



HAL
open science

Current trends and advances in analytical techniques for the characterization and quantification of biologically recalcitrant organic species in sludge and wastewater: A review

Sylvain Faixo, Nicolas Gehin, Stéphane Balayssac, Véronique Gilard, Sofiane Mazeghrane, Mathieu Haddad, Gilberte Gaval, Etienne Paul, Jean-Christophe Garrigues

► To cite this version:

Sylvain Faixo, Nicolas Gehin, Stéphane Balayssac, Véronique Gilard, Sofiane Mazeghrane, et al.. Current trends and advances in analytical techniques for the characterization and quantification of biologically recalcitrant organic species in sludge and wastewater: A review. *Analytica Chimica Acta*, 2021, 1152, pp.338284. 10.1016/j.aca.2021.338284 . hal-03195665

HAL Id: hal-03195665

<https://hal.science/hal-03195665v1>

Submitted on 9 Mar 2023

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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **Current trends and advances in analytical techniques for the characterization and**
2 **quantification of biologically recalcitrant organic species in sludge and wastewater: A**
3 **review**

4
5 Sylvain Faixo^{a,d}, Nicolas Gehin^b, Stéphane Balayssac^c, Véronique Gilard^c, Sofiane
6 Mazeghrane^a, Mathieu Haddad^e, Gilberte Gaval^a, Etienne Paul^d, Jean-Christophe Garrigues^{b,*}

7
8 ^a *SUEZ, Centre International de Recherche Sur l'Eau et l'Environnement (CIRSEE), 38 rue*
9 *du Président Wilson, 78230 Le Pecq, France*

10 ^b *Laboratoire des IMRCP, Université de Toulouse, CNRS UMR 5623, Université Paul*
11 *Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex 9, France*

12 ^c *Laboratoire SPCMIB, Université de Toulouse, CNRS UMR 5068, Université Paul Sabatier,*
13 *118 route de Narbonne, 31062 Toulouse Cedex 9, France*

14 ^d *TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France - 135 avenue de*
15 *Rangueil, 31077 Toulouse CEDEX 04*

16 ^e *Suez Treatment Infrastructure, 183 avenue du 18 juin 1940, 92500, Rueil Malmaison,*
17 *France*

18

19

20 * Corresponding author. E-mail address: garrigues@chimie.ups-tlse.fr

21

22

23 **Abstract**

24 The study of organic matter in wastewater is a major regulatory and environmental issue and
25 requires new developments to identify non-biodegradable refractory compounds, produced

26 mainly by thermal treatments. Recent advances linking physicochemical properties to
27 spectroscopic analyzes (UV, Fluorescence, IR) have shown that the refractory property is
28 favored by several physicochemical parameters: weight, hydrophobicity, aromaticity and
29 chemical functions. Currently, the most effective developments for the quantification of
30 refractory compounds are obtained with hyphenated methods, based on steric separation of
31 the macromolecular species by steric exclusion chromatography (SEC)/PDA/Fluorescence
32 systems. Hyphenated techniques using High Resolution Mass Spectrometry (HRMS), ultra-
33 high-resolution mass spectrometry with Fourier-transform ion cyclotron resonance mass
34 spectrometry (FT-ICR MS) and NMR have been developed to analyze macromolecules in
35 wastewater with minor sample preparation procedures. A particular class has been identified,
36 the melanoidins, generated by Maillard reactions between sugars, amino acids, peptides and
37 proteins present in wastewater and sludge, but low molecular weight compounds formed as
38 intermediates, such as ketones, aldehydes, pyrazines, pyridines or furans, are also recalcitrant
39 and are complex to identify in the complex matrices. The lack of available standards for the
40 study of these compounds requires the use of specific techniques and data processing.
41 Advances in chemometrics are obtained in the development of molecular or physicochemical
42 indices resulting from the data generated by the analytical detectors, such as aromaticity
43 calculated by $SUVA_{254}$ and determined by UV, fluorescence, molar mass, H/C ratio or
44 structural studies (measuring the amount of unsaturated carbon) given by hyphenated
45 techniques with SEC. It is clear that nitrogen compounds are widely involved in
46 refractoriness. New trends in nitrogen containing compounds characterization follow two
47 axes: through SEC/PDA/Fluorescence and HRMS/NMR techniques with or without
48 separation. Other techniques widely used in food or marine science are also being imported to
49 this study, as it can be seen in the use of "omics" methods, high-performance thin layer
50 chromatography (HPTLC) and chromatography at the critical condition, rounding out the

51 important developments around SEC. While improving the performance of stationary phases
52 is one of the challenges, it results in a fundamental understanding of the retention mechanisms
53 that today provide us with more information on the structures identified. The main objective
54 of this review is to present the spectroscopic and physicochemical techniques used to qualify
55 and characterize refractoriness with a specific focus on chemometric approaches.

56

57 **Keywords**

58 Melanoidins, recalcitrant organic compounds, wastewater, hydrothermal carbonization, size
59 exclusion chromatography, hyphenated techniques

60

61 **Contents**

62	1. Introduction	7
63	2. Trends in molecular characterization: dissolved organic matter and recalcitrant organic	
64	compounds	17
65	2.1 Recalcitrant species	17
66	2.2 Maillard reactions.....	21
67	2.3 Physicochemical and spectrophotometric analyses.....	24
68	2.3.1 Dissolved organic carbon (DOC) and nitrogen (DON).....	25
69	2.3.2 UV/Vis spectrophotometry.....	26
70	2.3.3 Fluorescence spectroscopy	27
71	2.3.4 Fourier transform infrared spectroscopy (FT-IR).....	31
72	3. Advances and trends in chromatographic and NMR approaches	34
73	3.1 Sample treatment.....	34
74	3.2 Reversed-phase high performance liquid chromatography (RP-HPLC) and size exclusion	
75	chromatography (SEC).....	39

76	3.3 Hyphenated techniques.....	48
77	3.4 High-resolution mass spectrometry.....	51
78	3.5 Nuclear magnetic resonance spectroscopy (NMR).....	63
79	4. Chemometrics	69
80	4.1 Cluster analysis (CA).....	70
81	4.2 Principal component analysis (PCA).....	71
82	4.3 Partial Least Squares (PLS) and PLS discriminant analysis (PLS-DA).....	72
83	5. Conclusions	75
84	Acknowledgements	79
85	References	80

86

87 **Abbreviations**

88 ACN: acetonitrile

89 AIC: akaike information criterion

90 ANOVA: analysis of variance

91 BBI: broad band inverse

92 CA: cluster analysis

93 HCA: hierarchical CA

94 CE: capillary electrophoresis

95 CMP: comprehensive multiphase

96 COD: chemical oxygen demand

97 COSY: correlation spectroscopy

98 CP/TOSS cross polarization/total sideband suppression

99 CP-MAS: cross polarization magic-angle spinning

100 DBE: double bound equivalent

101 DEPT: distortionless enhancement by polarization transfer

102 DOC: dissolved organic carbon

103 DOM: dissolved organic matter

104 rDOM: refractory DOM

105 DON: dissolved organic nitrogen

106 rDON: refractory DON

107 DOSY: diffusion-ordered spectroscopy

108 DP/MASS: direct polarization/magic angle spinning

109 DP/TOSS: direct polarization/ total sideband suppression

110 EEM: excitation emission matrices

111 FA: fulvic acid

112 FIFFF: flow field-flow fractionation

113 FT-ICR MS: fourier-transform ion cyclotron resonance mass spectrometry

114 HA: humic acid

115 HiA: hydrophilic acids

116 HiB: hydrophilic bases

117 HiN: hydrophilic neutrals

118 HIX: humification index

119 HMBC: heteronuclear multiple-bond correlation

120 HoA: hydrophobic acid

121 HoB: hydrophobic bases

122 HoN: hydrophobic neutrals

123 HPAEC-PAD: high-performance anion-exchange chromatography with pulsed amperometric

124 detection

- 125 HRMS: high-resolution mass spectrometry
- 126 HS: humic substances
- 127 HSQC: heteronuclear single quantum correlation
- 128 HTC: hydrothermal carbonization
- 129 IC: inorganic carbon
- 130 LCCC: liquid chromatography at the critical condition
- 131 LLE: liquid-liquid extraction
- 132 MRP: Maillard reaction products
- 133 MW: molecular weight
- 134 HMW: high MW
- 135 LMW: low MW
- 136 MWD: MW distribution
- 137 NOM: natural organic matter
- 138 OCD: organic carbon detection
- 139 OND: organic nitrogen detection
- 140 PCA: principal component analysis
- 141 PDA: photo diode array detector
- 142 PLS-DA: partial least squares discriminant analysis
- 143 OPLS-DA: orthogonal PLS-DA
- 144 POM: particulate organic matter
- 145 PRAM: polarity rapid assessment method
- 146 RAMP: ramped amplitude
- 147 RP: reversed phase
- 148 SFS: synchronous fluorescence spectroscopy

149 SPE: solid-phase extraction
150 SUVA₂₅₄: specific ultraviolet absorbance at 254nm
151 TC: total carbon
152 TCI: proton-optimized triple resonance NMR inverse probe
153 THP: thermal hydrolysis process
154 TKN: total Kjeldahl nitrogen
155 TN: total nitrogen
156 TOC: total organic carbon
157 TOCSY: total correlation spectroscopy
158 WWTP: wastewater treatment plant

159

160 **1. Introduction**

161 The development of new analytical methods for wastewater treatment to identify the
162 wide variety of often toxic compounds present and those specifically generated during
163 treatment processes, is a major present and future challenge. Wastewater treatment plants
164 (WWTP) are a necessary step in depollution because they prevent the discharge of
165 anthropogenic pollutants in natural waters. Prevention of some environmental issues as
166 eutrophication or chemical contamination is secured by the increasingly stringent discharge
167 concentration standards (depending on geographical area) for global parameters such as
168 chemical oxygen demand, COD (125 mg O₂.L⁻¹ in Europe), or total nitrogen, TN (10-15 mg
169 N.L⁻¹ in Europe) [1–4]. Most urban WWTPs use biological processes for water treatment to
170 reduce the concentration of dissolved compounds and suspended solids and the main waste
171 produced is composed of biomass called sludge, and this waste to be processed [5].

172 The increasing amount of sewage sludge produced from WWTPs represents a
173 significant challenge for wastewater facilities operations and their environmental footprint

174 [6]. Anaerobic digestion (AD) is thus implemented within various WWTPS with the purpose
175 of energy recovery, sludge volume reduction and biomass stabilization, as it can reduce the
176 organics, odor and pathogens, as well as generating biogas for the production of electricity
177 and heat. Pretreatment, and in particular the hydrolysis step of AD, will change the
178 composition of the sludge flocs by introducing organic substances from the cellular organelles
179 of the bacteria used in the processes. Thus, various pretreatment methods before AD have
180 been widely developed to promote hydrolysis efficiency and improve the AD performance,
181 including mechanical, thermal, chemical and biological treatments [7].

182 Thermal treatment is a technology that has recently been implemented in WWTPs to
183 improve sludge processing. Sludge is heated to temperatures of 140-200 °C to improve biogas
184 production through anaerobic digestion [8] and/or sludge dewaterability [9], thus reducing
185 overall costs [8–12]. The consequences of these thermal processes are the formation of
186 significant amounts of soluble molecules, which are not degraded by biological processes and
187 are defined as refractory dissolved organic matter (rDOM) [13]. Refractory compounds also
188 named recalcitrant or persistent, have been defined as molecules not degraded by the
189 biological treatment of wastewater. However, if the definition of refractory compounds is
190 quite relative and depends on the method of degradation used by the WWTP, refractoriness
191 has been well studied and standardized in the literature [14,15]. There is therefore an analytical
192 challenge linked to regulatory constraints on the concentration of these refractory compounds
193 in wastewater and to the constant development of sludge thermal treatment technologies
194 [2,16–18]. Refractory compounds in WWTPs may schematically originate from 3 sources:
195 initially present in the raw wastewater, formed during the main biological processes [19] and
196 formed during other processes like anaerobic digestion [20] or thermal processes [21–26].
197 The resulting refractory organic matter of various origins, then, represents a challenge when
198 following the WWTPs discharge standards. It appears that thermal treatments generate high

199 concentrations of refractory dissolved organic nitrogen (rDON), measured by total Kjeldahl
 200 nitrogen (TKN) (260 to 470 mg N.L⁻¹), and rDOM, measured by COD (5.25-12.5 g O₂.L⁻¹),
 201 depending on the temperature used in the process (120 – 208 °C). Table 1 gives values for the
 202 amount of refractory organic matter found by different authors and its potential impact on
 203 WWTP discharge quality.

204

205

206

207 Table 1. Quantification of rDOM and rDON produced by thermal treatments of sewage
 208 sludge.

Sludge treatment	Thermal treatment	rDOM in COD (g O ₂ .L ⁻¹)	rDON in TKN (mg N.L ⁻¹)	WWTP outlet	Reference
140-165 °C, 6bar	THP	N/A	N/A	Increase of TKN from 1.8 to 2.8 mg N.L ⁻¹	[21]
Thermal treatment, anaerobic and aerobic	THP	5.25	260	-	[22]
170 °C, anaerobic and aerobic	THP	Increase of rDOM from 2.6 to 8.4 % in the recycled stream.		Increase of COD from 52 to 63 mg O ₂ .L ⁻¹	[24]

170 °C, 7.9 bar,	THP		260	-	[27]
anaerobic					
120-170 °C, 25	THP	9 - 12.5	380 - 470	-	[28]
psi, anaerobic					

209

210 In addition to these refractory compounds formed during thermal processes, natural
 211 organic matter (NOM) is present in surface water or soil and is also found in wastewater. This
 212 organic matter can be separated between natural and anthropogenic species, and then in each
 213 class between soluble and solid aggregates. From the total organic matter present in
 214 wastewater samples, natural organic matter (NOM), DOM, rDOM and biodegradable DOM
 215 are identified by analytical methods before and after the different treatment processes as
 216 shown in Fig. 1. There are wide chemical families of rDOM (surfactants, pharmaceuticals,
 217 chemicals, Amadori compounds, melanoidins, etc.) and the chemical structures range from
 218 small molecules to polymers, as shown in Fig. 2, with the identified structures of humic
 219 substances (HS), polyphenols, Amadori compounds and melanoidins.

220 In Fig. 2, HS and phenols can come from both natural and anthropogenic sources.
 221 Maillard compounds and melanoidins are anthropogenic compounds. Structural studies of
 222 rDOM show that the majority of the refractory compounds are nitrogen compounds, forming
 223 the dissolved organic nitrogen (DON) and refractory dissolved organic nitrogen (rDON)
 224 families [21,23,24,33], as presented in Table 1.

225 Determinations of COD and total nitrogen (TN) are global measurements commonly
 226 used to evaluate the efficiency of wastewater treatment. However, they give global and
 227 nonspecific information about molecules. Therefore, COD and TN measurements are not
 228 relevant for understanding the processes of refractory compound formation, which limits the
 229 optimization strategies of the technologies involved such as thermal treatment or anaerobic

230 digestion or the evaluation of technologies able to degrade (chemical oxidation) or remove
231 (adsorption) these troublesome compounds. As the main goal up until the last decade was the
232 optimization of biogas production, very few authors worked on the characterization of the
233 structure of the rDOM formed. More stringent discharge standards and stronger human health
234 concerns have, however, led to an increasing interest in the characterization of rDOM
235 [22,23,34–36]. The rDOM found after thermal treatment has been studied in other industrial
236 fields that are developing identification methods that can be dedicated to wastewater, such as
237 food [37,38], animal waste [39], and surface water [40]. The analytical development
238 performed for the characterization of these rDOM will therefore be considered in this review.

239 The rDOM is made up of thousands of different compounds, each in small
240 concentrations leading to a large overall concentration, particularly in thermally-treated
241 sewage sludge [23]. Therefore, the use of advanced analytic methods is necessary, first to
242 fractionate the compounds and then to get accurate chemical information about the diverse
243 molecules that compose the rDOM. The rDOM is found in an extremely wide range of
244 molecular masses, physicochemical properties and concentrations. In particular, a large
245 variance is observed in aromaticity, hydrophobicity, and molar mass distribution for organic
246 molecules in the samples. Inorganic compounds are also present, some compounds are
247 complexed with anions, cations or small molecules such as phenols, adding still more
248 complexity, and thus making the matrix effect responsible for the analytical variance very
249 important. This justifies the use of complex, time consuming and expensive, hyphenated and
250 analytical techniques. This complexity in the matrix makes sample preparation a crucial step
251 for each analytical objective, in order to simplify matrices and split species before analysis. A
252 broad spectrum of spectroscopic techniques (IR, UV, fluorescence, mass spectrometry,
253 NMR.) is used to advance structural studies. Hyphenated techniques using the high-resolution
254 analyzers also provide an orthogonal experimental space for improving analytical resolution.

255 Table 2 summarizes the methods, mainly used for the analyses of rDOM, depending on the
256 structural information sought. The orthogonality will be a key element in the analytical chain,
257 notably for identifying certain classes of nitrogen compounds, aromatic compounds,
258 macromolecular compounds or intermediates of certain chemical reactions encountered in
259 thermal treatment processes. The choice of the components in the analytical chain of Table 2,
260 from sample preparation to data processing, will be guided by two main objectives:
261 quantifying and identifying structures. In both cases, the absence of the quantification
262 standards and isotopic standards needed for the mass spectrometry analysis of rDOM will be a
263 key point to be resolved. Chemometrics are increasingly used to overcome the lack of
264 standards, allowing a relative quantification and classification of isolated compounds in these
265 complex matrices, and they can add value to the standardization of the interpretation of
266 results. All the data must be linked to isolate the refractory species and identify the key
267 factors for a more detailed study. The hyphenated method's final objective is to correlate
268 COD with resolutive monitoring and speciation specifically for compounds generated during
269 the treatment processes [41].

270 In this review, the state of the art of all analytical methods dedicated to the study of
271 DOM will be considered, with a specific focus on rDOM. Due to the structural complexity of
272 rDOM and wastewater matrices, the main objective of this review is to present the
273 spectroscopic and physicochemical techniques used to qualify and characterize rDOM. A
274 large part of the review is devoted to new chemometrics developments, allowing a
275 classification of the isolated species based on key analytical parameters linked to rDOM
276 structure and properties. Advances in NMR and mass spectrometry will also be discussed,
277 including the identification in other matrices of methods that can be applied to the study of
278 wastewater rDOM.

279 Table 2: Analytical methods used for the analysis and characterization of recalcitrant compounds.

Methods	Advantages	Drawbacks	Structural information	Recalcitrant species
Dissolved organic carbon (DOC) and nitrogen (DON)	Fast Simple Inexpensive Normalized for quantification	No resolution Narrow monitoring scale	Soluble C and N DOC/DON ratio	No specificity
UV/Vis spectrophotometry	Fast Simple Inexpensive	Matrix effect UV-quenching Possible interferences No split Low resolution	Aromaticity (UV 254 nm) SUVA ₂₅₄ nm	Hydrophilic and hydrophobic rDOM
Fluorescence spectroscopy	Rather specific Good sensitivity	Possible interferences No split	Protein-like or humic-like peaks with specific Ex/Em wavelength couples	Serine and Tyrosine like rDOM Melanoidins
Fourier transform infrared spectroscopy	Inexpensive	Not specific	Functional groups	Melanoidins HA, FA

Nuclear magnetic resonance spectroscopy (liquid and solid state)	Holistic Structural information, Quantitative information without specific standard. Easy sample preparation for CP-MAS	Expensive Low sensitivity	Relative proportions of functional groups. Full structural characterization after isolation. Presence of specific nuclei (ex: ^{31}P).	Product from human activities (surfactants, phthalates) HA Protein-like
Reversed-phase high performance liquid chromatography	Fast Simple fractionation for 2 to 5 species	No discrimination between macromolecules, proteins, polysaccharides and micro-pollutants.	Macromolecular DOM and rDOM fractionation based on structural properties: C/O ratio, hydrophobicity	Pharmaceuticals HA Phenolic compounds Amadori compounds
Steric exclusion chromatography (SEC)	Fractionation controlled by molecular weight and other interactions. Wide variety of SEC columns for linear or branched polymers. Split of	High influence of the composition and pH of the buffer eluent. Low resolution of HPLC SEC columns.	Size distribution. Correlation with carbon and molecular nitrogen with the use of simple detectors	HA, FA Melanoidins Proteins Polysaccharides

	the salts, low small molecules and polymers.	Time consuming separations.		
Hyphenated techniques:	Combination of SEC		Amino-acids composition	Specific tyrosine- and serine-
- SEC/PDA/Fluorescence	fractionation and specificity of the fluorescence emission of known species		Aromaticity	like proteins.
	Relative quantification		Size distribution	HA, FA
			SUVA ₂₅₄ , UV 280 nm and UV 210nm of each peak	Melanoidins
- SEC/HRMS	Structural information	Matrix effect	Elemental analysis	HA, FA
	Macromolecule elemental analysis for FT-ICR MS	Multiple adducts in ESI	O/C and N/C ratio	Melanoidins
			Aromaticity	Proteins
			Double bond equivalents	Polysaccharides
			CH ₂ and C ₂ H ₄ O homolog series	Amadori compounds

280 **2. Trends in molecular characterization: dissolved organic matter and recalcitrant organic**
281 **compounds**

282 2.1 Recalcitrant species

283 First, part of DOM is natural organic matter that can contain recalcitrant compounds
284 such as HS, for example. Consistent review from [42] describes NOM, the analytical method
285 used and treatment options. In addition to natural recalcitrant compounds, many recalcitrant
286 organic species are produced by human activities such as landfill leachates [43] or food
287 industry [44]. Specifically, in WWTPs, low quantities of refractory compounds have always
288 been reported throughout the process, but recently-implemented thermal treatments form
289 additional, new recalcitrant compounds that can show biological activity and effect on human
290 health: Genotoxicity, cytotoxicity, antioxidative, antimicrobial activity [32]. Such effluents
291 can contain up to 100 times more refractory compounds than treated or natural waters and are
292 problematic for compliance with WWTP discharge standards. Furthermore, recalcitrant
293 compounds formed during thermal processes can contain nitrogen, which is the most
294 problematic element as its discharge standards are already low [23,45]. Lastly, while many
295 natural recalcitrant compounds are inert, some anthropogenic compounds may be hazardous
296 for the environment and/or human health [46,47].

297 Refractory compounds are remarkable in that they cannot be attributed to a simple
298 family or chemical class. The literature shows the development of a wide variety of analytical
299 methods that can be improved to optimize the detection limit, or resolution [48]. Depending
300 on the desired objectives, which may be regulatory, qualitative, quantitative or related to
301 structural analysis, the methods developed will involve a wide variety of processes and
302 detection systems.

303 Refractoriness is generally correlated by several physicochemical parameters such as
304 high molecular weight (HMW), hydrophobicity, aromaticity or chemical functions, especially

305 certain substitutions such as chlorinated aromatics. These characteristics can be found in the
306 humic substances (HS), humic acids (HA) and fulvic acids (FA) that constitute the majority of
307 rDOM species isolated in raw water, but also in many refractory compounds of anthropogenic
308 origin. For example, HA and FA are acidic organic polymers recalcitrant to biodegradation
309 [49,50]. Their complex molecular structures show a link between aromaticity, which can be
310 measured by UV absorption at 254 nm or the H/C ratio, and molar mass calculated for each
311 chromatographic peak identified in water samples, as shown on Fig. 3. It is, therefore,
312 essential to be able to access resolute analytical methods for the fractionation of refractory
313 compounds and to be able to measure other related properties such as aromaticity for access to
314 sample monitoring. Fig. 3 shows the classification possibilities for various HAs and FAs
315 isolated from surface water samples according to these 2 parameters. This classification
316 methodology will be used with other rDOM species for analytical developments allowing
317 rapid comparison of recalcitrant compounds in wastewater samples.

318 With the large variation of reactive compounds in raw wastewater and of their
319 physicochemical properties, the analytical challenge is present at all levels: sample
320 preparation, fractionation and structural analysis. For HS, structural heterogeneity is
321 demonstrated by the analytical approach following fractionation of specific physicochemical
322 data: molecular mass and hydrophobicity. Modelling of polymeric structures generated in a
323 glycine/glucose model showed the formation of three groups of macromolecules between 10
324 kDa and 20 kDa. The composition of the polymers will depend on the ratio, the type of
325 reactants and the reaction parameters, and can be determined using fractionation techniques.
326 In this model, the origin of nitrogen-containing polymers is attributed to specific Amadori
327 intermediate and the non-nitrogen-containing polymers to glucosone and deoxyglucosone
328 [52]. All these intermediates can be identified by GC/MS analysis of reaction product dialysis.
329 New developments using direct analysis or derivatization aim to gain better knowledge of the

330 reaction mechanisms involved through the structural study and quantification of the
 331 recalcitrant compounds. The optimization of treatment processes can also be managed on an
 332 analytical basis. Table 3 summarizes the analytical techniques developed for the study of the
 333 Maillard products that this review focuses on (see part 2.2).

334

335 Table 3. Analytical techniques used for the analysis of Maillard reaction products. Reprinted
 336 from [48].

Maillard reaction products (MRPs)	Analytical techniques
Amadori compounds (Direct analysis)	Column chromatography HPLC differential refractometry detection HPLC involving derivatization HPAEC coupled electrochemical and/or DAD FAB-MS ESI coupled HPLC and EC, EC coupled MS MALDI-TOF NMR LC-MS
Indirect analysis (2-FM-AA)	Ion-pair RP-HPLC CEC UV-detection HPLC-MS
Advanced Maillard products	HPLC-DAD
Carboxymethyl-lysine	RP-HPLC RP-HPLC o-phthalaldehyde pre- column derivatization GC-MS
Pyrraline	RP-HPLC
Cross linking products Lysine	LC-MS with ESI
Dimmers Arginine-lysine	LC-MS with ESI

	Ion-exchange chromatography FAB–MS
Other amino acid derivatives,	HPLC-coupled GC–MS
Argyrimidine	
Pyrazinones	HPLC with UV and fluorescence detection
Lysine aminoreductone	HPLC-DAD
Final stage MRPs	HPLC, NMR, MS, UV, IR spectrometry
General melanoidins	MALDI-TOF mass spectrometry

337

338

339 2.2 Maillard reactions

340 Thermal treatment with temperatures greater than 120 °C applied to sewage sludge
341 solubilizes organic matter and breaks bacteria cell walls [53,54] with the production of low
342 and HMW molecules. Some of them, such as the carbonyl and amino functional groups,
343 include reducing carbohydrates, amino acids, peptides and proteins which may react with
344 each other in special conditions. An analytical survey of sewage sludge shows that samples
345 composed of 15-30 % carbohydrates and 20-50 % proteins generate ideal conditions for
346 Maillard reactions [55–57]. Maillard reactions are non-enzymatic condensation, controlled by
347 temperature, that occurs between reducing sugars and proteins [58]. Refractory compounds in
348 thermally treated sewage sludge were modelled with a Maillard reaction scheme by
349 identifying specific Amadori intermediates [26,33]. Maillard reactions have been widely
350 studied in the food industry due the physicochemistry and especially the brown color of the
351 synthesized compounds at high temperature. It can be responsible for flavor adjustments in
352 many types of food [59,60]. Depending on the conditions, these reactions can also be the
353 source of a waste that requires specific treatment, like brewery stillage [44]. Maillard
354 reactions are a complex system of interconnected chemical reactions and are not yet well

355 understood. Generated compounds depend on many parameters such as the nature of the
356 reactants, pH and temperature [48,61]. Low molecular weight (LMW) compounds are formed
357 as intermediates through the earliest stages of Maillard reactions, including ketones,
358 aldehydes, pyrazines, pyridines and furans [48,62,63]. All these intermediates rapidly react in
359 an aldol-type condensation. The degradation of monosaccharides such as aldose sugars leads
360 to the formation of furaldehydes and furanones, such as furan-2-carboxaldehyde and furan-2-
361 aldehyde [64]. The subsequent steps involve dehydration, cyclization, retro-aldolization,
362 isomerization, rearrangement and condensation. In the last steps, all the intermediates react
363 with amino acids and peptides to form LMW monomers leading to the production of HMW
364 molecules, called melanoidins, by polymerization [65].

365 The complexity of the matrix, for the different industries interested in melanoidins, is
366 an analytical challenge and explains our lack of knowledge of these compounds. Melanoidins
367 are thought to be the cause of the formation of a brown color in food and wastewater samples
368 after thermal treatment [32]. This brown color formation is coupled with increased UV
369 absorbance and is largely observed in the liquid phase of thermally treated sludge, which is
370 also correlated to refractoriness [8,21,66,67]. The brown color can be linked to the decrease of
371 the soluble protein fraction in samples that may have reacted during Maillard reactions from
372 78 % to 68 % with an increase in treatment temperature from 135 °C to 190 °C [66]. The non-
373 biodegradable soluble chemical oxygen demand (COD) and UV absorbance were 2.7 and 1.7
374 times higher for the thermal treatment of sewage sludge at 165 °C versus 140 °C, linked to the
375 increase of synthetic melanoidin in solution [21]. Finally, melanoidins exhibited a strong IR
376 absorption band, notably at 1607 cm^{-1} , attributed to extensive conjugation, which can be
377 connected to the identification of UV bands shifting to high wavelengths. These results
378 indicate that spectrophotometric methods can be used for melanoidin quantification and
379 wastewater monitoring.

380 The core constant structure of melanoidins is formed by mixture of α -dicarbonyl
381 Maillard reaction intermediates, branched with amino compounds. Pyrolysis of the isolated
382 polymers produced typical Amadori products, such as pyrazines, pyrroles, pyridines and
383 furans. Longer thermal treatment time and higher temperature are correlated with HMW
384 melanoidins formation [32]. The molecular weight and nitrogen composition of the
385 recalcitrant polymers depend on the nature of reactants and of the industry's mixture of raw or
386 wastewater: > 30 kDa in bread [68], > 10 kDa in coffee [69], but the reaction conditions have
387 a major influence. In the case of sewage sludge, temperatures involved in some thermal
388 sludge treatments exceed 120 °C to be energetically efficient [8,12]. The final biological and
389 physicochemical properties of wastewater samples is correlated with analytical data from
390 melanoidins. The majority of the color (68 %), DOC (53 %) and UV absorbance at 254 nm
391 (54 %) were in the analytes fraction > 10 kDa, indicating a significant amount of aromaticity
392 and conjugated molecular structures in the HMW fractions of hydrolyzed sludge samples
393 from thermal hydrolysis processes [21]. But simple physicochemical parameters are not
394 significant for the specific identification of melanoidins, and high levels of soluble COD can
395 be found in the LMW fraction, below 5 kDa. HMW contributes to refractoriness for classic
396 biological WWTP treatments [70] and melanoidins are no exception. A major part (65 %) of
397 the total refractory organic carbon (rTOC) was above 10 kDa and 25 % was under 0.5 kDa for
398 sewage sludge treated at 208 °C [21,22].

399 Melanoidins contain significant nitrogen quantities and their refractoriness may affect
400 compliance with discharge standards. DON increases up to 50 % with thermal treatment of
401 sewage sludge, before anaerobic digestion [21], and up to 24 % after anaerobic digestion [28].
402 Regarding structural analyses, about 50 % of rDON is found in HMW fractions up to 10 kDa
403 [22]. It should be pointed out that the literature mainly depicts Maillard reactions as being
404 responsible for refractoriness generated during thermal treatments since, quantitatively, they

405 are probably the main cause of refractory compound formation [48]. Other sources can still
 406 be expected such as caramelization reactions [71] or the liberation of cell walls in recalcitrant
 407 components. Table 4 summarizes all available data, fully presented in Table S1
 408 (supplementary data), for the structure of compounds synthesized during Maillard reactions,
 409 as well as the Amadori intermediates that can be identified, depending on the progress of the
 410 reaction, the conditions and the reactants presents in the wastewater and sludge.

411 Table 4. Chemical species, identified after thermal processes on residual water and sludge.

Thermal process	Temperature (°C)	Identified chemical species	Range of MW (g.mol ⁻¹)	Reference
HTC	170	Aromatic, Heterocyclic, Pyrazines, Pyrimidines, Aldehydes, Ketones	60 - 282	[23]
HTC	200 - 320	Ketones, Heterocyclic, Carboxyl, Amide, Pyridines, Humic substances, Phenols	94 - > 10 000 (HS)	[34]
THP	172	Alkanes, Esters, Carboxyl, Amide, Pyrrolidines, Pyridines, Benzenoids, Peptides	127 - 572	[72]
HTC	208	Aldehydes, Ketones, Phenols, Heterocyclic, Pyrazines, Pyrimidines	73 - 182	[73]

412

413 Table 4 shows the wide structural variety of Amadori intermediates found during and
414 after thermal processes from the Maillard reactions. The very different physicochemical
415 properties (polarity, hydrophobicity, molar mass, volatility, chromophores, and fluorescence
416 properties) of these compounds will require specific analytical developments to identify and
417 quantify them all. This structural and functional diversity may also explain why HMW
418 compounds are not always identified in the studies cited. The techniques used to identify these
419 recalcitrant compounds may not be suitable for the identification of macromolecules (HS,
420 melanoidins).

421

422 2.3 Physicochemical and spectrophotometric analyses

423 DOM is heterogeneous in nature, chemical functions and properties. The
424 physicochemical methods that have been developed for wastewater analysis are designed to
425 identify and quantify certain chemical or species functions, such as proteins, polysaccharides
426 and charged colloids, and to understand the characteristics of the samples such as
427 biodegradability, the presence of recalcitrant compounds and the color of the water. The
428 interest of these physicochemical techniques lies in their short analysis times, the use of
429 simple equipment and the simple sample preparation and justifies that they are still widely
430 used today.

431

432

433 2.3.1 Dissolved organic carbon (DOC) and nitrogen (DON)

434 DOC is a concentration parameter representative of the total dissolved organic
435 molecules (NOM, DOM and rDOM). The DOC concentration is determined with a total
436 organic carbon (TOC) analyzer after sample filtration (0.45 to 0.22 μm). TOC measurements
437 provide information on total carbon (TC) and inorganic carbon (IC) concentrations. DOC is

438 calculated by subtracting IC from TC [74]. Biodegradable DOC is the DOC fraction that can
439 be metabolized by bacteria in wastewater, and recalcitrant species and recalcitrant DOC can
440 be identified by analyzing the non-biodegradable fractions [75].

441 DON refers to soluble organic compounds retrieved after filtration (0.45 to 0.22 μm),
442 that also contain nitrogen. The DON concentration is determined by subtracting the
443 concentrations of ammonia, nitrate and nitrite from the concentration of total nitrogen. Total
444 nitrogen concentrations in wastewater samples are determined using various colorimetric
445 methods such as the persulfate digestion method, nitrates using the dimethylphenol method,
446 nitrites using the diazotization method and ammonia using the salicylate method [22]. These
447 methods' yield is less accurate when the inorganic nitrogen concentrations are high. For the
448 recalcitrant compounds monitored in wastewater, melanoidins are the most intensively studied
449 complex, and heterogeneous nitrogen-containing Maillard reaction products. A first approach
450 to rDON and melanoidin quantification can be achieved using DOC/DON measurements and
451 the calculation of the carbon-on-nitrogen ratio (C/N). This ratio increase with the
452 concentration of different reducing sugars and glycine [21]. The removal of biodegradable
453 DOC is of great importance for the water research and water treatment industry and
454 normalization of rDOM concentrations vs DOC or DON are performed routinely. This field
455 of research is now able to focus on more specific species, such as specific melanoidins or
456 Amadori intermediates. In this case, it is necessary to add selective sample preparation
457 techniques upstream to ensure a more resolute fractionation.

458

459 2.3.2 UV/Vis spectrophotometry

460 UV/Vis spectrophotometry is an analytical method widely used for the
461 characterization of organic matter in all kinds of natural and wastewater. Used since the
462 1970s, it first served to model the composition of ocean waters and the fate of various

463 dissolved compounds: chlorophyll pigments and NOM. The UV-visible spectrum is studied
464 for the classification of samples from specific measurements of chlorophyll pigments at 430
465 nm and continuous analysis between 300 and 700 nm [76]. More broadly, UV/Vis
466 spectrophotometry provides structural information on the nature of the bonds present in the
467 sample (chromophores and specific chemical functions, sp^2 -hybridized carbon) and
468 quantitative information on the concentration of absorbent species in the spectral range
469 studied (proteins, nucleic acids, HS, etc.). Even if the NOM is represented by species that
470 have characteristic UV spectra, sample spectra are typically broad without any fractionation.
471 The comparison of samples is mainly performed by measurements at specific wavelengths
472 such as 254 nm, 210 nm or 280 nm, and ratios between these values [77]. The resulting
473 indices give insight into the aromaticity of DOM and its molecular weight can be discussed
474 [78]. Certain macromolecular DOM show a distinctive ultraviolet absorbance ratio index
475 (URI) calculated using the UVA_{210}/UVA_{254} ratio, due to their different compositional
476 proportions. URI values show the lowest values for HA (1.59, highest aromaticity),
477 intermediate for FA (1.88, intermediate aromaticity), and highest for proteins (13.5 for BSA,
478 lowest aromaticity) [79–81]. Specific UV absorbance at 254 nm ($SUVA_{254}$) is a powerful
479 index measured using the ratio of UV absorbance to DOC, correlated with the NOM species.
480 For these spectrophotometric measurements, it is necessary to control the turbidity of the
481 samples by filtration. Some ions in solution (nitrate, nitrite, iron) may have a matrix effect,
482 depending on their concentration [82].

483 The importance of sample absorbance is linked to biological response and the removal
484 of the colored fraction from the thermally-treated sludge liquid fraction with resins could
485 increase biodegradability from 33.2 to 57.8 % [83]. Most of the UV-quenching substances can
486 be found in HMW substances, which are presumed to be mostly refractory. For access to
487 these macromolecular species, a size fractionation sample preparation can be coupled to

488 UV/Vis spectrophotometry measurements [21,22,84]. The use of specific wavelengths does
489 not differentiate the rDOM chemical families, it is then necessary to combine several
490 physicochemical measures such as COD, DOC or TOC. SUVA₂₅₄ is a specific indicator of
491 aromaticity, which is generally a property of refractory compounds. With a SUVA₂₅₄ value
492 greater than 4, a sample is composed mainly of hydrophobic and aromatic NOM, with
493 HMWs, whereas a SUVA₂₅₄ value lower than 3 suggests a mixture of low molecular weight
494 hydrophilic NOM [85].

495

496 2.3.3 Fluorescence spectroscopy

497 Fluorescence spectroscopy provides useful data on the structure of soluble compounds
498 in water. Fluorescence in aquatic matrices has been extensively studied and sample
499 monitoring can be achieved in different areas such as WWTP influents and effluents, textiles
500 and the food industry. The specific excitation and emission wavelength values and intensity
501 can be correlated to specific structural functions, such as carboxylic acidity or the aromatic
502 character of amino acids which comprise macromolecules and recalcitrant compounds. It is a
503 rapid, sensitive technique requiring simple sample treatment with filtration, then pH
504 adjustment and small volumes for analysis. The most widely used methods for fluorescence
505 spectra measurements applied to wastewater matrices are excitation emission matrices
506 (EEMs) and synchronous fluorescence spectroscopy (SFS) [16]. An EEM is a 3D scan
507 resulting in a contour plot of emission wavelength as a function of the excitation wavelength
508 with fluorescence intensity in the third dimension. EEMs can provide specific fingerprint
509 profiles for fluorescent components comprising wastewaters [86].

510 This specificity makes it possible to assign certain compounds or functions to specific
511 Ex/Em wavelength couples, or EEM coordinates as presented in Fig. 4.

512

513 Fig. 4. presents the identification of 7 classes of recalcitrant compounds in a wastewater
514 sample, four corresponding to proteins (B1, B2, T1 and T2), two to HS (A and C) and one
515 corresponding to melanoidins (M). For wastewater, EEM contour plots are located in two
516 major groups. The first is made up of protein-like peaks: peak T1 (275-285 / 320-350 nm,
517 tryptophan-like), T2 (215-237 / 340-381 nm, tryptophan-like), B1 (275-310 / 305-320 nm,
518 tyrosine-like) and B2 (220-237 / 305-320 nm, tyrosine-like). The second is made up of humic-
519 like peaks: peak A (230-260 / 400-480 nm, fulvic-like), C (320-360 / 420-460 nm, humic-
520 like) and M (290-310 / 370-420 nm). Protein-like peaks T1 and T2 tend to be predominant in
521 urban wastewater due to microbial activity and the accumulation of biodegradable organic
522 matter [87]. Humic-like peaks (A and C) are usually the most abundant in natural water
523 samples due to the presence of FA and HA of microbial origin. Indices are calculated from the
524 EEM spectra: the fluorescence index (FI), that is the ratio of emission intensity at 450/500 nm
525 measured at an excitation wavelength of 370 nm [88], and the humification index (HIX)
526 determined as the ratio of fluorescence intensities for two integrated regions of the emissions
527 scans: Em 435-480 nm divided by Em 300-345 nm, at Ex 254 nm [89]. These indices are used
528 for comparative studies of effluents, depending on the agricultural, urban, or industrial nature
529 of the organic matter's origins. Biodegradability predictions are notably made using EEM
530 analyses performed in one hour instead of the conventional biodegradability tests that can take
531 between 5 and 30 days [90,91]. Melanoidins formed during thermal sewage sludge processes
532 have been studied using EEM in the literature and the related Em/Ex couples are given in
533 Table 5.

534

535 Table 5 – Ex/Em wavelength associated to melanoidins recalcitrant compounds.

Matrix	Peak 1 (nm)	Peak 2 (nm)	Reference
Synthetic melanoidin	360/440	230/440	[21]

Thermally treated sewage sludge	330/410	230/400	[21]
Thermally treated sewage sludge	320/400	280/350	[34]
Thermally treated sewage sludge	360/425		[92]
Thermally treated sewage sludge	320/420		[93]

536

537 SFS is performed on water samples to produce 2D graphs obtained by scanning both
538 the emission and excitation monochromators simultaneously at a constant wavelength
539 interval. The development of SFS was devoted to the analysis of specific micro-pollutants
540 such as polycyclic aromatic hydrocarbons [94]. Absorption and excitation mechanisms are
541 followed by electron relaxation to basic energy levels, with the emission of a photon-termed
542 fluorescence [95]. The specific excitation/emission wavelengths are characteristic of
543 particular molecular conformations and peak intensity is proportional to the concentration of
544 the fluorophore in solution. The precise excitation/emission signature depends on the
545 chemical structure and composition of wastewater samples, but matrix effects are important,
546 as well as pH and the temperature of the samples.

547 Compared to UV absorbance, fluorescence is more specific and some functions or
548 types of compounds can be measured at specific Ex/Em wavelength couples. Fluorescence
549 sensitivity is increased 10 to 1,000-fold over that of UV absorption spectroscopy. The last
550 advantage of fluorescence is its potential for discriminating NOM and rDOM, which absorb at
551 similar wavelengths but emit at specific values [87].

552

553 2.3.4 Fourier transform infrared spectroscopy (FT-IR)

554 FT-IR has been applied for NOM quantification and characterization, in wastewater,
555 for monitoring the sludge, or the composition of WWTP influents. This analytical method is
556 based on the determination of spectra of the samples in the infrared wavelength range. The

557 results are used to determine specific absorption bands for the identification of chemical
 558 functions and structures [96]. The samples must be completely dehydrated by heating (105 °C
 559 for 6 h) or freeze-drying to remove residual water before any measurement. This sample
 560 preparation protocol is incompatible with labile organic compounds. By analyzing for specific
 561 standards, sensitive methods are used to compare the intensity of absorption bands and
 562 specific spectra transformations to obtain a semi-quantitative analysis of NOM [97]. Spectral
 563 databases were reported for characteristic absorption bands of substances that most occur in
 564 wastewater [98]. Table 6 summarizes the main chemical structures and functions identified in
 565 sludge and wastewater, including reactive proteins, sugars, polysaccharides and amines,
 566 which induce melanoidins during thermal treatment.

567 Table 6. FT-IR Library for analysis of the activated sludge composition, from [97], copyright
 568 Elsevier.

Wavenumber (cm⁻¹)	Shape	Intensity	Bond	Compound(s)	Class
3650 - 3200	b	str	st O-H	Glycerol	Polyalcohol
	b	m		Glucose	Monosaccharide
	b	m		Lactose	Saccharide
	b	m		Sucrose	Saccharide
	b	m		Phenol	Phenols
3300 -2800	b	m	st NH ₃ (N ⁺ -H)	Albumin	Protein
	b	m		Ammonium sulfate	Ammonium
	b	m		Dry meat extract	Peptide

	b	m		Peptone G	Peptide
3000 -2840	s (m)	str	st C-H	Glycerol	Polyalcohol
	s (m)	m		Glucose	Monosaccharide
	s (m)	m		Oleic acid	Fats
	s (m)	w		Sucrose	Saccharide
2900	s	w		Lactose	Saccharide
2360	s	w	st NH ⁺	Ammonium sulfate	Ammonium
2075	s	str	C≡N	Potassium cyanide	Cyanides
2045	s	str	SC≡N	Potassium thiocyanate	Thiocyanides
1775	s	w	st C=O	Sodium carbonate	Carbonates
1710	s	str		Oleic acid	Fats
1680	s	str	st C=C	Phenol	Phenols
1655	s	str	st as COO ⁻	Albumin	Protein
1650 - 1575	s (m)	str		Dry meat extract	Peptide
1600	s	str		Peptone G	Peptide
1570	s	w	def NH ₂	Ammonium sulfate	Ammonium
1535	s	str	sy NH ₃ ⁺	Albumin	Protein
	s	str		Peptone G	Peptide
1515	s	str		Dry meat extract	Peptide

1460	s	str	st C=O	Sodium carbonate	Carbonates
	s	str		Acid sodium carbonate	Carbonates
1445 -1600	s	m	def C-H (skelet)	Glucose	Monosaccharide
	s	m		Lactose	Saccharide
	s	m		Sucrose	Saccharide
1440	s	m	def CH ₂	Oleic acid	Fats
1410	s	str	CH ₂	Glycerol	Polyalcohol
1400	s	str	S(=O) ₂ (SO ₄)	Ammonium sulfate	Ammonium
1400	s	str	def NH ₃	Dry meat extract	Peptide
1380	s	str	N=O	Sodium nitrate	Nitrate
1270	s	s	st NO ₂	Sodium nitrite	Nitrite
1255	s	m	def NH ₃	Peptone G	Peptide
1245	s	m		Albumin	Protein
1235	s	m	def c=O	Oleic acid	Fats
1135	s	s	st S=O (R-SO- OR)	Sodium sulfate	Sulfate
	s	s		Sodium sulfite	Sulfite
	s	s		Sodium thiosulfate	Thiosulfate
1100-1040	s (m)	s	st C-O	Glycerol	Polyalcohol

1100	s	s	st Si-O	Sodium silicate	Silicate
1100-1000	s (m)	s	st CO	Lactose	Saccharide
	s (m)	s		Sucrose	Saccharide
1075	s	w	S=O	Ammonium sulfate	Ammonium
1015	s	s	C-O (CH ₂ -OH)	Glucose	Monosaccharide

b broad band, *s* sharp band, *s (m)* multiple sharp band, *str* strong intensity band, *m* medium intensity band, *w* weak intensity band, *st* stretching vibration, *st* as stretching asymmetric vibration, *def* deforming vibration, *sy* symmetrical vibration, *skelet* skeletal vibration.

569

570 The study of activated sludge before and after thermal treatment by FT-IR helps to
 571 identify the origin of the recalcitrant compounds. The main limitations of the methodology
 572 are the need for samples dehydration, the impossibility of studying labile compounds, and
 573 bands overlapping observed for these complex matrices.

574

575 3. Advances and trends in chromatographic and NMR approaches

576 3.1 Sample treatment

577 The complexity of the composition of the matrices from surface water, wastewater,
 578 and samples from the various steps of the thermal treatment of sludge led to the development
 579 of particle sample treatments. The bulk DOM accumulating in water-soluble compounds is a
 580 heterogeneous mixture of HS, hydrophilic acids, proteins, lipids, polysaccharides, carboxylic
 581 acids, amino acids, chemicals and complex aggregates that vary in size, shape, composition
 582 and physicochemical properties. This complexity requires consideration of the heterogeneity

583 of the samples that induce variance related to the scale and physical properties of the matrices
584 (colloid concentration, viscosity, depth of sampling) [99]. Extraction and pre-concentration of
585 all the dissolved organic species are critical to ensuring precise characterization and
586 quantification. Water matrices can be fractionated using resin adsorbents in six classes:
587 aquatic HS, hydrophobic acid (HoA), hydrophobic bases (HoB), hydrophobic neutrals (HoN),
588 hydrophilic acids (HiA), hydrophilic bases (HiB), and hydrophilic neutrals (HiN), as
589 presented in Fig. 5. HoA are represented by humic and fluvic acids; HoB aromatic amines;
590 HoN hydrocarbons, pesticides and carbonyl compounds; HiA sugars acids, fatty acids and
591 hydroxyl acids; HiB proteins and amino sugars, HiN oligo and polysaccharides.

592 After filtration and the elution steps on three cation and anion resin columns, the
593 distribution of these six classes were quantified by UV absorbance at 285 nm and dissolved
594 organic carbon (DOC), and varied substantially depending on the kind of wastewater and the
595 type of treatment. The ratio calculated with UV absorbance vs DOC was also investigated on
596 each sample fraction and provided recalcitrant compound identification [100]. Amberlite
597 XAD resins, which are porous, spherical, macroreticular polystyrene polymers, are the most
598 commonly used sorbents, which can separate NOM into hydrophobic and hydrophilic
599 fractions [101,102]. Recalcitrant organic matter present in surface water can be determined
600 after this total DOM sample preparation procedure and is separated into aromatic or aliphatic
601 species and fractions [103]. The specific fractionation technique has been improved and
602 adapted to characterize HS in soils and their water-soluble fractions [104] by combining
603 liquid-liquid extractions (LLE), using several organic solvents and water, as well as size
604 exclusion preparative chromatography (prep HPSEC) on a Biosep SEC-2000 column. The
605 fractions obtained, containing specific polar and molar mass species, were then characterized
606 by several GC/MS and NMR techniques for humeomic structural studies. Prep HPSEC is a
607 chromatographic method used in water sample preparation to separate NOM molecules

608 according to their size, molecular weight and hydrodynamic volume [105]. The development
609 of high-performance SEC columns allowed the generalized development of prep HPSEC for
610 fractionation during sample treatment steps for the study of NOM and recalcitrant compounds
611 [106]. The use of solid phase extraction (SPE) cartridges prior to reversed-phase analysis,
612 enabling the NOM to be split using different organic solvents of decreasing polarity. This
613 technique has been used to isolate nitrosamine precursors, which are present in wastewater.
614 The use of serial SPE supports improves the resolution of specific toxicity and recalcitrant
615 compound fractions in water samples [107]. The polarity rapid assessment method (PRAM) is
616 a development dedicated to the study of NOM according to the polarity of the species. The
617 protocols consist of sampling different SPE cartridges before and after, using different organic
618 solvents. The selectivity of the SPE cartridges, as well as their porosity, are decisive for
619 splitting. PRAM allows elution of the same species at the same steps, without any matrix
620 variations (pH, ionic force, composition) of the wastewater samples [108]. It is possible to
621 split NOM and rDOM by membrane filtration, with microfiltration, ultrafiltration, and
622 nanofiltration processes with different molecular weight cutoffs of 0.5 K, 1 K, 3 K, 10 K, 30
623 K, and 100 K [109]. Fundamentals studies of the steric and electrostatic interactions between
624 NOM and membrane materials show that NOM is very complex in structure [79]. The split
625 process induces molecular volume and charges of the analytes, then pore size and charges of
626 the membrane. Membrane filtration is very effective for NOM fractionation based on
627 molecular structure and size, the main obstacle with this method is membrane fouling and the
628 matrix effect for concentrated samples [110].

629 Field-flow fractionation (FFF) and Flow field-flow fractionation (FIFFF) are used to
630 separate the NOM constituents by size distribution. In the FFF system, a field is applied to a
631 water sample, perpendicular to the direction of the flow pumped into a fine canal with a long
632 length, resulting in the separation of the macromolecular species. The FIFFF system consists

633 of a separation cell with an ultrafiltration membrane on the bottom. The pore size of the
634 membrane represents the lower-cutoff of the macromolecular analytes. A perpendicular field
635 is applied to provide retention of the analyte, due to the diffusion variation linked to the size
636 distribution [111].

637 The pH control of the samples and the concentration of divalent and trivalent metal
638 ions are necessary to guarantee the DOC physicochemical measurements and fluorescence
639 spectra [112]. It is often necessary to adjust the pH of all the samples either to that of the
640 mobile phase used in the analytical steps, or to set it at an acid value, generally between pH =
641 6 and pH = 4, by HCl additions. Removal of metal ions can be done by adding chelating agent
642 or cation exchange resin (chelex), then filtration. Molecular complexity affects some
643 analytical techniques when the detection limit by species is reached and requires
644 concentrating fractions or samples. The most commonly used concentration techniques are
645 pressure-reducing evaporation, freeze-drying, and more rarely electrodialysis and reverse
646 osmosis. Evaporation and freeze-drying techniques require the least amount of samples, but
647 may lose volatile compounds. Reverse osmosis is the technique that eliminates all insoluble
648 ions and particles and respect all organic species but requires a large amount of sample.
649 Electrodialysis eliminates ions and concentrates organic matter, but the cost of use is high
650 [113]. Sample preparation for highly heterogeneous wastewater matrices, especially after heat
651 treatment, must take into account changes in viscosity and foam formation. These
652 mechanisms, characteristic of the amphiphilic macromolecular species, synthesized by
653 thermal means, can alter the composition of fractions and samples. Since amphiphilic
654 constituents need to be isolated and determined, it is possible to quantify the foaming
655 potential of each sample using standardized methods to control homogeneity. The addition of
656 anti-foam is sometimes necessary [28]. The selection of a sampling method for the analysis of
657 NOM and rDOM in wastewater before and after treatment, (thermal, oxidation) is dependent

658 on the nature of the samples and the final objectives. Classically, wastewater sampling is
659 conducted in two directions, depending on the next analytical procedure: grab sampling and
660 simple filtration or centrifugation for measurements over a short time period (< 5 min) or
661 integrated sampling (filtration, fractionation, concentration, adjustments) extending over a
662 longer period of more than one hour.

663

664 3.2 Reversed-phase high performance liquid chromatography (RP-HPLC) and 665 size exclusion chromatography (SEC)

666 RP-HPLC can be used for DOM analysis, depending on the hydrophobicity of the
667 analytes. Typically, an RP-18 stationary phase is used, with an elution gradient mixture of
668 organic modifier and buffer, to split between 2 and 5 species, depending on the samples. The
669 pH and ionic strength of the gradient have a significant influence on the hydrophobic analytes
670 and fraction measurements [114]. RP-HPLC makes it possible to isolate micro-pollutants in
671 wastewater, but it does not discriminate between macromolecules, proteins and
672 polysaccharides.

673 Reversed-phase specific retention mechanisms are used for the fractionation of
674 specific DOM macromolecules based on structural properties, such as the CH₄ vs O ratio.
675 After optimizing the elution gradient, a correlation can be found between the retention time of
676 HS, its C/O ratio, and its molar mass. It is also possible to split HS species according to their
677 degree of unsaturation. The major interest in the use of reversed phases in HS separation is the
678 possibility of structurally separating different isobaric macromolecular species [115].

679

680 SEC is one of the most widely used chromatographic techniques for the investigation
681 of wastewater samples. The NOM, DOM and recalcitrant compounds in these samples are
682 composed of polymorphous and polyfunctional organic. Molecular structures are complex,

683 with a wide range of molecular sizes and functions and the result is a high variance in the
684 physicochemical properties of the solutes studied, with difficulties in developing
685 chromatographic methods with direct stationary phases or RP-HPLC, and retention
686 mechanisms related to hydrophobic or polar interactions.

687 Molecular weight distribution (MWD) and size distribution are specific fundamental
688 properties that need to be measured for DOM in order to understand the physical and
689 chemical characteristics and the reactivity of all organic analytes, especially in wastewater.
690 MWD and size distribution of soluble species in surface water or wastewater are key
691 structural parameters that have been linked to specific properties such as adsorption, metal
692 binding [116], hydrophobic organic partitioning and electrostatic effects [81]. MW is a key
693 factor for the study of recalcitrant compounds generated by DOM reactivity, the study of the
694 bioreactivity and biodegradation and optimization of the water treatment process. Molecules
695 with LMW have smaller molecular volume, allowing nanopore adsorption [117].

696 The fastest technique for studying MWD and size distribution is SEC, with the use of
697 high-performance permeation columns (HP-SEC). The development of SEC methods was
698 carried out to model the different wastewater treatment processes to obtain a profile of
699 samples representative of the DOM based on the properties of the species in solution [118].
700 Four main factors may influence SEC analyses for MWD measurements. The first is related to
701 DOM reactivity, particularly with the stationary phase and the resulting non-steric
702 interactions, these interactions can be minimized by optimizing the type of mobile phase,
703 ionic strength and pH to suppress charge interactions with the use of buffer eluent [119]. The
704 second is associated with data processing of the SEC chromatograms, with the baseline
705 definition and peak integration. Due to the complex mixture analysis, peak construction,
706 baseline recovery and MW cutoff affect the values of average MWs [120]. When the
707 analytical conditions have been set, the third parameter that will affect the measurements will

708 be the calibration of the SEC chromatogram. Calibration standards of the same molecular type
709 (similar MW range and charge distribution) as the water sample must be used. When these
710 standards are not available, as for recalcitrant compounds, the use of macromolecular
711 mixtures including polydextrans (PDX), proteins, oligosaccharides and blue dextran can be
712 used with LMW compounds (glycerol, acetone or methanol) for total diffusion determination
713 [121,122]. Finally, MWD and average MW can be influenced by the detection mode, which
714 changes the shape of the SEC chromatograms [123].

715 For a mixture of soluble molecules such as humic and fulvic substances and
716 recalcitrant compounds, MWD is defined by the number average molecular weight (M_n), the
717 weight-average molecular weight (M_w), and polydispersity (ρ). M_n and M_w were determined
718 using the following eq. (1) and eq. (2):

719

$$720 \quad M_n = \frac{\sum_{i=1}^n h_i}{\sum_{i=1}^n \frac{h_i}{M_i}} \quad (1)$$

$$721 \quad M_w = \frac{\sum_{i=1}^n h_i M_i}{\sum_{i=1}^n h_i} \quad (2)$$

722

723 where “ h_i ” is the peak height at elution volume “ T ” and M_i is the MW at elution volume i . The
724 M_w/M_n ratio gives ρ .

725 As said in section 2.2, compounds formed during thermal sewage sludge processes can
726 be of various MW, from a few Da (ketones, small aromatics, etc.) to 100 kDa (HMW
727 melanoidins), and each of them is more or less refractory. SEC coupled with organic carbon
728 detection (OCD) is often developed to study which MW species are degraded or removed
729 with various treatments such as precipitation, resins and biodegradation [83]. The specific
730 recalcitrant compounds generated by thermal processes showed variations in M_w from 2 kDa

731 to 200 kDa depending on the reactive DOM in the wastewater input [124]. But the question
732 arises as to the limits of measuring polydispersity by measuring the molecular size range for
733 DOM in general, and for charged electrolytes such as HA or melanoidins. By modelling, it is
734 possible to show that, due to the fact that a statistical linear segment can be composed of
735 several ordinary segments, LMW compounds will consist of a limited number of statistical
736 segments. But for HS, LMW species are not necessarily linear molecules, they can be rather
737 irregular particles with a significant influence of functional polydispersity at this small scale.
738 The result is a significant shift between molecular size and molecular weight. The higher the
739 molecular weight, the lower the drift, but there is an analytical bias in the measurement of the
740 molar mass of polymer charged with SEC [81].

741 Knowledge of the size distribution of the DOM is helpful in assessing the properties
742 and reactivity of the analytes. LMW compounds have a lower radius and molecular volume,
743 allowing them to interact with nanopores, excluding HMW molecules. They also have a
744 higher diffusion coefficient that leads to faster adsorption kinetics, and are more soluble in
745 water and can therefore be more bioavailable. Conversely, HMW compounds will be less
746 soluble in water and will have greater free energy of adsorption. They may generate higher
747 electrostatic fields and show higher aromaticity, leading to greater adsorption of hydrophobic
748 compounds [117]. The size distribution of the DOM allows to link the abundance of specific
749 fractions with their sedimentation property and origin in term of microbial biopolymers [125].
750 For nitrogen containing compounds, bioavailability is directly related to molar mass.
751 Degradation kinetics and bioavailability of isolated surface water DON can be linked to
752 specific molar fractions > 1 kDa and < 1 kDa [126]. To take these mechanisms into account, it
753 is possible to compare water samples by integrating a greater number chromatographic
754 descriptors in addition to the size distribution, calculated from the SEC chromatograms. For
755 HSs, 20 SEC descriptors including distribution coefficients and statistical moments that

756 describe the SEC-elution curve, allow structural discrimination for surface water sample
757 comparisons [127]. To correlate the size distribution, specifically with carbon and molecular
758 nitrogen, selective OCD and OND (organic nitrogen detector) detectors can be used.
759 SEC/OCD/OND is a particular development for surface water and wastewater, applied to the
760 study of DOM in order to characterize the SEC fractions by specific organic carbon OCD and
761 organic nitrogen OND detectors [51]. SEC/OCD/OND had already been used to describe
762 thermally treated sewage sludge [23,72,124]. As this technique separates biopolymers,
763 hydrophobic dissolved organic carbon, HS, building blocks and LMW acids and neutrals, its
764 use before and after biodegradation of thermally treated sludge gives information on which
765 fraction is refractory, but also provides some quantification. Studies showed that almost 80 %
766 of refractory matter could be found in humic-like substances (20-30 %), hydrophobic
767 fractions (20-30 %) and LMW neutrals (20-30 %) [23,72].

768 An analysis of SEC retentions is used to develop liquid chromatography at the critical
769 condition (LCCC). LCCC is a technique that provides easy identification of trace components
770 in polymer systems, distribution of functionality and distribution of molecular weight of the
771 polymers in copolymers. The critical condition corresponds to the chromatographic condition
772 in which Gibbs free-transfer energy for this specific species of the polymer is zero. The term
773 enthalpy and the term entropy in the Gibbs equation compensate for each other exactly. At the
774 critical conditions of a specific neutral polymer, variable molecular weight oligomers are
775 released at the same retention time. However, if a functionalized polymer is injected to the
776 critical conditions of the neutral lead, a separation is obtained that is entirely due to the results
777 of the distribution of functionality [128]. By combining LCCC and SEC, it is possible to
778 improve analytical resolution in the DOM study of surface water, especially for HS, with
779 LCCC providing information on functional polydispersity and SEC on the size range of each
780 functional fraction [129].

781

782 Table 7 summarizes the stationary phases and key chromatographic parameters used in
783 HPLC and SEC to identify the different chemical species of rDOM. The major advantage of
784 chromatographic techniques is to introduce separation of compounds based on physico-
785 chemical properties (hydrophobicity, molar mass). The limits of these methods appear at the
786 validation stage and the absence of standardization in the use of stationary phases, solvents
787 and buffer solutions.

Table 7. Stationary phase geometry, porosity and particle size, and analytical conditions for DOM and recalcitrant compound analysis in various matrices and industries.

Matrix/Industry	Column, geometry, porosity, particle size	Eluent	Detection	Compounds identified	Reference
Humic acids, soil and surface water, Water treatment	Asahipak GF 310 HQ 2x250 mm, 400 Å, 5 µm (Keystone Scientific)	ACN: water, acetate and phosphate buffer	UV: 200 nm	Humic acids	[129]
Glucose-Lysine model, food	TSK G3000 7.5x600 mm, 250 Å, 5 µm (Tosoh) YMC-Pack ODS-5-A 4.6x250 mm, 120 Å, 5 µm (YMC , GmbH)	0,05 M phosphate buffer, pH 7.2 ACN: water	UV : 280 ; 450 nm PDA	Melanoidin 1 spc. Maillard intermediates, 2 spcs.	[130]
Surface water, Water treatment	Nucleosil 100-3 C18 HD 100x2.1 mm, 100 Å, 3 µm	Water: MeOH (5 mM ammonium acetate) gradient	ESI-HRMS	Phenazone, byproducts	[131]

Instant coffee, food	PL aquagel-OH MIXED 7.5x300 mm, Macroporous, 8 µm (PL Laboratory)	0.1 M NaNO ₃	RI	Melanoidins 6 spcs	[132]
green coffee beans model, food	Hypersil Gold RP C18 100x2.1 mm, 175 Å, 1.9 µm (Thermo Scientific)	Water: ACN (0.1 % formic acid) gradient	PDA-ESI-MS	Chlorogenic acids, Chlorogenic acid–arabinose hybrids	[133]
Surface water model with HS, Water treatment	Gemini C18 50x4.60 mm, 110 Å, 3 µm (Phenomenex)	ACN: water, MeOH: water, modifiers	PDA	Sulfamerazine, byproducts,	[134]
Wastewater	LiChrosorb RP-18 125x4 mm, 100 Å, 5µm	MeOH: water (0.014 M oxalic acid)	UV: 270 nm	Trimethoprim	[135]
Instant coffee, food	Purospher STAR RP-8 EC 150x4.6 mm, 120 Å, 3 µm (Merck)	ACN: water (0.05 mM per-fluorooctanoic acid)	ESI-MS/MS UV: 280 nm	Acrylamide, Hydroxymethyl furfural, Antioxidants, melanoidins (after ultrafiltration)	[136]
Wastewater	Zorbax Bonus RP C18 2.1x150	Water: ACN (0.1%	UV : 285 nm	Dinitrobenzene for DOM	[80]

	mm, 80 Å, 5 µm (Agilent)	formic acid)		sorption characterization	
Surface water	Kinetex XB-C18 2.1x100 mm,	Water: MeOH (0.1%	ESI-MS/MS	Micropollutants, 39 spcs,	[137]
downstream industrial areas and urban WWTPs	100 Å, 1.7 µm (Phenomenex)	formic acid), gradient		correlation with EEMs	
Cocoa beans, Food	Accucore C18 3.0x150 mm, 80 Å, 2.6 µm Thermo Fisher	Water: ACN (0.1% formic acid), gradient	UV: 280, 320 nm ESI-HRMS	Soluble and bound to melanoidins phenolic compounds, 15 spcs	[138]

789 3.3 Hyphenated techniques

790 Mass spectrometry associated with GC or LC is generally the preferred technology for
791 investigating DOM because it has high recognition potential. Hyphenated methods covering
792 spectroscopic and spectrometric techniques for the analysis of DOM can follow different
793 strategies: scan targets, suspect objects, and non-targets as presented in Fig. 6, using mass
794 spectrometry to identify targeted or unknown molecules linked to a specific risk.

795 In the case of DOM, which needs to be comprehensively studied, analytes cannot be
796 studied with reference to commercial standards. The important structural variability will
797 require specific developments (sample preparation, fractionation, detection modes) to
798 establish resolute analytical profiles with the aim of replacing reference standards with
799 surrogate parameters.

800 The combination of SEC and PDA detection online is the most commonly used
801 method for DOM spectra analysis. $SUVA_{254}$ provides robust detection of aromatic structures,
802 but cannot discriminate different classes, such as HS, proteins or melanoidins. Specific
803 wavelengths can be analyzed for each chromatographic peak to correlate water quality and
804 SEC profiles. DOM with LMW (< 0.5 kDa) shows a peak area measured at 210 nm linked to
805 oxidized nitrogen concentrations and the concentrations of active bacteria in drinking water
806 [140]. UV spectral analyses in line with SEC are simple and fast to develop for qualitative
807 comparisons. To further the structural analyses, online fluorescence detection will improve
808 identification scores. In addition to humic and fulvic-like fluorescent compounds, a large
809 variety of protein-like compounds can be monitored, linked to MWD by their specific
810 tyrosine and tryptophan-like fluorescence. Hyphenated SEC/PDA/Fluorescence techniques
811 are used for the characterization of specific fractions for a complete surface water analysis as
812 shown in Fig. 7.

813 The specific protein fluorescence properties are related to the fact that of the 20 natural
814 amino acids, only 3 can emit fluorescence after excitation at a characteristic UV wavelength:
815 phenylalanine, tyrosine and tryptophan. Since these 3 amino acids are not frequent in natural
816 proteins, each of them produces a characteristic spectrum [95]. Phenylalanine has the lowest
817 wavelengths for excitation and emission, often incompatible with the solvents and buffers
818 used in SEC. Tryptophan has the highest quantum yield, and its fluorescence properties are
819 influenced by the environment (polarity, pH solvent type). For water and wastewater analysis,
820 tyrosine-like and tryptophan-like proteins are determined by selecting the specific
821 excitation/emission wavelengths of tyrosine (220/310 nm) or tryptophan (270/355 nm). Other
822 biopolymers can emit specific fluorescence, such as HA and FA (330/425 nm and 390/500
823 nm) [141]. The aromaticity of each SEC fraction can be quantified: The absorbance ratio at
824 210 nm vs 254 nm provided surrogate quantification of the aromatic character and
825 hydrophobic/hydrophilic properties of each MW species [142]. From these considerations,
826 reading Fig. 7 shows that the study of 7 different classes of samples – raw water and 6
827 samples from various treatments – is possible from the analysis of 6 SEC fractions. The left
828 column shows the chromatograms obtained for each of the 4 specific Em/Ex pairs of tyrosine-
829 like and tryptophan-like proteins, as well as UV absorption at 210 and 254 nm. The results
830 show that the ozonation process studied in the publication produces a poor decreases the
831 aromaticity of 6 fractions. HA and FA are indeed recalcitrant compounds to this process,
832 since the areas of the 6 SEC fractions are almost identical before and after treatment, while
833 some tryptophan like proteins are degraded.

834

835 3.4 High-resolution mass spectrometry

836

837 HRMS is used to improve detection resolution for the identification of low
838 concentrations of each of the soluble species isolated by a chromatographic technique. In
839 particular, the non-targeted studies presented Fig. 7 will involve basic analyses to identify the
840 structures. The complexity of the structures entails using different approaches to separate and
841 extract HRMS data. For the study of LMW compounds, RP18 separation is often sufficient.
842 The state of charge of the species and possible adducts should be taken into account. The
843 presence of many mineral ions in solution requires searching for multiple adducts $[M+H]^+$,
844 $[M+Na]^+$, $[M+K]^+$, $[M+NH_4]^+$, for monocharged species in ESI+ or $[M-H]^-$, $[M+Cl]^-$, $[M+Br]^-$
845 in ESI- mode. Wastewater analysis also shows complex species with divalent cations or more.
846 In reversed phase, HRMS hyphenation, using a simple aqueous eluent with an acid modifier
847 (formic acid), can be used to search for only protonated or deprotonated species with a simple
848 charge state of +1 or -1 [143]. LC-HRMS analysis of monoisotopic signals allows authors to
849 identify 2409 molecular features that provide probabilities of raw formulas of chemical
850 species by structural database comparison. The highest mass accuracy allows formulas to be
851 assigned with an error < 5 ppm. Compared to conventional COD, this method allows
852 molecular-scale monitoring for MW < 1 kDa. Hydrothermal liquefaction of wastewater from
853 sewage sludge shows the diversity of DOM, identified by an elemental analysis of HRMS
854 data [23,63]. Non-targeted HRMS studies allow structural analysis of DOM, with a specific
855 monitoring of the different chemical species, according to the steps of the treatment process.
856 More than 2,400 species identified by their empirical formulas, can be tracked to identify
857 compounds that are recalcitrant to biological treatment processes. The resolution of HRMS
858 systems, coupled with SEC or RP-HPLC systems, does not allow direct study of compounds
859 with MW > 1000 Da [143]. In targeted studies, more than 111,000 compounds potentially
860 present in the wastewater DOM can be investigated [144]. The majority of species found in
861 wastewater contained C, H, O elements accounting for 38–43 % of species detected in ESI-

862 mode, and 53-83 % of species detected in ESI+ mode. Species containing C, H, O, N, P, S,
863 and Cl were essentially found in ESI-, those with C, H, O, and N in ESI+. Very few
864 compounds with P, S, and Cl were detected in ESI+ [145]. Elemental analyses are used to
865 calculate specific indices to characterize the DOM [146]. It is possible to determine the H/C,
866 N/C, O/C ratios for all identified species and to graphically represent these ratios in a Van
867 Krevelen plot, which gives resolution mapping as presented Fig. 8.

868 These ratios can be determined for certain combinations of elements: CHO, CHON,
869 CHONS, including halogen. H/C vs O/C ratios identify the main chemical families found in
870 wastewater: 1 - lipids, 2 - proteins, 3 - amino sugars, 4 - carbohydrates, 5 - condensed
871 hydrocarbons, 6 - lignin, and 7 - tannin (Fig. 8). For each chemical family, it is possible to
872 follow the rDOM species that are always present after a treatment process. Fig. 8 also shows
873 that recalcitrant compounds are present in all chemical classes, but that some structures in
874 each class are less recalcitrant, especially when the H/C ratio is close to 2. The double bond
875 equivalent (DBE) measurements reflect the level of unsaturated compounds in wastewater.
876 With this classification, it is possible to monitor the specific organic nitrogen in wastewater
877 effluents and to identify the consumed, produced and recalcitrant compounds at the molecular
878 level, for a reactivity characterization of the different effluents [147,148]. The reactivity of
879 DOM species can also be linked to the formation of recalcitrant by-products, such as
880 nitrosodimethylamine, which is promoted by the presence of species identified by HRMS
881 with high hydrogen saturation ($H/C > 1.5$) [149]. To model the aromaticity of DOM, it is
882 possible to calculate a DBE minus oxygen (DBE-O) index in order to discriminate species
883 with carboxylic function (rich oxygen species) and the ratio between DBE and carbon number
884 (DBE/C), which gives a more accurate picture of the unsaturated carbonated skeleton of
885 macromolecules. DBE estimates the unsaturation of DOM for all class of compounds, that is a
886 major advance compared to $SUVA_{254}$, which only applies to UV-active molecules

887 [150,151,152]. Significant progress is made in the comparative study of wastewater, by
888 identifying the predominant species. In particular, among the raw DOM, HRMS allows to
889 isolate the characteristic signals of a molecular signature. In the study of a treatment process,
890 the authors assigned 4 m/z signals 297.1530, 311.1686, 325.1843, 339.1999 which have been
891 identified as anthropogenic alkyl benzene sulfonates [150,151]. The resolution of HRMS
892 systems, which has been constantly growing for more than 20 years, enables us to detect
893 different exact masses for around ten compounds of the same molar mass, with MW < 1000
894 Da, separated in the same chromatogram peak. Ultra high resolutions are needed to resolve
895 more eluted compounds with small MW or polymers. For MW < 1000 Da compounds, it is
896 possible to calculate the Kendrick mass defect (KMD) of each identified species by dividing
897 the nominal molar mass by the exact mass of a selected molecular fragment (CH_2 , $\text{C}_2\text{H}_4\text{O}$,
898 etc.) [148,146]. Compounds with the same number of heteroatoms, rings and double bonds,
899 but different numbers of CH_2 groups, will have identical KMD.
900 Homologous structure classification is then possible, as shown in Fig 9.

901 Fig. 9 shows the identification of 31 homologs with CH_2 and 38 with $\text{C}_2\text{H}_4\text{O}$ in and
902 out of treatment process, all of them were retained having KMD error < 5 ppm. Each graph
903 shows the Kendrick mass defect as a function of nominal Kendrick mass for all ions observed.
904 The homolog species, for example the members of the CH_2 series, have the same Kendrick
905 mass defect but different nominal Kendrick mass and are identified along the horizontal lines
906 on the graph.

907 Monitoring water samples for screening for known compounds and emerging
908 pollutants uses hyphenated methods with LC or GC and HRMS, with a search for the exact
909 masses of parent ions, precursors or degradation products identified in commercial or home-
910 made databases. In this study, 33 pharmaceuticals and 113 of their known transformation
911 products were identified and quantified, with a limit of quantification in order of $\text{ng}\cdot\text{L}^{-1}$ in

912 wastewaters. The main advance was obtained by differentiating between isobaric structural
913 isomers [153]. Recently, the characterization of DOM at the chemical formula level was
914 achieved by using ion mobility and FT-ICR tandem mass spectrometry. The orthogonal use of
915 HRMS and MS/MS is capable to isolate species with m/z difference of 0.036 [154]. For
916 general DOM studies, the sample preparation is limited to filtration. For certain classes of
917 analytes, such as emerging contaminants and pharmaceuticals, which are also recalcitrant
918 compounds, pH adjustments and SPE extraction for concentration and matrix simplification
919 are required [155]. Multi-residue HRMS analyses and quantification are possible for
920 wastewater study, with relative standard deviation $< 20\%$ [156], which allow to develop new
921 tools for evaluating molecular-scale treatment processes [157]. If the research involves the
922 identification of enantiomers, enantioselective separations can be developed, particularly in
923 supercritical mode, with tandem mass spectrometry for wastewater matrices. The interest of
924 SFC-MS is found in fast gradients (< 10 min). The method shows significant performance in
925 this wastewater matrix, with a linearity of calibrations in a wide range (0.005 to 1000 mg. L^{-1}
926 in ESI+ and up to 8000 mg. L^{-1} for ESI-), a recovery rate $> 70\%$ and a relative standard
927 deviation $< 20\%$ [158]. Some antibiotics, such as macrolides, and in particular azithromycin,
928 have been identified as recalcitrant compounds after the various stages of wastewater
929 treatment. Identification was conducted using reverse-phase liquid chromatography with an
930 ESI-ion trap detector in MS/MS mode [159]. Some UV absorbers used in cosmetics, personal
931 care products and paints are found in wastewater and are also recalcitrant compounds, such as
932 benzophenone derivatives, identified by reversed-phase liquid chromatography and tandem
933 mass spectrometry [160]. For an overall characterization of the refractory organic matter, an
934 initial ultrafiltration fractionation technique can be combined with a UV absorption or
935 fluorescence analysis on each fraction, followed by LC/MS and GC/MS analyses. This
936 identifies volatile compounds by extraction using a non-polar organic solvent and of those

937 which are more polar, and water-soluble. In this case, specific extraction methods are needed,
938 with long evaporation steps and recovery with compatible solvents [161]. Maillard precursors
939 and LMW reaction products are usually identified by reversed-phase liquid chromatography
940 with HRMS. The identification of known reducing sugars and amino acids can be achieved by
941 elemental analysis, with a high accuracy (< 4 ppm). For the identification of new Maillard
942 precursors, reaction models with a mixture of different sugars and amino compounds are
943 prepared. After reaction, the analysis of all the identified compounds with tandem mass
944 spectrometry can produce identified reaction products by fragmentation [162]. Maillard
945 reaction products and Amadori compounds, which are formed during the first steps of the
946 Maillard reaction, could be analysed by reversed-phase liquid chromatography with mass
947 spectrometry or tandem mass spectrometry, but the main problem is quantification in complex
948 matrices, given the lack of analytical and isotopic standards [48]. Melanoidins have been
949 extensively studied in the food industry. By combining SEC for splitting by MW, then
950 analyzing each fraction in reversed-phase liquid chromatography coupled with mass
951 spectrometry, it was possible to identify melanoidins with very HMW (> 200 kDa) and to
952 study the oligosaccharide moieties of each fraction. With this 2D technique, it is possible to
953 identify the m/z signal of the hexose unit (162 Da) and the different degrees of
954 oligosaccharide polymerization (DP), from DP2 (m/z : 341 Da) to DP4 (m/z : 665 Da) [163].
955 Melanoidins are negatively-charged complex polymeric structures of various molecular
956 weights, determined by various chromatographic and HRMS techniques [32,71]. The analysis
957 of melanoidin degradation products after hydrolysis, using high-performance anion-exchange
958 chromatography with pulsed amperometric detection (HPAEC-PAD) or high-performance
959 thin layer chromatography (HPTLC), provides structural information. By increasing the
960 degree of polymerization of carbohydrates used as reducing sugars in the Maillard reaction,
961 the release of monosaccharides increased from about 3 % to 95 %, depending on the

962 carbohydrate [31,164]. The study of SEC fractions of melanoidins by pyrolysis/GC/MS
963 provides identification of products. The origin of nitrogen-containing polymers can be
964 assigned to specific Amadori intermediates, as presented in Table S1 (supplementary data).
965 All accurate masses are understood to exact masses to identify raw formulas of the
966 intermediates [52]. The characterization of DOM in solution, generated in particular during
967 thermal treatments, can be achieved by ESI Fourier-transform ion cyclotron resonance mass
968 spectrometry (ESI-FT-ICR MS), coupled with SEC fractionation [23]. The very high
969 resolution of these detectors has been widely used to elucidate complex petroleum oil
970 structures. In the case of wastewater, the DOM can be characterized in a single analysis and
971 compounds classified according to structural criteria: composition in heteroatoms and notably
972 nitrogen, DBE, carbon number in the structure and all ratios related to these elementary
973 analyses. Depending on the species sought, it is possible to search for specific classes, as
974 shown in Fig. 10, presenting the preheat treatment and post-heat treatment fraction study, with
975 the production of HMW conjugate species after treatment. It should be noted that this
976 technique only identifies ionized species by the ESI probe and that species without ionisable
977 functions will not be identified. Quantification is only possible by using isotopic standards of
978 the species studied, in the absence of these standards, comparison of samples is done by
979 analysing relative abundances.

980 Fig. 10 clearly shows the increase in the relative abundance of all classes of nitrogen-
981 containing compounds after thermal treatment of sludge, observed by an increase in intensity
982 (red color) on the graph. There is also an increase in the number of carbon and DBE, showing
983 that higher molecular weight species are generated, and more conjugated (DBE shifted to
984 higher values).

985 Table 8 summarizes the stationary phases and eluent used in hyphenated methods for
986 DOM characterization. SEC and reversed-phase columns were largely used for the separation

987 and elemental analysis of LMW and HMW DOM and recalcitrant compoundsUV
988 spectrometry and fluorescence detection, widely used to characterize rDOM, do not allow
989 structural identification and, in particular, do not allow the study of specific species such as
990 nitrogen containing compounds. HRMS is widely used in global DOM studies and molecular-
991 scale characterization of rDOM. The elemental analysis, possible by HRMS, allows to
992 identify the molecular structures that are related to refractoriness. Important advances are
993 possible for targeted and non-targeted monitoring, allowing an optimization of wastewater
994 treatment processes. But HRMS studies are only possible, for the majority of detectors
995 currently available, for molecules of MW <1000 Da. This limitation explains the low number
996 of studies available for melanoidins. For these macromolecules, advances were obtained by
997 ESI-FT-ICR MS. For this last technique, it is the very high molecular diversity of species that
998 will limit the interpretation of mass spectra.

999 Table 8. Stationary phase geometry, porosity and particle size, analytical conditions and detection modes used in the different hyphenated
 1000 methods developed for DOM and recalcitrant compound analysis in various matrices and industries.

Matrix/Industry	Column, geometry, porosity, particle size	Eluent	Detection	Identified compounds	Reference
Humic acids, soil and surface water, water treatment	Asahipak GF 310 HQ 2x250 mm, 400 Å, 5 µm (Keystone Scientific)	ACN: water, acetate and phosphate buffer	UV: 200 nm	Humic acids	[129]
Glucose-Lysine model, food	TSK G3000 7.5x600 mm, 250 Å, 5 µm (Tosoh)	0.05 M phosphate buffer, pH 7.2	UV : 280 ; 450 nm	Melanoidin 1 spc.	[130]
	YMC-Pack ODS-5-A 4.6x250 mm, 120 Å, 5 µm (YMC, GmbH)	ACN: water	PDA	Maillard intermediates, 2 spcs.	
Surface water, water treatment	Nucleosil 100-3 C18 HD 100x2.1 mm, 100 Å, 3 µm	Water: MeOH (5 mM ammonium acetate) gradient	ESI-HRMS	Phenazone, by- products	[131]
Instant coffee, food	PL aquagel-	0.1 M NaNO ₃	RI	Melanoidins 6 spcs	[132]

	OH MIXED 7.5x300 mm, Macroporous, 8 µm (PL Laboratory)				
green coffee bean model, food	Hypersil Gold RP C18 100x2.1 mm, 175 Å, 1.9 µm (Thermo Scientific)	Water: ACN (0.1 % formic acid) gradient	PDA-ESI-MS	Chlorogenic acids, Chlorogenic acid– arabinose hybrids	[133]
Surface water model with HS, water treatment	Gemini C18 50x4.60 mm, 110 Å, 3 µm (Phenomenex)	ACN: water, MeOH: water, modifiers	PDA	Sulfamerazine, by- products	[134]
Wastewater	LiChrosorb RP-18 125x4 mm, 100 Å, 5 µm	MeOH: water (0.014 M oxalic acid)	UV: 270 nm	Trimethoprim	[135]
Instant coffee, food	Purospher STAR RP-8 EC 150x4.6 mm, 120 Å, 3 µm (Merck)	ACN: water (0.05 mM per-fluorooctanoic acid)	ESI-MS/MS UV: 280 nm	Acrylamide, Hydroxymethyl furfural, Antioxidants, melanoidins (after ultrafiltration)	[136]

Wastewater	Zorbax Bonus RP C18 2.1×150 mm, 80 Å, 5 µm (Agilent)	Water: ACN (0.1 % formic acid)	UV: 285 nm	Dinitrobenzene for DOM sorption characterization	[80]
Surface water downstream industrial areas and urban WWTPs	Kinetex XB-C18 2.1x100 mm, 100 Å, 1.7 µm (Phenomenex)	Water: MeOH (0.1 % formic acid), gradient	ESI-MS/MS	Micropollutants, 39 spcs, correlation with EEMs	[137]
Cocoa beans, food	Accucore C18 3.0×150 mm, 80 Å, 2.6 µm Thermo Fisher	Water: ACN (0.1 % formic acid), gradient	UV: 280, 320 nm ESI-HRMS	Soluble and bound to melanoidins phenolic compounds, 15 spcs	[138]
Wastewater	Acquity UPLC BEH C18 2.1x50 mm, 130 Å 1.7 µm Waters SAS	Water: ACN (0.1 % formic acid), gradient	ESI-HRMS	2409 compounds detected, Van Krevelen plots	[143]
Wastewater	Hypersil Gold PFP, 100x2.1 mm, 175 Å, 1.9 µm Thermo Scientific	Water: ACN (0.1 % formic acid), gradient	ESI-HRMS	33 pharmaceuticals and 113 by-products	[153]

1001 3.5 Nuclear magnetic resonance spectroscopy (NMR)

1002 NMR approaches are not yet extensively used in the field of complex mixture analysis
1003 in the environmental sciences and the characterisation of NOM. In 2019, a review [166]
1004 summarized the main input of NMR over the past twenty years on the characterization of the
1005 organic matter in water. Both liquid-state and solid-state NMR experiments can be performed
1006 using dedicated spectrometers and probes [167–170]. NMR spectroscopy and MS, has the
1007 advantage of providing structural information, based on elemental composition and chemical
1008 shift, on sample content and also emphasizes the holistic characterization of a sample. NMR
1009 can be used to observe the nuclei of elements with a nonzero spin quantum number, and
1010 several active nuclei can be detected, the most common being protons, carbon-13, fluorine-19,
1011 phosphorus-31 and, to a lesser extent, nitrogen-14 or-15. NMR spectroscopy is highly
1012 reproducible, robust, and quantitative by nature since the NMR signal response i.e. the
1013 integrated signal area, is proportional to the number of nuclei contributing to the signal [171].
1014 The fact that NMR is a non-selective method is an advantage for the analysis of mixtures to
1015 the extent that all the compounds present in the solution (provided they contain the nucleus
1016 under investigation) are detected simultaneously in a single run. It is also a non-destructive
1017 technique that can provide relevant information on the composition of the mixture from the
1018 analysis of different nuclei and through various NMR experiments. Analyses range from one
1019 dimensional experiments focusing on ^1H [172] or, for most studies, on the ^{13}C nucleus
1020 [167,169] to homonuclear and heteronuclear 2D experiments like ^1H - ^1H COSY or ^1H - ^{13}C
1021 HSQC commonly used in liquid-state NMR [167]. Beyond these common nuclei, it should be
1022 mentioned that the ^{15}N nucleus was also successfully investigated for organic matter
1023 characterization in solid-state NMR with CP-MAS (Cross Polarization Magic-Angle
1024 Spinning) experiments and DEPT (Distortionless Enhancement by Polarization Transfer)
1025 sequences in solutions [173,174]. Furthermore, studying the phosphorus nucleus can be useful

1026 to obtain the distribution of different phosphorus compound classes in environmental samples
1027 [175,176]. For instance, ^{31}P liquid and solid-state NMR were recently used to understand
1028 changes in phosphorus speciation after thermal hydrolysis pre-treatment of sludge. The
1029 analysis demonstrated that orthophosphate was the most abundant phosphate present [177].
1030 Lastly, other specific NMR approaches based on Low-Field NMR relaxometry techniques for
1031 probing the mobility of protons have been found useful in characterizing the water present in
1032 biomass [178,179].

1033 To the best of our knowledge, none of the studies reported characterization of
1034 recalcitrant compounds using NMR after a thermal degradation process. The main recent
1035 studies involving the characterization and analysis of recalcitrant compounds, however, are
1036 reported in Table 9. Depending on the nature of the samples, liquid- or solid-state NMR was
1037 carried out. The presence of recalcitrant compounds in Brazilian urban wastewater, which
1038 mainly include surfactants and phthalates, was investigated in 2015 [180] and 2017 [181]. In
1039 these studies, both untargeted-NMR profiling (^1H and HSQC), leading to quantitative
1040 estimation of some pollutants, and structural identification of recalcitrant compounds with
1041 conventional 1D and 2D NMR experiments were performed. Moreover, in the 2017 study
1042 [181], multivariate chemometric analyses from quantitative NMR data of the aromatic region
1043 were used to monitor the degradation of organic matter and treatment performances. In 2014,
1044 there was a report [182] on the characterization by several mono- and bi-dimensional NMR
1045 techniques (COSY, TOCSY, DOSY) of a refractory end product of HA subjected to a
1046 humeomics procedure i.e. previously separated in size fractions by preparative HPSEC
1047 chromatography. In a study performed in 2019 [183] the application of both solution-state
1048 (2D ^1H - ^{13}C HSQC) and solid-state (^{13}C -CPMAS) NMR is reported to characterize particulate
1049 organic matter (POM) in agricultural and food waste digesters collected from anaerobic
1050 digestion plants in Sweden. In this study, the combination of NMR data and multivariate

1051 chemometrics was used to characterize both the labile POM fraction in solid waste (lipids,
1052 starch-like carbohydrates) and recalcitrant POM structures (protein-like compounds, lignin,
1053 polysaccharides). The recording of ^{13}C NMR spectra with CP-MAS NMR is particularly well
1054 suited to analyzing soil samples with minimal sample preparation. The method can be used to
1055 assess the relative proportions of functional groups corresponding to a specific chemical shift
1056 area i.e. alkyl (δC : 0-47 ppm), O-alkyl (δC : 47-90 ppm), anomeric (δC : 90-108 ppm),
1057 aromatic (δC : 108-145 ppm), O,N-substituted aromatic (δC : 145-167 ppm), and carbonyl
1058 groups (δC : 167-220 ppm). A similar approach using CP-MAS was recently used to
1059 investigate the chemical composition of cultivated and forest soils in subtropical and
1060 temperate regions [184]. Advanced pulse sequences were used for the characterization of
1061 recalcitrant aromatic components in soil organic matter in order to study their changes in
1062 relation to land use and soil management in the Mollisol region of China [185]. The
1063 composition of recalcitrant humin from estuarine sediments was studied [186] by combining
1064 solid-state NMR and Comprehensive Multiphase-NMR (CMP-NMR). The advantage of
1065 CMP-NMR is its use in performing a full range of NMR experiments on the same unaltered
1066 natural sample, without prior extraction, containing liquid-, gel- and solid-like phases. This
1067 approach made it possible to conclude that the recalcitrant OM in estuarine sediments is
1068 mainly composed of proteins. All the NMR experiments reported above and in Table 9 for the
1069 characterization of recalcitrant compounds demonstrated that NMR can be used to (i) screen
1070 samples holistically (ii) quantitate samples (iii) determine the structure of recalcitrant species
1071 after chromatographic purification.

1072

1073

1074

1075

1076 Table 9: Summary of recent publications using NMR approaches for the investigation of recalcitrant compounds

Matrix	Recalcitrant compounds	NMR experiments	Spectrometer ^1H frequency, probe, temperature, specific procedure or data treatment	Reference
Liquid-state NMR				
Humic acid from volcanic soil	refractory end product of a humic acid	^1H , ^1H DOSY, COSY, TOCSY, ^1H - ^{13}C HSQC,	- 400 MHz, 5 mm BBI probe, 298 K - water suppression with presaturation or Watergate 3-9-19	[182]
Urban wastewater	linear alkylbenzene sulfonate surfactant homologs	^1H , COSY, HSQC, HMBC	- 600 MHz, 5 mm ^1H - $^{13}\text{C}/^{15}\text{N}$ TCI cryoprobe, 298 K - zgpcpr pulse sequence for water suppression - quantification with Eretic2 TM method	[180]
Urban wastewater	surfactants and phthalates	^1H , HSQC	- 600 MHz, 5 mm ^1H - $^{13}\text{C}/^{15}\text{N}$ TCI cryoprobe, 298 K - zgpcpr pulse sequence for water suppression - quantification with Eretic2 TM method - multivariate analysis: PCA, HCA, PLS-DA	[181]
Agricultural and food waste	cellulose-like and protein-like	^1H , HSQC	- 600 MHz, 5 mm BBO cryoprobe, 298 K - multivariate analysis: PCA, OPLS-DA	[183]

digesters substances		Solid-state NMR, CP-MAS		
Estuarine	methylene moieties	^1H , ^{13}C , DOSY,	- 500 MHz, 4 mm prototype CMP MAS $^1\text{H}/^{13}\text{C}/^{19}\text{F}/^2\text{H}$	[186]
sediments	from protein-like substances	TOCSY, NOESY ^{13}C RAMP	probe, 298 K - 500 MHz, 4 mm BB ^1H -X probe	
Soil organic matter	aromatic components	^{13}C CP/TOSS,	- 400 MHz, 4 mm double resonance probehead	[185]
Land (Mollisol)		^{13}C DP/TOSS, ^{13}C DP/MAS	- quantification of specific functional groups - multivariate analysis: PCA,	
Agricultural and food waste digesters	cellulose-like and protein-like substances	^{13}C	- 500 MHz, 4 mm MAS probe with ^1H decoupling - relative C distribution (% of total integrals) - multivariate analysis: PCA	[183]
Soil	carbonyl, phenolic, aromatic compounds	^{13}C	- 400 MHz, 4 mm standard bore probehead, 297 K - statistical analysis: Chi-square, AIC, ANOVA	[184]

1077 As described in section 2.2, melanoidins are refractory end-products of thermal
1078 treatment applied to sewage sludge. NMR experiments had not yet been described for
1079 characterization of melanoidins in the environmental field but were applied to follow their
1080 formation as well as to structural studies [187,188]. In agro-food applications, solution-state
1081 NMR was recently used to investigate the effect of roasting time on the chemical structure of
1082 coffee melanoidins [69] and to characterize them after fractionation by SEC in dulce de leche,
1083 a confectionary dairy product [189]. These studies, along with the bibliography dealing with
1084 the analysis of recalcitrant compounds with NMR, illustrate the feasibility of characterization
1085 of melanoidins from sludge samples after thermal treatment, an area that deserves further
1086 investigation.

1087 In the field of this review, one advantage of NMR is that it requires easy sample
1088 preparation for the analysis of raw samples. Indeed, solid samples can be analyzed after a
1089 drying step using CP-MAS NMR. For solution-state NMR, sample preparation requires
1090 dissolution or dilution in a deuterated solvent, commonly deuterated water for environmental
1091 samples. The only precaution to take for both solid and liquid samples is, if necessary, the
1092 removal of paramagnetic mineral species samples, in particular Fe-containing minerals. This
1093 step can be essential prior to NMR spectroscopy, paramagnetic species would otherwise
1094 induce signal loss and extensive peak broadening in the NMR spectra. It can be done through
1095 prior acidic treatment of the sample with HCl [183] or 2% hydrofluoric acid [190] for liquid-
1096 or solid- state, respectively. However, if NMR is to be applied for the complete structural
1097 characterization of a recalcitrant compound, a prior isolation with a set of chromatographic
1098 methods is required. Another key point of NMR is its ability to quantify compounds without
1099 the need for specific standard compounds, but this advantage should be linked to the main
1100 drawback of NMR, its intrinsic insensitivity. Currently, a limit of detection on the nanomole
1101 level can be considered using conventional 5 mm NMR tubes on a 500 MHz solution-state

1102 spectrometer for a pure compound. Environmental samples consist of a complex mixture of
1103 molecules, and quantification of their components is a real challenge. The quantification of
1104 organic compounds in wastewater samples at low concentrations (micromolar and, in some
1105 cases, nanomolar) has been reported, but this result can only be achieved through prior pre-
1106 concentration by HPLC-(UV/MS)-SPE-NMR with automatic sampling storage [180]. This
1107 example illustrates that, for the characterization of organic matter and recalcitrant compounds,
1108 the power of NMR will be increased if the method is used as an orthogonal and
1109 complementary method with a chromatographic approach. In the coming years, recent
1110 sensitivity improvements based on higher magnetic fields, better probes, advanced pulse
1111 sequences and hyperpolarization [191] may significantly reinforce the role of NMR in these
1112 applications. The combined approach of NMR with the aforementioned chromatographic
1113 and/or spectroscopic methods offers promising integrated analytical tools whose
1114 complementarity would be strengthened by the use of chemometric approaches.

1115

1116 **4. Chemometrics**

1117 Since the nature and properties of DOM in wastewater are extremely heterogeneous,
1118 comparative sample studies involve orthogonal analytical techniques that generate a large
1119 amount of information. Several steps are needed to process this data, from method validation
1120 to combined data analysis, to identify known structures, map samples, or quantify similarities
1121 or differences, seeking to limit the time it takes to process the data. Chemometrics is already
1122 known as a powerful support tool for analyzing, interpreting and solving environmental issues
1123 [192,193]. Due to the huge amounts of measurements obtained on miscellaneous collected
1124 samples, the correct data processing techniques and statistical methods are required but can be
1125 very tricky and difficult to select [194]. Multivariate statistical data analysis, such as cluster
1126 analysis (CA) or principal component analysis (PCA), as well as the discriminant analysis

1127 techniques, can be used as emphasized methods to find relevant patterns and trends in various
1128 analytical datasets. Chemometric approaches are useful for handling and interpreting large
1129 and complex data matrices and for highlighting the correlations between the variables
1130 evaluated. Many graphical representations of the results simplify their interpretation by
1131 reducing and categorizing variables into groups with similar characteristics.

1132

1133 4.1 Cluster analysis (CA)

1134 CA is an unsupervised statistical approach that classifies samples without prior
1135 knowledge of the dataset. This exploratory method organizes data into clusters by maximizing
1136 the similarities between samples within each group and by maximizing the dissimilarity
1137 between groups. Hierarchical cluster analysis (HCA) is the most common method in which
1138 agglomerative clusters are produced using the distance between samples as a measure of
1139 similarity [195]. Results are generally illustrated by a dendrogram where similarities of
1140 structures are expressed by distance depending on the species studied. In water analysis, CAs
1141 are frequently employed when physicochemical parameters are collected. Among the studies
1142 reported, we can cite the work by Singh *et al.* [196] in which 31 physicochemical parameters
1143 were analyzed over 5 years within a stretch of about 12 km from 26 wastewater drains that
1144 discharged into the Gomti River in India. The clusters identified provide information on
1145 sewage and were used to group some drains to be diverted and connected to individual
1146 sewage treatment plants. In another study, twelve geochemical parameters of water from
1147 South Chennai Lake in India were studied to identify the factors affecting the pre-monsoon
1148 and post-monsoon chemical composition of the water. The dendrogram analysis gave clusters
1149 depending of the pollution levels at the stations and was helpful in determining the effects of
1150 the monsoon on each station [197]. In the last example, 22 physicochemical parameters were

1151 used to evaluate the water quality of the Pisuerga River in Spain and to classify samples based
1152 on seasonal and temporal criteria [198].

1153 4.2 Principal component analysis (PCA)

1154 PCA is more commonly used than CA for data from the analytical techniques. PCA is
1155 also an unsupervised multivariate statistical method frequently used in exploratory data
1156 analysis. PCA simplifies the dataset by reducing the number of dimensions and creating new
1157 ones. The principal components are a linear combination of the original variables which
1158 explains most of the variance in the data matrix. The eigenvalues give the variance ratio used
1159 by each principal component. Depending on the normalized data, mean-centering or
1160 autoscaling, and the aim of the analysis, covariance or the correlation matrix is considered.
1161 Score plots (samples) and loading plots (variables) are the graphical representations of PCA
1162 used to determine the relationships between samples and/or variables [199]. PCA is
1163 meaningful for water sample analyses for data resulting from spectroscopy applications.
1164 Indeed, PCA is a powerful tool for analyzing complex matrices and determining the
1165 relationships and correlations between the variables obtained with the same analytical
1166 method. The use of a fast, non-destructive, user-friendly method, such as UV-visible, infrared
1167 and fluorescence spectroscopy, without complex sample preparation, may initially be
1168 preferred to the new and more sensitive methods of analysis such as GC, LC coupled with
1169 mass spectrometry or NMR, to explore biological wastewater treatment processes [200], for
1170 example, or for monitoring wastewater treatment efficiency [192,201]. Of course, PCA is also
1171 applied to the physicochemical parameters to identify the producers of water pollution in
1172 rivers [198] and in the air, or by combining these parameters with analytical results. For
1173 wastewater treatment monitoring, the intensity of specific wavelengths (240, 250, 254, 260,
1174 275 and 365 nm) is measured in UV-visible spectrometry and used to characterize the quality

1175 assessment [202]. The relationship between soil organic properties and the bands of IR
1176 spectra from the corresponding HA is established by PCA [203].

1177 Hyphenated methods with LC or GC for the separation part and NMR and mass
1178 spectrometry as the analyzer, usually require more complex, specific sample preparation (see
1179 sections 3.2, 3.3 and 3.4) for data recording and also dedicated data treatment before
1180 multivariate analysis. First, a peak alignment algorithm, such as icoshift [204,205], can be
1181 applied to synchronize peaks or signals on the chromatograms or NMR spectra due to the
1182 small variations of retention times or chemical shifts. Second, an automatic binning procedure
1183 is used to segment all data into small regions preferably containing only one signal, or else
1184 the selected signals are integrated manually depending on the information that needs to be
1185 determined [206]. PCA has been carried out to investigate variations in the NMR spectral
1186 dataset to characterize the labile POM fraction in solid wastes from anaerobic digestion plants
1187 in Sweden [183], the recalcitrant aromatic components in soil organic matter from
1188 Heilongjiang Province in China [185] or to monitor the degradation of organic matter from a
1189 sewage treatment plant in São Carlos-SP in Brazil [181]. Recently, PCA was used to evaluate
1190 datasets of two chromatographic methods coupled with mass spectrometry of drinking water
1191 samples, wastewater and river water samples from Warburg, Germany [207].

1192

1193 4.3 Partial Least Squares (PLS) and PLS discriminant analysis (PLS-DA)

1194 PLS is a multivariate statistical method and has some similarities with PCA. PLS is
1195 useful for modeling a relationship between a set of dependent variables (Y) and a set of
1196 independent variables (X). PLS regression attempts to maximize the covariance between
1197 matrices X and Y and the model generates latent variables in X that will predict the latent
1198 variables in Y [208]. This linear regression is known as PLS1 if it deals with only one Y

1199 response and PLS2 for several Y responses. A special form of PLS is PLS discriminant
1200 analysis (PLS-DA), wherein Y is an indicator, a value 0 or 1 for example, depending which
1201 class the object is a member of [209]. The score plot shows clustering, while the loading plot
1202 indicates the discriminating variables between groups. A cross-validation method or separate
1203 test sets are used to estimate the optimal number of latent variables and to guard against
1204 overfitting. For multivariate calibration, PLS is also widely used [210]. PLS has been applied
1205 to explore the possible relationship between the IR spectrum and SOC or the humification
1206 index based on HA absorbance at 465 nm [203]. Predictive models including PLS1 and PLS2
1207 were developed to find spectral wavelengths in UV-Vis spectra dataset in correlation with
1208 water quality parameters such as DOM and TOC. This approach was intended to predict the
1209 concentration of several water components and parameters and to be used to monitor the
1210 performances of wastewater treatment plants [202]. Spectral wavelengths were selected based
1211 on variable importance in the projection scores obtained using the model [211]. PLS has also
1212 been applied for the estimation of the characteristics of urban wastewater [212] and for fault
1213 detection in the wastewater treatment process [213]. In addition to wastewater quality
1214 parameters, the permeate of a membrane bioreactors has been successfully modelled by PLS
1215 on 2D-fluorescence dataset [201]. The discriminant version of PLS is a powerful statistical
1216 tool and has been employed to find seasonal variations in the composition of sewage [196]
1217 and to investigate which micropollutant signals in LC-HRMS were the most important in the
1218 discrimination for complex water matrices [207].

1219 The interest of chemical information extraction by data-driven means in the design of
1220 physicochemical indices related to the properties of DOM leads to the specific development
1221 of descriptors dedicated to the study of the biodegradability of DOM and to the study of
1222 recalcitrant compounds. To characterize DOM heterogeneity related to aromaticity, SEC-
1223 generated MWs can be combined with fluorescence spectra collected between 300 nm and

1224 480 nm (Em) for an excitation wavelength (Ex) set at 254 nm. By calculating all humification
1225 indices (HIX) for each emission spectrum (HIX_i), it is possible to calculate HIX-based MW
1226 values (MW_{HIX}) from the SEC chromatogram for each separated species. Water samples from
1227 different treatments can be compared based on the MW_{HIX} to identify refractory species and
1228 to study the evolution of aromaticity and biodegradability [193]. All of these approaches have
1229 also been applied to large microscopy datasets to monitor activated sludge in wastewater
1230 treatment processes [214]. Fig. 11 summarize the different chemometric methods used for
1231 DOM and recalcitrant compound characterization.

1232 Fig. 11 shows a dendrogram derived from a hierarchical CA for DOM samples (A)
1233 showing a link between marine samples, with a correlation of > 0.5 for offshore or estuary
1234 samples, and a correlation of > 0.1 between the 2 types of sampling. However, there is no link
1235 for samples taken in the river upstream of the estuary. To identify these links, formulas
1236 developed from 2,143 compounds from a library developed by the authors were searched in
1237 each sample. Then the O/C and H/C ratios and various indices (DBE, etc.) were calculated to
1238 characterize each sample. The hierarchical CA was constructed from these indices until the
1239 correlation coefficients were obtained in the matrix. To identify the key parameters, a PCA
1240 analysis was performed (B), showing that nearly 60 % of the variance is explained by PC1
1241 (58.3 %) related to the aromaticity of the compounds present in the water, calculated from the
1242 DBE and DBE/C ratios. Other chemometric analyses provide a classification of compounds
1243 by chemical families (C), using van Krevelen diagrams representing the H/C vs O/C ratios for
1244 each analyte. It is possible to identify each recalcitrant compound by comparing the graphs
1245 before and after a degradation process. All analytes are identified in their chemical family
1246 based on HRMS analyses. By determining the HIX indices for each peak identified in SEC, it
1247 is possible to represent an HIX chromatogram to identify the most condensed species present

1248 in surface water samples, represented in the figure by the highest peaks. These most
1249 condensed species show significant refractory properties.

1250

1251 **5. Conclusions**

1252 The identification, characterization and quantification of the different chemical species
1253 constituting DOM and rDOM in wastewater, will implement simple physicochemical or
1254 complex hyphenated techniques, depending on the objective. The value of physicochemical
1255 methods (DOC, DON, UV, IR) is found in the actual standardization and normalization and
1256 the simple sample preparation needed. But these methods are limited in the search for
1257 characterizations which will require a specific fractionation step (molar mass, charge,
1258 polarity) and a sensitive and resolute detection technique. The sample preparation step can
1259 also constitute a first level of splitting of this complex DOM. It is essentially ion exchange
1260 resins and membrane filtration techniques that are most used to split DOM and study
1261 recalcitrant macromolecular compounds after removal of low molecular weight fractions. For
1262 these techniques, no automation is envisaged and no standardization studies are available. Yet
1263 solute/membrane interactions are an important negative point to be studied for robustness. It
1264 appears that the sample preparation step is more oriented towards the analytical method that
1265 follows than towards a normalization of the treatment to guarantee the extraction of all the
1266 species present for identification and quantification. It therefore seems essential to develop
1267 studies from standards, such as melanoidins synthesized under specific conditions, or
1268 commercial HA and FA, to define the limits of each method of sample preparation. New
1269 developments appear in FFF and FIFFF and provide repeatable, robust fractionation.
1270 Miniaturization of this stage will certainly be a challenge to make it universally usable in
1271 routine.

1272 Recalcitrant identification can be achieved with specific physicochemical ratio such as
1273 SUVA₂₅₄, URI and DOC/DON ratios that provide information and quantify rDOM. Compared
1274 to UV absorbance, fluorescence is therefore more specific and sensitive for discrimination of
1275 recalcitrant compounds, but the limited number of transitions and the absence of specific data
1276 for particular chemical functions require the introduction of another analytical dimension
1277 complementary to fluorescence, such as FT-IR or a separation technique. FT-IR is widely
1278 used for sludge monitoring, since the sample must be completely dehydrated before analysis,
1279 which may limit the scope of application. Significant advances in the characterization of
1280 DOM and notably in the identification of rDOM and all families of recalcitrant compounds
1281 have been achieved using separation methods. The SEC technique is essentially the most used
1282 in the study of NOM, DOM and rDOM. Its main interest lies in the classification of analytes
1283 according to molecular weight distribution, which represents a constant property in the very
1284 large variability of the DOM. The most commonly used stationary phases are the 8 or 5 µm
1285 permeation columns. The main defect of the SEC technique is the existence of secondary
1286 interactions between analytes and stationary phases. It is then necessary to use perfectly
1287 characterized eluents with defined pH and concentration buffers to be optimized for the best
1288 resolution. The result is a lack of normalization for this technique, with each solution
1289 appearing to have better resolution under a particular condition. In addition, the standards
1290 available to calibrate SEC columns do not always correspond to the compounds studied, such
1291 as melanoidins, which do not have standardization. Since MWD is influenced by the mode of
1292 detection, the nature of the solutes, the buffer used and the range of separation of the
1293 stationary phase, it is sometimes necessary to combine the SEC with several spectroscopic
1294 methods, in order to develop hyphenated techniques suitable for the desired structural study.

1295 The fastest and cheapest advances in hyphenated techniques have been achieved with
1296 SEC/PDA/Fluorescence, which provides all the advantages of each element of the analytical

1297 chain in a single injection. It is thus possible to measure the main characteristic indices of
1298 NOM, DOM and rDOM for each chromatograph peak. Relative quantitative analysis is also
1299 possible. If the structural study is to be further developed, the use of HRMS is needed. We
1300 note an attempt to replace separation techniques with the use of very high-resolution MS (FT-
1301 ICR MS), which allows elementary analyses on small or HMW molecules, without
1302 degradation and without prior separation. The use of hyphenated HRMS methods, however,
1303 allows the use of more conventional HRMS detectors, with much success. HRMS techniques
1304 provide a global analysis (elementary, H/C ratio, O/C, N/C, DBE), online. For homologous
1305 compounds, it is also possible to differentiate species. In the literature, major advances in
1306 HRMS are found for small molecules with a MW < 1000 Da. Very few studies are available
1307 for native melanoids and macromolecules of rDOM. Yet the structural study of these
1308 molecules is the key to understanding refractory properties. Only high-precision frequency-
1309 derived data (NMR, HRMS and FT-ICR MS) can provide sufficient resolution to resolve the
1310 nonrepetitive complex structure of rDOM [216]. When using HRMS, the low
1311 chromatographic resolution of SEC was advantageous for spectra measurements of low and
1312 high molecular mass. A correlation between SEC calibration and elemental composition can
1313 provide information on the aromaticity and heteroatom composition of some moieties [159].
1314 On the other hand, the mass spectra are complex to analyze, and it is necessary to study the
1315 positive and negative ESI spectra. Depending on the species and adducts with the different
1316 anions and cations present in the wastewater, a very large number of signals will be generated
1317 for the same species. The matrix effect is strong, which limits quantitative analyses, especially
1318 since isotopic standards are non-existent for macromolecular rDOM. If there are some new
1319 developments in HPTLC, for example, or in Pyr-GC/MS, in libraries under construction –
1320 notably for the study of Amadori compounds – no 2D study has been developed to date for
1321 the analysis of rDOM and melanoidins. It is, however, possible to develop SEC separation

1322 with a second specific dimension based on hydrophobic interactions, as performed to separate
1323 specific proteins from gum arabic with amphiphilic properties [217]. New LCCC
1324 developments provide a superior response to conventional SEC chromatography,
1325 complementing a comparison on functional polydispersity of samples. Future 2D
1326 developments will be necessary, combining orthogonal retention mechanisms to connect with
1327 specific rDOM properties, to increase separation power before analysis using hyphenated
1328 techniques, to improve the quantitative response, to decrease the matrix effect, and to study
1329 various interactions (steric, hydrophobic, electrostatic). The "omics" studies modeled on
1330 humeomics, developed for the study of NOM, have not yet been developed for the
1331 characterization of recalcitrant compounds. While there are constant structures in HSs that
1332 provide validation of the methods developed, there is still no consensus around recalcitrant
1333 compounds and melanoidins from sludge and wastewater. It is essential to improve our
1334 structural knowledge of these compounds, as currently developed for the food industry, in
1335 order to propose new developments based on the study of reaction mechanisms for the
1336 synthesis and degradation of all rDOM present in wastewater. This new development will
1337 include advances in proteomics, applied to the study of rDOM, the aim being to understand
1338 the mechanisms related to refractory properties.

1339 For the structural determination of recalcitrant compounds, NMR could be considered
1340 as a method of choice. Nevertheless, the potential of NMR approaches in the area of this
1341 review has yet to be fully tapped. Indeed, the intrinsic insensitivity of NMR is not a limitation
1342 in all environmental applications, while its inherently quantitative nature and its holistic
1343 approach are its main strengths. Moreover, current and future NMR developments may
1344 enhance the technique's potential. Dissolution dynamic nuclear polarization is a promising
1345 approach, a versatile method for enhancing the sensitivity of NMR by boosting signal
1346 intensities by four to five orders of magnitude [218]. All of the techniques described in this

1347 review can be combined, leading to huge amounts of data to be processed and analyzed. The
1348 implementation of chemometric tools, such as CA, PCA and PLS multivariate approaches are
1349 essential for comprehensive data analysis. Finally, by combining all the developments
1350 described in each analytical field, linking a property to a set of physicochemical descriptors or
1351 to specific structural data, future developments will open up a new field in the mechanistic
1352 study of rDOM for theoretical knowledge and water treatment industry applications.

1353

1354 **Acknowledgements**

1355 The authors would like to thank the SUEZ group for its scientific contribution and financial
1356 support for this study.

1357

1358

1359 **References**

- 1360 [1] European commission, Ninth Report on the implementation status and the programmes
1361 for implementation (as required by Article 17) of Council Directive 91/271/EEC
1362 concerning urban waste water treatment, (2013).
- 1363 [2] W. Guo, J. Xu, J. Wang, Y. Wen, J. Zhuo, Y. Yan, Characterization of dissolved
1364 organic matter in urban sewage using excitation emission matrix fluorescence
1365 spectroscopy and parallel factor analysis, *Journal of Environmental Sciences*. 22 (2010)
1366 1728–1734. [https://doi.org/10.1016/S1001-0742\(09\)60312-0](https://doi.org/10.1016/S1001-0742(09)60312-0).
- 1367 [3] S. Lehtoranta, R. Vilpas, T.J. Mattila, Comparison of carbon footprints and
1368 eutrophication impacts of rural on-site wastewater treatment plants in Finland, *Journal*
1369 *of Cleaner Production*. 65 (2014) 439–446.
1370 <https://doi.org/10.1016/j.jclepro.2013.08.024>.

- 1371 [4] Z. Wang, W. Liu, N. Zhao, H. Li, Y. Zhang, W. Si-Ma, J. Liu, Composition analysis of
1372 colored dissolved organic matter in Taihu Lake based on three dimension excitation-
1373 emission fluorescence matrix and PARAFAC model, and the potential application in
1374 water quality monitoring, *Journal of Environmental Sciences*. 19 (2007) 787–791.
1375 [https://doi.org/10.1016/S1001-0742\(07\)60132-6](https://doi.org/10.1016/S1001-0742(07)60132-6).
- 1376 [5] P. Rajasulochana, V. Preethy, Comparison on efficiency of various techniques in
1377 treatment of waste and sewage water – A comprehensive review, *Resource-Efficient
1378 Technologies*. 2 (2016) 175–184. <https://doi.org/10.1016/j.reffit.2016.09.004>.
- 1379 [6] V.K. Tyagi, A.A. Khan, N.W. Jern, A. Khursheed, A.A. Kazmi, eds., Identification of
1380 recalcitrant compounds after anaerobic digestion with various sludge pretreatment
1381 methods, in: *Post Treatments of Anaerobically Treated Effluents*, IWA Publishing,
1382 2019: pp. 283–302. https://doi.org/10.2166/9781780409740_0283.
- 1383 [7] G. Zhen, X. Lu, H. Kato, Y. Zhao, Y.-Y. Li, Overview of pretreatment strategies for
1384 enhancing sewage sludge disintegration and subsequent anaerobic digestion: Current
1385 advances, full-scale application and future perspectives, *Renewable and Sustainable
1386 Energy Reviews*. 69 (2017) 559–577. <https://doi.org/10.1016/j.rser.2016.11.187>.
- 1387 [8] W.P.F. Barber, Thermal hydrolysis for sewage treatment: A critical review, *Water
1388 Research*. 104 (2016) 53–71. <https://doi.org/10.1016/j.watres.2016.07.069>.
- 1389 [9] M. Escala, T. Zumbühl, Ch. Koller, R. Junge, R. Krebs, Hydrothermal Carbonization as
1390 an Energy-Efficient Alternative to Established Drying Technologies for Sewage
1391 Sludge: A Feasibility Study on a Laboratory Scale, *Energy Fuels*. 27 (2013) 454–460.
1392 <https://doi.org/10.1021/ef3015266>.
- 1393 [10] J. Ariunbaatar, A. Panico, G. Esposito, F. Pirozzi, P.N.L. Lens, Pretreatment methods
1394 to enhance anaerobic digestion of organic solid waste, *Applied Energy*. 123 (2014)
1395 143–156. <https://doi.org/10.1016/j.apenergy.2014.02.035>.

- 1396 [11] D. Kim, K. Lee, K.Y. Park, Hydrothermal carbonization of anaerobically digested
1397 sludge for solid fuel production and energy recovery, *Fuel*. 130 (2014) 120–125.
1398 <https://doi.org/10.1016/j.fuel.2014.04.030>.
- 1399 [12] L. Wang, Y. Chang, A. Li, Hydrothermal carbonization for energy-efficient processing
1400 of sewage sludge: A review, *Renewable and Sustainable Energy Reviews*. 108 (2019)
1401 423–440. <https://doi.org/10.1016/j.rser.2019.04.011>.
- 1402 [13] P. Devos, M. Haddad, H. Carrère, Thermal Hydrolysis of Municipal sludge: Finding
1403 the Temperature Sweet Spot: A Review, *Waste Biomass Valor.* (2020).
1404 <https://doi.org/10.1007/s12649-020-01130-1>.
- 1405 [14] D.J. Barker, D.C. Stuckey, A review of soluble microbial products (SMP) in
1406 wastewater treatment systems, *Water Research*. 33 (1999) 3063–3082.
1407 [https://doi.org/10.1016/S0043-1354\(99\)00022-6](https://doi.org/10.1016/S0043-1354(99)00022-6).
- 1408 [15] Ph. Namour, M.C. Müller, Fractionation of organic matter from wastewater treatment
1409 plants before and after a 21-day biodegradability test: A physical-chemical method for
1410 measurement of the refractory part of effluents, *Water Research*. 32 (1998) 2224–2231.
1411 [https://doi.org/10.1016/S0043-1354\(97\)00428-4](https://doi.org/10.1016/S0043-1354(97)00428-4).
- 1412 [16] E.M. Carstea, J. Bridgeman, A. Baker, D.M. Reynolds, Fluorescence spectroscopy for
1413 wastewater monitoring: A review, *Water Research*. 95 (2016) 205–219.
1414 <https://doi.org/10.1016/j.watres.2016.03.021>.
- 1415 [17] M.-F. Dignac, P. Ginestet, D. Rybacki, A. Bruchet, V. Urbain, P. Scribe, Fate of
1416 wastewater organic pollution during activated sludge treatment: nature of residual
1417 organic matter, *Water Research*. 34 (2000) 4185–4194. [https://doi.org/10.1016/S0043-](https://doi.org/10.1016/S0043-1354(00)00195-0)
1418 [1354\(00\)00195-0](https://doi.org/10.1016/S0043-1354(00)00195-0).

- 1419 [18] V. Réveillé, L. Mansuy, É. Jardé, É. Garnier-Sillam, Characterisation of sewage
1420 sludge-derived organic matter: lipids and humic acids, *Organic Geochemistry*. 34
1421 (2003) 615–627. [https://doi.org/10.1016/S0146-6380\(02\)00216-4](https://doi.org/10.1016/S0146-6380(02)00216-4).
- 1422 [19] J. Chudoba, Quantitative estimation in cod units of refractory organic compounds
1423 produced by activated sludge microorganisms, *Water Research*. 19 (1985) 37–43.
1424 [https://doi.org/10.1016/0043-1354\(85\)90320-3](https://doi.org/10.1016/0043-1354(85)90320-3).
- 1425 [20] S. Azman, A.F. Khadem, J.B. van Lier, G. Zeeman, C.M. Plugge, Presence and Role of
1426 Anaerobic Hydrolytic Microbes in Conversion of Lignocellulosic Biomass for Biogas
1427 Production, *Critical Reviews in Environmental Science and Technology*. 45 (2015)
1428 2523–2564. <https://doi.org/10.1080/10643389.2015.1053727>.
- 1429 [21] J. Dwyer, D. Starrenburg, S. Tait, K. Barr, D.J. Batstone, P. Lant, Decreasing activated
1430 sludge thermal hydrolysis temperature reduces product colour, without decreasing
1431 degradability, *Water Research*. 42 (2008) 4699–4709.
1432 <https://doi.org/10.1016/j.watres.2008.08.019>.
- 1433 [22] A. Gupta, J.T. Novak, R. Zhao, Characterization of organic matter in the thermal
1434 hydrolysis pretreated anaerobic digestion return liquor, *Journal of Environmental
1435 Chemical Engineering*. 3 (2015) 2631–2636.
1436 <https://doi.org/10.1016/j.jece.2015.07.029>.
- 1437 [23] S. Hao, S. Ren, N. Zhou, H. Chen, M. Usman, C. He, Q. Shi, G. Luo, S. Zhang,
1438 Molecular composition of hydrothermal liquefaction wastewater from sewage sludge
1439 and its transformation during anaerobic digestion, *Journal of Hazardous Materials*. 383
1440 (2020) 121163. <https://doi.org/10.1016/j.jhazmat.2019.121163>.
- 1441 [24] V. Toutian, M. Barjenbruch, T. Unger, C. Loderer, C. Remy, Effect of temperature on
1442 biogas yield increase and formation of refractory COD during thermal hydrolysis of

- 1443 waste activated sludge, *Water Research*. (2019) 115383.
1444 <https://doi.org/10.1016/j.watres.2019.115383>.
- 1445 [25] M. Usman, S. Hao, H. Chen, S. Ren, D.C.W. Tsang, S. O-Thong, G. Luo, S. Zhang,
1446 Molecular and microbial insights towards understanding the anaerobic digestion of the
1447 wastewater from hydrothermal liquefaction of sewage sludge facilitated by granular
1448 activated carbon (GAC), *Environment International*. 133 (2019) 105257.
1449 <https://doi.org/10.1016/j.envint.2019.105257>.
- 1450 [26] D. McCarty, J. Gosset, P. McCarty, Effect of Thermal Pretreatment on Digestibility
1451 and Dewaterability of Organic Sludges, (1978) 14.
- 1452 [27] C.A. Wilson, C.T. Tanneru, S. Banjade, S.N. Murthy, J.T. Novak, Anaerobic Digestion
1453 of Raw and Thermally Hydrolyzed Wastewater Solids Under Various Operational
1454 Conditions, *Water Environment Research*. 83 (2011) 815–825.
1455 <https://doi.org/10.2175/106143011X12928814444934>.
- 1456 [28] M.J. Higgins, S. Beightol, U. Mandahar, R. Suzuki, S. Xiao, H.-W. Lu, T. Le, J. Mah,
1457 B. Pathak, H. DeClippeleir, J.T. Novak, A. Al-Omari, S.N. Murthy, Pretreatment of a
1458 primary and secondary sludge blend at different thermal hydrolysis temperatures:
1459 Impacts on anaerobic digestion, dewatering and filtrate characteristics, *Water Research*.
1460 122 (2017) 557–569. <https://doi.org/10.1016/j.watres.2017.06.016>.
- 1461 [29] M.N. Jones, N.D. Bryan, Colloidal properties of humic substances, *Advances in*
1462 *Colloid and Interface Science*. 78 (1998) 1–48. [https://doi.org/10.1016/S0001-](https://doi.org/10.1016/S0001-8686(98)00058-X)
1463 [8686\(98\)00058-X](https://doi.org/10.1016/S0001-8686(98)00058-X).
- 1464 [30] C.S. Dzah, Y. Duan, H. Zhang, N.A. Serwah Boateng, H. Ma, Latest developments in
1465 polyphenol recovery and purification from plant by-products: A review, *Trends in Food*
1466 *Science & Technology*. 99 (2020) 375–388. <https://doi.org/10.1016/j.tifs.2020.03.003>.

- 1467 [31] B. Cämmerer, V. Jalyschkov, L.W. Kroh, Carbohydrate structures as part of the
1468 melanoidin skeleton, *International Congress Series*. 1245 (2002) 269–273.
1469 [https://doi.org/10.1016/S0531-5131\(02\)00890-7](https://doi.org/10.1016/S0531-5131(02)00890-7).
- 1470 [32] H.-Y. Wang, H. Qian, W.-R. Yao, Melanoidins produced by the Maillard reaction:
1471 Structure and biological activity, *Food Chemistry*. 128 (2011) 573–584.
1472 <https://doi.org/10.1016/j.foodchem.2011.03.075>.
- 1473 [33] D.C. Stuckey, P.L. McCarty, The effect of thermal pretreatment on the anaerobic
1474 biodegradability and toxicity of waste activated sludge, *Water Research*. 18 (1984)
1475 1343–1353. [https://doi.org/10.1016/0043-1354\(84\)90002-2](https://doi.org/10.1016/0043-1354(84)90002-2).
- 1476 [34] H. Chen, Y. Rao, L. Cao, Y. Shi, S. Hao, G. Luo, S. Zhang, Hydrothermal conversion
1477 of sewage sludge: Focusing on the characterization of liquid products and their
1478 methane yields, *Chemical Engineering Journal*. 357 (2019) 367–375.
1479 <https://doi.org/10.1016/j.cej.2018.09.180>.
- 1480 [35] J. Stemann, A. Putschew, F. Ziegler, Hydrothermal carbonization: Process water
1481 characterization and effects of water recirculation, *Bioresource Technology*. 143 (2013)
1482 139–146. <https://doi.org/10.1016/j.biortech.2013.05.098>.
- 1483 [36] J.A. Villamil, A.F. Mohedano, J.J. Rodriguez, M.A. de la Rubia, Valorisation of the
1484 liquid fraction from hydrothermal carbonisation of sewage sludge by anaerobic
1485 digestion: Anaerobic digestion of the liquid fraction from HTC of sewage sludge, *J.*
1486 *Chem. Technol. Biotechnol.* 93 (2018) 450–456. <https://doi.org/10.1002/jctb.5375>.
- 1487 [37] A. Kaushik, S. Basu, V.S. Batra, M. Balakrishnan, Fractionation of sugarcane molasses
1488 distillery wastewater and evaluation of antioxidant and antimicrobial characteristics,
1489 *Industrial Crops and Products*. 118 (2018) 73–80.
1490 <https://doi.org/10.1016/j.indcrop.2018.03.040>.

- 1491 [38] D. Pant, A. Adholeya, Biological approaches for treatment of distillery wastewater: A
1492 review, *Bioresource Technology*. 98 (2007) 2321–2334.
1493 <https://doi.org/10.1016/j.biortech.2006.09.027>.
- 1494 [39] L. Yang, B. Si, M.A. Martins, J. Watson, H. Chu, Y. Zhang, X. Tan, X. Zhou, Y.
1495 Zhang, Improve the biodegradability of post-hydrothermal liquefaction wastewater
1496 with ozone: conversion of phenols and N-heterocyclic compounds, *Water Sci Technol*.
1497 2017 (2018) 248–255. <https://doi.org/10.2166/wst.2018.108>.
- 1498 [40] K. Brezinski, B. Gorczyca, Multi-spectral characterization of natural organic matter
1499 (NOM) from Manitoba surface waters using high performance size exclusion
1500 chromatography (HPSEC), *Chemosphere*. 225 (2019) 53–64.
1501 <https://doi.org/10.1016/j.chemosphere.2019.02.176>.
- 1502 [41] C. Wei, H. Wu, Q. Kong, J. Wei, C. Feng, G. Qiu, C. Wei, F. Li, Residual chemical
1503 oxygen demand (COD) fractionation in bio-treated coking wastewater integrating
1504 solution property characterization, *Journal of Environmental Management*. 246 (2019)
1505 324–333. <https://doi.org/10.1016/j.jenvman.2019.06.001>.
- 1506 [42] A. Bhatnagar, M. Sillanpää, Removal of natural organic matter (NOM) and its
1507 constituents from water by adsorption – A review, *Chemosphere*. 166 (2017) 497–510.
1508 <https://doi.org/10.1016/j.chemosphere.2016.09.098>.
- 1509 [43] F.M. da Costa, S.D.A. Daflon, D.M. Bila, F.V. da Fonseca, J.C. Campos, Evaluation of
1510 the biodegradability and toxicity of landfill leachates after pretreatment using advanced
1511 oxidative processes, *Waste Management*. 76 (2018) 606–613.
1512 <https://doi.org/10.1016/j.wasman.2018.02.030>.
- 1513 [44] G.D. Gebreeyessus, A. Mekonnen, E. Alemayehu, A review on progresses and
1514 performances in distillery stillage management, *Journal of Cleaner Production*. 232
1515 (2019) 295–307. <https://doi.org/10.1016/j.jclepro.2019.05.383>.

- 1516 [45] D. Zhang, Y. Feng, H. Huang, W. Khunjar, Z.-W. Wang, Recalcitrant dissolved
1517 organic nitrogen formation in thermal hydrolysis pretreatment of municipal sludge,
1518 Environment International. 138 (2020) 105629.
1519 <https://doi.org/10.1016/j.envint.2020.105629>.
- 1520 [46] P. Kumar, R. Chandra, Decolourisation and detoxification of synthetic molasses
1521 melanoidins by individual and mixed cultures of *Bacillus* spp., Bioresource
1522 Technology. 97 (2006) 2096–2102. <https://doi.org/10.1016/j.biortech.2005.10.012>.
- 1523 [47] M. Pham, L. Schideman, J. Scott, N. Rajagopalan, M.J. Plewa, Chemical and
1524 Biological Characterization of Wastewater Generated from Hydrothermal Liquefaction
1525 of *Spirulina*, Environ. Sci. Technol. 47 (2013) 2131–2138.
1526 <https://doi.org/10.1021/es304532c>.
- 1527 [48] R. Chandra, R.N. Bharagava, V. Rai, Melanoidins as major colourant in sugarcane
1528 molasses based distillery effluent and its degradation, Bioresource Technology. 99
1529 (2008) 4648–4660. <https://doi.org/10.1016/j.biortech.2007.09.057>.
- 1530 [49] F. Adani, M. Spagnol, P. Genevini, Biochemical Origin and Refractory Properties of
1531 Humic Acid Extracted from the Maize Plant, Biogeochemistry. 78 (2006) 85–96.
1532 <https://doi.org/10.1007/s10533-005-2902-7>.
- 1533 [50] W. Machado, J.C. Franchini, M. de Fátima Guimarães, J.T. Filho, Spectroscopic
1534 characterization of humic and fulvic acids in soil aggregates, Brazil, Heliyon. 6 (2020)
1535 e04078. <https://doi.org/10.1016/j.heliyon.2020.e04078>.
- 1536 [51] S.A. Huber, A. Balz, M. Abert, W. Pronk, Characterisation of aquatic humic and non-
1537 humic matter with size-exclusion chromatography – organic carbon detection – organic
1538 nitrogen detection (LC-OCD-OND), Water Research. 45 (2011) 879–885.
1539 <https://doi.org/10.1016/j.watres.2010.09.023>.

- 1540 [52] V.A. Yaylayan, E. Kaminsky, Isolation and structural analysis of maillard polymers:
1541 caramel and melanoidin formation in glycine/glucose model system, *Food Chemistry*.
1542 63 (1998) 25–31. [https://doi.org/10.1016/S0308-8146\(97\)00237-9](https://doi.org/10.1016/S0308-8146(97)00237-9).
- 1543 [53] C. Bougrier, J.P. Delgenès, H. Carrère, Effects of thermal treatments on five different
1544 waste activated sludge samples solubilisation, physical properties and anaerobic
1545 digestion, *Chemical Engineering Journal*. 139 (2008) 236–244.
1546 <https://doi.org/10.1016/j.cej.2007.07.099>.
- 1547 [54] C.A. Wilson, J.T. Novak, Hydrolysis of macromolecular components of primary and
1548 secondary wastewater sludge by thermal hydrolytic pretreatment, *Water Research*. 43
1549 (2009) 4489–4498. <https://doi.org/10.1016/j.watres.2009.07.022>.
- 1550 [55] D. Fytili, A. Zabaniotou, Utilization of sewage sludge in EU application of old and new
1551 methods—A review, *Renewable and Sustainable Energy Reviews*. 12 (2008) 116–140.
1552 <https://doi.org/10.1016/j.rser.2006.05.014>.
- 1553 [56] J. Jimenez, F. Vedrenne, C. Denis, A. Mottet, S. Déléris, J.-P. Steyer, J.A. Cacho
1554 Rivero, A statistical comparison of protein and carbohydrate characterisation
1555 methodology applied on sewage sludge samples, *Water Research*. 47 (2013) 1751–
1556 1762. <https://doi.org/10.1016/j.watres.2012.11.052>.
- 1557 [57] L. Mikkelsen, Physico-chemical characteristics of full scale sewage sludges with
1558 implications to dewatering, *Water Research*. 36 (2002) 2451–2462.
1559 [https://doi.org/10.1016/S0043-1354\(01\)00477-8](https://doi.org/10.1016/S0043-1354(01)00477-8).
- 1560 [58] L.C. Maillard, M.A. Gautier, Action des acides aminés sur les sucres ; formation des
1561 mélanoïdines par voie méthodique. *Compt. Rend. Seances. Acad. Sci.* 1912, 154, 66–
1562 68.
- 1563 [59] H. Cui, J. Yu, S. Xia, E. Duhoranimana, Q. Huang, X. Zhang, Improved controlled
1564 flavor formation during heat-treatment with a stable Maillard reaction intermediate

1565 derived from xylose-phenylalanine, *Food Chemistry*. 271 (2019) 47–53.
1566 <https://doi.org/10.1016/j.foodchem.2018.07.161>.

1567 [60] M.A.J.S. van Boekel, Formation of flavour compounds in the Maillard reaction,
1568 *Biotechnology Advances*. 24 (2006) 230–233.
1569 <https://doi.org/10.1016/j.biotechadv.2005.11.004>.

1570 [61] S.I.F.S. Martins, W.M.F. Jongen, M.A.J.S. van Boekel, A review of Maillard reaction
1571 in food and implications to kinetic modelling, *Trends in Food Science & Technology*.
1572 11 (2000) 364–373. [https://doi.org/10.1016/S0924-2244\(01\)00022-X](https://doi.org/10.1016/S0924-2244(01)00022-X).

1573 [62] G.P. Rizzi, Chemical structure of colored maillard reaction products, *Food Reviews*
1574 *International*. 13 (1997) 1–28. <https://doi.org/10.1080/87559129709541096>.

1575 [63] G.L.L. Scalone, P. Lamichhane, T. Cucu, N. De Kimpe, B. De Meulenaer, Impact of
1576 different enzymatic hydrolysates of whey protein on the formation of pyrazines in
1577 Maillard model systems, *Food Chemistry*. 278 (2019) 533–544.
1578 <https://doi.org/10.1016/j.foodchem.2018.11.088>.

1579 [64] J.M. Silván, J. van de Lagemaat, A. Olano, M.D. del Castillo, Analysis and biological
1580 properties of amino acid derivatives formed by Maillard reaction in foods, *Journal of*
1581 *Pharmaceutical and Biomedical Analysis*. 41 (2006) 1543–1551.
1582 <https://doi.org/10.1016/j.jpba.2006.04.004>.

1583 [65] J.E. Hodge, Dehydrated Foods, *Chemistry of Browning Reactions in Model Systems*, J.
1584 *Agric. Food Chem.* 1 (1953) 928–943. <https://doi.org/10.1021/jf60015a004>.

1585 [66] C. Bougrier, J.P. Delgenès, H. Carrère, Impacts of thermal pre-treatments on the semi-
1586 continuous anaerobic digestion of waste activated sludge, *Biochemical Engineering*
1587 *Journal*. 34 (2007) 20–27. <https://doi.org/10.1016/j.bej.2006.11.013>.

- 1588 [67] N. Miyata, M. Ike, K. Furukawa, M. Fujita, Fractionation and characterization of
1589 brown colored components in heat treatment liquor of waste sludge, *Water Research*.
1590 30 (1996) 1361–1368. [https://doi.org/10.1016/0043-1354\(95\)00315-0](https://doi.org/10.1016/0043-1354(95)00315-0).
- 1591 [68] C. Helou, P. Jacolot, C. Niquet-Léridon, P. Gadonna-Widehem, F.J. Tessier, Maillard
1592 reaction products in bread: A novel semi-quantitative method for evaluating
1593 melanoidins in bread, *Food Chemistry*. 190 (2016) 904–911.
1594 <https://doi.org/10.1016/j.foodchem.2015.06.032>.
- 1595 [69] B. Ripper, C.R. Kaiser, D. Perrone, Use of NMR techniques to investigate the changes
1596 on the chemical composition of coffee melanoidins, *Journal of Food Composition and*
1597 *Analysis*. 87 (2020) 103399. <https://doi.org/10.1016/j.jfca.2019.103399>.
- 1598 [70] R.M.W. Amon, R. Benner, Bacterial utilization of different size classes of dissolved
1599 organic matter, *Limnol. Oceanogr.* 41 (1996) 41–51.
1600 <https://doi.org/10.4319/lo.1996.41.1.0041>.
- 1601 [71] B. Čosović, V. Vojvodić, N. Bošković, M. Plavšić, C. Lee, Characterization of natural
1602 and synthetic humic substances (melanoidins) by chemical composition and adsorption
1603 measurements, *Organic Geochemistry*. 41 (2010) 200–205.
1604 <https://doi.org/10.1016/j.orggeochem.2009.10.002>.
- 1605 [72] D. Lu, F. Sun, Y. Zhou, Insights into anaerobic transformation of key dissolved organic
1606 matters produced by thermal hydrolysis sludge pretreatment, *Bioresource Technology*.
1607 266 (2018) 60–67. <https://doi.org/10.1016/j.biortech.2018.06.059>.
- 1608 [73] M.A. De la Rubia, J.A. Villamil, J.J. Rodriguez, R. Borja, A.F. Mohedano, Mesophilic
1609 anaerobic co-digestion of the organic fraction of municipal solid waste with the liquid
1610 fraction from hydrothermal carbonization of sewage sludge, *Waste Management*. 76
1611 (2018) 315–322. <https://doi.org/10.1016/j.wasman.2018.02.046>.

- 1612 [74] H.-C. Kim, M.-J. Yu, Characterization of aquatic humic substances to DBPs formation
1613 in advanced treatment processes for conventionally treated water, *Journal of Hazardous*
1614 *Materials*. 143 (2007) 486–493. <https://doi.org/10.1016/j.jhazmat.2006.09.063>.
- 1615 [75] M. Kawai, N. Nagao, N. Kawasaki, A. Imai, T. Toda, Improvement of COD removal
1616 by controlling the substrate degradability during the anaerobic digestion of recalcitrant
1617 wastewater, *Journal of Environmental Management*. 181 (2016) 838–846.
1618 <https://doi.org/10.1016/j.jenvman.2016.06.057>.
- 1619 [76] K.S. Baker, R.C. Smith, Bio-optical classification and model of natural waters. 21: Bio-
1620 optical model 2, *Limnol. Oceanogr.* 27 (1982) 500–509.
1621 <https://doi.org/10.4319/lo.1982.27.3.0500>.
- 1622 [77] J.-P. Croué, J.A. Leenher, U.S. GEOLOGICAL SURVEY, (2003) 9.
- 1623 [78] J.L. Weishaar, G.R. Aiken, B.A. Bergamaschi, M.S. Fram, R. Fujii, K. Mopper,
1624 Evaluation of Specific Ultraviolet Absorbance as an Indicator of the Chemical
1625 Composition and Reactivity of Dissolved Organic Carbon, *Environ. Sci. Technol.* 37
1626 (2003) 4702–4708. <https://doi.org/10.1021/es030360x>.
- 1627 [79] N. Her, G. Amy, J. Sohn, U. Gunten, UV absorbance ratio index with size exclusion
1628 chromatography (URI-SEC) as an NOM property indicator, *Journal of Water Supply:*
1629 *Research and Technology-Aqua*. 57 (2008) 35–44.
1630 <https://doi.org/10.2166/aqua.2008.029>.
- 1631 [80] M.B. Ahmed, Md.A.H. Jahir, C. Khoureshed, J.L. Zhou, H.H. Ngo, D.L. Nghiem, M.
1632 Moni, L. Sun, Sorptive removal of dissolved organic matter in biologically-treated
1633 effluent by functionalized biochar and carbon nanotubes: Importance of sorbent
1634 functionality, *Bioresource Technology*. 269 (2018) 9–17.
1635 <https://doi.org/10.1016/j.biortech.2018.08.046>.

- 1636 [81] J.C.M. de Wit, W.H. van Riemsdijk, L.K. Koopal, Proton binding to humic substances.
1637 1. Electrostatic effects, *Environ. Sci. Technol.* 27 (1993) 2005–2014.
1638 <https://doi.org/10.1021/es00047a004>.
- 1639 [82] R.A. Dobbs, R.H. Wise, R.B. Dean, The use of ultra-violet absorbance for monitoring
1640 the total organic carbon content of water and wastewater, *Water Research.* 6 (1972)
1641 1173–1180. [https://doi.org/10.1016/0043-1354\(72\)90017-6](https://doi.org/10.1016/0043-1354(72)90017-6).
- 1642 [83] V. Penaud, J.-P. Delgenès, R. Moletta, Characterization of Soluble Molecules from
1643 Thermochemically Pretreated Sludge, *J. Environ. Eng.* 126 (2000) 397–402.
1644 [https://doi.org/10.1061/\(ASCE\)0733-9372\(2000\)126:5\(397\)](https://doi.org/10.1061/(ASCE)0733-9372(2000)126:5(397)).
- 1645 [84] X. Liu, W. Wang, X. Gao, Y. Zhou, R. Shen, Effect of thermal pretreatment on the
1646 physical and chemical properties of municipal biomass waste, *Waste Management.* 32
1647 (2012) 249–255. <https://doi.org/10.1016/j.wasman.2011.09.027>.
- 1648 [85] J.K. Edzwald, J.E. Tobiason, Enhanced coagulation: US requirements and a broader
1649 view, (n.d.) 8.
- 1650 [86] F.J. Rodríguez-Vidal, M. García-Valverde, B. Ortega-Azabache, Á. González-
1651 Martínez, A. Bellido-Fernández, Characterization of urban and industrial wastewaters
1652 using excitation-emission matrix (EEM) fluorescence: Searching for specific
1653 fingerprints, *Journal of Environmental Management.* 263 (2020) 110396.
1654 <https://doi.org/10.1016/j.jenvman.2020.110396>.
- 1655 [87] R.K. Henderson, A. Baker, K.R. Murphy, A. Hambly, R.M. Stuetz, S.J. Khan,
1656 Fluorescence as a potential monitoring tool for recycled water systems: A review,
1657 *Water Research.* 43 (2009) 863–881. <https://doi.org/10.1016/j.watres.2008.11.027>.
- 1658 [88] D.M. McKnight, E.W. Boyer, P.K. Westerhoff, P.T. Doran, T. Kulbe, D.T. Andersen,
1659 Spectrofluorometric characterization of dissolved organic matter for indication of

1660 precursor organic material and aromaticity, *Limnol. Oceanogr.* 46 (2001) 38–48.
1661 <https://doi.org/10.4319/lo.2001.46.1.0038>.

1662 [89] A. Zsolnay, E. Baigar, M. Jimenez, B. Steinweg, F. Saccomandi, Differentiating with
1663 fluorescence spectroscopy the sources of dissolved organic matter in soils subjected to
1664 drying, *Chemosphere.* 38 (1999) 45–50. [https://doi.org/10.1016/S0045-6535\(98\)00166-](https://doi.org/10.1016/S0045-6535(98)00166-0)
1665 0.

1666 [90] J. Jimenez, Q. Aemig, N. Doussiet, J.-P. Steyer, S. Houot, D. Patureau, A new organic
1667 matter fractionation methodology for organic wastes: Bioaccessibility and complexity
1668 characterization for treatment optimization, *Bioresource Technology.* 194 (2015) 344–
1669 353. <https://doi.org/10.1016/j.biortech.2015.07.037>.

1670 [91] J. Jimenez, E. Gonidec, J.A. Cacho Rivero, E. Latrille, F. Vedrenne, J.-P. Steyer,
1671 Prediction of anaerobic biodegradability and bioaccessibility of municipal sludge by
1672 coupling sequential extractions with fluorescence spectroscopy: Towards ADM1
1673 variables characterization, *Water Research.* 50 (2014) 359–372.
1674 <https://doi.org/10.1016/j.watres.2013.10.048>.

1675 [92] Z.-X. Xu, H. Song, P.-J. Li, Z.-X. He, Q. Wang, K. Wang, P.-G. Duan, Hydrothermal
1676 carbonization of sewage sludge: Effect of aqueous phase recycling, *Chemical*
1677 *Engineering Journal.* (2019) 123410. <https://doi.org/10.1016/j.cej.2019.123410>.

1678 [93] J. Jimenez, Organic matter characterization with 3D fluorescence spectroscopy for
1679 anaerobic digestion modeling of wastewater treatment sludge, UNIVERSITE
1680 MONTPELLIER II SCIENCES ET TECHNIQUES DU LANGUEDOC, 2012.

1681 [94] D. Patra, A.K. Mishra, Recent developments in multi-component synchronous
1682 fluorescence scan analysis, *TrAC Trends in Analytical Chemistry.* 21 (2002) 787–798.
1683 [https://doi.org/10.1016/S0165-9936\(02\)01201-3](https://doi.org/10.1016/S0165-9936(02)01201-3).

- 1684 [95] J.R. Lakowicz, ed., Principles of Fluorescence Spectroscopy, Springer US, Boston,
1685 MA, 2006. <https://doi.org/10.1007/978-0-387-46312-4>.
- 1686 [96] Y. Pan, H. Li, X. Zhang, A. Li, Characterization of natural organic matter in drinking
1687 water: Sample preparation and analytical approaches, Trends in Environmental
1688 Analytical Chemistry. 12 (2016) 23–30. <https://doi.org/10.1016/j.teac.2016.11.002>.
- 1689 [97] M. Kowalski, K. Kowalska, J. Wiszniowski, J. Turek-Szytow, Qualitative analysis of
1690 activated sludge using FT-IR technique, Chem. Pap. 72 (2018) 2699–2706.
1691 <https://doi.org/10.1007/s11696-018-0514-7>.
- 1692 [98] E. Pretsch, P. Bühlmann, M. Badertscher, Structure Determination of Organic
1693 Compounds, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.
1694 <https://doi.org/10.1007/978-3-540-93810-1>.
- 1695 [99] R.L. Olsen, R.W. Chappell, J.C. Loftis, Water quality sample collection, data treatment
1696 and results presentation for principal components analysis – literature review and
1697 Illinois River watershed case study, Water Research. 46 (2012) 3110–3122.
1698 <https://doi.org/10.1016/j.watres.2012.03.028>.
- 1699 [100] A. Imai, T. Fukushima, K. Matsushige, Y.-H. Kim, K. Choi, Characterization of
1700 dissolved organic matter in effluents from wastewater treatment plants, Water
1701 Research. 36 (2002) 859–870. [https://doi.org/10.1016/S0043-1354\(01\)00283-4](https://doi.org/10.1016/S0043-1354(01)00283-4).
- 1702 [101] G.R. Aiken, E.M. Thurman, R.L. Malcolm, H.F. Walton, Comparison of XAD
1703 macroporous resins for the concentration of fulvic acid from aqueous solution, Anal.
1704 Chem. 51 (1979) 1799–1803. <https://doi.org/10.1021/ac50047a044>.
- 1705 [102] R. Lamsal, K.R. Montreuil, F.C. Kent, M.E. Walsh, G.A. Gagnon, Characterization and
1706 removal of natural organic matter by an integrated membrane system, Desalination. 303
1707 (2012) 12–16. <https://doi.org/10.1016/j.desal.2012.06.025>.

- 1708 [103] J. Zumstein, J. Buffle, Circulation of pedogenic and aquagenic organic matter in an
1709 eutrophic lake, *Water Research*. 23 (1989) 229–239. <https://doi.org/10.1016/0043->
1710 1354(89)90047-X.
- 1711 [104] A. Nebbioso, A. Piccolo, Advances in humeomics: Enhanced structural identification
1712 of humic molecules after size fractionation of a soil humic acid, *Analytica Chimica*
1713 *Acta*. 720 (2012) 77–90. <https://doi.org/10.1016/j.aca.2012.01.027>.
- 1714 [105] A.M. Posner, Importance of Electrolyte in the Determination of Molecular Weights by
1715 ‘Sephadex’ Gel Filtration, with Especial Reference to Humic Acid, *Nature*. 198 (1963)
1716 1161–1163. <https://doi.org/10.1038/1981161a0>.
- 1717 [106] C.W.K. Chow, R. Fabris, J. van Leeuwen, D. Wang, M. Drikas, Assessing Natural
1718 Organic Matter Treatability Using High Performance Size Exclusion Chromatography,
1719 *Environ. Sci. Technol.* 42 (2008) 6683–6689. <https://doi.org/10.1021/es800794r>.
- 1720 [107] M. Krauss, P. Longrée, E. van Houtte, J. Cauwenberghs, J. Hollender, Assessing the
1721 Fate of Nitrosamine Precursors in Wastewater Treatment by Physicochemical
1722 Fractionation, *Environ. Sci. Technol.* 44 (2010) 7871–7877.
1723 <https://doi.org/10.1021/es101289z>.
- 1724 [108] F.L. Rosario-Ortiz, S. Snyder, I.H. (Mel) Suffet, Characterization of the Polarity of
1725 Natural Organic Matter under Ambient Conditions by the Polarity Rapid Assessment
1726 Method (PRAM), *Environ. Sci. Technol.* 41 (2007) 4895–4900.
1727 <https://doi.org/10.1021/es062151t>.
- 1728 [109] J. Buffle, P. Deladoey, W. Haerdi, The use of ultrafiltration for the separation and
1729 fractionation of organic ligands in fresh waters, *Analytica Chimica Acta*. 101 (1978)
1730 339–357. [https://doi.org/10.1016/S0003-2670\(01\)93369-0](https://doi.org/10.1016/S0003-2670(01)93369-0).

- 1731 [110] A.D. Revchuk, I.H. (Mel) Suffet, Ultrafiltration separation of aquatic natural organic
1732 matter: Chemical probes for quality assurance, *Water Research*. 43 (2009) 3685–3692.
1733 <https://doi.org/10.1016/j.watres.2009.05.029>.
- 1734 [111] Z. Zhou, L. Guo, A critical evaluation of an asymmetrical flow field-flow fractionation
1735 system for colloidal size characterization of natural organic matter, *Journal of*
1736 *Chromatography A*. 1399 (2015) 53–64. <https://doi.org/10.1016/j.chroma.2015.04.035>.
- 1737 [112] M.J. Pullin, S.E. Cabaniss, Physicochemical variations in DOM-synchronous
1738 fluorescence: Implications for mixing studies, *Limnol. Oceanogr.* 42 (1997) 1766–
1739 1773. <https://doi.org/10.4319/lo.1997.42.8.1766>.
- 1740 [113] Y. Pan, W. Li, H. An, H. Cui, Y. Wang, Formation and occurrence of new polar
1741 iodinated disinfection byproducts in drinking water, *Chemosphere*. 144 (2016) 2312–
1742 2320. <https://doi.org/10.1016/j.chemosphere.2015.11.012>.
- 1743 [114] P. Egeberg, J. Alberts, Determination of hydrophobicity of NOM by RP-HPLC, and the
1744 effect of pH and ionic strength, *Water Research*. 36 (2002) 4997–5004.
1745 [https://doi.org/10.1016/S0043-1354\(02\)00228-2](https://doi.org/10.1016/S0043-1354(02)00228-2).
- 1746 [115] A.C. Stenson, Reversed-Phase Chromatography Fractionation Tailored to Mass
1747 Spectral Characterization of Humic Substances, *Environ. Sci. Technol.* 42 (2008)
1748 2060–2065. <https://doi.org/10.1021/es7022412>.
- 1749 [116] T. Tanaka, S. Nagao, Y. Sakamoto, T. Ohnuki, S. Ni, M. Senoo, Distribution
1750 Coefficient in the Sorption of Radionuclides onto Ando Soil in the Presence of Humic
1751 Acid: Influence of the Molecular Size of Humic Acid, *Journal of Nuclear Science and*
1752 *Technology*. 34 (1997) 829–834. <https://doi.org/10.1080/18811248.1997.9733748>.
- 1753 [117] S.E. Cabaniss, Q. Zhou, P.A. Maurice, Y.-P. Chin, G.R. Aiken, A Log-Normal
1754 Distribution Model for the Molecular Weight of Aquatic Fulvic Acids, *Environ. Sci.*
1755 *Technol.* 34 (2000) 1103–1109. <https://doi.org/10.1021/es990555y>.

- 1756 [118] N. Her, G. Amy, J. Chung, J. Yoon, Y. Yoon, Characterizing dissolved organic matter
1757 and evaluating associated nanofiltration membrane fouling, *Chemosphere*. 70 (2008)
1758 495–502. <https://doi.org/10.1016/j.chemosphere.2007.06.025>.
- 1759 [119] E. Hoque, M. Wolf, G. Teichmann, E. Peller, W. Schimmack, G. Buckau, Influence of
1760 ionic strength and organic modifier concentrations on characterization of aquatic fulvic
1761 and humic acids by high-performance size-exclusion chromatography, *Journal of*
1762 *Chromatography A*. 1017 (2003) 97–105.
1763 <https://doi.org/10.1016/j.chroma.2003.08.038>.
- 1764 [120] S. Mori, H. Maréchal, H. Suzuki, Evaluation of “Solvent-Peak Separation” Column for
1765 the Determination of Polymer Molecular Mass Averages by SEC, *International Journal*
1766 *of Polymer Analysis and Characterization*. 4 (1997) 87–97.
1767 <https://doi.org/10.1080/10236669708033940>.
- 1768 [121] I.V. Perminova, F.H. Frimmel, D.V. Kovalevskii, G. Abbt-Braun, A.V. Kudryavtsev,
1769 S. Hesse, DEVELOPMENT OF A PREDICTIVE MODEL FOR CALCULATION OF
1770 MOLECULAR WEIGHT OF HUMIC SUBSTANCES, *Water Research*. 32 (1998)
1771 872–881. [https://doi.org/10.1016/S0043-1354\(97\)00283-2](https://doi.org/10.1016/S0043-1354(97)00283-2).
- 1772 [122] J. Peuravuori, K. Pihlaja, Molecular size distribution and spectroscopic properties of
1773 aquatic humic substances, *Analytica Chimica Acta*. 337 (1997) 133–149.
1774 [https://doi.org/10.1016/S0003-2670\(96\)00412-6](https://doi.org/10.1016/S0003-2670(96)00412-6).
- 1775 [123] R. Rausa, auE. Mazzolari, V. Calemma, Determination of molecular size distributions
1776 of humic acids by high-performance size-exclusion chromatography, *Journal of*
1777 *Chromatography A*. 541 (1991) 419–429. [https://doi.org/10.1016/S0021-](https://doi.org/10.1016/S0021-9673(01)96014-5)
1778 [9673\(01\)96014-5](https://doi.org/10.1016/S0021-9673(01)96014-5).
- 1779 [124] M. Usman, S. Ren, M. Ji, S. O-Thong, Y. Qian, G. Luo, S. Zhang, Characterization and
1780 biogas production potentials of aqueous phase produced from hydrothermal

1781 carbonization of biomass - major components and their binary mixtures, Chemical
1782 Engineering Journal. (2020) 124201. <https://doi.org/10.1016/j.cej.2020.124201>.

1783 [125] N. Kawasaki, K. Matsushige, K. Komatsu, A. Kohzu, F.W. Nara, F. Ogishi, M. Yahata,
1784 H. Mikami, T. Goto, A. Imai, Fast and precise method for HPLC–size exclusion
1785 chromatography with UV and TOC (NDIR) detection: Importance of multiple detectors
1786 to evaluate the characteristics of dissolved organic matter, *Water Research*. 45 (2011)
1787 6240–6248. <https://doi.org/10.1016/j.watres.2011.09.021>.

1788 [126] M. Li, K. Li, K. Chen, C. Liu, Y. Ma, X. Wang, Size-based bioavailability of land-
1789 based DON and its impact on eutrophication of Jiaozhou bay, *Marine Pollution*
1790 *Bulletin*. 152 (2020) 110898. <https://doi.org/10.1016/j.marpolbul.2020.110898>.

1791 [127] A.V. Kudryavtsev, I.V. Perminova, V.S. Petrosyan, Size-exclusion chromatographic
1792 descriptors of humic substances, *Analytica Chimica Acta*. 407 (2000) 193–202.
1793 [https://doi.org/10.1016/S0003-2670\(99\)00814-4](https://doi.org/10.1016/S0003-2670(99)00814-4).

1794 [128] S.L. Phillips, L. Ding, M. Stegemiller, S.V. Olesik, Development of Water-Based
1795 Liquid Chromatography at the Critical Condition, *Anal. Chem*. 75 (2003) 5539–5543.
1796 <https://doi.org/10.1021/ac034456l>.

1797 [129] S.L. Phillips, S.V. Olesik, Initial Characterization of Humic Acids Using Liquid
1798 Chromatography at the Critical Condition Followed by Size-Exclusion
1799 Chromatography and Electrospray Ionization Mass Spectrometry, *Anal. Chem*. 75
1800 (2003) 5544–5553. <https://doi.org/10.1021/ac0344891>.

1801 [130] E.J. Kwak, Y.S. Lee, M. Murata, S. Homma, Effect of pH control on the intermediates
1802 and melanoidins of nonenzymatic browning reaction, *LWT - Food Science and*
1803 *Technology*. 38 (2005) 1–6. <https://doi.org/10.1016/j.lwt.2003.09.011>.

- 1804 [131] R. Rodil, J.B. Quintana, R. Cela, Transformation of phenazone-type drugs during
1805 chlorination, *Water Research*. 46 (2012) 2457–2468.
1806 <https://doi.org/10.1016/j.watres.2012.02.018>.
- 1807 [132] C.P. Passos, M.R. Cepeda, S.S. Ferreira, F.M. Nunes, D.V. Evtuguin, P. Madureira, M.
1808 Vilanova, M.A. Coimbra, Influence of molecular weight on in vitro immunostimulatory
1809 properties of instant coffee, *Food Chemistry*. 161 (2014) 60–66.
1810 <https://doi.org/10.1016/j.foodchem.2014.03.119>.
- 1811 [133] A.S.P. Moreira, M.A. Coimbra, F.M. Nunes, C.P. Passos, S.A.O. Santos, A.J.D.
1812 Silvestre, A.M.N. Silva, M. Rangel, M.R.M. Domingues, Chlorogenic acid–arabinose
1813 hybrid domains in coffee melanoidins: Evidences from a model system, *Food*
1814 *Chemistry*. 185 (2015) 135–144. <https://doi.org/10.1016/j.foodchem.2015.03.086>.
- 1815 [134] A.P.S. Batista, A.C.S.C. Teixeira, W.J. Cooper, B.A. Cottrell, Correlating the chemical
1816 and spectroscopic characteristics of natural organic matter with the photodegradation of
1817 sulfamerazine, *Water Research*. 93 (2016) 20–29.
1818 <https://doi.org/10.1016/j.watres.2015.11.036>.
- 1819 [135] F.C. Moreira, J. Soler, M.F. Alpendurada, R.A.R. Boaventura, E. Brillas, V.J.P. Vilar,
1820 Tertiary treatment of a municipal wastewater toward pharmaceuticals removal by
1821 chemical and electrochemical advanced oxidation processes, *Water Research*. 105
1822 (2016) 251–263. <https://doi.org/10.1016/j.watres.2016.08.036>.
- 1823 [136] C.P. Passos, K. Kukurová, E. Basil, P.A.R. Fernandes, A. Neto, F.M. Nunes, M.
1824 Murkovic, Z. Ciesarová, M.A. Coimbra, Instant coffee as a source of antioxidant-rich
1825 and sugar-free coloured compounds for use in bakery: Application in biscuits, *Food*
1826 *Chemistry*. 231 (2017) 114–121. <https://doi.org/10.1016/j.foodchem.2017.03.105>.
- 1827 [137] M.O. Barbosa, A.R. Ribeiro, N. Ratola, E. Hain, V. Homem, M.F.R. Pereira, L.
1828 Blaney, A.M.T. Silva, Spatial and seasonal occurrence of micropollutants in four

1829 Portuguese rivers and a case study for fluorescence excitation-emission matrices,
1830 Science of The Total Environment. 644 (2018) 1128–1140.
1831 <https://doi.org/10.1016/j.scitotenv.2018.06.355>.

1832 [138] J. Oracz, E. Nebesny, D. Żyżelewicz, Identification and quantification of free and
1833 bound phenolic compounds contained in the high-molecular weight melanoidin
1834 fractions derived from two different types of cocoa beans by UHPLC-DAD-ESI-HR-
1835 MSn, Food Research International. 115 (2019) 135–149.
1836 <https://doi.org/10.1016/j.foodres.2018.08.028>.

1837 [139] F. Hernández, J. Bakker, L. Bijlsma, J. de Boer, A.M. Botero-Coy, Y. Bruinen de
1838 Bruin, S. Fischer, J. Hollender, B. Kasprzyk-Hordern, M. Lamoree, F.J. López, T.L. ter
1839 Laak, J.A. van Leerdam, J.V. Sancho, E.L. Schymanski, P. de Voogt, E.A.
1840 Hogendoorn, The role of analytical chemistry in exposure science: Focus on the aquatic
1841 environment, Chemosphere. 222 (2019) 564–583.
1842 <https://doi.org/10.1016/j.chemosphere.2019.01.118>.

1843 [140] H. Huang, E. Sawade, D. Cook, C.W.K. Chow, M. Drikas, B. Jin, High-performance
1844 size exclusion chromatography with a multi-wavelength absorbance detector study on
1845 dissolved organic matter characterisation along a water distribution system, Journal of
1846 Environmental Sciences. 44 (2016) 235–243. <https://doi.org/10.1016/j.jes.2015.12.011>.

1847 [141] A. Ignatev, T. Tuhkanen, Step-by-step analysis of drinking water treatment trains using
1848 size-exclusion chromatography to fingerprint and track protein-like and humic/fulvic-
1849 like fractions of dissolved organic matter, Environ. Sci.: Water Res. Technol. 5 (2019)
1850 1568–1581. <https://doi.org/10.1039/C9EW00340A>.

1851 [142] A. Ignatev, T. Tuhkanen, Monitoring WWTP performance using size-exclusion
1852 chromatography with simultaneous UV and fluorescence detection to track recalcitrant

1853 wastewater fractions, *Chemosphere*. 214 (2019) 587–597.
1854 <https://doi.org/10.1016/j.chemosphere.2018.09.099>.

1855 [143] Y. Verkh, M. Rozman, M. Petrovic, A non-targeted high-resolution mass spectrometry
1856 data analysis of dissolved organic matter in wastewater treatment, *Chemosphere*. 200
1857 (2018) 397–404. <https://doi.org/10.1016/j.chemosphere.2018.02.095>.

1858 [144] Aalizadeh, R., Schymanski, E., Thomaidis, N., Williams, A., Glowacka, N., Cirka, L.,
1859 2020. Merged Norman Suspect List “SusDat” [WWW Document]. Norman Netw.
1860 [https://www.norman-](https://www.norman-network.com/sites/default/files/files/suspectListExchange/061120Update/susdat_2020-06-18-123341%20%283%29.xlsx)
1861 [network.com/sites/default/files/files/suspectListExchange/061120Update/susdat_2020-](https://www.norman-network.com/sites/default/files/files/suspectListExchange/061120Update/susdat_2020-06-18-123341%20%283%29.xlsx)
1862 [06-18-123341%20%283%29.xlsx](https://www.norman-network.com/sites/default/files/files/suspectListExchange/061120Update/susdat_2020-06-18-123341%20%283%29.xlsx), DOI: 10.5281/zenodo.2664077).

1863 [145] A.C. Maizel, C.K. Remucal, The effect of advanced secondary municipal wastewater
1864 treatment on the molecular composition of dissolved organic matter, *Water Research*.
1865 122 (2017) 42–52. <https://doi.org/10.1016/j.watres.2017.05.055>.

1866 [146] P. Aguilar-Alarcón, S.V. Gonzalez, M.A. Simonsen, A.R. Borrero-Santiago, J. Sanchís,
1867 A. Meriac, J. Kolarevic, A.G. Asimakopoulos, Ø. Mikkelsen, Characterizing changes
1868 of dissolved organic matter composition with the use of distinct feeds in recirculating
1869 aquaculture systems via high-resolution mass spectrometry, *Science of The Total*
1870 *Environment*. 749 (2020) 142326. <https://doi.org/10.1016/j.scitotenv.2020.142326>.

1871 [147] R. Mesfioui, N.G. Love, D.A. Bronk, M.R. Mulholland, P.G. Hatcher, Reactivity and
1872 chemical characterization of effluent organic nitrogen from wastewater treatment plants
1873 determined by Fourier transform ion cyclotron resonance mass spectrometry, *Water*
1874 *Research*. 46 (2012) 622–634. <https://doi.org/10.1016/j.watres.2011.11.022>.

1875 [148] B. Zhang, C. Shan, Z. Hao, J. Liu, B. Wu, B. Pan, Transformation of dissolved organic
1876 matter during full-scale treatment of integrated chemical wastewater: Molecular

1877 composition correlated with spectral indexes and acute toxicity, *Water Research*. 157
1878 (2019) 472–482. <https://doi.org/10.1016/j.watres.2019.04.002>.

1879 [149] M.J. Farré, A. Jaén-Gil, J. Hawkes, M. Petrovic, N. Catalán, Orbitrap molecular
1880 fingerprint of dissolved organic matter in natural waters and its relationship with
1881 NDMA formation potential, *Science of The Total Environment*. 670 (2019) 1019–
1882 1027. <https://doi.org/10.1016/j.scitotenv.2019.03.280>.

1883 [150] N. Cortés-Francisco, M. Harir, M. Lucio, G. Ribera, X. Martínez-Lladó, M. Rovira, P.
1884 Schmitt-Kopplin, N. Hertkorn, J. Caixach, High-field FT-ICR mass spectrometry and
1885 NMR spectroscopy to characterize DOM removal through a nanofiltration pilot plant,
1886 *Water Research*. 67 (2014) 154–165. <https://doi.org/10.1016/j.watres.2014.08.046>.

1887 [151] N. Cortés-Francisco, J. Caixach, Molecular Characterization of Dissolved Organic
1888 Matter through a Desalination Process by High Resolution Mass Spectrometry,
1889 *Environ. Sci. Technol*. 47 (2013) 9619–9627. <https://doi.org/10.1021/es4000388>.

1890 [152] C.A. Hughey, C.L. Hendrickson, R.P. Rodgers, A.G. Marshall, K. Qian, Kendrick
1891 Mass Defect Spectrum: A Compact Visual Analysis for Ultrahigh-Resolution
1892 Broadband Mass Spectra, *Anal. Chem*. 73 (2001) 4676–4681.
1893 <https://doi.org/10.1021/ac010560w>.

1894 [153] M. Llorca, D. Lucas, L. Ferrando-Climent, M. Badia-Fabregat, C. Cruz-Morató, D.
1895 Barceló, S. Rodríguez-Mozaz, Suspect screening of emerging pollutants and their
1896 major transformation products in wastewaters treated with fungi by liquid
1897 chromatography coupled to a high resolution mass spectrometry, *Journal of*
1898 *Chromatography A*. 1439 (2016) 124–136.
1899 <https://doi.org/10.1016/j.chroma.2015.10.077>.

- 1900 [154] D. Leyva, R. Jaffe, F. Fernandez-Lima, Structural Characterization of Dissolved
1901 Organic Matter at the Chemical Formula Level Using TIMS-FT-ICR MS/MS, *Anal.*
1902 *Chem.* 92 (2020) 11960–11966. <https://doi.org/10.1021/acs.analchem.0c02347>.
- 1903 [155] P. Gago-Ferrero, A.A. Bletsou, D.E. Damalas, R. Aalizadeh, N.A. Alygizakis, H.P.
1904 Singer, J. Hollender, N.S. Thomaidis, Wide-scope target screening of >2000 emerging
1905 contaminants in wastewater samples with UPLC-Q-ToF-HRMS/MS and smart
1906 evaluation of its performance through the validation of 195 selected representative
1907 analytes, *Journal of Hazardous Materials.* 387 (2020) 121712.
1908 <https://doi.org/10.1016/j.jhazmat.2019.121712>.
- 1909 [156] L. Vergeynst, A. Haeck, P. De Wispelaere, H. Van Langenhove, K. Demeestere, Multi-
1910 residue analysis of pharmaceuticals in wastewater by liquid chromatography–magnetic
1911 sector mass spectrometry: Method quality assessment and application in a Belgian case
1912 study, *Chemosphere.* 119 (2015) S2–S8.
1913 <https://doi.org/10.1016/j.chemosphere.2014.03.069>.
- 1914 [157] G. Nürenberg, U. Kunkel, A. Wick, P. Falås, A. Joss, T.A. Ternes, Nontarget analysis:
1915 A new tool for the evaluation of wastewater processes, *Water Research.* 163 (2019)
1916 114842. <https://doi.org/10.1016/j.watres.2019.07.009>.
- 1917 [158] D. Camacho-Muñoz, B. Kasprzyk-Hordern, K.V. Thomas, Enantioselective
1918 simultaneous analysis of selected pharmaceuticals in environmental samples by
1919 ultrahigh performance supercritical fluid based chromatography tandem mass
1920 spectrometry, *Analytica Chimica Acta.* 934 (2016) 239–251.
1921 <https://doi.org/10.1016/j.aca.2016.05.051>.
- 1922 [159] B. Loganathan, M. Phillips, H. Mowery, T.L. Jones-Lepp, Contamination profiles and
1923 mass loadings of macrolide antibiotics and illicit drugs from a small urban wastewater

1924 treatment plant, *Chemosphere*. 75 (2009) 70–77.
1925 <https://doi.org/10.1016/j.chemosphere.2008.11.047>.

1926 [160] N. Negreira, I. Rodríguez, M. Ramil, E. Rubí, R. Cela, Solid-phase extraction followed
1927 by liquid chromatography–tandem mass spectrometry for the determination of
1928 hydroxylated benzophenone UV absorbers in environmental water samples, *Analytica*
1929 *Chimica Acta*. 654 (2009) 162–170. <https://doi.org/10.1016/j.aca.2009.09.033>.

1930 [161] W. Bae, H. Won, B. Hwang, R.A. de Toledo, J. Chung, K. Kwon, H. Shim,
1931 Characterization of refractory matters in dyeing wastewater during a full-scale Fenton
1932 process following pure-oxygen activated sludge treatment, *Journal of Hazardous*
1933 *Materials*. 287 (2015) 421–428. <https://doi.org/10.1016/j.jhazmat.2015.01.052>.

1934 [162] Q.-Q. Du, S.-Y. Liu, R.-F. Xu, M. Li, F.-R. Song, Z.-Q. Liu, Studies on structures and
1935 activities of initial Maillard reaction products by electrospray ionisation mass
1936 spectrometry combined with liquid chromatography in processing of red ginseng, *Food*
1937 *Chemistry*. 135 (2012) 832–838. <https://doi.org/10.1016/j.foodchem.2012.04.126>.

1938 [163] K. Brudzynski, D. Miotto, Honey melanoidins: Analysis of the compositions of the
1939 high molecular weight melanoidins exhibiting radical-scavenging activity, *Food*
1940 *Chemistry*. 127 (2011) 1023–1030. <https://doi.org/10.1016/j.foodchem.2011.01.075>.

1941 [164] A. Hollnagel, L.W. Kroh, Degradation of Oligosaccharides in Nonenzymatic Browning
1942 by Formation of α -Dicarbonyl Compounds via a “Peeling Off” Mechanism, *J. Agric.*
1943 *Food Chem*. 48 (2000) 6219–6226. <https://doi.org/10.1021/jf9906127>.

1944 [165] N. Sudasinghe, B. Dungan, P. Lammers, K. Albrecht, D. Elliott, R. Hallen, T. Schaub,
1945 High resolution FT-ICR mass spectral analysis of bio-oil and residual water soluble
1946 organics produced by hydrothermal liquefaction of the marine microalga
1947 *Nannochloropsis salina*, *Fuel*. 119 (2014) 47–56.
1948 <https://doi.org/10.1016/j.fuel.2013.11.019>.

- 1949 [166] M. Derrien, S.R. Brogi, R. Gonçalves-Araujo, Characterization of aquatic organic
1950 matter: Assessment, perspectives and research priorities, *Water Research*. 163 (2019)
1951 114908. <https://doi.org/10.1016/j.watres.2019.114908>.
- 1952 [167] E.G. Alves Filho, L.M. Alexandre e Silva, A.G. Ferreira, Advancements in waste water
1953 characterization through NMR spectroscopy: review: Advancements in waste water
1954 characterization through NMR spectroscopy: review, *Magn. Reson. Chem.* 53 (2015)
1955 648–657. <https://doi.org/10.1002/mrc.4158>.
- 1956 [168] P. Conte, R. Spaccini, A. Piccolo, State of the art of CPMAS ¹³C-NMR spectroscopy
1957 applied to natural organic matter, *Progress in Nuclear Magnetic Resonance*
1958 *Spectroscopy*. 44 (2004) 215–223. <https://doi.org/10.1016/j.pnmrs.2004.02.002>.
- 1959 [169] J. Mao, X. Cao, D.C. Olk, W. Chu, K. Schmidt-Rohr, Advanced solid-state NMR
1960 spectroscopy of natural organic matter, *Progress in Nuclear Magnetic Resonance*
1961 *Spectroscopy*. 100 (2017) 17–51. <https://doi.org/10.1016/j.pnmrs.2016.11.003>.
- 1962 [170] L.N. Soucémarianadin, B. Erhagen, M.B. Nilsson, M.G. Öquist, P. Immerzeel, J.
1963 Schleucher, Two dimensional NMR spectroscopy for molecular characterization of soil
1964 organic matter: Application to boreal soils and litter, *Organic Geochemistry*. 113
1965 (2017) 184–195. <https://doi.org/10.1016/j.orggeochem.2017.06.019>.
- 1966 [171] C. Simmler, J.G. Napolitano, J.B. McAlpine, S.-N. Chen, G.F. Pauli, Universal
1967 quantitative NMR analysis of complex natural samples, *Current Opinion in*
1968 *Biotechnology*. 25 (2014) 51–59. <https://doi.org/10.1016/j.copbio.2013.08.004>.
- 1969 [172] A.I. Farounbi, P.K. Mensah, E.O. Olawode, N.P. Ngqwala, ¹H-NMR Determination of
1970 Organic Compounds in Municipal Wastewaters and the Receiving Surface Waters in
1971 Eastern Cape Province of South Africa, *Molecules*. 25 (2020) 713.
1972 <https://doi.org/10.3390/molecules25030713>.

- 1973 [173] N. Mahieu, D.C. Olk, E.W. Randall, Accumulation of heterocyclic nitrogen in humified
1974 organic matter: a ¹⁵N-NMR study of lowland rice soils, *Eur J Soil Science*. 51 (2000)
1975 379–389. <https://doi.org/10.1046/j.1365-2389.2000.00328.x>.
- 1976 [174] K.A. Thorn, L.G. Cox, ¹⁵N NMR spectra of naturally abundant nitrogen in soil and
1977 aquatic natural organic matter samples of the International Humic Substances Society,
1978 *Organic Geochemistry*. 40 (2009) 484–499.
1979 <https://doi.org/10.1016/j.orggeochem.2009.01.007>.
- 1980 [175] L.B. Staal, A.B. Petersen, C.A. Jørgensen, U.G. Nielsen, P.H. Nielsen, K. Reitzel,
1981 Extraction and quantification of polyphosphates in activated sludge from waste water
1982 treatment plants by ³¹P NMR spectroscopy, *Water Research*. 157 (2019) 346–355.
1983 <https://doi.org/10.1016/j.watres.2019.03.065>.
- 1984 [176] L. Wang, W. Amelung, J. Prietzel, S. Willbold, Transformation of organic phosphorus
1985 compounds during 1500 years of organic soil formation in Bavarian Alpine forests – A
1986 ³¹P NMR study, *Geoderma*. 340 (2019) 192–205.
1987 <https://doi.org/10.1016/j.geoderma.2019.01.029>.
- 1988 [177] X. Han, F. Wang, B. Zhou, H. Chen, R. Yuan, S. Liu, X. Zhou, L. Gao, Y. Lu, R.
1989 Zhang, Phosphorus complexation of sewage sludge during thermal hydrolysis with
1990 different reaction temperature and reaction time by P K-edge XANES and ³¹P NMR,
1991 *Science of The Total Environment*. 688 (2019) 1–9.
1992 <https://doi.org/10.1016/j.scitotenv.2019.06.017>.
- 1993 [178] N.D. Weiss, L.G. Thygesen, C. Felby, C. Roslander, K. Gourlay, Biomass-water
1994 interactions correlate to recalcitrance and are intensified by pretreatment: An
1995 investigation of water constraint and retention in pretreated spruce using low field
1996 NMR and water retention value techniques, *Biotechnol Progress*. 33 (2017) 146–153.
1997 <https://doi.org/10.1002/btpr.2398>.

- 1998 [179] B. Wu, K. Zhou, Y. He, X. Chai, X. Dai, Unraveling the water states of waste-activated
1999 sludge through transverse spin-spin relaxation time of low-field NMR, *Water Research*.
2000 155 (2019) 266–274. <https://doi.org/10.1016/j.watres.2019.02.031>.
- 2001 [180] E.G. Alves Filho, L. Sartori, L.M.A. Silva, B.F. Silva, P.S. Fadini, R. Soong, A.
2002 Simpson, A.G. Ferreira, Non-targeted analyses of organic compounds in urban
2003 wastewater: Non-targeted analyses of organic compounds in urban wastewater, *Magn.*
2004 *Reson. Chem.* 53 (2015) 704–710. <https://doi.org/10.1002/mrc.4169>.
- 2005 [181] E.G. Alves Filho, L.M. Alexandre e Silva, A.G. Ferreira, Screening Organic
2006 Compounds in Urban Wastewater Using a Hyphenated System and NMR Pattern
2007 Recognition, in: R.K. Harris, R.L. Wasylshen (Eds.), *EMagRes*, John Wiley & Sons,
2008 Ltd, Chichester, UK, 2017: pp. 173–186.
2009 <https://doi.org/10.1002/9780470034590.emrstm1537>.
- 2010 [182] A. Nebbioso, P. Mazzei, D. Savy, Reduced complexity of multidimensional and
2011 diffusion NMR spectra of soil humic fractions as simplified by Humeomics, *Chem.*
2012 *Biol. Technol. Agric.* 1 (2014) 24. <https://doi.org/10.1186/s40538-014-0024-y>.
- 2013 [183] S. Shakeri Yekta, M. Hedenström, B.H. Svensson, I. Sundgren, M. Dario, A. Enrich-
2014 Prast, N. Hertkorn, A. Björn, Molecular characterization of particulate organic matter
2015 in full scale anaerobic digesters: An NMR spectroscopy study, *Science of The Total*
2016 *Environment*. 685 (2019) 1107–1115. <https://doi.org/10.1016/j.scitotenv.2019.06.264>.
- 2017 [184] Y. Zhang, S. Liu, Y. Cheng, Z. Cai, C. Müller, J. Zhang, Composition of soil
2018 recalcitrant C regulates nitrification rates in acidic soils, *Geoderma*. 337 (2019) 965–
2019 972. <https://doi.org/10.1016/j.geoderma.2018.11.014>.
- 2020 [185] S.-H. Yao, Y.-L. Zhang, Y. Han, X.-Z. Han, J.-D. Mao, B. Zhang, Labile and
2021 recalcitrant components of organic matter of a Mollisol changed with land use and

2022 plant litter management: An advanced ¹³C NMR study, *Science of The Total*
2023 *Environment*. 660 (2019) 1–10. <https://doi.org/10.1016/j.scitotenv.2018.12.403>.

2024 [186] R. Mylotte, A. Sutrisno, H. Farooq, H. Masoom, R. Soong, M.H.B. Hayes, A.J.
2025 Simpson, Insights into the composition of recalcitrant organic matter from estuarine
2026 sediments using NMR spectroscopy, *Organic Geochemistry*. 98 (2016) 155–165.
2027 <https://doi.org/10.1016/j.orggeochem.2016.05.011>.

2028 [187] X. Fang, K. Schmidt-Rohr, Fate of the Amino Acid in Glucose–Glycine Melanoidins
2029 Investigated by Solid-State Nuclear Magnetic Resonance (NMR), *J. Agric. Food Chem.*
2030 57 (2009) 10701–10711. <https://doi.org/10.1021/jf9020587>.

2031 [188] G.F. Mohsin, F.-J. Schmitt, C. Kanzler, J. Dirk Epping, S. Flemig, A. Hornemann,
2032 Structural characterization of melanoidin formed from d-glucose and l-alanine at
2033 different temperatures applying FTIR, NMR, EPR, and MALDI-ToF-MS, *Food*
2034 *Chemistry*. 245 (2018) 761–767. <https://doi.org/10.1016/j.foodchem.2017.11.115>.

2035 [189] A. Rodríguez, P. Lema, M.I. Bessio, G. Moyna, L.A. Panizzolo, F. Ferreira, Isolation
2036 and Characterization of Melanoidins from Dulce de Leche, A Confectionary Dairy
2037 Product, *Molecules*. 24 (2019) 4163. <https://doi.org/10.3390/molecules24224163>.

2038 [190] J. Skjemstad, P. Clarke, J. Taylor, J. Oades, R. Newman, The removal of magnetic
2039 materials from surface soils - a solid state ¹³C CP/MAS NMR study, *Soil Res.* 32
2040 (1994) 1215. <https://doi.org/10.1071/SR9941215>.

2041 [191] P. Giraudeau, NMR-based metabolomics and fluxomics: developments and future
2042 prospects, *Analyst*. 145 (2020) 2457–2472. <https://doi.org/10.1039/D0AN00142B>.

2043 [192] N.Z. Shafii, A.S.M. Saudi, J.C. Pang, I.F. Abu, N. Sapawe, M.K.A. Kamarudin, H.F.M.
2044 Saudi, Application of chemometrics techniques to solve environmental issues in
2045 Malaysia, *Heliyon*. 5 (2019) e02534. <https://doi.org/10.1016/j.heliyon.2019.e02534>.

- 2046 [193] Y.K. Lee, M.-H. Lee, J. Hur, A new molecular weight (MW) descriptor of dissolved
2047 organic matter to represent the MW-dependent distribution of aromatic condensation:
2048 Insights from biodegradation and pyrene binding experiments, *Science of The Total*
2049 *Environment*. 660 (2019) 169–176. <https://doi.org/10.1016/j.scitotenv.2019.01.035>.
- 2050 [194] S. Mas, A. de Juan, R. Tauler, A.C. Olivieri, G.M. Escandar, Application of
2051 chemometric methods to environmental analysis of organic pollutants: A review,
2052 *Talanta*. 80 (2010) 1052–1067. <https://doi.org/10.1016/j.talanta.2009.09.044>.
- 2053 [195] N. Bratchell, Cluster analysis, *Chemometrics and Intelligent Laboratory Systems*. 6
2054 (1989) 105–125. [https://doi.org/10.1016/0169-7439\(87\)80054-0](https://doi.org/10.1016/0169-7439(87)80054-0).
- 2055 [196] K.P. Singh, A. Malik, D. Mohan, S. Sinha, V.K. Singh, Chemometric data analysis of
2056 pollutants in wastewater—a case study, *Analytica Chimica Acta*. 532 (2005) 15–25.
2057 <https://doi.org/10.1016/j.aca.2004.10.043>.
- 2058 [197] G. Loganathan, S. Krishnaraj, J. Muthumanickam, R. K, Chemometric and trend
2059 analysis of water quality of the South Chennai lakes: an integrated environmental
2060 study: Multivariate statistical analysis, *J. Chemometrics*. 29 (2015) 59–68.
2061 <https://doi.org/10.1002/cem.2664>.
- 2062 [198] M. Vega, R. Pardo, E. Barrado, L. Debán, Assessment of seasonal and polluting effects
2063 on the quality of river water by exploratory data analysis, *Water Research*. 32 (1998)
2064 3581–3592. [https://doi.org/10.1016/S0043-1354\(98\)00138-9](https://doi.org/10.1016/S0043-1354(98)00138-9).
- 2065 [199] A.C. Rencher, *Methods of Multivariate Analysis*, John Wiley & Sons, Inc., New York,
2066 NY, USA, 2002. <https://doi.org/10.1002/0471271357>.
- 2067 [200] D.P. Mesquita, C. Quintelas, A.L. Amaral, E.C. Ferreira, Monitoring biological
2068 wastewater treatment processes: recent advances in spectroscopy applications, *Rev*
2069 *Environ Sci Biotechnol*. 16 (2017) 395–424. [https://doi.org/10.1007/s11157-017-9439-](https://doi.org/10.1007/s11157-017-9439-9)
2070 9.

- 2071 [201] C.F. Galinha, G. Carvalho, C.A.M. Portugal, G. Guglielmi, R. Oliveira, J.G. Crespo,
2072 M.A.M. Reis, Real-time monitoring of membrane bioreactors with 2D-fluorescence
2073 data and statistically based models, *Water Science and Technology*. 63 (2011) 1381–
2074 1388. <https://doi.org/10.2166/wst.2011.195>.
- 2075 [202] S. Platikanov, S. Rodriguez-Mozaz, B. Huerta, D. Barceló, J. Cros, M. Batle, G. Poch,
2076 R. Tauler, Chemometrics quality assessment of wastewater treatment plant effluents
2077 using physicochemical parameters and UV absorption measurements, *Journal of*
2078 *Environmental Management*. 140 (2014) 33–44.
2079 <https://doi.org/10.1016/j.jenvman.2014.03.006>.
- 2080 [203] M.A. Jiménez-González, A.M. Álvarez, P. Carral, G. Almendros, Chemometric
2081 assessment of soil organic matter storage and quality from humic acid infrared spectra,
2082 *Science of The Total Environment*. 685 (2019) 1160–1168.
2083 <https://doi.org/10.1016/j.scitotenv.2019.06.231>.
- 2084 [204] F. Savorani, G. Tomasi, S.B. Engelsen, icoshift: A versatile tool for the rapid alignment
2085 of 1D NMR spectra, *Journal of Magnetic Resonance*. 202 (2010) 190–202.
2086 <https://doi.org/10.1016/j.jmr.2009.11.012>.
- 2087 [205] G. Tomasi, F. Savorani, S.B. Engelsen, icoshift: An effective tool for the alignment of
2088 chromatographic data, *Journal of Chromatography A*. 1218 (2011) 7832–7840.
2089 <https://doi.org/10.1016/j.chroma.2011.08.086>.
- 2090 [206] R.H. Jellema, A. Folch-Fortuny, M.M.W.B. Hendriks, Variable Shift and Alignment,
2091 in: *Comprehensive Chemometrics*, Elsevier, 2020: pp. 115–136.
2092 <https://doi.org/10.1016/B978-0-12-409547-2.14886-3>.
- 2093 [207] L.L. Hohrenk, M. Vosough, T.C. Schmidt, Implementation of Chemometric Tools To
2094 Improve Data Mining and Prioritization in LC-HRMS for Nontarget Screening of

2095 Organic Micropollutants in Complex Water Matrixes, *Anal. Chem.* 91 (2019) 9213–
2096 9220. <https://doi.org/10.1021/acs.analchem.9b01984>.

2097 [208] P. Geladi, B.R. Kowalski, Partial least-squares regression: a tutorial, *Analytica Chimica*
2098 *Acta.* 185 (1986) 1–17. [https://doi.org/10.1016/0003-2670\(86\)80028-9](https://doi.org/10.1016/0003-2670(86)80028-9).

2099 [209] D. Ballabio, V. Consonni, Classification tools in chemistry. Part 1: linear models. PLS-
2100 DA, *Anal. Methods.* 5 (2013) 3790. <https://doi.org/10.1039/c3ay40582f>.

2101 [210] R.G. Brereton, Introduction to multivariate calibration in analytical chemistry, *Analyst.*
2102 125 (2000) 2125–2154. <https://doi.org/10.1039/b003805i>.

2103 [211] I.-G. Chong, C.-H. Jun, Performance of some variable selection methods when
2104 multicollinearity is present, *Chemometrics and Intelligent Laboratory Systems.* 78
2105 (2005) 103–112. <https://doi.org/10.1016/j.chemolab.2004.12.011>.

2106 [212] M.-N. Pons, W. Jing, O. Potier, CHEMOMETRIC ESTIMATION OF
2107 WASTEWATER COMPOSITION FOR THE ON-LINE CONTROL OF
2108 TREATMENT PLANTS, *IFAC Proceedings Volumes.* 38 (2005) 49–54.
2109 <https://doi.org/10.3182/20050703-6-CZ-1902.02212>.

2110 [213] Ao Chen, Hongpeng Zhou, Yujian An, Wei Sun, PCA and PLS monitoring approaches
2111 for fault detection of wastewater treatment process, in: 2016 IEEE 25th International
2112 Symposium on Industrial Electronics (ISIE), IEEE, Santa Clara, CA, USA, 2016: pp.
2113 1022–1027. <https://doi.org/10.1109/ISIE.2016.7745032>.

2114 [214] D.P. Mesquita, A.L. Amaral, E.C. Ferreira, Activated sludge characterization through
2115 microscopy: A review on quantitative image analysis and chemometric techniques,
2116 *Analytica Chimica Acta.* 802 (2013) 14–28. <https://doi.org/10.1016/j.aca.2013.09.016>.

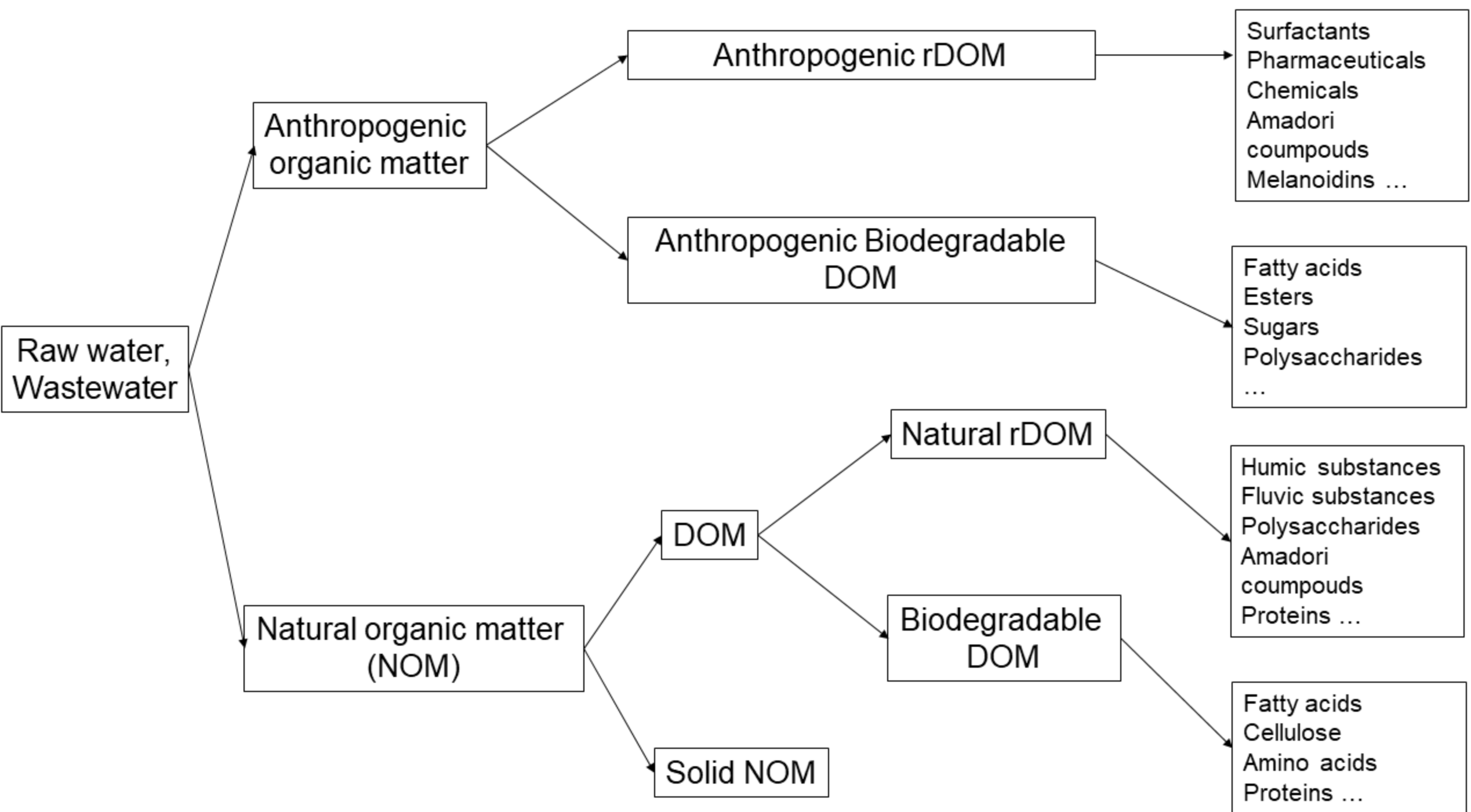
2117 [215] R.L. Sleighter, Z. Liu, J. Xue, P.G. Hatcher, Multivariate Statistical Approaches for the
2118 Characterization of Dissolved Organic Matter Analyzed by Ultrahigh Resolution Mass

2119 Spectrometry, Environ. Sci. Technol. 44 (2010) 7576–7582.
2120 <https://doi.org/10.1021/es1002204>.

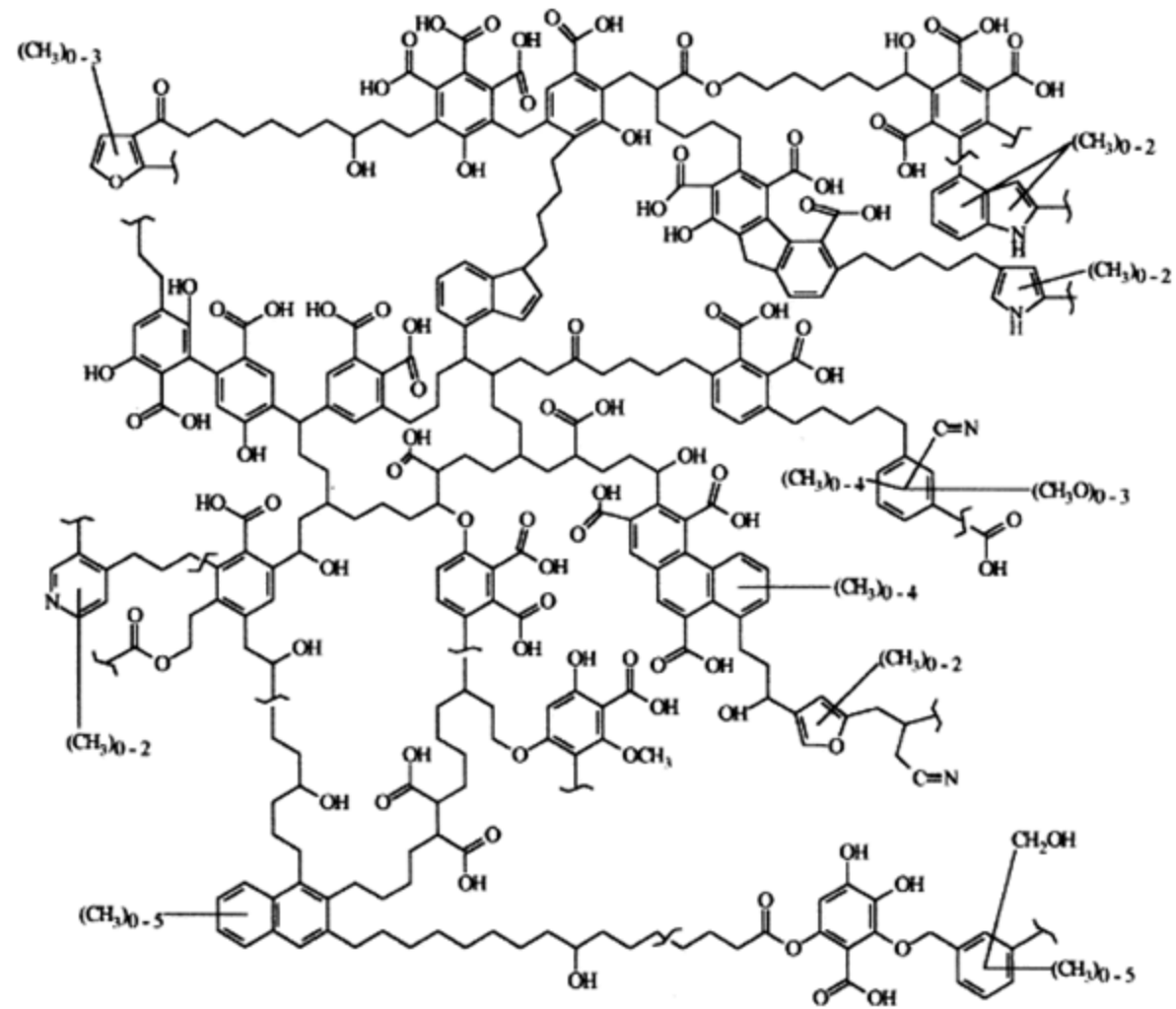
2121 [216] N. Hertkorn, C. Ruecker, M. Meringer, R. Gugisch, M. Frommberger, E.M. Perdue, M.
2122 Witt, P. Schmitt-Kopplin, High-precision frequency measurements: indispensable tools
2123 at the core of the molecular-level analysis of complex systems, Anal Bioanal Chem.
2124 389 (2007) 1311–1327. <https://doi.org/10.1007/s00216-007-1577-4>.

2125 [217] M. Atgié, J.C. Garrigues, A. Chennevière, O. Masbernat, K. Roger, Gum Arabic in
2126 solution: Composition and multi-scale structures, Food Hydrocolloids. 91 (2019) 319–
2127 330. <https://doi.org/10.1016/j.foodhyd.2019.01.033>.

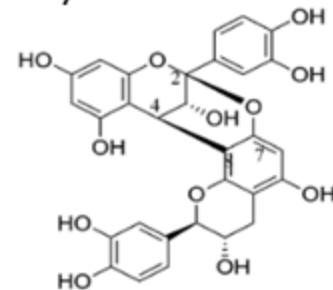
2128 [218] S. Jannin, J.-N. Dumez, P. Giraudeau, D. Kurzbach, Application and methodology of
2129 dissolution dynamic nuclear polarization in physical, chemical and biological contexts,
2130 Journal of Magnetic Resonance. 305 (2019) 41–50.
2131 <https://doi.org/10.1016/j.jmr.2019.06.001>.
2132



a) Humic substances



b) Phenol



Proanthocyanidin A1, epicatechin-(2β→7,4 β→8)-catechin

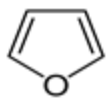
c) Maillard compounds



Pyrazine

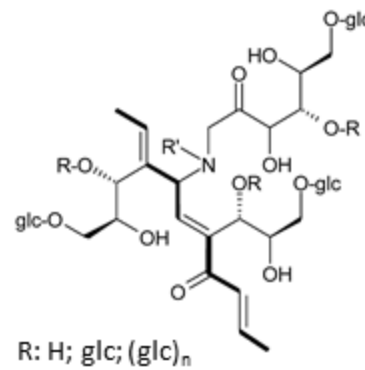


pyridine

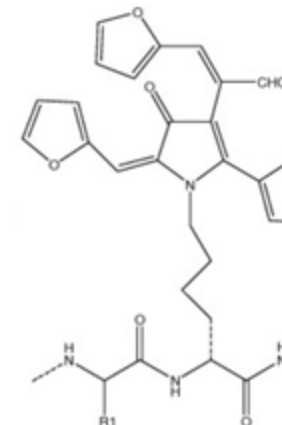


Furan

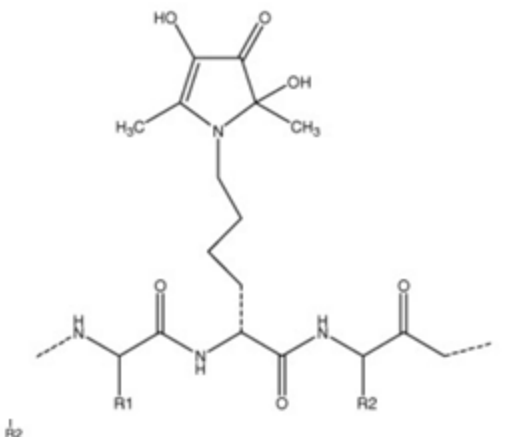
d) Melanoidins



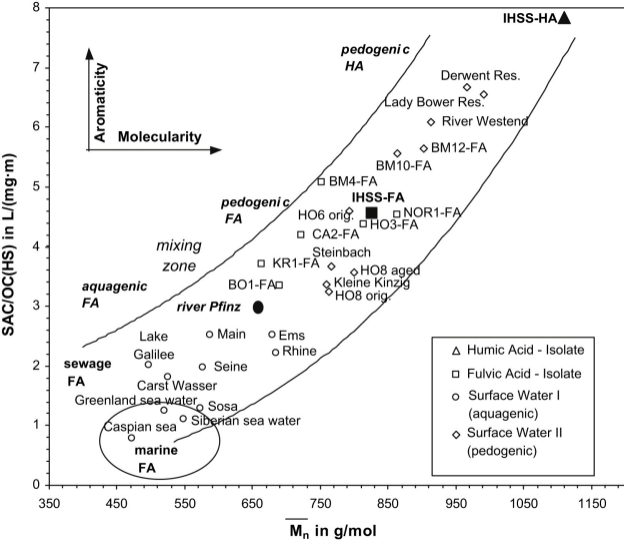
Melanoidin structure formed from 3-deoxyhexosuloses and ARP

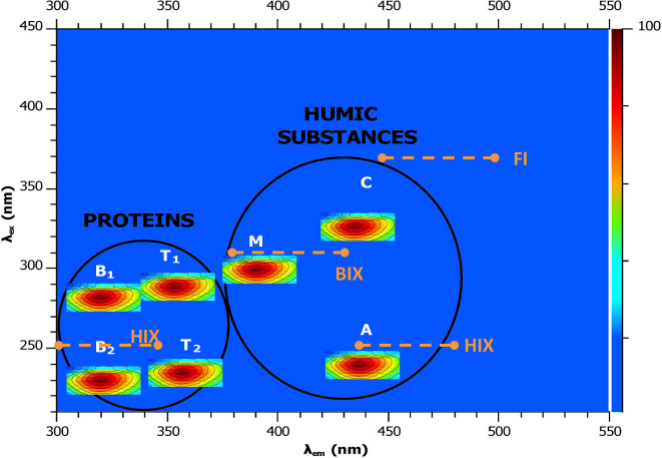


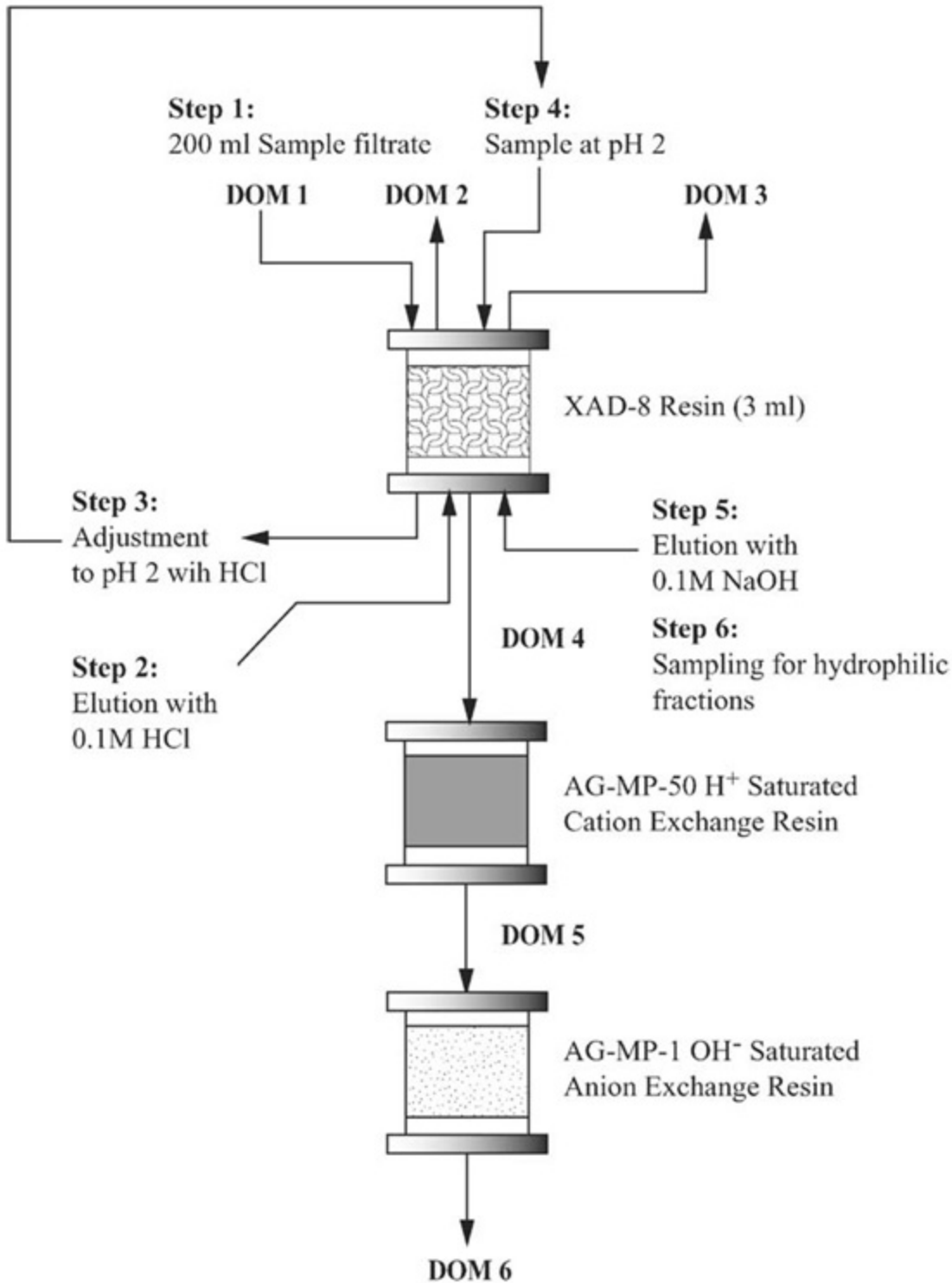
Trifuryl-pyrrolinone

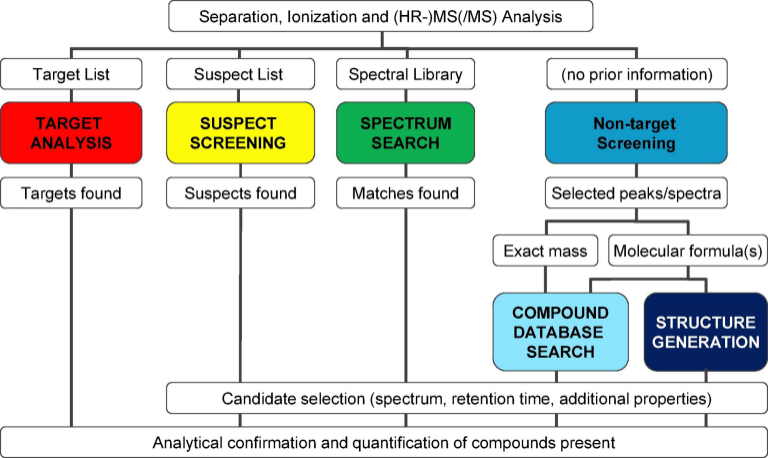


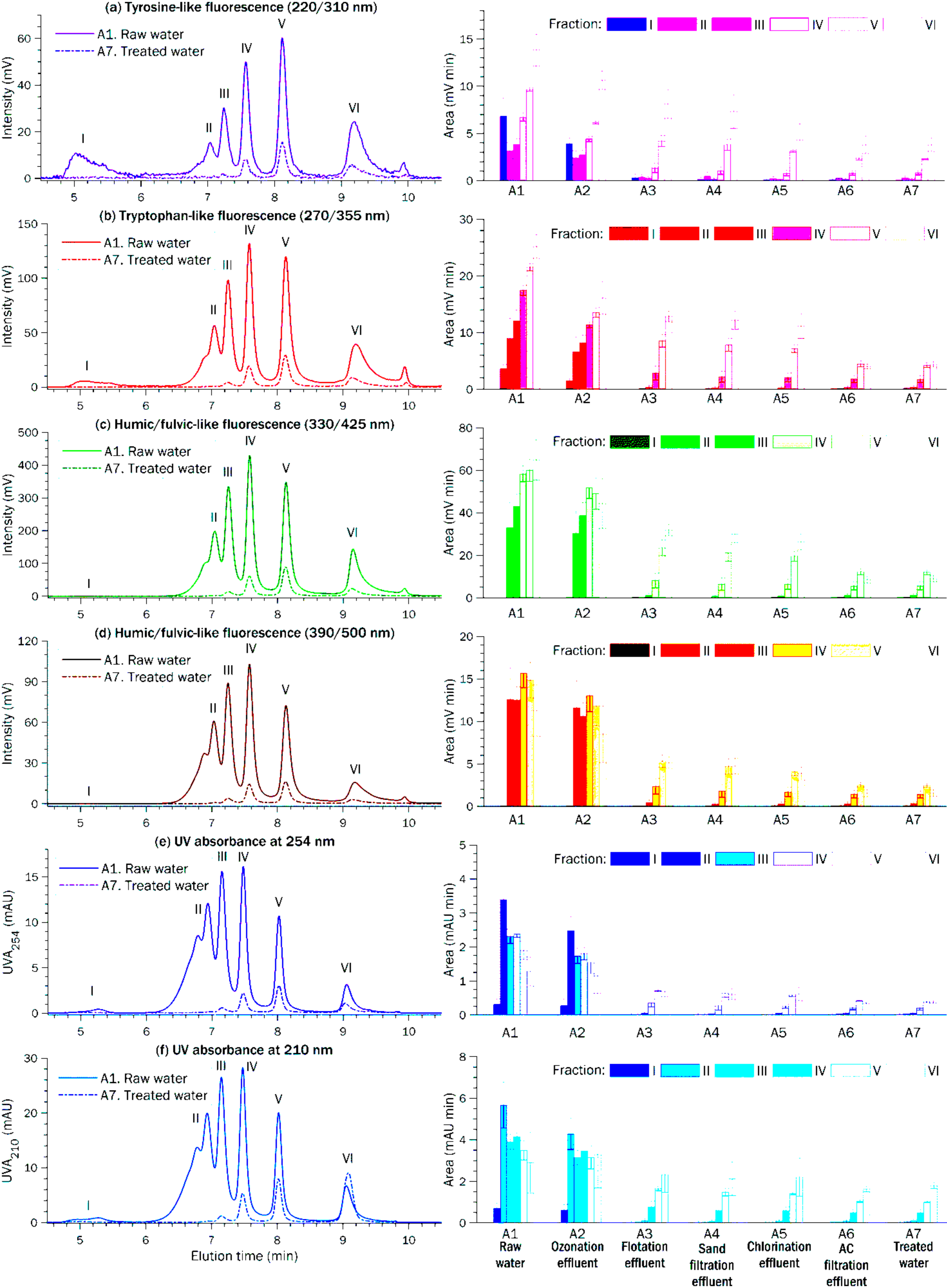
3) Pronyl-protein

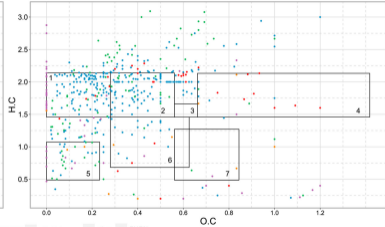
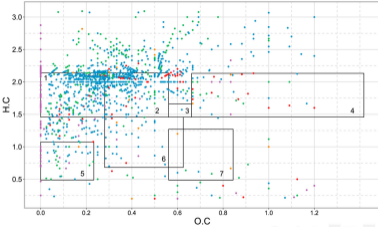






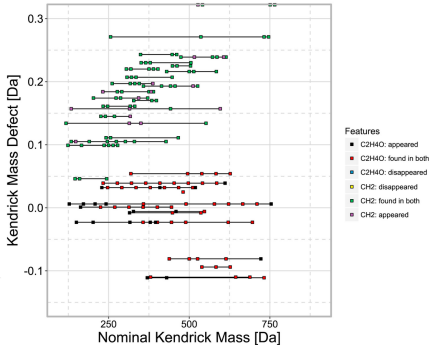
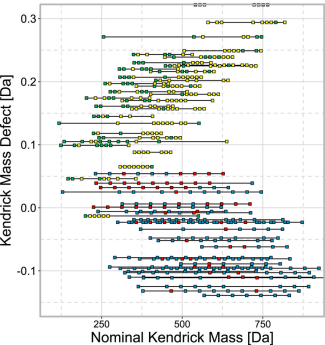




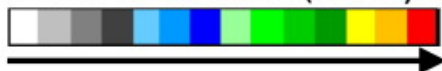


Type of matter

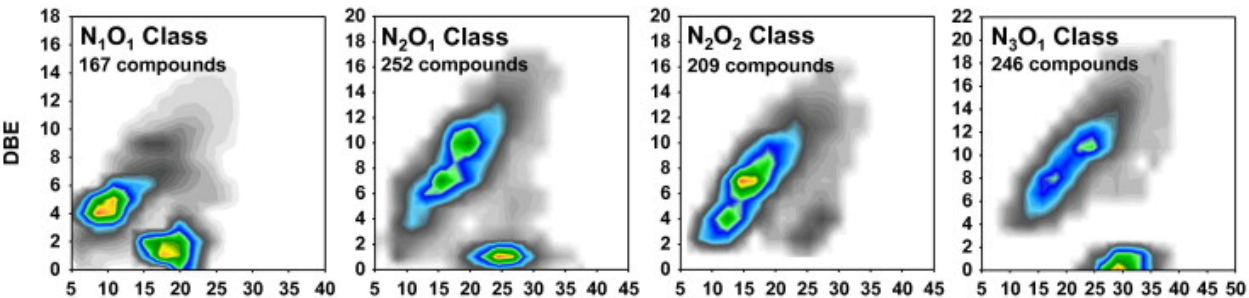
- CHO
- CHONS
- with Halogens
- other
- CHON



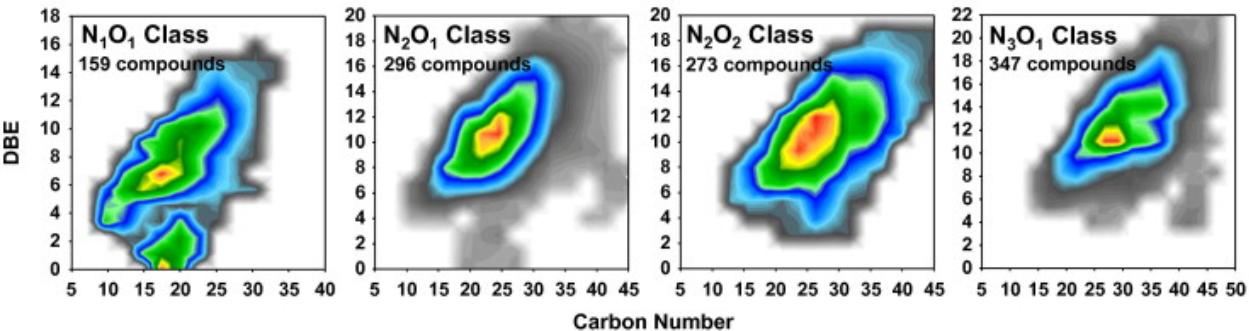
Relative Abundance (% total)

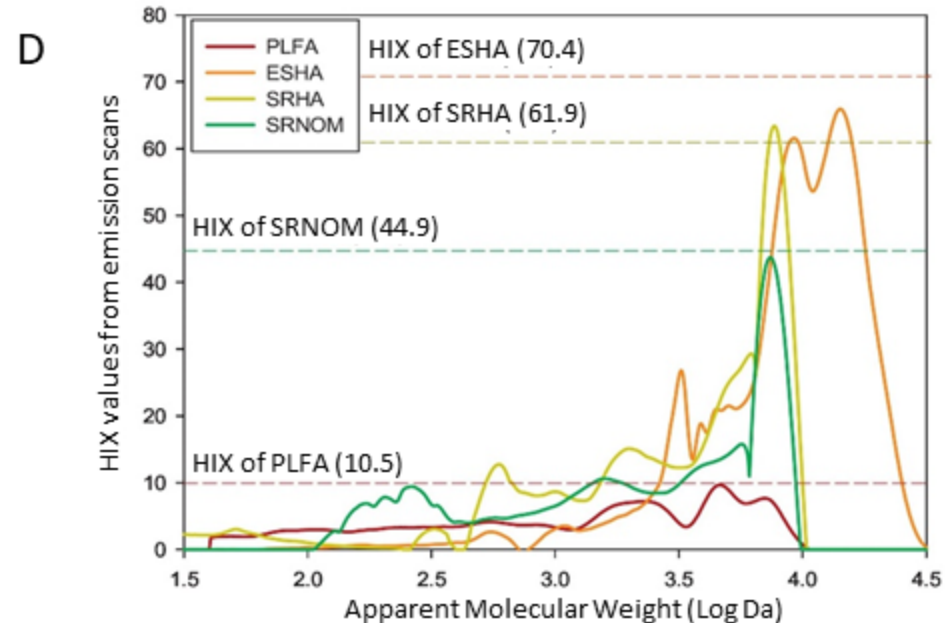
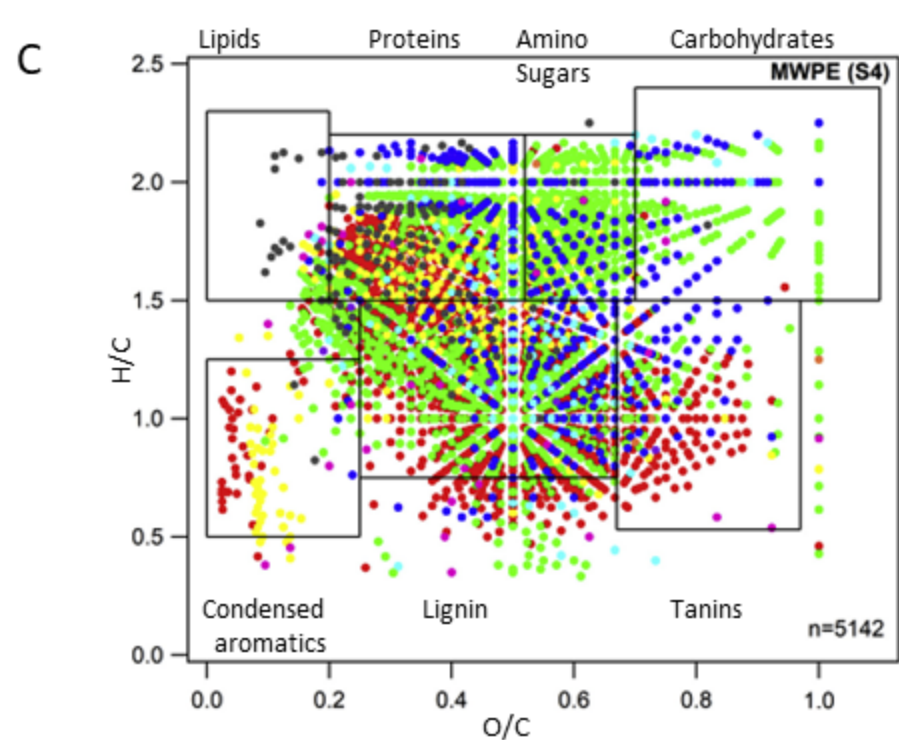
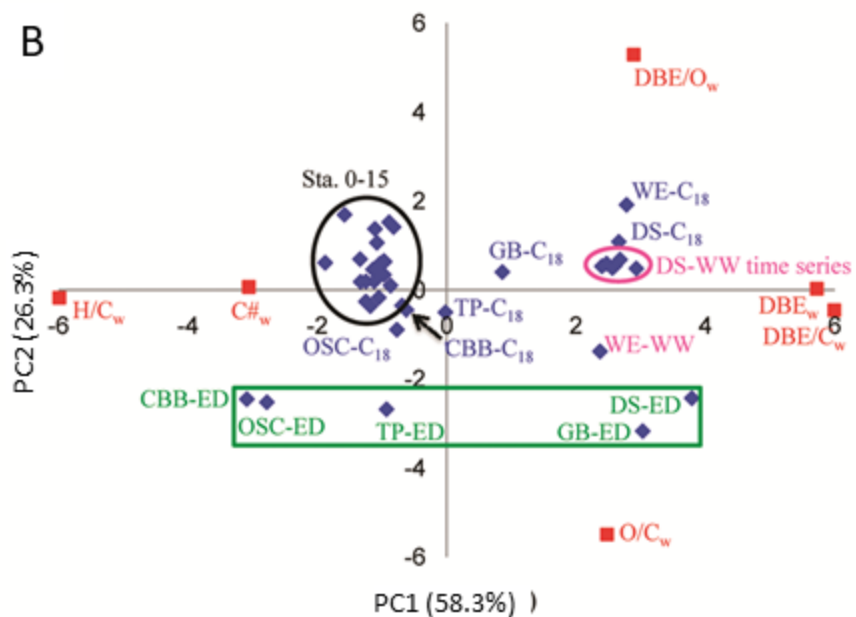
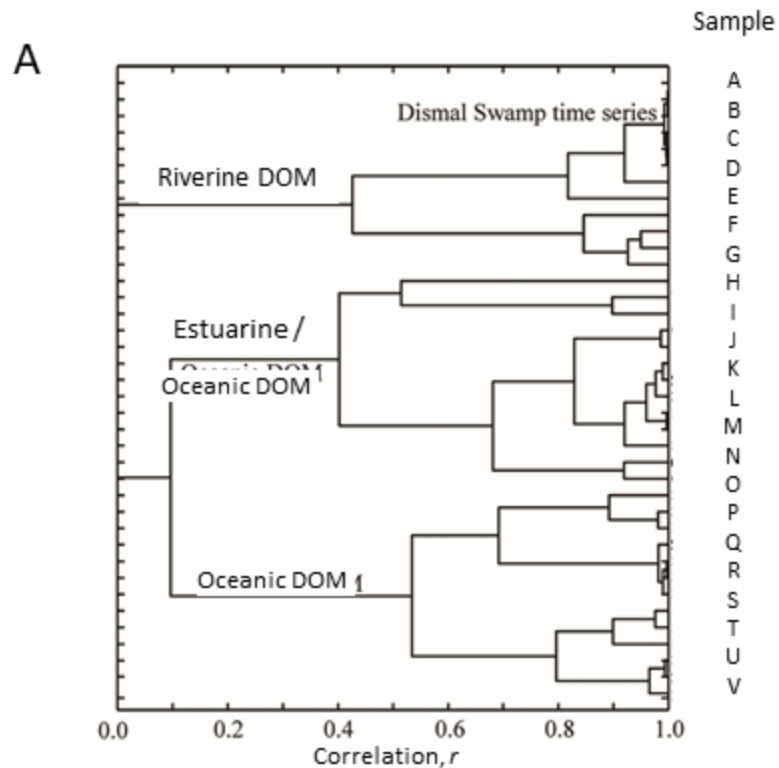


Aqueous Fraction



HTL Bio-Oil





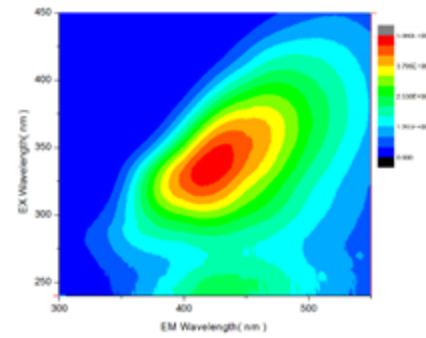
Wastewater sample

Analytical and Chemometrics

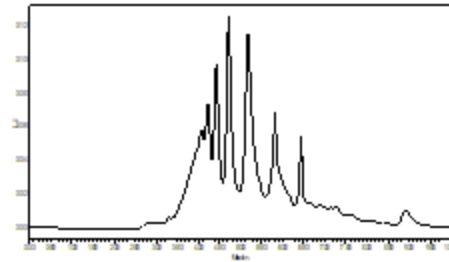
Chemical structure



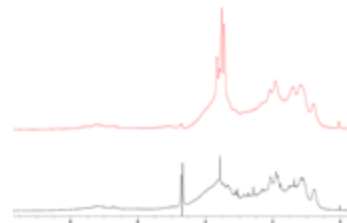
Fluo



SEC



NMR



MS

