Characterization of liquid-liquid extraction fractions from lignocellulosic biomass by high performance liquid chromatography hyphenated to tandem high-resolution mass spectrometry

Carole Reymond, Alexis Dubuis, Agnès Le Masle, Cyril Colas, Ludovic Chahen, Emilie Destandau, Nadège Charon

Characterization of liquid-liquid extraction fractions from lignocellulosic biomass by high performance liquid chromatography hyphenated to tandem high-resolution mass spectrometry (HPLC/MS^n)

Carole Reymond[a], Alexis Dubuis[a], Agnès Le Masle[a,*], Cyril Colas[b,c], Ludovic Chahen[a], Emilie Destandau[b], Nadège Charon[a]

[a] *IFP Energies nouvelles, Rond-point de l'échangeur de Solaize, BP 3, 69360 Solaize, France*
[b] *Institut de Chimie Organique et Analytique, Université d'Orléans, CNRS UMR 7311, Rue de Chartres, 45067 Orléans, France*
[c] *Centre de Biophysique Moléculaire, CNRS UPR 4301, Université d'Orléans, rue Charles Sadron, 45071 Orléans, France*
* *Corresponding author. E-mail address: agnes.le-masle@ifpen.fr (A. Le Masle)*

Abstract

The conversion of lignocellulosic biomass is a major challenge in the field of renewable energies and bio-based chemicals. The diversity of biomasses and processes leads to complex products having a wide range of polarities and molecular weights. Nowadays, the molecular description of these oxygenated matrices is still largely incomplete and new analytical strategies are required to have a better understanding of biomass products properties. The present study proposes a reliable protocol based on successive liquid-liquid extractions prior to high performance liquid chromatography hyphenated to high-resolution tandem mass spectrometry (HPLC/MS^n) using a linear ion trap-Fourier transform ion cyclotron resonance mass spectrometer (LTQ/FT-ICR). The protocol allowed to fractionate an industrial sample coming from the sulfuric acid-based pretreatment of a wheat straw into four key chemical families: carbohydrates, organic acids, phenols and neutral compounds. Each fraction was separately analyzed, which limited matrix effects during mass spectrometry ionization step. Electrospray and atmospheric pressure chemical ionization sources were used in both positive and negative modes in order to ionize and detect a maximum of compounds. Thanks to HPLC/MS^n, structures of heavy lignin-carbohydrate complexes (LCC) were elucidated (up to 600 g/mol) as well as carbohydrate oligomers having acid functionalities. Mono, di, tri and tetra-aromatic compounds coming from lignin were also detected. The results reported in this paper demonstrate the complexity of pretreated biomass samples and propose an analytical approach from sample simplification to data treatment in order to describe the biomass composition.

Keywords: Lignocellulosic biomass, liquid-liquid extraction, high performance liquid chromatography, high-resolution mass spectrometry, tandem mass spectrometry

1. Introduction

Global increase in energy consumption combined with a depletion of fossil fuel and global warming lead to a growing concern for renewable energy sources[1–3]. Lignocellulosic biomass also called biomass of second generation is a promising and renewable resource to produce biofuels and valuable bio-based chemicals [4]. Coming from various feedstocks such as forest residues, agricultural wastes or dedicated crops, lignocellulosic biomass is the most abundant material of plant cell walls. This composite material consists of linear or branched polymeric carbohydrates (cellulose and hemicelluloses) and aromatic macromolecules of lignin [5]. Several transformation routes are studied to convert this complex raw material into biofuels and bio-products [6–8]. As regards biochemical transformation route, biomass is first pretreated to increase the accessibility of cellulose to enzymes that are used to hydrolyze this polysaccharide to produce glucose. During this step, the hemicelluloses were dissolved and hydrolyzed, which leads to complex aqueous matrices made up of several hundreds of oxygenated compounds that are distributed over a wide range of polarities and/or molecular weights, potentially being heat sensitive and having several chemical groups for a single molecule. For all these reasons but also because of the low concentrations of some analytes, the complete characterization of pretreated biomass sample is still a challenge [9–12]. A better understanding of relationships between the products composition and reactivity is expected to support development of innovative and efficient processes.

Mono-dimensional gas chromatography (GC) or comprehensive two-dimensional GC×GC coupled with flame ionization detection and/or mass spectrometry (MS) are often used in literature to characterize samples originating from biomass. However, these techniques presents limitation for compounds with a molecular mass above 200 g/mol which usually need derivatization step [13–15]. High performance liquid chromatography (HPLC) techniques are relevant complementary approaches to GC, especially when dealing with thermal sensitive, polar and/or high molecular weight compounds. To go deeper in the characterization and structural elucidation, high-resolution mass spectrometry (HRMS) and multi-stage tandem mass spectrometry (MS^n) are required [16]. Most of publications using HRMS with or without HPLC mainly focus on lignin derivative compounds [17–21]. Recently, Jarrell *et al.* investigated lignin compounds produced from an organosolv switchgrass by HPLC/MS^n [22]. However, when considering very complex samples, HPLC may lack resolution leading to peaks co-elutions and thus enhancement of matrix effects in MS ionization. To simplify matrix composition and to limit co-elutions, a fractionation step prior to analysis should be performed. Among these analytical pretreatments, liquid-liquid extraction (LLE) is an attractive technique easy to implement to separate compounds of different polarities [23–25]. Moreover, it allows to recover the whole sample, which is critical to fully characterize biomass products. Methyltertiobutylether (MTBE) was already used to extract phenols, organic acids and neutral degradation products from biomass hydrolysates [23]. MTBE extraction allows to isolate highly hydrophilic compounds such as carbohydrates from other oxygenated molecules [12,26]. An interesting methodology based on successive LLE and pH modifications has been proposed by Kanaujia et al. to go further the fractionation of another kind of biomass product (fast pyrolysis oils) [24]. In this way, multi-step LLE appears to be a promising approach to fractionate biomass products into selective chemical groups [25].

In this study, a new protocol was set up to fractionate an aqueous biomass product coming from the sulfuric acid-based pretreatment of a wheat straw in four fractions according to the main chemical

families: carbohydrates, organic acids, phenols and neutral compounds. Then, HPLC/MS^n analyses using two ionization sources *i.e.* electrospray (ESI) and atmospheric pressure chemical ionization (APCI) were performed to obtain structural information. The aim of this study is to characterize the whole sample from carbohydrates to aromatic compounds coming from the hydrolysation of hemicellulose and the solubilisation of lignin.

2. Experimental

2.1. *Chemicals*

The reagents used for the LLE protocol (methyl *tert*-butyl ether (MTBE) and methanol) were HPLC grade purchased from Sigma-Aldrich (Saint-Quentin-Fallavier, France). Both sulfuric acid (H_2SO_4, 96%) and sodium hydroxide (NaOH, 0.5 mol/L) were obtained from Carlo-Erba reagent (Val de Rueil, France). Hydrochloric acid (HCl, 0.5 mol/L) was purchased from VWR (Fontenay sous Bois, France). Methanol and formic acid used for LC/MS analysis were MS grade purchased from VWR (Fontenay sous bois, France). Deionized water was produced by a Milli-Q water purifier (Millipore SAS, Molsheim, France).

2.2. *Sample*

The biomass sample investigated in this work was provided by IFP Energies nouvelles (Solaize, France). It corresponded to the water soluble fraction obtained by a sulfuric acid-based pretreatment of a wheat straw. Xylose, glucose, 5-HMF (5-(hydroxymethyl)furfural) and furfural were quantified at 80.4 g/L, 8.5 g/L, 0.311 g/L and 0.081 g/L respectively using reference methods [27,28]. A pH of 2.1 was measured using a combined pH electrode (Fisher Scientific, Illkirch, France). The entire sample (including particles) was introduced in the LLE protocol. Sample and fractions were filtered before HPLC/MS analysis using Macherey-Nagel PTFE membrane filters (pore size: 0.20 µm, diameter: 13 mm, Düren, Germany).

2.3. *Liquid-liquid fractionation*

A fractionation protocol corresponding to successive LLE was set up to generate four fractions from the aqueous biomass sample: carbohydrates (AQ1), organic acids (AQ3), phenols (ORG3) and neutral products (ORG2) (Figure 1). To do so, 4.5 mL of the sample were introduced and weighed in a 30 mL separating funnel. Then 1.5 mL of MTBE were added and weighed in the separating funnel. After 30 s of vigorous shaking, the two phases were left 5 min for equilibration. The organic and aqueous phases were separated in two flasks. The LLE extraction was performed two more times on the aqueous phase using each time 1.5 mL of MTBE (4.5 mL used in total). The resulting organic (ORG1) phases were combined and weighed to perform a mass balance. AQ1 was kept for analysis. ORG1 was fully reintroduced in the separating funnel and 4 mL of 0.5 mol/L sodium hydroxide solution were added. This volume may be adapted according to the sample acidity to reach pH 12 in the aqueous fraction (AQ2a). Organic (ORG2) and aqueous (AQ2a) phases were shaken and collected after equilibration. ORG2 was concentrated by evaporation under a nitrogen flow and weighed for mass balance. The dry extract was dissolved in 250 µL of methanol and transferred into a vial for analysis. AQ2a was acidified with 0.5 mol/L hydrochloric acid to reach pH 7 (AQ2b) under pH combined electrode control. AQ2b (9 mL) was then extracted in triplicate with MTBE (3x3 mL). AQ3 was collected for analysis. ORG3 was evaporated under nitrogen and the dry extract was dissolved in

250 µL of methanol. ORG3 was transferred into a vial for analysis. The entire LLE protocol was conducted at room temperature (20 ± 2°C).

2.4. *Liquid chromatography/high-resolution mass spectrometry*

All analyses were performed using an Agilent 1290 UHPLC system consisting of a binary pump, an autosampler, a temperature-controlled column compartment and a photodiode array (PDA) UV detector. The extra-column variance was measured to be equal to 9 µL². The chromatographic system was hyphenated to a Fourier-transform ion cyclotron resonance mass spectrometer (LTQ-FT-ICR Thermo Scientific; 7 T magnet). Ionization was carried out with the ESI and APCI sources, both working in positive or negative modes.

For HPLC/MS analysis, 1 µL of each sample was injected. Separation was achieved on a Kinetex C18 column (100 × 3 mm, particle size 2.6 µm, Phenomenex, France) at 30°C. The mobile phase solvents used were a 0.01% (v/v) formic acid in water (A) and a 0.01% (v/v) formic acid in methanol (B) at 600 µL/min. A linear gradient was used as follows: 0.0 - 4.0 min, 1% B; 4.0 - 30.0 min, from 1 to 99% B; 30.0 - 35.0 min, 99% B; 35.0 - 36.0 min, from 99 to 1% B; 36.0-40.0 min, 1% B. UV signal was recorded from 210 to 400 nm.

Ionization efficiency was optimized with a design of experiments approach according to a previous work [29]. For ESI source, a splitter was set up at the outlet of the HPLC column and at the entrance of the MS device to reduce the solvent flow (split ratio 1:2.2). To avoid excessive pressure in the UV cell, HPLC flow was split before PDA. Flow rates of sheath, auxiliary and sweep gases as well as vaporizer and transfer capillary temperatures, transfer capillary voltage, spray voltage and corona discharge values are summarized in Table 1. In order to have at least 10 points for each chromatographic peak, a resolving power of 12500 at *m/z* 400 was used for the FT-ICR mass spectrometer. The mass accuracy of the determinations was 2 ppm.

MS^n experiments were performed using the data dependent acquisition functionality. The most intense ion at each time was selected, subjected to isolation and then to fragmentation in the LTQ (linear ion trap in front of the FT-ICR mass spectrometer). An isolation window of 2 *m/z* was used and fragmentation was done with collision energy of 35%.

3. Results and discussion

3.1. *A selective LLE fractionation*

3.1.1. Validation of the LLE protocol

A fractionation protocol based on successive LLE was developed to improve HPLC/MS characterization of aqueous biomass samples. Four fractions, selective of targeted chemical families are produced from the whole biomass sample to limit co-elutions and to help chromatograms understanding. More precisely, high water-soluble compounds (*i.e.* carbohydrates) are expected to be recovered mainly in aqueous phase 1 (AQ1), non ionizable compounds at pH 12 mainly in organic phase 2 (ORG2), phenols (pKa > 9) mainly in organic phase 3 (ORG3) and carboxylic acids (pKa < 7) mainly in aqueous phase 3 (AQ3) as illustrated in Figure 1.

The first key step of the LLE protocol is to get on one hand an aqueous phase containing carbohydrates and, on the other hand, an organic phase containing all other oxygenated compounds.

MTBE was reported in previous studies as an efficient solvent to extract organic species from aqueous biomass samples [12,23]. To confirm these observations, partition experiments were realized in the MTBE/water solvent system buffered at pH 2.8 using 35 model molecules representative of the main oxygenated chemical families (Figure S1 and Table S1). These partition measurements demonstrated a good extraction of phenols, organic acids, aldehydes, ketones, esters and alcohols in ORG1 while very hydrophilic compounds such as carbohydrates and their derivatives were recovered in AQ1. The second step of the protocol consisted in addition of diluted sodium hydroxide solution in ORG1 to transfer ionizable compounds in aqueous phase 2 (AQ2) at pH 12, namely carboxylic acids and phenols. At this step, all non ionizable compounds at pH 12 (pKa > 14) remained in ORG2. Then, an acidification is performed on AQ2 to reach pH 7, intermediate between phenols pKa (9-10) and carboxylic acids pKa (4-5). Therefore, phenols are neutral at pH 7 and carboxylic acids under their ionic form, which allows a final MTBE extraction to separate re-protonated phenols in ORG3 from carboxylates in AQ3. Mass balances using model molecules were done on the global scheme to assess a controlled fractionation selectivity (data available in Supplementary information Figure S2). However, some neutral compounds (furfural, 5-HMF, ethanol, acetic acid, etc.) were partitioned between organic and aqueous phases and they were consequently detected in different LLE fractions.

3.1.2.HPLC-UV chromatograms of the biomass sample and LLE fractions

The LLE protocol was applied to an aqueous product coming from the sulfuric acid-based pretreatment of a wheat straw. A first extraction was performed to evaluate the mass repartition of main compounds contained in the sample. The resulting mass balance done on dry extracts of the four fractions revealed a significant proportion of carbohydrates and presence of minority products (AQ1: 89% w/w, ORG2: 2% w/w, ORG3: 4% w/w, AQ3: 3% w/w, losses: 2% w/w). The LLE protocol was repeated on the whole sample to produce the fractions used for analysis. Organic fractions were concentrated as described in the experimental section to improve the analysis of minor products. HPLC-UV chromatograms obtained for the whole sample and LLE fractions are presented in Figure 2.

HPLC-UV chromatogram of the entire sample presents about sixty peaks with very different intensities, revealing many co-elutions, and is therefore difficult to process without prior fractionation. All HPLC-UV chromatograms are overlaid to suggest effective fractionation of compounds in the LLE fractions. It is interesting to notice that some carbohydrate-type molecules in AQ1 carry chromophore groups responsible for their UV response. Hyphenation with high-resolution mass spectrometry is required to achieve the structural characterization of these polyfunctional compounds.

3.2. *Characterization of LLE fractions by HPLC/MS^n*

Despite some fractions represented a small part of the entire sample, their characterization is crucial. Indeed, during the biomass pretreatment which is the first step of the biochemical transformation, some inhibitors of the hydrolysis enzymatic and alcoholic fermentation are released. These inhibitors can have an impact on the yield of the reaction even at low concentration. Some of them are already known such as phenolic compounds or furans, however many of them are still unknown. In order to characterize all classes of components in the initial biomass sample, LLE fractions were analyzed by HPLC/MS^n using reverse phase chromatography with ESI and APCI ionizations in positive and negative modes (ESI-/+ and APCI-/+). Deprotonated $[M-H]^-$ molecules and protonated $[M+H]^+$ molecules are

mostly formed in negative and positive modes respectively. High-resolution mass spectrometry allowed to measure *m/z* ratio with an accuracy of 2 ppm. Molecular formulae were calculated considering carbon, hydrogen and oxygen between 1 and 100 atoms for each. For ESI+ and APCI+, nitrogen was also considered for molecular formulae calculations. In total for the four LLE fractions, 480 molecular formulae were measured in ESI-, 444 in ESI+, 109 in APCI- and 122 in APCI+. Although ESI- was the most suitable ionization mode for this sample by detecting the highest number of compounds, other detections were complementary to ESI- and thus required for a complete characterization of the sample. A Venn diagram is presented in supplementary information to illustrate this complementarity (Figure S3). To represent all detected components, van Krevelen diagrams were built, representing H/C *versus* O/C ratios in order to reveal areas corresponding to chemical classes (Figure 3).

AQ3, ORG2 and ORG3 covered really specific areas on the van Krevelen diagram, which demonstrates the specificity of LLE protocol. AQ3 and ORG3 were mainly centered on an area located at H/C and O/C ratios of 1 and 0.4, this area is known as phenolic zone [21]. ORG2 might be related to lipids with rather high H/C ratios and low O/C ratios [21]. AQ1 was distributed into two different zones, including an area typically attributed to carbohydrates (H/C and O/C ratios being higher than 1.5 and 0.8 respectively). Discussion of each fraction composition is detailed next. Most abundant ions belonging to each of the four fractions were subjected to isolation and CID (collision-induced dissociation) experiments. High-resolution mass spectra were measured to obtain elemental compositions of fragments.

3.2.1.Aqueous fraction 1

According to LLE protocol, a large part (89% w/w) of the initial sample is focused in AQ1. Components present in this fraction have a high solubility in water and might be related to carbohydrate compounds. According to UV chromatogram in Figure 2, two groups of peaks can be distinguished: a first group corresponds to the less retained compounds on the non-polar column, having retention times inferior to 10 min (less than 12% of methanol in the HPLC mobile phase), while a second group consists on the strongly compounds on the HPLC column, eluting after 10 min. So it was assumed that AQ1 contains carbohydrates owing unsaturated chemical functions, which explains observations about retention behavior on a non-polar HPLC stationary phase and possibility to see peaks from UV detection. HPLC/MS analysis of AQ1 using several ionization conditions (ESI+/- and APCI +/-) delivered numerous compounds that were represented on a van Krevelen diagram (Figure 3). The size of the circle represents the mass intensity of the measured ions. First of all, by using different colors for retained (black dots) and un-retained compounds (green and orange dots) on chromatographic column, a new chemical family of compounds which might correspond to lignin-carbohydrate complexes (LCC) can be clearly separated from other carbohydrates. This class of molecules which combine lignin derivatives and carbohydrate units, was already related in the literature [30,31]. Due to the phenol moiety, LCC have lower H/C and O/C ratios than 'conventional' carbohydrates like glucose or xylose. This distinction can also be highlighted when taking into account the double bond equivalent (DBE) of compounds (Figure S4). For the same molecular mass, a retained compound exhibits a higher DBE than a non-retained compound due to its aromatic ring(s) attached to the carbohydrate part. This additional type of information could help usefully to assume structural elucidation of compounds analyzed in AQ1.

Considering non lignin-carbohydrate compounds whose retention times are inferior to 10 min, several groups can be distinguished from Figure 4:

(1) some compounds detected in both negative and positive ionization modes are located on a line where H/C ratios are two times higher than the O/C ones (H/C = 2×O/C).

(2) Under this line H/C = 2×O/C, compounds were ionized only in negative mode (green dots), which might indicate the presence of carboxylic acid function. One may suggests detection of uronic acids in this area of the van Krevelen diagram.

(3) Above the line H/C = 2×O/C, nitrogen compounds are mainly detected (orange dots).

The line H/C = 2×O/C corresponds to dehydration reactions with loss of a H_2O molecule. Compounds located on this line may be reaction intermediates, such the ones coming from degradation of xylose to produced furfural or from glucose to produce 5-hydromethylfurfural (5-HMF). Oligomers constituted of several $C_xH_{2x}O_x$ carbohydrate units (xylose, glucose, mannose, etc.) are located on this dehydration line due to the loss of water to condense two sugar units. From a mathematical point of view, a polymer with infinity of xylose (or glucose) units has H/C and O/C ratios which tend to 1.60 and 0.80 respectively (1.67 and 0.83 for glucose series). These two points are indicated by points (1) and (2) on Figure 4. Moreover, these two specific points correspond to the first dehydration of xylose (or glucose) meaning that compounds on the dehydration line under these points are carbohydrate degradation products.

Fragmentations in negative and positive-ion modes ESI were performed from MS^2 to MS^7 experiments. Usual carbohydrates neutral losses such as H_2O (18 Da), $C_2H_4O_2$ (60 Da), $C_5H_8O_4$ (132 Da), $C_5H_{10}O_5$ (150 Da) or $C_6H_{10}O_5$ (162 Da) are shown in Figure S5. These fragments confirm the presence of carbohydrates in AQ1 which are specifically extracted by the LLE protocol. Loss of xylose can give $C_5H_8O_4$ or $C_5H_{10}O_5$ fragments in negative-ion mode according to how xylose is linked to the rest of the molecule. $C_5H_{10}O_5$ fragment might also come from specific fragmentation of hexose ($^{0,1}X^-$ according to Domon and Costello nomenclature [32]). Xylose losses were more abundant than glucose ones ($C_6H_{10}O_5$), which is consistent with the higher concentration of xylose (80.4 g/L) than glucose (8.5 g/L) in the initial sample. Loss of $C_3H_4O_2$ (72 Da) corresponds to the remaining part of xylose after a loss of $C_2H_4O_2$. Moreover, for retained components (retention times between 10 and 20 min), many losses of CO_2 and $^{\bullet}CH_3$ radical happened mainly after MS^3 fragmentation stage. CO_2 loss is well known neutral fragment from carboxylic acid and $^{\bullet}CH_3$ radical from methoxy group on aromatic ring. These observations were consistent with the hypothesis of the existence of lignin-carbohydrate complexes.

To illustrate the fragmentation process with some examples of proposed structures, fragmentations of the most intense peaks on base peak chromatogram (BPC) are detailed. Peaks having various retention times were chosen and located on the van Krevelen diagram dedicated to AQ1 (Figure 5). Domon and Costello nomenclature was used to depict the fragmentation spectrum.

For retained compounds (eluted after 10 min), the most intense peak using negative-ion mode ESI (Figure 5) but also the other detection modes was peak (b) eluted at 13.5 min which can be assessed to feruloyl xyloside (Figure 6). In negative-ion mode ESI, MS^2 and MS^3 consisted in the loss of the xylose part, leading to Y_0^- fragment, then fragmentations of the lignin part started with $^{\bullet}CH_3$ and CO_2

losses. Positive-ion mode ESI provided complementary information with a different fragmentation process: Z_0^+ fragment was formed, then methoxy group fragmentation consisted in a CH_4O loss instead of a $^{\bullet}CH_3$ radical loss. Last fragmentation step was dedicated to a loss of CO. Moreover, UV spectrum of compound b was the same as ferulic acid, which confirmed its identification. Compound b' had the same molecular formula and UV spectrum as compound b but its fragmentation was different. This isomer might correspond to a ferulic acid linked to a xylose unit on a different position (for example on the hydroxyl group of ferulic acid instead of carboxylic acid group). Other fragmentations confirmed the presence of LCC for retained compounds. Fragmentations of peaks a and c allowed to identify coumaroyl xyloside and feruloyl dixyloside respectively (Figure 6). The carbohydrate part was removed first and only afterwards the fragmentation of the phenolic part took place. The positions of the glycosidic bond and the bond between the phenolic and the carbohydrate parts could not be strictly determined. However, by considering all the fragments, the phenolic group might be link to the xylose on its position 4 for compounds (a), (ab) and (c).For the compound (c), three positions of the glycosidic bond were possible: 1-4, 1-3 and 1-2 (Figure S6). For some components and especially for molecules with a phenol moiety that is chemically complex (*ie* polyaromatic), no detailed structure could be proposed. As regards peak d, the carbohydrate part was a hexose unit while the phenolic part has a $C_{20}H_{19}O_7$ formula. MS^3 to MS^7 fragmentations gave precious information on the lignin part such as $^{\bullet}CH_3$ losses, which may indicate methoxy group on aromatic ring. Also, phenolic part has a DBE of 11 for only 20 atoms of carbon which should indicate two aromatic rings.

Concerning carbohydrates eluted at the beginning of the run (between 0 and 10 min), many compounds had similar fragments. Indeed, many of them are composed of an acidic sugar (such as methylglucuronic acid or glucuronic acid) attached to 1, 2 or 3 pentose and/or hexose units. MS^2 spectra of peaks α, β and γ are shown in Figure 7. Positions of the chemical groups were not taken into account here. Methylglucuronic acid group was confirmed by detection of a fragment at *m/z* 207.0510 which lost CH_4O corresponding to a methoxy group. Peaks δ, ε and ζ were identified as glucuronic acid with one, two and three pentose units respectively.

The highest intensity peak detected on BPC in negative-ion mode ESI (peak η, *m/z* 513.1469/$C_{19}H_{29}O_{16}$) had similar fragments than peak β and was a bit more retained on the HPLC stationary phase. It might have similar structure with an extra acetyl group ($COCH_3$) linked to a pentose.

For compounds eluted between 0 and 10 min, sugar acids were detected in negative-ion mode ESI. Many oligomers combining sugar acids and pentose and/or hexose units were present in AQ1. Moreover, on the van Krevelen diagram (Figure 5), same components with different numbers of pentose units are located on a same line which converges toward point (1). Thanks to this linear structuration and to the molecular formulae, it should be possible to make predictions on the structures of unknown molecules without fragmentation experiments, provided one compound has already been identified on the same line. For example, feruloyl xyloside and feruloyl dixyloside (peaks b and c respectively) were identified by the use of MS^n analyses. The line going through these two points and point (1) also goes through $C_{25}H_{34}O_{16}$ (H/C = 1.36; O/C = 0.34) and $C_{30}H_{42}O_{20}$ (H/C = 1.40; O/C = 0.66) molecules. Thus, they may be assessed to feruloyl trixyloside and tetraxyloside.

It has to be noticed that for this complex fraction, despite the use of chromatographic separation ahead the mass spectrometer, it remained some co-elutions especially at the beginning of the run for the very polar compounds. These co-elutions may lead to matrix effect and ion suppression in the atmospheric ion source. Thus, some compounds might to be not be detected. A dedicated chromatographic separation and ionization method should be optimized for this class of molecules.

3.2.2.Aqueous fraction 3

AQ3 is dedicated to carboxylic acids according LLE protocol. This fraction represents 3% (w/w) of our initial biomass sample. Despite the low mass percentage of this fraction, it was necessary to carry out the fractionation protocol to isolate minor products and to identify species that were not detected in presence of the major fraction (AQ1) due to matrix effects. On the van Krevelen diagram, components in this fraction are mainly centered around H/C of 1.0 and O/C of 0.4, meaning that they are composed of aromatic rings (Figure 8 (1)). Interaction with the apolar stationary phase (retention times between 10 and 20 min) are in concordance with the hypothesis of aromatic acids. These aromatic carboxylic acids can be separated between mono, di, tri and tetra-aromatic compounds by plotting the DBE value as a function of molecular mass (Figure 8 (2)). Few aliphatic carboxylic acids were also detected in this fraction. At the opposite of aromatic acids, they have a high H/C ratio and a low DBE. Moreover, furans such as 5-HMF were observed. Indeed because of their low log P, they were distributed in several LLE fractions. During fragmentation process in negative-ion mode ESI, principal losses were CO_2 coming from fragmentation of carboxylic acid function and $^{\bullet}CH_3$ radical from fragmentation of methoxy group attached to an aromatic ring (Figure 8 (3)). Among the 17 compounds submitted to fragmentation, 10 of them lost the carboxylic acid moiety as the major fragment during the first CID event. When methoxy functions are present on an aromatic ring, methyl radical losses are in competition with CO_2 loss. Radical methyl groups were removed one by one during each fragmentation stage. Therefore experimental data enable us to identify coumaric, syringic and ferulic acids as mono-aromatic compounds. For coumaric acid, exclusive loss of CO_2 was observed, whereas for ferulic and syringic acids additional losses of one and two methyl radicals were exhibited respectively. For few dimeric components, water was the most abundant loss in MS^2 experiment although CO_2 loss exhibited a lower intensity. This loss was associated to CH_2O loss which might indicate β-O-4 linkage [17]. However, fragmentation by CID shown limitations for untargeted structural identifications of di, tri and tetra-aromatic compounds. To propose confident structures, fragmentation should be done on target components in order to obtain the complete fragmentation tree (Table S2).

In addition, this fraction was highly connected with LCC measured in AQ1. Indeed, during LCC fragmentations, molecular formulae of the phenolic part free of the carbohydrate moiety were measured and these formulae were also measured in AQ3 in many cases. For example, for the peak d, phenolic part was $C_{20}H_{20}O_7$ (Figure 6); this fragment was also measured as deprotonated compounds in AQ3. UV spectrum of peak d was also similar to the peak corresponding to $C_{20}H_{20}O_7$ in AQ3 (not shown). This observation suggested that LCC are mainly composed of phenolic acid associated with carbohydrates. Since carbohydrates do not absorb in UV, comparison of LCC and phenolic acid UV spectrum should be a supplementary clue for identification in case of the complex was not fragmented.

3.2.3.Organic fraction 2

ORG2 represents a minor part (2% w/w) of our initial sample and is related to compounds that are neutral at pH >12. Mass between 90 and 600 *m/z* were detected, mainly in positive-ion mode ESI, which suggests carbonyl groups. On the van Krevelen diagram, components are distributed between 0.5 and 2.0 for H/C and between 0.1 and 0.5 for O/C but an important part is centered around H/C of 1.5 and O/C of 0.2 (Figure S7 (1)). The two most intense peaks observed on UV chromatogram (Figure 3) were identified as 5-HMF (t_R=5.2 min) and furfural (t_R=5.5 min). On the contrary to AQ3, DBE values were comprised between 2 and 8 for 95% of compounds (Figure S7 (2)). Low DBE associated to relative high H/C ratio was a first clue to attest the presence of aliphatic chains. This hypothesis was confirmed with fragmentation experiments and detection of C_4H_8 neutral loss which may come from the fragmentation of aliphatic chains from lipids (Figure S7 (3)). Losses of water were abundant and might come from hydroxyl groups. Moreover, CO losses occurring during the first fragmentation stage (MS^2) may indicate presence of carbonyl functions such as ketone and/or aldehyde groups [33]. Another abundant loss was C_2H_2O (ketene) which is known to be characteristic of acetyl group.

3.2.4. Organic fraction 3

The last fraction (ORG3) represents 4% (w/w) of our initial sample and is expected to be composed of phenols with a pKa varying between 7 and 12. Similarly to AQ3, the van Krevelen diagram of ORG3 is centered around H/C of 1 and O/C of 0.4, corresponding to an aromatic area (Figure S8 (1)). Graph representing DBE as a function of molecular mass also proves the phenolic nature of components in this fraction (Figure S8 (2)). Fragmentations of monoaromatic compounds allowed structural identification with a first methyl radical loss from methoxy group and then a combination of CO and CO_2 losses. Marcum *et al.* already reported a study on fragmentation of small molecules related to lignin which demonstrated that CO losses often occur after the loss of methyl radical during the fragmentation of methoxy group [19]. Hydroxybenzaldehyde, vanillin, syringaldehyde, acetosyringone and coniferaldehyde were identified as monomeric compounds in ORG3 (Table S3). For di, tri and tetra-aromatic compounds, fragmentation patterns depend on the linkage between units. For example, ion at *m/z* 273.0765 ($C_{15}H_{14}O_5$) was detected at two retention times (12.6 and 17.3 min) and had two different fragmentation patterns. The first one implies ($CO+H_2O$) fragment as the main MS^2 neutral loss, whereas the second one shows a methyl radical loss. In both cases, the di-aromatic compounds did not fragment in two mono-aromatic units indicating carbon-carbon bond between lignin units such as 5-5, β-β or β-5 links. At the opposite, tetra-aromatic ion at *m/z* 683.2144 ($C_{38}H_{36}O_{12}$) was fragmented in two di-aromatic fragments: $C_{18}H_{17}O_6$ and $C_{20}H_{19}O_6$. Then, $C_{18}H_{17}O_6$ ion fragment lose 3 methyl radical fragments during MS^3, MS^4 and MS^5 stages, which may correspond to the fragmentation of three methoxy group but also suppose carbon-carbon bound in the di-aromatic ion fragment.

In many case, ion fragment corresponding to guaiacol ($C_7H_8O_2$) was measured during the fragmentation of tri and tretra-aromatic compounds (Table S3).

4. Conclusion

In this work, a multi-technique analytical approach combining liquid-liquid extractions, LC/HRMS and multi-stage fragmentations, was developed to achieve an exhaustive characterization of biomass sample coming from the industrial pretreatment of wheat straw. First, this approach consisted in organizing the biomass sample in four fractions according to chemical family with a reliable liquid-liquid extraction protocol. Then, HPLC/MS^n experiments were performed using high-resolution LTQ-

FT-ICR/MS on the fractions. The combination of LLE protocol and LC separation highly limited the risk of co-elutions and thus ion suppression phenomenon which may occur in atmospheric pressure ionization source. In this way, the composition of the sample should be more representative than using direct introduction or LC/HRMS analysis. Also, isomers could be distinguish which is not possible by direct introduction mass spectrometry. High resolution mass spectrometry allowed to measure elemental compositions with a list of chemical formulae for each fractions. In order to go deeper in the understanding of the biomass composition, multi-stage fragmentations up to MS^7 were conducted. Typical fragments obtained were in accordance with fraction chemical families. By combining LLE fractions properties, retention time, HPLC-UV spectrum, molecular formula as well as structural information delivered by fragmentation experiments, structures were proposed for compounds up to 600 g/mol. To the best of our knowledge, for the first time carbohydrates with carboxylic acid function and heavy lignin-carbohydrate complexes were elucidated using $HPLC/MS^n$ method. The ability of this analytical approach to describe the main chemical families of a lignocellulosic biomass sample is a promising tool and should be applied on several biomass samples in order to progress in the comprehension of relationships between the products composition and reactivity. Moreover, this approach may also be useful during the optimization of pretreatment processes in order to observe the impact of the different process parameters (temperature, time reaction, etc.) on the chemical composition of liquid samples.

5. References

[1] J. Mohtasham, Review Article-Renewable Energies, Energy Procedia 74 (2015) 1289–1297. https://doi.org/10.1016/j.egypro.2015.07.774.

[2] I. Dincer, C. Acar, A review on clean energy solutions for better sustainability, Int. J. Energy Res. 39 (5) (2015) 585–606. https://doi.org/10.1002/er.3329.

[3] N. Abas, A. Kalair, N. Khan, Review of fossil fuels and future energy technologies, Futures 69 (2015) 31–49. https://doi.org/10.1016/j.futures.2015.03.003.

[4] F.H. Isikgor, C.R. Becer, Lignocellulosic biomass: A sustainable platform for the production of bio-based chemicals and polymers, Polym. Chem. 6 (25) (2015) 4497–4559. https://doi.org/10.1039/C5PY00263J.

[5] A. Brandt, J. Gräsvik, J.P. Hallett, T. Welton, Deconstruction of lignocellulosic biomass with ionic liquids, Green Chem. 15 (3) (2013) 550. https://doi.org/10.1039/c2gc36364j.

[6] P.F.H. Harmsen, W.J.J. Huijgen, L.M. Bermudez Lopez, R.R.C. Bakker, Literature review of physical and chemical pretreatment processes for lignocellulosic biomass, 2010 (accessed 3 November 2017).

[7] P. Gallezot, Conversion of biomass to selected chemical products, Chem. Soc. Rev. 41 (4) (2012) 1538–1558. https://doi.org/10.1039/c1cs15147a.

[8] P. Kumar, D.M. Barrett, M.J. Delwiche, P. Stroeve, Methods for Pretreatment of Lignocellulosic Biomass for Efficient Hydrolysis and Biofuel Production, Ind. Eng. Chem. Res. 48 (8) (2009) 3713–3729. https://doi.org/10.1021/ie801542g.

[9] L.J. Jönsson, B. Alriksson, N.-O. Nilvebrant, Bioconversion of lignocellulose: inhibitors and detoxification, Biotechnol. Biofuels 6 (1) (2013) 16. https://doi.org/10.1186/1754-6834-6-16.

[10] H.B. Klinke, A.B. Thomsen, B.K. Ahring, Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass, Appl. Microbiol. Biotechnol. 66 (1) (2004) 10–26. https://doi.org/10.1007/s00253-004-1642-2.

[11] C. Luo, D.L. Brink, H.W. Blanch, Identification of potential fermentation inhibitors in conversion of hybrid poplar hydrolyzate to ethanol, Biomass and Bioenergy 22 (2) (2002) 125–138. https://doi.org/10.1016/S0961-9534(01)00061-7.
[12] B. Du, L.N. Sharma, C. Becker, S.-F. Chen, R.A. Mowery, G.P. van Walsum, C.K. Chambliss, Effect of varying feedstock-pretreatment chemistry combinations on the formation and accumulation of potentially inhibitory degradation products in biomass hydrolysates, Biotechnol. Bioeng. 107 (3) (2010) 430–440. https://doi.org/10.1002/bit.22829.
[13] Y. Liu, Q. Shi, Y. Zhang, Y. He, K.H. Chung, S. Zhao, C. Xu, Characterization of Red Pine Pyrolysis Bio-oil by Gas Chromatography–Mass Spectrometry and Negative-Ion Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, Energy Fuels 26 (7) (2012) 4532–4539. https://doi.org/10.1021/ef300501t.
[14] J.H. Marsman, J. Wildschut, F. Mahfud, H.J. Heeres, Identification of components in fast pyrolysis oil and upgraded products by comprehensive two-dimensional gas chromatography and flame ionisation detection, J. Chromatogr. A 1150 (1-2) (2007) 21–27. https://doi.org/10.1016/j.chroma.2006.11.047.
[15] T. Sfetsas, C. Michailof, A. Lappas, Q. Li, B. Kneale, Qualitative and quantitative analysis of pyrolysis oil by gas chromatography with flame ionization detection and comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry, J. Chromatogr. A 1218 (21) (2011) 3317–3325. https://doi.org/10.1016/j.chroma.2010.10.034.
[16] M. Staš, J. Chudoba, D. Kubička, J. Blažek, M. Pospíšil, Petroleomic Characterization of Pyrolysis Bio-oils: A Review, Energy Fuels 31 (10) (2017) 10283–10299. https://doi.org/10.1021/acs.energyfuels.7b00826.
[17] H. Sheng, W. Tang, J. Gao, J.S. Riedeman, G. Li, T.M. Jarrell, M.R. Hurt, L. Yang, P. Murria, X. Ma, J.J. Nash, H.I. Kenttämaa, (-)ESI/CAD MSn Procedure for Sequencing Lignin Oligomers Based on a Study of Synthetic Model Compounds with β-O-4 and 5-5 Linkages, Anal. Chem. 89 (24) (2017) 13089–13096. https://doi.org/10.1021/acs.analchem.7b01911.
[18] B.C. Owen, L.J. Haupert, T.M. Jarrell, C.L. Marcum, T.H. Parsell, M.M. Abu-Omar, J.J. Bozell, S.K. Black, H.I. Kenttämaa, High-performance liquid chromatography/high-resolution multiple stage tandem mass spectrometry using negative-ion-mode hydroxide-doped electrospray ionization for the characterization of lignin degradation products, Anal. Chem. 84 (14) (2012) 6000–6007. https://doi.org/10.1021/ac300762y.
[19] C.L. Marcum, T.M. Jarrell, H. Zhu, B.C. Owen, L.J. Haupert, M. Easton, O. Hosseinaei, J. Bozell, J.J. Nash, H.I. Kenttämaa, A Fundamental Tandem Mass Spectrometry Study of the Collision-Activated Dissociation of Small Deprotonated Molecules Related to Lignin, ChemSusChem 9 (24) (2016) 3513–3526. https://doi.org/10.1002/cssc.201600678.
[20] E. Kiyota, P. Mazzafera, Sawaya, Alexandra C H F, Analysis of soluble lignin in sugarcane by ultrahigh performance liquid chromatography-tandem mass spectrometry with a do-it-yourself oligomer database, Anal. Chem. 84 (16) (2012) 7015–7020. https://doi.org/10.1021/ac301112y.
[21] J. Hertzog, V. Carré, Y. Le Brech, A. Dufour, F. Aubriet, Toward Controlled Ionization Conditions for ESI-FT-ICR-MS Analysis of Bio-Oils from Lignocellulosic Material, Energy Fuels 30 (7) (2016) 5729–5739. https://doi.org/10.1021/acs.energyfuels.6b00655.
[22] T.M. Jarrell, C.L. Marcum, H. Sheng, B.C. Owen, C.J. O'Lenick, H. Maraun, J.J. Bozell, H.I. Kenttämaa, Characterization of organosolv switchgrass lignin by using high performance liquid chromatography/high resolution tandem mass spectrometry using hydroxide-doped negative-

ion mode electrospray ionization, Green Chem 16 (5) (2014) 2713–2727. https://doi.org/10.1039/C3GC42355G.
[23] S.-F. Chen, R.A. Mowery, V.A. Castleberry, G.P. van Walsum, C.K. Chambliss, High-performance liquid chromatography method for simultaneous determination of aliphatic acid, aromatic acid and neutral degradation products in biomass pretreatment hydrolysates, J. Chromatogr. A 1104 (1-2) (2006) 54–61. https://doi.org/10.1016/j.chroma.2005.11.136.
[24] P.K. Kanaujia, D.V. Naik, D. Tripathi, R. Singh, M.K. Poddar, L.S.K. Konathala, Y.K. Sharma, Pyrolysis of Jatropha Curcas seed cake followed by optimization of liquid-liquid extraction procedure for the obtained bio-oil, Journal of Analytical and Applied Pyrolysis 118 (2016) 202–224. https://doi.org/10.1016/j.jaap.2016.02.005.
[25] Y. Wei, H. Lei, L. Wang, L. Zhu, X. Zhang, Y. Liu, S. Chen, B. Ahring, Liquid–Liquid Extraction of Biomass Pyrolysis Bio-oil, Energy Fuels 28 (2) (2014) 1207–1212. https://doi.org/10.1021/ef402490s.
[26] S.-F. Chen, High-Performance Liquid Chromatographic Methods for Quantitative Assessment of Degradation Products and Extractives in Pretreated Lignocellulose, 2007.
[27] A. Sluiter, B. Hames, R. Ruiz, C. Scarlata, J. Sluiter, and D. Templeton: NREL, Determination of Sugars, Byproducts, and Degradation Products in Liquid Fraction Process Samples: Laboratory Analytical Procedure (LAP); Issue Date: 12/08/2006.
[28] A. Sluiter, B. Hames, R. Ruiz, C. Scarlata, J. Sluiter, D. Templeton, and D. Crocker: NREL, Determination of Structural Carbohydrates and Lignin in Biomass: Laboratory Analytical Procedure (LAP) (Revised July 2011).
[29] C. Reymond, A. Le Masle, C. Colas, N. Charon, A rational strategy based on experimental designs to optimize parameters of a liquid chromatography-mass spectrometry analysis of complex matrices, Talanta 205 (2019) 120063. https://doi.org/10.1016/j.talanta.2019.06.063.
[30] T.-Q. Yuan, S.-N. Sun, F. Xu, R.-C. Sun, Characterization of lignin structures and lignin-carbohydrate complex (LCC) linkages by quantitative 13C and 2D HSQC NMR spectroscopy, J. Agric. Food Chem. 59 (19) (2011) 10604–10614. https://doi.org/10.1021/jf2031549.
[31] K.S. Boes, R.H. Narron, S. Park, N.R. Vinueza, Mass Spectrometry Exposes Undocumented Lignin-Carbohydrate Complexes in Biorefinery Pretreatment Stream, ACS Sustainable Chem. Eng. 6 (8) (2018) 10654–10659. https://doi.org/10.1021/acssuschemeng.8b01986.
[32] B. Domon, C.E. Costello, A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates, Glycoconjugate J 5 (4) (1988) 397–409. https://doi.org/10.1007/BF01049915.
[33] L.M. Amundson, V.A. Gallardo, N.R. Vinueza, B.C. Owen, J.N. Reece, S.C. Habicht, M. Fu, R.C. Shea, A.B. Mossman, H.I. Kenttämaa, Identification and Counting of Oxygen Functionalities and Alkyl Groups of Aromatic Analytes in Mixtures by Positive-Mode Atmospheric Pressure Chemical Ionization Tandem Mass Spectrometry Coupled with High-Performance Liquid Chromatography, Energy Fuels 26 (5) (2012) 2975–2989. https://doi.org/10.1021/ef2019098.

Figure captions

Table 1: Experimental conditions for ESI and APCI in positive and negative modes

Figure 1: A selective LLE fractionation for aqueous biomass samples

Figure 2: HPLC-UV chromatograms of the whole sample (black chromatogram) and LLE fractions obtained (AQ1: green, AQ3: purple, ORG2: blue, ORG3: red) at 254 nm

Figure 3: van Krevelen diagram using all detection modes (ESI+/- and APCI+/-) for each LLE fraction

Figure 4: van Krevelen diagram for AQ1. Colors were used to differentiate compounds eluted at the beginning of the run (with less than 12% of MeOH in the mobile phase in orange and green) and those eluted later (black and grey). Positive and negative ion-mode detections were also differentiate.

Figure 5: Base peak chromatogram of AQ1 in negative-ion mode ESI and associated van Krevelen diagram for CxHyOz compounds with red dots for the annotated peaks

Figure 6: Proposition of fragmentation scheme from MS^1 to MS^7 of compound (b) in negative-ion mode ESI and positive-ion mode ESI, compounds (a), (c) and (d) in negative-ion mode ESI for fraction AQ1. Relative intensity of each fragment is given in bracket

Figure 7: MS^2 mass spectra of peak α (red), β (blue) and γ (green) present in fraction AQ1 in negative-ion mode ESI

Figure 8: AQ3 van Krevelen diagram (1), DBE as a function of molecular mass (2) and losses in negative-ion mode ESI for MS^2 (red dots) until MS^7 (black outline circles) according to retention time (3)

Figure 1

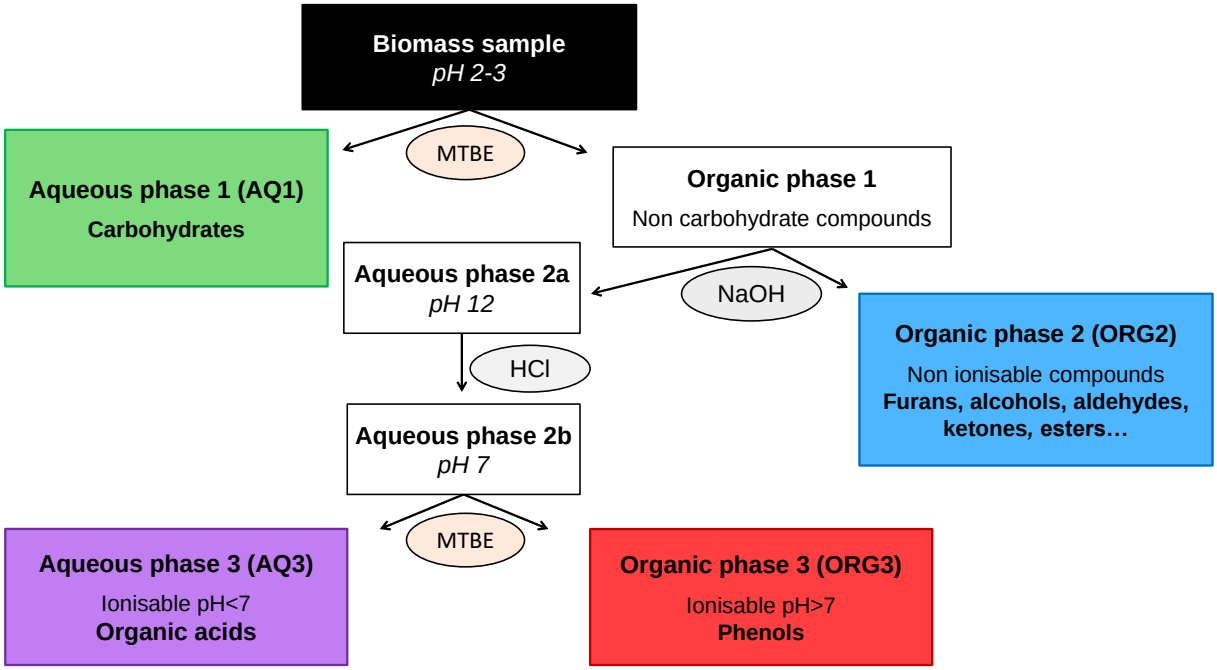

Biomass sample
pH 2-3
MTBE
Aqueous phase 1 (AQ1)
Carbohydrates
Organic phase 1
Non carbohydrate compounds
Aqueous phase 2a
pH 12
NaOH
Organic phase 2 (ORG2)
Non ionisable compounds
Furans, alcohols, aldehydes, ketones, esters…
HCl
Aqueous phase 2b
pH 7
MTBE
Aqueous phase 3 (AQ3)
Ionisable pH<7
Organic acids
Organic phase 3 (ORG3)
Ionisable pH>7
Phenols

Figure 2

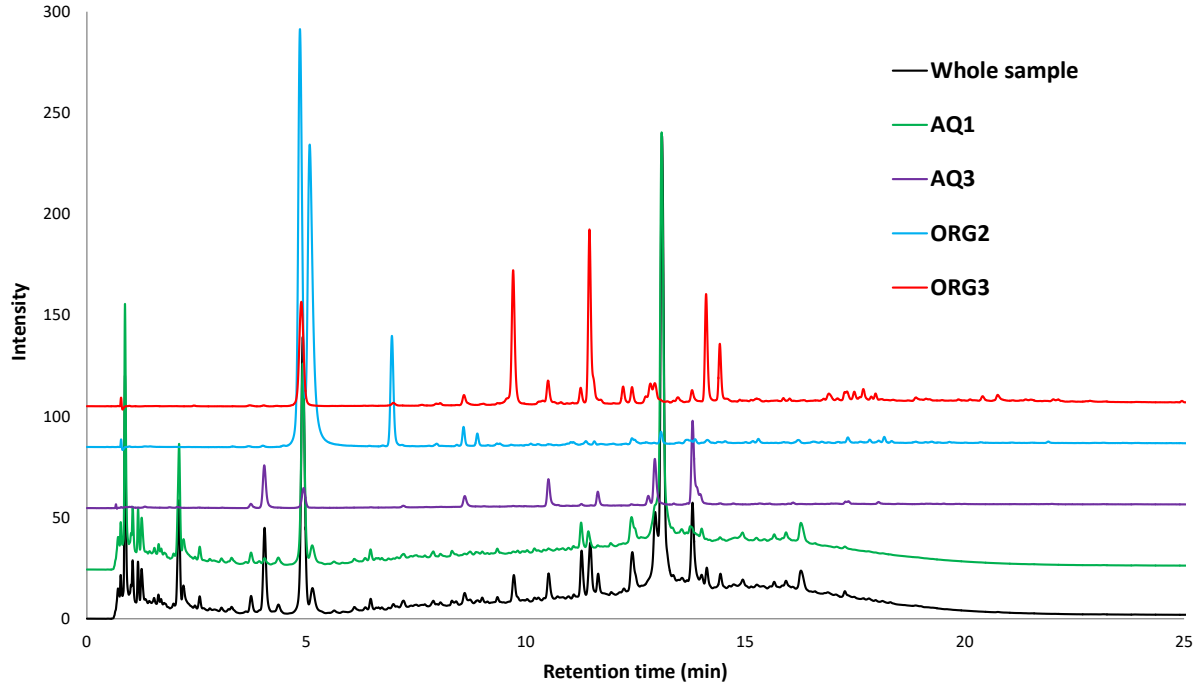

300
250
200
150
100
50
0
Intensity
0
5
10
15
20
25
Retention time (min)
Whole sample
AQ1
AQ3
ORG2
ORG3

Figure 3

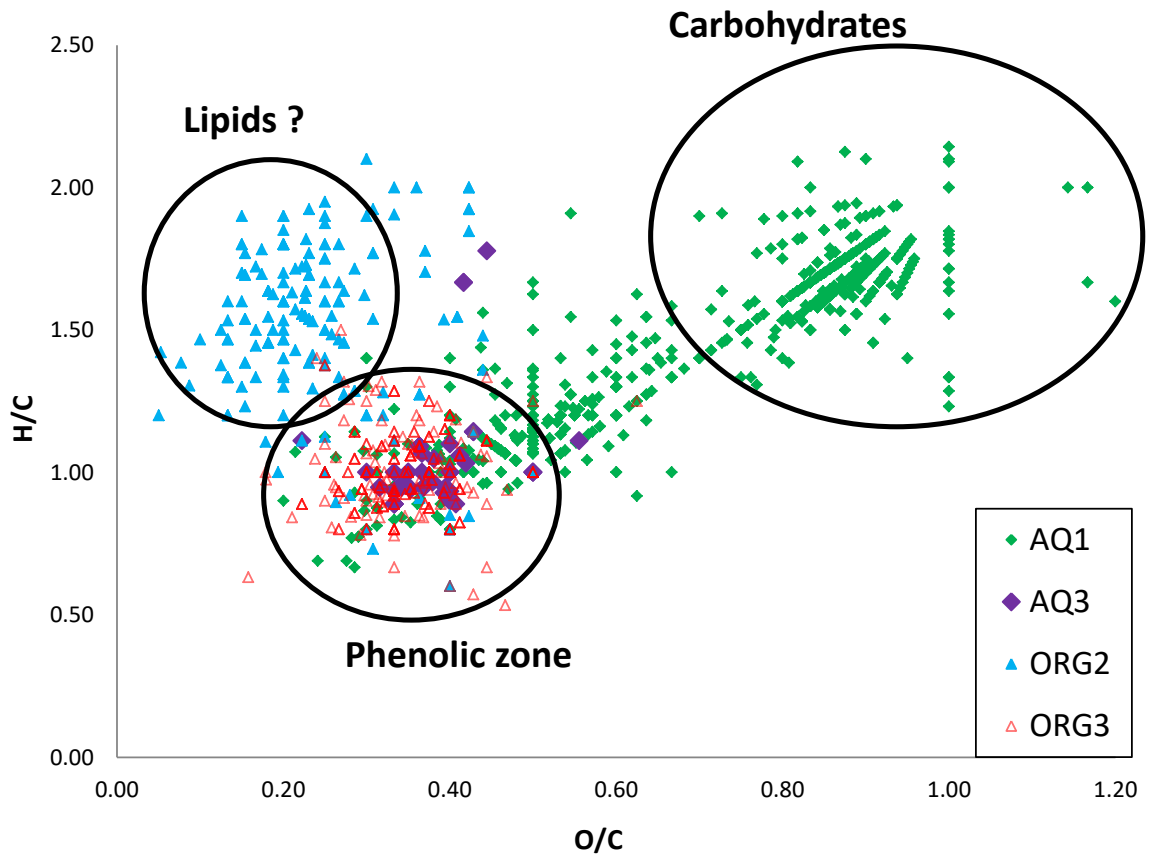

Carbohydrates
Lipids ?
Phenolic zone
AQ1
AQ3
ORG2
ORG3
H/C
O/C
2.50
2.00
1.50
1.00
0.50
0.00
0.00
0.20
0.40
0.60
0.80
1.00
1.20

Figure 4

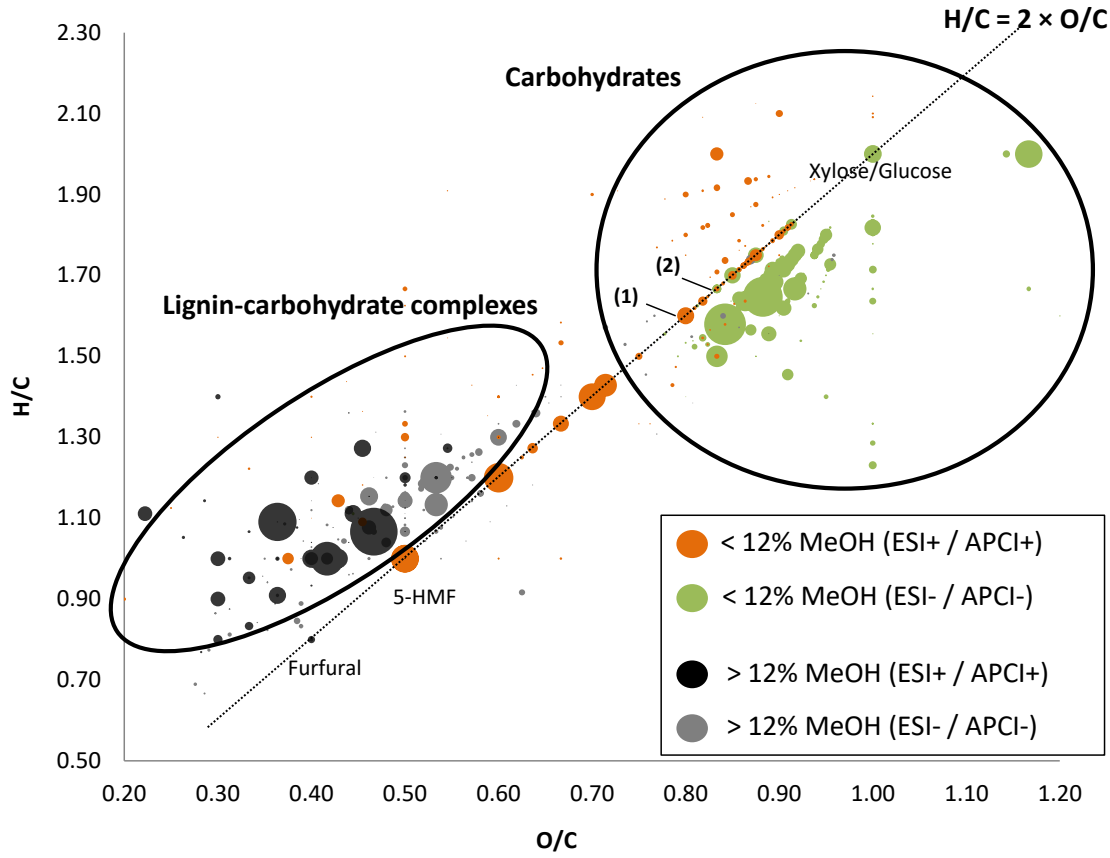

H/C = 2 × O/C
Carbohydrates
Xylose/Glucose
(2)
(1)
Lignin-carbohydrate complexes
5-HMF
Furfural
< 12% MeOH (ESI+ / APCI+)
< 12% MeOH (ESI- / APCI-)
> 12% MeOH (ESI+ / APCI+)
> 12% MeOH (ESI- / APCI-)
H/C
O/C
2.30
2.10
1.90
1.70
1.50
1.30
1.10
0.90
0.70
0.50
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00
1.10
1.20

Figure 5

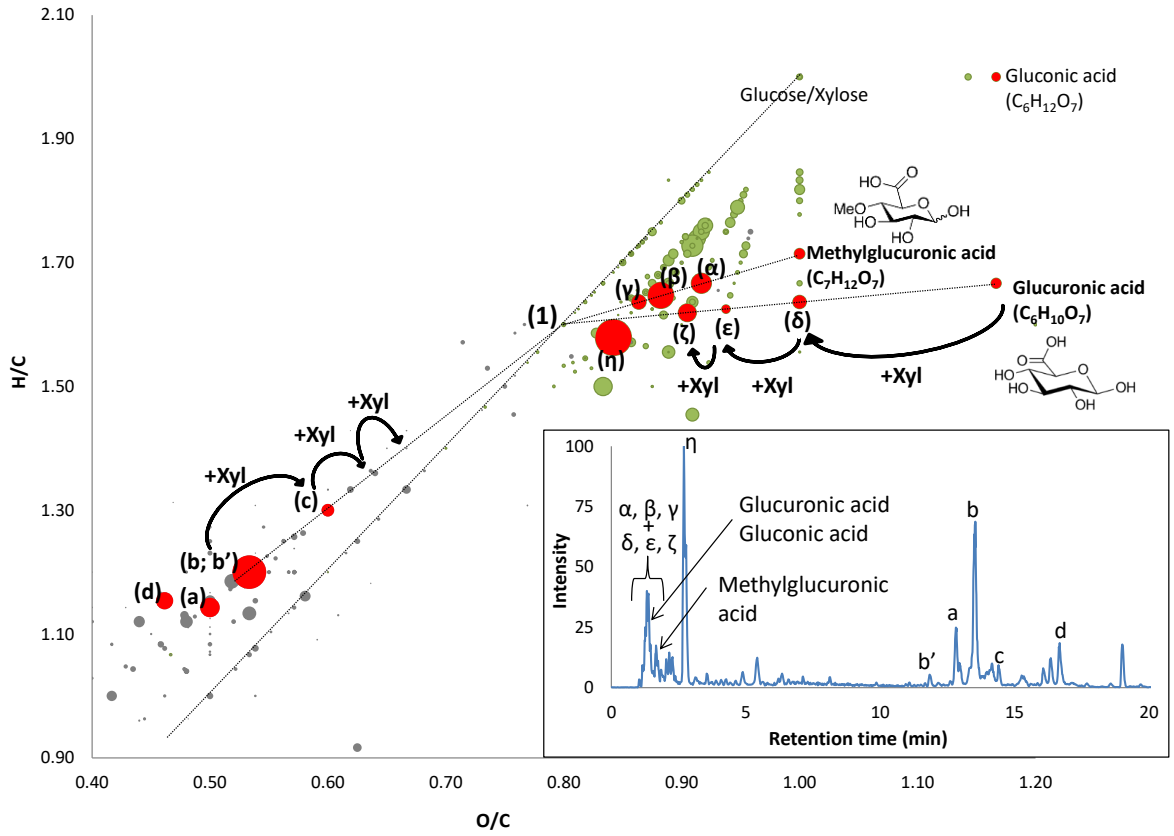

H/C
O/C
2.10
1.90
1.70
1.50
1.30
1.10
0.90
0.40
0.50
0.60
0.70
0.80
0.90
1.00
1.10
1.20
Glucose/Xylose
Gluconic acid
(C6H12O7)
Methylglucuronic acid
(C7H12O7)
Glucuronic acid
(C6H10O7)
(1)
(α)
(β)
(γ)
(δ)
(ε)
(ζ)
(η)
+Xyl
(a)
(b; b')
(c)
(d)
Intensity
Retention time (min)
α, β, γ
+
δ, ε, ζ
Glucuronic acid
Gluconic acid
Methylglucuronic
acid
a
b
b'
c
d
η

Figure 6

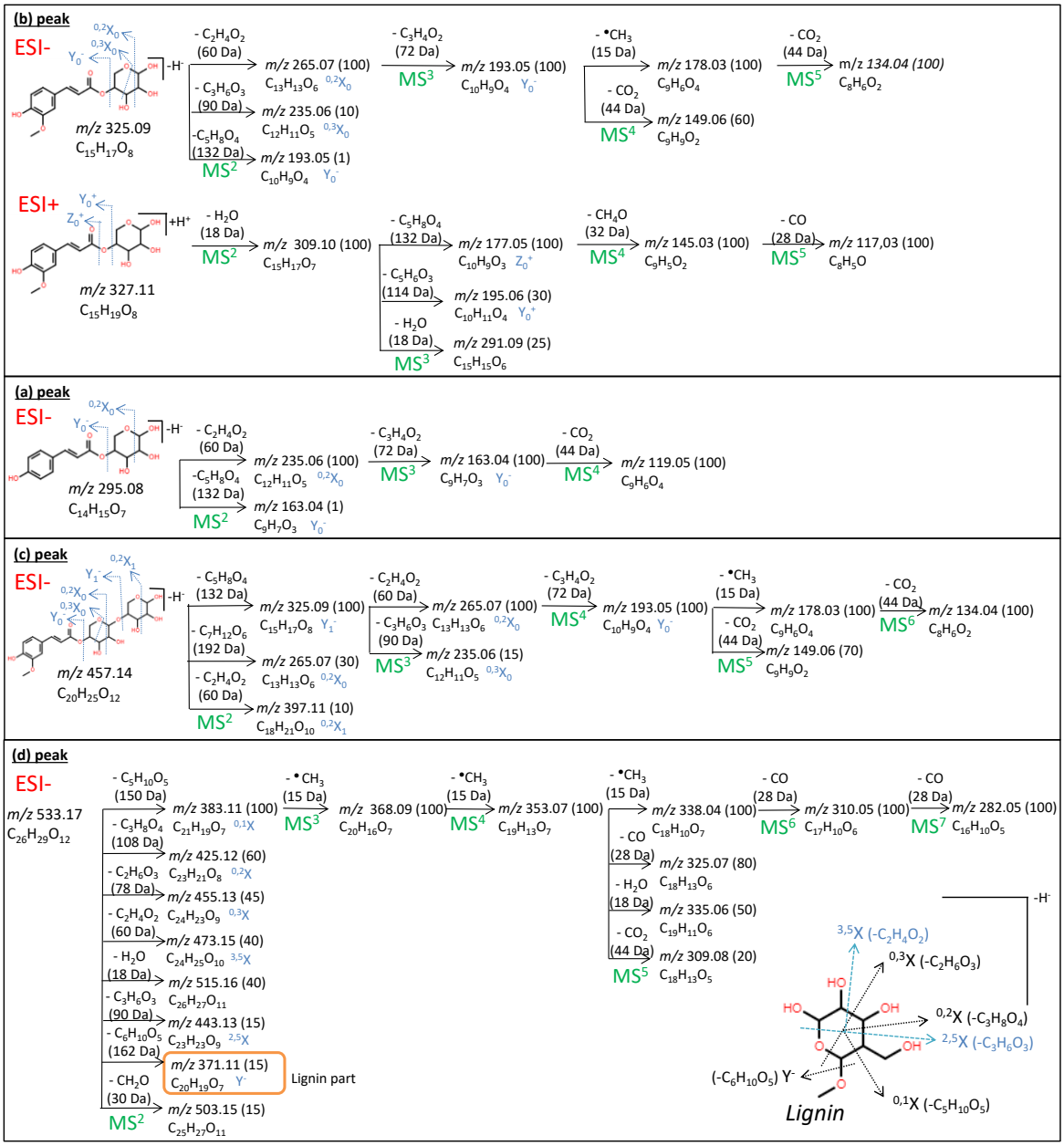

(b) peak
ESI-
m/z 325.09
C15H17O8
- C2H4O2 (60 Da)
m/z 265.07 (100)
C13H13O6
- C3H6O3 (90 Da)
m/z 235.06 (10)
C12H11O5
-C5H8O4 (132 Da)
m/z 193.05 (1)
C10H9O4
MS2
- C3H4O2 (72 Da)
m/z 193.05 (100)
C10H9O4
MS3
- •CH3 (15 Da)
m/z 178.03 (100)
C9H6O4
- CO2 (44 Da)
m/z 149.06 (60)
C9H9O2
MS4
- CO2 (44 Da)
m/z 134.04 (100)
C8H6O2
MS5
ESI+
m/z 327.11
C15H19O8
- H2O (18 Da)
m/z 309.10 (100)
C15H17O7
MS2
- C5H8O4 (132 Da)
m/z 177.05 (100)
C10H9O3
- C5H6O3 (114 Da)
m/z 195.06 (30)
C10H11O4
- H2O (18 Da)
m/z 291.09 (25)
C15H15O6
MS3
- CH4O (32 Da)
m/z 145.03 (100)
C9H5O2
MS4
- CO (28 Da)
m/z 117,03 (100)
C8H5O
MS5
(a) peak
ESI-
m/z 295.08
C14H15O7
- C2H4O2 (60 Da)
m/z 235.06 (100)
C12H11O5
-C5H8O4 (132 Da)
m/z 163.04 (1)
C9H7O3
MS2
- C3H4O2 (72 Da)
m/z 163.04 (100)
C9H7O3
MS3
- CO2 (44 Da)
m/z 119.05 (100)
C9H6O4
MS4
(c) peak
ESI-
m/z 457.14
C20H25O12
- C5H8O4 (132 Da)
m/z 325.09 (100)
C15H17O8
- C7H12O6 (192 Da)
m/z 265.07 (30)
C13H13O6
- C2H4O2 (60 Da)
m/z 397.11 (10)
C18H21O10
MS2
- C2H4O2 (60 Da)
m/z 265.07 (100)
C13H13O6
- C3H6O3 (90 Da)
m/z 235.06 (15)
C12H11O5
MS3
- C3H4O2 (72 Da)
m/z 193.05 (100)
C10H9O4
MS4
- •CH3 (15 Da)
m/z 178.03 (100)
C9H6O4
- CO2 (44 Da)
m/z 149.06 (70)
C9H9O2
MS5
- CO2 (44 Da)
m/z 134.04 (100)
C8H6O2
MS6
(d) peak
ESI-
m/z 533.17
C26H29O12
- C5H10O5 (150 Da)
m/z 383.11 (100)
C21H19O7
- C3H8O4 (108 Da)
m/z 425.12 (60)
C23H21O8
- C2H6O3 (78 Da)
m/z 455.13 (45)
C24H23O9
- C2H4O2 (60 Da)
m/z 473.15 (40)
C24H25O10
- H2O (18 Da)
m/z 515.16 (40)
C26H27O11
- C3H6O3 (90 Da)
m/z 443.13 (15)
C23H23O9
- C6H10O5 (162 Da)
m/z 371.11 (15)
C20H19O7
- CH2O (30 Da)
m/z 503.15 (15)
C25H27O11
MS2
Lignin part
- •CH3 (15 Da)
m/z 368.09 (100)
C20H16O7
MS3
- •CH3 (15 Da)
m/z 353.07 (100)
C19H13O7
MS4
- •CH3 (15 Da)
m/z 338.04 (100)
C18H10O7
- CO (28 Da)
m/z 325.07 (80)
C18H13O6
- H2O (18 Da)
m/z 335.06 (50)
C19H11O6
- CO2 (44 Da)
m/z 309.08 (20)
C18H13O5
MS5
- CO (28 Da)
m/z 310.05 (100)
C17H10O6
MS6
- CO (28 Da)
m/z 282.05 (100)
C16H10O5
MS7
-H-
3,5X (-C2H4O2)
0,3X (-C2H6O3)
0,2X (-C3H8O4)
2,5X (-C3H6O3)
0,1X (-C5H10O5)
(-C6H10O5) Y-
Lignin

Figure 7

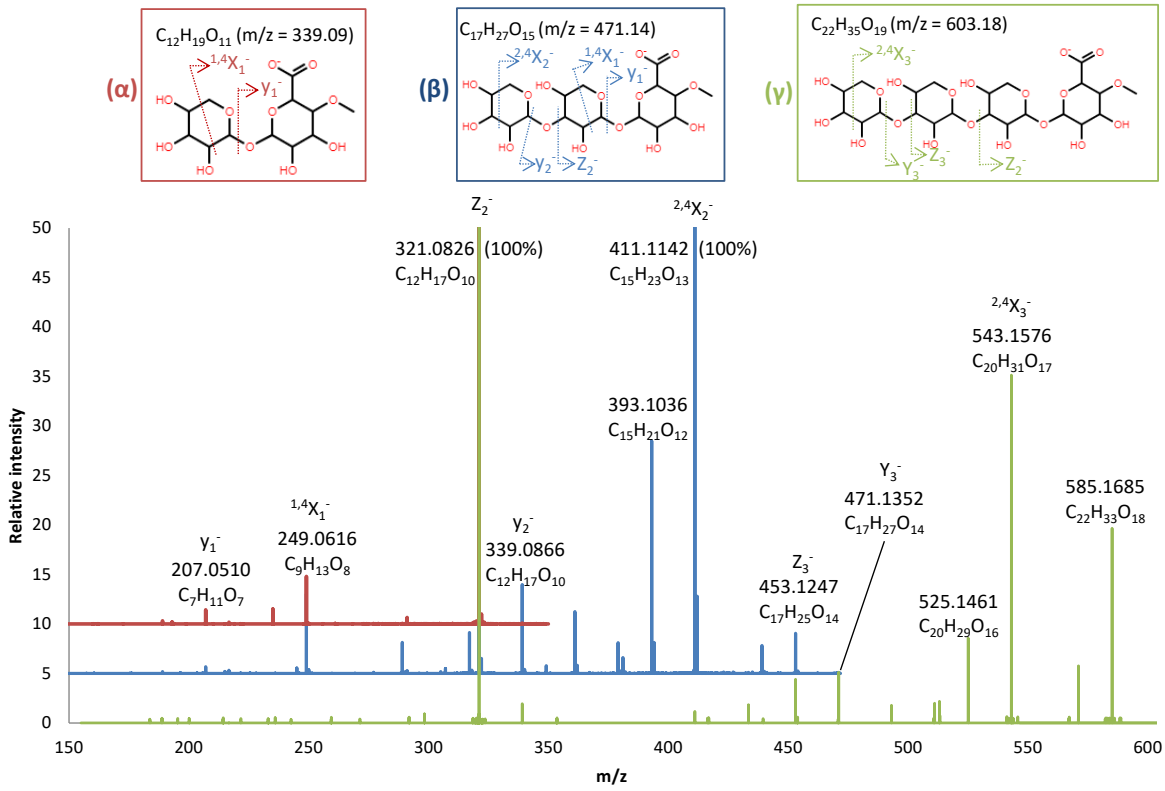
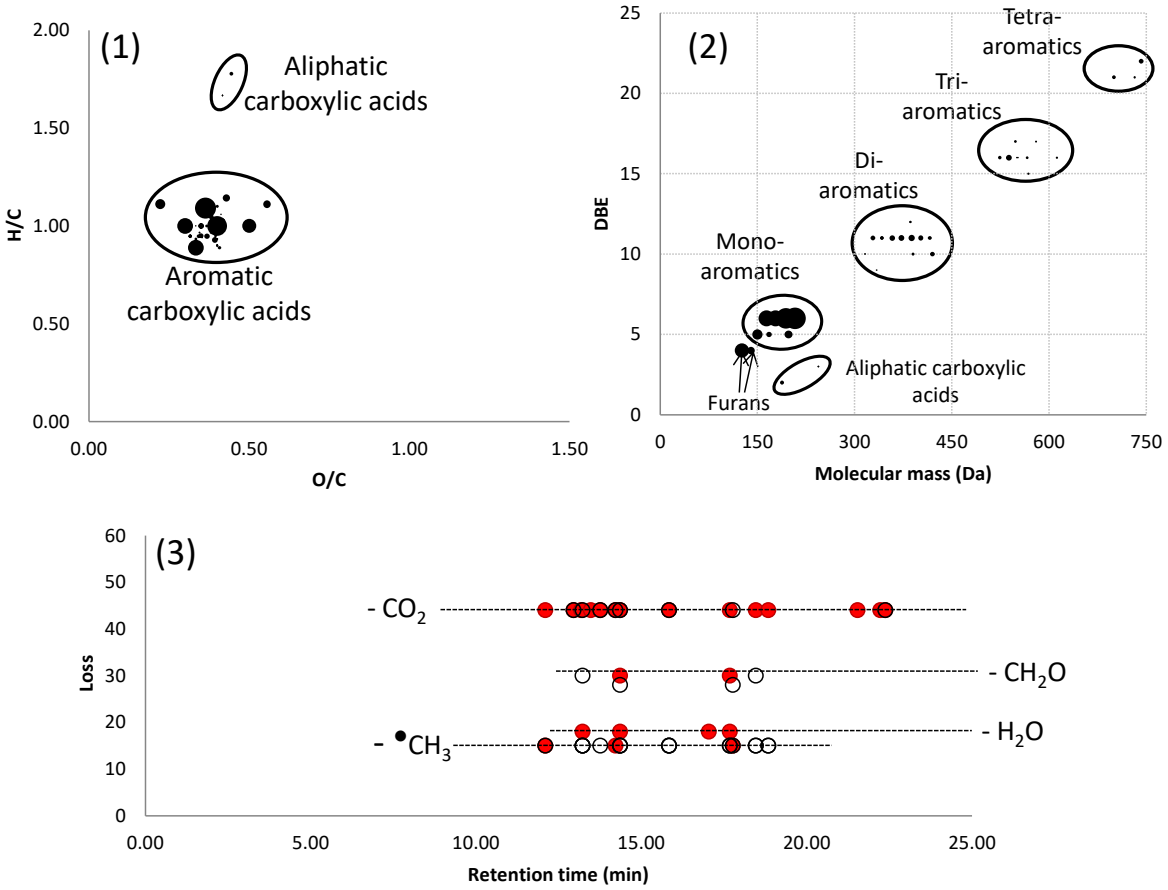

(α)
C12H19O11 (m/z = 339.09)
(β)
C17H27O15 (m/z = 471.14)
(γ)
C22H35O19 (m/z = 603.18)
Relative intensity
m/z
Z2-
321.0826 (100%)
C12H17O10
2,4X2-
411.1142 (100%)
C15H23O13
2,4X3-
543.1576
C20H31O17
393.1036
C15H21O12
Y3-
471.1352
C17H27O14
585.1685
C22H33O18
1,4X1-
249.0616
C9H13O8
y1-
207.0510
C7H11O7
y2-
339.0866
C12H17O10
Z3-
453.1247
C17H25O14
525.1461
C20H29O16

Figure 8


(1)
Aliphatic carboxylic acids
Aromatic carboxylic acids
H/C
O/C
2.00
1.50
1.00
0.50
0.00
0.00
0.50
1.00
1.50
(2)
Tetra-aromatics
Tri-aromatics
Di-aromatics
Mono-aromatics
Aliphatic carboxylic acids
Furans
DBE
Molecular mass (Da)
25
20
15
10
5
0
0
150
300
450
600
750
(3)
- CO2
- CH2O
- •CH3
- H2O
Loss
Retention time (min)
60
50
40
30
20
10
0
0.00
5.00
10.00
15.00
20.00
25.00