



**HAL**  
open science

## Centrifugal partition chromatography as a fractionation tool for the analysis of lignocellulosic biomass products by liquid chromatography coupled to mass spectrometry

Alexis Dubuis, Agnès Le Masle, Ludovic Chahen, Emilie Destandau, Nadège Charon

### ► To cite this version:

Alexis Dubuis, Agnès Le Masle, Ludovic Chahen, Emilie Destandau, Nadège Charon. Centrifugal partition chromatography as a fractionation tool for the analysis of lignocellulosic biomass products by liquid chromatography coupled to mass spectrometry. *Journal of Chromatography A*, 2019, 1597, pp.159-166. 10.1016/j.chroma.2019.03.031 . hal-02184672

**HAL Id: hal-02184672**

**<https://ifp.hal.science/hal-02184672>**

Submitted on 16 Jul 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1                   **Centrifugal partition chromatography as a fractionation tool for the analysis of**  
2                   **lignocellulosic biomass products by liquid chromatography coupled to mass**  
3                   **spectrometry**

4  
5 Alexis Dubuis<sup>a</sup>, Agnès Le Masle<sup>a,\*</sup>, Ludovic Chahen<sup>a</sup>, Emilie Destandau<sup>b</sup>, Nadège Charon<sup>a</sup>

6  
7 <sup>a</sup> IFP Energies nouvelles, Rond-point de l'échangeur de Solaize, BP 3, 69360 Solaize, France

8 <sup>b</sup> Institut de Chimie Organique et Analytique, Université d'Orléans, CNRS UMR 7311, 2 Rue de Chartres, 45067  
9 Orléans, France

10 \*Corresponding author. E-mail address: agnes.le-masle@ifpen.fr (A. Le Masle)

11  
12  
13 **Abstract**

14 The conversion of lignocellulosic biomass into biofuels and bio-products leads to oxygenated matrices  
15 having a wide range of polarities and molecular weights. A complete analytical characterization of  
16 these complex mixtures is necessary to improve conversion processes. In this study, an innovative  
17 centrifugal partition chromatography (CPC) protocol was developed to fractionate aqueous biomass  
18 samples with a MTBE-water solvent system, by mixing elution, displacement and extrusion modes in  
19 the same run. This new protocol was validated on model molecules and applied to the water soluble  
20 phase of a fast pyrolysis bio-oil. It demonstrated a promising separation with a relevant selectivity on  
21 the most significant chemical families of biomass samples: carbohydrates, furans, carboxylic acids and  
22 phenols. CPC fractions of the sample were collected and analyzed comprehensively by HPLC-UV/MS  
23 (with ESI negative and positive ionization modes). This CPC x LC approach allowed more accurate  
24 attributions on the 217 peaks detected. The use of different detection modes gave a complete view of  
25 the water soluble phase of a fast pyrolysis bio-oil through 2D maps. Molecular characterization was  
26 enhanced by independent information: CPC retention time, LC retention time, UV and MS spectra.  
27 Concomitance of these different chemical information is of precious help for unambiguous  
28 identification.

29 *Highlights:*

- 30                   • *CPC chemical fractionation of the water soluble phase of a fast pyrolysis bio-oil*  
31                   • *Mix of elution, displacement and extrusion modes*  
32                   • *Prediction of the mobile phase pH for an optimized separation*  
33                   • *Off-line comprehensive two-dimensional liquid chromatography*  
34                   • *Multi-detection 'fingerprints' of the sample*

35 *Keywords:* lignocellulosic biomass, fast pyrolysis bio-oil, centrifugal partition chromatography, mass  
36 spectrometry, off-line comprehensive two-dimensional liquid chromatography.

37 **1. Introduction**

38 Resources alternatives to fossil fuels will be needed in the near future to diversify energy sources and  
39 meet the rising global demand. Among renewable solutions, the conversion of lignocellulosic biomass  
40 into biofuels and bio-based chemicals is a promising way under development [1–3]. The raw material

41 of biomass contains three major constituents: cellulose and hemicelluloses, respectively linear and  
42 branched polysaccharides, and the lignin made of aromatic macromolecules [4,5]. Products resulting  
43 from biochemical or thermochemical conversion processes are complex matrices with many  
44 oxygenated chemical functions like carboxylic acids, phenols, aldehydes, ketones, carbohydrates,  
45 furans or alcohols. In general, these matrices have special features: a wide range of polarities and  
46 molecular weights, acidity and heat sensitivity [6]. In this context, analytical tools have a key role to  
47 play in order to better understand the composition and reactivity of biomass products.

48 Gas chromatography (GC) and comprehensive two-dimensional gas chromatography (GC x GC) were  
49 extensively used to identify and quantify volatile oxygenated compounds [7,8]. Liquid  
50 chromatography (LC) and supercritical fluid chromatography (SFC) were used to characterize  
51 compounds of higher molecular weight or thermally unstable [9,10]. On-line comprehensive two-  
52 dimensional LC (LC x LC) was applied successfully on aqueous biomass samples to enhance the peak  
53 capacity [11,12]. These complementary separation approaches and hyphenation to high resolution  
54 mass spectrometry (MS) provided a detailed molecular characterization of biomass products.  
55 However, despite the high resolution of these systems, a large part of compounds remains unknown  
56 due to co-elutions and the complexity of MS spectrum [13].

57 A different approach consists in fractionating the sample prior to analysis in order to raise co-elutions  
58 and improve its analysis with the existing techniques. Liquid-liquid extraction (LLE) has been widely  
59 used since it allows a simple separation by solubility differences and a full recovery of the sample with  
60 no irreversible adsorption on a solid phase. Previous studies have shown the value of a controlled LLE  
61 sample preparation to facilitate the analysis [14,15]. For instance, Kanaujia *et al.* demonstrated the  
62 interest of LLE to selectively classify lignocellulosic biomass degradation products into chemical  
63 families on the base of their solubility [16]. Such fractionation is of precious help to guide the  
64 identification of unknown products, or to choose between different structures for the same chemical  
65 formula.

66 To go further on, centrifugal partition chromatography (CPC) is a liquid-liquid separation technique  
67 based on solubility differences with a much higher efficiency than LLE. A solubility fractionation by  
68 CPC seems relevant since it provides a different selectivity than LC in reversed phase. Furthermore,  
69 CPC is compatible with heat sensitive compounds (performed at room temperature) and prevents  
70 sample lost (no irreversible adsorption on a solid phase). The CPC column is constituted by hundreds  
71 of extraction cells connected in series, filled with a stationary phase retained by the application of a  
72 centrifugal force while a mobile phase is pumped through it [17,18]. CPC was initially developed for  
73 the purification of natural products and then extended to other areas of application, including  
74 biorefining products analysis [19–21]. Recently, Le Masle *et al.* demonstrated the interest of a first  
75 dimensional CPC separation applied on a biomass sample prior to analysis by reversed phase liquid  
76 chromatography (RPLC) [22]. However, even if this approach allows to raise co-elutions, the  
77 characterization still remains difficult since the resulting 2D chromatograms shows no real chemical  
78 information or organization.

79 CPC can also be used to separate mixture of organic acids or phenols according to their pH in  
80 displacement mode, better known as ‘pH zone refining’. This technique consists in controlling the pH  
81 of the phases by adding additives to the solvent system [23–26]. Since biomass samples contain  
82 ionizable organic acids and phenols, CPC in displacement mode may be an interesting way of  
83 fractionation. Hamzaoui *et al.* proposed the CPC isolation of phenolic compounds from a crude bark  
84 extract using an interesting mix of elution, displacement and extrusion modes [27]. However the pH  
85 zone refining uses additives like salts at relatively high concentrations, which is detrimental for MS

86 analysis [28,29]. Moreover this mode is usually dedicated to products belonging to the same chemical  
87 class with close pKa and fractionation experiments are very time and solvent consuming [30].

88 In the present study, we propose to use CPC as a first dimension to fractionate any aqueous biomass  
89 sample composed of both neutral and ionizable compounds using an innovative operating combination  
90 of elution, displacement and extrusion modes. The objective is to develop a CPC protocol able to  
91 produce fractions representative of oxygenated chemical class and analyzable by RPLC/MS. The  
92 proposed method allows especially the fractionation of ionizable species without adding additives  
93 commonly used in conventional pH zone refining. This new CPC fractionation approach hyphenated  
94 to LC/MS was applied to the water soluble phase of a fast pyrolysis bio-oil.

## 95 **2. Experimental**

### 96 *2.1. Materials*

97 Methyltertbutylether (MTBE), methanol, acetonitrile were HPLC grade purchased from VWR  
98 (Fontenay sous Bois, France). Sodium hydroxide solution (NaOH, 0.5 mol/L) was obtained from  
99 Carlo-Erba reagent (Val de Reuil, France). Deionized water was produced by a Milli-Q water purifier  
100 (Millipore SAS, Molsheim, France). Formic acid, acetic acid, 5-(hydroxymethyl)furan-2-carbaldehyde  
101 (5-HMF), (2R,3S,4S,5R,6S)-2-(hydroxymethyl)-6-[2-(hydroxymethyl)phenoxy]oxane-3, 4,5-triol  
102 (salicin), (E)-3-(4-hydroxy-3-methoxy-phenyl)prop-2-enoic acid (ferulic acid), 2-methoxyphenol  
103 (guaiacol) and (2E)-3-phenylprop-2-enal (cinnamaldehyde) were purchased from Sigma-Aldrich  
104 (Saint-Quentin-Fallavier, France). 5-HMF, salicin, ferulic acid, guaiacol and cinnamaldehyde were  
105 used to prepare a model mix at 5 g/L in methanol. This solution was then diluted to 250 mg/L in  
106 deionized water to prepare the model mix 1 and 2 (designated by MM1 and MM2). Acetic acid was  
107 added to MM2 at 3.2 g/L.

108 The sample used in this study was the water soluble phase of a fast pyrolysis bio-oil from softwood  
109 sawdust provided by IFP Energies nouvelles (Solaize, France). The sample was filtered prior to CPC  
110 separation using INTERCHIM PTFE syringe filters (UPTIDISC, pore size: 0.20  $\mu\text{m}$ , diameter: 4 mm,  
111 Montluçon, France).

### 112 *2.2. CPC instrumentation*

113 CPC experiments were performed on a SCPC100 associated to a Spot Prep II from Armen Instrument,  
114 France (Gilson Purification, USA). The column used has an exact volume of 131 mL and the  
115 apparatus a total of 20 mL of dead volume. The system was equipped with a quaternary pump, an  
116 automatic sample injection valve with a 5 mL sample loop, a diode array detector (acquisition  
117 wavelength range: 200 – 400 nm) and a fraction collector. All data were managed using the Armen  
118 Glider CPC Control Software. Experiments were conducted at room temperature ( $20 \pm 2^\circ\text{C}$ ).

### 119 *2.3. Preparation of solvent system*

120 The biphasic solvent system was prepared by mixing in a separatory funnel MTBE and water with a  
121 50 : 50 volumetric ratio. The funnel was agitated, left 5 minutes for equilibration and the two phases  
122 were separated. The aqueous phase was selected as the mobile phase and the organic phase as the  
123 stationary phase (descending mode). The organic phase was kept for the separation (A). The pH of the  
124 mobile phase was adjusted by adding sodium hydroxide solution (0.5 mol/L) to the aqueous phase  
125 under the control of a combined pH electrode (Fisher Scientific, Illkirch, France). Three bottles of  
126 mobile phase were prepared: the first without adding sodium hydroxide to the aqueous phase (B), the  
127 second at the pH predicted by the model (C) and the third at pH 12 (D).

128        2.4. *CPC fractionation procedure*

129        CPC method conditions are described in Table 1.

130        The stationary phase ratio was measured at 77 %. UV signal was monitored at 254 nm. Fractions were  
131        collected every minute during the 45 minutes of the method and numbered consequently from 1 to 45.  
132        A rinsing method was used in ascending mode to neutralize residual hydroxide ions from the column.  
133        A pH between 7 and 8 was expected at the fraction collector by the end of the rinsing step.

134        2.5. *Determination of partition coefficient*

135        The partition coefficient  $K$  corresponds to the ratio of the solute concentrations  $[S]$  in the organic and  
136        in the aqueous phases (Equation 1):

137        
$$K = \frac{[S]_{org}}{[S]_{aq}} \text{ (Equation 1)}$$

138         $K$  was measured on model molecules in the MTBE-water system. The biphasic system was prepared  
139        as described in part 2.3. 5 mL of each phase and approximately 100 mg of standard were added to a 20  
140        mL flask. After vigorous shaking and 5 minutes of equilibration, 1 mL of each phase was transferred  
141        into vials. The upper phase was evaporated under nitrogen flux and 1 mL of methanol was added to  
142        the vials. Upper and lower phase were analyzed by HPLC-UV (method described in part 2.7).

143        In this study, the constant used to define the partition equilibrium of acid on its HA form between  
144        organic and aqueous phases was defined as  $K_D$  (Equation 2):

145        
$$K_D = \frac{[HA]_{org}}{[HA]_{aq}} \text{ (Equation 2)}$$

146        The pH of the aqueous phase was measured to determine the proportion of neutral and ionic species  
147        for  $K_D$  determination.

148        2.6. *Off-line comprehensive two-dimensional CPC x LC analysis*

149        A total of 45 CPC fractions were collected. pH was measured from fractions 1 to 35 with a combined  
150        pH electrode. All fractions were then evaporated to dryness under nitrogen and dissolved into 1 mL of  
151        methanol (to the same dilution factor for elution and extrusion). Fractions 6 and 7 were filtered using  
152        INTERCHIM PTFE syringe filters because of the presence of some insoluble particles (UPTIDISC,  
153        pore size: 0.20  $\mu\text{m}$ , diameter: 4 mm, Montluçon, France). The sample was prepared in the same  
154        conditions as CPC fractions for 1D LC chromatograms.

155        2.7. *High performance liquid chromatography coupled to mass spectrometry*

156        HPLC analyses were performed on a Shimadzu LC 20AD (Kyoto, Japan) composed of a binary pump,  
157        a SIL 20AC autosampler (10 °C), a CTO 20AC column oven (30 °C), a SPD M20A UV detector and a  
158        CBM 20A system controller. The injection volume was 5  $\mu\text{L}$ . Separation was achieved on a Kinetex  
159        C18 column (150 x 3 mm, 2.6  $\mu\text{m}$ , Phenomenex, Le Pecq, France) at 30 °C. Mobile phases used were  
160        water (A1) and acetonitrile (B1) both acidified with 0.025 % (v:v) of formic acid. The separation was  
161        performed at 0.6 mL/min with the following elution gradient: 0 – 6 min (100 % A1), 6 – 26 min (from  
162        0 % B1 to 40 % B1, normalized elution gradient slope of 2.2 %), 26 – 26.1 min (100 % A1) and 26.1 –  
163        32 min (100 % A1). The dwell time was measured at 1.1 min. UV signal was measured on the  
164        wavelength range 200 – 400 nm (Deuterium lamp, cell temperature of 30 °C, sampling frequency of  
165        12.5 Hz and time constant of 0.080 s). When only a UV detection signal was desired, fractions 1 to 35

166 can be directly injected by HPLC without evaporation step. MS detection was performed on a  
167 Shimadzu 2020 simple quadrupole. The flow rate was reduced to 150  $\mu\text{L}/\text{min}$  for MS detection using a  
168 post UV T-split. An Electrospray Ionization (ESI) source was used in negative and positive modes  
169 with an interface voltage of - 4.5 kV and + 4.5 kV respectively. Drying and nebulizing gas flow rates  
170 were respectively 15 L/min and 1.5 L/min. Temperatures were set to 200  $^{\circ}\text{C}$  for the heat block and 250  
171  $^{\circ}\text{C}$  for the desolvation line. The scan range was set to 100 – 800 m/z. Reproducibility of retention  
172 times and MS sensitivity were checked with standards injected regularly during sequences  
173 (Supplementary material). Data acquisition and treatment were managed with Shimadzu software  
174 Labsolution.

### 175 **3. Results and discussion**

#### 176 *3.1. LC/MS analysis of the water soluble phase of a fast pyrolysis bio-oil*

177 The aqueous biomass sample was analyzed by HPLC-UV/MS using generic conditions to have a first  
178 description of its complexity. Chromatograms obtained with UV and MS detection are presented in  
179 Figure 1.

180 The proposed reversed phase separation coupled to mass spectrometry offers a first molecular  
181 description of the sample. This direct analysis allows the detection of approximately 70 peaks with UV  
182 detection, 55 with ESI positive and 60 with ESI negative. All modes included, 139 peaks were counted  
183 manually in Figure 1, considering peaks with a signal-to-noise ratio greater than 3 (table in  
184 Supplementary material). UV and MS detections provide complementary profiles of the sample, with  
185 maxima positioned at different times of the chromatogram. However, the chromatograms present  
186 many co-elutions and are therefore difficult to process. A first dimensional separation is required to  
187 increase the peak capacity and simplify chromatograms.

#### 188 *3.2. CPC separation of the aqueous biomass sample*

##### 189 *3.2.1. Selection of the solvent system*

190 The biphasic system is essential in CPC because it represents the mobile and the stationary phases. In  
191 this study, the objective is to fractionate the entire sample rather than extracting a single compound,  
192 using elution and displacement modes. To do so, the biphasic system MTBE-water appears to be well  
193 adapted for many reasons. Firstly, this system is very stable and commonly used for both classical  
194 elution and pH zone refining applications [23,27]. Moreover, MTBE is able to extract a large number  
195 of organic species present in biomass products, including acids and phenolic compounds [31]. The  
196 organic phase is preferentially used as a stationary phase to retain ionizable chemical compounds.  
197 Additionally, the mobile phase has to be aqueous to modify the pH during the separation and ionize  
198 compounds retained at the head of the CPC column. Thus, compounds will be eluted with the aqueous  
199 mobile phase under their ionic form, depending on their pKa. This descending mode configuration is  
200 better for the system stability according to Ito [17] and is more suitable for coupling with RPLC  
201 (production of directly analyzable aqueous fractions).

##### 202 *3.2.2. Fractionation strategy*

203 In a classical CPC experiment with elution and extrusion modes, the separation relies only on the  
204 partition coefficient (K) of solutes in the biphasic system. In accordance with CPC usual  
205 recommendations, K should be within the range 0.5 – 3 in the elution zone to maximize peaks  
206 resolution and reduce peaks broadening [17]. The partition coefficient of a solute can be used to  
207 calculate the retention time, knowing CPC operating parameters (column volume, stationary phase  
208 retention ratio, flow rate) [18]. In this way, common chemical families present in biomass samples can

209 be positioned on a CPC chromatogram. Thus, in classical CPC elution/extrusion modes in descending  
210 mode, carbohydrates and their derivatives are eluted in the first fractions, close to the dead volume  
211 because of their high affinity to water. Then, furans and other small polar compounds are eluted ( $0.5 <$   
212  $K < 1.5$ ). For these chemical families, the elution mode allows an excellent fractionation. However,  
213 the less polar compounds like carboxylic acids, phenols or aldehydes or hybrids compounds are eluted,  
214 with more dispersed peaks in the rest of the elution ( $1.5 < K < 3$ ), and mainly in the extrusion fractions  
215 ( $K > 3$ ) without clear chemical fractionation. The proposed strategy in Figure 2 consists in adding a  
216 displacement mode to force ionizable species like carboxylic acids and phenols to elute, in a limited  
217 zone of elution, after the furans and before the extrusion mode.

218 The order of separation of the displaced compounds should mainly follow their pKa. It is important to  
219 notice that, for the rest of the publication, ‘hybrid’ species like phenolic carboxylic acids are  
220 considered as carboxylic acids since their lower pKa corresponds to the carboxylic acid function.

221 In this strategy, the selectivity between carboxylic acids and phenols has to be optimized and the key  
222 factor is the pH of the mobile phase at the intermediate step. This pH needs to be high enough to  
223 neutralize the total sample acidity and consequently, to allow the complete elution of carboxylic acids.  
224 However, if the chosen pH is too high, phenol will be ionized also and may co-elute with carboxylic  
225 acids, losing the interest of an intermediate step. Finally, the intermediate pH depends of the total  
226 acidity of the sample and has to be calculate prior to the CPC protocol.

### 227 3.2.3. Modelling approach

228 In this section are presented the main hypothesis considered to predict the elution of ionizable species  
229 during the displacement mode. The resulting model allows to estimate the sodium hydroxide quantity  
230 necessary to elute all acidic compounds of the sample within a limited time of elution. The  
231 corresponding operating parameter is the intermediate pH of the mobile phase, manually adjusted  
232 using a combined pH electrode. An exhaustive study of the phenomena at the liquid-liquid interface in  
233 pH zone refining was proposed by Kotland *et al.* [24]. Herein, the model developed does not take into  
234 account all hydrodynamic aspects involved. For the calculations, the CPC column was assimilated to a  
235 succession of LLE cells connected in series, considering a limited number of 100 theoretical plates,  
236 enough to have an accurate prediction of species behavior using the displacement mode. Each cell can  
237 be represented by the situation described in Figure 3.

238 In Figure 3, all ionizable species ( $3 < pKa < 11$ ) are symbolized by HA for the acidic form and A<sup>-</sup> for  
239 the conjugate base. The A<sup>-</sup> form is considered as totally insoluble in the organic phase. Regarding the  
240 pKa of carboxylic acids ( $3 < pKa < 4$ ) and considering that the pH intermediate would be certainly  
241 superior to 7, the reaction in the aqueous phase between the hydroxide ion HO<sup>-</sup> and HA is supposed  
242 total, with a constant  $K_{aq}$  (Equation 3):

$$243 K_{aq} = \frac{[A^-]_{aq}}{[HA]_{aq} * [HO^-]_{aq}} \text{ (Equation 3)}$$

244 The equilibrium defined by the  $K_D$  constant is displaced by the consumption of HA<sub>aq</sub> which is then  
245 transferred from cell to cell in its A<sup>-</sup><sub>aq</sub> form. Before starting this transfer of HA from the organic to the  
246 aqueous phase, the complete acidity of the sample introduced in the first cell needs to be neutralized.  
247 Therefore, a titration of the total amount of acid is necessary. However, small organic acids (formic  
248 acid, acetic acid) which are common degradation products in biomass samples do not require  
249 neutralization since they are no longer present in the CPC column during the intermediate step ( $K <$   
250  $1.5$ ). Consequently, these acids should be quantified independently and subtracted from the total  
251 acidity. The acid concentration obtained  $[HA]_{aq}$  is then implemented in the model.

252 To calculate the quantity of solute transferred, others CPC parameters are needed in the model: the  
253 stationary phase retention ratio, the column volume, the injection volume, and the flow rate. The  $K_D$   
254 was determined for the biphasic system MTBE-water using the LC-UV method described in section  
255 2.7 within the range 0.6 – 940 (respectively for the pyruvic and the ferulic acids, commonly present in  
256 aqueous biomass samples). By fixing all others parameters, the two limits  $K_D = 0.5$  and  $K_D = 2000$   
257 were evaluated in the model, and the  $K_D$  demonstrated a limited influence on the peak width. The  $K_D$   
258 value of the ferulic acid was considered by default for the model. In practice, the model uses a  
259 spreadsheet and the results are displayed with graphical representation of concentration profiles of  
260 each specie as a function of the time, at different theoretical plates (corresponding to different  
261 positions in the column). The elution of acidic compounds can be monitored with species A- at the  
262 final theoretical plate. The pH of the mobile phase can be finely adjusted by the user to control the  
263 width of the peak. In this study, a peak width of 10 minutes was fixed as the optimal response for the  
264 elution of acids, corresponding to the duration of the intermediate step.

### 265 3.3. Validation of the fractionation protocol

#### 266 3.3.1. Model molecules

267 The feasibility of the CPC protocol was validated on model molecules representative of the targeted  
268 chemical families. In biomass samples, there are numerous carboxylic acids species and they can be  
269 present in various concentrations. The acetic acid is one of the main degradation product of  
270 lignocellulosic biomass and its concentration is often largely superior to other carboxylic acids. In the  
271 water soluble phase of a pyrolysis bio-oil, the acetic acid was quantified at 3.4 g/L using a LC-UV  
272 reference method at 210 nm [32]. Acetic acid should be eluted before the displacement zone because  
273 of its high hydrophilicity ( $K < 1$ ) and therefore should not consume hydroxides ions of the mobile  
274 phase. However, it was important to measure the influence of high concentration of acetic acid on the  
275 separation of other carboxylic acids and phenols. Two mixes containing both 5 model molecules were  
276 prepared: one without acetic acid (MM1) and one with 3.2 g/L of acetic acid (MM2). A pH of 9.8 was  
277 predicted for the mobile phase to elute ferulic acid of MM1 and MM2 in 10 minutes. CPC  
278 chromatograms obtained for MM1 and MM2 are presented in Figure 4.

279 Salicin (1), hydroxymethylfurfural (2), guaiacol (4) and cinnamaldehyde (5) are eluted at the same  
280 position of chromatograms in Figure 4. Salicin is a glucoside (carbohydrate derivative) with a high  
281 solubility in water, this explains its elution within the dead volume (6 – 7 min).  
282 Hydroxymethylfurfural retention time (9 min) is consistent with its partition coefficient ( $K = 0.4$ ). The  
283 guaiacol is displaced at 30 minutes in the phenol zone of both chromatograms. The signal observed  
284 between 35 and 37 minutes is a disturbance due to the elution of the stationary phase during extrusion.  
285 The cinnamaldehyde is eluted at 40 minutes in the extrusion in accordance with its physicochemical  
286 properties ( $K > 3$ ,  $pK_a > 14$ ). Still, differences can be noticed between the two model mix  
287 chromatograms. The elution of acetic acid in MM2 results as anticipated in a pH drop of the aqueous  
288 mobile phase between 9 and 16 minutes. Acetic acid and hydroxymethylfurfural are co-eluted on the  
289 210 nm CPC chromatograms provided in supplementary material. The peak of ferulic acid in MM2 is  
290 delayed and broadened compared to MM1. Thus, displacement of the ferulic acid is less efficient for  
291 MM2 because hydroxides of the mobile phase are consumed by the remaining acetic acid in the  
292 column. The peak width of 10 minutes predicted was correct for the ferulic acid, but the risk is to elute  
293 a part of acids on phenols elution window. For this reason, the pH of the mobile phase predicted by the  
294 model should be increased of 0.1 – 0.2 point, to reduce acids peak width and prevent additional acidity  
295 coming from unknown acids of the sample in the elution zone. Acetic acid has also a positive effect on  
296 the separation by retaining more efficiently acids in CPC column head. Thereby, it avoids the elution  
297 of a small proportion of acids within the dead volume, as for MM1 in Figure 4 (see the baseline signal



298 between peaks 1 and 2). The high acid content in biomass samples acts like an acid buffer for CPC  
299 first cells, leading to the situation observed for MM2. These CPC tests realized on model molecules  
300 confirmed the good chemical selectivity of the protocol.

### 301 *3.3.2. Water soluble phase of a fast pyrolysis bio-oil*

302 The first dimensional CPC separation was then applied on the water soluble phase of a fast pyrolysis  
303 bio-oil. First, the complex sample was fractionated by CPC using the classical elution – extrusion  
304 mode without pH modification of the mobile phase (chromatogram available in supplementary  
305 material) to evaluate the separation of the different solute families only based on their partition  
306 coefficient K. A promising separation was obtained by elution for the first part of the chromatogram (0  
307 – 15 min) and a band was recovered in the extrusion. In these highly retained fractions some  
308 molecules carry ionizable functions (carboxylic acids, phenols), justifying the use of the displacement  
309 mode to elute them in the empty part of the chromatogram (15 – 35 min). Thus, the acidity of the  
310 sample was estimate at 4.4 mmol/L. Using the predictive model, the pH of the mobile phase should be  
311 10.4 to elute acids during the intermediate step. According to the model, a pH of 10 is not enough to  
312 elute acids and a pH of 11 reduces the width of the acid representative peak to 3 minutes, which may  
313 lead to co-elutions. CPC fractionations at pH 10, 10.5 and 11 for the intermediate step were  
314 investigated to assess the quality of these predictions. An additional experiment was done without  
315 intermediate step (equivalent to a direct step at pH 12). The aim was to find the practical best  
316 intermediate pH for this sample. Corresponding CPC chromatograms are given in Figure 5.

317 CPC chromatograms are well overlaid in the elution zone (0 – 17 min), showing the repeatability of  
318 the elution fractionation. pH measurements in the fractions confirmed the presence of acids eluted by  
319 their partition coefficients. Two drops in pH are observed: the first may be the formic acid (5 – 7 min,  
320 not retained with MTBE-water solvent system) and the second is the acetic acid (9 – 17min, as for the  
321 model mix experiments). In the displacement zone (17 – 27 min), chromatograms are very different  
322 depending on the intermediate pH used. This demonstrates the interest of pH modifications and the  
323 key role of the intermediate step. Best separations are obtained with the intermediate pH of 10.5 and  
324 11 as a band appears between 17 and 27 minutes. The tests realized at pH 10 and without intermediate  
325 step (directly pH 12) confirm the existence of an optimal pH. In fact, with an intermediate pH of 10, as  
326 predicted by the model, acidic compounds are not eluted in the band (17 – 27 min) but the peak at 28  
327 minutes is more important than those at pH 10.5 and 11. The reason is an insufficient amount of  
328 hydroxide ions supplied by the mobile phase to neutralize the acidity of the sample in the first cells of  
329 the CPC, which delays the displacement of all ionizable compounds to the pH 12 step. Though, the  
330 opposite effect is observed with a direct step at pH 12 (experiment without intermediate step), as the  
331 peak at 27 minutes is much higher, because it contains the majority of ionizable products. The  
332 experiment without intermediate step shows the presence of a peak at 20 minutes eluted in the  
333 displacement zone with its partition coefficient (peak more dispersed). In the phenols zone, a peak  
334 constituted of a less ionizable chemical family is present in all chromatograms after the main peak  
335 (between 29 and 32 minutes). No UV peak is detected in the extrusion, so pH 12 was enough to elute  
336 the majority of ionizable products retained in the stationary phase. The CPC fractionation was applied  
337 successfully on the water soluble phase of a pyrolysis bio-oil. The intermediate pH of 10.5 provides  
338 the best separation considering the band spreading of the intermediate step, in accordance with model  
339 prediction. Therefore, pH 10.5 fractions were selected to perform the second dimensional separation  
340 by RPLC-UV/MS.

### 341 *3.4. Off-line comprehensive two-dimensional CPC x LC analysis using UV and MS detections*

342 To go further in the molecular characterization of the sample, a second LC separation was performed  
343 comprehensively on CPC fractions. This two-dimensional approach benefits from the chemical  
344 organization of CPC and the high resolving power of HPLC, which has already demonstrated a  
345 promising separation on the initial sample (Figure 1). The limited use of additive with the  
346 displacement mode allows coupling with mass spectrometry. The objective of this discussion is to  
347 show the interest of using the CPC protocol as a first dimension in an off-line coupling with HPLC-  
348 UV/MS, taking advantage of the complementarity of the separation modes and three different  
349 detections: UV, MS (ESI negative, ESI positive). An evaporation step was performed on CPC  
350 fractions to improve MS detected signals. The 2D contour plots obtained are presented in Figure 6.

351 2D maps were represented using a home-made software also used to detect all local maxima with a  
352 minimum intensity equal to 3 times the signal to noise ratio. A total of 217 peaks are reported  
353 combining the three detection modes after manual checking to remove integrations due to the LC  
354 elution gradient (vertical lines in Figure 6). In comparison with the 139 peaks detected by 1D LC, this  
355 high number of detected peaks demonstrates the need for a first dimensional CPC separation and its  
356 complementarity with HPLC. Early CPC fractions (5 – 7 min) have a band in the UV map which  
357 remains difficult to process like in the 1D LC chromatogram. However, the elution of this highly  
358 water-soluble fraction (certainly mainly carbohydrates derivatives) greatly improved the rest of the  
359 map. In this way, many co-elutions were raised with the CPC separation, especially in the  
360 displacement zone. Some peaks were detected by MS in the extrusion, which could not be anticipated  
361 from the CPC-UV chromatogram (Figure 5). The absence of UV response in the range 200 – 400 nm  
362 reduces possible propositions for the chemical family of these unknown products (terpenoids, fatty  
363 acids, sterols, cutins...). 2D maps simplify the comparison of the three detections used, providing  
364 strong arguments for identification. On the 217 molecules, 112 were detected by UV, 99 by ESI  
365 negative and 70 by ESI positive (table of comparison provided in the Supplementary material). The  
366 multi-detection proposed provide additional arguments for qualitative analysis. This chemical  
367 information along with the CPC retention time are helpful for structural elucidation. An identification  
368 was proposed when molecules were detected with at least two modes and after a comparison with UV  
369 and MS standards spectrum (Table 2).

370 Molecules identified are mainly phenolic compounds from the displacement zone. The  
371 physicochemical properties of these products were compared to their CPC retention times to assess the  
372 chemical selectivity of the CPC fractionation. 5-HMF is logically eluted in the elution zone (11 min)  
373 in accordance with its negative log P (-0.1). Phenolic acids with a pKa of approximately 4 are eluted in  
374 fractions 21 and 22. In these first fractions, the separation is clearly driven by the acid function (pKa).  
375 From fractions 23 to 26, products are also mainly eluted in the ascending order of their pKa. However,  
376 hydrophobicity plays a major role on the separation of phenols which are not carrying an acid  
377 function. For instance, syringaldehyde (pKa = 7.2) is eluted in fraction 22 probably because of its low  
378 hydrophobicity (log P = 1.1). Peaks more dispersed on the CPC axis of the 2D map can be assigned to  
379 compounds eluted by their partition coefficient rather than displacement. 3,4-dihydroxybenzaldehyde  
380 (log P = 1.1) is the best example as it is eluted from fractions 16 to 24 on the UV map of Figure 6 (also  
381 visible on the CPC chromatogram of Figure 5). The elution order of phenols combined the properties  
382 pKa - log P and is therefore more difficult to understand than acids. These few products confirm the  
383 interest of the CPC protocol to enhance identification of a complex sample. The 2D approach  
384 combined with different detection modes is a powerful tool, providing complementary fingerprints of  
385 samples.

#### 386 4. Conclusion

387 A CPC protocol was developed to fractionate an aqueous biomass sample by chemical family with a  
388 MTBE-water solvent system, by mixing elution, displacement and extrusion modes. A modeling  
389 approach was used to determine the optimal pH for the aqueous mobile phase in the displacement  
390 zone. Protocol and model were validated on model molecules and applied to the water soluble phase  
391 of a fast pyrolysis bio-oil. The first dimensional CPC demonstrated a promising separation with a  
392 relevant selectivity on the most significant chemicals families of biomass samples: carbohydrates,  
393 furans, carboxylic acids and phenols. CPC fractions of the complex sample were collected and  
394 analyzed comprehensively by LC-UV/MS (ESI negative and positive). This two-dimensional  
395 approach allowed the detection of 217 peaks with more accurate attributions. The use of different  
396 detection modes gives a complete view of the sample through 2D contour plots. Molecular  
397 characterization is enhanced by CPC retention time, LC retention time, UV and MS spectra. 2D maps  
398 representations by chemical classes can also be used as fingerprints to compare samples coming from  
399 various feedstocks or processes.

## 400 **Acknowledgments**

401 The authors thank Pascal Duchene and Damien Leinekugel Le Cocq from IFP Energies nouvelles for  
402 their help with 2D data management and modelling calculations.

## 403 **5. References**

- 404 [1] A.K. Chandel, V.K. Garlapati, A.K. Singh, F.A.F. Antunes, S.S. da Silva, The path forward for  
405 lignocellulose biorefineries: Bottlenecks, solutions, and perspective on commercialization,  
406 *Bioresour. Technol.* 264 (2018) 370–381. <https://doi.org/10.1016/j.biortech.2018.06.004>.
- 407 [2] P.R. Seidl, A.K. Goulart, Pretreatment processes for lignocellulosic biomass conversion to  
408 biofuels and bioproducts, *Current Opinion in Green and Sustainable Chemistry* 2 (2016) 48–53.  
409 <https://doi.org/10.1016/j.cogsc.2016.09.003>.
- 410 [3] P. Gallezot, Conversion of biomass to selected chemical products, *Chem. Soc. Rev.* 41 (4) (2012)  
411 1538–1558. <https://doi.org/10.1039/c1cs15147a>.
- 412 [4] A. Brandt, J. Gräsvik, J.P. Hallett, T. Welton, Deconstruction of lignocellulosic biomass with  
413 ionic liquids, *Green Chem.* 15 (3) (2013) 550. <https://doi.org/10.1039/c2gc36364j>.
- 414 [5] P. McKendry, Energy production from biomass (part 1): overview of biomass, *Bioresour.*  
415 *Technol.* 83 (2002) 37–46.
- 416 [6] M. Staš, D. Kubička, J. Chudoba, M. Pospíšil, Overview of Analytical Methods Used for  
417 Chemical Characterization of Pyrolysis Bio-oil, *Energy Fuels* 28 (1) (2013) 385–402.  
418 <https://doi.org/10.1021/ef402047y>.
- 419 [7] P.K. Kanaujia, Y.K. Sharma, M.O. Garg, D. Tripathi, R. Singh, Review of analytical strategies in  
420 the production and upgrading of bio-oils derived from lignocellulosic biomass, *Journal of*  
421 *Analytical and Applied Pyrolysis* 105 (2014) 55–74. <https://doi.org/10.1016/j.jaap.2013.10.004>.
- 422 [8] B. Du, L.N. Sharma, C. Becker, S.-F. Chen, R.A. Mowery, G.P. van Walsum, C.K. Chambliss,  
423 Effect of varying feedstock-pretreatment chemistry combinations on the formation and  
424 accumulation of potentially inhibitory degradation products in biomass hydrolysates, *Biotechnol.*  
425 *Bioeng.* 107 (3) (2010) 430–440. <https://doi.org/10.1002/bit.22829>.
- 426 [9] J. Crepier, A. Le Masle, N. Charon, F. Albrieux, S. Heinisch, Development of a supercritical  
427 fluid chromatography method with ultraviolet and mass spectrometry detection for the  
428 characterization of biomass fast pyrolysis bio oils, *J. Chromatogr. A* 1510 (2017) 73–81.  
429 <https://doi.org/10.1016/j.chroma.2017.06.003>.
- 430 [10] T.M. Jarrell, C.L. Marcum, H. Sheng, B.C. Owen, C.J. O'Lenick, H. Maraun, J.J. Bozell, H.I.  
431 Kenttämaa, Characterization of organosolv switchgrass lignin by using high performance liquid

- 432 chromatography/high resolution tandem mass spectrometry using hydroxide-doped negative-ion  
433 mode electrospray ionization, *Green Chem* 16 (5) (2014) 2713–2727.  
434 <https://doi.org/10.1039/C3GC42355G>.
- 435 [11] A. Le Masle, D. Angot, C. Gouin, A. D'Attoma, J. Ponthus, A. Quignard, S. Heinisch,  
436 Development of on-line comprehensive two-dimensional liquid chromatography method for the  
437 separation of biomass compounds, *J. Chromatogr. A* 1340 (2014) 90–98.  
438 <https://doi.org/10.1016/j.chroma.2014.03.020>.
- 439 [12] D. Tomasini, F. Cacciola, F. Rigano, D. Sciarrone, P. Donato, M. Beccaria, E.B. Caramão, P.  
440 Dugo, L. Mondello, Complementary analytical liquid chromatography methods for the  
441 characterization of aqueous phase from pyrolysis of lignocellulosic biomasses, *Anal. Chem.* 86  
442 (22) (2014) 11255–11262. <https://doi.org/10.1021/ac5038957>.
- 443 [13] M. Staš, J. Chudoba, D. Kubička, J. Blažek, M. Pospíšil, Petroleomic Characterization of  
444 Pyrolysis Bio-oils: A Review, *Energy Fuels* 31 (10) (2017) 10283–10299.  
445 <https://doi.org/10.1021/acs.energyfuels.7b00826>.
- 446 [14] Y. Wei, H. Lei, L. Wang, L. Zhu, X. Zhang, Y. Liu, S. Chen, B. Ahring, Liquid–Liquid  
447 Extraction of Biomass Pyrolysis Bio-oil, *Energy Fuels* 28 (2) (2014) 1207–1212.  
448 <https://doi.org/10.1021/ef402490s>.
- 449 [15] S. Ren, X.P. Ye, A.P. Borole, Separation of chemical groups from bio-oil water-extract via  
450 sequential organic solvent extraction, *Journal of Analytical and Applied Pyrolysis* 123 (2017)  
451 30–39. <https://doi.org/10.1016/j.jaap.2017.01.004>.
- 452 [16] P.K. Kanaujia, D.V. Naik, D. Tripathi, R. Singh, M.K. Poddar, L.S.K. Konathala, Y.K. Sharma,  
453 Pyrolysis of *Jatropha Curcas* seed cake followed by optimization of liquid–liquid extraction  
454 procedure for the obtained bio-oil, *Journal of Analytical and Applied Pyrolysis* 118 (2016) 202–  
455 224. <https://doi.org/10.1016/j.jaap.2016.02.005>.
- 456 [17] Y. Ito, Golden rules and pitfalls in selecting optimum conditions for high-speed counter-current  
457 chromatography, *Journal of Chromatography A* 1065 (2) (2005) 145–168.  
458 <https://doi.org/10.1016/j.chroma.2004.12.044>.
- 459 [18] A. Berthod, K. Faure, Separations with a Liquid Stationary Phase: Countercurrent  
460 Chromatography or Centrifugal Partition Chromatography, *Analytical Separation Science* (2015)  
461 1177–1206. <https://doi.org/10.1002/9783527678129.assep046>.
- 462 [19] D.P. Ward, M. Cárdenas-Fernández, P. Hewitson, S. Ignatova, G.J. Lye, Centrifugal partition  
463 chromatography in a biorefinery context: Separation of monosaccharides from hydrolysed sugar  
464 beet pulp, *J. Chromatogr. A* 1411 (2015) 84–91. <https://doi.org/10.1016/j.chroma.2015.08.006>.
- 465 [20] M. Bojczuk, D. Żyżelewicz, P. Hodurek, Centrifugal partition chromatography - A review of  
466 recent applications and some classic references, *J. Sep. Sci.* 40 (7) (2017) 1597–1609.  
467 <https://doi.org/10.1002/jssc.201601221>.
- 468 [21] C.-S. Lau, K.A. Bunnell, E.C. Clausen, G.J. Thoma, J.O. Lay, J. Gidden, D.J. Carrier, Separation  
469 and purification of xylose oligomers using centrifugal partition chromatography, *J. Ind.  
470 Microbiol. Biotechnol.* 38 (2) (2011) 363–370. <https://doi.org/10.1007/s10295-010-0799-1>.
- 471 [22] A. Le Masle, S. Santin, L. Marlot, L. Chahen, N. Charon, Centrifugal partition chromatography a  
472 first dimension for biomass fast pyrolysis oil analysis, *Analytica Chimica Acta* (2018).  
473 <https://doi.org/10.1016/j.aca.2018.04.040>.
- 474 [23] Y. Ito, pH-zone-refining counter-current chromatography: origin, mechanism, procedure and  
475 applications, *J. Chromatogr. A* 1271 (1) (2013) 71–85.  
476 <https://doi.org/10.1016/j.chroma.2012.11.024>.
- 477 [24] A. Kotland, S. Chollet, J.-M. Autret, C. Diard, L. Marchal, J.-H. Renault, Modeling pH-zone  
478 refining countercurrent chromatography: a dynamic approach, *J. Chromatogr. A* 1391 (2015) 80–  
479 87. <https://doi.org/10.1016/j.chroma.2015.03.005>.

- 480 [25] A. Kotland, S. Chollet, C. Diard, J.-M. Autret, J. Meucci, J.-H. Renault, L. Marchal, Industrial  
481 case study on alkaloids purification by pH-zone refining centrifugal partition chromatography, *J.*  
482 *Chromatogr. A* 1474 (2016) 59–70. <https://doi.org/10.1016/j.chroma.2016.10.039>.
- 483 [26] M. Hamzaoui, J. Hubert, R. Reynaud, L. Marchal, A. Foucault, J.-H. Renault, Strong ion  
484 exchange in centrifugal partition extraction (SIX-CPE): effect of partition cell design and  
485 dimensions on purification process efficiency, *J. Chromatogr. A* 1247 (2012) 18–25.  
486 <https://doi.org/10.1016/j.chroma.2012.05.046>.
- 487 [27] M. Hamzaoui, J.-H. Renault, R. Reynaud, J. Hubert, Centrifugal partition extraction in the pH-  
488 zone-refining displacement mode: an efficient strategy for the screening and isolation of  
489 biologically active phenolic compounds, *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.*  
490 937 (2013) 7–12. <https://doi.org/10.1016/j.jchromb.2013.07.024>.
- 491 [28] M.C. García, The effect of the mobile phase additives on sensitivity in the analysis of peptides  
492 and proteins by high-performance liquid chromatography-electrospray mass spectrometry, *J.*  
493 *Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 825 (2) (2005) 111–123.  
494 <https://doi.org/10.1016/j.jchromb.2005.03.041>.
- 495 [29] N.B. Cech, C.G. Enke, Practical implications of some recent studies in electrospray ionization  
496 fundamentals, *Mass Spectrom. Rev.* 20 (6) (2001) 362–387. <https://doi.org/10.1002/mas.10008>.
- 497 [30] G. Dong, J. Xu, Y. Gu, Y. Wei, A general separation method of phenolic acids using pH-zone-  
498 refining counter-current chromatography and its application to oat bran, *J. Chromatogr. B Analyt.*  
499 *Technol. Biomed. Life Sci.* 992 (2015) 36–42. <https://doi.org/10.1016/j.jchromb.2015.04.024>.
- 500 [31] S.-F. Chen, R.A. Mowery, V.A. Castleberry, G.P. van Walsum, C.K. Chambliss, High-  
501 performance liquid chromatography method for simultaneous determination of aliphatic acid,  
502 aromatic acid and neutral degradation products in biomass pretreatment hydrolysates, *J.*  
503 *Chromatogr. A* 1104 (1-2) (2006) 54–61. <https://doi.org/10.1016/j.chroma.2005.11.136>.
- 504 [32] A. Sluiter, B. Hames, R. Ruiz, C. Scarlata, J. Sluiter, D. Templeton, and D. Crocker,  
505 Determination of Structural Carbohydrates and Lignin in Biomass: Laboratory Analytical  
506 Procedure (LAP): NREL/TP-510-42618, National Renewable Energy Laboratory (2008).

## 507 Figure captions

509 **Figure 1:** HPLC-UV/MS chromatograms of the water soluble phase of a pyrolysis bio-oil (254 nm,  
510 base peak chromatograms).

511 **Figure 2:** CPC fractionation strategy using a mix of elution, displacement and extrusion modes.

512 **Figure 3:** Simplified representation of liquid-liquid equilibria within each CPC cell.

513 **Figure 4:** CPC-UV chromatograms of the model mix MM1 without acetic acid and MM2 with acetic  
514 acid (254nm). 1. Salicin, 2. Hydroxymethylfurfural, 3. Ferulic acid, 4. Guaiacol, 5. Cinnamaldehyde.

515 **Figure 5:** CPC-UV chromatograms of the water soluble phase of a pyrolysis bio-oil (254nm)  
516 performed at pH 10, 10.5 and 11 for the intermediate step. An experiment was performed without  
517 intermediate step. pH measurement in fractions are represented with dotted lines.

518 **Figure 6:** CPC x LC contour plots of the water soluble phase of a pyrolysis bio-oil obtained with UV  
519 (254 nm) and MS detections (ESI negative, ESI positive). MS 2D maps were obtained from the base  
520 peak chromatograms.

521 **Table 1:** CPC conditions used in descending mode.

522 **Table 2:** Products identified in the water soluble phase of a pyrolysis bio-oil using UV and MS  
523 detections. pKa and log P were obtained from <https://chemicalize.com/#/calculation>.