Electronic Supporting Information (ESI)

Upgrading AquaSolv Omni (AqSO) biorefinery: access to highly ethoxylated lignins in high yields through reactive extraction (REx)

Davide Rigo[a],*, Nadine Kohlhuber[b], Lukas Fliri[a], Daryna Diment[a], Mijung Cho[a], Ivan Sumerskii[c], Michael Hummel[a], Antje Potthast[b], and Mikhail Balakshin[a],†

[a] Department of Bioproducts and Biosystems, Aalto University, Vuorimiehentie 1, Espoo, 02150 (Finland)

[b] Department of Chemistry, Institute of Chemistry of Renewable Resources, University of Natural Resources and Life Sciences Vienna (BOKU), Konrad-Lorenz-Strasse 24, Tulln3430, Austria

[c] Core Facility of Analysis of Lignocellulosics (ALICE), University of Natural Resources and Life Sciences, Vienna, Konrad-Lorenz-Strasse 24, A-3430, Tulln, Austria

†Deceased, 2022.

Corresponding author: *davide.rigo@unive.it*

Table of contents:

The reaction conditions used for reactive extraction (REx)

Reactive extraction has been carried out under different conditions. The effect of [EtOH], [H$_2$SO$_4$] and time were investigated. To facilitate the reader's understanding, **Table S1** summarizes the different conditions employed during the present investigation.

**Table S1.** Varied parameters during reactive extraction.

Entry	Reaction conditions		
	[EtOH] (%)	[H$_2$SO$_4$] (M)	t (h)
1			0.5
2	70		1
3			4
4		0.15	8
5	80		
6	90		4
7			
8	99	0.30	
9		0.60	0.5
10		1.2	

Solvent (EtOH:H$_2$O) + H$_2$SO$_4$ (cat.)

S-500

Figure S1. Experimental apparatus used to carry out REx.

## The effect of parameters on the composition of the residual solids (RS)

The effect of the reactive extraction conditions on the compositional analysis of the residual solids (RS) is reported in **Figure S2**. The effect of [H_2SO_4] is very subtle (**Figure S2**, top right). Similarly, an increase of the time does not significantly affect the composition of the RS, despite the "zero" time which is referred to the blank sample reported in **Table S2**. The heaviest effect was found for the [EtOH]: an increase in the ethanol concentration led to an increase in the lignin content in RS. Noteworthy, the xylan content was constant for each analyzed sample, not depending on the reactions conditions.
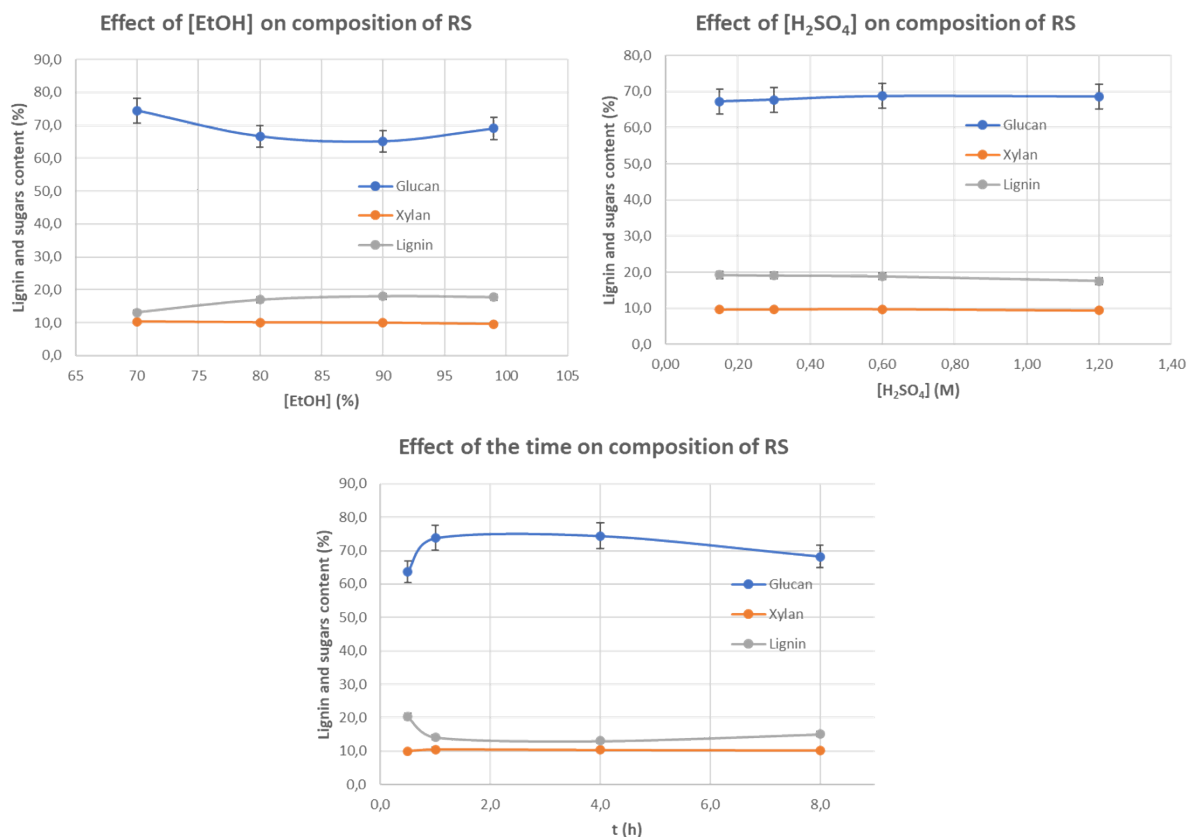
Figure S2. Effect of time (bottom), [EtOH] (top, left), and [H_2SO_4] (top, right) on the composition of the residual solids.

HSQC analysis of reactive extracted lignins (RELs)

Ethoxylation of phenolic OH. 2D HSQC analysis confirmed that phenolic OH groups were not etherified during reactive extraction. **Figure S3** shows that the characteristic signals are missing in the range δ_C/δ_H = 13.0-16.0/1.3-1.5 (red box) according to ChemDraw simulations.
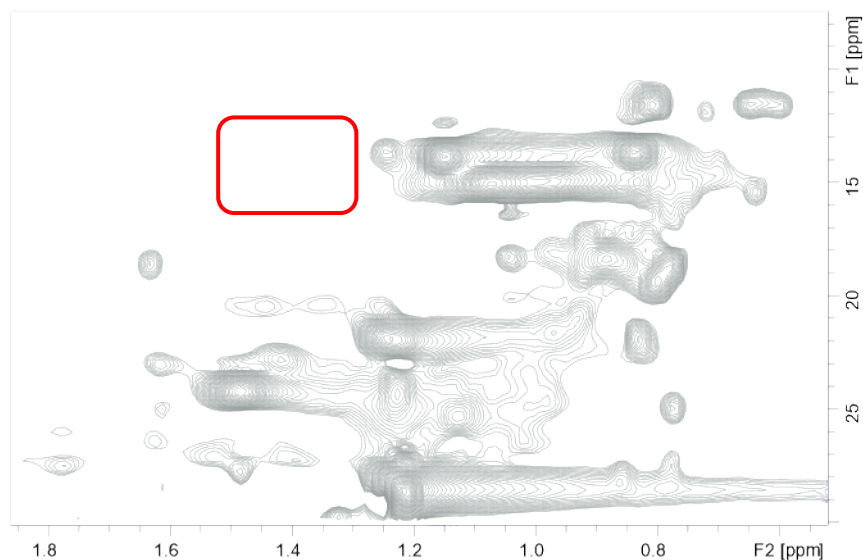
Figure S3. HSQC spectrum of REL with focus on the aliphatic region. No signal of phenolic EtO-groups was detected. Reaction conditions: t = 4h, [EtOH] = 99%, [H$_2$SO$_4$] = 0.15 M, T = reflux.

Other lignin units. Selected lignin samples were analyzed by 2D HSQC NMR. The analyzed samples together with the reaction conditions for their preparation are reported in **Table S2**. More details about the samples isolation are reported in the experimental section (main text).

Table S2. Labelling of reactive extracted lignin (RELs) selected samples.

| Label	Conditions[a]		
	[EtOH] (%)	[H$_2$SO$_4$] (M)	t (h)
Blank[b]	-	-	-
Sample 1[c]	99	-	4
Sample 2	99	1.2	0.5
Sample 3	99	0.15	4

[a]Conditions for the sample preparation. [b]Sample obtained from the direct extraction of hydrothermally treated solids with 70% aq. EtOH. [c]Sample isolated by rotary evaporation.

The HSQC analysis revealed that typical lignin moieties, like phenyl coumaran, resinol, vinyl-type linkages, and aromatic aldehydes, were present in RELs (**Figure S4**).
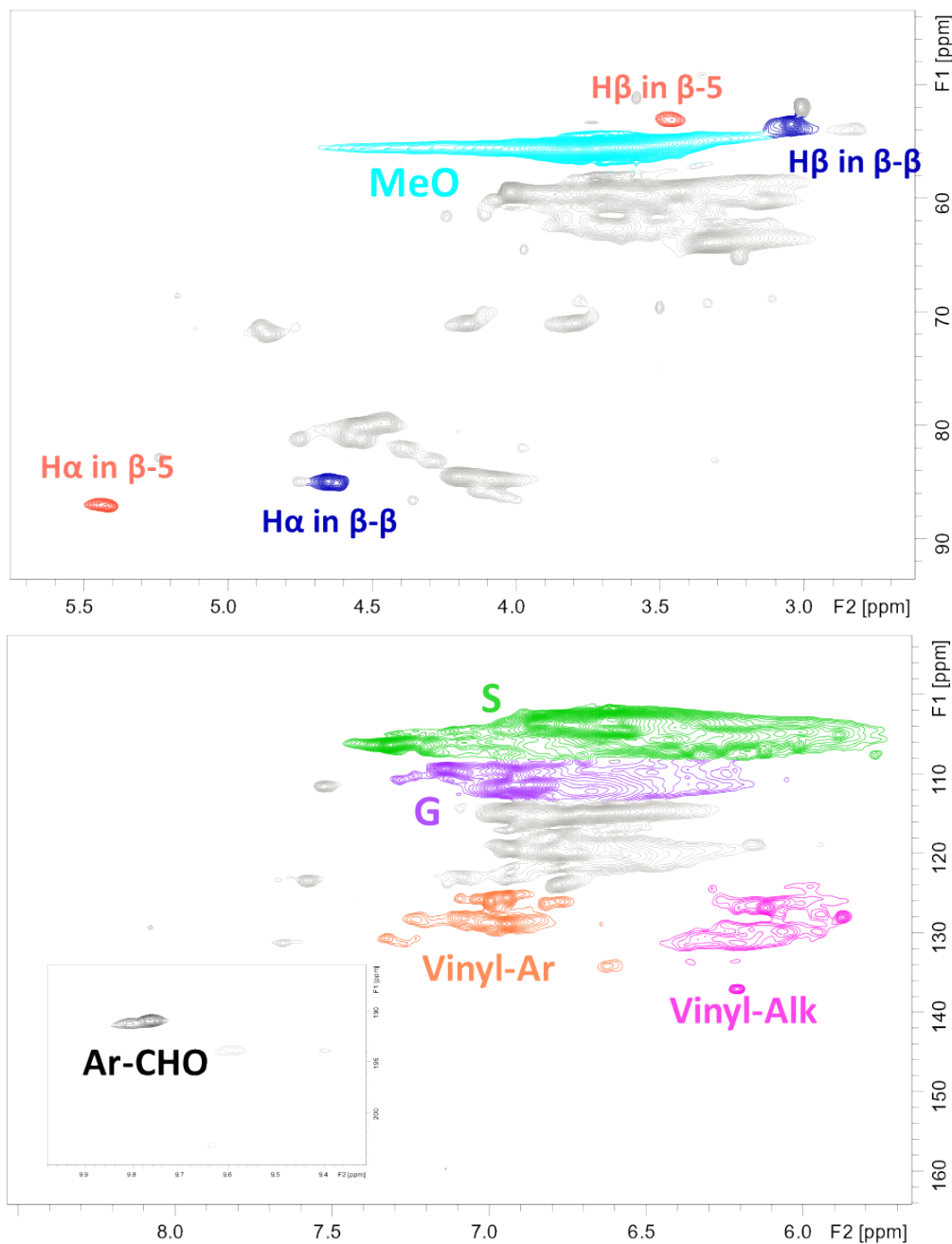
Figure S4. HSQC spectrum of REL. Top: oxygenated aliphatic region. Bottom: aromatic and aldehydes regions. Reaction conditions: t = 4h, [EtOH] = 99%, [H_2SO_4] = 0.15 M, T = reflux.

In addition, lignin carbohydrate complexes (LCCs) were detected as well (**Figure S5**).
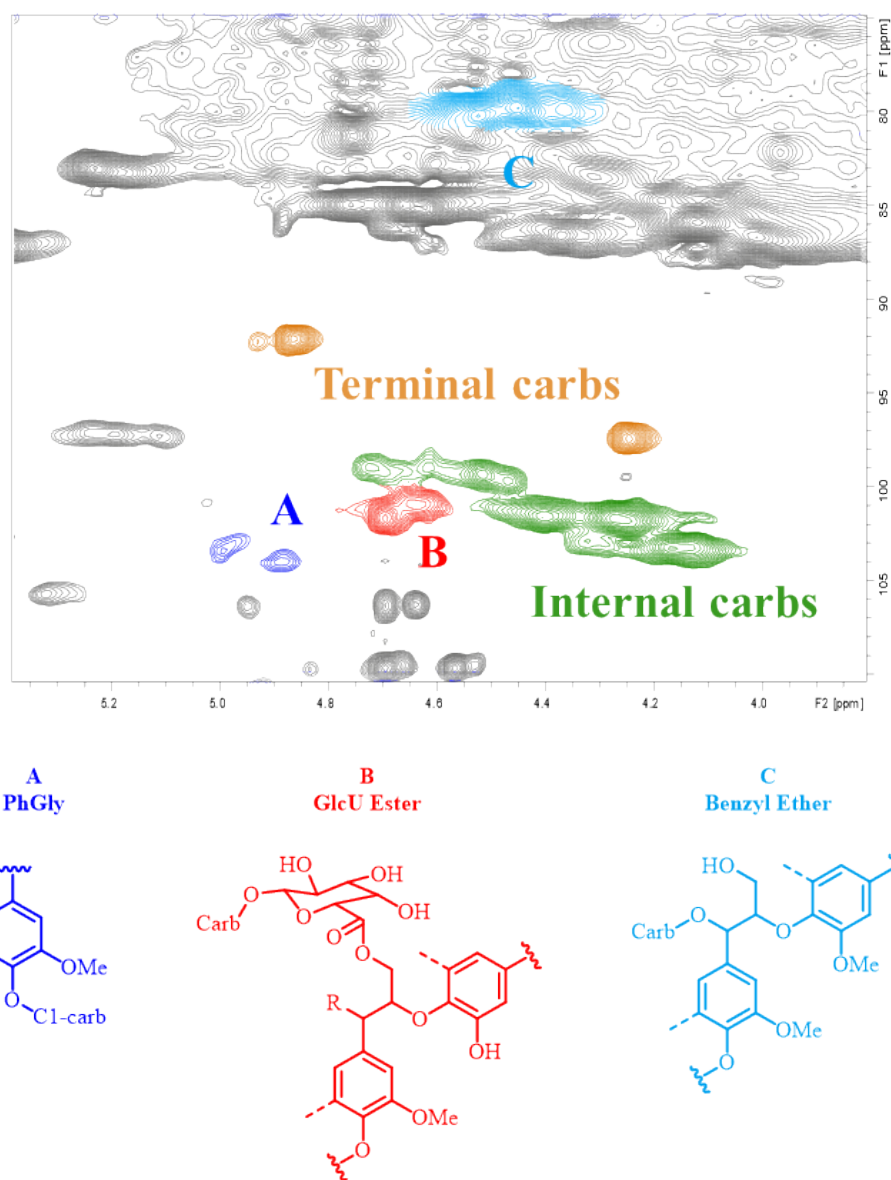

Figure S5. HSQC spectrum of REL with focus on the LCC linkages region. Reaction conditions: t = 4h, [EtOH] = 99%, [H$_2$SO$_4$] = 0.15 M, T = reflux.

Table S3 summarizes the results for the quantification of these moieties. The peaks were assigned based on previously reported data.[1] The average carbohydrate chain length (carbs DP) was calculated as the ratio between the terminal carbs and the total carbs.[2]

Table S3. Quantification of lignin moieties by HSQC NMR analysis expressed per 100 Ar. Only cross-peaks used for the quantification are listed. Details about the samples preparation are reported in **Table S1**.

Moiety	Integration range δ_C/δ_H (ppm)	Blank	Sample 1	Sample 2	Sample 3
S/G ratio		2.7	3.0	2.3	2.3
resinol	86.7–84.1/4.71–4.52	5.3	6.9	4.6	4.5
phenylcoumaran	89.4–85.4/5.72–5.29	2.7	3.5	3.0	3.1
Ar-CHO	192.5–189.2/10.06–9.55	0.8	3.5	0.5	0.5
GlcU Acid	98.1–96.2/5.34–5.04	0.3	0.4	0.2	0.2
GlcU Esters	101.5–100.0/4.72–4.59	0.2	0.3	0.3	0.5
PhGly	104.7–99.5/5.23–4.80	0.3	0.4	0.2	0.3
Term carb.		0.7	1.8	0.1	0.2
Internal carb.		5.1	6.7	3.6	5.1
Total carb.		5.8	8.5	3.7	5.3
carb. DP		8.3	4.7	37	26.5
Vinyl (total)	127.1–123.5/6.47–5.80 133.0–127.3/6.52–5.82 130.9–127.3/7.05–6.75 137.2–132.4/6.75–6.54 149.5–137.2/8.25–7.20	26.3	16.4	16.2	15.1

To conclude this section, a typical full HSQC spectrum of RELs is reported in **Figure S6**.
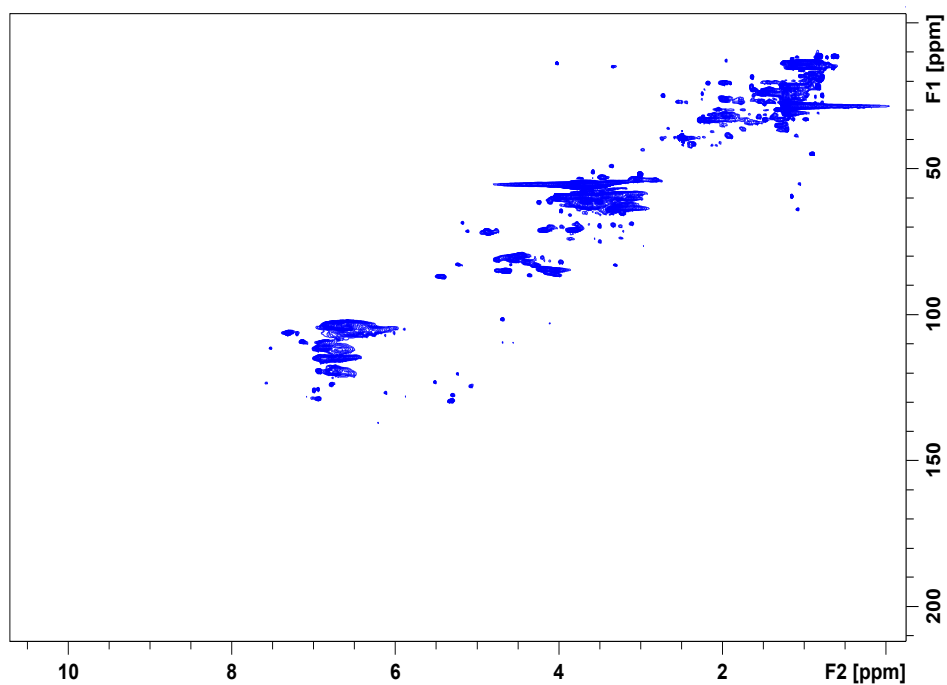
Figure S6. Tipycal HSQC spectrum of REL. Reaction conditions: t = 4h, [EtOH] = 99%, [H_2SO_4] = 0.15 M, T = reflux.

Quantitative ^{31}P NMR analysis of selected reactive extracted lignins (RELs)

The analysis was performed according to our previously reported optimized protocol.[3] Samples were selected as: i) blank sample obtained upon simple extraction with 70% EtOH; ii) sample prepared through REx under catalyst free conditions, with [EtOH] = 99%, and t = 4h; two samples with the highest DS obtained under iii) [EtOH] = 99%, t = 0.5h, [H$_2$SO$_4$] = 1.2 M and iv) [EtOH] = 99%, t = 4h, [H$_2$SO$_4$] = 0.15 M. Importantly, , the isolation of sample ii) could not be performed according to the extraction-precipitation used for the other samples, as no lignin precipitation occurred. Thus, its isolation was achieved by rotary evaporation. This is most likely related to a decrease in the lignin molecular weight over time during the extraction.[4]

Results have been corrected considering the EtO-groups and carbohydrates content in RELs as follows. The calculation for carbs is based on an "average" molar mass for carbohydrates derived from composition (%) by methanolysis and by assuming 2 OH/mol (monomer unit),[1,2] while the molar mass of EtO-groups has been considered as 44 g/mol, which is approximately the molecular weight of the CH$_3$CH$_2$- group.

$$x_{corr} = \frac{x \cdot [M_{Ar} + \dfrac{(carb \cdot M_{carb}) + (EtO \cdot M_{EtO})}{1000}]}{10}$$

Where:

x_{corr} = the OH/COOH content corrected for the presence of carbohydrates and ethoxy expressed per 100 Ar

x = the OH/COOH content in mmol/g

carb = carbohydrate content (per 100 Ar) determined by HSQC

EtO = ethoxy groups content (per 100 Ar) determined by HSQC

M_{Ar} = average molar mass of the aromatic ring (approx. 220 g/mol)

M_{carb} = average molar mass of the carbohydrate fraction based on carbohydrate composition (%) from methanolysis analysis (approx. 163 g/mol)[1,2]

M_{EtO} = molar mass of CH$_3$CH$_2$- group (approx. 44 g/mol)
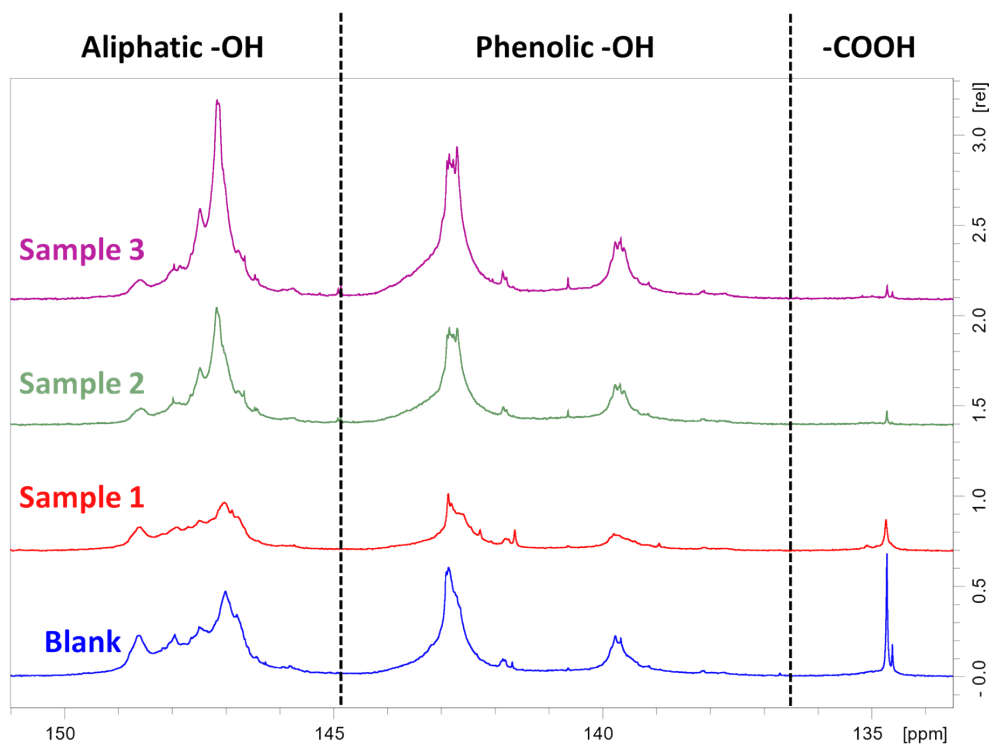
Figure S7. ^{31}P NMR spectra of the selected samples.

Quantitative ^{13}C NMR analysis of selected lignin samples

The quantitative ^{13}C NMR analysis has been performed according to our previously reported protocol[5] and the peaks assignment was based on previously reported data.[5–8] Considering the aim of the present paper, focus has been put in the quantification of ethoxy groups. The spectra are reported in **Figure S6** and the integration ranges of the moieties of interest are shown in **Table S5**. The amount of each specific moiety per 100 Ar has been calculated by setting the integral of the aromatic signals at δ_C in the range 100-160 equal to 600 (6 carbons each 1 aromatic → 600 carbons per 100 aromatics), implying that all the other units are expressed per 100 Ar. The moieties amount expressed in mol/g has been calculated as follows:

$$Moiety\ (mol/g) = \frac{{}^{I_x}/_{I_{IS}} \cdot mol_{IS}}{g_{sample}}$$

Where:

I_x is the integral value for the "x" moiety;

I_{IS} is the integral value for the internal standard set at 3;

mol_{IS} are the mol of internal standard put inside the NMR tube;

g_{sample} are the grams of lignin sample inside the NMR tube.

Table S5. Quantification of key lignin units by quantitative ^{13}C NMR analysis.

Entry	Moiety	Integration range (δ_c, ppm)	Sample 1[a]		Sample 2[b]	
			Per 100 Ar	mol/g	Per 100 Ar	mol/g
1	Aromatic	100.0-160.0	-		-	
2	MeO-groups	54.4-57.5	127.8	1.92	131.9	1.93
3	EtO-ethers	13.2-14.5	28.5[c]	0.33	21.9[c]	0.43
4	EtO-esters	14.5-18.2	12.3[c]	0.18	11.8[c]	0.20

[a]Conditions for the sample preparation: [EtOH] = 99%, [H$_2$SO$_4$] = 0.15 M, t = 4 h. [b]Conditions for the sample preparation: [EtOH] = 99%, [H$_2$SO$_4$] = 1.2 M, t = 0.5 h. [c]Data present also in the main text.
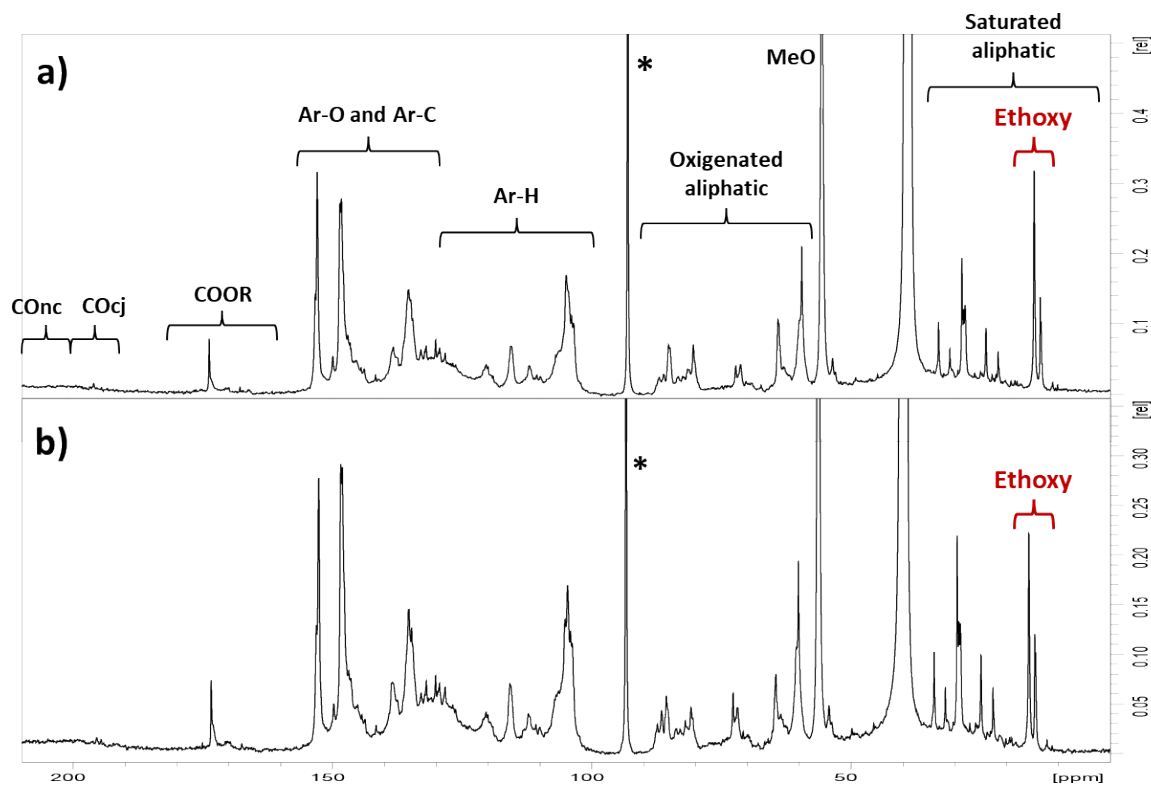
Figure S8. Quantitative ^{13}C NMR spectra of selected reactive extracted lignins (RELs). Reactive extraction conditions: a) [EtOH] = 99%, [H_2SO_4] = 0.15 M, t = 4h; b) [EtOH] = 99%, [H_2SO_4] = 1.2 M, t = 0.5h. Star indicates the signal of trioxane used as an internal standard.

Molar mass characterization of RELs

Table S6. Influence of the reaction time, ethanol, and sulfuric acid concentration on the statistical moments of the reactive extracted lignins.

Influence	Label	M_n [Da]	M_w [Da]	M_z [Da]	Đ
Time (0-8h)[b]	Blank[a]	1380	8110	47750	5.88
	REx-0.5h	2260	8850	37240	3.92
	REx-1h	2480	10750	48470	4.33
	REx-4h	2920	11780	59890	4.04
	REx-8h	2810	10660	43020	3.80
EtOH concentration (range 70-99%)[c]	Blank[a]	1380	8110	47750	5.88
	REx-EtOH70	2920	11780	59900	4.04
	REx-EtOH80	2640	7970	24260	3.02
	REx-EtOH90	2020	5810	15840	2.88
	REx-EtOH99	2860	4220	6230	1.48
H_2SO_4 concentration (range 0.15-1.2 M)[d]	REx (cat-free)	1070	3720	15510	3.77
	REx-0.15H	2580	4620	8640	1.79
	REx-0.30H	2490	3930	6330	1.58
	REx-0.60H	3080	4440	6350	1.44
	REx-1.2H	3020	4800	7720	1.59

[a]Sample obtained from the direct of extraction of S-500 with 70% aq. ethanol at room temperature. [b]Other conditions: $[H_2SO_4]$ = 0.15 M, [EtOH] = 70%. [c]Other conditions: $[H_2SO_4]$ = 0.15 M, t = 4h. [d]Other conditions: [EtOH] = 99%, t = 0.5h.
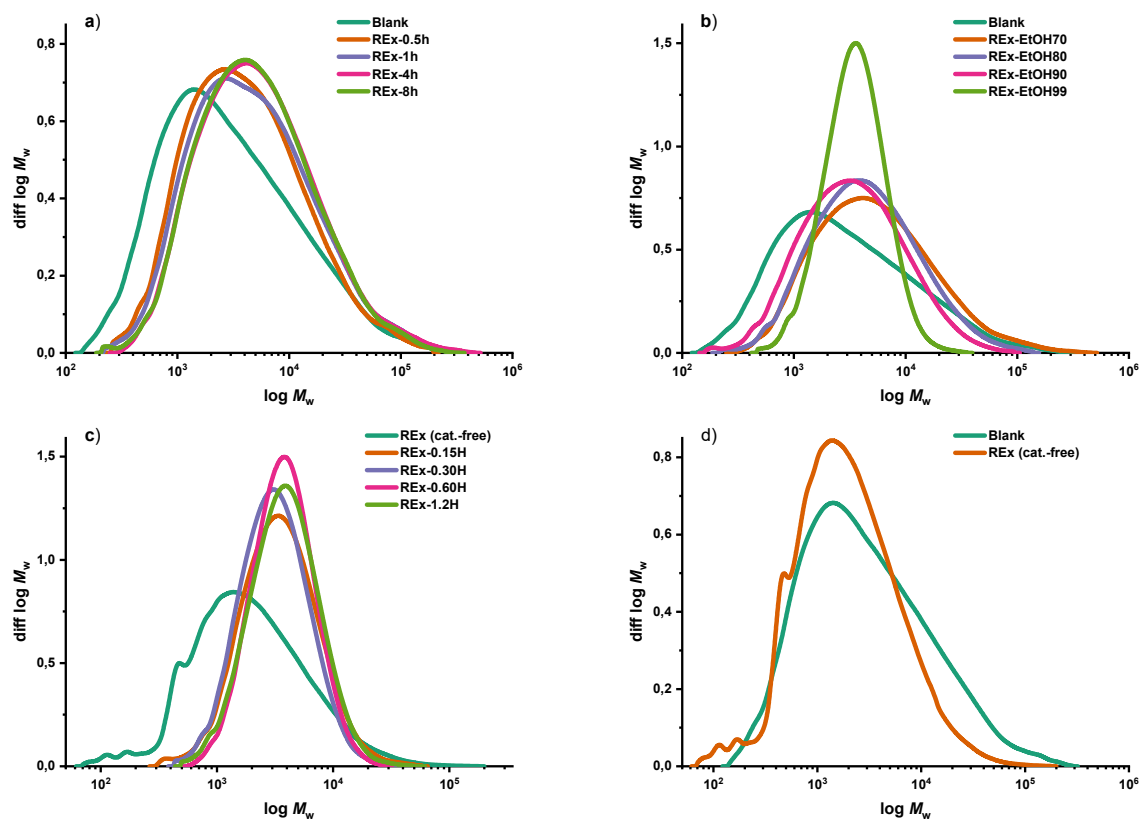
Figure S9. Influence of the reaction a) time; b) and d) ethanol concentration; c) sulfuric acid concentration on the molar mass distributions of the reactive extracted lignins.

Thermal properties of reactive extracted lignins
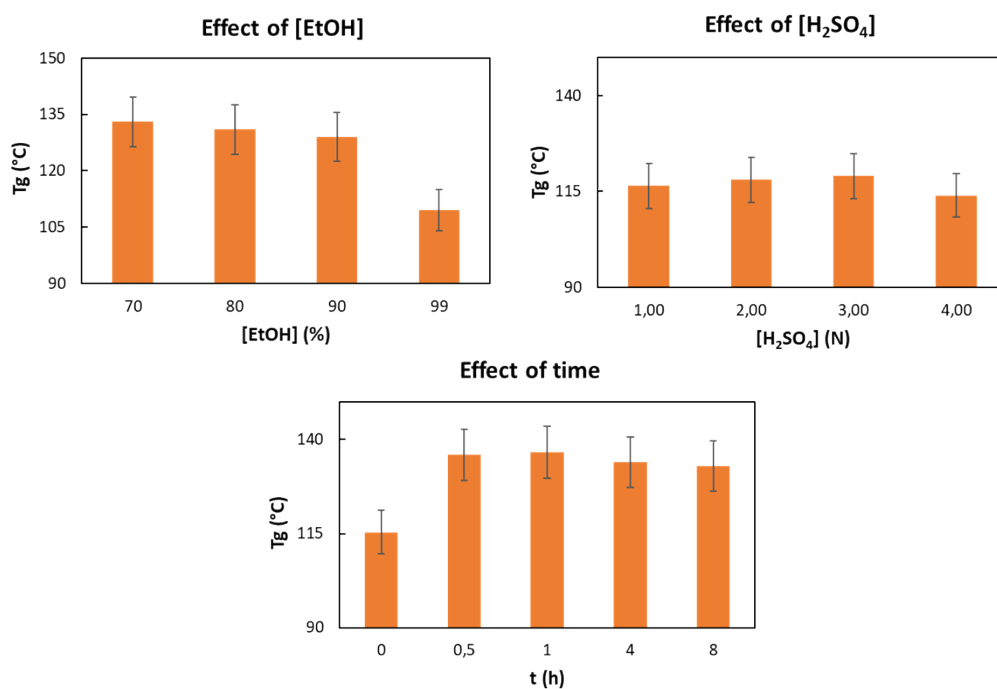
Figure S10. The effect of process variables on the thermal properties of RELs. Other conditions: $[H_2SO_4]$ = 0.15 M, t = 4h, T = reflux (top, left); [EtOH] = 99%, t = 0.5h, T = reflux (top, right); [EtOH] = 70%, $[H_2SO_4]$ = 0.15 M, T = reflux (bottom).
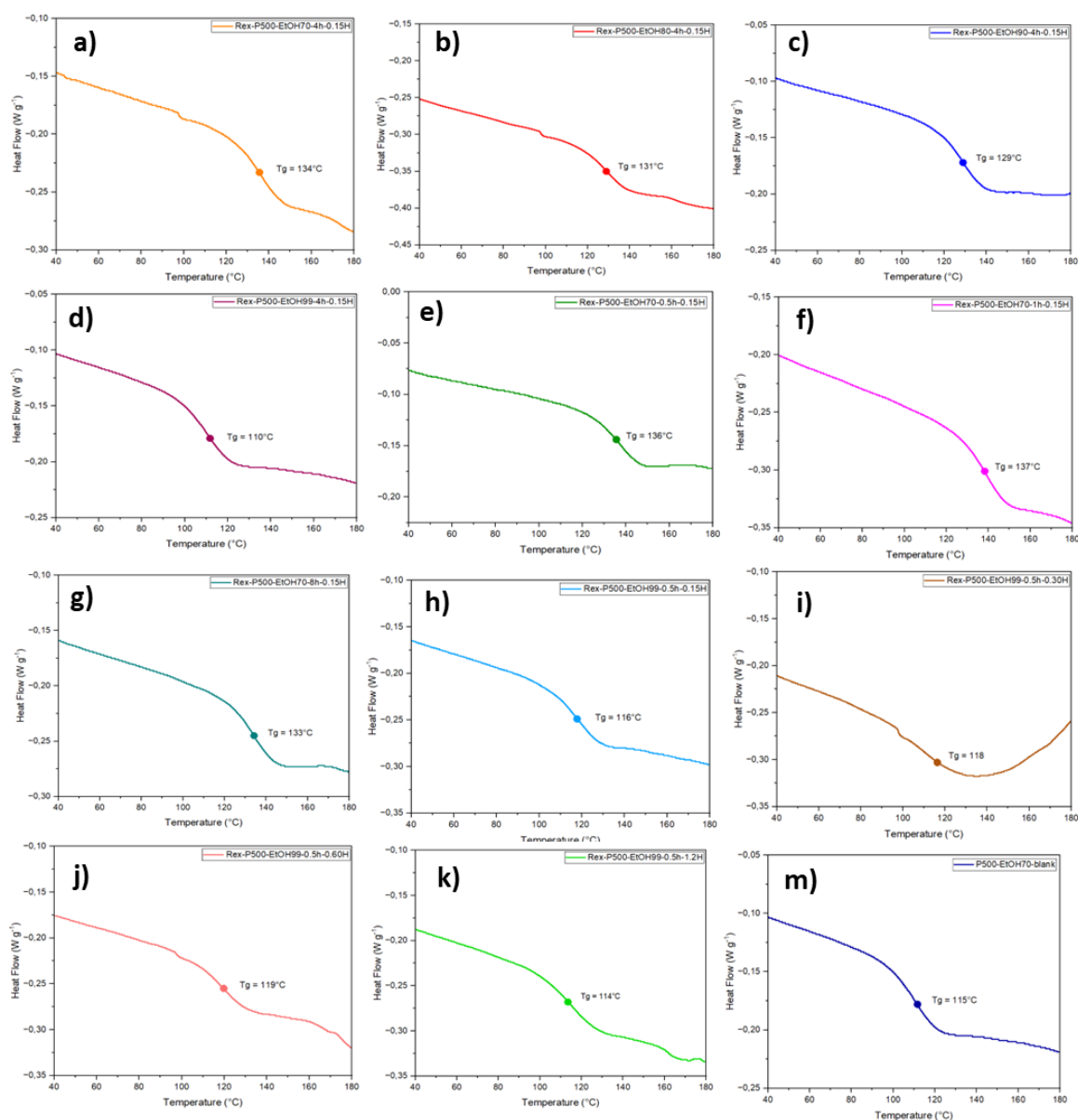
Figure S11. DSC curves of REx samples. T = reflux in all cases. Other conditions: **a)** $[H_2SO_4]$ = 0.15 M, t = 4h, [EtOH] = 70%; **b)** $[H_2SO_4]$ = 0.15 M, t = 4h, [EtOH] = 80%; **c)** $[H_2SO_4]$ = 0.15 M, t = 4h, [EtOH] = 90%; **d)** $[H_2SO_4]$ = 0.15 M, t = 4h, [EtOH] = 99%; **e)** $[H_2SO_4]$ = 0.15 M, t = 0.5h, [EtOH] = 70%; **f)** $[H_2SO_4]$ = 0.15 M, t = 1 h, [EtOH] = 70%; **g)** $[H_2SO_4]$ = 0.15 M, t = 8 h, [EtOH] = 70%; **h)** $[H_2SO_4]$ = 0.15 M, t = 0.5 h, [EtOH] = 99%; **i)** $[H_2SO_4]$ = 0.3 M, t = 0.5 h, [EtOH] = 99%; **j)** $[H_2SO_4]$ = 0.6 M, t = 0.5 h, [EtOH] = 99%; **k)** $[H_2SO_4]$ = 1.2 M, t = 0.5 h, [EtOH] = 99%; **m)** Sample obtained from the direct of extraction of S-500 with 70% aq. ethanol at room temperature.

**Ball milling and bleaching of the residual solids**

With the aim of demonstrating if ethoxylation occurred in the cellulose/hemicellulose and/or lignin fraction of the residual solids after reactive extraction, the residual lignin has been extracted. In order to extract the residual lignin, two approaches have been used: i) ball milling followed by solvent (dioxane) extraction and ii) bleaching with $NaClO_2$.

Ball milling. The residual solids were milled and extracted following previously reported procedures.[9] In detail, the residual solids (conditions: [EtOH] = 99%, [H_2SO_4] = 0.15 M, t = 4h) were subjected to 6 h milling in a planetary ball mill (Pulverisette 7, Fritsch, Germany) at 600 rpm with 9 ZrO_2 balls (diameter of 10 mm). The treated wood meal was collected and then extracted with aqueous dioxane (96%, v/v). The extracted milled solids were dried in a vacuum oven (T = 40 °C, p = 5 mbar), while the liquid dioxane mixture was rotary evaporated (T = 40 °C, p = 20 mbar) to obtain the residual milled lignin. Due to poor solubility, it was not possible to analyze the milled solids by solution state NMR,[10] while the milled lignin was analyzed by ^{1}H NMR. The spectrum is reported in **Figure S20**.

Bleaching. Bleaching was performed according to previously reported procedures. The residual solids (conditions: [EtOH] = 99%, [H_2SO_4] = 0.15 M, t = 4h) were treated with a mixture of 1.7% $NaClO_2$ and pH 4.5 acetate buffer solution for 24h. Once the reaction was complete, the bleached solids were filtered, exhaustively washed with deionized water and ethanol, and finally vacuum oven dried (T = 40 °C, p = 5 mbar). The bleached solids were analyzed by solution state NMR.[10]

**NMR analyses of the residual solids**

Solid state ^{13}C CP/MAS NMR:
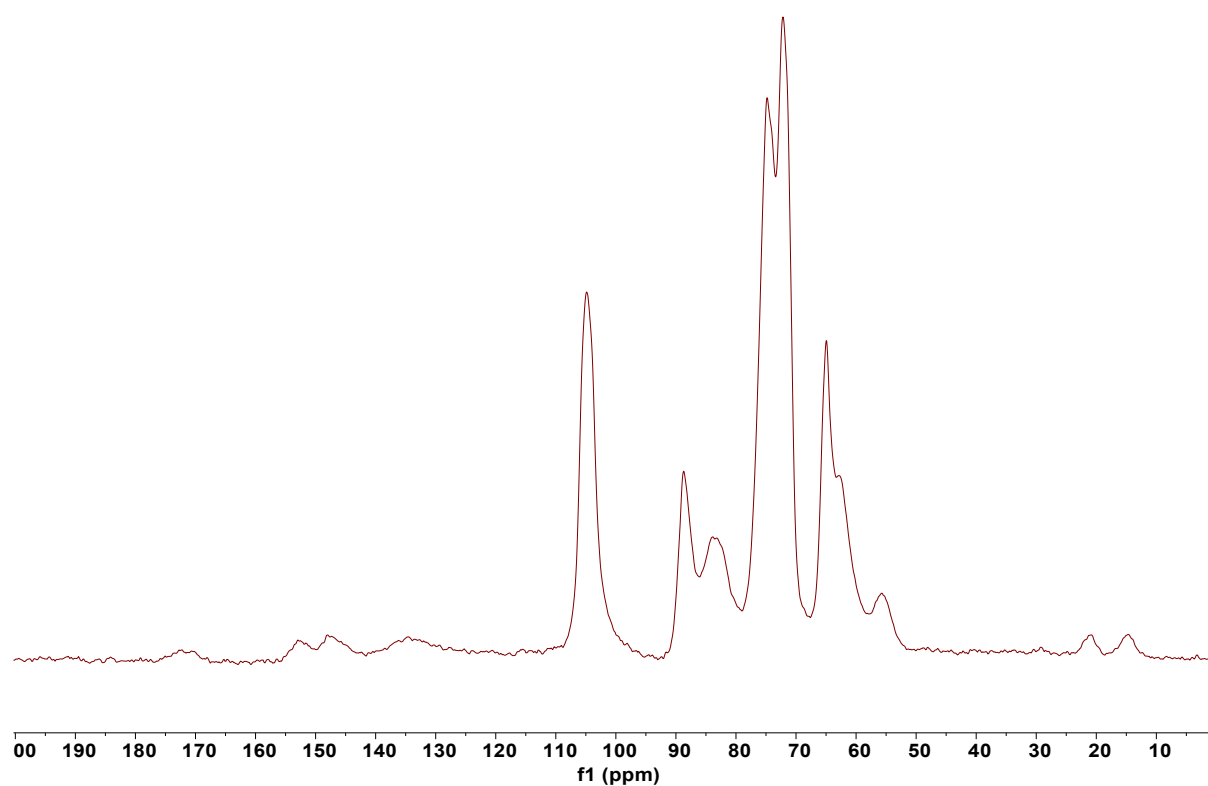
Figure S12. Solid state ^{13}C CP/MAS NMR spectrum (100 MHz) of the RS obtained after reactive extraction with the highest DS (Conditions: [EtOH] = 99%, [H_2SO_4] = 0.15 M, t = 4 h) before bleaching.
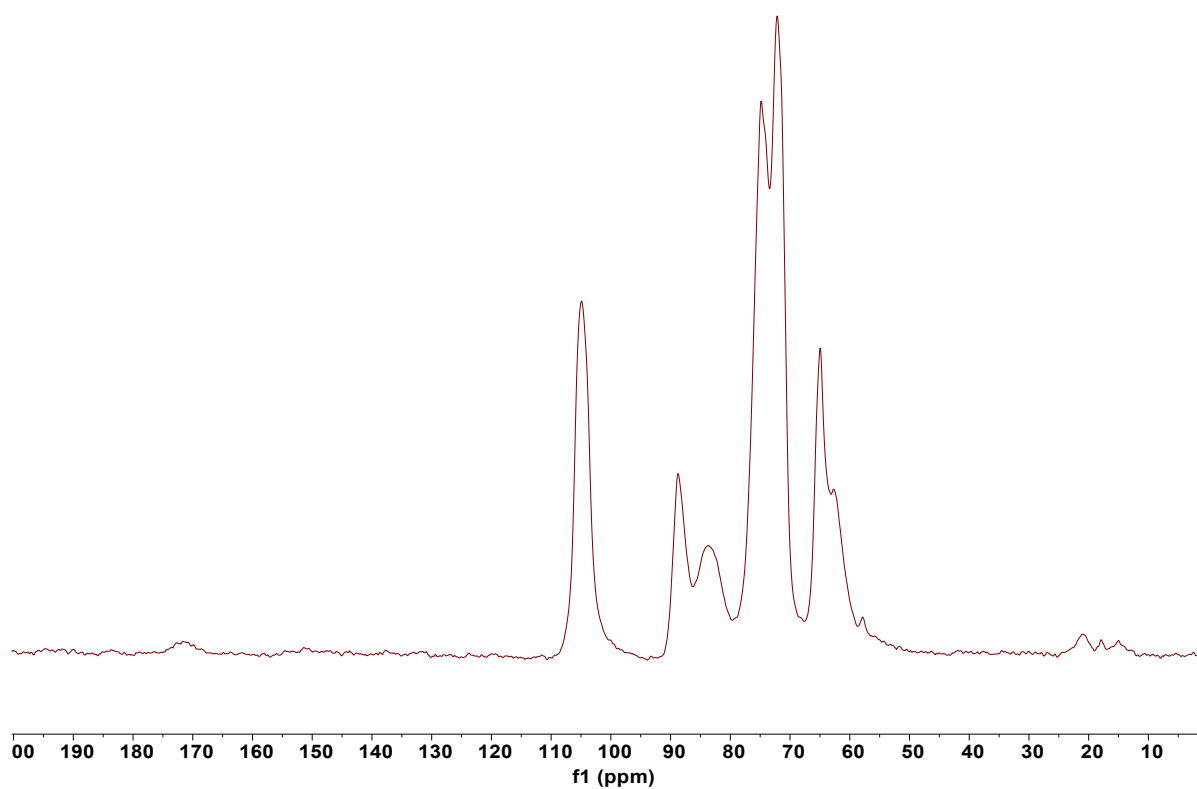
Figure S13. Solid state ^{13}C CP/MAS NMR spectrum (100 MHz) of the RS obtained after reactive extraction with the highest DS (Conditions: [EtOH] = 99%, [H$_2$SO$_4$] = 0.15 M, t = 4h) after bleaching.
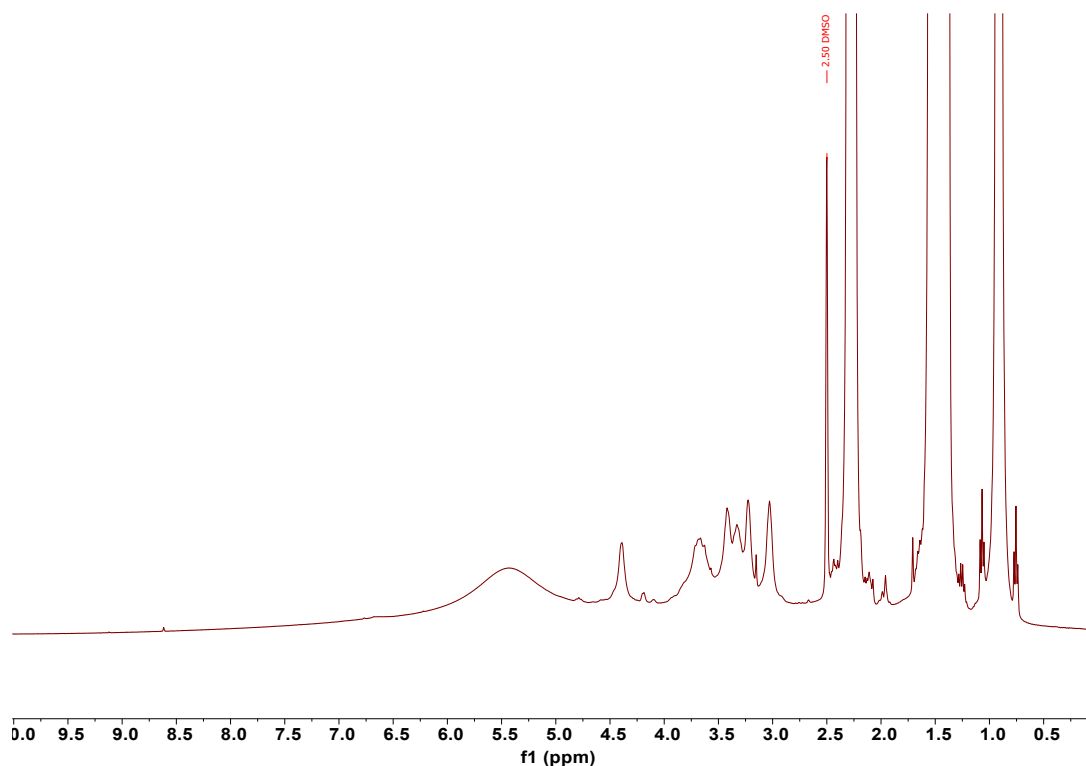
Solution state NMR:
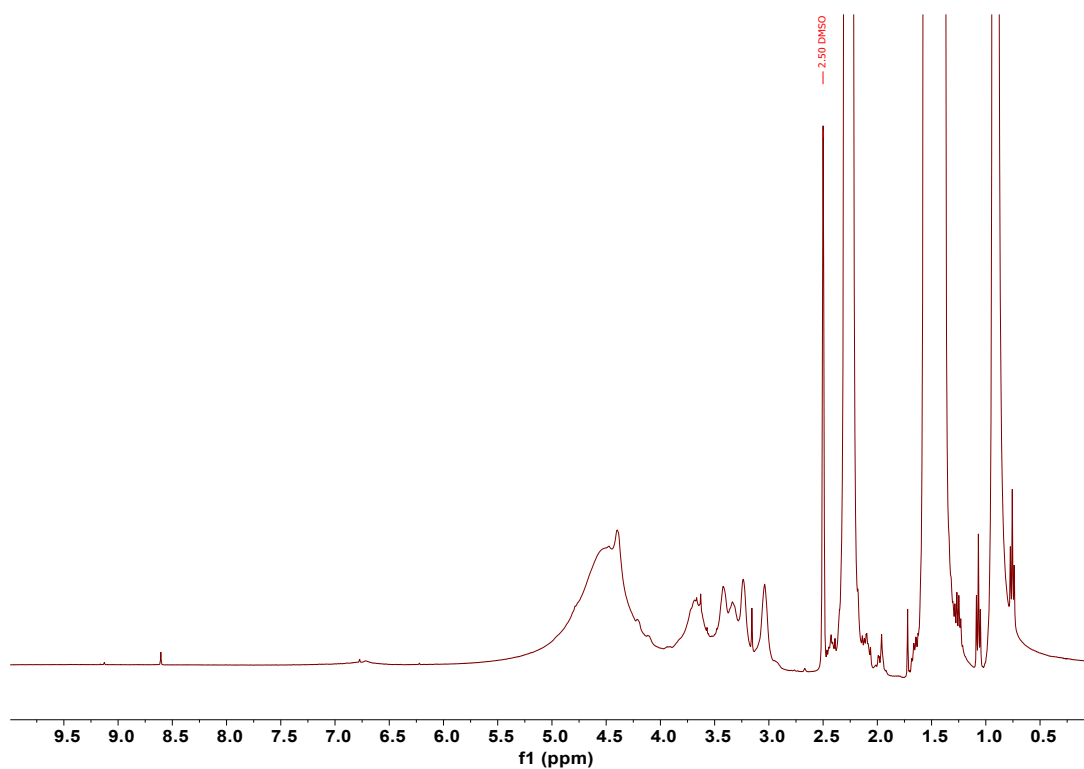
Figure S14. Quantitative ^{1}H NMR spectrum ([P$_{4444}$][OAc]:DMSO-d$_6$, 1:4 wt%; 400 MHz; 65°C) of the RS of the S-500 control sample (2.5 wt%). Some fractions remained insoluble.
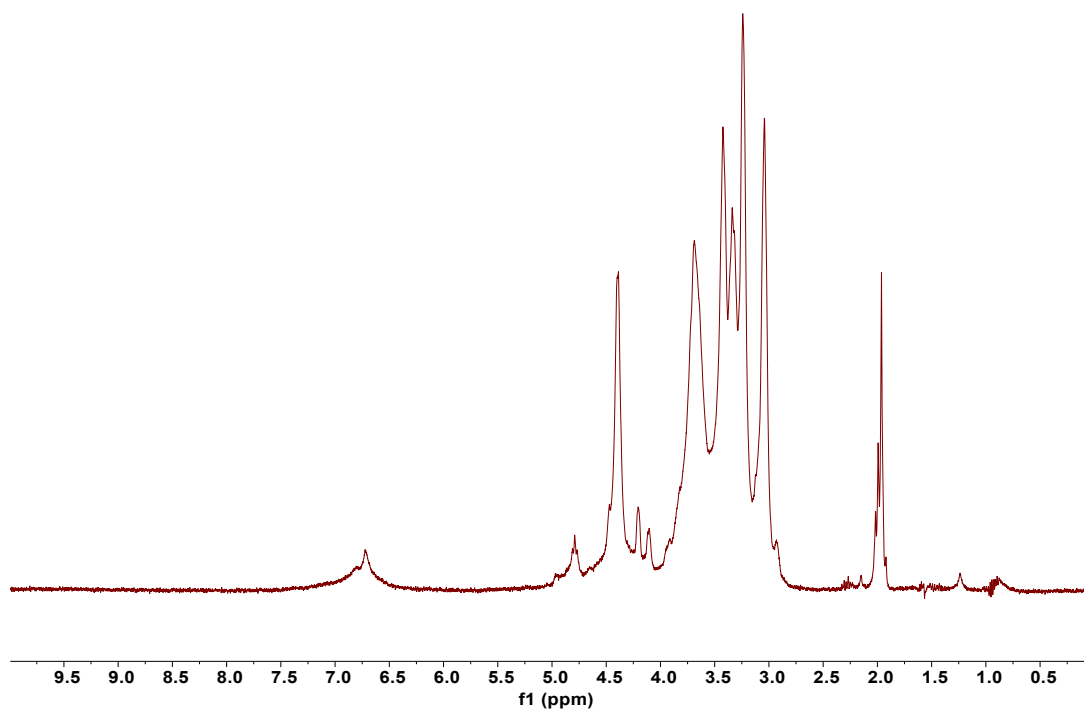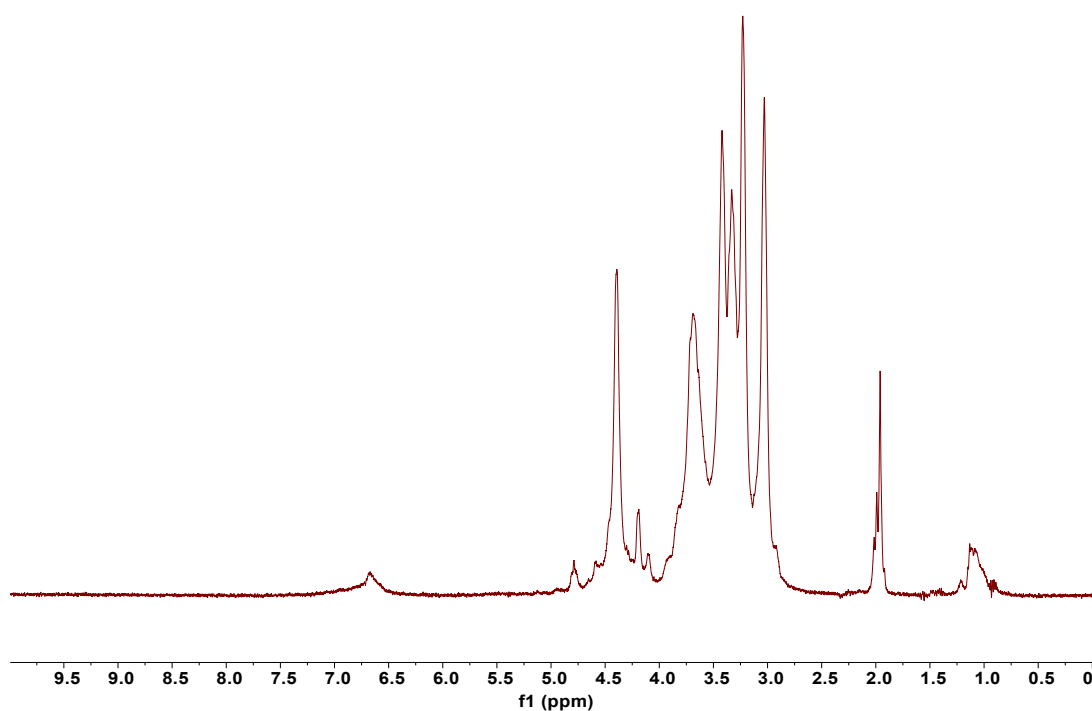

Figure S15. Diffusion edited ^{1}H NMR spectrum ([P$_{4444}$][OAc]:DMSO-d$_6$, 1:4 wt%; 400 MHz; 65°C) of the RS of the S-500 control sample (2.5 wt%). Some fractions remained insoluble.
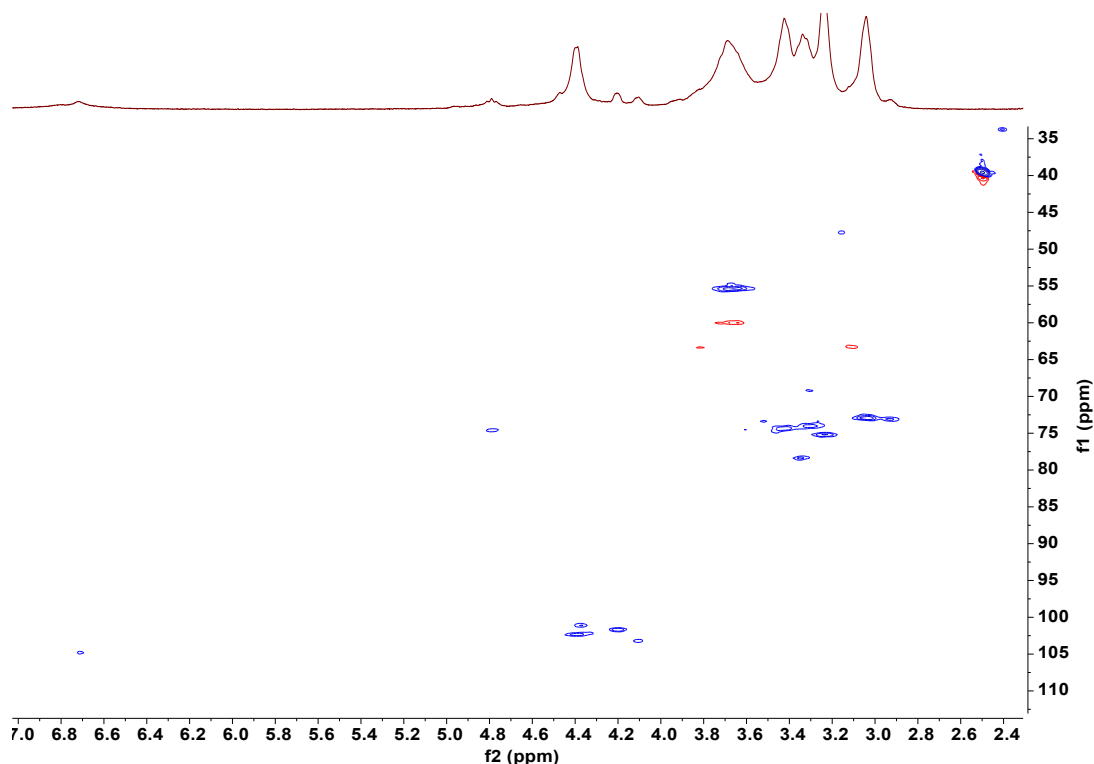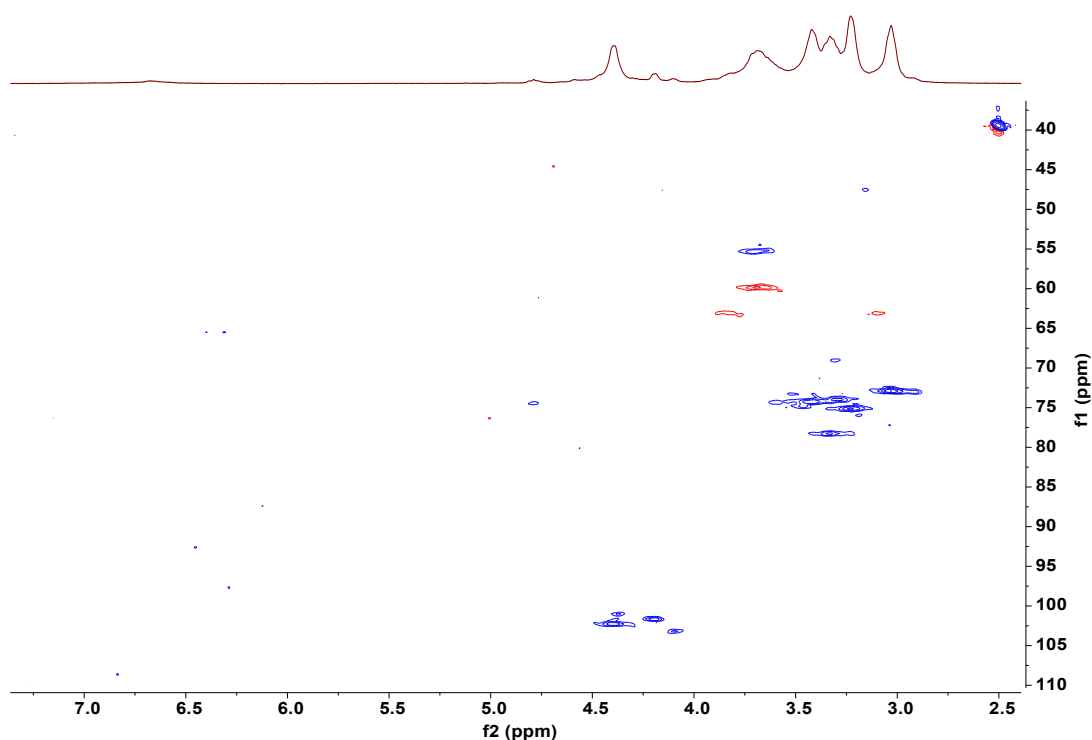
Figure S16. Multiplicity-edited HSQC spectrum ([P$_{4444}$][OAc]:DMSO-d$_6$, 1:4 wt%; 400 MHz ^{1}H frequency; 65°C) of the RS of the S-500 control sample (2.5 wt%).CH$_2$ resonances are shown in red, CH/CH$_3$ signals are shown in blue. Diffusion edited ^{1}H trace shown on top. Some fractions remained insoluble.
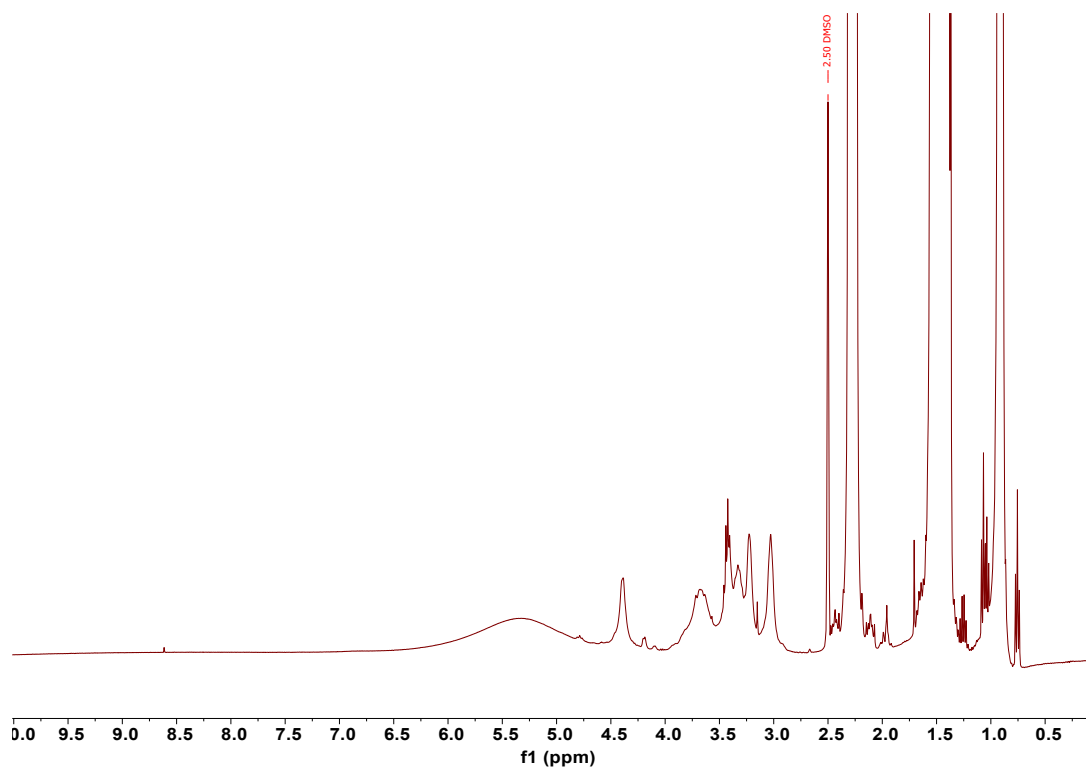

Figure S17. Quantitative ^{1}H NMR spectrum ([P$_{4444}$][OAc]:DMSO-d$_6$, 1:4 wt%; 400 MHz; 65°C) of the RS after reactive extraction (Conditions: [EtOH] = 99%, [H$_2$SO$_4$] = 0.15 M, t = 4 h) before bleaching. Some fractions remained insoluble.
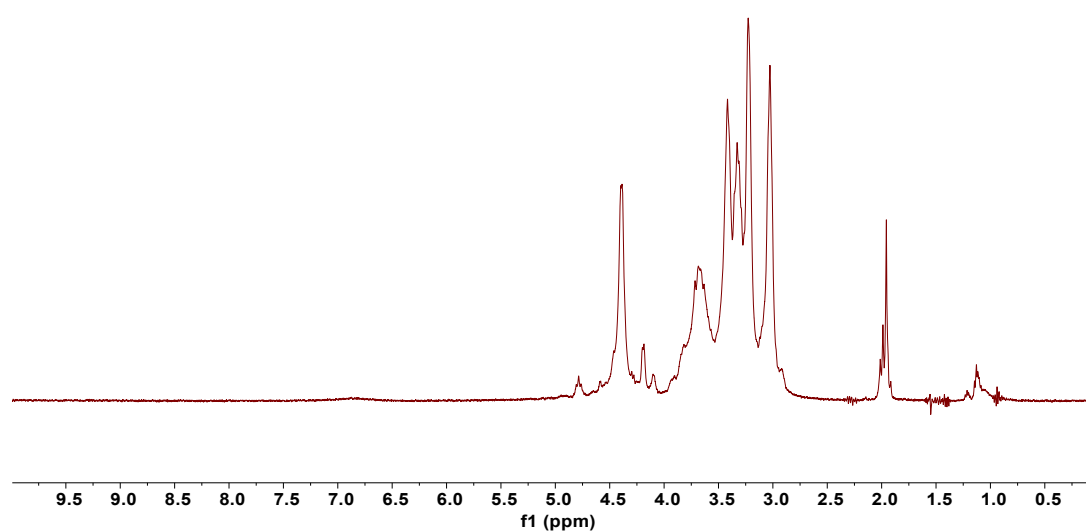
Figure S18. Diffusion edited ^{1}H NMR spectrum ([P$_{4444}$][OAc]:DMSO-d$_6$, 1:4 wt%; 400 MHz; 65°C) of the RS after reactive extraction (Conditions: [EtOH] = 99%, [H$_2$SO$_4$] = 0.15 M, t = 4 h) before bleaching. Some fractions remained insoluble.

Figure S19. Multiplicity-edited HSQC spectrum ([P$_{4444}$][OAc]:DMSO-d$_6$, 1:4 wt%; 400 MHz ^{1}H frequency; 65°C) of the RS after reactive extraction (Conditions: [EtOH] = 99%, [H$_2$SO$_4$] = 0.15 M, t = 4 h) before bleaching. CH$_2$ resonances are shown in red, CH/CH$_3$ signals are shown in blue. Diffusion edited ^{1}H trace shown on top. Some fractions remained insoluble.

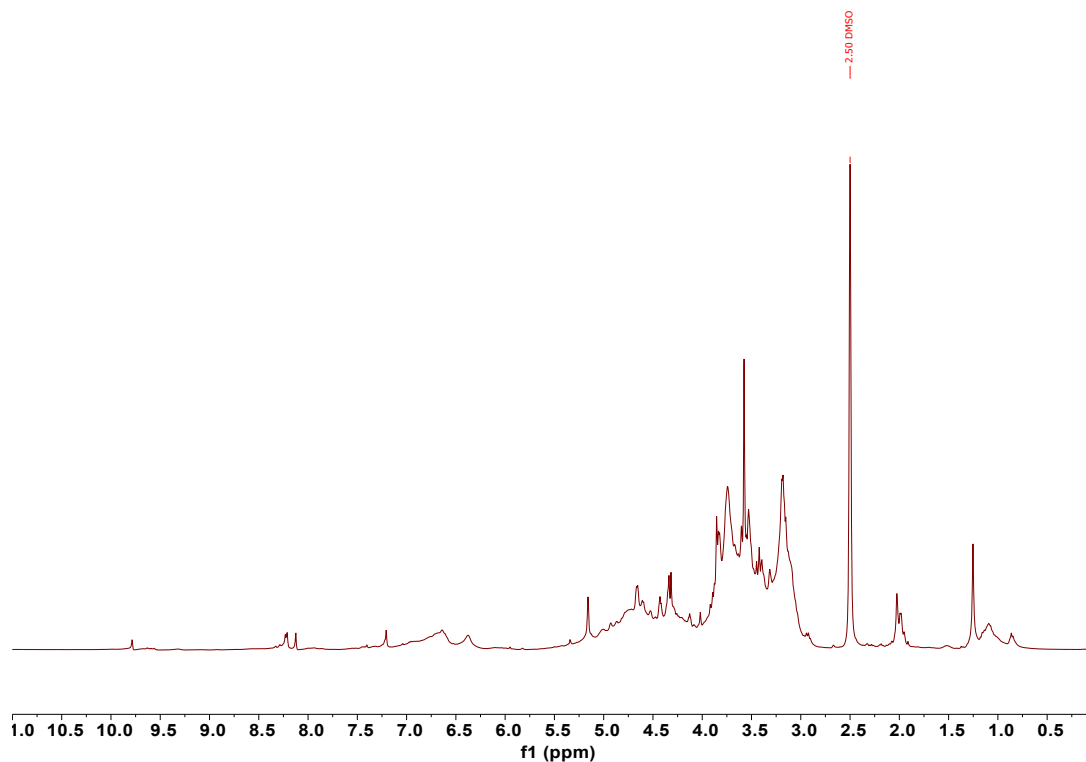**Figure S20.** Quantitative ^{1}H NMR spectrum ([P$_{4444}$][OAc]:DMSO-d$_6$, 1:4 wt%; 400 MHz; 65°C) of the RS after reactive extraction (Conditions: [EtOH] = 99%, [H$_2$SO$_4$] = 0.15 M, t = 4 h) after bleaching.

Figure S21. Diffusion edited ^{1}H NMR spectrum ([P$_{4444}$][OAc]:DMSO-d$_6$, 1:4 wt%; 400 MHz; 65°C) of the RS after reactive extraction (Conditions: [EtOH] = 99%, [H$_2$SO$_4$] = 0.15 M, t = 4 h) after bleaching.
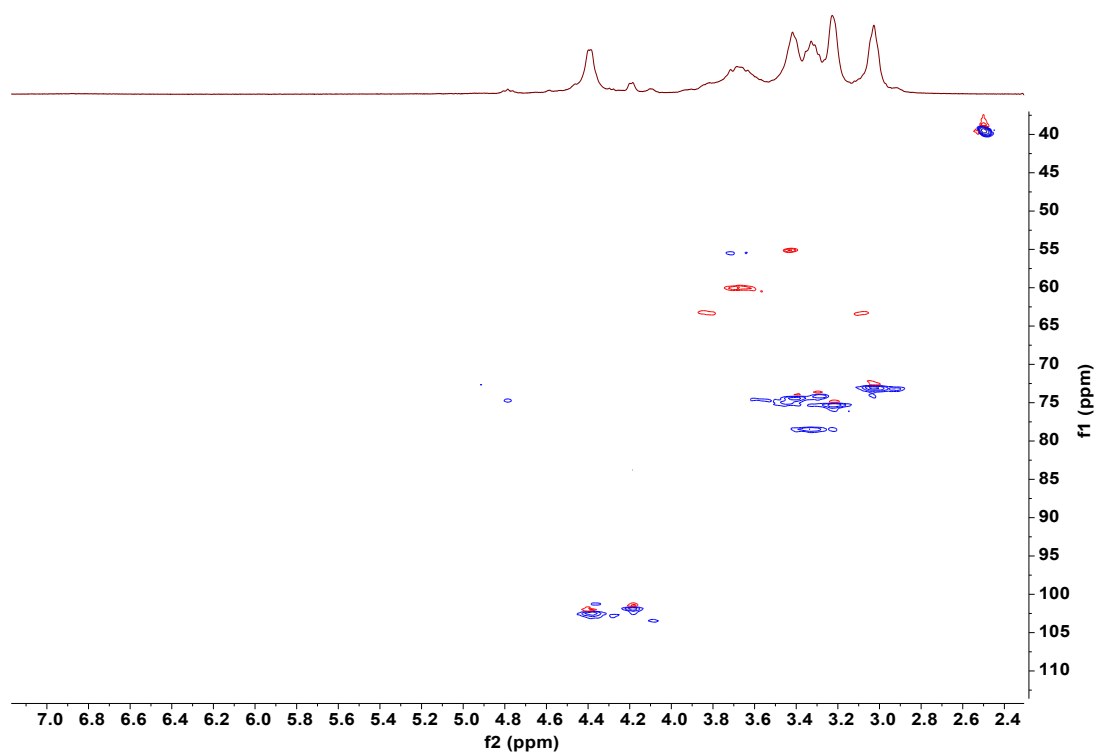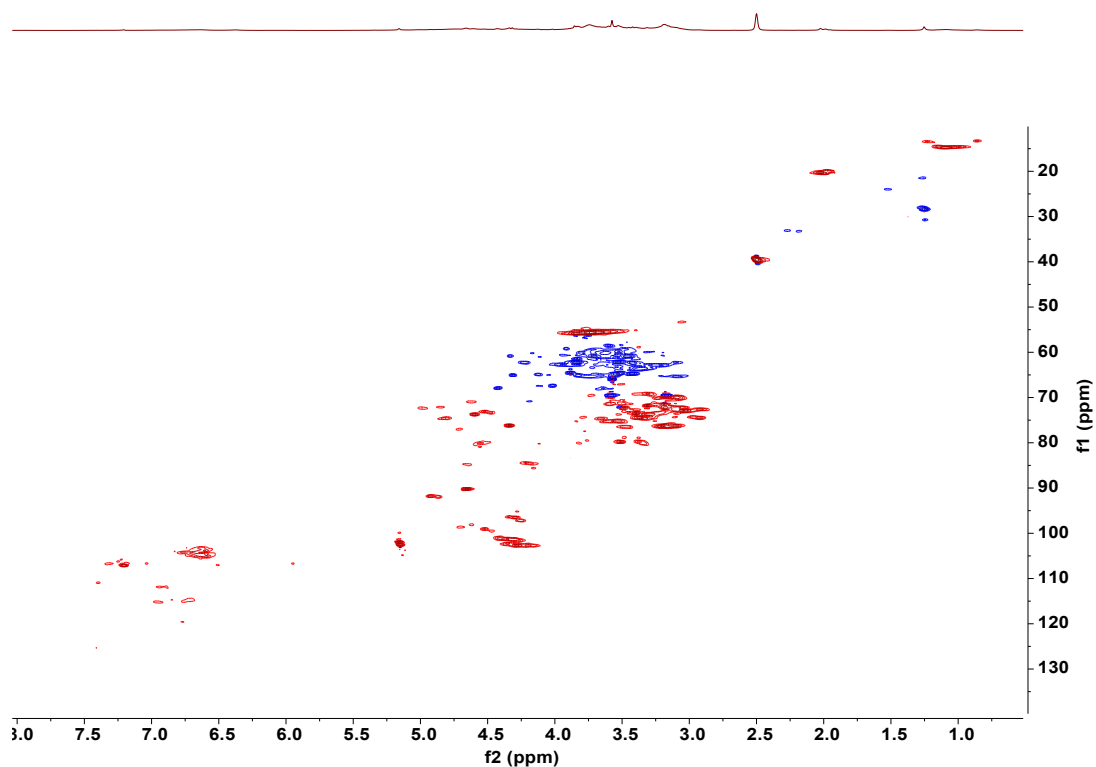
Figure S22. Multiplicity-edited HSQC spectrum ([P$_{4444}$][OAc]:DMSO-d$_6$, 1:4 wt%; 400 MHz ^{1}H frequency; 65°C) of the RS after reactive extraction (Conditions: [EtOH] = 99%, [H$_2$SO$_4$] = 0.15 M, t = 4 h) after bleaching.CH$_2$ resonances are shown in red, CH/CH$_3$ signals are shown in blue. Diffusion edited ^{1}H trace shown on top.

Figure S23. ^{1}H NMR spectrum (DMSO-d$_6$; 400 MHz; 65°C) of the lignins extracted by ball milling from the RS after reactive extraction (Conditions: [EtOH] = 99%, [H$_2$SO$_4$] = 0.15 M, t = 4h) after bleaching.

Figure S24. Multiplicity-edited HSQC spectrum (DMSO-d$_6$; 400 MHz ^{1}H frequency; 65°C) of the lignins extracted by ball milling from the RS after reactive extraction (Conditions: [EtOH] = 99%, [H$_2$SO$_4$] = 0.15 M, t = 4 h) after bleaching. CH$_2$ resonances are shown in blue, CH/CH$_3$ signals are shown in red. ^{1}H trace shown on top.

References:

1   P. Schlee, D. Tarasov, D. Rigo and M. Balakshin, *ChemSusChem*, 2023, **16**, e2023005.

2   D. Tarasov, P. Schlee, A. Pranovich, A. Moreno, L. Wang, D. Rigo, M. H. Sipponen, C. Xu, and M. Balakshin, *GreenChem*, 2022, **24**, 6639.

3   M. Balakshin and E. Capanema, *JWCT,* 2015, **35**, 220.

4   A. Tolbert, H. Akinosho and R. Khunsupat, A. K. Naskar, and A. J. Ragauskas, *Biofuels, Bioprod. Bioref.,* 2014, **8**, 836.

5   E. A. Capanema, M. Y. Balakshin and J. F. Kadla, *J Agric Food Chem*, 2004, **52**, 1850–1860.

6   M. Y. Balakshin, E. A. Capanema, R. B. Santos, H. M. Chang and H. Jameel, *Holzforschung*, 2016, **70**, 95–108.

7   M. Y. Balakshin and E. A. Capanema, *RSC Adv*, 2015, **5**, 87187–87199.

8   M. Y. Balakshin, E. A. Capanema, C. L. Chen and H. S. Gracz, *J Agric Food Chem*, 2003, **51**, 6116–6127.

9   M. Y. Balakshin, E. A. Capanema, R. B. Santos, H. M. Chang and H. Jameel, *Holzforschung*, 2016, **70**, 95–108.

10   L. Fliri, K. Heise, T. Koso, A. R. Todorov, D. Rico, S. Hietala, J. Fiskari, I. Kilpeläinen, M. Hummel and A. W. T. King, *Nat. Protoc.*, 2023, **18**, 2084.