Ultra-high performance supercritical fluid chromatography hyphenated to atmospheric pressure chemical ionization high resolution mass spectrometry for the characterization of fast pyrolysis bio-oils

Julien Crepier, Agnès Le Masle, Nadège Charon, Florian Albrieux, Pascal Duchêne, Sabine Heinisch

Ultra-high performance supercritical fluid chromatography hyphenated to atmospheric pressure chemical ionization high resolution mass spectrometry for the characterization of fast pyrolysis bio-oils.

AUTHORS : Julien CREPIER[(a)], Agnès LE MASLE[(a)], Nadège CHARON[(a)], Florian ALBRIEUX[(a)], Pascal DUCHENE[(a)], Sabine HEINISCH*[(b)]

[a] *IFP Energies nouvelles, Rond-point de l'échangeur de Solaize, BP 3, 69360 Solaize, France*

[b] *Université de Lyon, Institut des Sciences Analytiques, UMR 5280, CNRS, ENS Lyon, 5 rue de la Doua, 69100 Villeurbanne, France*

CORRESPONDENCE :

Sabine HEINISCH
E-mail : sabine.heinisch@univ-lyon1.fr
Phone.: +33 4 37 42 35 51

Agnès LE MASLE
E-mail : agnes.le-masle@ifpen.fr
Phone.: +33 4 37 70 23 91

ABSTRACT

Extensive characterization of complex mixtures requires the combination of powerful analytical techniques. A Supercritical Fluid Chromatography (SFC) method was previously developed, for the specific case of fast pyrolysis bio oils, as an alternative to gas chromatography (GC and GCXGC) or liquid chromatography (LC and LCxLC), both separation methods being generally used prior to mass spectrometry (MS) for the characterization of such complex matrices. In this study we investigated the potential of SFC hyphenated to high resolution mass spectrometry (SFC-HRMS) for this characterization using Negative ion Atmospheric Pressure Chemical ionization ((-)APCI) for the ionization source. The interface between SFC and (-)APCI/HRMS was optimized from a mix of model compounds with the objective of maximizing the signal to noise ratio. The main studied parameters included both make-up flow-rate and make-up composition. A methodology for the treatment of

APCI/HRMS data is proposed. This latter allowed for the identification of molecular formulae. Both SFC-APCI/HRMS method and data processing method were applied to a mixture of 36 model compounds, first analyzed alone and then spiked in a bio-oil. In both cases, 19 compounds could be detected. Among them 9 could be detected in a fast pyrolysis bio-oil by targeted analysis. The whole procedure was applied to the characterization of a bio-oil using helpful representations such as mass-plots, van Krevelen diagrams and heteroatom class distributions. Finally the results were compared with those obtained with a Fourier Transform ion-cyclotron resonance mass spectrometer (FT-ICR/MS).

KEY WORDS

Ultra-High Performance Supercritical Fluid Chromatography; High Resolution Mass Spectrometry; APCI source; Biomass fast pyrolysis; Bio-oil ; Complex samples

1. Introduction

Because of the necessity to develop new sources of energy for the future, the production of second generation biofuels from lignocellulosic biomass seems to be a promising option, implying different ways of conversion [1]. One of them (fast pyrolysis) consists in liquefying biomass by thermochemical process operating in the range of 400 to 450°C. This process results in bio-oils, very rich in oxygen compounds, corrosive and thermally unstable. For further uses as biofuels or bio-based products, upgrading is necessary which can be only achieved if a detailed characterization is available. Recent publications present a comprehensive overview of current analytical techniques used to characterize pyrolysis bio-oils [2–4]. It is pointed out in both papers that high resolution mass spectrometry (HRMS) has become the primary method for the analytical characterization of bio-oils, considering its potential to determine both the molecular weights and the elemental compositions of thousands of bio-oil compounds [3]. Electrospray ionization (ESI) and Atmospheric Chemical Ionization (APCI) are commonly applied ionization techniques, mostly operated in negative ionization mode. According to Stas et al. [3], a distinct advantage of negative-ion APCI is that it can detect some more unsaturated, less polar bio-oil compounds with higher carbon numbers and m/z range not detectable by negative-ion ESI. In spite of its impressive analytical power, two key issues arise from the use of HRMS as single analytical technique. Those include (i) the risk of matrix effects reducing the ionization yield and (ii) the impossibility of

differentiating the very large number of positional and structural isomers present in bio-oils. However both issues may be theoretically overcome if an appropriate separation technique is hyphenated to HRMS.

Compound identification by gas chromatography hyphenated with mass spectrometry (GC-MS) and quantification using gas chromatography and flame ionization detection (GC-FID) are commonly carried out [5]. Thanks to its high resolution power, GC and overall GCxGC make a valuable contribution to the detailed characterization of complex matrices such as bio-oils. Nearly 300 compounds could be identified by GC-MS or GCxGC-MS in fast pyrolysis bio-oils, providing a wide range of chemical families including aldehydes, ketones, aromatic esters, carboxylic acids, alcohols, carbohydrates, furans, pyrans, phenols, benzenediols, methoxyphenols, dimethoxyphenols [6]. However, without prior derivatization step, some problems may occur with molecular structures higher than around 200 g/mol including (i) low separation power with the presence of numerous coelutions even in GCxGC [4], (ii) very high retention for heavy compounds and (iii) thermal instability (e.g. for carbohydrates) leading to compound degradation in the injection unit. In addition, there may be some identification issues from usual data bases, in particular with polyfunctional and oxygenated compounds having a high number of carbon atoms. As a result, in spite of the high potential of GCxGC, alternative separation techniques are strongly required in order to provide a more comprehensive view on bio-oil composition.

Two-dimensional reversed phase liquid chromatography (RPLC) techniques were applied to the analysis of the aqueous fraction of a bio-oil. It was found that RPLC x RPLC had the potential of resolving up to 2000 peaks [7], highlighting the potential of this technique for the comprehensive analysis of the aqueous phase of the bio-oil. Both photo diode array (PDA) and MS detection were latter coupled to RPLC x RPLC and a more detailed analysis could be obtained [8]. Finally, an orthogonal separation system was also recently proposed involving both RPLC and, for the first time, SFC (RPLC x SFC) [9]. However in spite of promising results on the aqueous fraction, neither LC, nor SFC techniques were ever been applied to the characterization of the whole bio-oil sample.

In this context we guessed that SFC hyphenated to HRMS could be a more versatile analytical technique, able to provide more comprehensive information on bio-oil composition. In a previous work [10], a SFC-UV method was developed with a view to later analyzing the whole sample by SFC-HRMS. The optimization of the separation parameters was directly performed

on bio-oil sample in order to take into account the complexity of such a sample at the earliest stage of method development [10]. Because of CO_2 decompression at the outlet of SFC device, only atmospheric ionization sources such as ESI or APCI can be hyphenated to SFC. The use of SFC-ESI/MS was often reported in different application fields with simple quad or high resolution mass spectrometers [11–13]. The APCI source has been rarely used in SFC/MS but recently proposed for the analysis of natural non-polar compounds [14].

Our choice for the APCI source was directed by the presence of components with a very large variety of chemical and physical properties (polarity, molecular weight, chemical functionality, m/z range etc…). The selection of suitable interface parameters was based on an optimization procedure, presented in this study for the specific case of bio-oils. The large amount of data generated in SFC-HRMS for complex sample analysis makes the use of specific software necessary, especially for non-targeted analysis as for bio-oils. We therefore developed a home-made software for data processing. Its key features are presented here. The relevance of the whole approach is highlighted with a mixture of model compounds, analyzed alone and spiked in complex bio-oil matrix. The obtained results regarding bio-oil composition are discussed with the support of usual representations including mass-maps, van Krevelen diagrams and heteroatom class distribution. Finally, these results are compared to those obtained with a Fourier Transform ion-cyclotron resonance mass spectrometer (FT-ICR/MS) which is known to be the most powerful in terms of mass resolving power.

2. Materials and methods

2.1. Chemicals and sample preparation

Solvents (acetonitrile, methanol, water) were MS grade from Sigma Aldrich (Steinheim, Germany). Carbon dioxide SFC grade (99.97%) was purchased from Air Liquide (B50 bottle under pressure). Tetrahydrofuran (THF) was purchased from VWR (Fontenay sous bois, France).

36 model compounds were purchased from Sigma Aldrich (Steinheim, Germany). Their names and structures are listed in Table S1 of the supplementary information. The model mix was obtained by dissolving each compound in THF (200 mg/kg). The fast pyrolysis bio-oil was obtained from conifer sawdust, provided by IFP Energies nouvelles. It was diluted in THF (1/5 w/w) before analysis. The diluted bio oil was spiked with model compounds (200 mg/kg each).

2.2. UHPSFC-UV instrument and column

All experiments were performed on an Acquity UPC² instrument (Waters, Milford, MA, USA). Key parameters (stationary and mobile phases, back pressure, column temperature and gradient conditions) were optimized according to a procedure developed in a previous work and based on the maximization of peak capacity [10]. The mobile phase flow-rate was 1.4 mL/min. The organic solvent modifier was a mix of acetonitrile and water (98/2 v/v). The oven temperature was set at 30 °C. Back Pressure Regulator (BPR) was set at 150 bar. The injection volume was 1 µl. The column used was an Acquity BEH-2EP (100 x 3mm, 1.7µm). The mobile phase varied from 1% to 40 % of organic solvent modifier in 14 minutes. The injector needle was washed with 600 µL of methanol after each injection. The column outlet was connected to a photo-diode array detector (PDA) equipped with a 8µL high pressure UV cell (400 bars) with a path length of 10 mm. The detection wavelengths varied between 210 and 400 nm with a resolution of 1.2 nm. The sampling rate was set at 40 Hz. The instrument control was performed by Empower 3 software (Waters).

2.3. HRMS instrument

Mass spectra were obtained with an Ion Trap –Time of Flight (IT-ToF) instrument (Shimadzu, Kyoto, Japan) equipped with an atmospheric pressure chemical ionization (APCI) source operated in negative mode. The resolution was 9385 for m/z=520.9095. Mass error was 5 ppm with internal calibration and 20 ppm with external calibration using sodium formate clusters to enlarge the range of calibration from 45 to 928 Th. MS parameters were optimized in order to favor the detection of pseudo-molecular ions. Mass range was between 80 and 800 uma; accumulation time was set at 30 ms; nebulizing gas flow was 0.5L/min; drying gas pressure was 100 kPa, both APCI and CDL temperatures were set at 250 °C while the heat block temperature at 280 °C. . The optimization of the interface between SFC and MS is presented in the Result section.

2.4. MS data processing

The APCI source mode was selected for this study. Corresponding MS data represent a large amount of information and therefore require suitable data processing to achieve the identification of a maximum of compounds. Starting from raw data, the obtained

chromatograms with MS (total ion current) or ultra violet (UV) detection were not sufficient to get relevant information. Data were therefore represented using a 2D-colour plot (mass-map) with information on retention time (x-axis), mass over charge m/z ratio (y-axis) and intensity (color scale). Peak intensity was described by a logarithmical color gradient. However it is important to note that peak intensities should not be directly compared since ionization yields strongly depend on compound chemical structures. For each mass-map spot, there may be numerous possible structures. As a result HRMS data were processed with an in-house software (so-called SFC/MS software in the rest of the study), in order to get accurate mass measurement and hence a set of several formulae for each mass-map spot. This in-house software was developed with the objective of (i) drawing and comparing mass-maps, (ii) being as universal as possible and (iii) maintaining the whole control regarding further identification procedure. The file format is based on the widely used mzXML extension [15], allowing the use of a large range of chromatographic (LC, LCxLC, SFC) and mass spectrometry (ToF, Orbitrap, FT-ICR/MS…) systems. For molecular formula calculation, the following parameters were used: elemental composition $^{12}C_{1\text{-}50}$, $^{1}H_{1\text{-}100}$, $^{16}O_{0\text{-}20}$, $^{14}N_{0\text{-}1}$ (^{13}C were also taken into account) ; mass error inferior or equal to $\pm$ 20 ppm ; H/C ratio = 0.2-3.1, O/C ratio = 0-1.8 ; N/C ratio = 0-1.3. In case of several possible molecular formulae, the most likely one was selected so that a unique elemental composition ($C_cH_hO_oN_n$) was assigned to a given m/z value. For each molecular formula, a score was calculated based on both mass error and isotopic data (equally) and the molecular formula having the highest score was selected. In addition, due to the fact that the elution of a given compound can take a few seconds, the corresponding data were lumped together which could avoid the risk of double identification. To validate the identification procedure, a mixture of 36 model compounds (Table S1 of the supplementary information) was analyzed alone and spiked in a bio-oil in order to point out possible matrix effects which could hinder the identification procedure. The concentration of each compound was 200 mg/kg in tetrahydrofuran (THF). The objective was to find, in both cases, the correct molecular formula for each model compound.

2.5. FT-ICR/MS instrument

The FT-ICR/MS instrument used for comparison with SFC-HRMS analysis was a Thermo Scientific LTQ FT Ultra (Bremen, Germany) composed of a linear ion trap and an ioncyclotron resonance cell in a 7 Tesla superconducting magnet. Sample was diluted in methanol (1:50 ;

v:v) prior to the injection by infusion mode (5µL/min) and ionized by ACPI mode. The number of microscans were set at 8 and 50 scans were accumulated. Data treatment was achieved with an in-house software called KendrickInside. For molecular formula calculation, the following parameters were used: elemental composition $^{12}C_{1-50}$, $^{1}H_{1-100}$,$^{16}O_{0-20}$, $^{14}N_{0-1}$ (^{13}C were also taken into account) ; mass error lower or equal to + 5 ppm.

3. RESULTS AND DISCUSSION

3.1. Optimization of SFC-(-)APCI-HRMS interface

With an APCI ionization source as used in this study, the mobile phase is under atmospheric pressure when entering the source, which results in CO_2 decompression in the introduction capillary. The resulting CO_2 evaporation makes the compounds concentrated in the liquid solvent (co-solvent). There may be therefore a risk of sample precipitation in the capillary, especially when the concentration of organic modifier in the mobile phase is low, for instance in starting gradient conditions. To prevent this from occurring, an additional pump can be used to deliver an additional amount of liquid solvent. Such device (so called Isocratic Solvent Manager – ISM), as proposed by Waters for our UHPSFC instrument, enables to add the CO_2-miscible make-up solvent (i.e. methanol) to the mobile phase via a T-union. Fig.1 shows the interface configuration (delimited by a frame) which also includes a second zero-dead volume T-union designed to split the flow coming from the first T-union in such a way that a fraction of the total flow is directed towards BPR device and the other one towards MS. Adding a protic solvent is also intended to improve the ionization yield by promoting charge exchange. However, with such interface configuration and the present APCI-IT-ToF-MS instrument, the MS signal was not stable enough, suggesting that the amount of solvent entering the APCI source was too low. A second make-up pump had therefore to be added along with a third zero-dead volume T-union to increase the flow-rate entering the APCI source as further discussed. The following discussion presents a theoretical approach to explain the limitation encountered with the commercially available interface and the procedure we used to optimize the second make-up conditions (flow-rate and solvent composition).

The solvent flow-rate entering the ionization source should be adapted according to the ionization source specificity. That requires that its value could be reliably predicted, depending

227 on SFC parameters and interface conditions. Theoretically, it is possible to predict the solvent
228 flow-rate, knowing the pressure drop in the tubing, the flow-rates delivered by both SFC pump
229 and ISM, the tubing geometry and the concentration of organic solvent in the mobile phase.
230 According to the Poiseuille-Hagen law, the pressure drop in the tubing is given by

231 $$\Delta P = \frac{128\,\eta}{\pi} \times R \times F \qquad (1)$$

Where F, is the flow-rate through the tubing, R, a term taking into account the tubing dimensions ($R = L/d^4$, L and d being the tubing length and diameter respectively) and η, the viscosity of the fluid (i.e. the fluid composed of CO2 and organic solvents coming from both SFC and make-up pumps).

The total flow-rate, F_T, prior to the second T-union is given by

$$F_T = F_{MS} + F_W \quad (2)$$

where, F_{MS} and F_W are the flow-rates after the splitter, towards MS and the waste. F_T is also given by the sum of flow-rates entering the first T-union:

$$F_T = F_{SFC} + F_{Pump\,1} \quad (3)$$

where, F_{SFC} and $F_{pump\,1}$ are the flow-rates delivered by SFC pump and Pump #1 respectively.

As shown in Fig 1, the section between the second T-union and the ionization source, is composed of two different tubes (blue and red in Fig.1) connected by a zero dead volume union. The red one diameter being significantly larger than the blue one (175µm vs 50µm), the pressure drop involved may not be considered in the calculations. Considering the same pressure drop in the two paths located after the splitter (second T-union), F_{MS} can be calculated according to

$$F_{MS} = \frac{\frac{\Delta P_{BPR} \times \pi}{128\,\eta} + R_W \times F_T}{(R_w + R_{MS})} \quad (4)$$

ΔP_{BPR} is the back pressure due to BPR. R_W and R_{MS} (Eq.1) relate to the capillaries located between the second T-union and BPR and between the second T-union and MS inlet respectively (the pressure drop in the tube located between the third T-union and MS inlet was low enough to be not taken into account).

The fraction, X_s, of solvent after the first T-union is given by

$$X_s = \frac{X_{s,SFC} \times F_{SFC} + F_{Pump\,1}}{F_T} \quad (5)$$

where $X_{s,SFC}$ is the volume fraction of solvent in SFC mobile phase. Finally, by combining Eqs. 4 and 5, the predicted solvent flow-rate entering the MS source can be calculated according to

$$F_{S,MS} = \frac{X_{s,SFC} \times F_{SFC} + F_{pump\,1}}{F_T} \times \frac{\frac{\Delta P_{BPR} \times \pi}{128\,\eta} + R_W \times F_T}{(R_w + R_{MS})} + F_{Pump\,2} \quad (6)$$

With $F_{Pump\,2}$ being the solvent flow-rate delivered by Pump #2 (Fig.1). Eq.6 can be considered as valid provided that (i) the fluid viscosity can be accurately assessed, (ii) the fluid viscosity is constant along the tube located between the second and the third T-union in spite of CO_2

decompression; (iii) the tubing dimensions are reliable and (iv) the solvent fraction, X_s, is maintained after flow-splitting. Flow-rate predictions can be inaccurate if one or more of these conditions are not fulfilled. We therefore compared some experimental measures to the predicted values given by Eq.6 in order to assess the validity of this theoretical approach. The measures were carried out without Pump 2 ($F_{Pump\ 2} = 0$) with acetonitrile (ACN) as co-solvent and methanol as make-up solvent. The make-up flow was varied from 200 to 1500 µL/min. SFC mobile phase conditions were those optimized in a previous study [10] and described in the experimental section. Two different co-solvent concentrations were considered, corresponding to initial and final gradient compositions (i.e. 1% ACN and 40% ACN). Flow-rate measurements were performed according to a method previously described [16]. Fluid viscosity values were estimated based on experimental correlations proposed by Ouyang [17], recently applied to SFC-MS with methanol as co-solvent [16] and adapted to binary mixtures of acetonitrile and methanol. As illustrated in Fig.2, showing the variation of solvent flow entering MS with Pump #1 flow, experimental and predicted values are in very good agreement for the two studied co-solvent compositions (i.e. 1% ACN and 40% ACN), thereby validating our theoretical approach. Fig.2 also shows that an increase in the make-up flow (containing MeOH) or in ACN concentration in SFC mobile phase, increases the solvent flow-rate entering the ionization source. However both curves tend towards the same constant value of nearly 300 µL/min which was found to be the threshold value to get a stable signal with the APCI source. The first option to increase the solvent flow could be to change the restriction capillary dimensions (red one in Fig.1). As theoretically shown in Fig. 3a for a mobile phase composition of 1% ACN, a reduction of the capillary length from 75 to 45 cm should lead to an increase in solvent flow from 230 µL/min to 380 µL/min, for a make-up flow of 500 µL/min. Meanwhile, the split ratio increases from 0.45 to 0.72 as illustrated in Fig. 3b. For a given capillary length, Fig.3 clearly shows that increasing the make-up flow slightly increases the solvent flow entering MS but strongly decreases the split ratio and hence the signal intensity in case of mass flow dependent detectors such as APCI-MS as also discussed elsewhere [18]. In summary, the first option could be the use of a restriction capillary with 45 cm length (instead of 75 cm proposed in the commercial interface) at a make-up flow of 500µL/min.

The second option considered in the present study involved no change in the commercially available interface but the addition of a second make-up solvent prior to MS inlet (Pump #2).

The advantage of this second option lies in the fact that optimizing, for this second make-up, both solvent flow-rate and solvent composition, should provide more versatile solutions depending on the type of complex sample and also depending on the polarity of APCI ionization source. The selection of the type of solvent entering APCI source may be of first importance to make easier the charge exchange between analytes and nitrogen plasma around the Corona needle. Water is usually recommended as additional solvent to enhance the ionization yield with an APCI source. Accordingly, a mixture of water and MeOH was considered within a composition range between 35/65 and 65/35 (water/MeOH, V/V). The solvent flow-rate was studied in the range 100-300 µL/min. Both ranges were found to be suitable in terms of both signal intensity and signal stability from a preliminary study with 15 model compounds detected in (-)APCI/HRMS (see Table 1). Model compounds were selected according to published studies on bio oil matrices and so that their retention times covered the whole retention space. 9 experiments well distributed among the parameter space were carried out with the proposed commercial interface at a make-up #1 flow of 500µL/min. For each of the 15 compounds, the signal-to-noise ratio, obtained with a given set of conditions was normalized with respect to the 9 sets of conditions, thereby providing a radar plot and a corresponding delimited area as shown in Fig. 4a. The calculated response function represented the fraction of the space occupied by the colored area and therefore varied between 0 and 1. The response function was fitted with a polynomial function. The resulting response surface in Fig.4b shows that the highest response values correspond to low flow-rates and high water concentrations. The response surface is curved with minimum response values at intermediate solvent compositions (i.e. around 50% water) which supports the necessity to optimize. It is important to note that optimization results are expected to be fully dependent on the analytes and it is therefore essential to carefully choose model compounds in accordance with the studied complex matrix and, if possible, with their retention times well distributed among the separation space as done in the present study.

Finally, our optimized conditions consisted in keeping the proposed commercial interface with a make-up solvent #1 composed of methanol at a flow-rate of 500 µL/min and a make-up solvent #2 composed of 65% water and 35% MeOH at a flow-rate of 100µL/min.

3.1. SFC-(-)APCI/HRMS results for model compounds

The analysis of complex samples such as biomass fast pyrolysis bio-oils by SFC-HRMS generates a huge amount of data that are not easy to process without dedicated software. We therefore built our own SFC/MS software as described in Materials and methods Section. This software was designed to attribute a molecular formula to each mass peak detected during the SFC run. The applied procedure was carried out according to the following golden rules suggested by Kind et al. [19] for filtering molecular formulae obtained by accurate mass spectrometry: (i) use any information about the sample (e.g. the major elements present and their relative abundance); (ii) use isotopic distribution around pseudo molecular ion signal; (iii) limit the number of heteroatoms in the molecular formula; (iv) use the ratios H/C and heteroatom/C to reduce the number of possibilities. The SFC/MS software was challenged with the SFC(-)APCI/HRMS analysis of a mixture containing 36 model compounds (see Table S1 in Supplementary Information), first dissolved in THF and then spiked in a bio-oil sample in order to highlight possible matrix effects which could reduce the ionization yield and hence could alter the quality of information. The results are displayed in Figs. 5a and 5b respectively, with base peak chromatogram (BPC) at the bottom and mass map at the top. For model compounds alone (Fig.5a), 15 peaks can be observed, well distributed across the separation. However 19 peaks were detected in (-)APCI/HRMS, suggesting that some model compounds were not separated in SFC (i.e peaks #2, #3 and #4; #5 and #6; #8 and #9 as can be seen in Fig.5b). The mass map generated by SFC/MS software allowed to add a third dimension corresponding to the mass over charge ratio (m/z). From these data, a unique molecular formula was proposed for each of the 19 detected compounds. Compound names, retention times, measured masses, molecular formulae resulting from SFC/MS calculation and corresponding mass errors are listed in Table 1. It is interesting to notice that, for each detected molecule, the accurate mass measurement allowed to propose the expected molecular formula, thereby leading to unambiguous molecular identification. For the sample composed of model compounds spiked in the bio-oil, the same 19 molecules could be detected and their molecular formulae identified in spite of possible matrix effects due to the presence of a very large number of components in bio-oil samples. By showing no effect of the bio-oil matrix on the ionization yield, these results ensure the suitability of the proposed method for formula identification.

3.2. SFC-(-)APCI/HRMS results for a biomass fast pyrolysis oil

Similarly, a bio-oil was analyzed with the same optimized conditions, using the same procedure. The results in terms of BPC chromatogram and mass map are shown in Fig.5c. These results give some valuable insights:

(i) The mass range (m/z) seems to be mainly between 150 and 400 uma which points out the complementarity of SFC and GCxGC which is known to provide a mass range rather between 0 and 200 uma [6].

(ii) The separation space is well occupied by the components except in the first part of the chromatogram corresponding to the isocratic step. This is not supported by UV detection which allowed to observe a large number of peaks in this first part [10] (see Fig. S1 in supplementary Information). Such peaks detected in UV but not detected in (-)APCI correspond to components, such as furans or non-aromatic ketones that are not easily ionized in APCI source. A complementary analysis with positive ionization could bring additional information on compounds that are more prone to favor the formation of $[M + H]^+$ ions.

(iii) From MS spectra resulting from SFC-(-)APCI/HRMS bio-oil analysis, 1379 molecular formulae could be proposed by our LC/SFC software. Among them, those corresponding to a model compound detected were investigated and 12 molecular formulae were found. They are listed in Table 2 along with their corresponding information (mass errors, retention times of both model compounds and similar molecular formulae found in the bio-oil). The difference in retention times (Table 2) allowed us to assess the degree of fit that the bio-oil compound had relative to the model compound. Based on a difference lower than 0.1 min, 7 model compounds (numbered in Table 1) or their positional isomers were strongly suspected to be present in the studied bio-oil: isoeugenol (#4), methoxynaphtol (#6), vanillin (#7), coniferaldehyde (#8), catechol (#12), vanillic acid (#16) and sinapic acid (#19). Moreover 5 model compounds that could be detected either alone or spiked in the bio-oil could not be detected in the bio-oil at their expected retention times. However their molecular formulae were identified at retention times significantly different, suggesting the presence of structural isomers.

(iv) Such mass maps could be easily used as characteristic fingerprints of complex samples allowing for in depth comparison of different samples.

The presence of different structural and/or positional isomers in the bio-oil was confirmed by a list of molecular formulae (Table 3) that were identified at different retention times. This result supports the fact that SFC can be a powerful analytical tool to discriminate compounds

having the same molecular formulae but different retention times, which is not possible with any direct HRMS analysis in direct infusion mode (i.e. without prior separation).

The heteroatom class distribution, with oxygen families ranging from O_1 to O_{15} and nitrogen family O_xN_1, is presented in Fig.6 for three equal parts of the SFC separation. Such data representation is often used with HRMS analysis. As can be observed and already highlighted, very few components could be detected in the first part of the separation. Although relative abundance distributions strongly depend on ionization conditions as well as on bio-oil properties, it can be observed that O_{11} to O_{15} families (most oxygenated compounds) were mainly detected in the third part while O_2 to O_6 families were more intense in the second part which is consistent with expected retention in SFC on a polar stationary phase (i.e. Acquity BEH-EP)

3.1. Comparison of SFC-HRMS and FT-ICR/MS analysis of a biomass fast pyrolysis oil

In order to have a clear idea about how the proposed analytical technique can be complementary to modern HRMS techniques offering very high resolving power, we compared SFC-HRMS to FT-ICR/MS (Fourier transform-ion cyclotron resonance mass spectrometry) for the specific case of bio-oil analysis. Since several years, HRMS techniques alone are being increasingly used to describe the composition of biomass fast pyrolysis oils [3]. In particular, FT-ICR/MS has gained in interest over the past ten years, providing key information in terms of m/z ratios, molecular formulae and double bound equivalent (DBE) for compounds being detected essentially by electrospray ionization source, and more scarcely by APPI [20–25]. In this work, the studied bio-oil was also analyzed by FT-ICR/MS using an APCI source in negative mode and the resulting data were compared with those obtained in SFC-(-)APCI/HRMS. A very large number of peaks (*i.e.* 3949 identified molecular formulae) were detected by FT-ICR/MS, illustrating its huge sensitivity compared to SFC-IT-TOF/MS (i.e. 1379 identified molecular formulae). However it is interesting to notice that among all the molecular formulae identified by FT-ICR/MS and SFC/MS, only 835 were common to both techniques. That underlines the great benefit of SFC prior to HMRS which enables the separation of several positional or structural isomers (as shown in Table 3) while direct injection in FT-ICR/MS cannot differentiate them, leading to the same molecular formula if no additional structural data are provided.

It should be noted that, similarly to SFC-HRMS, the results obtained in FT-ICR/MS must only be used for qualitative analysis due to the dependence of the response factor on the compound. This also implies that any attempt to compare FT-ICR/MS and SFC-HRMS data must be done with caution. However the heteroatom class distributions might be compared in terms of their relative abundance. It appears in Fig.7 that both distributions are different although the ionization source (i.e. (-)APCI) was the same. Our (-)APCI/FT-ICR/MS results are quite consistent with reported studies dealing with (-)ESI/FT-ICR/MS in which distributions were focused on O_3 to O_8 families [22,26–28]. The comparison of both heteroatom class distributions (Fig.7) indicates that same ranges of O_x families are covered by FT-ICR/MS and SFC-HRMS, with a clear benefit of SFC-HRMS to specifically analyze molecules having low number of oxygen atoms (O_1-O_3), suggesting that SFC separation prior to HRMS detection greatly enhances the detection of such species by preventing from strong ion suppression which may occur when the whole bio-oil is directly introduced in FT-ICR/MS. Indeed some reported studies on different biomass products have proved that polar analytes are much more affected by matrix effects than nonpolar ones [29,30]. Furthermore the relative intensity for O_{12} to O_{15} families seems to be higher in SFC-HRMS than in FT-ICR. These results also suggest that a better ionization yield can be achieved in SFC-HRMS for these highly-oxygenated compounds, thereby still supporting the fact that the separation prior to HRMS can be very useful.

Another interesting way to present the results and to get relevant information about bio-oil composition consists in drawing van Krevelen diagram, based upon elemental formulae, in the form of a dot matrix representing H/C ratio versus O/C ratio (Fig.8). These ratios are characteristic of a compound class which can be identified by a delimited area in the diagram. As underlined by Stas et al. [3], this diagram can be used to evaluate (1) the abundance of compounds from different classes and (2) the correlation between compounds from different classes. Both van Krevelen diagrams derived from SFC-(-)APCI/MS (Fig. 8a) and FT-ICR/MS (Fig. 8b) data are in good agreement. Detected species are intensively focused within areas usually dedicated to phenolics (i.e. O/C = 0-0.6; H/C = 0.5-1.5) and carbohydrates (i.e. O/C = 0.6-1.1 ; H/C > 1.5). This shows that a high number of compounds exhibiting a medium polarity are present in the studied bio-oil and can be detected by (-) APCI.

4. Conclusion

This study presents the first detailed characterization of a bio-oil by SFC hyphenated to HRMS with negative ion APCI as ionization source. The interface between SFC and (-)APCI/HRMS was optimized for a specific commercial equipment with a procedure that can be applied in the future to any ionization source and any commercially available equipment provided that tubing geometry are known and model compounds are available.

As shown, this coupling can be a valuable technique for assessing bio-oil composition and an alternative and complement to more usual methods such as HRMS alone or GCxGC-MS. It was pointed out that some model compounds could not be detected by using the single negative ion APCI as ionization technique, suggesting that additional ionization techniques (i.e. APCI in positive mode and ESI in positive and negative modes) should be combined to achieve a more comprehensive bio-oil analysis.

In spite of very attractive analytical possibilities due to its very high resolving power, FT-ICR/MS alone cannot permit the distinction between positional and structural isomers which can be abundant in complex samples such as bio-oils as highlighted in this study. Moreover, a clear reduction of signal intensity, likely due to matrix effects, was pointed out in FT-ICR/MS.

Overall, SFC-HRMS is a very promising analytical tool for the analysis of complex chemical samples. The proposed mass-maps as characteristic fingerprints could be useful for in-depth comparison of complex samples. Finally, considering the ability of SFC to both separate isomers and reduce matrix effects, its hyphenation to high resolution mass spectrometry can provide an access to a large number of detailed data, mandatory to go further on complex sample characterization.

References

[1] A.V. Bridgwater, H. Hofbauer, S. van Loo, Thermal Biomass Conversion, 2009.

[2] C.M. Michailof, K.G. Kalogiannis, T. Sfetsas, D.T. Patiaka, A.A. Lappas, Advanced analytical techniques for bio-oil characterization, WIREs Energy Environ 5 (6) (2016) 614–639.

[3] 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.

[4] P.K. Kanaujia, Y.K. Sharma, U.C. Agrawal, M.O. Garg, Analytical approaches to characterizing pyrolysis oil from biomass, TrAC Trends in Analytical Chemistry 42 (2013) 125–136.

[5] N. Charon, J. Ponthus, D. Espinat, F. Broust, G. Volle, J. Valette, D. Meier, Multi-technique characterization of fast pyrolysis oils, Journal of Analytical and Applied Pyrolysis 116 (2015) 18–26.

[6] M. Stas, D. Kubic, J. Chudoba, M. Pospís, Overview of Analytical Methods Used for Chemical Characterization of Pyrolysis Bio-oil, Energy Fuels (28) (2014) 385–402.

[7] A. Le Masle, D. Angot, C. Gouin, A. D'Attoma, J. Ponthus, A. Quignard, S. Heinisch, Development of on-line comprehensive two-dimensional liquid chromatography method for the separation of biomass compounds, Journal of chromatography A 1340 (2014) 90–98.

[8] D. Tomasini, F. Cacciola, F. Rigano, D. Sciarrone, P. Donato, M. Beccaria, E.B. Caramão, P. Dugo, L. Mondello, Complementary Analytical Liquid Chromatography Methods for the Characterization of Aqueous Phase from Pyrolysis of Lignocellulosic Biomasses, Anal. Chem. 86 (22) (2014) 11255–11262.

[9] M. Sarrut, A. Corgier, G. Crétier, A. Le Masle, S. Dubant, S. Heinisch, Potential and limitations of on-line comprehensive reversed phase liquid chromatography×supercritical fluid chromatography for the separation of neutral compounds: An approach to separate an aqueous extract of bio-oil, Journal of chromatography A 1402 (2015) 124–133.

[10] J. Crepier, A. Le Masle, N. Charon, F. Albrieux, S. Heinisch, Development of a supercritical fluid chromatography method with ultraviolet and mass spectrometry detection for the characterization of biomass fast pyrolysis bio oils, Journal of Chromatography A 1510 (2017) 73-81.

[11] E. Lemasson, S. Bertin, P. Hennig, E. Lesellier, C. West, Comparison of ultra-high performance methods in liquid and supercritical fluid chromatography coupled to electrospray ionization - mass spectrometry for impurity profiling of drug candidates, Journal of chromatography A 1472 (2016) 117–128.

[12] L. Novakova, M. Rentsch, A. Grand-Guillaume Perrenoud, R. Nicoli, M. Saugy, J.-L. Veuthey, D. Guillarme, Ultra high performance supercritical fluid chromatography coupled with tandem mass spectrometry for screening of doping agents. II: Analysis of biological samples, Analytica Chimica Acta 853 (2015) 647–659.

513 [13] D. Spaggiari, F. Mehl, V. Desfontaine, A. Grand-Guillaume Perrenoud, S. Fekete, S. Rudaz,
514 D. Guillarme, Comparison of liquid chromatography and supercritical fluid
515 chromatography coupled to compact single quadrupole mass spectrometer for targeted
516 in vitro metabolism assay, Journal of chromatography A 1371 (2014) 244–256.
517 [14] J. Duval, C. Colas, V. Pecher, M. Poujol, J.-F. Tranchant, E. Lesellier, Hyphenation of ultra
518 high performance supercritical fluid chromatography with atmospheric pressure
519 chemical ionisation high resolution mass spectrometry: Part 1. Study of the coupling
520 parameters for the analysis of natural non-polar compounds, Journal of chromatography
521 A 1509 (2017) 132–140.
522 [15] P.G.A. Pedrioli, J.K. Eng, R. Hubley, M. Vogelzang, E.W. Deutsch, B. Raught, B. Pratt, E.
523 Nilsson, R.H. Angeletti, R. Apweiler, K. Cheung, C.E. Costello, H. Hermjakob, S. Huang, R.K.
524 Julian, E. Kapp, M.E. McComb, S.G. Oliver, G. Omenn, N.W. Paton, R. Simpson, R. Smith,
525 C.F. Taylor, W. Zhu, R. Aebersold, A common open representation of mass spectrometry
526 data and its application to proteomics research, Nature biotechnology 22 (11) (2004)
527 1459–1466.
528 [16] A. Grand-Guillaume Perrenoud, C. Hamman, M. Goel, J.-L. Veuthey, D. Guillarme, S.
529 Fekete, Maximizing kinetic performance in supercritical fluid chromatography using
530 state-of-the-art instruments, Journal of chromatography A 1314 (2013) 288–297.
531 [17] L.-B. Ouyang, New Correlations for Predicting the Density and Viscosity of Supercritical
532 Carbon Dioxide Under Conditions Expected in Carbon Capture and Sequestration
533 Operations, TOPEJ 5 (1) (2011) 13–21.
534 [18] D. Guillarme, V. Desfontaine, S. Heinisch, J.-L. Veuthey, Journal of Chromatography B
535 (2018) Submitted.
536 [19] T. Kind, O. Fiehn, Seven Golden Rules for heuristic filtering of molecular formulas
537 obtained by accurate mass spectrometry, BMC bioinformatics 8 (2007) 105.
538 [20] N.S. Tessarolo, R.V.S. Silva, G. Vanini, A. Casilli, V.L. Ximenes, F.L. Mendes, A. Rezende
539 Pinho, W. Romão, E.V.R. Castro, C.R. Kaiser, D.A. Azevedo, Characterization of thermal
540 and catalytic pyrolysis bio-oils by high-resolution techniques: 1 H NMR, GC × GC-TOFMS
541 and FT-ICR MS, Journal of Analytical and Applied Pyrolysis 117 (2016) 257–267.
542 [21] N.S. Tessarolo, R.C. Silva, G. Vanini, A. Pinho, W. Romão, E.V.R. Castro, D.A. Azevedo,
543 Assessing the chemical composition of bio-oils using FT-ICR mass spectrometry and

544 comprehensive two-dimensional gas chromatography with time-of-flight mass
545 spectrometry, Microchemical Journal 117 (2014) 68–76.
546 [22] E.A. Smith, S. Park, A.T. Klein, Y.J. Lee, Bio-oil Analysis Using Negative Electrospray
547 Ionization: Comparative Study of High-Resolution Mass Spectrometers and Phenolic
548 versus Sugaric Components, Energy Fuels 26 (6) (2012) 3796–3802.
549 [23] J.M. Jarvis, A.M. McKenna, R.N. Hilten, K.C. Das, R.P. Rodgers, A.G. Marshall,
550 Characterization of Pine Pellet and Peanut Hull Pyrolysis Bio-oils by Negative-Ion
551 Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry,
552 Energy Fuels 26 (6) (2012) 3810–3815.
553 [24] J.M. Jarvis, D.S. Page-Dumroese, N.M. Anderson, Y. Corilo, R.P. Rodgers, Characterization
554 of Fast Pyrolysis Products Generated from Several Western USA Woody Species, Energy
555 Fuels 28 (10) (2014) 6438–6446.
556 [25] I. Miettinen, M. Mäkinen, T. Vilppo, J. Jänis, Characterization of Phase-Separated Pine
557 Wood Slow Pyrolysis Oil by Negative-Ion Electrospray Ionization Fourier Transform Ion
558 Cyclotron Resonance Mass, Energy Fuels (2015).
559 [26] N. Sudasinghe, J.R. Cort, R. Hallen, M. Olarte, A. Schmidt, T. Schaub, Hydrothermal
560 liquefaction oil and hydrotreated product from pine feedstock characterized by
561 heteronuclear two-dimensional NMR spectroscopy and FT-ICR mass spectrometry, Fuel
562 137 (2014) 60–69.
563 [27] P.V. Abdelnur, B.G. Vaz, J.D. Rocha, M.B.B. de Almeida, M.A.G. Teixeira, R.C.L. Pereira,
564 Characterization of Bio-oils from Different Pyrolysis Process Steps and Biomass Using
565 High-Resolution Mass Spectrometry, Energy Fuels 27 (11) (2013) 6646–6654.
566 [28] I. Miettinen, S. Kuittinen, V. Paasikallio, M. Mäkinen, A. Pappinen, J. Jänis,
567 Characterization of fast pyrolysis oil from short-rotation willow by high-resolution Fourier
568 transform ion cyclotron resonance mass spectrometry, Fuel 207 (2017) 189–197.
569 [29] R. Bonfiglio, R.C. King, T.V. Olah, K. Merkle, The effects of sample preparation methods
570 on the variability of the electrospray ionization response for model drug compounds,
571 Rapid Communications in Mass Spectrometry 13 (12) (1999) 1175–1185.
572 [30] Lekh Nath Sharma, Identification and Quantitation of Potential Fermentation Inhibitors
573 in Biomass Pretreatment Hydrolysates Using High Performance Liquid Chromatography
574 in Combination with Ultraviolet Detection and Tandem Mass Spectrometry.
575

Figure captions

Figure 1 : Schematic representation of the interface used for hyphenation of SFC to APCI/IT-TOF/MS. The proposed commercial interface is delimited by the dotted frame. T1, T2 and T3 represent the 3 zero-dead volume T-unions.

Figure 2 : Variation of solvent flow entering MS source as a function of Pump #1 flow-rate for two different ACN compositions in the mobile phase (1%ACN and 40%ACN). Theoretical and experimental curves are represented by solid and dotted lines respectively. Conditions : Waters interface (see Fig.2); mobile phase flow-rate : 1.4mL/min; BPR 150 bar; 30°C.

Figure 3 : Theoretical variation of (a) solvent flow entering MS source and (b) split ratio, as a function of both Pump #1 flow-rate and restriction capillary length (i.d. 50µm) with 1%ACN as co-solvent. Same other conditions as in Fig.2

Figure 4 : Illustration of the response function calculation and its variation depending on solvent make-up #2 conditions. (a) Radar plots representing the normalized signal-to-noise ratio for 15 model compounds (see Table 1 for the numbering) obtained with a given set of conditions. The response function is the fraction of the space occupied by the blue colored area (b) Response function versus both the Pump #2 flow and the composition of solvent.

Figure 5 : Mass maps and Base Peak Chromatograms of (a) model mix; (b) spiked bio-oil sample and (c) bio-oil sample analyzed in SFC-(-)APCI/HRMS. Detected model compounds are numbered in the different figures. (see Table 1 for analytical results) . Chromatographic conditions are given in Materials and methods Section.

Figure 6 : Heteroatom class distributions for the first (blue), second (red) and third part (green) of the SFC separation derived from (-) APCI/HRMS mass spectra. Sample: fast pyrolysis bio-oil. SFC and MS conditions given in Materials and methods Section.

606 Figure 7 : Comparison of heteroatom class distributions between SFC-(-)APCI/HRMS and FT-
607 ICR/MS, both with negative ion APCI as ionization source. Conditions given in Materials and
608 methods Section.
609
610 Figure 8 : Comparison of the van Krevelen diagrams (H/C vs O/C) of a bio-oil, obtained from
611 (a) (-)APCI/FT-ICR/MS and (b) SFC-(-)APCI/HRMS data. Each dot corresponds to an identified
612 molecular formula with color related to its relative abundance. Fast pyrolysis bio-oil sample.
613 SFC and MS conditions given in Materials and methods Section.

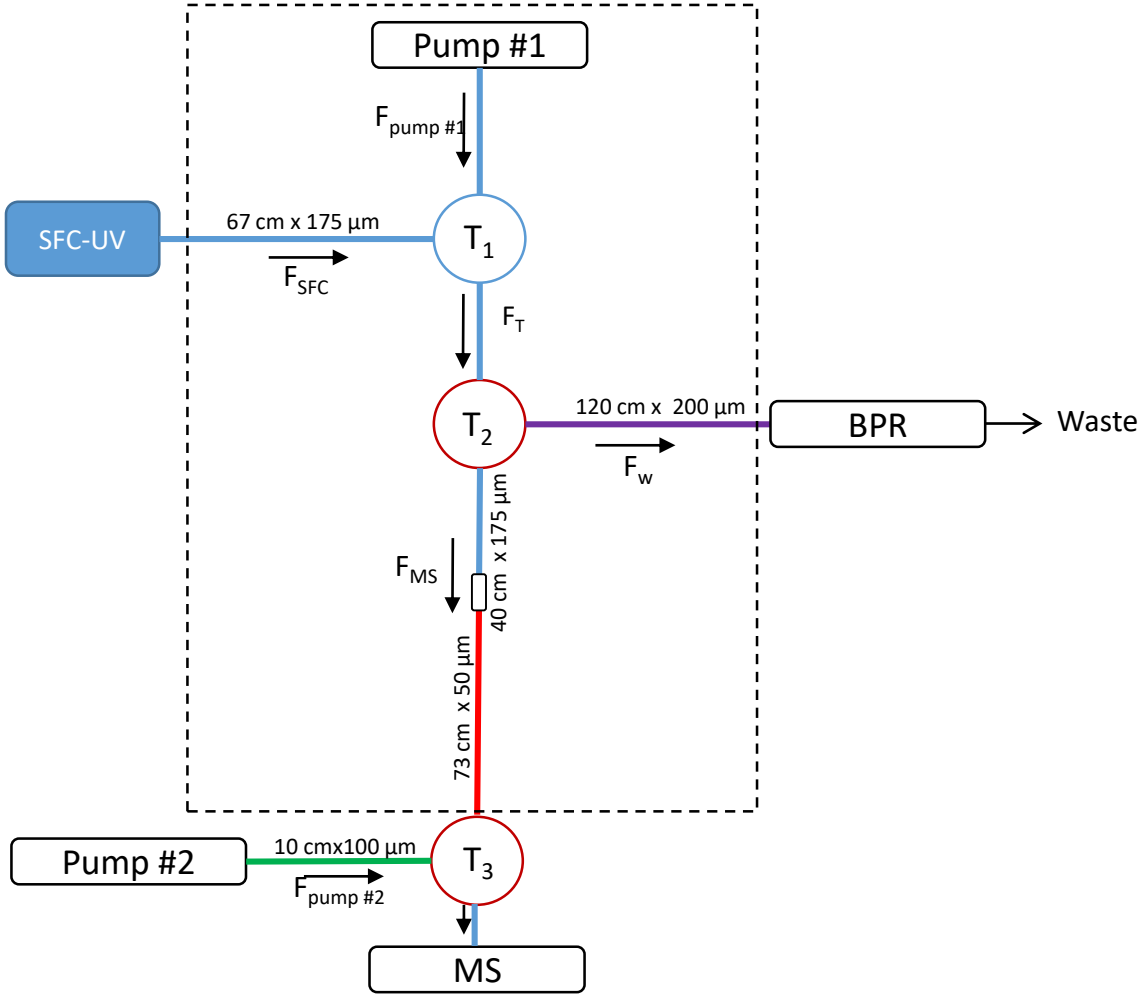

Pump #1
$F_{pump\ \#1}$
SFC-UV
67 cm x 175 µm
F_{SFC}
T_1
F_T
T_2
120 cm x 200 µm
BPR
Waste
F_w
F_{MS}
40 cm x 175 µm
73 cm x 50 µm
Pump #2
10 cmx100 µm
$F_{pump\ \#2}$
T_3
MS

Figure 1

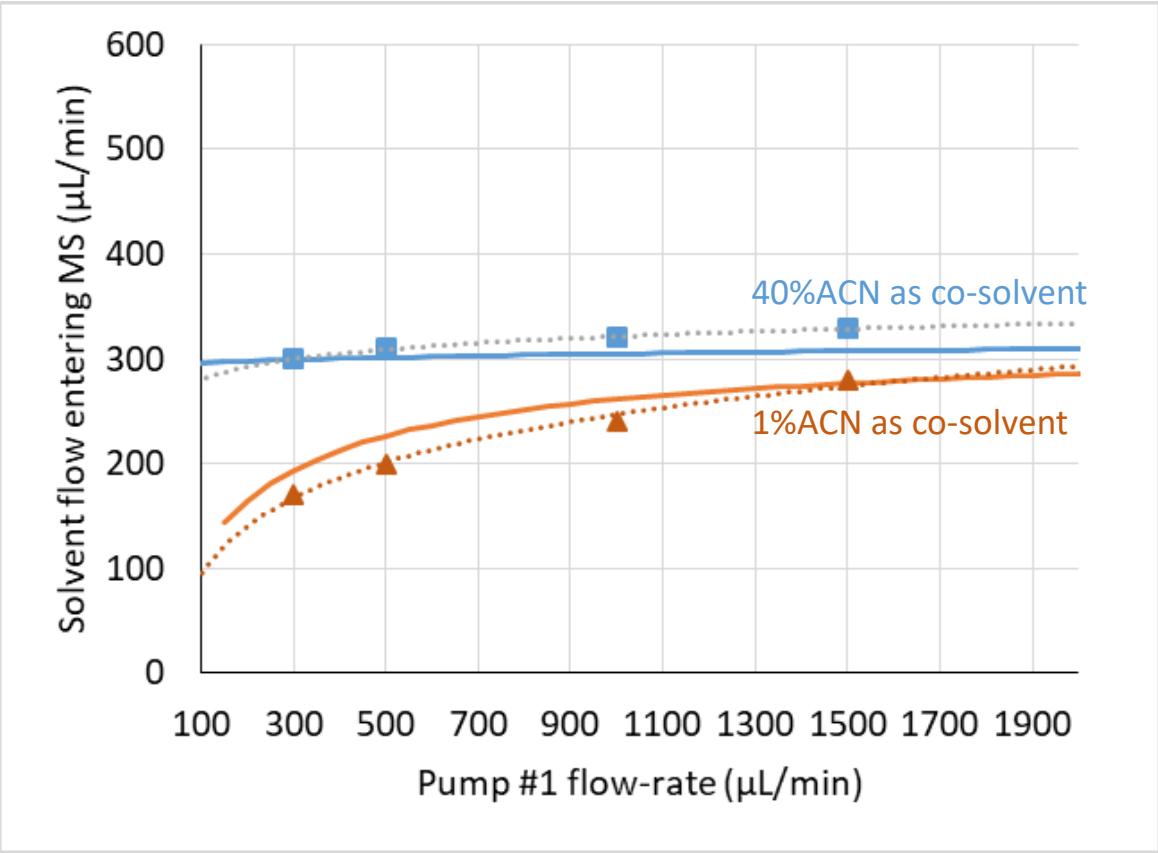
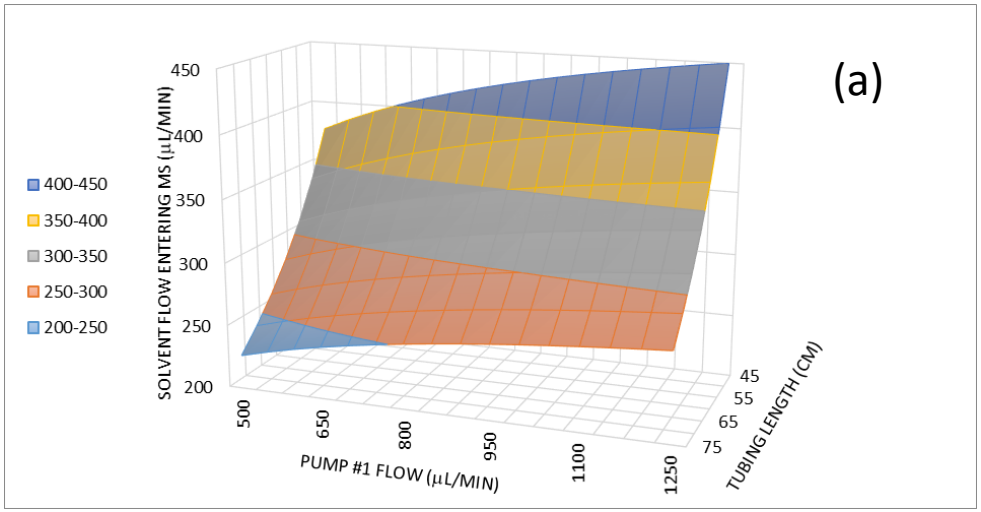

600
500
400
300
200
100
0
Solvent flow entering MS (µL/min)
40%ACN as co-solvent
1%ACN as co-solvent
100 300 500 700 900 1100 1300 1500 1700 1900
Pump #1 flow-rate (µL/min)

Figure 2


(a)
400-450
350-400
300-350
250-300
200-250
SOLVENT FLOW ENTERING MS (µL/MIN)
450
400
350
300
250
200
500
650
800
950
1100
1250
PUMP #1 FLOW (µL/MIN)
45
55
65
75
TUBING LENGTH (CM)

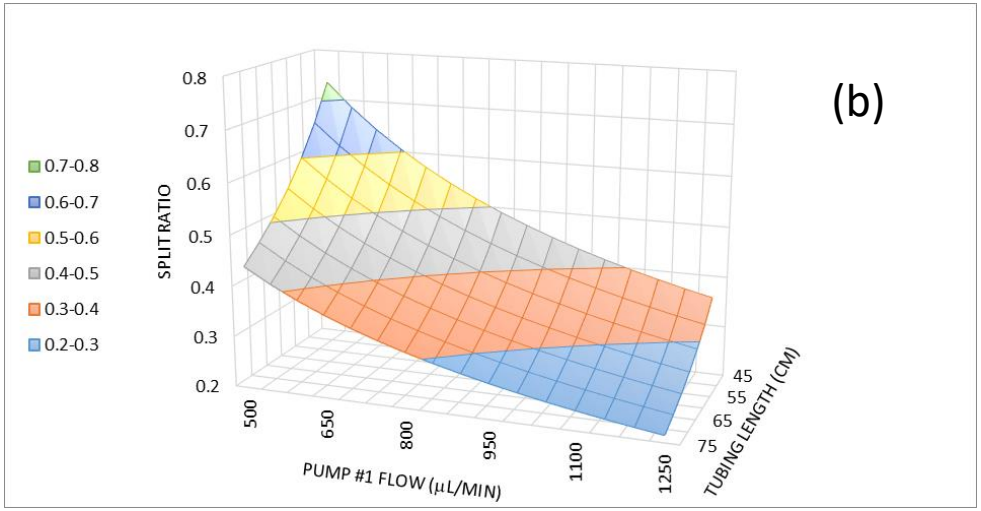

(b)
0.7-0.8
0.6-0.7
0.5-0.6
0.4-0.5
0.3-0.4
0.2-0.3
SPLIT RATIO
0.8
0.7
0.6
0.5
0.4
0.3
0.2
500
650
800
950
1100
1250
PUMP #1 FLOW (µL/MIN)
45
55
65
75
TUBING LENGTH (CM)

Figure 3

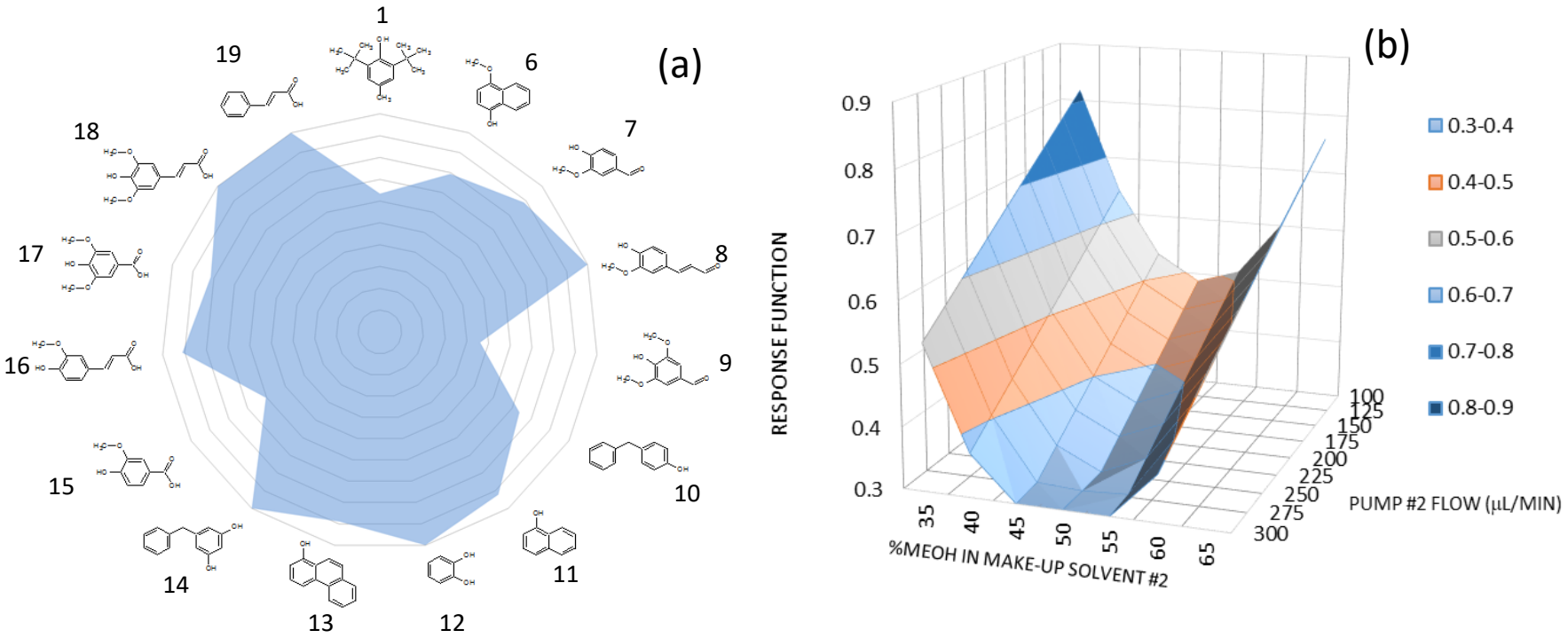

(a)
1
6
7
8
9
10
11
12
13
14
15
16
17
18
19
(b)
RESPONSE FUNCTION
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.3-0.4
0.4-0.5
0.5-0.6
0.6-0.7
0.7-0.8
0.8-0.9
%MEOH IN MAKE-UP SOLVENT #2
35
40
45
50
55
60
65
PUMP #2 FLOW (µL/MIN)
100
125
150
175
200
225
250
275
300

Figure 4

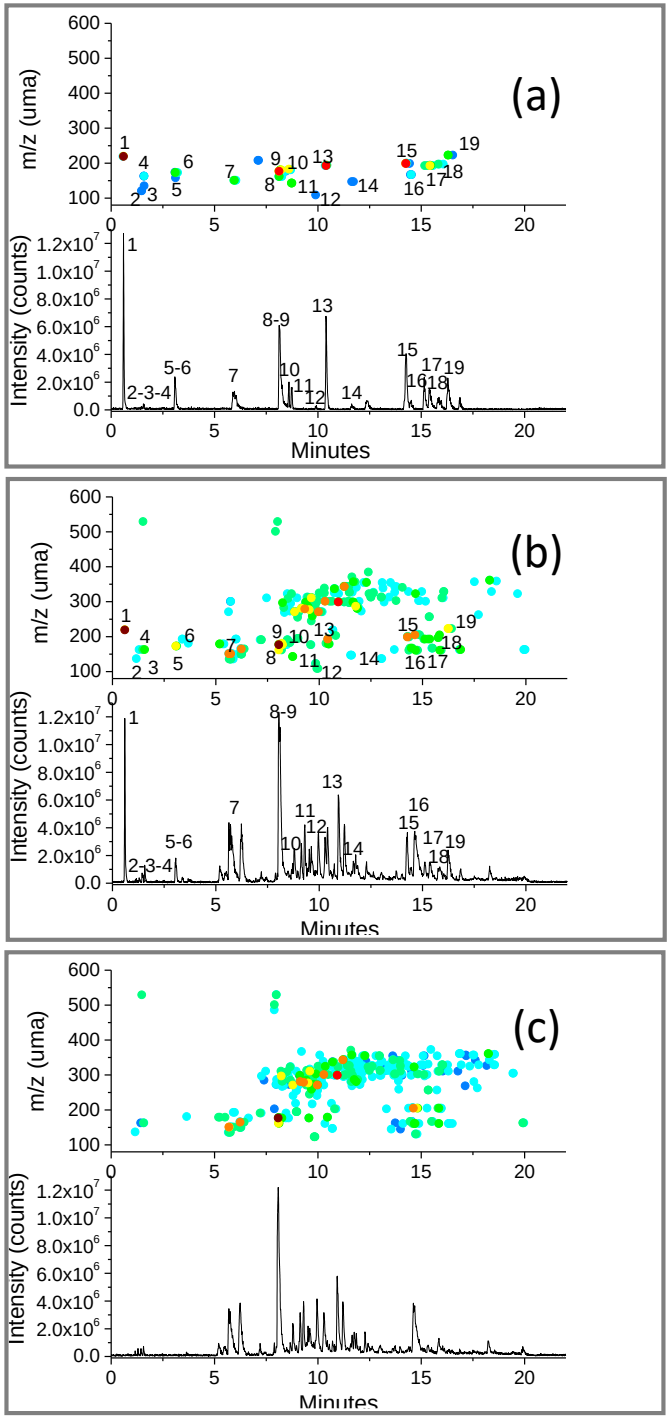

(a)
m/z (uma)
Intensity (counts)
Minutes
(b)
m/z (uma)
Intensity (counts)
Minutes
(c)
m/z (uma)
Intensity (counts)
Minutes

Figure 5

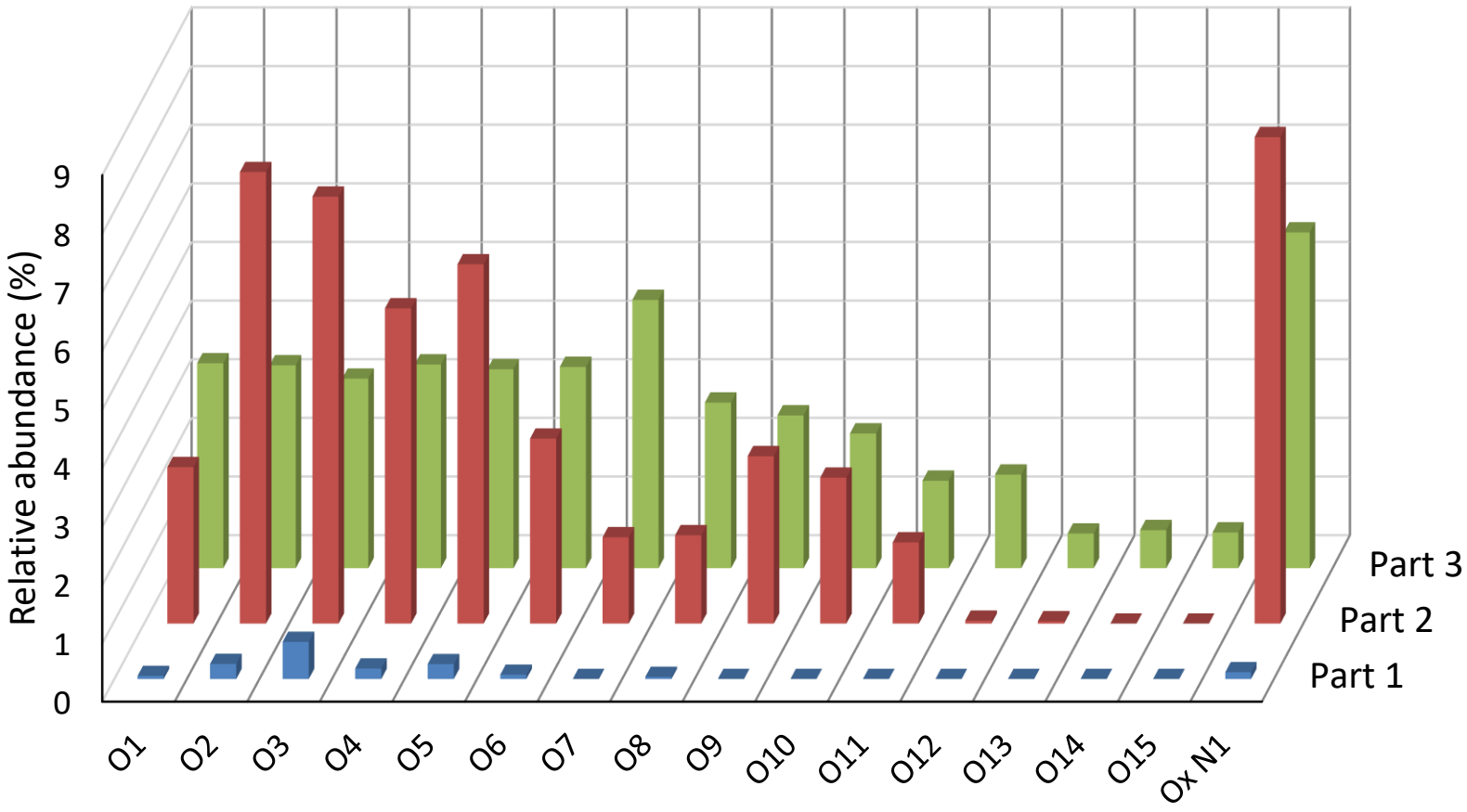

Relative abundance (%)
0
1
2
3
4
5
6
7
8
9
O1
O2
O3
O4
O5
O6
O7
O8
O9
O10
O11
O12
O13
O14
O15
Ox N1
Part 3
Part 2
Part 1

Figure 6

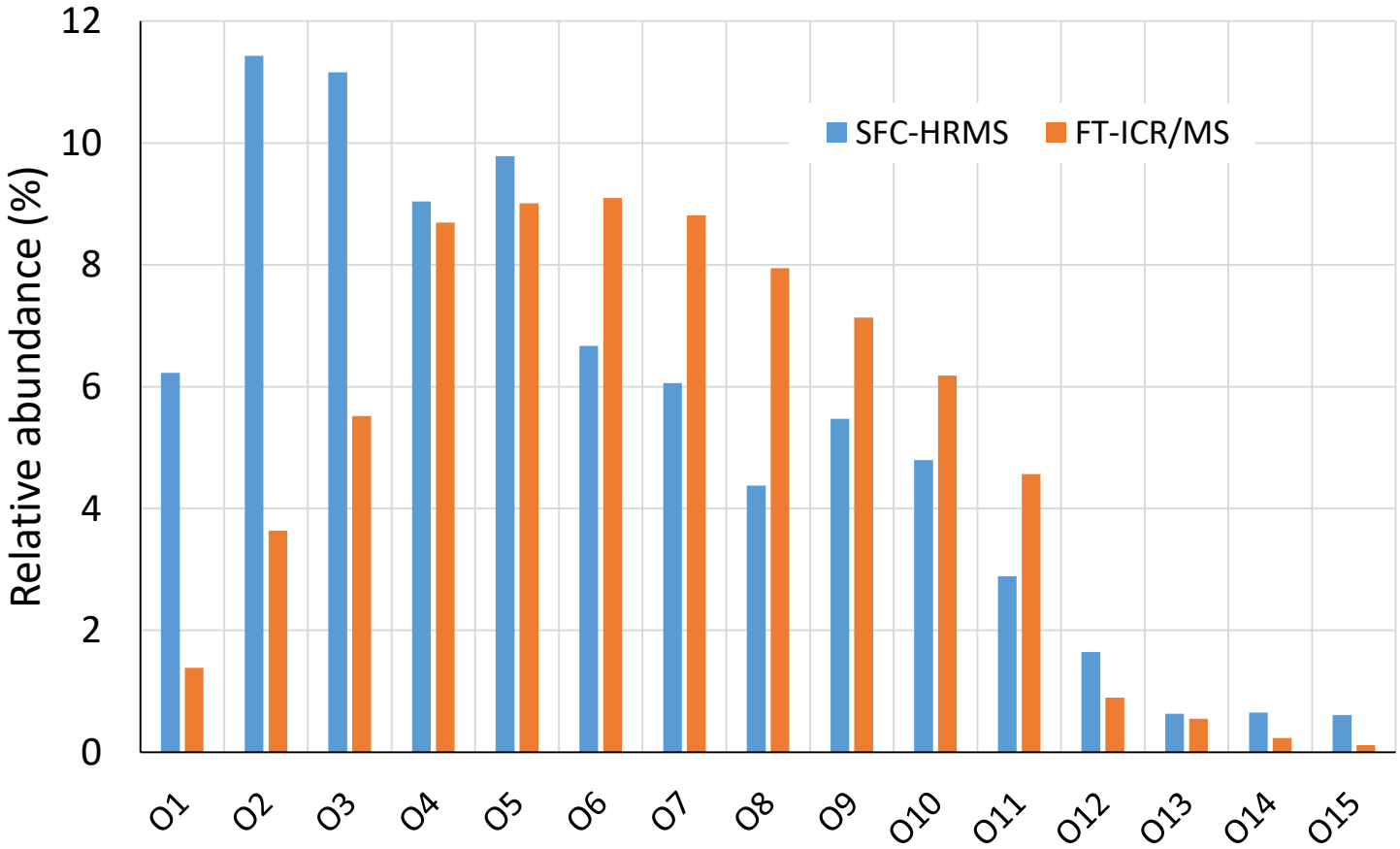

Relative abundance (%)
12
10
8
6
4
2
0
SFC-HRMS
FT-ICR/MS
O1
O2
O3
O4
O5
O6
O7
O8
O9
O10
O11
O12
O13
O14
O15

Figure 7

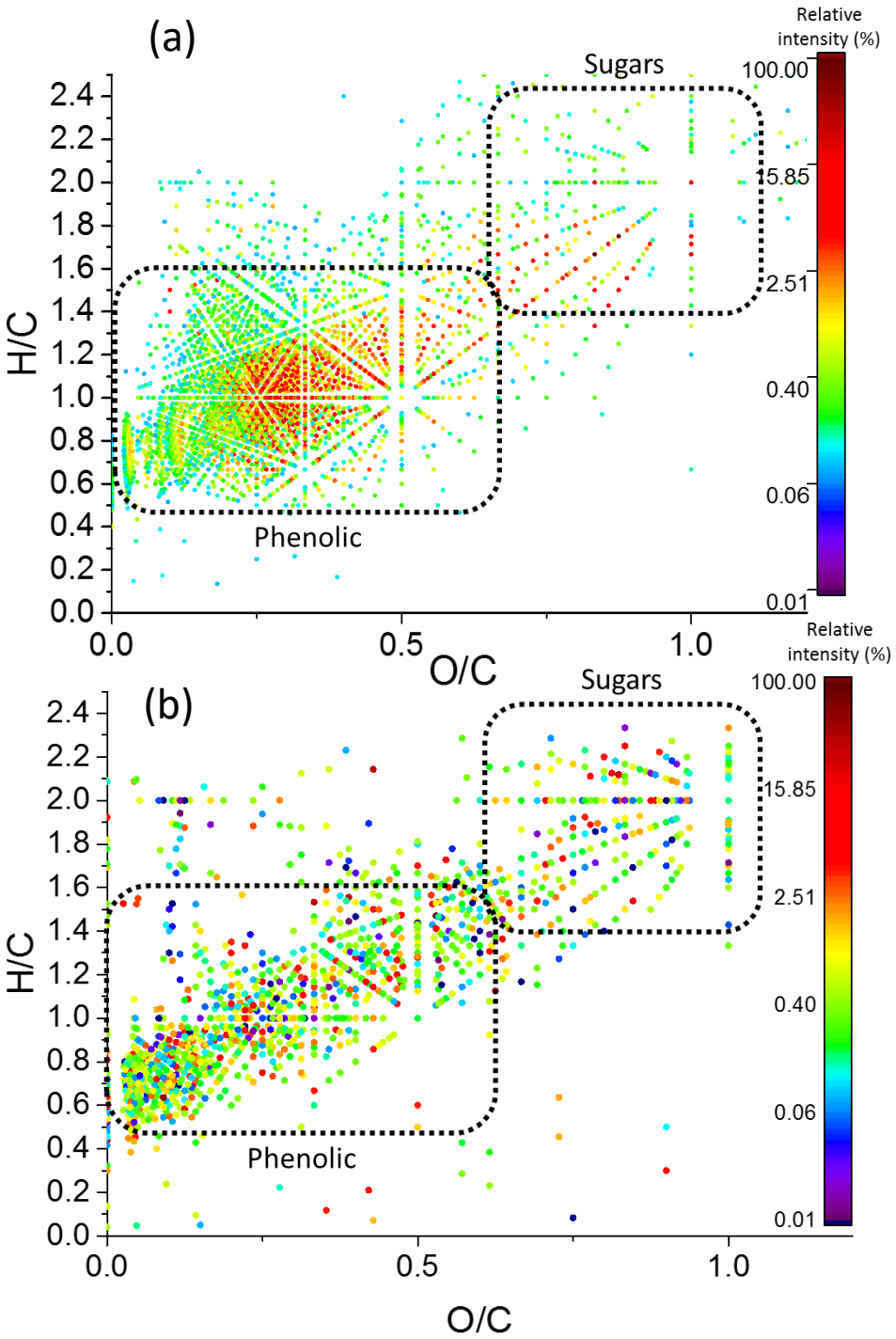

(a)
Sugars
Phenolic
Relative intensity (%)
100.00
15.85
2.51
0.40
0.06
0.01
H/C
O/C
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
0.0
0.5
1.0
(b)
Sugars
Phenolic
Relative intensity (%)
100.00
15.85
2.51
0.40
0.06
0.01
H/C
O/C

Figure 8

Table 1 : List of 19 (among 36) model compounds detected in SFC-(-)APCI-HRMS and their corresponding results obtained from SFC/MS software. See experimental section for SFC and MS conditions.

	IUPAC name	Usual name	Retention time (min)	Accurate weight [M-H]-	Molecular formula	Mass error Δm (ppm)
1	2,6-ditertbuthyl-4methylphénol	2,6-ditertbuthyl-4methylphénol	0.597	219.1765	$C_{15}H_{24}O_1$	-0.634
2	2,6-Dimethylphenol	Xylenol	1.478	121.0685	$C_8H_{10}O_1$	-14.773
3	2,4,6-trimethylphenol	Trimethyl phenol	1.595	135.079	$C_9H_{12}O_1$	-19.936
4	2-methoxy-4-[(E)-prop-1-enyl]phenol	Isoeugenol	1.595	163.0792	$C_{10}H_{12}O_2$	-18.015
5	1,3-Dimethoxy-2-hydroxybenzene	Syringol	2.532	153.0582	$C_8H_{10}O_3$	8.378
6	2-Methoxy-1-naphthol	Methoxy-naphtol	3.137	173.0628	$C_{11}H_{10}O_2$	19.051
7	4-hydroxy-3-méthoxybenzaldéhyde	Vanillin	6.062	151.0425	$C_8H_8O_3$	19.414
8	(E)-3-(4-hydroxy-3-methoxyphenyl)prop-2-enal	Coniferaldehyde	8.157	177.0589	$C_{10}H_{10}O_3$	8.937
9	4-hydroxy-3,5-diméthoxybenzaldéhyde	Syringaldehyde	8.217	181.056	$C_9H_{10}O_4$	5.345
10	4-Benzylphenol	Hydroxy-diphenylmethane	8.622	183.0841	$C_{13}H_{12}O_1$	13.991
11	Naphtalén-1-ol	Naphtol	8.757	143.0516	$C_{10}H_8O_1$	14.412
12	Benzene-1,2-diol	Catechol	9.893	109.03	$C_6H_6O_2$	-2.779
13	9-Phenanthrenol	Phenantrol	10.438	193.0691	$C_{14}H_{10}O_1$	17.670
14	(2E)-3-Phenylprop-2-enoic acid	Trans-cinnamic acid	11.625	147.0435	$C_9H_8O_2$	-11.242
15	4-benzylbenzene-1,3-diol	Benzyl-resorcinol	14.3	199.0779	$C_{13}H_{12}O_2$	-9.811
16	4-Hydroxy-3-methoxybenzoic acid	Vanillic acid	14.553	167.038	$C_8H_8O_4$	-14.861
17	(E)-3-(4-hydroxy-3-methoxy-phenyl)prop-2-enoic acid	Ferulic acid	15.418	193.0539	$C_{10}H_{10}O_4$	-8.974
18	4-Hydroxy-3,5-dimethoxybenzoic acid	Syringic acid	15.877	197.0459	$C_9H_{10}O_5$	1.792
19	3-(4-hydroxy-3,5-dimethoxyphenyl)prop-2-enoic acid	Sinapic acid	16.3	223.0656	$C_{11}H_{12}O_5$	15.256

Table 2: List of molecular formulae corresponding to model compounds (listed in Table 1) and identified in the bio-oil sample. The difference in retention times allows to assess the degree of fit that a compound found in the bio-oil has relative to a model compound.

		Model compounds		***Bio oil sample***		***Δ (tr)[(d)] (min)***
Molecular formula (n)[(a)]	***Accurate mass [M-H]⁻***	***tr[(b)]***	***Δm[(c)]***	***tr[(b)]***	***Δm[(c)]***	
$C_{10}H_{12}O_2$ (4)	163.0792	1.59	*18.01*	1.58	*13.16*	0.01
$C_{11}H_{10}O_2$ (6)	173.0628	3.05	*19.05*	3.10	*17.31*	0.05
$C_8H_8O_3$ (7)	151.0425	5.43	*19.41*	5.49	*17.42*	0.06
$C_{10}H_{10}O_3$ (8)	177.0589	8.11	*8.93*	8.06	*11.76*	0.05
$C_9H_{10}O_4$ (4)	181.0560	8.11	*5.34*	3.67	*15.83*	4.44
$C_6H_6O_2$ (12)	109.0300	9.90	*2.78*	9.90	*13.78*	0.00
$C_{14}H_{10}O_1$ (13)	193.0691	10.44	*17.67*	19.61	*18.07*	9.17
$C_9H_8O_2$ (14)	147.0435	11.60	*11.24*	10.68	*11.20*	0.92
$C_8H_8O_4$ (16)	167.0380	14.32	*14.86*	14.39	*9.08*	0.07
$C_{10}H_{10}O_4$ (17)	193.0539	15.42	*8.97*	5.94	*15.89*	9.48
$C_9H_{10}O_5$ (18)	197.0459	15.80	*1.79*	10.37	*11.40*	5.43
$C_{11}H_{12}O_5$ (19)	223.0656	16.26	*15.26*	16.31	*5.84*	0.05

(a) : model compound number (as in Table 1)
(b) : retention times (min)
(c) : mass error (ppm)
(d) : difference in retention times (min) between model compound and similar bio-oil molecular formula

Table 3: List of molecular formulae identified at several different retention times for a fast pyrolysis bio-oil in SFC-(-)APCI-HRMS. See experimental section for SFC and MS conditions.

Molecular formula	Retention times (min)				Unitary mass (uma)
$C_{10}H_{10}O_2$	10.34	14.71	16.51		162
$C_{10}H_{12}O_2$	1.56	19.91			164
$C_{10}H_{10}O_3$	8.13	8.46	6.61		178
$C_{10}H_{12}O_3$	5.19	10.47			180
$C_{11}H_{12}O_4$	6.37	14.68			208
$C_{16}H_{16}O_4$	8.78	9.95			272
$C_6H_{10}O_5$	14.67	15.84			162
$C_{20}H_{24}O_5$	11.2	13.62	15.33		344
$C_8H_{12}O_6$	13.31	15.95			204
$C_{14}H_{18}O_6$	11.62	11.85	17.54		282
$C_{16}H_{16}O_6$	9.9	14.81			304
$C_{20}H_{26}O_6$	16.89	18.24			362
$C_7H_{10}O_7$	14.61	14.78	15.8		206
$C_{12}H_{16}O_7$	9.4	16.74			272
$C_{13}H_{20}O_7$	12.15	17.95			288
$C_{13}H_{18}O_8$	12.1	13.19	15.36		302
$C_9H_{20}O_9$	8.81	15.01			272
$C_{12}H_{22}O_9$	8.56	10.77			310
$C_{10}H_{20}O_{10}$	9.12	13.01	14.51	15.7	300
$C_{10}H_{22}O_{10}$	11.34	13.89	15.19		302
$C_{11}H_{22}O_{10}$	10.87	12.45	15.58	16.63	314
$C_{11}H_{24}O_{10}$	11.3	18.24			316
$C_{12}H_{20}O_{10}$	8.38	10.97	14.65		324
$C_{12}H_{22}O_{10}$	10.35	12.69	14.34	16.21	326
$C_{13}H_{22}O_{10}$	10.7	10.77	12.18		338

Supplementary information for

Ultra-high performance supercritical fluid chromatography hyphenated to atmospheric pressure chemical ionization high resolution mass spectrometry for the characterization of fast pyrolysis bio-oils.

AUTHORS : Julien CREPIER[(a)], Agnès LE MASLE[(a)], Nadège CHARON[(a)], Florian ALBRIEUX[(a)], Pascal DUCHENE[a], Sabine HEINISCH[(b)]

[a]*IFP Energies nouvelles, Rond-point de l'échangeur de Solaize, BP 3, 69360 Solaize, France*

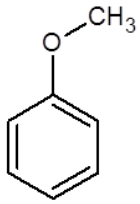
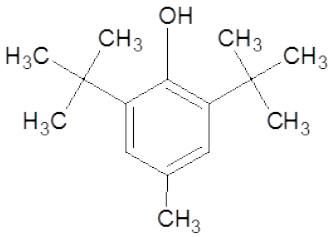
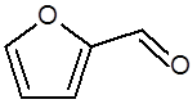
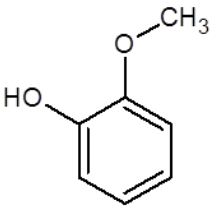
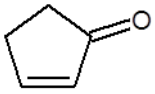
[b] *Université de Lyon, Institut des Sciences Analytiques, UMR 5280, CNRS, ENS Lyon, 5 rue de la Doua, 69100 Villeurbanne, France*

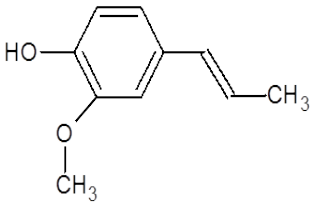
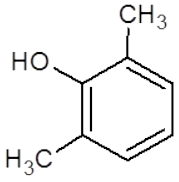
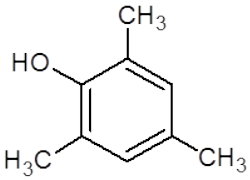
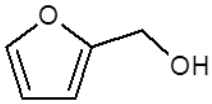
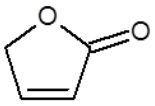
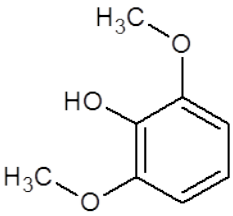
CORRESPONDENCE :

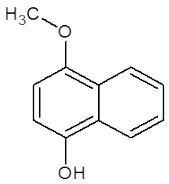
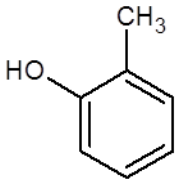
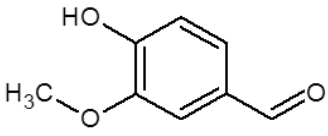
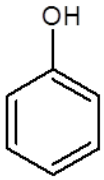
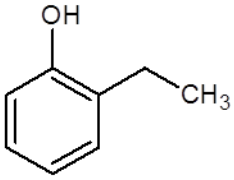
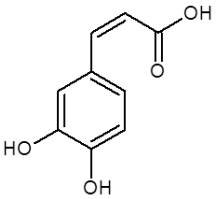
Sabine HEINISCH
E-mail : sabine.heinisch@univ-lyon1.fr
Phone.: +33 4 37 42 35 51

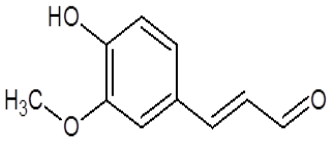
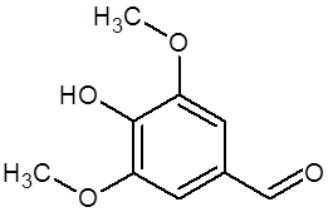
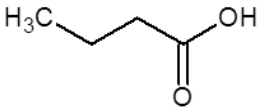
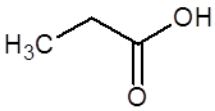
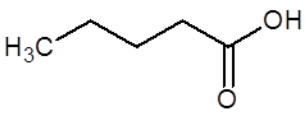
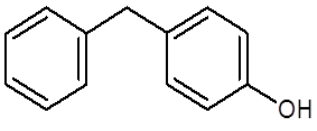
Agnès LE MASLE
E-mail : agnes.le-masle@ifpen.fr
Phone.: +33 4 37 70 23 91

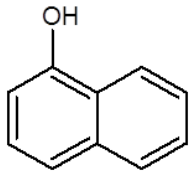
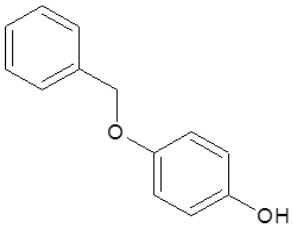
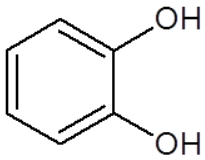
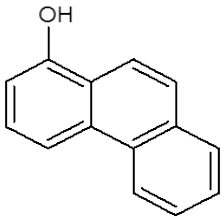
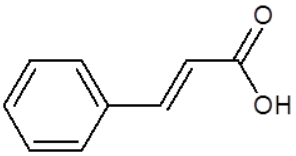
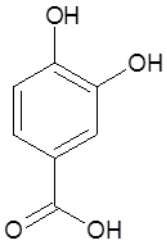
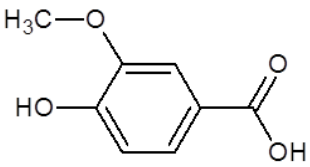
Table S1 : List of the 36 studied model molecules and their characteristics

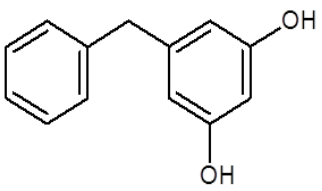
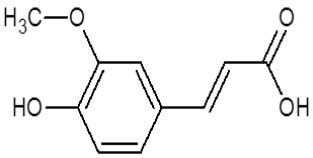
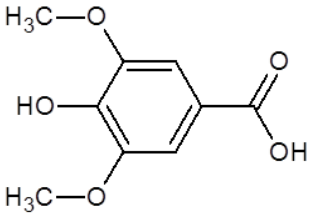
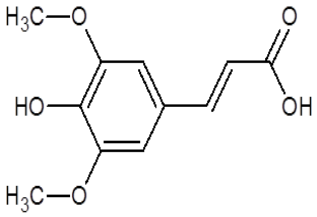
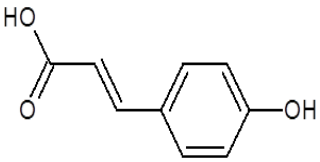
IUPAC name	**Structure**	**C**	**H**	**O**	**Accurate mass (uma)**
Furan		4	4	1	68.0262
Anisol		7	8	1	108.0575
2,6-ditertbuthyl-4methylphenol		15	24	1	220.1827
2-Furaldehyde		5	4	2	96.0211
2-Methoxyphenol		7	8	2	124.0524
2-Cyclopenten-1-one		5	6	1	82.0419

2-methoxy-4-[(E)-prop-1-enyl] phenol		10	12	2	164.0837
2,6-Dimethylphenol		8	9	1	121.0653
2,4,6-trimethylphenol		9	12	1	136.0888
2-Furylmethanol		5	6	2	98.0368
2(3H)-Furanone		4	4	2	84.0211
1,3-Dimethoxy-2-hydroxybenzene		8	10	3	154.0630

2-Methoxy-1-naphthol		11	10	2	174.0681
(4-Methylphenyl) methanol		7	8	1	108.0575
4-hydroxy-3-methoxy benzaldehyde		8	8	3	152.0473
Phenol		6	6	1	94.0419
2-Ethylphenol		8	10	1	122.0732
acid (E) 3-(3,4-dihydroxyphényl) prop-2-ènoïque		9	8	4	180.0423

(E)-3-(4-hydroxy-3-methoxyphenyl) prop-2-enal		10	10	3	178.0630
4-hydroxy-3,5-dimethoxy benzaldehyde		9	10	4	182.0579
Butanoic acid		4	8	2	88.0524
Propanoic acid		3	6	2	74.0368
Pentanoic acid		5	10	2	102.0681
4-Benzylphenol		13	12	1	184.0888

Naphtalen-1-ol		10	8	1	144.0575
4-(benzyloxy) phenol		13	12	2	200.0837
Benzene-1,2-diol		6	6	2	110.0368
9-Phenanthrenol		14	10	1	194.0732
(2E)-3-Phenylprop-2-enoic acid		9	8	2	148.0524
3,4-dihydroxybenzoic acid		7	6	4	154.0266
4-Hydroxy-3-methoxybenzoic acid		8	8	4	168.0423

4-benzylbenzene-1,3-diol		13	12	2	200.0837
(E)-3-(4-hydroxy-3-methoxy-phenyl)prop-2-enoic acid		10	10	4	194.0579
4-Hydroxy-3,5-dimethoxy benzoic acid		9	10	5	198.0528
3-(4-hydroxy-3,5-dimethoxyphenyl)prop-2-enoic acid		11	12	5	224.0685
3-(4-hydroxyphényl)-prop-2-enoic acid		9	8	3	164.0473

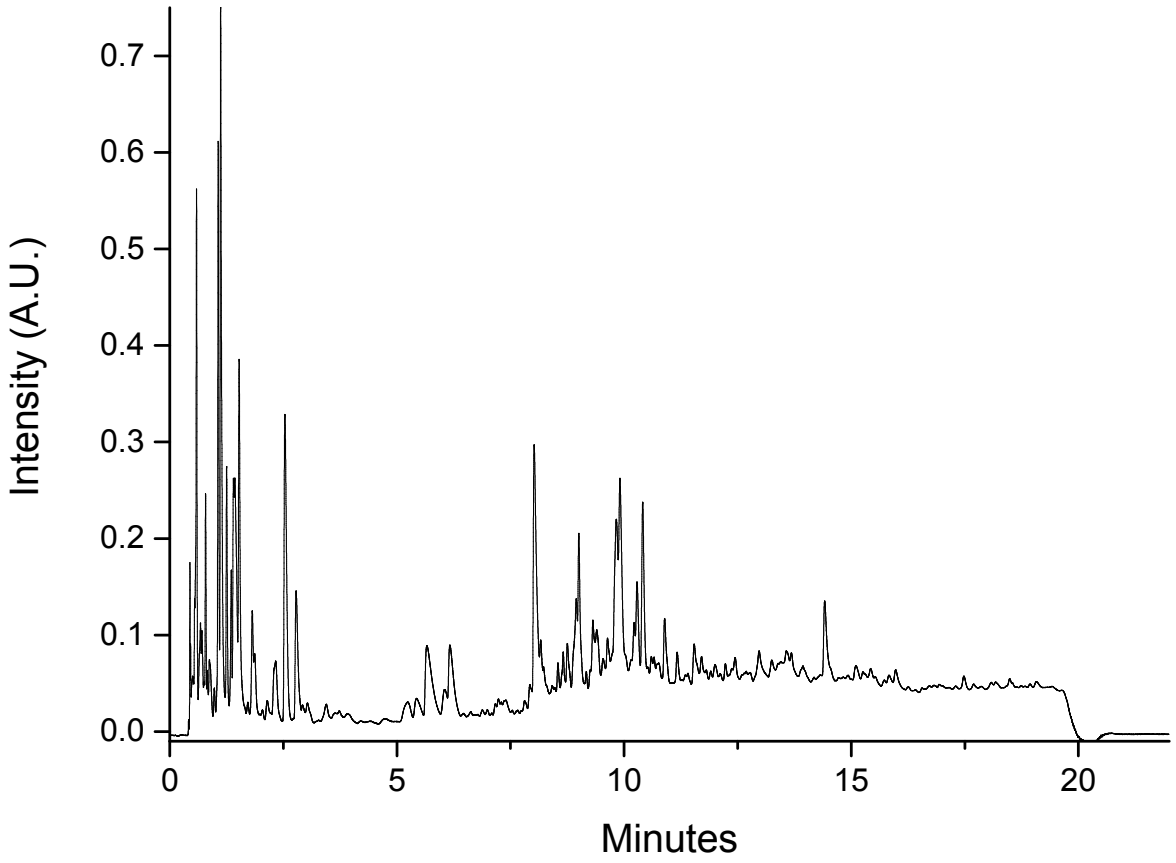


Figure S1 : SFC separation of a fast pyrolysis bio-oil under optimized conditions (stationary phase: Acquity BEH 2-EP, modifier: ACN/H_2O (98/2), temperature: 30°C, BPR pressure: 150 bar). UV detection (210nm).