, F. E.; Sinninghe Damsté, J. S. 2007. "Towards Calibration of the TEX_{86} Palaeothermometer for Tropical Sea Surface Temperatures in Ancient Greenhouse Worlds." *Organic Geochemistry* 38: 1537–1546.

Schouten, S.; Huguet, C.; Hopmans, E. C.; Kienhuis, M. V. M.; Sinninghe Damsté, J. S.2007. "Analytical Methodology for TEX_{86} Paleothermometry by High-Performance Liquid Chromatography/Atmospheric Pressure Chemical Ionization-Mass Spectrometry." *Analytical Chemistry* 79(7): 2940–2944.

Schouten, S.; Van Der Meer, M. T.J.; Hopmans, E. C.; Rijpstra, W.; Irene C.; Reysenbach, A. L.; Ward, D. M.; Sinninghe Damsté, J. S. 2007. "Archaeal and Bacterial Glycerol Dialkyl Glycerol Tetraether Lipids in Hot Springs of Yellowstone National Park." *Applied and Environmental Microbiology* 73(19): 6181–6191.

Schouten, S.; Hopmans, E. C.; Van Der Meer, J.; Mets, A.; Bard, E.; Bianchi, T. S.; Diefendorf, A.; Escala, M.; Freeman, K. H. *et al.* 2009. "An Interlaboratory Study of TEX_{86} and BIT Analysis Using High-Performance Liquid Chromatography-Mass Spectrometry." *Geochemistry, Geophysics, Geosystems* 10(3): 1–13.

Schouten, S.; Hopmans, E. C.; Rosell-Melé, A.; Pearson, A.; Adam, P.; Bauersachs, T.; Bard, E.; Bernasconi, S. M.; Bianchi, T. S., Brocks, J. J. Carlson, *et al.* 2013. "An Interlaboratory Study of TEX_{86} and BIT Analysis of Sediments, Extracts, and Standard Mixtures." *Geochemistry, Geophysics, Geosystems* 14(12): 5263–5285.

Schouten, S.; Hopmans, E. C.; Schefuß, E.; Sinninghe Damsté, J. S. 2002. "Distributional Variations in Marine Crenarchaeotal Membrane Lipids: A New Tool for Reconstructing Ancient Sea Water Temperatures?" *Earth and Planetary Science Letters* 204(1–2): 265–274.

Schouten, S., Hopmans, E. C.; Sinninghe Damsté, J. S. 2013. "The Organic Geochemistry of Glycerol Dialkyl Glycerol Tetraether Lipids: A Review." *Organic Geochemistry* 54: 19–61.

Schouten, S.; Van Der Meer, M. T. J.; Hopmans, E. C.; Rijpstra, W. I. C.; Reysenbach, A. L.: Ward, D. M.; Sinninghe Damsté, J. S. 2007. "Archaeal and Bacterial Glycerol Dialkyl Glycerol Tetraether Lipids in Hot Springs of Yellowstone National Park." *Applied and Environmental Microbiology* 73(19): 6181-6191.

Schroeter, N.; Lauterbach, S.; Stebich, M.; Kalanke, J.; Mingram, J.; Yildiz, C.; Schouten, S.; Gleixner, G. 2020. "Biomolecular Evidence of Early Human Occupation of a High-Altitude Site in Western Central Asia During the Holocene." *Frontiers in Earth Science* 8(February): 1–13.

Seemann, B.; Alon, T.; Tsizin, S.; Fialkov, A. B.; Amirav, A. 2015. "Electron Ionization LC-MS with Supersonic Molecular Beams - The New Concept, Benefits and Applications." *Journal of Mass Spectrometry* 50(11): 1252-1263.

Simpson, I. A.; Stephen J. D.; Bull, I. D.; Evershed, R. P. 1998. “Early Anthropogenic Soil Formation at Tofts Ness, Sanday, Orkney.” *Journal of Archaeological Science* 25(8): 729–746.

Sinninghe Damsté, J. S.; Schouten, S.; Hopmans, E. C.; Van Duin, A. C. T.; Geenevasen, J. A. J. 2002. “Crenarchaeol: The Characteristic Core Glycerol Dibiphytanyl Glycerol Tetraether Membrane Lipid of Cosmopolitan Pelagic Crenarchaeota.” *Journal of Lipid Research* 43(10): 1641-1651.

Sinninghe Damsté, J. S.; Rijpstra, W.; Irene C.; Hopmans, E. C.; den Uijl, M. J.; Weijers, J. W. H.; Schouten, S. 2018. “The Enigmatic Structure of the Crenarchaeol Isomer.” *Organic Geochemistry* 124: 22–28.

Sistiaga, A.; Berna, F.; Laursen, R.; Goldberg, P. 2014. “Steroidal Biomarker Analysis of a 14,000 Years Old Putative Human Coprolite from Paisley Cave, Oregon.” *Journal of Archaeological Science* 41: 813–817.

Sistiaga, A.; Mallol, C.; Galva´n, B.; Summons, R. E. 2014. “The Neanderthal Meal : A New Perspective Using Faecal Biomarkers.” 9(6): 6–11.

Sistiaga, A.; Wrangham, R.; Rothman, J. M.; Summons, R. E. 2015. “New Insights into the Evolution of the Human Diet from Faecal Biomarker Analysis in Wild Chimpanzee and Gorilla Faeces.” *PLoS ONE* 10(6): 15–17.

Sjövall, P.; Thiel, V.; Siljeström, S.; Heim, C.; Hode, T.; Lausmaa, J. 2008. “Organic geochemical microanalysis by time-of-flight secondary ion mass spectrometry (ToF-SIMS).” *Geostandards and Geoanalytical Research* 32(3): 267-277.

Skubic, C.; Vovk, I.; Rozman, D.; Križman, M. 2020. “Simplified LC-MS Method for Analysis of Sterols in Biological Samples.” *Molecules* 25(18): 1–10.

Snyder, L. R.; Kirkland, J. J. ; Glajch, J. L. 1997. *Practical HPLC Method Development*. 2nd ed.

John Wiley and Sons.

Song, Y.; Kenworthy, A. K.; Sanders, C. R. 2014. “Cholesterol as a co-solvent and a ligand for membrane proteins.” *Protein Science* 23(1): 1-22.

Speranza, E. D.; Colombo, M.; Heguilor, S.; Tatone, L. M.; Colombo, J. C. 2020. “Alterations in the Sterol Signature of Detritivorous Fish along Pollution Gradients in the Río de La Plata Basin (Argentina): From Plant to Sewage-Based Diet.” *Environmental Research* 184(March): 109351.

Spero, H. J.; Bijma, J.; Lea, D. W.; Bernis, B. E. 1997. “Effect of Seawater Carbonate Concentration on Foraminiferal Carbon and Oxygen Isotopes.” *Nature* 390(6659): 497–500.

Stefens, J. L.; Santos, J. H. Z. D.; Filho, J. G. M.; Da Silva, C. G. A.; Peralba, M. D. C. R. 2007. “Lipid Biomarkers Profile - Presence of Coprostanol: Recent Sediments from Rodrigo de Freitas Lagoon - Rio de Janeiro, Brazil.” *Journal of Environmental Science and Health - Part A Toxic/Hazardous Substances and Environmental Engineering* 42(11): 1553–1560.

Steinig, S.; Dummann, W.; Park, W.; Latif, M.; Kusch, S.; Hofmann, P.; Flögel, S. 2020. “Evidence for a Regional Warm Bias in the Early Cretaceous TEX_{86} Record.” *Earth and Planetary Science Letters* 539: 116184.

Stetter, K. O. 2006. “Hyperthermophiles in the History of Life.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 361(1474): 1837-1842.

Stolz, A.; Jooß, K.; Höcker, O.; Römer, J.; Schlecht, J.; Neusüß, C. 2019. “Recent Advances in Capillary Electrophoresis-mass Spectrometry: Instrumentation, Methodology and Applications.” *Electrophoresis* 40(1): 79-112.

Stott, A. W.; Evershed, R. P.; Jim, S.; Jones, V.; Rogers, J. M.; Tuross, N.; Ambrose, S. 1999. “Cholesterol as a New Source of Palaeodietary Information: Experimental Approaches and Archaeological Applications.” *Journal of Archaeological Science* 26(6): 705–716.

Sturt, H. F.; Summons, R. E., Smith, K.; Elvert, M.; Hinrichs, K. U. 2004. "Intact Polar Membrane Lipids in Prokaryotes and Sediments Deciphered by High-Performance Liquid Chromatography/Electrospray Ionization Multistage Mass Spectrometry - New Biomarkers for Biogeochemistry and Microbial Ecology." *Rapid Communications in Mass Spectrometry* 18(6): 617–28.

Sun, J.; Zhao, X. E.; Dang, J.; Sun, X.; Zheng, L.; You, J.; Wang, X. 2017. "Rapid and Sensitive Determination of Phytosterols in Functional Foods and Medicinal Herbs by Using UHPLC–MS/MS with Microwave-Assisted Derivatization Combined with Dual Ultrasound-Assisted Dispersive Liquid–Liquid Microextraction." *Journal of Separation Science* 40(3): 725–732.

Sutton, P. A. and Rowland, S. J. 2012. "High Temperature Gas Chromatography-Time-of-flight-Mass Spectrometry (HTGC-ToF-MS) for High-boiling Compounds." *Journal of Chromatography A* 1243: 69-80.

Thompson, D. H.; Kim, J. M.; Rananavare, S. B.; Wong, K. F.; Wheeler, J. J.; Humphry-Baker, R. 1992. "Tetraether Bolaform Amphiphiles as Models of Archaebacterial Membrane Lipids: Raman Spectroscopy, ^{31}P NMR, X-Ray Scattering, and Electron Microscopy." *Journal of the American Chemical Society* 114(23): 9035–9042.

Tierney, J. E. and Tingley, M. P. 2014. "A Bayesian, Spatially-Varying Calibration Model for the TEX_{86} Proxy." *Geochimica et Cosmochimica Acta* 127: 83-106.

Tort, L. F. L.; Iglesias, K.; Bueno, C.; Lizasoain, A.; Salvo, M.; Cristina, J.; Kandratavicius, N.; Pérez, L.; Figueira, R.; Bícego, M. C.; Taniguchi, S.; Venturini, N.; Brugnoli, E.; Colina, R.; Victoria, M. 2017. "Wastewater Contamination in Antarctic Melt-Water Streams Evidenced by Virological and Organic Molecular Markers." *Science of the Total Environment* 609: 225–231.

Trommer, G.; Siccha, M.; van der Meer, M. T. J.; Schouten, S.; Sinninghe Damsté, J. S.; Schulz, H.; Hemleben, C.; Kucera, M. 2009. "Distribution of Crenarchaeota Tetraether Membrane

Lipids in Surface Sediments from the Red Sea." *Organic Geochemistry* 40(6): 724-731.

Tsizin, S.; Bokka, R.; Keshet, U.; Alon, T.; Fialkov, A. B.; Tal, N.; Amirav, A. 2017. "Comparison of Electrospray LC–MS, LC–MS with Cold EI and GC–MS with Cold EI for Sample Identification." *International Journal of Mass Spectrometry* 422: 119-125.

Food and Drug administration (FDA) office of regulary affaris. 2020. "Guidance for Methods, Method Verification and Validation." ORA Laboratory Manual (ORA-LAB.5.4.5), 2nd edition. https://www.fda.gov/media/73920/download

Van de Vossenberg, J. L. C. M.; Driessen, A. J. M.; Konings, W. N. 1998. "The Essence of Being Extremophilic: The Role of the Unique Archaeal Membrane Lipids." *Extremophiles* 2(3): 163-170.

Veiga, P.; Juste, C.; Lepercq, P.; Saunier, K.; Béguet, F.; Gérard, P. 2005. "Correlation between Faecal Microbial Community Structure and Cholesterol-to-Coprostanol Conversion in the Human Gut." *FEMS Microbiology Letters* 242(1): 81–86.

Villanueva, L.; Sinninghe Damsté, J. S.; Schouten, S. 2014. "A Re-Evaluation of the Archaeal Membrane Lipid Biosynthetic Pathway." *Nature Reviews Microbiology* 12(6): 438–448.

Vors, C.; Joumard-Cubizolles, L.; Lecomte, M.; Combe, E.; Ouchchane, L.; Drai, J.; Raynal, K.; Joffre, F.; Meiller, L.; Le Barz, M.; Gaborit, P. *et al.* 2020. "Milk Polar Lipids Reduce Lipid Cardiovascular Risk Factors in Overweight Postmenopausal Women: Towards a Gut Sphingomyelin-Cholesterol Interplay." *Gut* 69(3): 487–501.

Walker, R. W; Wun, C. K.; Litsky, W.; Dutka, B.J. 1982. "Coprostanol as an Indicator of Fecal Pollution." *C R C Critical Reviews in Environmental Control* 12(2): 91–112. https://doi.org/10.1080/10643388209381695.

Wakehman, S.G. 1989. "Reduction of Stenols to Stanols in Particulate Matter Oxic-anoxic Boundaries in Sea Water." *Nature* 342: 787-790.

Wang, S.; Fan, J.; Xu, L.; Ye, K.; Shu, T.; Liu, S. 2019. “Enhancement of Antioxidant Activity in O/W Emulsion and Cholesterol-Reducing Capacity of Epigallocatechin by Derivatization with Representative Phytosterols.” *Journal of Agricultural and Food Chemistry* 67: 12461-12471.

Wang, Z.; Wang, Y.; Yu, T.; Hu, Z.; Wang, Y. 2021. “An LC–ESI/MS/MS Method for the Determination of Lupeol via Precolumn Derivatization and Its Application to Pharmacokinetic Studies in Rat Plasma.” *Biomedical Chromatography* 35(3): 1–8.

Weijers, J. W. H.; Schouten, S.; Spaargaren, O. C.; Sinninghe Damsté, J. S. 2006. “Occurrence and Distribution of Tetraether Membrane Lipids in Soils: Implications for the Use of the TEX_{86} Proxy and the BIT Index.” *Organic Geochemistry* 37(12): 1680–1693.

Weijers, J. W. H.; Schouten, S.; Hopmans, E. C.; Geenevasen, J. A. J.; David, O. R. P.; Coleman, J. M.; Pancost, R. D.; Sinninghe Damsté, J. S. 2006. “Membrane Lipids of Mesophilic Anaerobic Bacteria Thriving in Peats Have Typical Archaeal Traits.” *Environmental Microbiology* 8(4): 648-657.

Weijers, J. W. H.; Wiesenberg, G. L. B.; Bol, R.; Hopmans, E. C.; Pancost, R. D. 2010. “Carbon Isotopic Composition of Branched Tetraether Membrane Lipids in Soils Suggest a Rapid Turnover and a Heterotrophic Life Style of their Source Organism(s).” *Biogeosciences* 7(9): 2959-2973.

White, A. J.; Stevens, L. R.; Lorenzi, V.; Munoz, S. E.; Lipo, C. P.; Schroeder, S. 2018. “An Evaluation of Fecal Stanols as Indicators of Population Change at Cahokia, Illinois.” *Journal of Archaeological Science* 93(March): 129–134.

White, A. J.; Stevens, L. R.; Lorenzi, V.; Munoz, S. E.; Schroeder, S.; Cao, A.; Bogdanovich, T. 2019. “Fecal Stanols Show Simultaneous Flooding and Seasonal Precipitation Change Correlate with Cahokia’s Population Decline.” *Proceedings of the National Academy of Sciences of the United States of America* 116(12): 5461–5466.

Woese, C. R. 2004. “The Archaeal Concept and the World It Lives in: A Retrospective.” *Photosynthesis Research* 80: 361–372.

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): 5088-5090.

Wörmer, L.; Lipp, J. S.; Schröder, J. M.; Hinrichs, K. U. 2013. “Application of Two New LC-ESI-MS Methods for Improved Detection of Intact Polar Lipids (IPLs) in Environmental Samples.” *Organic Geochemistry* 59: 10–21.

Wormer, L.; Elvert, M.; Fuchser, J.; Lipp, J. S.; Buttigieg, P. L.; Zabel, M.; Hinrichs, K.-U. 2014. “Ultra-high-resolution Paleoenvironmental Records via Direct Laser-based Analysis of Lipid Biomarkers in Sediment Core Samples.” *Proceedings of the National Academy of Sciences* 111(44): 15669-15674.

Wörmer, L.; Wendt, J.; Alfken, S.; Wang, J. X.; Elvert, M.; Heuer, V. B.; Hinrichs, K. U. 2019. “Towards Multiproxy, Ultra-high resolution Molecular Stratigraphy: Enabling Laser-induced Mass Spectrometry Imaging of Diverse Molecular Biomarkers in Sediments.” *Organic Geochemistry* 127: 136-145.

Wuchter, C.; Schouten, S.; Coolen, M. J. L.; Sinninghe Damsté, J. S. 2004. “Temperature-Dependent Variation in the Distribution of Tetraether Membrane Lipids of Marine Crenarchaeota: Implications for TEX_{86} Paleothermometry.” *Paleoceanography* 19(4): 1–10.

Xie, S.; Lipp, J. S.; Wegener, G.; Ferdelman, T. G.; Hinrichs, K. U. 2013. “Turnover of Microbial Lipids in the Deep Biosphere and Growth of Benthic Archaeal Populations.” *Proceedings of the National Academy of Sciences of the United States of America* 110(15): 6010-6014.

Yang, H.; Xiao, W.; Jia, C.; Xie, S. 2014. “Paleoaltimetry Proxies Based on Bacterial Branched

Tetraether Membrane Lipids in Soils." *Frontiers of Earth Science* 9(1): 13–25.

Yoshinaga, M. Y.; Kellermann, M. Y.; Rossel, P. E.; Schubotz, F.; Lipp, J. S.; Hinrichs, K. U. 2011. "Systematic Fragmentation Patterns of Archaeal Intact Polar Lipids by High-Performance Liquid Chromatography/electrospray Ionization Ion-trap Mass Spectrometry." *Rapid Communications in Mass Spectrometry* 25(23): 3563-3574.

Yu, Y.; Li, G.; Wu, D.; Liu, J.; Chen, J.; Hu, N.; Wang, H.; Wang, P.; Wu, Y. 2020. "Thiol Radical-based Chemical Isotope Labelling for Sterols Quantitation through High Performance Liquid Chromatography-tandem Mass Spectrometry Analysis." *Analytica Chimica Acta* 1097: 110-119.

Zeng, Z.; Liu, X. L.; Wei, J. H.; Summons, R. E.; Welander, P. V. 2018. "Calditol-linked Membrane Lipids are Required for Acid Tolerance in Sulfolobus Acidocaldarius." *Proceedings of the National Academy of Sciences of the United States of America* 115(51): 12932-12937.

Zeng, Z.; Liu, X. L.; Farley, K. R.; Wei, J. H.; Metcalf, W. W.; Summons, R. E.; Welander, P. V. 2019. "GDGT Cyclization Proteins Identify the Dominant Archaeal Sources of Tetraether Lipids in the Ocean." *Proceedings of the National Academy of Sciences of the United States of America* 116(45): 22505–22511.

Zhao, X.; Yan, P.; Wang, R.; Zhu, S.; You, J.; Bai, Y.; Liu, H. 2016. "Sensitive Determination of Cholesterol and Its Metabolic Steroid Hormones by UHPLC – MS / MS via Derivatization Coupled with Dual Ultrasonic-Assisted Dispersive Liquid – Liquid Microextraction." 30: 147–154.

Zhu, C.; Lipp, J. S.; Wörmer, L.; Becker, K. W.; Schröder, J.; Hinrichs, K. U. 2013. "Comprehensive Glycerol Ether Lipid Fingerprints through a Novel Reversed Phase Liquid Chromatography-Mass Spectrometry Protocol." *Organic Geochemistry* 65: 53–62.

Zhu, Z. T.; Li, Y. M.; Guo, Y. T.; Sun, M.; Zhao, Y. F. 2006. “A New Fragmentation Rearrangement of the N-Terminal Protected Amino Acids Using ESI-MS/MS.” *Indian Journal of Biochemistry and Biophysics* 43(6): 372–376.

Zillig, W. 1991. “Comparative Biochemistry of Archaea and Bacteria.” *Current Opinion in Genetics and Development* 1(4): 544-551.

Zu, S.; Deng, Y.-Q.; Zhou, C.; Li, J.; Li, L.; Chen, Q.; Li, X.-F.; Zhao, H.; Gold, S.; He, J.; Li, X.; Zhang, C.; Yang, H.; Cheng, G.; Qin, C. F. 2020. “25-Hydroxycholesterol is a potent SARS-CoV-2 inhibitor.” *Cell Research* 30: 1043–1045.

Zubarev, R. A. and Makarov, A. 2013. “Orbitrap Mass Spectrometry.” *Analytical chemistry* 85: 5288-5296.