AARHUS UNIVERSITY

Coversheet

This is the accepted manuscript (post-print version) of the article.

Contentwise, the accepted manuscript version is identical to the final published version, but there may be differences in typography and layout.

How to cite this publication

Please cite the final published version:

René B. Madsen, Haofei Zhang, Patrick Biller, Allen H. Goldstein, and Marianne Glasius (2017). Characterizing Semivolatile Organic Compounds of Biocrude from Hydrothermal Liquefaction of Biomass. In *Energy & Fuels*, 2017, 31 (4), pp 4122–4134
DOI: <https://doi.org/10.1021/acs.energyfuels.7b00160>

Publication metadata

Title: Characterizing Semivolatile Organic Compounds of Biocrude from Hydrothermal Liquefaction of Biomass
Author(s): René B. Madsen, Haofei Zhang, Patrick Biller, Allen H. Goldstein, and Marianne Glasius
Journal: *Energy & Fuels*
DOI/Link: <https://doi.org/10.1021/acs.energyfuels.7b00160>
Document version: Accepted manuscript (post-print)

This document is the Accepted Manuscript version of a Published Work that appeared in final form in *Energy & Fuels*, 2017, 31 (4), pp 4122–4134. <https://doi.org/10.1021/acs.energyfuels.7b00160>, copyright © American Chemical Society after peer review and technical editing by the publisher.

General Rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

If the document is published under a Creative Commons license, this applies instead of the general rights.

1 **Characterizing Semi-volatile Organic Compounds of Bio-crude from** 2 **Hydrothermal Liquefaction of Biomass**

3 René B. Madsen¹, Haofei Zhang², Patrick Biller¹, Allen H. Goldstein², Marianne Glasius^{1*}

4 ¹ Interdisciplinary Nanoscience Center and Department of Chemistry, Aarhus University,
5 Langelandsgade 140, Aarhus, Denmark

6 ² Department of Environmental Science, Policy, and Management, University of California, Berkeley,
7 CA, USA

8 * Corresponding author: glasius@chem.au.dk

9 **Abstract**

10 Hydrothermal liquefaction of biomass produces a complex bio-crude, which can be further upgraded to
11 biofuel or chemicals, but there is a need for improved molecular understanding of product composition
12 and reaction pathways. This study extensively characterizes semi-volatile compounds in bio-crudes
13 from hydrothermal liquefaction (HTL) of microalgae (*N. gaditana*, *C. vulgaris*), macroalgae (*L.*
14 *hyperborea*), residue (Dried Distillers Grains with Solubles), and lignocellulosic (*M. x giganteus*). The
15 bio-crudes were analyzed using 2D gas chromatography coupled to time-of-flight mass spectrometry
16 with in-situ silylation. A total of 73 fatty acids were detected of which C₁₈ and C₂₀ compounds were
17 most diverse while palmitic acid was the single most abundant fatty acid. Multiple fatty acid amides
18 were detected in bio-crude from *N. gaditana* while being almost absent in samples from other lipid and
19 protein containing feedstocks. Several alkylated indoles and quinolines were observed in bio-crudes
20 from protein containing feedstocks. Monoglycerides, indanones, and alkylated benzenediols and
21 chromen-2-ones, which are typically not reported, were also detected. These results provide new
22 knowledge of a bio-crude fraction, which is difficult to characterize.

23 **Keywords:** Hydrothermal liquefaction; GCxGC-TOFMS; Online silylation; Biomass conversion;
24 Gas chromatography - mass spectrometry.
25

26 **1. Introduction**

27 Hydrothermal liquefaction (HTL) is a promising technique for conversion of wet biomasses in
28 particular into biofuels and value-added chemicals ¹⁻³. The process is carried out at conditions of 300-
29 400 °C and 200-300 bars where the unique properties of hot compressed water are utilized to obtain a
30 bio-crude with typically 10-20% oxygen.⁴ The process results in four complex product phases: a gas
31 phase, a bio-crude, an aqueous phase, and a solid residue.

32 Research into HTL has expanded in recent years to cover all aspects of the technology from
33 accessibility of feedstock to potential fates of products.⁵ The search for new feedstocks continues with
34 tested feedstocks now including residues, lignocellulosics, yeast, bacteria, and algae. ⁶⁻⁹ The desired
35 bio-crude product is either intended for use as a heavy fuel combustion or a drop-in fuel upon
36 upgrading.¹⁰ The product will thus require further processing based on comprehensive information of
37 its chemical composition in order to avoid additional expenses.

38 The bio-crude is a highly complex mixture of organic compounds of which most contain heteroatoms.^{7,}

39 ¹¹ A variety of advanced analytical techniques has been applied to investigate the composition of the
40 bio-crude, including Fourier-transform ion cyclotron resonance mass spectrometry (FTICR-MS)¹²,
41 nuclear magnetic resonance (NMR), size exclusion chromatograph (SEC)⁷, gas chromatography
42 coupled to mass spectrometry (GC-MS), and pyrolysis-GC-MS (py-GC-MS).¹³ Conventional GC-MS
43 is routinely used in molecular characterization of HTL bio-crude, often relying on automated library
44 searches of mass spectra. Many studies report relative peak areas of, either summed compound classes,

45 or single compounds based on the total ion chromatogram (TIC). The complex nature of bio-crudes
46 means that co-elution often occurs leading to overestimation of peak areas based on TICs. The samples
47 may also have varying fractions amenable to GC introducing variation to the data set based on relative
48 peak areas.

49 The compound classes reported from GC-MS analysis of bio-crudes include hydrocarbons, alcohols,
50 cyclic oxygenates, fatty acids, fatty acid derivatives, nitrogenated organic compounds, oxygenated
51 aromatics, small organic acids, and triterpenoids.¹⁴ Many of these compound classes are known to
52 suffer from matrix effects, discrimination and poor chromatographic performance in GC-MS analysis
53 which can increase the extent of co-elution and lead to poor repeatability, and thus complicating the
54 analysis.¹⁵ The use of silylating reagents can enhance chromatographic performance and the use of 2D
55 GC can be used to overcome co-elution and poor repeatability as recently applied to crude oils and
56 other complex petroleum mixtures.¹⁶

57 Emphasis is extending to include more compounds in contrast to earlier work which included only the
58 most abundant peaks.^{11, 17} The interpretations are almost exclusively based on relative peak areas,
59 which potentially lead to misinterpretations from incorrectly identified compounds, co-eluting peaks
60 and normalization with varying total peak area depending on the sample preparation method and GC-
61 MS settings. Recent analytical work is implementing derivatization to enhance chromatographic
62 performance with silylating reagents being the preferred method.^{18, 19} Performing silylation in solution
63 raises the question of derivatization efficiency and stability, which potentially introduces variation.²⁰

64 The analytical method of GC-MS is particularly valuable for characterization of volatile compounds
65 while identification becomes increasingly difficult with increasing molecular size as chromatographic
66 noise increases and fewer compounds are available in the database. Relatively limited work has been

done on characterization of semi-volatile compounds in bio-crude from HTL of biomass (semi-volatile defined as organic compounds with boiling points ranging from 240 °C to 400 °C). In this study we present an extensive characterization of semi-volatiles in HTL bio-crude from five highly varying biomasses, namely the microalgae *Chlorella vulgaris* (*C. vulgaris*) and *Nannochloropsis gaditana* (*N. gaditana*), the macroalgae *Laminaria hyperborea* (*L. hyperborea*), lignocellulosic *Miscanthus x giganteus* (*M. x giganteus*), and Dried Distillers Grains with Solubles (DDGS), a residue from bioethanol production. The resulting bio-crudes represent a wide range of semi-volatile compounds being produced at standard HTL process conditions. Bio-crudes were analyzed using highly efficient in-situ gas phase derivatization (silylation) coupled to 2D GC with time-of-flight MS detection (GCxGC-TOFMS). Development of the in-situ gas phase derivatization and its application with GCxGC-TOFMS has been described in detail elsewhere.^{21, 22} The efficiency of the gas phase derivatization means that compounds, which are typically not reported due to insufficient volatility or inefficient derivatization were detected here. The 2D set-up furthermore provides separation of most compounds. However, the method is limited by the fact that substantial silylating reagent is trapped requiring a long solvent delay meaning that the most volatile organic compounds are not detected. Furthermore, the bio-crude contains a heavy fraction, which is not amenable to GC. In order to make the results applicable for future studies, we have included information on Kovats retention indices of identified compounds.²³

2. Experimental

2.1 Chemicals, standards, and biomass

88 All standards were of GC grade and obtained from Sigma-Aldrich. Potassium carbonate was obtained
89 from Merck Chemicals. Chloroform and methanol were of HPLC grade and obtained from Sigma-
90 Aldrich. *M. x giganteus* was from Department of Agroecology, Aarhus University. DDGS was from
91 Lantmännen Agroetanol AB, Norrköping, Sweden. *C. vulgaris* was from a local health food store. *N.*
92 *gaditana* was from Lgem, Netherlands while *L. hyperborea* was obtained from The Scottish
93 Association for Marine Sciences (SAMS). The elemental distribution and biochemical composition of
94 the biomasses are presented in Supporting information (SI Table 1) along with the analytical methods.

95 **2.2 Hydrothermal liquefaction**

96 HTL reactions were carried out in 20 mL Swagelok batch reactors. Reactors were conditioned with
97 demineralized water and loaded with 10 mL biomass slurry containing 10 w/w% biomass, 2 w/w%
98 potassium carbonate (no catalyst added to *N. gaditana*), and 88 w/w% demineralized water. All HTL
99 reactions were carried out at 338 °C for 20 min in a fluidized sand bath, then quenched by rapid
100 cooling in a water bath, and subsequently vented. The aqueous phase was decanted into a centrifuge
101 tube. The tube was centrifuged at 6500 rpm for 5 min and the aqueous phase was decanted into a
102 preparative vial. The remaining content of the reactor was extracted with dichloromethane. This
103 organic phase was vacuum filtrated and subsequently the dichloromethane was evaporated under a
104 stream of nitrogen.

105 **2.3 GCxGC-TOFMS**

106 Bio-crudes were dissolved in 1:1 chloroform and methanol solution at approximately 15 mg ml⁻¹. The
107 sample solution (0.5 µl) was deposited on a quartz fiber filter together with 2 µl of internal standard
108 solution of deuterated even-numbered n-alkanes (C₁₂-C₃₆). The filter was thermally desorbed at 320 °C
109 (in Helium) using a thermal desorption system (TDS 3, Gerstel) with online gas phase derivatization by
110 *N*-methyl-*N*-(trimethylsilyl)-trifluoroacetamide (MSTFA).^{21, 22} Thus, compounds were either

111 derivatized as they evaporated from the filter or made volatile by derivatization on the filter. Only
112 compounds less volatile than MSTFA can be analyzed due to final purging and solvent delay. Desorbed
113 compounds were trapped in a focusing trap held at 30 °C. The compounds were then desorbed onto a
114 first dimension nonpolar column (60 m x 0.25 mm x 0.25 µm, Rxi-5Sil MS, Restek Corporation)
115 coupled to a second dimension polar column (1 m x 0.25 mm x 0.25 µm, Rtx-200MS, Restek
116 Corporation) by a dual-loop modulator (1.5 m x 0.25 mm, Rxi guard column) that is cooled with
117 refrigerated air and periodically heated by a pulse of hot air every 2 seconds. The oven temperature was
118 initially set at 40 °C, increasing at 3.5 °C min⁻¹ to 320 °C (hold time 5 min). The MS was operated in
119 electron ionization mode (70 ev) and data was collected at 100 Hz.

120 The data was acquired with Chemstation, and chromatograms were processed using GC Image.

121 Analytes were tentatively identified using NIST2011 based on fragmentation patterns.

122 Results were evaluated based on peak responses across the five different samples as responses are
123 expected to be within the linear range of the TOF-mass spectrometer.²⁴ This means that double peak
124 response is equal to twice as high concentration for the single compound; however, it does not extend
125 across individual compounds and the peak response for two different compounds is not necessarily
126 comparable which is especially true for derivatized and non-derivatized compounds. Furthermore, it is
127 our experience that compounds with increasing retention on the 2nd dimension (polar column) have
128 decreasing response factors leading to lower peak responses for more polar compounds such as
129 carbonyls and diketopiperazines (DKPs). Hence, peak abundance does not necessarily represent
130 relative mass concentrations and comparisons across compound classes should be especially cautious.

131

132 **3. Results and discussion**

133 3.1 Compound group analysis of bio-crude of algae

134 Microalgae have been extensively investigated as feedstocks for HTL because of the wet nature of the
135 biomass and the often high lipid content. Furthermore, microalgae do not need size reduction to obtain
136 stable slurries, making the feedstock easily amenable for continuous processing.⁵ The vast majority of
137 characterization studies of bio-crude from microalgae have relied on conventional one-dimensional
138 GC-MS identifying the most abundant peaks, typically palmitic acid, oleic acid or hexadecanamide
139 along with smaller peaks of indoles, phytane, pyrazines, pyrroles, and triterpenoids depending on the
140 feedstock and process conditions.^{7, 25, 26}

141 Figure 1 shows the GCxGC-TOFMS chromatogram of bio-crude from *N. gaditana*, a high protein,
142 moderate lipid containing microalgae in which the unsaturated fatty acid content is especially high.
143 The chromatogram displayed several regions of interest, which are seldomly reported for *N. gaditana*
144 bio-crude. The high lipid content is converted by HTL to saturated and unsaturated fatty acids of C₁₄,
145 C₁₆, C₁₈ and C₂₀, seen as four specific regions in the chromatograms. While the regions of C₁₆ and C₁₈
146 fatty acids are consistently reported in the literature, they are often identified as only palmitic acid and
147 oleic acid or just as fatty acids. This work identified the presence of myristic acid (C₁₄), two C₁₅ fatty
148 acids, four C₁₆ fatty acids, three C₁₇ fatty acids, nine C₁₈ fatty acids, 25 C₂₀ fatty acids, and lignoceric
149 acid (C₂₂). The presence of C₂₀ fatty acids is typically not reported in the literature because without
150 derivatization the C₂₀ fatty acids are non-amenable for GC analysis. Instead they condense in the inlet
151 with subsequent decomposition and potential artifacts in later runs.

152 Fatty acid amides, formed from condensation of free fatty acids with ammonium or amines²⁷, constitute
153 another region of interest. Most studies report only the presence of hexadecanamide^{6, 28} while more
154 recent work on model compounds has shown the presence of over 40 different fatty acid amides.²⁷
155 Barreiro, et al.²⁹ reported the presence of six different fatty acid amides in bio-crude from *N. gaditana*

156 while finding only four different fatty acids. This study identified the presence of 26 different fatty acid
157 amides from bio-crude of *N. gaditana*. The fatty acid amides ranged from tetradecanamide to
158 octadecanamide and were primary amides, methylated, dimethylated or pyrrolidine amides.
159 Diketopiperazines (DKP) have been reported from HTL of amino acids³⁰ and are formed from
160 dimerization of amino acids; however, they seem to form only at low temperature HTL and to be
161 further degraded at temperatures above 300 °C.³¹ DKPs were only found in minor amounts in bio-crude
162 of *N. gaditana* and it is likely that they are initially formed and then degraded.

163 Indoles constitute another class of nitrogen-containing compounds which is often found in bio-crude of
164 protein-rich feedstocks^{14, 26}. For bio-crude of *N. gaditana*, indoles were present only to a minor extent
165 in a region encompassing also quinolines. Indole is often reported as a major component from protein-
166 rich feedstock and is proposed to form from degradation of tryptophan or from reaction between phenol
167 and glycine. In this work multiple alkylated indoles were detected which indicates that they are likely
168 to form from secondary reactions between phenol and amino acids as they do not occur naturally in
169 biomass and alkylation from indole is likely to occur at the 1 and 3 position.

170 Several hydrocarbons were detected which included both saturated branched aliphatics and unsaturated
171 aliphatics. Unsaturated aliphatics were mainly C₁₆, C₁₈, or C₂₀ with either double bonds or triple bonds.
172 Microalgae are rich in phytol and previous studies have shown that unsaturated hydrocarbons are
173 produced during HTL while fatty acids have been shown to partially decarboxylate especially in
174 alkaline HTL.^{32, 33} The final quarter of the chromatogram was characterized by 18 different triterpenes
175 and triterpenoids in the form of choletane and ergostane derivatives. Typically, only a few of these
176 compounds are reported in the literature despite their apparent abundance in the chromatogram here,
177 where most of these are not even derivatized. The presence of six long-chain unsaturated fatty alcohols
178 was unique for bio-crude of *N. gaditana*. These were highly abundant C₃₀, C₃₁, and C₃₂ compounds.

179 *C. vulgaris* is a protein rich and low lipid containing microalgae, extensively investigated in HTL.^{17, 34}
180 The bio-crude of *C. vulgaris* gave only a small peak for arachidic acid (C₂₀) compared to the massive
181 number of C₂₀ fatty acids of *N. gaditana* (Fig. 2). Instead a substantial number of unsaturated C₁₆ and
182 C₁₈ fatty acids were observed which included 16 different C₁₆ fatty acids and 21 different C₁₈ fatty
183 acids (enlargement Fig. 2) along with palmitic acid and stearic acid. Several of these were
184 tetraunsaturated fatty acids, which have previously not been reported for bio-crude of HTL even though
185 they are naturally present in microalgae.³⁵ Most publications on *C. vulgaris* identify these regions as
186 being the most abundant ones. However, the entire regions are in most cases assigned to one or two
187 fatty acids, such as palmitic acid or oleic acid.

188 Fatty acid amides were almost completely absent (only four minor peaks) despite the high protein and
189 moderate lipid content of *C. vulgaris* and the high abundance of C₁₆ and C₁₈ fatty acids of the bio-
190 crude. Madsen, et al.³⁶ showed that high concentrations of pyrazines are formed in the aqueous phase
191 from HTL of *C. vulgaris* when the carbohydrate to protein ratio of the feedstock is close to 1, as the
192 degradation products of carbohydrate and protein react instead of forming fatty acid amides. This also
193 implies that Strecker degradation and Amadori rearrangement occurs faster than fatty acid amide
194 formation.

195 Partially degraded triglycerides in the form of monoglycerides of hexadecanoic acid and octadecanoic
196 acid constitute a group of compounds which are rarely reported in HTL bio-crude.³⁷ The main reason
197 for lack of detection of these compounds is that they are not amenable to conventional GC-MS due to
198 low vapor pressures, and the lower efficiency of silylation in solvents means that they would likely not
199 be detected without optimization of silylation conditions. It is interesting to note that monoglycerides
200 were not detected in bio-crude of *N. gaditana* even though it has similar abundances of C₁₆ and C₁₈
201 fatty acids.

202 Bio-crude of *C. vulgaris* contained a region of indole/quinoline similar to *N. gaditana* while the region
203 of hydrocarbons was more abundant in aliphatic hydrocarbons with triple bonds.

204 Few characterizations based on GC-MS of bio-crude from HTL of macroalgae have been reported.
205 Yang, et al. ³⁸ showed that bio-crude from *Enteromorpha prolifera* consisted of palmitic acid, cyclic
206 oxygenates, alkenes, and pyrazines. Anastasakis and Ross (2011) found that bio-crude of *L. saccharina*
207 consisted mainly of cyclic oxygenates, phenolics, indoles, and fatty acids.

208 To our knowledge extensive molecular characterization of HTL bio-crude of macroalgae has not been
209 published and this is the first study to report DKPs, indanones, benzenediols, monoglycerides and
210 triterpenoids. *L. hyperborea* has a high carbohydrate content which typically leads to formation of a
211 number of cyclic oxygenates not detected with the present analytical method.^{39, 40}

212 The most notable regions were C₁₄, C₁₆, and C₁₈ saturated and unsaturated fatty acids (Fig. 3).
213 However, the bio-crude also contained saturated pelargonic acid (C₉), capric acid (C₁₀), lauric acid
214 (C₁₂), pentadecyclic acid (C₁₅), margaric acid (C₁₇), and arachidic acid (C₂₀). A total of 14 different C₁₈
215 fatty acids were detected including four tetraunsaturated. High abundances of monoglycerides of C₁₆
216 and C₁₈ were observed along with a minor peak of C₁₄ monoglyceride. The monoglyceride abundance
217 was even higher than for bio-crude of *C. vulgaris*, despite the significantly lower lipid content of *L.*
218 *hyperborea*. Thus, the triglycerides of macroalgae seem to be significantly more resistant to hydrolysis
219 than those of microalgae.

220 Several minor peaks of DKPs were observed in bio-crude of *L. hyperborea* (no silylated DKPs were
221 observed), which were different from the DKPs of *N. gaditana* bio-crude. It has been proposed that
222 DKPs are formed from degradation of protein and recombination of amino acids ³³. However, the high
223 protein feedstock of *C. vulgaris* did not yield DKPs while the high protein feedstock of *N. gaditana* and

low protein feedstock of *L. hyperborea* did, which indicates that competitive reactions occur affecting the formation or degradation of DKPs.

A number of carbonyl compounds were detected for bio-crude of *L. hyperborea* including acetophenones and indanones which are formed by degradation of lignin and carbohydrates.^{41, 42} Bio-crude of *L. hyperborea* contained 23 different benzenediols in the form of alkylated catechol and hydroquinone, while microalgae had almost no benzenediols.

The sample also contained a number of hydrocarbons which compared to other samples consisted mostly of saturated aliphatics. Also the triterpenoid composition was different as it consisted of pregnenane derivatives for bio-crude of *L. hyperborea*.

3.2 Compound group analysis of bio-crude of DDGS

The DDGS used in this work is a residue from bio-ethanol production. Residues are often found in large quantities as wet streams making them highly attractive for HTL processing.

The most notable region in the chromatogram of this sample is the dense region of C₁₈ fatty acids (Fig. 4). A total of 18 peaks were detected here, including stearic acid (C₁₈) and several mono, di, tri, and tetraunsaturated C₁₈ fatty acids. The sample contained only minor amounts of palmitoleic acid (C₁₆) compared to bio-crude of *C. vulgaris*. The analysis also showed pelargonic acid (C₉), capric acid (C₁₀), lauric acid (C₁₂), myristic acid (C₁₄), pentadecylic acid (C₁₅), margaric acid (C₁₇), arachidic acid (C₂₀), behenic acid (C₂₂), lignoceric acid (C₂₄), and cerotic acid (C₂₆), similar to bio-crude of *C. vulgaris*. The abundance of monoglycerides was also analogous to *C. vulgaris*.

Furthermore, the region of fatty acid amides was almost completely absent, as for *C. vulgaris*, even though the feedstock consists of 33% protein. DDGS contains approximately the same ratio of carbohydrate to protein as *C. vulgaris* and similar reaction pathways for pyrazine formation were observed in the aqueous phase.³⁶ This supports our previous notion that degradation products of

247 carbohydrates react with degradation products of proteins, leading to an absence of fatty acid amides.

248 Additional similarity of bio-crudes from DDGS and *C. vulgaris* was observed for the indole/quinoline
249 region which was almost identical in terms of compounds.

250 The samples of DDGS and *L. hyperborea* contained comparable aldehyde/ketone and benzenediol
251 regions. However, for DDGS the aldehyde/ketone region only showed indenones, acetophenones, and
252 chromenones and thus no indanones. Instead an additional region with aromatic acids containing eight
253 cinnamic acids was present in the bio-crude from DDGS.

254 Furthermore, this sample only showed eight triterpenoids which were ergostane, cholestane, and
255 stigmastane derivatives, similar to *C. vulgaris*.

256 **3.3 Compound group analysis of bio-crude of *M. x giganteus***

257 Lignocellulosics are less often used for HTL for a number of reasons which include: necessary size
258 reduction, need for use of process water (can be overcome with water recycling), difficulties with
259 pumping, low growth rates, and high lignin contents leading to higher amounts of solid residue.^{5, 43}

260 However, their widespread abundance means that they are an accessible biomass with low demands for
261 nutrients. Energy crops, such as *M. x giganteus*, are of specific interest as they can grow on non-arable
262 land due to low nutrient requirements. Despite the abundance of grass feedstocks only few have been
263 subjected to HTL.^{44, 45}

264 Bio-crude of *M. x giganteus* was the most complex sample analyzed here (Fig. 5). Despite the low lipid
265 content of *M. x giganteus*, the most abundant peaks detected were for palmitic acid and oleic acid
266 although they were far less abundant than in other samples. Only three other C₁₈ fatty acids were
267 detected. The third and fourth most abundant peaks were monoglycerides of C₁₆ and C₁₈ fatty acids.
268 This is despite the fact that the feedstock was grinded to < 90 µm making them more accessible to
269 degradation at typical HTL process parameters. This indicates that the glycerides of lignocellulosics are

270 far more resistant to hydrolysis than in other feedstocks and process conditions can be optimized.
271 Phenolics and benzenediols are typically reported as the main components of bio-crude from
272 lignocellulosics. The responses of single compounds of this class were considerably lower than fatty
273 acids; however, the number of different compounds was the highest for all compound classes with 62
274 different phenolics and 30 different benzenediols.

275 The region of aldehyde/ketone (indanones, acetophenones, and chromenones) is typically not reported
276 for lignocellulosics. While it is the least abundant group of peaks in this study it represents one of the
277 most diverse ones, with 38 different compounds of cyclic oxygenates, indanones, indenones,
278 acetophenones, propriophenones, and chromenones.

279 **3.4 Comparative characterization of bio-crude**

280 In the following sections the most abundant compounds, based on peak response, of each compound
281 class are evaluated across the five different samples. A threshold peak response was used in order to
282 evaluate only the most abundant compounds within each compound class of fatty acids (>25),
283 monoglycerides (>25), alcohols (>3), phenolics (>15), benzenediols (>15), nitrogen containing
284 compounds (>10), hydrocarbons (>5), and carbonyls (>5).

285 Figure 6 shows the sum of normalized peak responses for the compound classes highlighted in section
286 3.1-3. Fatty acids are identified as the most abundant semi-volatile compound class in all samples
287 which is coherent with previous studies on HTL of microalgae, macroalgae, and DDGS while
288 phenolics and benzenediols are often considered the most abundant compound classes in bio-crudes of
289 lignocellulosics. However, it is well known that analysis of fatty acids suffer from matrix effects, active
290 site interactions, and poor chromatographic performance without derivatization. Alcohols are the
291 second most abundant semi-volatile compound class of *N. gaditana* and DDGS, while nitrogen-
292 containing compounds are equally abundant in *N. gaditana*; however, most of the nitrogen-containing

293 compounds are non-derivatized and it is likely that their concentrations are actually higher than those
294 of alcohols. It is interesting to note that the protein content of *N. gaditana* (46%), *C. vulgaris* (49%),
295 and DDGS (33%) are similar while the sum of peak responses for semi-volatile nitrogen containing
296 compounds are 3-4 times higher for *N. gaditana*, indicating that a change in reaction pathways occurs
297 based on the relative content of carbohydrates, proteins, and lipids of the feedstock.

298 Hydrocarbons, we found are mainly from algae feedstocks, but they are not derivatized and are
299 therefore expected to be found in higher concentrations than indicated by the peak responses presented
300 in this work. Semi-volatile carbonyls are a mixture of mainly non-derivatized indanones/indenones and
301 smaller abundances of oxygenated ketonized aromatics, which are derivatized. These are almost solely
302 detected for *L. hyperborea* and *M. x giganteus*. Phenolics and benzenediols are predominantly detected
303 in bio-crude from *M. x giganteus* while smaller amounts are found from DDGS and *L. hyperborea*.

304 In the following, individual compounds from fatty acid, monoglycerides, alcohols, phenolics,
305 benzenediols, nitrogen-containing compounds, and carbonyls are investigated. Hydrocarbons are listed
306 in Supporting Information (SI Table 2).

307 **3.4.1 Fatty acids and monoglycerides**

308 Fatty acids are the main components of most HTL bio-crudes since high lipid feedstocks result in high
309 bio-crude yield and further upgrading is accomplished in high yield as well. Experiments with model
310 compounds (e.g. sunflower oil) have shown almost 100% conversion to bio-crude.⁴⁶

311 The largest peak for all samples was observed for palmitic acid (hexadecanoic acid), and the largest
312 value was observed for bio-crude of *N. gaditana* (Table 1). Fatty acids with molecular weights less than
313 palmitic acid were generally at least three times more abundant in the sample of *N. gaditana*, except
314 myristic acid and lignoceric acid which were also highly abundant for *L. hyperborea*. The sample of *M.*
315 *x giganteus* contained 4 to 7 times less palmitic acid and equal amounts of stearic acid compared to the

316 other samples. However, it contained almost none of the otherwise highly abundant fatty acids of cis-9-
317 hexadecenoic acid, palmitoleic acid ((Z)-9-hexadecenoic acid), oleic acid, and trans-9-octadecenoic
318 acid. The bio-crude of *C. vulgaris* and *L. hyperborea* had similar abundances of cis-9-hexadecenoic
319 acid, palmitic acid, linoleic acid, oleic acid, and stearic acid; however, the diversity of the unsaturated
320 C₁₈ fatty acids was far larger for *C. vulgaris*. Abundances of these unsaturated C₁₈ fatty acids were
321 even higher for DDGS. A number of eicosatetraenoic acids and eicosapentaenoic acids were only
322 present for *N. gaditana* and not included in this analysis.

323 Monoglycerides are rarely reported from HTL bio-crudes. Bio-crudes from the high lipid feedstocks of
324 *N. gaditana* and DDGS had the lowest abundances of all monoglycerides while it was three times
325 higher for *C. vulgaris* bio-crude. Bio-crudes of *L. hyperborea* and *M. x giganteus* had up to 26 times
326 higher concentrations of α-glyceryl stearate in *L. hyperborea* bio-crude compared to *N. gaditana* bio-
327 crude which may be due to less accessible cell structures. Furthermore, the concentrations of oleic acid
328 and stearic acid are comparable for *L. hyperborea* bio-crude; however, neither α- or β-glyceryl oleate
329 (RI 2788 and 2744) could be detected. A very small peak of β-glyceryl oleate was detected for *C.*
330 *vulgaris*. This indicates that the glycerides containing saturated fatty acids may be less susceptible to
331 hydrolysis during the HTL process than those containing unsaturated fatty acids.

332 3.4.2 Alcohols

333 Alcohols are rarely reported from HTL studies because single alcohols are found in small amounts with
334 vapour pressures coinciding with a significant fraction of the bio-crude leading to co-elution. The
335 alcohols detected in this study were primarily derived from pigments or wax.

336 The most abundant alcohols were straight long chain and branched long chain alcohols (Table 2). The
337 bio-crude of *C. vulgaris* was the most diverse with regard to simple long chain primary alcohols albeit
338 none of these were highly abundant. The bio-crudes of *M. x giganteus* and *L. hyperborean* had the

339 highest abundances of the phytol-like compounds of 3,7,11,15-tetramethylhexadecan-1-ol and
340 3,7,11,15-tetramethyloctadecan-1-ol. Bio-crude of *L. hyperborea* also showed two abundant glycols.
341 The bio-crude of DDGS contained several different unknown saccharides, which in one case was
342 highly abundant. However, saccharides were not detected for bio-crudes of high carbohydrate
343 feedstock such as *L. hyperborea* and *M. x giganteus*. In general it would be expected that saccharides
344 are displaced to the aqueous phase.

345 The bio-crude of *N. gaditana* was the only one to have a specific alcohol region appointed in section
346 3.1. Dotricentenol was the most abundant of the six fatty alcohols determined and was the compound
347 with highest response for all alcohols across the samples.

348 Glycerol was detected in equally low abundance for all samples. However, it was not detected in *N.*
349 *gaditana*.

350 **3.4.3 Phenolics and benzenediols**

351 Phenolics and benzenediols are highly abundant compounds from HTL of especially carbohydrate and
352 lignin rich feedstocks and represent compounds of potentially high value. They show medium to high
353 activity towards hydrodeoxygenation and are a source of oxygen in the final bio-fuels.^{47, 48}

354 The bio-crudes of *N. gaditana* and *C. vulgaris* contained almost no phenolics or benzenediols although
355 two long chain alkyl benzenediols (3-((1E,3E)-pentdeca-1,3-dienyl)-1,2-benzenediol and ((3-((1E,3E)-
356 hexadeca-1,3-dienyl)-1,2-benzenediol) were detected for *N. gaditana* (Table 3). These compounds
357 were not part of NIST library and are identified based on retention times and fragmentation patterns. It
358 has previously been proposed that phenolics and benzenediols are partly formed from degradation of
359 carbohydrates which supports the absence of these compounds in the sample made from *N. gaditana*.
360 On the other hand, *C. vulgaris* contains 26.5 % carbohydrates (dry weight); however it is proposed that
361 the majority of intermediate products from carbohydrate degradation react with degradation products

362 from protein to produce nitrogen-containing aromatics. DDGS (35.0% carbohydrate) has
363 approximately the same biochemical composition as *C. vulgaris*; however the bio-crude from DDGS
364 has a much higher abundance of especially benzenediols (catechols and hydroquinones). Instead the
365 carbohydrates are degraded to indanones. The formation of phenolics is clearly favored by high lignin
366 contents with *M. x giganteus* showing high abundances of several simple phenolics, such as 3-
367 methoxyphenol, 2-ethylphenol, 3-ethylphenol, and ethylguaiacol. A series of phenolics containing the
368 structure of 3-phenol-prop-2-en-1-ol were detected for *M. x giganteus* bio-crude and are expected to be
369 degradation products from lignin. These compounds were not part of the NIST database.

370 Benzenediols can be found in three different forms, catechol (benzene-1,2-diol), hydroquinone
371 (benzene-1,4-diol), and resorcinol (benzene-1,3-diol). They are easily distinguished by their
372 approximate ratios of M^+ to 73 m/z which are 1:4 for catechol, 4:1 for hydroquinone, and 1:1 for
373 resorcinol. In this work, only catechols and hydroquinones were detected. The most abundant
374 benzenediol detected was 4-ethylcatechol which was only detected in *M. x giganteus*.

375 Trimethylsilylated 4-ethylcatechol was not present in the NIST library and has been identified from
376 comparison with literature.⁴⁹ None of the series of alkylated catechols and hydroquinones were present
377 in the NIST library and their identification is based on comparison of retention times and fragmentation
378 patterns with 4-ethylcatechol (See SI Fig. 1).

379 3.4.4 Nitrogen-containing compounds

380 The nitrogen-containing compounds are an important compound class as combustion of nitrogen-
381 containing bio-crude may produce nitrogen oxides. Fossil crude-oil contains only small amounts of
382 nitrogen and commercial catalysts are therefore unsuitable for denitrogenation making knowledge of
383 nitrogen-containing compounds important.⁴⁷

384 The nitrogen-containing compounds were mainly ring structures of either aromatics (indoles,
385 quinolones, imidazoles, or indolizines) or combined with amide bonds (pyrrolidinones or fatty acid
386 amide) and free amines were not present. By far the highest diversity, and in most cases also the
387 highest abundance, of nitrogen-containing compounds was found for bio-crudes of the high protein
388 feedstock of *N. gaditana* (Table 4). The ring structures were mostly made up of indoles which were
389 generally observed in all samples apart from *M. x giganteus*. The level of the most abundant indoles
390 was approximately equal in the bio-crudes from *N. gaditana*, *C. vulgaris*, *L. hyperborea*, and DDGS
391 whereas the remaining compounds were either not detected or in significantly lower abundance.
392 While several fatty acid amides have previously been detected from HTL of model compounds²⁷, they
393 are rarely reported from biomasses with the exception of hexadecanamide or *N*-
394 methylhexadecanamide. These compounds were also the most abundant fatty acid amides detected and
395 generally fatty acid amides were almost exclusively detected in *N. gaditana* bio-crude. It is noteworthy
396 that the feedstocks of *C. vulgaris* and DDGS contain approximately the same protein concentration as
397 *N. gaditana* while the lipid concentration is half and the carbohydrate concentration is double. This
398 implies that the proteins are degraded to amino acids at a sufficient rate and able to react with reducing
399 saccharides before the amino acids are deaminated or decarboxylated to form pyrazines. However, a
400 considerable amount of amino acids is also deaminated to form small organic acids dissolved in the
401 aqueous phase. When proteins are in excess the released ammonia, methylamine, or dimethylamine is
402 able to react with fatty acids to form fatty acid amides. It is interesting to note that only (Z)-octadec-9-
403 enamide was detected even though (E)-octadec-9-enoic acid was found in the same abundance as (Z)-
404 octadec-9-enoic acid. Furthermore, a total of only 26 fatty acid amides were detected even though 45
405 different ones should be possible.

406 3.4.5 Aldehydes and ketones

407 Carbonyls are typically reported as highly abundant in most HTL bio-crudes; however, they are nearly
408 always presented as cyclopentenones and mainly occur from lignocellulosic feedstocks. To our
409 knowledge, no other study has reported the presence of semi-volatile carbonyls in bio-crudes from
410 HTL of biomass. The most abundant carbonyls were largely indenones, acetophenones, and a series of
411 alkylated chromenones (Table 5). Indenones and chromenones are typically not reported from studies
412 of HTL which is mainly due to the fact that most studied feedstocks have been microalgae, macroalgae
413 and residues where lignin is either absent or found in only small amounts. The bio-crudes from
414 microalgae contained only small amounts of a single carbonyl in the form of 6,10,14-
415 trimethylpentadecan-2-one while DDGS bio-crude contained amounts of different indenones,
416 acetophenones, and alkylated chromenones. The diversity and abundance of these compounds were
417 significantly higher for *L. hyperborea* indicating that some of these compounds are also formed from
418 degradation of carbohydrates as the lignin content of *L. hyperborea* is lower than in DDGS. By far the
419 highest number of and the greatest abundance of most carbonyls were observed for bio-crudes of *M. x*
420 *giganteus*. Indenones are not silylated and therefore the response is not enhanced from derivatization,
421 making their responses even more significant and 2,3-dihydro-1H-inden-1-one is especially abundant.
422 A number of alkylated chromen-2-ones were detected which were not present in the NIST library.
423 Chromenones are widely distributed in plant material where they have a protective function following
424 stress events and antimicrobial activity⁵⁰ and they are therefore expected to be ambiguous to *M. x*
425 *giganteus*; however, their abundance should also be highly dependent on the growth conditions of the
426 feedstock. Chromenone were observed with the same abundance both in bio-crudes of *L. hyperborea*
427 and *M. x giganteus* and while they are ambiguous to plant materials they have not been extensively
428 reported for macroalgae. Furthermore, since *L. hyperborea* contains only minor amounts of lignin, it
429 indicates that chromenones in HTL bio-crude are formed from carbohydrate degradation. The alkylated

430 chromen-2-ones were identified from umbelliferone and a series of mass spectra with increments of 14
431 m/z and characteristic fragmentation patterns (SI Fig. 2).

432 **3.5 Reaction pathways**

433 A generalized scheme of reaction pathways for biomasses containing carbohydrates, lipids and proteins
434 based on the most abundant compounds is presented in Figure 7. The limited fraction characterized in
435 this work means that several secondary products are likely to be formed through additional pathways
436 observed by other studies, which are referenced throughout the section.

437 Lipids include fatty acids, monoglycerides, diglycerides, triglycerides, phospholipids, waxes, steroids,
438 and hydrophobic vitamins. Hydrolysis of glycerides leads to the formation of glycerol, free fatty acids,
439 and monoglycerides observed in this work and previously reported in other studies.³⁷ Unsaturated fatty
440 alcohols (C₃₀₋₃₂) were also observed, likely formed from hydrolysis of waxes.⁵¹ Triterpenoid
441 derivatives detected in this work and other studies are naturally occurring compounds in algae and
442 plants released from the cells during HTL.^{26, 51, 52} In the reaction pathways outlined, chlorophyll is
443 considered a lipid as it is extracted during the determination of the total lipid of the feedstock (other
444 pigments were not considered). Hydrolysis of chlorophyll leads to the release of phytol which has been
445 reported to be reduced to several pristene, phytene, and neophytadiene isomers, which were also
446 observed during this work.³²

447 Proteins are initially hydrolyzed to amino acids by cleavage of peptide bonds. Studies on HTL of pure
448 protein and amino acids show rapid degradation through deamination and decarboxylation producing
449 ammonia, carboxylic acids, carbon dioxide, and amines, which are predominantly water soluble
450 compounds and not observed in this work.⁵³ Depending on the side chain of the amino acids a number
451 of additional degradation and recombination products may be produced of which only a fraction are
452 accounted for in these reaction pathways.⁵⁴

453 In the presence of fatty acids, ammonia and amines can react through condensation to form fatty acid
454 amides. Previous studies focused specifically on fatty acid amides identified 39 amides from HTL of
455 municipal solid waste many of which the amine was longer than methylamine.²⁷ Most other studies
456 report only the presence of hexadecanamide while this work reports 26 different fatty acid amides
457 predominantly as primary, methyl, or pyrrolidine amides in bio-crude from *N. gaditana*. These amides
458 may be formed from ammonia, methylamine, and pyrrolidine, while the remaining amines most likely
459 react faster with other products.

460 Prior to deamination or decarboxylation, amino acids may condensate to form DKPs which can
461 decompose to pyrazines at temperature $> 300\text{ }^{\circ}\text{C}$.^{14, 37} Pyrazines are likely displaced to the aqueous
462 phase and are too volatile to be detected with the method used in this work.³⁶

463 The interaction of carbohydrates and proteins is often mentioned in HTL studies. Amino acids and
464 ammonia may react with saccharides in Strecker degradation or Amadori rearrangements ultimately
465 leading to formation of pyrazines.³⁶ It is also well established that hemicellulose initially degrades to
466 furfurals and furans which can react with ammonia to form hydroxypyridines.³⁷ However, these
467 compounds were not detected in this work. Another potential interaction is the reduction of Schiff
468 bases formed between ammonia and saccharides to form aniline. Both the amine and the ortho position
469 of aniline are nucleophilic and reaction with for example β -keto acids would lead to favorable six
470 membered ring formation and subsequent reduction to form quinolines. However, it should be noted
471 that the formation of aniline has not been shown in the literature.

472 Formation of stable five and six membered rings with subsequent reduction to obtain conjugated
473 double bonds or in many cases aromaticity is a common theme for products of carbohydrate and
474 protein degradation in HTL. Carbohydrates are especially noted for the degradation products of
475 alkylated cyclopentanones and cyclopent-2-enones, which are too volatile to be detected with the

method used while oxygenated aromatics in the form of phenolics, catechols, and hydroquinones are also abundant.⁵⁵ Cyclopent-2-enones can further react in either two consecutive aldol condensations or one aldol condensation followed by a nucleophilic enol reaction and subsequent reduction to form indenones. Cyclopentanones can also undergo aldol condensation and the α -position is stabilized by sp^2 orbitals of the ketone making the 3-position of the cyclopentanone nucleophilic leading to ring closure and reduction forming indanones. Indanones and indenones may also be formed from condensation of furfurals.⁴¹

Nucleophilic reaction of the ortho position of phenol with amino acids would add an electron withdrawing ketone making the already favorable 5-membered ring formation possible through nucleophilic attack of the amine and subsequent elimination of the hydroxyl group. Finally, reduction would lead to formation of alkylated indoles which are observed with a large alkyl variation. Both amino acids and saccharides also degrade to multiple small carboxylic acids which are mostly water soluble or too volatile to be detected with the method used and not included in the pathways.^{30, 55}

Chromen-2-ones are observed naturally in plant material being formed from cinnamic acid and it is likely that these are simply released from the biomass during the HTL process. Another possible source of formation could be the nucleophilic reaction between the hydroxyl group of resorcinols and β -keto acids (present in aqueous phase) followed by favorable six-membered ring formation through the para directing nucleophilic attack of phenol and subsequent reduction to form chromen-2-ones. While a number of alkylated benzenediols were observed none of these were from resorcinol, thus the pathway is questionable.

The most notable difference between bio-crude of lignocellulosic and other high carbohydrate feedstocks is the presence of multiple alkylated phenolics, catechols, and hydroquinones from which the diversity is much higher than reported from other studies.⁵⁶ The reaction network of lignin is

499 complex, consisting of breakages of polymeric ether bonds and demethoxylation producing a multitude
500 of phenolic compounds which to a large extent are similar to those of cellulose and hemicellulose.⁵⁵
501 These compounds consist of phenolics, catechols, hydroquinones, and acetophenones.

502 **4. Conclusion**

503 The composition of semi-volatile compounds in bio-crude from HTL of microalgae, macroalgae,
504 DDGS, and *M. x giganteus* was analyzed by GCxGC-TOFMS with in-situ silylation. Numerous C₁₆,
505 C₁₈, and C₂₀ fatty acids were detected, a complexity, which has previously not been reported. Several
506 monoglycerides were observed, especially in bio-crudes from low lipid containing feedstocks
507 highlighting the differences in accessibility of cell structures. In the presence of protein, fatty acid
508 amides may be obtained, which included 24 different compounds; most were primary, methyl, and
509 pyrrolidine amides present almost solely for protein and lipid rich feedstocks (*N. gaditana*).
510 Alkylated indoles and quinolines were abundant in bio-crudes from protein rich feedstocks. Indoles are
511 likely formed from phenolics and amino acids while quinolines may derive from reactions between
512 ammonia, saccharides, and β -keto acids.
513 Several diverse carbonyls were observed which included acetophenones, indanones/indenones, and
514 chromen-2-ones. Acetophenones were formed from lignin in *M. x giganteus*, indanones/indenones were
515 derived from further reaction of carbohydrate degradation products which may involve cyclic
516 oxygenates, while chromen-2-ones most likely occur naturally in the feedstock.
517 A diversity of alkylated benzenediols was observed from carbohydrate rich feedstocks (DDGS and *L.*
518 *hyperborea*) while their abundances were increased along with the presence of alkylated phenolics
519 when processing a lignin and carbohydrate rich feedstock (*M. x giganteus*).

520 This work sheds additional light on the complexity of a part of HTL bio-crude, which is difficult to
521 analyze and provides useful information on characterization of semi-volatile compounds of HTL bio-
522 crude.

523 Acknowledgements

524 We would like to acknowledgement Professor Bo Brummerstedt Iversen and his research group for
525 providing access to process equipment and CHNS-O analysis. We thank the Innovation Fund Denmark
526 (Grant No. 1305-00030B) and National Centre for Food and Agriculture (Denmark) for funding.

527

528

529 References

- 530 (1) Kang, S.; Li, X.; Fan, J.; Chang, J. *Renewable and Sustainable Energy Reviews* **2013**, *27*, 546-558
531 (2) Zhou, C.; Xia, X.; Lin, C.; Tong, D.; Beltramini, J. *Chem. Soc. Rev.* **2011**, *40*, 5588-5617
532 (3) Barreiro, D. L.; Prins, W.; Ronsse, F.; Brilman, W. *Biomass Bioenergy* **2013**, *53*, 113-127
533 (4) Peterson, A. A.; Vogel, F.; Lachance, R. P.; Fröling, M.; Antal, M. J.; Tester, J. W. *Energy Environ. Sci.* **2008**, *1*,
534 32-65
535 (5) Elliott, D. C.; Biller, P.; Ross, A. B.; Schmidt, A. J.; Jones, S. B. *Bioresour. Technol.* **2015**, *178*, 147-156
536 (6) Biller, P.; Ross, A. B. *Bioresour. Technol.* **2011**, *102*, 215-225
537 (7) Vardon, D. R.; Sharma, B. K.; Scott, J.; Yu, G.; Wang, Z.; Schideman, L.; Zhang, Y.; Strathmann, T. J. *Bioresour.*
538 *Technol.* **2011**, *102*, 8295-8303
539 (8) Valdez, P. J.; Nelson, M. C.; Feaeth, J. L.; Wang, H. Y.; Lin, X. N.; Savage, P. E. *Energy Fuels* **2014**, *28*, 67-75
540 (9) Karagöz, S.; Bhaskar, T.; Muto, A.; Sakata, Y. *Fuel* **2005**, *84*, 875-884
541 (10) Yeh, T. M.; Dickinson, J. G.; Franck, A.; Linic, S.; Thompson Jr, L. T.; Savage, P. E. *J. Chem. Technol.*
542 *Biotechnol.* **2012**, *88*, 13-24
543 (11) Pedersen, T. H.; Rosendahl, L. *Biomass Bioenergy* **2016**, *83*, 206-215
544 (12) Leonardis, I.; Chiaberge, S.; Fiorani, T.; Spera, S.; Battistel, E.; Bosetti, A.; Cesti, P.; Reale, S.; De Angelis, F.
545 *ChemSusChem* **2013**, *6*, 160-167
546 (13) Torri, C.; Barreiro, D. L.; Conti, R.; Fabbri, D.; Brilman, W. *Energy Fuels* **2016**, *30*, 1135-1144
547 (14) Villadsen, S. R.; Dithmer, L.; Forsberg, R.; Becker, J.; Rudolf, A.; Iversen, S. B.; Iversen, B. B.; Glasius, M.
548 *Energy Fuels* **2012**, *26*, 6988-6998
549 (15) Farajzadeh, M. A.; Nouri, N.; Khorram, P. *TrAC* **2014**, *55*, 14-23
550 (16) Worton, D. R.; Zhang, H.; Isaachman-VanWertz, G.; Chan, A. W. H.; Wilson, K. R.; Goldstein, A. H. *Environ.*
551 *Sci. Technol.* **2015**, *49*, 13130-13138
552 (17) Biller, P.; Riley, R.; Ross, A. B. *Bioresour. Technol.* **2011**, *102*, 4841-4848
553 (18) Díaz-Vázquez, L. M.; Rojas-Pérez, A.; Fuentes-Caraballo, M.; Robles, I. V.; Jena, U.; Das, K. C. *Front. Energy*
554 *Res.* **2015**, *3*, DOI 10.3389/fenrg.2015.00006.

555 (19) Zhu, Z.; Toor, S. S.; Rosendahl, L.; Chen, G. *Environ. Prog. Sus. Energy* **2014**, *33*, 737-743
 556 (20) Villas-Bôas, S. G.; Smart, K. F.; Sivakumaran, S.; Lane, G. A. *Metabolites* **2011**, *1*, 3-20
 557 (21) Zhang, H.; Worton, D. R.; Shen, S.; Nah, T.; Isaachman-VanWertz, G.; Wilson, K. R.; Goldstein, A. H. *Environ.*
 558 *Sci. Technol.* **2015**, *49*, 9768-9777
 559 (22) Isaachman, G.; Kreisberg, N. M.; Yee, L. D.; Worton, D. R.; Chan, A. W. H.; Moss, J. A.; Hering, S. V.;
 560 Goldstein, A. H. *Atmospheric Measurement Techniques Discussion* **2014**, *7*, 7495-7529
 561 (23) Babushok, V. I.; Linstrom, P. J.; Reed, J. J.; Zenkevich, I. G.; Brown, R. L.; Mallard, W. G.; Stein, S. E. *J.*
 562 *Chrom. A* **2007**, *1157*, 414-421
 563 (24) Harris, D. C. *Quantitative Chemical Analysis*. Freeman: New York, NY, 2010; p 580.
 564 (25) Ross, A. B.; Biller, P.; Kubacki, M. L.; Li, H.; Lea-Langton, A.; Jones, J. M. *Fuel* **2010**, *89*, 2234-2243
 565 (26) Brown, T. M.; Duan, P.; Savage, P. E. *Energy Fuels* **2010**, *24*, 3639-3646
 566 (27) Chiaberge, S.; Leonardis, I.; Fiorani, T.; Bianchi, G.; Cesti, P.; Bosetti, A.; Crucianelli, M.; Reale, S.; De
 567 Angelis, F. *Energy Fuels* **2013**, *27*, 5287-5297
 568 (28) Valdez, P. J.; Dickinson, J. G.; Savage, P. E. *Energy Fuels* **2011**, *25*, 3235-3243
 569 (29) Barreiro, D. L.; Riede, S.; Hornung, U.; Kruse, A.; Prins, W. *Algal Res.* **2015**, *12*, 206-212
 570 (30) Klingler, D.; Berg, J.; Vogel, H. J. *Supercritical Fluids* **2007**, *43*, 112-119
 571 (31) Yang, W.; Li, X.; Li, Z.; Tong, C.; Feng, L. *Bioresour. Technol.* **2015**, *196*, 99-108
 572 (32) Changi, S.; Brown, T. M.; Savage, P. E. *Chem. Eng. J.* **2012**, *189-190*, 336-345
 573 (33) Toor, S. S.; Rosendahl, L.; Rudolf, A. *Energy* **2011**, *36*, 2328-2342
 574 (34) Jazrawi, C.; Biller, P.; He, Y.; Montoya, A.; Ross, A. B.; Maschmeyer, T.; Haynes, B. S. *Algal Res.* **2015**, *8*, 15-
 575 22
 576 (35) Tokusoglu, Ö.; Ünal, M. K. *J. Food Sci.* **2003**, *68*, 1144-1148
 577 (36) Madsen, R. B.; Biller, P.; Jensen, M. M.; Becker, J.; Iversen, B. B.; Glasius, M. *Energy Fuels* **2016**, *30*, 10470-
 578 10483
 579 (37) Croce, A.; Battistel, E.; Chiaberge, S.; Spera, S.; De Angelis, F.; Reale, S. *ChemSusChem* **2016**, *9*, 1-12
 580 (38) Yang, W.; Li, X.; Liu, S.; Feng, L. *Energy Conversion Management* **2014**, *87*, 938-945
 581 (39) Xu, Y.; Duan, P.; Wang, F. *Fuel Process. Tech.* **2015**, *130*, 268-274
 582 (40) Anastasakis, K.; Ross, A. B. *Fuel* **2015**, *139*, 546-553
 583 (41) Deniel, M.; Haarlemmer, G.; Roubaud, A.; Weiss-Hortala, E.; Fages, J. *Sus. Energy Fuels* **2017**, DOI
 584 10.1039/C6SE00065G.
 585 (42) Yan, L.; Cui, Y.; Gou, G.; Wang, Q.; Jiang, M.; Zhang, S.; Hui, D.; Gou, J.; Zhou, Z. *Composite Part B* **2017**,
 586 *112*, 8-14
 587 (43) Akhtar, J.; Amin, N. A. S. *Renewable and Sustainable Energy Reviews* **2011**, *15*, 1615-1624
 588 (44) Kumar, S.; Gupta, R. B. *Energy Fuels* **2009**, *23*, 5151-5160
 589 (45) Biller, P.; Lawson, D.; Madsen, R. B.; Glasius, M. *Biomass Conversion Biorefinery* **2017**, DOI
 590 10.1007/s13399-016-0230-x.
 591 (46) Teri, G.; Luo, L.; Savage, P. E. *Energy & Fuels* **2014**, *28*, 7501-7509
 592 (47) Biller, P.; Sharma, B. K.; Kunwar, B.; Ross, A. B. *Fuel* **2015**, *159*, 197-205
 593 (48) Jensen, C. U.; Hoffmann, J.; Rosendahl, L. *Fuel* **2016**, *165*, 536-543
 594 (49) Tamburini, D.; Lucejko, J. J.; Ribechini, E.; Colombini, M. P. *J. Anal. Appl. Pyrolysis* **2016**, *118*, 249-258
 595 (50) Bourgaud, F.; Hehn, A.; Larbat, R.; Doerper, S.; Gontier, E.; Kellner, S.; Matern, U. *Phytochem. Rev.* **2006**, *5*,
 596 293-308
 597 (51) Volkman, J. K.; Barreiro, D. L.; Dunstan, G. A.; Jeffrey, S. W. *Org. Geochem.* **1992**, *18*, 131-138
 598 (52) Villaverde, J. J.; Domingues, R. M. A.; Freire, C. S. R.; Silvestre, A. J. D.; Pascoal Neto, C.; Ligerio, P.; Vega, A.
 599 *J. Agric. Food. Chem.* **2009**, *57*, 3626-3631

600 (53) Deniel, M.; Haarlemmer, G.; Roubaud, A.; Weiss-Hortala, E.; Fages, J. *Ren. Sus. Energy Rev.* **2016**, *54*, 1632-
601 1652
602 (54) Changi, S.; Zhu, M.; Savage, P. E. *ChemSusChem.* **2012**, *5*, 1743-1757
603 (55) Pedersen, T. H.; Rosendahl, L. *Biomass Bioenergy* **2015**, *83*, 206-215
604 (56) Pedersen, T. H.; Grigoros, I. F.; Hoffmann, J.; Toor, S. S.; Daraban, I. M.; Jensen, C. U.; Iversen, S. B.;
605 Madsen, R. B.; Glasius, M.; Arturi, K. R.; Nielsen, R. P.; Søgaaard, E. G.; Rosendahl, L. A. *Applied Energy* **2016**, *162*,
606 1034-1041

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626 Table 1. Normalized peak responses of the most abundant fatty acids (only peak responses > 25).

	RI	<i>N. gaditana</i>	<i>C. vulgaris</i>	<i>L. hyperborea</i>	DDGS	<i>M. x giganteus</i>
Capric acid (C ₁₀)	1480	40	15	14	14	ND
Lauric acid (C ₁₂)	1678	36	11	17	8	ND
Myristic acid (C ₁₄)	1879	523	50	255	16	15
Pentadecylic acid (C ₁₅)	1941	29	7	18	7	8
?-hexadecenoic acid (C ₁₆)	2051	ND	71	6	4	ND
?-hexadecenoic acid (C ₁₆)	2060	873	68	142	3	5
Palmitelaidic acid (C ₁₆)	2065	581	71	10	ND	ND
Palmitic acid (C ₁₆)	2084	1500	1098	778	1098	302
unknown fatty acid (C ₁₇) (branched)	2140	7	45	ND	ND	ND
?,?-octadecadienoic acid (C ₁₈)	2240	ND	29	ND	ND	ND
Linoleic acid (C ₁₈)	2246	53	23	17	67	6
Oleic acid (C ₁₈)	2254	271	479	443	841	29
trans-9-octadecenoic acid (C ₁₈)	2260	255	505	ND	816	13
Stearic acid (C ₁₈)	2280	62	439	357	289	224
?,?-octadecadienoic acid (C ₁₈)	2297	ND	95	10	410	ND
?,?-octadecatetraenoic acid (C ₁₈)	2298	ND	63	16	ND	ND
?,?-octadecadienoic acid (C ₁₈)	2301	5	ND	5	194	ND
?,?-octadecadienoic acid (C ₁₈)	2306	ND	61	9	310	ND
?,?-octadecadienoic acid (C ₁₈)	2311	ND	59	6	ND	ND
?,?-octadecadienoic acid (C ₁₈)	2319	ND	ND	ND	42	ND
?,?-octadecadienoic acid (C ₁₈)	2324	ND	18	ND	184	ND
?,?-octadecadienoic acid (C ₁₈)	2336	12	ND	ND	64	ND
?,?-octadecadienoic acid (C ₁₈)	2342	23	300	37	736	ND
?,?-octadecatetraenoic acid (C ₁₈)	2352	10	35	ND	11	ND
1,3-propanediol palmitate	2366	ND	ND	29	ND	28
1,3-propanediol stearate	2567	ND	ND	ND	ND	52
Lignoceric acid (C ₂₄)	2886	2	4	67	8	11
α -glyceryl myristate	2433	ND	ND	8	ND	ND
β -glyceryl palmitate	2593	ND	4	16	ND	44
α -glyceryl palmitate	2628	20	44	322	16	316
β -glyceryl stearate	2791	ND	3	14	ND	15
α -glyceryl stearate	2826	12	58	356	18	313

? means that the configuration is unknown

629 Table 2. Normalized peak responses of the most abundant alcohols (only peak responses > 3).

	RI	<i>N. gaditana</i>	<i>C. vulgaris</i>	<i>L. hyperborea</i>	DDGS	<i>M. x giganteus</i>
Phenylethanol	1242	ND	4	ND	4	ND
Phosphate	1287	ND	4	13	12	ND
Glycerol	1290	ND	5	6	5	4
Glycol compound	1530	ND	ND	10	ND	ND
Glycol compound	1545	ND	ND	56	ND	ND
Dodecan-1-ol	1591	1	2	8	1	1
Levoglucosan	1727	3	6	10	6	ND
Tetradecan-1-ol	1790	ND	16	14	ND	ND
3,7,11,15-tetramethylhexadec2-en-1-ol	1850	42	ND	ND	ND	ND
(1E)-o-methyloxyme-d-Glucose	1957	3	2	ND	ND	ND
Hexadecan-1-ol	1991	ND	18	ND	1	ND
Heptadecan-1-ol	2092	ND	5	ND	ND	ND
Octadecan-1-ol	2190	ND	2	3	1	ND
3,7,11,15-tetramethylhexadecan-1-ol	2379	ND	17	67	8	66
Saccharide	2407	ND	ND	ND	14	ND
Saccharide	2429	ND	ND	ND	13	ND
3,7,11,15-tetramethyloctadecan-1-ol	2581	ND	14	74	ND	120
Saccharide	2597	ND	ND	ND	325	ND
Saccharide	2606	ND	ND	ND	109	ND
?-triacontenol	3366	10	ND	ND	ND	ND
?-triacontenol	3375	50	ND	ND	ND	ND
?-hentriacontenol	3430	7	ND	ND	ND	ND
?-hentriacontenol	3434	15	ND	ND	ND	ND
?-dotriacontenol	3482	135	ND	ND	ND	ND
?-dotriacontenol	3500	344	ND	ND	ND	ND

630 ? means that the configuration is unknown

631

632

633

634

635

636 Table 3. Normalized peak responses of the most abundant phenolics and benzenediols (only peak
637 responses > 15).

Phenolics	RI	<i>N. gaditana</i>	<i>C. vulgaris</i>	<i>L. hyperborea</i>	DDGS	<i>M. x giganteus</i>
2-Methoxyphenol	1234	ND	ND	ND	ND	51
3-Methoxyphenol	1242	ND	ND	ND	ND	86
2-Ethylphenol	1248	2	6	ND	5	77
3-Ethylphenol	1254	5	4	4	7	72
2-Methoxy-5-methylphenol	1333	ND	ND	ND	ND	27
ethylguaiaicol	1412	ND	ND	ND	ND	125
2,6-Dimethoxyphenol	1419	ND	ND	ND	ND	33
4-ethyl-3-methylguaiaicol	1497	ND	ND	ND	ND	38
ethylsyringol	1570	ND	ND	ND	ND	34
4-Hydroxyphenylethanol	1592	8	ND	9	15	ND
4-propansyringol	1648	ND	ND	ND	ND	19
(Z)-3-(p-phenol)-prop-2-en-1-ol	1658	ND	ND	ND	4	22
(E)-3-(p-phenol)-prop-2-en-1-ol	1663	ND	ND	ND	11	47
(E)-3-(p-phenol)-prop-2-en-1-ol + C ₁	1716	ND	ND	ND	ND	21
(E)-3-(p-phenol)-prop-2-en-1-ol + C ₂	1730	ND	ND	ND	ND	24
coniferyl alcohol moiety	1782	ND	ND	ND	ND	33
coniferyl alcohol moiety	1800	ND	ND	ND	ND	25
4-(3-hydroxypropyl)-2-methoxyphenol	1843	ND	ND	ND	ND	16
Benzenediols						
1,2-Benzenediol	1328	ND	ND	ND	8	77
1,2-Benzenediol	1331	ND	ND	ND	6	30
4-Methylcatechol	1405	ND	ND	6	17	55
3-Methylcatechol	1415	ND	ND	0	5	11
Hydroquinone	1420	ND	ND	7	ND	3
3-ethylcatechol	1472	ND	ND	ND	3	13
4-ethylcatechol	1480	ND	ND	ND	57	256
?,?-dimethylcatechol	1482	ND	ND	ND	9	12
Methylhydroquinone	1482	ND	ND	17	ND	28
Catechol + C ₂	1483	ND	ND	22	ND	ND
Catechol + C ₃	1549	ND	ND	8	15	9
3-propylcatechol	1559	ND	ND	21	42	ND
Catechol + C ₄	1587	ND	ND	ND	12	28
Catechol + C ₄	1612	ND	ND	ND	8	10

Hydroquinone + C ₃	1613	ND	ND	ND	7	ND
5-ethyl-3-methoxy-1,2-benzenediol	1622	ND	ND	ND	ND	23
1,2,3-benzenetriol	1622	3	10	40	6	ND
Hydroquinone + C ₃	1664	ND	ND	13	5	ND
3-((1E,3E)-tetradeca-1,3-dienyl)-1,2-benzenediol (A)	2431	ND	ND	ND	ND	11
3-((1E,3E)-pentadeca-1,3-dienyl)-1,2-benzenediol (B)	2476	9	ND	ND	ND	15
(E)-((3-(hexadec-1-enyl)-1,2-benzenediol	2516	ND	ND	ND	ND	31
((3-((1E,3E)-hexadeca-1,3-dienyl)-1,2-benzenediol (C)	2541	20	ND	ND	ND	10
(Z)-((3-(hexadec-1-enyl)-1,2-benzenediol (A)	2551	ND	ND	ND	ND	40
(E)-((3-(octadec-1-enyl)-1,2-benzenediol (B)	2609	ND	ND	ND	ND	13
3-docosyl-1,2-benzenediol	2643	ND	ND	ND	ND	12

638 ? means that the configuration is unknown

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653 Table 4. Normalized peak responses of the most abundant nitrogen containing compounds (only peak
654 responses > 10).

		<i>N. gaditana</i>	<i>C. vulgaris</i>	<i>L. hyperborea</i>	DDGS	<i>M. x giganteus</i>
<i>n</i> -butyl-pyrrolidin-2-one	1365	15	6	ND	8	0
<i>n</i> -pentylsuccinimide	1398	15	ND	ND	ND	ND
3-methylquinoline	1403	21	9	ND	5	ND
7-methylindole	1414	28	21	20	16	ND
Indole	1459	18	25	ND	5	ND
1H-Pyrido[3,4-b]indole, 2,3,4,9-tetrahydro-1-(1-methylethyl)-	1468	23	45	60	53	ND
5-Isopropyl-2,4-imidazolidinedione	1494	ND	ND	14	ND	ND
(5R,8R,8aS)-8-Methyl-5-butyl-1,2,3,5,8,8a-hexahydroindolizine	1494	18	ND	ND	ND	ND
2,5-dimethylindole	1620	9	11	3	4	ND
3-ethan-2-olindole	1921	16	13	12	19	3
<i>N</i> -methyltetradecanamide	2044	13	ND	ND	ND	ND
(<i>E</i>)-9-hexadecenamide	2196	23	ND	ND	ND	ND
(<i>Z</i>)-9-hexadecenamide	2199	13	ND	ND	ND	ND
Hexadecanamide	2218	68	5	ND	2	ND
(<i>E</i>)- <i>N</i> -methyl-hexadec-9-enamide	2232	34	ND	ND	ND	ND
(<i>Z</i>)- <i>N</i> -methyl-hexadec-9-enamide	2237	17	ND	ND	ND	ND
<i>N</i> -methylhexadecanamide	2256	92	34	ND	ND	ND
(<i>Z</i>)-9-octadecenamide	2275	19	ND	ND	ND	ND
Octadecanamide	2298	40	18	ND	ND	ND
Eicosanamide	2440	28	ND	ND	ND	ND
Unknown fatty acid amide	2478	9	ND	ND	ND	15
Unknown eicosenamide	2529	20	ND	ND	ND	ND
Unknown fatty acid amide	2551	33	ND	ND	ND	ND
Hexadecanoic acid, pyrrolidide	2628	26	ND	ND	4	ND

655

656

657

658

659 Table 5. Normalized peak responses of the most abundant carbonyls (only peak responses > 5).

		<i>N. gaditana</i>	<i>C. vulgaris</i>	<i>L. hyperborea</i>	DDGS	<i>M. x giganteus</i>
2,2,5,5-tetramethylcyclopent-3-en-1-one	1210	ND	ND	ND	2	7
2,3,4,5,6,7-hexahydro-1H-Inden-1-one	1282	ND	ND	ND	ND	7
2,3-dihydro-1H-Inden-1-one	1303	ND	ND	11	3	57
2,3-dihydro-2-methyl-1H-Inden-1-one	1329	ND	ND	31	9	23
2,3-dihydro-3-methyl-1H-Inden-1-one	1341	ND	ND	ND	4	8
7-methylindan-1-one	1419	ND	ND	ND	ND	20
2-hydroxyacetophenone	1435	ND	ND	ND	2	8
6-Methylindan-1-one	1436	ND	ND	ND	ND	17
2,3-dihydro-3,3-dimethyl-1H-Inden-1-one	1445	ND	ND	8	4	8
2,3-dihydro-2,2-dimethyl-1H-Inden-1-one	1457	ND	ND	21	ND	12
4-hydroxyacetophenone	1494	ND	ND	ND	4	25
7-hydroxychromen-2-one	1693	ND	ND	6	1	16
Homovanillic acid	1697	ND	ND	ND	ND	13
2,5-dihydroxyacetophenone	1718	ND	ND	ND	ND	12
2,4-dihydroxyacetophenone	1733	ND	ND	ND	ND	8
Unknown 7-hydroxychromen-2-one	1748	ND	ND	ND	ND	32
?-methyl-7-hydroxychromen-2-one	1760	ND	ND	4	ND	8
?-methyl-7-hydroxychromen-2-one	1761	ND	ND	20	ND	17
?-methyl-7-hydroxychromen-2-one	1777	ND	ND	8	ND	7
Unknown carbonyl	1785	ND	ND	ND	ND	8
6,10,14-trimethylpentadecan-2-one	1873	6	13	5	ND	ND
Unknown chromen-2-one	1874	ND	ND	8	ND	8
Unknown chromen-2-one	1884	ND	ND	11	ND	ND
Methyl vanillactate moiety	2473	ND	ND	ND	ND	8
Methyl vanillactate moiety	2534	ND	ND	ND	ND	20

660 ? means that the configuration is unknown

661

662

663

664

665

Figure 1. A) TIC of *N. gaditana* bio-crude. B) TIC of *N. gaditana* bio-crude with compound markers. Dark blue – nitrogen ring structures, Black – hydrocarbons, Pink – fatty acid amides, Yellow – fatty acids, Light yellow – alcohols, Purple – triterpenoids. Primary and secondary axes show retention on first and second column, respectively. Circles reflect the abundance of the peak.

Figure 2. A) TIC of *C. vulgaris* bio-crude. B) TIC of *C. vulgaris* bio-crude with compound class markers. Dark blue – nitrogen ring structures, Black – hydrocarbons, Pink – fatty acid amides, Yellow – fatty acids, Light yellow – alcohols, Purple – triterpenoids. Primary and secondary axes show retention on first and second column, respectively. Circles reflect the abundance of the peak.

Figure 3. A) TIC of *L. hyperborea* bio-crude. B) TIC of *L. hyperborea* bio-crude with compound markers. Dark blue – nitrogen ring structures, Black – hydrocarbons, Yellow – fatty acids, Light yellow – alcohols, Purple – triterpenoids, Dark green – benzenediols, Light green - carbonyls. Primary and secondary axes show retention on first and second column, respectively. Circles reflect the abundance of the peak.

Figure 4. A) TIC of DDGS bio-crude. B) TIC of DDGS bio-crude with compound markers. Dark blue – nitrogen containing ring structures, Black – hydrocarbons, Yellow – fatty acids, Light yellow – alcohols, Purple – triterpenoids, Dark green – benzenediols, Light green – carbonyls, Red - phenolics. Primary and secondary axes show retention on first and second column, respectively. Circles reflect the abundance of the peak.

Figure 5. A) TIC of bio-crude of *M. x giganteus* bio-crude. B) TIC of bio-crude of *M. x giganteus* bio-crude with compound markers. Dark blue – nitrogen ring structures, Black – hydrocarbons, Yellow – fatty acids, Light yellow – alcohols, Purple – triterpenoids, Dark green – benzenediols, Light green – carbonyls, Red - phenolics. Primary and secondary axes show retention on first and second column, respectively. Circles reflect the abundance of the peak.

711

712 *Figure 6. Normalized peak responses of fatty acids, alcohols, nitrogen-containing compounds,*
 713 *hydrocarbons, carbonyls, phenolics, and benzenediols detected in bio-crudes of N. gaditana, C.*
 714 *vulgaris, L. hyperborea, DDGS, and M. x giganteus.*

715

716

717

718 *Fig. 7. Reaction pathways for degradation of lipids, proteins, carbohydrates, and lignin during HTL.*

719

720

721

722