

# Direct evidence of $\alpha$ ester linkages between lignin and glucuronoxylan that reveal the robust heteropolymeric complex in plant cell walls

Hiroshi Nishimura (✉ [nishimura.hiroshi.j03@kyoto-u.jp](mailto:nishimura.hiroshi.j03@kyoto-u.jp))

Kyoto University

**Kazuma Nagata**

Kyoto University

**Takashi Nagata**

Kyoto University

**Masato Katahira**

Kyoto University

**Takashi Watanabe**

Kyoto University

---

## Article

### Keywords:

**Posted Date:** February 8th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1327348/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

Nature Plants

**Direct evidence of  $\alpha$  ester linkages between lignin and glucuronoxylan that reveal the robust heteropolymeric complex in plant cell walls**

Hiroshi Nishimura\*<sup>1</sup>, Kazuma Nagata<sup>2</sup>, Takashi Nagata<sup>2</sup>, Masato Katahira<sup>2</sup> and Takashi Watanabe\*\*<sup>1</sup>

<sup>1</sup>Research Institute for Sustainable Humanosphere (RISH), Kyoto University, Uji 611-0011, Japan.

<sup>2</sup>Institute of Advanced Energy (IAE), Kyoto University, Uji 611-0011, Japan.

\* nishimura.hiroshi.j03@kyoto-u.jp

\*\* twatanab@rish.kyoto-u.ac.jp

**One Sentence Summary:**

Conclusive evidence of  $\alpha$ -ester linkages between lignin and glucuronoxylan was obtained by advanced multi-dimensional NMR.

18 **Abstract**

19 All vascular plants have lignin that interacts with polysaccharides in cell walls; however, the  
20 chemical structure of covalent lignin-carbohydrate (L-C) linkages remains unresolved because of  
21 the small number of L-C linkage relative to the polymer main chain. Sample isolation that  
22 predominantly maintains intact L-C linkages followed by advanced multi-dimensional NMR  
23 showed that the critical L-C linkage between lignin and xylan involves the  $\alpha$ -position of the  $\beta$ -  
24 aryl ether unit of lignin and the sixth position of the glucuronic acid residue of 4-*O*-methyl  
25 glucuronoxylan. The L-C ester linkages occur primarily at the  $\alpha$ -position of lignin, supporting a  
26 theoretical synthetic pathway through a quinone methide intermediate. In contrast, the  
27 conventional  $\gamma$ -L-C ester was undetectable, indicating that this linkage is not the primary type.  
28 Instead, the  $\gamma$ -acetyl-substituted lignin coexisted. The results on primary L-C linkages contribute  
29 to plant physiology and biology and advance research into biomass conversion.

30

31 Structural elucidation of lignin-carbohydrate (L-C) linkages is a central challenge for  
32 understanding the mechanism of plant cell wall formation and for strategic molecular breeding of  
33 plants with controlled degradability and rigidity, which are key factors considered in plant  
34 biomass conversion and wood utilisation. Here we show the principal ester linkage between  
35 lignin and xylan (hemicellulose) involves the  $\alpha$ -position of lignin  $\beta$ -ether units rather than the  
36 previously proposed  $\gamma$ -position. Glucuronoxylan forms a covalent bond with lignin to yield an  
37 amorphous three-fold screw conformation. In lignin-carbohydrate complexes (LCCs), acetyl-  
38 substitution at the 3-position of xylan and glucuronic-substitution at the 2-position of xylan is the  
39 main side-chain substitution (Fig. 1). We discuss the agreement between a theoretical synthetic  
40 pathway via the quinone methide intermediate and the difficulty of rearranging  $\alpha$ - to  $\gamma$ - in natural  
41 macromolecules.

42 Plant biomass is the most abundant natural polymer on earth. Conversion and use of plant  
43 biomass to produce value-added chemicals, biofuels, and materials are indispensable to a  
44 sustainable society<sup>1</sup>. The plant cell wall is principally composed of cellulose, hemicelluloses,  
45 and lignin. Lignin is an aromatic polymer that is universally conserved in vascular plants. The  
46 plant cell wall formation starts with the biosynthesis of cellulose microfibrils. Hemicelluloses  
47 are assembled into microfibrils with different orientations. Lignin is later deposited into the  
48 polysaccharide matrix by dehydrogenative polymerisation of monolignols. Lignin is a  
49 hydrophobic polymer that binds covalently to hydrophilic polysaccharides, hemicelluloses, to  
50 form an amphipathic structure called the lignin-carbohydrate complex (LCC). This complex  
51 gives the plant cell walls rigidity by interfacing the polymers with different physicochemical  
52 properties<sup>2</sup>. LCCs have potential use as amphiphilic polymers, UV absorbers<sup>3-4</sup>, and bio-active

53 substances<sup>5-7</sup> in practical applications. Anti-virus<sup>8-10</sup> activities of LCCs have interesting  
54 physiological activities.

55 Ester bonds between glucuronoxylan and lignin are the major L-C linkage in plant cell walls,  
56 especially in hardwood and herbaceous plants<sup>2</sup>. The complex structure of hemicelluloses and  
57 lignin give rise to the robustness and flexibility of plant cell walls<sup>11,12</sup>. Bundles of cellulose  
58 microfibrils are coated with xylan-lignin complexes and cross-linked by glucomannan in wood  
59 cell walls<sup>12,13</sup>. Acetyl groups distributed in the xylan main chain increase affinity to cellulose by  
60 controlling the helical structure of xylan. The two-fold screw xylan is spatially close to  
61 cellulose, adheres to the hydrophilic surface of cellulose microfibrils, and has the same rigidity  
62 as cellulose. Conversely, three-fold screw xylan is more amorphous and probably interacts with  
63 lignin<sup>14-17</sup>. The hydrophobicity of acetyl groups also increases affinity to lignin.

64 Previous studies have revealed covalent linkages between glucuronoxylan and lignin, e.g., IR<sup>18</sup>,  
65 alkaline degradation<sup>19</sup>, sodium borohydride reduction<sup>20</sup>, and DDQ oxidation<sup>21,22</sup>. Recently,  
66 microbial glucuronic acid esterases were used to cleave the L-C ester bonds to facilitate  
67 disintegration of plant cell walls in biomass conversion applications<sup>23-30</sup>. Using specific  
68 enzymes that cleave the ester bond is an attractive approach for both characterising L-C bonds  
69 and biomass conversion; however, definitive data defining the binding sites have not been  
70 reported.

71 Two-dimensional (2D) NMR is a promising approach to identify covalent linkages between  
72 lignin and carbohydrates. Currently, there are four different L-C bonds,  $\alpha$ -ether (benzyl ether),  
73  $\alpha$ -ester (benzyl ester),  $\gamma$ -ester, and phenyl glycoside, have been reported based on <sup>1</sup>H-<sup>13</sup>C  
74 heteronuclear single quantum coherence (HSQC) correlations<sup>31-37</sup>. As for ester-type L-C  
75 linkages, Balakshin et al.<sup>31,32</sup> proposed a significant presence of  $\gamma$ -esters instead of  $\alpha$ -esters in

76 pine LCC. In contrast, Yuan et al.<sup>37</sup> reported that the  $\gamma$ -ester signal in 2D NMR data of poplar  
77 LCC was indeterminable. The conclusions of these studies are based on a comparison of  
78 chemical shifts between model compounds and LCC fractions. However, the chemical shifts of  
79 the relevant signals are not identical to those of synthetic model compounds. Furthermore, the  
80 possibility of misassignment of the signals because of severe overlap with signals from other  
81 unidentified compounds cannot be excluded. Thus, the purity, mild LCC sample preparation  
82 conditions, and long-range correlated NMR experiments are necessary to solve these issues, as  
83 reported for  $\beta$ -ether bonds between lignin and acetylglucomannan<sup>38</sup>.

84

## 85 **Results**

86 This study determined the principal ester linkages between lignin and xylan (hemicellulose) at  
87 the  $\alpha$ -position of lignin  $\beta$ -ether units. A conceptual drawing of lignocellulosic fibre in hardwood  
88 cell walls focusing on the covalent L-C  $\alpha$ -ester linkage and a chemical structural model are  
89 shown in Fig. 1.

90 **LCC sample preparation.** Fractionation procedures under mild conditions are critical because  
91 L-C ester linkages in extracted LCCs are susceptible to denaturation, including acyl migration.  
92 Here, we avoided conditions, such as strong mechanical milling, metal ion contaminations, high  
93 temperature, and strong acid or alkaline treatments. The isolation scheme is shown in Extended  
94 Data Fig. 1. (see method section and caption in Extended Data Fig. 1). Wood powder was ball-  
95 milled in an agate pot, filled with nitrogen under the control of 60°C (Extended Data Fig. 2).  
96 The crude extracted "LCC-DMSO" was treated with hemicellulase. The soluble liquid fraction  
97 "LCC-EL" showed the L-C NMR signals followed by purification, however, too small quantities  
98 for advanced NMR experiments. Fortunately, "LCC-ER-AL80", extracted with methanol/water  
99 (8/2, v / v) from the precipitate fraction "LCC-ER", was an ideal LCC separated using the  
100 solubility differences.

101 **Determination of the L-C  $\alpha$ -ester linkage structure by multi-dimensional NMR.** The LCC-  
102 ER-AL80 was analysed by  $^1\text{H}$ - $^{13}\text{C}$  HSQC NMR (Fig. 2). We succeeded in assigning signals to  
103  $\alpha$ -ester-type LCC by multi-dimensional NMR analysis (Extended Data Table. 1). Signals related  
104 to the L-C linkage between the  $\alpha$ -position of the  $\beta$ -aryl ether unit of the lignin and the 6' of the  
105 glucuronic acid residue of 4'-*O*-methyl glucuronoxylan are shown in Fig. 2b.

106 The  $\alpha$ -CH correlation of  $\alpha$ -ester-type LCC in HSQC spectra is composed of three signals, I- $\alpha$   
107 ( $\delta_{\text{H}}/\delta_{\text{C}} = 5.88 \text{ ppm}/74.4 \text{ ppm}$ ), II- $\alpha$  ( $\delta_{\text{H}}/\delta_{\text{C}} = 6.08 \text{ ppm}/75.2 \text{ ppm}$ ), and III- $\alpha$  ( $\delta_{\text{H}}/\delta_{\text{C}} = 6.00$   
108  $\text{ppm}/75.5 \text{ ppm}$ ) (Fig. 3a). The  $\beta$ -CH correlations of  $\alpha$ -ester-type LCC overlapped with the signals  
109 from the  $\beta$ -position of the lignin  $\beta$ -O-4' subunit in the HSQC spectrum (Fig. 3a). We succeeded  
110 in assigning the  $\beta$ -position signals accurately using a 3D TOCSY-HSQC spectrum, I- $\beta$  ( $\delta_{\text{H}}/\delta_{\text{C}} =$   
111  $4.48 \text{ ppm}/82.9 \text{ ppm}$ ) (Fig. 3b), II- $\beta$  ( $\delta_{\text{H}}/\delta_{\text{C}} = 4.28 \text{ ppm}/83.7 \text{ ppm}$ ) (Fig. 3c), and III- $\beta$  ( $\delta_{\text{H}}/\delta_{\text{C}} =$   
112  $4.28 \text{ ppm}/83.9 \text{ ppm}$ ) (Fig. 3d).

113 The  $\gamma$ -CH<sub>2</sub> correlations of  $\alpha$ -ester-type LCC were determined from 3D TOCSY-HSQC (Fig.  
114 3e) and 2D <sup>1</sup>H-<sup>13</sup>C HMBC (HMBC) (Fig. 3f) spectra as  $\gamma_1$  ( $\delta_{\text{H}}/\delta_{\text{C}} = 3.55 \text{ ppm}/62.0 \text{ ppm}$ ).  $\gamma_2$   
115 ( $\delta_{\text{H}}/\delta_{\text{C}} = 3.83 \text{ ppm}/62.0 \text{ ppm}$ ) was assigned using HSQC-TOCSY and HMBC spectra. In the  
116 HMBC spectrum, correlation signals between the  $\alpha$ -proton of the L-C ester to the  $\beta$ - and  $\gamma$ -  
117 carbons of the lignin  $\beta$ -aryl ether unit and syringyl aromatic ring (S<sub>1</sub> and S<sub>2,6</sub>) were observed  
118 (Fig. 3f). The aromatic ring proximity to the  $\alpha$ -position was mainly the syringyl lignin.  
119 Although weak, a signal representing a correlation between LC $\alpha$ -est-II $\alpha$  and the guaiacyl lignin  
120 unit was also observed (Fig. 3f).

121 A significant correlation between the  $\alpha$  proton of  $\alpha$ -ester-type LCC and the carbonyl carbon ( $\delta_{\text{C}}$   
122  $= 169.7 \text{ ppm}$ ) was observed (Fig. 3f). This carbonyl carbon resonates upfield when compared  
123 with the carbonyl carbon of the acetylated derivative of lignin ( $\delta_{\text{C}} \approx 170.2 \text{ ppm}$ ). Combined with  
124 the correlation data described in Fig. 6, this correlation was determined to be the carbonyl carbon  
125 at the 6'-position of glucuronic acid.

126 The 4'-*O*-methyl glucuronoxylan moiety of  $\alpha$ -ester-type LCC was assigned by through-bond  
127 correlations from 1' to 5', in the 3D TOCSY-HSQC spectrum (Fig. 4 **a-f**, see caption). A critical  
128 correlation indicating the presence of the  $\alpha$ -ester L-C linkage is the correlation between the  $\alpha$ -CH  
129 of the lignin  $\beta$ -*O*-4' subunit and the ester carbonyl (C=O) at the 6-position of the glucuronic acid  
130 residue in 4'-*O*-methyl-glucuronoxylan (Fig. 5). Starting from the  $\alpha$ -CH correlation signal of  $\alpha$ -  
131 ester-type LCC shown in the HSQC spectrum (Fig. 5**a**), the  $\alpha$ -CH proton ( $\delta_{\text{H}} = 5.9$  and 6.0 ppm)  
132 was correlated to the carbonyl carbon of glucuronoxylan, 6'-GlcA ( $\delta_{\text{C}} = 169.7$  ppm) in the  
133 HMBC spectrum (Fig. 5**b**). Additionally, the L-C ester carbonyl (C=O,  $\delta_{\text{C}} = 169.7$  ppm)  
134 correlated to 5'-CH of glucuronate (5'-GluA,  $\delta_{\text{H}} = 4.80$  ppm) in the HMBC spectrum (Fig. 5**b**),  
135 which confirms the  $\alpha$ -ester L-C linkage to 4'-*O*-methyl-glucuronoxylan.

136 Notably, the chemical shift (4.80 ppm) of the 5'-GluA hydrogen was near identical to the  
137 chemical shift of the acetyl-substituted xylan at the 3-position (Xyl-3-Ac), as shown in Fig. 5**a**.  
138 We should discriminate between  $\alpha$ -ester-type LCC and acetyl-substituted xylan at the 3-position  
139 (Xyl-3-Ac) before drawing the final conclusion. Correlations derived from  $^2J_{\text{CH}}$  and  $^3J_{\text{CH}}$  spin-  
140 couplings are observed in HMBC, whereas, in 1,1-ADEQUATE, only  $^2J_{\text{CH}}$  spin-couplings are  
141 observed<sup>39</sup>. The correlation signal between the hydrogen of 5'-GluA and ester carbonyl (C=O)  
142 was observed in the 1,1-ADEQUATE spectrum (Fig. 5**c**), which confirms  $\alpha$ -ester-type LCC and  
143 excludes the possibility of the alternative acetyl-substituted xylan at the 3-position. Correlations  
144 ( $\sim\delta_{\text{H}}/\delta_{\text{C}} = 4.8$  ppm/170.4 ppm) derived from the acetyl-substituted xylan at the 3-position (Xyl-3-  
145 Ac) as a  $^3J_{\text{CH}}$  coupling were observed in the HMBC spectrum (Fig. 5**b**).

146  $\gamma$ -Ester-type LCC was proposed by Balakshin et al. The signals at  $\delta_{\text{H}}/\delta_{\text{C}} = 62$ –65 ppm/4.0–4.5  
147 ppm were attributed to  $\gamma$ -ester-type LCC<sup>32</sup>. For LCC-ER-AL80, correlations between protons in

148 the region of 4.0–4.5 ppm and carbonyl carbons ( $\delta_c = 170.2$  ppm) were observed (Fig. 5b, peaks  
149 indicated by beige arrow). This observation suggests the presence of  $\gamma$ -ester-type LCC;  
150 however, this is incorrect, as explained below. The correlations between protons in the 4.0–4.5  
151 ppm region and carbonyl carbons (170.2 ppm) are not observed in 1,1-ADEQUATE (Fig. 5c).  
152 This observation indicates that the correlations observed in the HMBC spectrum (Fig. 5b) are not  
153 derived from  $^2J_{CH}$  spin couplings but  $^3J_{CH}$  spin couplings. Therefore, these correlations observed  
154 in the HMBC spectrum (Fig. 5b, beige arrow) do not originate from  $\gamma$ -ester-type LCC in which  
155  $^2J_{CH}$  spin couplings are present, but from  $\gamma$ -acetyl lignin in which  $^3J_{CH}$  spin coupling should be  
156 present. Thus, the presence of  $\gamma$ -ester type LCC is refuted at the detectable level of long-range  
157 correlated NMR.

158 The above results showed that we successfully assigned the ester linkage between the  $\alpha$ -  
159 position of the  $\beta$ -O-4' subunit of lignin and the 6-position of the glucuronic acid residue in 4'-O-  
160 methyl-glucuronoxylan.

161

## 162 **Discussion**

163 Plant secondary cell walls are a highly complex composite of cellulose, hemicelluloses (mainly  
164 xylan and glucomannan), and lignin. The supramolecular network of these macromolecules  
165 renders the cell walls with sufficient mechanical strength and rigidity. LCCs serve as a  
166 compatibiliser between hydrophobic lignin and hydrophilic polysaccharides in the cell wall  
167 network. Thus, L-C linkages restrict the degradation of plant cell walls by chemical and  
168 enzymatic reactions. Considerable research has been devoted to developing versatile pre-  
169 treatment methods and genetically engineered plants with cell wall structures suitable for  
170 biomass conversion. The efficient conversion of plant biomass based on an accurate  
171 understanding of the cell wall structure is essential for lignocellulosic biorefinery.

172 Lignin binds to polysaccharides through covalent and non-covalent linkages that have been  
173 studied in detail using solid-state NMR<sup>16, 17</sup>, and the interaction between cellulose and xylan  
174 probably involves a hydrophobic surface formed by a two-fold helical screw symmetry with a  
175 regular pattern of xylan substituted by acetate or glucuronic acid<sup>17</sup>. In stark contrast to the strong  
176 interaction between xylan and lignin, direct interactions between cellulose and lignin are weak  
177 and not a major factor responsible for the strength of cell walls. Solid-state NMR and simulation  
178 results showed that lignin is closely associated with xylan by electrostatic and hydrophobic  
179 interactions and that a distorted xylan structure is advantageous for lignin binding<sup>16, 17</sup>.

180 Previous studies with model compounds have assigned the linkage between glucuronic acid and  
181 the lignin dimer<sup>40-42</sup>; however, NMR chemical shifts of signals representing these model  
182 compounds differ from those of the natural polymer. Recently, an ester-type L-C linkage at the

183  $\alpha$ -position was reported<sup>43</sup>. However, the correlation with the carbonyl carbon in the HMBC  
184 spectrum was around 196 ppm, indicating a correlation with an aldehyde rather than glucuronic  
185 acid. Thus, this correlation is not derived from  $\alpha$ -ester type LCC and represents a different  
186 chemical moiety.

187 Here, we clearly showed the existence of L-C linkages between the  $\alpha$ -position of the  $\beta$ -O-4'  
188 lignin unit, which is the main structure of lignin, and the 6-position of glucuronic acid of  
189 glucuronoxylan. Preventing artificial denaturation of the sample during the ball-milling step was  
190 critical in the sample preparation process. Thus, we performed ball-milling under mild  
191 conditions by maintaining the temperature under 60°C, replacing the air with a nitrogen  
192 atmosphere after evacuation, and using an agate stone pot. The crude sample, LCC-DMSO, gave  
193 rise to the C–H correlation signal at the  $\alpha$ -position of the L-C-ester in the HSQC spectrum,  
194 indicating that  $\alpha$ -ester-type LCC exists in the natural wood cell wall (Extended Data Fig. 3).

195 Interestingly, this L-C-ester signal was not seen in samples treated by vibro-ball-milling in a  
196 stainless-steel pot for 24 h, indicating mild mechanical processing and non-denaturing extraction  
197 and fractionation are critical for isolating  $\alpha$ -ester-type LCC.

198 The  $\beta$ -position correlation signal of LC-ester overlapped with other signals. Thus, a 3D  
199 TOCSY-HSQC spectrum was used to overcome this issue. We succeeded in identifying the  $\alpha$ -  
200 position and  $\beta$ -position of at least three LC-ester-types with different chemical shifts (Fig. 3a-d)  
201 by examining the  $\alpha$ (H)-position of the LC-ester in the 3D TOCSY-HSQC spectrum. Differences  
202 in chemical shifts arose primarily from differences in diastereomers (*threo* and *erythro*). HMBC  
203 correlations (Fig. 3f) demonstrated that the  $\alpha$ -position of the L-C-ester locates proximal to the  
204 syringyl lignin aromatic unit. The HMBC correlation between the  $\alpha$ -position H of the LC-ester

205 and the carbon at the 1-position of the lignin aromatic ring resonated at 134.7 ppm ( $\delta_{\text{H}}/\delta_{\text{C}} = 6.00$   
206 ppm/134.7 ppm) in addition to the syringyl nucleus (S1) found at 132.7 ppm ( $\delta_{\text{H}}/\delta_{\text{C}} = 5.89$   
207 ppm/132.7 ppm), derived from the phenolic terminal unit or the guaiacyl nucleus.

208 Xylans are substituted at the 2- or 3-position with acetyl groups, and some are acetylated at  
209 both the 2- and 3-positions. The pattern of this modification has been identified to occur at even  
210 numbers, and the two-fold screw structure plays an essential role in the interaction with  
211 cellulose<sup>16, 17, 44</sup>. HSQC spectra (Fig. 2, Extended Data Fig. 3a, c) of LCC-ER-AL80 clearly  
212 showed a lower frequency of acetylation of glucuronoxylan than the crude extracts, LCC-DMSO  
213 (Extended Data Fig. 3b). In addition, acetylation occurred primarily at the 3-position of xylan,  
214 whereas modification of the glucuronic acid side chain occurred at the 2-position of xylan  
215 (Extended Data Fig. 3c).

216 This xylan substituent pattern supports the structure that facilitates the formation of threefold  
217 helical screw conformations<sup>45</sup>. From molecular dynamics simulations, acetylation at the 2-  
218 position stabilises the two-fold helical screw conformation, and xylan interacts tightly with  
219 cellulose, whereas acetylation at the 3-position takes an amorphous structure as the three-fold  
220 helical screw conformation and xylan interacts readily with lignin<sup>45</sup>.

221 Ester bonds between lignin and the 4-*O*-methyl glucuronic acid residue in xylan are formed by  
222 nucleophilic addition of the carboxyl group to a quinone methide intermediate formed by the  
223 radical coupling of  $\beta$ - and phenoxy radicals of monolignols and those from growing lignin  
224 chains. As expected from the resonance structure, the  $\alpha$ -position of the quinone methide is  
225 electrophilic and reacts with the nucleophile, the carboxyl group of 4-*O*-methyl-glucuronic acid.

226 Therefore, the 4-*O*-methyl-glucuronic acid residue ester bonds are formed exclusively at the  $\alpha$ -  
227 position of phenyl propane units. An acidic pH favours the reaction of quinone methide with  
228 glucuronic acid<sup>46</sup>. Biosynthesis of secondary cell walls starts by forming cellulose and  
229 hemicelluloses. Xylan has two and three helical structures. Two-fold extended xylan accounts  
230 for ~70 mol% of the polyose bridges in cellulose microfibrils of plant cell walls<sup>16</sup>. Lignin is  
231 deposited with a considerable delay after forming the tightly packed hemicellulose-cellulose  
232 matrix and associates with the three-fold xylan helix through electrostatic interactions.  
233 Therefore, it is unlikely that all ester L-C bonds in secondary cell walls migrate from the  $\alpha$ - to  $\gamma$ -  
234 position in the tightly packed polysaccharide bridges, which are finally fulfilled with lignin.  
235 Migration of glucuronate from the  $\alpha$ - to  $\gamma$ -position was first reported by Toikka<sup>40, 41</sup> based on the  
236 study of synthetic L-C ester model compounds. However, this hypothesis was proposed by an *in*  
237 *vitro* solution reaction using a low molecular mass model. As discussed, in plant cell walls, the  
238 aromatic polymer, lignin, is densely deposited in the high molecular weight polysaccharide  
239 network where xylan chains fold onto the surface of cellulose fibrils through hydrogen bonding,  
240 which makes migration far more difficult than that observed in solution reactions of model  
241 compounds (Fig. 6).

242 The conclusive NMR assignments herein clearly showed that the  $\alpha$ -ester rather than the  $\gamma$ -ester  
243 is the major bond between glucuronoxylan and lignin, as expected from the biosynthetic pathway  
244 of secondary cell walls, critically revealing that previous studies reporting the  $\gamma$ -ester may  
245 include misassigned NMR signals, which were conducted only by comparison of chemical shifts  
246 with model compounds, or migration after or during the extraction process of LCCs.

247

248 **Conclusions**

249 Multi-dimensional NMR showed that the L-C  $\alpha$ -ester linkage in LCC is mediated by lignin and  
250 glucuronoxytan. The linkage was formed between the  $\alpha$ -position of the  $\beta$ -O-4' lignin unit and  
251 the 6-position of the 4-O-methyl glucuronic substituent of glucuronoxytan. The dominant  
252 proximal lignin unit was syringyl, and the primary acetyl substituent of xylan in LCC was at the  
253 3-position. Analysis of HMBC and 1,1-ADEQUATE spectra revealed clear evidence of L-C  
254 linkages at the  $\alpha$ -position of lignin  $\beta$ -O-4' units. However, no correlation was observed at the  $\gamma$ -  
255 position. Instead,  $\gamma$ -acetyl-substituted lignin was identified. We have elucidated that the major  
256 ester-type LCC is the  $\alpha$ -ester instead of the  $\gamma$ -ester, which amends conventional principles.  
257 Understanding the covalent L-C-ester linkage contributes to plant sciences and facilitates  
258 research into biomass conversion, refineries, biotechnology, and enzymology.

259

260 **Acknowledgements**

261 This study was supported in part by JSPS KAKENHI (16H06210, 21H02258), JST-Mirai  
262 Program (JPMJMI21EG) from Japan Science and Technology Agency, Japan Association for  
263 Chemical Innovation Foundation, The Kyoto University Foundation, mission research program  
264 (M5-2-4), and ADAM collaboration program of RISH, ZE joint usage research program of IAE,  
265 Kyoto University. We thank Edanz (<https://jp.edanz.com/ac>) for editing a draft of this  
266 manuscript.

267

268 **Author Contributions**

269 H.N., M.K., and T.W. designed the research. H.N., T.W., K.N., and T.N. performed the  
270 research. H.N. and T.W. wrote the manuscript. T.N. and M.K. refined the manuscript.

271

272 **References**

- 273 1. Baral NR, Sundstrom ER, Das L, Gladden J, Eudes A, Mortimer JC, *et al.* Approaches for More Efficient  
 274 Biological Conversion of Lignocellulosic Feedstocks to Biofuels and Bioproducts. *ACS Sustain Chem Eng*  
 275 2019, **7**(10): 9062-9079.  
 276
- 277 2. Koshijima T, Watanabe T. *Association between lignin and carbohydrates in wood and other plant tissues,*  
 278 *Springer Series in Wood Science.* Springer, 2003.  
 279
- 280 3. Nanbu T, Shimada J, Kobayashi M, Hirano K, Koh T, Machino M, *et al.* Anti-UV activity of lignin-  
 281 carbohydrate complex and related compounds. *In Vivo* 2013, **27**(1): 133-139.  
 282
- 283 4. Sakagami H, Sheng H, Okudaira N, Yasui T, Wakabayashi H, Jia J, *et al.* Prominent Anti-UV Activity and  
 284 Possible Cosmetic Potential of Lignin-carbohydrate Complex. *In Vivo* 2016, **30**(4): 331-339.  
 285
- 286 5. Kushida T, Makino T, Tomomura M, Tomomura A, Sakagami H. Enhancement of Dectin-2 gene  
 287 expression by lignin-carbohydrate complex from *Lentinus edodes* mycelia extract (LEM) in a mouse  
 288 macrophage-like cell line. *Anticancer Res* 2011, **31**(4): 1241-1248.  
 289
- 290 6. Sakagami H, Kushida T, Oizumi T, Nakashima H, Makino T. Distribution of lignin-carbohydrate complex  
 291 in plant kingdom and its functionality as alternative medicine. *Pharmacol Ther* 2010, **128**(1): 91-105.  
 292
- 293 7. Tsuji R, Ikado K, Fujiwara D. Modulation of Innate Immunity by lignin-Carbohydrate, a Novel TLR4  
 294 Ligand, Results in Augmentation of Mucosal IgA and Systemic IgG Production. *Int J Mol Sci* 2017, **19**(1).  
 295
- 296 8. Li R, Ouda R, Kimura C, Narita R, Nishimura H, Fujita T, *et al.* Conversion of Beech Wood into Antiviral  
 297 Lignin-Carbohydrate Complexes by Microwave Acidolysis. *Acs Sustain Chem Eng* 2021, **9**(28): 9248-  
 298 9256.  
 299
- 300 9. Lee JB, Yamagishi C, Hayashi K, Hayashi T. Antiviral and immunostimulating effects of lignin-  
 301 carbohydrate-protein complexes from *Pimpinella anisum*. *Biosci Biotechnol Biochem* 2011, **75**(3): 459-  
 302 465.  
 303
- 304 10. Sakagami H, Kawano M, Thet MM, Hashimoto K, Satoh K, Kanamoto T, *et al.* Anti-HIV and  
 305 immunomodulation activities of cacao mass lignin-carbohydrate complex. *In Vivo* 2011, **25**(2): 229-236.  
 306
- 307 11. Giummarella N, Pu YQ, Ragauskas AJ, Lawoko M. A critical review on the analysis of lignin carbohydrate  
 308 bonds. *Green Chem* 2019, **21**(7): 1573-1595.  
 309
- 310 12. Terashima N, Yoshida M, Hafren J, Fukushima K, Westermark U. Proposed supramolecular structure of  
 311 lignin in softwood tracheid compound middle lamella regions. *Holzforschung* 2012, **66**(8): 907-915.  
 312
- 313 13. Cosgrove DJ, Jarvis MC. Comparative structure and biomechanics of plant primary and secondary cell  
 314 walls. *Front Plant Sci* 2012, **3**: 204.  
 315
- 316 14. Busse-Wicher M, Grantham NJ, Lyczakowski JJ, Nikolovski N, Dupree P. Xylan decoration patterns and  
 317 the plant secondary cell wall molecular architecture. *Biochem Soc Trans* 2016, **44**(1): 74-78.  
 318
- 319 15. Simmons TJ, Mortimer JC, Bernardinelli OD, Poppler AC, Brown SP, deAzevedo ER, *et al.* Folding of  
 320 xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nat Commun* 2016, **7**: 13902.  
 321
- 322 16. Kang X, Kirui A, Dickwella Widanage MC, Mentink-Vigier F, Cosgrove DJ, Wang T. Lignin-  
 323 polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. *Nat Commun* 2019,  
 324 **10**(1): 347.  
 325

- 326 17. Terrett OM, Lyczakowski JJ, Yu L, Iuga D, Franks WT, Brown SP, *et al.* Molecular architecture of  
327 softwood revealed by solid-state NMR. *Nat Commun* 2019, **10**(1): 4978.  
328
- 329 18. Wang P.Y. BHI. Uronic acid ester groups in some softwoods and hardwoods *Tappi* 1967(50): 123-124.  
330
- 331 19. Das NN, Das S.C., Mukherjee, A. K. On the ester linkage between lignin and 4-*O*-methyl-D-glucurono-D-  
332 xylan in jute fiber (*Corchorus capsularis*). *Carbohydr Res* 1984, **127**: 345-348.  
333
- 334 20. Takahashi N, Koshijima, T. Ester linkages between lignin and glucuronoxylan in a lignin-carbohydrate  
335 complex from beech (*Fagus crenata*) wood. *Wood Sci Technol* 1988, **22**: 177-189.  
336
- 337 21. Imamura T, Watanabe T, Kuwahara M, Koshijima T. Ester Linkages between Lignin and Glucuronic-Acid  
338 in Lignin-Carbohydrate Complexes from *Fagus Crenata*. *Phytochemistry* 1994, **37**(4): 1165-1173.  
339
- 340 22. Watanabe T, Koshijima T. Evidence for an Ester Linkage between Lignin and Glucuronic-Acid in Lignin  
341 Carbohydrate Complexes by DDQ-Oxidation. *Agric Biol Chem* 1988, **52**(11): 2953-2955.  
342
- 343 23. d'Errico C, Jorgensen JO, Krogh KBRM, Spodsberg N, Madsen R, Monrad RN. Enzymatic Degradation of  
344 Lignin-Carbohydrate Complexes (LCCs): Model Studies Using a Fungal Glucuronoyl Esterase from  
345 *Cerrena Unicolor*. *Biotechnol Bioeng* 2015, **112**(5): 914-922.  
346
- 347 24. Baath JA, Giummarella N, Klaubauf S, Lawoko M, Olsson L. A glucuronoyl esterase from *Acremonium*  
348 *alcalophilum* cleaves native lignin-carbohydrate ester bonds. *FEBS Lett* 2016, **590**(16): 2611-2618.  
349
- 350 25. Huttner S, Klaubauf S, de Vries RP, Olsson L. Characterisation of three fungal glucuronoyl esterases on  
351 glucuronic acid ester model compounds. *Appl Microbiol Biotechnol* 2017, **101**(13): 5301-5311.  
352
- 353 26. Huynh H, Ishii N, Matsuo I, Arioka M. A novel glucuronoyl esterase from *Aspergillus fumigatus*-the role  
354 of conserved Lys residue in the preference for 4-*O*-methyl glucuronoyl esters. *Appl Microbiol Biotechnol*  
355 2018, **102**(5): 2191-2201.  
356
- 357 27. Mosbech C, Holck J, Meyer AS, Agger JW. The natural catalytic function of CuGE glucuronoyl esterase in  
358 hydrolysis of genuine lignin-carbohydrate complexes from birch. *Biotechnol Biofuels* 2018, **11**: 71.  
359
- 360 28. Mosbech C, Holck J, Meyer A, Agger JW. Enzyme kinetics of fungal glucuronoyl esterases on natural  
361 lignin-carbohydrate complexes. *Appl Microbiol Biotechnol* 2019, **103**(10): 4065-4075.  
362
- 363 29. Lin MI, Hiyama A, Kondo K, Nagata T, Katahira M. Classification of fungal glucuronoyl esterases (FGEs)  
364 and characterization of two new FGEs from *Ceriporiopsis subvermispota* and *Pleurotus eryngii*. *Appl*  
365 *Microbiol Biotechnol* 2018, **102**(22): 9635-9645.  
366
- 367 30. Ernst HA, Mosbech C, Langkilde AE, Westh P, Meyer AS, Agger JW, *et al.* The structural basis of fungal  
368 glucuronoyl esterase activity on natural substrates. *Nat Commun* 2020, **11**(1): 1026.  
369
- 370 31. Balakshin M, Capanema E, Gracz H, Chang HM, Jameel H. Quantification of lignin-carbohydrate linkages  
371 with high-resolution NMR spectroscopy. *Planta* 2011, **233**(6): 1097-1110.  
372
- 373 32. Balakshin MY, Capanema EA, Chang HM. MWL fraction with a high concentration of lignin-carbohydrate  
374 linkages: Isolation and 2D NMR spectroscopic analysis. *Holzforschung* 2007, **61**(1): 1-7.  
375
- 376 33. Lawoko M, Henriksson G, Gellerstedt G. Structural differences between the lignin-carbohydrate complexes  
377 present in wood and in chemical pulps. *Biomacromolecules* 2005, **6**(6): 3467-3473.  
378
- 379 34. Miyagawa Y, Kamitakahara H, Takano T. Fractionation and characterization of lignin-carbohydrate  
380 complexes (LCCs) of *Eucalyptus globulus* in residues left after MWL isolation. Part II: Analyses of xylan-  
381 lignin fraction (X-L). *Holzforschung* 2013, **67**(6): 629-642.

- 382  
383 35. Miyagawa Y, Mizukami T, Kamitakahara H, Takano T. Synthesis and fundamental HSQC NMR data of  
384 monolignol beta-glycosides, dihydromonolignol beta-glycosides and p-hydroxybenzaldehyde derivative  
385 beta-glycosides for the analysis of phenyl glycoside type lignin-carbohydrate complexes (LCCs).  
386 *Holzforschung* 2014, **68**(7): 747-760.  
387
- 388 36. Tarasov D, Leitch M, Fatehi P. Lignin-carbohydrate complexes: properties, applications, analyses, and  
389 methods of extraction: a review. *Biotechnol Biofuels* 2018, **11**: 269.  
390
- 391 37. Yuan TQ, Sun SN, Xu F, Sun RC. Characterization of Lignin Structures and Lignin-Carbohydrate  
392 Complex (LCC) Linkages by Quantitative C-13 and 2D HSQC NMR Spectroscopy. *J Agr Food Chem*  
393 2011, **59**(19): 10604-10614.  
394
- 395 38. Nishimura H, Kamiya A, Nagata T, Katahira M, Watanabe T. Direct evidence for  $\alpha$  ether linkage between  
396 lignin and carbohydrates in wood cell walls. *Scientific Reports* 2018, **8**(1): 6538.  
397
- 398 39. Hilton BD, Martin GE. Investigation of the experimental limits of small-sample heteronuclear 2D NMR. *J*  
399 *Nat Prod* 2010, **73**(9): 1465-1469.  
400
- 401 40. Toikka M, Sipilä J, Teleman A, Brunow G. Lignin-carbohydrate model compounds. Formation of lignin  
402 methyl arabinoside and lignin methyl galactoside benzyl ethers via quinone methide intermediates. *J Chem*  
403 *Soc Perk T 1* 1998(22): 3813-3818.  
404
- 405 41. Toikka M, Brunow G. Lignin-carbohydrate model compounds. Reactivity of methyl 3-O-( $\alpha$ -L-  
406 arabinofuranosyl)- $\beta$ -D-xylopyranoside and methyl  $\beta$ -D-xylopyranoside towards a  $\beta$ -O-4-quinone  
407 methide. *J Chem Soc Perk T 1* 1999(13): 1877-1883.  
408
- 409 42. Nylander F, Sunner H, Olsson L, Christakopoulos P, Westman G. Synthesis and enzymatic hydrolysis of a  
410 diaryl benzyl ester model of a lignin-carbohydrate complex (LCC). *Holzforschung* 2016, **70**(5): 385-391.  
411
- 412 43. Sapouna I, Lawoko M. Deciphering lignin heterogeneity in ball milled softwood: unravelling the synergy  
413 between the supramolecular cell wall structure and molecular events. *Green Chem* 2021, **23**(9): 3348-  
414 3364.  
415
- 416 44. Grantham NJ, Wurman-Rodrich J, Terrett OM, Lyczakowski JJ, Stott K, Iuga D, *et al.* An even pattern of  
417 xylan substitution is critical for interaction with cellulose in plant cell walls. *Nat Plants* 2017, **3**(11): 859-  
418 865.  
419
- 420 45. Gupta M, Rawal TB, Dupree P, Smith JC, Petridis L. Spontaneous rearrangement of acetylated xylan on  
421 hydrophilic cellulose surfaces. *Cellulose* 2021, **28**(6): 3327-3345.  
422
- 423 46. Sipilä J, Brunow G. On the Mechanism of Formation of Non-Cyclic Benzyl Ethers During Lignin  
424 Biosynthesis Part 3. The Reactivity of a  $\beta$ -O-4-Type Quinone Methide with Methyl- $\alpha$ -D-glucopyranoside  
425 in Competition with Vanillyl Alcohol. The Formation and the Stability of Benzyl Ethers between Lignin  
426 and Carbohydrates. 1991, **45**(s1): 3-8.

427

428

429 **Methods**

430 The LCC samples were carefully prepared under mild conditions. Samples were kept below  
431 60°C throughout the fractionation and extraction procedures. Wiley-milled wood powder was  
432 prepared from the sapwood timber of Japanese beech wood (*Fagus crenata*, *buna* in Japanese).  
433 The isolation scheme is shown in Extended Data Fig. 1. To eliminate extractives, obtained beech  
434 wood powder was extracted twice with an ethyl alcohol-toluene mixture (1:2, v/v) with a ratio of  
435 5.3 L per 1 kg wood at 60°C for 24 h. Next, pectin was removed by extraction with 0.25 w/v%  
436 aqueous potassium acetate with a ratio of 10 L per 1 kg wood at 60°C for 24 h. The wood  
437 powder was then washed with pure water and acetone, followed by drying. The pre-extracted  
438 wood powder was subjected to planetary ball-milling by PULVERISETTE 6 (Fritsch Japan Co.  
439 Ltd., Kanagawa, Japan) with an overpot system, as shown in Extended Data Fig. 2. The pot was  
440 evacuated, sealed, and replaced with nitrogen gas using an overpot system after evacuation.  
441 Heat damage was prevented by inserting a 1.2-min cooling period between 1 min milling  
442 intervals to ensure the internal temperature remained below 60°C, which was checked by an  
443 irreversible thermography patch (NiGK Co., Saitama, Japan, shown in Extended Data Fig. 2c).  
444 For the milling conditions, 1 g wood powder and 100 g of zirconia beads ( $\Phi$  3 mm) were placed  
445 in an agate pot with a capacity of 80 mL. The sample was then milled for 1 min at 550 rpm and  
446 this 1 min milling procedure was repeated 180 times, i.e., 3 h (net).

447 **Preparation of LCC samples.** The ball-milling wood (20 g) was extracted with DMSO (40  
448 mL) in twice with stirring followed by lyophilisation. The extracts (2.60 g) termed "LCC-  
449 DMSO" was treated with a polysaccharide-degrading enzyme, hemicellulase cellulysin GM5  
450 (HBI Enzymes Inc., Hyogo, Japan), derived from *Aspergillus niger*, which was partially purified  
451 by ammonium sulfate precipitation and ultrafiltration (Amicon, MWCO 10 kDa, Merck Japan

452 Co. Ltd.) before use. The enzyme had 719 kU/g xylanase activity when beech wood-derived  
453 xylan (Nacalai Tesque, Kyoto, Japan) was used as the substrate.

454 Enzymatic treatment was conducted under the following conditions: LCC-DMSO (2.60 g), acetic  
455 acid buffer (pH 4.0, 50 mM, 65 mL), and 1 g w/v% enzyme at 50°C for 80 h with stirring. After  
456 the reaction, the LCC-DMSO was partly hydrolysed and washed with water three times followed  
457 by centrifugation to separate supernatant and precipitates, termed soluble liquid fraction "LCC-  
458 EL" and the residue "LCC-ER". The LCC-ER (2.3 g) was obtained after lyophilisation. A  
459 fraction of the obtained sample was dissolved in DMSO-*d*<sub>6</sub> for NMR analysis. The LCC-EL was  
460 purified by chromatography using a hydrophobic interaction/size exclusion gel TOYOPEARL  
461 HW50 (Tosoh Bioscience LLC, Tokyo, Japan). LCC-ER (2.2 g) was extracted with 60 mL  
462 methanol/water (v/v, 8/2) twice at room temperature for 1 h with 100 kHz ultrasound irradiation  
463 (VS-100III, AS ONE Co. Ltd, Osaka, Japan). The precipitate and supernatant were centrifuged,  
464 and the supernatant was collected, followed by evaporation and lyophilisation to obtain a 0.1-g  
465 purified LCC sample termed "LCC-ER-AL80".

466 **NMR experiments.** Samples were dissolved in 0.5 mL of DMSO-*d*<sub>6</sub> and transferred to an NMR  
467 tube (DMS-005TB, Shigemi Co. Ltd., Tokyo, Japan). NMR spectra were recorded at 313 K  
468 using a Bruker AVANCE III 600 MHz spectrometer with a 5-mm TCI cryoprobe. A 1D <sup>1</sup>H  
469 NMR spectrum of the sample was acquired using an acquisition time (AQ) of 0.9 s, an interscan  
470 delay (D1) of 2.0 s, and 32 scans. The two-dimensional (2D) <sup>1</sup>H-<sup>13</sup>C HSQC (HSQC) spectrum  
471 was recorded using an adiabatic HSQC experiment (Bruker pulse program, hsqccetgpsisp2.2)  
472 with the following parameters: AQ, 0.10 s; D1, 2.0 s; a <sup>1</sup>H spectral width of 16 ppm, and 165  
473 ppm for <sup>13</sup>C with 2048 × 512 complex points; 32 scans per increment. The <sup>1</sup>J<sub>CH</sub> used was 145

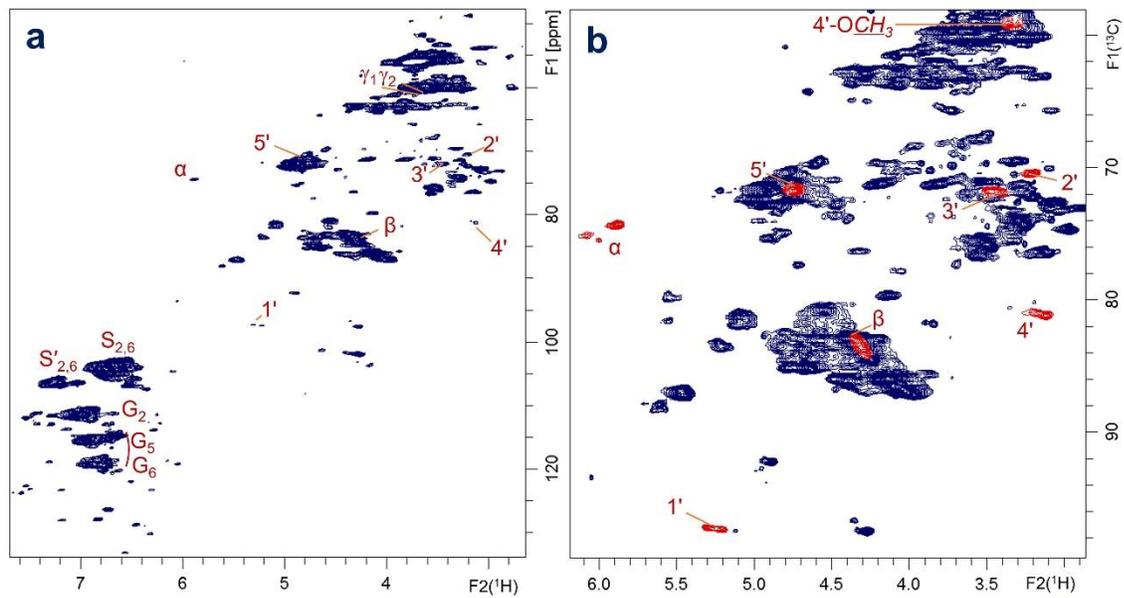
474 Hz. Before Fourier transformation, the data matrices were zero-filled to 1024 points in the  $^{13}\text{C}$   
475 dimension and processed with a Gaussian window function in the  $^1\text{H}$  dimension and a sine-bell  
476 square window function in the  $^{13}\text{C}$  dimension. The multiplicity-edited HSQC was recorded  
477 using the Bruker pulse program, hsqcedetgp. The 2D  $^1\text{H}$ - $^{13}\text{C}$  HMBC (HMBC) spectrum was  
478 recorded using a magnitude-mode ge-2D HMBC with a low-pass  $J$ -filter (Bruker pulse program,  
479 hmbcgpndqf). The parameters for this 2D HMBC were: AQ, 0.12 s; D1, 1.2 s; a  $^1\text{H}$  spectral  
480 width of 14 ppm, and 150 ppm for  $^{13}\text{C}$  with  $2048 \times 512$  complex points; 48 scans per increment.  
481 The 2D HMBC spectra were processed with a sine-bell window function in  $^1\text{H}$  and  $^{13}\text{C}$   
482 dimensions. The three-dimensional (3D)  $^1\text{H}$ - $^{13}\text{C}$  TOCSY-HSQC spectrum was recorded using  
483 the Bruker pulse program, mlevhsqcetgp3d with the following parameters: AQ, 0.07 s; D1, 1.0 s;  
484 a spectral window of 12 ppm for the  $^1\text{H}$  direct dimension, 58 ppm for  $^{13}\text{C}$ , and 5.5 ppm for the  
485 indirect  $^1\text{H}$  dimension with  $1024 \times 384 \times 64$  complex points; 8 scans per increment; 100 ms  
486 TOCSY spinlock. The 3D spectrum was processed with a sine-bell square window function in  
487 all dimensions. 1,1-ADEQUATE spectrum was recorded using the Bruker pulse program,  
488 adeq1letgprdsp with the following parameters: AQ, 0.11 s; D1, 1.5 s; a  $^1\text{H}$  spectral window of  
489 16 ppm, and 80 ppm for  $^{13}\text{C}$  with  $2048 \times 196$  complex points; 192 scans per increment. The data  
490 matrices were zero-filled to 1024 points in the  $^{13}\text{C}$  dimension and processed with a Gaussian  
491 window function in the  $^1\text{H}$  dimension and a sine-bell square window function in the  $^{13}\text{C}$   
492 dimension. Data processing was performed using standard Bruker Topspin-NMR software (ver.  
493 3.1). The solvent (DMSO) peak was used as an internal chemical shift reference point ( $\delta_{\text{H}}/\delta_{\text{C}} =$   
494 2.49 ppm/39.6 ppm).

495

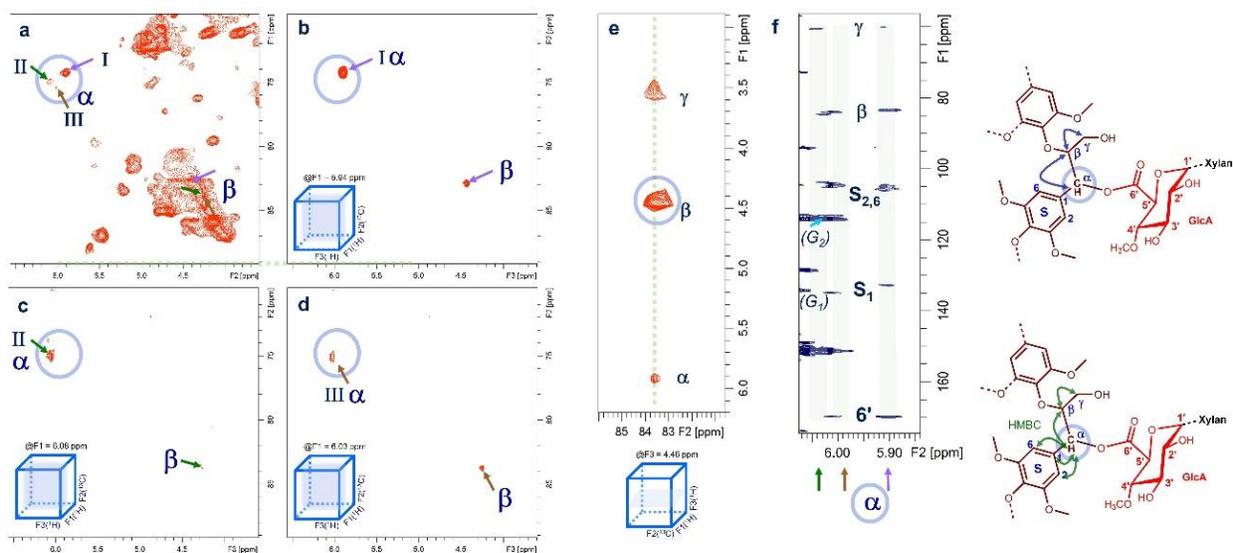


496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506

**Fig. 1 | Conceptual diagram of a lignocellulosic fibre in hardwood cell walls and the elucidated structural model of the  $\alpha$ -ester lignin-xylan linkage of lignin-carbohydrate complex (LCC).** **a**, This study determined the principal ester linkages between lignin and xylan (hemicellulose) at the  $\alpha$ -position of the lignin  $\beta$ -ether units, as shown in red. Glucurono-xylan covalently bonds with lignin, with an amorphous three-fold screw conformation. In contrast, glucuronoxylan is associated with crystalline cellulose by non-covalent bonds with a two-fold screw conformation. **b**, The ester bond between the 6-position of 4'-O-methyl glucuronate of glucuronoxylan and the  $\alpha$ -(benzyl) position of lignin  $\beta$ -O-4' units are the primary L-C cross-linked structure of lignocellulose.

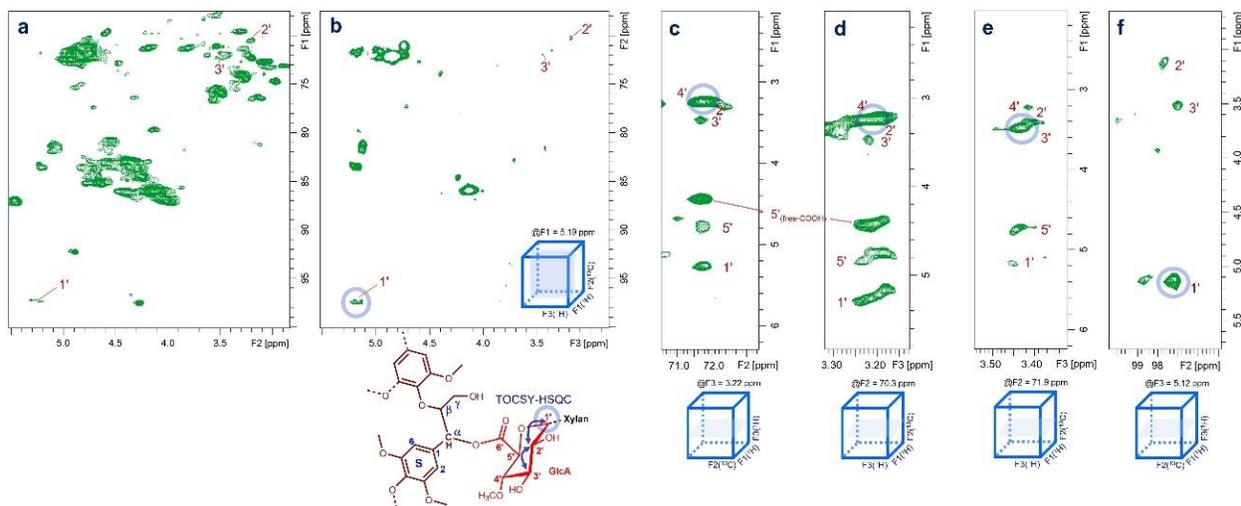


507  
 508 **Fig. 2 | The 2D  $^1\text{H}$ - $^{13}\text{C}$  HSQC spectrum of the  $\alpha$ -ester-type LCC. a, Full, and a magnified (b) region of the**  
 509 **spectrum. Signal assignments originating from  $\alpha$ -ester L-C linkages are labelled in (a) and (b). Particular signals**  
 510 **are labelled and highlighted in red in (b).**  
 511



512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

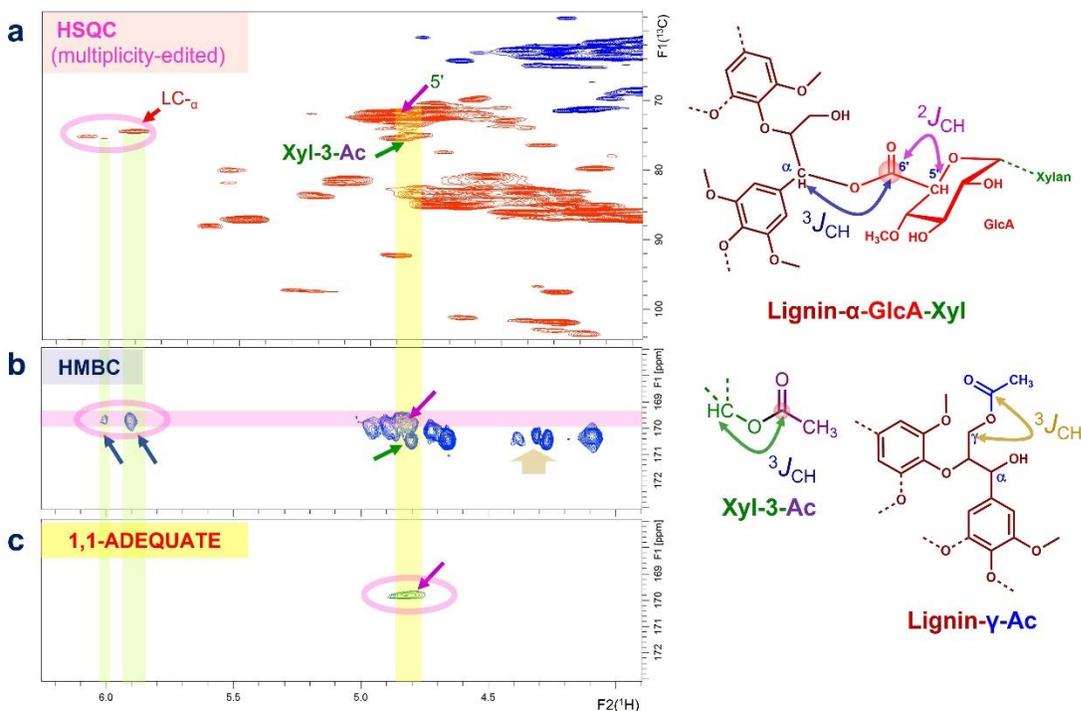
**Fig. 3 | Reliable signal attribution of the  $\alpha$ -ester-type LCC by multi-dimensional correlation NMR.** The  $\alpha$  and  $\beta$  correlations in the 2D  $^1\text{H}$ - $^{13}\text{C}$  HSQC (a) and 3D TOCSY-HSQC spectra (b–d) for the lignin  $\beta$ -aryl ether unit in the  $\alpha$ -ester-type LCC. Three signals were observed for the  $\alpha$ -position of the lignin  $\beta$ -aryl ether unit of L-C ester linkages (circled in a). The extracted planes at each  $\alpha$ -position in the 3D TOCSY-HSQC spectrum identified each  $\beta$ -position (b–d). Correlations labelled as I, II, and III are distinguished by arrows (I: purple (b); II: green (c); III: brown (d)). In each case, a signal from a TOCSY correlation from the  $\alpha$ -position to the  $\beta$ -position was observed, which was unresolved in the 2D  $^1\text{H}$ - $^{13}\text{C}$  HSQC spectrum. e, The extracted plane from the 3D TOCSY-HSQC at the  $\beta$  position of the lignin  $\beta$ -aryl ether unit of L-C ester linkages. TOCSY correlations to the  $\alpha$ - and  $\gamma$ -positions were observed. f, In the 2D  $^1\text{H}$ - $^{13}\text{C}$  HMBC spectrum, correlation signals between the  $\alpha$ -proton of the L-C ester to the  $\beta$ - and  $\gamma$ -carbons of the lignin  $\beta$ -aryl ether unit and syringyl aromatic ring ( $S_1$  and  $S_{2,6}$ ) were observed. The aromatic ring proximity to the  $\alpha$ -position is mainly the syringyl lignin, and signals from the guaiacyl units were very weak.



526  
527

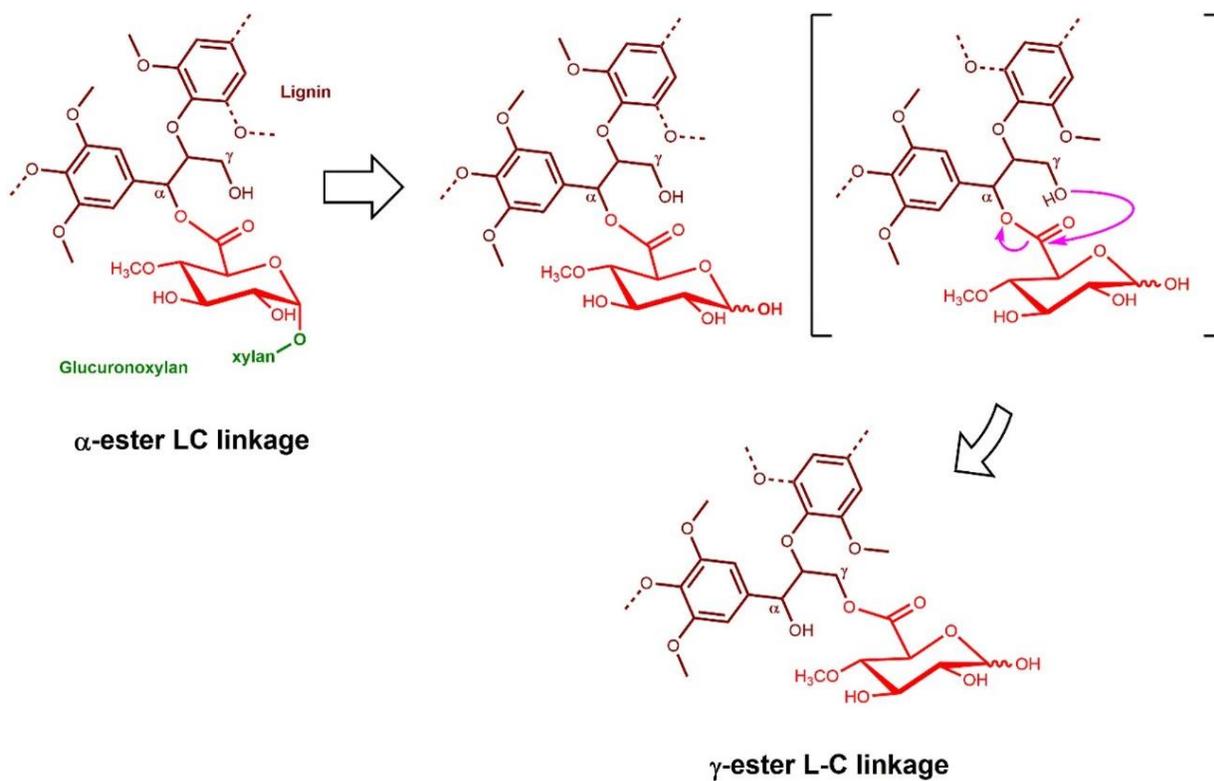
528 **Fig. 4 | Correlation signals of 4'-O-methyl glucuronate moiety from L-C ester linkages in the LCC sample.** 2D  
 529  $^1\text{H}$ - $^{13}\text{C}$  HSQC spectrum (a) and selected 2D planes from the 3D TOCSY-HSQC spectrum (b-f). b, Extracted plane  
 530 from the 3D TOCSY-HSQC spectrum showing the anomeric hydrogen at the 1' position of the glucuronic acid (5.19  
 531 ppm). TOCSY correlations with GlcA2 and GlcA3 were observed, and the 2D  $^1\text{H}$ - $^{13}\text{C}$  HSQC spectrum (a) displays  
 532 the same spectral region for comparison. d, Extracted region of the 3D TOCSY-HSQC showing the signal from the  
 533 hydrogen at GlcA2 (3.20 ppm), and extracted region in (e) showing the correlation to the carbon in GlcA2 (70.3  
 534 ppm). TOCSY correlation signals from GlcA1 to GlcA5 were observed in each case. The chemical shifts of the  
 535 signals arising from the hydrogens in GlcA2 and GlcA4 partially overlap. e, Extracted region from the 3D TOCSY-  
 536 HSQC showing the signal arising from the carbon (71.9 ppm) of GlcA3. TOCSY correlation signals from GlcA1 to  
 537 GlcA5 were observed. f, Extracted region from the 3D TOCSY-HSQC showing correlation signals for the hydrogen  
 538 in GlcA1. TOCSY correlation signals from GlcA1 to GlcA2 and GlcA3 were observed.

539  
540



541  
542

543 **Fig. 5 | Clear evidence of an  $\alpha$ -ester-type LCC using 2D 1,1-ADEQUATE and HMBC spectra.** The  
544 multiplicity-edited HSQC spectrum (a), HMBC spectrum (b), and 1,1-ADEQUATE spectrum (c) are shown. b, In  
545 the HMBC spectrum, there is a correlation between the carbonyl carbon at 169.6 ppm and the  $\alpha$ -position of LCC  
546 shown in the HSQC spectrum (a). This carbon ( $\delta_{\text{C}}$  169.6 ppm) correlates with the signal at  $\delta_{\text{H}}$  4.8 ppm in (b). a,  
547 The acetyl-substituted xylan at the 3-position (Xyl-3-Ac) and the glucuronate moiety at the 5-position in LC  
548 glucuronoxytan have almost the same  $\delta_{\text{H}}$  of 4.8 ppm. c, In the 1,1-ADEQUATE spectrum, only the correlation  
549 signal derived from  $^2J_{\text{CH}}$  is detected. Therefore, the correlation signal with the carbonyl group of the ester type LC  
550 bond was clearly distinguished from acetyl-substituted xylan at the 3-position. Signals at  $\delta_{\text{H}}$  4.3–4.4 ppm in (b)  
551 represent correlations between the proton at the  $\gamma$  position and the carbonyl carbon. No 1,1-ADEQUATE  
552 correlation derived from the  $\gamma$ -ester-type LCC was observed and was identified as a  $\gamma$ -acetyl substituted lignin unit.



553  
 554  
 555  
 556  
 557  
 558  
 559

**Fig. 6 | Proposed rearrangement reaction of  $\alpha$ -ester LCC to  $\gamma$ -ester LCC.** Original L-C linkage at the  $\alpha$ -position of the lignin aryl ester unit on LCC migrates to the  $\gamma$ -position. The migration may arise from artificial degradation; partial bond-cleavage of glucuronoxylan in LCC during sample preparation, e.g., ball-milling and heat, chemical, and enzymatic treatments.

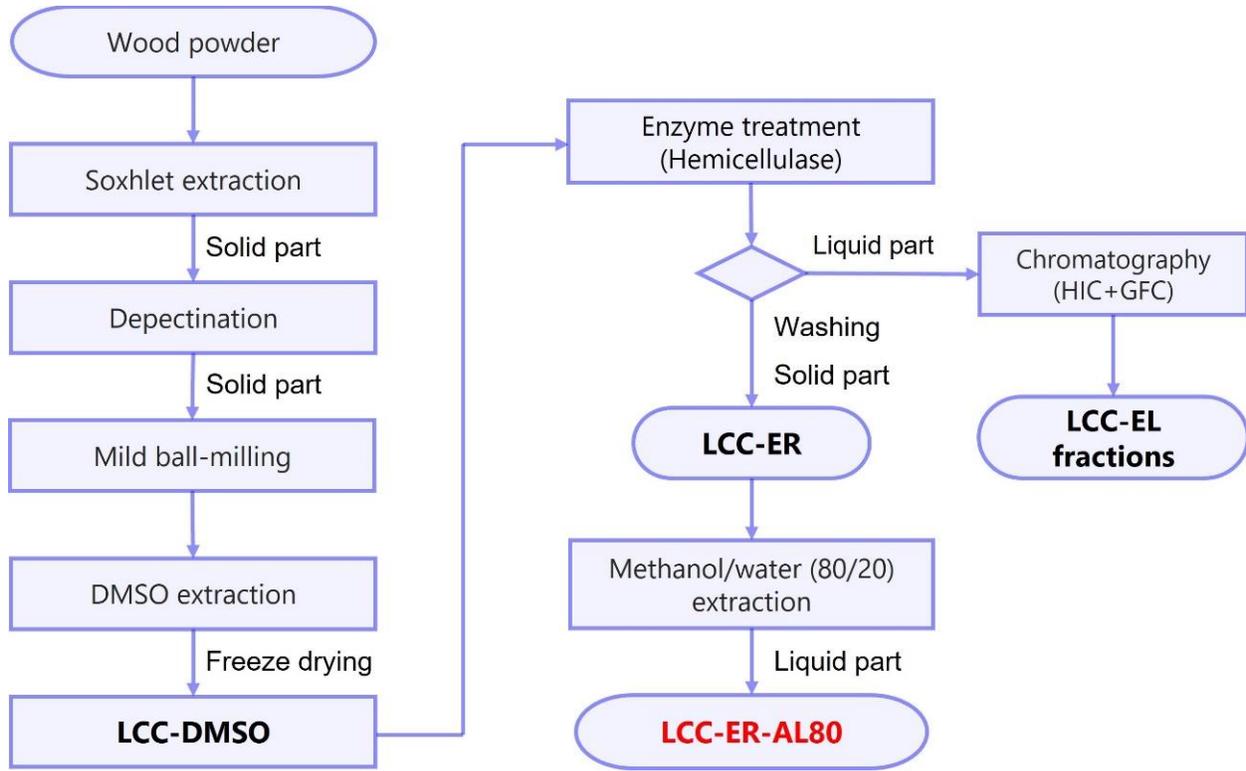
**Extended Table 1 | NMR signal assignments of the  $\alpha$ -ester-type LCC.**

Labels	$\delta_H$ (ppm)	$\delta_C$ (ppm)	Assignments
I- $\alpha$	5.88	74.4	$\alpha$ -CH of $\beta$ -O-4' substructures in $\alpha$ -ester-type LCC linked to esterified syringyl units
II- $\alpha$	6.08	75.2	$\alpha$ -CH of $\beta$ -O-4' substructures in $\alpha$ -ester-type LCC linked to free phenolic syringyl units
III- $\alpha$	6.00	75.5	$\alpha$ -CH of $\beta$ -O-4' substructures in $\alpha$ -ester-type LCC linked to guaiacyl I units
I- $\beta$	4.48	82.9	$\beta$ -CH of $\beta$ -O-4' substructures in $\alpha$ -ester-type LCC linked to esterified syringyl units
II- $\beta$	4.28	83.7	$\beta$ -CH of $\beta$ -O-4' substructures in $\alpha$ -ester-type LCC linked to free phenolic syringyl units
III- $\beta$	4.28	83.9	$\beta$ -CH of $\beta$ -O-4' substructures in $\alpha$ -ester-type LCC linked to guaiacyl I units
$\gamma_1, \gamma_2$	3.55, 3.83	61.0/62.0	$\gamma$ -CH <sub>2</sub> of $\beta$ -O-4' substructures in $\alpha$ -ester-type
1'-GluA	5.19 (5.13–5.31)	97.2–97.5	1'-CH (anomeric position) in 4'-O-methyl glucuronic acid in $\alpha$ -ester-type LCC
2'-GluA	3.20	70.3	2'-CH in 4'-O-methyl glucuronic acid in $\alpha$ -ester-type LCC
3'-GluA	3.43	71.9	3'-CH in 4'-O-methyl glucuronic acid in $\alpha$ -ester-type LCC
4'-GluA	3.12–3.24	81.13	4'-CH in 4'-O-methyl glucuronic acid in $\alpha$ -ester-type LCC
5'-GluA	4.80	71.7	5'-CH in 4'-O-methyl glucuronic acid in $\alpha$ -ester-type LCC
6'-GluA	–	169.7	6'-C=O in 4'-O-methyl glucuronic acid in $\alpha$ -ester-type LCC
CH <sub>3</sub> -4'-GluA	3.32	59.2	Methoxy CH <sub>3</sub> in 4'-O-methyl glucuronic acid moiety of glucuronoxylan
S <sub>2,6</sub>	6.61–670	103.4–105.1	C <sub>2</sub> -H <sub>2</sub> and C <sub>6</sub> -H <sub>6</sub> in syringyl units
S' <sub>2,6</sub>	7.19–7.32	105.9–106.7	C <sub>2</sub> -H <sub>2</sub> and C <sub>6</sub> -H <sub>6</sub> in $\alpha$ -oxidized syringyl units
S <sub>1</sub>	–	132.7/134.7	C <sub>1</sub> in syringyl units (etherified units/free phenolic units)
G <sub>2</sub>	6.99	111.3	C <sub>2</sub> -H <sub>2</sub> in guaiacyl units (centre of width in signals)
G <sub>5</sub>	6.95	115.4	C <sub>5</sub> -H <sub>5</sub> (containing minor C <sub>6</sub> -H <sub>6</sub> in part) in guaiacyl units (centre of width in signals)
G <sub>6</sub>	6.75	119.3	C <sub>6</sub> -H <sub>6</sub> in guaiacyl units (centre of width in signals)

**Abbreviations:**

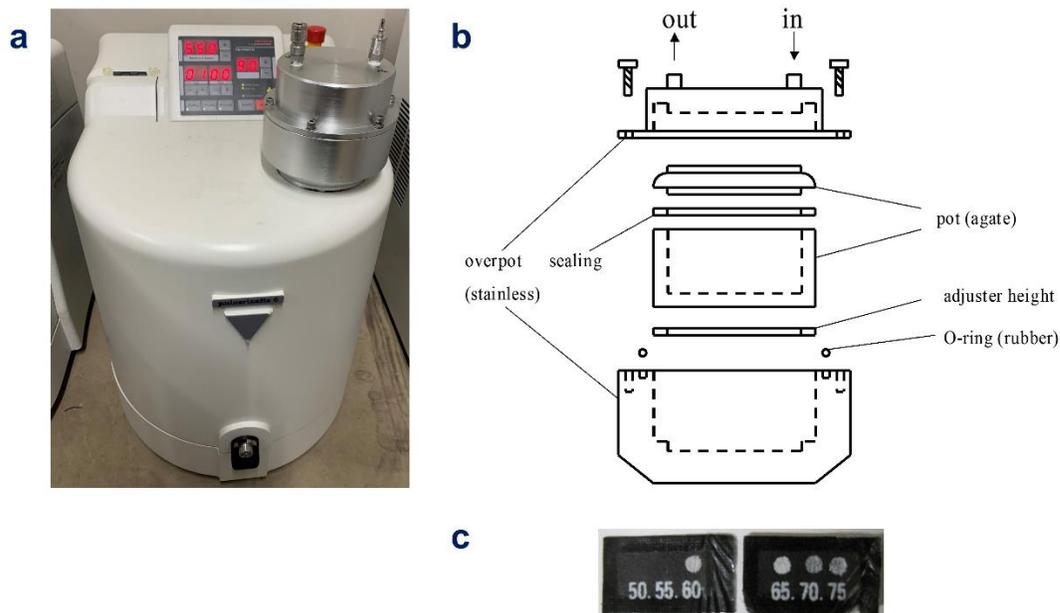
LCC,  $\alpha$ -**ester** (benzyl **ester**)-type lignin-carbohydrate complex; G, guaiacyl lignin units; Chemical shift values are taken at the centre of the signals

563



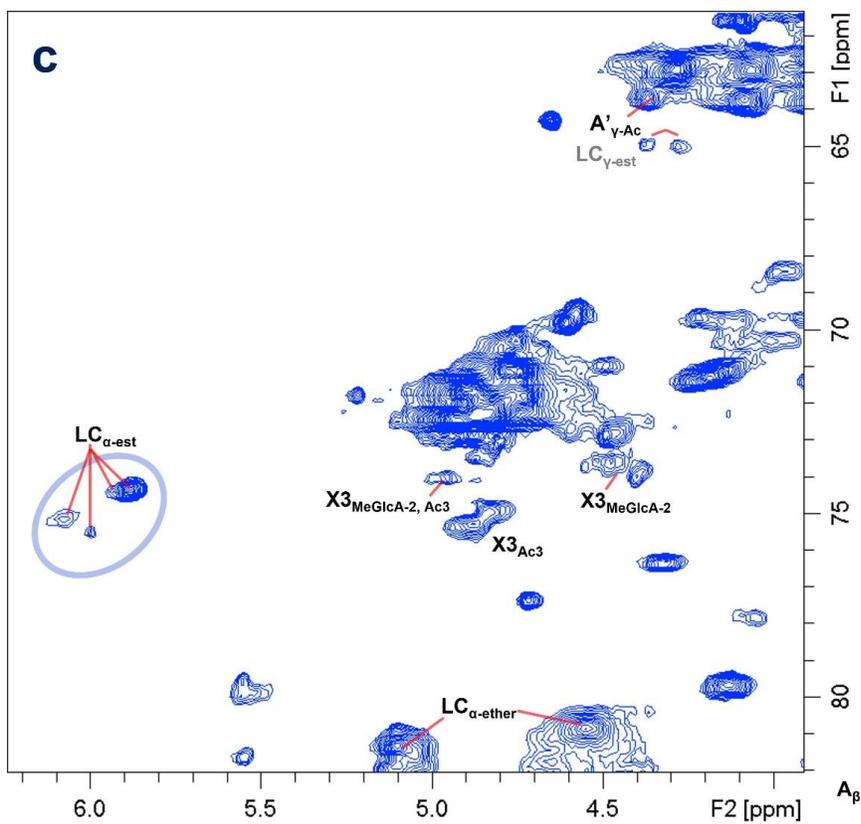
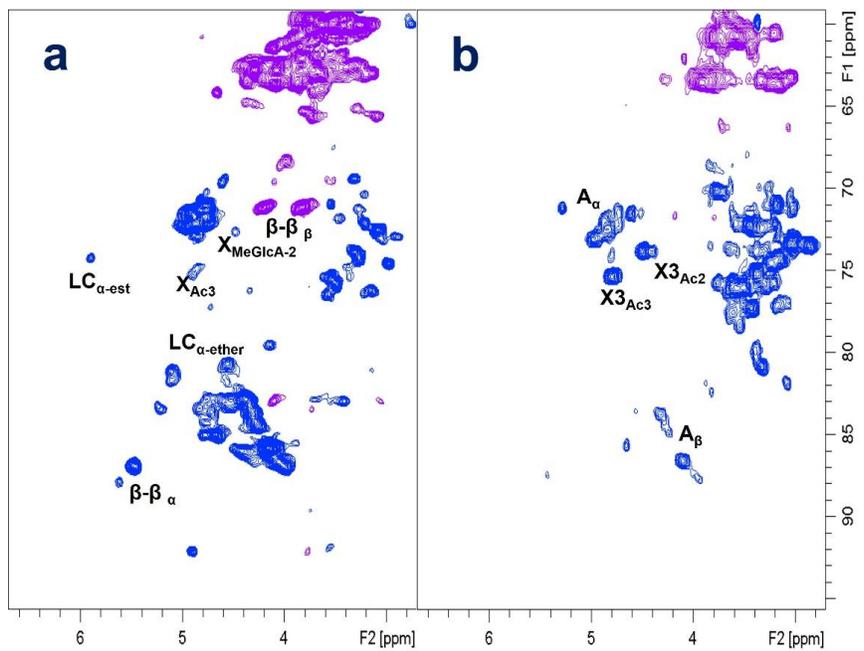
564  
565  
566  
567

Extended Data Fig. 1 | A flow diagram of the LCC sample fractionation process.



569  
570  
571  
572  
573  
574  
575  
576  
577

**Extended Data Fig. 2 | Planetary ball-milling treatment performed under mild conditions for LCC sample preparation.** **a**, Photograph of the planetary ball-milling equipment, PULVERISETTE 6 and the overpot system (shown on the device). **b**, A schematic of the overpot system. The agate pot vessel is set in the overpot. **c**, Photograph of the thermo label after ball-milling, showing that the treatment was conducted at 60°C. The temperature was monitored during ball-milling, and irreversible thermal labels were pasted on the bottom of the agate pot.

579  
580

581 **Extended Data Fig. 3 | NMR data of LCC-ER-AL80 and LCC-DMSO.** Multiplicity-edited 2D  $^1\text{H}$ - $^{13}\text{C}$  NMR  
 582 spectra of (a) LCC-ER-AL80 extracted by methanol/water, followed by enzymatic treatment, and (b) crude extracts  
 583 of LCC-DMSO. c, Expanded region of the 2D  $^1\text{H}$ - $^{13}\text{C}$  HSQC spectrum of LCC-ER-AL80.

584