

## MIT Open Access Articles

*From ether to acid: A plausible degradation pathway of glycerol dialkyl glycerol tetraethers*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**Citation:** Liu, Xiao-Lei, et al. "From Ether to Acid: A Plausible Degradation Pathway of Glycerol Dialkyl Glycerol Tetraethers." *Geochimica et Cosmochimica Acta*, vol. 183, June 2016, pp. 138–52. © 2016 Elsevier Ltd.

**As Published:** <http://dx.doi.org/10.1016/J.GCA.2016.04.016>

**Publisher:** Elsevier BV

**Persistent URL:** <http://hdl.handle.net/1721.1/117014>

**Version:** Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

**Terms of use:** Creative Commons Attribution-NonCommercial-NoDerivs License



1 From ether to acid: a plausible degradation pathway of glycerol  
2 dialkyl glycerol tetraethers

3

4 Xiao-Lei Liu <sup>a, b, \*</sup>, Daniel Birgel <sup>c</sup>, Felix J. Elling <sup>a, #</sup>, Paul A. Sutton <sup>d</sup>, Julius S. Lipp <sup>a</sup>,  
5 Rong Zhu <sup>a, §</sup>, Chuanlun Zhang <sup>e</sup>, Martin Könneke <sup>a</sup>, Jörn Peckmann <sup>c, f</sup>, Steven J.  
6 Rowland <sup>d</sup>, Roger E. Summons <sup>b</sup>, Kai-Uwe Hinrichs <sup>a</sup>

7

8

9 <sup>a</sup> Organic Geochemistry Group, MARUM Center for Marine Environmental Sciences and  
10 Department of Geosciences, University of Bremen, 28359 Bremen, Germany

11 <sup>b</sup> Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of  
12 Technology, 77 Massachusetts Avenue, Cambridge, MA 02139-4307, USA

13 <sup>c</sup> Institute of Geology, University of Hamburg, 20146 Hamburg, Germany

14 <sup>d</sup> School of Geography, Earth and Environmental Sciences, University of Plymouth,  
15 Plymouth, Devon PL4 8AA, U.K.

16 <sup>e</sup> State Key Laboratory of Marine Geology, Tongji University, Shanghai 200092, China

17 <sup>f</sup> Department of Geodynamics and Sedimentology, Center for Earth Sciences, University  
18 of Vienna, 1090 Vienna, Austria

19 <sup>#</sup> Present address, Harvard University, 20 Oxford St, Cambridge, MA 02138, USA

20 <sup>§</sup> Present address, ETH Zurich, Universitätsstrasse 16, 8092 Zürich, Switzerland

21

22 \* Corresponding author, E-mail: xlliu@mit.edu

23

24

25 **Abstract**

26

27 Glycerol dialkyl glycerol tetraether (GDGT) are ubiquitous microbial lipids with  
28 extensive demonstrated and potential roles as paleoenvironmental proxies. Despite the  
29 great attention they receive, comparatively little is known regarding their diagenetic fate.  
30 Putative degradation products of GDGTs, identified as hydroxyl and carboxyl derivatives,  
31 were detected in lipid extracts of marine sediment, seep carbonate, hot spring sediment  
32 and cells of the marine thaumarchaeon *Nitrosopumilus maritimus*. The distribution of  
33 GDGT degradation products in environmental samples suggests that both biotic and  
34 abiotic processes act as sinks for GDGTs. More than a hundred newly recognized  
35 degradation products afford a view of the stepwise degradation of GDGT via 1) ether  
36 bond hydrolysis yielding hydroxyl isoprenoids, namely, GDGTol (glycerol dialkyl  
37 glycerol triether alcohol), GMGD (glycerol monobiphytanyl glycerol diether), GDD  
38 (glycerol dibiphytanol diether), GMM (glycerol monobiphytanol monoether) and bpdol  
39 (biphytanediol); 2) oxidation of isoprenoidal alcohols into corresponding carboxyl  
40 derivatives and 3) chain shortening to yield C<sub>39</sub> and smaller isoprenoids. This plausible  
41 GDGT degradation pathway from glycerol ethers to isoprenoidal fatty acids provides the  
42 link to commonly detected head-to-head linked long chain isoprenoidal hydrocarbons in  
43 petroleum and sediment samples. The problematic C<sub>80</sub> to C<sub>82</sub> tetraacids that cause  
44 naphthenate deposits in some oil production facilities can be generated from H-shaped  
45 glycerol monoalkyl glycerol tetraethers (GMGTs) following the same process, as  
46 indicated by the distribution of related derivatives in hydrothermally influenced  
47 sediments.

48

## 49 **1. INTRODUCTION**

50

51 Lipid biomarkers have the potential of providing valuable information regarding the  
52 composition of ancient ecosystems and paleoenvironmental conditions throughout most  
53 of the sedimentary record (e.g., Peters et al., 2004). For most of the frequently used  
54 biomarkers, such as steroids, hopanoids, pigments and their derivatives, the post-  
55 depositional structural transformations are rather well constrained (e.g., Mackenzie et al.,  
56 1982; Innes et al., 1997; Peters et al., 2004; Brocks and Schaeffer, 2008) and contribute  
57 to the foundation of the geological biomarker concept that links geomolecules to their  
58 biological precursors. An exception is the intensively studied group of isoprenoidal  
59 glycerol dialkyl glycerol tetraether (GDGT) lipids produced by Archaea. Despite their  
60 prominent use as molecular proxies for the reconstruction of paleoenvironmental  
61 conditions (e.g., Ingalls and Pearson, 2013; Schouten et al., 2013) and, in their intact  
62 polar form, for the ecology of extant archaeal communities (e.g., Lipp and Hinrichs, 2009;  
63 Liu et al., 2011; Meador et al., 2015; Yoshinaga et al., 2015), our understanding of their  
64 post-depositional behavior is fragmentary.

65

66 Three principal archaeal lipid categories are observed in environmental and geological  
67 samples: (i) intact polar lipids (IPL) as building blocks of the cellular membrane,  
68 consisting of a glycerolalkylether backbone and a polar headgroup which, in most  
69 instances is glycosidic (Sturt et al., 2004, Lipp and Hinrichs, 2009), (ii) the corresponding  
70 core glycerol alkyl ethers derived from hydrolytic cleavage of the polar headgroups on  
71 timescales of days to several tens of millennia, depending on depositional conditions and

72 enzymatic activity (Harvey et al., 1986; Xie et al., 2013), and (iii) degradation products  
73 of core glycerol alkyl ethers that occur as hydrocarbons (Moldowan and Seifert, 1979),  
74 alcohols (Schouten et al., 1998; Saito and Suzuki, 2010) and carboxylic acids (Meunier-  
75 Christmann, 1988; Schouten et al., 2003; Birgel et al., 2008a).

76

77 The second group, i.e., the core lipids, is among the most extensively studied biomarker  
78 class in the last decade (Ingalls and Pearson, 2013; Schouten et al., 2013), in particular  
79 the GDGT derivatives. GDGTs accumulate in cold and moderately heated aquatic  
80 sediments with seemingly little molecular alteration and remain intact in sediments over  
81 tens of millions of years (e.g., Kuypers et al., 2001). GDGTs, including the bacterial non-  
82 isoprenoidal types (Weijers et al., 2009; Liu et al., 2012b), are among the most prominent  
83 lipids in marine sediments and soils. Their ubiquity and abundance result from both the  
84 widespread distribution of their producing, largely uncultured, microbes and their  
85 relatively high recalcitrance caused by the ether-linkages. Within the domain Archaea,  
86 GDGTs are taxonomically widely distributed and probably produced by members of all  
87 phyla (Pearson and Ingalls, 2013; Schouten et al., 2013).

88

89 Thermal diagenesis (Rowland, 1990) and hydrous pyrolysis (Pease et al., 1998)  
90 experiments have shown the generation of isoprenoidal hydrocarbons from fresh archaeal  
91 cultures. The exact fate of GDGTs is not clear, but they seem to be the most plausible  
92 precursors of a wealth of compounds of putative archaeal origin found in thermally  
93 mature formations. These compounds include head-to-head linked C<sub>32</sub> to C<sub>40</sub> isoprenoid  
94 hydrocarbons in petroleum samples (Moldowan and Seifert, 1979), biphytanediols

95 (Schouten et al., 1998; Saito and Suzuki, 2010) and biphytanediols (Meunier-  
96 Christmann, 1988; Birgel et al., 2008a) in recent sediments and rock samples of possibly  
97 diagenetic and/or biogenic origin; however these diols or diacids have never been  
98 detected in any archaeal cell extracts. Another group of recently discovered, widespread  
99 compounds includes a series of glycerol ether derivatives, the glycerol dibiphytanol  
100 diethers (GDDs; Knappy et al., 2012; Liu et al., 2012a). Although the occurrence of core  
101 GDDs and their glycosidic intact polar lipids in archaeal cell extracts suggests that they  
102 play a role in archaeal lipid biosynthesis (Liu et al., 2012a; Meador et al., 2014), a  
103 diagenetic contribution of these lipids in natural settings cannot be ruled out (e.g., Yang  
104 et al., 2014).

105

106 Another conspicuous compound series of putative archaeal origin are the so-called ‘H-  
107 shaped’ or ‘ARN’ C<sub>80</sub>-C<sub>82</sub> isoprenoidal tetracarboxylic acids found in certain petroleum  
108 types (Lutnaes et al., 2006, 2007) and believed to contribute significantly to the  
109 problematic naphthenate deposits formed during oil processing (e.g., Baugh et al., 2004;  
110 2005). The archaeal lipids that are structurally related to the C<sub>80</sub> tetraacids are ‘H-shaped’  
111 glycerol monoalkyl glycerol tetraethers (H-GMGTs) found in thermophilic archaeal taxa  
112 (Morii et al., 1998; Schouten et al., 2008a), although H-GMGT-0 may also have non-  
113 thermophilic origins (Schouten et al. 2013). Whether the C<sub>80</sub> tetraacids are degradation  
114 products of H-GMGTs or actually bio-surfactants directly synthesized by Archaea living  
115 in the crude oil remains ambiguous (Lutnaes et al., 2006, 2007). To date, isoprenoidal  
116 tetracarboxylic acids have not been detected in archaeal cells.

117

118 Based on the distribution patterns of newly identified series of GDGT degradation  
119 products in sedimentary samples (hydrothermally overprinted sediments from the  
120 Guaymas Basin and a hot spring in China, Miocene seep carbonates, and marine  
121 subsurface sediments), cell extracts and hydrolysis experiments, here we construct a  
122 precursor-product reaction network from GDGTs to alcoholic and carboxylic acid  
123 biphytane derivatives. Additionally, through the identification of five types of putative  
124 intermediates in the sediment from Guaymas Basin, we provide strong support for the  
125 hypothesis that isoprenoidal C<sub>80</sub> to C<sub>82</sub> tetraacids are derived from step-wise degradation  
126 of 'H-shaped' GMGTs.

127

## 128 **2. MATERIALS and METHODS**

129

### 130 **2.1. Sample collection and preparation**

131

132 *N. maritimus* strain SCM1 was grown aerobically at 28 °C and pH 7.5 in 8.5 l HEPES-  
133 buffered Synthetic Crenarchaeota Medium (1.5 mM NH<sub>4</sub>Cl; Könneke et al., 2005;  
134 Martens-Habbena et al., 2009). The medium was inoculated with 5% of a mid-  
135 logarithmic phase pre-culture of *N. maritimus*. Biomass was harvested early and late in  
136 the growth phase as well as early and late in the stationary phase (one batch for each  
137 time-point) using cross-flow filtration (Elling et al., 2014). Purity of the culture was  
138 checked daily by phase contrast microscopy. Growth was monitored by measuring nitrite  
139 formation photometrically (Stickland and Parsons, 1972) and by counting 2%  
140 formaldehyde-fixed, SYBR Green I stained cells (Lunau et al., 2005) at the beginning

141 and the end of the experiment. Lipids were extracted from each batch following a  
142 modified Bligh and Dyer protocol as described previously (Sturt et al., 2004; Elling et al.,  
143 2014).

144

145 The Marmorito seep carbonate samples (Marmorito limestone; see also Peckmann et al.,  
146 1999) were taken close to the village of Marmorito, in the Monferrato hills close to  
147 Torino, Italy. The Marmorito limestone is composed of dolomite and calcite, and is  
148 embedded in Miocene strata chiefly consisting of siliciclastic sediments that were  
149 deposited in a marine shelf environment. Rock samples from the vicinity Marmorito,  
150 comprising not only the so-called Marmorito limestone, are among the early examples  
151 where methane-seepage, fossil chemosynthetic benthic communities, molecular fossils,  
152 and methane-related carbonate precipitation have been described (e.g., Clari et al., 1988,  
153 1994; Peckmann et al., 1999; Thiel et al., 1999). Apart from the characteristic molecular  
154 fossils of the consortium mediating the anaerobic oxidation of methane, biomarkers of  
155 aerobic methanotrophic bacteria were also found (Peckmann et al., 1999; Birgel and  
156 Peckmann, 2008). A detailed list of compounds identified in the Marmorito limestone  
157 and a description of the applied decalcification procedure can be found in Birgel and  
158 Peckmann (2008) and references therein. Lipid extraction was performed as described  
159 in Birgel et al. (2006).

160

161 The Guaymas Basin sediment sample was retrieved during the *RV Atlantis* cruise AT15-  
162 56 to the Guaymas Basin, Gulf of California, Mexico, during *Alvin* dive 4568 (November  
163 22 to December 6 2009, 27° 00.449' N, 111° 24.347' W). The sample came from an oil-



164 impregnated hydrothermally active area, where sedimentary temperature steeply  
165 increased from ~3 °C to ~100 °C within 35 cm (Gutierrez et al., 2015). The sample was  
166 surface sediment (0-4 cm depth) with the highest temperature at time of sampling  
167 reaching ~12 °C. However, due to the dynamic nature of the hydrothermal activity and  
168 the upward flux of fluids in the Guaymas Basin (Pearson et al., 2005), it is likely that the  
169 sample has been previously heated to higher temperatures and/or contains extractable  
170 organic matter formed in deeper layers at higher temperatures. Wet sediment (15-20 g)  
171 was extracted using the modified Bligh and Dyer method (Sturt et al., 2004). To  
172 minimize the heavy background of oil contaminants, an aliquot of the total lipid extract  
173 (TLE) was cleaned with the Hybrid SPE<sup>®</sup>-Phospholipid cartridges before LC-MS  
174 analysis according to the method described by Zhu et al. (2013a).

175

176 Two further sediment samples, including a marine subsurface sediment (Leg201-1227;  
177 Hole 1227A, mixed from five samples: 1227A-2H2-65-75cm, 1227A-2H5-83-93cm,  
178 1227A-3H2-55-65cm, 1227A-11H2-118-128m, 1227A-13H3-0-15cm; spanning from  
179 8.1-113.6 m below sea floor) and a hot spring sediment (T-15, see sample description in  
180 SI) were also extracted with the Bligh and Dyer method for lipid analysis.

181

## 182 **2.2. Lipid analysis**

183

184 For lipid analysis by normal phase liquid chromatography (NP-LC), an aliquot of TLE of  
185 each sample was dissolved in *n*-hexane/isopropanol (99.5:0.5 v/v) for injection.

186 Compound separation was performed on a Dionex Ultimate 3000 RS UHPLC system

187 (Thermo Scientific, Bremen, Germany) at 50 °C, following the recently developed  
188 tandem column protocol (Becker et al., 2013) using two ACQUITY UPLC® BEH Hilic  
189 Amide columns (2.1 x 150 mm, 1.7 µm, Waters). Solvent gradient was programmed for a  
190 constant flow rate of 0.5 mL min<sup>-1</sup> and a linear increase from 3% B to 20% B in 20  
191 minutes, and then linearly to 50% B at 35 minutes, after then up to 100% B at 45 min,  
192 holding for 6 minutes, finally back to 3% B for 9 minutes to re-equilibrate the column,  
193 where A was *n*-hexane and B was *n*-hexane/isopropanol (90:10). Detection was achieved  
194 with a Bruker Maxis accurate-mass quadrupole time-of-flight (qTOF) mass spectrometer  
195 (Bruker Daltonik, Bremen, Germany) coupled to the UHPLC via an atmospheric pressure  
196 chemical ionization (APCI) interface run in positive ion mode. APCI source parameters  
197 were as follows: corona current 3500 nA, nebulizer gas pressure 5 bar, drying gas flow 8  
198 L min<sup>-1</sup>, drying gas (N<sub>2</sub>) temperature 160 °C, vaporizer temperature 400 °C. The scan  
199 range was 150-2000 *m/z* at a rate of 2 Hz. Lipids were identified based on accurate mass  
200 (better than 1 ppm), retention times and diagnostic fragments and under consideration of  
201 general GDGT mass spectral features (e.g., Liu et al., 2012b), and quantified by  
202 measurement of [M+H]<sup>+</sup> responses, with a extraction window of individual ion  
203 chromatograms of ± 0.01 *m/z* units.

204

205 For the detection of carboxylic acid derivatives, reversed phase (RP) LC-MS (cf. Zhu et  
206 al., 2013b) was applied with the same LC system using an ACE3 C<sub>18</sub> column (2.1 × 150  
207 mm, 3 µm; Advanced Chromatography Technologies Ltd., Aberdeen, Scotland) coupled  
208 with a guard cartridge and maintained at 45 °C, and the same mass spectrometer (qTOF)  
209 equipped with an electrospray ionization (ESI) source and operated in positive mode

210 (Bruker Daltonik, Bremen, Germany). An aliquot of TLE of each sample was dissolved  
211 in methanol prior to injection. Separation of compounds was achieved isocratically with  
212 100% eluent A for 10 min, followed by a rapid gradient to 24% B in 5 min, and then a  
213 slow gradient to 65% B in 55 min at a flow rate of 0.2 mL min<sup>-1</sup>, where the eluent A was  
214 methanol/formic acid/14.8 M NH<sub>3(aq)</sub> (100:0.04:0.10, v/v/v) and B was 2-propanol/formic  
215 acid/14.8 M NH<sub>3(aq)</sub> (100:0.04:0.10, v/v/v). After each run, the column was washed with  
216 90% B for 10 min and subsequently re-equilibrated with 100% A for another 10 min. The  
217 ESI-MS conditions were set as capillary voltage 4500 V, nebulizing gas (N<sub>2</sub>) pressure 0.8  
218 bar, and dry gas (N<sub>2</sub>) 4 L min<sup>-1</sup> at a temperature of 200 °C.

219

## 220 **3. RESULTS**

221

### 222 **3.1. Structural elucidation of products of GDGT breakdown**

223

224 Table 1 provides an overview of 19 distinct compound classes observed in this study.  
225 Structural assignments are based on formulae established from accurate mass  
226 measurements, chromatographic behavior and tandem mass spectrometry (MS<sup>2</sup>) (Fig. 1, 2  
227 and 3 and S1a-d contain data showing detailed MS<sup>2</sup> fragmentation patterns). Many  
228 compounds remain only tentatively identified. Identifications of the hydroxyl derivatives,  
229 including GDGTol (glycerol dialkyl glycerol triether alcohol), GMGD (glycerol  
230 monobiphytanyl glycerol diether), GDD (glycerol dibiphytanol diether), GMM (glycerol  
231 monobiphytanol monoether) and bpdol (biphytanediol), were additionally confirmed by  
232 examination of products released after acidic degradation of purified GDGT-0

233 (compound compositions are given in Fig. S2). The identification of the H-tetrol was  
234 additionally supported by the co-injection of an H-tetrol mixture synthesized by reduction  
235 of the corresponding acids, which were isolated and characterized by NMR spectroscopy  
236 previously (Lutnaes et al., 2007; Fig. S3, preparation described in SI). C<sub>39</sub> isoprenoids  
237 were assigned based on chromatographic behavior, molecular formulae generated from  
238 accurate mass measurements and characteristic fragment ions (MS<sup>2</sup> fragmentation  
239 patterns are shown in Fig. S1a). For example, the two isomers of GDGTol-0 (highlighted  
240 with green circles in Fig. 3e) eluted earlier than those of C<sub>40/39</sub>-GDGTA-0 (two pink solid  
241 line circles in Fig. 3e). In addition, GDGTol-0 and C<sub>40/39</sub>-GDGTA-0 also have  
242 measurably different molecular masses, [M+H]<sup>+</sup> of m/z: 1320.3312 and 1320.2930,  
243 respectively and these afford calculated formulae of C<sub>86</sub>H<sub>175</sub>O<sub>7</sub><sup>+</sup> for GDGTol-0 and  
244 C<sub>85</sub>H<sub>171</sub>O<sub>8</sub><sup>+</sup> for C<sub>40/39</sub>-GDGTA-0. To add weight to this assignment, the fragment ion of  
245 m/z: 651.6245 in the MS<sup>2</sup> of C<sub>40/39</sub>-GDGTA-0 compared to m/z: 665.6391 of GDGTA-0  
246 represents the loss of methylene from the biphytane (Fig. S1a).

247

### 248 **3.2. Terminally hydroxylated biphytanyl derivatives**

249

250 Hydrolysis of the glycerol ether bonds in isoprenoidal GDGT generates five types of  
251 biphytane-based alcohols, with and without the glycerol backbones, namely, GDGTol,  
252 GMGD, GDD, GMM and bpdol (compound structures are given in Fig. S2). For  
253 example, we demonstrate their structural relationships to acyclic caldarchaeol, GDGT-0,  
254 with a parallel glycerol configuration, a distribution observed in a seep carbonate sample  
255 from Marmorito (Fig. 1a and b). Constitutional isomers, which are labeled as peak ‘a’

256 and 'b' in the chromatogram of GDGTol, GMGD and GMM in Fig. 1, represent different  
257 combinations of *sn*-2 or *sn*-3 glycerol ether bonding. Degradation products of the ring-  
258 containing GDGTs are more numerous, due to different possible arrangements of a ring  
259 on the biphytanyl chains. A detailed isomeric study of GDGTol, GMGD and GMM and  
260 their implications will be discussed in our following works. As for H-GMGTs, which  
261 have two biphytane chains linked by a bis-methylene C-C bond at position C<sub>20</sub> (Lutnaes  
262 et al., 2006; 2007) the cleavage of two glycerol units results in a C<sub>80</sub> H-tetrol, instead of  
263 two biphytanediols (bpdioles). These hydroxyl derivatives of GDGT were either all, or at  
264 least some, present in our sample set comprising extracts of marine sediment, seep  
265 carbonate, hydrothermal vent, hot spring sediment and archaeal cell extract (Table 1).

266

### 267 **3.3. Terminally-carboxylated biphytanyl derivatives**

268

269 Carboxylic acids, corresponding to each of hydroxyl derivatives mentioned above, could  
270 logically be generated by oxidation of each terminal alcohol of the biphytane. These were  
271 also detected in the analyzed samples and are subsequently referred to as GDGTA for the  
272 carboxyl analogue of GDGTol and GMMA for GMM (Fig. 1a, b and Table 1). For  
273 isoprenoids possessing two primary hydroxyl groups, such as GDD and bpdiole, both  
274 mono- and dicarboxyl analogues were detected, with the monocarboxyl derivative of  
275 GDD termed GDDA, the dicarboxyl derivative GDDAA, as well as biphytane mono- and  
276 diacids. In the case of GMGD, which does not possess a primary biphytane-bound  
277 hydroxyl group, no acid derivatives were detected. In the case of H-tetrol, which contains  
278 four primary hydroxyl groups on its two linked bpdioles, we detected mono-, di-, tri- and

279 tetraacids in an oily sediment sample collected from a hydrothermal vent site in Guaymas  
280 Basin (Fig. 2).

281

282 Under APCI conditions, the ionization patterns of biphytane monoacid/monool  
283 (bpmonoacid/ol) and biphytanedi acids (bpdiacid) differed from bpdliols. A protonated  
284 molecular ion  $[M+H]^+$  was usually the major ion in mass spectra of bpdliols, while the  
285 acids produced more complex mass spectra in our analyses. Dehydrated molecular ions  
286 and unknown adducts were formed during the ionization of carboxyl derivatives (Fig. 1b  
287 and S1d). The main adduct ions detected were  $[M+H+42]^+$  for bpmonoacid/ol and  
288  $[M+H+84]^+$  for bpdliacs (Fig. S1d).

289

### 290 **3.4. C<sub>39</sub> and C<sub>38</sub> isoprenoids**

291

292 In addition to the series of hydroxylated and carboxylated C<sub>40</sub> biphytane derivatives, we  
293 identified some with shortened alkyl chains such as the C<sub>39</sub> analogues. For example, in  
294 most analyzed samples analogues with one methylene unit less than GDGTol and  
295 GDGTA, here termed C<sub>40/39</sub>-GDGTol and C<sub>40/39</sub>-GDGTA (Arabic numbers represent the  
296 carbon number of the two isoprenoidal chains in the molecule), always co-occurred at  
297 lower abundance with their C<sub>40/40</sub> analogues (Fig. 3b, e). C<sub>39</sub> derivatives were also  
298 detected as analogues of GDD, bpdliol and GMM (Fig. 3a, c and d). Additionally, we  
299 detected signals that we attribute to the pseudo-homologue of C<sub>38</sub> isoprenoids, as  
300 GDGTol and GDD (Fig. 3a and b). Within the group of GDDs reduced by two C-atoms,  
301 we observed both C<sub>39/39</sub> and C<sub>40/38</sub> derivatives (Fig. 3a). Three isomers of the acyclic and

302 four isomers of crenarchaeol-related GDD were detected (Fig. 3a). In previous studies  
303  $\delta^{13}\text{C}$  depleted  $\text{C}_{39}$  head-to-head linked isoprenoids were detected in Cretaceous (Sandy et  
304 al., 2012) and Carboniferous seep carbonates (Birgel et al., 2008b); those compounds  
305 presumably represent the hydrocarbon derivatives of the  $\text{C}_{39}$  functionalized compounds  
306 found in this study. Interestingly, no derivatives of  $\text{C}_{80}$  H-tetrol and tetraacid with  
307 reduced carbon chains were detected in our samples.

308

## 309 **4. DISCUSSION**

310

### 311 **4.1. The occurrence of GDGT degradation derivatives in environmental samples**

312

313 The distributions of major degradation derivatives of GDGT-0 and H-GMGT-0 were  
314 compared in four representative environmental samples (Fig. 4), which include modern to  
315 late Miocene marine subsurface sediment (Leg201-1227), modern hot spring (T-15) and  
316 hydrothermally heated sediment (Guaymas Basin 4568), and the Miocene Marmorito  
317 seep carbonate. Distinct patterns of degradation derivatives in these four types of samples  
318 from different environments and of different age (modern sediments and ancient  
319 carbonate rock) reflect variable degrees of degradation and preservation.

320

#### 321 ***4.1.1. Distributions of regular GDGT degradation products in environmental samples***

322

323 The distributions of GDGT degradation products across our sample set suggest that the  
324 biological sources of GDGTs, as well as the debris depositional histories, influence the

325 diagenetic trajectory of GDGTs. By way of example, the degradation derivatives of  
326 GDGT-0 are around twice as abundant as their precursor in the marine sediment sample  
327 of Leg 201-1227, in which bpdial-0 comprises over half of all of the detected degradation  
328 products (Fig. 4). However, the GDGT degradation derivatives in Miocene seep  
329 carbonates exhibited a lower overall relative abundance, but with a greater variety and  
330 portion of labile components, such as the carboxyl derivatives (Table 1 and Fig. 4).

331 Regular isoprenoidal GDGTs (as compared to H-GMGTs) preserved in marine sediments  
332 are primarily derived from planktonic archaea, dwelling in the water column. In contrast,  
333 those in seep carbonates receive a larger *in-situ* contribution from benthic communities  
334 engaged in anaerobic oxidation of methane (AOM), as confirmed by low  $\delta^{13}\text{C}$  values of  
335 biphytanic diacids (cf. Birgel et al., 2008a). The higher proportion of degradation  
336 derivatives compared to their GDGT precursors in deep subsurface sediment apparently  
337 results from a mild to moderate degradation process, but degradation lasted longer than at  
338 the Marmorito seep site and was favored by extensive transportation. The extraordinarily  
339 good preservation of labile compounds in seep carbonates can be attributed to significant  
340 *in-situ* GDGT production combined with co-eval carbonate formation and resulting early  
341 lithification within methane seepage systems (e.g., Peckmann et al., 1999; Birgel et al.,  
342 2008a), rather than representing a time-integrated pattern with input from various  
343 sedimentary or sedimentary/planktonic archaea (cf. Feng et al., 2014; Birgel et al., 2008a).

344

345 The relative abundance of degradation products in Guaymas basin sediment is over 40%  
346 of all GDGT-derived compounds (Fig. 4). Both biodegradation and thermal diagenesis  
347 can be potential sources of degradation products, due to the presence of both active



348 microbial communities and high temperature hydrothermal fluids in Guaymas basin  
349 sediments (Teske et al., 2014; Gutierrez et al., 2015). In contrast, there are only low  
350 abundances of GDD and bpdial (< 10%) detected as GDGT degradation products in the  
351 hot spring sediment (Fig. 4).

352

353 In addition to the contrasting patterns of the different environments, there is a  
354 compositional discrepancy between each class of GDGT product within the same sample.  
355 For example, previous studies that documented the presence of bpdials and bpdialic acids in  
356 environmental samples have shown that the ring distributions and carbon isotopic  
357 compositions differ for biphytanes released from coexisting GDGTs (Schouten et al.,  
358 1998; Birgel et al., 2008a; Saito and Suzuki, 2010). The Marmorito limestone samples  
359 (Fig. 5; Birgel et al. (2008a) provide results on hydrocarbons and bpdialic acids), show  
360 distinct ring distribution patterns for GDGT, bpdial, bpdialic acid/ol and bpdialic acid;  
361 GDGTs are dominated by GDGT-0 and crenarchaeol, while the relative abundance of  
362 tricyclic biphytane derivatives derived from crenarchaeol gradually decreases for the  
363 hydroxyl to carboxyl products. Multiple inputs, combined with selective preservation  
364 could cause such distributional differences. The ring distribution of GDGTs preserved in  
365 sediments may reflect a mixed contribution from both planktonic and benthic species and  
366 is frequently dominated by compounds from planktonic sources (Wuchter et al., 2005;  
367 Huguet et al., 2007; Lengger et al., 2012). Planktonic and sedimentary archaeal  
368 communities could contribute different lipids. Further, compared to biphytanyl products  
369 derived from benthic species within the sediment, those derived from water column will  
370 have experienced a very different transportation history. This might explain why the

371 tricyclic diacid derived from planktonic archaea are rare, while the acyclic, mono- and  
372 bicyclic diacids with origins from sedimentary methanotrophic archaea, are more  
373 common and more <sup>13</sup>C-depleted in the seep carbonate (isotopic data published in Birgel  
374 et al., 2008a).

375

#### 376 *4.1.2. The degradation of H-GMGTs to H-tetrols and H-tetraacids*

377

378 H-GMGT-0 and its hydroxylated degradation products are generally present in the  
379 samples from low-temperature and hydrothermal environments (Table 1 and Fig. 4),  
380 while compounds with multiple cycloalkyl groups and additional methylations are  
381 restricted to the hydrothermal samples alone. In low-temperature samples, H-GMGT-0 is  
382 far more abundant than H-GMD-0 (acyclic H-shaped glycerol monoalkyl diether) and  
383 tetrol-0 (Fig. 4). The apparently different patterns of H-GMGT-0 and its derivatives,  
384 compared to GDGT-0 and its highly abundant degradation products, point to  
385 contributions of distinct species and/or different diagenetic pathways.

386

387 We detected H-tetrol and H-tetraacids, and their partially oxidized ‘H-shaped’ mono-, di-  
388 and triacid intermediates, in the oily Guaymas Basin sediment (Fig. 2). The detection of  
389 these intermediates likely implies enzymatic modification of existing ‘H-shaped’ GDGTs  
390 by living organisms. H-tetraacids accumulated in naphthenate deposits in oil well  
391 infrastructure and pipelines and in crude oils, have distributions usually dominated by 4-8  
392 cyclopentyl rings and 1 or 2 additional methyl substitutions (Lutnaes et al., 2007; Sutton  
393 & Rowland, 2014). Although H-GMGT-0 is known to be present in a broad range of

394 mesophilic environments (Schouten et al., 2008a), those with multiple cyclizations and  
395 methylations have been only detected in hyperthermophilic archaeal species (Schouten et  
396 al., 2008b; Knappy et al., 2011; Liu et al., 2012b). H-GMGTs existed in both hot spring  
397 and Guaymas Basin sediments analyzed. However, H-tetraacids were only found in the  
398 oil-impregnated Guaymas Basin sediment (Table 1). In addition to the diagenetic  
399 contribution, Lutnaes et al. (2006) also speculated that these tetraacids might be produced  
400 by thermophilic, oil-degrading archaea as biosurfactants to facilitate their metabolism.  
401 Additional studies would be required to test this rigorously. Given the occurrence of  
402 biphytane, C<sub>39</sub> and smaller head-to-head linked isoprenoids in the geological record, C<sub>80</sub>  
403 based H-shaped isoprenoidal hydrocarbons might also exist. Whilst these are not  
404 detectable with conventional GC-MS, they would be amenable to high temperature GC-  
405 MS (cf. Sutton & Rowland, 2012)

406

#### 407 **4.2. The occurrence of GDGT degradation derivatives in archaeal cell extracts**

408

409 In a previous study concerning the characterization of GDD, Liu et al. (2012a) observed  
410 the existence of core GDD in the cell extracts of the methanogen *Methanothermococcus*  
411 *thermolithotrophicus*. More recently, Meador et al. (2014) reported detection of both  
412 core and monoglycosidic GDD (1G-GDD) in a culture of the planktonic ammonia  
413 oxidizer *N. maritimus*. Here we extend these observations and report newly identified  
414 biphytane derivatives in fresh biomass collected at different growth phases of cell  
415 extracts of the archaeon *N. maritimus*. Core lipids accounted for 8.2-15% of the total  
416 detected lipids during different growth phases (Elling et al., 2014). We selected the most

417 abundant component, crenarchaeol, to illustrate the distribution of its hydroxyl and  
418 carboxyl derivatives in different growth phases. GDD-Cren. and bpdial-Cren. were the  
419 only detected hydroxyl components during early exponential growth (Fig. 6); they jointly  
420 accounted for less than 1% of all crenarchaeol-based derivatives. In the late growth phase,  
421 however, the relative abundances of GDD-Cren. and bpdial-Cren. were almost doubled,  
422 while the acid (GDDA-Cren.) and a C<sub>39</sub> component (C<sub>40/39</sub>-GDD-Cren.) emerged as well.  
423 GDD-Cren., depleted by two methylene units (C<sub>39/39</sub> and C<sub>40/38</sub>-GDD-Cren.), together  
424 with GDGTol-Cren. were only observed in the stationary phase.

425

426 Although these derivatives occurred in cell extracts, they cannot be simply attributed as  
427 intermediates of GDGT biosynthesis. For example, the carboxyl and C<sub>39</sub> based  
428 derivatives, such as GDDA and C<sub>40/39</sub>-GDD, are more likely further oxidized products of  
429 GDD. For example, archaeal biphytanyl moieties, synthesized via the mevalonic acid  
430 (MVA) or methylerythritol phosphate (MEP) pathways, should have a carbon number  
431 that is a multiple of five. Accordingly, the most plausible formation pathway for the C<sub>38</sub>  
432 and C<sub>39</sub> moieties is oxidation and decarboxylation of C<sub>40</sub> precursors. Formation of C<sub>38</sub>  
433 and C<sub>39</sub> moieties via degradation of C<sub>40</sub> biphytanyl moieties is also consistent with the  
434 higher relative abundance of GDDA and C<sub>40/39</sub>-GDD during the late stationary phase (Fig.  
435 6). The loss of one methylene unit represents an  $\alpha$ -oxidation step. A similar well-known  
436 enzymatic  $\alpha$ -oxidation of isoprenoids is that of phytanic acid (C<sub>20</sub>) to pristanic acid (C<sub>19</sub>)  
437 by a wide range of organisms (e.g., Rontani and Volkman, 2003; Jansen and Wanders,  
438 2006, and other studies cited therein). As in phytanic acid, the C<sub>3</sub> methyl group in  
439 biphytanic acid (C<sub>40</sub>) derivatives prevents an initial  $\beta$ -oxidation mechanism; instead,

440 these compounds undergo  $\alpha$ -oxidation to yield C<sub>39</sub> based carboxyl isoprenoids. In such a  
441 scenario the production of GDD and GDGTol consisting of C<sub>38</sub> derivatives would require  
442 two successive  $\alpha$ -oxidation steps. This is inconsistent with the reported  $\alpha$ -oxidation of  
443 phytanic acid, which is followed by hydroxylation and then  $\beta$ -oxidation (e.g. Jansen and  
444 Wanders, 2006, and other studies cited therein). Therefore, elucidation of the degradation  
445 process leading to a C<sub>38</sub> isoprenoid requires further study. Furthermore,  $\alpha$ - as well as  $\beta$ -  
446 oxidation would only result in carboxyl derivatives, and could not explain the occurrence  
447 of hydroxyl analogues, such as C<sub>40/39</sub>-GDD and C<sub>40/39</sub>-GDGTol (Fig. 3). Although the  
448 detection of carboxyl derivatives and C<sub>38</sub> and C<sub>39</sub> based isoprenoids in a metabolic active  
449 culture of *N. maritimus* may imply an intracellular modification of existing C<sub>40</sub> based  
450 lipids, probably as a result of cell senescence, their contribution is less than 2% of the  
451 core lipid fraction, or approximately 0.01% of the entire lipidome. The exact mechanisms  
452 responsible for the remarkable proportions (up to 70% in the marine subsurface sediment,  
453 Leg201-1227, Fig. 4) of degradation products in environmental samples remain  
454 unresolved. The presence of degradation products in archaeal cell extracts, as well as  
455 their increase towards later growth and stationary phases, suggests that some are formed  
456 rapidly and probably via enzymatic catalysis. On the other hand, the high proportion and  
457 diversity of degradation products in the hydrothermally influenced settings also leaves  
458 room for an additional role of abiotically-mediated chemical degradation reactions, at  
459 least for some of the speculated steps leading via catagenesis from GDGTs to biphytanyl  
460 hydrocarbons (cf. Rowland, 1990).

461

### 462 **4.3. Evidence of analogous behavior of non-isoprenoidal GDGTs**

463 Non-isoprenoidal GDGTs, such as the hybrid isoprenoidal/branched GDGT (IB-GDGT)  
464 and branched GDGTs, are known to be widely distributed in various environments,  
465 although their exact structures and biological source(s) remain unknown (Liu et al.,  
466 2012b). The degradation pathways described here are not limited to archaeal GDGTs. For  
467 example, in one of the seep carbonate samples, which contains high abundances of  
468 branched and IB-GDGTs, there are non-isoprenoidal GDDs, GMMs and diols detected.  
469 These possibly represent degradation products of the branched and IB-GDGTs in the  
470 same deposit (Fig. S4 and S5). Our analytical methods also reveal the presence of  
471 carboxyl derivatives of these lipids. Non-isoprenoidal GDGTol, GDGTA, GMGD, and  
472 their corresponding products with reduced carbon chains (loss of one or two methylene  
473 units) were not identified, however, and we attribute this to their low overall abundance  
474 in the analyzed samples.

475

## 476 **5. CONCLUSIONS**

477

478 As with other lipid classes, the intact archaeal tetraethers released from the cells of  
479 defunct archaea into various depositional settings, are subjected to diagenesis in  
480 sediments. We detected three major classes of GDGT degradation products comprising  
481 biphytanyl molecules with terminal hydroxyl, terminal carboxyl and shortened carbon  
482 chains. A hypothetical scheme of the GDGT degradation pathway to rationalize such a  
483 pathway is illustrated in Fig. 7. The labile polar head groups of intact GDGTs, as they  
484 occur in living cells, are initially lost to produce the more recalcitrant core GDGTs.  
485 Hydrolysis of the different ether bonds then, as we suggest, results in discrete series of

486 hydroxyl derivatives composed of one or two glycerol and biphytanol units. Oxidation of  
487 each terminal hydroxyl functional group may then generate related carboxyl products.  
488 Elimination of the C<sub>1</sub> carbon via  $\alpha$ -oxidation and subsequent  $\beta$ -oxidation steps, will likely  
489 convert the C<sub>40</sub> biphytanyl based compounds into shortened isoprenoids, such as C<sub>40/39</sub>-  
490 GDGTA and C<sub>40/39</sub>-GDGTol. Further chemical or biological alteration of these  
491 intermediates may over geological time, result in the C<sub>40</sub>, C<sub>39</sub> and smaller head-to-head  
492 linked isoprenoidal hydrocarbons reported to occur in petroleum and marine deposits (e.g.  
493 Moldowan and Seifert, 1979; Birgel et al., 2008b).

494

495 A multitude of diagenetic processes will lead to the degradation of GDGTs in sediments.  
496 However, we also observed the accumulation of C<sub>39</sub> based carboxyl isoprenoids in the  
497 later growth and stationary phases of an archaeal cell culture. This implies that enzymatic  
498 pathways for degradation of GDGTs also exist, very likely, as a response to substrate  
499 limitation, senescence or cell lysis. Additional studies will be required to study this facet  
500 in greater detail.

501

502 The detection of H-GMGT, H-GMD, H-tetrols and their further oxidized carboxyl  
503 intermediates in the oil-contaminated sediments of Guaymas Basin (Fig. 2) elucidated a  
504 formation pathway from H-GMGT to H-tetraacids under anoxic conditions.

505 Metagenomic data for samples from the Guaymas Basin and especially sediment with oil  
506 impregnation may help to elucidate the origin of H-tetraacids further.

507

508 In various geological settings these hydroxyl and carboxyl derivatives co-occur with their  
509 GDGT precursors, including both isoprenoidal and non-isoprenoidal GDGTs. Our  
510 analysis of their distributions in environmental samples and archaeal cell extracts  
511 represents an initial effort to document the possible diagenetic pathways of GDGTs and  
512 to bring about the same level of understanding that we have for steroids and triterpenoids  
513 (e.g., Peters et al., 2004).

514



515

516 **ACKNOWLEDGEMENTS**

517

518 X.-L.L. was funded by European Union's Seventh Framework Programme–“Ideas”

519 Specific Programme, ERC grant agreement No. 247153 (Advanced Grant DARCLIFE;

520 PI K.-U.H.), the Simons Foundation Collaboration on the Origins of Life (SCOL).

521 Cultivation experiments and lipid analyses at the University of Bremen were funded by

522 the Deutsche Forschungsgemeinschaft through the Gottfried Wilhelm Leibniz Prize

523 awarded to K.-U.H. (Hi 616-14-1). SJR and PAS were funded by a European Research

524 Council Advanced Investigators Award to SJR for project OUTREACH (agreement no.

525 228149).

526

527

528

529 **REFERENCES**

530

531 Baugh T.D., Wolf N.O., Mediaas H., Vindstad J.E., Grande K. (2004) Characterization of  
532 a Calcium Naphthanate Deposit -The ARN Acid Discovery. Preprints/American  
533 Chemical Society, Division of Petroleum Chemistry **49**, 274–276.

534 Baugh T.D., Grande K.V., Mediaas H., Vindstad J.E., Wolf N.O. (2005) In The  
535 Discovery of High Molecular Weight Naphthenic Acids (ARN Acid) Responsible for  
536 Calcium Naphthenate Deposits. SPE International Symposium on Oilfield Scale,  
537 Aberdeen, U.K., May 11–12 2005; SPE 93011.

538 Becker K.W., Lipp J.S., Zhu C., Liu X.-L., Hinrichs K.-U. (2013) An improved method  
539 for the analysis of archaeal and bacterial ether core lipids. *Org. Geochem.* **61**, 34–44.

540 Birgel, D., Thiel, V., Hinrichs, K.-U., Elvert, M., Campbell, K.A., Reitner, J., Farmer,  
541 J.D., Peckmann, J. (2006) Lipid biomarker patterns of methane-seep microbialites  
542 from the Mesozoic convergent margin of California. *Org. Geochem.* **37**, 1289–1302.

543 Birgel, D., Peckmann, J. (2008) Aerobic methanotrophy at ancient marine methane seeps:  
544 a synthesis. *Org. Geochem.* **39**, 1659–1667.

545 Birgel D., Elvert M., Han X., Peckmann J. (2008a) <sup>13</sup>C-depleted biphytanic diacids as  
546 tracers of past anaerobic oxidation of methane. *Org. Geochem.* **39**, 152–156.

547 Birgel, D., Himmler, T., Freiwald, A., Peckmann, J. (2008b) A new constraint on the  
548 antiquity of anaerobic oxidation of methane: Late Pennsylvanian seep limestones  
549 from southern Namibia. *Geology* **36**, 543–546.

550 Brocks J.J., Schaeffer P. (2008) Okenane, a biomarker for purple sulfur bacteria  
551 (Chromatiaceae), and other new carotenoid derivatives from the 1640 Ma Barney  
552 Creek Formation. *Geochim. Cosmochim. Acta* **72**, 1396–1414.

553 Clari P., Gagliardi C., Governa M.E., Ricci B., Zuppi G.M. (1988) I calcari de  
554 Marmorito: Una testimonianza di processidiagene- tici in presenza di metano. *Boll*  
555 *Museo Regionale Sci Naurali Torino* **6**, 197–216.

556 Clari P., Fornara L., Ricci B., Zuppi G.M. (1994) Methane-derived carbonates and  
557 chemosymbiotic communities of Piedmont (Miocene, northern Italy): an update.  
558 *Geo-Marine Lett.* **14**, 201–209.

559 Elling F.J., Könneke M., Lipp J.S., Becker K.W., Gagen E.J., Hinrichs K.-U. (2014)  
560 Effects of growth phase on the membrane lipid composition of the thaumarchaeon  
561 *Nitrosopumilus maritimus* and their implications for archaeal lipid distributions in  
562 the marine environment. *Geochim. Cosmochim. Acta* **141**, 579–597.

563 Gutierrez, T., Biddle, J.F., Teske, A., Aitken, M.D. (2015) Cultivation-dependent and  
564 cultivation-independent characterization of hydrocarbon-degrading bacteria in  
565 Guaymas Basin sediments. *Front. Microbiol.* **6**, doi: 10.3389/fmicb.2015.00695

- 566 Harvey, H.R., Fallon, R.D., Patton, J.S. (1986) The effect of organic matter and oxygen  
567 on the degradation of bacterial membrane lipids in marine sediments. *Geochim.*  
568 *Cosmochim. Acta* **50**, 795–804.
- 569 Huguet, C., Schimmelmann, A., Thunell, R., Lourens, L.J., Sinninghe Damsté, J.S.,  
570 Schouten, S. (2007) A study of the TEX<sub>86</sub> paleothermometer in the water column and  
571 sediments of the Santa Barbara Basin, California. *Paleoceanography* **22**,  
572 doi:10.1029/2006PA001310
- 573 Innes, H.E., Bishop, A.N., Head, I.M., Farrimond, P. (1997) Preservation and diagenesis  
574 of hopanoids in Recent lacustrine sediments of Priest Pot, England. *Org. Geochem.*  
575 **26**, 565-576.
- 576 Jansen G.A., Wanders R.J.A. (2006) Alpha-Oxidation. *Biochim Biophys Acta – Mol. Cell*  
577 **1763**, 1403–1412.
- 578 Könneke M., Bernhard A. E., de la Torre J. R., Walker C. B., Waterbury J. B., Stahl D.  
579 A. (2005) Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature*  
580 **437**, 543–546.
- 581 Knappy, C.S., Nunn, C.E.M., Morgan, H.W., Keely, B.J. (2011) The major lipid cores of  
582 the archaeon *Ignisphaera aggregans*: implications for the phylogeny and  
583 biosynthesis of glycerol monoalkyl glycerol tetraether isoprenoid lipids.  
584 *Extremophiles* **15**, 517–528.
- 585 Knappy C.S., Keely B.J. (2012) Novel glycerol dialkanol triols in sediments:  
586 transformation products of glycerol dibiphytanyl glycerol tetraether lipids or  
587 biosynthetic intermediates? *Chem. Commun.* **48**, 841–843.
- 588 Kuypers, M. M. M., Blokker, P., Erbacher, J., Kinkel, H., Pancost, R. D., Schouten, S. and  
589 Sinninghe Damsté, J. S. (2001) Massive expansion of marine Archaea during a  
590 mid-Cretaceous oceanic anoxic event. *Science* **293**, 92-95.
- 591 Lengger, S.K., Hopmans, E.C., Reichart, G.-J., Nierop, K.G.J., Damsté, J.S.S., Schouten,  
592 S. (2012) Intact polar and core glycerol dibiphytanyl glycerol tetraether lipids in the  
593 Arabian Sea oxygen minimum zone. Part II: Selective preservation and degradation  
594 in sediments and consequences for the TEX<sub>86</sub>. *Geochim. Cosmochim. Acta* **98**, 244–  
595 258.
- 596 Lipp J.S., Hinrichs K.-U. (2009) Structural diversity and fate of intact polar lipids in  
597 marine sediments. *Geochim. Cosmochim. Acta* **73**, 6816–6833.
- 598 Liu, X.-L., Lipp, J.S., Hinrichs, K.-U. (2011) Distribution of intact and core GDGTs in  
599 marine sediments. *Org. Geochem.* **42**, 368–375.
- 600 Liu X.-L., Lipp J.S., Schröder J.M., Summons R.E., Hinrichs K.-U. (2012a) Isoprenoid  
601 glycerol dialkanol diethers: A series of novel archaeal lipids in marine sediments.

- 602 *Org. Geochem.* **43**, 50–55.
- 603 Liu X.-L., Summons R.E., Hinrichs K.-U. (2012b) Extending the known range of  
604 glycerol ether lipids in the environment: structural assignments based on tandem  
605 mass spectral fragmentation patterns. *Rapid Commun. Mass Spectrom.* **26**, 2295–  
606 2302.
- 607 Lunau M., Lemke A., Walther K., Martens-Habbena W., Simon M. (2005) An improved  
608 method for counting bacteria from sediments and turbid environments by  
609 epifluorescence microscopy. *Environ. Microbiol.* **7**, 961–968.
- 610 Lutnaes B.F., Brandal Ø., Sjöblom J., Krane J. (2006) Archaeal C<sub>80</sub> isoprenoid tetraacids  
611 responsible for naphthenate deposition in crude oil processing. *Org. Biomol. Chem.* **4**,  
612 616.
- 613 Lutnaes B.F., Krane J., Smith B.E., Rowland S.J. (2007) Structure elucidation of C<sub>80</sub>, C<sub>81</sub>  
614 and C<sub>82</sub> isoprenoid tetraacids responsible for naphthenate deposition in crude oil  
615 production. *Org. Biomol. Chem.* **5**, 1873.
- 616 Mackenzie, A.S., Brassell, S.C., Eglinton, G., Maxwell, J.R. (1982) Chemical fossils –  
617 the geological fate of steroids. *Science* **217**, 491–504.
- 618 Martens-Habbena W., Berube P.M., Urakawa H., de la Torre J.R., Stahl D.A. (2009)  
619 Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and  
620 Bacteria. *Nature* **461**, 976–979.
- 621 Meador T.B., Zhu C., Elling F.J., Könneke M., Hinrichs K.-U. (2014) Identification of  
622 isoprenoidglycosidic glycerol dibiphytanoldiethers and indications for their  
623 biosynthetic origin. *Org. Geochem.* **69**, 70–75.
- 624 Meador, T.B., Bowles, M., Lazar, C., Zhu, C., Teske, A., Hinrichs, K.-U. (2015) The  
625 archaeal lipidome in estuarine sediment dominated by members of the Miscellaneous  
626 Crenarchaeotal Group. *Environ. Microbiol.* **17**, 2441–2458.
- 627 Meunier-Christman C. (1988) Géochimie organique de phosphates et schistes bitumineux  
628 marocains: étude du processus de phosphatogenèse. PhD thesis, University of  
629 Strasbourg, p. 133.
- 630 Moldowan J.M., Seifert W.K. (1979) Head-to-head linked isoprenoid hydrocarbons in  
631 petroleum. *Science* **204**, 169–171.
- 632 Morii, H., Eguchi, T., Nishihara, M., Kakinuma, K., König, H., Koga, Y. (1998) A novel  
633 ether core lipid with H-shaped C<sub>80</sub>-isoprenoid hydrocarbon chain from the  
634 hyperthermophilic methanogen *Methanothermus fervidus*. *Biochim. Biophys. Acta*  
635 **1390**, 339–345.
- 636 Peckmann, J., Thiel, V., Michaelis, W., Clari, P., Gaillard, C., Martire, L., Reitner, J.  
637 (1999) Cold seep deposits of Beauvoisin (Oxfordian; southeastern France) and  
638 Marmorito (Miocene; northern Italy): microbially induced authigenic carbonates. *Int.*

- 639 *J. Earth Sci.* **88**, 60–75.
- 640 Pearson, A., Seewald, J.S., Eglinton, T.I. (2005) Bacterial incorporation of relict carbon  
641 in the hydrothermal environment of Guaymas Basin. *Geochim. Cosmochim. Acta* **69**,  
642 5477–5486.
- 643 Pearson, A., Ingalls, A.E. (2013) Assessing the use of Archaeal lipids as marine  
644 environmental proxies. *Annu. Rev. Earth Planet. Sci.* **41**, 359–384.
- 645 Pease, T.K., Van Vleet, E.S., Barre, J.S., Dickins, H.D. (1998) Degradation of glyceryl  
646 ethers by hydrous and flash pyrolysis. *Org. Geochem.* **29**, 979–988.
- 647 Peters, K. E., Walters, C. C., Moldowan, J. M. (2004) The biomarker guide, 2nd edn.  
648 Cambridge, UK: Cambridge University Press.
- 649 Rontani J.-F., Volkman J.K. (2003) Phytol degradation products as biogeochemical  
650 tracers in aquatic environments. *Org. Geochem.* **34**, 1–35.
- 651 Rowland, S.J. (1990) Production of acyclic isoprenoid hydrocarbons by laboratory  
652 maturation of methanogenic bacteria. *Org. Geochem.* **15**, 9–16.
- 653 Sandy, M. R., Lazăr, I., Peckmann, J., Birgel, D., Stoica, M., Roban, R. D. (2012)  
654 Methane-seep brachiopod fauna within turbidites of the Sinaia Formation, Eastern  
655 Carpathian Mountains, Romania. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **323-325**,  
656 42-59.
- 657 Saito H., Suzuki N. (2010) Distribution of acyclic and cyclic biphytanediols in recent  
658 marine sediments from IODP Site C0001, Nankai Trough. *Org. Geochem.* **41**, 1001–  
659 1004.
- 660 Schouten S., Hoefs M.J., Koopmans M.P., Bosch H.-J., Sinninghe Damsté J.S. (1998)  
661 Structural characterization, occurrence and fate of archaeal ether-bound acyclic and  
662 cyclic biphytanes and corresponding diols in sediments. *Org. Geochem.* **29**, 1305–  
663 1319.
- 664 Schouten, S., Wakeham, S. G., Hopmans, E. C., Sinninghe Damsté, J. S. (2003)  
665 Biogeochemical evidence that thermophilic archaea mediate the anaerobic oxidation  
666 of methane. *Appl. Environ. Microbiol.* **69**, 1680-1686.
- 667 Schouten, S., Baas, M., Hopmans, E.C., Sinninghe Damsté, J.S. (2008a) An unusual  
668 isoprenoid tetraether lipid in marine and lacustrine sediments. *Org. Geochem.* **39**,  
669 1033–1038.
- 670 Schouten, S., Baas, M., Hopmans, E.C., Reysenbach, A.-L., Sinninghe Damsté, J.S.  
671 (2008b) Tetraether membrane lipids of *Candidatus “Aciduliprofundum boonei”*, a  
672 cultivated obligate thermoacidophilic euryarchaeote from deep-sea hydrothermal  
673 vents. *Extremophiles* **12**, 119–124.
- 674 Schouten, S., Hopmans, E.C., Sinninghe Damsté, J.S. (2013) The organic geochemistry  
675 of glycerol dialkyl glycerol tetraether lipids: A review. *Org. Geochem.* **54**, 19–61.

- 676 Stickland J.D.H., Parsons T.R.A. (1972) A Practical Handbook of Seawater Analysis.,  
677 Fisheries Research Board of Canada, Ottawa.
- 678 Sturt H. F., Summons R. E., Smith K., Elvert M., Hinrichs K.-U. (2004) Intact polar  
679 membrane lipids in prokaryotes and sediments deciphered by high-performance  
680 liquid chromatography/electrospray ionization multistage mass spectrometry - new  
681 biomarkers for biogeochemistry and microbial ecology. *Rapid Commun. Mass  
682 Spectrom.* **18**, 617–628.
- 683 Sutton, P.A., Rowland, S.J. (2012) High temperature gas chromatography–time-of-flight-  
684 mass spectrometry (HTGC–ToF-MS) for high-boiling compounds. *J. Chromatogr.  
685 A.* **1243**, 69–80.
- 686 Sutton, P.A., Rowland, S.J. (2014) Determination of the content of C<sub>80</sub> tetraacids in  
687 petroleum. *Energy Fuels* **28**, 5657–5669.
- 688 Teske, A., Callaghan, A.V., LaRowe, D.E. (2014) Biosphere frontiers of subsurface life  
689 in the sedimented hydrothermal system of Guaymas Basin. *Front. Microbiol.* **5**,  
690 doi:10.3389/fmicb.2014.00362
- 691 Thiel, V., Peckmann, J., Seifert, R., Wehrung, P., Reitner, J., Michaelis, W. (1999)  
692 Highly isotopically depleted isoprenoids: molecular markers for ancient methane  
693 venting. *Geochim. Cosmochim. Acta* **63**, 3959–3966.
- 694 Weijers, J.W.H., Schouten, S., Hopmans, E.C., Geenevasen, J.A.J., David, O.R.P.,  
695 Coleman, J.M., Pancost, R.D., Sinninghe Damsté, J.S. (2006) Membrane lipids of  
696 mesophilic anaerobic bacteria thriving in peats have typical archaeal traits. *Environ.  
697 Microbiol.* **8**, 648–657.
- 698 Weijers, J.W.H., Schouten, S., van den Donker, J.C., Hopmans, E.C., Sinninghe Damsté,  
699 J.S. (2007) Environmental controls on bacterial tetraether membrane lipid  
700 distribution in soils. *Geochim. Cosmochim. Acta* **71**, 703–713.
- 701 Wuchter, C., Schouten, S., Wakeham, S.G., Sinninghe Damsté, J.S. (2005) Temporal and  
702 spatial variation in tetraether membrane lipids of marine Crenarchaeota in  
703 particulate organic matter: Implications for TEX<sub>86</sub> paleothermometry.  
704 *Paleoceanography* **20**, doi:10.1029/2004PA001110
- 705 Xie, S., Lipp, J.S., Wegener, G., Ferdelman, T.G., Hinrichs, K.-U. (2013) Turnover of  
706 microbial lipids in the deep biosphere and growth of benthic archaeal populations.  
707 *Proc. Natl. Acad. Sci. USA* **110**, 6010–6014.
- 708 Yang H., Pancost R.D., Tang C., Ding W., Dang X., Xie S. (2014) Distributions of  
709 isoprenoid and branched glycerol dialkanol diethers in Chinese surface soils and a  
710 loess–paleosol sequence: Implications for the degradation of tetraether lipids. *Org.  
711 Geochem.* **66**, 70–79.

712 Yoshinaga, M.Y., Lazar, C.S., Elvert, M., Lin, Y-S., Zhu, C., Heuer, V.B., Teske, A.,  
713 Hinrichs, K.-U. (2015) Possible roles of uncultured archaea in carbon cycling in  
714 methane-seep sediments. *Geochim. Cosmochim. Acta*, **164**, 35-52.

715 Zhu, R., Evans, T.W., Wörmer, L., Lin, Y.-S., Zhu, C., Hinrichs, K.-U. (2013a) Improved  
716 sensitivity of sedimentary phospholipid analysis resulting from a novel extract  
717 cleanup strategy. *Org. Geochem.* **65**, 46–52.

718 Zhu C., Lipp J.S., Wörmer L., Becker K.W., Schröder J., Hinrichs K.-U. (2013b)  
719 Comprehensive glycerol ether lipid fingerprints through a novel reversed phase liquid  
720 chromatography–mass spectrometry protocol. *Org. Geochem.* **65**, 53–62.

721

722

723

724

Compounds	Abbreviation	Structural illustration	Distribution				
			Leg201-1227	Marmorito seep carbonate	Guaymas Basin 4568	Hot spring T-15	<i>N. maritimus</i> late stat. phase
glycerol dialkyl glycerol tetraether	GDGT		+	+	+	+	+
glycerol dialkyl glycerol triether alcohol	GDGTol		+	+	+	n.d.	+
glycerol dialkyl glycerol triether acid	GDGTA		+	+	n.d.	n.d.	n.d.
glycerol dibiphytanol diether	GDD		+	+	+	+	+
glycerol dibiphytanol diether monoacid	GDDA		+	+	+	n.d.	+
glycerol dibiphytanol diether diacid	GDDAA		n.d.	+	n.d.	n.d.	n.d.
glycerol monobiphytanyl glycerol diether	GMGD		+	+	+	n.d.	n.d.
glycerol monobiphytanol monoether	GMM		+	+	+	n.d.	n.d.
glycerol monobiphytanol monoether acid	GMMA		n.d.	+	n.d.	n.d.	n.d.
biphytanic diol	bpdiol		+	+	+	+	+
biphytanic monoacid	bpmonoacid/ol		+	+	n.d.	n.d.	n.d.
biphytanic diacid	bpdiacid		n.d.	+	n.d.	n.d.	n.d.
H-shaped glycerol monoalkyl glycerol tetraether	H-GMGT		+	+	+	+	n.d.
H-shaped glycerol monoalkyl diether	H-GMD		+	+	+	+	n.d.
H-shaped C <sub>80</sub> tetrol	H-tetrol		+	+	+	+	n.d.
H-shaped C <sub>80</sub> monoacid	H-monoacid		n.d.	n.d.	+	n.d.	n.d.
H-shaped C <sub>80</sub> diacid	H-diacid		n.d.	n.d.	+	n.d.	n.d.
H-shaped C <sub>80</sub> triacid	H-triacid		n.d.	n.d.	+	n.d.	n.d.
H-shaped C <sub>80</sub> tetraacid	H-tetraacid		n.d.	n.d.	+	n.d.	n.d.

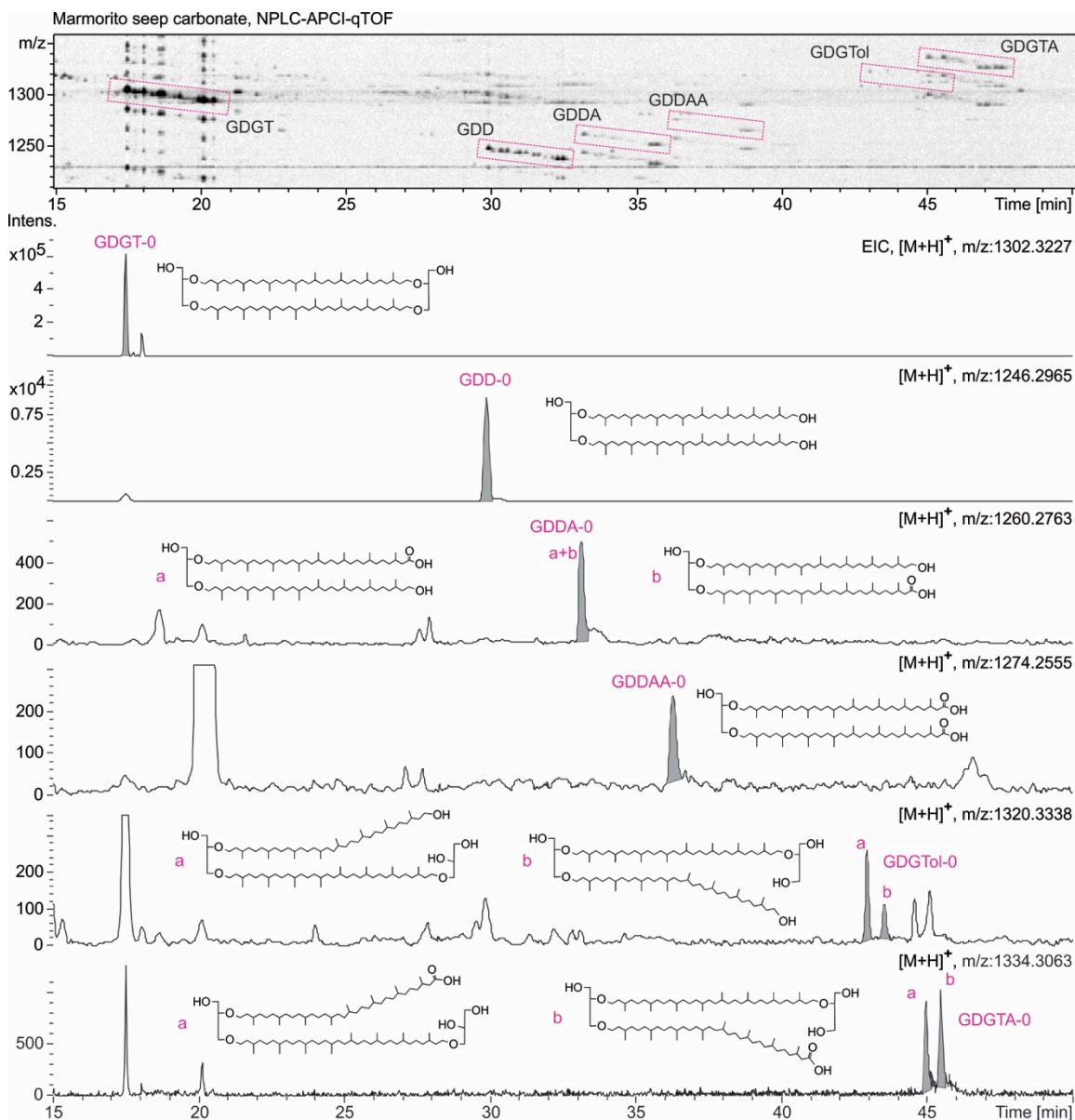


726 **Table 1.** Compound classes discussed in the paper. Illustrated structures show only the  
727 acyclic biphytane derivatives. Constitutional isomers are not included. Representative  
728 samples showing the distribution of GDGT and degradation products in various  
729 environment settings and cell. '+' compound detected, 'n.d.' compound not detected.

730

731

732



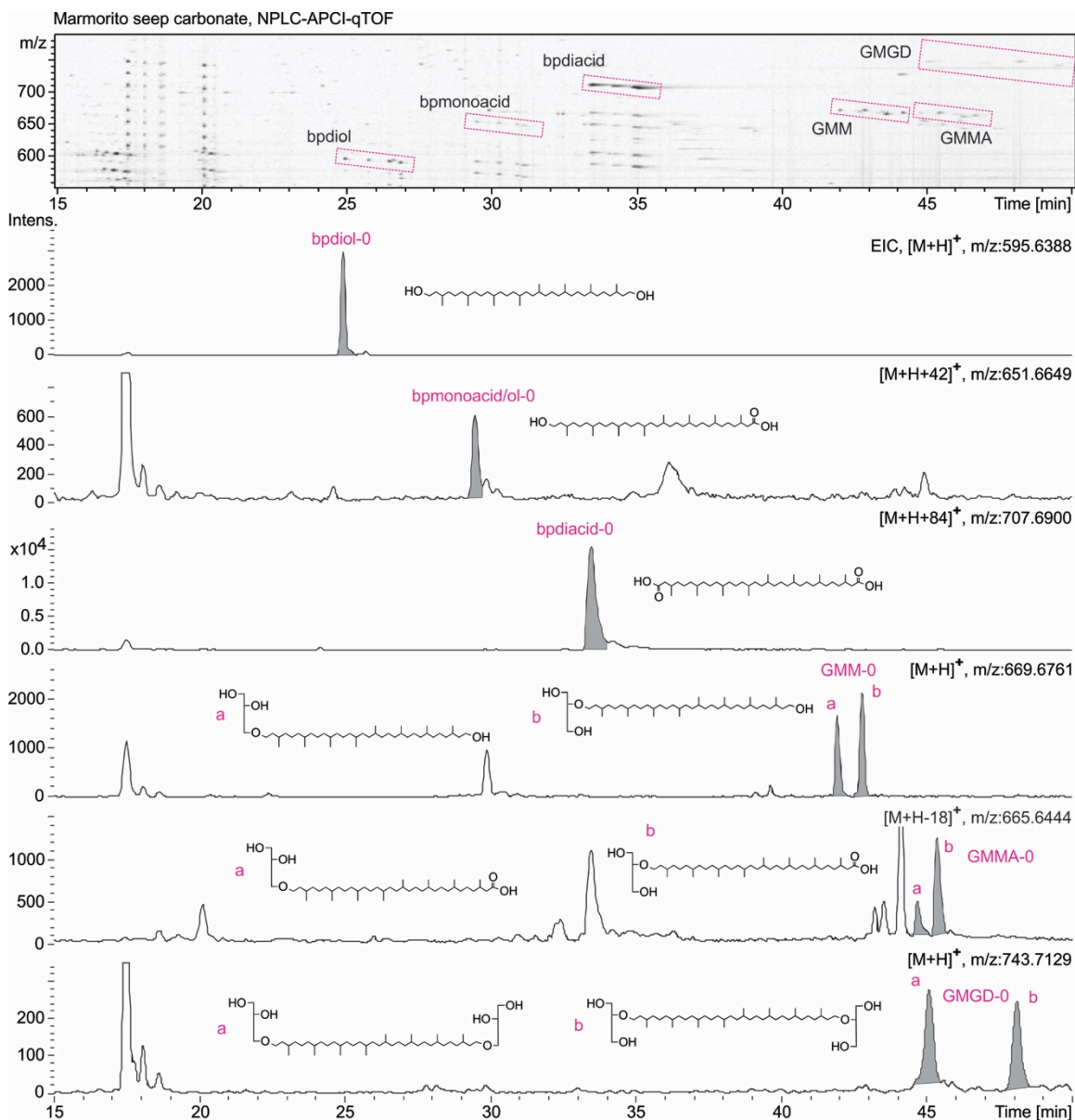
733

734 **Fig. 1a**

735 Density maps and extracted ion chromatograms (EICs) of NPLC-APCI-qTOF showing  
 736 the detection of isoprenoidal GDGT-0 and its hydroxylated and carboxylated derivatives  
 737 with two (Fig. 1a) or one biphytane units (Fig. 1b) in the Marmorito seep carbonate.

738 Multiple isomers were observed for GDDA, GDGTol, GDGTA, GMM and GMGD and  
 739 are labeled as 'a' and 'b'. The isomeric composition of GMGD, 'a' and 'b', may provide  
 740 insights regarding the regioisomerism of GDGTs and will be subject of a future report.

741



742

743

744 Fig. 1b

745 Density maps and EICs of NPLC-APCI-qTOF showing the detection of isoprenoidal

746 GDGT-0 and hydroxylated and carboxylated derivatives with two (Fig. 1a) or one,

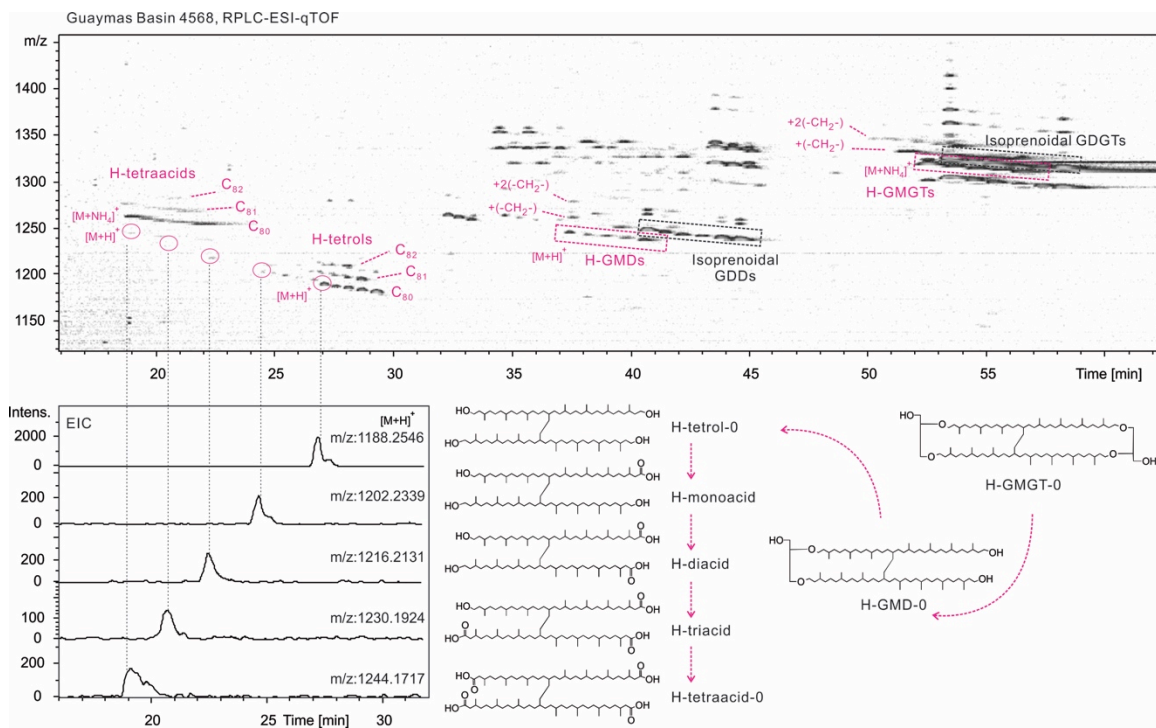
747 biphytane units (Fig. 1b) in the Marmorito seep carbonate. Multiple isomers were

748 observed for GDDA, GDGTol, GDGTA, GMM and GMGD and are labeled as 'a' and

749 'b'. The isomeric composition of GMGD, 'a' and 'b', may provide insights regarding the

750 regioisomerism of GDGTs.

751



752

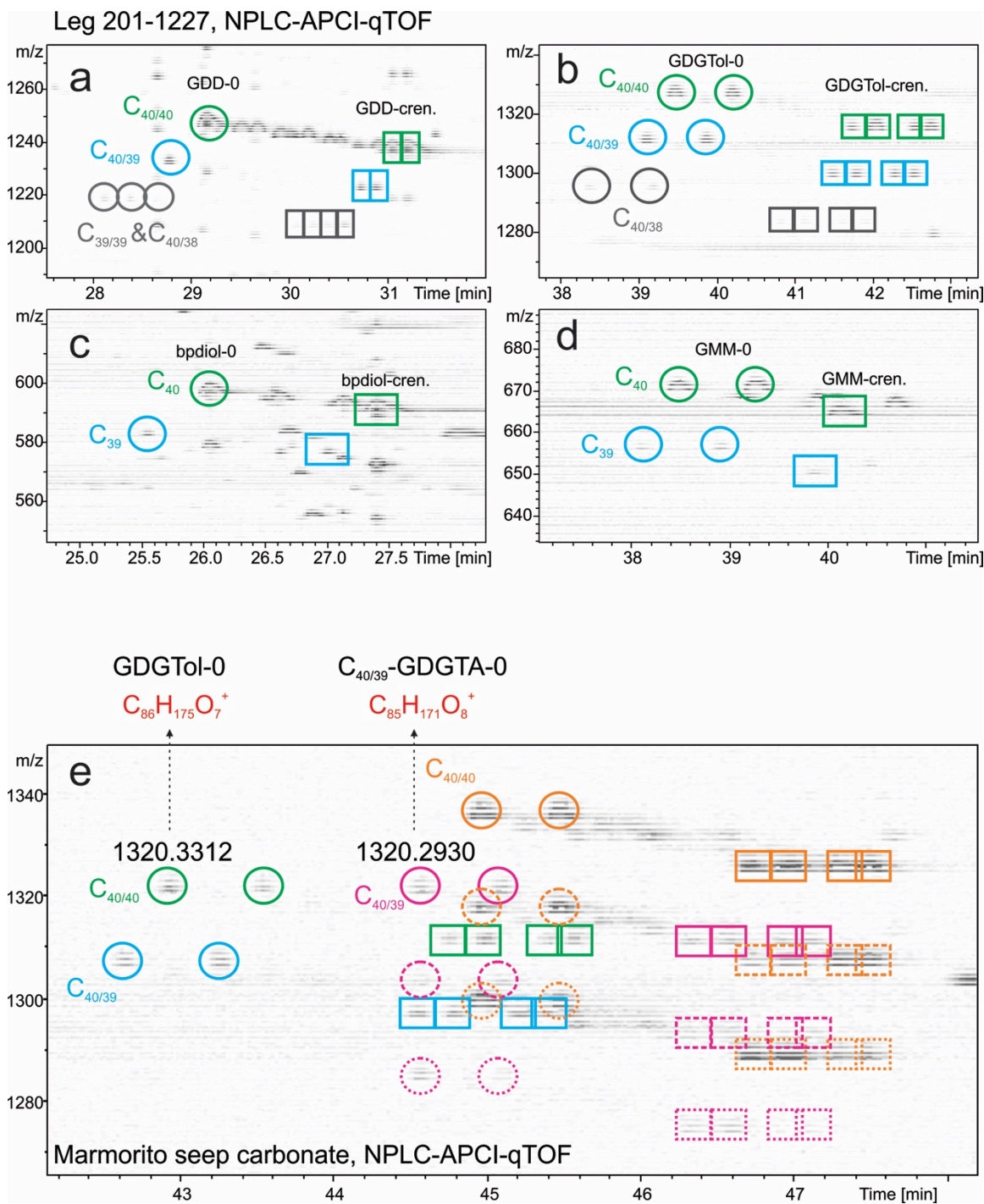
753

754 Fig. 2 Density maps and EICs derived from RPLC-ESI-qTOF analysis showing the  
 755 detection of H-GMGs and their hydroxylated and carboxylated derivatives in a  
 756 hydrothermal sediment from the Guaymas Basin. Stepwise oxidation of the four hydroxyl  
 757 groups in the H-tetrols possibly generated C<sub>80</sub> mono-, di-, tri- and H-tetraacids. Mono-  
 758 and di-methylated H-GMGs, H-GMDs and corresponding C<sub>81</sub>, C<sub>82</sub> H-tetrols and H-  
 759 tetraacids were also detected. Molecular structures of the intermediates are suggested.  
 760 Isomers of monoacid, diacid and triacid may exist but could not be separated with the  
 761 present LC methods.

762

763

764



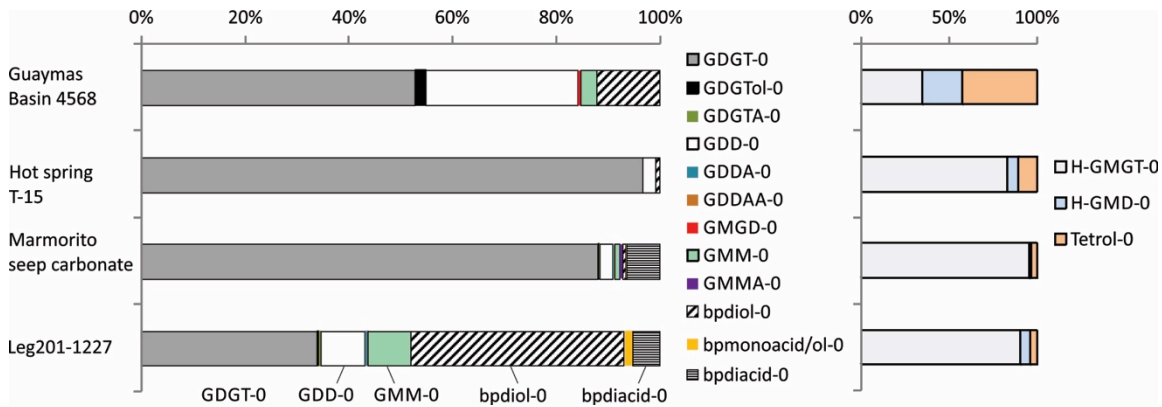
765

766 Fig. 3 Density maps of NPLC-APCI-qTOF showing in one marine sediment (Leg 201-  
 767 1227) the detection of GDD, GDGTol, biphytane diol, GMM and their co-occurring  
 768 further degradation derivatives with  $C_{39}$  and  $C_{38}$  based isoprenoids, (a-d). Highlighted are  
 769 the most dominant components including the GDGT-0 (open circle) and crenarchaeol-  
 770 related (rectangle) compounds. Different compound classes are also color coded.  $C_{40}$

771 hydroxyl derivatives are in green, C<sub>39</sub> in blue and the C<sub>38</sub> related in grey, C<sub>40</sub> carboxyl  
772 derivatives are in orange and C<sub>39</sub> in pink. Multiple isomers occur due to different ether  
773 bonding and ring arrangements. The occurrence of C<sub>40/39</sub>-GDGTA was shown in one  
774 sample of Marmorito seep carbonate (e). The molecular ion of C<sub>40/39</sub>-GDGTA-0 and  
775 GDGTol-0 gives similar but distinguishable masses, m/z: 1320.2930 and 1320.3312,  
776 respectively. Under APCI, dehydrated ions of carboxyl derivatives occurred (highlighted  
777 ions with dashed lines).

778

779



780

781

782 Fig. 4 Relative abundances of diagenetic derivatives of GDGT-0 (left panel) and H-

783 GMGT-0 (right panel) in four representative environmental samples.

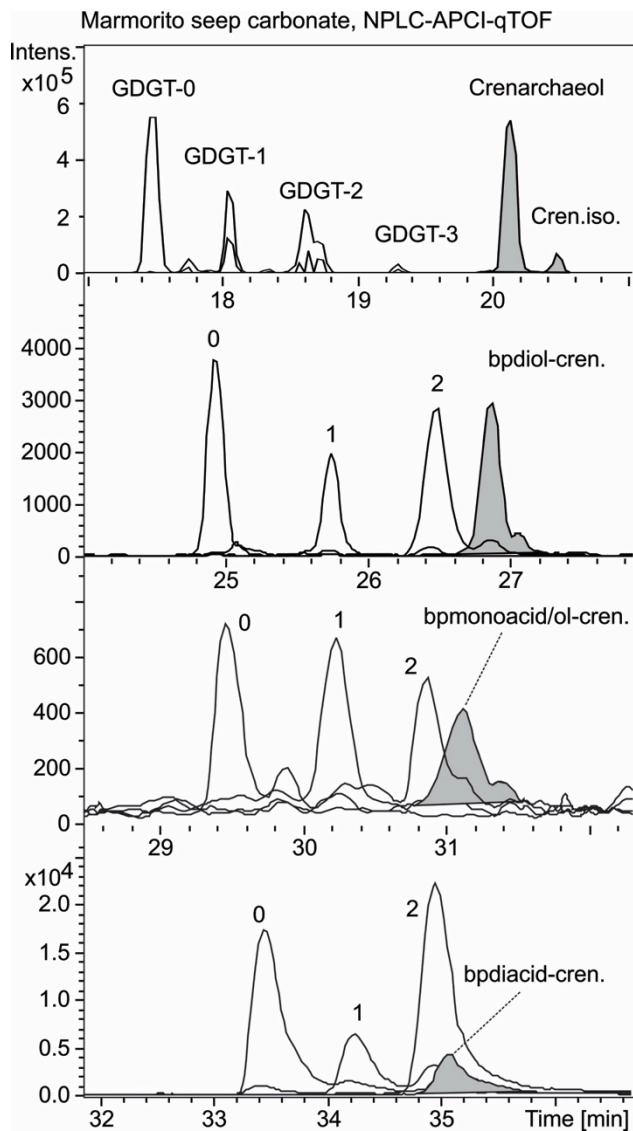
784

785

786

787



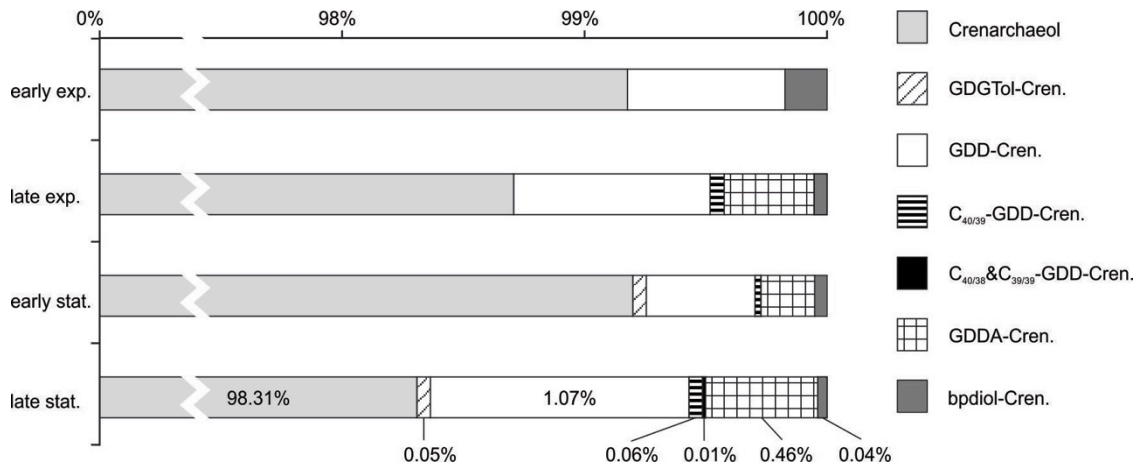


788

789 Fig. 5 Extracted ion chromatograms from NPLC-APCI-qTOF analysis, showing the  
 790 distribution of GDGT, bpdial, bpmonoacid/ol and bpdiaacids in a seep carbonate,  
 791 Marmorito. Peaks of crenarchaeol and its related derivatives are shaded; crenarchaeol  
 792 derivatives have decreased abundances in the acid fractions.  
 793



794



795

796

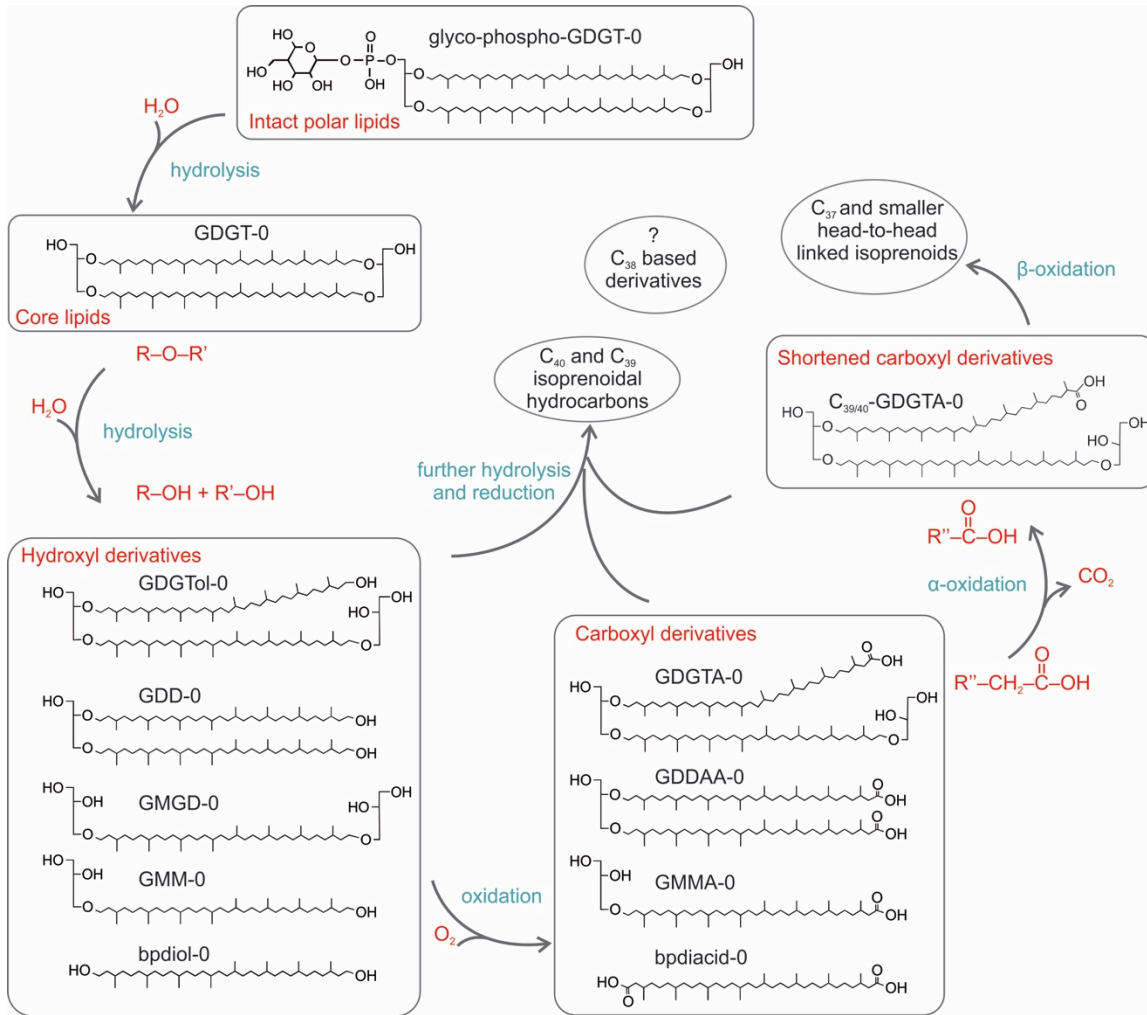
797

798 Fig. 6 The relative abundance of core crenarchaeol and its related hydroxyl, carboxyl and

799 C<sub>39</sub> and C<sub>38</sub> based derivatives in different growth phases of *N. maritimus*.

800

801



802

803 Fig. 7 Hypothetical scheme showing the suggested diagenetic pathways of GDGTs.

804 GDGT-0 and its related IPL and degradation derivatives are used as an example. In the

805 flow chart, R, R' and R'' represent different alkyl chains. Intermediate components, such

806 as GDDA, bpdmonoacid/ol are not included. The formation of C<sub>38</sub> based isoprenoids is not

807 clear and thus is labeled with '?'.

808

809

810 **Supplementary information**

811

812 **Sample preparation**

813

814 **1. Hot spring sediment T-15**

815 Sediment was collected from the bottom of a geothermal wells located in Ruidian (Jinze  
816 hot spring: N 25.44138°, E98.46004°, Elevation:1740m. Tengchong County, Yunnan  
817 Province, China), with the water pH of 6.71 and temperature of 80.6 °C. Roughly 5 g of  
818 freeze-dried sediment was extracted in an ultrasonic bath with methanol (MeOH; twice),  
819 methanol-dichloromethane (DCM; 1:1, v/v; twice) and finally DCM (twice). All  
820 supernatants were collected in a flask and completely dried under N<sub>2</sub>. The total lipid  
821 extract was fractionated over a pre-activated silica gel chromatography into apolar  
822 (alkane) and polar (intact polar GDGTs and GDGT core lipids) fractions using *n*-hexane-  
823 DCM (9:1, v/v) and DCM-MeOH (1:1, v/v) as eluents respectively. The polar fraction  
824 was used for the lipid analysis in this study.

825

826 **2. H-tetrol standards**

827 An oilfield calcium naphthenate deposit was Soxhlet extracted sequentially  
828 (dichloromethane/*n*-hexane, 1/1, v/v, 300 mL, 8h; toluene/acetone, 3/2, v/v, 250 mL, 6h;  
829 propan-2-ol/dichloromethane, 4/1, v/v, 250mL, 6h) to remove interstitial oil and the  
830 residue dried (overnight, 70°C) before acidification by heating (70°C, overnight) in a  
831 capped and sealed vial with hydrochloric acid (3 mL) and cyclohexane (3 mL). The  
832 acidified organic fraction was recovered by extraction into diethylether (DEE; 3 x 5 mL)  
833 by mixing (vortex 10 s) and centrifugation (2500 rpm/3 min). Solvent was removed from

834 the decanted supernatant under blow-down ( $N_2$ ,  $70^\circ C$ ) prior to dilution in DEE/0.1%  
835 ammonia (8 mL), loading on a pre-conditioned (1% aqueous ammonia, 20 mL; water, 20  
836 mL; DEE, 10 mL) SAX solid phase extraction cartridge (Sigma-Aldrich Company Ltd.,  
837 Dorset, UK; DSC-SAX, 12 mL, 2 g) and sequentially eluting the cartridge with 20 mL  
838 volumes of DEE, dichloromethane and DEE/2% formic acid (FA). Solvent was removed  
839 from the DEE/2% FA fraction ('acid fraction') under blow-down ( $N_2$ ,  $70^\circ C$ ) and an  
840 aliquot per-trimethylsilylated with BSTFA/1% TMCS (Sigma-Aldrich Company Ltd.,  
841 UK; ca. 50  $\mu L$ ,  $70^\circ C$ , 1 hr) before dilution in cyclohexane for analysis using high  
842 temperature gas chromatography (HTGC; Fig. SX1), or per-methylation for infusion  
843 electrospray ionization/mass spectrometry (ESI/MS; Fig. SX2). An aliquot of the isolated  
844 acid fraction was diluted in DEE for infrared spectroscopy (Fig. SX3).

845

846 The isolated acid fraction was reduced to alcohols by treatment with lithium aluminum  
847 hydride (LithAl; Sigma-Aldrich Company Ltd., UK; 1M in DEE). Dried acid fraction  
848 was dissolved in a small volume (ca. 1 mL) dry DEE (sodium wire) and transferred with  
849 washings to a small three-neck flask with Teflon flea and fitted with a Subaseal, calcium  
850 chloride drying tube condenser and a glass stopper, all over a magnetic stirring block.  
851 LithAl solution (ca. 3 mL) was pumped into the flask under nitrogen through the  
852 Subaseal whilst stirring, additional dry DEE (ca. 5mL) was added through the condenser  
853 to break up formed solids, before placing a bowl of warm water under the flask for 10  
854 minutes. After the reaction vessel had cooled, wet DEE (prepared by mixing DEE and  
855 water in separating funnel and drawing off aqueous phase) was added (10 mL, drop-wise  
856 at first) to hydrolyse unreacted hydrides, followed by drop-wise addition of sulphuric

857 acid (10% aqueous; ca. 1 mL) to destroy remaining LithAl. The contents of the flask  
858 were transferred with washings (ca. 3 mL, 10% H<sub>2</sub>SO<sub>4(aq)</sub>) to a separating funnel and the  
859 organic phase ('alcohol fraction') collected after washing with water (3 x 3 mL). Solvent  
860 was removed from the alcohol fraction by blow-down (N<sub>2</sub>, 70°C) and aliquots prepared  
861 for HTGC, ESI/MS and IR spectroscopy as above. All solvents were HPLC grade  
862 (Rathburn Chemicals Ltd., Walkerburn, UK or Fisher Scientific UK Ltd., Loughborough,  
863 UK) or LC/MS grade (Chromasolv®; Sigma-Aldrich Company Ltd., Dorset, UK), water  
864 was Elga Maxima (18.2 mΩ; Elga Ltd., Buckinghamshire, UK).

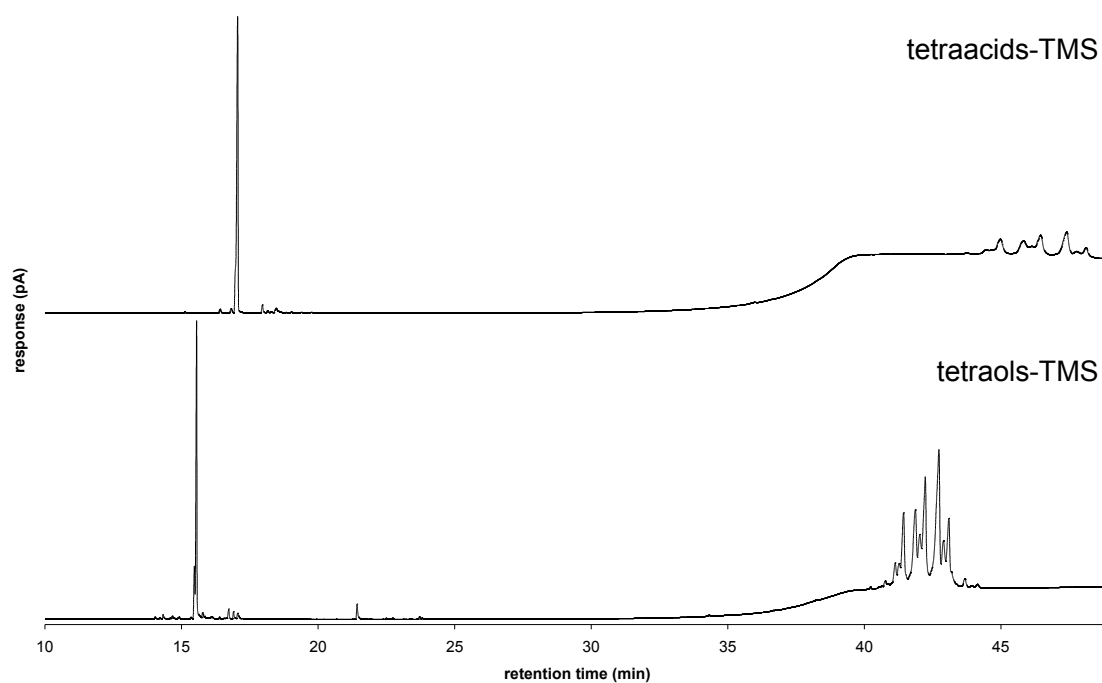
865

866 Reduction of the tetraacids to tetrols was monitored using HTGC (Fig. SX1), ESI/MS  
867 (Fig. SX4 and SX5) and IR spectroscopy (Fig. SX3). The HTGC system comprised an  
868 Agilent 6890 GC fitted with cool-on-column inlet (+3°C track oven mode; 0.5 µL manual  
869 injection), flame ionization detector (435°C) and Varian VF-5ht Ultimet column (15 m  
870 x 0.25 mm x 0.1 µm) with helium carrier gas (1 ml min<sup>-1</sup>, constant flow) and oven  
871 programme from 40 – 430°C at 10°C min<sup>-1</sup> with 10 min hold.

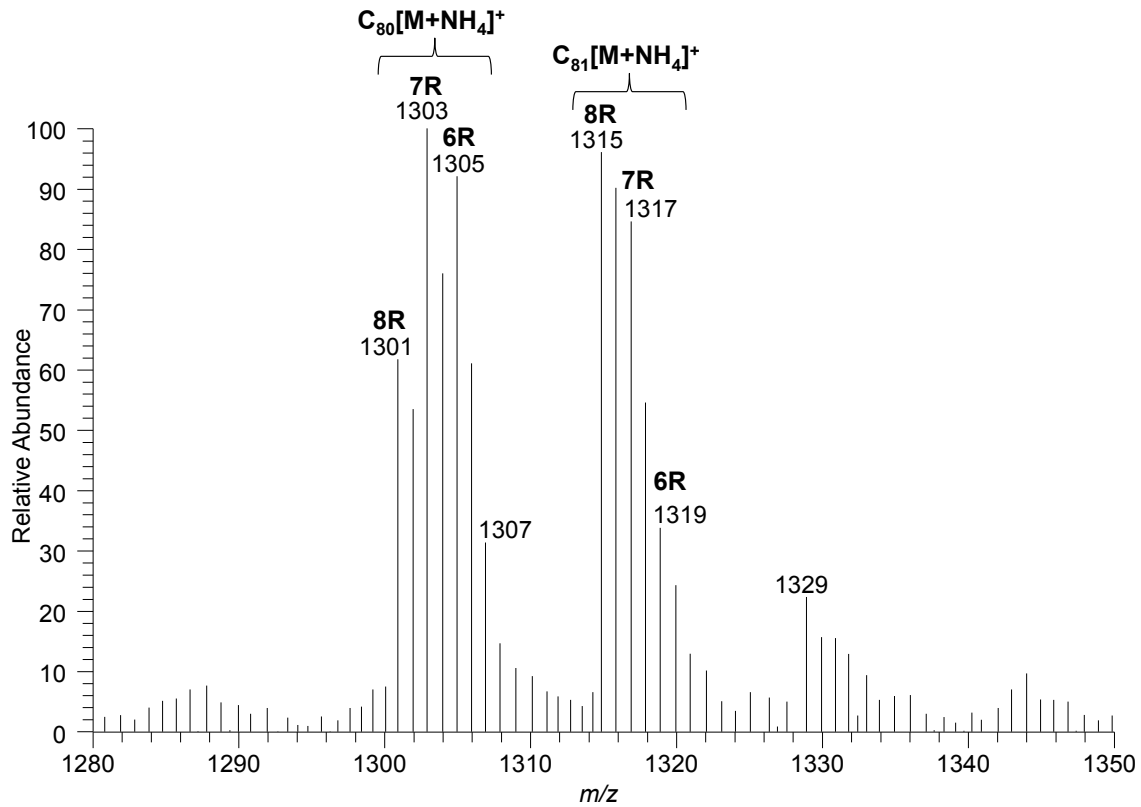
872

873 Infusion ESI/MS was carried out in positive ionization mode using a Finnigan Mat  
874 LCQ™ (ThermoFinnigan, San Jose, CA, USA) with ESI interface. Samples were diluted  
875 in propan-2-ol/10mM ammonium acetate and infused at 3 µL min<sup>-1</sup> with a Hamilton  
876 (Reno, CA, USA) 1725N (250 µL) syringe using the built-in syringe pump. Mass spectral  
877 data were acquired (and averaged over 1 minute) and processed using Xcalibur software.  
878 Typical instrument parameters were: source voltage (±) 4.5 kV; capillary voltage (±) 60  
879 V; capillary temperature 200 °C; nitrogen sheath gas flow rate 24 (arbitrary units).

880 Instrumental parameters were optimised on the most abundant ion using the autotune  
881 function.  
882  
883 Infrared spectroscopy was carried out using a Bruker Alpha Platinum ATR (Bruker Optik  
884 GmbH, Ettlingen, Germany) by measuring 32 sample scans (resolution  $4\text{ cm}^{-1}$ ;  
885 transmittance) and recording data between  $4000 - 375\text{ cm}^{-1}$ . Background comprised 32  
886 scans without sample.  
887



888  
889 Figure SX1. High temperature gas chromatograms of trimethylsilylated tetraacids  
890 obtained from an oilfield deposit (upper) and lower the trimethylsilylated tetraols  
891 obtained following  $\text{LiAlH}_4$  reduction (lower).  
892



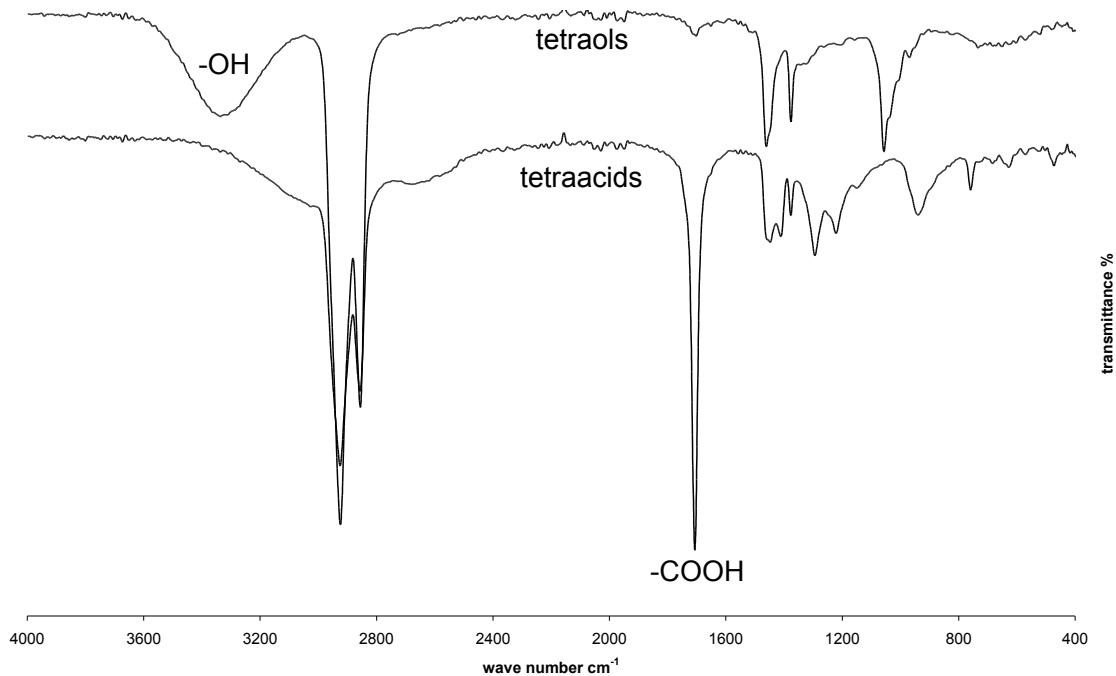
893

894 Figure SX2. Averaged mass spectrum from HPLC-electrospray ionization (+ve) mass  
 895 spectrometry of per-methylated esters of tetraacids obtained from an oilfield deposit (nR  
 896 refers to the number of cyclopentyl rings in the molecule)

897

898

899



900

901 Figure SX3. FTIR transmittance spectra of tetraacids isolated from an oilfield deposit

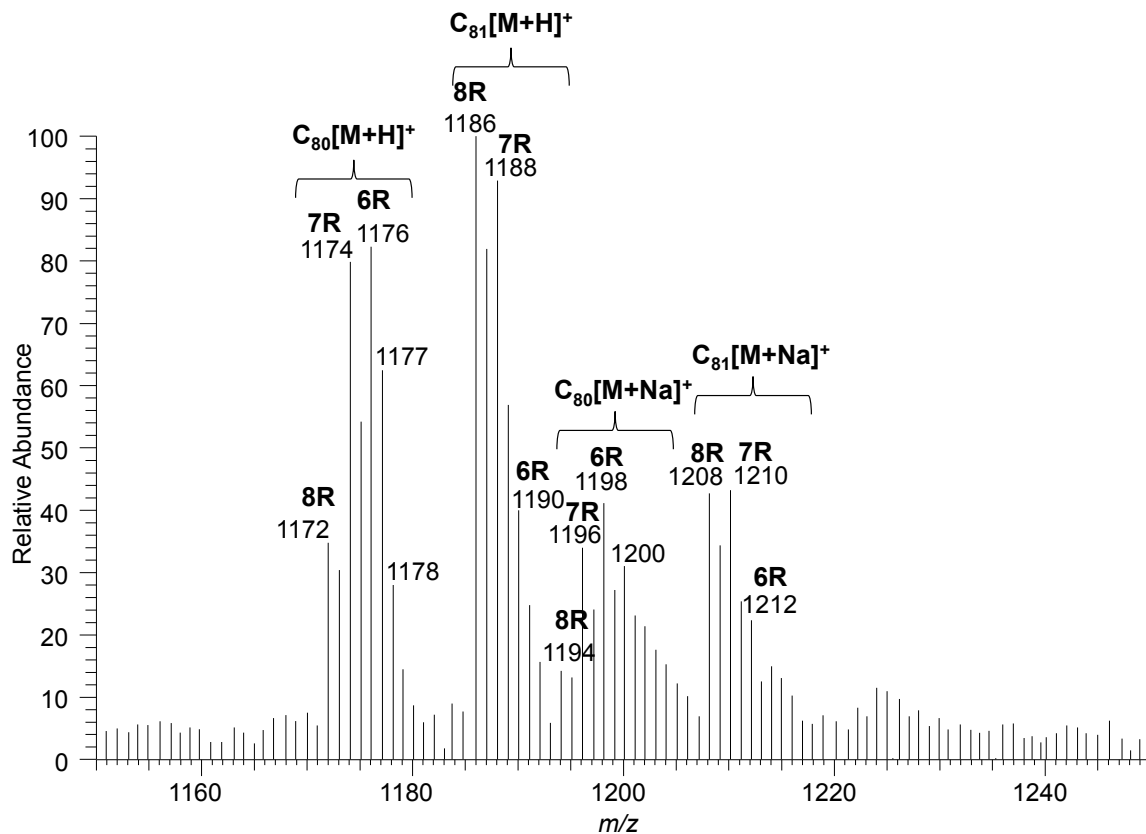
902 (lower) and tetraols produced from the tetraacids by LithAl reduction (upper).

903

904

905

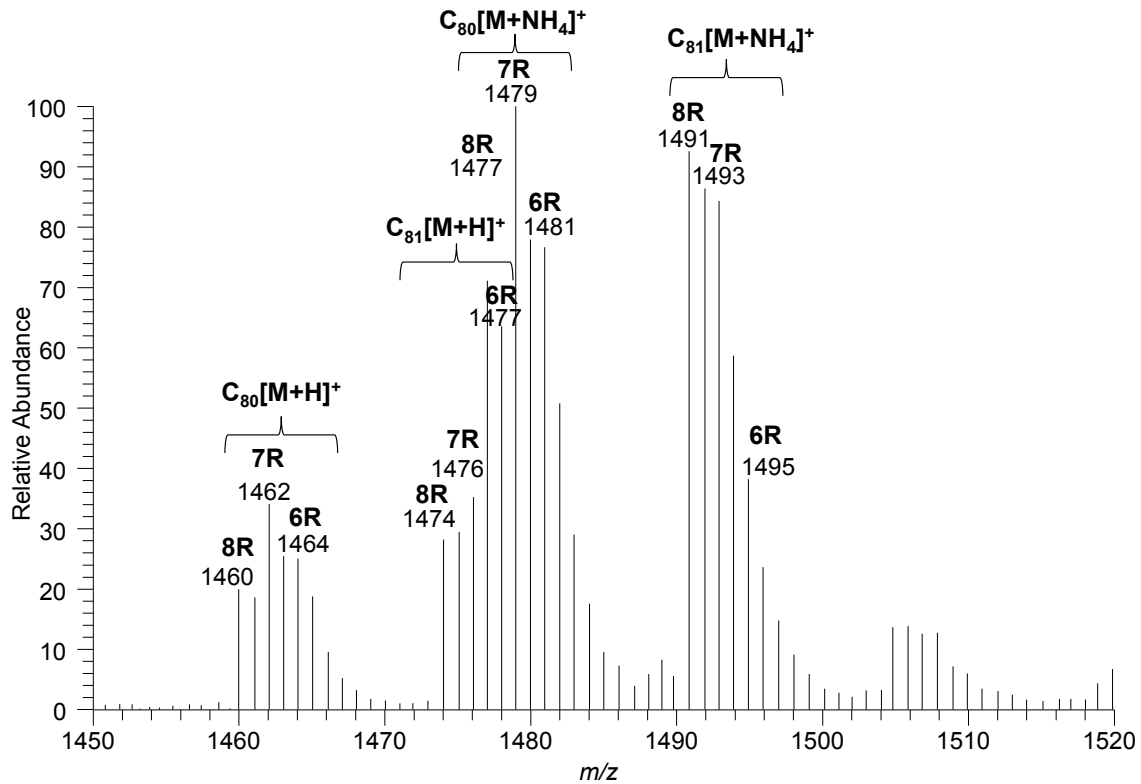




906

907 Figure SX4. Infusion electrospray ionization (+ve) mass spectrum of underivatized  
 908 tetraols obtained from the LithAl reduction of tetraacids from an oilfield deposit (nR  
 909 refers to the number of cyclopentyl rings in the molecule).

910



911

912 Figure SX5. Infusion electrospray ionization (+ve) mass spectrum of per-

913 trimethylsilylated tetraols obtained from the LithAl reduction of tetraacids from an

914 oilfield deposit (nR refers to the number of cyclopentyl rings in the molecule).

915

916

### 917 3. Acid hydrolysis of GDGT-0 standard

918 GDGT-0 standard was isolated with a semi-preparative LC protocol (as described in Zhu

919 et al., 2014) from acid-hydrolyzed biomass of *Archaeoglobus fulgidus*. Strong acid

920 hydrolysis was then performed with 10% methanolic HCl and GDGT-0 standard under

921 70 °C for 96 hours. After dried with a N<sub>2</sub> flow the treated sample was dissolved in *n*-

922 hexane for LC-APCI-MS analysis.

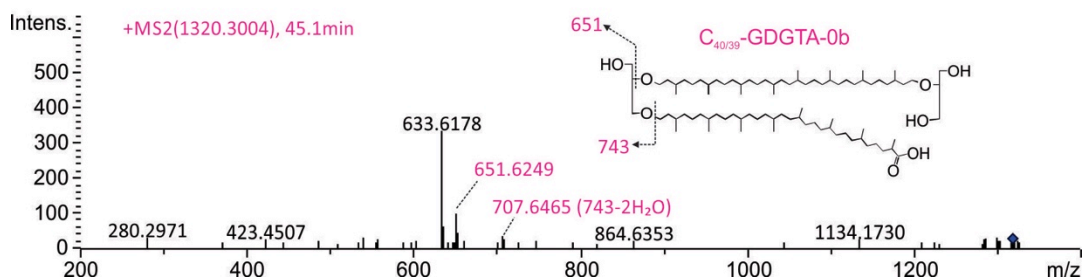
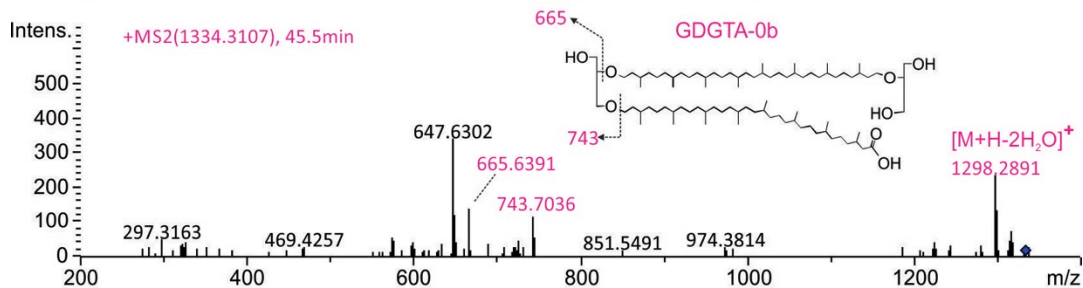
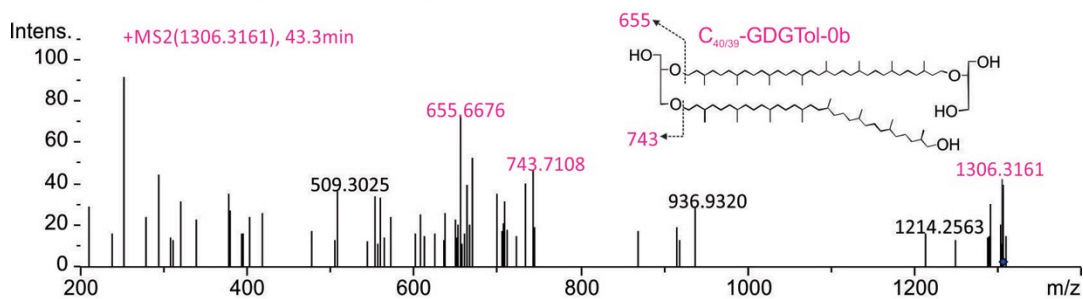
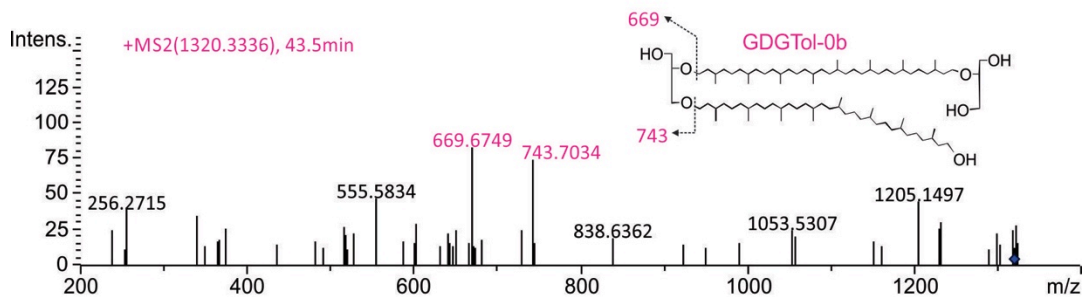
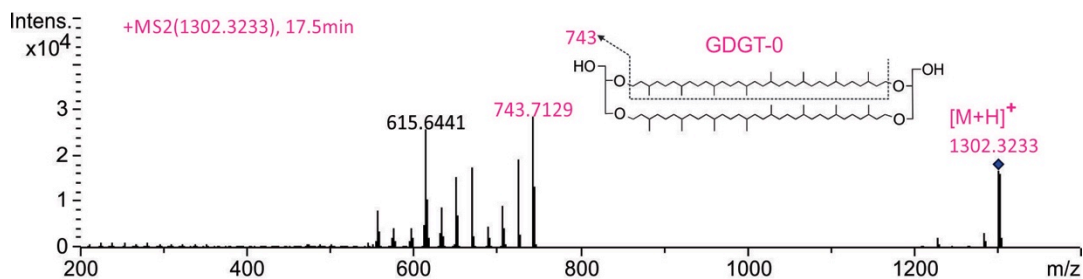
923

924

925 **Reference:**

926 Zhu, C., Meador, T.B., Dummann, W., Hinrichs, K.-U., 2014. Identification of unusual  
927 butanetriol dialkyl glycerol tetraether and pentanetriol dialkyl glycerol tetraether lipids in  
928 marine sediments. *Rapid Commun. Mass Spectrom.* 28, 332–338.

929

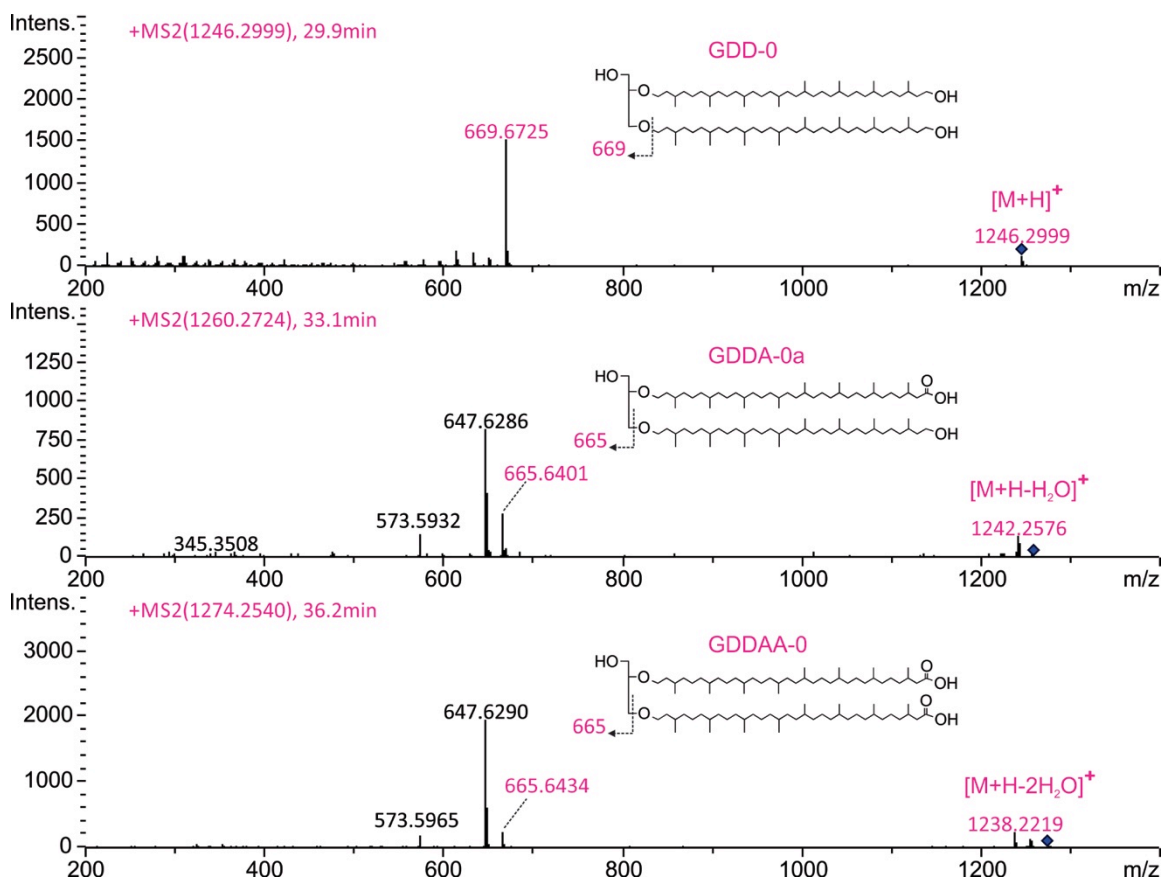


930

931

932 **Figure S1a.** MS<sup>2</sup> fragmentation patterns supporting the identification of GDGT-0,  
 933 GDGTol-0 and GDGTA-0 in Fig. 1a, and C<sub>40/39</sub>-GDGTol-0 and C<sub>40/39</sub>-GDGTA-0 in Fig.  
 934 3e.

935



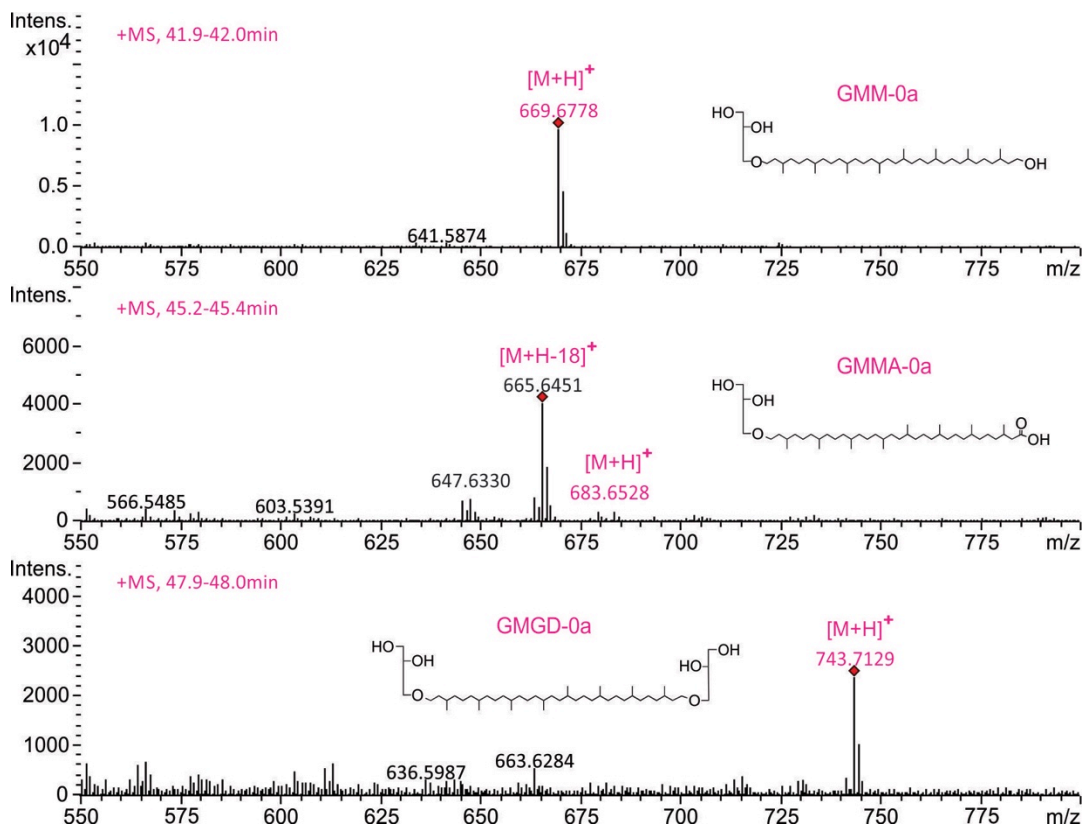
936

937

938 **Figure S1b.** MS<sup>2</sup> fragmentation patterns supporting the identification of GDD-0, GDDA-  
939 0 and GDDAA-0 in Fig. 1a.

940

941



942

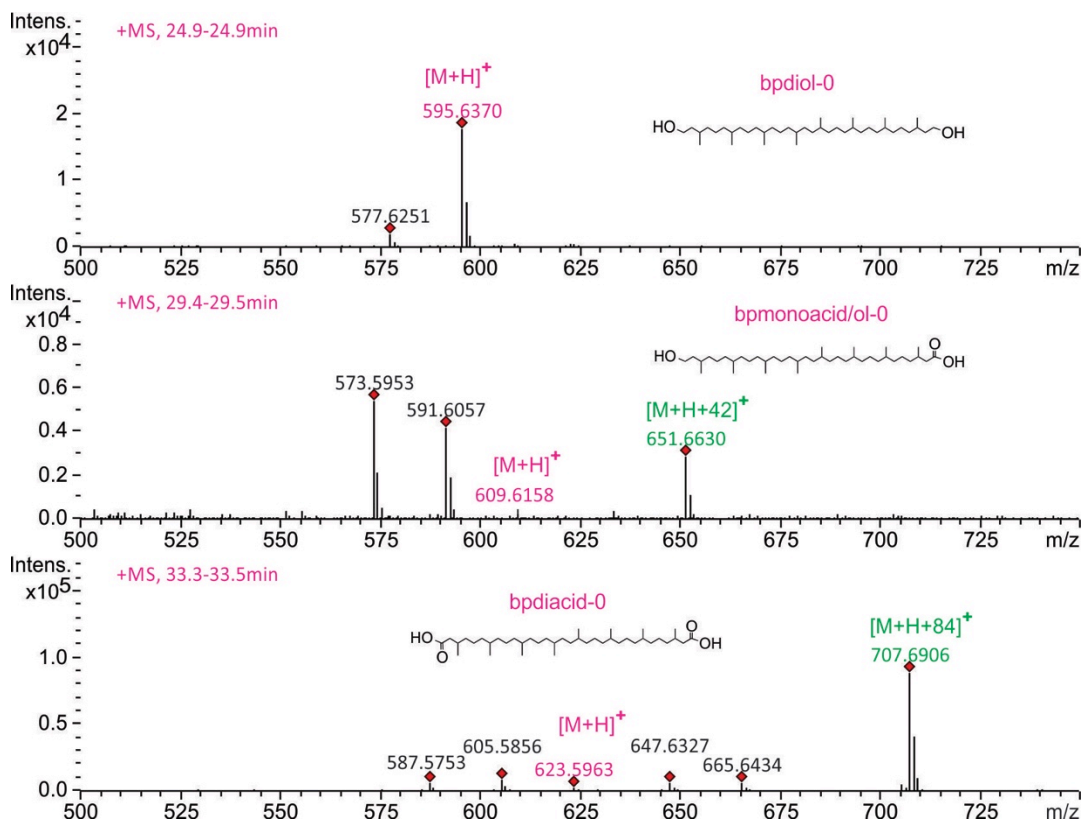
943

944 **Figure S1c.** MS<sup>1</sup> ions supporting the identification of GMM-0, GMMA-0 and GMGD-0  
 945 in Fig. 1b.

946

947

948



949

950 **Figure S1d.** MS<sup>1</sup> ions supporting the identification of bpdial-0, bpmmonoacid/ol-0 and  
 951 bpdiaacid-0 in Fig. 1b.

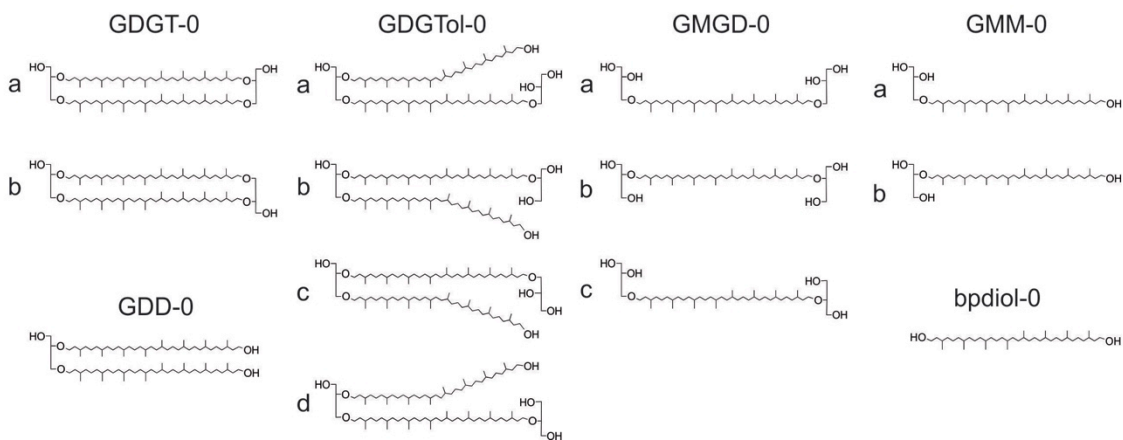
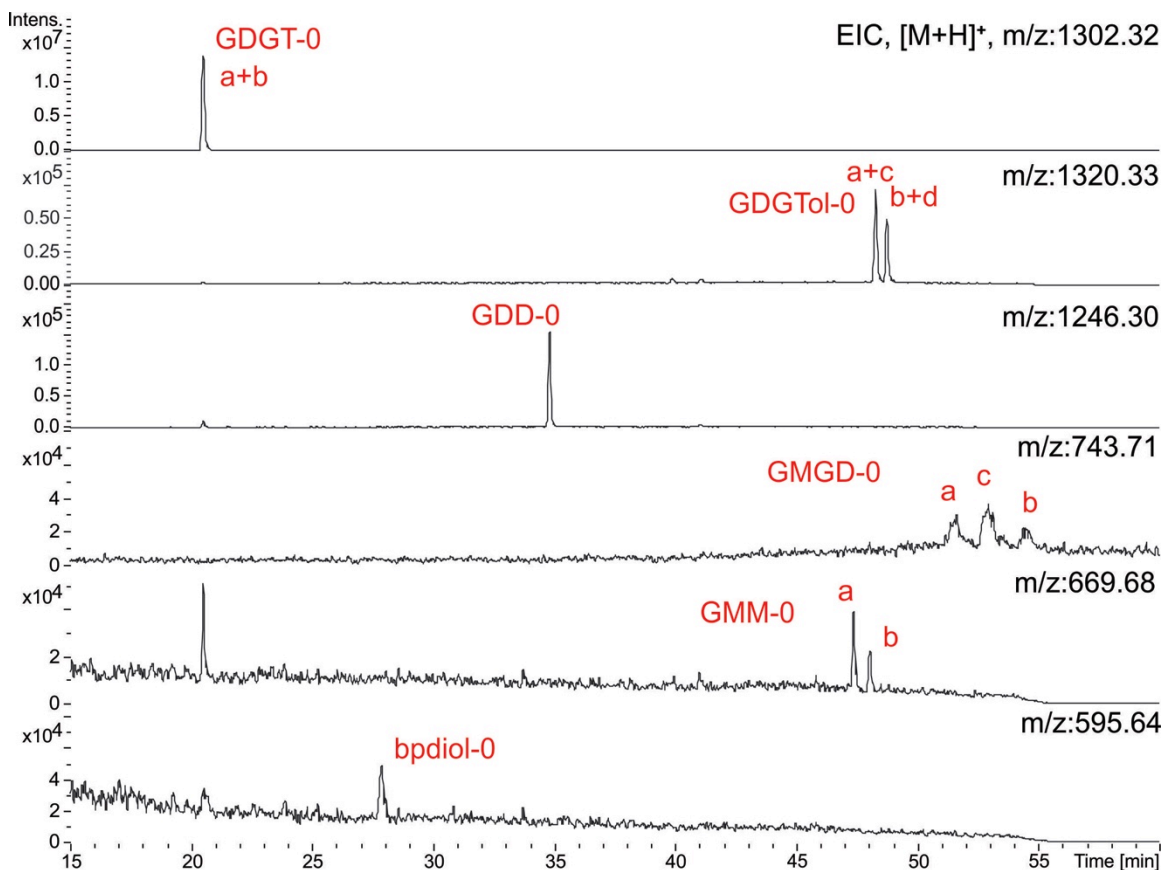
952

953

954

955

956



957

958

959 **Figure S2**

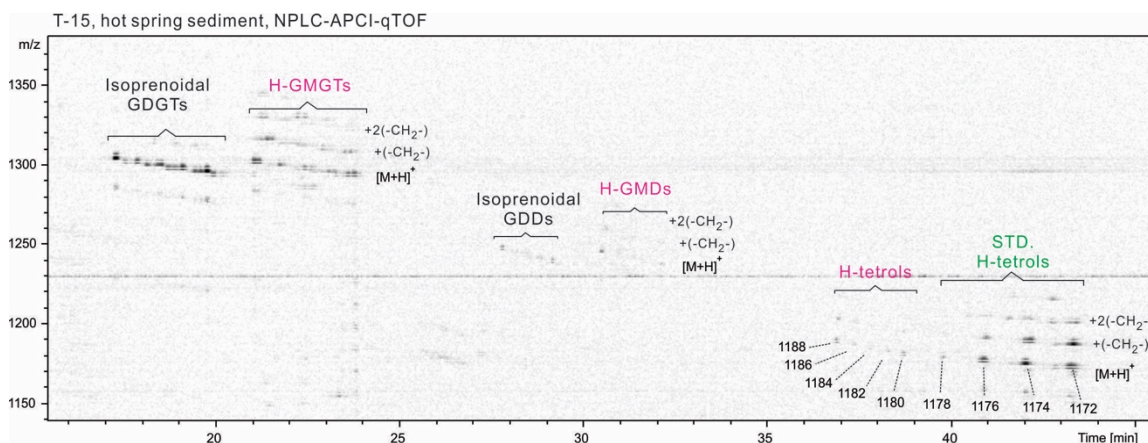
960 Major hydroxyl derivatives released from GDGT-0 by chemical degradation, a mild ether  
 961 cleavage conducted by adding 1000ng GDGT-0 into 1mL of 10% HCl in methanol, and



962 heated to 70 °C for 96 hours. The composition of GMGDs indicated a nearly 1:1 mixture  
963 of parallel and anti-parallel GDGT-0.

964

965



966

967

### 968 **Figure S3**

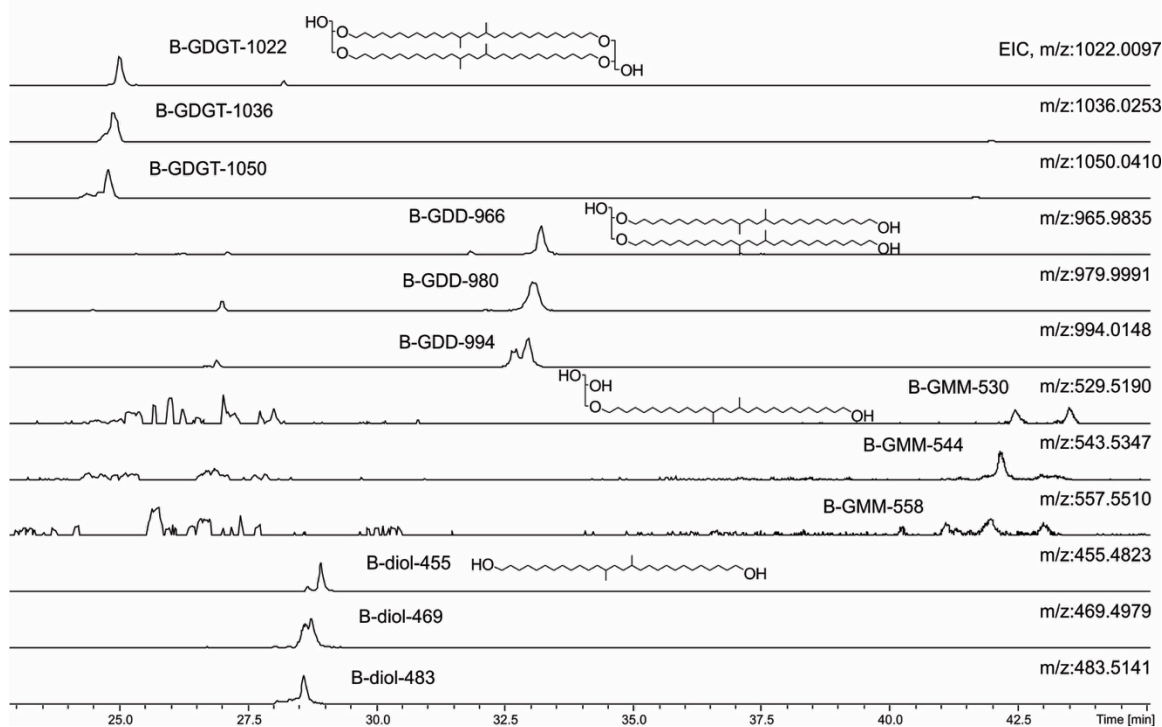
969 Density maps of NPLC-APCI-qTOF showing the occurrence of H- GDGTs and their  
970 degradation derivatives in the hot spring sediment, T-15. The identification of H-tetrols  
971 are confirmed by their similar chromatographic behavior with added standards (green  
972 color text).

973

974

975

Marmorito seep carbonate, NPLC-APCI-qTOF



976

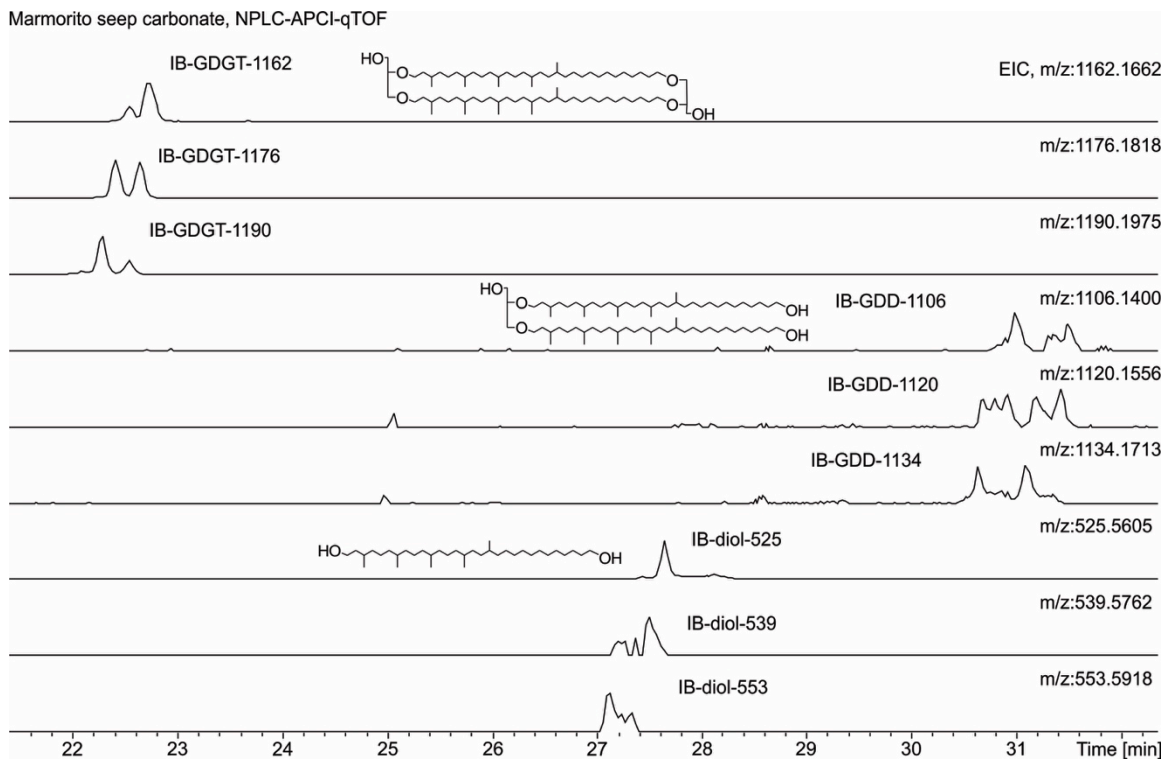
977 **Figure S4**

978 Extracted ion chromatograms of NPLC-APCI-qTOF showing the occurrence of branched  
979 GDGTs and derivatives in Marmorito seep carbonate.

980

981

982



983

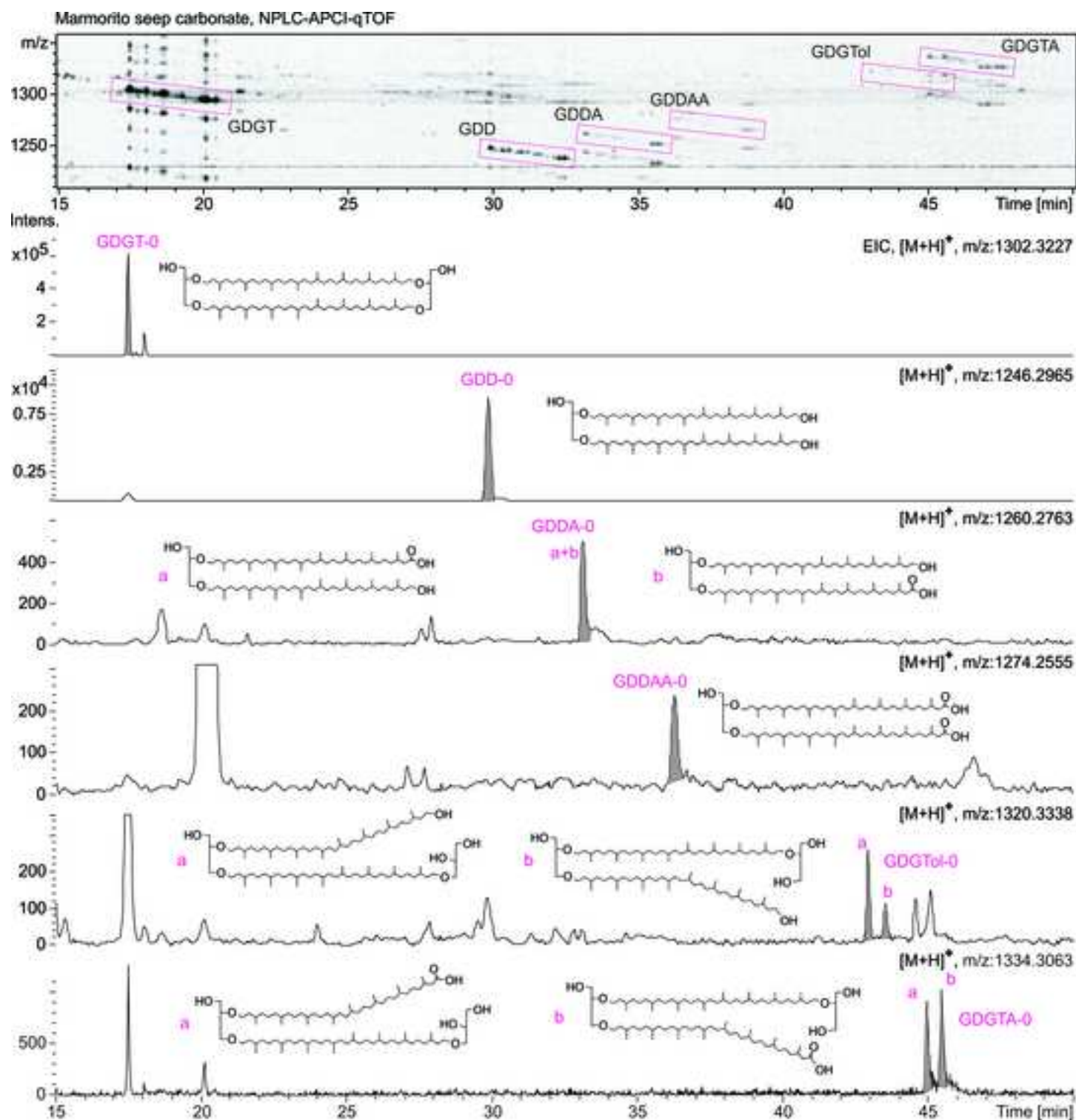
984 **Figure S5**

985 Extracted ion chromatograms of NPLC-APCI-qTOF showing the occurrence of IB-  
986 GDGTs and derivatives in Marmorito seep carbonate.

987

Compounds	Abbreviation	Structural illustration	Distribution				
			Leg201-1227	Marmorito seep carbonate	Guaymas Basin 4568	Hot spring T-15	<i>N. maritimus</i> late stat. phase
glycerol dialkyl glycerol tetraether	GDGT		+	+	+	+	+
glycerol dialkyl glycerol triether alcohol	GDGTol		+	+	+	n.d.	+
glycerol dialkyl glycerol triether acid	GDGTA		+	+	n.d.	n.d.	n.d.
glycerol dibiphytanol diether	GDD		+	+	+	+	+
glycerol dibiphytanol diether monoacid	GDDA		+	+	+	n.d.	+
glycerol dibiphytanol diether diacid	GDDAA		n.d.	+	n.d.	n.d.	n.d.
glycerol monobiphytanyl glycerol diether	GMGD		+	+	+	n.d.	n.d.
glycerol monobiphytanol monoether	GMM		+	+	+	n.d.	n.d.
glycerol monobiphytanol monoether acid	GMMA		n.d.	+	n.d.	n.d.	n.d.
biphytanic diol	bpdiol		+	+	+	+	+
biphytanic monoacid	bpmonoacid/ol		+	+	n.d.	n.d.	n.d.
biphytanic diacid	bpdiacid		n.d.	+	n.d.	n.d.	n.d.
H-shaped glycerol monoalkyl glycerol tetraether	H-GMGT		+	+	+	+	n.d.
H-shaped glycerol monoalkyl diether	H-GMD		+	+	+	+	n.d.
H-shaped C <sub>80</sub> tetrol	H-tetrol		+	+	+	+	n.d.
H-shaped C <sub>80</sub> monoacid	H-monoacid		n.d.	n.d.	+	n.d.	n.d.
H-shaped C <sub>80</sub> diacid	H-diacid		n.d.	n.d.	+	n.d.	n.d.
H-shaped C <sub>80</sub> triacid	H-triacid		n.d.	n.d.	+	n.d.	n.d.
H-shaped C <sub>80</sub> tetraacid	H-tetraacid		n.d.	n.d.	+	n.d.	n.d.

Figure

[Click here to download high resolution image](#)

Figure

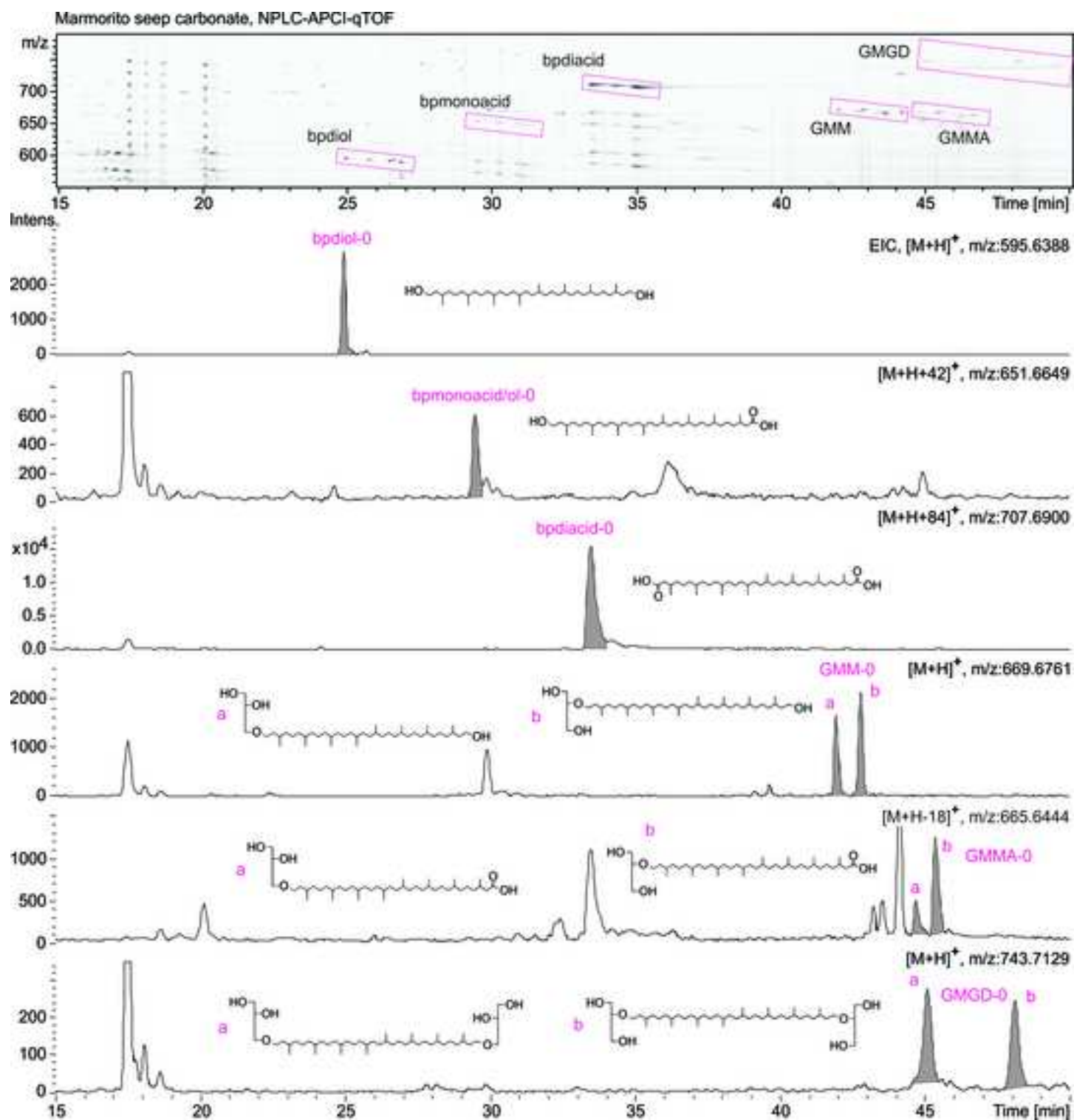
[Click here to download high resolution image](#)



Figure  
[Click here to download high resolution image](#)

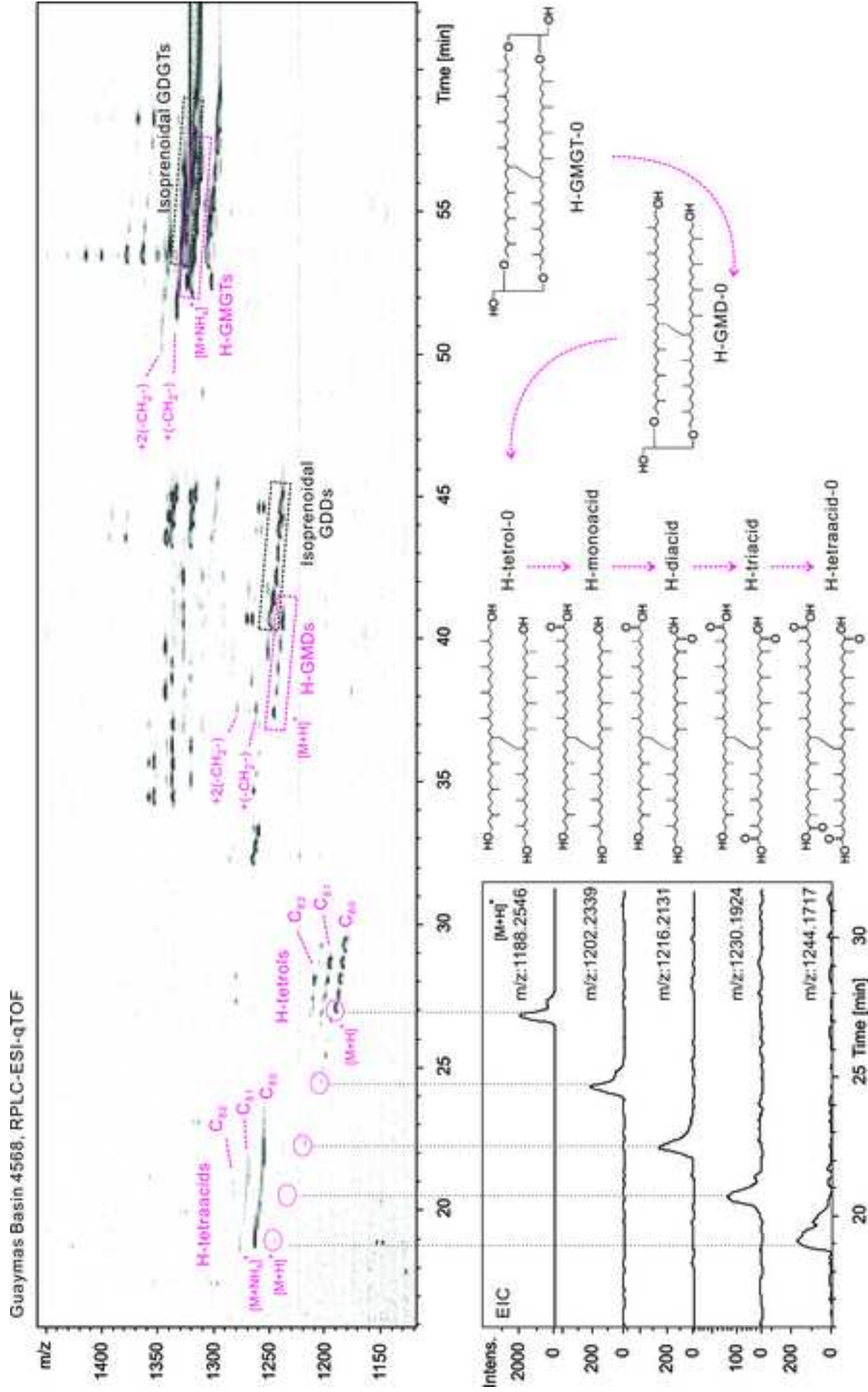
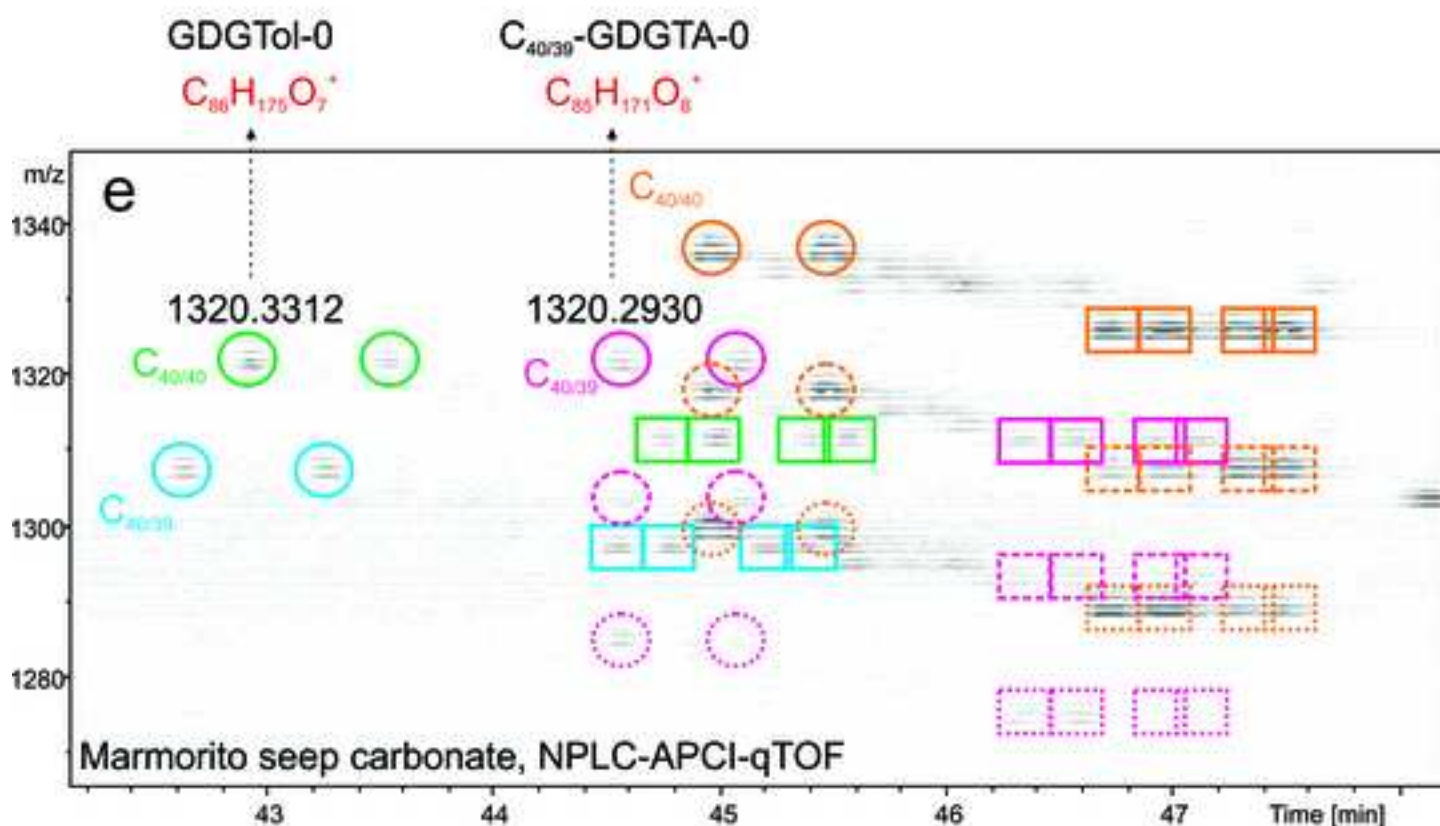
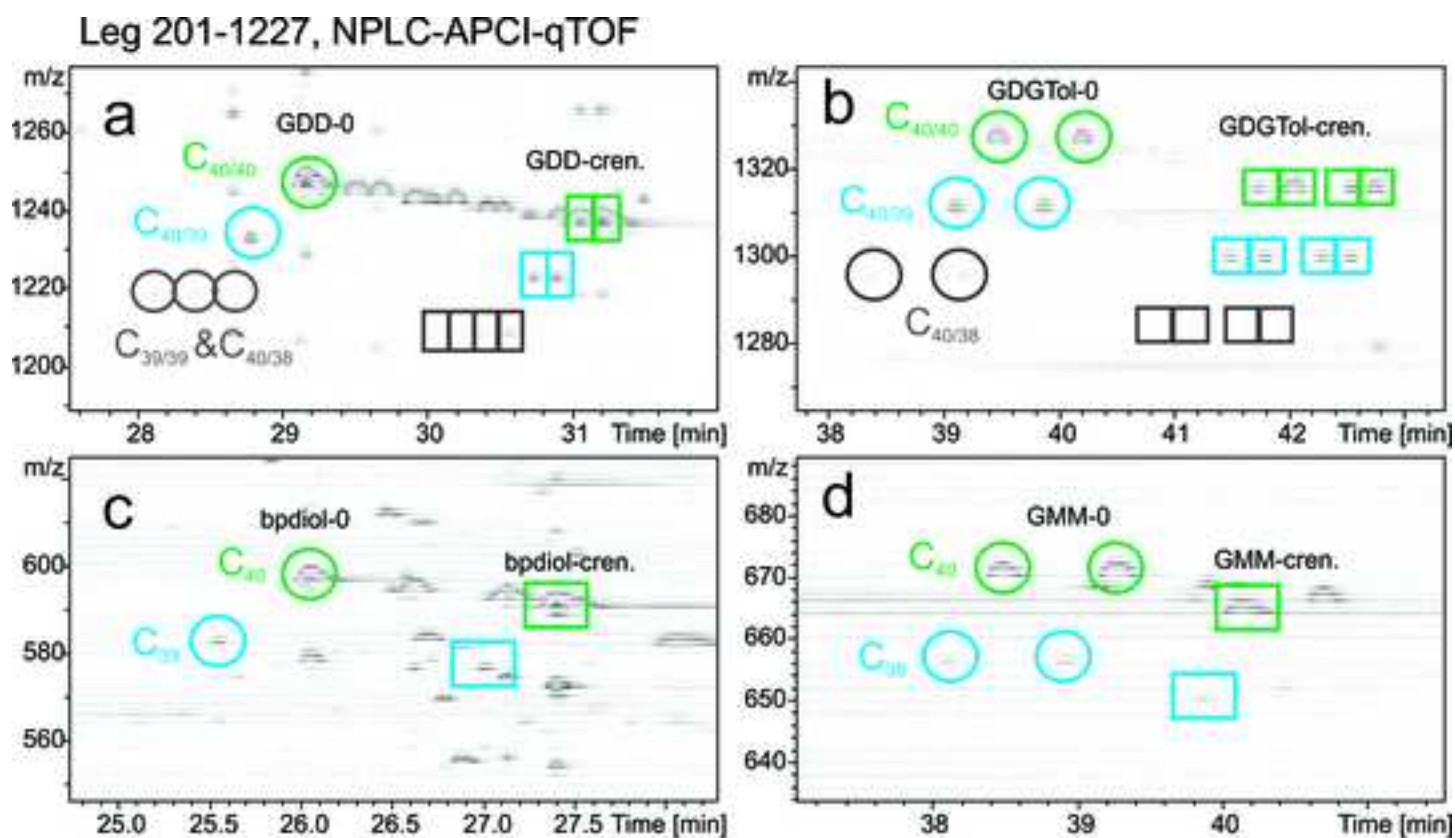


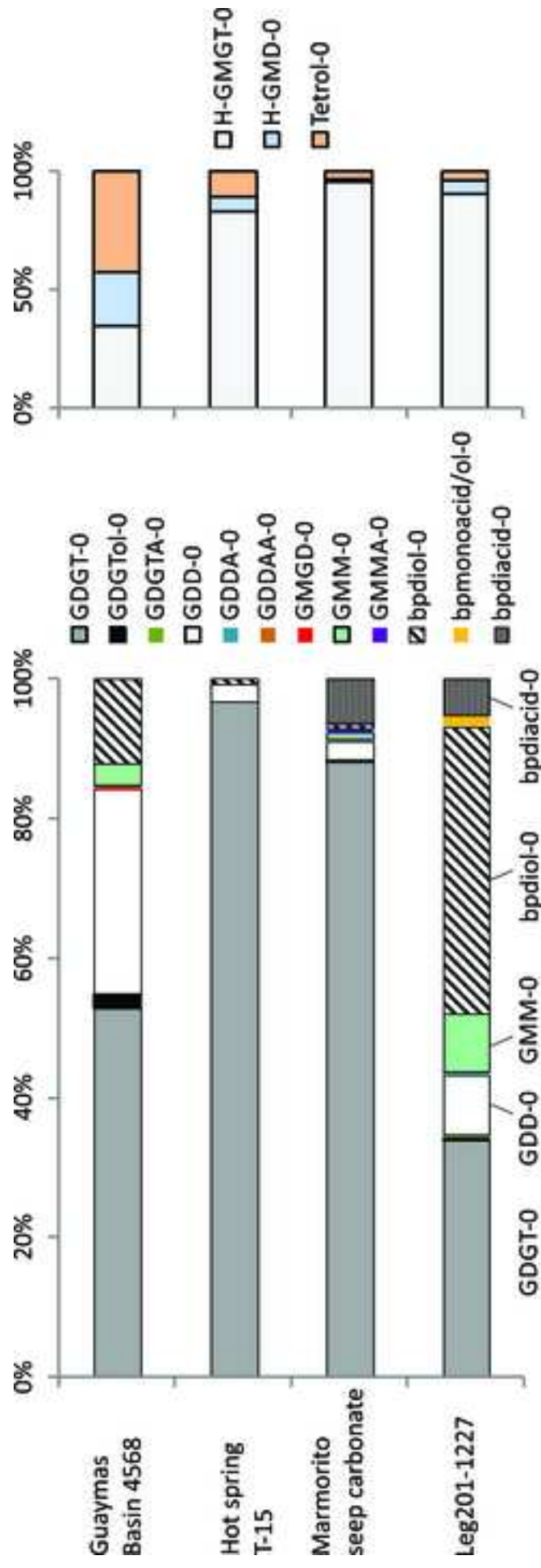
Figure  
[Click here to download high resolution image](#)





Figure

[Click here to download high resolution image](#)



Figure

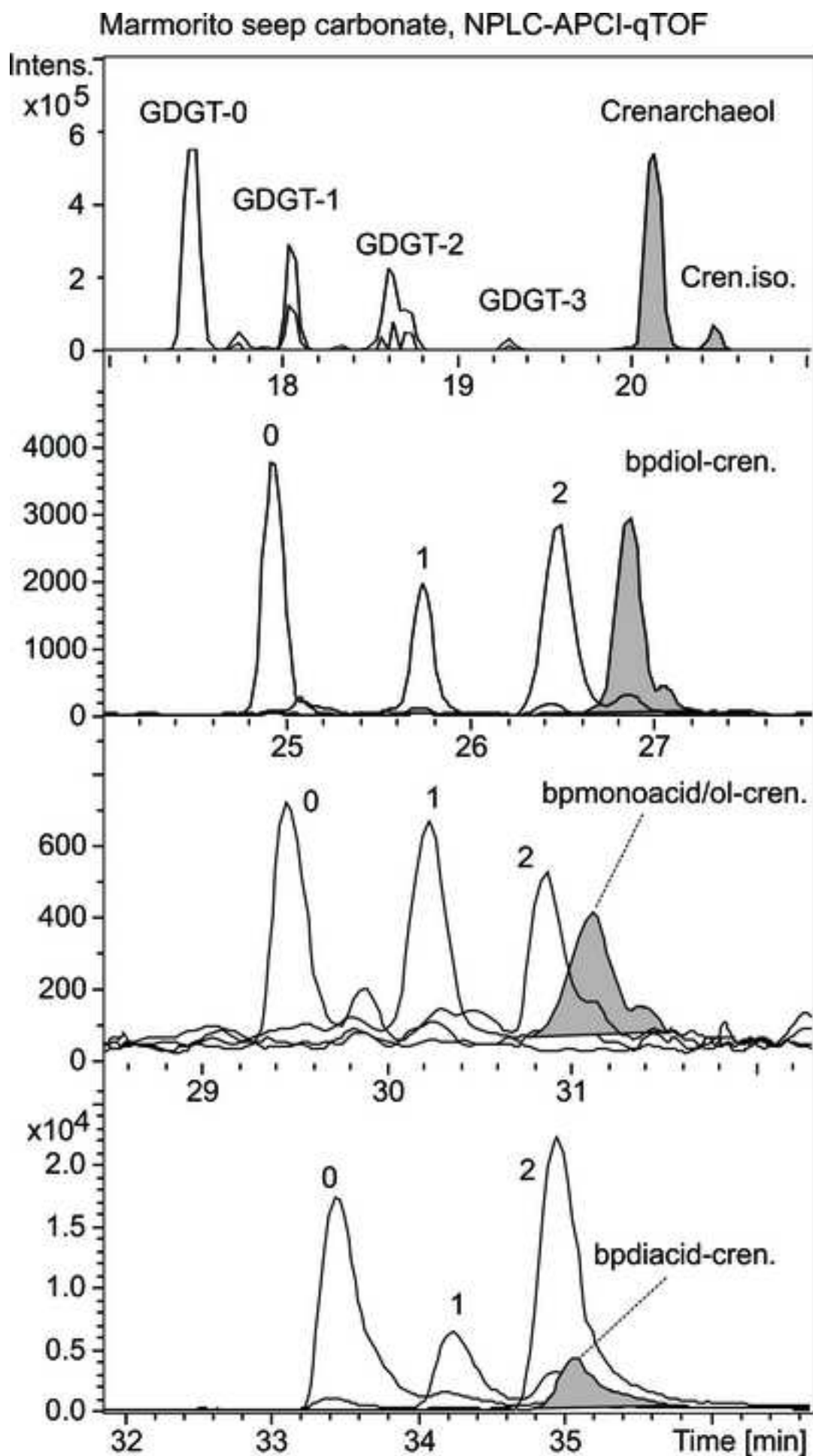
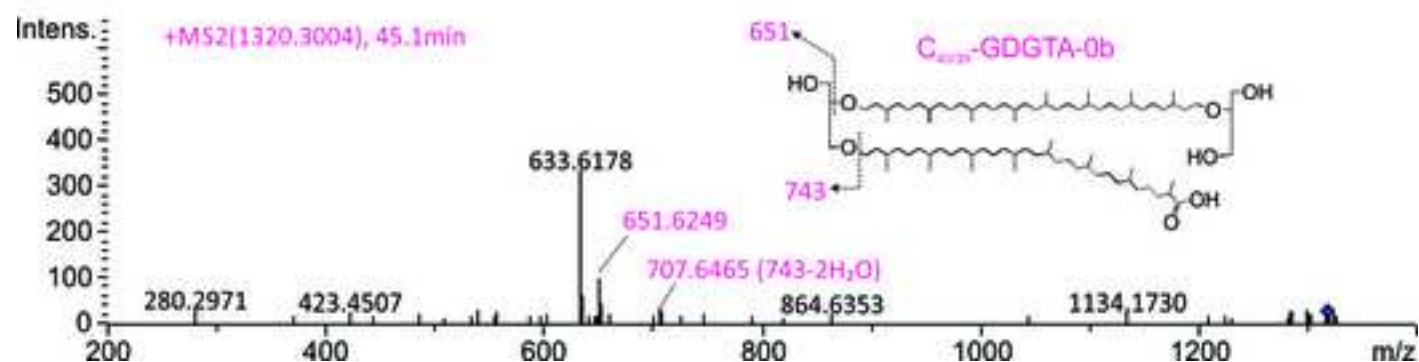
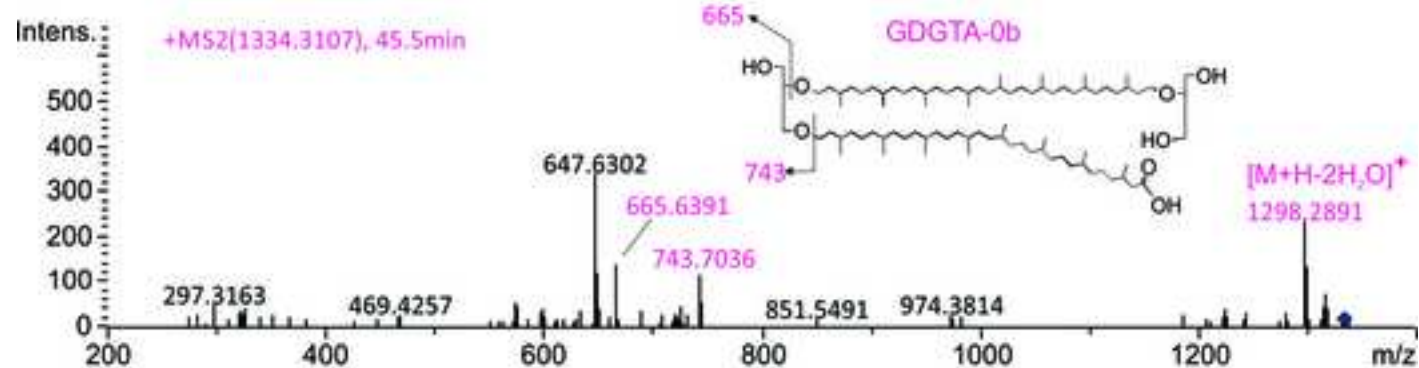
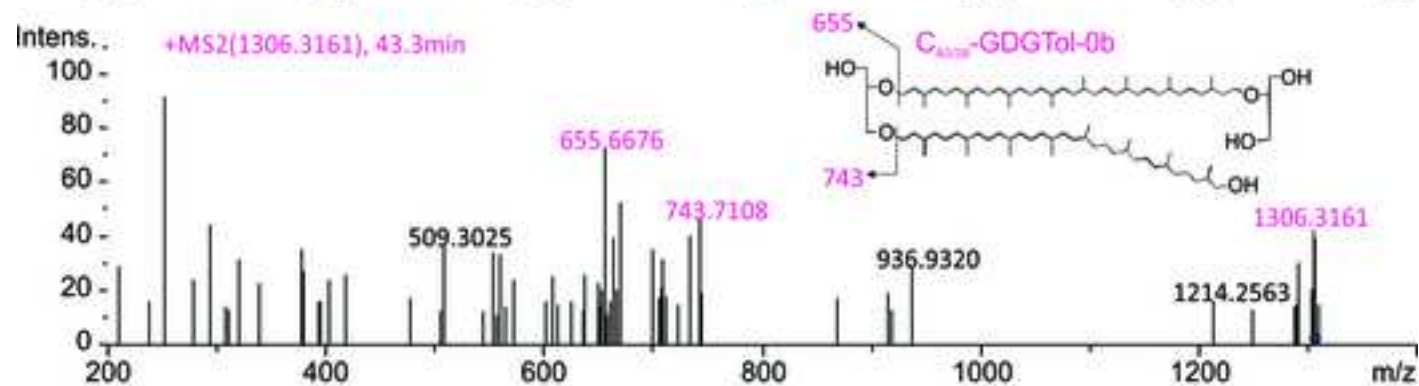
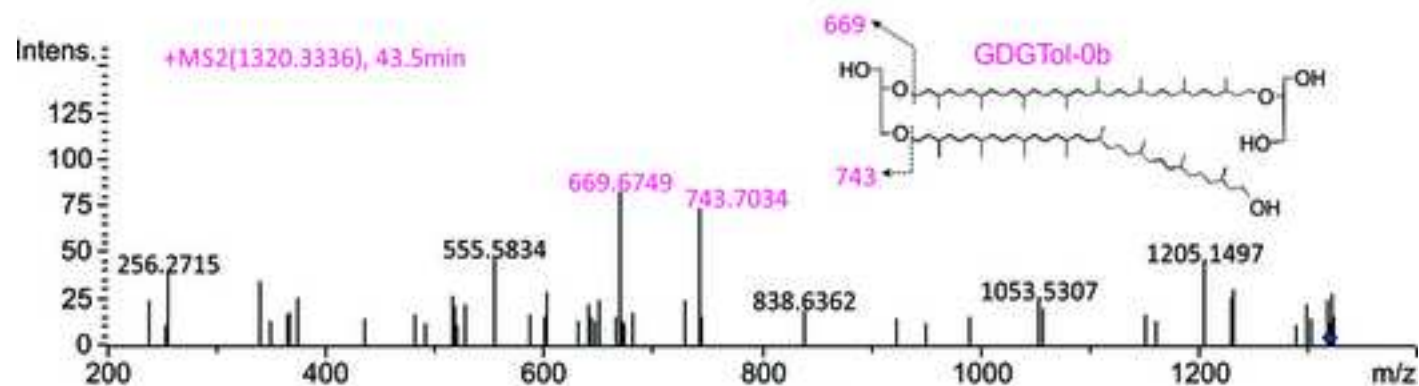
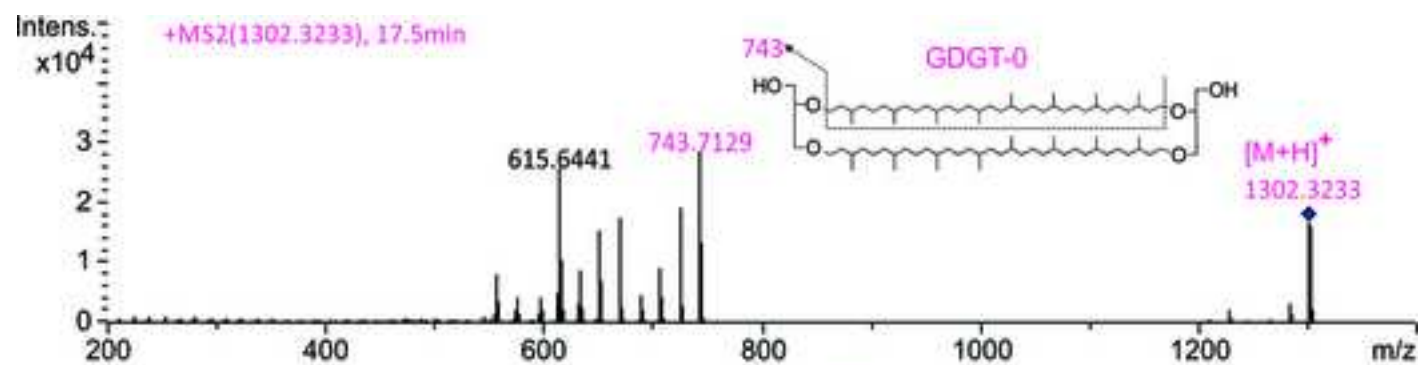
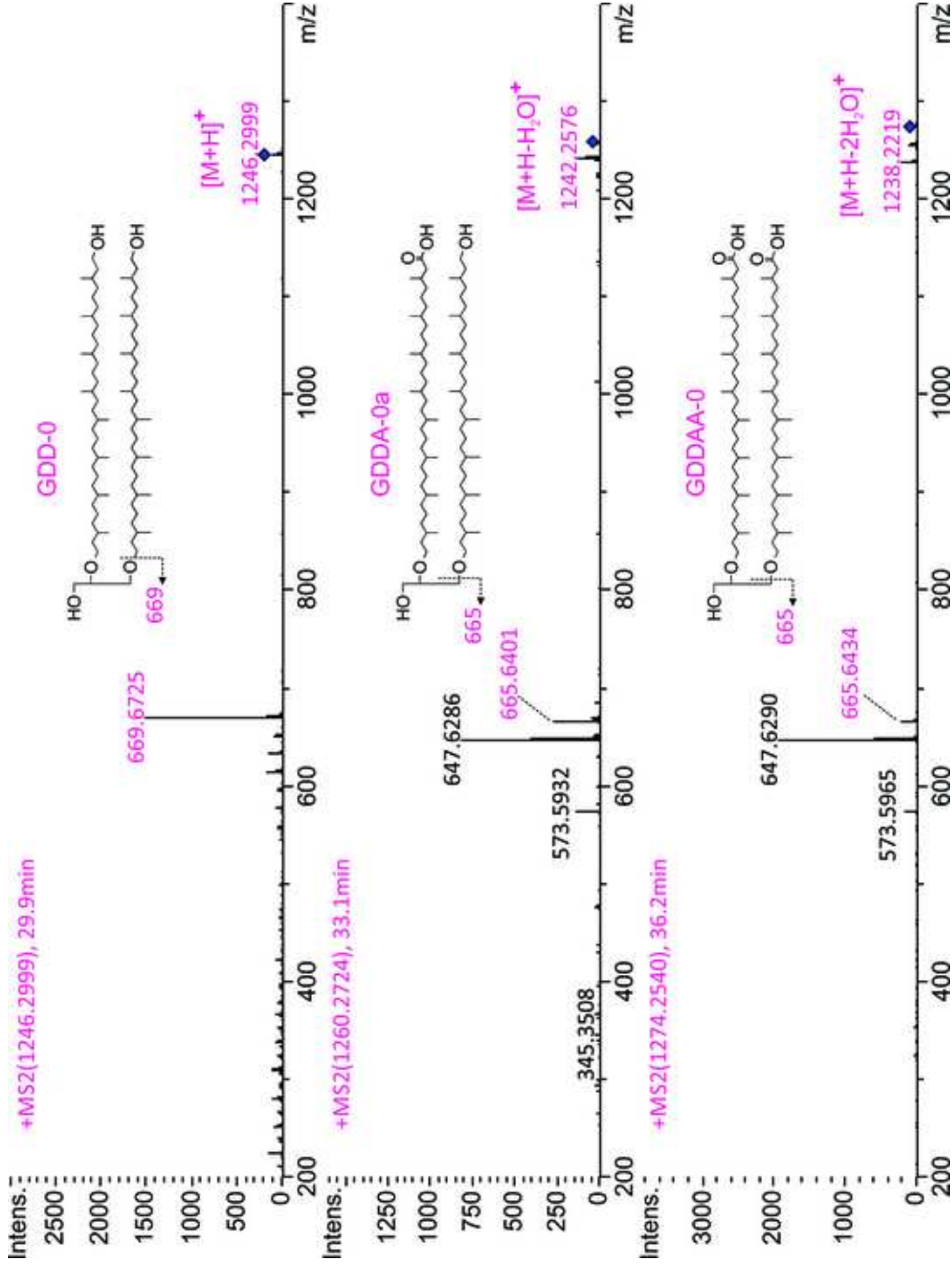
[Click here to download high resolution image](#)

Figure  
[Click here to download high resolution image](#)







Figure

[Click here to download high resolution image](#)

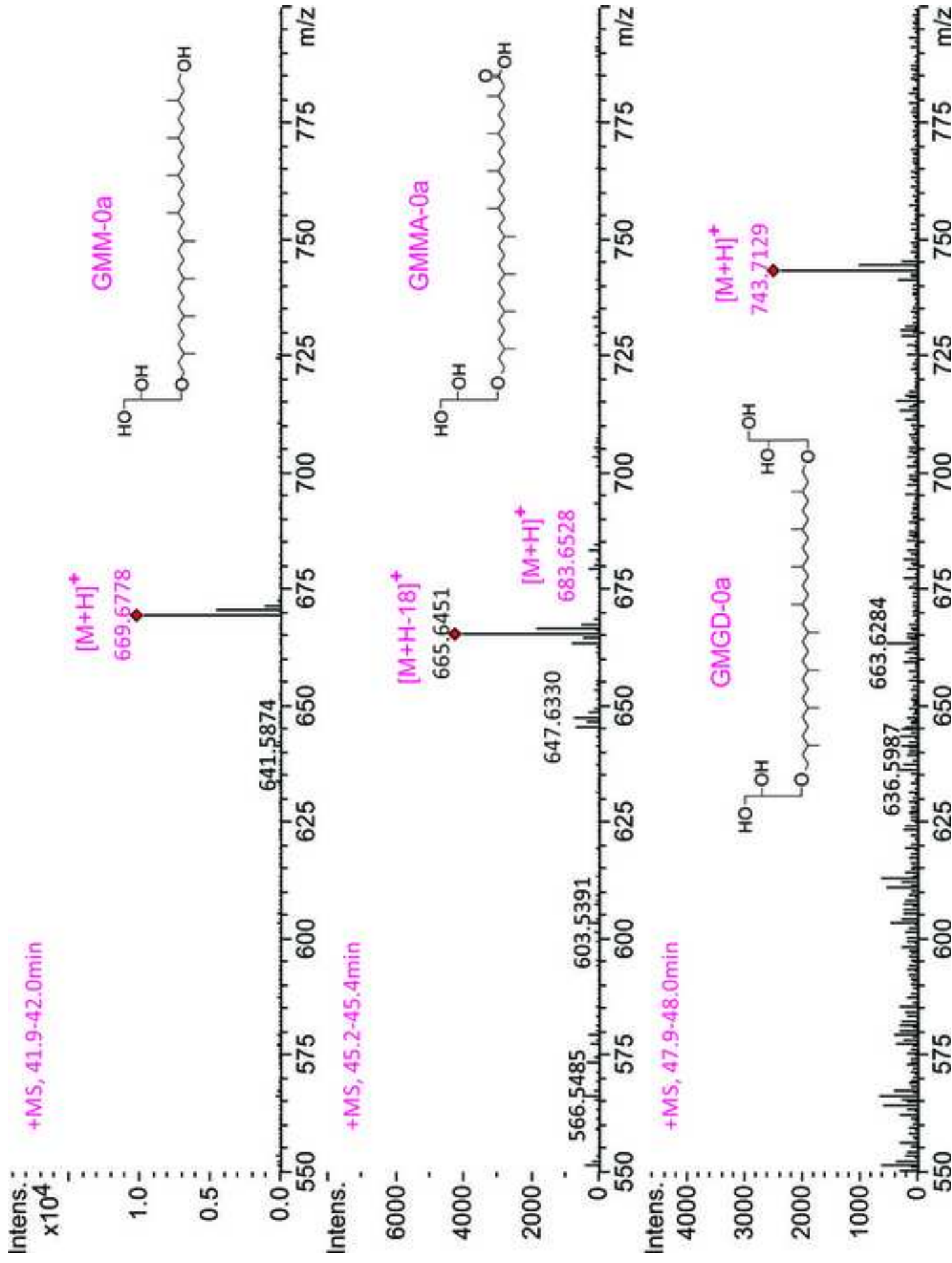


Figure  
[Click here to download high resolution image](#)

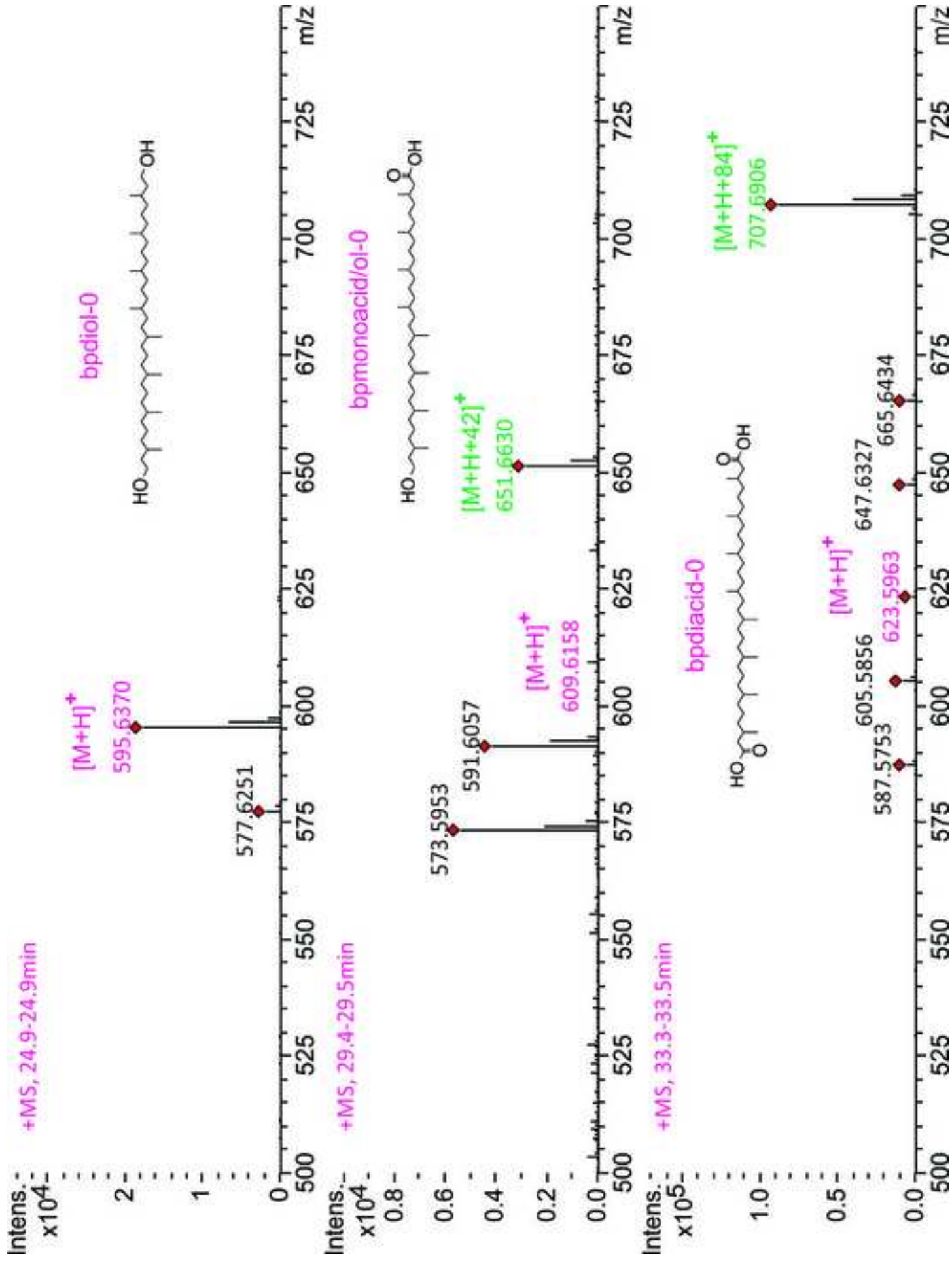
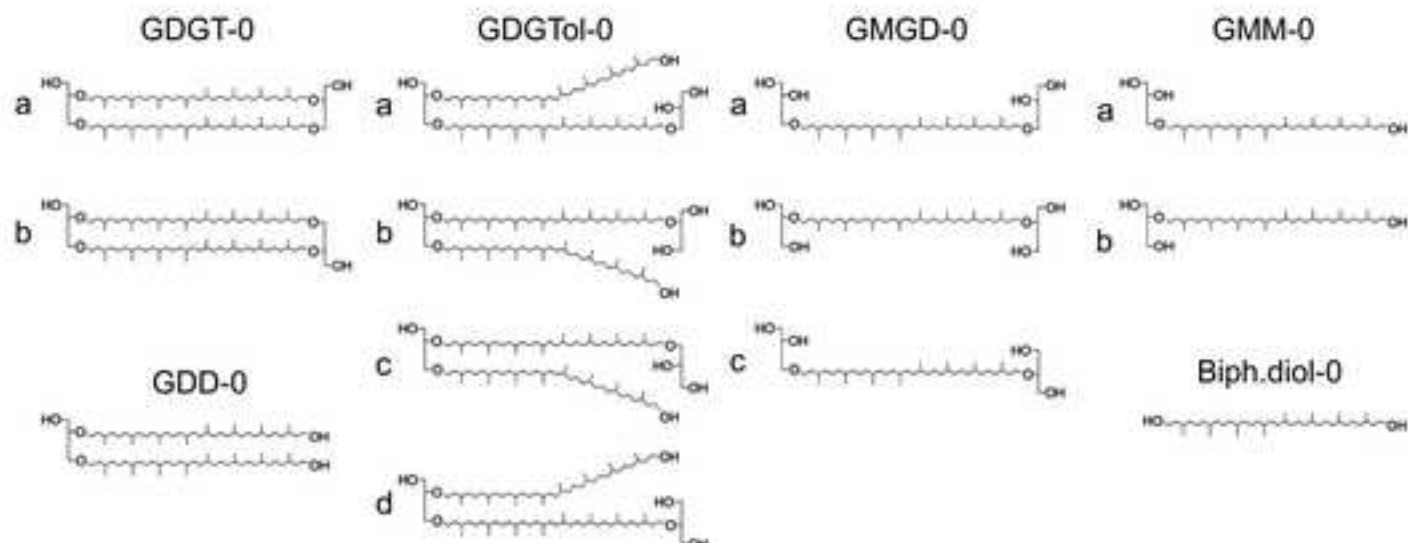
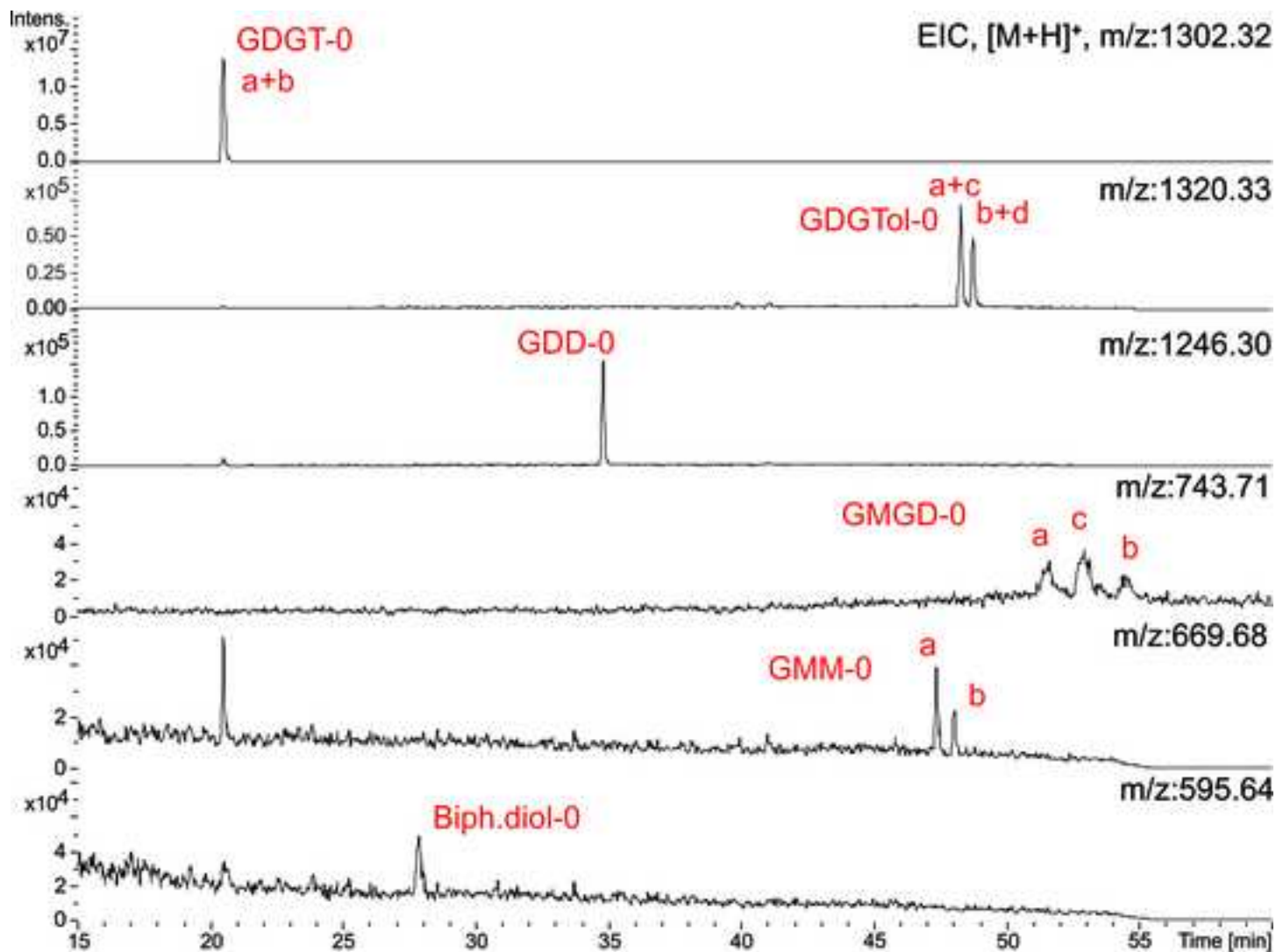
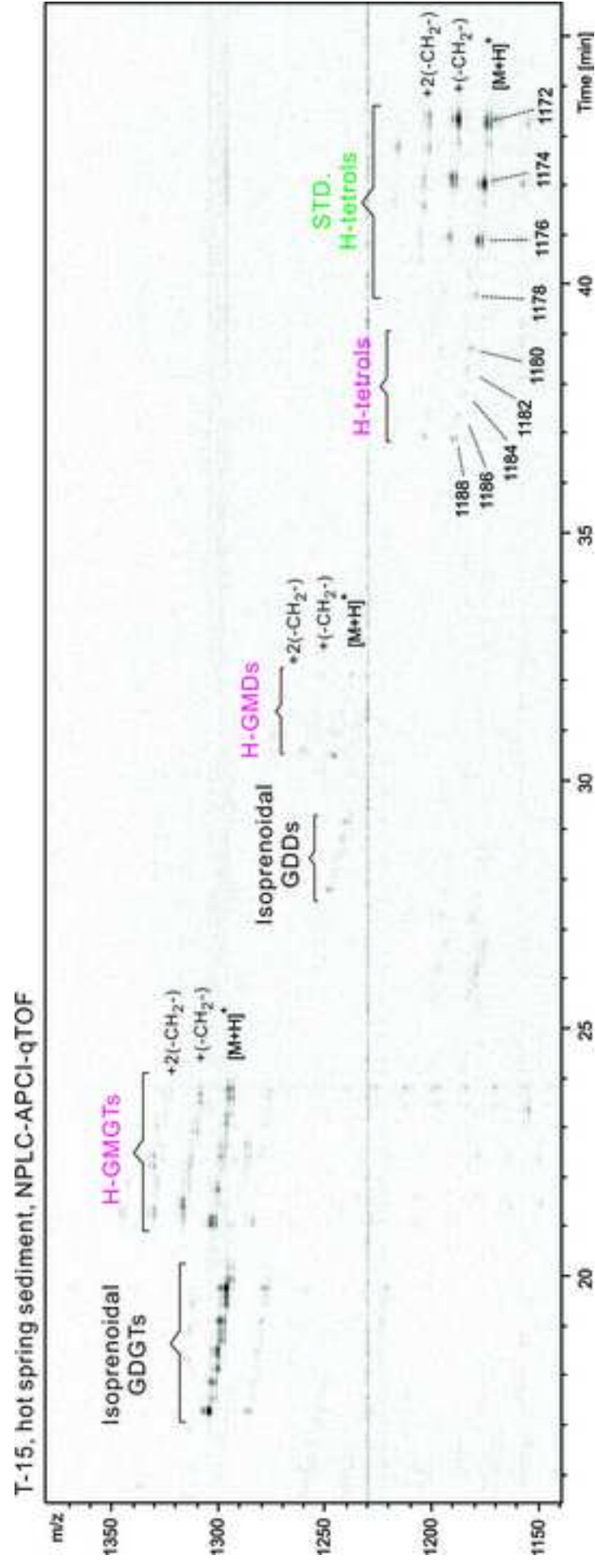


Figure  
[Click here to download high resolution image](#)







Marmorito seep carbonate, NPLC-APCI-qTOF

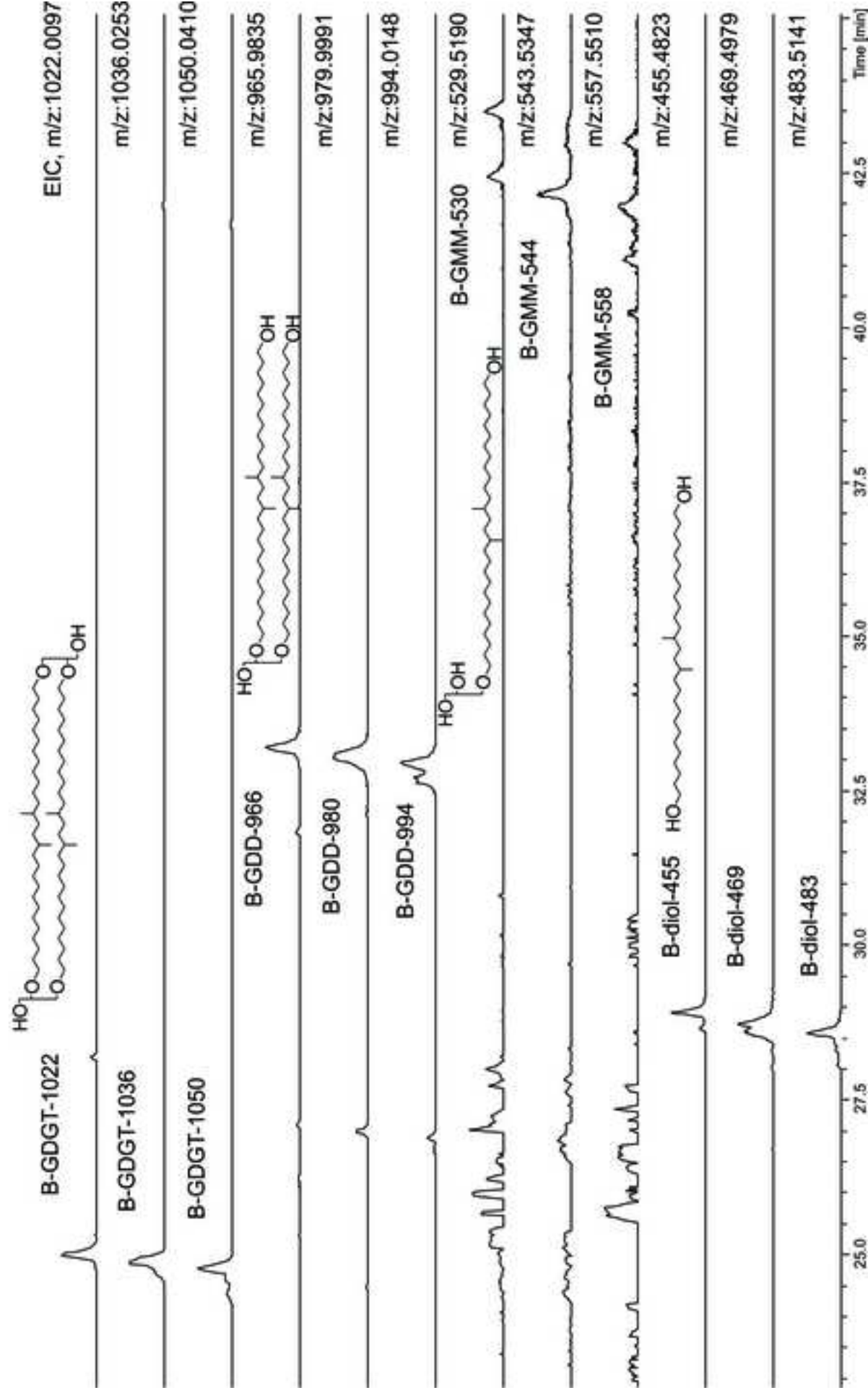


Figure  
[Click here to download high resolution image](#)

