



HAL
open science

A powerful two-dimensional chromatography method for the non-target analysis of depolymerised lignin

Eliise Tammekivi, Magali Batteau, Dorothee Laurenti, Hugo Lilti, Karine Faure

► **To cite this version:**

Eliise Tammekivi, Magali Batteau, Dorothee Laurenti, Hugo Lilti, Karine Faure. A powerful two-dimensional chromatography method for the non-target analysis of depolymerised lignin. *Analytica Chimica Acta*, 2024, 1288, pp.342157. 10.1016/j.aca.2023.342157. hal-04396928

HAL Id: hal-04396928

<https://hal.science/hal-04396928v1>

Submitted on 16 Jan 2024

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.

A powerful two-dimensional chromatography method for the non-target analysis of depolymerised lignin

Eliise Tammekivi^a, Magali Batteau^a, Dorothée Laurenti^b, Hugo Lilti^b, Karine Faure^{a*}

^aUniversité Claude Bernard Lyon 1, ISA UMR 5280, CNRS, 5 rue de la Doua, 69100 Villeurbanne, France

^bUniversité Claude Bernard Lyon 1, IRCELYON UMR 5256, CNRS, 2 Av. Albert Einstein, 69626 Villeurbanne, France

*Corresponding author: karine.faure@isa-lyon.fr, Phone number 0437423686

Abstract

Background: Lignin is an abundant natural polymer obtained as a by-product from the fractionation of lignocellulosic biomass. In the name of a circular economy, lignin should be valorised into valuable chemicals or biomaterials and utilised instead of petrochemicals. For the development of efficient valorisation processes, the structural characterisation of lignin can be highly beneficial. However, this is an arduous task, as the isolated (and sometimes processed) lignin mainly consists of various neutral monomers but also oligomers. In addition, the material contains isomers, which can be especially problematic to separate and identify.

Results: We present a powerful off-line comprehensive two-dimensional (2D) chromatography method combining liquid chromatography (LC), supercritical fluid chromatography (SFC), and high-resolution mass spectrometry for the non-target analysis of depolymerised lignin. The implementation of a 1-aminoanthracene column in the second dimension enabled a class separation of potential lignin monomers, dimers, trimers, and tetramers with an additional separation based on the number of hydroxyl groups and steric effects. The pentafluorophenyl column in the first dimension additionally improved the separation based on hydrophobicity. The comparison of off-line 2D LC×SFC to 1D SFC showed that besides the overall improved performance, the first method is also superior for the separation of isomers. Advanced data analysis methods (MS-DIAL, SIRIUS, and Feature-Based Molecular Network) were integrated into the non-target workflow to rapidly visualise and study the detected compounds, which proved to be especially beneficial for the characterisation of the separated isomers.

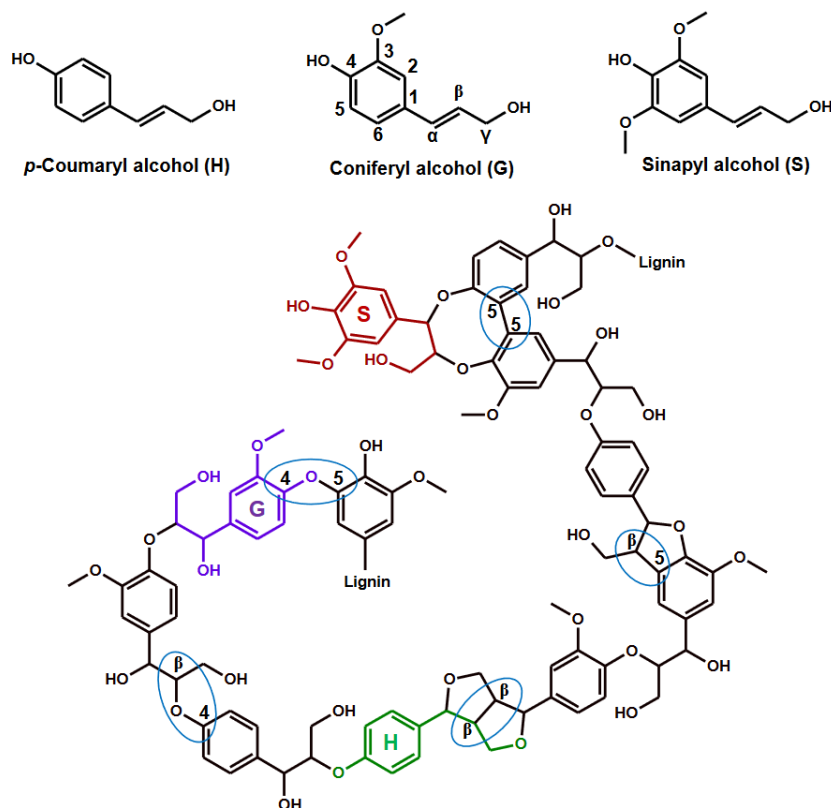
Significance: The method yielded the first 2D LC plot demonstrating a classification of lignin compounds, which can be applied to compare various lignin sources and processing methods. In addition, the technique demonstrated improved separation of compounds, including isomers, which was especially beneficial as 77% of the detected compounds had at least one isomer in the same lignin sample.

Abbreviations: ¹D, ²D: first and second dimension, respectively; 1D, 2D: one-dimensional and two-dimensional chromatography, respectively; MN: molecular network; FBMN: feature-based molecular network; HRMS: high resolution mass spectrometry; DDA: data-dependent acquisition; DIA: data-independent acquisition; MtBE: methyl *tert*-butyl ether; DEA: diethylamine; 1-AA: 1-aminoanthracene; BPC: base peak chromatograms; EIC: extracted ion chromatograms; GNPS: Global Natural Products Social Molecular Networking; DBE: Double Bond Equivalent; log*P*: octanol-water partition coefficient

39 1. Introduction

40 Lignin is an abundant natural polymer that is by far the largest natural resource for aromatic compounds.
41 Nevertheless, it is considered to be a low-value residue and is largely burned to power the pulp and
42 paper industry, where lignin is obtained by the fractionation of cellulose from lignocellulosic biomass.¹
43 However, due to the legislations to decrease carbon dioxide emissions, there is an increasing demand to
44 use less petroleum and petrochemicals by replacing them with alternative natural resources. One
45 alternative could be lignin, as it has potential to be valorised into valuable aromatic compounds (e.g.,
46 vanillin, phenol), by incorporating them into materials (e.g., coatings, resins, thermoplastics, battery
47 electrodes), or transformed into additives for biofuel.^{1,2} As the first step in the pulp or cellulosic ethanol
48 industry, lignin is separated from the lignocellulosic biomass, which also contains cellulose and
49 hemicellulose. Depending on the final aim and valorisation process, the isolated lignin can be
50 depolymerised to a varying extent into smaller oligomers and/or monomers.³ However, for constructing
51 efficient valorisation processes, thorough structural characterisation of the product and/or intermediates
52 can be highly beneficial.^{1,4}

53 Structurally, the polymeric lignin consists of randomly coupled three aromatic subunits (syringyl,
54 guaiacyl, and *p*-coumaryl) which, most commonly, are connected through β -O-4, β -5, β - β , 5-5, or 5-O-
55 4 linkages (Figure 1). The proportions of the monomeric subunits as well as the amount and type of the
56 linkages depend on the botanical origin of the lignin, thus yielding lignin polymers with highly varying
57 structures. In addition, largely because of the reactive hydroxyl groups that can lead to condensation and
58 repolymerisation, it is very difficult to isolate lignin without inducing modifications in the structure.
59 Therefore, because of the variety in the original structure, the influence of the fractionation process, and
60 variations in depolymerisation, the structural analysis of lignin and products is highly challenging.³



61
62 *Figure 1.* An exemplary partial lignin structure and its main units – *p*-coumaryl (H, green), guaiacyl (G, purple),
63 and syringyl (S, red) shown in the structure. Examples of the main linkages (β -O-4, β -5, β - β , 5-5, 5-O-4) are
64 highlighted by blue ovals. Above the lignin structure are the main subunits that by radical coupling form the lignin
65 structure.^{3,5}

66 One widely used method that can aid with obtaining detailed structural information about the
67 depolymerisation products is chromatography. Because of the complex structure of depolymerised
68 lignin, two-dimensional (2D) chromatography is applied more and more, as it has demonstrated higher
69 peak capacity and improved selectivity compared to one-dimensional (1D) chromatography.⁶⁻⁸ 2D gas
70 chromatography (2D GC, largely as GC×GC) is nowadays considered to be a standard method for the
71 analysis of processed biomass.⁹ However, with 2D GC commonly only lignin monomers and dimers are
72 detected, as larger oligomers are not volatile enough for GC.^{4,7} Volatility is less of an issue with liquid
73 chromatography (LC) and, indeed, 2D LC has been successfully applied for the analysis of poorly
74 volatile or thermally unstable compounds, such as carotenoids¹⁰ or bioactive polyphenols¹¹. However,
75 2D LC has been much less applied for the analysis of biomass, as it is often problematic to hyphenate
76 two LC dimensions having non-correlated separation mechanisms (in other words, high orthogonality)
77 such as NPLC×RPLC.⁴ In the case of depolymerised lignin, the sample generally consists of neutral
78 compounds, for which 2D LC methods combining reversed-phase LC (RPLC) and supercritical fluid
79 chromatography (SFC) have demonstrated good orthogonality.¹² The studies by Sun *et al.*⁶ and Sarrut
80 *et al.*¹³ have already demonstrated some first advantages of applying LC×SFC for the analysis of
81 monomeric neutral compounds. However, to obtain more information about the separation and detected
82 compounds, mass spectrometry (MS) detection is compulsory. Together with fragmentation, high
83 resolution MS (HRMS) is an indispensable tool to perform non-target analysis, where the aim is to detect
84 and annotate essentially as many analytes as possible without a prior analyte selection. Currently, only
85 a small proportion of the compounds can be identified, as the non-target acquisition and data analysis
86 remain a challenge.¹⁴ Applying 2D separation before the MS detection can be highly beneficial, as the
87 improved separation will reduce the complexity caused by the number of co-eluting ions and at the same
88 time decrease ion suppression, thus improving ionisation. These aspects can lead to identification with
89 higher confidence, detection of minor components, and a more reliable quantification when suitable
90 standards are available.¹⁵

91 For efficient and repeatable data analysis, more and more computational tools are emerging. In this
92 work, we applied MS-DIAL¹⁶ for peak picking, MS/MS deconvolution, peak alignment, and cleaning
93 of the MS/MS spectra. This software can use both data-dependent acquisition (DDA) and data-
94 independent acquisition (DIA) data as the input, which, after treatment, can be straightforwardly
95 exported in a format suitable for the Molecular Network (MN) analysis. The Global Natural Products
96 Social Molecular Networking (GNPS) platform organises detected features together based on their
97 MS/MS spectral similarities.¹⁷ This has proven to be a powerful approach for the discovery of novel
98 components also in natural products.¹⁸ In recent years, the combined utilisation of 2D LC and MN for
99 the non-target analysis of natural products has increased. One of the major areas is the analysis of
100 traditional Chinese medicine using off-line methods combining RPLC with NPLC, HILIC, or mix-mode
101 LC.¹⁹⁻²² As far as we know, Qu *et al.*²³ and Wei *et al.*²⁴ have been the only ones to apply off-line LC×SFC
102 together with MN for the analysis of complex samples, in their case, traditional Chinese medicine.
103 However, in our work, we collected approximately three times more fractions from the first dimension
104 (1D) to decrease the under-sampling effect that substantially decreases the effective peak capacity of an
105 off-line 2D method.²⁵

106 In the present study, the aim was to apply a powerful off-line comprehensive LC×SFC-HRMS/MS
107 method for the non-target analysis of depolymerised lignin, which also includes oligomers higher than
108 dimers. Particular focus was given to the separation and study of isomers with MN, as depolymerised
109 lignin is expected to contain a high number of isomers. To our knowledge, this is the first study to
110 combine a comprehensive 2D LC method with MS and apply it to the analysis of lignin.

111

112 2. Materials and methods

113 2.1 Chemicals and sample preparation

114 LC-MS grade methanol (MeOH) and methyl *tert*-butyl ether (MtBE) of HPLC grade were purchased
115 from Sigma-Aldrich (Darmstadt, Germany), and formic acid (FA, 95-97%) from Aldrich. LC-MS grade
116 acetonitrile (ACN) was purchased from Honeywell Riedel-de-Haën (Seelze, Germany) and pressurised
117 liquid CO₂ (99.995%) from Air Liquide (Pierre Bénite, France). Water suitable for MS was produced
118 by using an Elga water purification system (Veolia water STI, Le Plessis Robinson, France).

119 Based on previously published studies,^{6,26-28} 33 phenolic standards were selected to represent
120 monomeric compounds that can be produced from lignin depolymerisation. All the following standards
121 were purchased from Sigma-Aldrich (with numbering corresponding to Table S1 in the Supplementary
122 Information - SI): anisole (1); 2,3-benzofuran (2); 2,3-dihydrobenzofuran (3); eugenol (4); 2,4,6-
123 trimethylphenol (5); 2,6-dimethylphenol (6); 2-phenylethanol (7); *trans*-isoeugenol (8); syringol (9);
124 2,5-dimethylphenol (10); 2-ethylphenol (11); *o*-cresol (12); *m*-cresol (13); 4-ethylphenol (14); *p*-cresol
125 (15); phenol (16); acetovanillone (17); 3,4-dimethylbenzoic acid (18); acetosyringone (19);
126 syringaldehyde (20); 4-hydroxybenzaldehyde (21); 4-hydroxyacetophenone (22); coniferaldehyde (23);
127 4-methylcatechol (24); catechol (25); vanillyl alcohol (26); 4-hydroxybenzoic acid (27); hydroquinone
128 (28); ferulic acid (29); *p*-coumaric acid (30); 3,4-dihydroxyphenylacetic acid (31); pyrogallol (32); and
129 caffeic acid (33). Other characteristics of the standards are presented in Table S1 in the SI.

130 2.2 Sample preparation

131 From each monomer standard, single-compound stock solutions of 1 mg mL⁻¹ in methanol were prepared
132 and stored at 5 °C. From these stock solutions, five mix solutions containing five or six standards were
133 prepared and diluted with MeOH with a final concentration of 0.1 mg mL⁻¹ for each monomer. To ease
134 the MS interpretation, the standards were divided into mixtures (A, B, C, D, E, and F) so that monomers
135 having the same molecular formula were in different mix solutions (Table S1 in the SI).

136 The depolymerised lignin sample was prepared by the following procedure, which will be presented in
137 more detail in future publications focusing on the catalytic depolymerisation of lignin. Previously dried
138 SODA P1000 lignin was depolymerised in a batch reactor in the presence of a CoMoS/Al₂O₃ catalyst
139 and isobutanol at 280°C and 60 bar. The bio-oil was recovered by solubilisation in tetrahydrofuran
140 followed by filtration to remove the catalyst and evaporation of the solvents. The yielding solid mixture
141 was separated into two fractions by solubilisation in acetone and precipitation of the heavier fraction
142 with *n*-heptane. After centrifugation, the supernatant (lighter fraction) was recovered and evaporated to
143 dryness. From this oily solid, a solution of 30 mg mL⁻¹ was prepared in MeOH, sonicated for 10 min in
144 an ultrasonic bath, and the supernatant was collected for the chromatographic analyses.

145 2.3 Apparatus

146 The screening of columns for the SFC dimension was carried out using an Agilent (Santa Clara, CA,
147 USA) 1260 Infinity II UHPSFC system. A more detailed overview of the experimental conditions used
148 for the screening of columns is presented in Table S2 in the SI. The following columns, a Torus DEA
149 (diethylamine, 1.7 µm, 3 mm x 100 mm), a Torus DIOL (1.7 µm, 3 mm x 100 mm), a Torus 1-AA (1-
150 aminoanthracene, 1.7 µm, 3 mm x 100 mm), and a CORTECS HILIC (1.6 µm, 2.1 mm x 50 mm) were
151 purchased from Waters (Milford, MA, USA).

152 For the ¹D LC separation in the off-line LC×SFC analysis, an Agilent 1260 Infinity II HPLC system
153 was used, equipped with a binary solvent delivery pump (G7111A), an autosampler (G7129A) with a
154 900 µL loop, a thermostated column compartment (G7116A, up to 90 °C), a diode array detector (DAD)
155 with a 1 µL flow-cell (G7117C), and a fraction collector with a capacity of 120 fractions (G1364F), all

156 from Agilent Technologies. The dwell volume was determined experimentally to be 2.65 mL. The
157 instrument was controlled by using the Agilent Openlab CDS Chemstation software. A Kinetex F5
158 column (5 μm , 4.6 mm x 150 mm) was purchased from Phenomenex (Torrance, CA, USA). The column
159 geometry was selected to allow the collection of fractions with a reasonable volume to be handled. A
160 miVac centrifugal vacuum concentrator (Genevac, Ipswich, UK) was used to evaporate the fractions to
161 dryness. For the ^2D SFC separation, a Waters UPC² system equipped with a binary solvent delivery
162 pump, an autosampler with a 10 μL loop, a thermostated column compartment (up to 90 $^\circ\text{C}$), a DAD
163 with an 8.4 μL flow-cell, a backpressure regulator (BPR), and a make-up solvent pump all from Waters.
164 The instrument was controlled by using Waters Empower CDS software. A 5 cm long Torus 1-AA
165 column (1.7 μm , 3 mm x 50 mm) was purchased from Waters.

166 For the accurate-mass tandem mass spectrometry detection, an Agilent 6560 Ion Mobility Quadrupole
167 Time-of-Flight (Q-ToF) instrument equipped with an Agilent Jet stream electrospray ionisation (ESI)
168 source was used. Agilent MassHunter Qualitative Analysis 10.0 was used for the extraction of base peak
169 chromatograms (BPC) and extracted ion chromatograms (EIC). The 2D plots of the off-line LC \times SFC
170 measurement were visualised by using a Matlab (Mathworks) code developed in-house by S. Heinisch
171 and F. Rouvière.

172

173 2.4 Methods

174 2.4.1 Off-line 2D LC \times SFC-Q-ToF-MS/MS conditions

175 The ^1D LC separation was performed using the Kinetex F5 column with H_2O containing 0.1 vol.% of
176 FA (**A**) and ACN containing 0.1 vol.% of FA (**B**). The column temperature was 30 $^\circ\text{C}$ and the injection
177 volume was 10 μL . The flow rate was 2.1 mL min^{-1} and the gradient was the following: 1% to 100% of
178 **B** in 9.0 min, held at 100% of **B** for 0.77 min (one column dead time), decreased back to 1% of **B** in
179 0.77 min, and equilibrated for 3.58 min. The Agilent fraction collector enabled the collection of 120
180 fractions into chromatographic vials with 250 μL inserts every 0.085 min starting after the column dead
181 time. Therefore, the fractions consisted of the effluent eluting during gradient time and dwell time. The
182 variability induced by the collection process is around 1%. After collection, the fractions were
183 evaporated to dryness. Then, a volume equivalent to the previously collected volume of MtBE was
184 added to the residue and the vial was vortexed for a few seconds.

185 The ^2D SFC separation was performed using the 5 cm Torus 1-AA column with supercritical CO_2 (**A**)
186 and methanol as co-eluent (**B**). The column temperature was 40 $^\circ\text{C}$ and the injection volume was 7 μL .
187 The flow rate was 1.4 mL min^{-1} , which was determined as the maximum flow rate accessible at 50% of
188 **B**. The gradient was the following: 5% to 50% of **B** in 3.6 min, decreased back to 5% of **B** in 0.2 min,
189 and held for 1.0 min to equilibrate the column. The BPR temperature and pressure were set to 70 $^\circ\text{C}$ and
190 140 bar, respectively. The DAD signal was registered at 210 nm and 254 nm.

191 The ^2D SFC column was hyphenated to the MS by using the “*preBPR splitter with a sheath pump*”
192 configuration, described in more detail by Guillarme *et al.*²⁹ MeOH, with a flow rate of 0.5 mL min^{-1} ,
193 was used as the make-up solvent. The ESI was performed in the negative ionisation mode under the
194 following parameters: drying and sheath gas temperature 350 $^\circ\text{C}$ with a flow of 11 L min^{-1} , nebuliser
195 pressure 40 psi, capillary voltage 3.5 kV, and fragmentor voltage 185 V. The range of 90-1700 m/z was
196 set for the mass scan. All ions fragmentation (AIF) mode with two experiments – collision energy 0 eV
197 and 20 eV – were applied, which led to HRMS data for both MS^1 and MS^2 , respectively. The acquisition
198 rate was 3 spectra per second.

199 2.4.2 1D SFC-Q-ToF-MS/MS conditions

200 For the 1D SFC-Q-ToF-MS/MS analysis, the SFC method was modified to provide the optimal
201 separation that a corresponding 1D method can give, therefore enabling a fair comparison between the
202 2D and 1D separation (see section 3.7). Therefore, the longer 10 cm Torus 1-AA column and a longer
203 gradient time (30 min) were selected. Because of these changes, other modifications had to be made. As
204 the slope of the separation was much lower, the analytes eluted at a lower final eluent concentration
205 (35% instead of 50%). In addition, because of the increased pressure caused by the longer column, the
206 flow rate had to be decreased to 1.0 mL min⁻¹ corresponding to the maximum flow at 35% of co-eluent.
207 Therefore, the following gradient was applied: 5% to 35% of **B** (MeOH) in 30.0 min, 35% to 5% of **B**
208 in 0.5 min, and equilibrated for 2.2 min. The other aspects of the 1D analysis (instrumentation, MS
209 parameters, etc.) were the same as in the 2D SFC component of the off-line LC×SFC-HRMS analysis.

210 2.4.3 Advanced data analysis

211 For the more detailed non-target analysis, a combination of feature detection, visualisation, and
212 annotation tools was applied. First, the raw Agilent MassHunter data files (.D) were converted to .ABF
213 (Analysis Base File) format using ABFConverter (Reifycs Inc., Tokyo, Japan). The data was processed
214 with MS-DIAL software version 4.9 (RIKEN Centre for Sustainable Resource Science, Kanagawa,
215 Japan) using the AIF multiple collision energies option. The mass tolerance was set at 0.01 Da for MS¹
216 and 0.025 Da for MS². For retention time, the range of 0.2-3.8 min was defined. For peak detection, an
217 amplitude of 20000 was selected as the minimum peak height with a mass slice width of 0.05 Da. For
218 deconvolution with the MS²Dec algorithm, 1000 was set as the MS/MS abundance cut-off and 0.5 as
219 the sigma window value. The ion species [M-H]⁻ and [M-H₂O-H]⁻ were defined. For peak alignment,
220 the last fraction (fraction 120) was selected as the reference with a retention time (*t_R*) tolerance of 0.2 min
221 and MS¹ tolerance of 0.02 Da. Removal of features based on blank filtering was selected. The results of
222 the MS-DIAL processing were exported by generating an MS/MS spectral summary file (.MGF) and
223 feature quantification table (.txt). The results were also imported to the MS-FINDER software³⁰ to
224 determine the formula (mass tolerance ± 3 ppm) for each detected feature. The feature was excluded
225 from the files when *a*) there was no match with any C_xH_yO_z formula, *b*) the DBE value corresponding
226 to the formula was less than 4, *c*) the feature was likely a fragment as it had a similar chromatogram
227 with a higher *m/z* value (indicated by MS-DIAL), or *d*) the feature was also detected in consecutive
228 fraction(s) at the same retention time. In the last case, only the feature with the most intensive signal
229 (representing the maxima of the peak) was kept. Here, this procedure was done manually; however,
230 automatic “demodulation” processes are under development to convert 2D data into 1D data that can be
231 analysed with various novel softwares.³¹ In addition, intensive background signals that were present in
232 almost all MS² spectra (*m/z* 146.96, 188.94, and 119.95) were removed from the MS-DIAL output files.
233 The edited output files were combined into one MGF and one .txt file containing all possible lignin
234 features. Both files were used to construct the Feature-Based Molecular Network (FBMN) on the GNPS
235 (gnps.ucsd.edu) platform.³²

236 Because of the size of the combined files, they were first uploaded to the MassIVE repository
237 (massive.ucsd.edu), after which they could be imported into the FBMN workflow on GNPS. MS-DIAL
238 was selected as the Quantification Table Source and 0.02 Da was set as the mass tolerance for both
239 precursor and fragment ions. The following values were selected as the network parameters: 0.6 as the
240 minimum cosine score for a pair of consensus MS² spectra, 4 as the minimum number of shared fragment
241 ions, 10 as the maximum number of nodes for one single node, and 0 as the maximum connected
242 component size, which allows the connection of unlimited number of nodes in a single network. The
243 default values were kept for the other input values. After running the FBMN workflow, the output
244 network files were downloaded from the GNPS platform and imported to the Cytoscape software
245 (version 3.9.1.) for visual editing.

246 For the structural characterisation of the detected features, the combined MGF file was imported and
247 analysed in the SIRIUS software (version 5.6.3).^{33,34} By using this software, it is possible to propose the
248 molecular formula (SIRIUS) of a feature and elucidate its structure based on the fragmentation data
249 (CSI:FingerID³⁵). In the software, the MS² mass accuracy was set to 10 ppm, [M-H]⁻ adduct was
250 selected, and C, H, and O elements were allowed for the search of the molecular formula. The same [M-
251 H]⁻ adduct was selected for the fingerprint prediction and all the databases present in the software were
252 applied for database search. In addition, the CANOPUS³⁶⁻³⁸ function in the same software was applied
253 to predict the compound class.

254 2.4.4 Peak capacity and coverage of the 2D separation space

255 Some figures of merit were calculated to characterise the performance of the 2D LC method for the
256 analysis of this kind of depolymerised lignin sample. The most widely demonstrated values are peak
257 capacity and orthogonality, which have proven to be suitable for the comparison of different 2D
258 separation methods. However, as different methods can be used for the calculation of the values, we will
259 briefly present the approaches used in this study.

260 The peak capacity (n_C) for a gradient separation can be individually estimated for both dimensions using
261 the corresponding gradient time (t_g) and average peak width as 4σ (w) in time units according to Eq. 1:

$$262 \quad n_C = \frac{t_g}{w} \quad (1)$$

263 However, for 2D separations, it is necessary to include the under-sampling correction factor (β) to take
264 into account the peak broadening caused by the under-sampling during the 1D (in the formula noted as
265 D_1) fraction collection, which can be done by Eq. 2 presented in the work by Li *et al.*³⁹:

$$266 \quad n_{C,corrected}^{D1} = \frac{n_C^{D1}}{\beta} = \frac{n_C^{D1}}{\sqrt{1+3.35 \times (\frac{t_s}{w_1})^2}} \quad (2)$$

267 where t_s is the sampling time. Under ideal conditions, the peak capacity of a 2D method can be calculated
268 according to Eq. 3:

$$269 \quad n_{C,theoretical}^{2D} = n_C^{D1} \times n_C^{D2} \quad (3)$$

270 When combining Eq. 1, 2, and 3, the equation for effective 2D peak capacity can be obtained, which is
271 a more realistic value, as it also takes into account the effect of under-sampling.

$$272 \quad n_{C,effective}^{2D} = \frac{n_C^{D1} \times n_C^{D2}}{\sqrt{1+3.35 \times (\frac{t_s \times n_C^{D1}}{t_g^1})^2}} \quad (4)$$

273 The orthogonality was estimated using the convex hull approach.⁴⁰

274

275 3. Results and discussion

276 3.1 Screening of columns for the 2D SFC dimension

277 The selection of the suitable column for the 2D SFC dimension was made based on the analysis of 33
278 monomer standards, which were selected to represent possible lignin monomers. Four columns were
279 screened using generic gradient elution (5% – 50% of MeOH, normalised gradient slope of 3%, the
280 other parameters can be seen in Table S2). With the Cortecs HILIC column, the majority of the standards
281 co-eluted at dead time. In contrast, the Torus DEA column had very strong retention for some analytes

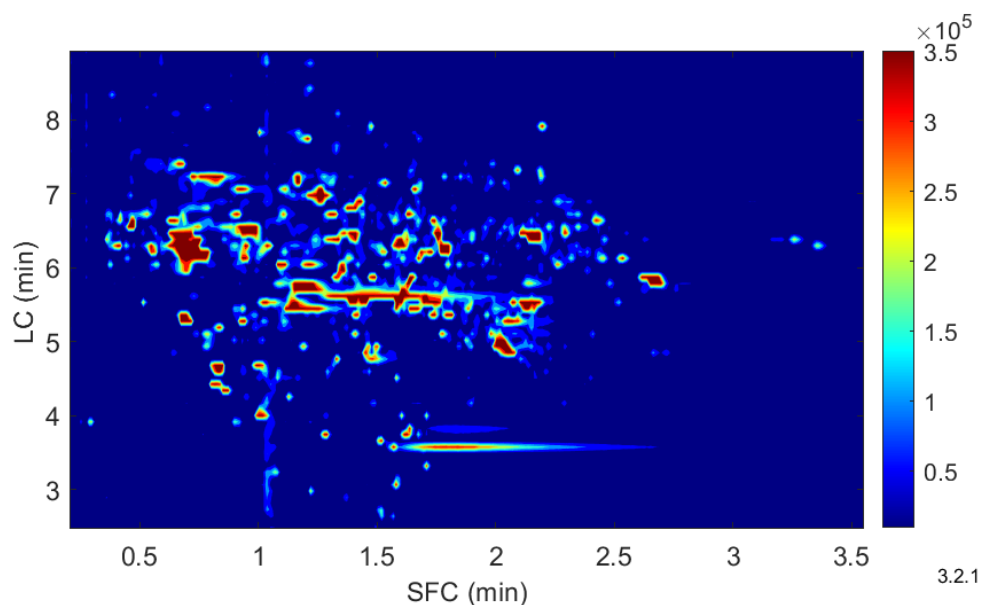
282 (pyrogallol; 3,4-dihydroxyphenylacetic acid; caffeic acid), causing peak broadening and difficulties with
283 the detection of these compounds on the chromatogram. The strong retention of phenolic acids was also
284 noted by Prothmann *et al.*⁴¹ with Torus DEA and Torus 2-picolylamine columns. DIOL and 1-AA both
285 exhibited a good separation over a similar elution range with a ΔC_e (difference of mobile phase
286 composition at the elution of the most and least retained compounds) value of 22.3% for DIOL and
287 24.7% for 1-AA using the 33 monomer standards (Figure S1 in the SI). However, the 1-AA column has
288 demonstrated a better separation of lignin monomers from dimers due to the enhanced selectivity
289 provided by the π - π interactions between the double bonds of lignin compounds and the aromatic
290 anthracene stationary phase.⁴² As we expected that the depolymerised lignin sample involved in our
291 study contains other compounds besides monomers, we selected the 1-AA column to investigate if this
292 mechanism also aids the separation of higher oligomers in the 2D setup.

293

294 3.2 General characterisation of the off-line LC×SFC separation

295 To find the suitable elution ranges for the two dimensions, the depolymerised lignin sample was
296 separately analysed with RPLC and SFC. The gradient times, flow rates, and other relevant factors for
297 both dimensions were selected based on calculations that are presented in our previous study on the
298 optimisation of an off-line LC×SFC method.²⁵ However, in the present work, we were able to collect
299 120 fractions instead of 80, which is beneficial for obtaining a higher effective peak capacity for the 2D
300 method. Indeed, using Eq. 4, a very high effective peak capacity value – 3218 – was calculated for the
301 LC×SFC method, demonstrating the high separation efficiency of the developed 2D method for the
302 analysis of depolymerised lignin.

303 Figure 2 presents the obtained LC×SFC-HRMS plot using the BPC signal. Based on the convex hull
304 approach, we calculated the coverage of the separation space to be 62%. Considering that the sample
305 consists largely of neutral compounds with similar structures, the obtained value can be considered good.
306 When taking into account the partial occupation of the separation space, the corrected peak capacity for
307 the 2D method is 1995, which nowadays is still an unreachable value with 1D separations. In addition,
308 as can be seen in Figure 2, there is no clustering of peaks on the upward or downward diagonals,
309 demonstrating the low correlation and suitability of the selected dimensions for the analysis of neutral
310 compounds present in depolymerised lignin.



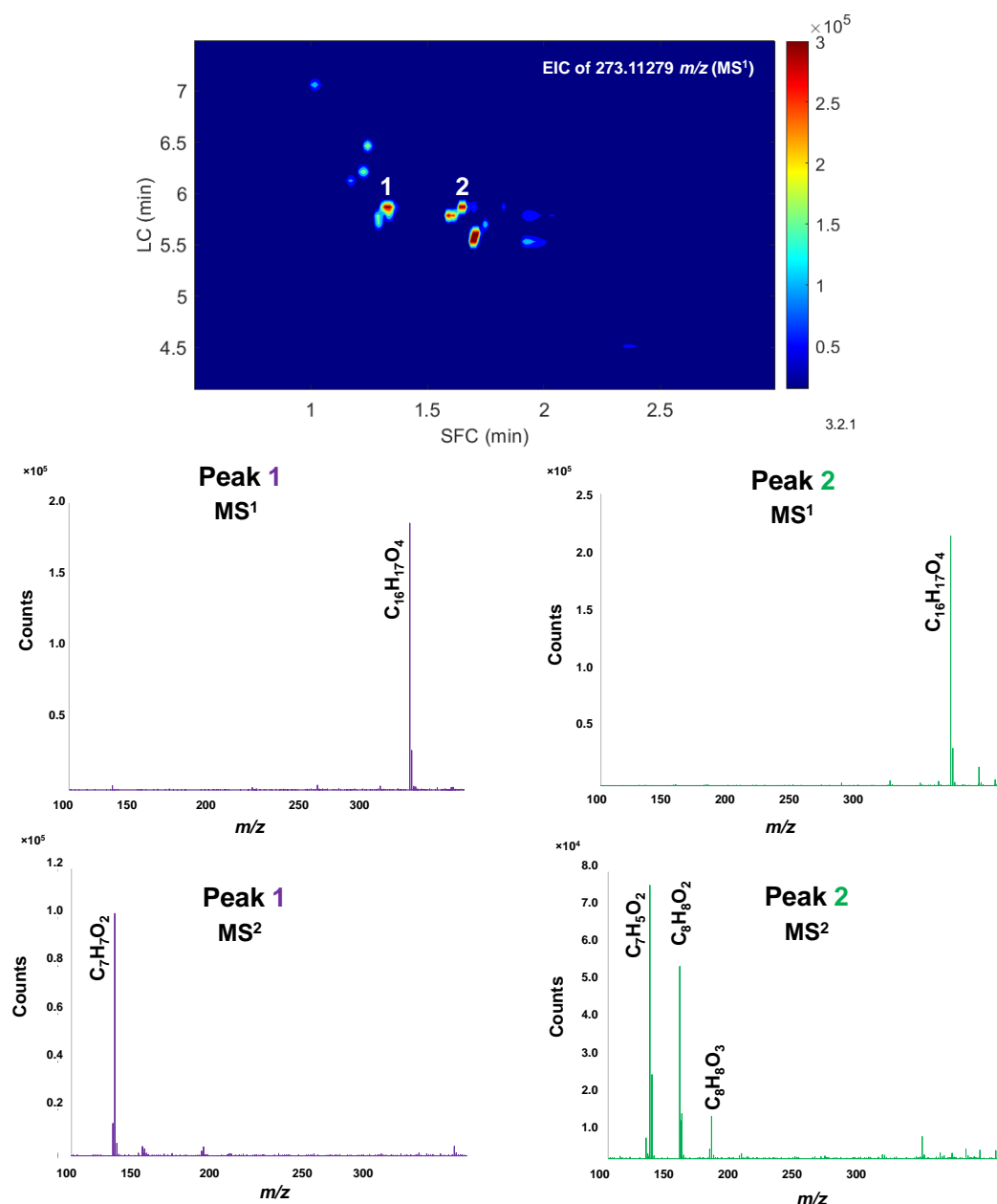
311

312 *Figure 2.* LC×SFC-HRMS BPC contour plot (MS^1) of depolymerised lignin.

313

314 3.3 MS-DIAL for the off-line LC×SFC-HRMS/MS data analysis

315 Depending on the data and research question at hand, different data analysis methods can be helpful in
316 interpreting the measured tandem mass spectra. For the data obtained with the technique used in this
317 study, an example is presented in Figure 3, showing the 2D plot for deprotonated $C_{16}H_{18}O_4$ and
318 exemplary MS^1 and MS^2 spectra for two separated peaks. It can be seen that these two compounds are
319 isomers with different fragmentation patterns and as the MS^1 and MS^2 spectra are clean of interfering
320 signals and the MS^1 contains mainly just the pseudo-molecular ion, it was decided to apply mass-spectral
321 software for the data analysis. The main advantage of using MS analysis software is the increased speed
322 of the data analysis, which is especially beneficial in this study as the 2D analysis generated 120 data
323 files. In addition, it can yield MS/MS data analysable with other approaches (here SIRIUS and FBMN).
324 In this study, MS-DIAL software was first applied, which detects features and associated fragments by
325 fitting the chromatograms of product ions to a precursor by using retention time.¹⁶



326

327 *Figure 3.* Example of mass spectra measured with the LC×SFC-HRMS method. In the first row is the LC×SFC-
328 HRMS EIC contour plot (MS^1) of 273.11279 (± 5 ppm) m/z , which corresponds to the exact mass of deprotonated
329 $C_{16}H_{18}O_4$. In the second and third row are the spectra that represent exemplary MS^1 and MS^2 spectra corresponding
330 to the two separated isomers (Peak 1 and Peak 2), respectively.

331

332 In addition to true features originating from the compounds of the sample, MS-DIAL can also pick up
333 so-called false features. These can be signal noise, system contamination, or in-source fragmentation.⁴³
334 Even with ESI, which is considered to be the softest ionisation technique, some compounds can have
335 in-source fragmentation.⁴⁴ Therefore, when there are many analytes eluting simultaneously, it becomes
336 more and more difficult to distinguish true features from fragments (false features), especially when the
337 analytes are very similar. In this study, we manually checked features for which the MS-DIAL software
338 showed an alert “Chromatogram similar” and removed the features with lower m/z values. Those
339 features are likely in-source fragments, because their peak shapes were similar to the peak shape of the
340 molecular ion and the m/z value corresponded mainly to the loss of CH_3 or CO .

341 MS-DIAL has been shown to enable the deconvolution of DIA MS/MS spectra for compounds that co-
342 elute to some extent.⁴⁵ However, the precursors have to have different exact masses and for an efficient
343 deconvolution, there has to be at least two data-point difference between the maxima of the
344 chromatographic peaks.⁴⁶ Recent advancements such as CorrDec⁴⁶ (which has been added to MS-DIAL
345 version 3.22 and up) have been shown to enable the deconvolution of even completely co-eluting
346 compounds. However, this approach requires that the analysis of the sample would be repeated at least
347 ten times, which would make the off-line 2D LC analysis of one sample dreadfully long. Also, for the
348 time being, there is no evidence that it would be able to deconvolute isomers, which signifies that
349 powerful chromatographic separation is still needed before the MS detection.

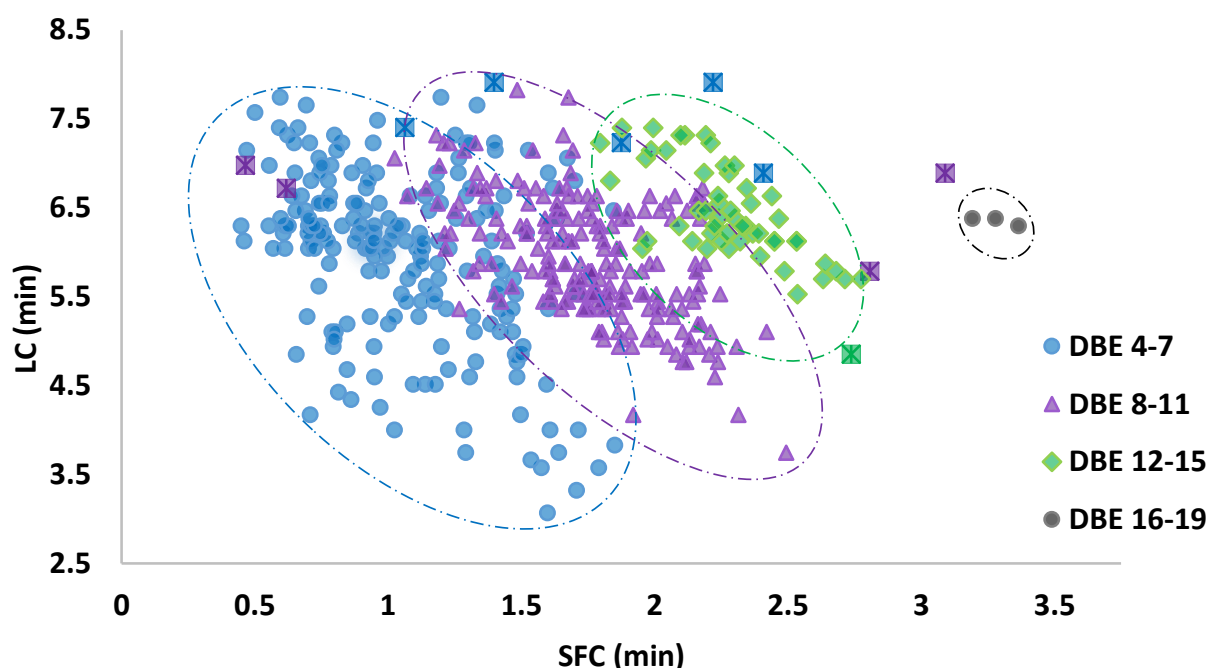
350

351 3.4 Classification of the detected lignin compounds

352 All of the 120 data files (in .ABF format) obtained from the LC×SFC-HRMS/MS measurement were
353 loaded to MS-DIAL and analysed with the parameters described in section 2.4. After combining the
354 results, 471 pseudo-molecular ions corresponding to a $C_xH_yO_z$ formula with $DBE \geq 4$ were detected in
355 the depolymerised lignin sample (Table S3 in the SI). DBE (Double Bond Equivalent, also Ring Double
356 Bond Equivalent aka RDBE or RDB is sometimes used for the naming) is a value that can be used to
357 classify detected compounds, including lignin depolymerisation products. Depending on the
358 experimental conditions of the depolymerisation and on other additional sample pre-treatment
359 processes, such as fractionation, the lignin sample can contain monomers (analytes with one aromatic
360 core), dimers (two aromatic cores), trimers (three aromatic cores), and so on. Looking at the structure
361 of lignin in Figure 1, the DBE value for monomers should generally stay in the range of 4-7, for dimers
362 in the range of 8-11, for trimers 12-15, and so forth. As the MS-DIAL software yielded exact masses
363 for all of the features, it was possible to find the molecular formula and, from that, the DBE value. For
364 the analysed sample, the three most retained compounds in the SFC dimension also had the highest DBE
365 value (17, generally corresponding to tetramers). After those compounds, there was no UV signal,
366 confirming that the sample did not contain oligomers higher than tetramers.

367 Based on the DBE value, the detected features were plotted on an LC×SFC plot (Figure 4) with the same
368 retention times as was seen in Figure 2. As hypothesised, the detected features are clustered together
369 based on their DBE value corresponding possibly to a mono-, di-, tri-, or tetramer. This phenomenon
370 can again be explained by the π - π interactions between the depolymerised lignin aromatic ring(s) and
371 the anthracene moiety of the 1-AA stationary phase. This kind of classification can be used to detect
372 outliers for further investigation. In Figure 4, the compounds that are likely not lignin compounds are
373 marked with crossed-out squares, as their structural candidates and compound classes in SIRIUS were

374 far from potential lignin compounds (e.g., aliphatic compounds, terpenoids). This classification can also
375 help with comparing samples to each other without the need for a detailed structural study. This kind of
376 visual comparison based on grouping on 2D plots is a very common approach in 2D GC but much less
377 explored in 2D LC.⁴



378
379 *Figure 4.* LC x SFC-HRMS plot (MS^1) of $C_xH_yO_z$ compounds ($DBE \geq 4$) detected with MS-DIAL in the
380 depolymerised lignin sample. The same data that was used in Figure 2 is also presented here. The ovals with dashed
381 lines represent the clusters of lignin compounds with DBE 4-7 (blue dots, likely monomers), 8-11 (purple triangles,
382 likely dimers), 12-15 (green diamonds, likely trimers), and 16-19 (black dots, likely tetramers). The outliers that
383 are presumably not lignin compounds are presented with crossed-out squares.

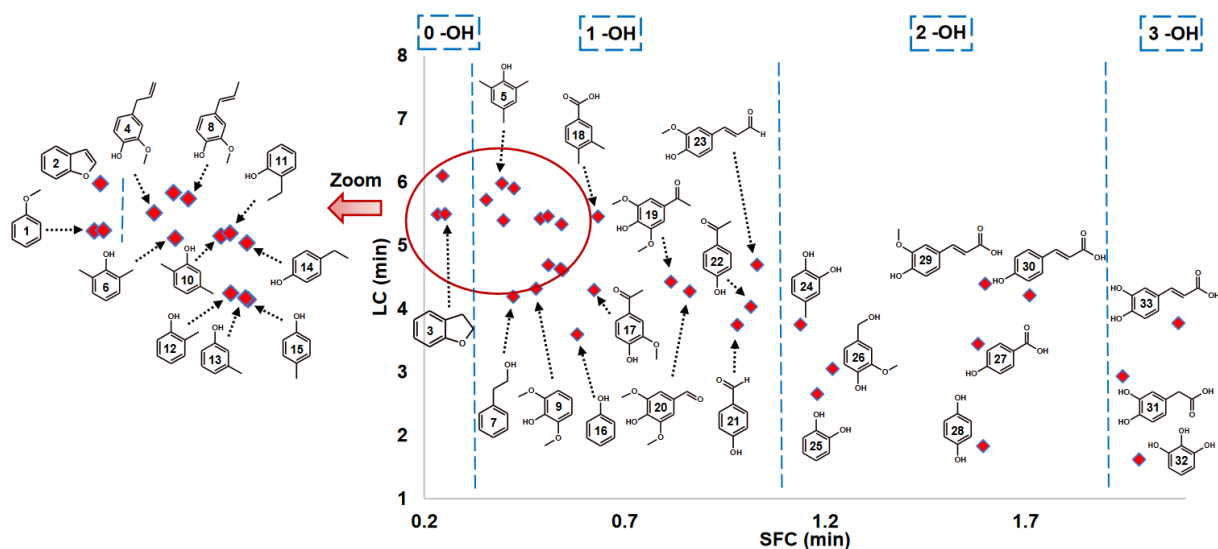
384 At the same time, one must be careful, as there can also be lignin compounds that have a DBE value
385 corresponding to a higher class of the presented DBE ranges. For example, as was shown in the SI of a
386 publication by Prothmann *et al.*⁵, additional double bonds in the structure of a dimer can yield a DBE
387 value that classically corresponds to a trimer. Therefore, for the sample analysed in our work, we studied
388 the potential structural candidates proposed by the SIRIUS software based on the comparison of MS/MS
389 pattern to see if the candidates belong to the class implied by the DBE value or not (Table S4 in the SI).
390 This was merely done to see if the candidates matched mainly mono-, di-, tri-, or tetramers, not to
391 tentatively identify all the structures of the detected compounds, as this is not the aim of this work. In
392 the case of compounds with a DBE value of 4-7, 185 compounds had one aromatic ring, 9 corresponded
393 more to unsaturated aliphatic compounds, and 3 had candidates matching to a similar extent to
394 compounds both having and not having an aromatic ring. With DBE value of 8-11, there were 202
395 compounds whose candidates all matched dimers and 1 that matched with monomers. With DBE range
396 12-15, 24 compounds matched largely trimers, 8 had candidates mostly corresponding to dimers, and
397 13 had both dimers and trimers as equally potential matches. For DBE 16-19, 1 compound matched
398 tetramers and 1 compound could be either trimer or tetramer. Therefore, under the classical DBE ranges,
399 94% had indeed highest matches with monomers, almost 100% with dimers, 53% with trimers, and one
400 out of two to tetramers. This suggests that some of the compounds classified by the DBE value could be
401 indeed smaller by one aromatic ring. To determine how much misclassification using DBE value can
402 cause, it would be necessary to analyse oligomer standards (which are scarce, as mentioned before) or
403 have a good match in the SIRIUS software.

404 The SIRIUS software uses structural fingerprints that have been predicted from the structures of
405 potential candidates found in chemical databases. For this, the corresponding structure has to be present
406 in one of the databases (such as PubChem, COCONUT) utilised by the software. In the case of the
407 detected compounds, over 95% had the first match with a compound that was present also in the
408 PubChem database. Therefore, to get an idea how well the PubChem database covers lignin compounds,
409 we searched for the lignin oligomers tentatively identified in the study by Prothmann *et al.*⁵ As can be
410 seen in Table S5 in the SI, out of the 35 compounds, only one had an exact stereo isotope match in
411 PubChem. For 12 compounds, there were no isomers matching at least with 90% of structural similarity.
412 These included all those dimers that had a DBE value classically corresponding to trimers. This
413 demonstrates, that the databases contain only a small part of the potential structures, as lignin oligomers
414 with all of the potential structures caused by the isolation and depolymerisation are quite understudied
415 compounds. However, some oligomers had up to 69 isomers present in the PubChem database, which
416 allows for some compounds to have a match at least with a structurally similar structure. In addition, for
417 a lot of the potential tri- and tetramers and for some dimers detected in our study, the match factor in
418 SIRIUS was quite low (around 50% or less). Therefore, for more reliable structural characterisation of
419 lignin compounds, specific lignin databases, further fragmentation studies, or manual elucidation of the
420 MS/MS spectra would be necessary. However, applying the SIRIUS software still enabled us to
421 demonstrate the benefits of using the 2D chromatographic method for the analysis of lignin before
422 MS/MS, which was the aim of our study. In addition, the classification based on DBE value can still
423 help to determine the possible class of monomer/oligomer, detect outliers for further investigation, and
424 compare samples at large.

425

426 3.5 Investigation of the separation mechanism

427 Although the analysis of the depolymerised lignin sample demonstrated the classification of lignin
428 compounds based on their DBE values on the 2D plot (Fig. 4), the analysis of monomeric standards aids
429 to study the separation mechanisms more in-depth. Therefore, the same 33 monomer standards that were
430 used for column screening were analysed with the optimised LC and SFC methods. In Figure 5, the
431 placements of the monomer standards on the 2D LC×SFC-HRMS plot are presented (the names of
432 compounds are shown in section 2.1 and Table S1). By studying the retention of similar standards 31
433 and 33, it can be seen that an additional double bond in the structure increases the retention in the SFC
434 dimension. Comparison of the retention of standards 4 and 8, which differ only in the position of the
435 double bond, demonstrates that a higher number of conjugated double bonds increases retention with
436 the 1-AA column. Once again, this is caused by the π - π interactions, which are stronger in the case of
437 larger conjugated structures.⁴⁷



438

439 *Figure 5.* 2D LC×SFC-HRMS plot of the monomer standards. The plot was constructed by combining the 1D LC
 440 and 1D SFC retention times of the standards. The names of the compounds corresponding to the numbering are
 441 shown in section 2.1 and Table S1. Dashed blue lines separate roughly the sections based on the number of
 442 hydroxyl groups in the monomer structure.

443 Another mechanism that plays a major role in the separation is the hydrogen bonding interaction. As
 444 can be seen by the vertical dashed blue lines in Figure 5, in the case of the analysed standards, the SFC
 445 dimension can be roughly sectioned based on the number of hydroxyl groups in the structure of the
 446 monomers. The more hydroxyl groups are in the structure of the compound, the stronger the hydrogen
 447 bonding interaction with the 1-AA polar sites, which in turn leads to increased retention. For example,
 448 standard 32 (3 –OH) has a higher retention time than 25 (2 –OH), but 25 has higher retention than 16 (1
 449 –OH). The same behaviour can also be seen for more complex monomers, such as 29 (2 –OH) and 23
 450 (1 –OH). However, this polar interaction can be sterically hindered by other functional groups present
 451 in the structure of the compound. For example, standard 6 eluted earlier in the SFC dimension than its
 452 positional isomer 10 because, in the structure of 6, both methyl groups are next to the hydroxyl group,
 453 weakening the polar interaction. Retention of standards 25 and 28 demonstrates the same effect even for
 454 compounds that have just –OH functional groups. The steric effect on SFC retention behaviour can also
 455 be observed for non-isomeric standards, such as 9 and 16, 29 and 30. Although we presume that the
 456 same interactions affect the separation of oligomers, without standards, we cannot propose a sectioning
 457 for oligomers.

458 The utilisation of the RPLC dimension in ¹D enables additional separation based on hydrophobicity.
 459 This can be seen, for example, when studying the log*P* (octanol-water partition coefficient) values and
 460 retention of standard 14 compared to 13 and 15 in the direction of the LC axis. Based on the publication
 461 by Smith *et al.*⁴⁸, for standard 14, a log*P* value of 2.50 has been found, which is in correspondence with
 462 the higher retention in LC compared to standards 13 and 15, with both having a log*P* value of 1.97.
 463 Moreover, the F5 (pentafluorophenyl propyl) phase is known for its combination of hydrophobic,
 464 aromatic, and steric separation mechanisms which also contributed to the improved selectivity of
 465 isomeric species such as monomers 4 and 8. Therefore, the developed 2D LC×SFC method enables the
 466 separation of compounds based on hydrophobicity, the number of hydroxyl groups, steric effects,
 467 position and number of double bonds. Having this kind of various interactions can be especially
 468 beneficial for the (improved) separation of positional isomers.

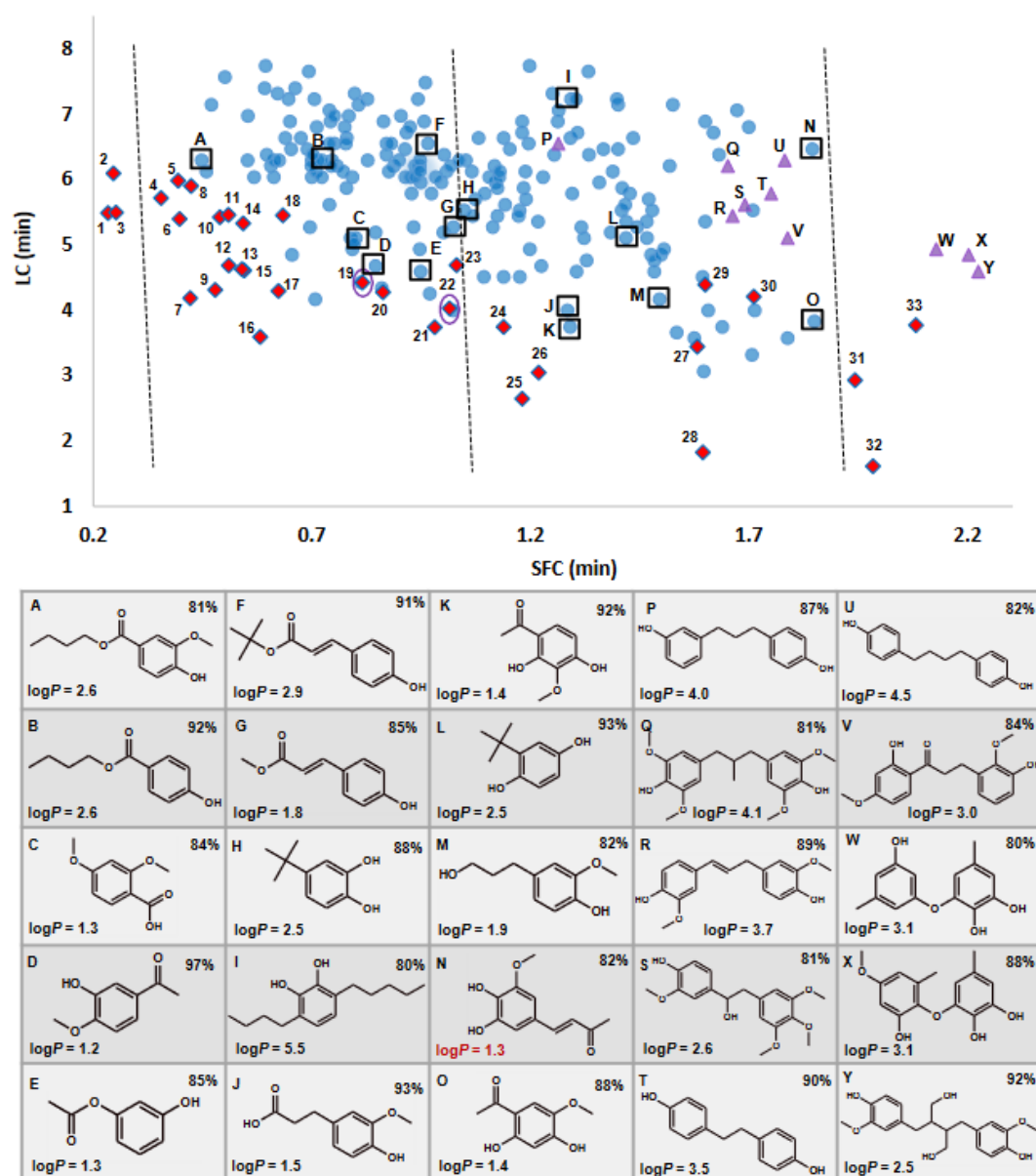
469

470 3.6 Elucidation of the detected compounds in the depolymerised lignin sample

471 The most straightforward way to determine whether a compound is present in the depolymerised lignin
472 sample is to use reference standards, such as the same 33 standard compounds selected based on
473 previous studies. Thanks to the multi-dimensional LC×SFC-HRMS/MS method, the elucidation of
474 compounds can be made based on four criteria – retention times in RPLC and SFC, accurate mass, and
475 fragmentation pattern. As has been proposed by Schymanski *et al.*⁴⁹, this yields a confidence level of 1,
476 and as the authors say, the application of an orthogonal method (as we have done here by applying both
477 LC and SFC) confirms this confidence level. For determining the t_R reproducibility of the SFC
478 dimension, Mix A containing monomer standards nr 4, 5, 11, 16, 20, and 28 was analysed during three
479 days, each with a repetition of five. Using the EIC signal of each molecular ion, a within-lab
480 reproducibility value of 0.012 min was calculated (3-fold the mean of t_R standard deviations). Because
481 of the fraction collection, one sampling time was selected as the reproducibility of the 1D RPLC
482 dimension. Therefore, the following criteria were applied: *a*) t_R in RPLC (± 0.085 min); *b*) t_R in SFC
483 (± 0.012 min); *c*) accurate mass of precursor (± 3 ppm); and *d*) MS² match over 90% in SIRIUS.
484 Interestingly, only two of the 33 standards (19: acetosyringone and 22: 4-hydroxyacetophenone) were
485 identified within the lignin sample, matching all criteria (overlay can be seen in Figure 6). No other
486 standard matched even the first three criteria. This demonstrates that because of the complexity of a
487 natural depolymerised sample, it is very difficult to obtain corresponding standards for identification
488 without performing some preliminary structural characterisation.

489 In the interest of obtaining more information about the detected compounds in the depolymerised lignin
490 sample, SIRIUS software was used. In addition to the software provided match scores with potential
491 candidates, knowing also the retention mechanisms in the two dimensions aids to give additional
492 assessment for the proposed candidates. As hydrophobicity is the driving force for separation in the
493 RPLC dimension, the $\log P$ value of the candidate has to be reasonable in relation to the t_R . In the SFC
494 dimension, when using the 1-AA stationary phase, the number and position of hydroxyl groups have an
495 effect on the retention, which can aid with the structural characterisation of similar, or even isomeric,
496 compounds. Overall, these aspects can help evaluate the suggested candidates and propose some
497 tentative structures. Therefore, in Figure 6, the tentative structures for some of the monomers and dimers
498 detected in the depolymerised lignin sample are proposed.

499 In the case of monomeric compounds, it was easier to find correlations and exclude illogical candidates
500 because of their simpler structure. A similar rough sectioning based on the number of hydroxyl groups
501 that was seen with standards was also seen with the monomeric compounds (black dashed lines in Figure
502 6). As there are no compounds in the area of none or three –OH groups, we can presume that all
503 monomers had at least one or two hydroxyl group(s). Compared to the monomeric standards, the
504 monomers in the real-life sample (blue dots in Figure 6) had generally higher retention in the RPLC
505 dimension, which suggests that those compounds are more hydrophobic. These two observations were
506 confirmed by studying their possible structures with the SIRIUS software. For example, as seen in
507 Figure 6, the possible structures of compounds A, B, F, and I all have more nonpolar side chains that
508 lead to higher hydrophobicity, which in turn leads to higher retention in RPLC. This can aid with the
509 separation of isomers, as can be seen with compound G, which has the same formula as standard 23
510 (C₁₀H₁₀O₃). These compounds showed similar retention in SFC but good separation in RPLC. In the
511 SFC dimension, isomers K and O had highly similar fragmentation patterns, leading to the suggestion
512 of the same around 10 positional isomers for both compounds with a SIRIUS match factor over 80%.
513 However, studying the position on the 2D plot suggests that these compounds have similar $\log P$ values
514 but different hindrances of the –OH groups. Using these characteristics, we propose tentative structures
515 for the two compounds, where the –OH groups are more hindered in compound K and less in compound
516 O. In terms of $\log P$ value, for the presented compounds only N has a very low value. This again suggest
517 that the real structure of the detected compound is absent from the used databases.



518

519 *Figure 6.* LC x SFC-HRMS plot of monomer standards (red diamonds with numbers) and some of the compounds
 520 detected in the depolymerised lignin sample: all compounds having DBE 4-7 (blue dots) and some of the
 521 compounds having DBE 8-11 (purple triangles). For monomer standards, the retention data was obtained from 1D
 522 LC and 1D SFC analyses and combined into a 2D plot. For the selected lignin compounds, the retention times are
 523 the same as in Figures 2 and 4. The dashed black lines indicate the rough sectioning for monomeric compounds
 524 based on the number of hydroxyl groups. The monomers that could be identified in the real-world sample by using
 525 corresponding standards are highlighted with purple ovals. Black rectangles and letters A to Y highlight some
 526 unknowns for which we propose one of the possible structures that had a SIRIUS match over 80%, reasonable
 527 computed logP value and number/position of hydroxyl groups. The logP values were obtained from the SIRIUS
 528 software and correspond to the XLogP values calculated using the Chemical Development Kit CDK. The letters
 529 of the compounds correspond to the following ID numbers shown in Table S3: A – ID 279; B – ID 256; C – ID
 530 60; D – ID 30; E – ID 25; F – ID 345; G – ID 77; H – ID 117; I – ID 437; J – ID 11; K – ID 6; L – ID 59; M – ID
 531 14; N – ID 324; O – ID 9; P – ID 346; Q – ID 270; R – ID 108; S – ID 148; T – ID 166; U – ID 280; V – ID 66;
 532 W – ID 47; X – ID 39; Y – ID 28.

533 For dimers and higher oligomers, the interpretation of elution order is more complex as there is an
 534 interplay of the interaction factors, none of which is largely dominating. As those compounds have more
 535 aromatic rings than monomers, the elution in the SFC dimension is shifted to higher retention times (as
 536 can be seen by the classifications in Figure 4). The higher number of double bonds also leads to a higher

537 probability of having longer conjugated systems in addition to less straightforwardly interpretable steric
538 effects, meaning that the –OH zones were not as clear. The absence of oligomer structures from
539 databases also amplifies the complexity. However, the same ordering of compounds based on interaction
540 mechanisms that were seen for monomeric compounds was also seen for some of the possible dimers.
541 Figure 6 shows some examples of tentative dimer structural elucidation. Compound U had higher
542 retention in the RPLC dimension than compound T, presumably because of the longer alkane chain.
543 Compound P is a similar compound to T and U, but because of the lower retention in the SFC dimension,
544 we presume that the positions of the two –OH groups have to be located so that there are more steric
545 effects impeding the polar interactions. In the structure of compound P, probably the positions of the –
546 OH groups are located so that there is even more hindrance; however, for this compound, there were
547 only a few isomers available in the databases used by the SIRIUS software.

548 The structures of the tentatively annotated monomers and dimers have quite unusual structures
549 containing saturated side chains and linkages. However, the sample analysed in this work was obtained
550 from the partial depolymerisation of lignin that was performed under hydrogen atmosphere, which can
551 cause reduction of double bonds. Of course, to confirm these conclusions about the structures of
552 potential dimers, again standards or databases including more lignin compounds would be needed.

553 Overall, involving an annotation software in the data analysis helped to elucidate the structures of
554 unknown compounds, which in turn demonstrated the usefulness of the 2D chromatographic separation
555 method for the non-target analysis of complex mixtures. Besides powerful separation, the combined
556 LC×SFC-HRMS/MS method also helps to reduce the number of candidates for each feature by
557 incorporating the logic of the retention mechanisms.

558

559 3.7 Comparison of the 2D LC×SFC method to 1D SFC

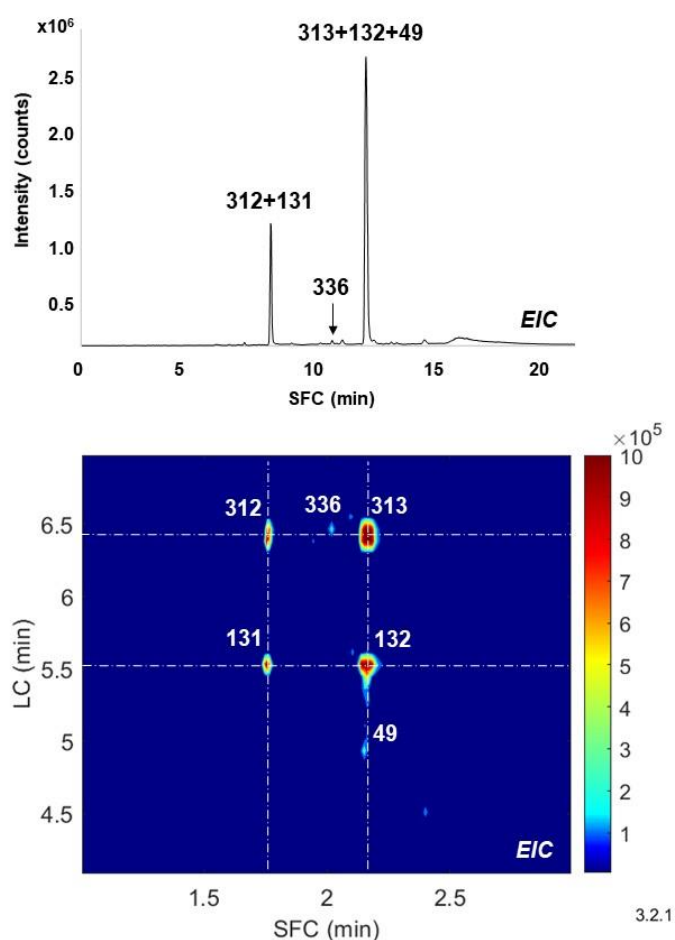
560 The developed LC×SFC method was compared to a long (gradient time 30 min) 1D SFC separation of
561 the same depolymerised lignin sample. For the 1D method, SFC was selected instead of RPLC to have
562 the same ionisation conditions as the 2D method, thus ensuring a fair comparison independent of
563 ionisation (in some cases, SFC has contributed to an improved ionisation).¹² As discussed before, the
564 corrected peak capacity of the LC×SFC method was 1995. For the 1D SFC separation of the
565 depolymerised lignin sample, a much lower value – 155 – was found for the peak capacity. Therefore,
566 by using the 2D setup, more than 12-fold higher peak capacity could be obtained for this sample.

567 3.7.1 Separation of isomers

568 In addition to miss-annotations caused by unwanted in-source fragmentation, the poor chromatographic
569 separation of isomers can also lead to false annotations.⁴⁴ Besides observing a decrease in co-elutions
570 for compounds with different accurate masses, isomers also had better separation in 2D. This can be
571 especially beneficial for the analysis of depolymerised lignin, as out of the detected $C_xH_yO_z$ compounds
572 with DBE ≥ 4 , almost 77% had at least one isomer (361 features out of 471) detected in the same sample
573 (Table S3 in the SI). For example, for deprotonated $C_{18}H_{20}O_5$, there were even 12 isomers detected in
574 the sample. In terms of discriminating isomers, one of the following criteria has to be fulfilled. The
575 straightforward case is when the retention times of the isomers are sufficiently different, thus giving
576 MS^2 spectra corresponding only to one of the compounds.¹⁸ However, when the isomers co-elute, the
577 MS^2 spectra will have a mixture of product ions originating from two compounds. In that case,
578 researchers have demonstrated that the co-eluting compounds can be differentiated when *i*) they have
579 characteristic fragmentation or *ii*) corresponding standards can be used to elucidate their mass
580 spectrometric fragmentation pathways.^{50–53} If one of these criteria is not met, then not even highly
581 powerful data analysis tools can currently aid with the deconvolution.⁵¹ In the case of target analysis
582 with standards, it is possible to use on-line LC×LC with heart-cutting mode allowing the known co-

583 elutions in ¹D to be separated with a shallower gradient in the ²D.⁵³ However, multiple heart-cutting
584 mode is technically limited in on-line configuration and so the detection of co-eluting isomers is
585 especially complex in the case of non-target analysis as it is impossible to detect (or even know) co-
586 elution of isomers when there is no evidence to suspect their presence in the 1D separation.
587 Comprehensive 2D chromatography analysis involving a second dimension with high isomer resolution
588 requires a shallow ²D gradient that so far only off-line configuration can provide.

589 Although, in some cases, SFC has shown to have a more efficient separation of isomers than RPLC,⁵⁴
590 in our study, we found that both LC and SFC dimensions are necessary for the analysis of depolymerised
591 lignin. For example, in Figure 7, when extracting the *m/z* value of 341.10263, only two major peaks are
592 seen on the 1D chromatogram, whereas the LC×SFC plot has four major peaks. As can be seen by
593 looking at the parallel lines on the EIC of the 2D plot in Figure 7, in the case of 1D analysis, there would
594 be co-elution of these isomers, regardless of whether RPLC or SFC would be used. Using MS-DIAL for
595 the 1D-SFC data also detected the same two peaks as two spots because the software spots precursor
596 ions peaks by monitoring two continuous data axes: retention time (*t_R*) and accurate mass (*m/z*).¹⁶
597 However, applying the developed 2D LC×SFC method, it was possible to detect six isomers for this
598 formula. This included a compound with a much lower intensity (ID 49) compared to its isomers that
599 would have co-elution when only the SFC dimension would be used (ID 313 and 132). Therefore, using
600 the 2D separation method yields MS data that corresponds to the separated isomer (and not a
601 combination of ions from the co-elution) and also enables the detection of much less intense isomers.



602

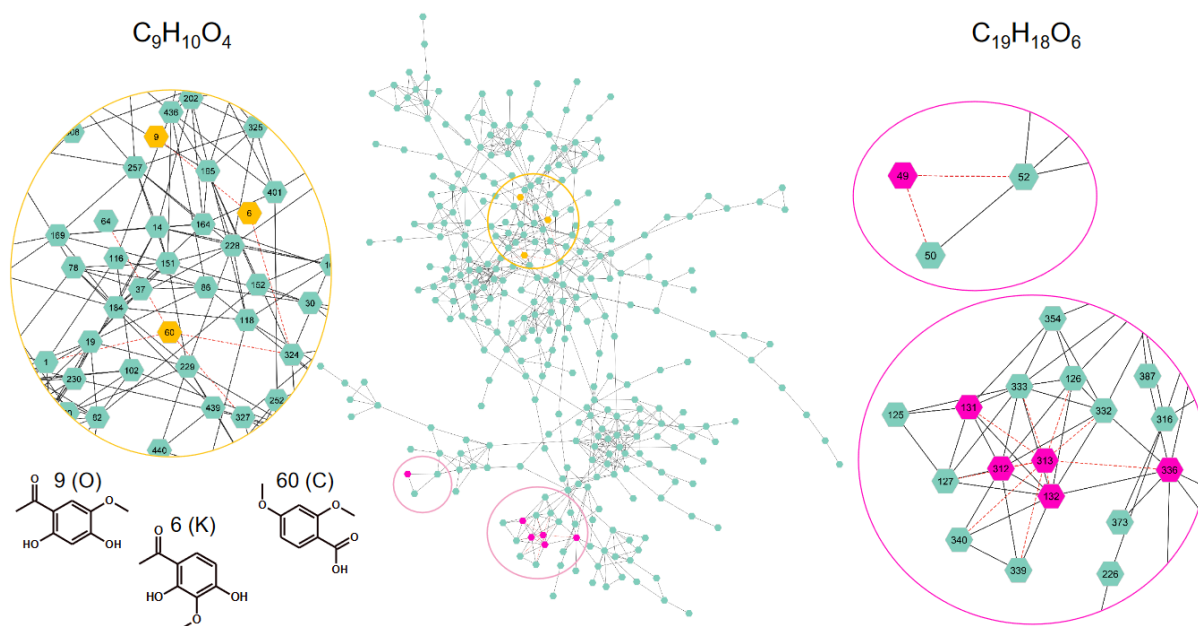
603 *Figure 7.* Comparison of 1D SFC-HRMS (top row) to 2D LC×SFC-HRMS (bottom row) based on the separation
604 of isomers. A value of 341.10263 *m/z* (± 5 ppm) was used to construct the EICs, which corresponds to deprotonated
605 C₁₉H₁₈O₆ and DBE of 11. The ID numbers of detected isomers (312, 131, 336, 313, 132, and 49) correspond to
606 the numbering presented in Table S3 in the SI. The BPC of the 1D SFC analysis is shown in Figure S2 in the SI.

607

608 3.8 FBMN – Feature-Based Molecular Network

609 As the final step, Molecular Network (MN) was applied. MN connects together features that have similar
610 fragmentation; thus, we can obtain more information about their structural similarities. For LC×SFC
611 data, MN has been used by Wei *et al.*²⁴, where precursor ions within a defined mass tolerance range and
612 nearly identical MS² spectra were merged into a single node. This was beneficial for the removal of
613 features that are caused by the collection of the same analyte into consecutive fractions. However, MN
614 would also fuse together those isomers with similar MS² that have *i*) different retention time in ²D or *ii*)
615 the same retention time in ²D but baseline separation in ¹D (features in non-consecutive fractions). This
616 issue can be overcome by using FBMN (a type of MN), which enables the visualisation of features with
617 the same accurate mass by also including MS¹ information, such as *t_R*.³² Such an approach can be
618 especially beneficial for the analysis of depolymerised lignin, as the sample analysed in our study
619 contained a very high number (approx. 77%) of isomers.

620 In Figure 8, the FBMN for the depolymerised lignin sample analysed with the 2D LC×SFC-HRMS/MS
621 method is presented. With the applied generic GNPS parameters, 318 compounds out of the 471 were
622 connected in the main network (shown in Figure 8), and the other compounds were in clusters where
623 there were less than five nodes. In orange are highlighted all features corresponding to deprotonated
624 C₉H₁₀O₄, which were noted with O (ID 9), K (ID 6), and C (ID 60) in Figure 6. It was seen that feature
625 9 is directly connected to feature 6 (red dashed line), which supports the proposition that these
626 compounds are positional isomers, as was discussed in section 3.6. Although compound 60 (noted as C
627 in Figure 6) also has the same formula as the previous two compounds, it is connected to feature 6
628 through another non-isomer feature. This is consistent with the structural candidates as feature 60 seems
629 to have different side groups and is probably not a positional isomer to 9 and 6.



630

631 *Figure 8.* Feature-Based Molecular Network (FBMN) of the depolymerised lignin sample analysed with the
632 LC×SFC-HRMS/MS technique, and visualised in Cytoscape. All the detected features corresponding to
633 deprotonated C₉H₁₀O₄ and C₁₉H₁₈O₆ are shown in orange and pink, respectively. All the edges showing all
634 connections for features with ID 9, 6, 60, 49, and 313 are highlighted with red dashed lines. For the compounds
635 highlighted with orange, the tentative structural candidates are presented, where the number corresponds to the
636 feature ID (Table S3 in SI) and the letter to the notation used in Figure 6.

637 Applying the same argumentation, we can also study the location of other isomers, for example, isomers
638 corresponding to deprotonated $C_{19}H_{18}O_6$, which was also the example in Figure 7. In Figure 8, all
639 isomers matching this formula are highlighted with pink colour. As can be seen with the red dashed
640 lines, 313 is directly connected to four of the isomers, demonstrating that five out of six isomers have
641 similar MS^2 patterns. This suggests that all except compound 49 are highly similar, possibly positional,
642 isomers. Therefore, FBMN has demonstrated to rapidly define and visualise if the isomers are
643 structurally highly similar (including positional isomers) or more different from each other.

644

645 4. Conclusions

646 In this work, a powerful comprehensive off-line 2D RPLC×SFC-HRMS/MS method was applied for
647 the non-target analysis of depolymerised lignin. The method enabled the detection of 471 $C_xH_yO_z$
648 compounds with $DBE \geq 4$, where the highest DBE value was 17 (classically corresponding to tetramers).
649 In addition, the 2D plot showed a classification of the lignin compounds based on their DBE range
650 largely corresponding to the number of aromatic rings in the structure. This kind of classification can be
651 used to find outliers (as was demonstrated here) or to compare different samples to each other.

652 The analysis of 33 monomeric standards enabled to investigate the 2D separation mechanism in more
653 detail. It was seen that the 2D LC×SFC separation was based on hydrophobicity, number of hydroxyl
654 groups, steric effects, and position/number of double bonds in the structure of the analyte. Understanding
655 the separation mechanism helped to characterise the LC×SFC-HRMS/MS results of the depolymerised
656 lignin and evaluate the structural candidates proposed by the SIRIUS software. In addition, the 2D
657 approach yielded clean MS^1 and MS^2 spectra that could be more easily analysed with mass spectral
658 analysis softwares (in this study, MS-DIAL, SIRIUS, and FBMN). However, it was seen that there is a
659 need for additional developments to improve the structural characterisation of lignin compounds with
660 annotation softwares.

661 Compared to a long 1D SFC method, the 2D technique provided an improved separation of compounds,
662 including isomers. In terms of lignin analysis, this is highly beneficial, as 77% of the detected features
663 had at least one corresponding isomer detected in the same sample. The FBMN additionally helped to
664 study the possible structural candidates of the isomers based on their MS^2 spectral similarities. The
665 established strategy could also suit the (non-target) analysis of other isomer-rich mixtures with neutral
666 compounds. Therefore, we have demonstrated that powerful chromatographic methods are needed (such
667 as the one applied here) for a more reliable identification based on mass spectral data. In relation to
668 isomers, this could be especially beneficial as the annotation of co-eluting isomers is still an unsolved
669 challenge.

670

671 Acknowledgments

672 This work was funded by the French National Research Agency (project BIOPOLIOL, ANR-21-CE43-
673 0026). The authors would like to thank Dr. Pilleriin Peets and Ph.D. candidate Helen Sepman for fruitful
674 discussions about the data analysis involving MS-DIAL and SIRIUS softwares. The anonymous
675 reviewers are also acknowledged for their constructive comments and help in improving the manuscript.

- 677 (1) Constant, S.; Wienk, H. L. J.; Frissen, A. E.; Peinder, P. de; Boelens, R.; Es, D. S. van; Grisel, R.
678 J. H.; Weckhuysen, B. M.; Huijgen, W. J. J.; Gosselink, R. J. A.; Bruijninx, P. C. A. New
679 Insights into the Structure and Composition of Technical Lignins: A Comparative
680 Characterisation Study. *Green Chem.* **2016**, *18* (9), 2651–2665.
681 <https://doi.org/10.1039/C5GC03043A>.
- 682 (2) Sethupathy, S.; Murillo Morales, G.; Gao, L.; Wang, H.; Yang, B.; Jiang, J.; Sun, J.; Zhu, D.
683 Lignin Valorization: Status, Challenges and Opportunities. *Bioresour. Technol.* **2022**, *347*,
684 126696. <https://doi.org/10.1016/j.biortech.2022.126696>.
- 685 (3) Bertella, S.; Luterbacher, J. S. Lignin Functionalization for the Production of Novel Materials.
686 *Trends Chem.* **2020**, *2* (5), 440–453. <https://doi.org/10.1016/j.trechm.2020.03.001>.
- 687 (4) Tammekivi, E.; Geantet, C.; Lorentz, C.; Faure, K. Two-Dimensional Chromatography for the
688 Analysis of Valorisable Biowaste: A Review. *Anal. Chim. Acta* **2023**, 341855.
689 <https://doi.org/10.1016/j.aca.2023.341855>.
- 690 (5) Prothmann, J.; Spégel, P.; Sandahl, M.; Turner, C. Identification of Lignin Oligomers in Kraft
691 Lignin Using Ultra-High-Performance Liquid Chromatography/High-Resolution Multiple-Stage
692 Tandem Mass Spectrometry (UHPLC/HRMSn). *Anal. Bioanal. Chem.* **2018**, *410* (29), 7803–
693 7814. <https://doi.org/10.1007/s00216-018-1400-4>.
- 694 (6) Sun, M.; Sandahl, M.; Turner, C. Comprehensive On-Line Two-Dimensional Liquid
695 Chromatography × supercritical Fluid Chromatography with Trapping Column-Assisted
696 Modulation for Depolymerised Lignin Analysis. *J. Chromatogr. A* **2018**, *1541*, 21–30.
697 <https://doi.org/10.1016/j.chroma.2018.02.008>.
- 698 (7) Joffres, B.; Nguyen, M. T.; Laurenti, D.; Lorentz, C.; Souchon, V.; Charon, N.; Daudin, A.;
699 Quignard, A.; Geantet, C. Lignin Hydroconversion on MoS₂-Based Supported Catalyst:
700 Comprehensive Analysis of Products and Reaction Scheme. *Appl. Catal. B Environ.* **2016**, *184*,
701 153–162. <https://doi.org/10.1016/j.apcatb.2015.11.005>.
- 702 (8) Abdelaziz, O. Y.; Brink, D. P.; Prothmann, J.; Ravi, K.; Sun, M.; García-Hidalgo, J.; Sandahl, M.;
703 Hulteberg, C. P.; Turner, C.; Lidén, G.; Gorwa-Grauslund, M. F. Biological Valorization of Low
704 Molecular Weight Lignin. *Biotechnol. Adv.* **2016**, *34* (8), 1318–1346.
705 <https://doi.org/10.1016/j.biotechadv.2016.10.001>.
- 706 (9) Staš, M.; Auersvald, M.; Vozka, P. Two-Dimensional Gas Chromatography Characterization of
707 Pyrolysis Bio-Oils: A Review. *Energy Fuels* **2021**, *35* (10), 8541–8557.
708 <https://doi.org/10.1021/acs.energyfuels.1c00553>.
- 709 (10) Donato, P.; Cacciola, F.; Mondello, L.; Giuffrida, D.; Beccaria, M.; Dugo, P.; Domenica
710 Mangraviti; Ivana Bonaccorsi; Margita Utczas. Analysis of the Carotenoid Composition and
711 Stability in Various Overripe Fruits by Comprehensive Two-Dimensional Liquid
712 Chromatography. *LCGC Eur.* **2016**, *29* (5), 252–257.
- 713 (11) Oliveira Lago, L.; Swit, P.; Moura da Silva, M.; Telles Biasoto Marques, A.; Welke, J.; Montero,
714 L.; Herrero, M. Evolution of Anthocyanin Content during Grape Ripening and Characterization of
715 the Phenolic Profile of the Resulting Wine by Comprehensive Two-Dimensional Liquid
716 Chromatography. *J. Chromatogr. A* **2023**, *1704*, 464131.
717 <https://doi.org/10.1016/j.chroma.2023.464131>.
- 718 (12) Bulet-Parendel, M.; Faure, K. Opportunities and Challenges of Liquid Chromatography Coupled
719 to Supercritical Fluid Chromatography. *TrAC Trends Anal. Chem.* **2021**, *144*, 116422.
720 <https://doi.org/10.1016/j.trac.2021.116422>.
- 721 (13) Sarrut, M.; Corgier, A.; Crétier, G.; Le Masle, A.; Dubant, S.; Heinisch, S. Potential and
722 Limitations of On-Line Comprehensive Reversed Phase Liquid Chromatography × supercritical
723 Fluid Chromatography for the Separation of Neutral Compounds: An Approach to Separate an
724 Aqueous Extract of Bio-Oil. *J. Chromatogr. A* **2015**, *1402*, 124–133.
725 <https://doi.org/10.1016/j.chroma.2015.05.005>.
- 726 (14) Nijssen, R.; Blokland, M. H.; Wegh, R. S.; de Lange, E.; van Leeuwen, S. P. J.; Berendsen, B. J.
727 A.; van de Schans, M. G. M. Comparison of Compound Identification Tools Using Data
728 Dependent and Data Independent High-Resolution Mass Spectrometry Spectra. *Metabolites* **2023**,
729 *13* (7), 777. <https://doi.org/10.3390/metabo13070777>.

- 730 (15) Caño-Carrillo, I.; Gilbert-López, B.; Montero, L.; Martínez-Piernas, A. B.; García-Reyes, J. F.;
731 Molina-Díaz, A. Comprehensive and Heart-Cutting Multidimensional Liquid Chromatography–
732 Mass Spectrometry and Its Applications in Food Analysis. *Mass Spectrom. Rev.* n/a (n/a).
733 <https://doi.org/10.1002/mas.21845>.
- 734 (16) Tsugawa, H.; Cajka, T.; Kind, T.; Ma, Y.; Higgins, B.; Ikeda, K.; Kanazawa, M.; VanderGheynst,
735 J.; Fiehn, O.; Arita, M. MS-DIAL: Data-Independent MS/MS Deconvolution for Comprehensive
736 Metabolome Analysis. *Nat. Methods* **2015**, *12* (6), 523–526. <https://doi.org/10.1038/nmeth.3393>.
- 737 (17) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.;
738 Watrous, J.; Kapon, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.;
739 Meehan, M. J.; Liu, W.-T.; Crüsemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderón,
740 M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C.-C.; Floros, D. J.; Gavilan,
741 R. G.; Kleigrew, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen,
742 C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.;
743 Gerwick, L.; Liaw, C.-C.; Yang, Y.-L.; Humpf, H.-U.; Maansson, M.; Keyzers, R. A.; Sims, A.
744 C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; Boya P, C. A.;
745 Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.;
746 O’Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson,
747 H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.;
748 McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.;
749 Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.;
750 Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodríguez, A. M. C.;
751 Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P.-M.; Phapale, P.;
752 Nothias, L.-F.; Alexandrov, T.; Litaudon, M.; Wolfender, J.-L.; Kyle, J. E.; Metz, T. O.; Peryea,
753 T.; Nguyen, D.-T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Müller, R.; Waters, K. M.; Shi, W.; Liu,
754 X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B. Ø.; Pogliano, K.; Linington, R. G.; Gutiérrez,
755 M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N. Sharing and
756 Community Curation of Mass Spectrometry Data with Global Natural Products Social Molecular
757 Networking. *Nat. Biotechnol.* **2016**, *34* (8), 828–837. <https://doi.org/10.1038/nbt.3597>.
- 758 (18) Olivon, F.; Apel, C.; Retailleau, P.; M. Allard, P.; L. Wolfender, J.; Touboul, D.; Roussi, F.;
759 Litaudon, M.; Desrat, S. Searching for Original Natural Products by Molecular Networking:
760 Detection, Isolation and Total Synthesis of Chloroaustralasines. *Org. Chem. Front.* **2018**, *5* (14),
761 2171–2178. <https://doi.org/10.1039/C8QO00429C>.
- 762 (19) Pan, H.; Zhou, H.; Miao, S.; Cao, J.; Liu, J.; Lan, L.; Hu, Q.; Mao, X.; Ji, S. An Integrated
763 Approach for Global Profiling of Multi-Type Constituents: Comprehensive Chemical
764 Characterization of Lonicerae Japonicae Flos as a Case Study. *J. Chromatogr. A* **2020**, *1613*,
765 460674. <https://doi.org/10.1016/j.chroma.2019.460674>.
- 766 (20) Wang, M.; Xu, X.; Wang, H.; Wang, H.; Liu, M.; Hu, W.; Chen, B.; Jiang, M.; Qi, J.; Li, X.;
767 Yang, W.; Gao, X. A Multi-Dimensional Liquid Chromatography/High-Resolution Mass
768 Spectrometry Approach Combined with Computational Data Processing for the Comprehensive
769 Characterization of the Multicomponents from Cuscuta Chinensis. *J. Chromatogr. A* **2022**, *1675*,
770 463162. <https://doi.org/10.1016/j.chroma.2022.463162>.
- 771 (21) Zhang, Z.; Xu, Y.; Shen, A.; Fu, D.; Liu, D.; Liu, Y.; Liang, X. Offline Two-Dimensional
772 Normal-Phase × Reversed-Phase Liquid Chromatography Coupled with High-Resolution Mass
773 Spectrometry for Comprehensive Analysis of Chemical Constituents in Euphorbia Kansui. *J.*
774 *Chromatogr. A* **2023**, *1693*, 463897. <https://doi.org/10.1016/j.chroma.2023.463897>.
- 775 (22) Dai, Y.; Zhang, K.; Xiong, L.; Wang, L.; Guo, Z.; Yang, J.; Wu, A.; Wu, J.; Zeng, J.
776 Comprehensive Profiling of Sanguisorba Officinalis Using Offline Two-Dimensional Mixed-
777 Mode Liquid Chromatography × Reversed-Phase Liquid Chromatography, Tandem High-
778 Resolution Mass Spectrometry, and Molecular Network. *J. Sep. Sci.* **2022**, *45* (10), 1727–1736.
779 <https://doi.org/10.1002/jssc.202200013>.
- 780 (23) Qu, B.; Liu, Y.; Shen, A.; Guo, Z.; Yu, L.; Liu, D.; Huang, F.; Peng, T.; Liang, X. Combining
781 Multidimensional Chromatography-Mass Spectrometry and Feature-Based Molecular Networking
782 Methods for the Systematic Characterization of Compounds in the Supercritical Fluid Extract of
783 Tripterygium Wilfordii Hook F. *Analyst* **2023**, *148* (1), 61–73.
784 <https://doi.org/10.1039/D2AN01471H>.

- 785 (24) Wei, W.; Hou, J.; Yao, C.; Bi, Q.; Wang, X.; Li, Z.; Jin, Q.; Lei, M.; Feng, Z.; Wu, W.; Guo, D. A
786 High-Efficiency Strategy Integrating Offline Two-Dimensional Separation and Data Post-
787 Processing with Dereplication: Characterization of Bufadienolides in Venenum Bufonis as a Case
788 Study. *J. Chromatogr. A* **2019**, *1603*, 179–189. <https://doi.org/10.1016/j.chroma.2019.06.037>.
- 789 (25) Teboul, E.; Tammekivi, E.; Batteau, M.; Geantet, C.; Faure, K. Off-Line Two-Dimensional
790 Separation Involving Supercritical Fluid Chromatography for the Characterization of the
791 Wastewater from Algae Hydrothermal Liquefaction. *J. Chromatogr. A* **2023**, *1694*, 463907.
792 <https://doi.org/10.1016/j.chroma.2023.463907>.
- 793 (26) Sipilä, K.; Kuoppala, E.; Fagernäs, L.; Oasmaa, A. Characterization of Biomass-Based Flash
794 Pyrolysis Oils. *Biomass Bioenergy* **1998**, *14* (2), 103–113. [https://doi.org/10.1016/S0961-](https://doi.org/10.1016/S0961-9534(97)10024-1)
795 [9534\(97\)10024-1](https://doi.org/10.1016/S0961-9534(97)10024-1).
- 796 (27) Branca, C.; Giudicianni, P.; Di Blasi, C. GC/MS Characterization of Liquids Generated from
797 Low-Temperature Pyrolysis of Wood. *Ind. Eng. Chem. Res.* **2003**, *42* (14), 3190–3202.
798 <https://doi.org/10.1021/ie030066d>.
- 799 (28) Le Masle, A.; Angot, D.; Gouin, C.; D'Attoma, A.; Ponthus, J.; Quignard, A.; Heinisch, S.
800 Development of On-Line Comprehensive Two-Dimensional Liquid Chromatography Method for
801 the Separation of Biomass Compounds. *J. Chromatogr. A* **2014**, *1340*, 90–98.
802 <https://doi.org/10.1016/j.chroma.2014.03.020>.
- 803 (29) Guillarme, D.; Desfontaine, V.; Heinisch, S.; Veuthey, J.-L. What Are the Current Solutions for
804 Interfacing Supercritical Fluid Chromatography and Mass Spectrometry? *J. Chromatogr. B* **2018**,
805 *1083*, 160–170. <https://doi.org/10.1016/j.jchromb.2018.03.010>.
- 806 (30) Tsugawa, H.; Kind, T.; Nakabayashi, R.; Yukihiro, D.; Tanaka, W.; Cajka, T.; Saito, K.; Fiehn,
807 O.; Arita, M. Hydrogen Rearrangement Rules: Computational MS/MS Fragmentation and
808 Structure Elucidation Using MS-FINDER Software. *Anal. Chem.* **2016**, *88* (16), 7946–7958.
809 <https://doi.org/10.1021/acs.analchem.6b00770>.
- 810 (31) Montero, L.; Meckelmann, S. W.; Kim, H.; Ayala-Cabrera, J. F.; Schmitz, O. J. Differentiation of
811 Industrial Hemp Strains by Their Cannabinoid and Phenolic Compounds Using LC × LC-HRMS.
812 *Anal. Bioanal. Chem.* **2022**, *414* (18), 5445–5459. <https://doi.org/10.1007/s00216-022-03925-8>.
- 813 (32) Nothias, L.-F.; Petras, D.; Schmid, R.; Dührkop, K.; Rainer, J.; Sarvepalli, A.; Protsyuk, I.; Ernst,
814 M.; Tsugawa, H.; Fleischauer, M.; Aicheler, F.; Aksenov, A. A.; Alka, O.; Allard, P.-M.; Barsch,
815 A.; Cachet, X.; Caraballo-Rodriguez, A. M.; Da Silva, R. R.; Dang, T.; Garg, N.; Gauglitz, J. M.;
816 Gurevich, A.; Isaac, G.; Jarmusch, A. K.; Kameník, Z.; Kang, K. B.; Kessler, N.; Koester, I.;
817 Korf, A.; Le Gouvellec, A.; Ludwig, M.; Martin, H. C.; McCall, L.-I.; McSayles, J.; Meyer, S. W.;
818 Mohimani, H.; Morsy, M.; Moyne, O.; Neumann, S.; Neuweger, H.; Nguyen, N. H.; Nothias-
819 Esposito, M.; Paolini, J.; Phelan, V. V.; Pluskal, T.; Quinn, R. A.; Rogers, S.; Shrestha, B.;
820 Tripathi, A.; van der Hoof, J. J. J.; Vargas, F.; Weldon, K. C.; Witting, M.; Yang, H.; Zhang, Z.;
821 Zubeil, F.; Kohlbacher, O.; Böcker, S.; Alexandrov, T.; Bandeira, N.; Wang, M.; Dorrestein, P. C.
822 Feature-Based Molecular Networking in the GNPS Analysis Environment. *Nat. Methods* **2020**, *17*
823 (9), 905–908. <https://doi.org/10.1038/s41592-020-0933-6>.
- 824 (33) Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A. A.; Melnik, A. V.; Meusel, M.;
825 Dorrestein, P. C.; Rousu, J.; Böcker, S. SIRIUS 4: A Rapid Tool for Turning Tandem Mass
826 Spectra into Metabolite Structure Information. *Nat. Methods* **2019**, *16* (4), 299–302.
827 <https://doi.org/10.1038/s41592-019-0344-8>.
- 828 (34) Böcker, S.; Dührkop, K. Fragmentation Trees Reloaded. *J. Cheminformatics* **2016**, *8* (1), 5.
829 <https://doi.org/10.1186/s13321-016-0116-8>.
- 830 (35) Dührkop, K.; Shen, H.; Meusel, M.; Rousu, J.; Böcker, S. Searching Molecular Structure
831 Databases with Tandem Mass Spectra Using CSI:FingerID. *Proc. Natl. Acad. Sci.* **2015**, *112* (41),
832 12580–12585. <https://doi.org/10.1073/pnas.1509788112>.
- 833 (36) Djoumbou Feunang, Y.; Eisner, R.; Knox, C.; Chepelev, L.; Hastings, J.; Owen, G.; Fahy, E.;
834 Steinbeck, C.; Subramanian, S.; Bolton, E.; Greiner, R.; Wishart, D. S. ClassyFire: Automated
835 Chemical Classification with a Comprehensive, Computable Taxonomy. *J. Cheminformatics*
836 **2016**, *8* (1), 61. <https://doi.org/10.1186/s13321-016-0174-y>.
- 837 (37) Kim, H. W.; Wang, M.; Leber, C. A.; Nothias, L.-F.; Reher, R.; Kang, K. B.; van der Hoof, J. J.
838 J.; Dorrestein, P. C.; Gerwick, W. H.; Cottrell, G. W. NPClassifier: A Deep Neural Network-

- 839 Based Structural Classification Tool for Natural Products. *J. Nat. Prod.* **2021**, *84* (11), 2795–
840 2807. <https://doi.org/10.1021/acs.jnatprod.1c00399>.
- 841 (38) Dührkop, K.; Nothias, L.-F.; Fleischauer, M.; Reher, R.; Ludwig, M.; Hoffmann, M. A.; Petras,
842 D.; Gerwick, W. H.; Rousu, J.; Dorrestein, P. C.; Böcker, S. Systematic Classification of
843 Unknown Metabolites Using High-Resolution Fragmentation Mass Spectra. *Nat. Biotechnol.*
844 **2021**, *39* (4), 462–471. <https://doi.org/10.1038/s41587-020-0740-8>.
- 845 (39) Li, X.; Stoll, D. R.; Carr, P. W. Equation for Peak Capacity Estimation in Two-Dimensional
846 Liquid Chromatography. *Anal. Chem.* **2009**, *81* (2), 845–850. <https://doi.org/10.1021/ac801772u>.
- 847 (40) Semard, G.; Peulon-Agasse, V.; Bruchet, A.; Bouillon, J.-P.; Cardinaël, P. Convex Hull: A New
848 Method to Determine the Separation Space Used and to Optimize Operating Conditions for
849 Comprehensive Two-Dimensional Gas Chromatography. *J. Chromatogr. A* **2010**, *1217* (33),
850 5449–5454. <https://doi.org/10.1016/j.chroma.2010.06.048>.
- 851 (41) Prothmann, J.; Sun, M.; Spégel, P.; Sandahl, M.; Turner, C. Ultra-High-Performance Supercritical
852 Fluid Chromatography with Quadrupole-Time-of-Flight Mass Spectrometry (UHPSFC/QTOF-
853 MS) for Analysis of Lignin-Derived Monomeric Compounds in Processed Lignin Samples. *Anal.*
854 *Bioanal. Chem.* **2017**, *409* (30), 7049–7061. <https://doi.org/10.1007/s00216-017-0663-5>.
- 855 (42) Prothmann, J.; Palmer, S.; Turner, C.; Sandahl, M. Separation of Monomeric and Dimeric
856 Phenolic Compounds in Lignosulphonate Lignin on Different Stationary Phases Using Ultrahigh-
857 Performance Supercritical Fluid Chromatography. *J. Chromatogr. A* **2021**, *1653*, 462408.
858 <https://doi.org/10.1016/j.chroma.2021.462408>.
- 859 (43) Guo, J.; Huan, T. Turning Metabolomics Data Processing from a “Black Box” to a “White Box.”
860 *LCGC Suppl.* **2022**, *40* (s9), 20–22.
- 861 (44) Xu, Y.-F.; Lu, W.; Rabinowitz, J. D. Avoiding Misannotation of In-Source Fragmentation
862 Products as Cellular Metabolites in Liquid Chromatography–Mass Spectrometry–Based
863 Metabolomics. *Anal. Chem.* **2015**, *87* (4), 2273–2281. <https://doi.org/10.1021/ac504118y>.
- 864 (45) Wasito, H.; Causon, T.; Hann, S. Alternating In-Source Fragmentation with Single-Stage High-
865 Resolution Mass Spectrometry with High Annotation Confidence in Non-Targeted Metabolomics.
866 *Talanta* **2022**, *236*, 122828. <https://doi.org/10.1016/j.talanta.2021.122828>.
- 867 (46) Tada, I.; Chaleckis, R.; Tsugawa, H.; Meister, I.; Zhang, P.; Lazarinis, N.; Dahlén, B.; Wheelock,
868 C. E.; Arita, M. Correlation-Based Deconvolution (CorrDec) To Generate High-Quality MS2
869 Spectra from Data-Independent Acquisition in Multisample Studies. *Anal. Chem.* **2020**, *92* (16),
870 11310–11317. <https://doi.org/10.1021/acs.analchem.0c01980>.
- 871 (47) Jumaah, F.; Plaza, M.; Abrahamsson, V.; Turner, C.; Sandahl, M. A Fast and Sensitive Method
872 for the Separation of Carotenoids Using Ultra-High Performance Supercritical Fluid
873 Chromatography–Mass Spectrometry. *Anal. Bioanal. Chem.* **2016**, *408* (21), 5883–5894.
874 <https://doi.org/10.1007/s00216-016-9707-5>.
- 875 (48) Smith, C. J.; Perfetti, T. A.; Morton, M. J.; Rodgman, A.; Garg, R.; Selassie, C. D.; Hansch, C.
876 The Relative Toxicity of Substituted Phenols Reported in Cigarette Mainstream Smoke. *Toxicol.*
877 *Sci.* **2002**, *69* (1), 265–278. <https://doi.org/10.1093/toxsci/69.1.265>.
- 878 (49) Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J.
879 Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating
880 Confidence. *Environ. Sci. Technol.* **2014**, *48* (4), 2097–2098. <https://doi.org/10.1021/es5002105>.
- 881 (50) Karatt, T. K.; Nalakath, J.; Perwad, Z.; Albert, P. H.; Abdul Khader, K. K.; Syed Ali Padusha, M.;
882 Laya, S. Mass Spectrometric Method for Distinguishing Isomers of Dexamethasone via Fragment
883 Mass Ratio: An HRMS Approach. *J. Mass Spectrom.* **2018**, *53* (11), 1046–1058.
884 <https://doi.org/10.1002/jms.4279>.
- 885 (51) Koelmel, J. P.; Li, X.; Stow, S. M.; Sartain, M. J.; Murali, A.; Kemperman, R.; Tsugawa, H.;
886 Takahashi, M.; Vasiliou, V.; Bowden, J. A.; Yost, R. A.; Garrett, T. J.; Kitagawa, N. Lipid
887 Annotator: Towards Accurate Annotation in Non-Targeted Liquid Chromatography High-
888 Resolution Tandem Mass Spectrometry (LC-HRMS/MS) Lipidomics Using a Rapid and User-
889 Friendly Software. *Metabolites* **2020**, *10* (3), 101. <https://doi.org/10.3390/metabo10030101>.
- 890 (52) Menicatti, M.; Pallecchi, M.; Bua, S.; Vullo, D.; Di Cesare Mannelli, L.; Ghelardini, C.; Carta, F.;
891 Supuran, C. T.; Bartolucci, G. Resolution of Co-Eluting Isomers of Anti-Inflammatory Drugs
892 Conjugated to Carbonic Anhydrase Inhibitors from Plasma in Liquid Chromatography by Energy-

- 893 Resolved Tandem Mass Spectrometry. *J. Enzyme Inhib. Med. Chem.* **2018**, *33* (1), 671–679.
894 <https://doi.org/10.1080/14756366.2018.1445737>.
- 895 (53) Pua, A.; Goh, R. M. V.; Ee, K.-H.; Huang, Y.; Liu, S. Q.; Lassabliere, B.; Yu, B. Improving
896 Resolution of Isomeric Flavonoids and Their Glycosides Using Two-Dimensional Liquid
897 Chromatography Coupled With High-Resolution Mass Spectrometry. *Chromatographia* **2021**, *84*
898 (5), 507–515. <https://doi.org/10.1007/s10337-021-04027-w>.
- 899 (54) Ganzera, M.; Zwerger, M. Analysis of Natural Products by SFC – Applications from 2015 to
900 2021. *TrAC Trends Anal. Chem.* **2021**, *145*, 116463. <https://doi.org/10.1016/j.trac.2021.116463>.
901