



HAL
open science

Predictive Models of Biohydrogen and Biomethane Production Based on the Compositional and Structural Features of Lignocellulosic Materials

Florian Monlau, Cécilia Sambusiti, Abdellatif Barakat, Xinmei Guo, Eric Latrille, Eric Trably, Jean-Philippe Steyer, Hélène Carrère

► **To cite this version:**

Florian Monlau, Cécilia Sambusiti, Abdellatif Barakat, Xinmei Guo, Eric Latrille, et al.. Predictive Models of Biohydrogen and Biomethane Production Based on the Compositional and Structural Features of Lignocellulosic Materials. *Environmental Science and Technology*, 2012, 46 (21), pp.12217 - 12225. 10.1021/es303132t . hal-02646678

HAL Id: hal-02646678

<https://hal.inrae.fr/hal-02646678>

Submitted on 8 Aug 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.

1 **Predictive models of biohydrogen and biomethane production based on the**
2 **compositional and structural features of lignocellulosic materials**

3 Florian MONLAU¹, Cecilia SAMBUSITI², Abdellatif BARAKAT³, Xin Mei GUO¹, Eric
4 LATRILLE¹, Eric TRABLY¹, Jean-Philippe STEYER¹, H el ene CARRERE^{1*}

5 ¹INRA, UR050, Laboratoire de Biotechnologie de l'Environnement, Avenue des Etangs, 11100
6 Narbonne, France.

7 ² Politecnico di Milano, DIIAR, Environmental Section, Piazza L. da Vinci, 32, 20133, Milano,
8 Italy.

9 ³ INRA, UMR IATE 1208, Ing enierie des Agro polym eres et Technologies Emergentes, 2, place
10 Pierre Viala F- 34060 Montpellier, France.

11
12
13 * phone: +33 4 68 42 51 68, fax +33 4 68 42 51 60; Email: helene.carrere@supagro.inra.fr
14

15 **KEYWORDS:** lignocellulosic biomass, structural features, anaerobic digestion, dark
16 fermentation, crystallinity

17
18 **Abstract:**

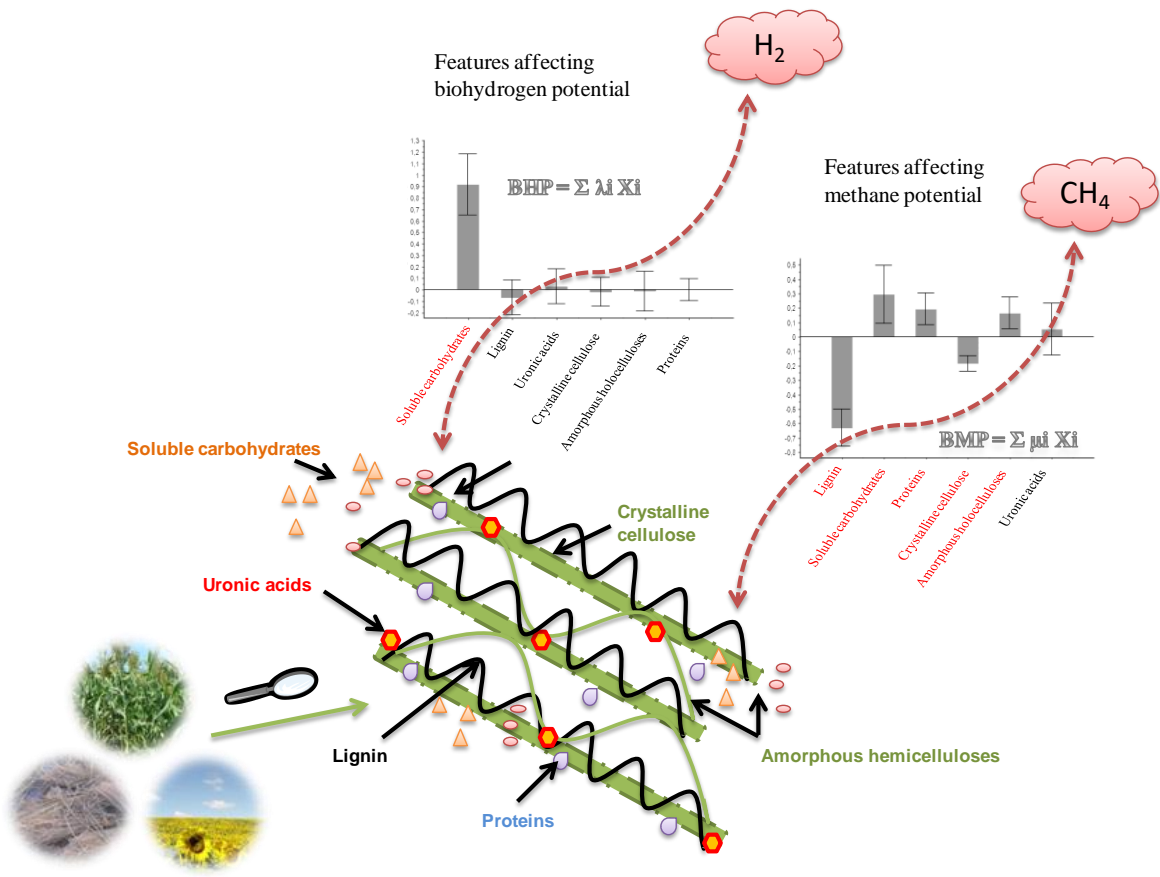
19 In an integrated biorefinery concept, biological hydrogen and methane production from
20 lignocellulosic substrates appears to be one of the most promising alternatives to produce
21 energy from renewable sources. However lignocellulosic substrates present compositional and
22 structural features that can limit their conversion into biohydrogen and methane. In this study,
23 biohydrogen and methane potentials of twenty lignocellulosic residues were evaluated.
24 Compositional (lignin, cellulose, hemicelluloses, total uronic acids, proteins and soluble
25 sugars) as well as structural features (crystallinity) were determined for each substrate. Two
26 predictive Partial Least Square (PLS) models were built to determine which compositional
27 and structural parameters affected biohydrogen or methane production from lignocellulosic
28 substrates, among proteins, total uronic acids, soluble sugars, crystalline cellulose, amorphous
29 holocelluloses and lignin. Only soluble sugars had a significant positive effect on biohydrogen

30 production. Besides, methane potentials correlated negatively to the lignin contents. In a
 31 lower extent, crystalline cellulose showed also a negative impact on methane potentials. In
 32 contrast, soluble sugars, proteins and amorphous hemicelluloses increased the methane
 33 production. These findings will help to develop further pretreatment strategies for enhancing
 34 both biohydrogen and methane production.

35

36 **Graphical abstract:**

37



38

39

40

41

42 **Introduction:**

43 Development of new technologies for renewable energy generation such as biohydrogen and
44 biomethane from lignocellulosic materials in a concept of integrated biorefinery appears as a
45 very promising alternative to fossil fuels ^{1,2}. Worldwide, the lignocellulosic biomass was
46 evaluated about 200 billion tons annually ³. The use of lignocellulosic biomass, and more
47 particularly agricultural residues as bioenergy sources is interesting because of (i) its
48 renewability, (ii) it provides additional incomes to farmers, (iii) it uses the non edible part
49 (stalks, leaves) of the plants and thus does not enter in competition with food and (iv) it permits
50 to treat residues which are often burnt in the field creating environmental pollution ^{3,4}.

51 Methane is produced by a biological process in four steps (hydrolysis, acidogenesis, acetogenesis
52 and methanogenesis) so-called anaerobic digestion. Biohydrogen is produced by dark
53 fermentation which consists of an intermediate stage of anaerobic digestion where the last step of
54 methanogenesis does not occur. Biohydrogen and methane can be produced using undefined
55 mixed microbial cultures ^{2,5}. Mixed cultures are easier to use than pure cultures as they do not
56 require aseptic conditions and can convert a large range of feedstocks into biohydrogen or
57 methane ^{6,7}. Nevertheless, in the case of biohydrogen production with mixed cultures, it is
58 necessary to apply heat shock or chemical pretreatments in order to block the conversion of
59 acetate or hydrogen and carbon dioxide into methane ⁵.

60 Lignocellulosic substrates are composed of three main fractions: lignin, cellulose and
61 hemicelluloses. Contrary to lignin, the holocelluloses, ie cellulose and hemicelluloses, can be
62 converted into biohydrogen and methane ². Nevertheless, lignocellulosic substrates present
63 structural features that limit the accessibility of holocelluloses to microorganisms and thus their
64 conversion to biohydrogen or methane.

65 Only few studies have attempted to give some insights on the effect of compositional and
66 structural features of lignocellulosic substrates on biohydrogen and methane production ⁸⁻¹². The

67 correlations found in the literature between the composition of lignocellulosic residues and
68 biohydrogen or methane production are summarized in Table 1.

69 For hydrogen production, Guo et al., (2012) recently showed a good correlation ($R^2 = 0.89$)
70 between biohydrogen yields and soluble carbohydrates extracted under mild acidic conditions (2
71 N hydrochloric acid) ¹². The main bottleneck of using lignocellulosic biomass in dark
72 fermentation processes is to convert holocelluloses into fermentable sugars ¹³. Recently, Yuan et
73 al. (2011) showed that hydrogen production from wheat straw was well correlated with the
74 degradation of cellulose and hemicelluloses into fermentable sugars ¹⁴. However, knowledge
75 about the effect of compositional and structural features on biohydrogen potentials remains very
76 limited.

77 According to Gunaseelan (2007), methane potentials can be predicted from five main chemical
78 constituents (total soluble carbohydrate, acid detergent fibers (ADF), lignin/ADF, nitrogen and
79 ash) which accounted for 90% of the total variation in methane potentials ($R^2=0.90$) ⁸. Negative
80 correlations were also found between lignin contents and biochemical methane potentials for
81 manure and energy crops ($R^2=0.88$) ¹¹. Similarly, Buffiere et al. (2006) showed a negative
82 correlation between anaerobic biodegradability and the sum of cellulose and lignin contents ¹⁰. In
83 contrast, Eleazer et al. (1997) reported that methane potentials from several municipal solid
84 wastes correlated positively to the sum of cellulose and hemicelluloses contents ¹⁵. In all these
85 studies, lignin seemed to be the main restrictive factor for methane production, likely by limiting
86 the microbial accessibility to holocelluloses during the fermentative process ^{2,16}. Overall, the
87 effect of other compositional features, especially cellulose, is still not clear and sometimes
88 contradictory between the different studies.

89 Except models established by Gunasselan (2007 and 2009), all models previously described were
90 built with only one or two compositional characteristics ^{8,9}. Moreover, only compositional
91 features have been considered and the effect of structural characteristics such as cellulose

92 crystallinity have not been investigated yet. Indeed, cellulose presents both crystalline and
93 amorphous parts and the crystalline one prevents cell penetration by micro-organisms or
94 extracellular enzymes ¹⁷. Other compositional characteristics such as the presence of pectin
95 (polymer of uronic acids) have not been considered in models. Recently, Pakarinen et al. (2012)
96 showed that pectin removal can significantly increase enzymatic hydrolysis of lignocellulosic
97 substrates ¹⁸.

98 Information about the influence of compositional and structural features on fermentative
99 processes is thus limited especially for biohydrogen production, and sometimes results are
100 contradictory. So the determination of compositional (lignin, holocelluloses, uronic acids and
101 soluble fractions) and structural (crystallinity of cellulose) characteristics appears essential to
102 understand the limitation of lignocellulosic material conversion into biohydrogen or methane.
103 Moreover, this study can be valuable to obtain guidelines for establishing further pretreatment
104 strategies to improve biohydrogen and methane production from lignocellulosic residues.

105 The objectives of this study were: (1) to characterize the compositional (cellulose,
106 hemicelluloses, lignin, uronic acids, proteins, soluble carbohydrates) and structural features
107 (crystallinity of cellulose) of various lignocellulosic substrates, (2) to evaluate their biohydrogen
108 and methane potentials and (3) to develop multilinear PLS models for predicting biohydrogen
109 and methane potentials from their compositional and structural features.

110

111 **2. Materials and methods**

112 **2.1. Lignocellulosic materials**

113 The substrates used in this study were selected among various lignocellulosic residues, biomass
114 crops and carbohydrate-rich substrates, but no lipid-rich substrate was considered. They
115 corresponded to rice straw, giant reed (stalks and leaves), three varieties of sunflower stalks (1, 2,
116 and 3), sunflower bark, sunflower oil cakes, maize (stalks, leaves and cobs), Jerusalem artichoke
117 (stalks, leaves and tubers), and six varieties of sorghum (1: seed sorghum stalks, 2: biomass
118 sorghum, 4: forage sorghum, 3,5, and 6: sweet sorghum). All substrates were milled into particles of
119 2 mm using a cutting milling Restch, SM 100. The substrates were analyzed for Total Solids (TS)
120 and Volatile Solids (VS) (Table 1) according to the APHA standard method¹⁹.

121 **2.2. Chemical composition**

122 Soluble sugars (glucose and fructose) from starch, sucrose and inulin were extracted using a mild
123 acid hydrolysis method²⁰. Samples (200 mg) were hydrolyzed at 121°C, for 1h, with 0.2% H₂SO₄.
124 The supernatant was filtrated with nylon filters (20 µm) and released carbohydrates (glucose and
125 fructose) were quantified by High-Pressure Liquid Chromatography (HPLC) method coupled to
126 refractometric detection. The analysis was done with a combined Water/Dionex system (Ultimate
127 3000), using a Biorad HPX-87P column at 85°C. The eluent corresponded to deionized water under
128 a flow rate of 0.6 mL min⁻¹. The system was calibrated with glucose and fructose standards (Sigma-
129 Aldrich®).

130 Structural-carbohydrates (glucose, xylose, arabinose, uronic acids) from cellulose, hemicelluloses
131 and pectins were measured using a strong acid hydrolysis method adapted from Effland et al.
132 (1977)²¹. Samples (200 mg) were first hydrolyzed with 12 M H₂SO₄ acid for 2 h at room
133 temperature, then diluted to reach a final acid concentration of 1.5 M and kept at 100°C for 3 h. The

134 insoluble residue was separated from the supernatant by filtration on fibreglass paper (GFF,
135 WHATMAN). This insoluble residue was washed with 50 mL of deionized water and then placed
136 in a crucible. The crucible and the paper fibreglass were dried at 100°C during 24 h to determine by
137 weighing the amount of Klason lignin. The supernatant was further filtrated with nylon filters (20
138 µm) and analyzed for quantification of monomeric carbohydrates. All monosaccharides (glucose,
139 xylose, arabinose, uronic acids) were analyzed by HPLC coupled to refractometric detection. The
140 analysis was carried out with a combined Water/Dionex system (Ultimate 3000), using a Biorad
141 HPX-87H column at 50°C. The eluent corresponded to 0.005 M H₂SO₄ under a flow rate of 0.3 mL
142 min⁻¹. A refractive index detector (Waters 2414) was used to quantify the carbohydrates. The
143 system was calibrated with glucose, xylose, arabinose, and uronic acids (galacturonic and
144 glucuronic) standards (Sigma–Aldrich®). Thereafter, cellulose and hemicelluloses contents were
145 estimated as follows (equation 1 and 2):

$$146 \text{ Cellulose (\% TS)} = \text{Glucose (\%TS)} / 1.11 \quad (1)$$

$$147 \text{ Hemicelluloses (\% TS)} = [\text{Xylose (\%TS)} + \text{Arabinose (\%TS)}] / 1.13 \quad (2)$$

148 where 1.11 is the conversion factor for glucose-based polymers (glucose) to monomers and 1.13 is
149 the conversion factor for xylose-based polymers (arabinose and xylose) to monomers according to
150 Petersson et al. (2007) ²².

151

152 **2.3. Crystallinity measurement assessment**

153 Fourier Transform Infrared Spectroscopy (FTIR) spectroscopy was used to determine the
154 crystallinity of lignocellulosic materials. FTIR spectra were collected in the 4000–600 cm⁻¹ range
155 using a Nexus 5700 spectrometer (ThermoElectron Corp.) with built-in diamond ATR single
156 reflection crystal and with a cooled MCT detector. Spectra were recorded in absorption mode at
157 4 cm⁻¹ intervals with 64 scans, at room temperature. Three spectra were recorded for each sample

158 and all spectra pre-treatments were analyzed using Omnic v7.3 software. Among the different
159 FTIR bands, the bands at 1430 and 898 cm^{-1} are sensitive to the amount of crystalline cellulose and
160 amorphous cellulose respectively ²³. The bands ratio H 1430/ H 898 commonly called Lateral Order
161 Indice (LOI) can be used to determine the amount of crystalline cellulose. Using equations 3 and 4,
162 the crystalline cellulose content was estimated (equation 5).

163

$$164 \text{ Cellulose} = \text{Crystalline cellulose} + \text{Amorphous cellulose} \quad (3)$$

$$165 \text{ LOI} = \text{Crystalline cellulose} / \text{Amorphous cellulose} \quad (4)$$

$$166 \text{ Crystalline cellulose (IR)} = \text{Cellulose} \times \text{CrI}_{\text{IR}} \quad (5)$$

$$167 \text{ Where } \text{CrI}_{\text{IR}} = \text{LOI} / (1 + \text{LOI})$$

168 To validate the use of FT-IR spectra to assess cellulose crystallinity, crystallinity was also
169 determined by a more common technology, as X-ray diffraction, on eight lignocellulosic substrates
170 (giant reed stalks, sunflower stalks 1, maize stalks, rice straw, sorghum 1, Jerusalem artichoke
171 stalks, maize cobs and sunflower oil cakes). X-ray measurements were performed in a Philips
172 Analytical X-diffractometer, using Cu Ka radiation at $\lambda = 0.1540 \text{ nm}$ (40 kV, 40 mA). The
173 measurements were carried out on powder compacted to small mats. DRX data were collected at 2θ
174 angle range from 5° to 50° with a step interval of 0.02° . The degree of crystallinity was expressed
175 as a percentage of crystallinity index (% CrI). The equation used to calculate the CrI was previously
176 described by Segal et al. (1959) in the following form²⁴:

$$177 \text{ CrI}_{\text{DRX}} = (I_{002} - I_{\text{am}}) / I_{002} * 100 \quad (6)$$

178 where I_{002} corresponds to the counter reading at peak intensity at a 2θ angle of 22° and I_{am} the
179 counter reading at peak intensity at 2θ angle of 16° in cellulose. $I_{002} - I_{\text{am}}$ corresponds to the
180 intensity of the crystalline peak and I_{002} is the total intensity after subtraction of the background
181 signal measured without cellulose ²⁵. Crystalline cellulose was determined using the equation 7:

182 Crystalline cellulose (DRX) = Cellulose x CrI_{DRX} (7)

183 A good correlation ($R^2 = 0.93$) was found between crystalline cellulose values determined by DRX
184 and FTIR (Figure 1). However, the amounts of crystalline cellulose determined by FTIR were
185 higher than DRX, likely because CrI_{IR} measurements corresponded only to approximated values.
186 Indeed, although 1430 and 898 cm^{-1} bands are sensitive to the amount of crystalline cellulose and
187 amorphous cellulose, respectively, and each band contains contributions from both crystalline and
188 amorphous regions. Therefore, FTIR measurements must be considered as relative values and the
189 FTIR method was only used to compare crystalline cellulose contents from different lignocellulosic
190 materials.

191

192 **2.4. Biohydrogen and methane production**

193 **2.4.1 BioHydrogen Potential (BHP)**

194 BHP experiments were carried out in batch mode at 37°C. The volume of each flask was 600 mL,
195 with a working volume of 400 mL. A quantity of 3.5 g VS of substrate was initially introduced in
196 each flask. Then, 200 mL of MES (2-[N-morpholino] ethane sulfonic acid, 50 mmol.L^{-1}) buffer and
197 3 mL of seed sludge of an anaerobic digester (as inoculum) (final concentration of 225 mg-COD.L^{-1})
198 were added to the flask. The inoculum was first treated at 90°C for 15 minutes to inhibit the
199 activity of methanogens and enrich in hydrogen producing bacteria. No additional nutrient medium
200 solution was added. The initial pH value was adjusted to 5.5 with NaOH 2 N or 37 % HCl. The
201 headspace of the flasks was flushed with nitrogen gas to reach anaerobic conditions. The
202 experimental procedure ended when the pressure in the flask headspace started to drop off
203 indicating hydrogen consumption. Each experiment was performed in duplicates.

204

205 **2.4.2 Biochemical Methane Potential (BMP)**

206 Lignocellulosic substrates were digested anaerobically in batch anaerobic flasks at 35°C during 40
207 days. The volume of each flask was 600 mL, with a working volume of 400 mL. Each flask
208 contained: macroelements (NH_4Cl , 26 g.L⁻¹; KH_2PO_4 , 10 g.L⁻¹; MgCl_2 , 6 g.L⁻¹; CaCl_2 , 3 g.L⁻¹),
209 oligoelements (FeCl_2 , 2 g.L⁻¹; CoCl_2 , 0.5 g.L⁻¹; MnCl_2 , 0.1 g.L⁻¹; NiCl_2 , 0.1 g.L⁻¹; ZnCl_2 , 0.05 g.L⁻¹;
210 H_3BO_3 , 0.05 g.L⁻¹; Na_2SeO_3 , 0.05 g.L⁻¹; CuCl_2 , 0.04 g.L⁻¹; Na_2MoO_4 , 0.01 g.L⁻¹), bicarbonate
211 buffer (NaHCO_3 , 50 g.L⁻¹), an anaerobic sludge at 5 g VS.L⁻¹ and the substrate at 5 g TS.L⁻¹. Once
212 the flasks were prepared, a degasification step with nitrogen gas was carried out to obtain anaerobic
213 conditions. The bottles were closed with air impermeable red butyl rubber septum-type stoppers.
214 Bottles were incubated at 35°C and each experiment was carried out in duplicates.

215 **2.4.3 Gas analysis**

216 Biogas volume was monitored continuously with a water displacement method. Acidified water (pH
217 =2) was used to minimize dissolution of carbon dioxide. All volumes were expressed under
218 temperature and pressure standard conditions. The gas composition (O_2 , CO_2 , CH_4 , H_2 and N_2) was
219 analysed using a gas chromatograph (Clarus 580, Perkin Elmer) equipped with two columns, a
220 molecular sieve (Molsieve, 5Å) and a thermal conductivity detector (TCD). One column
221 (RtMolsieve) was used to separate H_2 , O_2 , N_2 and CH_4 , and the second one (RtQBond) was used to
222 separate CO_2 from other gases. The calibration was carried out with a standard gas (Linde TM)
223 composed of 25 % CO_2 , 2 % O_2 , 10 % N_2 and 5 % H_2 and 58 % CH_4 .

224

225 **2.5 Partial least square regression**

226 PLS (Partial Least Square) models were developed using Unscrambler Version 10.2 software
227 (CAMO software, A/S, Oslo, Norway). This method is particularly adapted for data with highly
228 correlated variables. PLS models were used in full cross validation so-called leave-one-out cross
229 validation procedure. This is a model validation method in which one sample is left out iteratively,

230 a new calibration model is built, and then the sample that was left out is predicted using this model
231 ²⁶. The iteration is continued until all samples are left once out of the calibration set. The prediction
232 performances of the models were evaluated by calculating the coefficient of determination (R^2) and
233 the root mean square error of the calibration data set (RMSEPC). High R^2 and low RMSEPC values
234 indicate a good predictive robustness of the model. PLS models built were then tested on an
235 independent set, and the root mean square error of independent validation set (RMSEPiv) was
236 calculated to define the quality of the model. The RMSEP was defined as follow:

$$237 \quad RMSEP = \sqrt{\frac{\sum_1^n (\hat{y}_i - y_i)^2}{n}} \quad (8)$$

238 where: \hat{y}_i is the prediction value of the sample i in a calibration data set (or independent validation
239 set); y_i , is the measured BHP or BMP value of the sample i in a calibration data set (or in a
240 independent validation set) and n is the number of samples in calibration data set (or independent
241 validation set).

242

243 **3. Results and discussion**

244 3.1 Compositional and structural characteristics of the lignocellulosic substrates

245 Soluble sugars (SolSu), uronic acids (Ua), proteins (Pro), hemicelluloses (Hem), cellulose (Cell),
246 and lignin (Lig) contents of twenty lignocellulosic substrates are presented in Table 2, in % of TS.
247 Soluble sugars (non structural carbohydrates like starch, sucrose and inulin) were mainly present in
248 sorghum substrates (ranging from 8.2 to 22.8 %, except for sorghum 1). Gunaseelan (2007) noticed
249 as well a high content of soluble carbohydrates up to 23 % of VS in sorghum bicolor roots ⁸.
250 According to Thuesombat et al. (2007), Jerusalem artichoke presents 70-90% of inulin (linear poly-
251 fructose chain) which explains the high values of soluble sugars found in Jerusalem artichoke stalks
252 and tubers, ie 32.9 % and 59.1 % per TS respectively ²⁷. Proteins content ranged from 2.3 %

253 (sunflower stalk 2) to 29.7 % (sunflower oil cakes). This result is consistent with Raposo et al.,
254 (2008) who evaluated a protein content of 31 % per TS in sunflower oil cakes ²⁸. Uronic acids
255 (galacturonic and glucuronic) which originated from both hemicelluloses and pectins were also
256 quantified. Uronic acids contents ranged from 0.2 % (giant reed and Jerusalem artichoke stalks) to 7
257 % (sunflower stalks 1). Concerning the holocelluloses fraction, hemicelluloses content ranged from
258 5 % (Jerusalem artichoke tubers) to 34.6 % (maize cobs) and cellulose contents ranged from 5.4 %
259 (Jerusalem artichoke bulbs) to 33.1 % (giant reed stalks). Crystalline cellulose and amorphous
260 holocelluloses expressed in % TS using FTIR spectra are presented in Table 2. The crystalline
261 cellulose content ranged from 2.5 % for Jerusalem artichoke bulbs to 16.3 % for giant reed stalks.
262 The content of amorphous holocelluloses, which is the sum of amorphous cellulose and
263 hemicelluloses, ranged from 7.5 % (Jerusalem artichoke tubers) to 50.3 % (maize cobs). Finally,
264 lignin content ranged from 12.3 % (Jerusalem artichoke tubers) to 35 % (sunflower stalks bark).
265 Moreover on a same plant, lignin content was found higher in stalks than in leaves, except for giant
266 reeds that presented almost similar lignin contents. Similar trends were observed with 14.1 % and
267 18.4 % of lignin for wheat straw leaves and stalks, respectively ²⁹. The range of the variables values
268 (% per TS) is relatively high to permit to screen a wide range of compositional and structural
269 features (Table 2).

270

271 3.2 Biological hydrogen potential (BHP) and biological methane potential (BMP) tests

272 Hydrogen and methane potentials of lignocellulosic substrates are presented in Figure 2.
273 Biohydrogen potentials ranged from 1.6 (\pm 0.1) mL H₂ g⁻¹TS (sunflower stalk bark) to 120 (\pm 11)
274 mL H₂ g⁻¹TS (Jerusalem artichoke tubers). Similarly to Jerusalem artichoke tubers, Jerusalem
275 artichoke stalks presented an interesting biohydrogen potential as 62 (\pm 6) mL H₂ g⁻¹ TS were
276 produced. Except for sorghum 1, high sorghum hydrogen potentials from 23 (\pm 1) mL H₂ g⁻¹ TS to
277 64 (\pm 14) mL H₂ g⁻¹ TS were observed for sorghum substrates. Similar hydrogen yields were

278 reported on sweet sorghum stalks with 52 mL H₂ g⁻¹ VS³⁰. Sunflower stalks were found to produce
279 low hydrogen potentials: 1.8 (± 0.9), 2.1 (± 0.7) and 2.5 (± 0.5) mL H₂ g⁻¹ TS for sunflower stalks
280 numbers 3, 2, and 1, respectively. With similar lignocellulosic residues, slightly lower hydrogen
281 yields of 1 mL H₂ g⁻¹ VS and 3.16 mL H₂ g⁻¹ VS were observed for wheat straw and cornstalks,
282 respectively^{31,32}.

283 In addition, methane potentials ranged from 155 (± 2) mL CH₄ g⁻¹ TS (sunflower stalks bark) to 300
284 (± 14) mL CH₄ g⁻¹ TS (Jerusalem artichoke tubers). Such results are in agreement with literature
285 data as Dinuccio et al. (2010) found methane potentials of 317 mL CH₄ g⁻¹ VS for maize residues,
286 229 mL CH₄ g⁻¹ VS for barley straw and 195 mL CH₄ g⁻¹ VS for rice straw³³. Besides Jerusalem
287 artichoke tubers, interesting methane production of 230 (± 18) and 260 (± 4) mL CH₄ g⁻¹ TS were
288 observed for Jerusalem artichoke stalks and leaves, respectively. All sorghum substrates present
289 methane potentials higher than 210 (± 33) mL CH₄ g⁻¹ TS. Maize leaves and sunflower oil cakes
290 also led to good methane potentials with respectively 235 (± 3) and 244 (± 9) mL CH₄ g⁻¹ TS. Low
291 methane potentials were observed for the different varieties of sunflower stalks as 167 (± 27), 172
292 (± 5), and 175 (± 9) mL CH₄ g⁻¹ TS for sunflower stalks numbers 3, 1, and 2, respectively.
293 Moreover, on a same plant, the leaves appeared to have higher methane potentials than stalks. As an
294 example, methane potentials of 170 (± 22) and 210 (± 13) mL CH₄ g⁻¹ TS were respectively
295 observed for giant reed stalks and leaves, respectively. Overall, all results were lower than 480 mL
296 CH₄ / kg TS which is the theoretical methane potential of lignocellulosic substrates as proposed by
297 Frigon and Guiot (2010)³⁴. Some biodegradable parts are indeed not accessible during anaerobic
298 digestion of lignocellulosic substrates likely due to the compositional and structural characteristics
299 that limit the accessibility of microorganisms to holocelluloses, as previously suggested by Triolo et
300 al. (2011)¹¹.

301

302 3.3 PLS models

303 One of the main objectives of this study was to identify the compositional and structural features
304 affecting both biohydrogen and methane production from lignocellulosic residues, such as lignin
305 (Lig), amorphous holocelluloses (Am), crystalline cellulose (Cri), protein (Pro), uronic acids (Ua)
306 and soluble sugars (SolSu) contents. PLS models were built on eighteen lignocellulosic substrates
307 and an independent validation set of two substrates (sorghum 1 and sorghum 6) was used to validate
308 the PLS models. Table 2 shows the range values of the variables (Lig, Am, Cri, Pro, Ua, SolSu) in
309 which PLS models are relevant and should not be extrapolated out of these ranges. In particular, the
310 models should not be applied to lipid-rich substrates. Although these models are valid to estimate
311 biohydrogen or methane yields in relation to compositional and structural features of lignocellulosic
312 biomass, they provide no information about substrate degradation rates. Other abiotic and biotic
313 factors such as pH, particle size, accessible surface area, porosity, moisture content,... were not
314 considered and may also impact biohydrogen and methane yields.

315

316 3.3.1 Compositional and structural features affecting biohydrogen production

317 PLS analysis led to equation 9 as a multi-linear model for biohydrogen potentials. The quality of the
318 model to predict hydrogen potential was confirmed by a high R^2 (0.87) and a low value of RMSEPC
319 (11.6 mL H₂ g⁻¹ TS).

$$320 \text{ BHP (mL H}_2 \text{ g}^{-1} \text{ TS)} = 19.43 + 1.84 * \text{SolSu (g.g}^{-1} \text{ TS)} - 0.36 * \text{Lig (g.g}^{-1} \text{ TS)} + 0.53 \text{Ua (g.g}^{-1} \text{ TS)} \\ 321 - 0.14 * \text{Cri (g.g}^{-1} \text{ TS)} - 0.05 \text{Am (g.g}^{-1} \text{ TS)} - 0.02 * \text{Pro (g.g}^{-1} \text{ TS)} \quad (9)$$

322 This model was validated using a set of two independent samples (sorghum 1 and sorghum 6)
323 which were not included in the calibration data set. Results are presented in Table 3. Hydrogen
324 potentials of 9.2 and 45.9 mL H₂ g⁻¹ TS were predicted compared to 9.7 and 37.9 mL H₂ g⁻¹ TS
325 measured respectively for sorghum 1 and 6. The REMSEPIV was calculated on the validation data
326 set and a promising result of 5.7 mL H₂ g⁻¹ TS was observed showing the high accuracy of the
327 model.

328 Another interest of the PLS models is to determine which variables significantly impact the
329 predicted variable. Centred and reduced weighted regression coefficients for hydrogen potentials
330 are shown in Figure 3a. A strong positive correlation was found between hydrogen potentials and
331 soluble sugars (SolSu) whereas all other studied variables (Lig, Am, Cri, Ua, Pro) had no significant
332 impact. Considering the correlation of hydrogen production versus only soluble carbohydrates, a
333 high correlation of $R^2 = 0.95$ was observed (data not shown). These results are in accordance with
334 Zhang et al., (2007) who suggested that hydrogen yield enhancement was due to an increase of
335 soluble sugar content of the substrate ³². Recently, Guo et al. (2011) found a similar correlation (R^2
336 = 0.87) between biohydrogen potentials and carbohydrates extracted under mild conditions (2 N
337 hydrochloric acid) ¹². Shi et al. (2012) suggested as well that the enhancement of hydrogen yields
338 nearly coincided with an increase in water soluble sugars available from alkali pretreated and raw
339 sweet sorghum stalks, which were 2.23 and 0.86 g L⁻¹, respectively ³⁰. Accordingly Pan et al.
340 (2011) showed that hydrolysis of cellulosic biomass led to an enhancement of hydrogen production
341 due to an increase of soluble compounds that were much easier to be degraded ³⁵. In addition, our
342 results showed that proteins did not affect hydrogen potentials within the range of studied protein
343 contents (< 30% TS – see Table 2). This is consistent with Guo et al. (2012) who reported that hydrogen
344 potentials were lower in the case of protein rich-substrates, but mainly because of their lower contents in
345 carbohydrates¹².

346 In addition, pH is an important factor that can affect biohydrogen production. Although the optimal
347 pH for hydrogen production from carbohydrates is rather acidic (about 4.5 - 6), alkaline pH (about
348 8.5-11) are more favourable for proteins-rich substrates ^{36,37}. In our study, the initial pH was set up
349 at 5.5 that can explain the absence of significant effect of the protein content. The absence of
350 significant positive correlation between amorphous holocelluloses and hydrogen potentials can be
351 explained by the poor efficiency of H₂-producing bacteria to assimilate directly cellulosic materials.

352 To achieve high yields of hydrogen from lignocellulosic substrates, an hydrolysis step is therefore
353 required³⁸.

354 355 3.3.2 Compositional and structural features affecting methane production

356 PLS analysis led to equation 10 as a multi-linear model for methane potentials. The quality of the
357 model to predict methane potential was confirmed by a high R^2 (0.88) and a low value of RMSEPC
358 (14.9 mL CH₄ g⁻¹ TS).

$$359 \text{BMP (mL CH}_4 \text{ g}^{-1} \text{ TS)} = 303.14 - 4.53*\text{Lig (g.g}^{-1} \text{ TS)} + 0.77*\text{SolSu (g.g}^{-1} \text{ TS)} + 1.28*\text{Pro (g.g}^{-1} \\ 360 \text{TS)} - 1.59*\text{Cri (g.g}^{-1} \text{ TS)} + 0.61\text{Am (g.g}^{-1} \text{ TS)} + 1.33\text{Ua (g.g}^{-1} \text{ TS)} \quad (10)$$

361 This model was validated using an independent set of two samples (sorghum 1 and sorghum 6)
362 which were not included in the calibration data set. Results are presented in Table 3. Errors of 0.2 %
363 and 4 % between the experimental and predictive methane potentials were observed for sorghum 1
364 and sorghum 6, respectively. The REMSEPIV was calculated on the validation data set, and a result
365 of 6.7 mL CH₄ g⁻¹ TS was observed showing the high accuracy of the models. Centred and reduced
366 regression coefficients for the prediction of methane potentials are presented in Figure 3b. In this
367 case, lignin, crystalline cellulose, soluble sugars, amorphous holocelluloses and proteins contents
368 were found to have a significant effect on methane potentials. A strong negative correlation was
369 found between the lignin content and the methane production which is in agreement with other
370 reported studies^{11,39-41}. Kobayashi et al. (2004) showed a strong negative correlation ($R^2 = 0.95$)
371 between the amount of methane produced and the amount of lignin of steam exploded bamboo⁴¹.
372 Triolo et al. (2011) also found a high negative correlation ($R^2 = 0.88$) between the lignin content
373 and methane potentials of energy crops and manure¹¹. However, our results led to a weak
374 correlation (R^2 of 0.82, data not shown) when considering only lignin content and methane

375 potentials. Consequently, anaerobic biodegradation of lignocellulosic materials into methane is not
376 only related to the lignin content, as suggested elsewhere ¹¹.

377 PLS regression showed that crystalline cellulose had also a negative impact on methane production
378 but in a lower extent than lignin. Zhu et al. (2009) showed that lignin content and crystallinity are
379 the two dominant parameters affecting negatively the digestibility of lignocellulosic substrates.
380 Moreover, they suggested that cellulose crystallinity could have a higher influence on short time
381 hydrolysis, whereas lignin content could have a higher impact on long time hydrolysis ⁴².

382 Additionally, a significant positive correlation was found between methane potentials and the contents
383 in soluble sugars, proteins and amorphous hemicelluloses in our study. According to Hayashi et al.
384 (2005), the readily accessible regions (amorphous regions) of the lignocellulosic biomass are more
385 efficiently hydrolyzed during enzymatic hydrolysis, resulting in the accumulation of crystalline
386 cellulose ⁴³. Similarly, Scherer et al., (2000) showed that the most degradable part of spent grains
387 corresponded to their soluble and hemicelluloses fractions, while cellulose and lignin were slightly
388 degraded ⁴⁴. Besides, giving a quick tool to predict biohydrogen and methane potentials from
389 lignocellulosic substrates, the PLS models built in this study are also valuable to give directions
390 towards the development of pretreatments strategies of lignocellulosic residues for enhancing both
391 biohydrogen and methane production. Pretreatments leading to the solubilisation of hollocelluloses
392 might be recommended for enhancing biohydrogen production whereas delignification, hollocelluloses
393 solubilisation and reducing crystalline cellulose may be recommended for methane production.

394 **Acknowledgements**

395 The authors are grateful to ADEME, the French Environment and Energy Management Agency, for
396 financial support in the form of F. Monlau's PhD grant and to Dr Solhy Abderrahim (INANOTECH
397 Rabat) for his help in DRX analysis.

398

- 400 (1) Kaparaju, P.; Serrano, M.; Thomsen, A. B.; Kongjan, P.; Angelidaki, I. Bioethanol,
401 biohydrogen and biogas production from wheat straw in a biorefinery concept. *Bioresource Technol* **2009**,
402 *100*, 2562-2568.
- 403 (2) Monlau, F.; Barakat, A.; Trably, E.; Dumas, C.; Steyer, J.-P.; Carrere, H. Lignocellulosic
404 materials into Biohydrogen and Biomethane: impact of structural features and pretreatment. *Crit. Rev. Env.*
405 *Sci. Tec.* **2012**, *in press*, DOI: 10.1080/10643389.2011.604258.
- 406 (3) Ren, N.; Wang, A.; Cao, G.; Xu, J.; Gao, L. Bioconversion of lignocellulosic biomass to
407 hydrogen: Potential and challenges. *Biotechnol. Adv.* **2009**, *27*, 1051-1060.
- 408 (4) Sain, M.; Panthapulakkal, S. Bioprocess preparation of wheat straw fibers and their
409 characterization. *Ind. Crop. Prod.* **2006**, *23*, 1-8.
- 410 (5) Guo, X. M.; Trably, E.; Latrille, E.; Carrere, H.; Steyer, J. P. Hydrogen production from
411 agricultural waste by dark fermentation: A review. *Int. J. Hydrogen Energ.* **2010**, *35*, 10660-10673.
- 412 (6) Li, D. M.; Chen, H. Z. Biological hydrogen production from steam-exploded straw by
413 simultaneous saccharification and fermentation. *Int. J. Hydrogen Energ.* **2007**, *32*, 1742-1748.
- 414 (7) Ntaikou, I.; Antonopoulou, G.; Lyberatos, G. Biohydrogen production from biomass and wastes
415 via dark fermentation: a review. *Waste and Biomass Valorization* **2010**, *1*, 21-39.
- 416 (8) Gunaseelan, V. N. Regression models of ultimate methane yields of fruits and vegetable solid
417 wastes, sorghum and napiergrass on chemical composition. *Bioresource Technol* **2007**, *98*, 1270-1277.
- 418 (9) Gunaseelan, V. N. Predicting ultimate methane yields of *Jatropha curcus* and *Morus indica* from
419 their chemical composition. *Bioresource Technol* **2009**, *100*, 3426-3429.
- 420 (10) Buffiere, P.; Loisel, D.; Bernet, N.; Delgenes, J. P. Towards new indicators for the prediction of
421 solid waste anaerobic digestion properties. *Water Sci. Technol.* **2006**, *53*, 233-241.
- 422 (11) Triolo, J. M.; Sommer, S. G.; Moller, H. B.; Weisbjerg, M. R.; Jiang, X. Y. A new algorithm to
423 characterize biodegradability of biomass during anaerobic digestion: Influence of lignin concentration on
424 methane production potential. *Bioresource Technol* **2011**, *102*, 9395-9402.
- 425 (12) Guo, X. M. Biohydrogen production and metabolic pathways in dark fermentation related to the
426 composition of organic solid waste. *PhD Thesis, University of Montpellier 2*, **2012**.
- 427 (13) Cheng, C. L.; Lo, Y. C.; Lee, K. S.; Lee, D. J.; Lin, C. Y.; Chang, J. S. Biohydrogen production
428 from lignocellulosic feedstock. *Bioresource Technol* **2011**, *102*, 8514-8523.
- 429 (14) Yuan, X. Z.; Shi, X. S.; Zhang, P. D.; Wei, Y. L.; Guo, R. B.; Wang, L. S. Anaerobic
430 biohydrogen production from wheat stalk by mixed microflora: Kinetic model and particle size influence.
431 *Bioresource Technol* **2011**, *102*, 9007-9012.
- 432 (15) Eleazer, W. E.; Odle, W. S.; Wang, Y. S.; Barlaz, M. A. Biodegradability of municipal solid
433 waste components in laboratory-scale landfills. *Environ. Sci. Technol.* **1997**, *31*, 911-917.
- 434 (16) Klimiuk, E.; Pokój, T.; Budzynski, W.; Dubis, B. Theoretical and observed biogas production
435 from plant biomass of different fibre contents. *Bioresource Technol* **2010**, *101*, 9527-9535.
- 436 (17) Fan, L.-T.; Lee, Y.-H.; Gharpuray, M. M. The nature of lignocellulosics and their pretreatments
437 for enzymatic hydrolysis. *Adv. Biochem. Eng.* **1982**, *23*, 156-187.
- 438 (18) Pakarinen, A.; Zhang, J.; Brock, T.; Maijala, P.; Viikari, L. Enzymatic accessibility of fiber
439 hemp is enhanced by enzymatic or chemical removal of pectin. *Bioresource Technol* **2012**, *107*, 275-281.
- 440 (19) (APHA), A. P. H. A. Standard Methods for the Examination of Water and Wastewater, 20th ed.
441 **1998**.
- 442 (20) Nguyen S.; Sophonputtanaphoca, S.; Kim, E.; Penner, M. Hydrolytic methods for the
443 quantification of fructose equivalents in herbaceous biomass. *Appl. Biochem. Biotechnol.* **2009**, *158*, 352-
444 361.
- 445 (21) Effland, M. J. Modified Procedure to Determine Acid-Insoluble Lignin in Wood and Pulp.
446 *Tappi* **1977**, *60*, 143-144.
- 447 (22) Petersson, A.; Thomsen, M. H.; Hauggaard-Nielsen, H.; Thomsen, A. B. Potential bioethanol
448 and biogas production using lignocellulosic biomass from winter rye, oilseed rape and faba bean. *Biomass*
449 *Bioenerg.* **2007**, *31*, 812-819.
- 450 (23) Spiridon, I.; Teaca, C. A.; Bodirlau, R. Structural Changes Evidenced by Ftir Spectroscopy in
451 Cellulosic Materials after Pre-Treatment with Ionic Liquid and Enzymatic Hydrolysis. *BioResources* **2010**,
452 *6*, 400-413.

- 453 (24) Segal, L.; Creely, J. J.; Martin Jr, A. E.; Conrad, C. M. An Empirical Method for Estimating the
454 Degree of Crystallinity of Native Cellulose Using the X-Ray Diffractometer. *Textile Res. J.* **1959**, *29*, 786-
455 794.
- 456 (25) Park, S.; Baker, J. O.; Himmel, M. E.; Parilla, P. A.; Johnson, D. K. Cellulose crystallinity
457 index: measurement techniques and their impact on interpreting cellulase performance. *Biotechnol.*
458 *Biofuels* **2010**, *3*, 10.
- 459 (26) Raju, C. S.; Lokke, M. M.; Sutaryo, S.; Ward, A. J.; Moller, H. B. NIR Monitoring of Ammonia
460 in Anaerobic Digesters Using a Diffuse Reflectance Probe. *Sensors* **2012**, *12*, 2340-2350.
- 461 (27) Thuesombat, P.; Thanonkeo, P.; Laopaiboon, L.; Laopaiboon, P.; Yunchalard, S.;
462 Kaewkannetra, P.; Thanonkeo, S. The batch ethanol fermentation of jerusalem artichoke using
463 *saccharomyces cerevisiae*. *Sci. Tech. J.* **2007**, *7*.
- 464 (28) Raposo, F.; Borja, R.; Rincon, B.; Jimenez, A. M. Assessment of process control parameters in
465 the biochemical methane potential of sunflower oil cake. *Biomass Bioenerg.* **2008**, *32*, 1235-1244.
- 466 (29) Fukushima, R. S.; Hatfield, R. D. Comparison of the Acetyl Bromide Spectrophotometric
467 Method with Other Analytical Lignin Methods for Determining Lignin Concentration in Forage Samples.
468 *J. Agr. Food Chem.*, **2004**, *52*, 3713-3720.
- 469 (30) Shi, X. X.; Song, H. C.; Wang, C. R.; Tang, R. S.; Huang, Z. X.; Gao, T. R.; Xie, J. Enhanced
470 bio-hydrogen production from sweet sorghum stalk with alkalization pretreatment by mixed anaerobic
471 cultures. *Int. J. Energy Res.* **2010**, *34*, 662-672.
- 472 (31) Fan, Y. T.; Zhang, Y. H.; Zhang, S. F.; Hou, H. W.; Ren, B. Z. Efficient conversion of wheat
473 straw wastes into biohydrogen gas by cow dung compost. *Bioresource Technol* **2005**, *97*, 500-505.
- 474 (32) Zhang, M. L.; Fan, Y. T.; Xing, Y.; Pan, C. M.; Zhang, G. S.; Lay, J. J. Enhanced biohydrogen
475 production from cornstalk wastes with acidification pretreatment by mixed anaerobic cultures. *Biomass*
476 *Bioenerg.* **2007**, *31*, 250-254.
- 477 (33) Dinuccio, E.; Balsari, P.; Gioelli, F.; Menardo, S. Evaluation of the biogas productivity potential
478 of some Italian agro-industrial biomasses. *Bioresource Technol* **2010**, *101*, 3780-3783.
- 479 (34) Frigon, J. C.; Guiot, S. R. Biomethane production from starch and lignocellulosic crops: a
480 comparative review. *Biofuels Bioprod. Bior.* **2010**, *4*, 447-458.
- 481 (35) Pan, C. M.; Ma, H. C.; Fan, Y. T.; Hou, H. W. Bioaugmented cellulosic hydrogen production
482 from cornstalk by integrating dilute acid-enzyme hydrolysis and dark fermentation. *Int. J. Hydrogen Energ.*
483 **2011**, *36*, 4852-4862.
- 484 (36) Cai, M. L.; Liu, J. X.; Wei, Y. S. Enhanced biohydrogen production from sewage sludge with
485 alkaline pretreatment. *Environ. Sci. Technol.* **2004**, *38*, 3195-3202.
- 486 (37) Xiao, B. Y.; Han, Y. P.; Liu, J. X. Evaluation of biohydrogen production from glucose and
487 protein at neutral initial pH. *Int. J. Hydrogen Energ.* **2010**, *35*, 6152-6160.
- 488 (38) Saratale, G. D.; Chen, S. D.; Lo, Y. C.; Saratale, R. G.; Chang, J. S. Outlook of biohydrogen
489 production from lignocellulosic feedstock using dark fermentation - a review. *J. Sci. Ind. Res.* **2008**, *67*,
490 962-979.
- 491 (39) Chandler, J. A.; Jewell, W. J.; Gossett, J.M.; Van Soest, P.J.; Robertson, J.B. Predicting
492 methane fermentation biodegradability. *Biotechnol. Bioeng. Symp.* **1980**, 93-107.
- 493 (40) Monlau, F.; Barakat, A.; Steyer, J. P.; Carrère, H. Comparison of seven types of thermo-
494 chemical pretreatments on the structural features and anaerobic digestion of sunflower stalks. *Bioresource*
495 *Technol* **2012**, *120*, 241-247.
- 496 (41) Kobayashi, F.; Take, H.; Asada, C.; Nakamura, Y. Methane production from steam-exploded
497 bamboo. *J. Biosci. Bioeng.* **2004**, *97*, 426-428.
- 498 (42) Zhu, L.; O'Dwyer, J. P.; Chang, V. S.; Granda, C. B.; Holtzapple, M. T. Multiple linear
499 regression model for predicting biomass digestibility from structural features. *Bioresource Technol.* **2010**,
500 *101*, 4971-4979.
- 501 (43) Hayashi, N.; Kondo, T.; Ishihara, M. Enzymatically produced nano-ordered short elements
502 containing cellulose I[β] crystalline domains. *Carbohydr. Polym.* **2005**, *61*, 191-197.
- 503 (44) Scherer, P. A.; Vollmer, G. R.; Fakhouri, T.; Martensen, S. Development of a methanogenic
504 process to degrade exhaustively the organic fraction of municipal "grey waste" under thermophilic and
505 hyperthermophilic conditions. *Water Sci. Technol.* **2000**, *41*, 83-91.
- 506
507

508 **Captions tables and figures**

509

510 Table 1: Correlations found in the literature between the compositional characteristics of
511 lignocellulosic substrates and biohydrogen or methane production

512

513 Table 2: Compositional and structural features of lignocellulosic substrates and validity range of PLS
514 models. Values correspond to the means of two replicates of independent values \pm standard deviations
515 (error bars).

516

517 Table 3: External validation of the PLS models for biohydrogen and methane potentials

518

519 Figure 1: Correlation between crystalline celluloses determined by IR and DRX (expressed in % TS)

520 Figure 2: Biochemical biohydrogen and methane potentials of lignocellulosic substrates. Values
521 correspond to the means of two replicates of independent values \pm standard deviations (error bars).

522

523 Figure 3: Centred and reduced regression coefficients with their 95% confidence intervals for the
524 prediction of biohydrogen potentials (a) and methane potentials (b).

525

526

527

528

529

530 Table 1: Correlations found in the literature between the compositional characteristics of lignocellulosic substrates and biohydrogen or
531 methane production

Fermentation process	Biomass used	Compositional features	Equation	References
Biohydrogen	organic solid substrates (n=21)	Soluble Carbohydrates (Carb)	$BHP (mL H_2 \cdot g^{-1} TS) = 1.31 + 199.46 \text{ Carb}$	12
Methane	Manure (n=10), Energy crops (n=10)	Lignin (Lig)	$BMP (L CH_4 \cdot kg^{-1} VS) = -1.67 * Lig + 421.7$	11
Methane	Raw and thermo-chemically pretreated sunflower stalks (n= 8)	Lignin (Lig)	$BMP (L CH_4 \cdot kg^{-1} VS) = -0.65 * Lig + 379.8$	40
Methane	Lignocellulosic residues (n=7)	Soluble carbohydrates (Carb), acid detergent fiber (ADF), Protein (Pro), Lignin (Lig), Ash (A)	$BMP (L CH_4 \cdot kg^{-1} VS) = 0.18 + 0.48 * CaRB + 0.2 * ADF - 0.003 * Lig/ADF + 2.8 Pro - 0.83 * A$	8
Methane	Lignocellulosic residues (n=12)	Soluble carbohydrates (Carb), acid detergent fiber (ADF), Protein (N), Ash (A), Lipids (F)	$BMP (L CH_4 \cdot kg^{-1} VS) = 0.045 + 1.23 * Carb + 0.24 * Pro + 1.51 * F - 0.68 * ADF - 0.81 * Cell - 6.1 * A$	9
Methane	Lignocellulosic residues (n=15)	Lignin (Lig)	$Biodegradability (\%MV) = 0.83 - 1.82 * Lig$	39
Methane	Municipal solid waste (n=2), agricultural residues (n=2), manure (n=4), vegetables (n=6)	Lignin (Lig), cellulose (Cell)	$Biodegradability (\%DCO) = 0.87 - 1.03 (\% \text{ lignin} + \% \text{ cellulose})$	10

532

533 Table 2: Compositional and structural features of lignocellulosic substrates and validity range of PLS models. Values correspond to the means of
 534 two replicates of independent values \pm standard deviations (error bars).
 535
 536

Substrates	% TS	Chemical composition (% TS)						FT-IR spectra		
		SolSu	Pro	Ua	Hem	Cell	Lig	LOI	Cri (% TS)	Am (% TS)
Rice straw	0.96	0.8	5.3 (\pm 0.2)	0.6 (\pm 0.1)	18.8 (\pm 0.7)	26.2 (\pm 0.5)	27 (\pm 2.6)	0.85	12	33
Giant reed stalks	0.99	0.3	4.3 (\pm 0.7)	0.2 (\pm 0.0)	18.5 (\pm 0.7)	33.1 (\pm 1.3)	24.5 (\pm 0.1)	1.03	16.8	34.8
Giant reed leaves	0.99	2.9	8 (\pm 0.1)	0.7 (\pm 0.1)	8 (\pm 0.1)	20.9 (\pm 0.6)	25.4 (\pm 0.1)	0.95	10.2	28.4
Sunflower stalks 1	0.94	0	4.8 (\pm 0.1)	7.0 (\pm 0.6)	15.6 (\pm 0.3)	31 (\pm 1.6)	29.2 (\pm 1.6)	1.22	17	29.6
Sunflower stalks 2	0.96	0	2.3 (\pm 0.4)	3.9 (\pm 0.4)	14.3 (\pm 2.4)	31.2 (\pm 3.1)	27.7 (\pm 0.2)	1.2	17	28.4
Sunflower stalks 3	0.96	0	4.3 (\pm 0.7)	2.4 (\pm 0.2)	14.3 (\pm 0.7)	31.2 (\pm 0.7)	30 (\pm 1.7)	1.18	16.9	28.6
Sunflower stalks bark	0.97	0	2.8 (\pm 0.4)	1.7 (\pm 0.3)	13.5 (\pm 0.2)	27.4 (\pm 0.4)	35 (\pm 0.4)	1.1	14.4	26.5
Sunflower oil cakes	0.94	5.2	29.7 (\pm 3.4)	1.4 (\pm 0.2)	8.2 (\pm 0.2)	5.1 (\pm 0.3)	22.3 (\pm 2.8)	0.96	3.8	12.1
Maize stalks	0.99	0.4	7.4 (\pm 0.1)	0.7 (\pm 0.1)	21.2 (\pm 0.6)	27.1 (\pm 0.9)	23.2 (\pm 0.1)	1.14	14.5	33.9
Maize leaves	0.99	0.3	6.7 (\pm 0.6)	1.0 (\pm 0.2)	28.6 (\pm 3.3)	30.9 (\pm 3.1)	20.4 (\pm 0.6)	1.03	15.7	43.8
Maize cobs	0.96	0.2	4.3 (\pm 0.3)	0.7 (\pm 0.1)	34.6 (\pm 1.4)	29.8 (\pm 1.2)	19.2 (\pm 1.0)	0.89	14	50.3
Jerusalem artichoke stalks	0.96	32.9	2.8 (\pm 0.3)	0.2 (\pm 0.0)	8.8 (\pm 3.1)	9.6 (\pm 3.1)	20.3 (\pm 0.0)	0.95	4.7	13.7
Jerusalem artichoke leaves	0.94	2.6	12.4 (\pm 0.3)	0.7 (\pm 0.1)	4.7 (\pm 0.6)	8.8 (\pm 1.4)	12.9 (\pm 1.3)	1.22	4.8	8.6
Jerusalem artichoke tubers	0.98	59.1	10.4 (\pm 0.2)	1.5 (\pm 0.2)	5 (\pm 0.0)	5.4 (\pm 0.3)	12.3 (\pm 0.1)	1.1	2.8	7.5
Sorghum 1	0.95	0.4	4.6 (\pm 0.1)	0.9 (\pm 0.1)	26.1 (\pm 0.1)	29.1 (\pm 0.3)	22.5 (\pm 1.6)	1.09	15.2	40
Sorghum 2	0.91	15.4	6.5 (\pm 0.0)	0.6 (\pm 0.1)	19.4 (\pm 1.3)	22.2 (\pm 1.5)	21.4 (\pm 0.3)	1.05	11.4	28.3
Sorghum 3	0.91	18.5	8.1 (\pm 0.0)	1.0 (\pm 0.0)	20.9 (\pm 1.6)	20.1 (\pm 1.7)	18.5 (\pm 0.9)	0.98	10.3	27.9
Sorghum 4	0.94	8.2	8.2 (\pm 0.0)	0.6 (\pm 0.0)	21.7 (\pm 0.2)	18.3 (\pm 5.8)	20.7 (\pm 3.0)	1.03	9.3	27.7
Sorghum 5	0.92	22.8	6.9 (\pm 0.0)	0.6 (\pm 0.0)	20 (\pm 1.2)	19.7 (\pm 0.2)	19.8 (\pm 1.3)	1.11	10.4	26.2
Sorghum 6	0.88	21.3	6.2 (\pm 0.0)	0.6 (\pm 0.0)	18.5 (\pm 0.8)	18.1 (\pm 0.1)	21.3 (\pm 0.0)	1.1	9.5	24.4
<i>Validity range</i>		<i>0-59.1</i>	<i>2.3-29.7</i>	<i>0.2-7</i>	<i>4.7-34.6</i>	<i>5.4-33.1</i>	<i>12.3-35</i>		<i>2.5-16.3</i>	<i>7.5-50.3</i>

537

538

539

Table 3: External validation of the PLS models for biohydrogen and methane potentials

Independent samples	PLS model for biohydrogen potentials				PLS model for methane potentials			
	BHP measured mL H ₂ .g ⁻¹ TS	BHP predicted mL H ₂ .g ⁻¹ TS	Errors	RMSEPiv	BMP measured mL CH ₄ .g ⁻¹ TS	BMP predicted mL CH ₄ .g ⁻¹ TS	Errors	RMSEPiv
Sorghum 1	9.7	9.2	4.5%	5.7	209.5	209	0.2%	6.7
Sorghum 6	37.9	45.9	21.1%		240	230.5	4.0%	

540

541

542

543

544

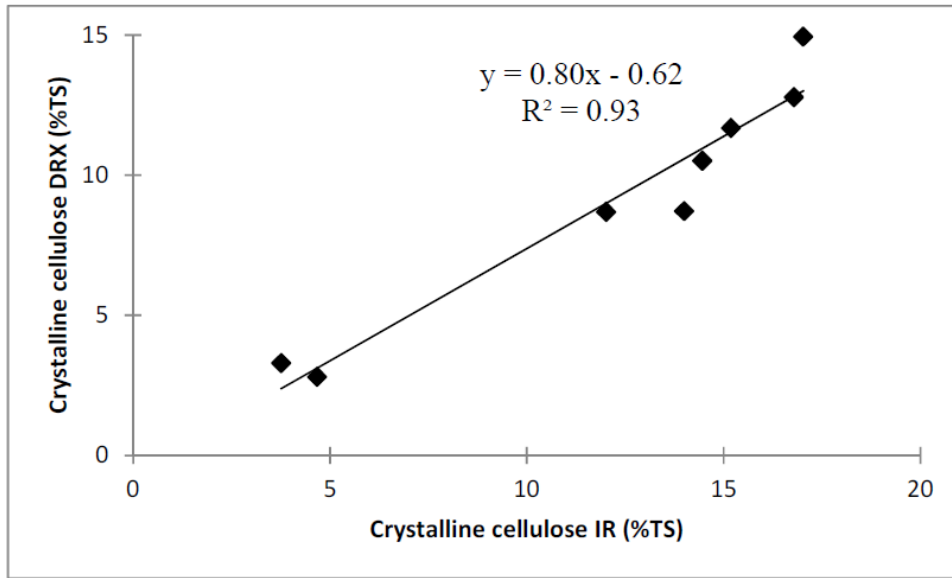


Figure 1: Correlation between crystalline celluloses determined by IR and DRX (expressed in % TS)

545

546

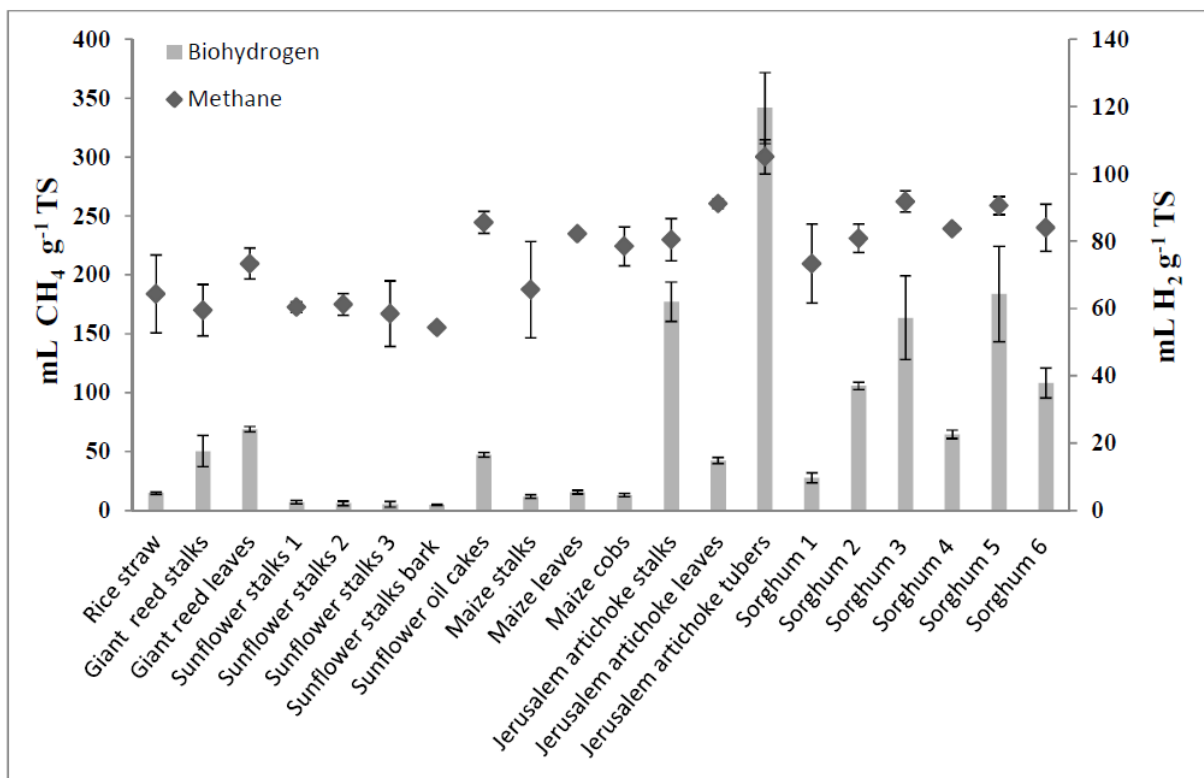


Figure 2: Biochemical biohydrogen and methane potentials of lignocellulosic substrates. Values correspond to the means of two replicates of independent values \pm standard deviations (error bars).

547

548

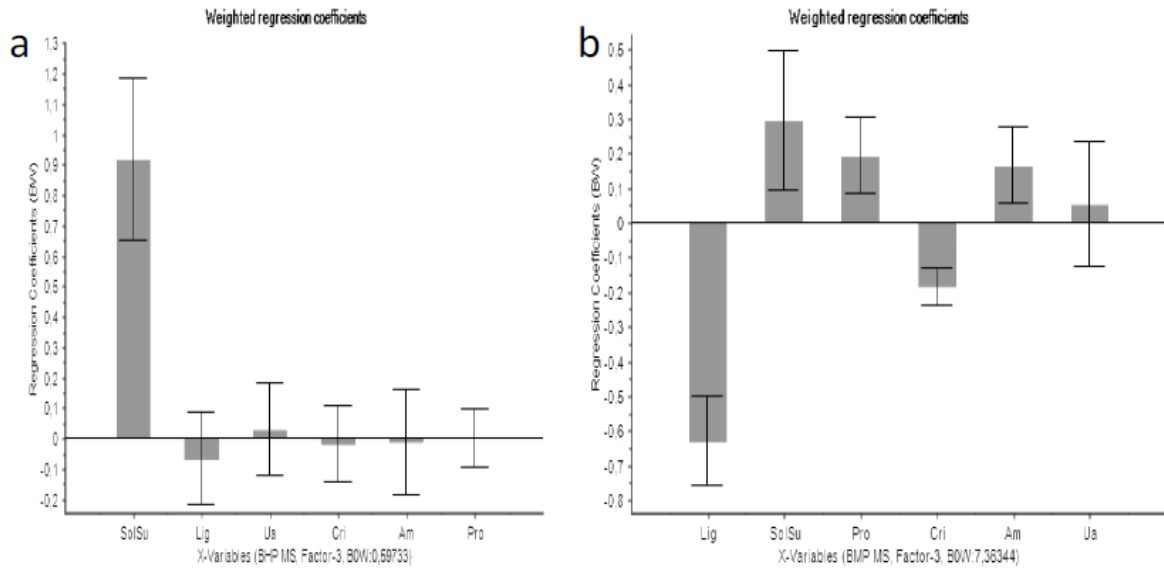


Figure 3: Centred and reduced regression coefficients with their 95% confidence intervals for the prediction of biohydrogen potentials (a) and methane potentials (b).

SolSu= soluble carbohydrates, Lig=lignin Ua= uronic acids, Cri= Crystalline cellulose, Am= amorphous holocelluloses, Pro=proteins

549

550

