Supplementary Information

Jinze Dou[1*]; Jincheng Wang[2]; Sami Hietala[3]; Dmitry V. Evtuguin[4]; Tapani Vuorinen[1*]; Jian Zhao[2]

[1]Department of Bioproducts and Biosystems, Aalto University, Espoo, Finland

[2]State Key Laboratory of Microbial Technology, Shandong University, Qingdao, China

[3]Department of Chemistry, University of Helsinki, Helsinki, Finland

[4]CICECO/Department of Chemistry, University of Aveiro, Aveiro, Portugal

Corresponding author (and address of all authors):

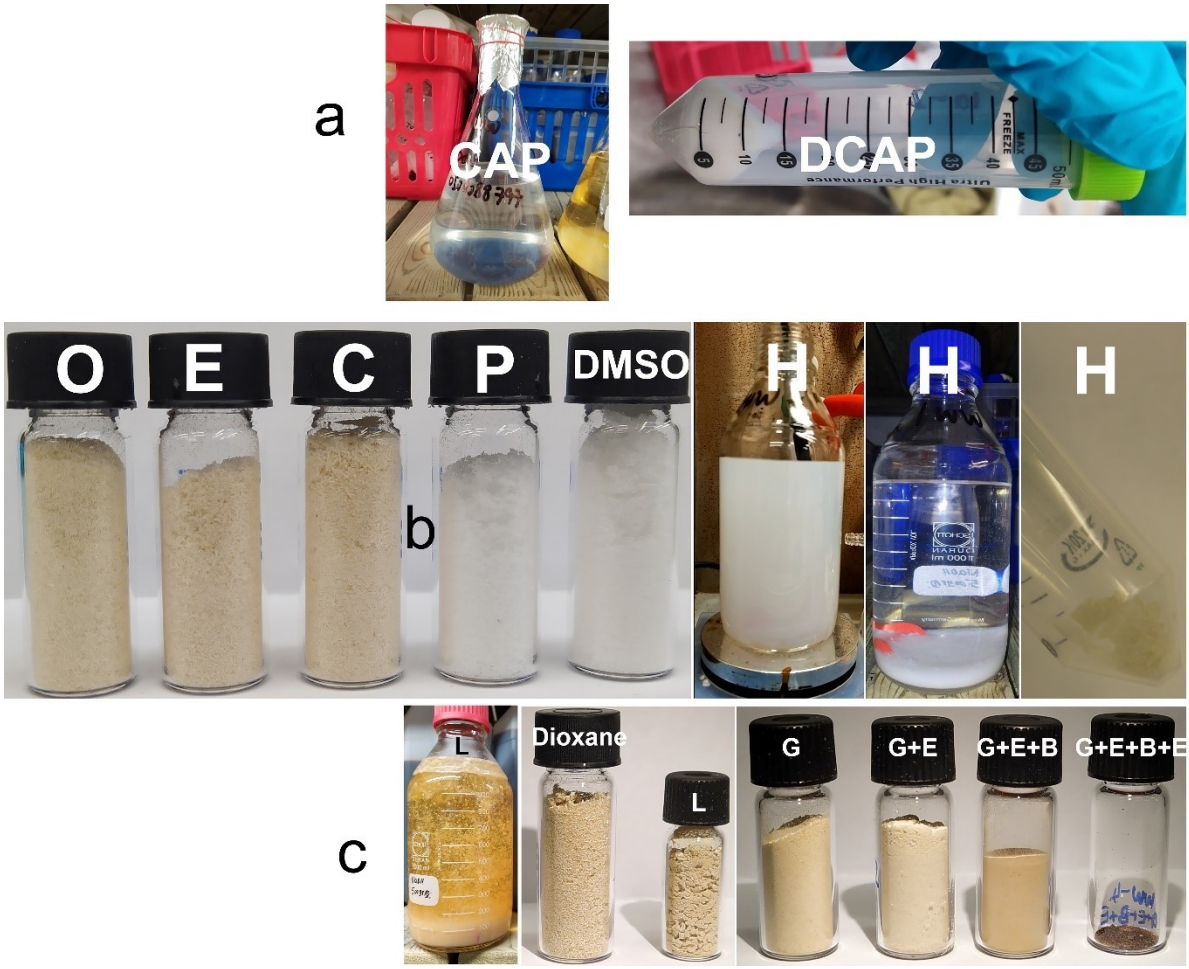
Dr. Jinze Dou (jinze.dou@aalto.fi), Department of Bioproducts and Biosystems, Aalto University

Prof. Tapani Vuorinen (tapani.vuorinen@aalto.fi), Department of Bioproducts and Biosystems, Aalto University

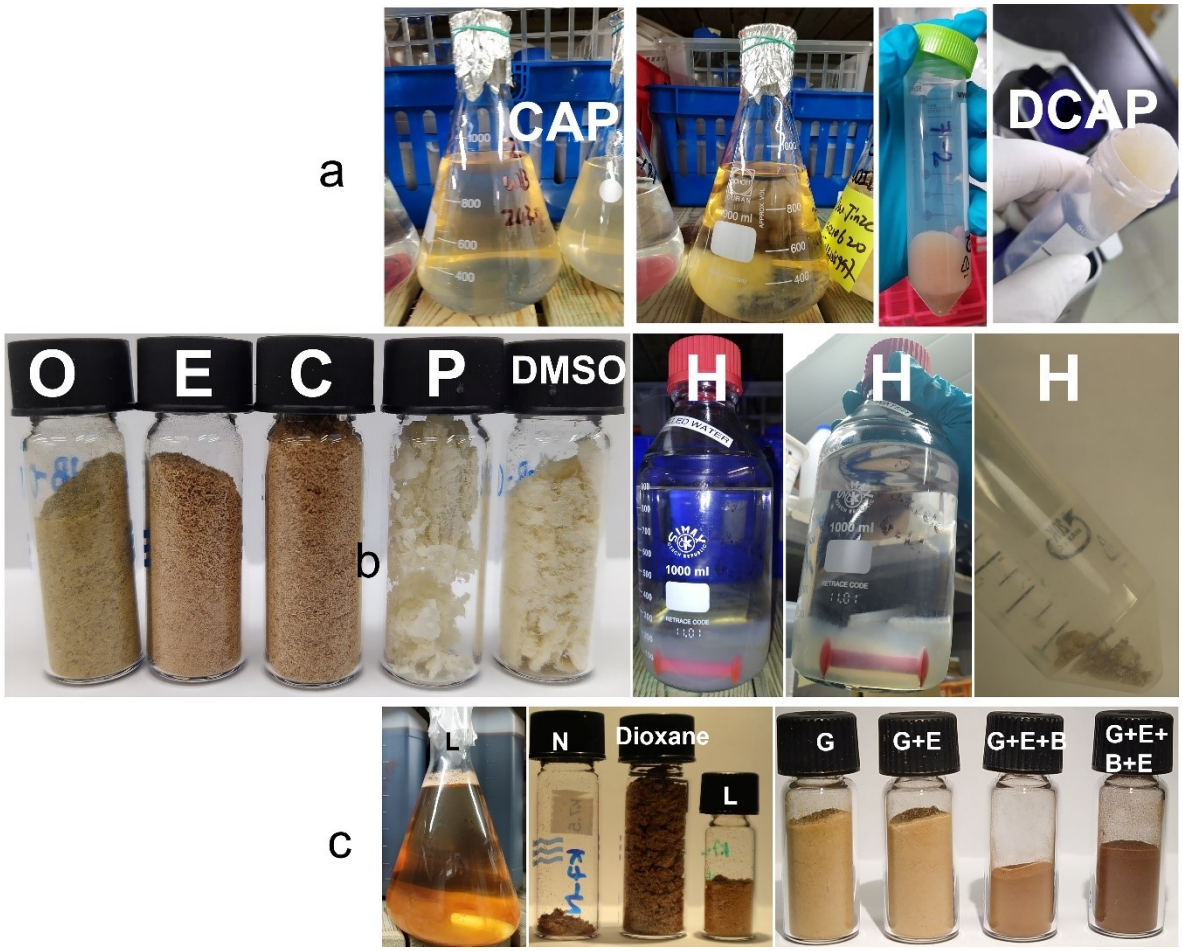
Lead Contact's email address: jinze.dou@aalto.fi

Supplementary Table 1. **Comparative methodologies of hemicellulose purification from switchgrass, wood (or wood pulp) and bark based on the conventional strategy (wood-based) and our strategy (bark).**

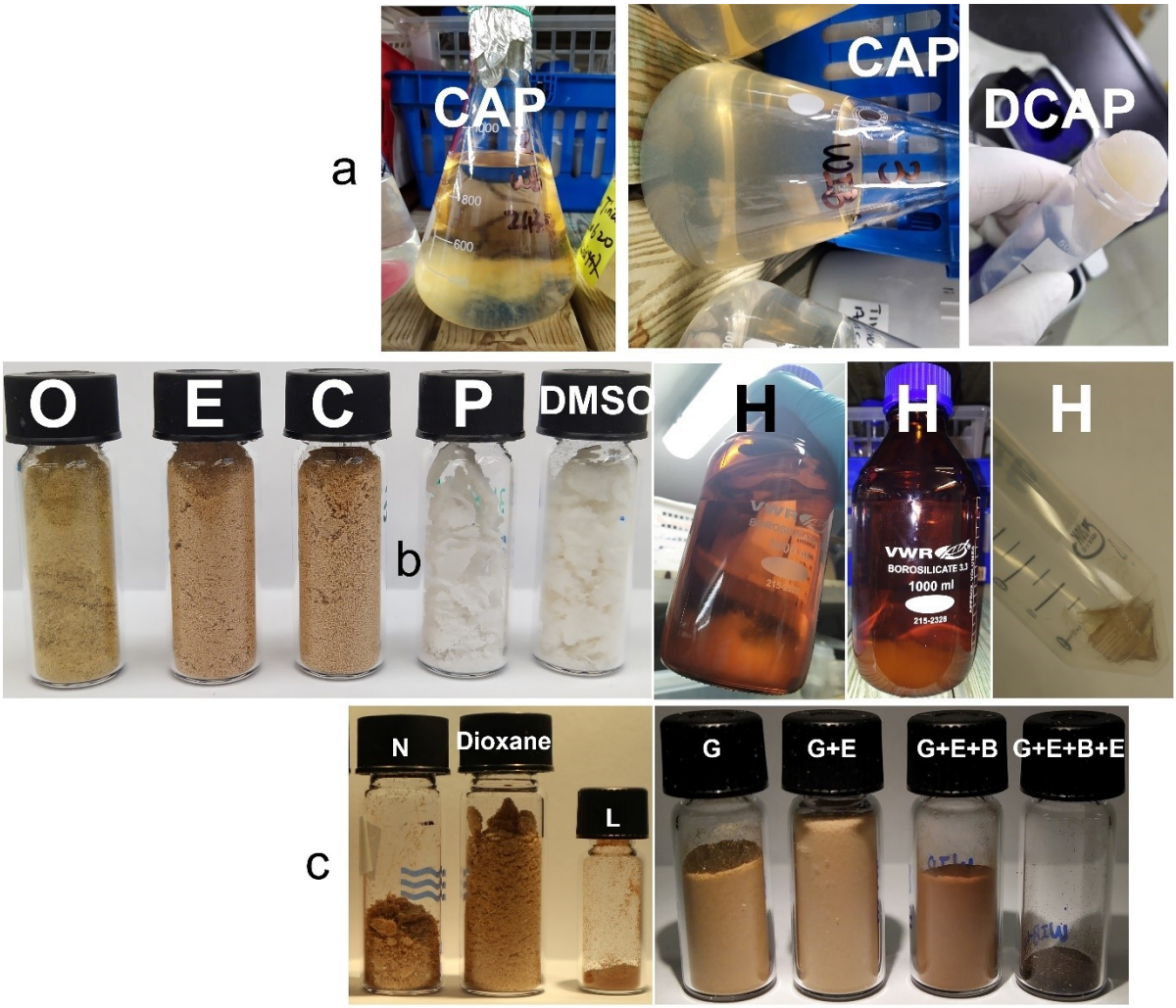
	methodology	mechanism	substrate	conditions	ref.
conventional strategy	sequential extraction with hot water and alkaline	cleaving ester bonds such as O-acetyl groups that is attached to xylan backbone	thermo-mechanical pulp from eucalyptus wood	2% NaOH, 90°C, 2.5 h	Sun et al. 2014[9]
	delignification with peracetic acid and followed with DMSO extraction	cleaving lignin's ester bonds	eucalyptus wood	85°C, 30 min, pH 4–5 (delignification with peracetic acid)	Evtugin et al. 2003[10]
	high pressure and hot water extraction	acetic acid formation promotes hemicellulose degradation	sulfite pulp from spruce and pine	hot water: 200-240°C, 5-60 min, 30-50 bars	Borrega et al. 2013[11]
	cellulolytic enzyme-aided extraction	hydrolytic cellulase selectively degrade cellulose	extracted and peracetic acid delignified switchgrass	ball milling and hydrolysis with 30 mg/g cellulase 37°C, 72 h	Ding et al. 2019[12]
our strategy	multi-stage pre-treatment followed by peracetic acid delignification and DMSO extraction	elimination of highly peracetic acid reactive components (i.e. extractives; pectin; protein; and tannins)	willow wood; willow bark (whole bark; inner bark; fibre bundle; and parenchyma)	multiple extraction (DCM; acetone; water; citric acid; pepsin; 0.1M NaOH) and peracetic acid delignification	present study



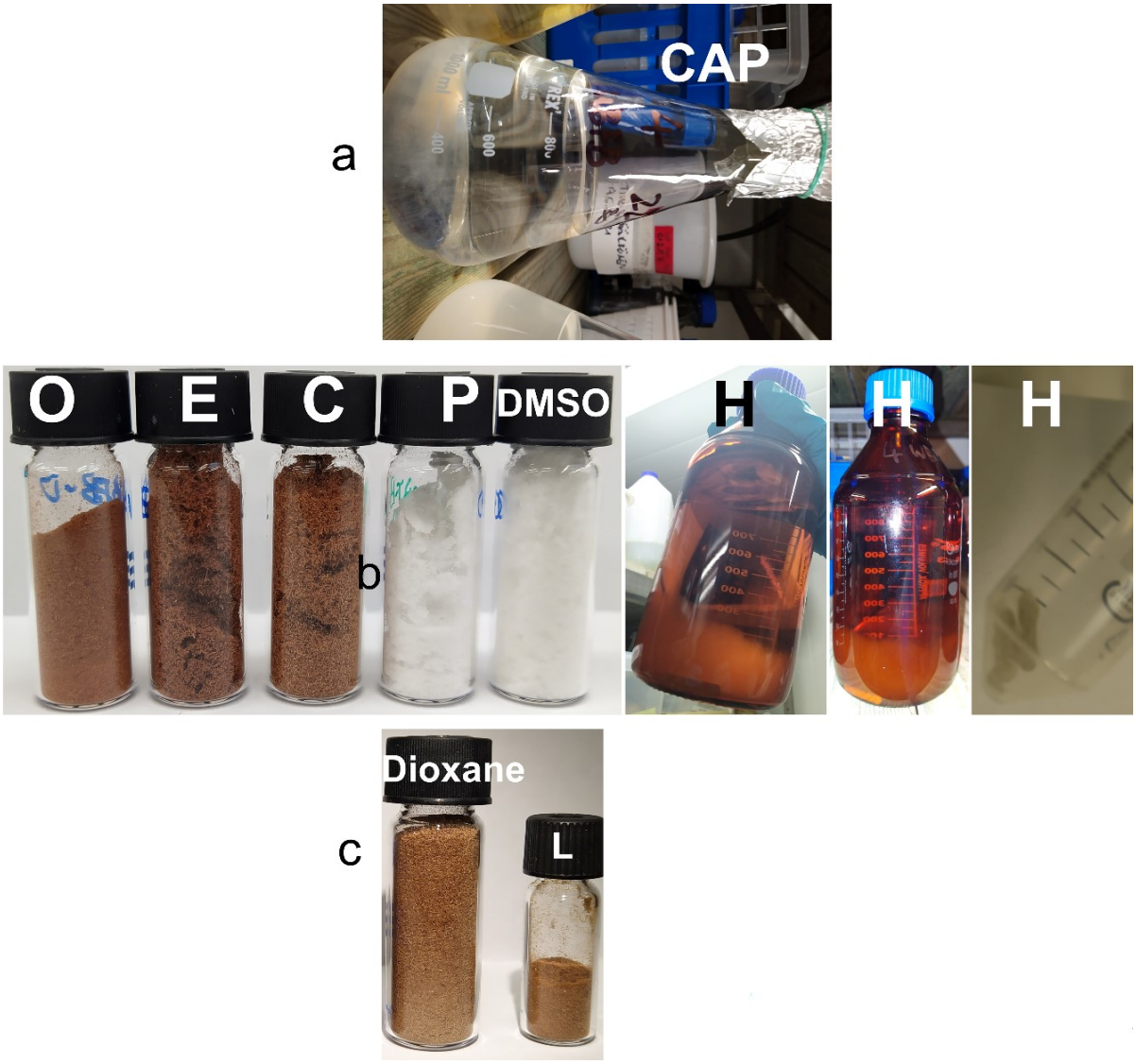
Supplementary Fig. 1 Photographs of the original (O), solid residues after the Soxhlet extraction (E), solid residues after the citric acid treatment (C) from the willow wood (WW). **a** Pectin purification -> citric acid treated pectin (CAP) and dialysis purified pectin (DCAP). **b** Hemicellulose recovery -> solid residues after the peracetic acid treatment (P) and the solid residues after DMSO extraction (DMSO) along with the purified hemicellulose (H). **c** Dioxane lignin recovery -> solid residues after the multiple extraction (citric acid; pepsin; 0.1M NaOH; dioxane-water) (Dioxane), recovered palish yellow color of dioxane lignin (L) in both precipitated form (under +5°C temperature) and solid form, and the reference lignin samples (WCW and CEL) from the previous study[15] (G: ground sample; G+E: solid residues after grinding and extraction; G+E+B (WCW): solid residues after the grinding, extraction, and ball milling; G+E+B+E (CEL): solid residues after grinding, extraction, ball milling, and cellulytic enzymatic purification).



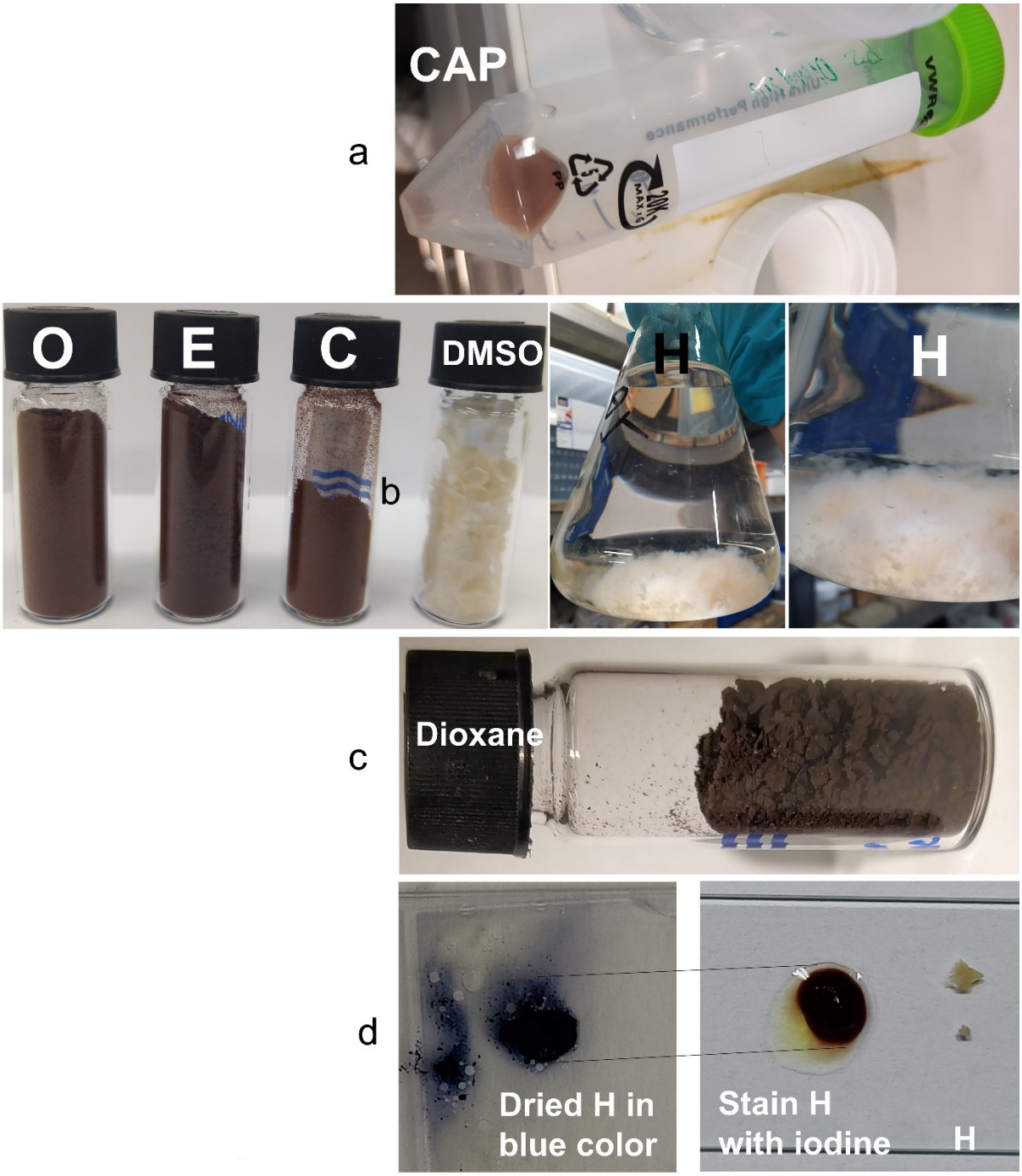
Supplementary Fig. 2 Photographs of the original (O), solid residues after the Soxhlet extraction (E), solid residues after the citric acid treatment (C) from the willow bark (WB). **a** Pectin purification -> citric acid treated pectin (CAP) and dialysis purified pectin (DCAP). **b** Hemicellulose recovery -> solid residues after the peracetic acid treatment (P) and the solid residues after DMSO extraction (DMSO) along with the purified hemicellulose (H). **c** Dioxane lignin recovery -> solid residues after the multiple extraction (citric acid; pepsin; 0.1M NaOH; dioxane-water) (Dioxane), recovered dioxane lignin (L) in both precipitated form (under +5°C temperature) and solid form, and the reference lignin samples (WCW and CEL) from the previous study[15] (G: ground sample; G+E: solid residues after the grinding and extraction; G+E+B (WCW): solid residues after the grinding, extraction, and ball milling; G+E+B+E (CEL): solid residues after grinding, extraction, ball milling, and cellulytic enzymatic purification).



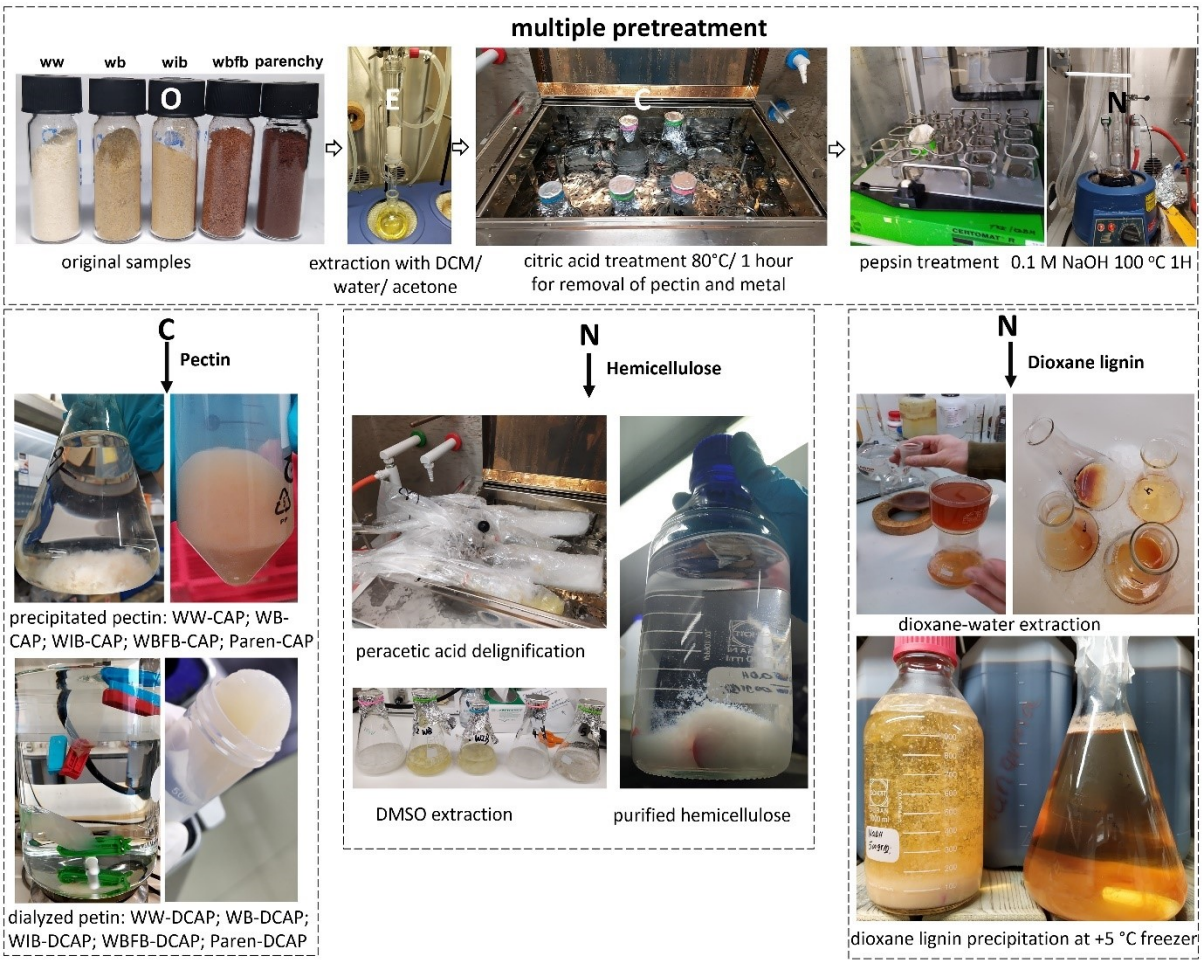
Supplementary Fig. 3 Photographs of the original (O), solid residues after the Soxhlet extraction (E), solid residues after the citric acid treatment (C) from the willow inner bark (WIB). a Pectin purification -> citric acid treated pectin (CAP) and dialysis purified pectin (DCAP). **b** Hemicellulose recovery -> solid residues after the peracetic acid treatment (P) and the solid residues after DMSO extraction (DMSO) along with the purified hemicellulose (H). **c** Dioxane lignin recovery -> solid residues after the multiple extraction (citric acid; pepsin; 0.1M NaOH; dioxane-water) (Dioxane), recovered dioxane lignin (L) in both precipitated form (under +5°C temperature) and solid form, and the reference lignin samples (WCW and CEL) from the previous study[15] (G: ground sample; G+E: solid residues after the grinding and extraction; G+E+B (WCW): solid residues after the grinding, extraction, and ball milling; G+E+B+E (CEL): solid residues after grinding, extraction, ball milling, and cellulytic enzymatic purification).



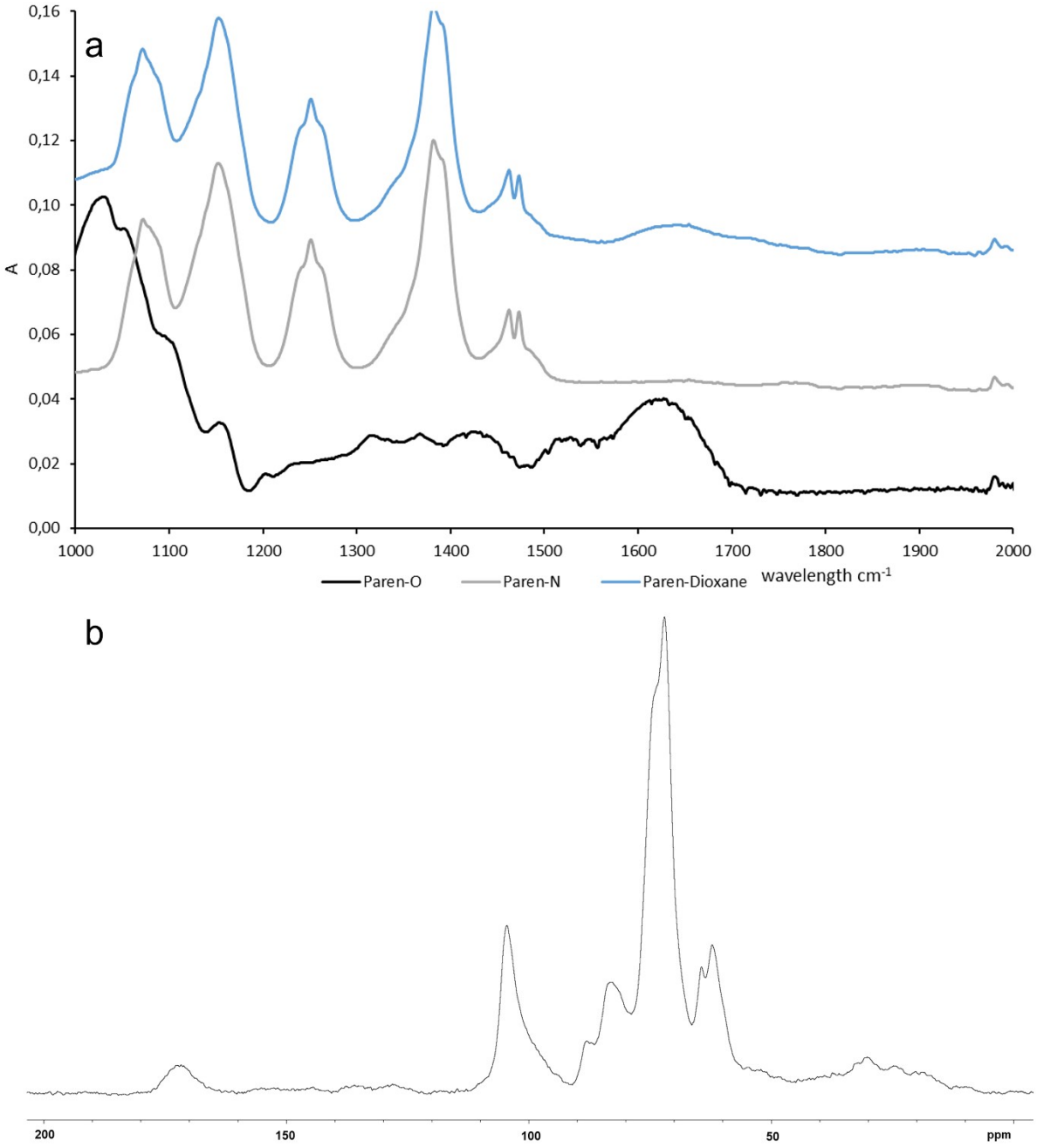
Supplementary Fig. 4 Photographs of the original (O), solid residues after the Soxhlet extraction (E), solid residues after the citric acid treatment (C) from the willow bark fiber bundle (WBFB). **a** Pectin purification -> citric acid treated pectin (CAP) and dialysis purified pectin (DCAP). **b** Hemicellulose recovery -> solid residues after the peracetic acid treatment (P) and the solid residues after DMSO extraction (DMSO) along with the purified hemicellulose (H). **c** Dioxane lignin recovery -> solid residues after the multiple extraction (citric acid; pepsin; 0.1M NaOH; dioxane-water) (Dioxane), recovered dioxane lignin (L).



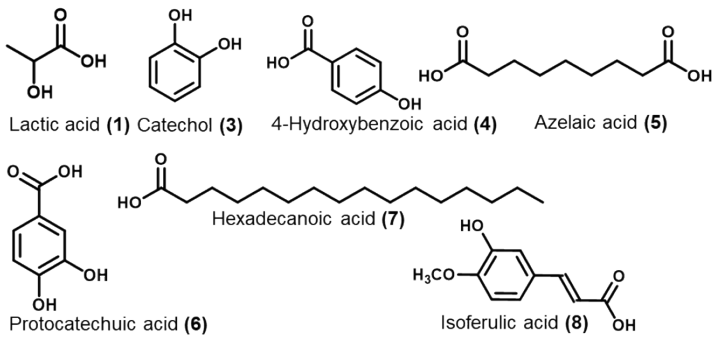
Supplementary Fig. 5 Photographs of the original (O), solid residues after the Soxhlet extraction (E), solid residues after the citric acid treatment (C) from parenchymatous tissues (parenchyma). **a** Pectin purification -> citric acid treated pectin (CAP). **b** Hemicellulose recovery -> solid residues after DMSO extraction (DMSO) along with the purified hemicellulose (H); **c** Dioxane lignin recovery -> solid residues after the multiple extraction (citric acid; pepsin; 0.1M NaOH; dioxane-water) (Dioxane); **d** Starch test with Lugol's Iodine (blue color was formed immediately when a drop of Lugol's solution was dropped on the hemicellulose that is on top of the microscopy slide).



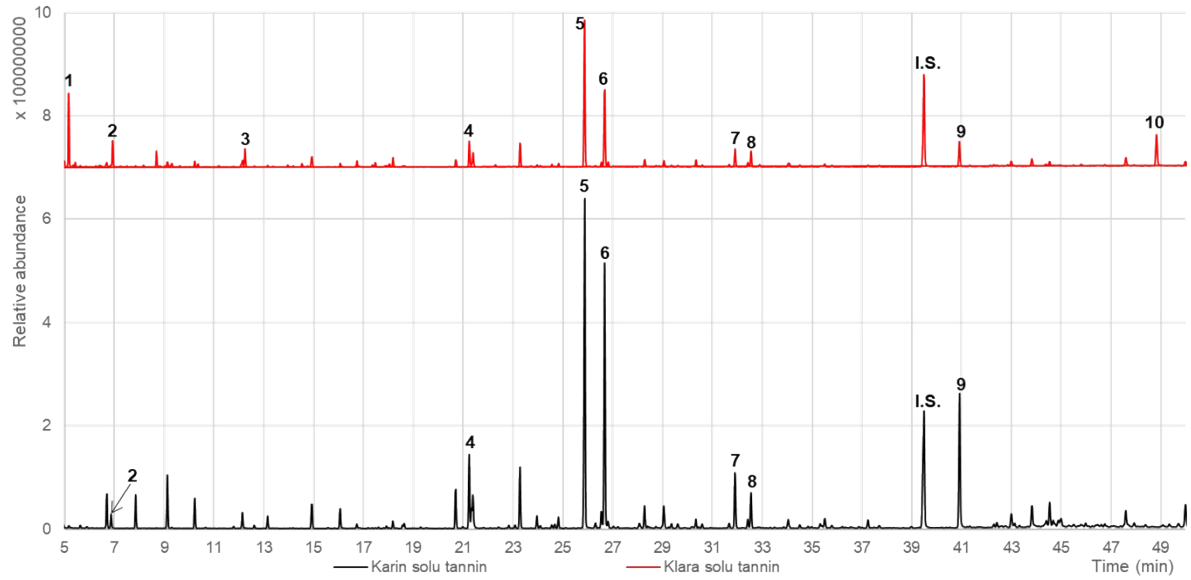
Supplementary Fig. 6 Fractionation scheme of the recovery of pectin; hemicellulose; and dioxane lignin from willow wood (WW); bark (WB); inner bark (WIB); fiber bundle (WBFB) and parenchyma tissues (parenchyma).



Supplementary Fig. 7 Chemical characteristics of the purified parenchyma tissues. a FT-IR. **b** CPMAS NMR spectrum. Parenchyma-N and Parenchyma-Dioxane were included as references for FT-IR comparison.

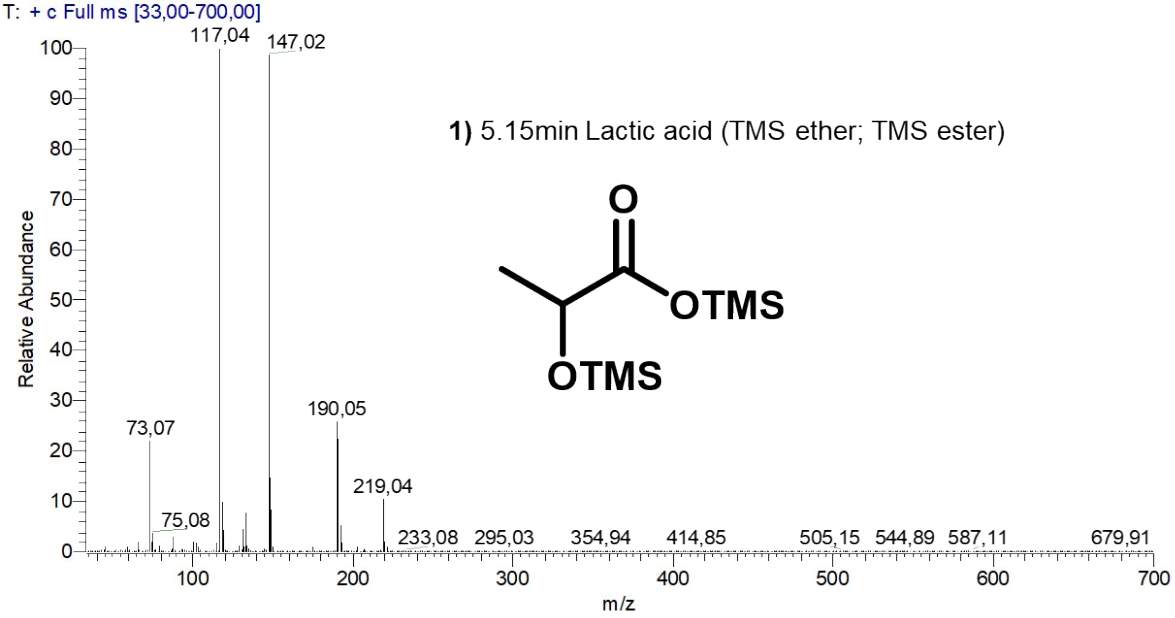


Supplementary Fig. 8 Detected diethyl ether-soluble 0.1M NaOH hydrolysable tannin-like substances (HTS) from Karin and Klara bark based on GC-MS total-ion chromatogram. Mass spectra of these trimethylsilylated compounds were summarized at **Supplementary Fig. 10**.



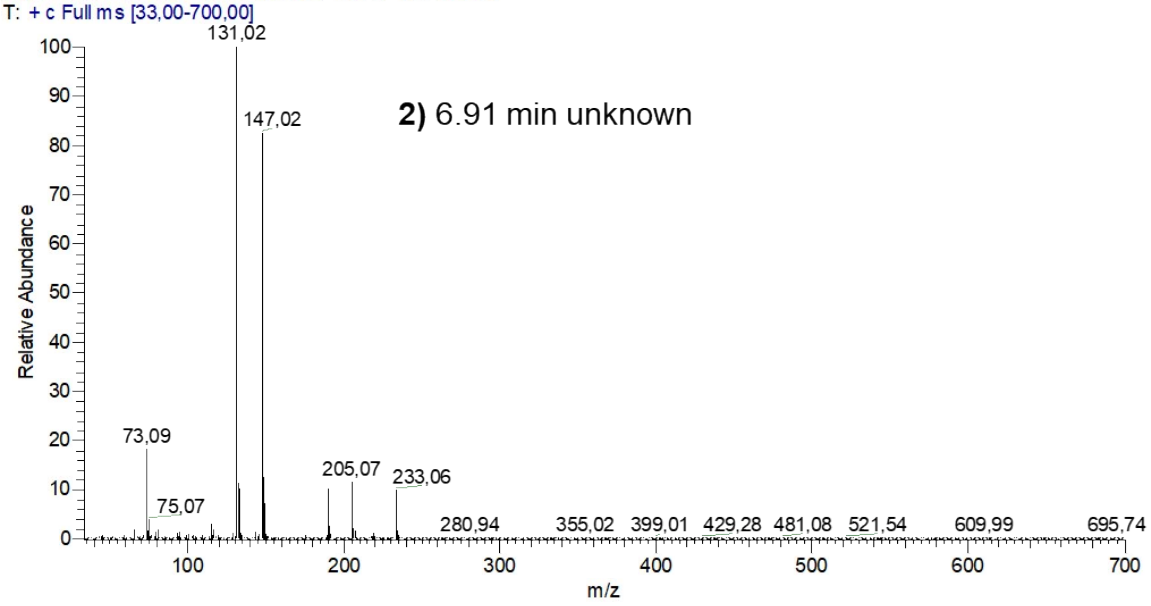
Supplementary Fig. 9 GC-MS total-ion chromatogram of the diethyl ether-soluble HTS from Karin and Klara bark.

Klara03Sol #95-98 RT: 5,13-5,20 AV: 4 NL: 1,84E7
T: + c Full ms [33,00-700,00]



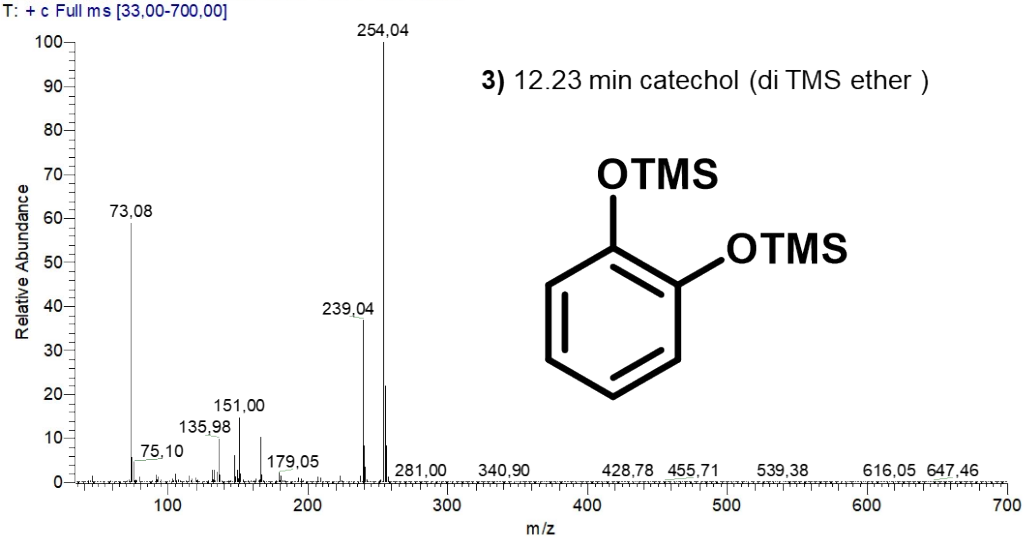
1) 5.15min Lactic acid (TMS ether; TMS ester)

Klara03Sol #173-176 RT: 6,89-6,96 AV: 4 NL: 8,08E6
T: + c Full ms [33,00-700,00]



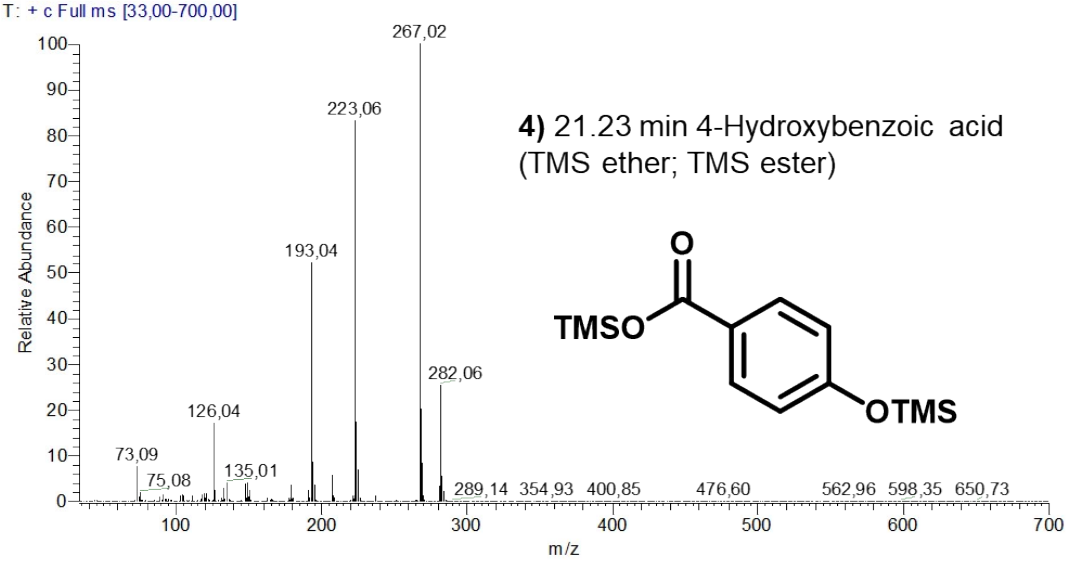
2) 6.91 min unknown

Klara03Sol #408-411 RT: 12,21-12,28 AV: 4 NL: 5,13E6
T: + c Full ms [33,00-700,00]



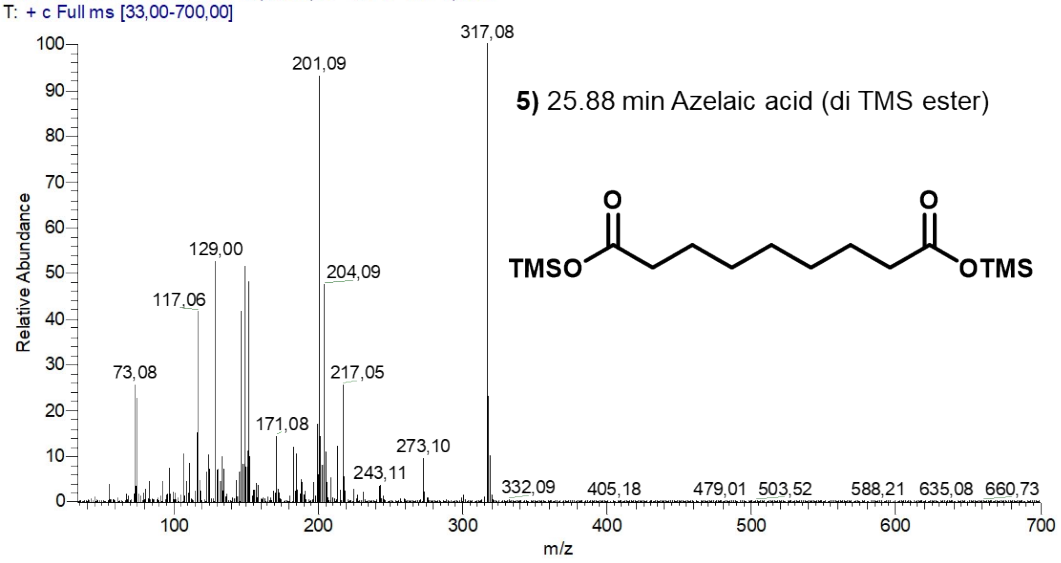
3) 12.23 min catechol (di TMS ether)

Klara03Sol #806-809 RT: 21,21-21,28 AV: 4 NL: 6,27E6
T: + c Full ms [33,00-700,00]



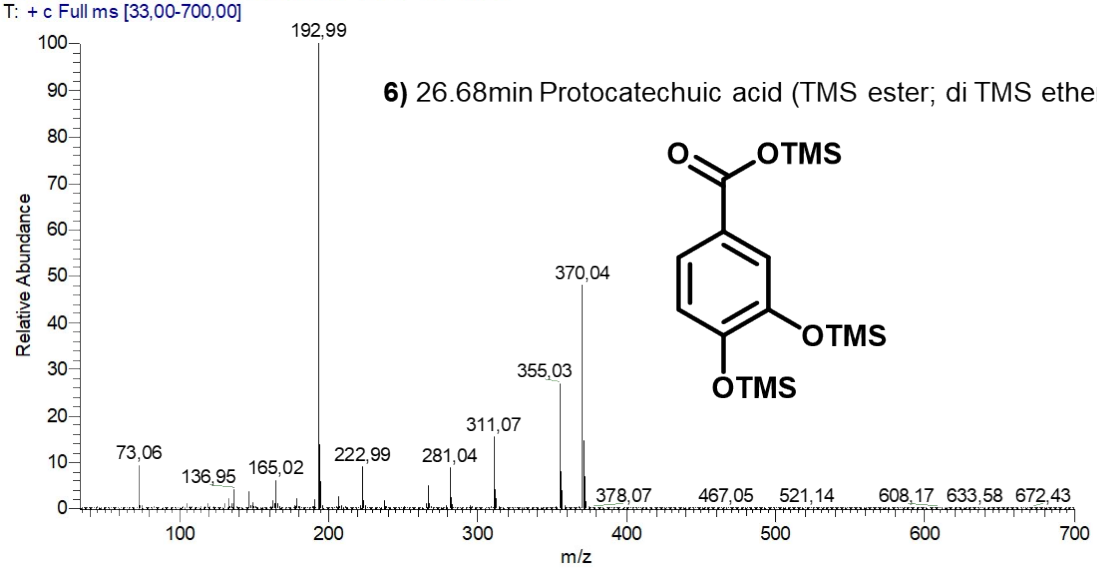
4) 21.23 min 4-Hydroxybenzoic acid (TMS ether; TMS ester)

Karin03Sol #1011-1014 RT: 25,84-25,91 AV: 4 NL: 3,62E7
T: + c Full ms [33,00-700,00]



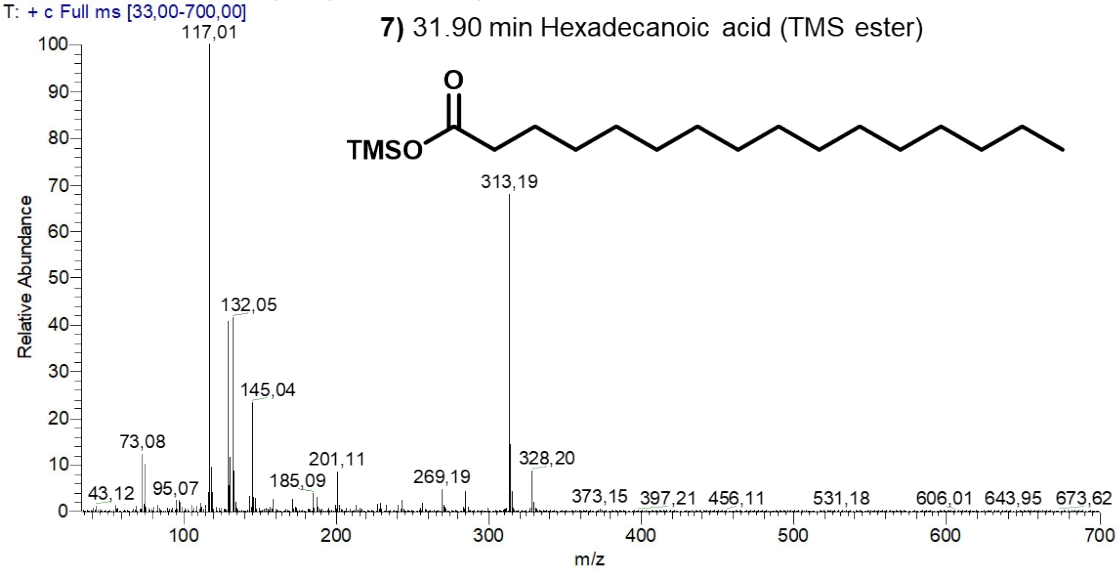
5) 25.88 min Azelaic acid (di TMS ester)

Karin03Sol #1046-1050 RT: 26,63-26,72 AV: 5 NL: 6,78E7
T: + c Full ms [33,00-700,00]



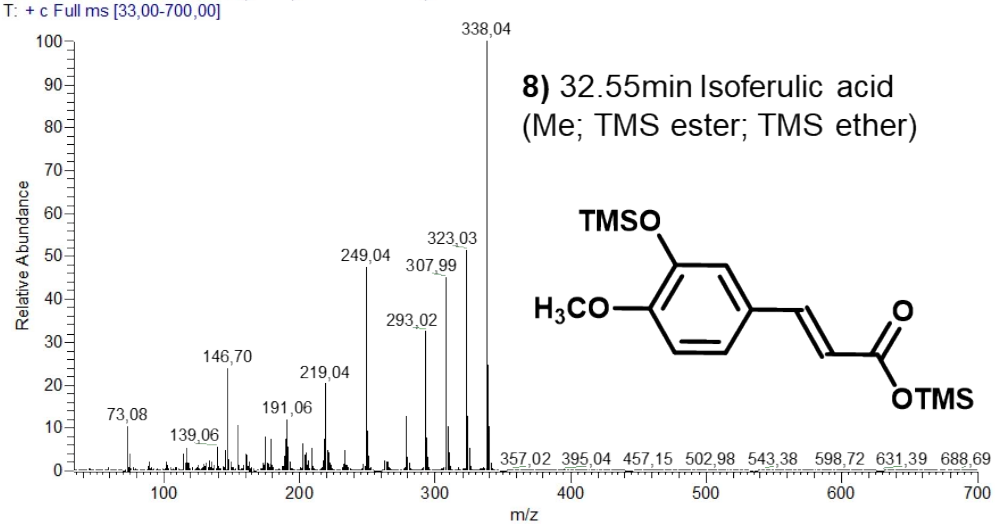
6) 26.68min Protocatechuic acid (TMS ester; di TMS ether)

Karin03Sol #1277-1281 RT: 31,85-31,94 AV: 5 NL: 1,12E7
T: + c Full ms [33,00-700,00]

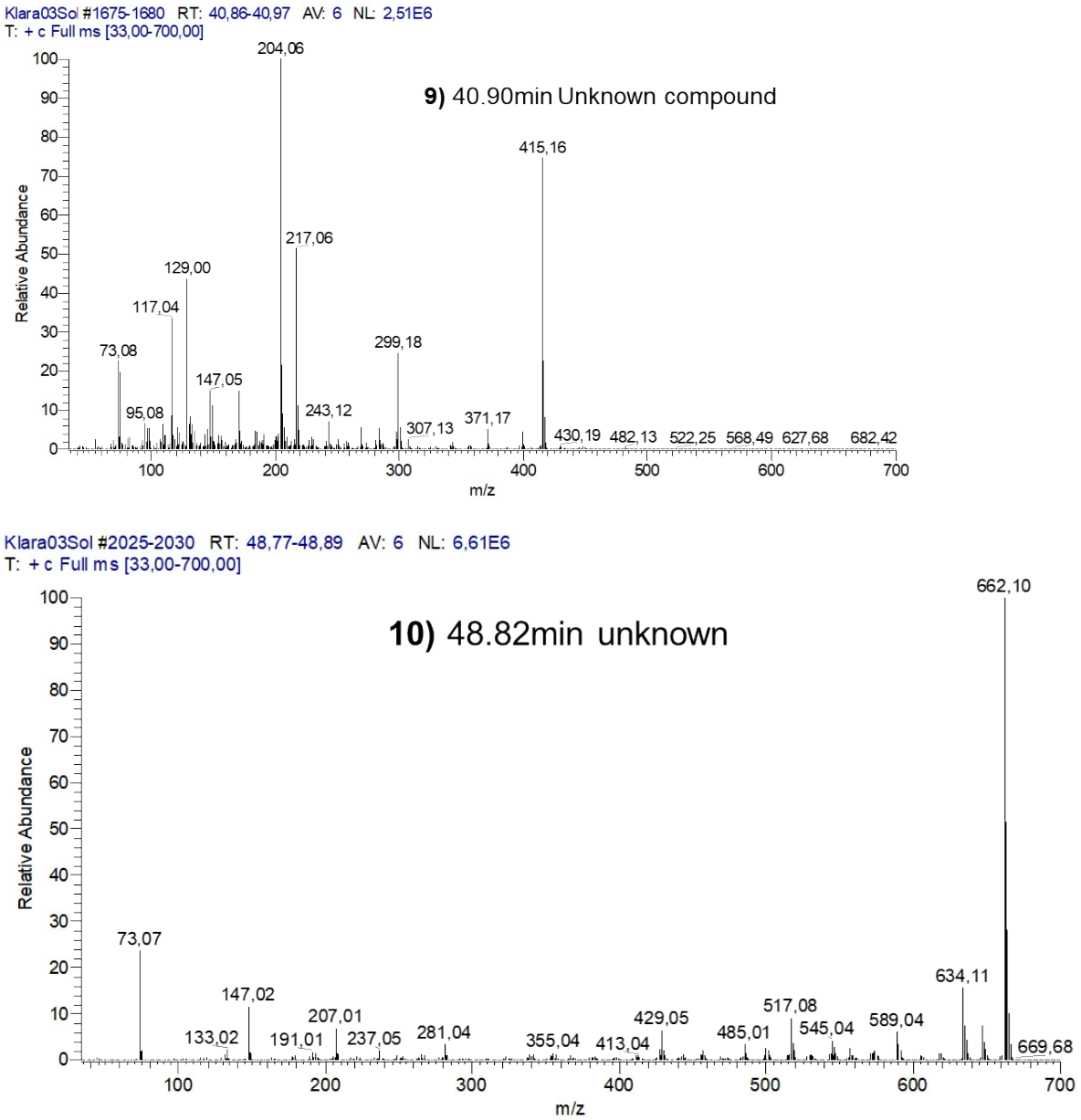


7) 31.90 min Hexadecanoic acid (TMS ester)

Karin03Sol #1306-1309 RT: 32,51-32,58 AV: 4 NL: 6,24E6
T: + c Full ms [33,00-700,00]



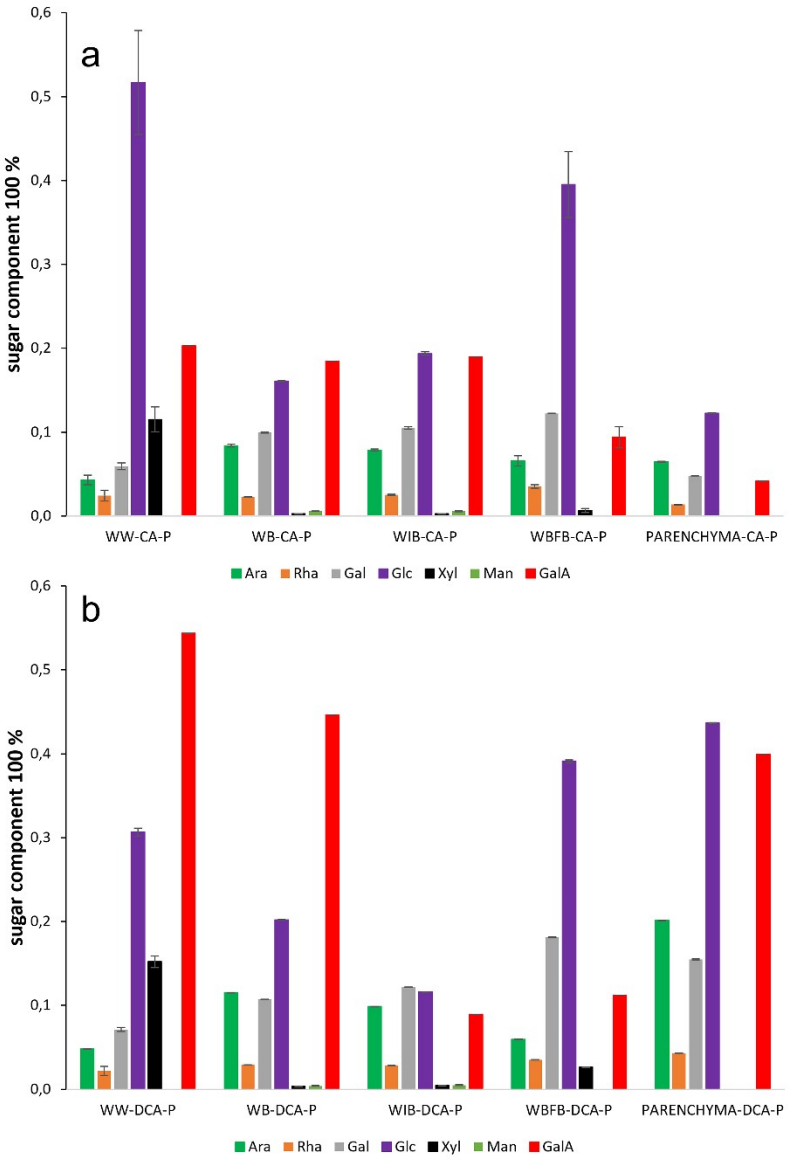
8) 32.55min Isoferulic acid
(Me; TMS ester; TMS ether)



Supplementary Fig. 10 Mass spectra of trimethylsilylated diethyl ether-soluble HTS detected from Karin and Klara bark.

Supplementary Table 2. Main diethyl ether soluble HTS (supplementary Fig. 8) obtained from Karin/ Klara bark. MS spectra were referenced with literature and publicly available databases [DB]: NIST Chemistry WebBook [N].

Retention time, min	Peak number	Compounds, detected as methyl ester (Me) and/or TMS ester/ether (TMS) derivatives	Characteristic fragments m/z	MW of TMS derivatives	DB/ Reference
Aromatic compound					
12.23	3	Catechol (di TMS ether)	254; 239; 151; 136; 75; 73	254	[N]
21.23	4	4-Hydroxybenzoic acid (TMS ether; TMS ester)	282; 267; 223; 193; 126; 73	282	[N]
26.68	6	Protocatechuic acid (TMS ester; di TMS ether)	370; 355; 311; 281; 223; 193; 73	370	[N]
32.55	8	Isoferulic acid (Me; TMS ester; TMS ether)	338; 323; 308; 293; 249; 219; 191; 146; 73	338	[N]
Fatty acids					
Saturated fatty acids (Mono-carboxylic acid)					
31.90	7	Hexadecanoic acid (TMS ester)	328; 313; 269; 145; 132; 117; 75; 73	328	[N]; [Branco et al. 2020][50]
Saturated fatty acids (alpha, omega-dicarboxylic acid)					
25.87	5	Azelaic acid (di TMS ester)	317; 217; 201; 171; 129; 117; 75; 73	332	[N]
2-Hydroxyacids (aliphatic)					
5.15	1	Lactic acid (TMS ether; TMS ester)	219; 190; 147; 117; 75; 73	234	[N]
unknown compound					
6.91	2	Unknown compound	233; 205; 147; 131; 75; 73	n.d.	n.d.
40.90	9	Unknown compound	415; 299; 217; 204; 129; 117; 73	n.d.	n.d.
48.82	10	Unknown compound	662; 634; 147; 73	n.d.	n.d.



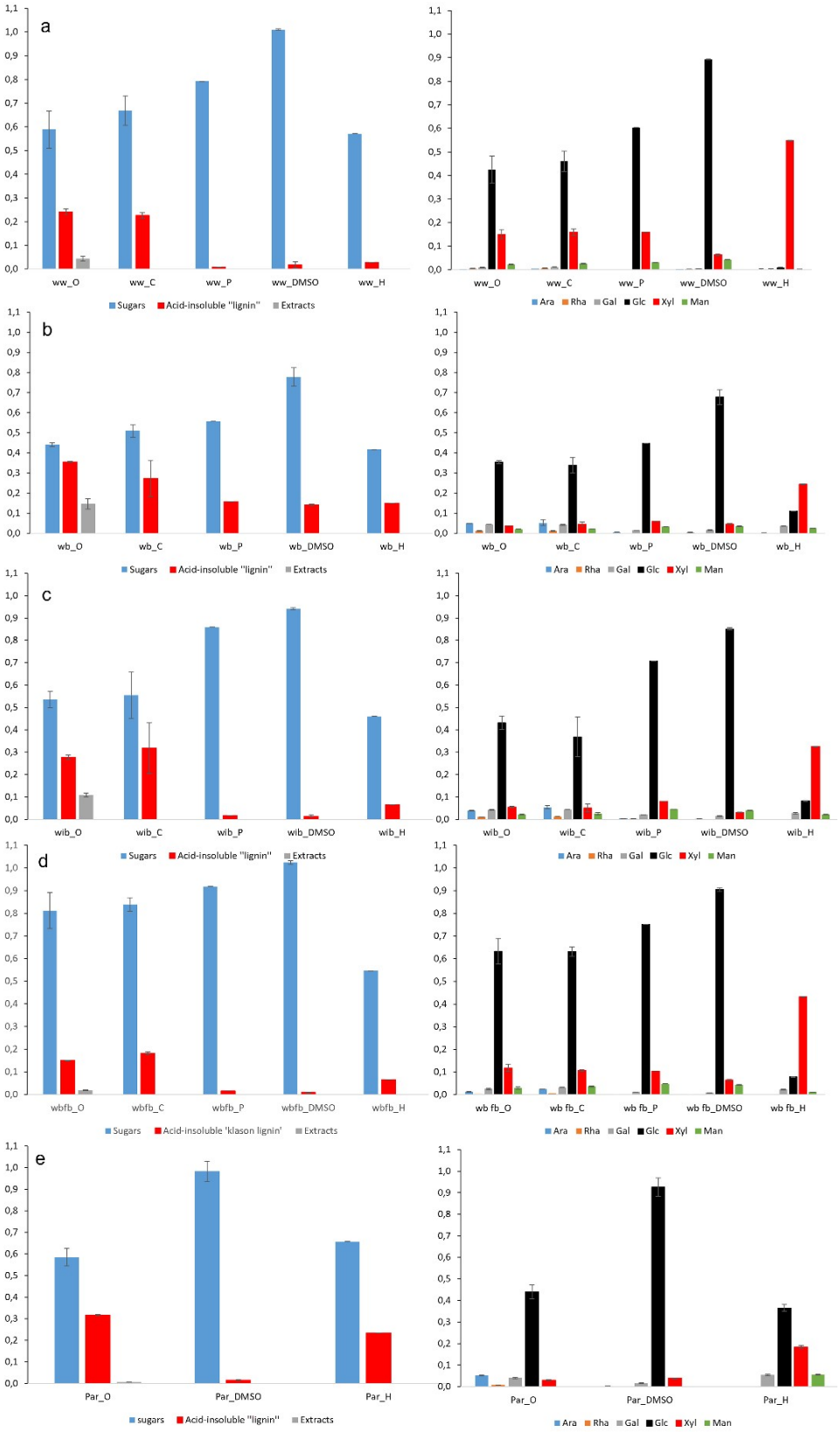
Supplementary Fig. 11 Carbohydrate composition (% of anhydro sugars in the monosaccharide) from five different samples. a Purified pectin (CAP). **b** Dialysis treated pectin (DCAP) that is collected through the scheme (**Supplementary Fig. 6**).

Supplementary Table 3. **Structural features of the ethanol precipitated pectin after dialysis (DCA-P) from five different samples: wood; bark; inner bark; fibre bundle; and parenchyma tissues.** The relative amount of linkage patterns of each neutral sugar was estimated from volume integrals of $^{1}H–^{13}C$ correlation contours (**Fig. 2**).

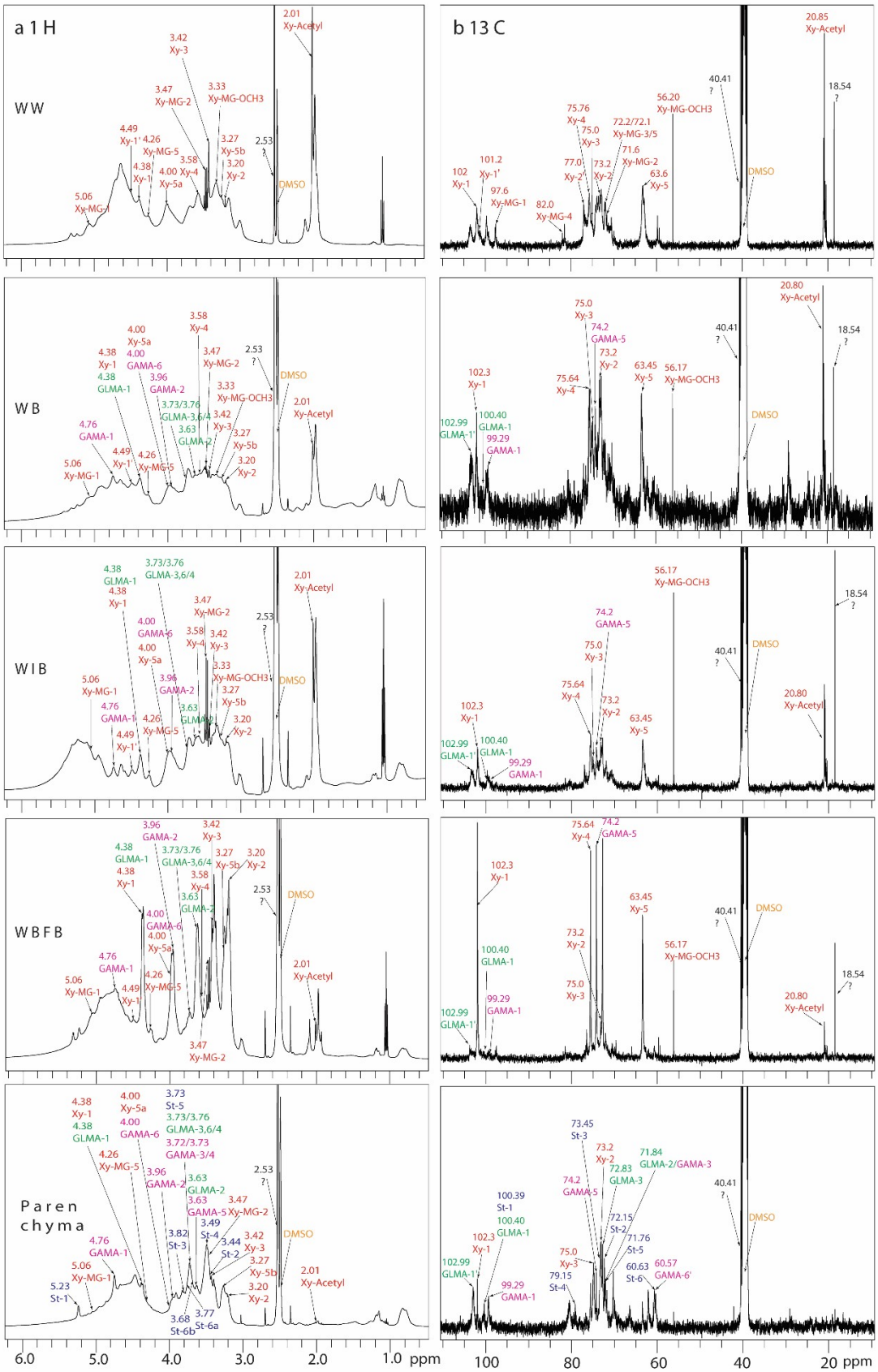
linkages	abbreviation	WW-DCAP	WB-DCAP	WIB-DCAP	WBFB-DCAP	Parenchy-DCAP
Arabinose Interunit Linkages (%)						
→5)-a-Araf-(1→	A1,5	61	98	82	90	78
→2,3,5)-a-Araf-(1→	A2,3,5	0	2	7	0	13
→2,5)-a-Araf-(1→	A2,5	17	0	11	10	9
β-Araf-(1→	A1	22	0	0	0	0
Rhamnose Interunit Linkages (%)						
→2)-a-Rhap-(1→	R1,2	n.d.	50	57	n.d.	n.d.
→2,4)-a-Rhap-(1→	R1,2,4	n.d.	50	43	n.d.	n.d.
Galactose Interunit Linkages (%)						
β-Galp-(1→	Ga	13	21	6	16	0
→4)-β-Galp-(1→	Ga1,4	87	79	94	84	100

Supplementary Table 4. Chemical shift assignments for dialyzed citric acid extracted pectin (DCA-P) (Supplementary Fig. 6) of willow wood (WW) and bark (WB; WIB; WBFB; and Parenchyma), referenced to TSP-*d4* (δC/δH, 0/ 0 ppm).

Residue	Label[1]	C-1 *H-1*	C-2 *H-2*	C-3 *H-3*	C-4 *H-4*	C-5 *H-5;5′*	C-6 *H-6;6′*	$\underline{C}H_3CO$	$CH_3\underline{C}O$	$\underline{C}H_3O$
→4)-α-Gal*p*A-6-*O*Me-(1→	Gal*p*A (*OMe*)	102.9 *5.4*	70.8 *3.7*	71.6 *4.11*	82.1 *4.31*	73.5 *5.1*				55.8 *3.81*
→4)-α-Gal*p*A-3-*O*Ac-(1→	Gal*p*A (OAc)	102.9 *5.4*	*71.6* *4.11*	*73.7* *5.14*	*n.d.* *n.d.*	73.5 *5.1*		23.5 *2.19*		
→3, 4)-α-Gal*p*A-(1→	Gal*p*A1,3,4	101.5 5.08	*n.d.* *n.d.*	*n.d.* *n.d.*	*n.d.* *n.d.*	*n.d.* *n.d.*				
→4)-α-Gal*p*A-(1→	Gal*p*A	102.9 *5.4*	69.3 *3.83*	*69.7* *3.9*	*83.2* *4.27*	74.2 *4.73*				
→5)-α-Ara*f*-(1→	$A_{1,5}$	110.6 *5.09*	84.2 *4.15*	79.5 *3.98*	85.4 *4.22*	69.4 *3.81*				
→3)-α-Ara*f*-(1→	$A_{1,3}$	109.4 *5.26*	84.2 *4.31*	87 *4.07*	84.5 *4.23*	64.1 *3.84;3.73*				
→2,3,5)-α-Ara*f*-(1→	$A_{2,3,5}$	110.1 *5.18*	87.96 *4.32*	83.3 *4.26*	85.2 *4.3*	69.2 *3.95*				
→2,5)-α-Ara*f*-(1→	$A_{2,5}$	110.1 *5.18*	89.6 4.17	n.d. n.d.	n.d. n.d.	n.d. n.d.				
β-Ara*f*-(1→	A_1-β	102.7 *5.42*	79.7 *4.03*	76.5 *3.98*	85.2 *3.91*	65.5 *3.82*				
α-Ara*f*-(1→	A_1-α	*112.3* *5.25*	79.7 *4.03*	76.5 *3.98*	85.2 *3.91*	65.5 *3.82*				
→2)-a-Rhap-(1→	R1,2	112.2 5.26	78.3 4.18	69.7 3.89	73.4 3.43	72.9 3.67	19.6 1.26			
→2,4)-a-Rhap-(1→	R1,2,4	112.2 5.26	78.3 4.18	72.7 4.06	84 3.79	69.9 3.87	19.6 1.26			
β-Gal*p*-(1→	Ga	106.1 *4.52*	74.7 *3.53*	*75* *3.65*	n.d. *n.d.*	77.6 *3.72*	63.6 *3.84*			
→4)-β-Gal*p*-(1→	Ga1,4	107.4 *4.65*	75.9 *3.69*	76.4 *3.78*	80.7 *4.17*	79.8 *3.67*	63.7 *3.83*			
→4)-α-D-Glc*p*-(1→	St1,4	102.5 *5.41*	74.5 *3.64*	76.5 *3.98*	79.7 *3.66*	74.2 *3.84*	63.7 *3.83*			
α-Xyl-(1→	X	101.5 *5.08*	71.6 *3.82*	79.4 *3.84*	70.1 *4.34*	63.1 (5a); 5b overlap *3.46 (5a);5b overlap*				



Supplementary Fig. 12 Overall chemical composition (% of the dry mass) (left) and carbohydrate composition (% of anhydro sugars in the monosaccharide) (right) from five different samples. a Willow wood. **b** Willow bark. **c** Willow inner bark. **d** Willow bark fibre bundle. **e** Parenchyma tissues that are collected from their original form (O), solid residues after the citric acid treatment (C), solid residues after the peracetic acid delignification (P), solid residues after the DMSO extraction (DMSO), and the recovered hemicellulose (H). (**Supplementary Fig. 6)**



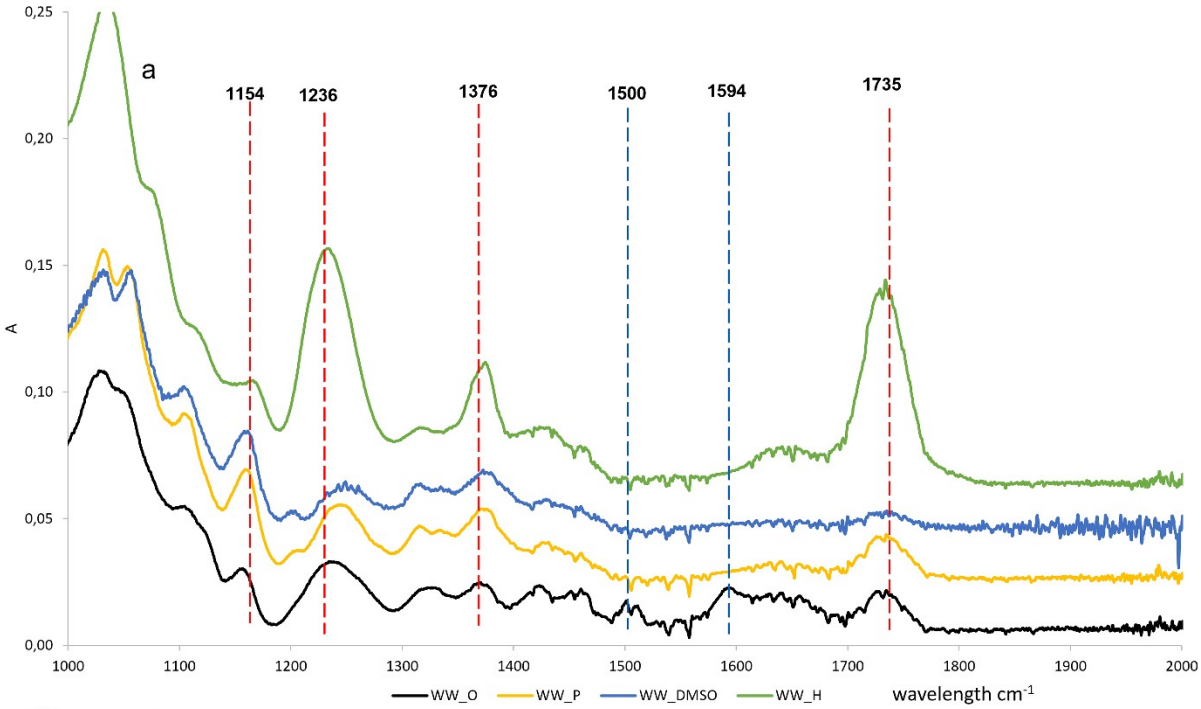
Supplementary Fig. 13 NMR spectrum of the purified hemicellulose from wood (WW); bark (WB); inner bark (WIB); fiber bundle (WB FB) and parenchyma tissues (Parenchyma). a ^{1}H. **b** ^{13}C. For full spectrum of the ^{13}C spectra, see **Supplementary Fig. 15**. For assignments, see **Supplementary Table 5–6**.

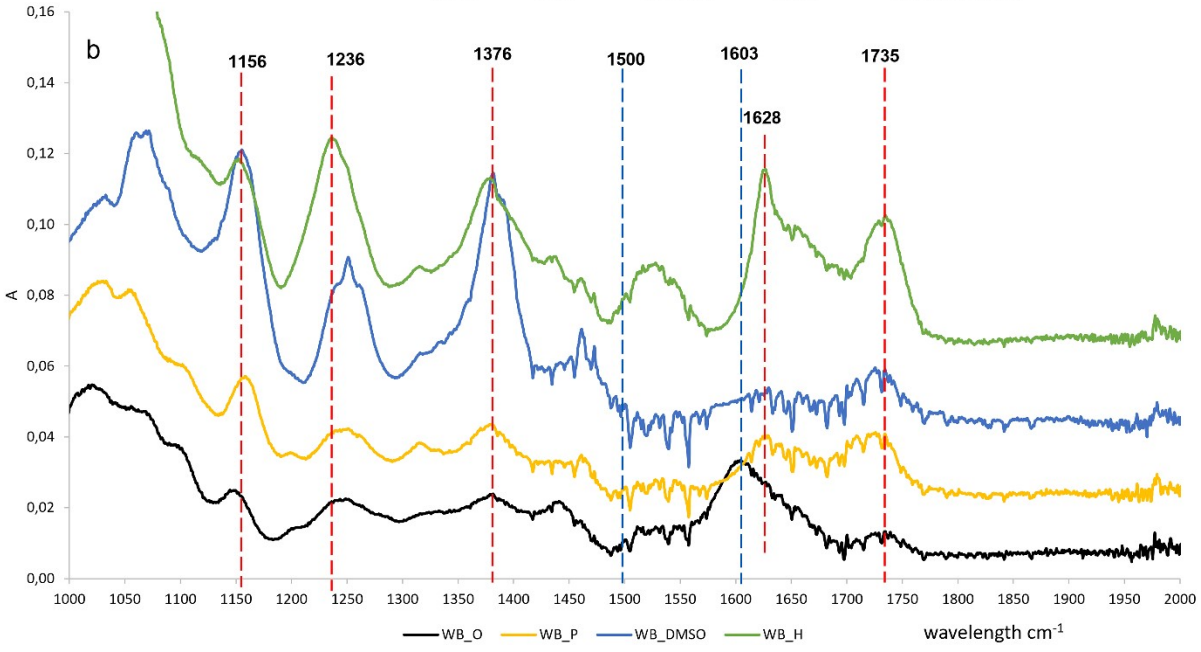
Supplementary Table 5. **Assignments of the ^{1}H NMR spectra of willow hemicelluloses (WW; WB; WIB; WBFB; Parenchyma) in comparison to the literatures[9, 34] and authentic starch.**

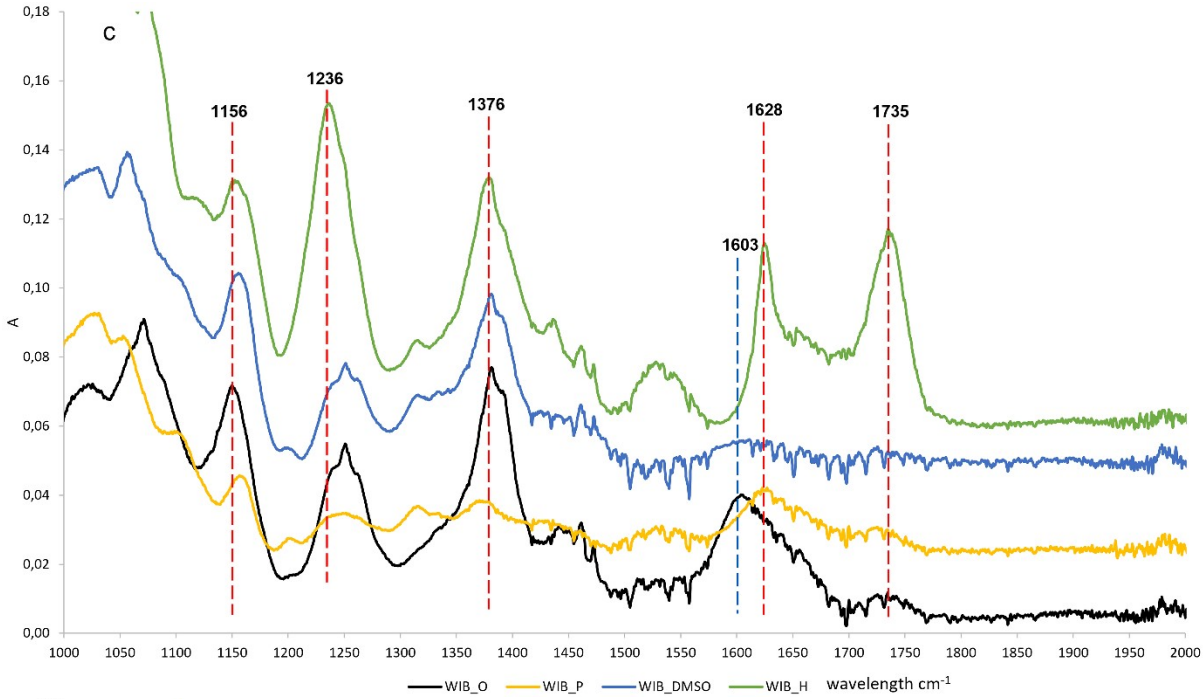
	abbreviation	WW	WB	WIB	WBFB	parenchyma	Yuan et al. 2010[34]	Sun et al. 2013[9]	authentic starch
(1-4)-β-D-Xylp (1)	Xy-1	4.38	4.38	4.38	4.38	4.38	4.41	4.28	
(1-4)-β-D-Xylp (5b)	Xy-5b	4.00	4.00	4.00	4.00	4.00	4.04	3.92	
(1-4)-β-D-Xylp (4)	Xy-4	3.58	3.58	3.58	3.58	—	3.72	3.61	
(1-4)-β-D-Xylp (3)	Xy-3	3.42	3.42	3.42	3.42	3.42	3.50	3.34	
(1-4)-β-D-Xylp (5a)	Xy-5a	3.27	3.27	3.27	3.27	3.27	3.32	3.22	
(1-4)-β-D-Xylp (2)	Xy-2	3.20	3.20	3.20	3.20	3.20	3.23	3.13	
4-O-Me-α-D-Glucuronic acid (C1H1)	Xy-MG-1	5.06	5.06	5.06	5.06	5.06	5.22	5.13	
4-O-Me-α-D-Glucuronic acid (C5H5)	Xy-MG-5	4.26	4.26	4.26	4.26	4.26	—	4.15	
4-O-Me-α-D-Glucuronic acid (C2H2)	Xy-MG-2	3.47	3.47	3.47	3.47	3.47	—	3.49	
methoxy group	Xy-MG-OCH3	3.33	3.33	3.33	—	—	3.40	3.34	
acetyl CH3 (-C=O CH3)	Xy-Acetyl	2.01	2.01	2.01	2.01	2.01	—	—	
(1-4)-β-D-Xylp-1-O-(4-OMe-a-D-GlcpA)	Xy-1'	4.49	4.49	4.49	4.49	—	4.58	—	
Starch (1)	St-1					5.23			5.228
Starch (2)	St-2					3.44			3.44
Starch (3)	St-3					3.82			3.82
Starch (4)	St-4					3.49			3.49
Starch (5)	St-5					3.73			3.73
Starch (6a)	St-6a					3.77			3.77
Starch (6b)	St-6b					3.68			3.68

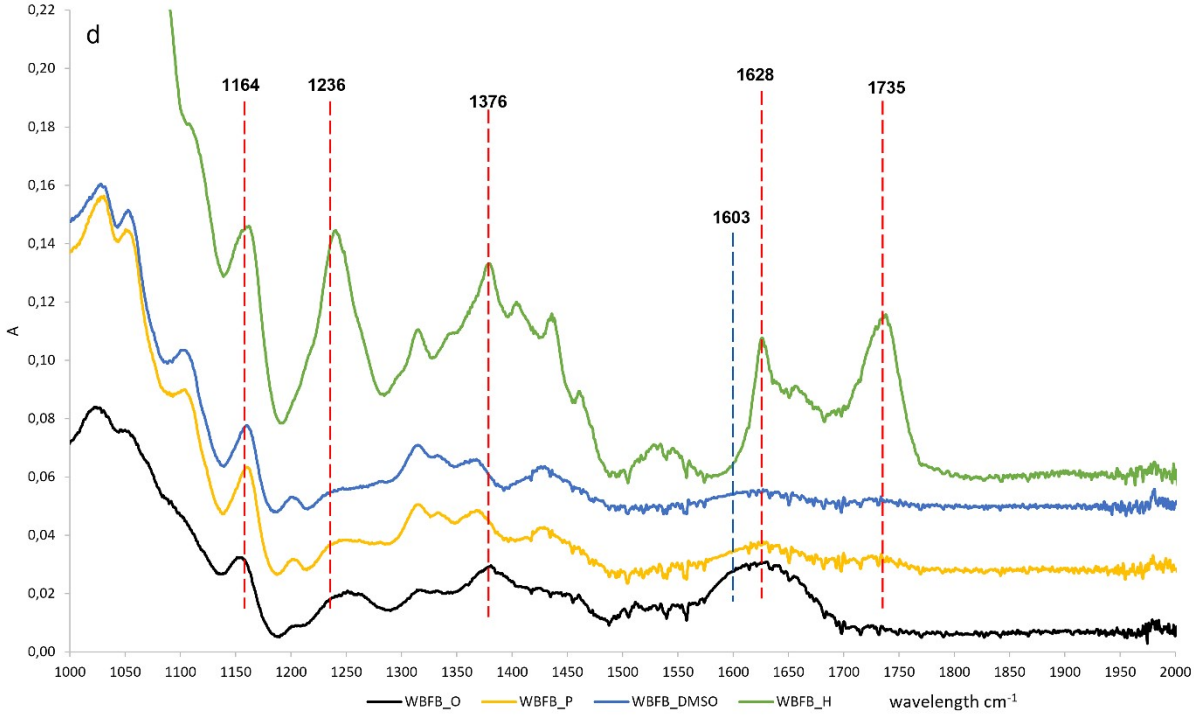
Supplementary Table 6. Assignments of the ^{13}C NMR spectra of willow hemicelluloses (WW; WB; WIB; WBFB; Parenchyma) in comparison to the literatures[9,34,39,38] and authentic starch.

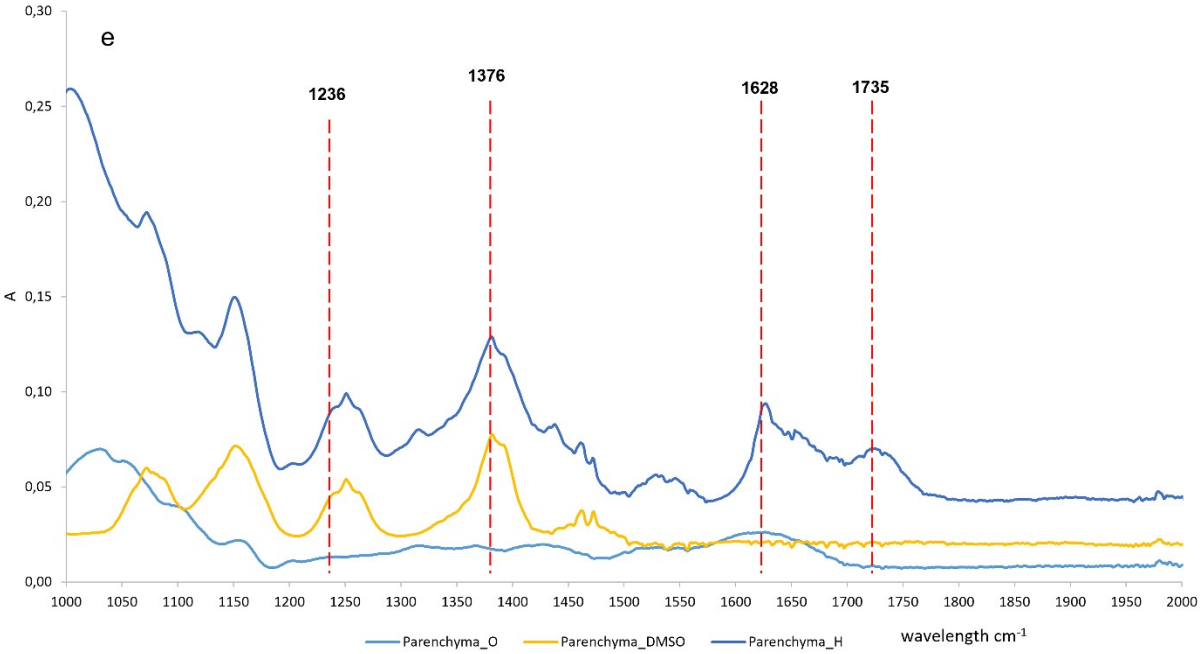
	abbreviation	WW	WB	WIB	WBFB	parenchyma	Yuan et al. 2010[34]	Sun et al. 2014[9]	Nobre et al. 2018[39]	Nishimura et al. 2018[38]	authentic starch
4-O-Me-α-D-Glucuronic acid(C6H6) GlcpA	Xy-MG-6	169.30	169.30	169.30	169.30	–	173.4	177.1	–	–	
(1-4)-β-D-Glcp (C1H1) or glucomannan C1'/H1' (glucose)	GLMA-1'	–	102.99	102.99	102.99	102.99	–	–	–	103.1	
(1-4)-β-D-Xylp (C1H1)	Xy-1	102.00	102.00	102.00	102.00	102.00	101.73	102	–	–	
glucomannan (1) (mannose branch side)	GLMA-1	–	100.40	100.40	100.40	100.40	–	–	–	100.6	
(1-4)-β-D-Xylp (C1'H1')	Xy-1'	101.20	–	–	101.20	101.20	–	–	–	–	
galactomannan (1)	GAMA-1	–	99.29	99.29	99.29	99.29	–	–	100.62	–	
4-O-Me-α-D-Glucuronic acid (C1H1) GlcpA	Xy-MG-1	97.60	–	–	–	–	97.6	97.5	–	–	
4-O-Me-α-D-Glucuronic acid (C4H4) GlcpA	Xy-MG-4	82.00	–	–	–	–	82.49	82.7	–	–	
(1-4)-β-D-Xylp (C2'H2')	Xy-2'	77.00	77.00	77.00	77.00	77.00	–	–	–	–	
(1-4)-β-D-Xylp (4)	Xy-4	75.76	75.76	75.76	75.76	75.76	76.42	76.0	–	–	
(1-4)-β-D-Xylp (3)	Xy-3	75.00	75.00	75.00	75.00	75.00	73.73	74.9	–	–	
galactomannan (5)	GAMA-5		74.20	74.20	74.20	74.20	–	–	74.05	–	
(1-4)-β-D-Xylp (2)	Xy-2	73.20	73.20	73.20	73.20	73.20	72.76	73.3	–	–	
glucomannan (3)	GLMA-3	–	–	–	–	72.83	–	–	–	71.0	
4-O-Me-α-D-Glucuronic acid (C3H3) GlcpA	Xy-MG-3	72.20	72.20	72.20	72.20	–	–	–	–	–	
4-O-Me-α-D-Glucuronic acid (C5H5) GlcpA	Xy-MG-5	72.10	72.10	72.10	72.10	–	72.27	72.2	–	–	
glucomannan (2)	GLMA-2 or GAMA-3	–	–	–	–	71.84	–	–	–	70.5	
4-O-Me-α-D-Glucuronic acid (C2H2) GlcpA	Xy-MG-2	71.60	71.60	71.60	71.60	–	71.33	–	–	–	
(1-4)-β-D-Xylp (5)	Xy-5	63.60	63.60	63.45	63.45	63.60	63.02	63.4	–	–	
branched galactose of GAMA (6)	GAMA-6'	–	–	–	–	60.57	–	–	61.80	–	
4-O-Me-α-D-Glucuronic acid (methoxyl) GlcpA	Xy-MG-OCH3	56.20	56.20	56.17	56.17	56.20	–	–	–	–	
unknown	unknown	40.41	40.41	40.41	40.41	40.41	–	–	–	–	
acetyl CH3	Xy-Acetyl	20.85	20.83	20.83	20.83	–	–	–	–	–	
unknown	unknown	18.54	18.54	18.54	18.54	–	–	–	–	–	
Starch (1)	St-1					100.39					100.40
Starch (2)	St-2					72.15					72.13
Starch (3)	St-3					73.45					73.45
Starch (4)	St-4					79.15					79.16
Starch (5)	St-5					71.76					71.80
Starch (6)	St-6					60.63					60.62


a
1154
1236
1376
1500
1594
1735
A
wavelength cm-1
WW_O
WW_P
WW_DMSO
WW_H


b
1156
1236
1376
1500
1603
1628
1735
A
wavelength cm-1
WB_O
WB_P
WB_DMSO
WB_H


c
1156
1236
1376
1603
1628
1735
A
WIB_O
WIB_P
WIB_DMSO
WIB_H
wavelength cm-1

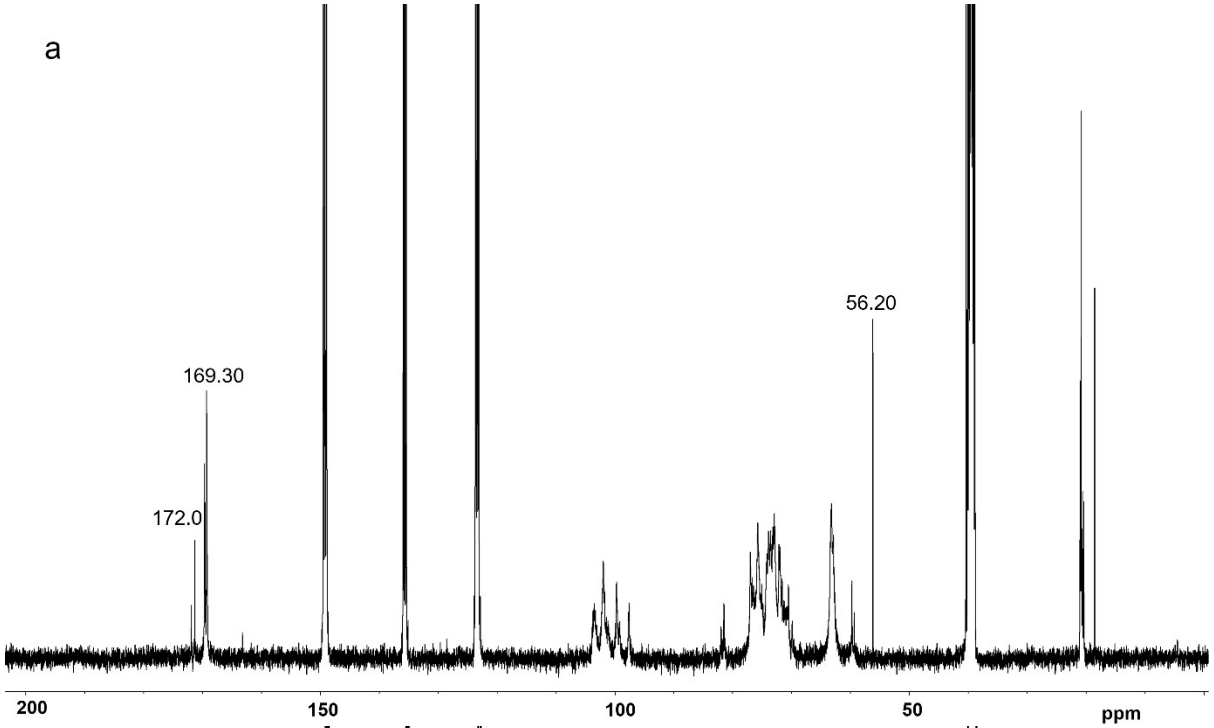

d
1164
1236
1376
1603
1628
1735
A
WBFB_O
WBFB_P
WBFB_DMSO
WBFB_H
wavelength cm-1

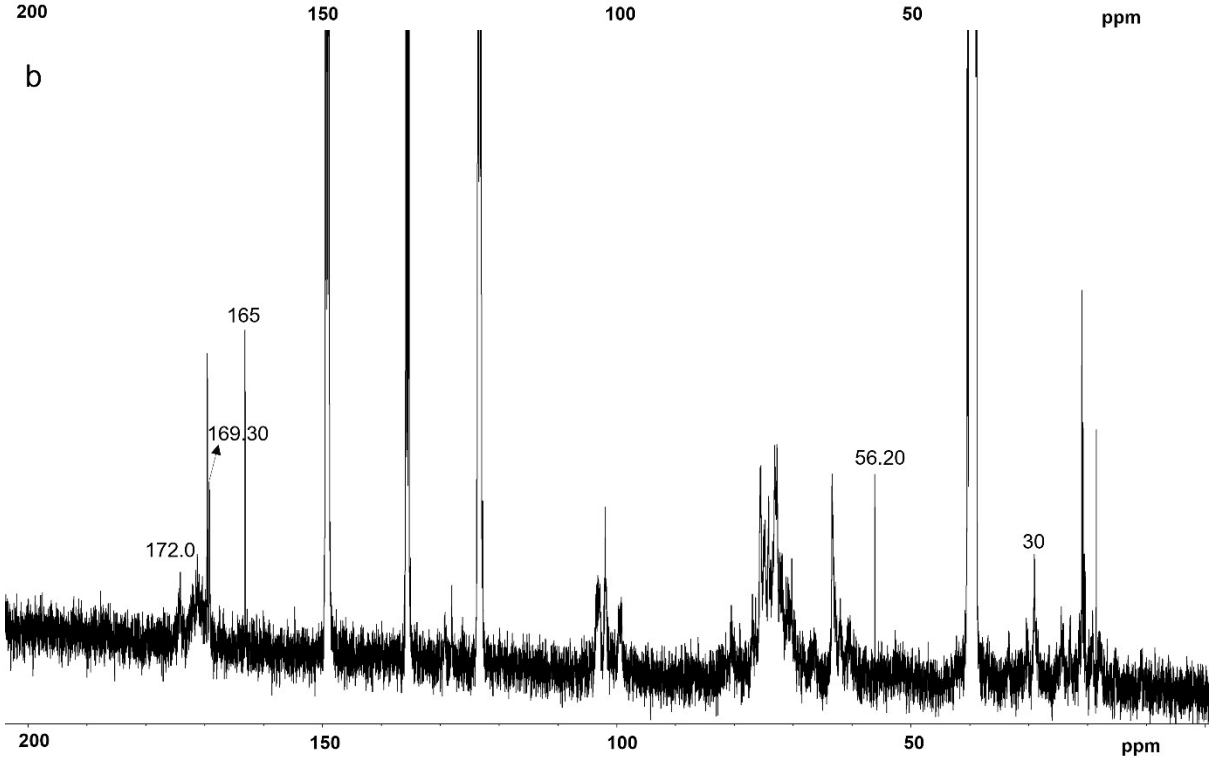


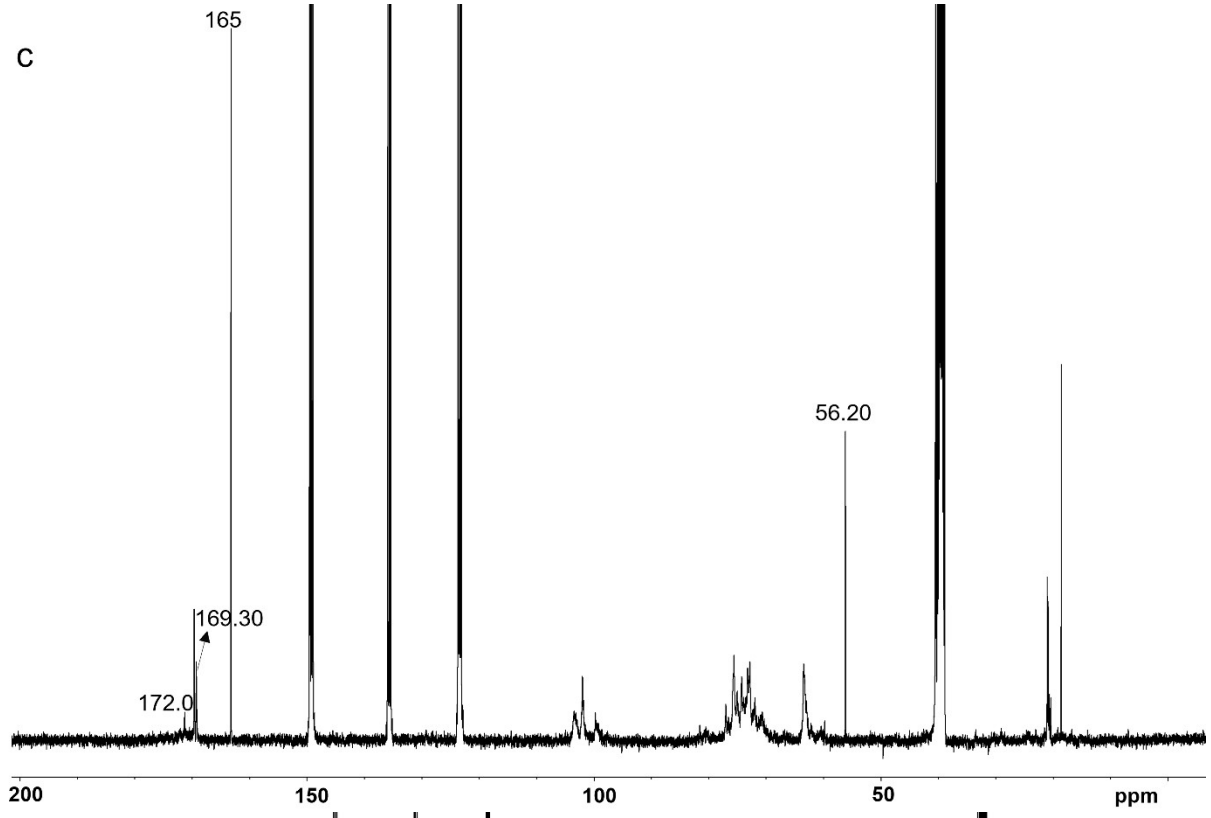
Supplementary Fig. 14 FT-IR spectrum of five different samples. a Willow wood. **b** Willow bark. **c** Willow inner bark. **d** Willow bark fibre bundle. **e** Parenchyma tissues that are collected from their original form (O), solid residues after peracetic acid delignification (P), solid residues after the DMSO extraction (DMSO), and the recovered hemicellulose (H). The assignment of FT-IR spectrum is summarized at **Supplementary Table 7**.

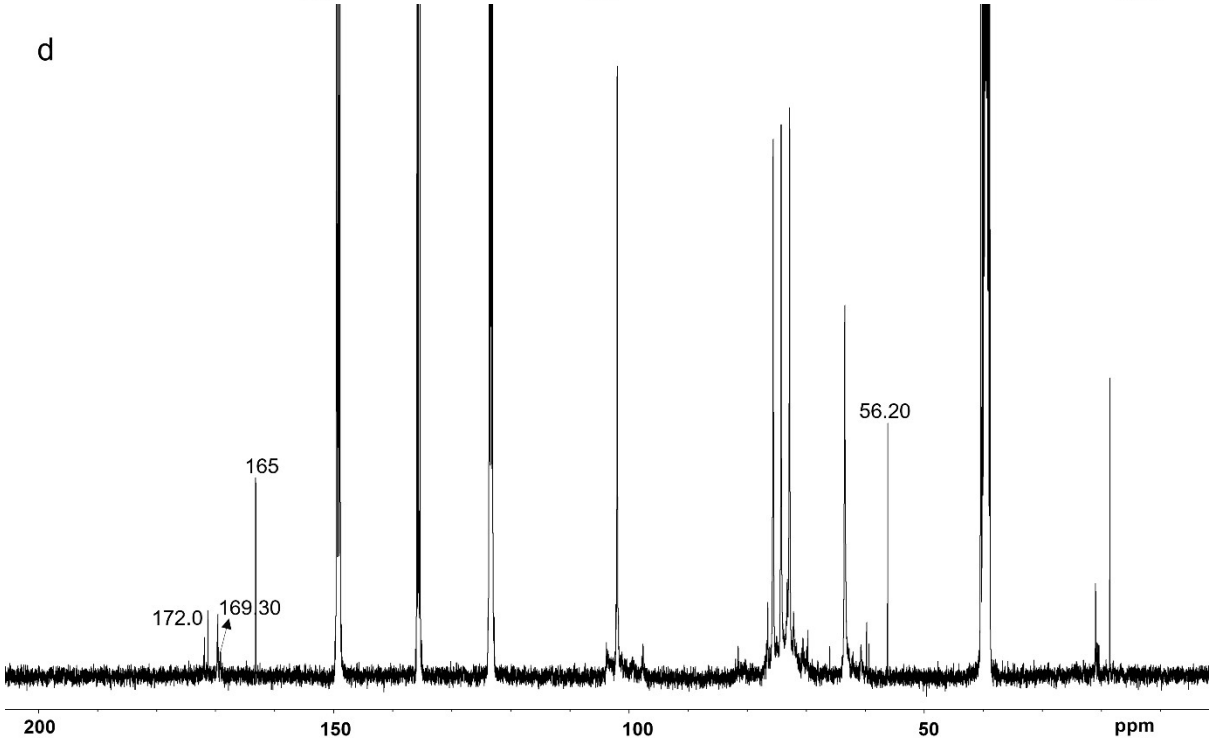
Supplementary Table 7. The main functional group assignment of recovered hemicellulose fractions from willow wood and bark in FT-IR spectra based on the literature[12].

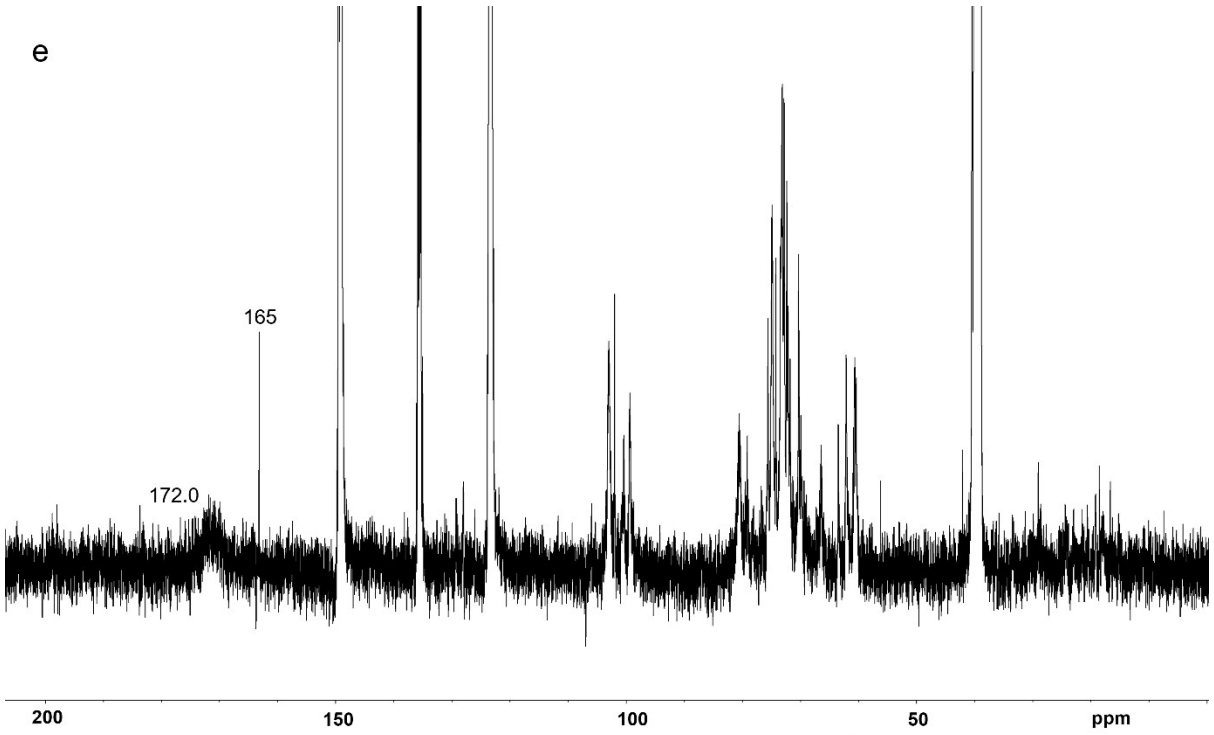
wave numbers (cm^{-1})	functional group	compounds
1154; 1156; 1164	-C-O-C fragment	arabinosyl side branches, pyranose ring skeletal (hemicellulose)
1236	- COOH vibration stretching of glucuronic acid (Xy-MG)	glucuronic acid (hemicellulose)
1376	C-CH3 stretching; C–H vibration of polysaccharides	cellulose
1500	lignin aromatics	lignin
1594 or 1603	lignin aromatics	lignin
1628	C=O stretching carboxylation or from the contribution of protein amide group	protein
1735	C=O stretching of hemicellulose	hemicellulose

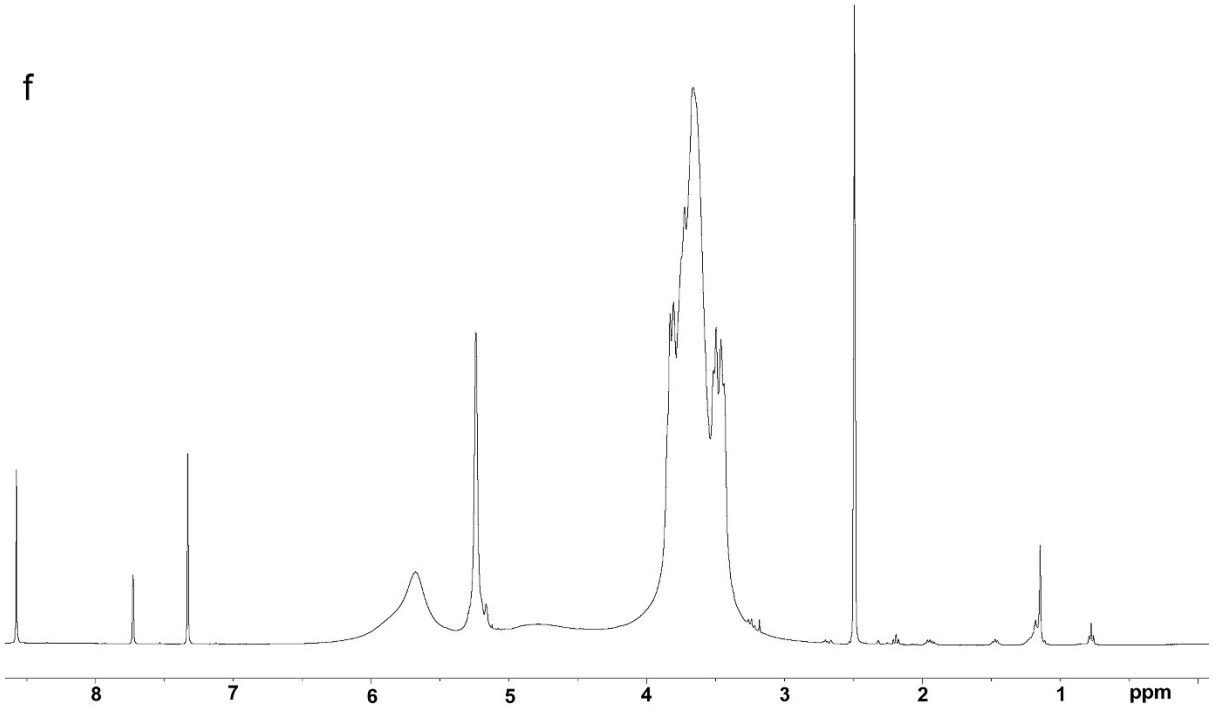

a
56.20
169.30
172.0
200
150
100
50
ppm

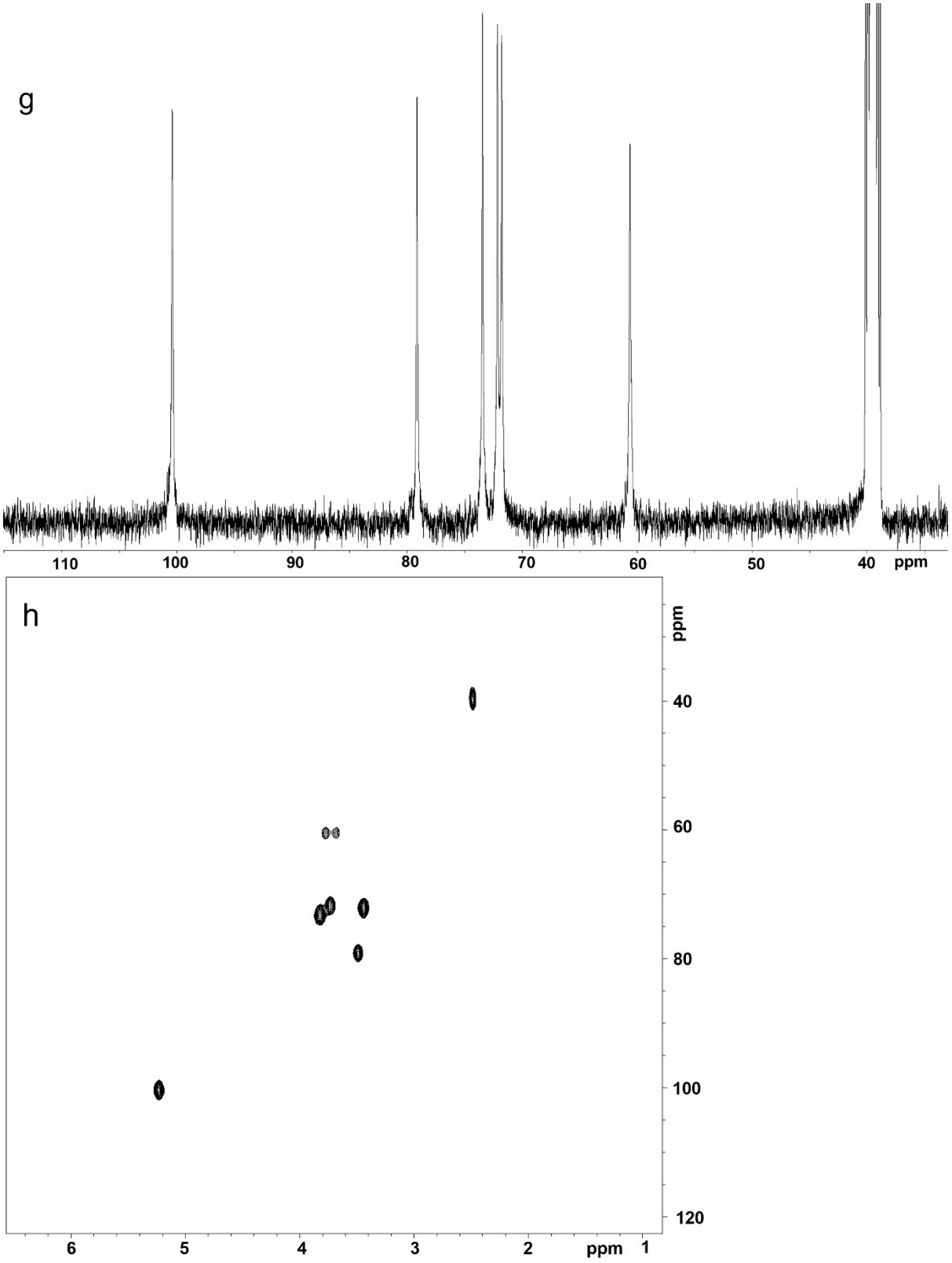

b
165
169.30
56.20
172.0
30
200
150
100
50
ppm


c
165
56.20
169.30
172.0
200
150
100
50
ppm


d
56.20
165
172.0
169.30
200
150
100
50
ppm


e
165
172.0
200
150
100
50
ppm

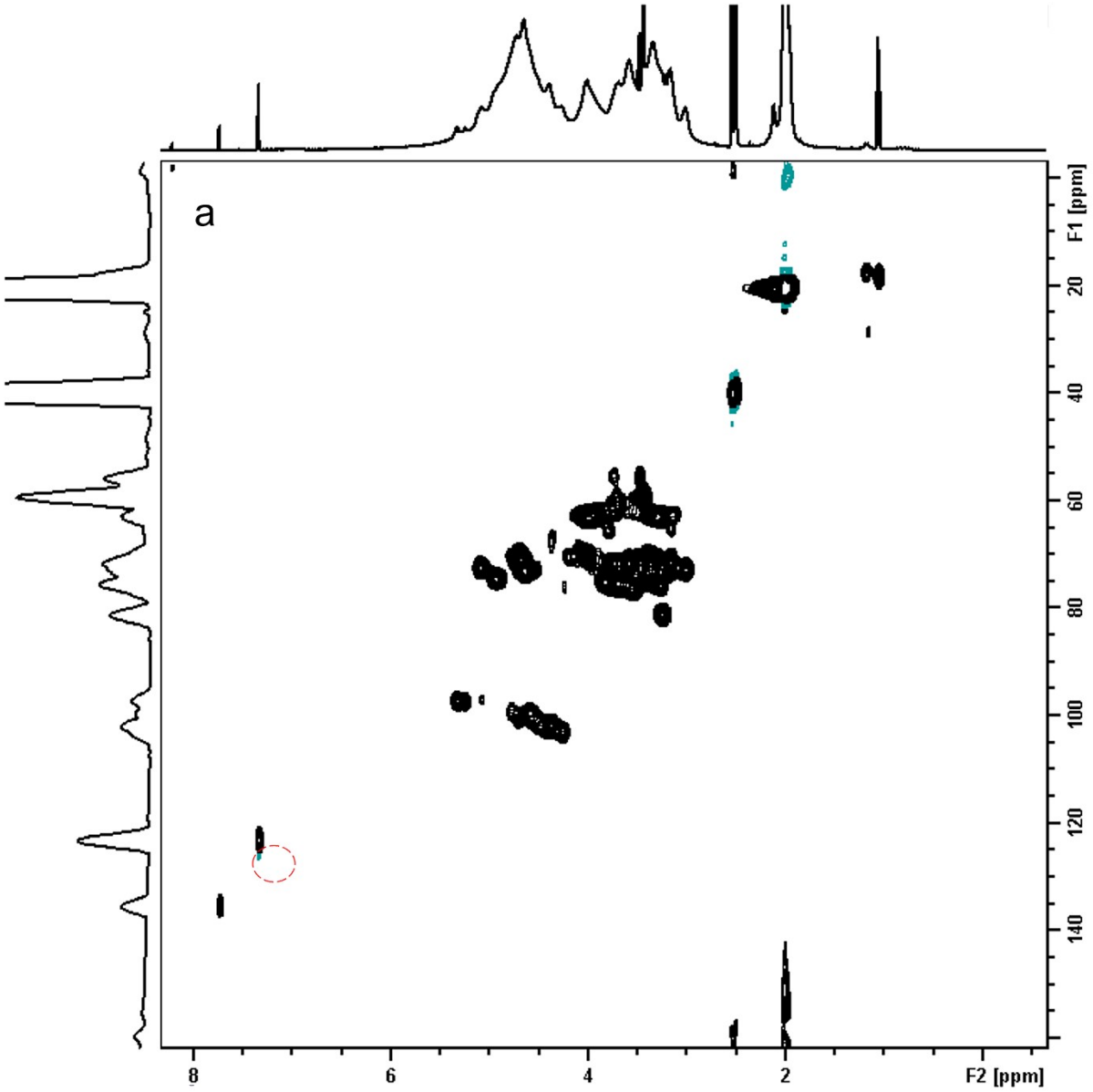

f
8
7
6
5
4
3
2
1
ppm

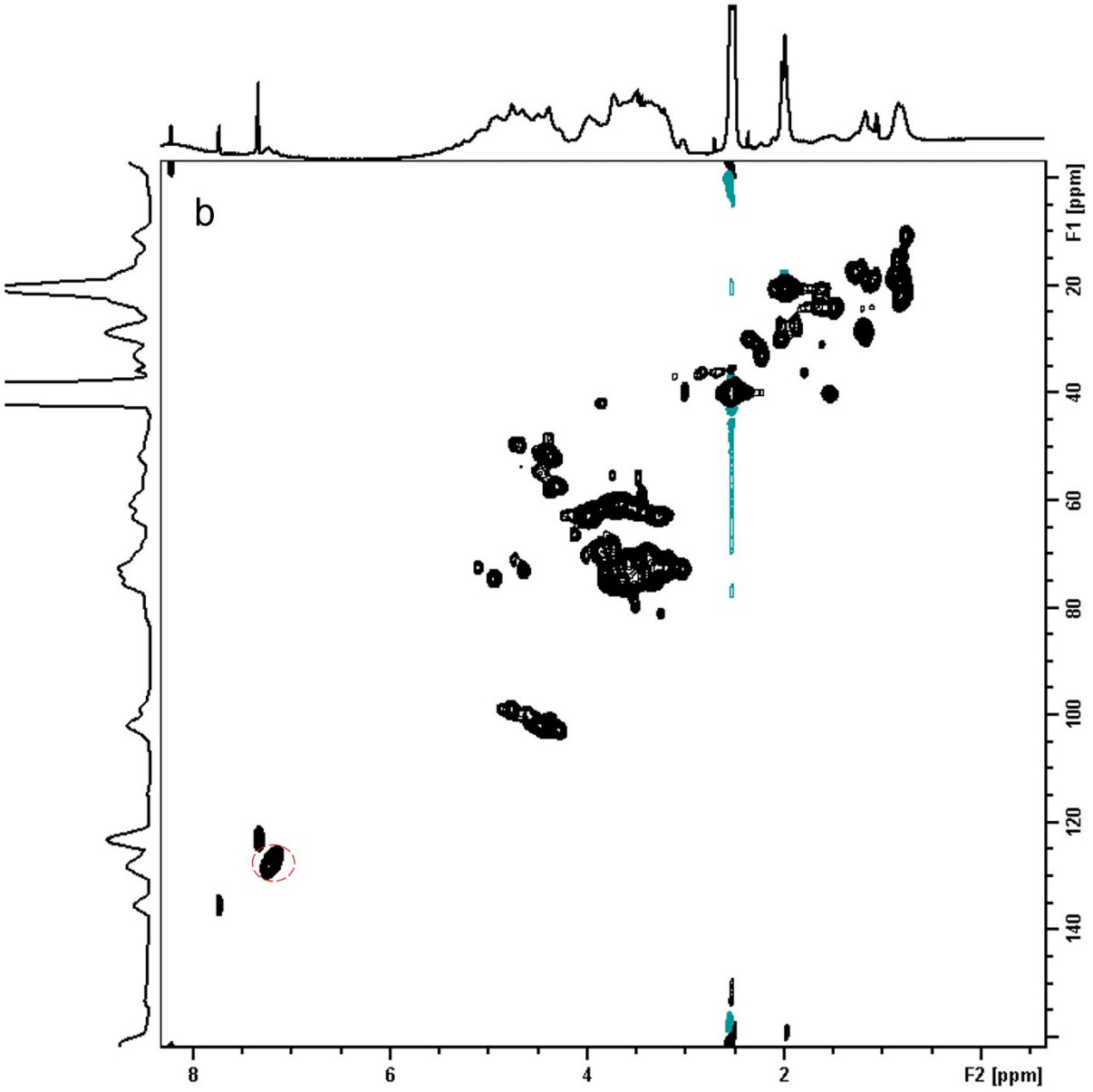


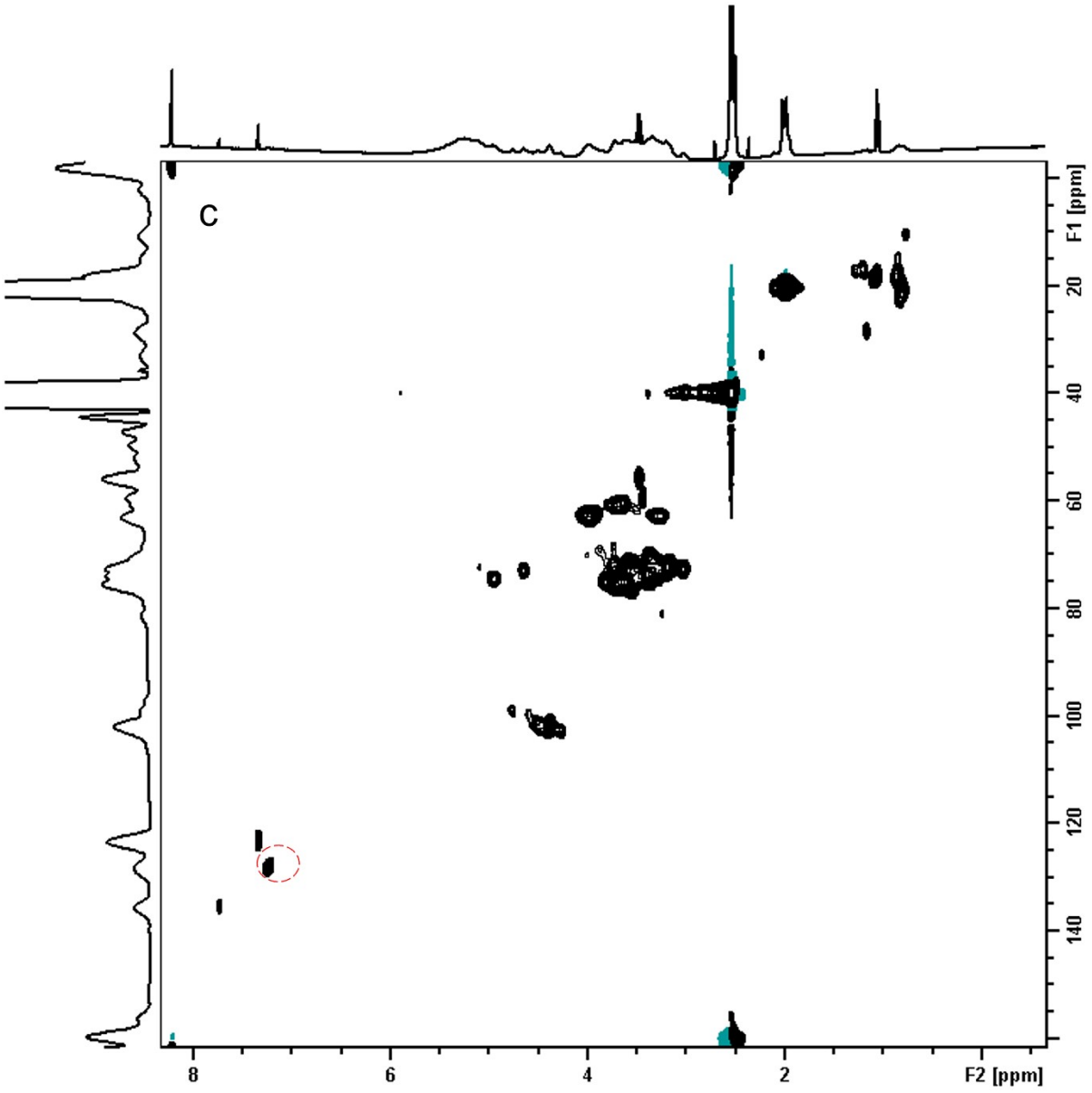
Supplementary Fig. 15 NMR spectrum of the purified hemicellulose (H) and authentic starch in the solvent of DMSO-*d6*/ pyridine-*d5* (v/v, 4/1). a ^{13}C of Willow wood (integral ratio of Xy-MG-OCH3/ Xy-MG-6 = 1/2.8323). **b** ^{13}C of willow bark (integral ratio of Xy-MG -OCH3/ Xy-MG-6 = 1/9.0533). **c** ^{13}C of willow inner bark (integral ratio of Xy-MG -OCH3/ Xy-MG-6 = 1/0.6501). **d** ^{13}C of willow bark fiber bundle (integral ratio of Xy-MG -OCH3/ Xy-MG-6 = 1/0.3737). **e** ^{13}C of parenchyma tissues. **f** ^{1}H NMR of authentic starch. **g** ^{13}C NMR of authentic starch. **h** HSQC NMR of authentic starch.

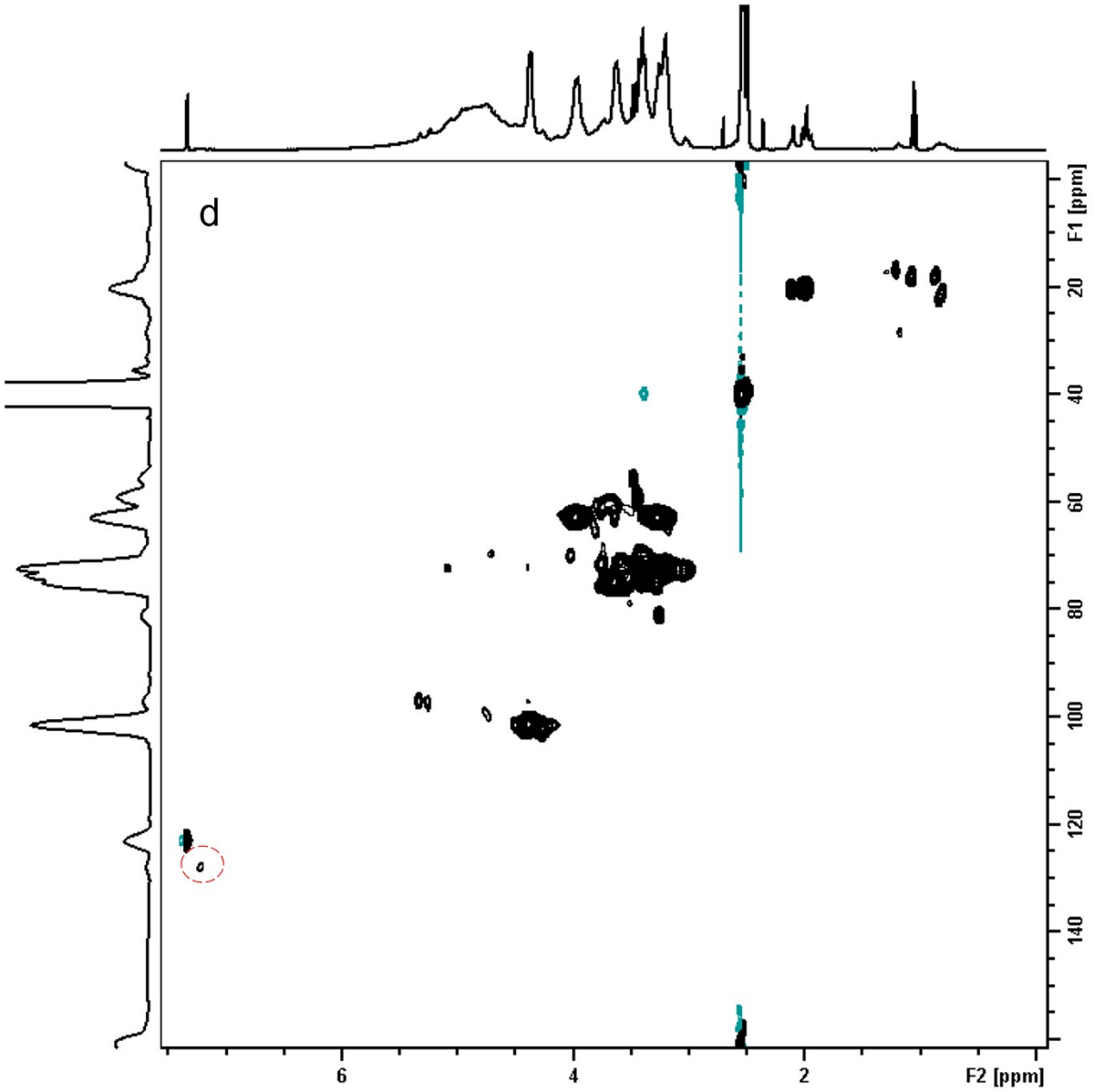
Supplementary Table 8. Assignments of the $^{1}H–^{13}C$ correlation peaks in the 2D-HSQC spectra of willow hemicelluloses (WW; WB; WIB; WBFB; and Parenchyma) in comparison to the literatures[9,12,34,39,38] and authentic starch.

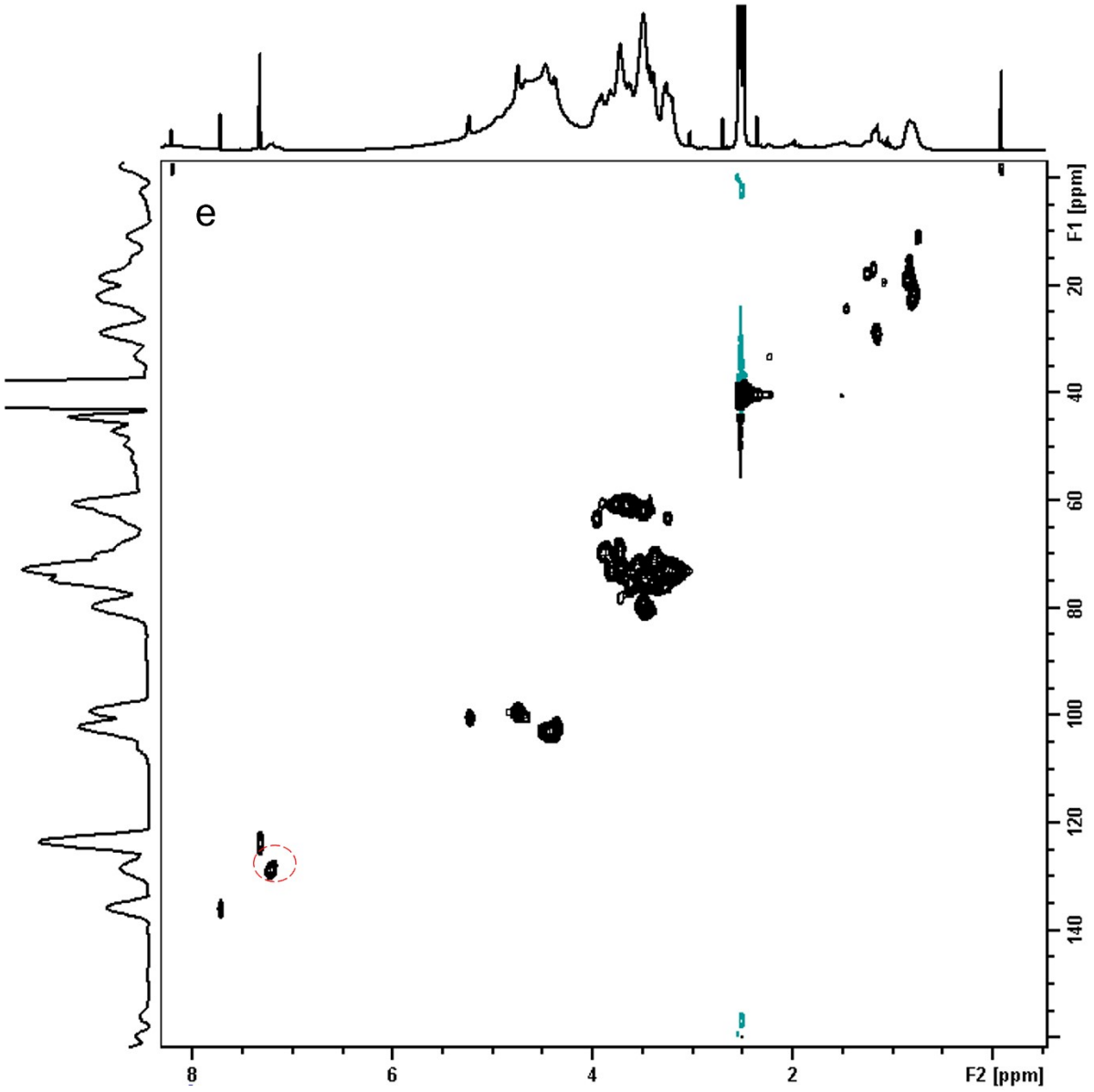
full name	abbreviations	WW (solvent of DMSO-*d6*/ pyridine-*d5*)	WB (DMSO-*d6*/ pyridine-*d5*)	WIB (DMSO-*d6*/ pyridine-*d5*)	WBFB (DMSO-*d6*/ pyridine-*d5*)	parenchyma (DMSO-*d6*/ pyridine-*d5*)	Nishimura et al. 2018[38] (DMSO-*d6*)	Nobre et al. 2018[39] (D2O)	Ding et al. 2019[12] (D2O)	Yuan et al. 2010[34] (D2O)	Sun et al. 2014[9] (D2O)
acetyl CH3	Xy-Acetyl	20.5/1.99	20.5/1.99	20.5/1.99	20.5/1.99	20.5/1.99	—	—	22.7/1.9	—	—
(1-4)-β-D-Xylp (5b)	Xy-5b	63.0/3.98	63.0/3.98	63.0/3.98	63.0/3.98	63.0/3.98	—	—	62.3/4.2	63.02/4.04	63.1/3.98
(1-4)-β-D-Xylp (5a)	Xy-5a	62.9/3.29	62.9/3.29	62.9/3.29	62.9/3.29	62.9/3.29	—	—	62.5/3.4	63.02/3.32	63.1/3.28
(1-4)-β-D-Xylp (2)	Xy-2	72.4/3.16	72.4/3.16	72.4/3.16	72.4/3.16	72.4/3.16	—	—	72.2/3.2	72.76/3.23	73.0/3.19
(1-4)-β-D-Xylp (3)	Xy-3	71.8/3.36	71.8/3.36	71.8/3.36	71.8/3.36	71.8/3.36	—	—	75.0/3.4	73.73/3.50	74.5/3.41
(1-4)-β-D-Xylp (4)	Xy-4	75.5/3.68	75.5/3.68	75.5/3.68	75.5/3.68	75.5/3.68	—	—	75.7/3.8	76.42/3.72	76.0/3.67
(1-4)-β-D-Xylp (C1H1)	Xy-1	101.9/4.37	101.9/4.37	101.9/4.37	101.9/4.37	101.9/4.37	—	—	101.1/4.4	101.73/4.41	—
2-O-acetyl-β-D-Xylp (2)	Xy-2-O-Ac-2	73.1/4.63	73.1/4.63	73.1/4.63	—	—	—	—	73.6/4.6	—	—
2-O-acetyl-β-D-Xylp (1)	Xy-2-O-Ac-1	99.6/4.57	99.6/4.57	99.6/4.57	—	—	—	—	99.4/4.6	—	—
3-O-acetyl-β-D-Xylp (1)	Xy-3-O-Ac-1	101.9/4.39	101.9/4.39	101.9/4.39	—	—	—	—	101.9/4.3	—	—
3-O-acetyl-β-D-Xylp (3)	Xy-3-O-Ac-3	74.66/4.92	74.66/4.92	74.66/4.92	—	—	—	—	—	—	—
2.3-di-O-Ac-b-D-Xylp	Xy-2,3-di-O-Ac-1	—	—	—	—	99.09/4.75	—	—	—	—	—
4-O-Me-α-D-Glucuronic acid(C1H1) GlcpA	Xy-MG-1	97.2/5.31	—	—	—	—	—	—	97.5/5.3	97.60/5.22	97.4/5.19
4-O-Me-α-D-Glucuronic acid(C2H2) GlcpA	Xy-MG-2	71.72/3.57	—	—	—	—	—	—	—	71.33/3.52	71.5/3.46
4-O-Me-α-D-Glucuronic acid(C3H3) GlcpA	Xy-MG-3	71.81/3.74	—	—	—	—	—	—	—	72.18/3.70	72.2/3.60
4-O-Me-α-D-Glucuronic acid(C4H4) GlcpA	Xy-MG-4	81.3/3.23	—	—	—	—	—	—	81.7/3.23	82.49/3.16	82.4/3.11
4-O-Me-α-D-Glucuronic acid(C5H5) GlcpA	Xy-MG-5	70.18/4.01	—	—	—	—	—	—	—	72.27/4.25	72.1/4.22
4-O-Me-α-D-Glucuronic acid (methoxyl) GlcpA	Xy-MG-OCH3	55.8/3.46	55.8/3.46	55.8/3.46	55.8/3.46	—	—	—	55.6/3.8	—	—
galactomannan (1)	GAMA-1	—	98.9/4.76	98.9/4.76	98.9/4.76	—	—	100.62/4.74	—	—	—
galactomannan (2)	GAMA-2	—	70.2/4.02	70.2/4.02	70.2/4.02	—	—	70.59/4.12	—	—	—
galactomannan (3)	GAMA-3	—	71.73/3.77	71.73/3.77	71.73/3.77	—	—	71.83/3.79	—	—	—
galactomannan (4)	GAMA-4	—	76.06/3.78	76.06/3.78	76.06/3.78	—	—	77.36/3.86	—	—	—
galactomannan (5)	GAMA-5	—	74.02/3.65	74.02/3.65	74.02/3.65	—	—	74.05/3.64	—	—	—
galactomannan (6)	GAMA-6	—	66.6/4.12	66.6/4.12	66.6/4.12	—	—	67.22/3.96	—	—	—
branched galactose of GAMA (1)	GAMA-1'	—	97.3/5.34	97.3/5.34	97.3/5.34	97.3/5.34	—	99.48/5.02	—	—	—
branched galactose of GAMA (2)	GAMA-2'	—	67.7/3.74	67.7/3.74	67.7/3.74	67.7/3.74	—	68.55/3.80	—	—	—
branched galactose of GAMA (3)	GAMA-3'	—	71.1/3.84	71.1/3.84	71.1/3.84	71.1/3.84	—	70.10/3.92	—	—	—
branched galactose of GAMA (4)	GAMA-4'	—	70.9/4.02	70.9/4.02	70.9/4.02	70.9/4.02	—	69.96/4.00	—	—	—
branched galactose of GAMA (5)	GAMA-5'	—	71.92/3.86	71.92/3.86	71.92/3.86	71.92/3.86	—	71.42/3.89	—	—	—
branched galactose of GAMA (6)	GAMA-6'	—	60.88/3.67	60.88/3.67	60.88/3.67	60.88/3.67	—	61.80/3.76	—	—	—
glucomannan C1'/H1' (glucose)	GLMA-1'	103.1/4.25	103.1/4.25	103.1/4.25	103.1/4.25	102.6/4.45	100.5/4.54	—	—	—	102.1/4.34
glucomannan (1) (mannose branch side)	GLMA-1	—	101.2/4.49	101.2/4.49	101.2/4.49	101.2/4.49	100.6/4.46	—	—	—	—
glucomannan (2)	GLMA-2	—	71.82/3.59	71.82/3.59	71.82/3.59	71.82/3.59	70.5/3.65	—	—	—	—
glucomannan (3)	GLMA-3	—	72.82/3.72	72.82/3.72	72.82/3.72	72.82/3.72	71.0/3.62	—	—	—	—
glucomannan (4)	GLMA-4	—	76.97/3.76	76.97/3.76	76.97/3.76	76.97/3.76	78.0/3.76	—	—	—	—
glucomannan (5)	GLMA-5	—	69.3/3.87	69.3/3.87	69.3/3.87	69.3/3.87	69.0/3.64	—	—	—	—
glucomannan (6)	GLMA-6	—	68.7/3.73	68.7/3.73	68.7/3.73	68.7/3.73	68.1/3.48	—	—	—	—
Starch (1)	St-1					100.08/ 5.228					
Starch (2)	St-2					72.05/3.44					
Starch (3)	St-3					73.07/3.82					
Starch (4)	St-4					79.03/3.49					
Starch (5)	St-5					71.69/3.73					
Starch (6a)	St-6a					60.44/3.77					
Starch (6b)	St-6b					60.33/3.68					


a
F1 [ppm]
20
40
60
80
100
120
140
8
6
4
2
F2 [ppm]

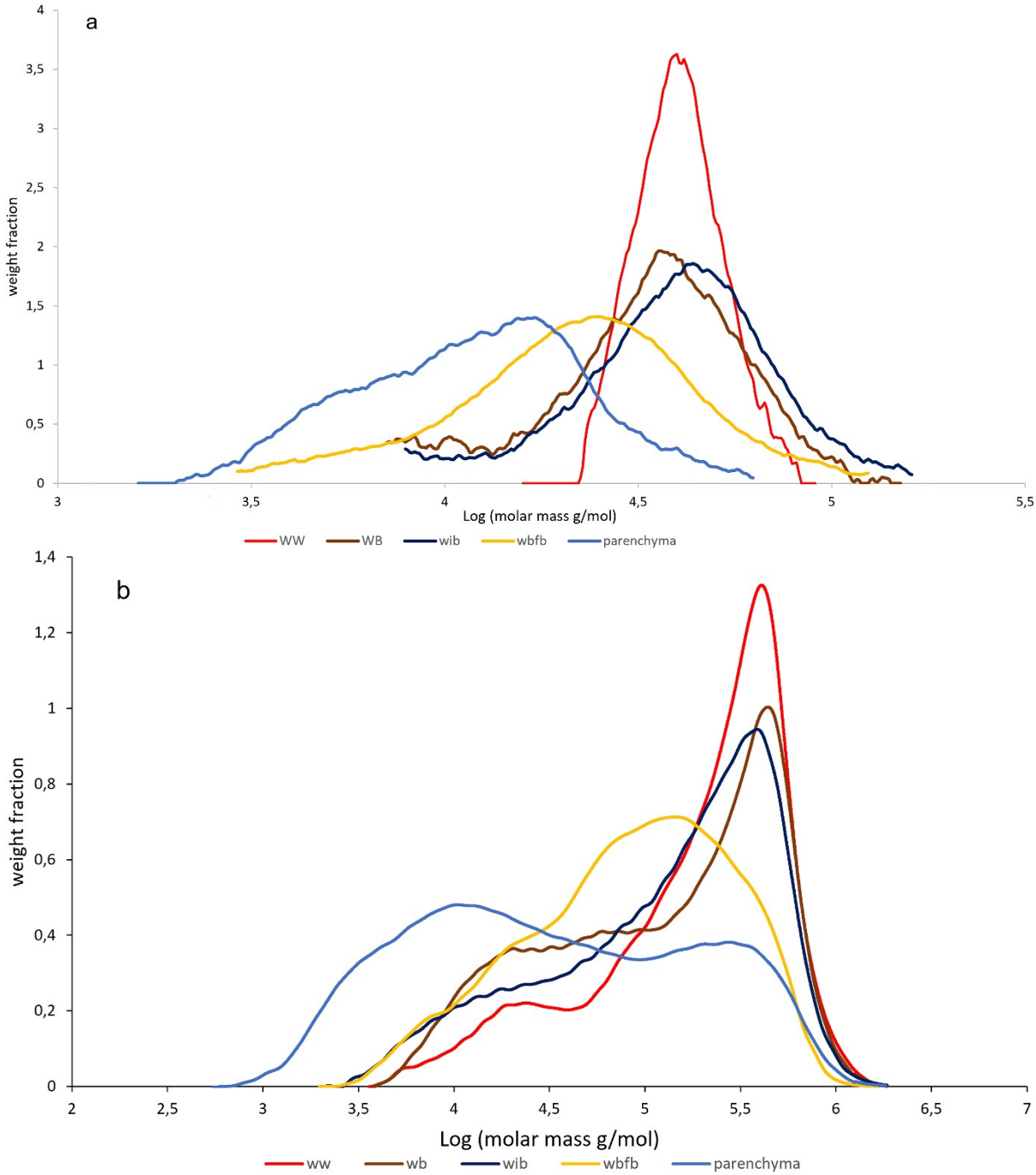

b
F1 [ppm]
20
40
60
80
100
120
140
8
6
4
2
F2 [ppm]


C
F1 [ppm]
20
40
60
80
100
120
140
8
6
4
2
F2 [ppm]


d
F1 [ppm]
20
40
60
80
100
120
140
6
4
2
F2 [ppm]



Supplementary Fig. 16 2D heteronuclear single quantum coherence (HSQC) NMR spectrum (δC/δH, 0–160/0–8.3 ppm) of the purified hemicellulose. **a** Willow wood. **b** Willow bark. **c** Willow inner bark. **d** Willow bark fiber bundle. **e** Parenchyma tissues. The phenylalanine protein (δC/δH, 129/7.20 ppm)[5] was identified and highlighted in red circle.



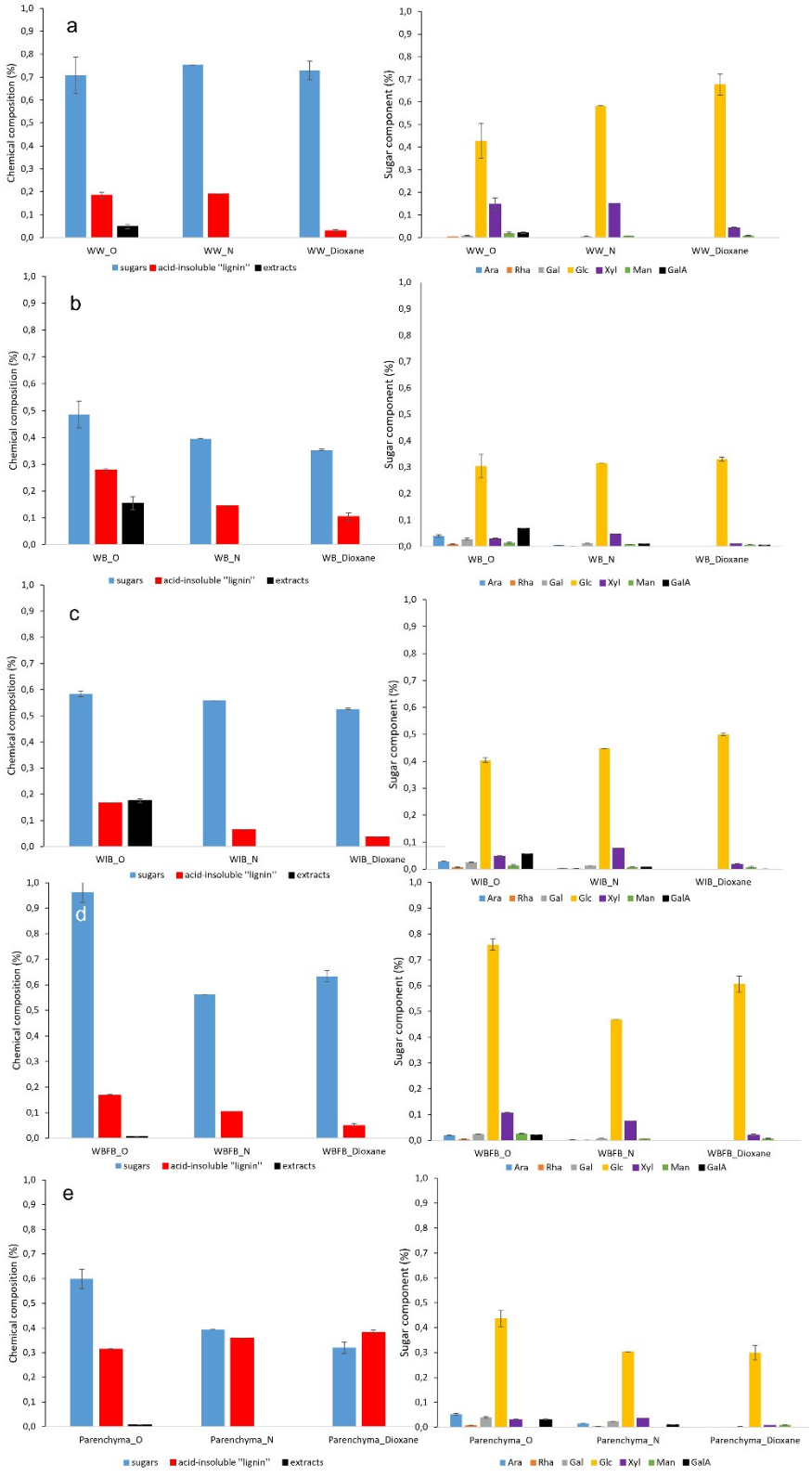
Supplementary Fig. 17 SEC molecular weight (Mw) spectrum distributions of willow samples (WW; WB; WIB; WBFB; and Parenchyma tissues) from the fractionation scheme supplementary Fig. 6. a Hemicellulose. **b** Purified ‘‘cellulose’’ (extracted as a side fraction from the DMSO purification of hemicellulose). For their real molecular weight (MW) values, see **Supplementary Table 9**.

Supplementary Table 9. **SEC molecular weight (Mw) spectrum distributions that is measured by gel-permeation chromatography (GPC). a** Purified hemicellulose. **b** Purified cellulose.

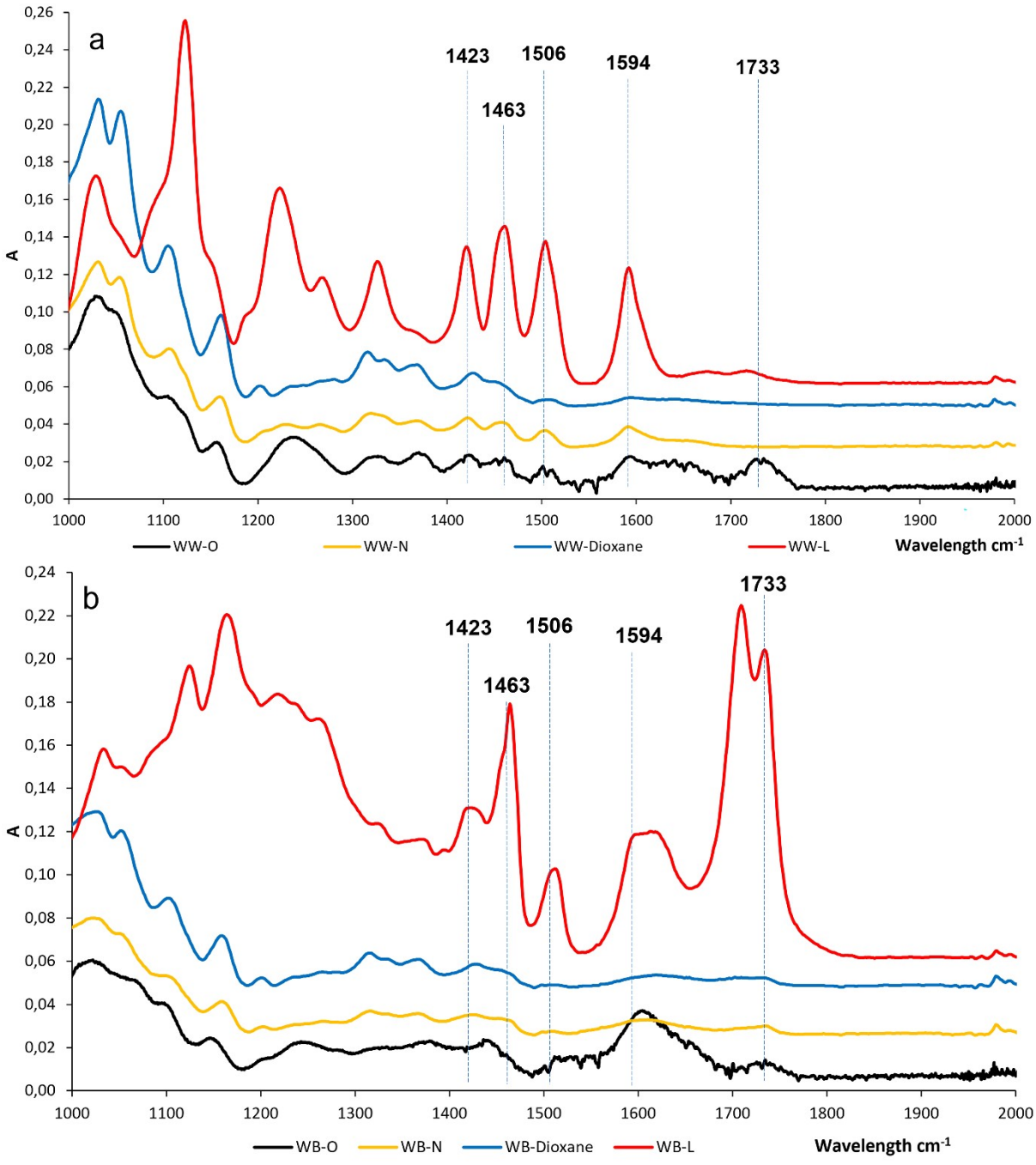
	hemicellulose					cellulose				
	Mn (kDa)	Mw (kDa)	Mz (kDa)	Mp (kDa)	PD	Mn (kDa)	Mw (kDa)	Mz (kDa)	Mp (kDa)	PD
WW	39.0	41.6	44.6	39.8	1.1	69.1	280.0	455.3	348.2	4.1
WB	27.4	38.5	49.1	35.7	1.4	47.4	238.8	446.6	402.3	5.0
WIB	32.9	45.8	59.9	43.8	1.4	48.9	231.9	428.1	335.1	4.8
WBFB	16.2	27.1	40.8	24.7	1.7	35.4	151.5	325.0	152.5	4.3
Parenchyma	9.3	14.4	21.1	17.0	1.6	36.4	262.8	400.8	35.4	8.9

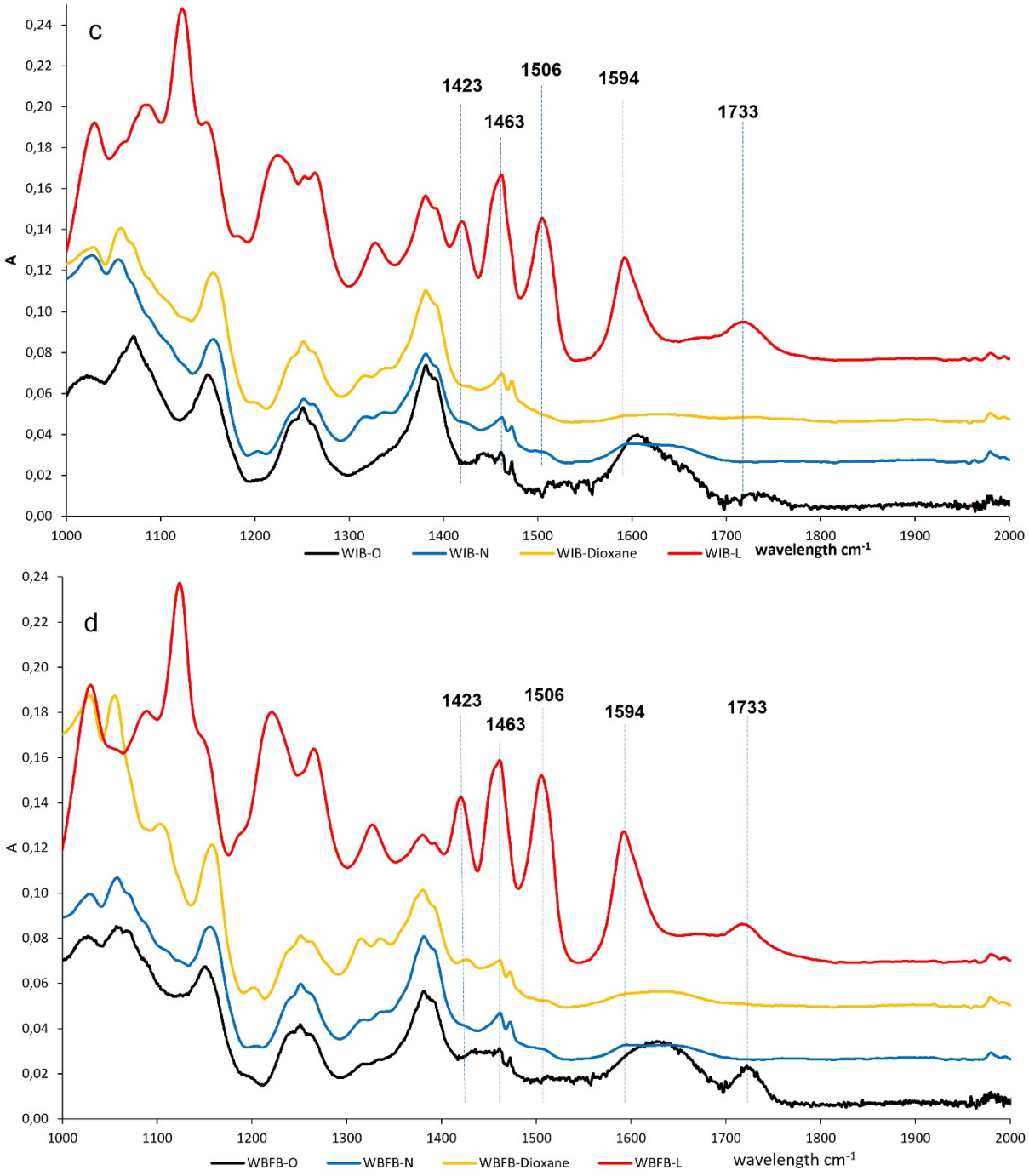
Supplementary Table 10. Determination of the ‘‘Klason Lignin’’ (100 %) from the 1) original biomass (as received); 2) ‘‘all-pretreatment’’ refers to the solid residue after extraction treatment (DCM/ acetone/ water), pepsin treatment, and 0.1M NaOH treatment; 3) refers to the solid residue after ‘‘all-pretreatment’’ plus the dioxane-water extraction treatment. Calculation of ‘klason lignin’ removal (100%) = (Klason lignin (2)-Klason lignin (3))/ Klason lignin (2). Standard deviations are included in the parenthesis.

	1) original	2) after ‘‘all-pretreatment’’	3) after ‘‘dioxane-extraction’’	‘‘klason lignin’’ removal
WW	0.19 (0.01)	0.19	0.03 (0.004)	0.84
WB	0.28 (0.01)	0.16	0.12 (0.014)	0.22
WIB	0.17 (0.001)	0.07	0.04 (0.001)	0.42
WBFB	0.17 (0.01)	0.10	0.05 (0.008)	0.52



Supplementary Fig. 18 Overall chemical composition (% of the dry mass) (left) and carbohydrate composition (% of anhydro sugars in the monosaccharide from the dry mass) (right) from five different samples. **a** Willow wood. **b** Willow bark. **c** Willow inner bark. **d** Willow bark fiber bundle. **e** Parenchyma tissues that are collected from their original form (O), 0.1M NaOH treated solid residue (N), and solid residues after the dioxane extraction (Dioxane) (**Supplementary Fig. 6)**. For abbreviations, see **Supplementary Fig. 1–5**. Standard deviations are included as the error bars.


a
1423
1506
1594
1733
1463
A
Wavelength cm-1
WW-O
WW-N
WW-Dioxane
WW-L
b
1423
1506
1594
1733
1463
A
Wavelength cm-1
WB-O
WB-N
WB-Dioxane
WB-L



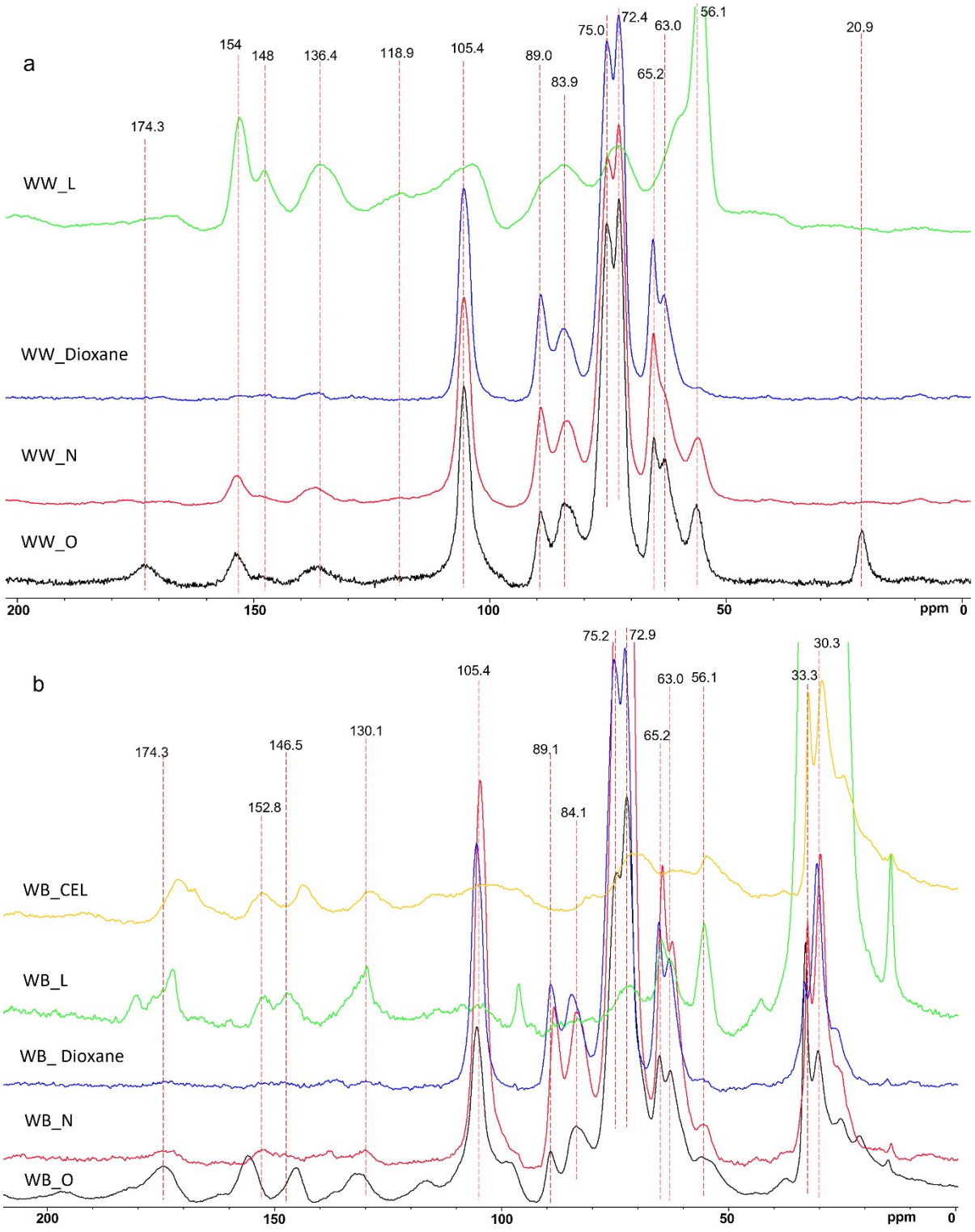
Supplementary Fig. 19 FT-IR spectrum of four different samples that is collected from their original form (O), 0.1M NaOH treated solid residue (N), solid residues after the dioxane/ water extraction (Dioxane), and the recovered dioxane lignin (L) (**Supplementary Fig. 6**). **a** Wood. **b** Bark. **c** Inner bark. **d** Fiber bundle. For abbreviations, see **Supplementary Fig. 1–5**.

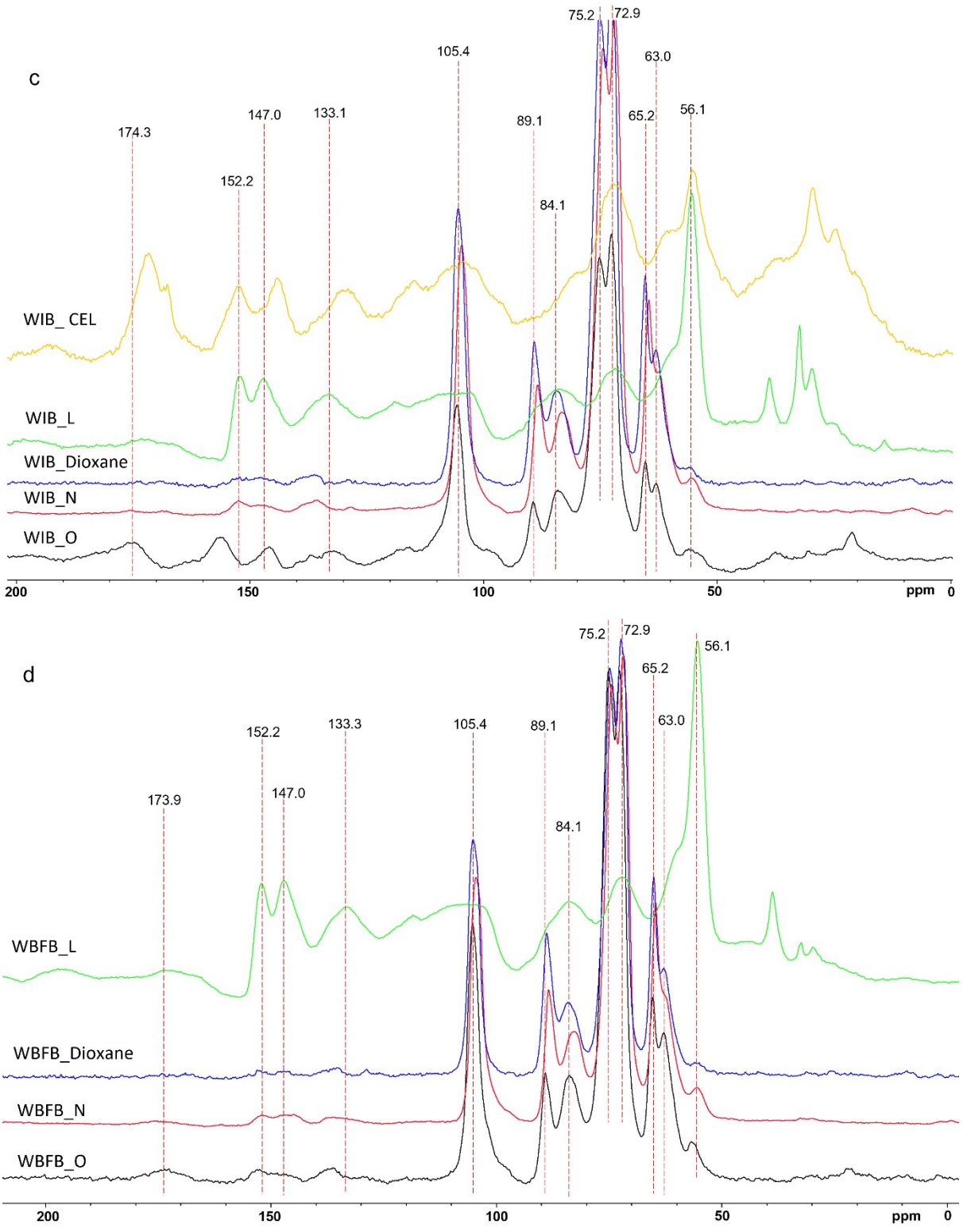
Supplementary Table 11. Assignments of the dioxane lignin in the CP/MAS NMR spectra from willow wood (WW) and bark (WB; WIB; WBFB; and Parenchyma).

^{13}C shift [ppm]	assignment
20.9	C4-C8 of the interflavonoid linkages of tannin or some aliphatic fats (aliphatic chains); extractive fats
30.3	aliphatic methylenes linked to carboxylic moieties in suberin
33.2	—OCH methylene adjacent to ester groups in suberin
56.1	M—COOCH3 in pectin
62.6 and 65.1	C6 carbon of non-crystalline and crystalline cellulose
72.40	C2,3,5 of cellulose
75.0	C2,3,5 of cellulose and hemicellulose
83.9	C4 carbon of non-crystalline cellulose
89.0	C4 carbon of crystalline cellulose
105.4	C1 carbon of cellulose
118.9	C2,6 aromatic carbons of syringyl in lignin
136.4	C1 aromatic carbons of syringyl in lignin
148	C4 aromatic carbons of syringyl in lignin
154	C3 and C5 aromatic carbons of syringyl in lignin
174.3	carboxyl groups of hemicelluloses or suberin

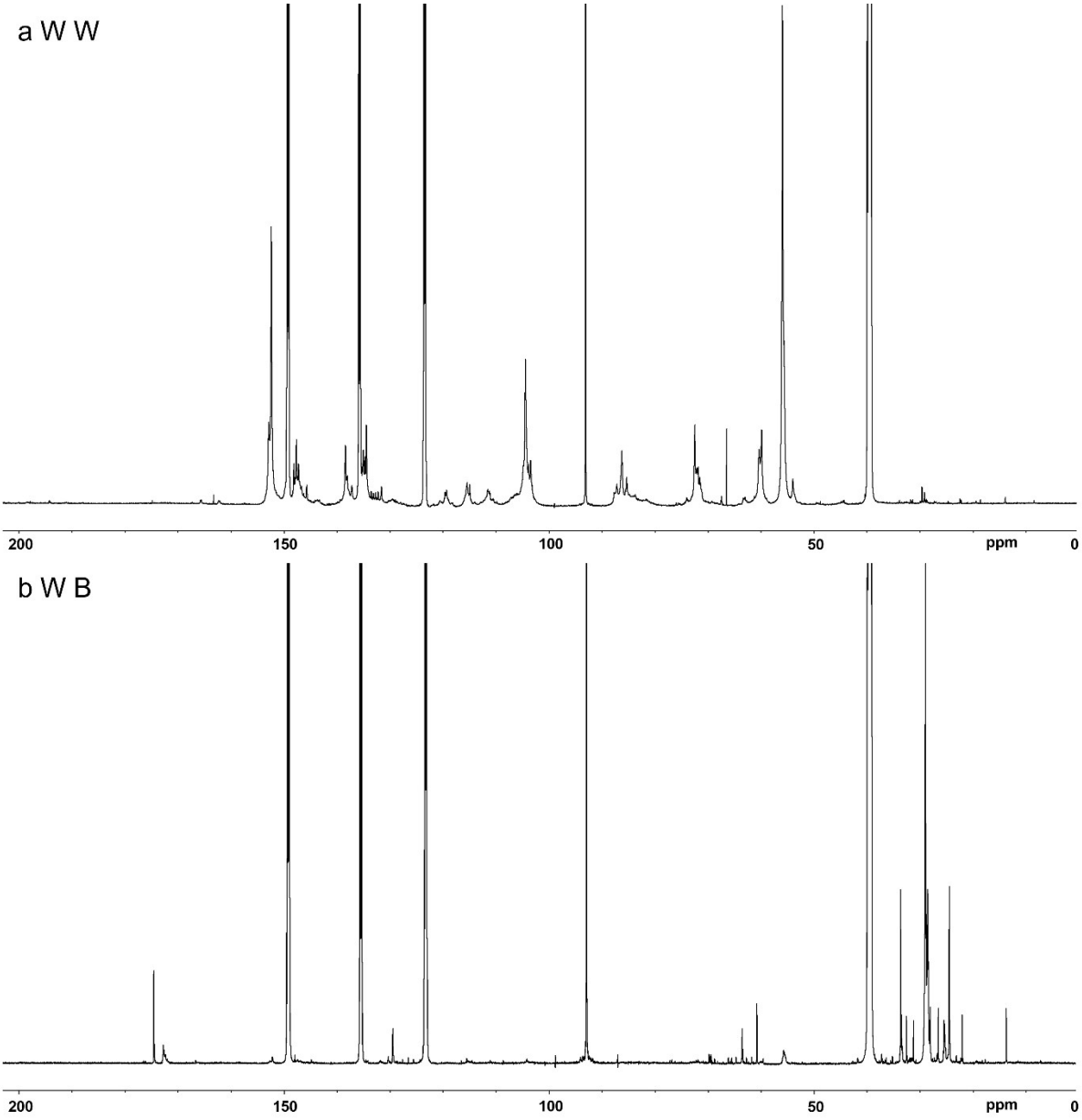
Supplementary Table 12. Assignments of the $^{13}C–^{1}H$ correlation peaks in the 2D HSQC spectra of dioxane lignin (Supplementary Fig. 5) from willow wood (WW) and bark (WB; WIB; WBFB; and Parenchyma).

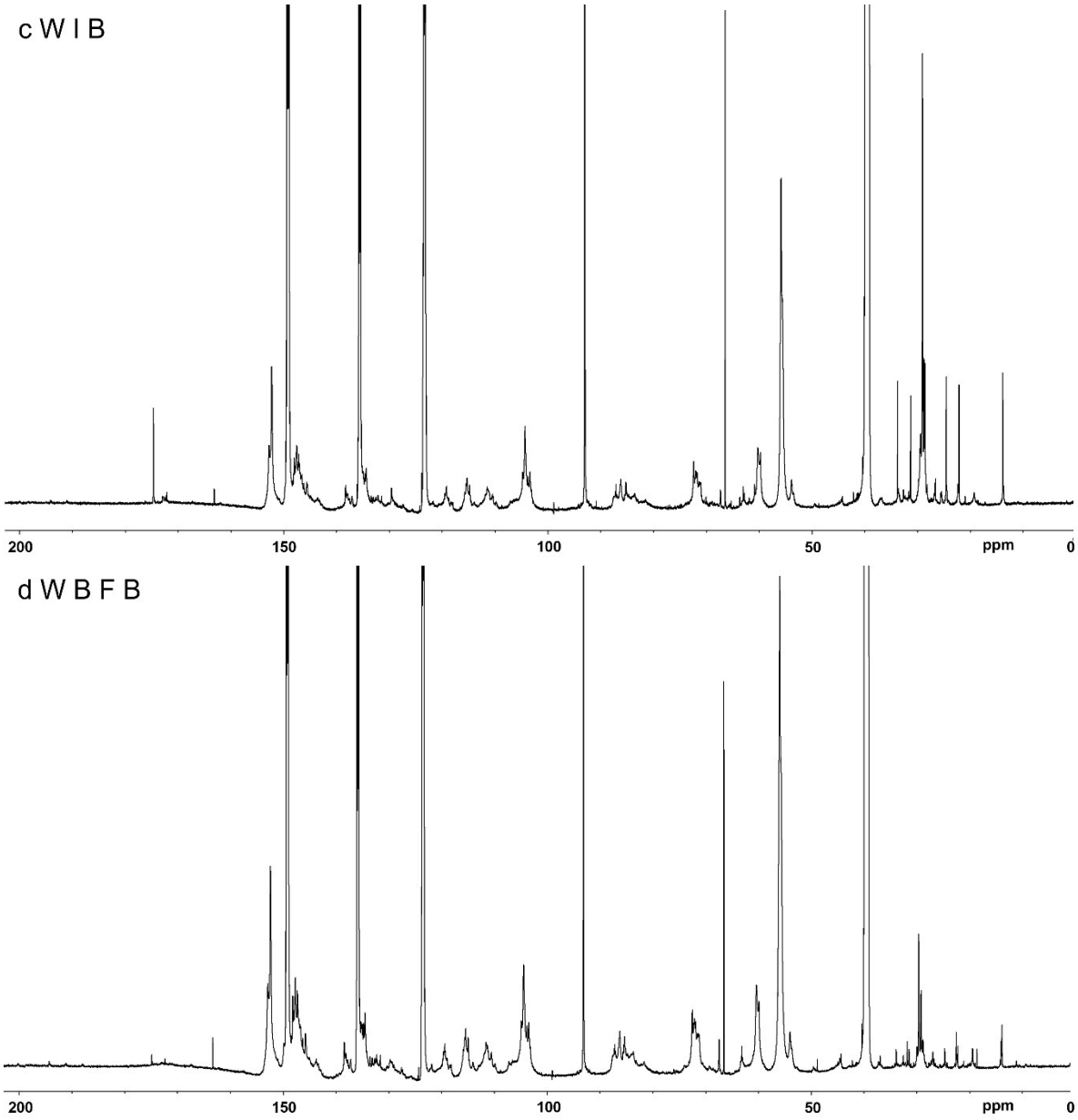
label	δC/δ H (ppm)	assignment
B_β	53.3/3.51	$C_\beta–H_\beta$ in phenylcoumaran substructures (**B**)
C_β	53.8/3.10	$C_\beta–H_\beta$ in β–β′ resinol substructures **(C)**
$–OCH_3$	55.6/3.74	C–H in methoxyls
A_γ	59.8/3.76 and 58.9/3.52	$C_\gamma–H_\gamma$ in γ-hydroxylated β-O-4′ substructures **(A)**
B_γ	62.9/3.77	$C_\gamma–H_\gamma$ in phenylcoumaran substructures **(B)**
C_γ	71.2/3.87 and 71.1/4.22	$C_\gamma–H_\gamma$ in β–β′ resinol substructures **(C)**
A_α	71.9/5.03	$C_\alpha–H_\alpha$ in β-O-4′ substructures **(A)** linked to S-unit
$A_{\beta(H/G)}$	84.0/4.45	$C_\beta–H_\beta$ in β-O-4′ substructures **(A)** linked to H- and G- unit
C_α	85.1/4.72	$C_\alpha–H_\alpha$ in β–β′ resinol substructures **(C)**
$A_{\beta(S)}$	86.1/4.25	$C_\beta–H_\beta$ in β-O-4′ substructures linked **(A)** to S unit
B_α	87.2/5.59	$C_\alpha–H_\alpha$ in phenylcoumaran substructures **(B)**
$S_{2,6}$	104.2/6.80 and 106.5/7.40	$C_2–H_2$ and $C_6–H_6$ in etherified syringyl units **(S)**
G_2	110.8/6.95 and 111.3/7.10	$C_2–H_2$ in guaiacyl units **(G)**
FA_β	114.28/ 6.51	$C_\beta–H_\beta$ in ferulic acid substructures **(FA)**
G_5/G_6	114.8/6.80 and 119.1/6.90	$C_5–H_5$ and $C_6–H_6$ in guaiacyl units **(G)**
FA_α	144.39/ 7.61	$C_\alpha–H_\alpha$ in ferulic acid substructures **(FA)**


a
174.3
154
148
136.4
118.9
105.4
89.0
83.9
75.0
72.4
65.2
63.0
56.1
20.9
WW_L
WW_Dioxane
WW_N
WW_O
200
150
100
50
ppm
0
b
174.3
152.8
146.5
130.1
105.4
89.1
84.1
75.2
72.9
65.2
63.0
56.1
33.3
30.3
WB_CEL
WB_L
WB_ Dioxane
WB_N
WB_O
200
150
100
50
ppm
0

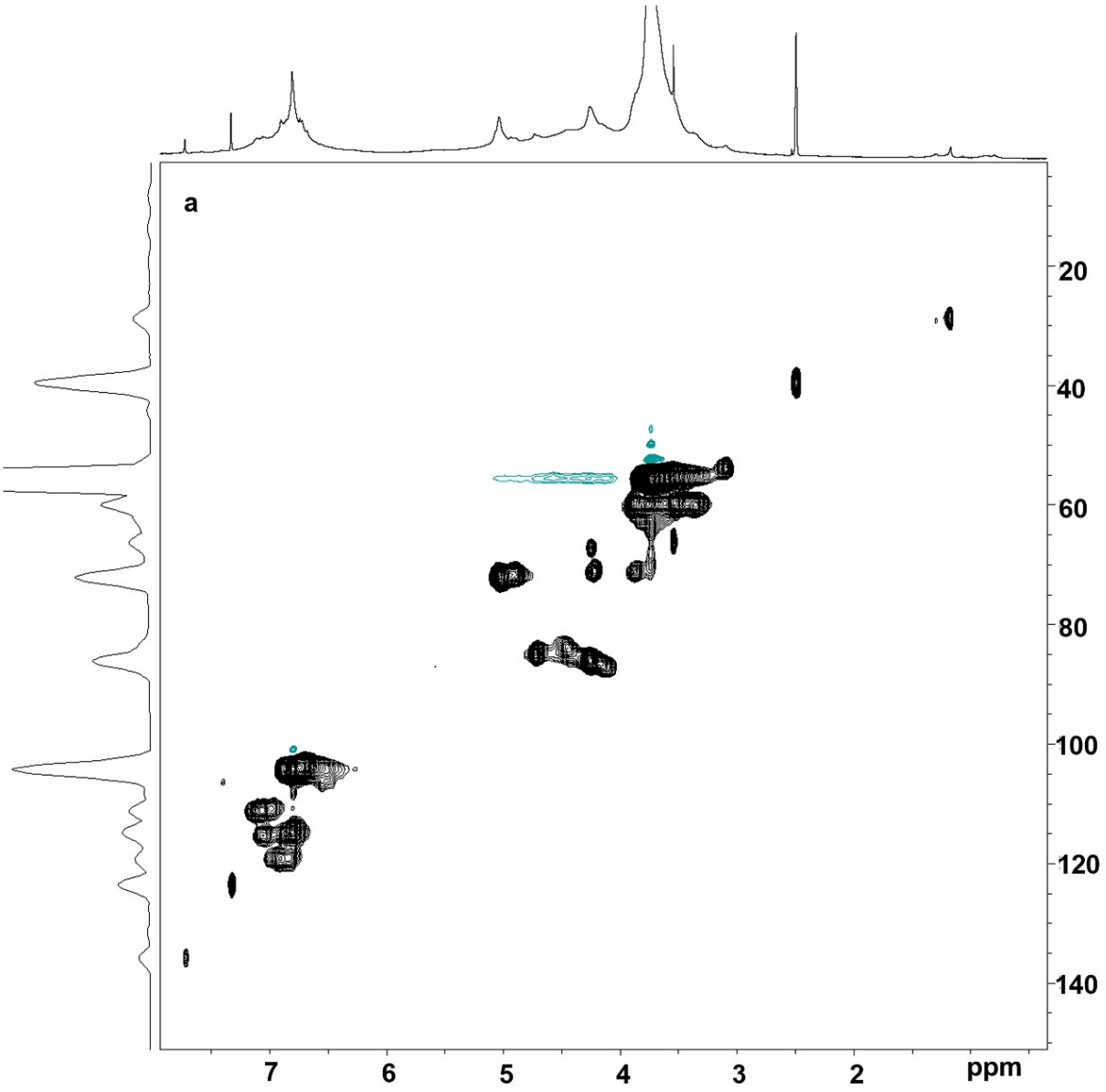


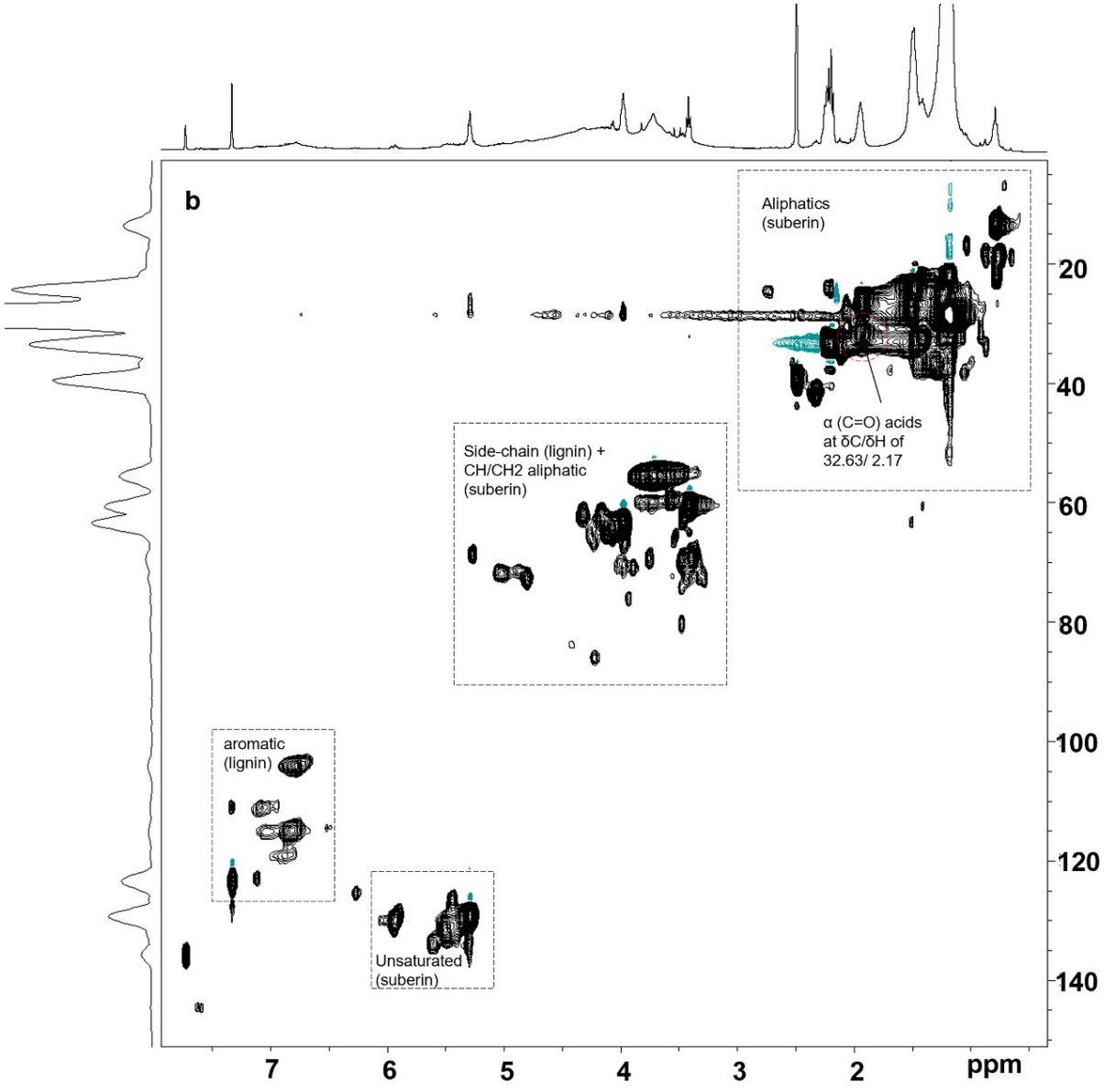
Supplementary Fig. 20 CPMAS NMR spectrums of four different samples that are collected from their original form (O), 0.1M NaOH treated solid residue (N), solid residues after the dioxane extraction (Dioxane), and the recovered dioxane lignin (L). a Wood. **b** Bark. **c** Inner bark. **d** Fiber bundle. The cellulytic enzyme lignin (CEL) is also included as the reference for comparison.[15]

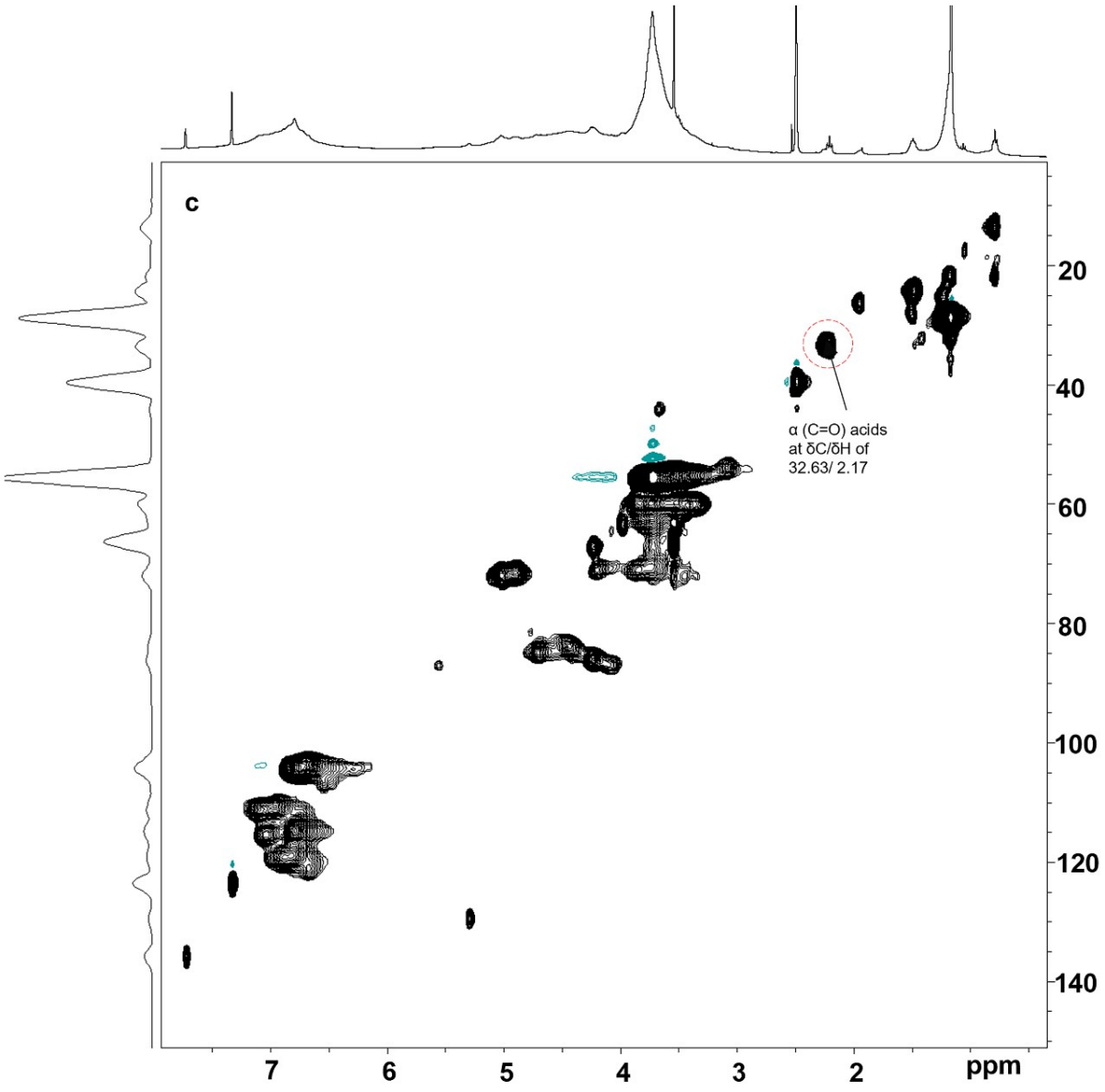

a W W
200
150
100
50
ppm
0
b W B
200
150
100
50
ppm
0

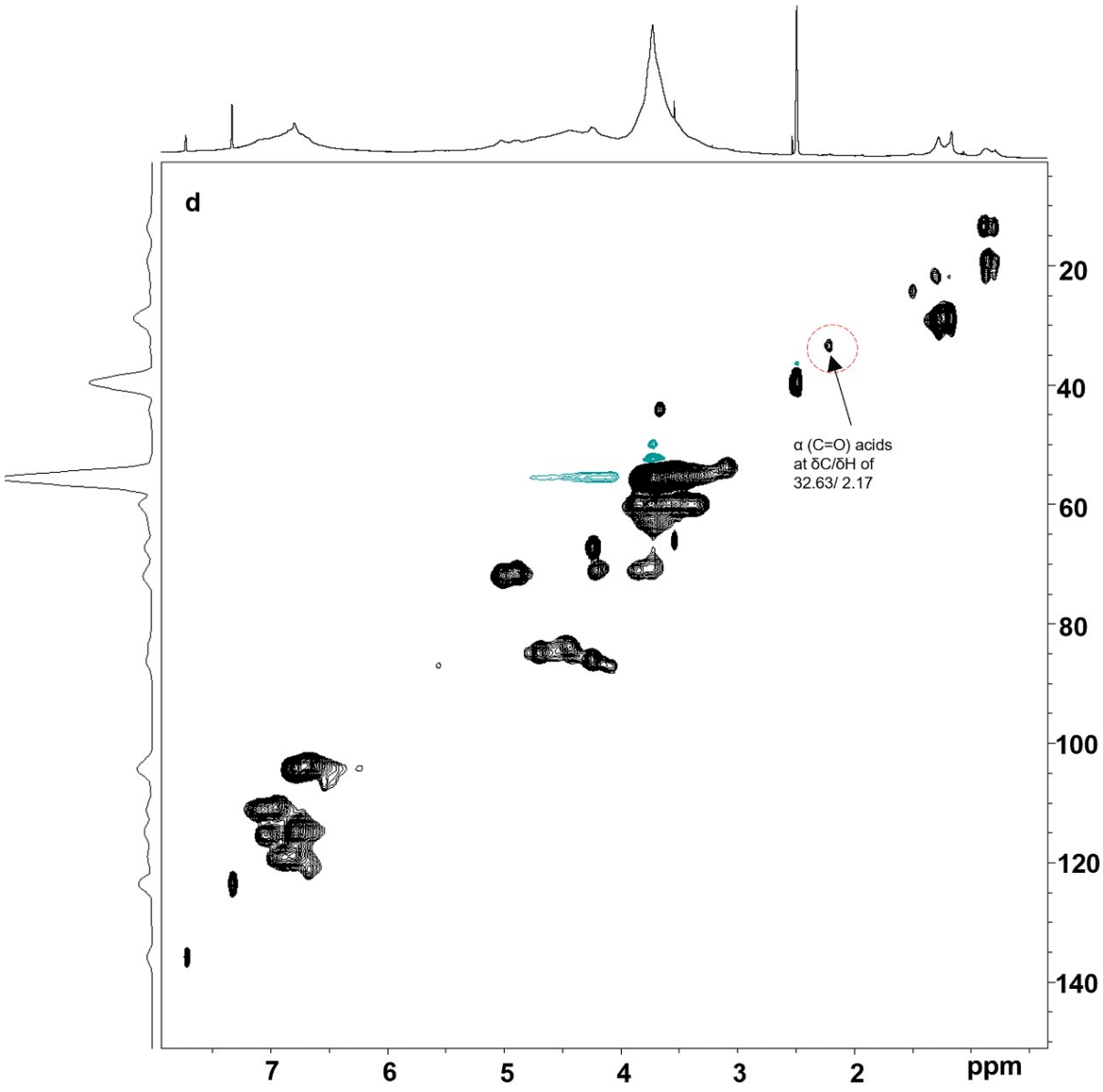


Supplementary Fig. 21 Quantitative liquid state ^{13}C NMR spectrum of the recovered dioxane lignin (DMSO-*d*6/pyridine-*d*5, v/v 4/1; 1,3,5-trioxane as the internal standard). **a** Willow wood (WW). **b** Willow bark (WB). **c** Willow inner bark (WIB). **d** Willow bark fibre bundle (WB FB).

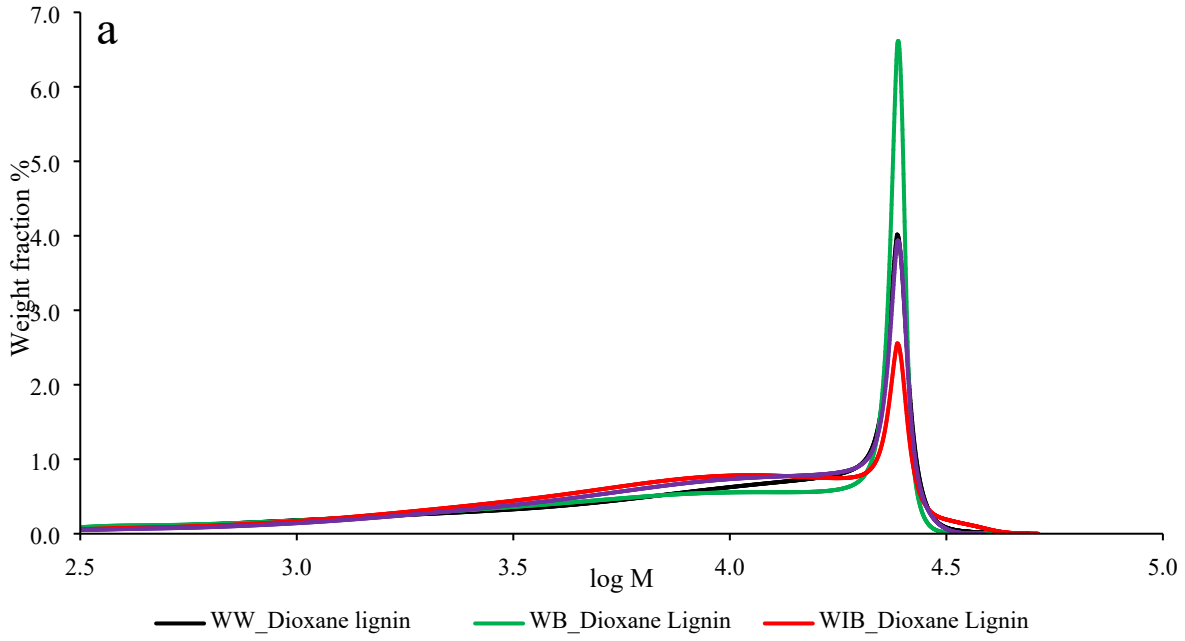

a
20
40
60
80
100
120
140
7
6
5
4
3
2
ppm

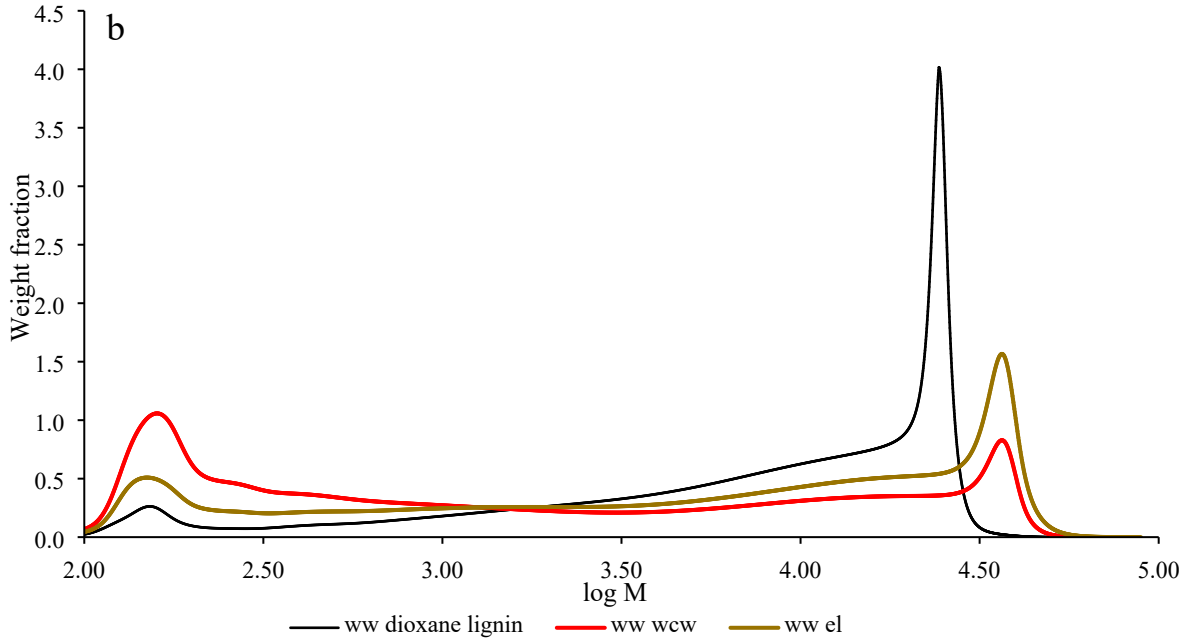

b
Aliphatics
(suberin)
α (C=O) acids
at δC/δH of
32.63/ 2.17
Side-chain (lignin) +
CH/CH2 aliphatic
(suberin)
aromatic
(lignin)
Unsaturated
(suberin)
20
40
60
80
100
120
140
7
6
5
4
3
2
ppm

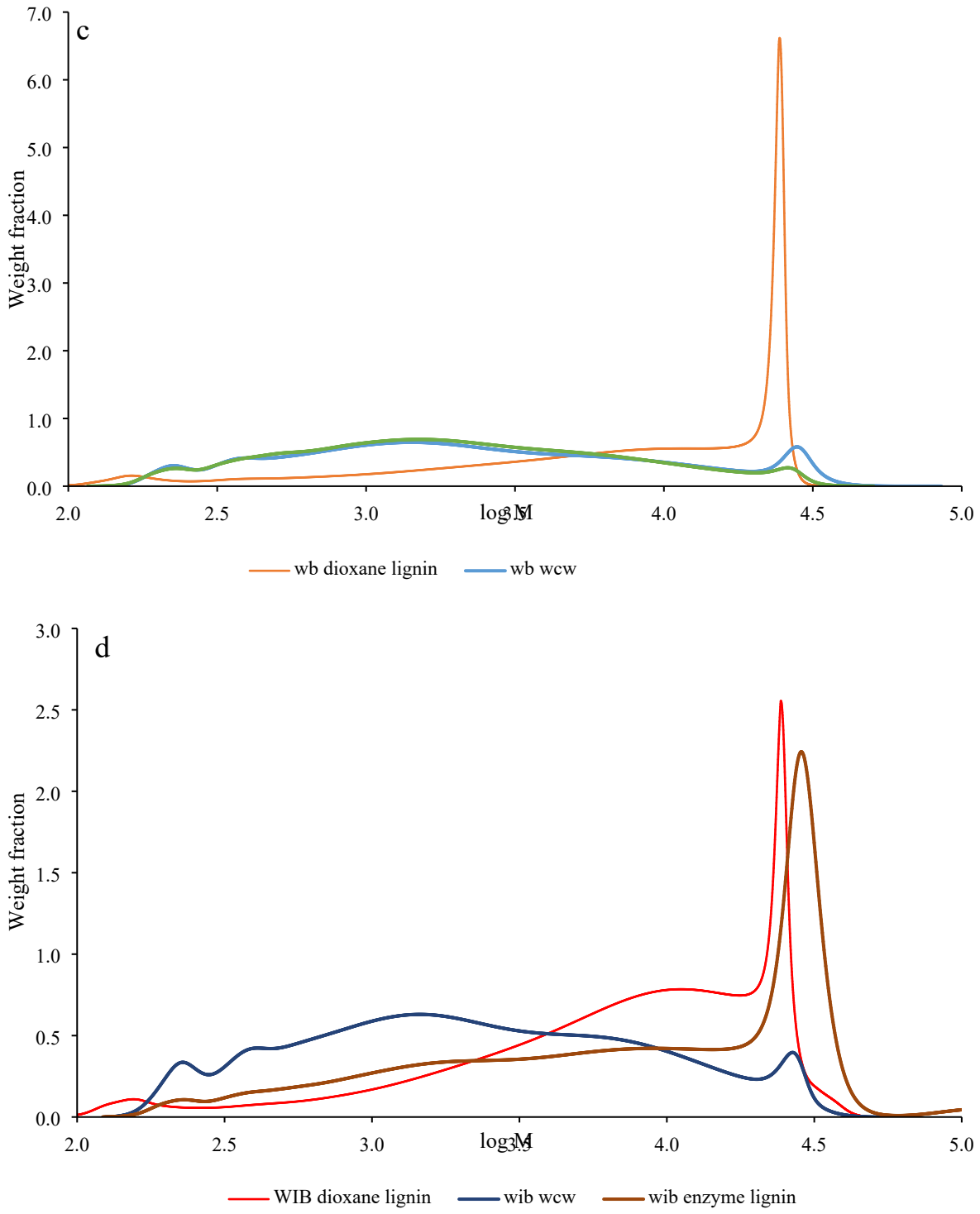

c
α (C=O) acids
at δC/δH of
32.63/ 2.17
20
40
60
80
100
120
140
7
6
5
4
3
2
ppm



Supplementary Fig. 22 HSQC NMR spectrum of the recovered dioxane lignin (DMSO-*d*6/pyridine-*d*5, v/v 4/1). a Willow wood (WW). **b** Willow bark (WB). **c** Willow inner bark (WIB). **d** Willow bark fibre bundle (WB FB).


a
Weight fraction %
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
2.5
3.0
3.5
4.0
4.5
5.0
log M
WW_Dioxane lignin
WB_Dioxane Lignin
WIB_Dioxane Lignin


b
Weight fraction
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
2.00
2.50
3.00
3.50
4.00
4.50
5.00
log M
ww dioxane lignin
ww wcw
ww el



Supplementary Fig. 23 The SEC elution profile of different lignin. **a** Willow dioxane lignin (WW; WB; WIB; WBFB). **b** Different WW lignin (Dioxane lignin, CEL, and WCW). **c** Different WB lignin (Dioxane lignin, CEL, and WCW). **d** Different WIB lignin (Dioxane lignin, CEL, and WCW). For abbreviations, see **Supplementary Fig. 1–5**.

Supplementary Table 13. **Average molecular weight (g/mol) of the dioxane lignin fractions in comparison to the whole cell wall (WCW) and cellulytic enzyme lignin (CEL)[15] fractions from both hybrids of Klara and Karin.** Mn: number-average molecular weight; Mw: weight-average molecular weight which can be averaged according to the weight of molecules; Mp: peak maximum molar mass.

source	sample	abbreviations	Mn	Mw	PDI	Mp
This study	Klara WW dioxane Lignin	WW_L	1575	12260	7.8	24330
This study	Klara WB dioxane Lignin	WB_L	1951	12740	6.5	24470
This study	Klara WIB dioxane Lignin	WIB_L	2187	11130	5.1	24370
This study	Klara WB FB dioxane Lignin	WBFB_L	2613	12370	4.7	24440
Dou et al. 2017[15]	Karin wood ww-wcw	WW_WCW	365.2	7768	21.3	161.4
Dou et al. 2017[15]	Karin wood ww-el	WW_CEL	699.3	12940	18.5	36400
Dou et al. 2017[15]	Karin bark WCW	WB_WCW	1059	5769	5.4	1437
Dou et al. 2017[15]	Karin bark enzyme lignin	WB_CEL	1027	4348	4.2	1524
Dou et al. 2017[15]	Karin inner bark WCW	WIB_WCW	1036	5040	4.9	1469
Dou et al. 2017[15]	Karin inner bark enzyme Lignin	WIB_CEL	2295	16340	7.1	28540

Supplementary Table 14. The light green coloured cells indicate the tailored enzyme cocktail after screening and selection based on structural features of pectin and hemicellulose of willow wood (WW). Screening % refers to the percentages of candidate enzymes that can be ruled out theoretically based on structural characteristics of the substrate (i.e., WW). Literatures refer to **Table 4**.

Substrate (screening %)	Specific area	Cleaving glycosidic bond	Enzyme	EC number	CAZy family	Abbreviation	Characteristics on the preferred structural features of the substrate
Pectin (84)	HG	cleave α-(1,4)-glycosidic bonds of Gal*p*A backbone	polygalacturonase	3.2.1.67	GH28	Exo-PG	prefer non-esterified HG, show decreasing activity with increasing degree of methyl-esterification
				3.2.1.15	GH28	Endo-PG	
			polymethyl galacturonase	3.2.1.-	GH28	Endo-PMG	HG with at least 75% of the carboxyl groups are esterified
			pectate lyase	4.2.2.2	PL1/PL2/PL3/PL9/PL10	Endo-PGL	low methyl-esterified HG/non-esterified HG
			pectin lyase	4.2.2.10	PL1	Endo-PNL	highly methyl-esterified HG
			xylogalacturonan hydrolase	3.2.1.-	GH28	Endo-XGH	cleaves the β-Xyl*p* substituted Gal*p*A backbone, prefers to act between two xylosidated GalA units
	HG (side chain)	active on the side chains of HG	pectin methyl esterase	3.1.1.11	CE8	PME	esterified HG
			pectin/rhamnogalacturonan acetyl esterase	3.1.1.6	CE12	PAE, RGAE	acetylated HG/ acetylated RG
				3.1.1.86	CE16	PAE	
	RG-I	cleave α-(1,2) glycosidic bonds between D-Gal*p*A and L-Rha*p*	rhamnogalacturonan I hydrolase	3.2.1.171	GH28	Endo-RGH	activity is not hindered by galactose substitutions on the RG-I backbone
		cleave α-(1,4) glycosidic bonds between Rha*p* and Gal*p*A	rhamnogalacturonan lyase	4.2.2.23	PL4	Endo-RGL	activity is increased by removal of acetyl groups and Ara side chains, decreased by removal of Gal side chains
				4.2.2.23	PL11	Endo-RGL	prefer RG-I backbone having galactan side chains without acetyl substitution
		cleave glycosidic bonds of arabinan/galactan	endo-α-L-1,5-arabinanase	3.2.1.99	GH43	Endo-Ara	random hydrolyzing α-(1,4)-glycosidic bonds of arabinan
			endo-β-1,4-galactanase	3.2.1.89	GH53	Endo-Gal	random hydrolyzing β-(1,4)-glycosidic bonds of arabinogalactan I
			endo-β-1,6-galactanase	3.2.1.164	GH30	Endo-Gal	random hydrolyzing β-(1,6)-glycosidic bonds of arabinogalactan II
hemicellulose (70)	xylan (main chain)	cleave β-(1,4)-glycosidic bonds of xylan backbone	endo-1,4-β-xylanase	3.2.1.8	GH10	Endo-Xyn	require two unsubstituted consecutive Xyl*p*, higher tolerance to substituents than GH10
					GH11		prefer unsubstituted xylan, require three unsubstituted consecutive Xyl*p*
			arabinoxylanase	3.2.1.-	GH5		hydrolyzing arabinoxylan but does not attack unsubstituted xylan
			glucuronoxylanase	3.2.1.136	GH30		hydrolyzing glucuronoxylan but does not attack unsubstituted xylan
	xylan (side chains)	active on the side chains of xylan	acetyl xylan esterase	3.1.1.72	CE1	AcXE	action on 2,3-di-O-acetylated Xyl*p* in xylan
					CE3		deacetylated for O-2 or O-3 monoacetylated Xylp in xylan
					CE4		deacetylated effectively for singly 2- or 3-O-acetylated Xyl*p* residues in xylan
					CE5		prefer O-2 monoacetylated Xyl*p* in xylan
					CE6		deacetylated for 2,3-di-O-acetylated Xyl*p* in xylan
			α-L-arabinofuranosidase	3.2.1.55	GH43	Abf	hydrolyzing (1→3)-α-l-arabinofuranosyl residues of doubly substituted xylopyranosyl residues in arabinoxylan
					GH51		release (1 → 2) and (1 → 3)-α-L-arabinofuranosyl residues from mono-substituted xylopyranosyl residues in arabinoxylan
					GH54		specific for O-3 linked arabinofuranosyl residues, Xyl*p* is mono-substituted or di-substituted
					GH62		removing arabinofuranosyl residues linked in position O-2 or O-3 of mono-substituted xylose units in arabinoxylan
			α-glucuronidase	3.2.1.139	GH115		removing 4-O-methyl glucuronic acid side groups from polymeric xylan
	mannan	cleave β-(1,4)-glycosidic bonds of mannan backbone	endo-β-1,4-mannanase	3.2.1.78	GH5	Man	prefer glucomannan
					GH26		prefer galactomannan
		cleave T-α-(1,6)-glycosidic bonds of galactomannan	α-galactosidase	3.2.1.22	GH27		removing the α-1,6-linked d-galactopyranosyl substituents attached to the mannan backbone

Supplementary Table 15. The light green colored cells indicate the tailored enzyme cocktail after screening and selection based on structural features of pectin and hemicellulose of willow bark (WB). Screening % refers to the percentages of candidate enzymes that can be ruled out theoretically based on structural characteristics of the substrate (i.e., WB). Literatures refer to **Table 4**.

Substrate (screening %)	Specific area	Cleaving glycosidic bond	Enzyme	EC number	CAZy family	Abbreviation	Characteristics on the preferred structural features of the substrate
Pectin (58)	HG	cleave α-(1,4)-glycosidic bonds of Gal*p*A backbone	polygalacturonase	3.2.1.67	GH28	Exo-PG	prefer non-esterified HG, show decreasing activity with increasing degree of methyl-esterification
				3.2.1.15	GH28	Endo-PG	
			polymethyl galacturonase	3.2.1.-	GH28	Endo-PMG	HG with at least 75% of the carboxyl groups are esterified
			pectate lyase	4.2.2.2	PL1/PL2/PL3/PL9/PL10	Endo-PGL	low methyl-esterified HG/non-esterified HG
			pectin lyase	4.2.2.10	PL1	Endo-PNL	highly methyl-esterified HG
			xylogalacturonan hydrolase	3.2.1.-	GH28	Endo-XGH	cleaves the β-Xyl*p* substituted Gal*p*A backbone, prefers to act between two xylosidated GalA units
	HG (side chain)	active on the side chains of HG	pectin methyl esterase	3.1.1.11	CE8	PME	esterified HG
			pectin/rhamnogalacturonan acetyl esterase	3.1.1.6	CE12	PAE, RGAE	acetylated HG/ acetylated RG
				3.1.1.86	CE16	PAE	
	RG-I	cleave α-(1,2) glycosidic bonds between D-Gal*p*A and L-Rha*p*	rhamnogalacturonan I hydrolase	3.2.1.171	GH28	Endo-RGH	activity is not hindered by galactose substitutions on the RG-I backbone
		cleave α-(1,4) glycosidic bonds between Rha*p* and Gal*p*A	rhamnogalacturonan lyase	4.2.2.23	PL4	Endo-RGL	activity is increased by removal of acetyl groups and Ara side chains, decreased by removal of Gal side chains
				4.2.2.23	PL11	Endo-RGL	prefer RG-I backbone having galactan side chains without acetyl substitution
		cleave glycosidic bonds of arabinan/galactan	endo-α-L-1,5-arabinanase	3.2.1.99	GH43	Endo-Ara	random hydrolyzing α-(1,4)-glycosidic bonds of arabinan
			endo-β-1,4-galactanase	3.2.1.89	GH53	Endo-Gal	random hydrolyzing β-(1,4)-glycosidic bonds of arabinogalactan I
			endo-β-1,6-galactanase	3.2.1.164	GH30	Endo-Gal	random hydrolyzing β-(1,6)-glycosidic bonds of arabinogalactan II
hemicellulose (35)	xylan (main chain)	cleave β-(1,4)-glycosidic bonds of xylan backbone	endo-1,4-β-xylanase	3.2.1.8	GH10	Endo-Xyn	require two unsubstituted consecutive Xyl*p*, higher tolerance to substituents than GH10
					GH11		prefer unsubstituted xylan, require three unsubstituted consecutive Xyl*p*
			arabinoxylanase	3.2.1.-	GH5		hydrolyzing arabinoxylan but not attack unsubstituted xylan
			glucuronoxylanase	3.2.1.136	GH30		hydrolyzing glucuronoxylan but not attack unsubstituted xylan
	xylan (side chains)	active on the side chains of xylan	acetyl xylan esterase	3.1.1.72	CE1	AcXE	action on 2,3-di-O-acetylated Xyl*p* in xylan
					CE3		deacetylated for O-2 or O-3 monoacetylated Xyl*p* in xylan
					CE4		deacetylated effectively for singly 2- or 3-O-acetylated Xyl*p* residues in xylan
					CE5		prefer O-2 monoacetylated Xyl*p* in xylan
					CE6		deacetylated for 2,3-di-O-acetylated Xyl*p* in xylan
			α-L-arabinofuranosidase	3.2.1.55	GH43	Abf	hydrolyzing (1→3)-α-l-arabinofuranosyl residues of doubly substituted xylopyranosyl in arabinoxylan
					GH51		release (1 → 2) and (1 → 3)-α-L-arabinofuranosyl residues from mono-substituted xylopyranosyl residues in arabinoxylan
					GH54		specific for O-3 linked arabinofuranosyl residues, Xyl*p* is mono-substituted or di-substituted
					GH62		remove arabinofuranosyl residues linked in O-2 or O-3 of mono-substituted xylose units in arabinoxylan
			α-glucuronidase	3.2.1.139	GH115		remove 4-O-methyl glucuronic acid side groups from polymeric xylan
	mannan	cleave β-(1,4)-glycosidic bonds of mannan	endo-β-1,4-mannanase	3.2.1.78	GH5	Man	prefer glucomannan
					GH26		prefer galactomannan
		cleave T-α-(1,6)-glycosidic bonds of galactomannan	α-galactosidase	3.2.1.22	GH27		remove the α-1,6-linked d-galactopyranosyl substituents attached to the mannan backbone

Supplementary Table 16. The light green colored cells indicate the tailored enzyme cocktail after screening and selection based on structural features of pectin and hemicellulose of willow bark fiber bundle (WBFB). Screening % refers to the percentages of candidate enzymes that can be ruled out theoretically based on structural characteristics of the substrate (i.e., WBFB). Literatures refer to **Table 4**.

Substrate (screening %)	Specific area	Cleaving glycosidic bond	Enzyme	EC number	CAZy family	Abbreviation	Characteristics on the preferred structural features of the substrate
Pectin (63)	HG	cleave α-(1,4)-glycosidic bonds of Gal*p*A backbone	polygalacturonase	3.2.1.67	GH28	Exo-PG	prefer non-esterified HG, show decreasing activity with increasing degree of methyl-esterification
				3.2.1.15	GH28	Endo-PG	
			polymethyl galacturonase	3.2.1.-	GH28	Endo-PMG	HG with at least 75% of the carboxyl groups are esterified
			pectate lyase	4.2.2.2	PL1/PL2/PL3/PL9/PL10	Endo-PGL	low methyl-esterified HG/non-esterified HG
			pectin lyase	4.2.2.10	PL1	Endo-PNL	highly methyl-esterified HG
			xylogalacturonan hydrolase	3.2.1.-	GH28	Endo-XGH	cleaves the β-Xyl*p* substituted Gal*p*A backbone, prefers to act between two xylosidated GalA units
	HG (side chain)	active on the side chains of HG	pectin methyl esterase	3.1.1.11	CE8	PME	esterified HG
			pectin/rhamnogalacturonan acetyl esterase	3.1.1.6	CE12	PAE, RGAE	acetylated HG/ acetylated RG
				3.1.1.86	CE16	PAE	
	RG-I	cleave α-(1,2) glycosidic bonds between D-Gal*p*A and L-Rha*p*	rhamnogalacturonan I hydrolase	3.2.1.171	GH28	Endo-RGH	activity is not hindered by galactose substitutions on the RG-I backbone
		cleave α-(1,4) glycosidic bonds between Rha*p* and Gal*p*A	rhamnogalacturonan lyase	4.2.2.23	PL4	Endo-RGL	activity is increased by removal of acetyl groups and Ara side chains, decreased by removal of Gal side chains
				4.2.2.23	PL11	Endo-RGL	prefer RGI backbone having galactan side chains without acetyl substitution
		cleave glycosidic bonds of arabinan/galactan	endo-α-L-1,5-arabinanase	3.2.1.99	GH43	Endo-Ara	random hydrolyzing α-(1,4)-glycosidic bonds of arabinan
			endo-β-1,4-galactanase	3.2.1.89	GH53	Endo-Gal	random hydrolyzing β-(1,4)-glycosidic bonds of arabinogalactan I
			endo-β-1,6-galactanase	3.2.1.164	GH30	Endo-Gal	random hydrolyzing β-(1,6)-glycosidic bonds of arabinogalactan II
hemicellulose (88)	xylan (main chain)	cleave β-(1,4)-glycosidic bonds of xylan backbone	endo-1,4-β-xylanase	3.2.1.8	GH10	Endo-Xyn	require two unsubstituted consecutive Xylp, higher tolerance to substituents than GH10
					GH11		prefer unsubstituted xylan, require three unsubstituted consecutive Xyl*p*
			arabinoxylanase	3.2.1.-	GH5		hydrolyzing arabinoxylan but does not attack unsubstituted xylan
			glucuronoxylanase	3.2.1.136	GH30		hydrolyzing glucuronoxylan but does not attack unsubstituted xylan
	xylan (side chains)	active on the side chains of xylan	acetyl xylan esterase	3.1.1.72	CE1	AcXE	action on 2,3-di-O-acetylated Xyl*p* in xylan
					CE3		deacetylated for O-2 or O-3 monoacetylated Xyl*p* in xylan
					CE4		deacetylated effectively for singly 2- or 3-O-acetylated Xyl*p* residues in xylan
					CE5		prefer O-2 monoacetylated Xyl*p* in xylan
					CE6		deacetylated for 2,3-di-O-acetylated Xyl*p* in xylan
			α-L-arabinofuranosidase	3.2.1.55	GH43	Abf	hydrolyzing (1→3)-α-l-arabinofuranosyl residues of doubly substituted xylopyranosyl residues in arabinoxylan
					GH51		releasing (1 → 2) and (1 → 3)-α-L-arabinofuranosyl residues from mono-substituted xylopyranosyl in arabinoxylan
					GH54		specific for O-3 linked arabinofuranosyl, Xyl*p* is mono-substituted or di-substituted
					GH62		removing arabinofuranosyl residues linked in O-2 or O-3 of mono-substituted xylose units in arabinoxylan
			α-glucuronidase	3.2.1.139	GH115		removing -MG side groups from polymeric xylan
	mannan	cleave β-(1,4)-glycosidic bonds of mannan backbone	endo-β-1,4-mannanase	3.2.1.78	GH5	Man	prefer glucomannan
					GH26		prefer galactomannan
		cleave T-α-(1,6)-glycosidic bonds of galactomannan	α-galactosidase	3.2.1.22	GH27		removing the α-1,6-linked d-galactopyranosyl substituents attached to the mannan backbone

Supplementary Table 17. The light green colored cells indicate the tailored enzyme cocktail after screening and selection based on structural features of pectin and hemicellulose of parenchyma tissues (Parenchyma). Screening % refers to the percentages of candidate enzymes that can be ruled out theoretically based on structural characteristics of the substrate (i.e., Parenchyma). Literatures refer to **Table 4**.

Substrate (screening %)	Specific area	Cleaving glycosidic bond	Enzyme	EC number	CAZy family	Abbreviation	Characteristics on the preferred structural features of the substrate
Pectin (63)	HG	cleave α-(1,4)-glycosidic bonds of Gal*p*A backbone	polygalacturonase	3.2.1.67	GH28	Exo-PG	prefer non-esterified HG, show decreasing activity with increasing degree of methyl-esterification
				3.2.1.15	GH28	Endo-PG	
			polymethyl galacturonase	3.2.1.-	GH28	Endo-PMG	HG with at least 75% of the carboxyl groups are esterified
			pectate lyase	4.2.2.2	PL1/PL2/PL3/PL9/PL10	Endo-PGL	low methyl-esterified HG/non-esterified HG
			pectin lyase	4.2.2.10	PL1	Endo-PNL	highly methyl-esterified HG
			xylogalacturonan hydrolase	3.2.1.-	GH28	Endo-XGH	cleaves the β-Xyl*p* substituted Gal*p*A backbone, prefers to act between two xylosidated GalA units
	HG (side chain)	active on the side chains of HG	pectin methyl esterase	3.1.1.11	CE8	PME	esterified HG
			pectin/rhamnogalacturonan acetyl esterase	3.1.1.6	CE12	PAE, RGAE	acetylated HG/ acetylated RG
				3.1.1.86	CE16	PAE	
	RG-I	cleave α-(1,2) glycosidic bonds between D-Gal*p*A and L-Rha*p*	rhamnogalacturonan I hydrolase	3.2.1.171	GH28	Endo-RGH	activity is not hindered by galactose substitutions on the RG-I backbone
		cleave α-(1,4) glycosidic bonds between Rha*p* and Gal*p*A	rhamnogalacturonan lyase	4.2.2.23	PL4	Endo-RGL	activity is increased by removal of acetyl groups and Ara side chains, decreased by removal of Gal side chains
				4.2.2.23	PL11	Endo-RGL	prefer RGI backbone having galactan side chains without acetyl substitution
		cleave glycosidic bonds of arabinan/galactan	endo-α-L-1,5-arabinanase	3.2.1.99	GH43	Endo-Ara	random hydrolyzing α-(1,4)-glycosidic bonds of arabinan
			endo-β-1,4-galactanase	3.2.1.89	GH53	Endo-Gal	random hydrolyzing β-(1,4)-glycosidic bonds of arabinogalactan I
			endo-β-1,6-galactanase	3.2.1.164	GH30	Endo-Gal	random hydrolyzing β-(1,6)-glycosidic bonds of arabinogalactan II
hemicellulose (65)	xylan (main chain)	cleave β-(1,4)-glycosidic bonds of xylan backbone	endo-1,4-β-xylanase	3.2.1.8	GH10	Endo-Xyn	require two unsubstituted consecutive Xyl*p*, higher tolerance to substituents than GH10
					GH11		prefer unsubstituted xylan, require three unsubstituted consecutive Xyl*p*
			arabinoxylanase	3.2.1.-	GH5		hydrolyzes arabinoxylan but does not attack unsubstituted xylan
			glucuronoxylanase	3.2.1.136	GH30		hydrolyzes glucuronoxylan but does not attack unsubstituted xylan
	xylan (side chains)	active on the side chains of xylan	acetyl xylan esterase	3.1.1.72	CE1	AcXE	action on 2,3-di-O-acetylated Xyl*p* in xylan
					CE3		deacetylated for O-2 or O-3 monoacetylated Xyl*p* in xylan
					CE4		deacetylated effectively for singly 2- or 3-O-acetylated Xyl*p* residues in xylan
					CE5		prefer O-2 monoacetylated Xyl*p* in xylan
					CE6		deacetylated for 2,3-di-O-acetylated Xyl*p* in xylan
			α-L-arabinofuranosidase	3.2.1.55	GH43	Abf	hydrolyse (1→3)-α-l-arabinofuranosyl residues of doubly substituted xylopyranosyl residues in arabinoxylan
					GH51		release (1 → 2) and (1 → 3)-α-L-arabinofuranosyl residues from mono-substituted xylopyranosyl residues in arabinoxylan
					GH54		specific for O-3 linked arabinofuranosyl residues, Xylp is mono-substituted or di-substituted
					GH62		removing arabinofuranosyl residues linked in position O-2 or O-3 of mono-substituted xylose units in arabinoxylan
			α-glucuronidase	3.2.1.139	GH115		removing 4-O-methyl glucuronic acid side groups from polymeric xylan
	mannan	cleave β-(1,4)-glycosidic bonds of mannan backbone	endo-β-1,4-mannanase	3.2.1.78	GH5	Man	prefer glucomannan
					GH26		prefer galactomannan
		cleave T-α-(1,6)-glycosidic bonds of galactomannan	α-galactosidase	3.2.1.22	GH27		remove the α-1,6-linked d-galactopyranosyl substituents attached to the mannan backbone

Supplementary Table 18. The summarized characterization techniques that were applied at this study for all samples of WW; WB; WIB; WBFB; Parenchyma.

category	code	notes	HPAEC	FT-IR	$^{1}H/^{13}C$ nmr	HSQC nmr	CP/MAS nmr	staining	GPC (SEC)	HPLC	GC-MS
original	O	raw sample	—	—			—	—		—	
multiple treatment	E	solid residues after extraction with DCM/ acetone/ water									
	C	solid residues after citric acid extraction	—								
	N	solid residues after 0.1M NaOH treatment	—	—			—			—	
purified substances	P	solid residues after peracetic acid treatment	—	—						—	
	HTS	hydrolysable tannin-like substances									—
	CAP; DCA-P	purified pectin; dialysis treated pectin	—			—			—		
	DMSO	solid residues after DMSO extraction	—	—							
	H	purified hemicellulose	—	—	—	—		— (parenchyma alone)	—		
	Dioxane	solid residues after dioxane lignin purification	—	—			—			—	
	L	purified dioxane lignin		—	—	—	—		—		
	WCW; CEL	Dou et al. 2018[15]					—		—		