1

2

Naphthenic acids in oil sands process waters: Identification by conversion of the acids or esters to hydrocarbons

6

7 Michael J. Wilde[a,1] and Steven J. Rowland*[a]

8

9 [a] Petroleum and Environmental Geochemistry Group, Biogeochemistry Research Centre, Plymouth
10 University, PL4 8AA, UK
11
12 * Corresponding author: fax: +44(0)1752 584710, e-mail: srowland@plym.ac.uk
13
14 [1] Present address: Department of Chemistry, University of Leicester, Leicester, UK, LE1 7RH;
15 mjw77@le.ac.uk
16

17

18

19 **Abstract**

20 Concerns over the toxicity associated with 'naphthenic' acids (NA) within oil produced

21 waters and oil sands process waters (OSPW), whether justified or not, have increased the

22 need for the structural elucidation of NA. Certainly, oil sands mining operations, such as

23 those in Alberta, Canada, result in the production of large volumes of OSPW contaminated

24 with NA. Monitoring NA, predicting the toxicities and accounting for the toxicity of residual

25 NA after clean-up treatments, have all been hindered by the lack of NA identifications.

26 Recently, a method involving the conversion of the esters of petroleum NA to the

27 corresponding hydrocarbons before analysis by comprehensive two-dimensional gas

28 chromatography-mass spectrometry (GC×GC-MS), resulted in the most comprehensive study

29 of bicyclic NA from petroleum to date. Here, we present results of the analysis by GC×GC-

30 MS of NA extracted from OSPW after conversion of the acids or ester derivatives to the

31 corresponding hydrocarbons. The identifications presented include novel alicyclic, aromatic

32 and sulphur-containing hydrocarbons and thus, by inference, of the corresponding acids. This

33 supports and significantly extends, previous identifications of OSPW NA as their methyl

34 esters and can now be used to better inform environmental monitoring programs and toxicity

35 studies.

36

39

40

41 **1. Introduction**

42 The identification of individual petroleum or 'naphthenic' acids (NA) has proved challenging

43 for analytical and environmental geochemists for over a century (e.g. Hell and Medinger

44 (1874)), owing to the complexity and polarity of the isolated acid mixtures. Early research

45 focused on the role of NA as geochemical markers, their interfacial and corrosive properties

46 and industrial uses of their metal salts (Seifert, 1975). However, the structural identification

47 and monitoring of NA has gained considerable interest in recent years, due to concerns over

48 the presence of NA in oil produced waters and oil sands process waters (OSPW) and their

49 associated toxicities (Thomas et al., 2009), particularly at the concentrations reported in the

50 OSPW storage lagoons (e.g. reviewed by Jones et al. (2011); Yue et al. (2015a) and

51 references therein).

52 Thus, intensive surface mining of bituminous oil sands, such as those of Alberta, Canada,

53 produces large volumes of OSPW. Despite recycling of the caustic extraction water, large

54 volumes of process water, contaminated with water-soluble organics from the bitumen, such

55 as NA, as well as suspended solids, continue to accumulate and are stored in tailings ponds

56 and lagoons (Kean, 2009).

57 Undiluted OSPW have been shown to be toxic to numerous organisms at the concentrations

58 found in tailings ponds and since this toxicity has been largely attributed to the NA content

59 (He et al., 2012; Marentette et al., 2015), considerable effort has been made to develop

60 methods for monitoring the extremely complex distributions of NA found in OSPW (Frank et

61 al., 2016). High and ultra-high resolution mass spectrometry methods have revealed that

62 OSPW contain an extensive range of NA (Yi et al., 2015), beyond so-called 'naphthenic' or

63 cyclic non-aromatic species (Rowland et al., 2014; Barrow et al., 2015). Many analytical

64 methods also include a chromatographic separation step (e.g. GC or UPLC) (Ortiz et al.,

65   2014; Brunswick et al., 2016; Headley et al., 2016), but most still lack sufficient separation

66   for identification of individual NA in the OSPW mixtures.

67   However, the use of comprehensive two-dimensional gas chromatography-mass spectrometry

68   (GC×GC–MS) for the analysis of NA as methyl esters has led to much better separations of

69   NA mixtures in OSPW. Thus, several bicyclic, tri- and pentacyclic diamondoid acids and

70   diacids have been identified by comparison of GC×GC retention times and mass spectra with

71   those of authentic compounds (Rowland et al., 2011b; Rowland et al., 2011d; Lengger et al.,

72   2013). However, the number of identifications possible was still limited by the complexity of

73   the acid mixtures, the sparsity of data in current mass spectral databases for relevant esters

74   and the poor availability of reference compounds for confirmation (Wilde et al., 2015).

75   Recently, a complementary method for the identification of NA was reported in which, after

76   calibration of the method with model acids, NA were first converted to the corresponding

77   hydrocarbons (Wilde and Rowland, 2015). This approach, based on refinements of chemical

78   transformation methods used in early investigations of NA (Seifert et al., 1969), coupled with

79   analysis of the reduced acids (viz: hydrocarbons) by GC×GC-MS, resulted in the

80   identification of over 40 individual bicyclic acids in a commercial NA mixture isolated from

81   petroleum. Identification of the bicyclanes and thus, by inference, of the original bicyclic

82   acids, was achieved by mass spectral comparison: many more published mass spectra exist

83   for reference hydrocarbons than for acids or esters (Denisov et al., 1977a; Denisov et al.,

84   1977b; Denisov et al., 1977c; Golovkina et al., 1984; Petrov, 1987). This made comparisons

85   of reference mass spectra with those of the unknowns, much more feasible than for the esters.

86   Following the success of the conversion approach on petroleum NA, the current investigation

87   aimed to apply the same method to the even more complex NA extracts of some OSPW.

88 **2. Methods**

89 Two NA extracts from OSPW tailings ponds from Alberta, Canada, were analysed. Sample

90 #1 was an NA extract from OSPW collected from the West In-Pit (WIP) tailings pond in

91 2009, and prepared as described previously (Frank et al., 2006). Briefly, a cleaned-up

92 concentrated naphthenate solution was obtained by passing the re-dissolved precipitated acids

93 from OSPW through a diethylaminoethyl-cellulose column (Frank et al., 2006). Sample #2

94 was OSPW from a second tailings pond. This OSPW was filtered, acidified and extracted in

95 DCM to provide a crude NA extract (cf. Pereira et al. (2013)).

96 **2.1. Instrumentation**

97 Comprehensive GC×GC-MS analyses were conducted on an Agilent 7890A gas

98 chromatograph (Agilent Technologies, Wilmington, DE) fitted with a Zoex ZX2 two stage

99 cooled loop GC×GC modulator and secondary oven (Houston, TX, USA) interfaced with a

100 BenchTOFdx™ time-of-flight mass spectrometer with an electron impact ionisation source

101 (Markes International, Llantrisant, Wales, UK). The primary column was a 60 m × 0.25 mm

102 × 0.25 μm Rxi®-1ms (Restek, Bellefonte, USA), followed by a 1 m x 0.1 mm deactivated

103 fused silica modulation loop. The secondary column was a 2.5 m × 0.1 mm × 0.1 μm BPX50

104 (SGE, Melbourne, Australia). Helium was used as carrier gas and the flow was kept constant

105 at 1.0 mL/min. 1 μL samples were injected at 275 °C splitless. Samples were analysed using

106 two temperature programmes (referred to as condition sets A and B). Condition set A was

107 described previously (Wilde and Rowland, 2015) and are given in the Supporting

108 Information. Condition set B involved the primary oven programmed from 40 °C, held for

109 1.3 min, then heated to 200 °C at 1 °C/min, to 280 °C at 5 °C/min and to 320 °C at 10 °C/min

110 and then held for 5 min. The secondary oven was programmed to track the primary oven at

111 20 °C above. The hot jet was programmed to start 20 °C above the primary oven and finish

112 100 °C above the primary oven over the period of the run; programmed from 60 °C, held for

113   1 min, then heated to 200 °C at 1.1 °C/min, to 280 °C at 3 °C/min and then to 400 °C at 5

114   °C/min. Modulation periods of 4 and 6 s were used. The MS transfer line and ion source

115   temperature were 290 °C and 300 °C, respectively.

116   Data processing was conducted using GC Image^{TM} v2.3 (Zoex, Houston, TX, USA). The

117   CLIC (Computer Language for Identifying Chemicals) expression tool within GC Image^{TM}

118   v2.3 was used for the determination of the presence or absence of compounds. CLIC

119   expressions are a powerful tool for applying mass spectral filters and produce advanced

120   extracted ion chromatograms (EICs) with additional constraints (e.g. Figure 1) (Reichenbach

121   et al., 2005). Some mass spectral matches were made using NIST Search MS 2.0. Reference

122   mass spectra obtained from the literature were input into a local NIST library using the NIST

123   librarian tool to allow mass spectral match quality values to be determined; typically >85 %.

124   **2.2. Derivatisation and Fractionation of NA (Scheme 1)**

125   Sample #1 was first derivatised by heating with BF_3-methanol (BF_3-MeOH) and the esters

126   fractionated by argentation (Ag-Ion) chromatography using a method described previously

127   (Scarlett et al., 2013) and outlined in the Supporting Information. Three fractions were used

128   for conversion to hydrocarbons; an 'alicyclic' fraction (F2) eluting with 100% hexane; an

129   'aromatic' fraction (F5) eluting with 95 %:5 % hexane:ether and an 'aromatic/sulphur'-

130   containing fraction (F7) eluting later with 95 %:5 % hexane:ether. The NA extract of Sample

131   #2 was not derivatised or fractionated prior to the conversion to the hydrocarbons: i.e. the

132   free acids were reduced to the hydrocarbons, avoiding the esterification step (Scheme 1).

133   **2.3. Conversion of Acids and Esters to Hydrocarbons**

134   Sample #1 (methyl esters) and sample #2 (free acids) were converted to hydrocarbons via the

135   three-step transformation developed and described previously (Scheme 1) (Wilde and

136   Rowland, 2015). Briefly, the acids and ester derivatives were reduced using lithium

137   aluminium hydride (LAH) to the primary alcohols. The alcohol products were derivatised to

138   the p-toluenesulfonate esters using tosyl chloride (TsCl) in the presence of 4-

139   (dimethylamino)pyridine (DMAP) and triethylamine (TEA). The tosylate products underwent

140   a 'SuperHydride®' reduction with excess lithium triethylborohydride (LiEt3BH) to produce

141   the hydrocarbons. To retain any volatile, low molecular weight products from conversion of

142   sample #2 (shown by preliminary GC-MS to contain more abundant low molecular weight,

143   volatile esters), the sample was concentrated to ~1.0 mL using a Kuderna-Danish apparatus,

144   followed by a simple silica chromatography clean-up step (Wilde and Rowland, 2015). The

145   masses obtained from the fractionation procedures and the yields of the conversion reactions

146   are detailed in Tables S1 – S3 in the Supporting Information.

147   **3.   Results and Discussion**

148   Fractionation of the NA extract of sample #1 by Ag-Ion chromatography (as methyl esters;

149   Scheme 1), produced 'alicyclic' (36 %), 'aromatic' (17 %) and 'aromatic/sulphur' (4 %) sub-

150   fractions of OSPW. These contrasted with the proportions of sub-fractions of a commercial

151   NA extract obtained from the refinement of petroleum (80.0 %, 6.4 % and 0.9 %

152   respectively) (Wilde and Rowland, 2015) but were consistent with the results of previous

153   fractionation studies of OSPW NA analysed by liquid chromatography with ion mobility and

154   high resolution mass spectrometry (Huang et al., 2016).

155   **3.1. Identification of NA after reduction to hydrocarbons**

156   **3.1.1.   Adamantane Acids**

157   Previously, tricyclic 'diamondoid' or adamantane acids, including diacids, have been

158   identified in OSPW and oil sands composite tailings pore water, as their methyl ester

159   derivatives using GC×GC-MS (Rowland et al., 2011b; Bowman et al., 2014; Lengger et al.,

160   2015). Therefore there is precedent for the identification of adamantane acids within some

7

161 OSPW NA. Thus, if the conversion of the more complex OSPW NA (free acids or esterified

162 acids) herein had been successful, analysis of the reduced acid products should lead to the

163 identification of the corresponding alkyl adamantane hydrocarbons.

164 The mass spectra of alkyl adamantanes are very distinctive and many reference mass spectra

165 are available in the literature: diamondoid hydrocarbons are commonly used as biomarkers in

166 petroleum geochemistry (Peters et al., 2005). The spectra of alkyl adamantanes usually

167 display a dominant or base peak ion corresponding to the loss of the largest alkyl group as a

168 radical, with the molecular ion for isomers substituted at tertiary carbons displaying a

169 relatively low intensity molecular ion (approx. $5 - 20$ %), compared with isomers substituted

170 at secondary carbon positions (approx. $30 - 50$ %) (Golovkina et al., 1984). The GC elution

171 order of many dimethyl-, ethyl-, methylethyl- and trimethyl adamantane isomers is

172 documented (Wingert, 1992; Wei et al., 2006; Wang et al., 2013): typically, alkyl

173 adamantanes substituted at the secondary carbon positions elute later in GC (and GC×GC)

174 than isomers substituted at the tertiary positions (Petrov, 1987).

175 The high complexity of the NA extracts from OSPW has hindered previous identifications

176 and even extracted ion monitoring of the GC×GC data can still result in very complex

177 chromatograms. However, use of CLIC expressions (cf. Wilde and Rowland (2015)), derived

178 from the key mass spectral features for alkyl adamantanes, simplified the resulting

179 chromatograms, displaying more clearly a series of isomers identified as mono- and poly-

180 substituted alkyl adamantanes by comparison with reference mass spectral data and known

181 elution orders (Figures 1, S1 and S2).

182 Assignment of the two ethyl isomers of adamantane in the reduced sample #2, after

183 comparison with the spectra reported by Polyakova et al. (1973) (Ad-II and V; Figure 1; C-F)

184 also allowed identification of numerous dimethyl- isomers (Figures S1 and S2) based on the

185 comparison of the retention positions, relative to the ethyl isomers. For example, the only two

186 dimethyladamantane isomers reported to elute between the two ethyl isomers are the cis- and

187 trans- isomers of 2,4-dimethyladamantane (Wei et al., 2006; Wang et al., 2013). In the

188 present study this assignment was supported by the high intensity molecular ions (50 – 60 %)

189 observed in the mass spectra of peaks Ad-III and –IV (Figure S1).

190 The earliest eluting peak (Ad-I) was assigned as 1,2-dimethyladamantane. The absence of

191 1,3-dimethyladamantane, expected to elute before the 1,2-dimethyl- isomer was consistent

192 with the absence of 3-methyladamantane-1-carboxylic acid methyl ester, after comparison of

193 the retention time and mass spectrum of the reference acid methyl ester with the original NA

194 methyl esters. The adamantane ethanoic acids were also identified in the methyl esters of

195 sample #2 by comparison of data with those for reference compounds; once more indicating

196 the complementary nature of the analysis of both the esters and hydrocarbons by GC×GC-

197 MS. The identification of these acids as the hydrocarbons in the reduced non-methylated,

198 unfractionated OSPW free NA (#2), also showed that the acids are not artefacts of

199 esterification.

200 The reduced samples #1 (reduced methyl esters of NA) and #2 (reduced, unesterified free

201 NA) also showed series of peaks displaying molecular ions at m/z 178, corresponding to C_{13}

202 tricyclic hydrocarbons, many of which had mass spectra matching the reference spectra of

203 trimethyl- and methylethyl- adamantanes (Figures 1 and S2). The additional clarification of

204 peaks provided by the CLIC expression chromatogram, allowed comparison of the retention

205 positions of the C_{13} adamantane isomers with those reported in the literature to aid

206 identification (Figure 1; A and B).

207 These results confirmed that the conversion of OSPW NA mixtures to the corresponding

208 hydrocarbons was successful, despite the increased complexity compared to the petroleum

209   acid mixtures analysed previously (Wilde and Rowland, 2015). Thus, due to the widely

210   available reference spectra and the known elution orders of numerous alkyl adamantanes,

211   conversion of the OSPW NA as acids or esters, to the corresponding hydrocarbons,

212   confirmed the presence of known acids, the presence of acids that were only tentatively

213   assigned previously and also resulted in the identification of numerous new acids.

214   ### 3.1.2. Bicyclic Acids

215   Alicyclic bicyclic acids are a major class of OSPW NA. High- and ultra-high resolution mass

216   spectrometric techniques often show ions corresponding to $C_nH_{2n-4}O_2$ acid species with

217   carbon numbers (n) ranging from C_{9-20}, as the most abundant $C_nH_{2n+z}O_2$ species (z, hydrogen

218   deficiency due to cyclicity) (Barrow et al., 2010). Analysis of commercial and OSPW NA

219   methyl esters by GC×GC-MS has shown that the ions detected by high resolution mass

220   spectrometry alone represent hundreds of different structural isomers (Damasceno et al.,

221   2014; Wilde et al., 2015). Based on the limited knowledge of bicyclic acids, a well-accepted

222   screening assay reported some bicyclic acids to be the most toxic of those studied (Jones et

223   al., 2011; Yue et al., 2015b). Scott et al. (2008) showed the ozonation of OSPW resulted in

224   the removal of most bicyclic acids ($C_nH_{2n-4}O_2$ species); however, little is known about the

225   residual acids, or of any transformation products.

226   Investigations into the structural identification of bicyclic acids resulted in the first

227   identification of novel bridged bicyclic acids, along with several bicyclic fused and

228   terpenoid-derived acids (Wilde and Rowland, 2015; Wilde et al., 2015). In the current

229   investigation, analysis of the 'alicyclic' acids as methyl esters of sample #1 and the

230   unfractionated acid methyl esters of sample #2, revealed homologous series of isomers with

231   mass spectra displaying molecular ions consistent with that of C_{11-16} and C_{10-15} bicyclic acids

232   respectively, as reported as typical of OSPW NA (Hindle et al., 2013). Careful concentration

233   of the hydrocarbon product of sample #2, following reduction of the unfractionated NA,

10

234   using a Kuderna-Danish apparatus and subsequent analysis by GC×GC-MS showed

235   homologous series of C_{10-15} bicyclanes (Figure S3; A and B).

236   Despite the complexity of sample #2, comparison of the mass spectra of those bicyclanes

237   which were not hindered by co-elution effects, resulted in the identification of

238   bicyclo[3.2.1]octane, bicyclo[3.3.0]octane and bicyclo[4.3.0]nonane alkyl derivatives (Figure

239   2 and S3). For example, the earliest eluting isomer (**bi-I**) was assigned as 1,4-

240   dimethylbicyclo[3.2.1]octane (Figure 2; A and B) after comparison with reference mass

241   spectra (Denisov et al., 1977a). The same bicyclane, and by inference the same acid, was

242   identified previously in petroleum-derived NA (Wilde and Rowland, 2015).

243   Comparison of the mass spectrum of isomer bi-II with the mass spectra of alkyl

244   bicyclo[3.3.0]octanes resulted in the identification of 1,2-dimethylbicyclo[3.3.0]octane

245   (Figure 2; C and D) (Denisov et al., 1977c). This assignment was supported by its retention

246   position relative to 1,4-dimethylbicyclo[3.2.1]octane (bi-I) which matched the elution order

247   of C_{10} bicyclanes reported previously (Piccolo et al., 2010) and complemented the

248   identification of several bicyclo[3.3.0]octane acids in petroleum-derived NA (Wilde and

249   Rowland, 2015). However, this was the first identification of a bicyclo[3.3.0]octane acid in

250   NA with a substituent at a bridgehead carbon.

251   The isomer bi-III displayed a mass spectrum matching those of *cis-cis* and *trans-cis* 3-

252   methylbicyclo[4.3.0]nonane (Figure 2; E and F) (Denisov et al., 1977b). Numerous

253   methylbicyclo[4.3.0]nonanes were identified in the petroleum NA (Wilde and Rowland,

254   2015) with the isomer assigned as the *cis-cis*/*trans-cis* 3-methyl- isomer eluting earliest of all

255   the 2- and 3-methyl- isomers. The mass spectra of the later eluting isomers bi-IV to VI were

256   difficult to assign. The isomer bi-IV was tentatively assigned as 2,6-

257   dimethylbicyclo[3.2.1]octane (Figure S3; C and D) and the mass spectra of the remaining

11

258   bicyclanes, e.g. bi-V and VI (Figure S3; E and F), did not match any of the reference spectra,

259   which encompassed an extensive range of bicyclane isomers.

260   Petroleum NA and NA extracted from OSPW have often been reported to possess

261   significantly different compositions (Grewer et al., 2010). This is often observed as

262   differences in the HPLC retention times of the unresolved NA 'humps' (OSPW NA eluting

263   slightly earlier) (Han et al., 2008). The differences are also reflected in the increased

264   resistance of OSPW NA to biodegradation; suggested to be because the OSPW NA possess

265   more highly branched carboxylated side chains (Misiti et al., 2014; Brown and Ulrich, 2015).

266   However, this speculation has never been supported by the identification of the acids. The

267   presence of aromatics in OSPW NA, may have an important role in the differences in

268   retention time and toxicity observed for different NA mixtures (Jones et al., 2012), but

269   aromaticity obviously does not explain the observations made for these alicyclic acids.

270   A complementary explanation suggested here is that the OSPW NA examined herein instead

271   possess a higher proportion of condensed, bridged structures, substituted at different positions

272   e.g. bridgehead carbon (Figure 2; A-D). Differences between the bicyclic acids, identified as

273   the corresponding bicyclanes, in petroleum NA and OSPW NA included the position of

274   substituents (e.g. bridgehead substituted isomers present in OSPW, Figure 2; C), as well as

275   the presence of some unknown bicyclic acid isomers detected in OSPW, but not detected in

276   petroleum NA (e.g. Figure S3). The suggestion is supported by the fact that bridged

277   compounds elute earlier than fused ring acids (e.g. decalin acids) and is reinforced by the

278   identification of bicyclo[2.2.1]heptane acids in OSPW and the reported elution order of

279   reference acids (Wilde et al., 2015). Bicyclo[2.2.1]heptane acids are the earliest eluting of

280   any the bicyclic acids previously analysed, which were not detected in petroleum NA as the

281   methyl esters, or corresponding bicyclanes (Wilde and Rowland, 2015; Wilde et al., 2015).

282   These structural differences may also help explain the differences in toxicity of some NA

283   mixtures.

284   Despite recent advances in spectrometric studies on OSPW and NA (Headley et al., 2016),

285   few examples exist which have provided structural evidence for the observed compositional

286   differences between different NA mixtures based on the identification of individual acids

287   (Swigert et al., 2015). Following the identifications herein; representative NA can now be

288   used to quantify the relative proportions of bicyclic species by high resolution methods;

289   relevant acids can be synthesised and the toxicities measured (Yue et al., 2015b).

3.1.3. Non-Diamondoid Tricyclic Acids

291   Apart from the adamantane acids, few tricyclic acids have been identified in OSPW to date.

292   The mass spectra of isomers tt-I to -VI all displayed fragmentation patterns characteristic of

293   those of tricyclic terpane hydrocarbons (Figure S4). Ekweozor and Strausz (1982) reported

294   the identification of a series of C_{19-30} tricyclic terpanes present in bitumen from the Athabasca

295   oil sands with a cheilanthane 'core', with the branched alkyl side chain substituted at position

296   14 (Figure S4; F). Based on the molecular ions, in the mass spectra of the hydrocarbons tt-I to

297   –VI, the series included two C_{20} (tt-I and -II), two C_{21} (-III and -IV), one C_{23} (-V) and one C_{24}

298   (-VI) compound (Figure S4). The structures were proposed to possess cheilanthane 'cores',

299   methyl- substituted in the 13-position with a branched alkyl chain substituted in the 14-

300   position. However, the exact position of the alkyl chain, or carboxylated side chain in the

301   original acids, is a subject of debate (Cyr and Strausz, 1983).

3.1.4. Tetracyclic and Pentacyclic Acids

303   The presence of tricyclic diamondoid acids in OSPW was proposed to be, at least partly, due

304   to the biotransformation of the corresponding hydrocarbons (Rowland et al., 2011b).

305   Pentacyclic diamondoid hydrocarbons, such as diamantane and alkyl diamantanes have also

306   been reported in various crude oils and were reported to be much more resistant to

307   biodegradation compared to the adamantane homologues (Wang et al., 2006).

308   However, some pentacyclic diamondoid acids have also been detected in OSPW (Rowland et

309   al., 2011d). The occurrence of these acids was suggested to occur from the biodegradation of

310   the corresponding alkyl diamantanes and that their presence was evidence for biodegradation

311   (Rowland et al., 2011d). Following the successful conversion and identification of the alkyl

312   adamantanes herein (and thus alkyl adamantane acids), the hydrocarbons of both reduced

313   samples #1 (reduced esters of NA) and #2 (reduced free acid NA) were examined for isomers

314   of penta- and tetracyclic 'diamondoids'.

315   Reference spectra for diamantane hydrocarbons, whilst fewer than those of the adamantanes,

316   are as distinctive, due the stability of the cage-like core, (e.g. displaying base peak ions due to

317   the loss of an alkyl radical and intense molecular ions for isomers with the alkyl substituent

318   on a secondary carbon position). Comparison of the reference mass spectra of 1-, 3- and 4-

319   methyldiamantanes and their GC elution order (Wingert, 1992; Wang et al., 2013), allowed

320   the identification of all three isomers in sample #2 and of 3-methyldiamantane in sample #1

321   (Figure S5; B).

322   Extracted ion monitoring of the key fragment ions revealed a series of isomers assigned as

323   alkyl diamantanes (diA-I to diA-XXV) ranging from C_{15-18} based on the observed molecular

324   ions (Figure S5). Interpretation of the mass spectra of peaks diA-I to -XXV (examples are

325   shown in Figure S5; B-H), resulted in the assignment of isomers of ethyl- (D), dimethyl- (E),

326   ethylmethyl- (F and G) and dimethylethyl- (H) diamantanes. These results provided strong

327   supporting evidence for the array of highly substituted alkyl diamantane acids previously

328   speculated to be present in OSPW (Rowland et al., 2011d).

14

329 Tetracyclic acids were also tentatively assigned previously based on mass spectral

330 interpretation of the spectra of the methyl esters (Rowland et al., 2011d). The tetracyclic

331 acids were proposed to possess ring-opened diamantane structures (Figure S6; A). However,

332 comparison of the retention positions and mass spectra of synthetic reference acids with those

333 of NA within OSPW, did not confirm their identification (Rowland et al., 2011a).

334 Series of peaks were observed in the hydrocarbon products obtained from both samples #1

335 and #2 herein, with spectra with molecular and base peak ions consistent with C_{14-18}

336 tetracyclic hydrocarbons, and some additional C_{13} isomers in the reduced sample #2 (Figures

337 S7 and S8). The mass spectra were very similar to those of alkyl adamantanes and alkyl

338 diamantanes, dominated by a base peak ion presumably corresponding to the loss of the

339 largest alkyl substituent and suggesting a highly stable core (Figures S7 and S8). These

340 hydrocarbons were presumed to originate from tetracyclic acids, such as those previously

341 tentatively assigned (Rowland et al., 2011d). Subsequent examination of data for the original

342 acid methyl esters indeed showed isomers across the same carbon number range, with mass

343 spectra similar to those previously reported in other OSPW samples (Figure S9) (Rowland et

344 al., 2011d).

345 The lowest carbon number isomers (C_{13-14}) displayed base peak ions at m/z 161 (Figure S7;

346 B-D). Assuming the compounds possess 'adamantanoid-like' structures, this indicated a C_{12}

347 'core'; m/z 161 (100 %) corresponding with a $C_{12}H_{17}^{+}$ highly stable cation. The only alkyl

348 (tetracyclic) 'adamantanoid' structures with a C_{12} core, previously assigned in crude oils, are

349 alkyl 2,4-cyclopentano-adamantanes (Petrov, 1987). Other stable, non-adamantanoid

350 tetracyclic structures, such as bridged or highly condensed cyclopentyl structures would not

351 allow for isomers with carbon numbers as low as C_{13} and the few reference mass spectra

352 available were significantly different (Vorob'eva et al., 1986).

353   Analysis of the reduced acid, hydrocarbon products herein therefore provides the first

354   evidence for a firmer assignment of the tetracyclic acids as possessing cyclopentano-

355   adamantane cores e.g. 2,4-cyclopentano-adamantane acids (Figure S6; B and C), at least for

356   some of the higher carbon number homologues. Subsequent re-examination of the

357   unfractionated OSPW NA as their acid methyl esters, did indeed reveal esters with mass

358   spectra consistent with those of the corresponding cyclopentano-adamantane acids (Figure

359   S9).

3.1.5. Monoaromatic Acids

361   Fraction 5 from the argentation fractionation of sample #1, was selected for conversion to the

362   hydrocarbons since it was rich in aromatic acids, as indicated by infrared spectroscopy

363   (FTIR), strong UV absorption and adsorption to silver ions in argentation chromatography

364   (Jones et al., 2012). The 'aromatic' acids have been shown to have similar acute toxicity to

365   the 'alicyclic' acids but only produce a weak estrogenic effect in zebrafish larvae (Scarlett et

366   al., 2013).

367   Rarely cited early investigations of petroleum NA, also showed that aromatic acids were

368   present (Knotnerus, 1957), but the use of the term 'naphthenic' to describe the acids appears

369   to have led most recent workers to overlook the relevance of the aromatic species. Even with

370   this realisation (Rowland et al., 2011c; Jones et al., 2012; Bowman et al., 2014), very few

371   aromatic acids have been firmly identified in OSPW to date.

372   However, using the mass spectral characteristics of aromatic hydrocarbons to aid assignments

373   herein, series of peaks were assigned as monoaromatic species (Figure S10). The base peak

374   ions in the mass spectra of isomer ma-I to -IV (e.g. m/z 105, 119, 133) suggested these

375   consisted mainly of di-, tri- and tetramethylalkylbenzenes (McLafferty, 1963). The presence

376   of many fragment ions (e.g. m/z 175, 161 and 147) suggested they did not possess straight n-

377 alkyl chain substituents, since the spectra were dissimilar from those reported (Ji-Zhou et al.,

378 1993). The lack of intense ions at *m/z* 106, 120 and 134 (from γ-hydrogen transfer as opposed

379 to benzylic or β-cleavage) also implied that either, the alkyl chain was not methyl-substituted

380 at the γ-position or, that both ortho- positions on the benzene ring were substituted

381 (Sinninghe Damsté et al., 1988). Some of the mass spectra closely resembled those reported

382 for monoaromatic hydrocarbons identified as alkylbenzenes obtained from artificially

383 matured melanoidins (Larter et al., 1981). This evidence implied that the corresponding

384 alkylbenzene alkanoic acids were OSPW NA.

385 Since alkylbenzenes are common pollutants and conceivably might also be hydrocarbon

386 contaminants in the OSPW, the corresponding acid methyl esters were searched for in the

387 esterified OSPW NA, based on the expected retention differences observed between the acid

388 methyl esters and the hydrocarbons (Figure S11). The acids (methyl esters) were indeed

389 present. The similarity in mass spectra between the alkylbenzene hydrocarbons and acids

390 (e.g. Figure S12) indicated that the carboxylate moiety was at the end of an alkyl side chain

391 and not directly substituted onto the benzene ring, consistent with the current understanding

392 of bacterial degradation of branched alkyl hydrocarbons (Misiti et al., 2014).

393 Recently, the methyl esters of C_{11} indane-2-acetic acid and tetralin-2-carboxylic acids were

394 identified, following derivatization, in oil sands composite tailings pore water (Bowman et

395 al., 2014). Such low molecular weight acids were not present in the OSPW aromatic fraction

396 of sample #1 under study herein and hence were not observed in the hydrocarbon reduction

397 products. Nonetheless, series of larger C_{14-20} branched alkyl indanes and tetralins were

398 tentatively identified (mb-I-VI) by comparison of the spectra in the NIST mass spectral

399 library and with available literature mass spectra (Figure S13) (Booth et al., 2006). These

400 identifications imply that the corresponding indane and tetralin alkanoic acids, not known

401 previously, were present in OSPW.

402 Numerous tricyclic monoaromatic acids were present in OSPW NA, though the only acid to

403 be assigned to date is dehydroabietic acid (DHAA) (Jones et al., 2012). One compound (tm-I)

404 with a mass spectrum which was similar to that of the corresponding hydrocarbon,

405 dehydroabietane, was identified in the OSPW hydrocarbons obtained from reduction of the

406 'aromatic' fraction of sample #1 (Figure S14; A and B). However, the spectrum of tm-I was

407 even more similar to that of the C_{19} and isomeric C_{20} hydrocarbons 13,14-dimethyl- and 13-

408 methyl-14-ethylpodocarpa-8,11,13-triene (Figure S14; C and D) (Azevedo et al., 1990). This

409 was supported by a subsequent retro-examination of the aromatic acid methyl esters of this

410 fraction, which failed to reveal the presence of DHAA methyl ester in this OSPW sample (*cf.*

411 unlike a different OSPW sample examined by Jones et al. (2012)).

412 Identification of the corresponding alkyl substituted podocarpa-8,11,13-triene acid has not

413 been previously reported in OSPW NA. However, Azevedo et al. (1994) reported the

414 presence of a series of C_{19-30} tricyclic C-ring monoaromatic carboxylic acids possessing the

415 same structure, with the alkanoate chain in the 14-position in Tasmanian tasmanite (marine

416 type shale).

3.1.6. Diaromatic Sulphur-Containing Acids

418 Fraction 7 from the fractionation of sample #1, was selected because it is a sulphur-rich,

419 aromatic fraction of OSPW NA for conversion to hydrocarbons (West et al., 2014). The latter

420 authors West et al. (2014) tentatively assigned five C_{16-18} diaromatic sulphur-containing acids

421 as possessing methyl substituted dibenzothiophene or dimethylnaphthothiophenes (or

422 dimethyldibenzothiopyrans) with methyl branched propanoate acid side chains. Despite

423 comparison of the GC×GC retention times and mass spectra with those of reference

424 dibenzothiophene acid methyl esters and rigorous mass spectral interpretation, none were

425 firmly identified in that study.

426   Examination of the 'aromatic/sulphur'-containing fraction of acid methyl esters analysed

427   herein by GC×GC-MS indeed revealed the five C_{16-18} isomers reported by West et al. (2014).

428   Following reduction, the corresponding sulphur-containing hydrocarbons were detected,

429   possessing mass spectra similar to those of the original acid methyl esters (e.g. dominated by

430   base peak ions at m/z 197, 211 and 225), corresponding to the loss of $\cdot C_3H_7$ (Figure 3 and

431   S15). Application of a CLIC expression (Figure 3; B) enabled clear mass spectra to be

432   obtained for interpretation and comparison with reference spectra and compounds. The only

433   structure possible for the alkanoate side chain of these isomers, consistent with the loss of M-

434   87 and the absence of a radical cation produced from a McLafferty rearrangement in the mass

435   spectra of the methyl esters (West et al., 2014), along with the consistent loss of $\cdot C_3H_7$ by

436   benzylic cleavage in the mass spectra of the hydrocarbons, was that of a dibenzothiophene (or

437   naphthothiophene) containing an isobutyl side chain. The retention position and mass

438   spectrum of synthesised 4-isobutyldibenzothiophene matched that of the dbt-I isomer,

439   confirming its identification (Figure 3; C and D), and by inference, the identification of

440   dibenzothiophene-4-isobutanoic acid in OSPW NA.

441   West et al. (2014) also reported the presence of sulphur-containing acids with 8 and 9 double

442   bond equivalents (DBE), eluting slightly earlier than those compounds possessing 10 DBE

443   now identified as dibenzothiophene acids. Examination of the F7 hydrocarbon product

444   herein, revealed a series of isomers with spectra displaying molecular ions consistent with

445   diaromatic sulphur-containing hydrocarbons with 8 DBE (e.g. m/z 214, 228, 242 and 256),

446   eluting slightly earlier than the 10 DBE dibenzothiophenes (e.g. isomer nat-1 in Figure 3; A).

447   The mass spectra of these isomers (nat-1 to -V) were dominated by intense base peak ions at

448   m/z 171 and 185, with very few fragment ions observed (Figure S16). This indicated that the

449   compounds were highly condensed, with the base peak ion most likely formed via benzylic

450   cleavage of an alkyl substituent. Comparison of the mass spectrum of the isomer nat-I, with

19

451 reference spectra in the NIST MS Library, showed that the spectrum was similar to that of 2-

452 methyl-2H-naphtho[1,8-bc]thiophene (Figure S16; A and B). The NIST mass spectrum

453 showed a loss of M-15, attributed to fragmentation of the methyl substituent, via benzylic

454 cleavage, as proposed previously for the fragmentation of alkyl 2H-naphtho[1,8-

455 bc]thiophenes (Figure S16; A) (Porter, 1985).

456 The mass spectrum of nat-I displayed a similar fragmentation pattern and ions to that of the

457 NIST reference spectrum. However, with the base peak ion of nat-I corresponding to the loss

458 of M-43 instead, this was attributed to the loss of a propyl group $\cdot C_3H_7$ (Figure S16; B). The

459 mass spectra of naphtho[1,8-bc]thiophene acids would be expected to display similar mass

460 spectral fragmentation patterns (due to the stability of the condensed structure) with

461 additional radical cations due to McLafferty rearrangement of the alkanoate side chain,

462 involving the hydrogen on the carbon adjacent to the sulphur atom. Subsequent re-

463 examination of the OSPW F7 NA methyl esters indeed revealed a series of peaks with mass

464 spectra displaying fragmentation patterns similar to those of the nat-I to -V hydrocarbons,

465 with additional even mass ions, corresponding to M-74, M-88 and M-102 ions (Figure S17),

466 attributed to neutral losses from hydrogen rearrangement on the alkanoate side chain. These

467 observations, coupled with the spectral interpretation; similarity with the reference spectrum

468 (Figure S16; A and B); retention position relative to the dibenzothiophenes and precedent for

469 the presence of sulphur-containing acids with 9 DBE confirmed by GC×GC with sulphur

470 chemiluminescence detection and high resolution mass spectrometry (West et al., 2014), led

471 to the tentative assignment of naphtho[1,8-bc]thiophene acids in OSPW NA herein.

472 Sulphur-containing hydrocarbons possess the same nominal mass as some aromatic

473 hydrocarbons and could be misassigned when analysed by unit resolution mass spectrometric

474 techniques (Hegazi and Andersson, 2007). However, interpretation and comparison of the

475   mass spectra of such hydrocarbons, clearly ruled them out as possibilities for those discussed

476   above.

4. Conclusions

478   The identifications presented herein of individual hydrocarbons and thus acids in OSPW,

479   included alicyclic, aromatic and sulphur-containing NA. The assignments were consistent

480   with, but expanded upon, previous evidence of such acids in other matrices, identified as the

481   esters (Rowland et al., 2011b; Rowland et al., 2011d; West et al., 2014; Wilde et al., 2015)

482   and were supported by the identification of multiple isomers of each acid type. The acids

483   identified can perhaps be used to better inform environmental monitoring programs as well as

484   for geochemical and toxicity studies. The same approach may have wide applicability to

485   other polluted waters and functionalised organic compounds in sediments and petroleum.

Supplementary Material

487   Mass spectra of all identified hydrocarbons and a summary of all identifications made are

488   provided in the Supplementary Material (Table S4 and Figures S1-S17).

Acknowledgements

499 **References**

500 Azevedo, D.A., de Aquino Neto, F.R., Simoneit, B.R.T., 1990. Mass spectrometric
501   characteristics of a novel series of ring-c monoaromatic tricyclic terpanes found in
502   Tasmanian tasmanite. Organic Mass Spectrometry 25, 475-480.
503 Azevedo, D.A., Neto, F.R.A., Simoneit, B.R.T., 1994. Extended saturated and monoaromatic
504   tricyclic terpenoid carboxylic acids found in Tasmanian tasmanite. Organic
505   Geochemistry 22, 991-1004.
506 Barrow, M.P., Peru, K.M., Fahlman, B., Hewitt, L.M., Frank, R.A., Headley, J.V., 2015.
507   Beyond Naphthenic Acids: Environmental Screening of Water from Natural Sources
508   and the Athabasca Oil Sands Industry Using Atmospheric Pressure Photoionization
509   Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. Journal of the
510   American Society for Mass Spectrometry 26, 1508-1521.
511 Barrow, M.P., Witt, M., Headley, J.V., Peru, K.M., 2010. Athabasca Oil Sands Process
512   Water: Characterization by Atmospheric Pressure Photoionization and Electrospray
513   Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry.
514   Analytical Chemistry 82, 3727-3735.
515 Booth, A.M., Sutton, P.A., Lewis, C.A., Lewis, A.C., Scarlett, A., Chau, W., Widdows, J.,
516   Rowland, S.J., 2006. Unresolved Complex Mixtures of Aromatic Hydrocarbons:
517   Thousands of Overlooked Persistent, Bioaccumulative, and Toxic Contaminants in
518   Mussels. Environmental Science & Technology 41, 457-464.
519 Bowman, D.T., Slater, G.F., Warren, L.A., McCarry, B.E., 2014. Identification of individual
520   thiophene-, indane-, tetralin-, cyclohexane-, and adamantane-type carboxylic acids in
521   composite tailings pore water from Alberta oil sands. Rapid Communications in Mass
522   Spectrometry 28, 2075-2083.
523 Brown, L.D., Ulrich, A.C., 2015. Oil sands naphthenic acids: A review of properties,
524   measurement, and treatment. Chemosphere 127, 276-290.
525 Brunswick, P., Hewitt, L.M., Frank, R.A., Kim, M., van Aggelen, G., Shang, D., 2016. A
526   traceable reference for direct comparative assessment of total naphthenic acid
527   concentrations in commercial and acid extractable organic mixtures derived from oil
528   sands process water. Journal of Environmental Science and Health, Part A, 1-7.
529 Cyr, T.D., Strausz, O.P., 1983. The structures of tricyclic terpenoid carboxylic acids and their
530   parent alkanes in the Alberta oil sands. Journal of the Chemical Society, Chemical
531   Communications 0, 1028-1030.
532 Damasceno, F.C., Gruber, L.D., Geller, A.M., de Campos, M.C.V., Gomes, A.O., Guimarães,
533   R.C., Péres, V.F., Jacques, R.A., Caramão, E.B., 2014. Characterization of naphthenic
534   acids using mass spectroscopy and chromatographic techniques: study of technical
535   mixtures. Analytical Methods 6, 807-816.
536 Denisov, Y.V., Matveyeva, I.A., Sokolova, I.M., Petrov, A.A., 1977a. Mass-spectrometric
537   study of hydrocarbons of bicyclo[3.2.1]octane series. Petroleum Chemistry U.S.S.R.
538   17, 85-93.
539 Denisov, Y.V., Sokolova, I.M., Petrov, A.A., 1977b. Mass Spectrometric study of
540   hydrocarbons of the bicyclo[4.3.0]octane series (In Russian). Neftekhimiya 17, 491-
541   497.
542 Denisov, Y.V., Vorob'eva, N.S., Petrov, A.A., 1977c. Mass Spectrometric study of
543   hydrocarbons of the bicyclo[3.3.0]octane series (In Russian). Neftekhimiya 17, 656-
544   662.

545 Ekweozor, C.M., Strausz, O.P., 1982. 18,19-Bisnor-13βH, 14-αh-cheilanthane: a novel
546   degraded tricyclic sesterterpenoid-type hydrocarbon from the athabasca oil sands.
547   Tetrahedron Letters 23, 2711-2714.
548 Frank, R.A., Kavanagh, R., Burnison, B.K., Headley, J.V., Peru, K.M., Der Kraak, G.V.,
549   Solomon, K.R., 2006. Diethylaminoethyl-cellulose clean-up of a large volume
550   naphthenic acid extract. Chemosphere 64, 1346-1352.
551 Frank, R.A., Milestone, C.B., Rowland, S.J., Headley, J.V., Kavanagh, R.J., Lengger, S.K.,
552   Scarlett, A.G., West, C.E., Peru, K.M., Hewitt, L.M., 2016. Assessing spatial and
553   temporal variability of acid-extractable organics in oil sands process-affected waters.
554   Chemosphere 160, 303-313.
555 Golovkina, L.S., Rusinova, G.V., Petrov, A.A., 1984. Mass Spectrometry of Saturated
556   Hydrocarbons. Russian Chemical Reviews 53, 870-887.
557 Grewer, D.M., Young, R.F., Whittal, R.M., Fedorak, P.M., 2010. Naphthenic acids and other
558   acid-extractables in water samples from Alberta: What is being measured? Science of
559   the Total Environment 408, 5997-6010.
560 Han, X., Scott, A.C., Fedorak, P.M., Bataineh, M., Martin, J.W., 2008. Influence of
561   Molecular Structure on the Biodegradability of Naphthenic Acids. Environmental
562   Science & Technology 42, 1290-1295.
563 He, Y., Patterson, S., Wang, N., Hecker, M., Martin, J.W., El-Din, M.G., Giesy, J.P.,
564   Wiseman, S.B., 2012. Toxicity of untreated and ozone-treated oil sands process-
565   affected water (OSPW) to early life stages of the fathead minnow (Pimephales
566   promelas). Water Research 46, 6359-6368.
567 Headley, J.V., Peru, K.M., Barrow, M.P., 2016. Advances in mass spectrometric
568   characterization of naphthenic acids fraction compounds in oil sands environmental
569   samples and crude oil—A review. Mass Spectrometry Reviews 35, 311-328.
570 Hegazi, A.H., Andersson, J.T., 2007. Limitations to GC-MS Determination of Sulfur-
571   Containing Polycyclic Aromatic Compounds in Geochemical, Petroleum, and
572   Environmental Investigations. Energy & Fuels 21, 3375-3384.
573 Hell, C., Medinger, E., 1874. Ueber das Vorkommen und die Zusammensetzung von Säuren
574   im Rohpetroleum. Berichte der deutschen chemischen Gesellschaft 7, 1216-1223.
575 Hindle, R., Noestheden, M., Peru, K., Headley, J., 2013. Quantitative analysis of naphthenic
576   acids in water by liquid chromatography–accurate mass time-of-flight mass
577   spectrometry. Journal of Chromatography A 1286, 166-174.
578 Huang, R., Chen, Y., Gamal El-Din, M., 2016. Silver-Ion Solid Phase Extraction Separation
579   of Classical, Aromatic, Oxidized, and Heteroatomic Naphthenic Acids from Oil Sands
580   Process-Affected Water. Environmental Science & Technology 50, 6433-6441.
581 Ji-Zhou, D., Vorkink, W.P., Lee, M.L., 1993. Origin of long-chain alkylcyclohexanes and
582   alkylbenzenes in a coal-bed wax. Geochimica et Cosmochimica Acta 57, 837-849.
583 Jones, D., Scarlett, A.G., West, C.E., Rowland, S.J., 2011. Toxicity of individual naphthenic
584   acids to *Vibrio fischeri*. Environmental Science & Technology 45, 9776-9782.
585 Jones, D., West, C.E., Scarlett, A.G., Frank, R.A., Rowland, S.J., 2012. Isolation and
586   estimation of the 'aromatic' naphthenic acid content of an oil sands process-affected
587   water extract. Journal of Chromatography A 1247, 171-175.
588 Kean, S., 2009. Eco-Alchemy in Alberta. Science 326, 1052.
589 Knotnerus, J., 1957. The chemical constitution of the higher naphthenic acids. Journal of the
590   Institute of Petrol 43, 307-312.
591 Larter, S.R., Solli, H., Douglas, A.G., 1981. Phytol-containing Melanoidins and their Bearing
592   on the Fate of Isoprenoid Structures in Sediments, in: Bjoroy, M. (Ed.), International
593   Meeting on Organic Geochemistry. John Wiley & Sons, Bergen, Norway, pp. 513-
594   523.

595 Lengger, S.K., Scarlett, A.G., West, C.E., Frank, R.A., Hewitt, L.M., Milestone, C.B.,
596   Rowland, S.J., 2015. Use of the distributions of adamantane acids to profile short-
597   term temporal and pond-scale spatial variations in the composition of oil sands
598   process-affected waters. Environmental Science: Processes & Impacts.
599 Lengger, S.K., Scarlett, A.G., West, C.E., Rowland, S.J., 2013. Diamondoid diacids ('O4'
600   species) in oil sands process-affected water. Rapid Communications in Mass
601   Spectrometry 27, 2648-2654.
602 Marentette, J.R., Frank, R.A., Bartlett, A.J., Gillis, P.L., Hewitt, L.M., Peru, K.M., Headley,
603   J.V., Brunswick, P., Shang, D., Parrott, J.L., 2015. Toxicity of naphthenic acid
604   fraction components extracted from fresh and aged oil sands process-affected waters,
605   and commercial naphthenic acid mixtures, to fathead minnow (Pimephales promelas)
606   embryos. Aquatic Toxicology 164, 108-117.
607 McLafferty, F.W., 1963. Mass spectrometry of organic ions. Academic Press, New York.
608 Misiti, T.M., Tezel, U., Pavlostathis, S.G., 2014. Effect of Alkyl Side Chain Location and
609   Cyclicity on the Aerobic Biotransformation of Naphthenic Acids. Environmental
610   Science & Technology 48, 7909-7917.
611 Ortiz, X., Jobst, K.J., Reiner, E.J., Backus, S.M., Peru, K.M., McMartin, D.W., O'Sullivan,
612   G., Taguchi, V.Y., Headley, J.V., 2014. Characterization of Naphthenic Acids by Gas
613   Chromatography-Fourier Transform Ion Cyclotron Resonance Mass Spectrometry.
614   Analytical Chemistry 86, 7666-7673.
615 Pereira, A.S., Bhattacharjee, S., Martin, J.W., 2013. Characterization of Oil Sands Process-
616   Affected Waters by Liquid Chromatography Orbitrap Mass Spectrometry.
617   Environmental Science & Technology 47, 5504-5513.
618 Peters, K.E., Walters, C.C., Moldowan, J.M., 2005. The Biomarker Guide: Biomarkers and
619   Isotopes in the Environment and Human History. Cambridge University Press.
620 Petrov, A.A., 1987. Petroleum Hydrocarbons. Springer-Verlag, Berlin Heidelberg.
621 Piccolo, L., Nassreddine, S., Toussaint, G., Geantet, C., 2010. Discussion on "A
622   comprehensive two-dimensional gas chromatography coupled with quadrupole mass
623   spectrometry approach for identification of C10 derivatives from decalin" by C.
624   Flego, N. Gigantiello, W.O. Parker, Jr., V. Calemma [J. Chromatogr. A 1216 (2009)
625   2891]. Journal of Chromatography A 1217, 5872-5873.
626 Polyakova, A.A., Khramova, E.V., Bagrii, Y.I., Tsitsugina, N.N., Lukashenko, I.M., Frid,
627   T.Y., Sanin, P.I., 1973. Mass spectrometric study of alkyladamantanes. Petroleum
628   Chemistry U.S.S.R. 13, 1-10.
629 Porter, Q.N., 1985. Mass Spectrometry of Heterocyclic Compounds. Wiley.
630 Reichenbach, S.E., Kottapalli, V., Ni, M., Visvanathan, A., 2005. Computer language for
631   identifying chemicals with comprehensive two-dimensional gas chromatography and
632   mass spectrometry. Journal of Chromatography A 1071, 263-269.
633 Rowland, S.J., Clough, R., West, C.E., Scarlett, A.G., Jones, D., Thompson, S., 2011a.
634   Synthesis and mass spectrometry of some tri-and tetracyclic naphthenic acids. Rapid
635   Communications in Mass Spectrometry 25, 2573-2578.
636 Rowland, S.J., Pereira, A.S., Martin, J.W., Scarlett, A.G., West, C.E., Lengger, S.K., Wilde,
637   M.J., Pureveen, J., Tegelaar, E.W., Frank, R.A., Hewitt, L.M., 2014. Mass spectral
638   characterisation of a polar, esterified fraction of an organic extract of an oil sands
639   process water. Rapid Communications in Mass Spectrometry 28, 2352-2362.
640 Rowland, S.J., Scarlett, A.G., Jones, D., West, C.E., Frank, R.A., 2011b. Diamonds in the
641   rough: Identification of individual naphthenic acids in oil sands process water.
642   Environmental Science & Technology 45, 3154-3159.

643 Rowland, S.J., West, C.E., Scarlett, A.G., Jones, D., Boberek, M., Pan, L., Ng, M., Kwong,
644   L., Tonkin, A., 2011c. Monocyclic and monoaromatic naphthenic acids: synthesis and
645   characterisation. Environmental Chemistry Letters 9, 525-533.
646 Rowland, S.J., West, C.E., Scarlett, A.G., Jones, D., Frank, R.A., 2011d. Identification of
647   individual tetra- and pentacyclic naphthenic acids in oil sands process water by
648   comprehensive two-dimensional gas chromatography/mass spectrometry. Rapid
649   Communications in Mass Spectrometry 25, 1198-1204.
650 Scarlett, A.G., Reinardy, H.C., Henry, T.B., West, C.E., Frank, R.A., Hewitt, L.M., Rowland,
651   S.J., 2013. Acute toxicity of aromatic and non-aromatic fractions of naphthenic acids
652   extracted from oil sands process-affected water to larval zebrafish. Chemosphere 93,
653   415-420.
654 Scott, A.C., Zubot, W., MacKinnon, M.D., Smith, D.W., Fedorak, P.M., 2008. Ozonation of
655   oil sands process water removes naphthenic acids and toxicity. Chemosphere 71, 156-
656   160.
657 Seifert, W.K., 1975. Carboxylic acids in petroleum and sediments. Fortschritte der Chemie
658   Organischer Naturstoffe 32, 1-49.
659 Seifert, W.K., Teeter, R.M., Howells, W.G., Cantow, M.J.R., 1969. Analysis of crude oil
660   carboxylic acids after conversion to their corresponding hydrocarbons. Analytical
661   Chemistry 41, 1638-1647.
662 Sinninghe Damsté, J.S., Kock-van Dalen, A.C., de Leeuw, J.W., 1988. Identification of long-
663   chain isoprenoid alkylbenzenes in sediments and crude oils. Geochimica et
664   Cosmochimica Acta 52, 2671-2677.
665 Swigert, J.P., Lee, C., Wong, D.C.L., White, R., Scarlett, A.G., West, C.E., Rowland, S.J.,
666   2015. Aquatic hazard assessment of a commercial sample of naphthenic acids.
667   Chemosphere 124, 1-9.
668 Thomas, K.V., Langford, K., Petersen, K., Smith, A.J., Tollefsen, K.E., 2009. Effect-Directed
669   Identification of Naphthenic Acids As Important in Vitro Xeno-Estrogens and Anti-
670   Androgens in North Sea Offshore Produced Water Discharges. Environmental
671   Science & Technology 43, 8066-8071.
672 Vorob'eva, N.S., Zemskova, Z.K., Pekh, T.I., Petrov, A.A., 1986. Diterpenoid tetracyclic
673   hydrocarbons of petroleum. Petroleum Chemistry U.S.S.R. 26, 69-76.
674 Wang, G., Shi, S., Wang, P., Wang, T.G., 2013. Analysis of diamondoids in crude oils using
675   comprehensive two-dimensional gas chromatography/time-of-flight mass
676   spectrometry. Fuel 107, 706-714.
677 Wang, Z., Yang, C., Hollebone, B., Fingas, M., 2006. Forensic Fingerprinting of
678   Diamondoids for Correlation and Differentiation of Spilled Oil and Petroleum
679   Products. Environmental Science & Technology 40, 5636-5646.
680 Wei, Z., Michael Moldowan, J., Dahl, J., Goldstein, T.P., Jarvie, D.M., 2006. The catalytic
681   effects of minerals on the formation of diamondoids from kerogen macromolecules.
682   Organic Geochemistry 37, 1421-1436.
683 West, C.E., Scarlett, A.G., Tonkin, A., O'Carroll-Fitzpatrick, D., Pureveen, J., Tegelaar, E.,
684   Gieleciak, R., Hager, D., Petersen, K., Tollefsen, K.-E., Rowland, S.J., 2014.
685   Diaromatic sulphur-containing 'naphthenic' acids in process waters. Water Research
686   51, 206-215.
687 Wilde, M.J., Rowland, S.J., 2015. Structural Identification of Petroleum Acids by Conversion
688   to Hydrocarbons and Multidimensional Gas Chromatography-Mass Spectrometry.
689   Analytical Chemistry 87, 8457-8465.
690 Wilde, M.J., West, C.E., Scarlett, A.G., Jones, D., Frank, R.A., Hewitt, L.M., Rowland, S.J.,
691   2015. Bicyclic naphthenic acids in oil sands process water: Identification by

692   comprehensive multidimensional gas chromatography–mass spectrometry. Journal of
693   Chromatography A 1378, 74-87.
694 Wingert, W.S., 1992. G.c.-m.s. analysis of diamondoid hydrocarbons in Smackover
695   petroleums. Fuel 71, 37-43.
696 Yi, Y., Birks, S.J., Cho, S., Gibson, J.J., 2015. Characterization of organic composition in
697   snow and surface waters in the Athabasca Oil Sands Region, using ultrahigh
698   resolution Fourier transform mass spectrometry. Science of the Total Environment
699   518–519, 148-158.
700 Yue, S., Ramsay, B.A., Brown, R.S., Wang, J., Ramsay, J.A., 2015a. Identification of
701   Estrogenic Compounds in Oil Sands Process Waters by Effect Directed Analysis.
702   Environmental Science & Technology 49, 570-577.
703 Yue, S., Ramsay, B.A., Wang, J., Ramsay, J., 2015b. Toxicity and composition profiles of
704   solid phase extracts of oil sands process-affected water. Science of the Total
705   Environment 538, 573-582.

706

Scheme 1: Route of conversion of the OSPW NA samples #1 and #2, to the corresponding

hydrocarbons.
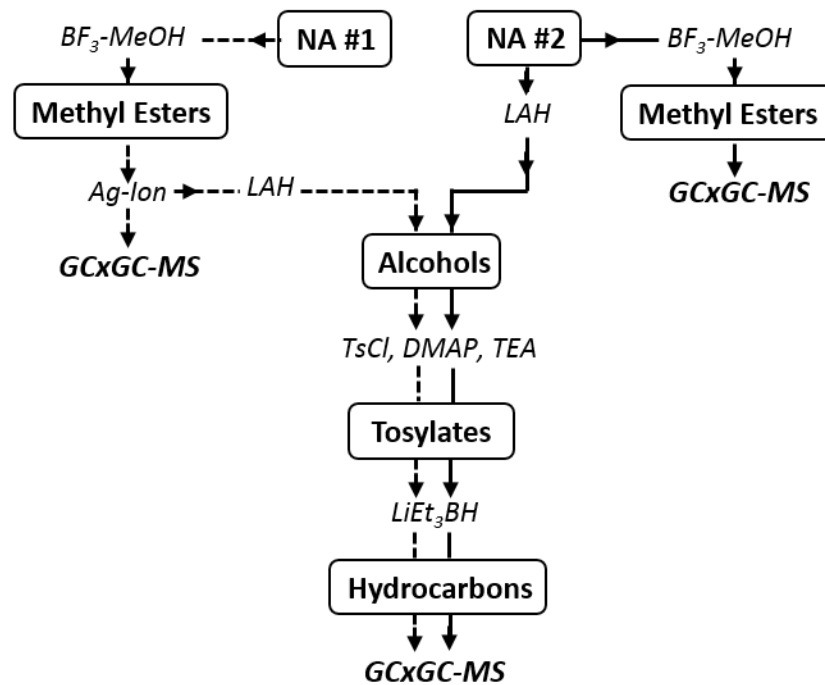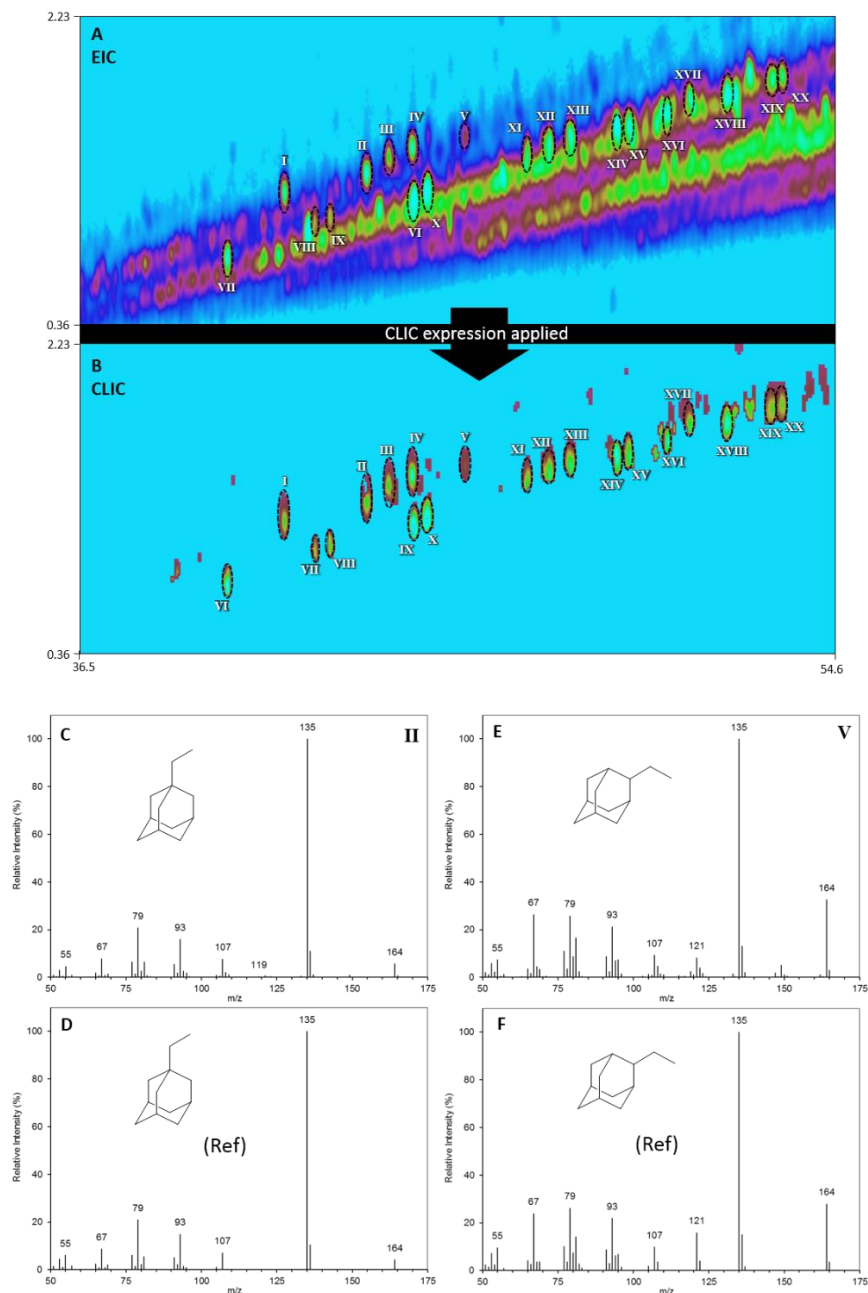
Figure 1: Comparison of (A) an EIC (*m/z* 135, 149, 163 and 178) and (B) a simplified chromatogram after a CLIC expression was applied of the reduced unfractionated OSPW sample (#2), clearly showing the presence of three dimethyl- (Ad-I, -III and -IV), (C-F) two ethyl- (-II and -V), eight trimethyl- (-VI-VIII, -X-XIII and -XVII) and seven ethylmethyl-adamantane isomers (-IX, -XIV-XVI and -XVIII-XX), based on comparison with reference mass spectra and known elution order (Polyakova et al., 1973; Wingert, 1992; Wang et al., 2013).
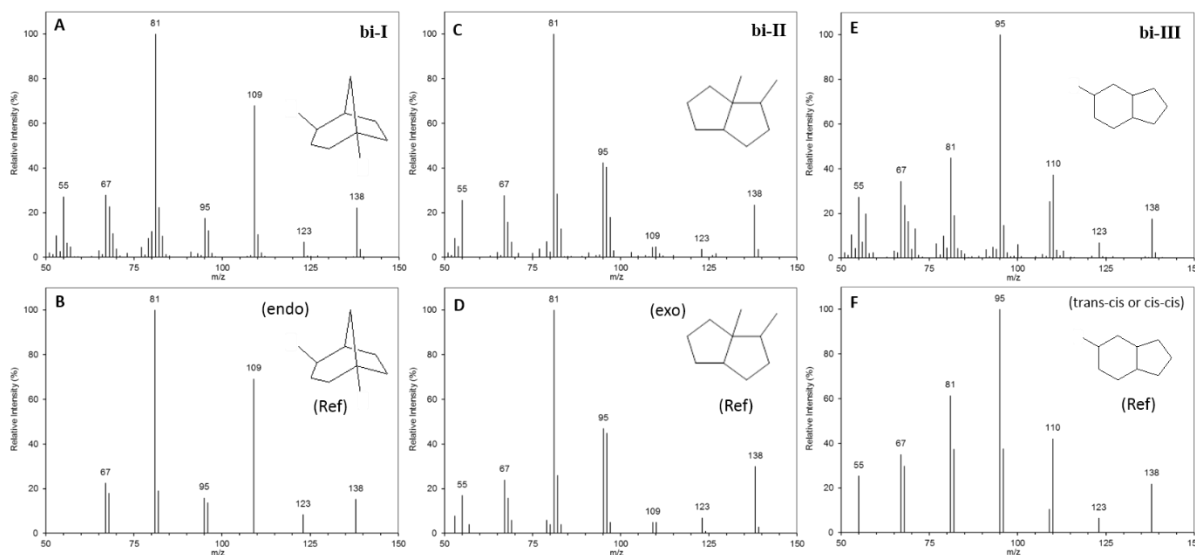
Figure 2: Identification of bridged and fused bicyclanes by mass spectral comparison with reference spectra and elution order (Denisov et al., 1977a; Denisov et al., 1977b; Denisov et al., 1977c).
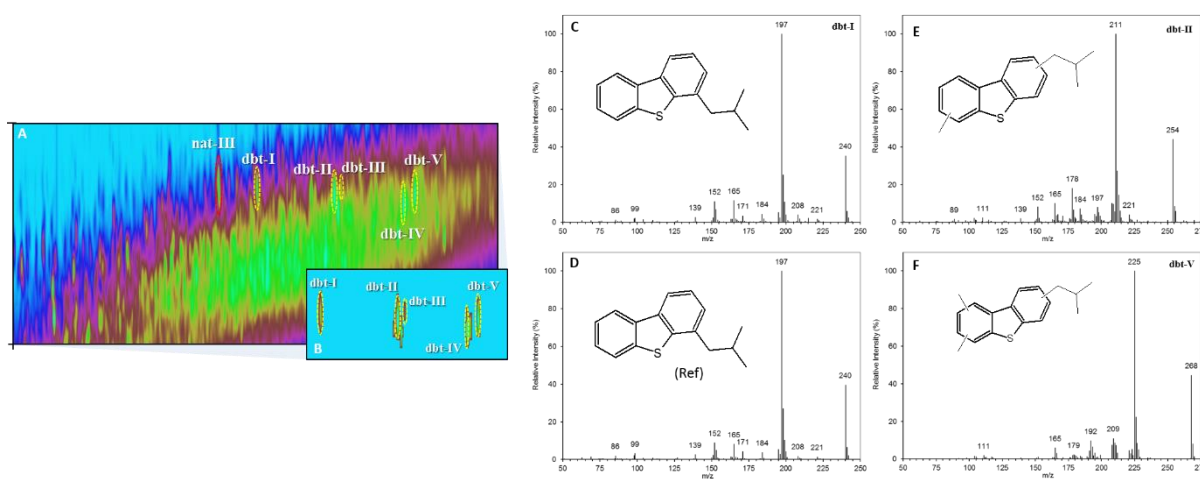
Figure 3: (A-B) TIC and CLIC EIC of reduced 'aromatic/sulphur' fraction (sample #1), showing isomers dbt-I to -V assigned as the reduced hydrocarbons of the five methyl esters reported by West et al. (2014). (C) Identification of dbt-I after comparison with (D) synthesised 4-isobutyldibenzothiophene and (E-F) assignment of dbt-II and −V as methyl- and dimethyl- isobutyldibenothiophenes.

SUPPORTING INFORMATION FOR PUBLICATION

Naphthenic acids in oil sands process waters: Identification by conversion of the acids or esters to hydrocarbons

Michael J. Wilde† and Steven J. Rowland*

*Petroleum and Environmental Geochemistry Group, Biogeochemistry Research Centre, Plymouth University,

PL4 8AA, UK, fax: +44(0)1752 584710, e-mail: srowland@plym.ac.uk

†Department of Chemistry, University of Leicester, Leicester, UK, LE1 7RH; mjw77@le.ac.uk

SUPPORTING INFORMATION TABLE OF CONTENTS

EXPERIMENTAL DETAILS

TABLES

FIGURES

REFERENCES (P. S-23)

SUPPORTING INFORMATION EXERPIMENTAL DETAILS

4.1. GCxGC-MS

The GC×GC-MS conditions set A involved the primary oven programmed from 30 °C, held for 1 min, then heated to 120 °C at 5 °C/min, to 220 °C at 0.8°C/min, to 280 °C at 5 °C/min and to 320 °C at 10 °C/min and then held for 10 min. The secondary oven was programmed to track the primary oven at 40 °C above. The hot jet was programmed to start 30 °C above the primary oven temperature until 150 °C, it was then ramped to 260 °C at 1.3 °C/min and then to 400 °C at 4 °C/min. The modulation period was set 4 and 6 s. The GC×GC-MS conditions set B involved the primary oven programmed from 40 °C, held for 1.3 min, then heated to 200 °C at 1 °C/min, to 280 °C and 5 °C/min and to 320 °C at 10 °C/min and then held for 5 min. The secondary oven was programmed to track the primary oven at 20 °C above. The hot jet was programmed to start 20 °C above the primary oven and finish 100 °C above the primary oven over the period of the run; programmed from 60 °C, held for 1 min, then heated to 200 °C at 1.1 °C/min, to 280 °C at 3 °C/min and then to 400 °C at 5 °C/min. The modulation period was set at 4 and 6 s.

4.2. Derivatisation and fractionation of sample #1

A concentrated naphthenate solution, extracted from oil sands process-affected water (OSPW), which had undergone a prior clean-up procedure involving weak anion exchange chromatography, was received from Environment Canada, Burlington, CA. The extraction and clean-up procedure was developed and reported by Frank et al. (2006).

The free NA (537 mg), extracted from the concentrated OSPW naphthenate solution as described by Jones et al. (2012), was dissolved 14 % BF3-methanol complex solution (Sigma Aldrich) and heated at 70 °C for 3 hours. The methylated solution was washed with water (Chromasolv® HPLC Grade, Sigma Aldrich) and the methyl esters extracted with hexane (HPLC Grade, Rathburns Chemical Ltd.). The methyl ester solution was dried over anhydrous Na_2SO_4 ($\geq$ 99.0 %, Sigma Aldrich), filtered and evaporated to dryness under N_2 at 40 °C.

The large scale fractionation was performed twice, on two quantities of sample #1. The mass of OSPW NA methyl esters loaded onto the phase in the first fractionation was 296 mg, and 300 mg in the second fractionation. A summary of the fractions for each fractionation is given in Tables S1 and S2.

TABLES

Table S-1: Summary of the fractions collected in the first Ag-Ion fractionation of the OSPW NA methyl esters.

Fraction	Eluent	Mass of eluate / mg
Wash	Hexane	0.4
1	100 % hexane	20.1
2	100 % hexane	76.8
3	100 % hexane	7.4
4	100 % hexane	1.7
5	5 % diethyl ether : 95 % hexane	61.3
6	5 % diethyl ether : 95 % hexane	30.8
7	5 % diethyl ether : 95 % hexane	9.2
8	10 % diethyl ether : 90 % hexane	11.5
9	100 % diethyl ether	29.6
10	100 % methanol	52.1
Total		300.9
Total (excluding F10)		248.8

Table S-2: Summary of the fractions collected in the second Ag-Ion fractionation of the OSPW NA methyl esters.

Fraction	Eluent	Mass of eluate / mg
Wash	Hexane	0.3
1	100 % hexane	7.7
2	100 % hexane	96.7
3	100 % hexane	16.7
4	100 % hexane	2.1
5	5 % diethyl ether : 95 % hexane	50.0
6	5 % diethyl ether : 95 % hexane	38.2
7	5 % diethyl ether : 95 % hexane	12.9
8	10 % diethyl ether : 90 % hexane	15.1
9	100 % diethyl ether	36.1
10	100 % methanol	222.9
Total		498.4
Total (excluding F10)		275.5

Table S-3: Summary of the masses and yields for the conversion of the Ag-Ion fractions 2, 5 and 7 of the derivatised OSPW NA and of the underivatised, unfractionated OSPW NA samples, to hydrocarbons.

Starting Reactant		Starting mass and yield / mg (approx. percentage yield / %)					
		LiAlH$_4$ Reduction		Tosylation		Super-Hydride® Reduction	
Derivatised (# sample number)	Fraction	NA methyl esters / mg	Alcohols / mg	Alcohols / mg	Tosylates / mg	Tosylates / mg	Hydrocarbons / mg
Yes (#1)	2	10[†]	9.8 (104 %[‡])	9.8	13.3 (82 %[‡])	12.2	20.7[*]
Yes (#1)	5	25[†]	23.0 (97 %[‡])	23.0	27.9 (77 %[‡])	27.9	17.2[*]
Yes (#1)	7	10[†]	9.1 (96 %[‡])	9.1	10.5 (74 %[‡])	10.5	6.7[*]
No (#2)	Whole	30[†]	25.0 (88 %[‡])	25.0	32.3 (78 %[‡])	32.0	-

[†] approximate masses, as methyl esters

[‡] approximate yield based on average molecular weight of acid methyl esters as 250 g/mol in F2, 286 g/mol in F5, 294 g/mol in F7 and 246 g/mol in the whole underivatised OSPW NA

[*] hydrocarbon product concentrated by N$_2$ blowdown resulting in evaporative losses; THF and boroxin residue present

[-] Kuderna-Danish apparatus used for concentration to reduce volatile compound losses, so no weight was recorded

34

35

Table S-4: Summary of naphthenic acids within OSPW identified as hydrocarbons after reduction of the esters and analysis by GCxGC-MS

Assigned label	Type	Hydrogen Deficiency of hydrocarbon or NA (z)	Name of hydrocarbon equivalent	Identification
Ad-I	Tricyclic (diamondoid)	-6	1,2-dimethyladamantane	Mass spectral interpretation, Golovkina et al. (1984); mass spectral comparison, Polyakova et al. (1973); GC elution order, Wingert (1992), Wei et al. (2006), Wang et al. (2013); previous assignment of NA in OSPW and comparison of retention position and mass spectra with reference NA, Rowland et al. (2011a), Bowman et al. (2014), Lengger et al. (2013)
Ad-II	Tricyclic (diamondoid)	-6	1-ethyladamantane	
Ad-III	Tricyclic (diamondoid)	-6	cis-2,4-dimethyladamantane	
Ad-IV	Tricyclic (diamondoid)	-6	trans-2,4-dimethyladamantane	
Ad-V	Tricyclic (diamondoid)	-6	2-ethyladamantane	
Ad-VI-VIII, X-XIII, XVII	Tricyclic (diamondoid)	-6	trimethyladamantane isomers	
Ad-IX, XIV-XVI, XVIII-XX	Tricyclic (diamondoid)	-6	ethylmethyladamantane isomers	
bi-I	Bicyclic	-4	1,4-dimethylbicyclo[3.2.1]octane	Mass spectral comparison, Denisov et al. (1977a); GC elution order Piccolo et al. (2010); previous assignment in petroleum NA and comparison of retention position and mass spectra with reference NA, Wilde and Rowland (2015), Wilde et al. (2015)
bi-II	Bicyclic	-4	1,2-dimethylbicyclo[3.3.0]octane	Mass spectral comparison, Denisov et al. (1977c); GC elution order Piccolo et al. (2010); previous assignment of isomers in petroleum NA and comparison of retention position and mass spectra with reference NA, Wilde and Rowland (2015), Wilde et al. (2015)
bi-III	Bicyclic	-4	3-methylbicyclo[4.3.0]nonane	Mass spectral comparison, Denisov et al. (1977b); GC elution order Piccolo et al. (2010); previous assignment of isomers in petroleum NA and comparison of retention position and mass spectra with reference NA, Wilde and Rowland (2015), Wilde et al. (2015)
bi-IV	Bicyclic	-4	2,6-dimethylbicyclo[3.2.1]octane	Mass spectral comparison, Denisov et al. (1977a); GC elution order Piccolo et al. (2010); previous assignment of isomers in petroleum NA, Wilde and

|				Rowland (2015) (tentative)
bi-V and VI	Bicyclic	-4	C_{10} dimethylbicyclooctane isomers	Mass spectral comparison, Denisov et al. (1977a) (tentative)
tt-I and II	Tricyclic	-6	C_{20} tricylic terpane; cheilanthane isomers	Mass spectral interpretation and mass spectral comparison, Hall and Douglas (1981), Philp (1985), Chicarelli et al. (1988); previous evidence of terpenoid NA in oil sands, Cyr and Strausz (1983)
tt-III and IV	Tricyclic	-6	C_{21} tricyclic terpane; cheilanthane isomers	
tt-V	Tricyclic	-6	C_{23} tricyclic terpane; cheilanthane isomer	
tt-VI	Tricyclic	-6	C_{24} tricyclic terpane; cheilanthane isomer	
diA-I	Pentacyclic	-10	3-methyldiamantane	Previous assignment of NA isomers in OSPW, Rowland et al. (2011b); mass spectral interpretation, Golovkina et al. (1984) and mass spectral comparison, Kuraš and Hála (1970), Musayev et al. (1983); GC elution order, Wingert (1992), Wang et al. (2013) (higher homologues are tentative)
diA-II	Pentacyclic	-10	ethyldiamantane	
diA-III-IX	Pentacyclic	-10	dimethyl- and ethyldiamantane isomers	
diA-X-XVI	Pentacyclic	-10	ethylmethyldiamantane isomers	
diA-XVII-XXV	Pentacyclic	-10	dimethylethyldiamantane isomers	
tAd-I	Tetracyclic	-8	methyl-2,4-cyclopentano-adamantane	Mass spectral interpretation and evidence of corresponding NA in OSPW, Petrov (1987), Vorob'eva et al. (1986) (higher homologues are tentative)
tAd-II and III	Tetracyclic	-8	ethyl-2,4-cyclopentano-adamantane isomers	
tAd-IV	Tetracyclic	-8	dimethyl-cyclopentano-adamantane	
tAd-V	Tetracyclic	-8	ethylmethyl-cyclopentano-adamantane	
tAd-VI	Tetracyclic	-8	trimethyl-cyclopentano-adamantane	
tAd-VII-IX	Tetracyclic	-8	C_{16-18} alkylcyclopentano-adamantane isomers	
ma-I-IV	Monocylic monoaromatic	-8	C_{16-18} alkylbenzene isomers	Mass spectral interpretation and comparison, Larter et al. (1981), Sinninghe Damsté et al. (1988), Ji-Zhou et al. (1993); evidence of corresponding NA in OSPW
mb-I-VI	Bicyclic monoaromatic	-10	C_{14-19} alkylindane and tetralin isomers	
tm-I	Tricyclic monoaromatic	-12	13-methyl-14-ethylpodocarpa-8,11,13-triene	Mass spectral comparison, Azevedo et al. (1990), Azevedo et al. (1992) and NIST MS library; previous assignment of monoaromatic terpenoid NA in tasmanite, Azevedo et al. (1994)
dbt-I	Sulphur-containing tricyclic diaromatic	-18	4-isobutyldibenzothiophene	Comparison of retention position with synthetic reference hydrocarbon, Wilde (2015); mass spectral interpretation and mass spectral comparison, Andersson et al. (2001), Bobinger et al. (1999), (Andersson and Schade, 2004), Li et al. (2014); previous assignment of NA

				isomers in OSPW, West et al. (2014)
dbt-II and III	Sulphur-containing tricyclic diaromatic	-18	methyl-isobutyldibenzothiophene isomers	Mass spectral interpretation and mass spectral comparison, Andersson et al. (2001), Bobinger et al. (1999), Andersson and Schade (2004), Li et al. (2014); previous assignment of NA isomers in OSPW, West et al. (2014)
dbt-IV and V	Sulphur-containing tricyclic diaromatic	-18	dimethyl-isobutyldibenzothiophene	
nat-I-V	Sulphur-containing tricyclic diaromatic	-16	alkyl 2H-naphtho[1,8-bc]thiophenes isomers	Mass spectral interpretation and comparison, (Porter, 1985), (Hawthone and Porter, 1968) and NIST MS library

38

39   **FIGURES**
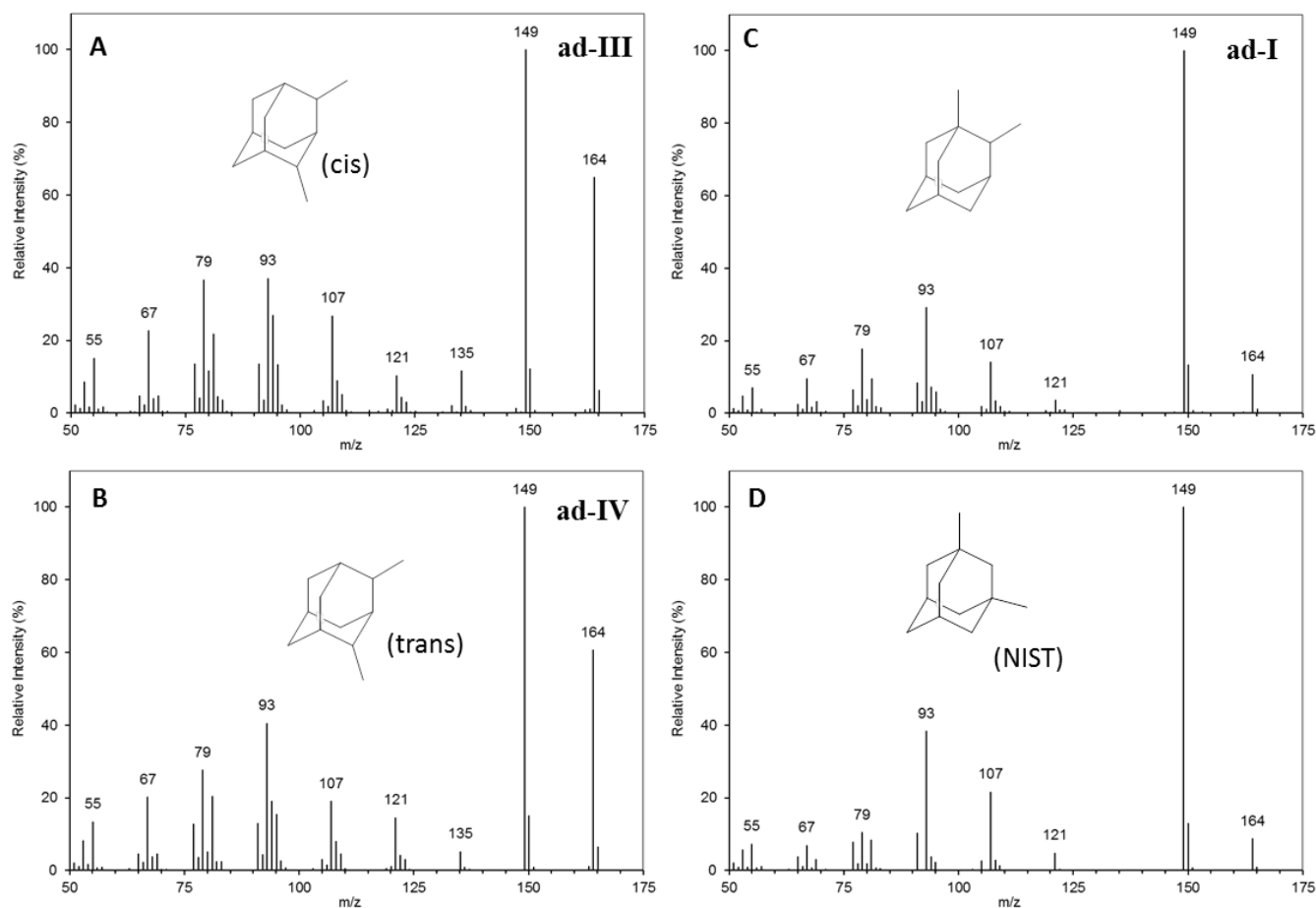
40

41   Figure S-1. Mass spectra of isomers ad-I, III and IV (labels refer to components in EIC in Figure 1 in main text) assigned as
42   dimethyladamantanes including (A and B) two isomers of 2,4-dimethyladamantane (III and IV) and (C) 1,2-
43   dimethyladamantane (I) by comparison of the known GC elution order of alkyl adamantanes relative to 1- and 2-
44   ethyladamantane (II and V) and by comparison with (D) the reference mass spectrum of 1,3-dimethyladamantane
45   (Polyakova et al., 1973; Wingert, 1992; Wang et al., 2013).

46
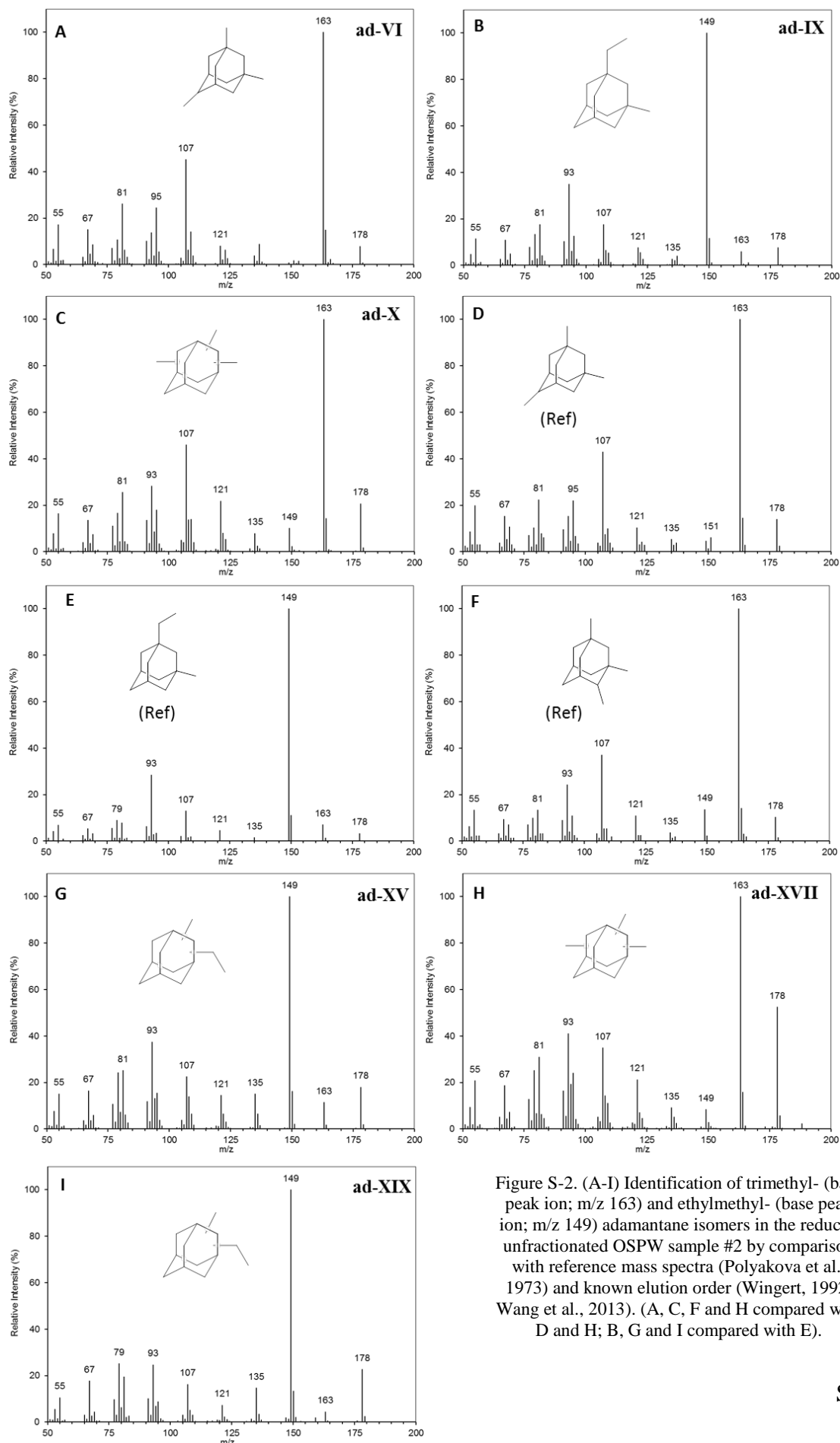
Figure S-2. (A-I) Identification of trimethyl- (base peak ion; m/z 163) and ethylmethyl- (base peak ion; m/z 149) adamantane isomers in the reduced unfractionated OSPW sample #2 by comparison with reference mass spectra (Polyakova et al., 1973) and known elution order (Wingert, 1992; Wang et al., 2013). (A, C, F and H compared with D and H; B, G and I compared with E).
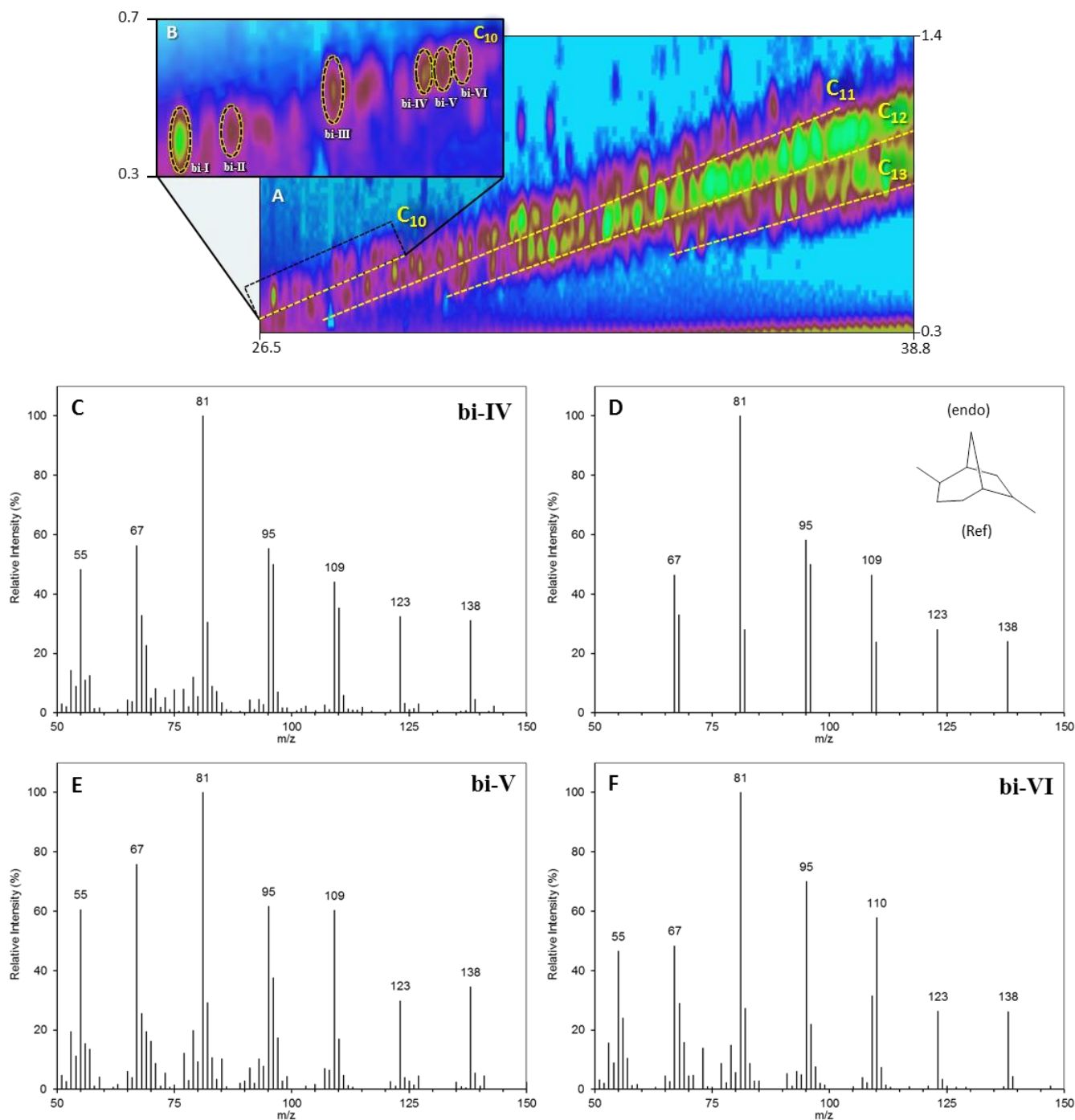
48
49
50　Figure S-3. Mass spectra of C_{10} bicyclanes (bi-IV to VI) in the reduced unfractionated NA (sample #2). (A) Isomer bi-IV
51　assigned after comparison with (B) the reference mass spectrum of endo-endo-2,6-dimethylbicyclo[3.2.1]octane. (C and D)
52　Mass spectra of isomers bi-V and VI which did not match any reference spectra and were significantly different to those
53　within reduced petroleum-derived acids previously reported (Wilde and Rowland, 2015).
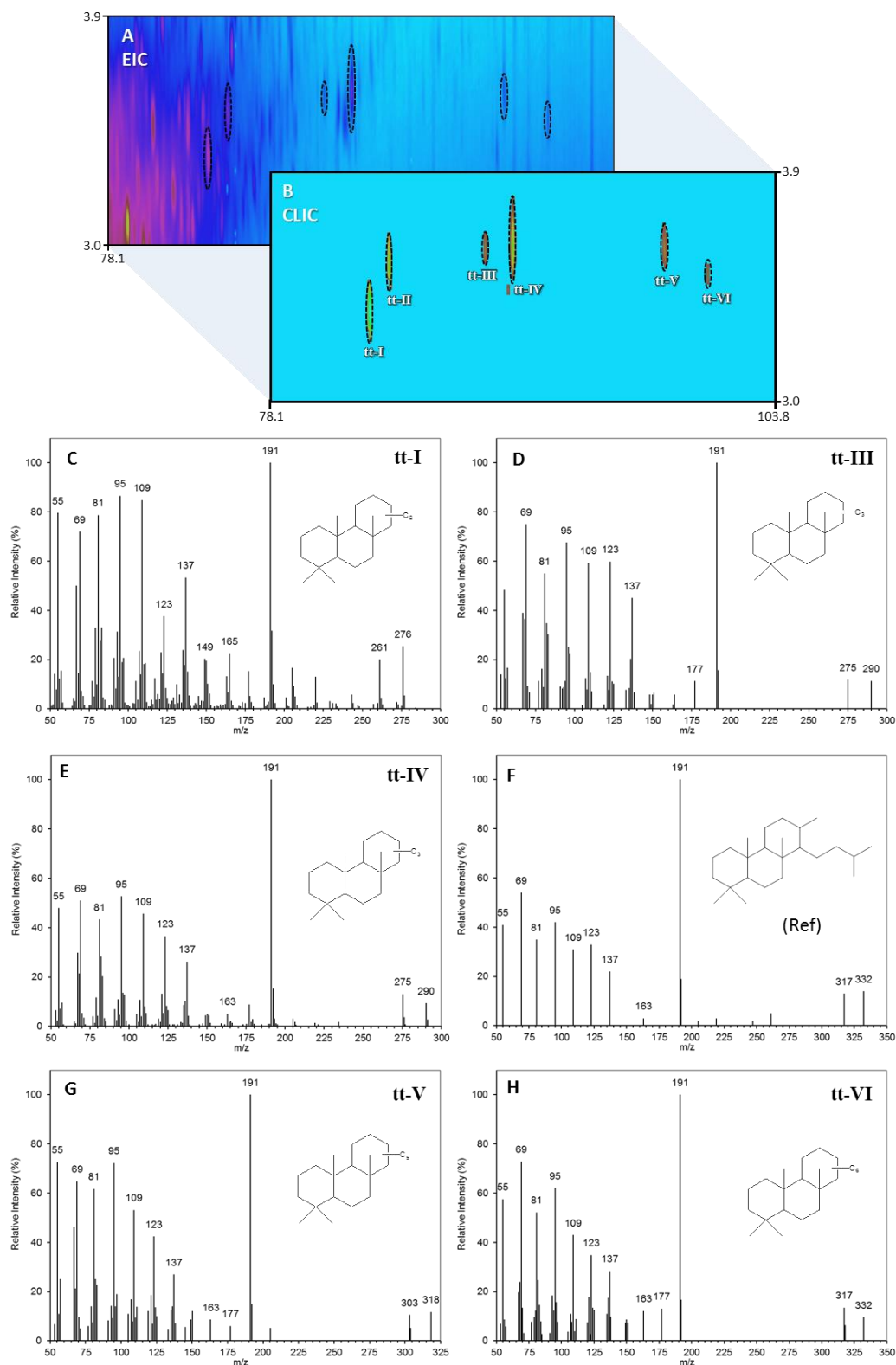
Figure S-4. (A) EIC (m/z 191) and (B) CLIC expression chromatogram revealing series of isomers assigned as (C-H) C_{20} (tt-I and II), C_{21} (tt-III and -IV), C_{23} (tt-V) and C_{24} (tt-VI) tricyclic terpanes by comparison with reference mass spectra (Hall and Douglas, 1981; Cyr and Strausz, 1983; Philp, 1985; Chicarelli et al., 1988). (All mass spectra share similar fragmentation pattern to reference mass spectrum F, particularly H and F).
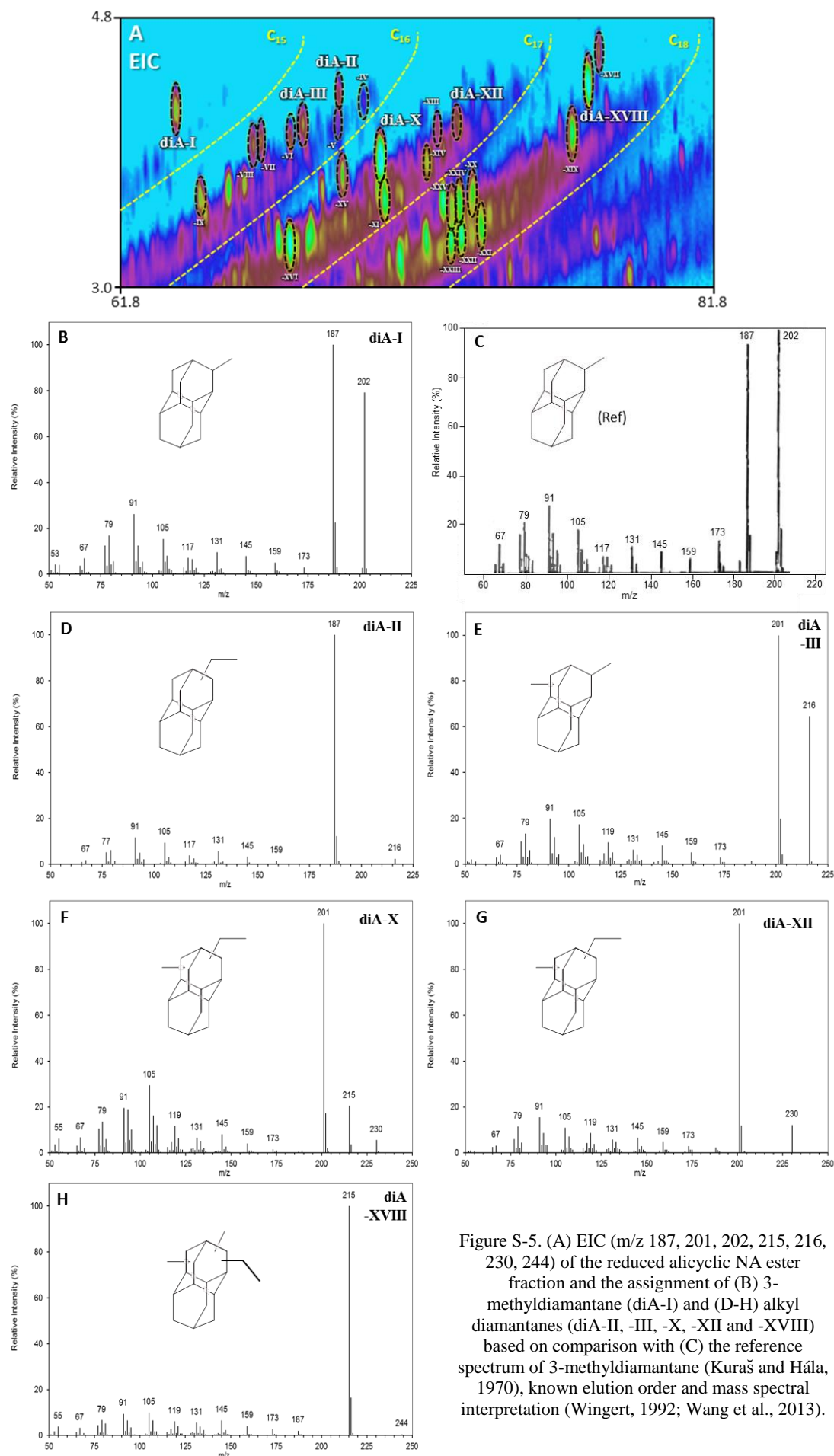
Figure S-5. (A) EIC (m/z 187, 201, 202, 215, 216, 230, 244) of the reduced alicyclic NA ester fraction and the assignment of (B) 3-methyldiamantane (diA-I) and (D-H) alkyl diamantanes (diA-II, -III, -X, -XII and -XVIII) based on comparison with (C) the reference spectrum of 3-methyldiamantane (Kuraš and Hála, 1970), known elution order and mass spectral interpretation (Wingert, 1992; Wang et al., 2013).

(A)   **(B)**   **(C)**   **(D)**
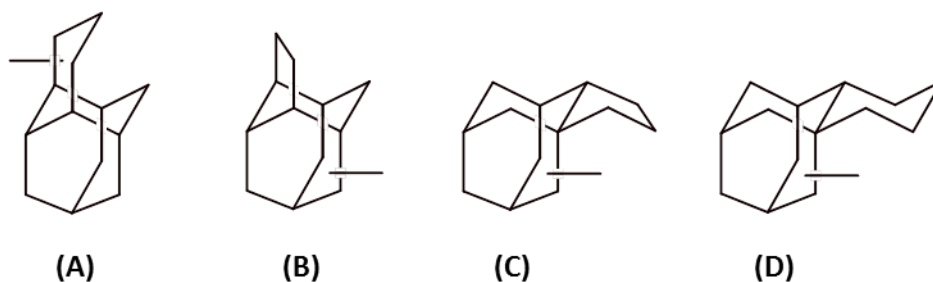
Figure S-6. Examples of substituted tetracyclic diamondoid compounds showing the structure of the skeletal core, including (A) 2,4-cyclohexano-adamantane (tetracyclic ring-opened diamantane), (B) 2,4-cyclopentano-adamantane, (C) 1,2-cyclopentano-adamantane and (D) 1,2-cyclohexano-adamantane.
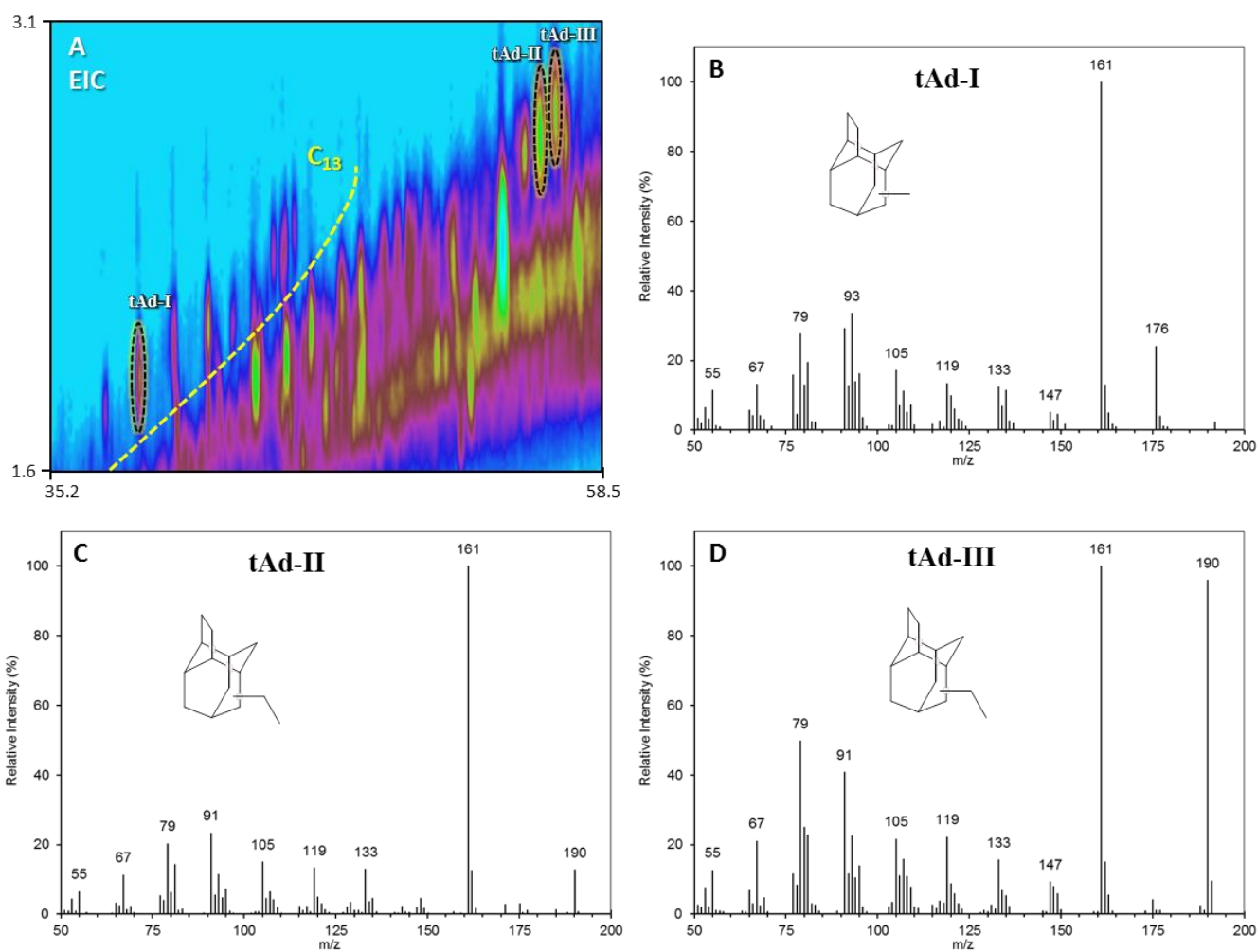
Figure S-7. (A) EIC (m/z 161, 175, 176, 189, 190, 203, 217) of the reduced sample #2, showing components tAd-1 to –III with (B-D) mass spectra containing molecular ions corresponding to C_{13} and C_{14} tetracyclic hydrocarbons, tentatively assigned as alkyl 2,4-cyclopentano-adamantanes.

72
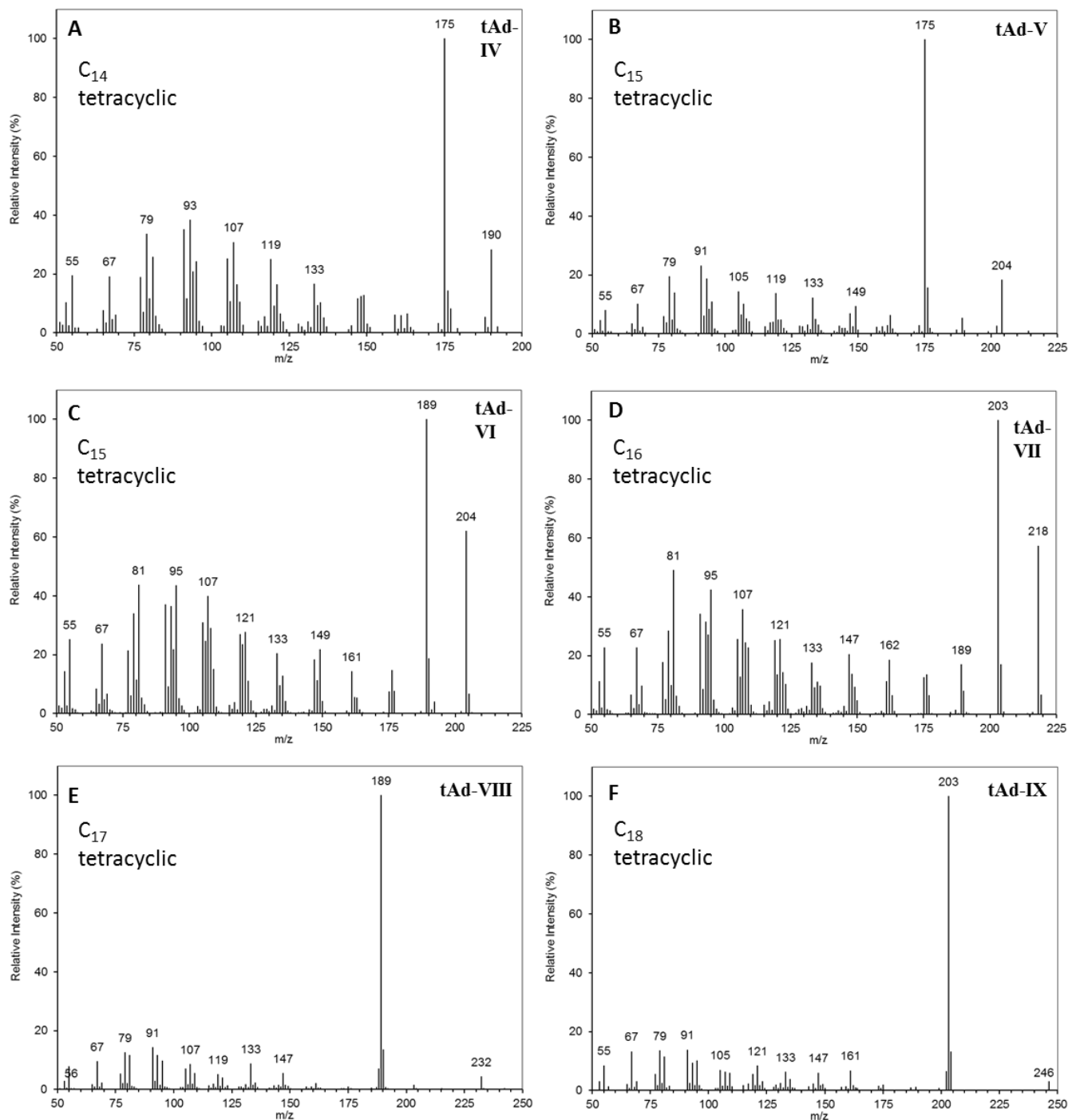
Figure S-8. (A-F) Mass spectra of C_{14-18} alkyl tetracyclic hydrocarbons (isomers tAd-IV to – IX) in the reduced sample #2, postulated to possess adamantanoid structures (e.g. Figure S-6; A-D).
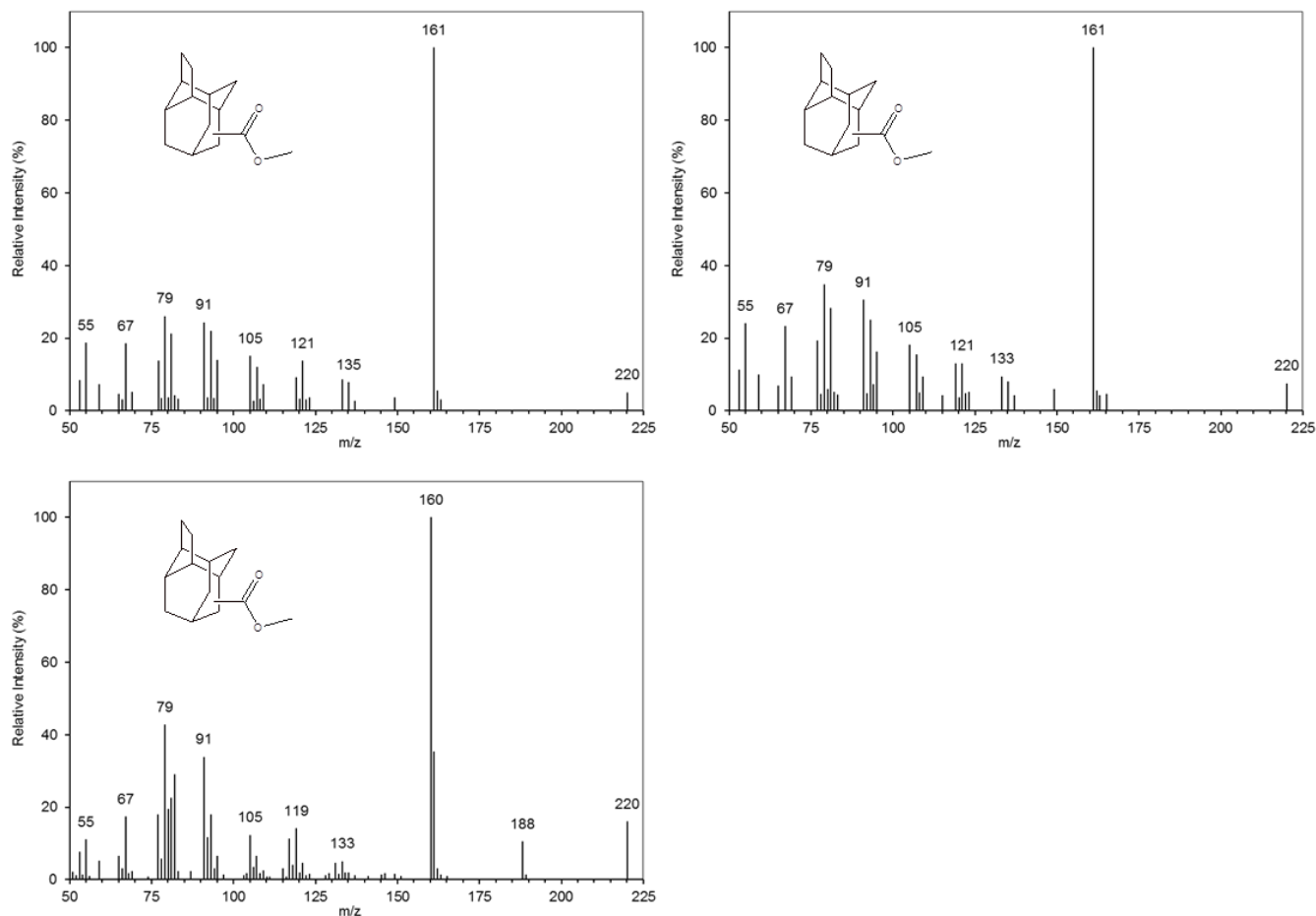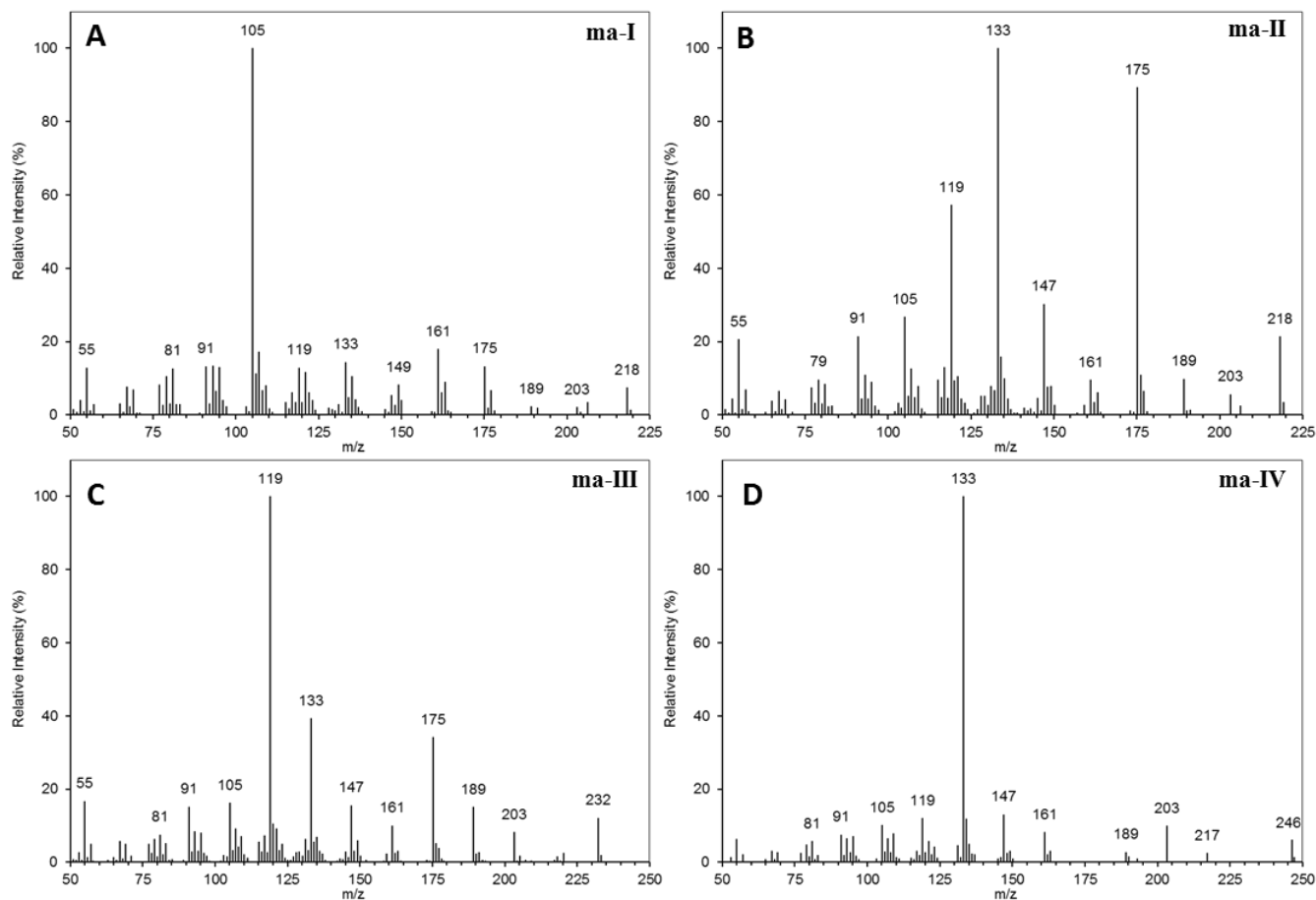
Figure S-9. Electron ionisation mass spectra of components in sample #2 NA methyl esters tentatively assigned as cyclopentano-adamantane acids after examination of the mass spectra of the corresponding hydrocarbons.
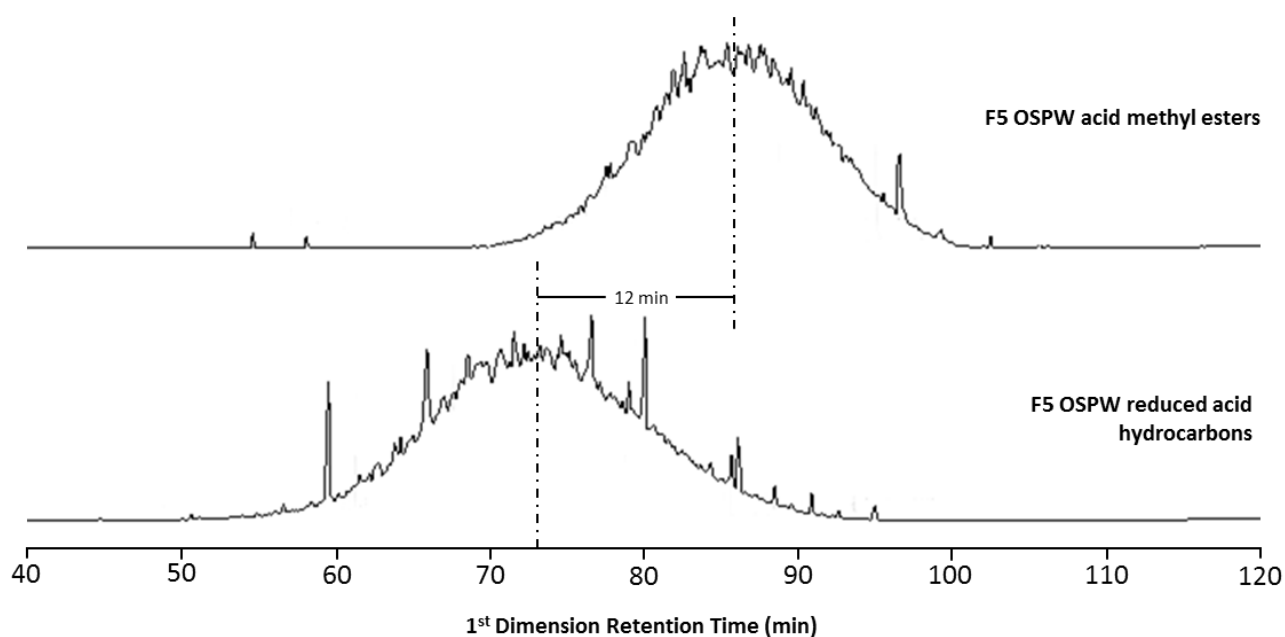
81

Figure S-10. (A-D) Example mass spectra of components assigned as C_{16-18} alkylbenzenes within the reduced aromatic fraction of sample #1.
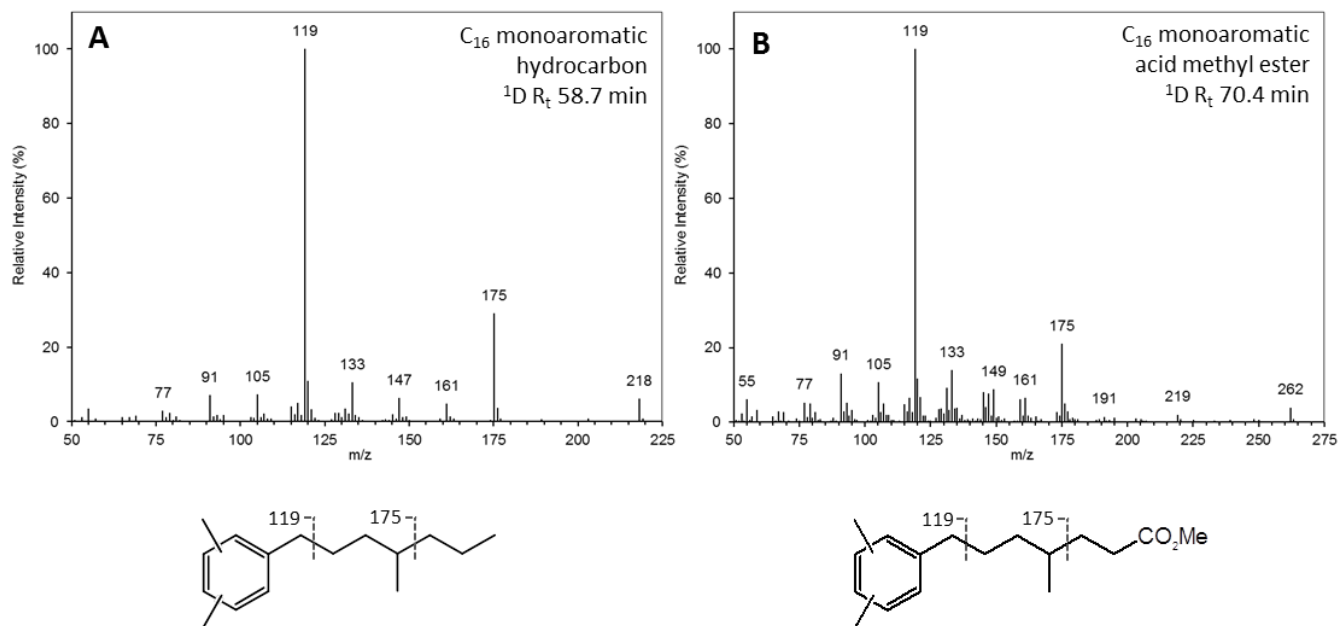
82
83

84

85

Figure S-11. Comparison of the GC-MS TIC chromatograms before (top) and after (bottom) reduction of the 'aromatic' fraction of sample #1 acid methyl esters, to the corresponding hydrocarbons.

86
87

88

89 Figure S-12. (A) Mass spectrum of a C_{16} monoaromatic hydrocarbon in the reduced aromatic fraction of sample #1, assigned
90   as a dimethyl branched alkylbenzene, based on comparison with reference spectra and mass spectral interpretation and (B)
91   the mass spectrum of a C_{16} monoaromatic acid methyl ester, assigned as the corresponding alkylbenzene acid methyl ester
92   before reduction of the aromatic fraction. Structures given are speculative, to demonstrate key fragment ions observed.

93

94
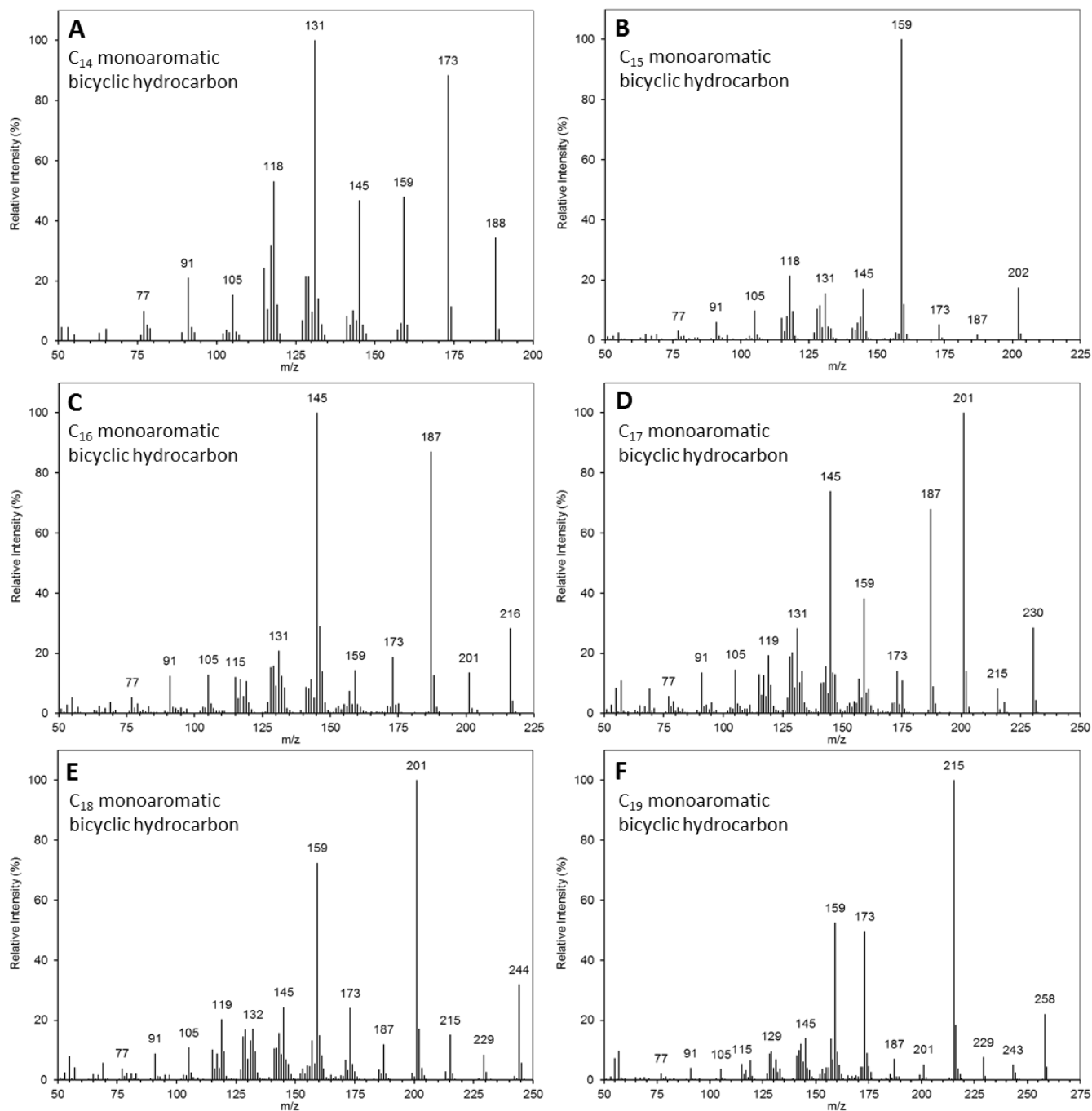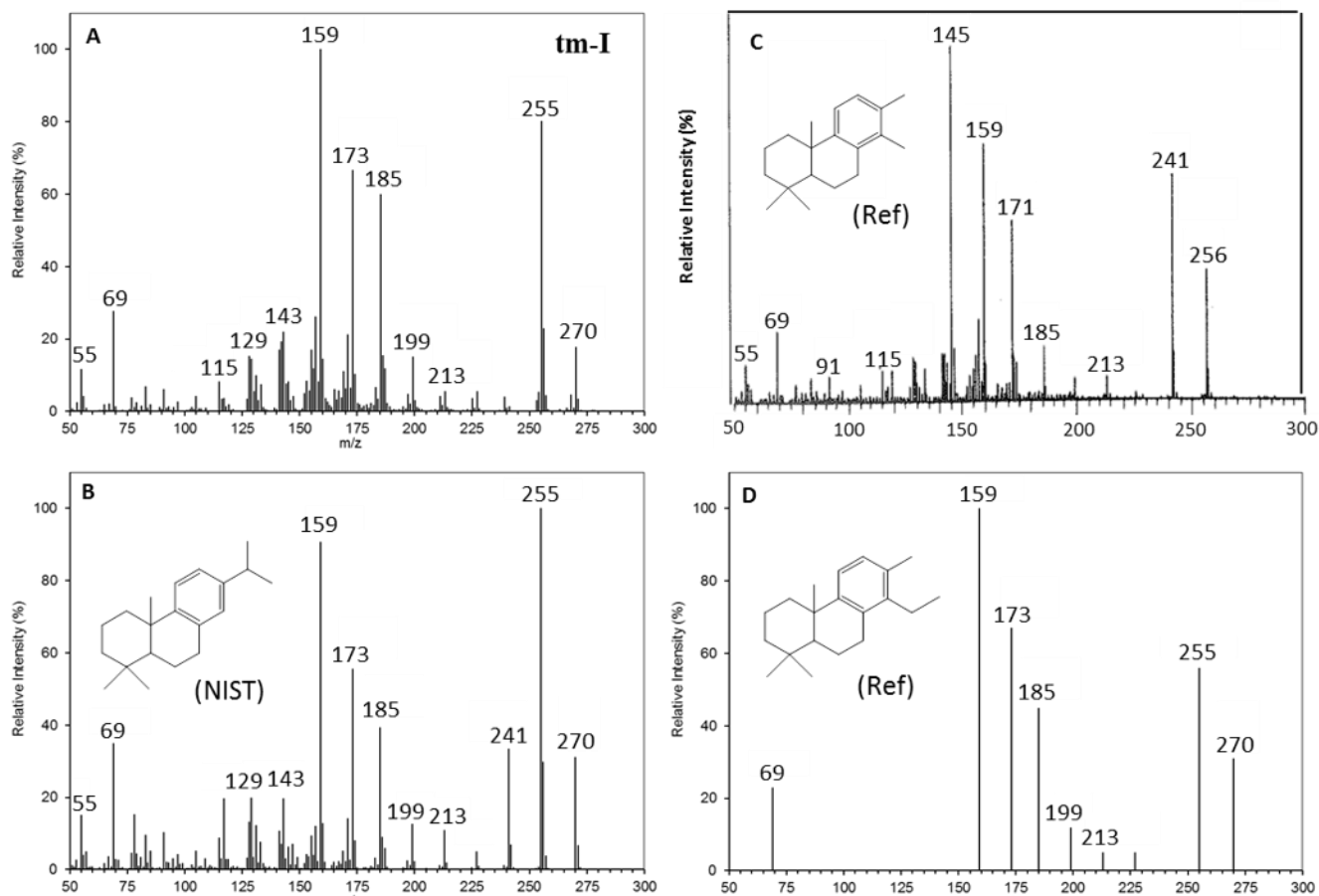
Figure S-13. (A-F) Mass spectra of C_{14-19} monoaromatic bicyclic hydrocarbons tentatively assigned based on mass spectral comparison and interpretation.

Figure S-14. (A) Mass spectrum of isomer tm-I, a C20 tricyclic monoaromatic hydrocarbon assigned as 13-methyl-14-ethylpodocarpa-8,11,13-triene after comparison with the mass spectra of (B) dehydroabietane plotted from the NIST MS Library, (C) 13,14-dimethylpodocarpa-8,11,13-triene reimaged from the spectrum reported by Azevedo et al. (1992) and (D) 13-methyl-14-ethylpodocarpa-8,11,13-triene replotted from the tabulated values reported by Azevedo et al. (1990).
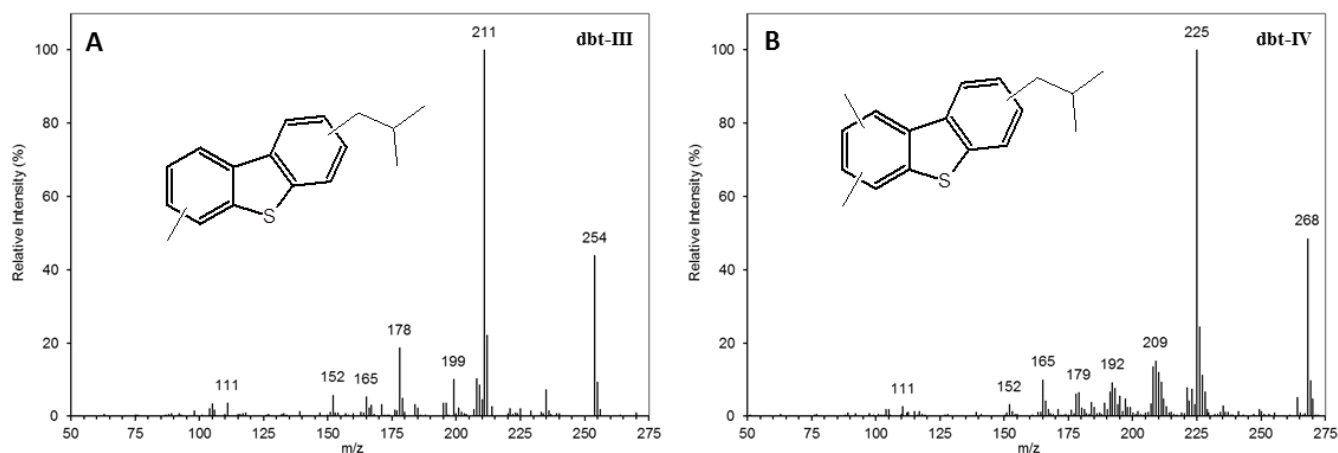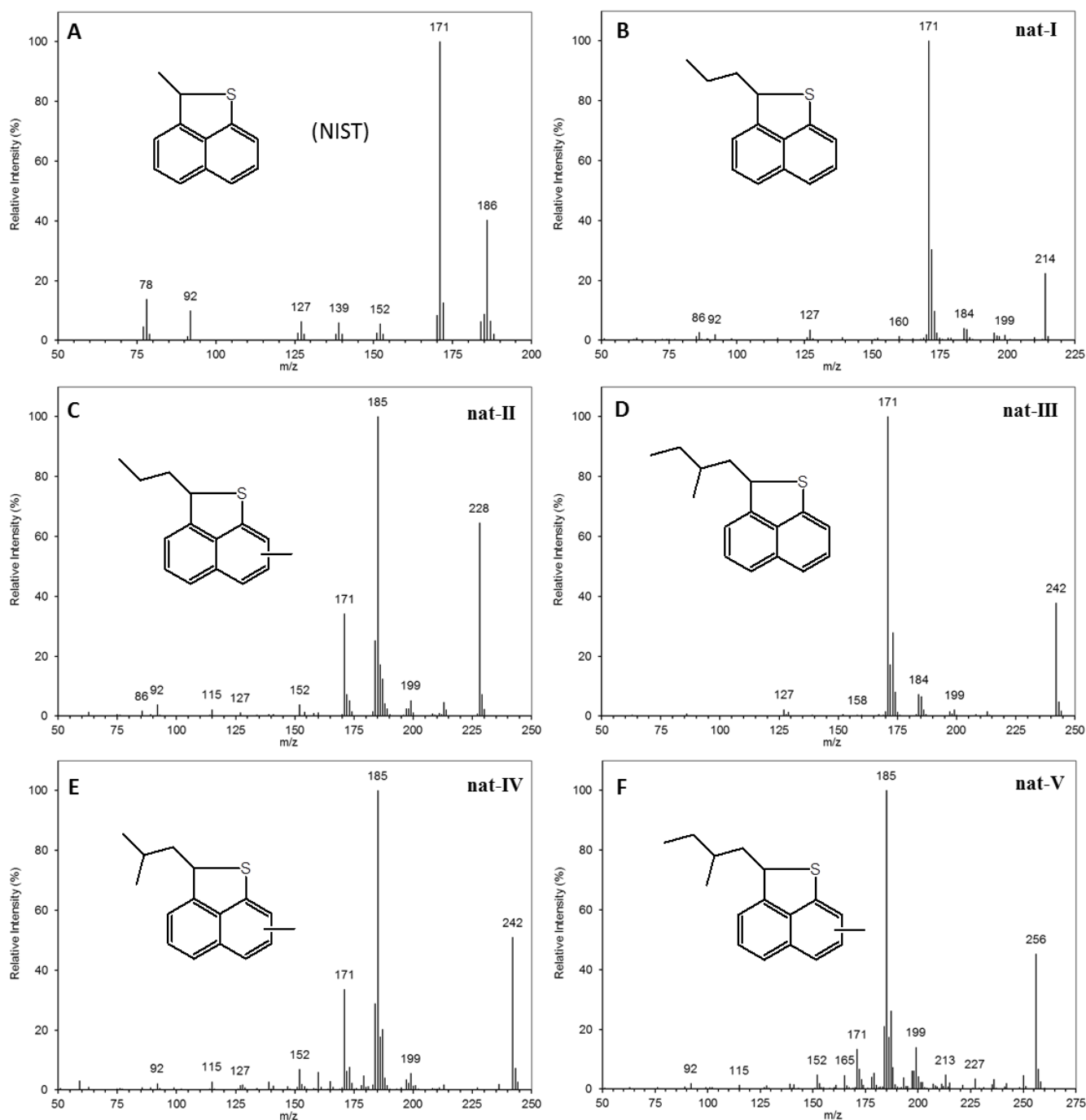
110   Figure S-15. (A and B) Assignment of dbt-III and -IV as methyl- and dimethyl- isobutyldibenzothiophenes.

111

112

113

Figure S-16. Mass spectra of a series of isomers (nat-I to -V) in the reduced 'aromatic, sulphur' fraction of sample #1,
tentatively assigned, after comparison with (A) the NIST spectrum of 2-methyl-2H-naphtho[1,8-bc]thiophene and their
retention positions relative to authentic 4-propyl- and isobutyl- dibenzothiophene, as (B-F) 2-alkyl (C3-5) substituted 2H-
naphtho[1,8-bc]thiophenes and methyl 2H-naphtho[1,8-bc]thiophenes.

118

119

120

121
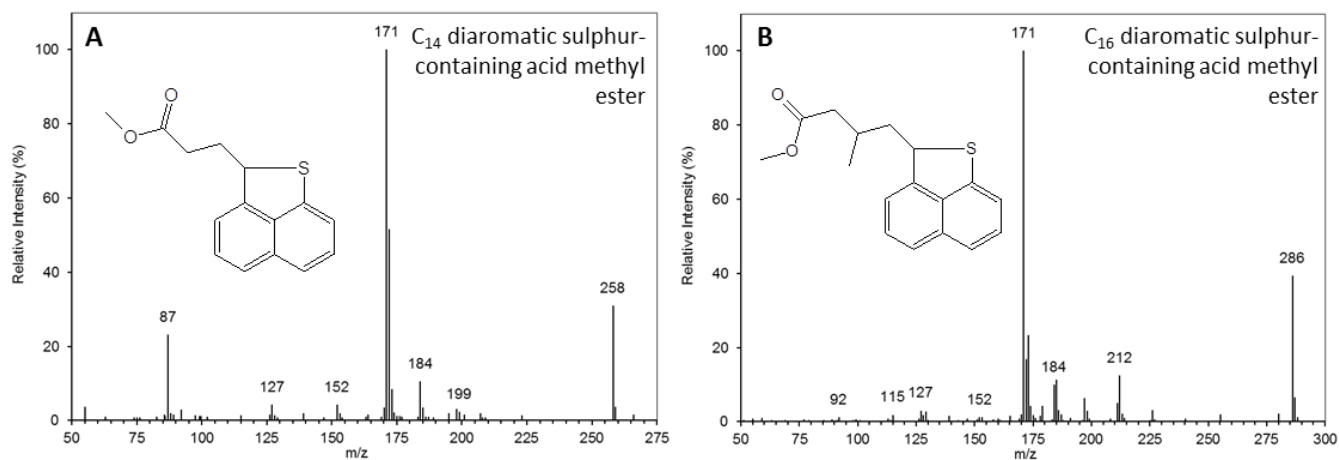
Figure S-17. (A and B) Example mass spectra of a C_{14} and C_{16} acid methyl ester in the 'aromatic, sulphur' fraction of sample #1, with molecular ions and retention positions consistent with condensed tricyclic, diaromatic sulphur-containing acid methyl esters, tentatively assigned as naphtho[1,8-bc]thiophene acids.

126

127

128

129

130

131

132

133

REFERENCES

Andersson, J.T., Schade, T., 2004. Higher alkylated dibenzothiophenes in some crude oils and hydrodesulfurized fuels. Prepr. Pap.-Am. Chem. Soc., Div. Fuel Chem 49, 338-340.

Andersson, J.T., Schräder, W., Traulsen, F., Werlich, S., 2001. Synthesis of Seven Trimethyldibenzothiophenes. Polycyclic Aromatic Compounds 18, 351-360.

Azevedo, D.A., Aquino Neto, F.R., Simoneit, B.R.T., Pinto, A.C., 1992. Novel series of tricyclic aromatic terpanes characterized in Tasmanian tasmanite. Organic Geochemistry 18, 9-16.

Azevedo, D.A., de Aquino Neto, F.R., Simoneit, B.R.T., 1990. Mass spectrometric characteristics of a novel series of ring-c monoaromatic tricyclic terpanes found in Tasmanian tasmanite. Organic Mass Spectrometry 25, 475-480.

Azevedo, D.A., Neto, F.R.A., Simoneit, B.R.T., 1994. Extended saturated and monoaromatic tricyclic terpenoid carboxylic acids found in Tasmanian tasmanite. Organic Geochemistry 22, 991-1004.

Bobinger, S., Traulsen, F., Andersson, J.T., 1999. Dibenzothiophene in Crude Oils: Products from the Photochemical Degradation. Polycyclic Aromatic Compounds 14, 253-263.

Bowman, D.T., Slater, G.F., Warren, L.A., McCarry, B.E., 2014. Identification of individual thiophene-, indane-, tetralin-, cyclohexane-, and adamantane-type carboxylic acids in composite tailings pore water from Alberta oil sands. Rapid Communications in Mass Spectrometry 28, 2075-2083.

Chicarelli, M.I., Neto, F.R.A., Albrecht, P., 1988. Occurrence of four stereoisomeric tricyclic terpane series in immature Brazilian shales. Geochimica et Cosmochimica Acta 52, 1955-1959.

Cyr, T.D., Strausz, O.P., 1983. The structures of tricyclic terpenoid carboxylic acids and their parent alkanes in the Alberta oil sands. Journal of the Chemical Society, Chemical Communications 0, 1028-1030.

Denisov, Y.V., Matveyeva, I.A., Sokolova, I.M., Petrov, A.A., 1977a. Mass-spectrometric study of hydrocarbons of bicyclo[3.2.1]octane series. Petroleum Chemistry U.S.S.R. 17, 85-93.

Denisov, Y.V., Sokolova, I.M., Petrov, A.A., 1977b. Mass Spectrometric study of hydrocarbons of the bicyclo[4.3.0]octane series (In Russian). Neftekhimiya 17, 491-497.

Denisov, Y.V., Vorob'eva, N.S., Petrov, A.A., 1977c. Mass Spectrometric study of hydrocarbons of the bicyclo[3.3.0]octane series (In Russian). Neftekhimiya 17, 656-662.

Frank, R.A., Kavanagh, R., Burnison, B.K., Headley, J.V., Peru, K.M., Der Kraak, G.V., Solomon, K.R., 2006. Diethylaminoethyl-cellulose clean-up of a large volume naphthenic acid extract. Chemosphere 64, 1346-1352.

Golovkina, L.S., Rusinova, G.V., Petrov, A.A., 1984. Mass Spectrometry of Saturated Hydrocarbons. Russian Chemical Reviews 53, 870-887.

Hall, P.B., Douglas, A.G., 1981. The distribution of cyclic alkanes in two lacustrine deposits, in: Bjoroy, M. (Ed.), International Meeting on Organic Geochemistry. John Wiley & Sons, Bergen, Norway, pp. 576-587.

Hawthone, D., Porter, Q., 1968. Naphtho[1,8-bc]thiophens. II. Mass spectrometry. Australian Journal of Chemistry 21, 171-183.

Ji-Zhou, D., Vorkink, W.P., Lee, M.L., 1993. Origin of long-chain alkylcyclohexanes and alkylbenzenes in a coal-bed wax. Geochimica et Cosmochimica Acta 57, 837-849.

Jones, D., West, C.E., Scarlett, A.G., Frank, R.A., Rowland, S.J., 2012. Isolation and estimation of the 'aromatic' naphthenic acid content of an oil sands process-affected water extract. Journal of Chromatography A 1247, 171-175.

182 Kuraš, M., Hála, S., 1970. The use of a gas chromatograph-mass spectrometer for the
183 analysis of complex hydrocarbon mixtures. Journal of Chromatography A 51, 45-57.
184 Larter, S.R., Solli, H., Douglas, A.G., 1981. Phytol-containing Melanoidins and their Bearing
185 on the Fate of Isoprenoid Structures in Sediments, in: Bjoroy, M. (Ed.), International Meeting
186 on Organic Geochemistry. John Wiley & Sons, Bergen, Norway, pp. 513-523.
187 Lengger, S.K., Scarlett, A.G., West, C.E., Rowland, S.J., 2013. Diamondoid diacids ('O4'
188 species) in oil sands process-affected water. Rapid Communications in Mass Spectrometry
189 27, 2648-2654.
190 Li, M., Wang, T.G., Shi, S., Liu, K., Ellis, G.S., 2014. Benzo[b]naphthothiophenes and alkyl
191 dibenzothiophenes: Molecular tracers for oil migration distances. Marine and Petroleum
192 Geology 57, 403-417.
193 Musayev, I.A., Bagrii, Y.I., Kurashova, E.K., Zaikin, V.G., Sanin, P.I., 1983. Diamantane
194 and 4-methyldiamantane in naphthalane and Russian crude oils. Petroleum Chemistry
195 U.S.S.R. 23, 182-185.
196 Petrov, A.A., 1987. Petroleum Hydrocarbons. Springer-Verlag, Berlin Heidelberg.
197 Philp, R.P., 1985. Fossil fuel biomarkers: applications and spectra. Elsevier.
198 Piccolo, L., Nassreddine, S., Toussaint, G., Geantet, C., 2010. Discussion on "A
199 comprehensive two-dimensional gas chromatography coupled with quadrupole mass
200 spectrometry approach for identification of C10 derivatives from decalin" by C. Flego, N.
201 Gigantiello, W.O. Parker, Jr., V. Calemma [J. Chromatogr. A 1216 (2009) 2891]. Journal of
202 Chromatography A 1217, 5872-5873.
203 Polyakova, A.A., Khramova, E.V., Bagrii, Y.I., Tsitsugina, N.N., Lukashenko, I.M., Frid,
204 T.Y., Sanin, P.I., 1973. Mass spectrometric study of alkyladamantanes. Petroleum Chemistry
205 U.S.S.R. 13, 1-10.
206 Porter, Q.N., 1985. Mass Spectrometry of Heterocyclic Compounds. Wiley.
207 Rowland, S.J., Scarlett, A.G., Jones, D., West, C.E., Frank, R.A., 2011a. Diamonds in the
208 rough: Identification of individual naphthenic acids in oil sands process water. Environmental
209 Science & Technology 45, 3154-3159.
210 Rowland, S.J., West, C.E., Scarlett, A.G., Jones, D., Frank, R.A., 2011b. Identification of
211 individual tetra- and pentacyclic naphthenic acids in oil sands process water by
212 comprehensive two-dimensional gas chromatography/mass spectrometry. Rapid
213 Communications in Mass Spectrometry 25, 1198-1204.
214 Sinninghe Damsté, J.S., Kock-van Dalen, A.C., de Leeuw, J.W., 1988. Identification of long-
215 chain isoprenoid alkylbenzenes in sediments and crude oils. Geochimica et Cosmochimica
216 Acta 52, 2671-2677.
217 Vorob'eva, N.S., Zemskova, Z.K., Pekh, T.I., Petrov, A.A., 1986. Diterpenoid tetracyclic
218 hydrocarbons of petroleum. Petroleum Chemistry U.S.S.R. 26, 69-76.
219 Wang, G., Shi, S., Wang, P., Wang, T.G., 2013. Analysis of diamondoids in crude oils using
220 comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry. Fuel
221 107, 706-714.
222 Wei, Z., Michael Moldowan, J., Dahl, J., Goldstein, T.P., Jarvie, D.M., 2006. The catalytic
223 effects of minerals on the formation of diamondoids from kerogen macromolecules. Organic
224 Geochemistry 37, 1421-1436.
225 West, C.E., Scarlett, A.G., Tonkin, A., O'Carroll-Fitzpatrick, D., Pureveen, J., Tegelaar, E.,
226 Gieleciak, R., Hager, D., Petersen, K., Tollefsen, K.-E., Rowland, S.J., 2014. Diaromatic
227 sulphur-containing 'naphthenic' acids in process waters. Water Research 51, 206-215.
228 Wilde, M.J., 2015. On the bicyclic acids of petroleum, Chemistry. Plymouth University.

229   Wilde, M.J., Rowland, S.J., 2015. Structural Identification of Petroleum Acids by Conversion
230   to Hydrocarbons and Multidimensional Gas Chromatography-Mass Spectrometry. Analytical
231   Chemistry 87, 8457-8465.
232   Wilde, M.J., West, C.E., Scarlett, A.G., Jones, D., Frank, R.A., Hewitt, L.M., Rowland, S.J.,
233   2015. Bicyclic naphthenic acids in oil sands process water: Identification by comprehensive
234   multidimensional gas chromatography–mass spectrometry. Journal of Chromatography A
235   1378, 74-87.
236   Wingert, W.S., 1992. G.c.-m.s. analysis of diamondoid hydrocarbons in Smackover
237   petroleums. Fuel 71, 37-43.

238

239