Systematic Diffusion Ordered Spectroscopy for the Selective Determination of Molecular Weight in real Lignins and Fractions arising from Base-Catalyzed Depolymerization Reaction mixtures.

Alfonso Cornejo,† Íñigo García-Yoldi,† Irantzu Alegria-Dallo,‡ Rebeca Galilea-Gonzalo,† Karina Hablich,† David Sánchez,‡ Eduardo Otazu,‡ Ibai Funcia,‡ María J. Gil,† Víctor Martínez-Merino†*

† Institute for Advanced Materials (INAMAT)-Dpt. of Sciences, Campus de Arrosadia, Universidad Pública de Navarra, E31006 Pamplona, Spain

‡ National Renewable Energy Centre (CENER), Av. Ciudad de la Innovación 7, E31261 Sarriguren, Spain

Corresponding author: alfonso.cornejo@unavarra.es

KEYWORDS

Biorefinery lignin; Base catalyzed depolymerization; DOSY spectroscopy; Selective mass determination.

ABSTRACT

The valorization of biorefinery downstream lignin fractions is a key issue in order to increase the sustainability of Second Generation Biofuels. The development of reliable methodologies for the selective determination of apparent masses of the poly-(hydroxy)-aromatic ethers arising from lignin depolymerization reaction is crucial. Diffusion Ordered Spectroscopy (DOSY) has been tested to estimate the molecular weight in biorefinery downstream lignins and base catalyzed depolymerization reaction mixtures. Excellent correlation was found in the calibration of molecular weight and diffusion coefficients with standards. DOSY permitted the selective estimation of the apparent masses of different fractions in the lignin and in the depolymerization reaction mixtures, providing a more profound knowledge of the reaction mixture composition than traditional Size Exclusion Chromatography (SEC). Excellent correlations have been achieved in the estimation of the apparent masses of poly-(hydroxy)-aromatic ethers between SEC and DOSY. This permits a reliable estimation of the molecular weight of different fractions in the lignin and in the depolymerization product, which is essential for their further applications.

INTRODUCTION

In recent years, Second Generation Biofuels, produced in biorefineries from lignocellulosic biomass, have emerged as a promising alternative[1,2] to replace fossil fuels. This lignocellulosic material is widely available at relatively low cost, as it is produced from forestry and agricultural residues, without having to compete with human nutrition. Generally speaking, , the treatment of biomass raw materials in biorefineries is focused on the separation of its main constituent fractions and is comprised of two processes:[3] i) pretreatment to increase the accessibility of cellulose, hemicelluloses and lignins, ii) enzymatic hydrolysis of the cellulosic components to monomeric sugars that are fermented to ethanol or butanol. Downstream lignin fraction is mainly burned to produce energy,[2] as its efficient valorization has not been developed yet.

Lignin represents about 30% mass in softwoods, while in hardwoods this share is in the range of 20-25%.[4] Lignin is a complex polymer comprised of three monomeric phenols, coumaryl, sinapyl and coniferyl alcohols, which, through plant cell synthesis, become linked by β-O-4, β-5, β-β, or C-C bonds.

Lignin fraction is a promising alternative for replacing fossil-fuels derived chemicals. Indeed, lignin is deemed as the largest source of aromatic building blocks. The development of efficient lignin depolymerization methodologies can contribute, not only to the economic and environmental viability of biorefineries, but also to the development of new sustainable strategies for the synthesis of aromatics, as an alternative to petrochemistry industry. This has boosted the development of new depolymerization methodologies like hydrogenolysis, oxidation or hydrodeoxygenation.[5–7] However, most of these strategies are conceived for technical or paper industry derived lignins

Depolymerization of downstream lignin fraction from the biorefineries is even more challenging. The acidic thermochemical pretreatment modifies lignin structure by the formation of C-C bonds which renders these lignins even more recalcitrant.[6] Base catalyzed depolymerization (BCD) is a valuable option for lignin valorization,[8] as the basic medium improves lignin solubility *via* formation of phenolate anions.[9] Although BCD is a well-established methodology it is still far from being completely optimized. Up to 18 w% yield in low molecular weight compounds has been reported using basic zeolite as catalyst,[10] but monomer yields are below 10 w% in most cases. The origin of these

relative low yields is, partly, the high reaction temperature 240 ºC -300 ºC,[6,8,11–13] which boost repolymerization of reactive monomeric phenolate anions causing a decrease in monomer yields. Repolymerization can be prevented by the addition of capping agents such as boric acid or phenol.[14]

Besides monomeric phenols, a fraction of poly-(hydroxy)-aromatic ethers is produced on BCD. This fraction can find application in the preparation of polyurethanes or resins, among others.[15,16] Further application of these oligomeric fractions is strongly dependent on their average or number molecular weight, M_w or M_n respectively. M_w and M_n are determined by Size Exclusion Chromatography (SEC). Nevertheless, during the last years, it was found that SEC determination is very sensitive to the interaction between the stationary phase and the analytes[17] and can present large standard deviations even using refraction index detectors.[18] Furthermore, SEC is unable to discriminate between poly-(hydroxy)-aromatic ethers and other fractions. Downstream biorefinery lignin present impurities, such as carbohydrates, whose degradation products cannot be discriminated by SEC leading to misinterpretation of SEC chromatograms and, particularly, M_w and M_n.[19]

Nuclear Magnetic Resonance (NMR) spectroscopy has become a powerful tool in the characterization of lignin and bio-oils. Many studies include systematic Heteronuclear Single Quantum Coherence (HSQC-NMR)[20–26] or even ^{31}P NMR in lignin characterization.[20,23,27,28] Diffusion Ordered Spectroscopy (DOSY-NMR), is a pseudo 2D experiment that correlates chemical shifts and the diffusion coefficient, D, at a given temperature. The Stokes-Einstein equation relates D coefficient to the hydrodynamic radium of the diffusate.[29] Therefore, DOSY-NMR can provide similar information than SEC.[20] Consequently, DOSY allows correlating chemical shift and molecular weight and can be envisaged as a NMR-chromatographical tool.[30] DOSY spectroscopy has been successfully used in the analysis of proteins,[31] in organometallic chemistry[32] and even in the study of the surface chemistry of nanoparticles.[33] Some studies already addressed the potential of DOSY showing the separation of phenol, guaiacol and 2,6-dimethoxyphenol from a mixture prepared from commercial standards.[30] More recently, good correlations of the apparent masses measured by SEC and DOSY were obtained for lignin fractions with narrow molecular weight distributions.[34,35] However, few studies

were published using conventional ^{1}H-DOSY or ^{31}P-DOSY[23,36,37] to estimate the molecular weight in depolymerization reaction mixtures using DOSY.

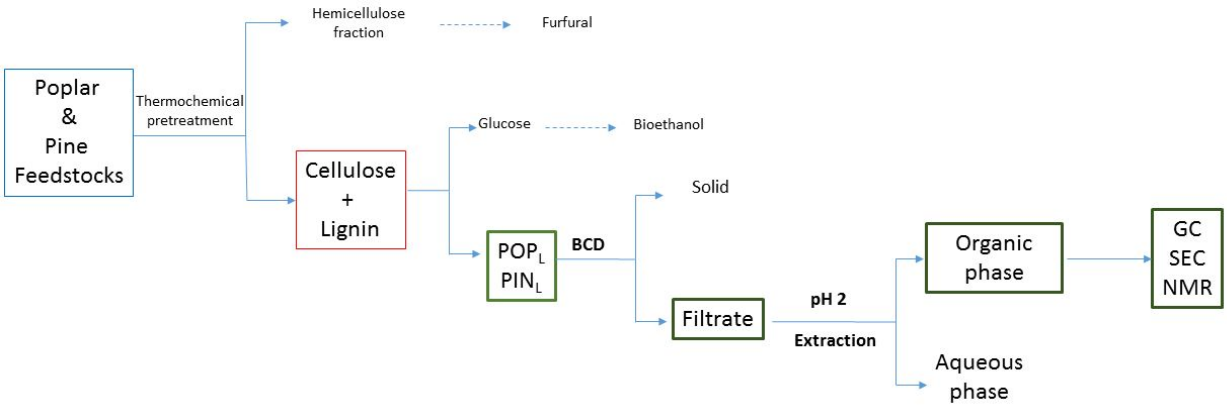
In this paper, we show the results of BCD for lignins obtained after thermochemical pretreatment of poplar and pine wood followed by polysaccharide deconstruction *via* enzymatic hydrolysis. The results of the calibration of diffusion coefficients and the molecular masses of two different families of standards as a function of their interaction with the solvent are presented. Finally, DOSY is systematically applied in the selective determination of the apparent molecular masses of the different fractions that are present in lignin and BCD reaction mixtures.

EXPERIMENTAL SECTION

All the chemicals were purchased from Fisher Chemicals and Sigma-Aldrich and used as received. 2-Phenoxy-1-phenylethanol (**PE**) was obtained as described elsewhere.[38] Biomass raw materials were processed at CENER facilities of Biorefinery and Bioenergy Centre (BIO2C), located in Aoiz (Spain). Poplar (*Populus sp.)* and pine (*Pinus radiata)* feedstocks were processed as described in a previous study[39] to obtain the starting solids $\mathbf{POP_L}$ and $\mathbf{PIN_L}$ respectively. For the samples used in this study, thermochemical

pretreatment was carried out using sulfuric acid at 4 w% against the dry weight of the total solid content at 192 ºC with a residence time of 5 minutes. Lignin content in $\mathbf{POP_L}$ and $\mathbf{PIN_L}$ was determined according to NREL/TP 510-42618. The ratio of syringyl (**S**) and guaiacyl (**G**) units, S/G ratio, was determined as described in the literature.[40,41]

Depolymerization reactions were carried out in an Autoclave Engineers equipped with PID controller, using a total volume of 50 mL in a 100 mL stainless steel vessel. Reaction time started on reaching the target temperature. The resulting solutions were filtered and the liquid was acidified to pH 2 upon addition of 1M HCl. 25 mL of a solution of bromobenzene (0.0015 M) in ethyl acetate, and additional ethyl acetate when necessary for the extraction, were then added and the organic phase was separated. 1.0 mL of the organic phase was taken for the quantification of monomeric phenols by GC-FID. The rest was evaporated to dryness and the yield of the bio-oil was gravimetrically calculated (Scheme 1).



Scheme 1. Overall procedure used in this study for obtaining BCD reaction mixtures.

The yield of monomeric phenols was quantified by GC-FID on an Agilent 6890 equipped with a capillary column HP-5MS [(5%-phenyl)-methylpolysiloxane, 60 m × 0.32 mm] with helium as carrier gas. The temperature program started at 50 °C and followed by heating to 120 ºC, 280 ºC, and 300 °C at 10 °C min^{-1}. During the program, temperature was held at 120 °C for 5 min, at 280 °C (8 min), and at 300 °C (2 min). Monomer yields gather the yield of 27 compounds, whose response factors in FID were calibrated using bromobenzene as internal standard. The monomers were grouped in phenols, guaiacols, syringols, cresols and vanillins (Table S1). Weight yield, w%, is referred to the weight of known monomers in the actual amount of lignin in the starting solid (**POP_L** or **PIN_L**). The

identity of the monomeric phenols was confirmed by GC-MS on a Shimadzu GC-2010 equipped with Shimadzu QP-2010 detector with the same analysis conditions.

Size exclusion chromatography (SEC) measurements were made on an Agilent 1100 equipped with RID detector. **PIN$_L$** and **POP$_L$** were analyzed using two coupled PolarGel-M columns (300 x 7.5 mm) and PolarGel-M guard (50 x 7.5 mm) as stationary phase and 0.1 % LiBr in DMF at 40 ºC at 0.7 mL·min $^{-1}$ as eluent. Base catalyzed depolymerization (BCD) reaction mixtures were analyzed using Coupled HR-5 and HR-1 Styragel columns (Waters) at 30 ºC as stationary phase and tetrahydrofuran as mobile phase at 1 mL·min^{-1} flow.

NMR measurements were made at 300 K on a Bruker Ascend III spectrometer equipped with a PABBO 5 probe, at 400 MHz and 101 MHz for ^{1}H and ^{13}C respectively and were processed using Bruker Topspin 3.2 software. NMR samples were prepared at 1 % w/v in DMSO-d_6 and referenced using residual signal at 2.50 ppm and 39.52 ppm for ^{1}H and ^{13}C measurements.[42] **PIN$_L$** and **POP$_L$** were acetylated for NMR measurements as described elsewhere.[20] ^{1}H experiments were run using the *zg30* pulse program at 16 scans. Two-dimensional ^{1}H- ^{13}C correlation was carried out using HSQC and HSQC-

TOCSY that run *hsqcetgpsi2* and *hsqcdietgpsisp.2* pulse program in an echo-antiecho acquisition mode with D1 = 1.48 s and D1= 2.0 s respectively. DOSY was run using *stebpgp1s* pulse program in QF acquisition mode. Diffusion delay (*d20*) in *stebpgp1s* was optimized using residual DMSO signal keeping gradient pulse length (*p30*) constant at 1000 μs, resulting in 160-170 ms. Each pseudo-2D experiment consisted on a series of 16 spectra. DOSY experiments were made in DMSO-d_6 at 300 K and at fixed low concentrations, 1 % w/v. The accuracy of the gradient was checked by the determination of the diffusion coefficient of residual DMSO in DMSO-d_6. DOSY analysis were processed using TOPSPIN 3.2 software from Bruker. Once F2 were phased, automatic baseline correction was run using a 5^{th} grade polynomial function. Using the T1/T2 relaxation module installed in the module permitted FID for the first spectrum (2 % gradient) to be extracted and to perform manual integration. The integration regions were exported to the relaxation module where decay values were fitted by area using the *vargrad* preinstalled function using a 5.35 G/mm as gradient calibration constant. Graphical processing was run using 'Dynamic Center v. 2.6.1' software from Bruker. Unless otherwise stated, Stejskal-Tanner equation was fitted using the intensity obtained after the peak peaking of

the first spectrum corresponding to 2% gradient. NOESY experiments were carried out using the *noesyphrv* sequence with random mixing time (285 ms – 315 ms) to prevent coupling artifacts. ^{31}P measures were made using a proton decoupled experiment running the *zgpg30* pulse program. Derivatization prior to ^{31}P measurements was carried out following the protocol described by Pu.[28]

RESULTS AND DISCUSSION

Base Catalyzed Depolymerization (BCD)

Two different feedstocks, pine (*Pinus radiata)* and poplar (*Populus sp.)* were chosen as soft and hard-wood models respectively, given their natural abundance in Navarre forestry. Feedstock chips were subjected to acidic thermochemical pretreatment followed by enzymatic hydrolysis to produce lignin rich solids, which were then named as **PIN$_L$** and **POP$_L$** for pine and poplar downstream lignin respectively.[39] Lignin content in these solids was 69 w% and 51 w% respectively (Table 1). The rest of the solids corresponded mainly to saccharides, 24 w% and 39 w% respectively. Lignin contents in these solids are lower than in technical lignins[12] but within the range of biorefinery downstream lignin fractions.[8] The syringyl/guaiacyl ratio (S/G) was 1.5 for **POP$_L$**, and 0.1 for **PIN$_L$**. This S/G value for

PIN_L is not consistent with most of the literature which reveals that no syringyl units are present in pine wood.[5,6] Coniferyl alcohols constitute approximately 90% of softwood lignin, although many exceptions are known.[43] Indeed, it has been stated that S/G ratio can be as high as 0.25 in *Pinus Pinaster* bark.[44] This causes an increase in S/G ratio when bark is used, as in this case.

Parameter	POP_L	PIN_L
Glycan (w%)	38.8	23.2
Xylan-mannan (w%)	-	0.8
Lignin AI (w%)	51.1	69.1
S (w%)	0.74	0.33
M_w (Da)	5216	2950
M_n(Da)	514	435
Dispersity	10.1	6.8
S/G ratio	1.5	0.1

Table 1. Characterization of downstream lignins from poplar (POP_L) and pine (PIN_L)

The low solubility of **PIN$_L$** and **POP$_L$** was critical in HSQC analysis. Although acetylation increased the solubility of the corresponding solids this was not complete. Thus, HSQC on acetylated **PIN$_L$** and **POP$_L$** presented only a few signals in the aliphatic and methoxy regions while none could be detected in the aromatic region, even using long acquisition times. In both cases, cross-peaks corresponding to the methoxy moiety turned up at *ca.* 3.75 ppm-55.4 ppm together with cross-peaks corresponding to C_β and C_γ in β-O-4 bonds. **PIN$_L$** was slightly soluble in DMSO that allowed HSQC-TOCSY analysis of the soluble fraction. Besides the cross-peaks corresponding to the methoxy moiety that turned up at *ca.* 3.75 ppm-55.4 ppm, the H-C cross peak that can correspond to the anomeric carbon of saccharides could be detected. This suggested that the detected signals in the range 3.0-5.0 ppm/ 60-100 ppm corresponded to H-C and long-distance correlations from partially degraded cellulose (Figure S1).

Optimization of the BCD reaction conditions was done using **POP$_L$** and NaOH as a base. Other bases like LiOH, KOH and CsOH[45] and organic bases were also tested but no improvements were found compared to NaOH (data not shown). Fixed NaOH

concentrations, 0.25 M (1 % w/v; pH 13.4), NaOH/solid ratios 25:1 and reaction temperature in the range of 150 ºC to 300 ºC during 240 min were tested.

A steady increase of known monomer yields from 3.2 w% to 9.1 w% was observed from 150 ºC to 200 ºC (Figure 1). Monomer yields steadily decreased thereafter due to repolymerization as did the soluble fraction. The effect of reaction time was also studied at different temperatures (Table S2). Monomer yields increased with longer reaction times at 200 ºC or lower temperatures and decreased at higher temperatures. Best overall reaction yields were obtained at 200 ºC, reaching 9.5 w% and 9.7 w% after 300 min and 400 min respectively (Table S2). It is worth noting that the product distribution at a given temperature in the range 150 ºC-225 ºC was constant regardless of the reaction time (Table S2, Figure 1). At low reaction temperatures, syringols are the predominant monomers with noticeable amounts of vanillins (44 % and 14 % at 150 ºC respectively). A decrease in syringols and vanillins together with an increase of guaiacols was observed in the range 175 ºC- 250 ºC. Monomeric phenols distribution was much more sensitive to reaction times at higher temperature. Increasing reaction times to 240 min caused the total disappearance of vanillins and syringols beyond 250 ºC and a drastic reduction of

guaiacols, which is counterbalanced with an increase of catechols and phenols to 14 %-19 % and 65 %-71 % respectively. The high reactivity of the carbonyl group explains the decrease in vanillins with increasing temperatures. The acidic thermochemical pretreatment of the biomass is probably in the origin of the increase in guaiacols and the decrease in syringols. It has been described that the acidic pretreatment causes the formation of C-C bonds between aromatics.[8,46] Syringyl units may form less C-C bounds upon acidic pretreatment and, consequently, at low temperatures are more easily released from lignin structures than guaiacols. Furthermore, it has also been stated that demethoxylation of syringols may occur with increasing temperature leading to the formation of syringols.[13]

Optimal reaction conditions for $\mathbf{POP_L}$ (200 ºC, 240-360 min) were used for $\mathbf{PIN_L}$ but phenolic monomer yield drastically decreased to 4.1 w%. As expected, guaiacols were the main products with 81 %, followed by phenols with only 6 %, and no guaiacols were detected. Lower monomer yield can be attributed to higher recalcitrance of the lignin due to higher cross-linking upon acidic thermal pretreatment.

Optimal conditions for **POP$_L$** were much milder compared with other studies with biorefinery[8] and Organosolv[11,45] lignins. This is a significant improvement as energy requirements are lower and the use of capping agents is avoided.

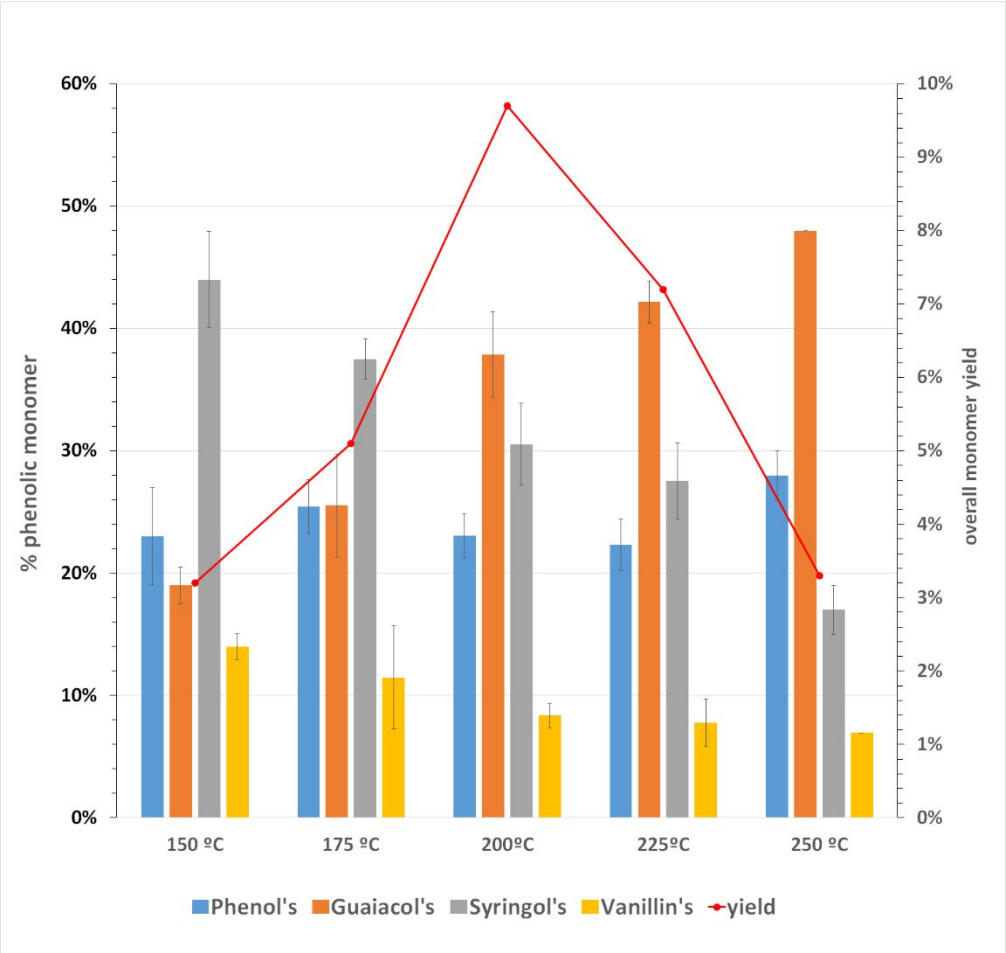


Figure 1. Monomer distribution and overall yields (w%) in BCD of **POP$_L$** in the range 150 ºC -250 ºC. Overall yields correspond to the maximum yield at given temperatures. Error bars are given at 95 % confidence level.

Six representative samples from BCD (Table 2) were analyzed by HSQC. In all cases, spectra presented cross-peaks in the range of 3.60-3.85 ppm/ 55-56 ppm as the major signals. These H-C cross-peaks corresponded to the presence of methoxy groups. BCD samples **1-4** from **POP$_L$** presented, in all cases similar spectra in the aromatic region where all the cross-peaks could be assigned to phenyl, guaiacyl and syringyl units (Figure S2). As expected, only guaiacyl moieties were detected in sample **6**, which came from **PIN$_L$** (Figure S2). Besides methoxy groups, C_γ-H_γ signals corresponding to β-O-4 bonds, together with C_β-H_β signals, were detected in the aliphatic region.

Signals at *ca.* 2.30 ppm -35 ppm and 1.25ppm - 29 ppm correspond to saturated aliphatic chains. These signals may correspond to α and β positions in aliphatic chains bounded to guaiacyl or syringyl units[47] from the poly-(hydroxy)-aromatic ether fraction. Additionally, the corresponding cross-peaks from saccharides previously detected in **PIN$_L$** vanished whereas H-C cross-peaks in the range 2.0-2.5 ppm/20-30 ppm appeared. These H-C cross-peaks can be assigned to C_α and C_β in carbonyl compounds, which is consistent with the degradation of cellulose that accompanies lignin in **POP$_L$** and **PIN$_L$**. Indeed, it is known that carbohydrates are degraded through peeling reactions in basic

conditions, providing aliphatic and cyclic ketones.[19] This behavior explains the presence of signals in the range of 1.0-1.5 ppm/ 10-25 ppm that correspond to long alkyl chains. The presence of aliphatic compounds, such as nonanol, 5-hydroxy-4-octanone or cyclohexanone derivatives, was suggested by GC-MS, but quantification of these products is beyond the scope of this study.

Size Exclusion Chromatography (SEC)

Neither **PIN_L** nor **POP_L** can be completely solubilized in DMF for SEC measurement. This low solubility is mainly attributed to the formation of C-C bonds between phenolic units upon acidic pre-treatment.[8] SEC analysis for **POP_L** and **PIN_L** (Figure S3) and after BCD were at least intriguing. M_w values, 2948 Da and 5216 Da for **POP_L** and **PIN_L**, decreased to *ca.* 1000 Da after BCD, that is rather consistent. However, M_n values were contradictory, since they were slightly higher after BCD (*ca.* 700 Da) than in the starting materials (514 Da and 435 Da for **POP_L** and **PIN_L** respectively). The reason for the low M_n values lies in the higher solubility in DMF of the low molecular weight fractions in the samples that may correspond to partially degraded cellulose residues after enzymatic hydrolysis. Apparent mass values in the crude reaction mixtures are more realistic

because of the total solubility of the samples consisting mainly of poly-(hydroxy)-aromatic ethers (Figure 2, Figure S4). SEC analysis showed the evolution of BCD with reaction time. When BCD of **POP$_L$** was run at 200 ºC it could be observed that mass distribution shifted to lower molecular weights (Figure 2, left) at longer reaction times. Two maxima were observed at apparent masses of 500 Da and 250 Da, whose relative intensity also increased with reaction time.

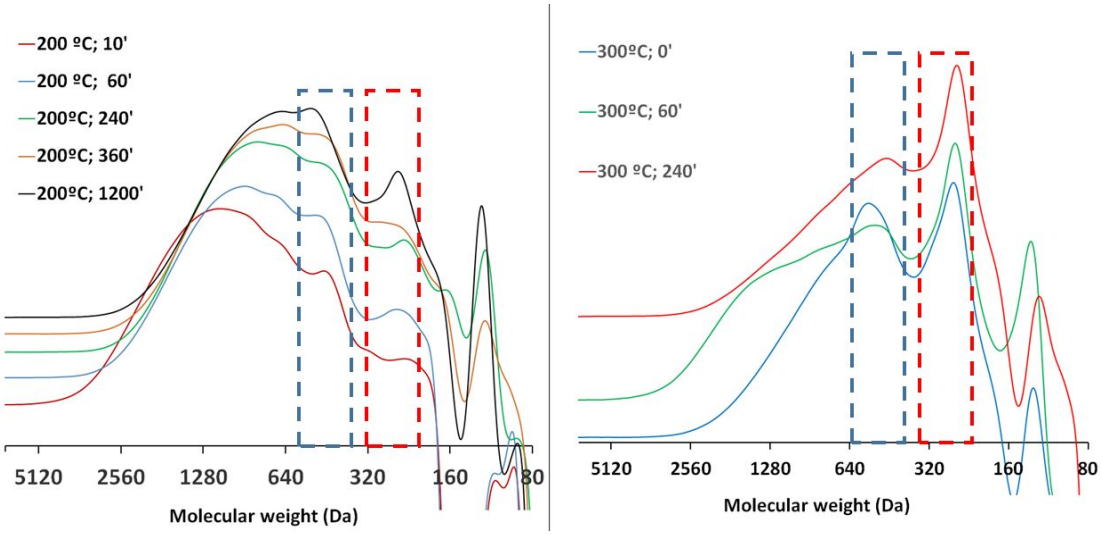


Figure 2. Normalized SEC for BCD of **POP$_L$** at left) 200 ºC and right) 300 ºC

An increase of the relative intensity of the signal at *ca*. 130 Da was also observed, corresponding to monomers with longer reaction time. A similar profile was noticed after ramping (0 min) when the reaction was run at 300 ºC. Longer reaction times produced a

slight increase of the peak at 250 Da and a broadening of the curve at higher molecular weight.

DOSY calibration

Some studies on small organometallic molecules and aggregates had already correlated the diffusion coefficient (D) of diffusates with its molecular weight (MW) as $MW^{1/3}$,[29] but this relation can only be used in small molecules.[48] However, the molecular weight of phenolic poly-(hydroxy)- ethers in BCD reaction mixtures are in the range of 200-1500 Da, according to SEC measurements. Assuming a spherical shape for oligomers, log D and log MW must correlate according to the Mark-Houwink equation,[36] as has been already reported for a series of β-O-4 vanillin oligomers.[23]

DOSY spectroscopy is sensitive to the concentration of the diffusate, artifacts arising from changes of viscosity or convection phenomena in the solvent,[29] which makes the determination of molecular weight not trivial. Therefore, all experiments were made using DMSO-d_6 at 300 K and at a fixed 1 % w/v diffusate concentration. These conditions prevented convection phenomena in the solvent and allowed free diffusion of the analyte. In a similar approach to that of SEC,[17] two calibration curves for log MW *vs* log D were

prepared according to different diffusate-solvent interactions. The first one corresponded to linear polystyrene (**PS**) standards used in SEC calibrations. The second one corresponded to poliethylenglycol (**PEG**),[49] phenolic monomers, 2-phenoxy-1-phenylethanol (**PE**) and guaiacylglycerol-β-guaiacyl ether (**VG**). The first curve accounted for intermolecular dispersion forces whereas the second accounted for strong Van der Waals interactions and, occasionally, hydrogen bonds with residual water.

Excellent correlation was found using **PS** (r^2= 0.999, see SI). In the case of the second curve, correlation was somewhat worse (r^2= 0.988, see SI) with a loss in linearity in the monomeric phenol region. Small changes due to interaction with residual H_2O must be taken into account in low molecular weight phenols. Indeed, NOESY measurements in phenol, vanillin and guaiacol showed weak interaction between one *orto*-hydrogen to the hydroxy group and residual water in the deuterated solvent (Figure 3 and Figure S6). These interactions produced an increase in the hydrodynamic radii of these molecules. Thus, when these water molecules were included in the molecular weights of the monomeric phenols excellent correlation between log D and log MW was observed (r^2=0.996, n=10, s.e,=0.040, $F_{1,8}$ = 1880.3, p<0.001, SI.8). This last curve permits the

estimation of molecular weight with less than 6% error in discrete molecules and lower than 15% for PEG polymers (MW ≤ 8000 Da).

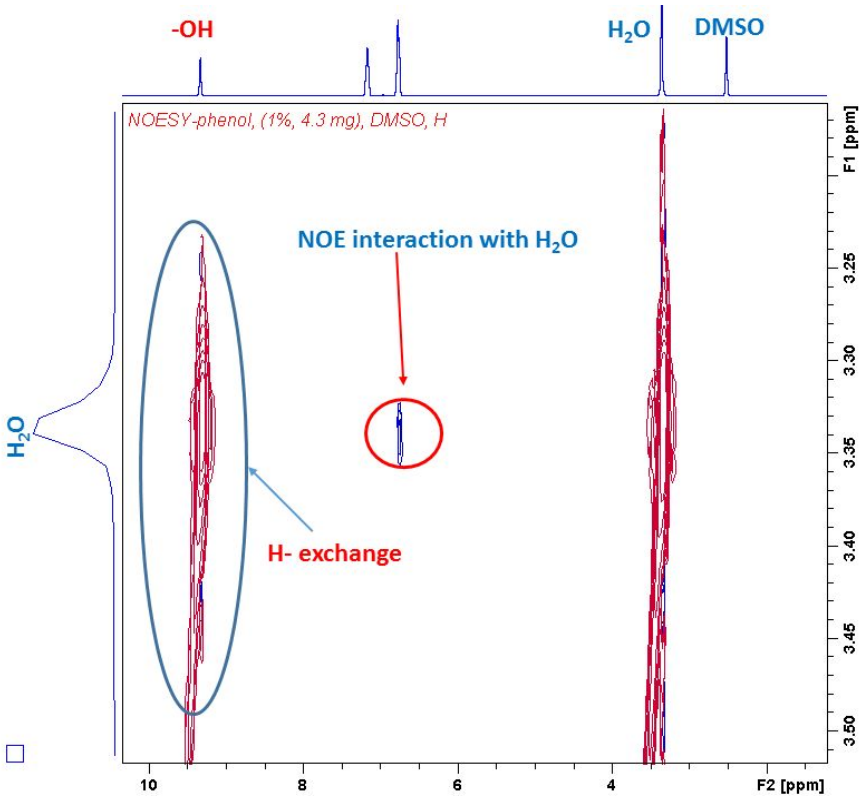


Figure 3. NOESY spectra for phenol in DMSO-d_6

DOSY analysis

The most important feature of DOSY spectrometry is that it permits the diffusion coefficient and the chemical shifts to be correlated. Therefore, DOSY allows using exclusively those signals corresponding to the aromatic hydrogen atoms to determine the molecular weight of the aromatic oligomeric fraction, disregarding other signals that are

not interesting (*i.e.* aliphatic signals). This is particularly advantageous in our case, where the starting solids presented an important amount of saccharides that may produce side-products.

PIN_L, **POP_L** and six representative reaction mixtures from BCD were analyzed using DOSY (Table 2). Average diffusion coefficients were calculated by fitting the Stejskal-Tanner equation using the intensity of the area of the integration corresponding to the aromatic region and two aliphatic regions (Table S.3). Therefore, apparent mass values should correspond to M_n values obtained by SEC. Experimental and predicted areas showed excellent correlations ($r^2 > 0.99$, Table S3) in all cases except in the case of **POP_L** due to the low solubility of the sample.

The solubility of **PIN_L** and **POP_L** in DMSO-d_6 was very low, which complicated the analysis. Acetylation of both solids and dissolution in DMSO-d_6 at 75 ºC overnight was inefficient, as part of the solids remained insoluble. Most of the signals in DOSY spectra corresponded to aliphatic hydrogen atoms. Aliphatic signals in the range of 4.0- 5.0 ppm were assigned to the cellulosic fraction that is already present in the lignin fraction and that cannot be discriminated using SEC. It is worth noting that these traces presented

significantly lower apparent masses than those from the aromatic region. Apparent mass in the aromatic region in **POP$_L$** (Table 2) was determined graphically as the signal integration was very low and fitting was inconsistent. Apparent masses in the aromatic region for **PIN$_L$** were 1274 Da and 928 Da for **PS**, and **PEG**-phenols (**PS**, **PEG**) calibration curves respectively. In the case of **POP$_L$**, apparent masses in the aromatic region were 7239 Da, and 6664 Da respectively. These values are much higher than those measured by SEC in **PIN$_L$** and **POP$_L$** (435 Da and 514 Da respectively) which is due to the higher solubility of the aliphatic and the saccharide fractions. Apparent masses for the saccharide region (4.5 – 3.0 ppm) ranged from 443 Da and 6132 Da in **POP$_L$** and from 250 Da to 2200 Da in **PIN$_L$** which is in line with the SEC measurements (Figure 2 and Figure S4).

Diffusion coefficients for samples **1**-**6** were higher than those measured in **POP$_L$**, **PIN$_L$** (Table 2, Figure 4) and **PEG-1500** too, evidencing lignin depolymerization. At first sight, it could also be observed that diffusion coefficients are strongly different in the aromatic and in the aliphatic regions of the spectra. Thus, in the case of **1** diffusion projections associated to aromatic ^{1}H NMR signals turned up in the range of -9.46 log(m^2/s) and -

9.92 log(m^2/s), thus in the range of 242 Da – 1300 Da and 146 Da-950 Da according to **PS** and **PEG**. Diffusion coefficients were averaged to -9.73 log(m^2/s) resulting in apparent masses of 658 Da (**PS**), which are in the same order of magnitude as those measured by SEC, 719 Da. Some of the traces presented similar log D as monomeric phenols or **PE**, that is to say, in the range of phenolic monomers and dimers (Figure 4). Signals corresponding to proton exchange with residual water could also be observed at low fields and -9.35 log (m^2/s).

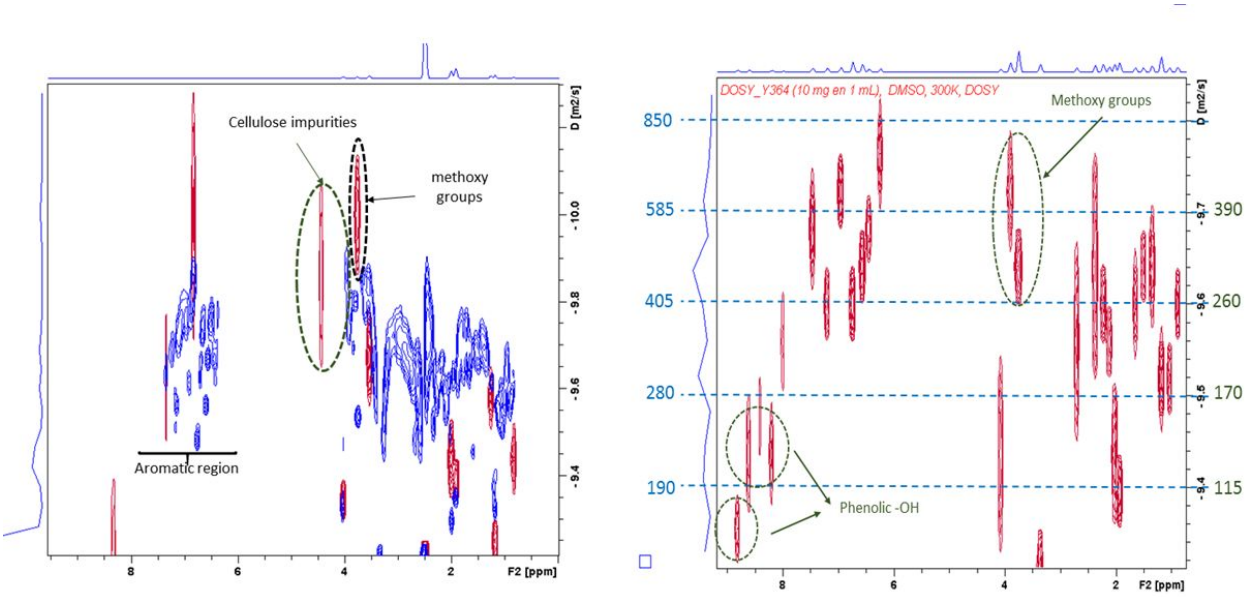


Figure 4. Left) DOSY for $\mathbf{POP_L}$ (red) and sample **1** (blue). Right) DOSY for sample **1**. Green and blue figures correspond to molecular weights obtained with **PEG** and **PS**

calibration respectively. This spectrum was obtained after fitting the intensities of manually defined integration regions.

A similar observation can be made for sample **4** (Figure 5). When the aromatic traces were analyzed it could be observed that most diffusion traces for the aromatic moieties ranged between 455 Da and 1045 Da according to **PS** calibration. Traces corresponding to low apparent masses, 250 Da to 340 Da were observed in samples **3** and **4**, which reveals the presence of dimeric and even monomeric fractions according to **PEG** calibration. The average apparent mass, which can be compared with M_n measured by SEC, in the aromatic region corresponded to 824 Da and 592 Da (**PS**, **PEG**) for **2**, 648 Da and 453 Da for **3** and 513 Da, 349 Da for **4**, which is, at least, in range with the values measured by SEC. Average apparent masses of 764 Da and 544 Da (**PS**, **PEG**) were found for sample **6** obtained from $\mathbf{PIN_L}$ with masses ranging between 565 Da and 1163 Da. Nevertheless, in sample **6** traces with apparent masses under 565 Da were less intense than in **1**, **3** and **4**, which suggests that lignin depolymerization proceeded to a greater extent in poplar derived lignin, $\mathbf{POP_L}$, under these conditions. SEC analysis for **6** also

showed a broad distribution with most of the oligomers at masses beyond 800 Da (Figure 2 and Figure S4).

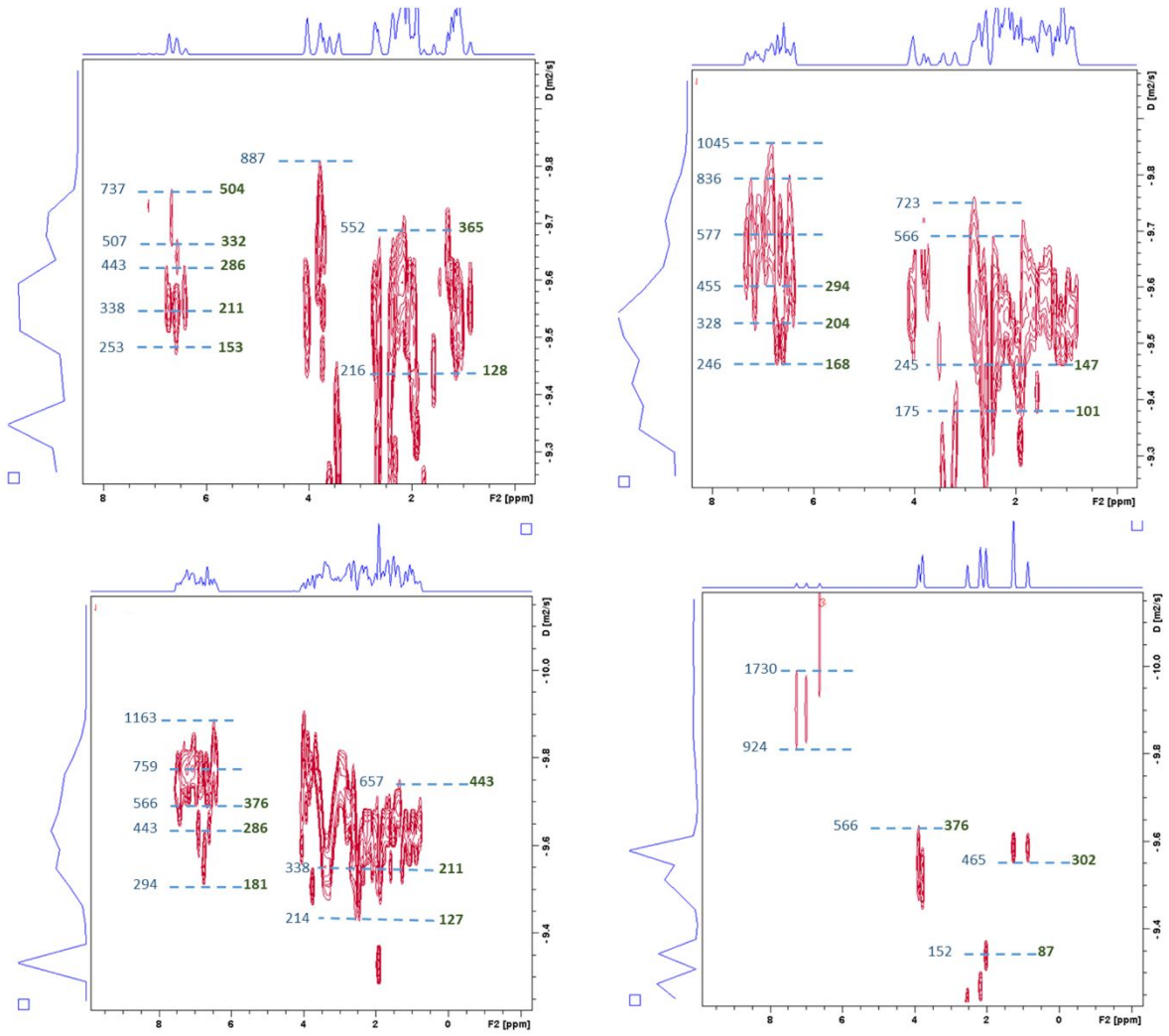


Figure 5. DOSY spectra for top left) sample **3**; top right) sample **4**; bottom left) sample **6;** bottom right) sample **5.** Green and blue figures correspond to molecular weights obtained with **PEG** and **PS** calibration respectively.

POP$_L$ was also reacted under microwave irradiation at 200 ºC during 60 min, sample **5,** and almost negligible monomer yields were measured. Apparent masses determined by SEC were in line with the former although two prominent peaks turned up in SEC at 120 Da and 190 Da. It is noteworthy, that no signal was detected in the aromatic region by HSQC nor by HSQC-TOCSY, but in the aliphatic region, cross-peaks corresponding to C_α from carbonyl and aliphatic chains were predominant (Figure S5). When sample **5** was analyzed using DOSY, it was difficult to find representative diffusion projections in the aromatic region (Figure 5). These aromatic traces presented apparent values in the range of 924 Da to 1730 Da showing that lignin depolymerization did not occur under these conditions. Diffusion traces in the aliphatic region were, however, intense with apparent masses ranging from 87 Da to 376 Da according to **PEG**, which showed that, although no depolymerization of lignin took place, degradation of the cellulosic fraction did. Samples **2**, **3** and **5** were derivatized with 2-chloro-4,4,5,5-tetramethyl-1,3,2-dioxaphospholane (**TMDP**) and analyzed using ^{31}P NMR (Figure S7) and it could be seen that the amount of aromatic hydroxyl groups in **5** is much lower than in samples **2** and **3**,

which confirmed the low amount of phenolic –OH in the reaction mixture under in the reaction mixture after microwave irradiation, as suggested by DOSY and HSQC-TOCSY.

Samples **1-4** and **6** showed representative traces in the aliphatic region too. The diffusion coefficients were significantly higher than those of the aromatic region. This proved that the aliphatic traces were not part of the poly-(hydroxy)-aromatic ethers fraction but of the degradation products from the cellulosic residue. Apparent masses for the aliphatic traces were, in most of the cases, in the range of 200 Da- 500 Da according to **PS** calibration or 130 Da – 300 Da according to **PEG** calibration. Integration values in the ^{1}H NMR spectra are higher in this region than in the aromatic, which suggests that peeling reaction of saccharides is more intense than lignin depolymerization in the starting solids. Hence, DOSY spectra can be interpreted, as in the case of SEC chromatograms, as a bimodal distribution. Furthermore, DOSY allowed selective determination of the molecular masses corresponding to the reaction products providing additional information. This is particularly important when dealing with downstream biorefinery lignins whose purity is relatively low.

Apparent mass values measured by DOSY were, in samples **1-4** and **6**, in the same range than M_n values determined by SEC. Small differences can be found between the apparent masses of the aromatic regions which can be attributed to the presence of the degraded saccharide fraction and the different nature of the used techniques. Indeed, these differences can even be found using the well-established SEC with refraction index detectors[18] as a consequence of the calibration with **PS**.[17]

Sample	T (ºC)	t (min)	SEC			Aromatic region			Aliphatic region 1			Aliphatic region		
			M_n (Da)	M_w (Da)	n	PS (Da)	PEG (Da)	10^{10}·D (m^2/s)	PS (Da)	PEG (Da)	10^{10}·D (m^2/s)	PS (Da)	PEG (Da)	10^{10}·D (m^2/s)
POP_L[a]	-	-	435	2950	6.8	7239	6664	0.421	4102	3540	0.599	292	186	3.09
1	200	240	719	1051	1.5	658	461	1.86	482	326	2.26	418	278	2.47
2	200	360	785	1112	1.4	824	592	1.62	330	213	2.86	428	285	2.43
3	250	60	869	1768	2	648	453	1.88	353	231	2.74	401	265	2.53
4	300	10	615	959	1.6	513	349	2.18	345	225	2.78	389	256	2.58

5[b]	200	300	874	1246	1.4	-			94	53	6.22	248	156	3.41
PIN$_L$[a]	-	-	514	5216	10.1	1274	928	1.24	681	478	1.83	242	151	3.47
6	200	210	720	1173	1.6	764	544	1.70	494	335	2.23	520	355	2.16

a) Acetylated sample was used in NMR measurements. b) Sample was prepared using microwave irradiation during 60 minutes

Table 2. Apparent masses estimated by SEC and NMR DOSY for the starting solids and samples **1-6**

DOSY spectroscopy is a valuable tool for the selective determination of apparent masses in the different fractions arising from BCD of downstream biorefinery lignins. Nevertheless, as in the case of SEC, attention must be paid to diffusate-diffusate and diffusate-solvent interactions which are of paramount importance. Nice correlations between log D and log MW were found when the nature of the diffusate-solvent interactions are similar (*ie.* **PS** and phenols-**PEG**). Discrepancies in the mass values were found in low mass molecules because of significant differences in the hydrodynamic radii that arise from the interaction of monomeric phenols with the residual H_2O in the sample through the formation of hydrogen bonds. Despite the good correlation found for monomeric phenols and **PEG**, the apparent mass values provided by DOSY for most reaction mixtures are closer to M_n when **PS** calibration was used. The increase in the molecular weight of the poly-(hydroxy)-aromatic chain renders these diffusates more hydrophobic and therefore more similar to **PS** than to **PEG** or monomeric phenols with regard to their interaction with DMSO.

CONCLUSIONS

Careful optimization of BCD reaction conditions provided excellent phenolic monomer yields (> 9.7 w%) at low reaction temperature (200 ºC) using 25:1 base/solid ratio Phenolic monomer distribution were tuned with temperature and reaction time. DOSY spectroscopy provided excellent log D *vs* log MW correlation in standards with similar intermolecular interactions and was successfully used in the determination of the molecular weight of poly-(hydroxy)-aromatic ethers in lignins and bio-oil samples. Additional information provided by DOSY could be invaluable in order to discriminate the molecular weight of different fractions in lignins and the bio-oils arising from BCD.

ASSOCIATED CONTENT

The Supporting Information is available free of charge on the ACS Publications website: Monomer groups and retention times. · Optimization of the BCD reaction conditions. · HSQC-NMR spectra and SEC chromatograms. · Calibration curve of logMW and logD using DOSY for PS. · Calibration curve of log MW and log D using DOSY for PEG and monomeric phenols, VG and PE. · NOESY spectra for phenol, vanillin and guaiacol. · Integration regions used in the determination of averaged diffusion coefficients. · ^{31}P NMR spectra for derivatized samples.

AUTHOR INFORMATION

Corresponding Author

Alfonso Cornejo - *Institute for Advanced Materials (INAMAT)-Dpt. of Sciences, Campus de Arrosadia, Universidad Pública de Navarra, E31006 Pamplona, Spain*

Authors:

Íñigo García-Yoldi - *Institute for Advanced Materials (INAMAT)-Dpt. of Sciences, Campus de Arrosadia, Universidad Pública de Navarra, E31006 Pamplona, Spain*

Irantzu Alegria-Dallo - *National Renewable Energy Centre (CENER), Av. Ciudad de la Innovación 7, E31261 Sarriguren, Spain*

Rebeca Galilea-Gonzalo - *Institute for Advanced Materials (INAMAT)-Dpt. of Sciences, Campus de Arrosadia, Universidad Pública de Navarra, E31006 Pamplona, Spain*

Karina Hablich - *Institute for Advanced Materials (INAMAT)-Dpt. of Sciences, Campus de Arrosadia, Universidad Pública de Navarra, E31006 Pamplona, Spain*

David Sánchez - *National Renewable Energy Centre (CENER), Av. Ciudad de la Innovación 7, E31261 Sarriguren, Spain*

Eduardo Otazu- *National Renewable Energy Centre (CENER), Av. Ciudad de la Innovación 7, E31261 Sarriguren, Spain*

Ibai Funcia - *National Renewable Energy Centre (CENER), Av. Ciudad de la Innovación 7, E31261 Sarriguren, Spain*

María J. Gil - *Institute for Advanced Materials (INAMAT)-Dpt. of Sciences, Campus de Arrosadia, Universidad Pública de Navarra, E31006 Pamplona, Spain*

Víctor Martínez-Merino - I*nstitute for Advanced Materials (INAMAT)-Dpt. of Sciences, Campus de Arrosadia, Universidad Pública de Navarra, E31006 Pamplona, Spain*

Notes

The authors declare no competing financial interests.

ACKNOWLEDGMENTS

Financial support from The Government of Navarra under Project “PC036-037 Biovalorización” and UNED-Fundación La Caixa “LCF/PR/PR15/51100007”is sincerely acknowledged. Íñigo García-Yoldi is grateful for the support from The Government of

Navarra via the concession of a " Technology Training Grant ("Beca de Formación de Tecnólogos").

REFERENCES

(1) Cao, L.; Yu, I. K. M.; Liu, Y.; Ruan, X.; Tsang, D. C. W.; Hunt, A. J.; Ok, Y. S.; Song, H.; Zhang, S. Lignin Valorization for the Production of Renewable Chemicals: State-of-the-Art Review and Future Prospects. *Bioresour. Technol.* **2018**, *269*, 465–475. DOI 10.1016/j.biortech.2018.08.065.

(2) Xu, C.; Arancon, R. A. D.; Labidi, J.; Luque, R. Lignin Depolymerisation Strategies: Towards Valuable Chemicals and Fuels. *Chem. Soc. Rev.* **2014**, *43* (22), 7485–7500. DOI 10.1039/c4cs00235k.

(3) Kumar, L.; Chandra, R.; Chung, P. A.; Saddler, J. Can the Same Steam Pretreatment Conditions Be Used for Most Softwoods to Achieve Good, Enzymatic Hydrolysis and Sugar Yields? *Bioresour. Technol.* **2010**, *101* (20), 7827–7833. DOI 10.1016/j.biortech.2010.05.023.

(4) Calvo-Flores, F. G.; Dobado, J. A. Lignin as Renewable Raw Material.

ChemSusChem **2010**, *3* (11), 1227–1235. DOI 10.1002/cssc.201000157.

(5) Li, C.; Zhao, X.; Wang, A.; Huber, G. W.; Zhang, T. Catalytic Transformation of Lignin for the Production of Chemicals and Fuels. *Chem. Rev.* **2015**, *115* (21), 11559–11624. DOI 10.1021/acs.chemrev.5b00155.

(6) Schutyser, W.; Renders, T.; Van Den Bosch, S.; Koelewijn, S. F.; Beckham, G. T.; Sels, B. F. Chemicals from Lignin: An Interplay of Lignocellulose Fractionation, Depolymerisation, and Upgrading. *Chem. Soc. Rev.* **2018**, *47* (3), 852–908. DOI 10.1039/c7cs00566k.

(7) Upton, B. M.; Kasko, A. M. Strategies for the Conversion of Lignin to High-Value Polymeric Materials: Review and Perspective. *Chem. Rev.* **2016**, *116* (4), 2275–2306. DOI 10.1021/acs.chemrev.5b00345.

(8) Katahira, R.; Mittal, A.; McKinney, K.; Chen, X.; Tucker, M. P.; Johnson, D. K.; Beckham, G. T. Base-Catalyzed Depolymerization of Biorefinery Lignins. *ACS Sustain. Chem. Eng.* **2016**, *4* (3), 1474–1486. DOI 10.1021/acssuschemeng.5b01451.

(9) Otromke, M.; White, R. J.; Sauer, J. Hydrothermal Base Catalyzed Depolymerization and Conversion of Technical Lignin – An Introductory Review. *Carbon Resour. Convers.* **2019**, *2* (1), 59–71. DOI 10.1016/j.crcon.2019.01.002.

(10) Chaudhary, R.; Dhepe, P. L. Solid Base Catalyzed Depolymerization of Lignin into Low Molecular Weight Products. *Green Chem.* **2017**, *19* (3), 778–788. DOI 10.1039/c6gc02701f.

(11) Erdocia, X.; Prado, R.; Corcuera, M. Á.; Labidi, J. Base Catalyzed Depolymerization of Lignin: Influence of Organosolv Lignin Nature. *Biomass Bioenerg.* **2014**, *66*, 379–386. DOI 10.1016/j.biombioe.2014.03.021.

(12) Fernández-Rodríguez, J.; Erdocia, X.; Sánchez, C.; González Alriols, M.; Labidi, J. Lignin Depolymerization for Phenolic Monomers Production by Sustainable Processes. *J. Energy Chem.* **2017**, *26* (4), 622–631. DOI 10.1016/j.jechem.2017.02.007.

(13) Lavoie, J. M.; Baré, W.; Bilodeau, M. *Depolymerization of Steam-Treated Lignin for the Production of Green Chemicals*; Elsevier Ltd, 2011; Vol. 102.

DOI 10.1016/j.biortech.2011.01.010.

(14) Toledano, A.; Serrano, L.; Labidi, J. Improving Base Catalyzed Lignin Depolymerization by Avoiding Lignin Repolymerization. *Fuel* **2014**, *116*, 617–624. DOI 10.1016/j.fuel.2013.08.071.

(15) Li, J.; Zhang, J.; Zhang, S.; Gao, Q.; Li, J.; Zhang, W. Alkali Lignin Depolymerization under Eco-Friendly and Cost-Effective NaOH/Urea Aqueous Solution for Fast Curing Bio-Based Phenolic Resin. *Ind. Crops Prod.* **2018**, *12*, 25–33. DOI 10.1016/j.indcrop.2018.04.027.

(16) Mahmood, N.; Yuan, Z.; Schmidt, J.; Tymchyshyn, M.; Xu, C. Hydrolytic Liquefaction of Hydrolysis Lignin for the Preparation of Bio-Based Rigid Polyurethane Foam. *Green Chem.* **2016**, *18* (8), 2385–2398. DOI 10.1039/c5gc02876k.

(17) Andrianova, A. A.; Yeudakimenka, N. A.; Lilak, S. L.; Kozliak, E. I.; Ugrinov, A.; Sibi, M. P.; Kubátová, A. Size Exclusion Chromatography of Lignin: The Mechanistic Aspects and Elimination of Undesired Secondary Interactions. *J. Chromatogr. A*

2018, *1534*, 101–110. DOI 10.1016/j.chroma.2017.12.051.

(18) Ritter, A.; Schmid, M.; Affolter, S. Determination of Molecular Weights by Size Exclusion Chromatography (SEC) - Results of Round Robin Tests. *Polym. Test.* **2010**, *29* (8), 945–952. DOI 10.1016/j.polymertesting.2010.08.002.

(19) Green, J. W.; Pearl, I. A.; Hardacker, K. W.; Andrews, B. D.; Haigh, F. C. Peeling Reaction in Alkaline Pulping. *TAPPI* **1977**, *60* (10), 120–125.

(20) Hu, G.; Cateto, C.; Pu, Y.; Samuel, R.; Ragauskas, A. J. Structural Characterization of Switchgrass Lignin after Ethanol Organosolv Pretreatment. *Energy and Fuels* **2012**, *26* (1), 740–745. DOI 10.1021/ef201477p.

(21) Kuznetsov, B. N.; Chesnokov, N. V.; Sudakova, I. G.; Garyntseva, N. V.; Kuznetsova, S. A.; Malyar, Y. N.; Yakovlev, V. A.; Djakovitch, L. Green Catalytic Processing of Native and Organosolv Lignins. *Catal. Today* **2018**, *309*, 18–30. DOI 10.1016/j.cattod.2017.11.036.

(22) Mansfield, S. D.; Kim, H.; Lu, F.; Ralph, J. Whole Plant Cell Wall Characterization

Using Solution-State 2D NMR. *Nat. Protoc.* **2012**, *7* (9), 1579–1589. DOI 10.1038/nprot.2012.064.

(23) Savy, D.; Mazzei, P.; Drosos, M.; Cozzolino, V.; Lama, L.; Piccolo, A. Molecular Characterization of Extracts from Biorefinery Wastes and Evaluation of Their Plant Biostimulation. *ACS Sustain. Chem. Eng.* **2017**, *5* (10), 9023–9031. DOI 10.1021/acssuschemeng.7b01928.

(24) Wang, H.; Pu, Y.; Ragauskas, A.; Yang, B. From Lignin to Valuable Products–Strategies, Challenges, and Prospects. *Bioresour. Technol.* **2019**, *271*, 449–461. DOI 10.1016/j.biortech.2018.09.072.

(25) Zeng, J.; Helms, G. L.; Gao, X.; Chen, S. Quantification of Wheat Straw Lignin Structure by Comprehensive NMR Analysis. *J. Agric. Food Chem.* **2013**, *61* (46), 10848–10857. DOI 10.1021/jf4030486.

(26) Van Den Bosch, S.; Renders, T.; Kennis, S.; Koelewijn, S. F.; Van Den Bossche, G.; Vangeel, T.; Deneyer, A.; Depuydt, D.; Courtin, C. M.; Thevelein, J. M.; Schutyser, W.; Sels, B.F. Integrating Lignin Valorization and Bio-Ethanol

Production: On the Role of Ni-Al_2O_3 Catalyst Pellets during Lignin-First Fractionation. *Green Chem.* **2017**, *19* (14), 3313–3326. DOI 10.1039/c7gc01324h.

(27) Beauchet, R.; Monteil-Rivera, F.; Lavoie, J. M. Conversion of Lignin to Aromatic-Based Chemicals (L-Chems) and Biofuels (L-Fuels). *Bioresour. Technol.* **2012**, *121*, 328–334. DOI 10.1016/j.biortech.2012.06.061.

(28) Pu, Y.; Cao, S.; Ragauskas, A. J. Application of Quantitative ^{31}P NMR in Biomass Lignin and Biofuel Precursors Characterization. *Energy Environ. Sci.* **2011**, *4* (9), 3154–3166. DOI 10.1039/c1ee01201k.

(29) Li, D.; Kagan, G.; Hopson, R.; Williard, P. G. Formula Weight Prediction by Internal Reference Diffusion-Ordered NMR Spectroscopy (DOSY). *J. Am. Chem. Soc.* **2009**, *131* (15), 5627–5634. DOI 10.1021/ja810154u.

(30) Ge, W.; Zhang, J. H.; Pedersen, C. M.; Zhao, T.; Yue, F.; Chen, C.; Wang, P.; Wang, Y.; Qiao, Y. DOSY NMR: A Versatile Analytical Chromatographic Tool for Lignocellulosic Biomass Conversion. *ACS Sustain. Chem. Eng.* **2016**, *4* (3), 1193–1200. DOI 10.1021/acssuschemeng.5b01259.

(31) S. Price, W. NMR Gradient Methods in the Study of Proteins. *Annu. Reports Sect. "C" (Physical Chem.)* **2000**, *96* (2), 3. DOI 10.1039/b000773k.

(32) Zuccaccia, D.; Macchioni, A. An Accurate Methodology to Identify the Level of Aggregation in Solution by PGSE NMR Measurements. *Organometallics* **2005**, *24*, 3476–3486. DOI 10.1021/om050145k CCC:

(33) Cros-Gagneux, A.; Delpech, F.; Nayral, C. C.; Cornejo, A.; Coppel, Y.; Chaudret, B. Surface Chemistry of InP Quantum Dots: A Comprehensive Study. *J. Am. Chem. Soc.* **2010**, *132* (51), 18147–18157. DOI 10.1021/ja104673y.

(34) Montgomery, J. R. D.; Lancefield, C. S.; Miles-Barrett, D. M.; Ackermann, K.; Bode, B. E.; Westwood, N. J.; Lebl, T. Fractionation and DOSY NMR as Analytical Tools: From Model Polymers to a Technical Lignin. *ACS Omega* **2017**, *2* (11), 8466–8474. DOI 10.1021/acsomega.7b01287.

(35) Montgomery, J. R. D.; Bazley, P.; Lebl, T.; Westwood, N. J. Using Fractionation and Diffusion Ordered Spectroscopy to Study Lignin Molecular Weight. *Chem. Open* **2019**, 601–605. DOI 10.1002/open.201900129.

(36) Ronnols, J.; Jacobs, A.; Aldaeus, F. Consecutive Determination of Softwood Kraft Lignin Structure and Molar Mass from NMR Measurements. *Holzforschung* **2017**, *71*, 563–570. DOI 10.1515/hf-2016-0182.

(37) Savy, D.; Mazzei, P.; Roque, R.; Nuzzo, A.; Bowra, S.; Santos, R. Structural Recognition of Lignin Isolated from Bioenergy Crops by Subcritical Water: Ethanol Extraction. *Fuel Process. Technol.* **2015**, *138*, 637–644. DOI 10.1016/j.fuproc.2015.07.004.

(38) Kruger, J. S.; Cleveland, N. S.; Zhang, S.; Katahira, R.; Black, B. A.; Chupka, G. M.; Lammens, T.; Hamilton, P. G.; Biddy, M. J.; Beckham, G. T. Lignin Depolymerization with Nitrate-Intercalated Hydrotalcite Catalysts. *Acs Catal.* **2016**, *6* (2), 1316–1328. DOI 10.1021/acscatal.5b02062.

(39) Cornejo, A.; Alegria-Dallo, I.; García-Yoldi, I.; Sarobe, I.; Sánchez, D.; Otazu, E.; Funcia, I.; Gil, M. J.; Martínez-Merino, V. Pretreatment and Enzymatic Hydrolysis for the Efficient Production of Glucose and Furfural from Wheat Straw, Pine and Poplar Chips. *Bioresour. Technol.* **2019**, *288*, 121583.

DOI 10.1016/j.biortech.2019.121583.

(40) Ohra-Aho, T.; Gomes, F. J. B.; Colodette, J. L.; Tamminen, T. S/G Ratio and Lignin Structure among Eucalyptus Hybrids Determined by Py-GC/MS and Nitrobenzene Oxidation. *J. Anal. Appl. Pyrolysis* **2013**, *101*, 166–171. DOI 10.1016/j.jaap.2013.01.015.

(41) Pinto, P. C. R.; Da Silva, E. A. B.; Rodrigues, A. E. Comparative Study of Solid-Phase Extraction and Liquid-Liquid Extraction for the Reliable Quantification of High Value Added Compounds from Oxidation Processes of Wood-Derived Lignin. *Ind. Eng. Chem. Res.* **2010**, *49* (23), 12311–12318. DOI 10.1021/ie101680s.

(42) Fulmer, G. R.; Miller, A. J. M.; Sherden, N. H.; Gottlieb, H. E.; Nudelman, A.; Stoltz, B. M.; Bercaw, J. E.; Goldberg, K. I. NMR Chemical Shifts of Trace Impurities: Common Laboratory Solvents, Organics, and Gases in Deuterated Solvents Relevant to the Organometallic Chemist. *Organometallics* **2010**, *29* (9), 2176–2179. DOI 10.1021/om100106e.

(43) Zakzeski, J.; Bruijnincx, P. C. A.; Jongerius, A. L.; Weckhuysen, B. M. The Catalytic

Valorization of Lignin for the Production of Renewable Chemicals. *Chem. Rev.* **2010**, *110* (6), 3552–3599. DOI 10.1021/cr900354u.

(44) Lourenço, A.; Pereira, H. No Titl. In *Lignin Trends and Applications*; Poletto, M., Ed.; Intech Open: Rijeka, 2018, pp. 65–99. DOI 10.5772/intechopen.71208.

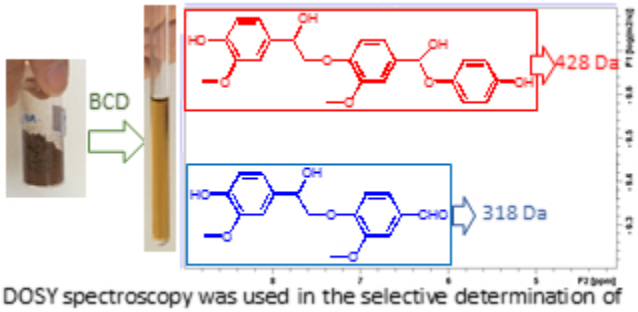
(45) Toledano, A.; Serrano, L.; Labidi, J. Organosolv Lignin Depolymerization with Different Base Catalysts. *J. Chem. Technol. Biotechnol.* **2012**, *87* (11), 1593–1599. DOI 10.1002/jctb.3799.

(46) Moxley, G.; Gaspar, A. R.; Higgins, D.; Xu, H. Structural Changes of Corn Stover Lignin during Acid Pretreatment. *J. Ind. Microbiol. Biotechnol.* **2012**, *39* (9), 1289–1299. DOI 10.1007/s10295-012-1131-z.

(47) Li, K.; Al-Rudainy, B.; Sun, M.; Wallberg, O.; Hulteberg, C.; Tunå, P. Membrane Separation of the Base-Catalyzed Depolymerization of Black Liquor Retentate for Low-Molecular-Mass Compound Production. *Membranes,* **2019**, *9* (8), 102. DOI 10.3390/membranes9080102.

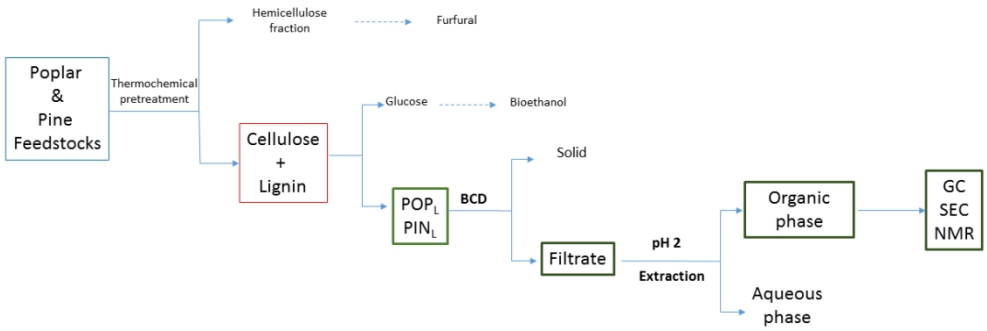
(48) Macchioni, A.; Ciancaleoni, G.; Zuccaccia, C.; Zuccaccia, D. Determining Accurate Molecular Sizes in Solution through NMR Diffusion Spectroscopy. *Chem. Soc. Rev.* **2008**, *37* (3), 479–489. DOI 10.1039/b615067p.

(49) Chen, A.; Wu, D.; Johnson, C. S. Determination of Molecular Weight Distribution for Polymers by Diffusion-Ordered NMR. *J. Am. Chem. Soc.* **1995**, *117*, 7965–7970. DOI 10.1021/ja00135a015.



DOSY spectroscopy was used in the selective determination of apparent mass of lignins and fractions arising from BCD.

84x47mm (96 x 96 DPI)



Scheme 1

325x111mm (96 x 96 DPI)

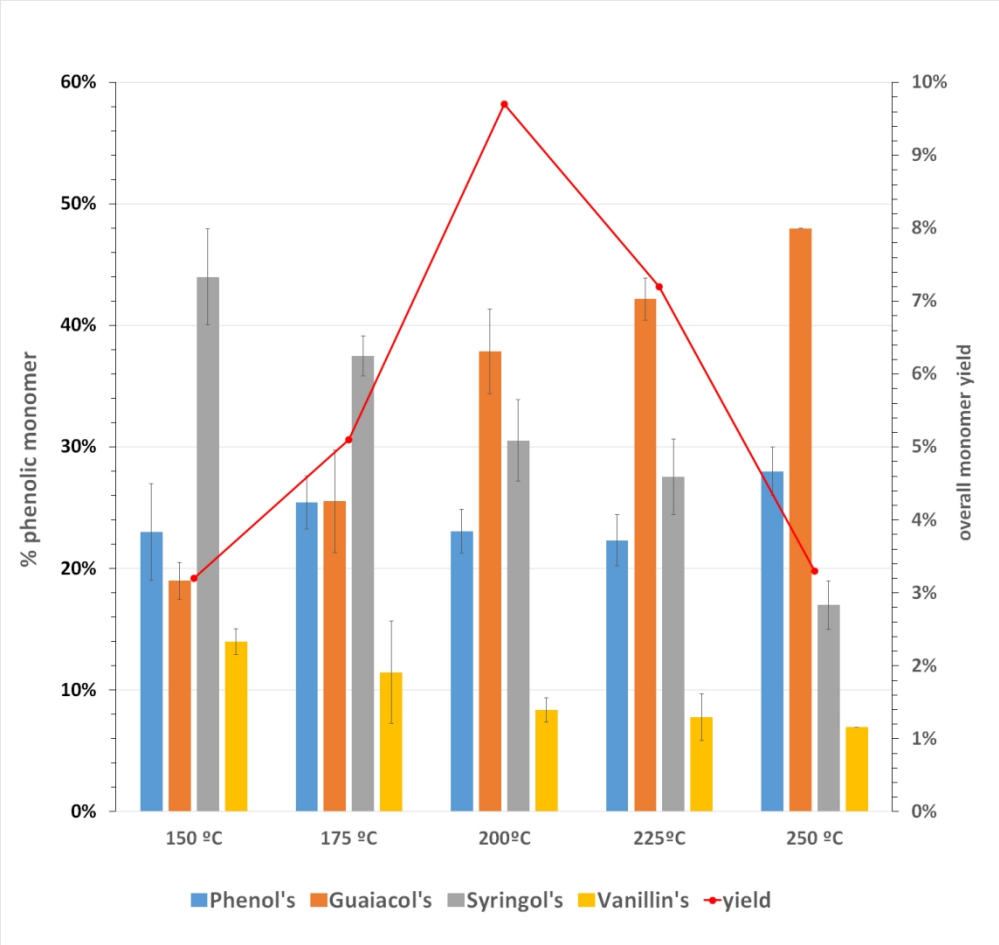


Figure 1. Monomer distribution and overall yields (w%) in BCD of POPL in the range 150 ºC -250 ºC. Overall yields correspond to the maximum yield at a given temperatures. Error bars are given at 95 % confidence level.

403x381mm (96 x 96 DPI)

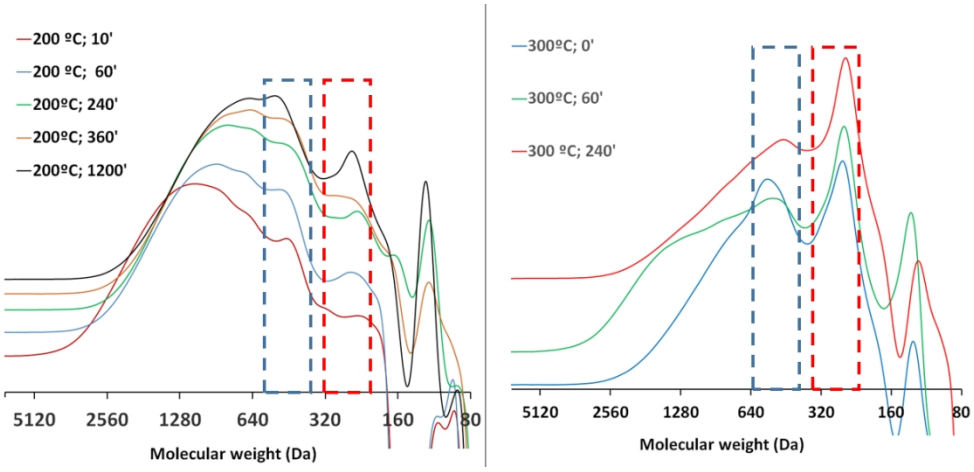


Figure 2. Normalized SEC for BCD of POPL at left) 200 ºC and right) 300 ºC

336x156mm (96 x 96 DPI)

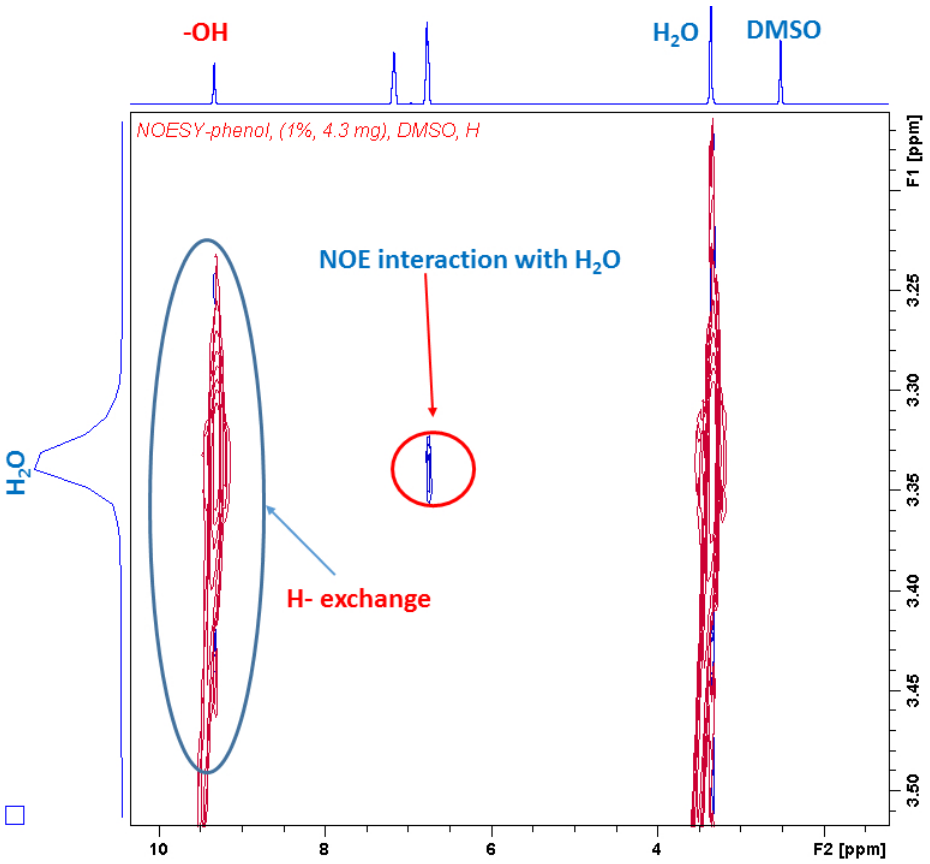


Figure 3. NOESY spectra for phenol in DMSO-d6

220x188mm (96 x 96 DPI)

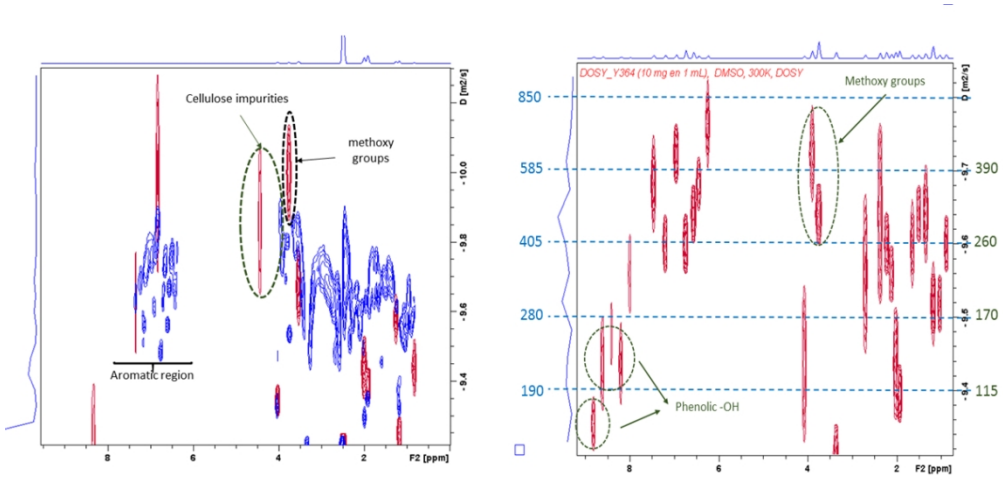


Figure 4. Left) DOSY for POPL (red) and sample 1 (blue). Right) DOSY for sample 1. Green and blue figures correspond to molecular weights obtained with PEG and PS calibration respectively. This spectrum was obtained after fitting the intensities of manually defined integration regions.

319x153mm (96 x 96 DPI)

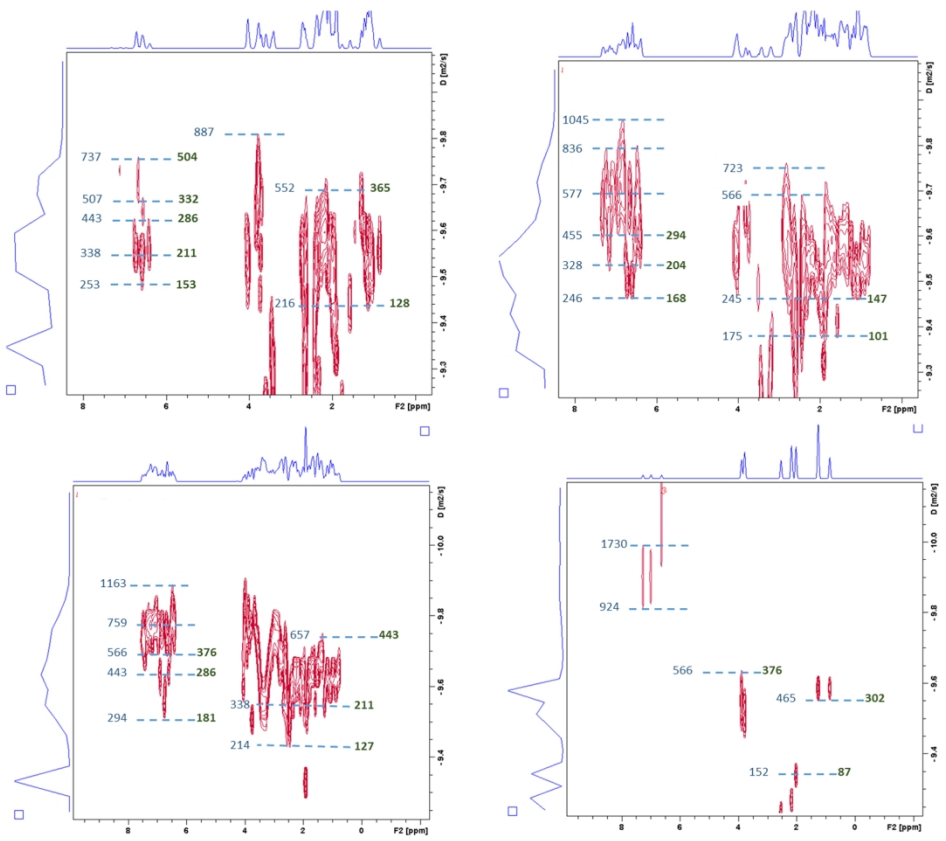


Figure 5. DOSY spectra for top left) sample 3; top right) sample 4; bottom left) sample 6; bottom right) sample 5. Green and blue figures correspond to molecular weights obtained with PEG and PS calibration respectively.

604x524mm (96 x 96 DPI)

Supporting information

Systematic Diffusion Ordered Spectroscopy for the Selective Determination of Molecular Weight in real Lignins and Fractions arising from Base-Catalyzed Depolymerization Reaction mixtures.

Alfonso Cornejo,† Íñigo García-Yoldi, † Irantzu Alegria-Dallo, ‡ Rebeca Galilea-Gonzalo, † Karina Hablich, † David Sánchez, ‡ Eduardo Otazu, ‡ Ibai Funcia, ‡ María J. Gil, † Víctor Martínez-Merino†*

† Institute for Advanced Materials (INAMAT)-Dpt. of Sciences, Campus de Arrosadia, Universidad Pública de Navarra, E31006 Pamplona, Spain

‡ National Renewable Energy Centre (CENER), Av. Ciudad de la Innovación 7, E31261 Sarriguren, Spain

Corresponding autor:

alfonso.cornejo@unavarra.es

Total number of pages: 16

Total number of Figures: 7

Total number of Tables: 3

TABLE OF CONTENTS

1. Monomer Groups and retention times

Table S1. Monomer groups and retention times.

Group	Compound	Retention time (min)
Internal Standard	Bromobenzene	8.67
Phenols	Phenol	9.22
Phenols	o-cresol	11.06
Phenols	p-cresol/m-cresol	11.55/11.88
Guaiacols	Guaiacylglycerol-β-guaiacyl ether	12.12
Vanillins	Vanillic acid	12.12
Guaiacols	Guaiacol	12.26
Phenols	2-ethylphenol	13.22
Phenols	3-ethylphenol	14.03
Phenols	4-ethylphenol	14.10
Catechols	Catechol	14.89
Others	3,4-dimethoxyphenol	14.90
Guaicols	4-methylguaiacol	14.95
Guaicols	4-ethylguaiacol	17.18
Catechols	4-methylcatechol	17.33
Syringols	Syringic acid	18.83
Syringols	2,6 dimethoxyphenol (Syringol)	18.88
Guaicols	4-Allylguaiacol, Eugenol	19.12
Phenols	4-hydroxybenzaldehyde	19.25
Guaicols	4-Propylguaiacol	19.33
Catechols	4-ethylcatechol	19.57
Vanillins	Vanillin	20.04
Others	Diphenylether	20.43
Guaicols	Guaiacylketone	22.40
Vanillins	Syringaldehyde	24.50
Vanillins	Acetosyringone	25.60
Syringols	4-methylsyringol	20.82
Vanillins	Acetovanillin	21.79

2. Optimization of the BCD reaction conditions

Table S2. Optimization of the BCD reaction conditions using 0.25 M NaOH (1 % w/v; pH 13.4) and NaOH/ POP_L 25:1. Error is given at 95 % confidence level.

T[a]	t (min)	Soluble fraction (%w)	Monomer (%w yield)	Phenol's (%)	Guaiacol's (%)	Catechol's (%)	Syringol's (%)	Vanillin's (%)
150	60	58	2.2	23	20	0	45	14
	120	63	2.3	27	21	0	40	13
	240	32	3.2	18	18	0	48	15
				22.7 ± 4.0	19.4 ± 1.5	0.0	43.9 ± 3.9	13.8 ± 1.0
175	0	53	1.4	25	16	0	37	21
175	60	61	4.9	24	29	0	39	7
175	120	67	6.8	24	26	0	40	10
175	180	67	8.2	20	29	0	37	9
175	240	39	5.1	27	27	0	37	8
175	300	50	7.1	27	30	0	34	11
				24.5 ± 2.2	26.1 ± 4.2	0.0	37.4 ± 2.0	11.1 ± 4.2
200	0	42	3.1	26	26	0	37	10
200	60	36	6.4	19	36	0	37	8
200	120	40	5.4	22	37	0	34	7
200	180	41	6.6	24	38	0	31	8
200	240	59	9.1	22	36	0	31	11
200	300	61	9.5	23	38	0	29	10
200	360	64	9.7	19	37	0	34	9
200	480	42	6.4	28	44	0	21	7
200	1200	34	5.4	23	48	0	23	6
				22.8 ± 1.9	37.9 ± 3.9	0.0	30.8 ± 3.7	8.5 ± 1.1
200 (PIN_L)	240		4.1	6	81	0	0	3
225	60	67	7.8	22	41	0	31	7
225	60	51	8.2	22	41	1	29	7
225	180	42	6.6	25	45	0	23	7
225	240	54	7.2	20	43	0	27	11
				22.3 ± 2.1	42.2 ± 1.7	0.2 ± 0.4	27.6 ± 3.1	7.8 ± 1.9
250	60	40	6.5	30	48	0	16	6
250	120	48	5.1	27	48	0	18	7
				28.1 ± 2.3	48.4 ± 0.1	0.0	16.8 ± 2.0	6.8 ± 0.5
275	0	51	7.5	19	40	0	31	9
275	60	57	4.9	43	35	8	5	10
275	120	36	2.4	59	26	6	3	6
275	240	58	5.9	65	19	17	0	0
300	0	40	4.4	50	33	6	2	9
300	60	35	3.2	55	18	15	5	6

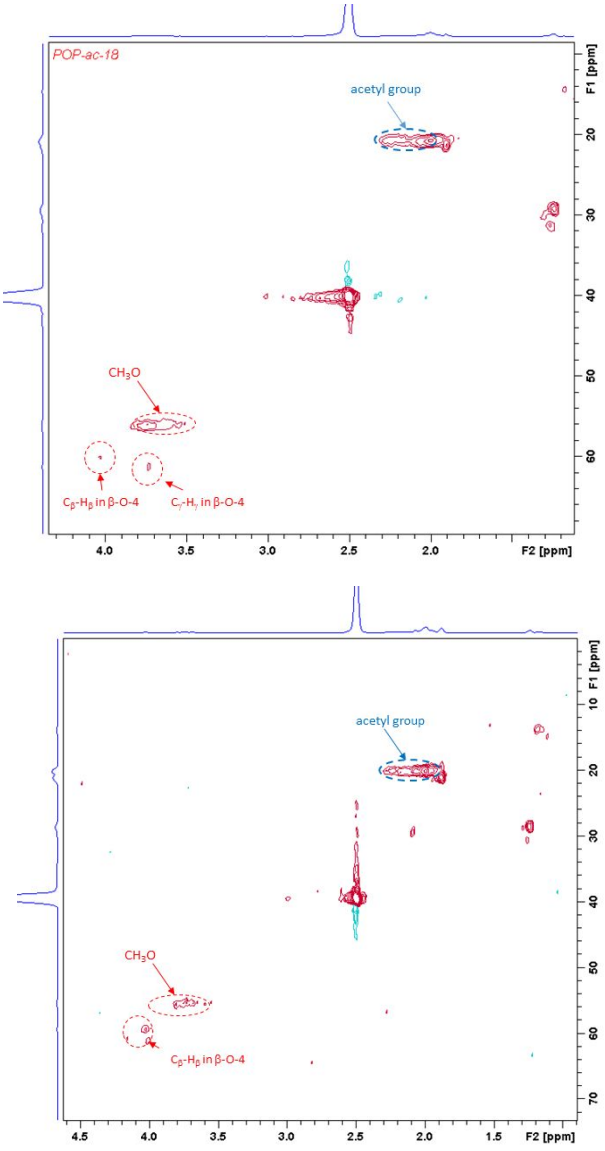
300	120	44	3.4	66	18	12	3	0
300	240	46	2.5	71	14	14	0	0

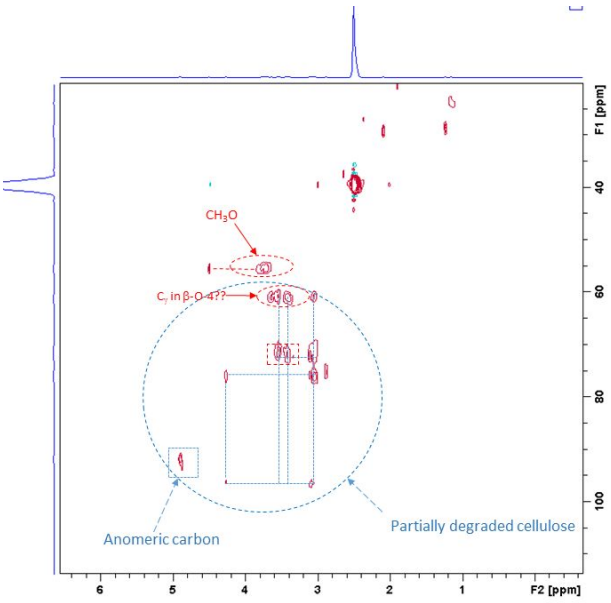
.

S accounts for syringyl groups; G account for guaiacyl groups; P accounts for phenyl units

3. HSQC-NMR spectra and SEC chromatograms

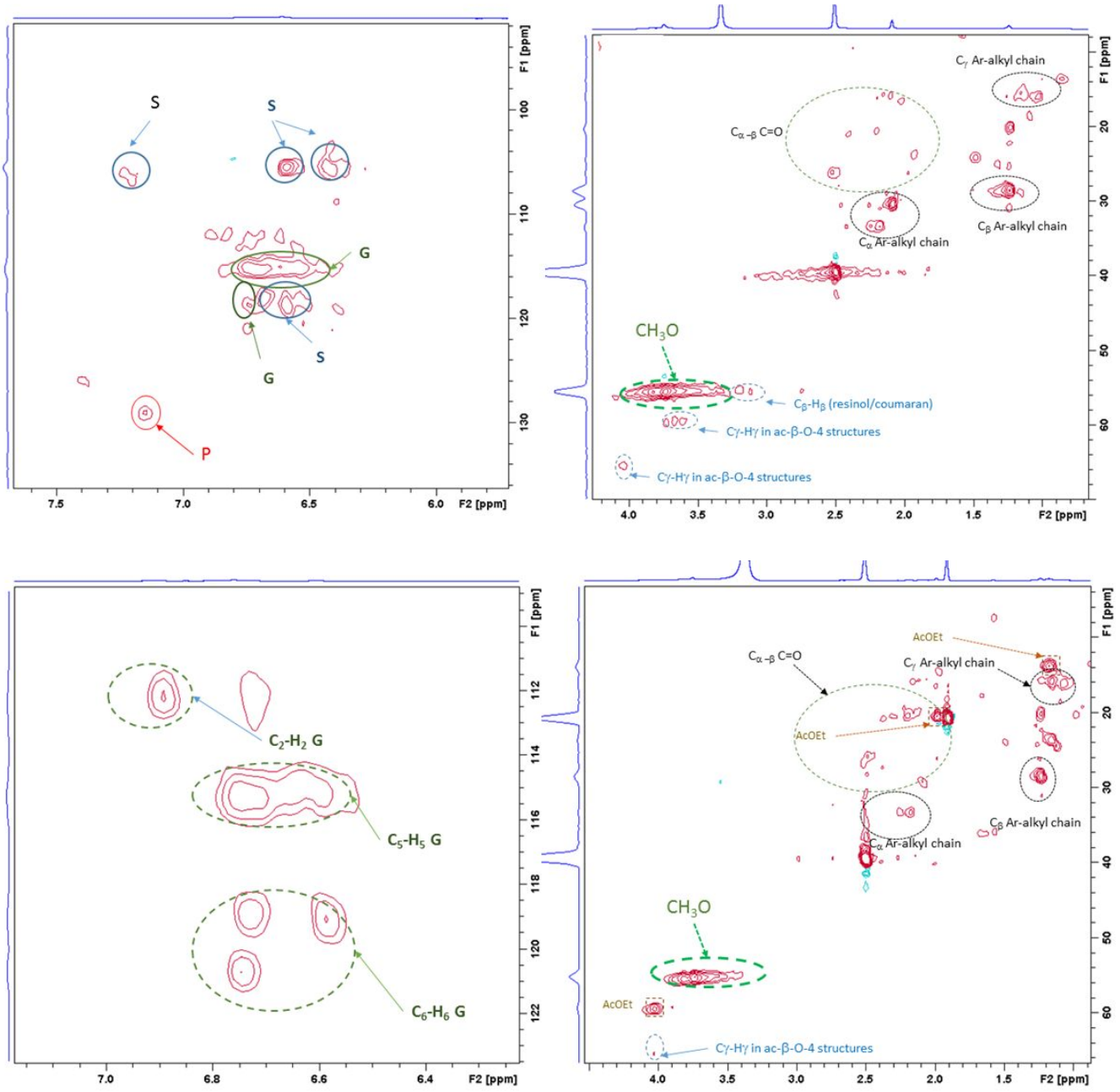
Figure S1. HSQC spectra for acetylated POP_L (top), acetylated PIN_L (middle) and HSQC-TOCSY for PIN_L




CH3O
Cγ in β-O-4??
Anomeric carbon
Partially degraded cellulose
F1 [ppm]
F2 [ppm]
40
60
80
100
6
5
4
3
2
1

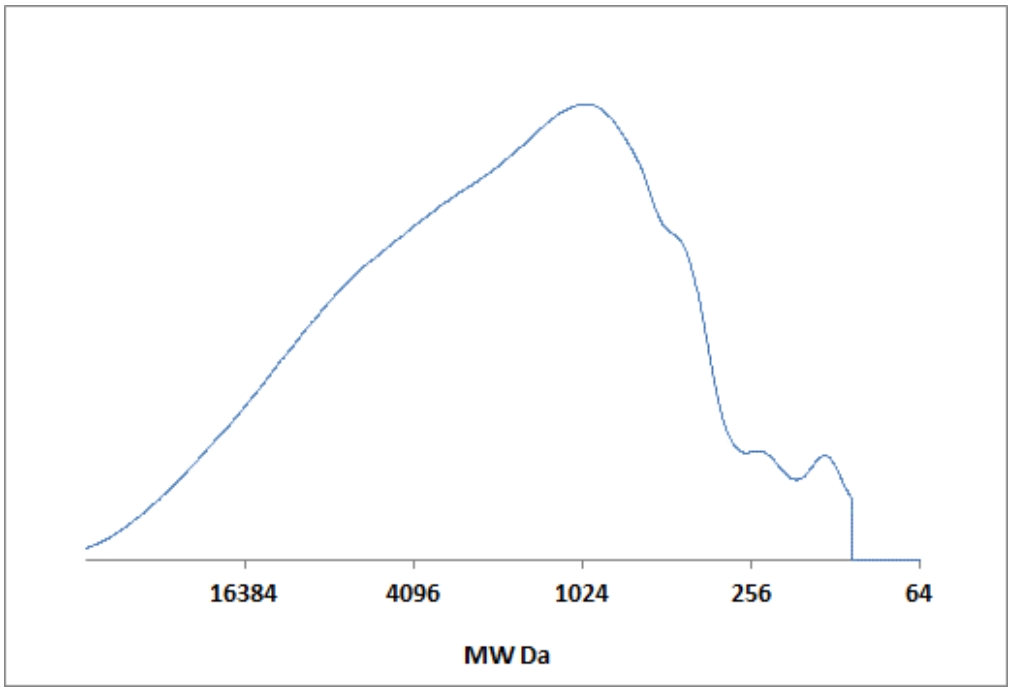
3. HSQC-NMR spectra and SEC chromatograms

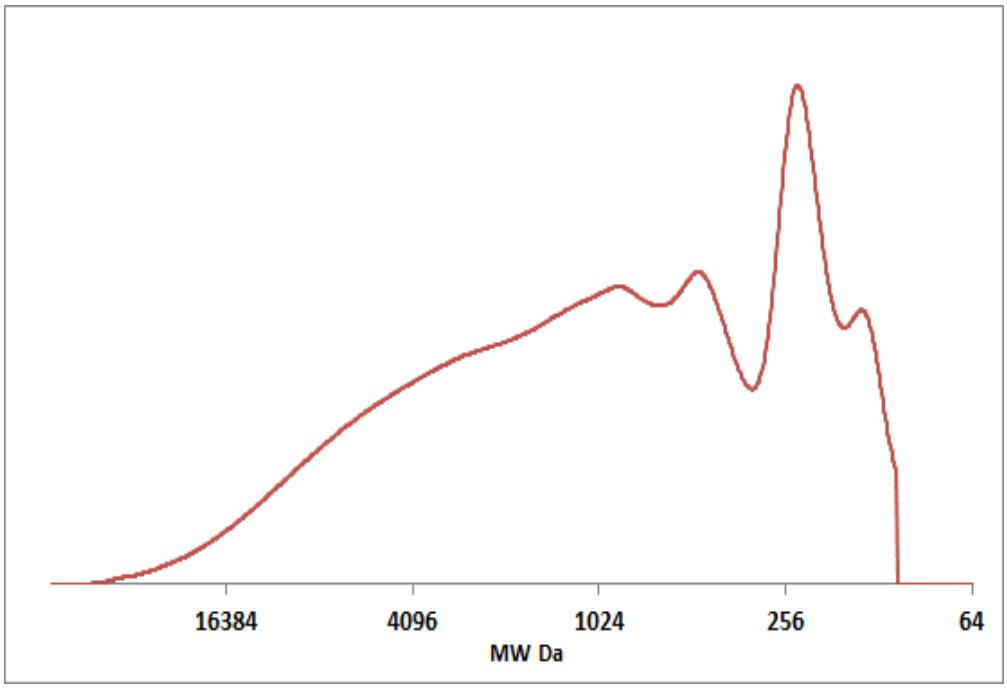
Figure S2. HSQC for sample **1** (top) and sample **6** (bottom)



3. HSQC-NMR spectra and SEC chromatograms

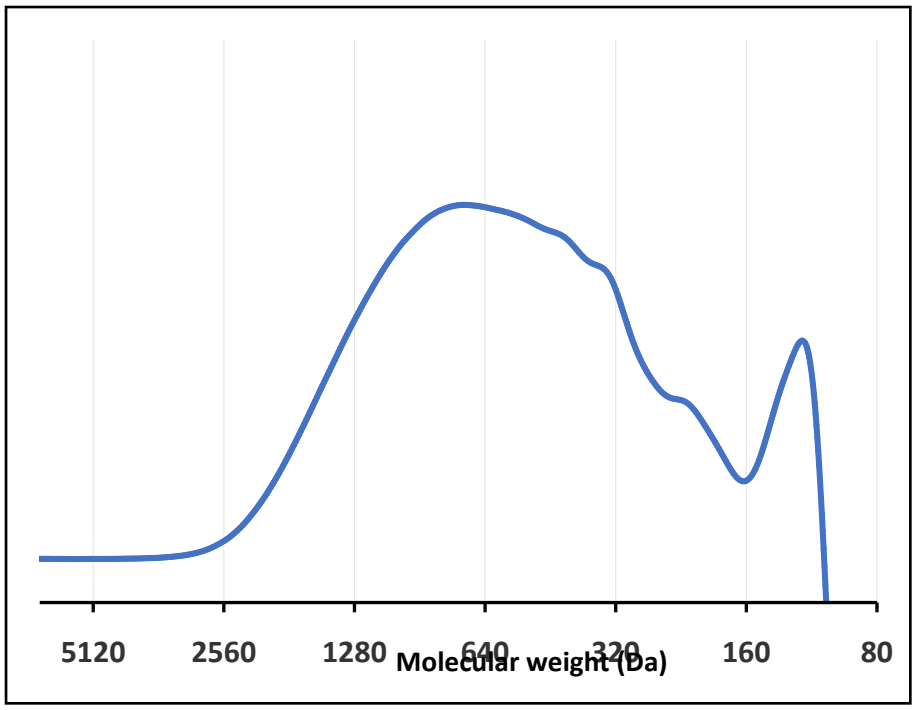
Figure S3. top) SEC for POP_L; bottom) SEC for PIN_L





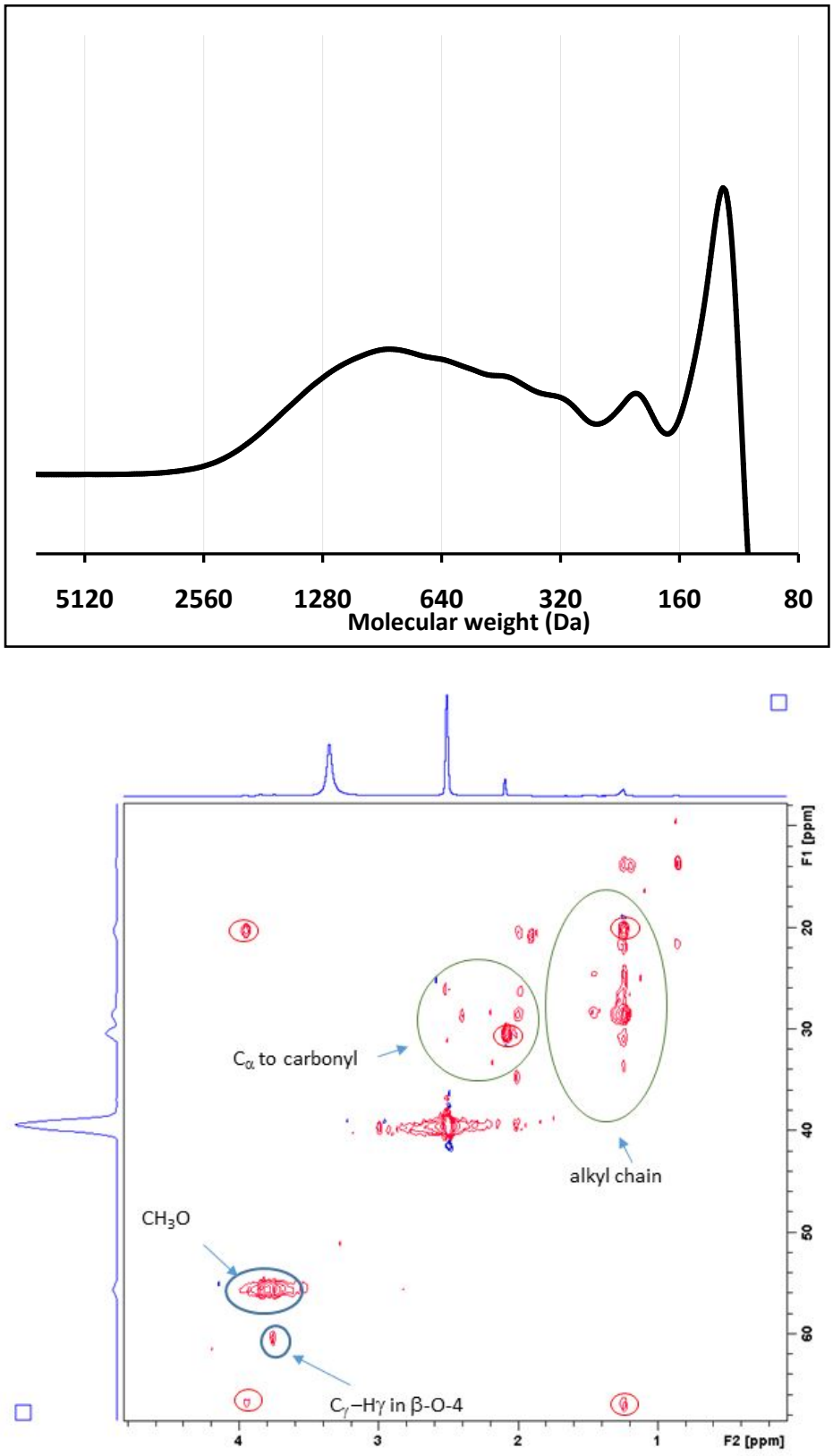
3. HSQC-NMR spectra and SEC chromatograms

Figure S4. SEC for BCD of PIN_L at 200 ºC for 240 min, sample 6.



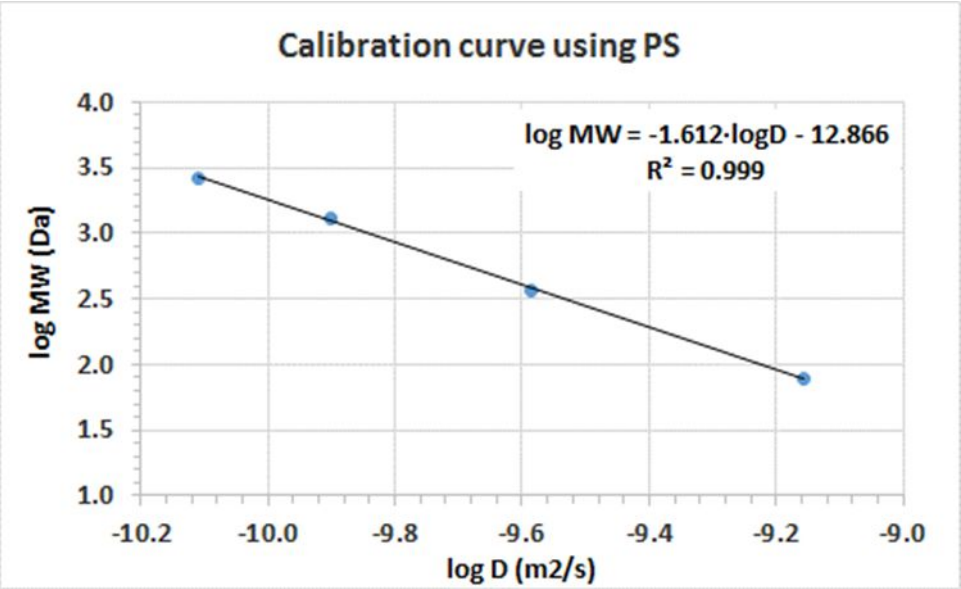
3. HSQC-NMR spectra and SEC chromatograms

Figure S5. Normalized SEC (top) and HSQC-TOCSY(bottom) for BCD of POP_L at 200 ºC for 30 min under microwave irradiation., sample **5**



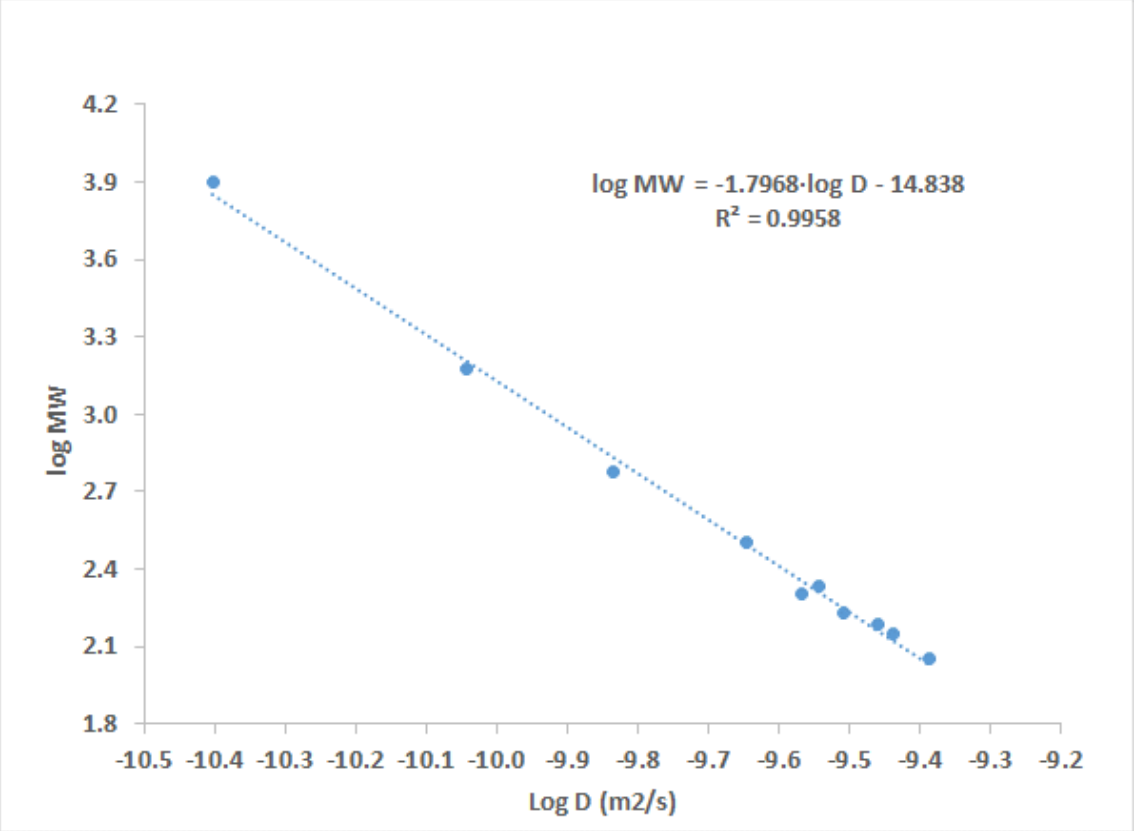
4. Calibration curve of logMW and logD using DOSY for PS

Compound	MW (Da)	log MW	log D ($m^2 \cdot s^{-1}$)
DMSO	78	1.892	-9.155
PS 370	370	2.568	-9.586
PS 1300	1300	3.114	-9.900
PS 2600	2600	3.415	-10.109



5. Calibration curve of log MW and log D using DOSY for PEG and monomeric phenols, VG and PE.

	MW (Da)	LogD (m_2/s)	log MW real	Predicted MW	MW error	MW % error
Guaiacol+H_2O	142	-9.440	2.153	133.4	-8.7	-6.1
Vanillin+H_2O	170	-9.509	2.231	177.9	7.7	4.5
Phenol+H_2O	112	-9.386	2.050	106.9	-5.2	-4.6
Syringol	154	-9.460	2.188	145.2	-8.9	-5.8
VG	320	-9.647	2.506	314.9	-5.5	-1.7
PE	214	-9.543	2.331	204.7	-9.6	-4.5
PEG200	200	-9.569	2.301	228.0	28.0	14.0
PEG600	600	-9.836	2.778	688.2	88.2	14.7
PEG1500	1500	-10.044	3.176	1627.5	127.5	8.5
PEG8000	8000	-10.403	3.903	7188.5	-811.5	-10.1



Multiple Linear Regression lunes, 18 de noviembre de 2019 16:21:21

Data source: Data 1 in Lignina2_calibracionDOSY_refDMSO.JNB

Log MW_5 = -14,837 - (1,797 $LogD_{dmso}$)

N = 10 Missing Observations = 1

R = 0,998 Rsqr = 0,996 Adj Rsqr = 0,995

Standard Error of Estimate = 0,040

	Coefficient	Std. Error	t	P	VIF
Constant	-14,837	0,401	-36,960	<0,001	
$LogD_{dmso}$	-1,797	0,0414	-43,363	<0,001	1,000

Analysis of Variance:

	DF	SS	MS	F	P
Regression	1	3,013	3,013	1880,343	<0,001
Residual	8	0,0128	0,00160		
Total	9	3,025	0,336		

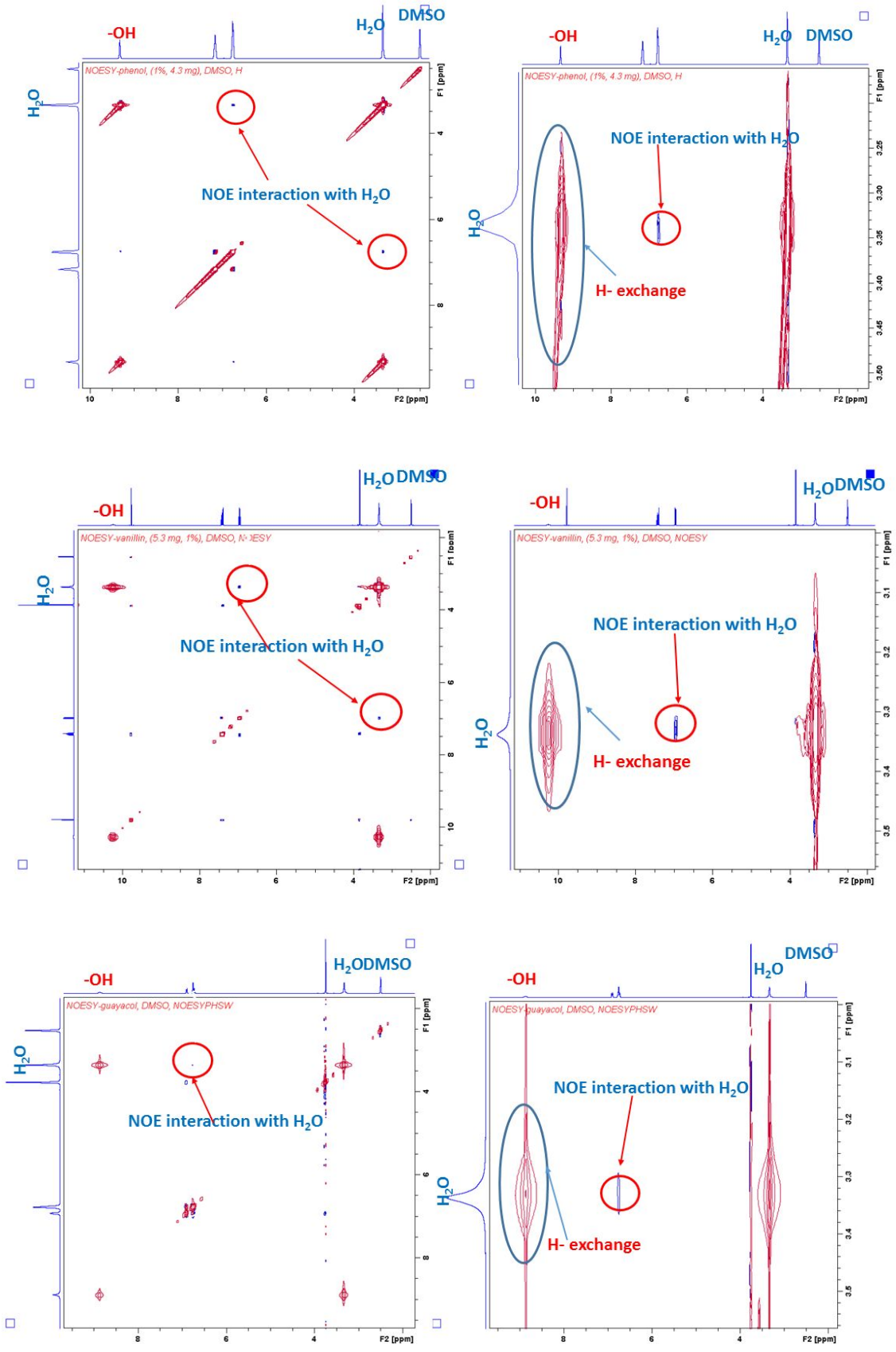
The dependent variable Col 7 can be predicted from a linear combination of the independent variables:

	P
$LogD_{dmso}$	<0,001

All independent variables appear to contribute to predicting $LogD_{dmso}$ ($P < 0.05$).
Normality Test (Shapiro-Wilk) Passed (P = 0,189)
Constant Variance Test (Spearman Rank Correlation): Passed (P = 0,199)
Power of performed test with alpha = 0,050: 1,000

6. NOESY spectra for phenol (top) vanillin (middle) and guaiacol (bottom)

Figure S6. NOESY spectra for phenol (top) vanillin (middle) and guaiacol (bottom).



7. Integration regions used in the determination of averaged diffusion coefficients

$$ln\left(\frac{I}{I_0}\right) = -\gamma^2\delta^2 {G_z}^2 D\left[\Delta + \left(\frac{4\delta}{3} + 3\tau/2\right)\right]$$

I_0: intensity at very low gradient value

I: intensity at a given gradient value

γ: gyromagnetic ratio

δ: length of the bipolar gradient pulse

Gz: gradient strength

Δ: time between pulses (d20)

τ: gradient ringdown delay

Table S3. Integration regions used in the determination of averaged diffusion coefficients. R^2 correspond to the correlation between experimental and calculated values for I.

	Aromatic		Aliphatic 1		Aliphatic 2	
Sample	Chemical shift (ppm)	R^2	Chemical shift (ppm)	R^2	Chemical shift (ppm)	R^2
POP_L	7.474 - 6.357	0.510	3.946 - 3.609	0.796	2.384 - 1.104	0.990
1	7.690 - 6.842	0.997	2.435 - 2.017	0.997	1.617 - 0.946	0.998
2	7.649 - 6.249	0.999	2.426 - 1.951	0.984	1.873 - 0.831	0.996
3	7.730 - 6.155	0.998	2.402 - 1.958	0.994		nm
4	7.663 - 6.33	1.000	2.453 - 1.951	0.995	1.851 - 0.726	0.999
5	nm	nm	3.178 - 2.720	0.999	2.384 - 0.593	0.997
PIN_L	8.134 - 6.424	0.991	4.714 - 3.730	0.990	2.249 - 0.647	0.987
6	7.717 - 6.31	0.999	2.411 - 1.966	0.997	1.832 - 0.74	0.999

8. ^{31}P NMR spectra for derivatized samples

Figure S7. ^{31}P spectra for derivatized samples 3 (top), 2 (middle),and 5 (bottom)