



HAL
open science

Multimethod Approach for Extensive Characterization of Gallnut Tannin Extracts

Aude Watrelot, Christine Le Guernevé, Hélène Hallé, Emmanuelle Meudec, Pascale Williams, Bertrand Robillard, François Garcia, Celine Poncet-Legrand, Veronique Cheynier

► **To cite this version:**

Aude Watrelot, Christine Le Guernevé, Hélène Hallé, Emmanuelle Meudec, Pascale Williams, et al.. Multimethod Approach for Extensive Characterization of Gallnut Tannin Extracts. *Journal of Agricultural and Food Chemistry*, inPress, 68 (47), pp.13426-13438. 10.1021/acs.jafc.9b08221 . hal-02888982

HAL Id: hal-02888982

<https://hal.inrae.fr/hal-02888982v1>

Submitted on 11 Jul 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.

This document is confidential and is proprietary to the American Chemical Society and its authors. Do not copy or disclose without written permission. If you have received this item in error, notify the sender and delete all copies.

Multi-method approach for extensive characterization of gallnut tannin extracts

Journal:	<i>Journal of Agricultural and Food Chemistry</i>
Manuscript ID	jf-2019-082217.R1
Manuscript Type:	Article
Date Submitted by the Author:	n/a
Complete List of Authors:	WatreLOT, Aude; INRAE, SPO; Iowa State University College of Agriculture and Life Sciences Le Guernevé, Christine; INRAE, SPO Halle, H��l��ne; INRAE, SPO Meudec, Emmanuelle; INRAE, SPO Veran, Frederic; INRAE, SPO Williams, Pascale; INRAE, SPO Robillard, Bertrand; Institut Œnologique de Champagne Garcia, Fran��ois ; Univ Montpellier, SPO Poncet-Legrand, C��line; INRA, Umr Spo Cheynier, Veronique ; INRAE, SPO

SCHOLARONE™
Manuscripts

1 Multi-method approach for extensive characterization of gallnut tannin extracts

2

3 Aude A. Watrelot¹, Christine Le Guernevé¹, Hélène Hallé¹, Emmanuelle Meudec¹, Frédéric
4 Véran¹, Pascale Williams¹, Bertrand Robillard², François Garcia¹, Céline Poncet-Legrand¹,
5 Véronique Cheynier^{1*}

6

7 ¹SPO, INRAE, Univ Montpellier, Montpellier SupAgro, 2 Place Pierre Viala, 34060 Montpellier,
8 France.

9

10 ²Institut Œnologique de Champagne, Z.I. de Mardeuil, BP 25, 51201 Epernay cedex, France

11

12 *corresponding author:

13 V. Cheynier

14 Phone: +33 (0)4 99 61 22 98

15 Fax: +33 (0)4 99 61 28 57

16 veronique.cheynier@inra.fr

17

¹present address: Iowa State University, Department of Food Science and Human Nutrition, 536 Farm House Lane, Ames, IA 50011-1054, USA. (watrelot@iastate.edu)

18 ABSTRACT

19 Gallotannins extracted from gallnuts are commonly added to wine to improve its properties. They
20 consist of mixtures of galloylester derivatives of glucose. However, their composition and
21 properties are not well established. In this study, methods based on liquid chromatography coupled
22 to UV-visible detection and mass spectrometry, size exclusion chromatography and 1D (^{31}P) and
23 2D (^1H DOSY, ^{31}P TOCSY, $^1\text{H}/^{13}\text{C}$ HSQC, HMBC) NMR spectroscopy have been implemented
24 for extensive chemical characterization of three commercial gallnut tannin extracts. Differences in
25 the proportions of the different constituents (gallic, digallic, trigallic acids, galloylglucose
26 derivatives) and in the structure and molecular weight distributions of gallotannins were
27 demonstrated between the three extracts, with chains containing 8.5, 12.2, and 12.4 galloyl groups
28 in average for TAN A, B1, and B2, respectively. The antioxidant capacities of the extracts,
29 evaluated using the ABTS method, were similar and related mostly to their total tannin content,
30 with only limited impact of the tannin composition.

31

32 Keywords: enological tannins, gallotannins, polygalloylglucose, ^{31}P NMR, ^1H , 2D $^1\text{H}/^{13}\text{C}$, ^1H
33 DOSY NMR, UPLC-DAD-ESI-MS, size exclusion chromatography, molecular weight
34 distribution, antioxidant properties.

35

36

37

38

39

40 INTRODUCTION

41 Oenological tannins are extracted from different botanical sources with water and/or alcohol. They
42 are commonly used in winemaking to stabilize color, to facilitate the clarification of wine and
43 musts, promote partial precipitation of proteins and prevent protein haze, and contribute to the
44 antioxidant protection of wines, but they must not modify the wine color nor bring further aroma,
45 according to the International Oenological Codex (OIV-Oeno 613-2019;
46 <http://www.oiv.int/public/medias/6881/oiv-oeno-613-2019-en.pdf>). Among the large polyphenol
47 group, tannins are defined by their capacity to precipitate alkaloids and proteins.¹ They comprise
48 condensed tannins which are flavanol oligomers and polymers found in fruits such as grape, apple,
49 or persimmon, as well as in tea and quebracho, and hydrolysable tannins. The latter are subdivided
50 in two categories, both composed of a polyol core and galloyl groups. Gallotannins, extracted from
51 gallnuts of *Quercus infectoria* (Turkish or Alep gall) or *Rhus semialata* (Chinese gall) and from
52 tara (*Caesalpinia spinosa*) consist of polygalloyl esters of glucose and quinic acid, respectively,
53 whereas ellagitannins, found in oak and chestnut, contain hexahydroxydiphenoyl (HHDP) units
54 formed by oxidative linkage of galloyl groups.^{2,3}

55 According to the international code of enological practices
56 (<http://www.oiv.int/public/medias/6558/code-2019-en.pdf>), tannin extracts sold as enological
57 tannins must contain 65 % polyphenols. In the earlier versions of the code, the method imposed
58 for estimation of tannin content was based on spectrophotometry measurement at 280 nm,
59 converted to gallic acid equivalent, but this has been recently replaced by gravimetric analysis of
60 the water soluble material adsorbed on a polyvinylpyrrolidone (PVPP) solid phase extraction
61 column. PVPP is a cross-linked synthetic polymer of polyvinylpyrrolidone known to bind
62 polyphenols through mechanisms similar to protein tannin interactions. This method thus relies on

63 the assumption that adsorption by PVPP is specific of tannins, like other common assays based on
64 precipitation with protein such as bovine serum albumin⁴ or with methylcellulose⁵. These
65 properties are related to the use of tannins in preventing protein haze⁶ and to astringency
66 perception,^{7,8} which is attributed to formation of insoluble complexes between tannins and salivary
67 proteins.⁹⁻¹¹

68 Comparison of the official OIV methods performed in a recent study confirmed that they provide
69 different results and all lack specificity.⁶ Liquid chromatography analysis of gallic acid or methyl
70 gallate released after acid hydrolysis¹² or methanolysis,¹³ respectively, is also commonly used to
71 quantify gallotannins. However, these methods do not allow the determination of gallotannin
72 composition in commercial tannin extracts.

73 Gallotannins extracted from Chinese galls were shown to be composed of a pentagalloylglucose
74 core to which three or four galloyl groups were linked by depsidic linkages, forming up to a
75 nonagalloylglucose.¹⁴ The presence of longer polygalloyl chains in the tannin structure has also
76 been suggested.^{14,15} Chromatographic analysis provided evidence of heterogeneous mixtures of
77 gallotannins composed of a glucose linked to up to 12 galloyl groups in Chinese galls,¹⁶⁻¹⁸ and up
78 to 8 in Turkish galls,^{16,18} The latter was described as a mixture of galloylglucoses based on
79 tetragalloylglucose cores¹⁹ or on both tetragalloylglucose and pentagalloylglucose cores.²⁰
80 Moreover, depside gallic acid oligomers up to the octamer have also been detected in a commercial
81 gallnut extract.¹²

82 Gallotannins from gallnut extracts thus seem to be very polydisperse in a given extract and the
83 chemical structure of these tannins as well as the position of linkages between glucose and gallic
84 acid residues is not fully identified and characterized. In addition, different gallnut extracts differ
85 in their tannin composition and thus potentially in their properties. Indeed, protein precipitation

86 capacity was found to vary with the composition of the gallotannin mixture and generally to
87 increase with the degree of galloylation.^{21,22} Moreover, pentagalloylglucose is highly antioxidant
88 compared to a monogalloylglucose, suggesting that the antioxidant activity increases with the
89 number of galloyl units.^{23,24} However, gallic acid itself showed higher antioxidant activity than
90 pentagalloylglucose.²⁵ In a recent study, measurement of the antioxidant activity by diverse
91 methods showed 2 to 3 fold differences between five commercial gallnut extracts but their
92 chemical composition was not elucidated and no structure-activity relationship was identified.⁶

93 The aim of this study is to extensively characterize three commercial gallnut tannin extracts using
94 a combination of analytical methods as well as determine their antioxidant properties. UPLC-
95 DAD-MS analysis enabled detection of the species present in the sample. Structural and
96 quantitative information was obtained by 1D ¹H, ³¹P NMR and 2D ¹H/¹³C spectroscopy and size
97 distribution was obtained by 2D ¹H DOSY NMR spectroscopy and gel permeation
98 chromatography (GPC). The antioxidant capacities of the gallnut tannin extracts were measured
99 by an ABTS method.

100

101 MATERIAL AND METHODS

102 Chemicals

103 Acetonitrile, water, methanol, acetic acid and formic acid were LC-MS grade from VWR ProLabo
104 (Fontenay sous Bois, France). Gallic acid, 2,2'-azino-bis(3-ethyl-benzothiazoline-6-sulfonic acid)
105 diammonium salt (ABTS), L-tartaric acid, ethanol, sodium hydroxide, potassium persulfate, β-
106 glucogallin, 1,2,3,4,6-pentagalloylglucose (PGG), methyl gallate, deuterated dimethylsulfoxide
107 (DMSO-d₆), deuterated pyridine (pyridine-d₅), deuterated chloroform (CDCl₃), 2-chloro-4,4,5,5-

108 tetramethyl-1,3,2-dioxaphospholane (cTMDP), lithium chloride, and hydrochloric acid 37 % were
109 provided by Sigma-Aldrich (Saint Louis, MO, USA), and N,N-dimethylformamide by Carl Roth
110 GmbH (Karsruhe, Germany).

111 Three commercial gallotannin extracts (TAN A, TAN B1, TAN B2), available in the powder form
112 and sold to the wine industry with claims related to their antioxidant properties were analyzed.
113 Sample A was Tanal W2 (Ajioto Onmnicem, Wetteren, Belgium). Samples B1 and B2 correspond
114 to two different batches of the same commercial brand (Tan'Activ GC, SilvaTeam, San Michele
115 Mondovì, Italy). All three were stored in the dark in a desiccator until analysis.

116 **Neutral Sugar Characterization by Gas Chromatography**

117 Neutral sugar composition was determined as alditol acetates after hydrolysis, reduction and
118 acetylation of 1 mg of gallnut tannin extracts as previously described.²⁶ The alditol acetates were
119 quantified by gas chromatography with flame ionization detection (GC-FID) (GC 2010 Plus
120 Shimadzu) using a DB225 (30 m × 0.25 mm ID, 0.25 µm film) capillary column and hydrogen 5.6
121 B50 as the carrier gas. Allose and myo-inositol were used as internal standards for the neutral
122 sugar quantification.

123 **Compound Identification and Quantification by UPLC-DAD-MS**

124 UPLC-DAD-MS analyses were performed using an Acquity UPLC-DAD system (Waters,
125 Milford, MA) equipped with an Amazon X mass spectrometer (Bruker Daltonics, Bremen,
126 Germany), piloted by Hystar software. For direct analysis of the gallnut tannin extracts, they were
127 dissolved in triplicate in a solution of water: methanol (50:50, v/v) containing 1 % formic acid at
128 a concentration of 5 g/L and vortexed. The extracts were centrifuged at 15000 rpm (Mikro 220R,
129 Hettich, Germany) for 15 min at 15 °C prior to be injected (0.5 µL) to the UPLC system. The

130 gallnut tannin extracts were also submitted to an acidic methanolysis. 5 mg of extract were
131 dissolved in 2.5 mL of acidified methanol 0.6 N HCl. The methanolysis reaction was performed
132 at 120 °C for 260 min then samples were placed in ice to cool down to room temperature. The
133 samples were transferred into 5 mL volumetric flasks and the volume was adjusted with water.
134 The depolymerized samples were then injected to the UPLC system after centrifugation at 15,000
135 rpm for 15 min at 15 °C. The column was a BEH C18 (1.7 μm ×1 mm×150 mm) equipped with a
136 0.2 μm prefilter (Waters, Milford, MA) placed in an oven at 35 °C. The flow rate was 0.08 mL/min
137 and the run time was 45 min. The gradient conditions consisting in solvent A (water with 1 %
138 formic acid) and solvent B (acetonitrile with 1 % formic acid) was as follows: 0 min (1.3 % B),
139 6.5 min (10 % B), 12 min (20 % B), 27 min (50 % B), 32 min (99.5 % B), 40 min (1.3 % B). The
140 detection was both with the DAD from 250 to 600 nm and with the MS equipped with an
141 electrospray-ionization source and an ion trap analyzer. The MS detection was operated in the
142 negative ion mode with the full scan mode under the following conditions: a scan between m/z
143 150 to 2000, the target mass set at m/z 400, a capillary tension of 4500 V, the nebulizer gas (N_2)
144 at 10 psi, and N_2 as the desolvation gas at 5 L/min and 200 °C. The collision energy for
145 fragmentation used for the MS2 experiments was set at 1 eV.

146 The quantification of the constituents in the crude extracts and in the extracts after methanolysis
147 was based on the calibration curve of standards based upon the peak area at 280 nm. Digallic and
148 trigallic acids and mono-, di-, tri-, and poly-galloylglucose were expressed as gallic acid
149 equivalent. After methanolysis, methyl gallate and methyl di-gallate were expressed as methyl
150 gallate equivalent.

151 **Determination of Molecular Size Distribution by Gel Permeation Chromatography**

152 The molecular size distribution of gallnut tannin extracts was carried out by gel permeation
153 chromatography according to a previously published method²⁷ slightly modified as follows. The
154 extracts were dissolved in the mobile phase (2.5 g/L) consisting in dimethylformamide with 1 %
155 acetic acid (v/v), 5 % water (v/v) and 0.15 M lithium chloride, in triplicate. After vortex then
156 centrifugation at 15000 rpm for 15 min at 15 °C, 20 µL of sample were injected to the high-
157 performance liquid chromatography system. The Agilent HPLC 1260 Infinity II system equipped
158 with a diode array detector consisted in two Phenogel columns connected in series (300 mm × 7.8
159 mm, 5 µm 50 Å and 300 mm × 7.8 mm, 5 µm 1000 Å) protected by a guard column (50 mm × 7.8
160 mm) from the same material placed at 60 °C. The isocratic solvent was dimethylformamide
161 containing 1 % acetic acid (v/v), 5 % water (v/v) and 0.15 M lithium chloride filtered at 0.2 µm,
162 at a flow rate of 0.8 mL/min, and the run time was 35 min. The detection was carried out at 280
163 nm and commercial gallic acid, β-glucogallin, and 1,2,3,4,6-pentagalloylglucose were injected
164 under the same conditions and used as size distribution standards.

165 NMR Spectroscopy

166 All NMR experiments were conducted on an Agilent DD2 500 MHz spectrometer (Agilent
167 Technologies, Santa Clara, Calif., USA), operating at 500.02, 202.43 and 175.74 MHz for proton,
168 phosphorus-31, and carbon-13 nuclei respectively.

- 169 • ¹H 1D, 2D DOSY, and ¹H/¹³C 2D NMR analysis

170 1D and 2D experiments were performed using a 5mm indirect detection probe equipped with a
171 gradient coil, at 25°C, on tannin extracts (20 mg) dissolved in 500 µL DMSO-d₆. 1D ¹H and 2D
172 heteronuclear ¹H/¹³C HSQC and HMBC were obtained using classical pulse sequences. DOSY
173 measurements were carried out using the DgcsteSL pulse sequence of VNMRJ4.1 library. Data
174 were acquired with an array of 32 gradients with amplitudes ranging from 3.0 to 50.0 Gcm⁻¹ in

175 equal steps of gradient squared, a total diffusion encoding gradient of 3 ms and a diffusion time of
176 0.2 s, giving a residual signal intensity of about 5% of the original intensity. For each gradient
177 step, 512 transients were recorded with 64 K complex data points. Free induction decays were
178 zero-filled to 128k points and apodized with a combination of exponential and Gaussian functions
179 set at -0.8 and +0.8 Hz respectively to enhance spectrum resolution. After phase correction, 2D
180 DOSY spectra were constructed from peak height measurement, with a discrete multiexponential
181 fitting at two components and a correction for the effects of pulsed field gradient non-uniformity,
182 using the DOSY module of VNMRJ4.2 software.

183 • Phosphitylation procedure

184 The phosphitylation procedure was adapted from that described earlier.^{28,29} An amount of 1 to 5
185 mg of model compounds (gallic acid, β -glucogallin, pentagalloylglucose) depending on their
186 molecular weights or 20 mg of gallnut tannin samples were dissolved in 400 μ L of anhydrous
187 NMR solvent consisting in pyridine- d_5 : $CDCl_3$ (1.6: 1; v/v), containing cholesterol (100 mM) as
188 an internal standard (IS) for the absolute quantification and chromium acetylacetonate, Cr-(acac)₃
189 (3 μ M) as a relaxation agent allowing to shorten the relaxation times of the phosphorus nuclei.
190 After complete dissolution, 100 μ L of cTMDP was added dropwise under continuous agitation.
191 The mixture was then transferred to a NMR tube. For each tannin extract, three independent
192 phosphitylated samples were prepared and analyzed. ³¹P spectra were recorded using a 5mm direct
193 gradient probe.

194 • Quantitative 1D ³¹P spectra

195 1D ³¹P NMR spectra were obtained using a standard pulse sequence with a 90° pulse and inverse-
196 gated ¹H decoupling. Spin-lattice relaxation time experiment realized on phosphitylated tannin

197 extract has shown that the highest T1 value was 1.5 s (measured for the carboxyl group of gallic
198 acid), which was taken into account to set the recycle delay to 15 s, i.e. more than 7 times higher
199 than the longest T1 value, allowing a quite complete relaxation of spins between scans. For the
200 tannin samples, 1D spectra of 512 scans were recorded to quantify with enough precision small
201 peaks. The NMR spectra were processed using Mnova 14.1.0 software. The data were zero-filled
202 to 64 K. A line broadening of 0.1 Hz and a drift correction were applied prior to Fourier
203 transformation and phase correction. The baseline was corrected using a cubic spline function
204 before peak area integration. The phosphitylated hydroxyl signal area of cholesterol was used as
205 internal reference for the determination of the absolute concentrations. All chemical shifts were
206 reported relative to the reaction product of water with Cl-TMDP, which gave a sharp signal in
207 pyridine/CDCl₃ at 132.2 ppm.

208 • 2D ³¹P TOCSY NMR spectra

209 2D ³¹P TOCSY spectra were obtained with a classic TOCSY pulse sequence using a mixing time
210 of 80 ms, 128 increments of 1 K complex data points and 64 scans per increments. The data set
211 was zero-filled to 256 x 1 K points prior to Fourier transformation. A squared sinusoidal window
212 function was used in both dimensions.

213 **Determination of the Antioxidant Capacity using the ABTS Method**

214 The antioxidant capacity of the three extracts was determined using the ABTS method.³⁰ Briefly,
215 a solution of 7 mM ABTS was mixed with a solution of 2.45 mM potassium persulfate in ethanol:
216 water (1:1, v/v) to produce ABTS⁺. The solution was then placed at 22 °C for 24 hours, and then
217 diluted in ethanol: water (1:1, v/v) to produce an absorbance of 0.70 ± 0.02 at 734 nm. For tannin
218 analysis, 1 mL of diluted ABTS⁺ solution and 0.1 mL of tannin solution sample (0.01 g/L in

219 acidified hydroalcoholic solution (12% (v/v) ethanol, 33 mM L-tartaric acid and pH 3.6 whose
220 value was adjusted with 1 M NaOH)) were mixed and incubated for 6 min before measuring the
221 absorbance at 734 nm. The percentage of inhibition was calculated by the following equation for
222 each tannin extract sample:

$$223 \quad \% \text{ inhibition} = \frac{(A_{734} \text{ ABTS}^+ - A_{734} \text{ Tannin})}{A_{734} \text{ ABTS}^+} \times 100$$

224 Where A_{734} corresponds to the absorbance at 734 nm.

225 All the experimental measurements were performed in triplicate using a SAFAS UV mc2
226 spectrophotometer (Monaco).

227 **Statistical Analysis**

228 ANOVA type II sum of squares analyses were performed using RStudio, version 3.5.1, with the
229 FactoMineR, agricolae and car packages. A Tukey LSD test with a p value of 0.05 was performed
230 using XLStat for the ABTS results.

231

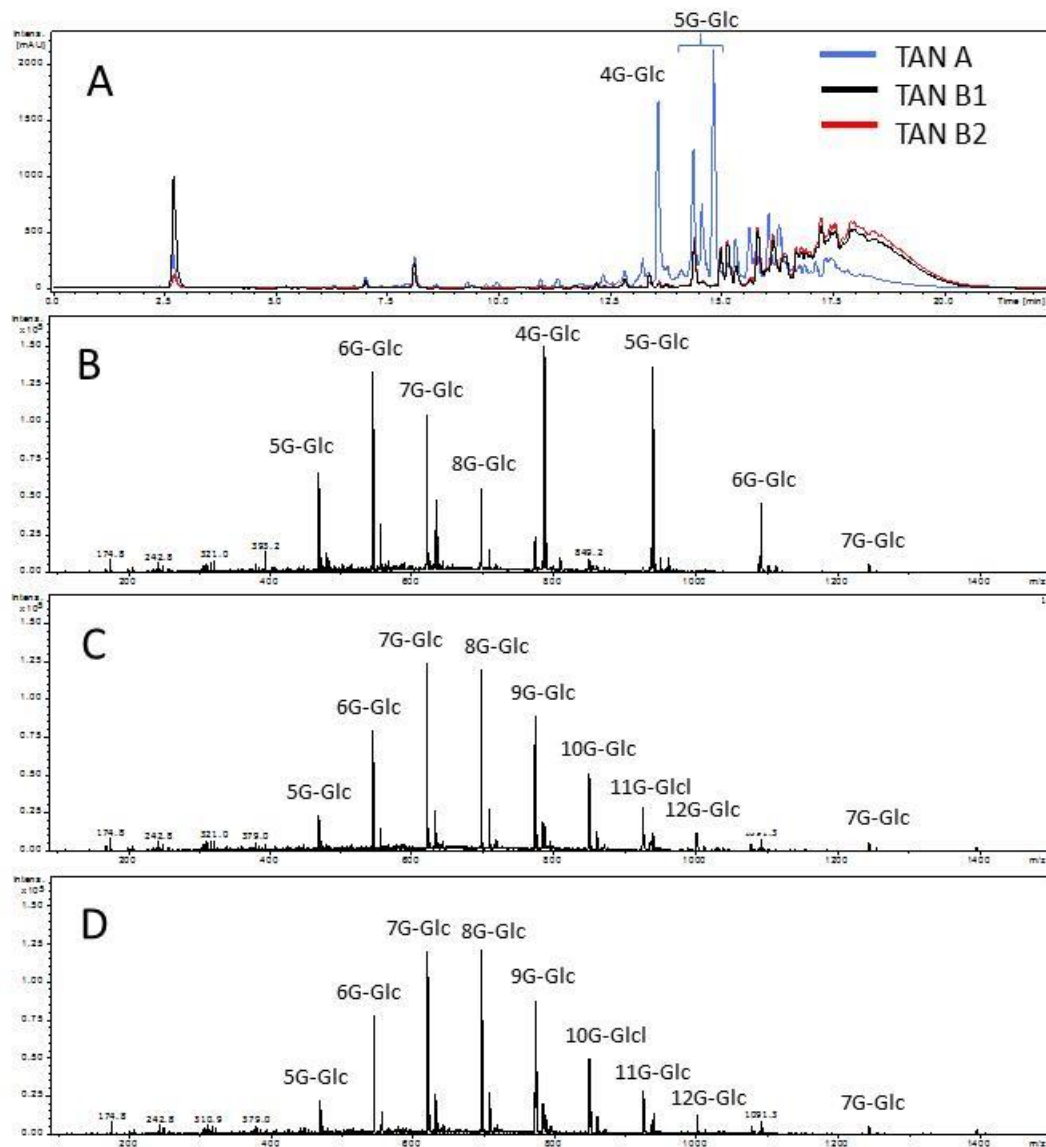
232 **RESULTS AND DISCUSSION**

233 **Identification and Quantification of the Constituents of Gallnut Tannin Extracts**

234 • UPLC-DAD-MS

235 Concentrations evaluated by integration of the total peak area at 280 nm after HPLC separation
236 were 1.06 ± 0.01 , 1.14 ± 0.02 , 1.13 ± 0.01 g gallic acid equivalent per g of extract (corresponding
237 to 790, 822 and 875 mg/g pentagalloylglucose equivalent per g of sample), for TAN A, B1 and
238 B2, respectively. The LC-UV profile of TAN A appeared very different from those of TAN B1
239 and B2, the former showing a number of well resolved peaks, while the latter are characterized by
240 a large hump eluted later (Figure 1). Such differences have been observed earlier by normal phase

241 HPLC analysis, showing that Turkish gall extracts are mixtures of penta and hexagalloylglucose,
242 while Chinese gall extracts contain larger gallotannins.^{16,18} Information on constitutive molecules
243 has also been obtained by mass spectrometry. In Turkish gall tannin, gallotannins composed of a
244 single glucose linked to up to 8 gallic acids were detected by MALDI-TOF analysis, with the most
245 abundant signals corresponding to tri- and tetra-galloylglucose.³¹ However, it should be
246 emphasized that relative intensities in the mass spectra do not reflect proportions of tannins of
247 different molecular weights, as peak intensities decrease when molecular weight increases³²⁻³⁴
248 which explains the discrepancy between LC-UV profiles and MS data. The most polar compounds
249 identified by LC-MS in all samples were gallic acid (m/z 169), digallic acid (m/z 321, two isomers
250 eluted at 7.2 and 8.5 min), and trigallic acid (m/z 473). Low intensity signals detected at m/z 331,
251 483, 635, 625, and 777 were attributed to monogalloylglucose (tentatively identified as β -
252 glucogallin from its retention time), digalloylglucose, trigalloylglucose, tetragallic acid and
253 pentagallic acid, respectively. Additional minor signals at m/z 495, 647 and 799 detected with
254 higher intensity in TAN B1 were tentatively attributed to di-, tri- and tetra-galloylquinic acids
255 which are present in Tara tannins but usually not detected in gallnut extracts.



256

257 *Figure 1. HPLC trace at 280 nm of TAN A (blue), TAN B1 (black) and TAN B2 (red) (A). Negative*
 258 *mode ESI MS spectra (150-1500 m/z) recorded after HPLC-DAD-MS analysis of TAN A (B), TAN*
 259 *B1 (C) and TAN B2 (D). G-Glc: galloyl-glucose*

260 The concentration of gallic acid was much higher in TAN B1 than in TAN A and TAN B2 while
 261 monogalloylglucose was found in the highest concentration in TAN A (Table 1). The
 262 concentrations of digallic acid and trigallic acids were also different between the three samples

263 with TAN A > TAN B1 > TAN B2, but digallic acid was overestimated by LC-UV since its major
264 isomer was coeluted with digalloylglucose.

265 Major peaks detected in TAN A were identified as tetragalloylglucose (eluted at 13.7 min and
266 detected as singly charged $[M-H]^-$ ion at m/z 787 and doubly charged $[M-2H]^{2-}$ ion at m/z 393) and
267 pentagalloylglucose ($[M-H]^-$ at m/z 939, $[M-2H]^{2-}$ at m/z 469, 3 isomers, eluted at 14.6, 14.8, and
268 15 min), followed by lower intensity peaks attributed to hexa- ($[M-H]^-$ at m/z 1091, $[M-2H]^{2-}$ at
269 m/z 545.5), hepta- (m/z 1243 and 621.5) and octa- ($[M-2H]^{2-}$ at m/z 697.5) galloylglucose. Some
270 of these compounds were also present in TAN B1 and TAN B2 while others, including
271 tetragalloylglucose, were found in significant amounts only in TAN A. The first
272 pentagalloylglucose isomer, eluted at 14.5 minutes, present in all samples and detected mostly as
273 the doubly charged ion at m/z 469, was identified as 1,2,3,4,6-penta-galloylglucose by comparison
274 of its retention time and spectral characteristics with those of the standard. The other two isomers,
275 specific of TAN A, were detected mostly as the singly charged ion at m/z 939. Their UV spectra
276 exhibited a shoulder at 300 nm, which is characteristic of depside bonds,³⁵ suggesting that they
277 are based on a tetra-galloylglucose core substituted by a fifth galloyl moiety through a depside
278 bond. The different pentagalloylglucose isomers also showed different fragmentation patterns.
279 Fragmentation of 1,2,3,4,6-penta-galloylglucose yielded fragments at m/z 769, and 617,
280 corresponding to successive losses of gallic acid and water molecules (-152-18) through cleavage
281 of the glucose ester bond. The same fragmentation was obtained from the ion at m/z 787 in TAN
282 A (617, 787-152-18), consistent with the tetragalloylglucose core. In contrast, the major fragment
283 ion obtained from the other pentagalloylglucose isomers was at m/z 787 (-152), arising from
284 cleavage of a depside bond. Similarly, the tetragalloylglucose isomers present in trace amounts in

285 TAN B1 and B2 yielded a fragment ion at m/z 635 (787-152), suggesting that they are mono-, di-
286 , or tri-galloylglucose derivatives containing depside bonds.

287 Signals corresponding to hexa- and hepta- galloylglucose were also more polydisperse and less
288 polar in TAN A than in TAN B, indicating that they are based on different structures. Signals
289 corresponding to larger galloylglucose polymers (from 8 to 10 galloyl groups) were more
290 polydisperse in all samples, reflecting a diversity of substitution patterns.

291 The total concentration of gallotannins determined by UPLC-UV was significantly higher in TAN
292 B2 than in B1 and A (Table 1). The proportion of lower molecular weight oligomers (tetra- and
293 penta-galloylglucose) was much higher in TAN A (Fig 1A) while TAN B1 and B2 contained
294 mostly higher molecular weight tannins eluted as a hump at the end of the chromatogram. The MS
295 spectrum of the hump showed singly charged, doubly charged and triply charged ions that
296 correspond to a polymeric series consisting of a glucose molecule linked to up to twelve (m/z 1001
297 $[M-2H]^{2-}$), eighteen (m/z 1457 $[M-2H]^{2-}$) and seventeen (m/z 1381 $[M-2H]^{2-}$) galloyl units in TAN
298 A, TAN B1, and TAN B2, respectively (Figure S1 in the Supporting Information). Another set of
299 masses ($\Delta 22$ amu, $\Delta 11$ amu, respectively for monocharged and doubly charged ions) which likely
300 correspond to sodium adducts was observed with lower intensity for each galloylglucose polymer.

301 • UPLC-DAD-MS after methanolysis

302 After methanolysis under acidic conditions, the majority of gallic acids involved in gallotannins
303 formed methyl gallate (around 600 mg/g) and methyl digallate (around 50 mg/g) (data not shown).
304 Detection of methyl digallate and of m/z values corresponding to di-, tri- and tetra- and penta-
305 galloylglucose (m/z 483, 635, 787, 939) indicated that methanolysis was incomplete. Moreover,
306 di-, tri- and tetra- and penta-galloylglucose methyl esters (m/z 345, 497, 649, 801, 953) were also

307 detected, showing that some glucose was linked to gallic acid through ether rather than ester bonds,
308 as proposed earlier.^{14,36} Further analysis of the LC-MS data showed that several isomers of the
309 methyl esters of galloylglucose ethers were present in all samples. Their proportions in TAN A
310 were different from those observed in TAN B1 and B2. Moreover, these structures were
311 significantly more abundant in TAN B than in TAN A and in TAN B2 than in TAN B1 (Figures
312 S2 and S3 in supporting information).

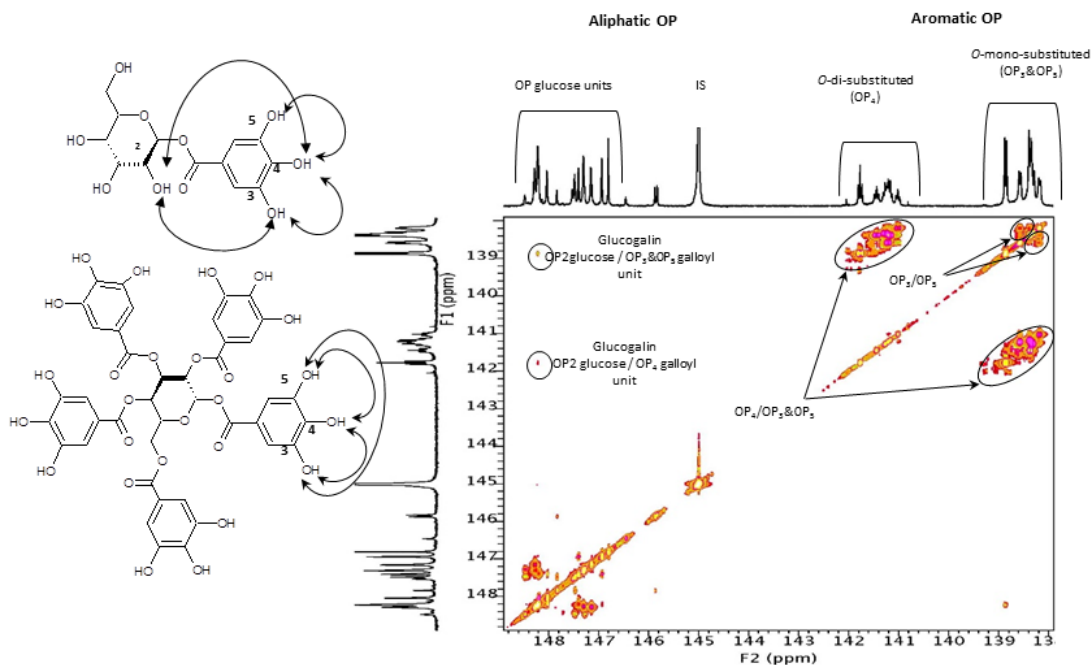
313 • NMR Spectroscopy

314 Further results upon composition of the tannin extracts were obtained from ³¹P NMR spectroscopy
315 analysis after phosphitylation of the free hydroxylic groups. This method has been shown to give
316 both qualitative and quantitative information on samples from various origins such as tannins,
317 lignins, olive oils, biomass and the chemical shifts of functionalized groups are well
318 documented.^{28,29,37-42} It has also provided evidence of both *meta* and *para*-depside linkages in
319 Chinese and Turkish gall extracts.²⁹ ³¹P NMR signals arising from the three classes of
320 phosphitylated functional groups appear in three well-separated chemical shift regions: the
321 aliphatic hydroxyl groups are observed in the most deshielded region between ~146-149 ppm, the
322 carboxyl groups give the most shielded signals, between ~134.5-135.5 ppm, and the aromatic
323 hydroxyl groups cover a wider range between ~137-143 ppm.

324 More precise assignment of the aromatic hydroxyl peaks was performed with the help of ³¹P
325 TOCSY scalar correlations using a mixture of gallotannin standards (Figure 2). P-P long-range
326 scalar correlations are indeed observed between the mono-*ortho*-substituted non-equivalent OP3 /
327 OP5 located in the sub-region ~137.5 to 138.7 ppm and between the mono-*ortho*-substituted OP3
328 / OP5 and the di-*ortho*-substituted OP4 found between ~140.5-141.5 ppm. These attributions are
329 in agreement with those indicated by Melone et al.²⁸ As expected, phosphitylated glucose hydroxyl

330 peaks appear in the most deshielded region between ~146.0-148.2 ppm. It is worth noting that
331 correlations occurred between glucose OP2 and aromatic phosphitylated hydroxyls OP3/OP5 and
332 OP4 of β -glucogallin.

333



334

335 *Figure 2: ³¹P 1D and 2D TOCSY NMR spectra of a mixture of glucose, β-glucogalin and PGG in*
 336 *pyridine-d₅:CDCl₃ after phosphitylation. The numbering system of phenolic compounds and the*
 337 *TOCSY correlations are shown on the left.*

338

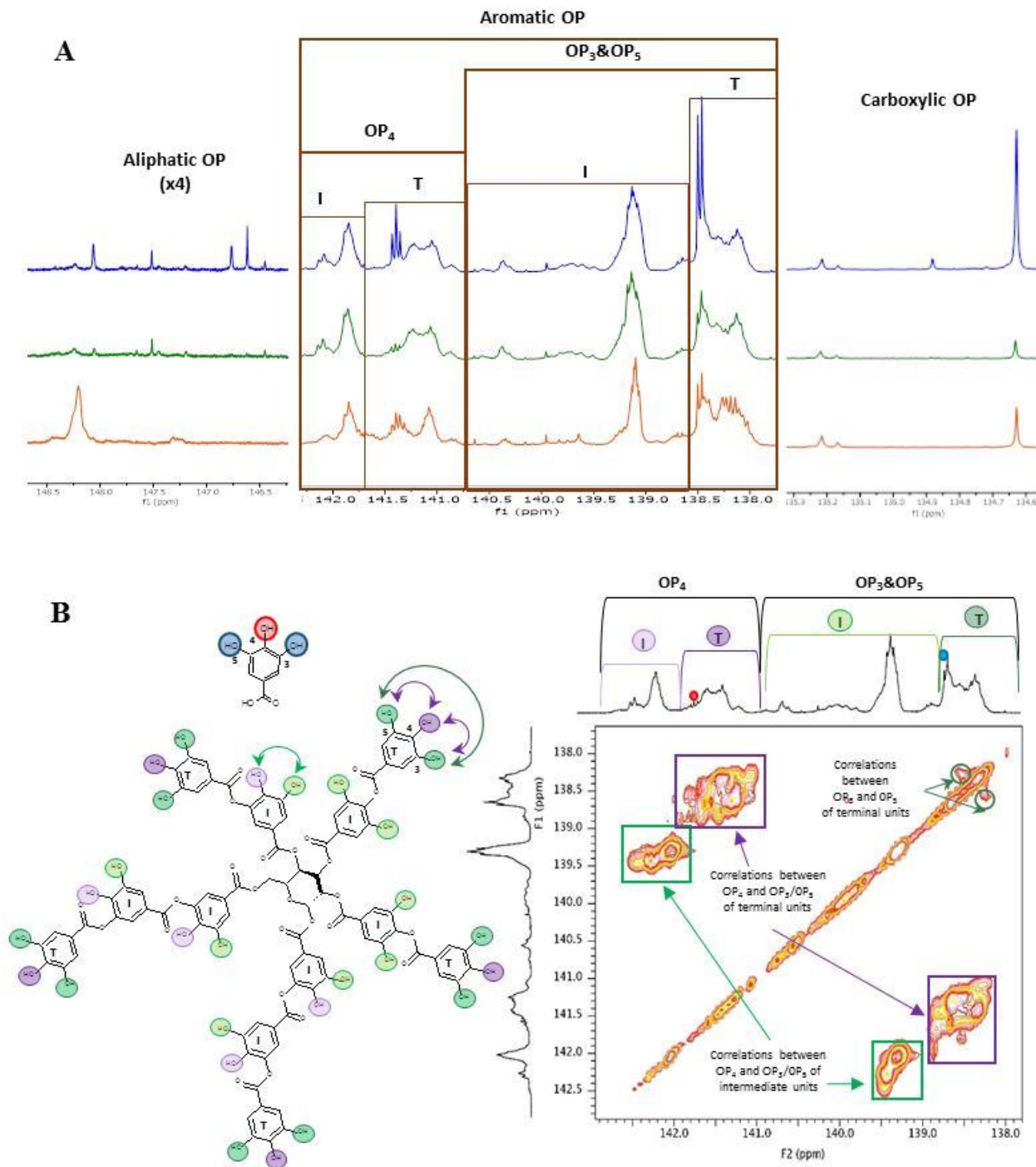
339 The aromatic region of the ³¹P NMR spectra of gallnut tannin extracts after phosphitylation
 340 reaction was much more crowded (Figure 3 A). Signals were attributed using both chemical shifts
 341 and ³¹P TOCSY correlations (Figure 3 B). Terminal unit OP₃ / OP₅ and OP₄ were easily attributed
 342 since their chemical shifts are similar to those of gallotannin standards (between ~137.8 to 138.6
 343 ppm and ~140.7-141.55 ppm respectively) and TOCSY correlations are observed between non-
 344 equivalent OP₃ / OP₅ and between OP₃ / OP₅ and OP₄. Additional more deshielded signals found

18

345 in the aromatic region from ~138.6 to 140.7 ppm and from 141.55 to 142.4 ppm were
346 unambiguously attributed respectively to OP3 /OP5 and OP4 of internal units, in accordance to
347 the presence of correlations between these two groups of signals. It is worth noting that the integral
348 area of internal OP3/OP5 signals exceeded that of internal OP4 signals, meaning that some of the
349 former correspond to OP3/OP5 of internal units substituted in the 4 position. In the chemical shift
350 range of carboxylic groups, three peaks were observed in all samples. Beside the carboxylic peak
351 of free gallic acid easily identified at 134.45 ppm, two other peaks appeared at 135.02 and 135.07
352 ppm. A fourth peak was also present in TAN B1 spectrum at 134.73 ppm (Figure 3 A). The peaks
353 arising from the aliphatic groups were very different depending on the tannin extract. The ³¹P
354 spectrum of TAN A exhibited mainly a broad and intense phosphitylated aliphatic peak whereas
355 the peaks in this region in TAN B1 and B2 spectra were numerous and sharper, with higher
356 intensity in TAN B1 than in TAN B2.

357 The signal areas were integrated to get more insight on the composition of the tannin extracts
358 (Table 2). Free gallic acid concentrations were about 0.1, 0.36 and 0.04 mmol/g (17, 61, 7 mg/g)
359 in TAN A, TAN B1, and TAN B2, respectively, consistent with the values determined by UPLC-
360 DAD (23, 73, and 9 mg/g). The other carboxyl signals at 135.7 and 135.02 ppm integrated for
361 lower concentrations (about 0.01 to 0.04 mmol/g) likely arise from the di and tri-gallic acids also
362 detected by UPLC-DAD-MS. Concentrations calculated for trigallic acid (8.5, 7.2 and 4.2 mg/g,
363 in TAN A, B1 and B2, respectively) are consistent with those determined by HPLC (Table 1).
364 Concentrations calculated for digallic acid (14.2, 13.5 and 7.4 mg/g, in TAN A, B1 and B2,
365 respectively) are lower than those determined by HPLC, due to coelution of digallic acid with
366 digalloylglucose (Table 1).

367



368
 369 *Figure 3: (A) ^{31}P 1D NMR spectra of TAN B1 (blue), TAN B2 (green) and TAN A (red) in pyridine-*
 370 *d_5 : CDCl_3 after phosphitylation. (I= Internal units, T=Terminal units). (B) ^{31}P 1D and 2D TOCSY*
 371 *NMR spectra of TAN B2 in pyridine- d_5 : CDCl_3 after phosphitylation, region of aromatic signals.*

372 *The numbering system of polyphenol compounds and the TOCSY correlations are shown on the*
373 *left.*

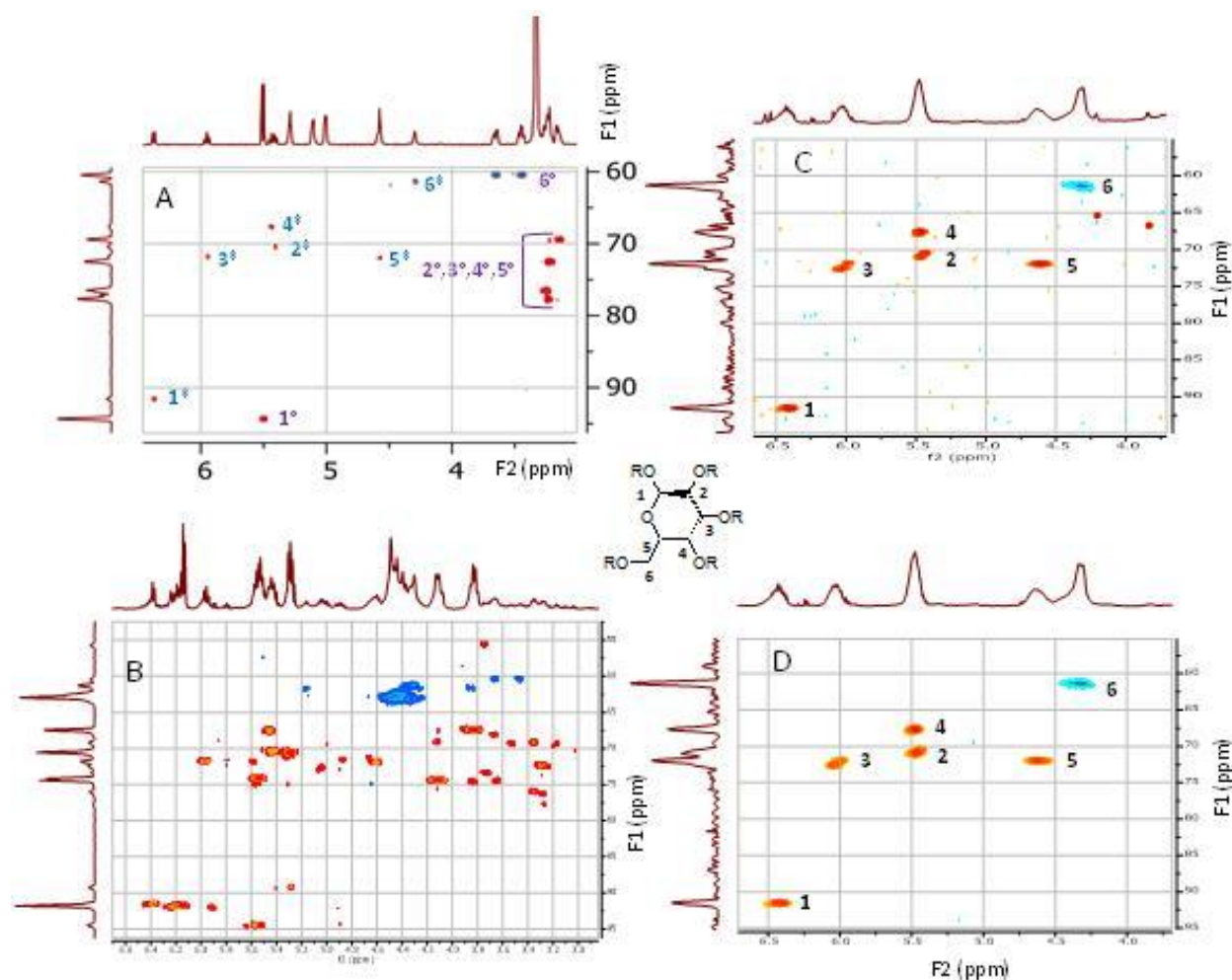
374

375 Terminal and internal units were quantified from the integration of phosphitylated phenolic
376 signals. The molarity of terminal units was calculated from the sum of OP4 and OP3 & OP5 signal
377 integrations divided by three whereas that of intermediate units was determined from the sum of
378 OP4 and OP3 & OP5 signal integrations divided by two to take into account the presence of both
379 *para*- and *meta*-depside bonds. The molarity of *meta*-depside bonds is directly obtained from the
380 integration value of the internal unit OP4 signals and that of *para*-depside bonds is deduced from
381 the difference between the concentration of intermediate units and that of *meta*-depside bond
382 internal units. The percentage of *para*-depside bonds calculated for TAN B samples was ~39%
383 whereas it was only about 21% for TAN A. Another difference between these extracts was the
384 ratio between the internal and the terminal units of the poly-galloylglucose molecules which was
385 close to 1.1 for TAN A whereas it was about 1.5 for both TAN B extracts. This means that the
386 gallotannin molecules of TAN A were mainly constituted of glucose substituted by chains of two
387 galloyl units, whereas the TAN B extracts contain longer galloyl chains. The molarity calculated
388 from the aliphatic signals of TAN B extracts was very low (less or equal to 0.01mmol/g), meaning
389 that glucotannin molecules of these two tannins were mainly constituted of penta-O-substituted
390 glucose moieties. In contrast, the intense broad signal observed in TAN A spectrum integrates for
391 about 0.5 mmol/g, indicating that the glucose hydroxyls were only partially substituted by galloyl
392 units. The ratio between the poly-galloylglucose terminal units and the hydroxyl glucose
393 concentrations was close to four, confirming that the gallotannin molecules of TAN A were mainly
394 constituted by tetra-O-substituted glucose moieties, in agreement with the UPLC-DAD-MS data.

395 The average number of gallic units per glucose residue was also calculated from the sum of the
396 total unit concentration divided by the concentration of the terminal units, then multiplied by the
397 number of substituted glucose hydroxyls, that was four in the case of TAN A and five for TAN B1
398 and B2. It provided a mean number of galloyl units per glucose moiety of 8.5, 12.2 and 12.4 for
399 TAN A, TAN B1 and TAN B2, respectively, consistent with chains of two galloyl units in TAN
400 A and longer in TAN B. The total concentrations of gallotannins calculated from the ^{31}P NMR
401 data were 725, 749 and 780 mg/g in TANA, TAN B1, and TANB2, respectively, amounting to
402 765, 829, and 799 mg/g, when taking into account gallic, digallic and trigallic acids.

403 1D ^1H and 2D $^1\text{H}/^{13}\text{C}$ analysis performed on the three samples confirmed the results obtained by
404 ^{31}P NMR. 2D $^1\text{H}/^{13}\text{C}$ HSQC spectra of the three tannin extracts showed typical chemical shifts of
405 galloyl and glucose units. As expected, ^1H and ^{13}C aromatic signals appear in more deshielded
406 regions than those of PGG (not shown) which indicate the presence of galloyl units linked by
407 depside bonds.²⁵ Aliphatic signals in the HSQC spectra of TAN B1 and B2 are characteristic of
408 glucoside moieties substituted at their five hydroxyl positions as observed for PGG (Figure 4 C,
409 D).

410



411

412 *Figure 4. Aliphatic region of 1D ^1H and 2D HSQC $^1\text{H}/^{13}\text{C}$ spectra. (A) Standard mixture of β -*
 413 *glucogallin ($^\circ$) and PGG (*); (B) Tan A; (C) TAN B1; (D) TAN B2. Attributed protons and carbons*
 414 *of glucoside moieties are noted in the spectra according to the glucoside numbering scheme.*

415

416 In contrast, aliphatic signals of TAN A (Figure 4 B) are much more numerous and dispersed in
 417 accordance with the presence of glucose core units partly and non-uniformly substituted. In the
 418 2D $^1\text{H}/^{13}\text{C}$ HMBC spectra (not shown), most of the carbonyl signals are located in the region
 419 characteristic of ester groups (~ 160 - 166 ppm). Four COOH groups are observed between 166.5-
 420 168ppm in TAN B1 spectrum and three in both TAN B2 and TAN A spectra in accordance with

23

421 signals detected in the ^{31}P spectra. The COOH signal of free gallic acid is easily identified by
422 comparison with a standard. Due to the low intensity of the other COOH peaks and the spectra
423 crowding, it was not possible to get more insights upon the structure of the molecules
424 corresponding to these minor peaks.

- 425 • GC analysis of neutral sugars

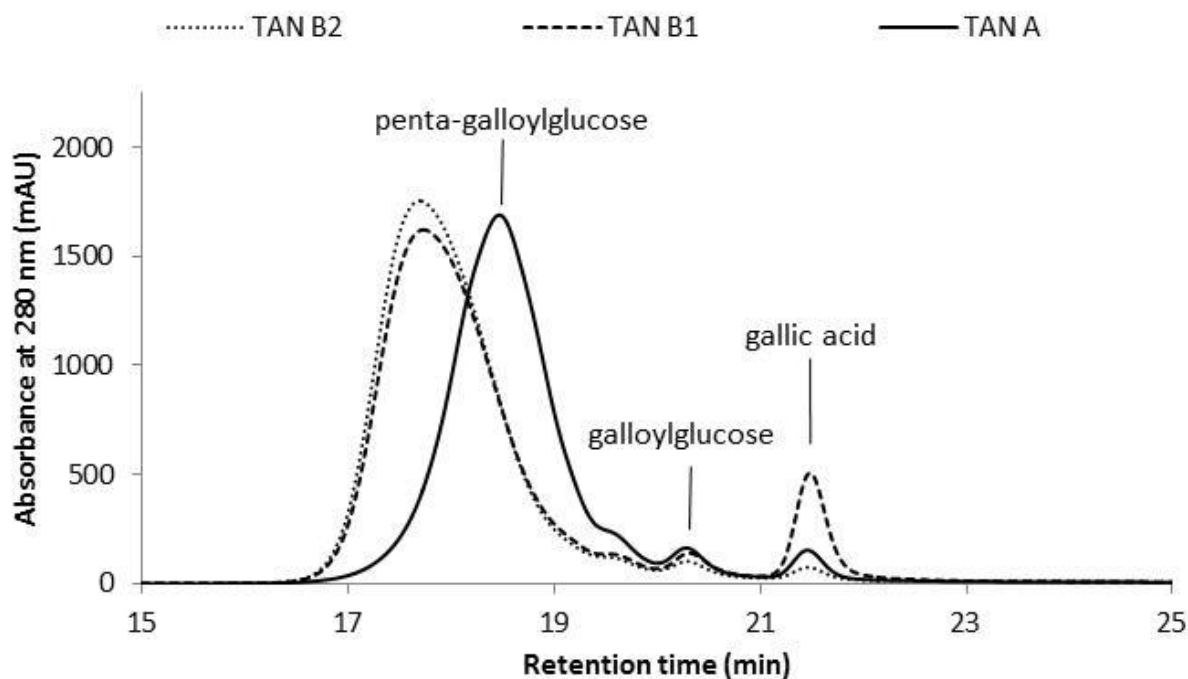
426 Glucose, arising from galloylglucose derivatives, was the only monosaccharide detected by GC
427 analysis in all samples. Its concentration was evaluated at 47.3 ± 2.2 , 23.4 ± 5.2 , and 16.1 ± 1.4
428 mg/g respectively in TAN A, TAN B1 and TAN B2.

429

430 **Characterization of the Molecular Weight Distribution of Tannins**

- 431 • Gel Permeation Chromatography

432 Gel permeation chromatography performed on the three gallotannin extracts showed the presence
433 of three peaks detected at 280 nm (Figure 5). The last one, eluted at 21.5 min, corresponds to the
434 retention time of gallic acid. Its peak area was higher in TAN B1 than in the two other tannin
435 extracts, as observed by LC-DAD-MS and NMR. The second peak, also present in small amount
436 in the three extracts, coeluted with β -glucogallin (20.3 min) and thus likely corresponds to digallic
437 acid as β -glucogallin was detected by LC-MS only in trace amounts. The larger peak detected in
438 TAN A showed the same retention time as pentagalloylglucose (18.5 min), while the major peak
439 detected in TAN B1 and B2, was eluted much earlier (17.7 min), thus showing a molecular size
440 larger than that of pentagalloylglucose. These GPC profiles are similar to those reported earlier for
441 Turkish and Chinese gall extracts.²⁹ The peak intensity was slightly higher in B2 than in B1 but
442 the shapes looked similar, suggesting similar size distributions.



443
 444 *Figure 5 Molecular size distribution of gallnut tannin extracts (TAN A and TAN B2 vs. B1 lots at*
 445 *2.5 g/L) detected at 280 nm by gel permeation chromatography. The retention times of gallic acid,*
 446 *β -glucogallin, and pentagalloylglucose standards are reported on the chromatogram.*

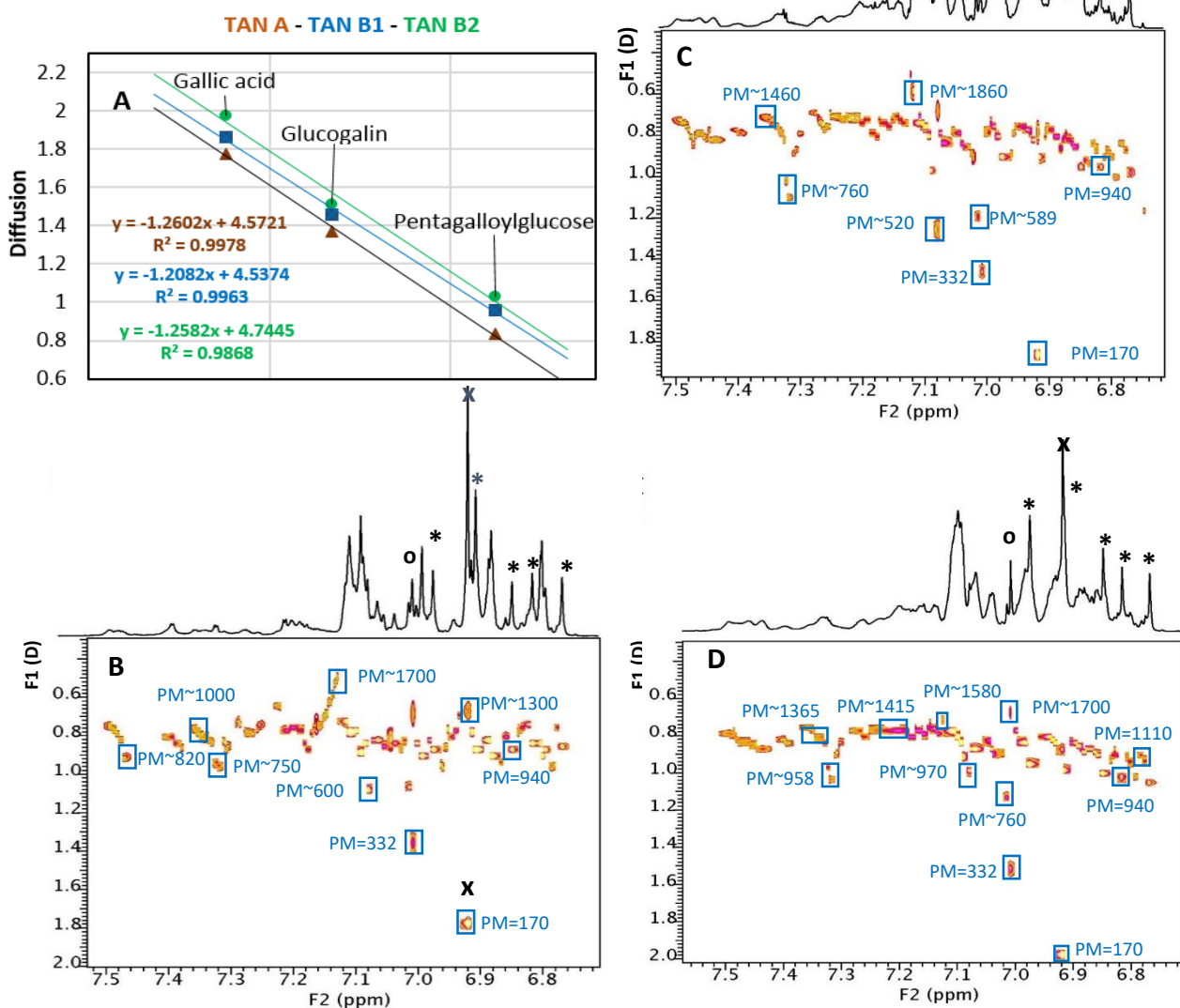
447

- 448 • ^1H DOSY NMR

449 2D ^1H DOSY spectra which separate NMR signals of compounds in mixtures according to their
 450 diffusion coefficients have been performed to get additional information on the polydispersity of
 451 the molecules in the tannin extracts. These data also allowed calculation of the molecular weight
 452 of unknown molecules using a calibration curve obtained from gallotannin standards (gallic acid,
 453 β -glucogallin, pentagalloylglucose) added to the extracts in solution. If peaks are intense and
 454 resolved enough, molecular weight of small gallotannin molecules (from ~170 to ~940) can be
 455 calculated with accuracy since the correlation coefficient of the calibration curve obtained from

456 the logarithm of diffusion coefficients ($\log D$) and the logarithm of the molecular weights (\log
 457 PM) of the standards is > 0.9 (Figure 6 A). Molecular weights of larger polymers were also
 458 calculated from this regression line but with uncertainties as no reference molecule of higher
 459 molecular weight gallotannins was available.⁴³ Molecular weights evaluated from the diffusion
 460 coefficients (Figure 6 B, C, D) ranged broadly from 500 (~digalloylglucose) to 1860 (~11
 461 galloylglucose) for all tannin extracts.

462



463

464

465

466

467

468

469

470 *Figure 6. A, calibration curves obtained from gallotannin standards (gallic acid, glucogallin and*
471 *pentagalloylglucose) added to the extracts in solution. Aromatic region of ¹H 2D DOSY NMR*
472 *spectra of TAN A (B), TAN B1 (C) and TAN B2 (D) extracts with added gallotannin standards in*
473 *DMSO d6 (x: gallic acid, o : β-glucogallin, * : pentagalloylglucose signals).*

474

475 Taken together, these data indicate that the three gallnut tannin extracts contain similar levels of
476 gallic acid derivatives but have different compositions, as summarized in Table 3. The
477 concentrations determined by UPLC-DAD using gallic acid as the calibration standard were higher
478 than 1000 mg/g in all three samples and thus overestimated. However, when expressed as
479 pentagalloylglucose equivalent, the concentrations determined by UPLC-DAD were only slightly
480 higher than those determined by ³¹P NMR spectroscopy after phosphorylation, and very good
481 congruence was obtained between the two methods both for the total gallic acid derivatives and
482 for gallotannins.

483 The large signal corresponding to an aliphatic OH group in the ³¹P NMR spectrum as well as
484 UPLC-DAD-MS data showing the presence of tetragalloylglucose and its derivatives indicate that
485 TAN A is mostly based on a tetra-O-substituted glucose core while TAN B1 and B2 are
486 pentagalloylglucose derivatives. These structural characteristics have been described earlier for
487 Turkish and Chinese gallnut extracts, respectively.^{29,44} LC-MS, GPC, and DOSY NMR analysis
488 enabled detection of a large range of molecular weights in all tannin samples. GPC analysis showed
489 that polymers were larger in TAN B1 and B2 than in TAN A, in agreement with the average
490 molecular weight calculated from integration of the ³¹P NMR data (average number of galloyl
491 groups evaluated at of 8.5, 12.2 and 12.4 respectively, for TAN A, TAN B1, and TAN B2). The
492 average chain length determined for TANA confirms the results obtained earlier for Turkish

493 gallnut extracts.²⁹ To the best of our knowledge, that of Chinese tannins is reported here for the
494 first time. The average degrees of galloylation of 12.2 and 12.4 determined for TAN B1 and TAN
495 B2, as well as the detection by UPLC-MS of galloylglucose structures containing up to 12, 18 and
496 17 galloyl groups, respectively in TAN A, TAN B1, and TAN B2, indicate the presence of long
497 polygalloyl chains in the tannin structure, as proposed earlier.^{14,15} This is also supported by
498 previous work showing the presence of gallic polymers up to the octamer,¹² likely arising from
499 cleavage of the glucose ester bonds of the original galloylglucose structure. Nevertheless, higher
500 molecular weight tannins were poorly detected by UPLC-ESI-MS, as shown earlier for
501 proanthocyanidin polymers³³ and by DOSY NMR. Sugar analysis by GC after hydrolysis and
502 derivatization also indicated that TAN A contains higher glucose contents than TAN B, consistent
503 with the lower degree of galloylation shown by the other methods. However, values obtained using
504 this method were not coherent with those determined by NMR. Reaction yields of 72% and 34%
505 were obtained respectively for β -glucogallin and pentagalloylglucose standards, confirming that
506 the method is not suitable for glucose ester derivatives and especially polymers.

507 Some differences in the composition of TAN B1 and B2 were also demonstrated, suggesting
508 differences in the preparation of these two batches. Thus, TAN B1 contained significantly higher
509 concentration of gallic, digallic, and trigallic acids and lower amounts of gallotannins than TAN
510 B2. It also contained an unknown low molecular weight aliphatic compound detected by ¹H DOSY
511 NMR (data not shown). The UPLC-DAD-MS data also showed significantly higher intensities of
512 tri- and tetra-galloylglucose and of mass signals attributed to minor compounds such as
513 galloylquinic derivatives in TAN B1.

514 Finally, the structures appeared more complex than expected. In particular, NMR provided
515 evidence of para-substituted gallic acid units in addition to the well-established meta-depside

516 structures. These structures were more common in TAN B than in TAN A, which may be related
517 to their longer galloyl chains. Detection of methyl galloylglucose derivatives by UPLC-MS
518 analysis after methanolysis indicated that some of the glucose units were linked to gallic acid
519 through ether rather than ester bonds, as proposed earlier.^{14,36} Different isomers of these structures
520 were predominant after methanolysis of TAN A and TAN B, reflecting further structural
521 differences.

522 **Antioxidant Capacity**

523 The antioxidant capacity of the three extracts was measured by the ABTS method. The percentage
524 of inhibition was significantly lower for TAN B1 at 35.82 ± 0.63 compared to TAN A and B2
525 (39.95 ± 0.77 and 38.84 ± 0.63 , respectively). This might be related to differences in the
526 proportions of gallotannins and lower molecular weight compounds such as gallic acid, reflecting
527 some gallotannin degradation during the tannin production process in TAN B1. However, the large
528 structural differences observed between TAN A and TAN B2 did not impact their antioxidant
529 properties measured by the ABTS method, which seem mostly related to the total tannin content.

530 The three tannin extracts analyzed in the present study showed much higher tannin contents than
531 some other commercial gallnut extracts described in the literature (100% vs 46 to 51% in Vignault
532 et al,⁶ as determined by UV spectrophotometry). The presence of other compounds in the gallnut
533 tannin extracts may explain the discrepancies between tannin composition and antioxidant
534 properties reported by these authors.⁶

535 Thus, comparison of our results with the literature data confirm that commercial gallnut extracts
536 exhibit large variations in their tannin content and composition and emphasize the need for detailed
537 characterization of these products in the frame of quality control and selection for specific

538 applications. The large structural differences evidenced in our study had very little impact on
539 radical scavenging activity which seems primarily driven by gallotannin concentration. However,
540 their impact on other properties, including organoleptic properties and color stabilizing effect
541 remain to be investigated.

542 ACKNOWLEDGEMENTS

543 The authors would like to thank Stéphanie Carrillo for technical assistance in the GPC analysis.

544 FUNDING SOURCES

545 Funding from FEDER and Région Grand-Est, in the frame of “Programme opérationnel
546 FEDER/FSE/IEJ 2014-2020” is gratefully acknowledged.

547 **Supporting information**

548 Zooms of the negative ion mode ESI MS spectrum recorded after UPLC-DAD-MS analysis of
549 TAN B2 showing the region 500-1500 m/z and mass spectra of the doubly charged 16G-glc,
550 17G-glc and triply charged 18G-glc.

551 Negative ion mode ESI MS spectra (90-2000 m/z) recorded after UPLC-DAD-MS analysis of TAN A, TAN
552 B1, and TAN B2, after methanolysis. UPLC profile at 280 and extracted ion chromatograms at m/z 497
553 (Me-2G-glc), 649 (Me-3G-glc), and 801 (Me-4G-glc).

554

555 REFERENCES

- 556 (1) Haslam, E.; Cai, Y. Plant Polyphenols (Vegetable Tannins): Gallic Acid Metabolism. *Nat. Prod. Rep.*
557 **1994**, *11* (0), 41–66.
- 558 (2) Quideau, S.; Feldman, K. S. Ellagitannin Chemistry. *Chem. Rev.* **1996**, *96* (1), 475–504.
- 559 (3) Arapitsas, P. Hydrolyzable Tannin Analysis in Food. *Food Chem.* **2012**, *135* (3), 1708–1717.

- 560 (4) Harbertson, J. F.; Kennedy, J. A.; Adams, D. O. Tannin in Skins and Seeds of Cabernet Sauvignon,
561 Syrah, and Pinot Noir Berries during Ripening. *Am. J. Enol. Vitic.* **2002**, *53* (1), 54–59.
- 562 (5) Sarneckis, C. J.; Damberg, R. G.; Jones, P.; Mercurio, M.; Herderich, M. J.; Smith, P. A.
563 Quantification of Condensed Tannins by Precipitation with Methyl Cellulose: Development and
564 Validation of an Optimised Tool for Grape and Wine Analysis. *Aust. J. Grape Wine Res.* **2006**, *12* (1),
565 39–49.
- 566 (6) Vignault, A.; González-Centeno, M. R.; Pascual, O.; Gombau, J.; Jourdes, M.; Moine, V.; Iturmendi,
567 N.; Canals, J. M.; Zamora, F.; Teissedre, P.-L. Chemical Characterization, Antioxidant Properties and
568 Oxygen Consumption Rate of 36 Commercial Oenological Tannins in a Model Wine Solution. *Food*
569 *Chem.* **2018**, *268*, 210–219.
- 570 (7) Kennedy, J. A.; Ferrier, J.; Harbertson, J. F.; Peyrot des Gachons, C. Analysis of Tannins in Red Wine
571 Using Multiple Methods: Correlation with Perceived Astringency. *Am. J. Enol. Vitic.* **2006**, *57* (4),
572 481–485.
- 573 (8) Mercurio, M. D.; Smith, P. A. Tannin Quantification in Red Grapes and Wine: Comparison of
574 Polysaccharide- and Protein-Based Tannin Precipitation Techniques and Their Ability to Model
575 Wine Astringency. *J. Agric. Food Chem.* **2008**, *56* (14), 5528–5537.
- 576 (9) de Freitas, V.; Mateus, N. Protein/Polyphenol Interactions: Past and Present Contributions.
577 Mechanisms of Astringency Perception. *Curr. Org. Chem.* **2012**, *16* (6), 724–746.
- 578 (10) Scollary, G. R.; Pasti, G.; Kallay, M.; Blackman, J.; Clark, A. C. Astringency Response of Red Wines:
579 Potential Role of Molecular Assembly. *Trends Food Sci. Technol.* **2012**, *27* (1), 25–36.
- 580 (11) Soares, S.; Brandao, E.; Mateus, N.; de Freitas, V. Sensorial Properties of Red Wine Polyphenols:
581 Astringency and Bitterness. *Crit. Rev. Food Sci. Nutr.* **2017**, *57* (5), 937–948.
- 582 (12) Moutounet, M.; Leaute, B.; Delbos, C.; Doco, T.; Meudec, E.; Souquet, J.-M. Analyse de la
583 composition de tanins oenologiques. *Rev. Fr. Oenologie* **2004**, No. 208, 22–27.

- 584 (13) Lei, Z.; Jervis, J.; Helm, R. F. Use of Methanolysis for the Determination of Total Ellagic and Gallic
585 Acid Contents of Wood and Food Products. *J. Agric. Food Chem.* **2001**, *49* (3), 1165–1168.
- 586 (14) Armitage, R.; Bayliss, G. S.; Gramshaw, J. W.; Haslam, E.; Haworth, R. D.; Jones, K.; Rogers, H. J.;
587 Searle, T. 360. Gallotannins. Part III. The Constitution of Chinese, Turkish, Sumach, and Tara
588 Tannins. *J. Chem. Soc.* **1961**, 1842–1853.
- 589 (15) Haslam, E.; Haworth, R. D.; Mills, S. D.; Rogers, H. J.; Armitage, R.; Searle, T. 359. Gallotannins. Part
590 II. Some Esters and Depsides of Gallic Acid. *J. Chem. Soc.* **1961**, 1836–1842.
- 591 (16) Beasley, T.; Ziegler, H.; Bell, A. Determination and Characterization of Gallotannin by High-
592 Performance Liquid-Chromatography. *Anal. Chem.* **1977**, *49* (2), 238–243.
- 593 (17) Nishizawa, M.; Yamagishi, T.; Nonaka, G.; Nishioka, I. Tannins and Related-Compounds .5. Isolation
594 and Characterization of Polygalloylglucoses from Chinese Gallotannin. *J. Chem. Soc.-Perkin Trans. 1*
595 **1982**, No. 12, 2963–2968.
- 596 (18) Verzele, M.; Delahaye, P. Analysis of Tannic Acids by High-Performance Liquid-Chromatography. *J.*
597 *Chromatogr.* **1983**, *268* (3), 469–476.
- 598 (19) Armitage, R.; Haworth, R.; Haslam, E.; Searle, T. Gallotannins .6. Turkish Gallotannin. *J. Chem. Soc.*
599 **1962**, No. SEP, 3808-.
- 600 (20) Nishizawa, M.; Nishioka, I.; Yamagishi, T.; Nonaka, G. Tannins and Related-Compounds .9. Isolation
601 and Characterization of Polygalloylglucoses from Turkish Galls (*Quercus-Infectoria*). *J. Chem. Soc.-*
602 *Perkin Trans. 1* **1983**, No. 5, 961–965.
- 603 (21) Verzele, M.; Delahaye, P.; Vandamme, F. Determination of the Tanning Capacity of Tannic Acids by
604 High-Performance Liquid-Chromatography. *J. Chromatogr.* **1986**, *362* (3), 363–374.
- 605 (22) Hagerman, A.; Robbins, C.; Weerasuriya, Y.; Wilson, T.; McArthur, C. Tannin Chemistry in Relation
606 to Digestion. *J. Range Manag.* **1992**, *45* (1), 57–62.

- 607 (23) Potier, P.; Maccario, V.; Giudicelli, M.-B.; Queneau, Y.; Dangles, O. Gallic Esters of Sucrose as a New
608 Class of Antioxidants. *Tetrahedron Lett.* **1999**, *40* (17), 3387–3390.
- 609 (24) Karas, D.; Ulrichová, J.; Valentová, K. Galloylation of Polyphenols Alters Their Biological Activity.
610 *Food Chem. Toxicol.* **2017**, *105*, 223–240.
- 611 (25) Luo, F.; Fu, Y.; Xiang, Y.; Yan, S.; Hu, G.; Huang, X.; Huang, G.; Sun, C.; Li, X.; Chen, K. Identification
612 and Quantification of Gallotannins in Mango (*Mangifera Indica* L.) Kernel and Peel and Their
613 Antiproliferative Activities. *J. Funct. Foods* **2014**, *8*, 282–291.
- 614 (26) Albersheim, P.; Nevins, D. J.; English, P. D.; Karr, A. A Method for the Analysis of Sugars in Plant
615 Cell-Wall Polysaccharides by Gas-Liquid Chromatography. *Carbohydr. Res.* **1967**, *5* (3), 340–345.
- 616 (27) Kennedy, J. A.; Taylor, A. W. Analysis of Proanthocyanidins by High-Performance Gel Permeation
617 Chromatography. *J. Chromatogr. A* **2003**, *995* (1), 99–107.
- 618 (28) Melone, F.; Saladino, R.; Lange, H.; Crestini, C. Tannin Structural Elucidation and Quantitative P-31
619 NMR Analysis. 1. Model Compounds. *J. Agric. Food Chem.* **2013**, *61* (39), 9307–9315.
- 620 (29) Melone, F.; Saladino, R.; Lange, H.; Crestini, C. Tannin Structural Elucidation and Quantitative 31P
621 NMR Analysis. 2. Hydrolyzable Tannins and Proanthocyanidins. *J. Agric. Food Chem.* **2013**, *61* (39),
622 9316–9324.
- 623 (30) Catarina Guedes, A.; Amaro, H. M.; Giao, M. S.; Xavier Malcata, F. Optimization of ABTS Radical
624 Cation Assay Specifically for Determination of Antioxidant Capacity of Intracellular Extracts of
625 Microalgae and Cyanobacteria. *Food Chem.* **2013**, *138* (1), 638–643.
- 626 (31) Pizzi, A.; Pasch, H.; Rode, K.; Giovando, S. Polymer Structure of Commercial Hydrolyzable Tannins
627 by Matrix-Assisted Laser Desorption/Ionization-Time-of-Flight Mass Spectrometry. *J. Appl. Polym.*
628 *Sci.* **2009**, *113* (6), 3847–3859.

- 629 (32) Mane, C.; Sommerer, N.; Yalcin, T.; Cheynier, V.; Cole, R. B.; Fulcrand, H. Assessment of the
630 Molecular Weight Distribution of Tannin Fractions through MALDI-TOF MS Analysis of Protein-
631 Tannin Complexes. *Anal. Chem.* **2007**, *79* (6), 2239–2248.
- 632 (33) Fulcrand, H.; Mané, C.; Preys, S.; Mazerolles, G.; Bouchut, C.; Mazauric, J.-P.; Souquet, J.-M.;
633 Meudec, E.; Li, Y.; Cole, R. B.; et al. Direct Mass Spectrometry Approaches to Characterize
634 Polyphenol Composition of Complex Samples. *Phytochemistry* **2008**, *69* (18), 3131–3138.
- 635 (34) Yanagida, A.; Kanda, T.; Shoji, T.; Ohnishi-Kameyama, M.; Nagata, T. Fractionation of Apple
636 Procyanidins by Size-Exclusion Chromatography. *J. Chromatogr. A* **1999**, *855* (1), 181–190.
- 637 (35) Arapitsas, P.; Menichetti, S.; Vincieri, F. F.; Romani, A. Hydrolyzable Tannins with the
638 Hexahydroxydiphenoyl Unit and the M-Depsidic Link: HPLC-DAD-MS Identification and Model
639 Synthesis. *J. Agric. Food Chem.* **2007**, *55* (1), 48–55.
- 640 (36) Verzele, M.; Delahaye, P.; Dewaele, C. A New Structural Formula for the Polygalloyl Glucose Tannic
641 Acids. *Bull. Soc. Chim. Belg.* **1984**, *93* (2), 151–152.
- 642 (37) Balakshin, M.; Capanema, E. On the Quantification of Lignin Hydroxyl Groups with P-31 and C-13
643 Nmr Spectroscopy. *J. Wood Chem. Technol.* **2015**, *35* (3), 220–237.
- 644 (38) Ben, H.; Chen, X.; Han, G.; Shao, Y.; Jiang, W.; Pu, Y.; Ragauskas, A. J. Characterization of Whole
645 Biomasses in Pyridine Based Ionic Liquid at Low Temperature by P-31 NMR: An Approach to
646 Quantitatively Measure Hydroxyl Groups in Biomass As Their Original Structures. *Front. Energy*
647 *Res.* **2018**, *6*, UNSP 13.
- 648 (39) Crestini, C.; Lange, H.; Bianchetti, G. Detailed Chemical Composition of Condensed Tannins via
649 Quantitative P-31 NMR and HSQC Analyses: Acacia Catechu, Schinopsis Balansae, and Acacia
650 Mearnsii. *J. Nat. Prod.* **2016**, *79* (9), 2287–2295.
- 651 (40) Dais, P.; Spyros, A. P-31 NMR Spectroscopy in the Quality Control and Authentication of Extra-
652 Virgin Olive Oil: A Review of Recent Progress. *Magn. Reson. Chem.* **2007**, *45* (5), 367–377.

- 653 (41) Filpponen, I.; Argyropoulos, D. S. Determination of Cellulose Reactivity by Using Phosphitylation
654 and Quantitative P-31 NMR Spectroscopy. *Ind. Eng. Chem. Res.* **2008**, *47* (22), 8906–8910.
- 655 (42) Meng, X.; Crestini, C.; Ben, H.; Hao, N.; Pu, Y.; Ragauskas, A. J.; Argyropoulos, D. S. Determination
656 of Hydroxyl Groups in Biorefinery Resources via Quantitative (PNMR)-P-31-N- Spectroscopy. *Nat.*
657 *Protoc.* **2019**, *14* (9), 2627–2647.
- 658 (43) Neufeld, R.; Stalke, D. Accurate Molecular Weight Determination of Small Molecules via DOSY-
659 NMR by Using External Calibration Curves with Normalized Diffusion Coefficients. *Chem. Sci.* **2015**,
660 *6* (6), 3354–3364.
- 661 (44) Buziashvili, I. Sh.; Komissarenko, N. F.; Kovalev, I. P.; Gordienko, V. G.; Kolesnikov, D. G. The
662 Structure of Gallotannins. *Chem. Nat. Compd.* **1973**, *9* (6), 752–755.
- 663
- 664

665 FIGURE CAPTIONS

666 Figure 2. UPLC trace at 280 nm of TAN A (blue), TAN B1 (black) and TAN B2 (red) (A).

667 Negative ion mode ESI MS spectra (150-1500 m/z) recorded after UPLC-DAD-MS analysis of

668 TAN A (B), TAN B1 (C) and TAN B2 (D). G-Glc: galloyl-glucose

669 Figure 2: ^{31}P 1D and 2D TOCSY NMR spectra of a mixture of glucose, β -glucogalin and PGG in

670 pyridine- d_5 : CDCl_3 after phosphitylation. The numbering system of phenolic compounds and the

671 TOCSY correlations are shown on the left.

672 Figure 3: (A) ^{31}P 1D NMR spectra of TAN B1 (blue), TAN B2 (green) and TAN A (red) in

673 pyridine- d_5 : CDCl_3 after phosphitylation. (I= Internal units, T=Terminal units). (B) ^{31}P 1D and

674 2D TOCSY NMR spectra of TAN B2 in pyridine- d_5 : CDCl_3 after phosphitylation, region of

675 aromatic signals. The numbering system of polyphenol compounds and the TOCSY correlations

676 are shown on the left.

677 Figure 4. Aliphatic region of 1D ^1H and 2D HSQC $^1\text{H}/^{13}\text{C}$ spectra. (A) Standard mixture of β -

678 glucogallin ($^\circ$) and PGG (*); (B) Tan A; (C) TAN B1; (D) TAN B2. Attributed protons and carbons

679 of glucoside moieties are noted in the spectra according to the glucoside numbering scheme.

680 Figure 5. Molecular size distribution of gallnut tannin extracts (TAN A and TAN B2 vs. B1 lots

681 at 2.5 g/L) detected at 280 nm by gel permeation chromatography. The retention times of gallic

682 acid, β -glucogallin, and pentagalloylglucose standards are reported on the chromatogram.

683 Figure 6. A, calibration curves obtained from gallotannin standards (gallic acid, glucogallin and

684 pentagalloylglucose) added to the extracts in solution. Aromatic region of ^1H 2D DOSY NMR

685 spectra of TAN A (B), TAN B1 (C) and TAN B2 (D) extracts with added gallotannin standards in

686 $\text{DMSO}-d_6$ (x: gallic acid, o: β -glucogallin, *: pentagalloylglucose signals).

687

688 *Table 1. Concentration of gallic acid, digallic acid, trigallic acid, β -glucogallin and poly-*
 689 *galloylglucose expressed in mg/g gallic acid equivalent.*

	gallic acid	digallic acid ¹	trigallic acid	β -glucogallin	gallotannins ²
TAN A	22.92 \pm 0.67 ^b	26.92 \pm 0.88 ^c	9.83 \pm 0.26 ^c	0.71 \pm 0.02 ^c	987.38 \pm 38.27 ^a
TAN B1	73.27 \pm 0.90 ^c	20.27 \pm 0.30 ^b	7.17 \pm 0.09 ^b	0.12 \pm 0.01 ^a	1027.69 \pm 18.65 ^a
TAN B2	8.91 \pm 0.10 ^a	12.54 \pm 0.26 ^a	4.73 \pm 0.14 ^a	0.17 \pm 0.01 ^b	1094.17 \pm 10.16 ^b

690 ¹two isomers, the least polar coeluted with digalloylglucose. ²estimated from total peak area at

691 280 nm minus peak areas of gallic acid, digallic acid, trigallic acid, and β -glucogallin.

692 Different superscripts indicate statistical differences between samples.

693

694 *Table 2. Integral values and compound contents of gallnut tannin extracts calculated from ³¹P 1D*
 695 *NMR spectra after phosphitylation*

	TAN A	TAN B1	TAN B2	Integration ranges in ppm
Integral values in mmol/g				
(OP3+OP5)T	4.30 ± 0.22	4.49 ± 0.19	3.80 ± 0.16	138.71-138.00
(OP4)T	2.12 ± 0.10	2.24 ± 0.12	2.09 ± 0.11	141.75-140.88
(OP3+OP5)I	2.83 ± 0.14	3.87 ± 0.16	4.00 ± 0.01	140.87-138.72
(OP4)I	1.83 ± 0.03	1.70 ± 0.09	1.74 ± 0.05	142.57-141.75
Concentrations in mmol/g				
Gallic acid	0.10 ± 0.01	0.35 ± 0.02	0.04 ± 0.00	134.70-134.54
Digallic acid	0.044 ± 0.002	0.042 ± 0.002	0.023 ± 0.002	135.09-135.05
Trigallic acid	0.018 ± 0.002	0.015 ± 0.001	0.009 ± 0.000	135.04-135.00
OP of glucose	0.52 ± 0.02			148.20-147.85
T ^a	2.14 ± 0.10	2.24 ± 0.05	1.96 ± 0.09	
I ^b	2.33 ± 0.08	2.79 ± 0.12	2.87 ± 0.03	
T*	1.98 ± 0.1	1.83 ± 0.04	1.89 ± 0.09	
I*	2.19 ± 0.08	2.66 ± 0.12	2.79 ± 0.03	
I meta ^c	1.83 ± 0.03	1.70 ± 0.09	1.74 ± 0.05	
I para ^d	0.5 ± 0.06	1.09 ± 0.05	1.13 ± 0.02	
Glucose ^e	0.50 ± 0.16	0.37 ± 0.01	0.38 ± 0.02	
Number of galloyl units per glucose moieties ^f	8.48 ± 1.00	12.24 ± 0.05	12.39 ± 0.27	
gallotannins in mg/g ^g	725 ± 28	749 ± 26	780 ± 20	
Mass balance in mg/g ^h	765 ± 30	829 ± 27	799 ± 20	

696 *Means and standard deviations are calculated from the results of three independent experiments.*

697 *The formula used are: a: ((OP3+OP5)T+(OP4)T)/3, b: ((OP3/OP5)I+(OP4)I)/2, c: (OP4)I,*

698 **d:** *I* - *I* meta, **e:** $T/4$ for TAN A, $T/5$ for TANB, **f:** $4(I+T)/(T)$ for TAN A, $5(I+T)/(T)$ for TAN B, **g:**
699 Sum of (molarity*molecular weight) of *T*, *I* and glucose moieties taking into account the loss of
700 H_2O in the ester bonds, **h:** Sum of gallotannins, mono-, -di- and tri-gallic acids. *I*= Internal units,
701 *T*=Terminal units, $T^*=T$ minus mono-, di- and tri-gallic acid terminal units, $I^*=I$ minus di- and
702 tri-gallic acid intermediate units

703

704 *Table 3. Summary of the data generated by the different methods*

		UPLC-DAD ¹	UPLC-MS	NMR	GPC
Gallic acid (mg/g)	A	22.9 ± 0.7		17 ± 1.7	B1>>A>B2
	B1	73.3 ± 0.9		59.5 ± 3.4	
	B2	8.9 ± 0.1		6.8 ± 0.1	
Digallic acid (mg/g)	A	26.9 ± 0.9		14.2 ± 0.6	
	B1	20.3 ± 0.3		13.5 ± 0.6	
	B2	12.5 ± 0.3		7.4 ± 0.6	
Trigallic acid (mg/g)	A	9.83 ± 0.26		8.5 ± 0.9	
	B1	7.17 ± 0.09		7.1 ± 0.5	
	B2	4.73 ± 0.14		4.3 ± 0.2	
β-glucogallin (mg/g)	A	0.71 ± 0.02			
	B1	0.12 ± 0.01			
	B2	0.17 ± 0.01			
Gallotannins (mg/g)	A	790 ± 66		725 ± 28	B2 > B1
	B1	822 ± 14		749 ± 26	
	B2	875 ± 8		780 ± 20	
Glucose substitution	A	tetragalloylglucose			
	B1	pentagalloylglucose			
	B2	pentagalloylglucose			
Average number of galloyl groups	A			8.5	5G-glc > 5G-glc > 5G-glc
	B1			12.2	
	B2			12.4	
Max number of galloyl groups detected	A		12		
	B1		17		
	B2		18		
Major species detected	A	4G-glc, 5G-glc			
	B1	7G-glc, 8G-glc			
	B2	7G-glc, 8G-glc (different isomers)			
Additional information	A		galloylglucose ethers (as methyl esters detected after methanolysis)	unknown low molecular weight aliphatic compound in B1	
	B1			proportion of p-depside bonds > in B	
	B2				

705 ¹ Gallic, digallic, trigallic acids, and β -glucogallin in mg/g gallic acid equivalent; gallotannins in mg/g
706 pentagalloylglucose equivalent.

707

708

709 Graphic for Table of Content

710

711

712

