Electronic Supplementary Information for:

The synthesis and analysis of lignin-bound Hibbert ketone structures in technical lignins

D. M. Miles-Barrett†, A. R. Neal†, C. Hand, J. R. D. Montgomery, I Panovic, O. S. Ojo, C. S. Lancefield, D. B. Cordes, A. M. Z. Slawin, T. Lebl and N. J. Westwood[a]*

[a] School of Chemistry and Biomedical Sciences Research Complex, University of St. Andrews and EaStCHEM, St. Andrews, Fife, Scotland, KY16 9ST (UK) * E-mail: njw3@st-andrews.ac.uk

General Information

Chemical reagents were obtained from Sigma-Aldrich, Fisher Scientific and Arcos Organics and were used as received unless specified. Douglas fir and beech sawdusts were purchased from Hot Smoked (Useful Stuff Ltd.). All reactions conducted under inert conditions were carried out in flame dried glassware under a positive atmosphere of nitrogen. The dry solvents used were obtained from a solvent purification system (MBraun, SPS-800).

Thin-layer chromatography was conducted with glass backed TLC plates. Developed plates were air dried and viewed with a UV lamp (254 & 365 nm); where required the plates were developed with $KMnO_4$ solution. Column chromatography using silica was performed with Davisil® silica gel (40 - 63 µm, 230-400 mesh, VWR) with a glass column. Mass spectrometry data was acquired through the University of St Andrews School of Chemistry mass spectrometry service or EPSRC Swansea Mass Spectrometry Service. IR spectra were obtained on a Shimadzu IRAffinity-1 Fourier Transform IR spectrophotometer as thin films. IR analysis was carried out using IResolution v1.50 with only characteristic peaks reported. ^{1}H NMR and ^{13}C NMR was performed on a Bruker Ascend 400 MHz, Bruker Avance 500 MHz, Bruker Avance III 500 MHz or Bruker Ascend 700 MHz spectrometer with solvent peak used as internal standard. Multiplicities reported as following: s = singlet, d = doublet, t = triplet, q = quartet and m = multiplet and J values are reported in Hz. NMR spectra were processed on a MestReNova 10.0 Mac/ 9.0 Windows version or TopSpin 3.5 Mac version. Coloured 2D HSQC NMR were processed using Adobe Illustrator CS6 Mac version.

GPC analysis was carried out using a Shimadzu HPLC/GPC system equipped with a CBM-20A communications bus, DGU-20A degassing unit, LC-20AD pump, SIL-20A auto-sampler, CTO-20A column oven and SPD 20A UV-Vis detector. Samples were analysed using a Phenogel 5 um 50A (300 x 7.8 mm) and Phenogel 5 µm 500A (300 x 7.8 mm) columns connected in series and eluted with inhibitor free THF (1 mL/min) with a column oven temperature of 30 ^{o}C.

Lignin 2D HSQC experiments

Oven-dried lignin samples (80 mg) are dissolved in 0.65 mL of δ-solvent (DMSO) in a 1.5 mL eppendorf tube and subjected to sonication for 10 minutes at 30 °C. Samples are centrifuged at 6000 RPM for 5 minutes. In all cases no significant amount of precipitate was formed during centrifugation. Supernatant is filtered through a 0.45 µM syringe filter into an over dried NMR

tube. 2D HSQC NMR spectra were acquired using a Bruker Ascend 700 MHz (w/ cryoprobe (CPP TCI probe)) or Bruker Avance III 500 MHz (BBFO+ probe) using NMR methods described by Tran *et al.*[S1] The only difference between the protocol used by Tran *et al.*[S1] and this work is that the Bruker pulse sequence 'hsqcetgpsp.3' has been replaced with 'hsqcetgpsp.2', all other parameters are identical.

Lignin 2D HMBC experiments

The samples prepared for 2D HSQC experiments were used. Two-dimensional HMBC experiments were recorded on a Bruker 700 MHz Ascend III spectrometer equipped with an inverse gradient 1H/13C/14N triple resonance cryoprobe using the hmbcetgpl3nd sequence from the Bruker library. Spectra were acquired using 4096 x 256 data points; sweep widths of 14.0029 (^{1}H, F_2) and 239.9964 ppm (^{13}C, F_1), ns: 12.

Lignin 2D HSQC-TOCSY experiments

Two-dimensional HSQC-TOCSY experiments were recorded on a Bruker 700 MHz Ascend III spectrometer equipped with an inverse gradient 1H/13C/14N triple resonance cryoprobe using the hsqcdietgpsisp.2 sequence from the Bruker library. Spectra were acquired using 1024 x 256 data points; sweep widths of 15.9408 (^{1}H, F_2) and 80.0008 ppm (^{13}C, F_1), relaxation delay of 1.0 s and a TOCSY mixing time of 60 ms. The F_1 axis was centred on 80.0 ppm.

GC-MS/FID Analysis

Gas chromatography-mass spectrometry (GC-MS) was performed using a Thermo Scientific Trace GC Ultra 3.0 w/ AS3000 auto-sampler equipped with a DSQ II 3.0 (EI/CI+) mass spectrometer, a Resteck Rtx-1 (crossbond dimethyl polysiloxane) column (30 m x 0.25 mm) with a 0.25 μM-film using helium as a carrier gas. The standard method is a 1 μL injection, a split ratio of 20:1, a helium flow of 1.5 mL/s with a temperature profile starting with 50 °C 2-minute isotherm followed by a 15°C/min ramp for 16.5 minutes finishing at 300 °C (held for 2 minutes).

Gas chromatography-flame ionisation detector (GC-FID) was performed using a Thermo Scientific Trace 1300 series GC w/ TriPlus 100LS auto-sampler equipped with flame ionisation detector, a Restek Rtx-25 Amine column (30 m x 0.25 mm) with 0.5 μM-film using helium as

carrier gas. The standard method for analysis is a 1 μL injection, a split ratio of 50:1, a helium flow of 2 mL/s with a temperature profile starting with 60 °C 5-minute isotherm followed by a 10 °C/min ramp for 31 minutes finishing at 320 °C (held for 5 minutes).

General Experimental Procedures

A: Grignard addition to aldehyde

To a stirred solution of the aldehyde (1 eq.) in dry THF at 0 °C was added vinylmagnesium bromide solution (1.0 M in THF, 1.2 eq.). The reaction was warmed to room temperature and stirred until TLC analysis indicated no starting material remained. The reaction was quenched with a saturated solution of aqueous ammonium chloride. The solution was diluted with ethyl acetate and washed sequentially with saturated aqueous ammonium chloride solution, water and brine. The organic layer was dried over magnesium sulfate and concentrated *in vacuo*.

B: Upjohn dihydroxylation of allylic alcohol

To a stirred solution of the allylic alcohol (1 eq.) and NMO (1.5 – 1.90 eq.) in THF: H_2O (9:1) was added OsO_4 (2.5% by wt. in ^{t}Butanol). The reaction was stirred until TLC analysis indicated no starting material remained before quenching with saturated aqueous Na_2SO_3 solution. The mixture was extracted with ethyl acetate and washed sequentially with saturated aqueous Na_2SO_3 solution, water and saturated brine solution. The organic layer was dried over sodium sulfate and concentrated *in vacuo*. The crude product was purified on silica (20-40% acetone in petroleum ether) to afford the desired triol.

C: Acidolysis of triol

A stirred solution of the triol in dioxane: 2M HCl (9:1; 0.1 $MolL^{-1}$) was heated to reflux for the specified time. The reaction mixture was allowed to cool and concentrated under reduced pressure. Purification on silica gel (25-50 % acetone in petroleum ether) afforded the desired ketone.

D: TBS protection of phenol

To a stirred solution of the phenol (1 eq.); imidazole (1.1 - 2 eq.) and DMAP (0.05 eq.) in DCM (0.2 $molL^{-1}$) was added TBS-Cl. The reaction was left to stir at room temperature until TLC analysis indicated no starting material remained. The reaction was quenched with a saturated solution of aqueous ammonium chloride. The organic layer was washed sequentially with

water and brine before drying over magnesium sulfate and concentrating under reduced pressure.

Dioxasolv extraction of lignin

Following a literature procedure[S2]; To Douglas fir/ beech sawdust was added 1,4-dioxane: 0.5/2/4M HCl (9:1, 8 ml/g). This was heated at reflux for 1 hr and then filtered after cooling. The filter cake was washed with a mixture of 1,4-dioxane: 0.5/2/4M HCl (9:1). The filtrate was concentrated *in vacuo* giving a brown/ red gum. The gum-like material was then dissolved in a minimum amount of acetone: H_2O (9:1) and added dropwise to vigorously stirred H_2O (10x volume). The precipitate was collected and dried in a desiccator over $CaCl_2$. The dried precipitate was then dissolved in acetone: MeOH (9:1) and added dropwise to diethyl ether (10x volume). Precipitated dioxasolv lignin was further re-precipitated from acetone: MeOH (9:1) into ethyl acetate (10x volume) and dried in a vacuum oven overnight at 40 °C prior to analysis.

Note: Without further precipitation after diethyl ether precipitation, lignin samples were found to contain large quantities of low molecular weight contaminants. Subsequent re-precipitation into ethyl acetate rectified this problem and removed most/all low molecular weight contaminants.

Model reactions with $M(OTf)_x$

To a sealed tube was added the model compound (50 mg), 1,2-ethanediol (1 wt. equivalent) and 1,4-dioxane (1.5 mL). $M(OTf)_x$ (5 wt.%) was added and the tube sealed and heated to 140 °C for 15 minutes. Upon cooling, the mixture was concentrated *in vacuo* to give a gum-like residue. This residue was extracted with toluene: DCM (9:1, 2 x 10 mL) and the combined organic layers were washed with water (2 x 10 mL), dried over Na_2SO_4 and concentrated *in vacuo* prior to analysis.

Depolymerisation of Lignin using $M(OTf)_x$

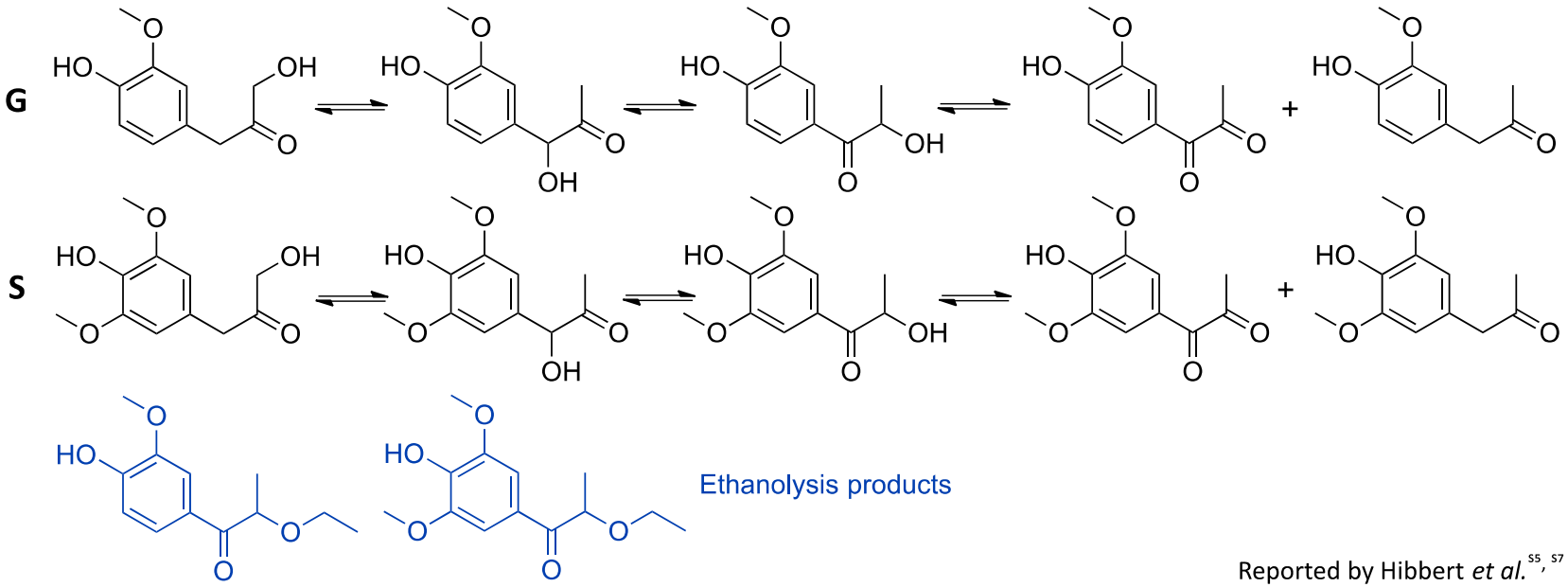
To a sealed tube containing DF lignin (100 mg) was added ethylene glycol (1 wt. equivalent) and 1,4-dioxane (3 mL). $M(OTf)_x$ (5 wt.%) was added and the sealed tube was heated to 140

°C for 15 minutes. Upon cooling, the mixture was concentrated *in vacuo* to give a gum-like residue. This residue was extracted with toluene: DCM (9:1, 4 x 5 mL) and the combined organic extracts were washed with water (2 x 10 mL), dried over Na_2SO_4 and concentrated *in vacuo*. For GC analysis, the total organic fraction and internal standard were dissolved in DCM (1 mL), passed through a 0.45 μM syringe filter and used directly for analysis.

GPC analysis of dioxasolv lignins

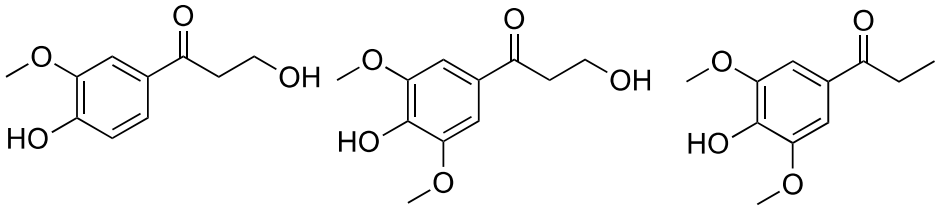
For GPC analysis, samples were acetylated using pyridine: acetic anhydride (1:1 v/v, 25 mg in 2 mL) for 16 hours. Samples were concentrated *in vacuo* using azeotropic distillation with ethanol (3 x 5 mL), toluene (3 x 5 mL) and $CHCl_3$ (3 x 5 mL). Samples were further dried in a vacuum oven for 16 hours before dissolving in HPLC grade (inhibitor free) THF (10 mg/ mL) and passing through a 0.45 μM syringe filter prior to analysis.

The Hibbert ketone family derived from the acidolysis of lignin



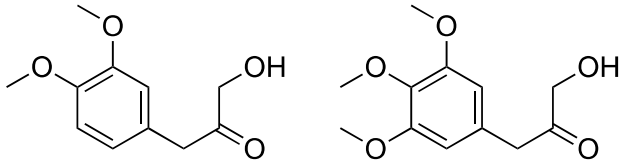
Reported by Hibbert *et al.*[S5, S7]

Other ketones derived from lignin using alternative depolymerisation methods *



Reported by Lancefield *et al.,* Angew. Chem. Int. Ed. Engl., 2014, 2-7 [S2]

Other models of the Hibbert ketones

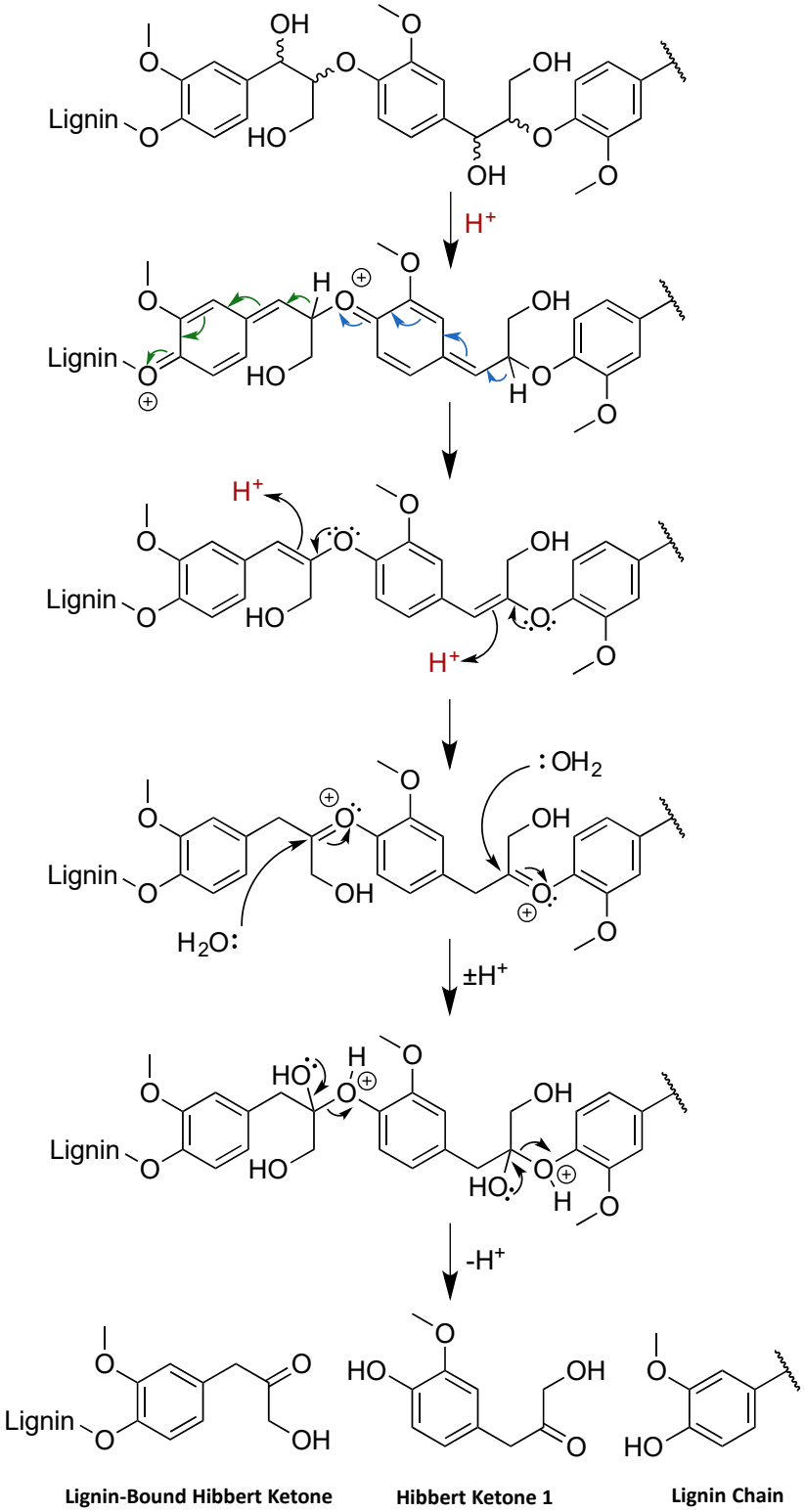


3 **4**

Methylated Hibbert ketones, models for Lignin-bound HK reported in this study

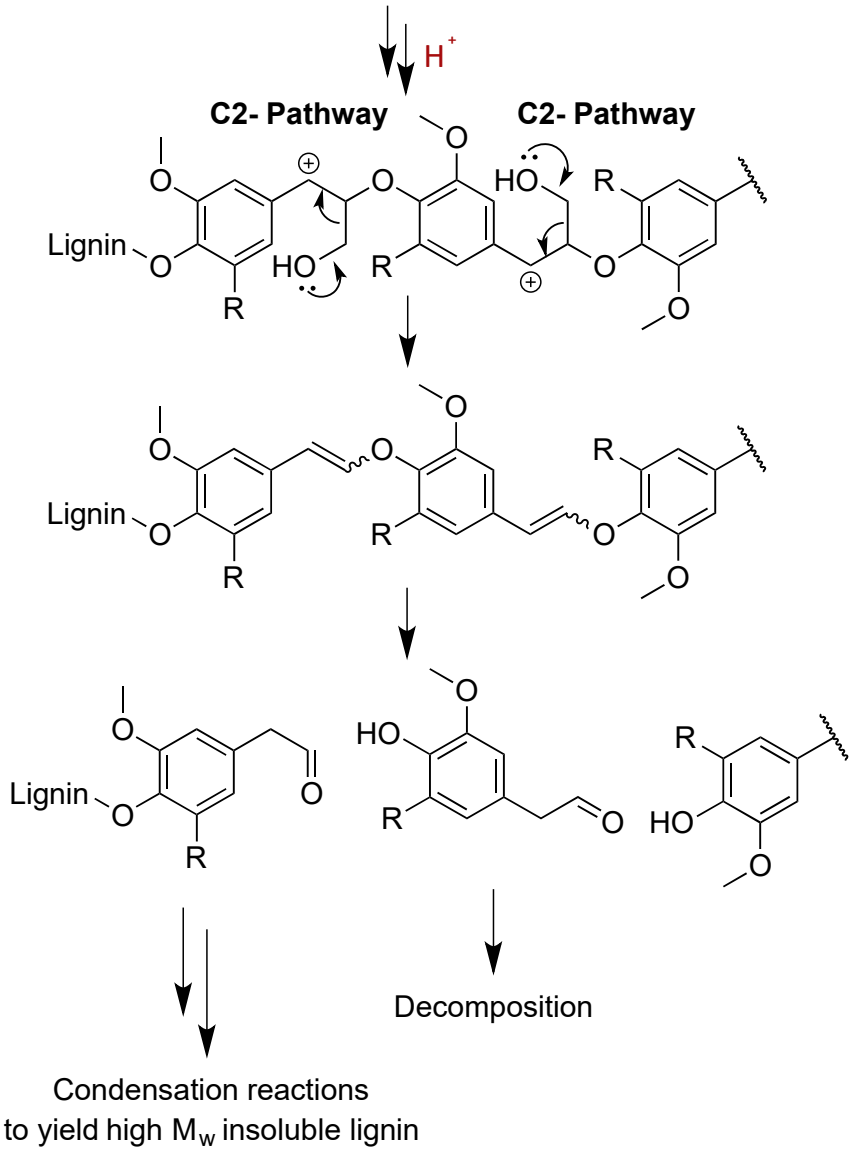
Figure S1: The Hibbert ketone family and other ketones derived from lignin. * Hibbert *et al.*[S3-S4] reported these structures but to the best of our knowledge they were not isolated from the acidolysis of lignins.

Proposed C3-C3, C2-C2 and C2-C3 degradation pathways

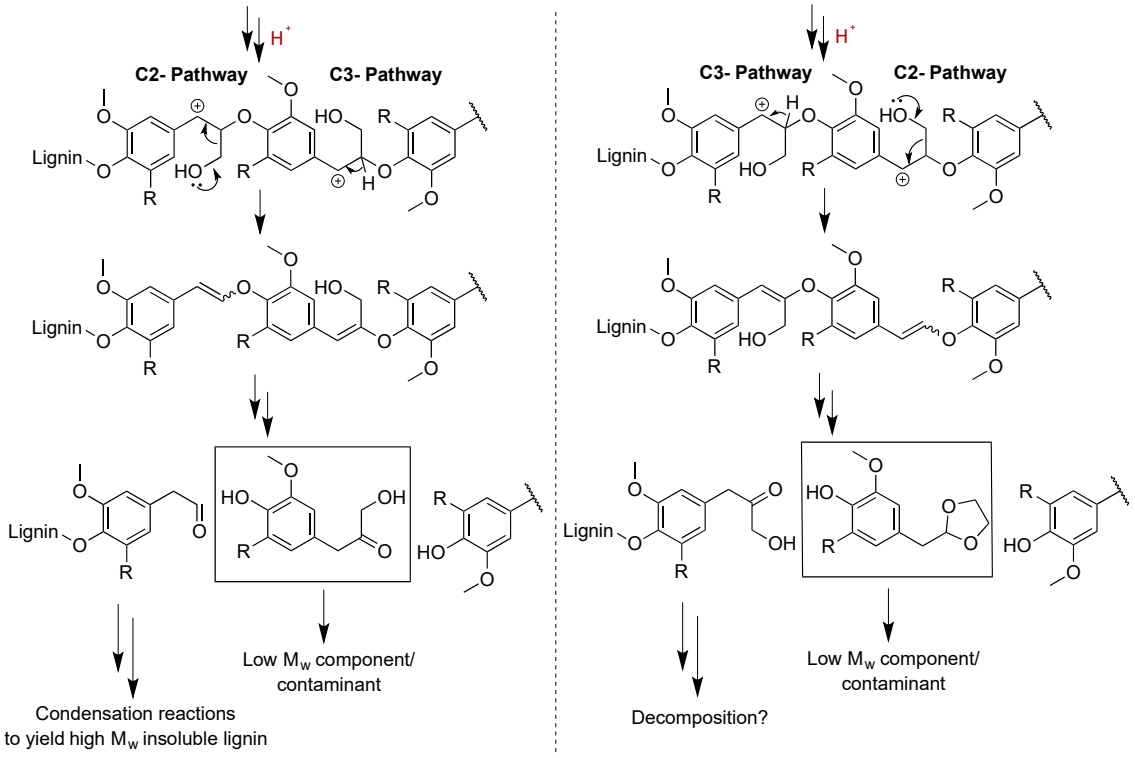


Scheme S1: Proposed Mechanism for generation of LBHK, HK **1** and the cleaved lignin chain during acidolysis of lignin/ lignocellulosic biomass.

From the acidolysis of lignin, generation of a LBHK structure and HK **1** or **2** requires two consecutive β-O-4 units. Generation of a LBHK structure alone does not require two consecutive β-O-4 units as cleavage of one β-O-4 linkage can generate a LBHK unit and release the rest of the lignin chain. Within the plant whole-cell wall, β-O-4 linkages are the most common linkage (softwoods: 45-50%, hardwoods: 60-62%, grasses: 74-84%) making the probability of two β-O-4's being in succession higher for hardwoods over softwoods.[S8]



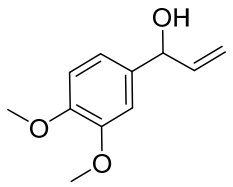
Scheme S2: Possible fates of lignin from a secondary carbocation to give C2- fragmentation products. R=H/OMe.



Scheme S3: Possible fates of lignin from a secondary carbocation to give C2-C3/ C3-C2 fragmentation products. R=H/OMe.

Experimental procedures for the synthesis of G- and S- non-phenolic Hibbert Ketones

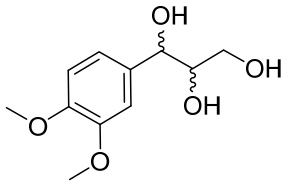
1-(3,4-dimethoxyphenyl)prop-2-en-1-ol (S1)



Followed general procedure **A** using: 3,4-dimethoxybenzaldehyde (**5**) (4.00 g, 24.1 mmol, 1 eq.); vinylmagnesium bromide solution (1.0 M in THF, 26.5 mL, 26.5 mmol, 1.1 eq.) and THF (120 mL) to yield **S1** as a yellow amorphous solid in quantitative yield.

^{1}H NMR (500 MHz; $CDCl_3$) δ 6.91 (dd, *J* 3.6, 1.8, 1H), 6.89 (d, *J* 1.8, 1H), 6.84 (d, *J* 8.1, 1H), 6.05 (ddd, *J* 17.1, 10.3, 5.9, 1H), 5.35 (dt, *J* 17.2, 1.5, 1H), 5.20 (dt, *J* 10.3, 1.4, 1H), 5.16 (s, 6-1H), 3.89 (s, 1H), 3.87 (s, 1H).Spectroscopic data was consistent with the literature.[S9]

1-(3,4-dimethoxyphenyl)propane-1,2,3-triol (6)

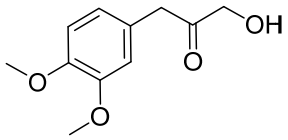


Followed general procedure **B** with **S1** (24.1 mmol, 1 eq.); NMO (4.23 g, 36.1 mmol, 1.5 eq.) and OsO_4 (0.2 mL) in THF: water (108 mL :10 mL) and the mixture was left to stir for 16 hours. Purification on silica afforded the desired triol **6** (2.14 g; 9.39 mmol; 39% - over 2 steps) as a colourless thick oil (3.0:1 mixture of diastereoisomers - deduced by relative integrals of peaks at 4.72 ppm and 4.56 ppm in ^{1}H NMR.)

IR (FTIR)ν_{max}: 3340 (b, O-H str), 2939 (C-H str), 1597, 1512 cm^{-1}; **HRMS** (NSI+) *m/z* [M + Na^+] calcd. for $C_{11}H_{16}O_6Na^+$ 251.0890, found 251.0890; **^{1}H NMR** (500 MHz, $CDCl_3$) δ 6.90 – 6.87 (m, 1H), 6.85 – 6.81 (m, 1H), 6.80 – 6.77 (m, 1H), 4.72 (d, *J* = 5.2 Hz, 0.75H), 4.56 (d, *J* = 7.1 Hz, 0.25H), 3.83 – 3.81 (m, 6H), 3.78 – 3.75 (m, 0.75H), 3.74 – 3.71 (m, 0.25H), 3.69 – 3.64 (m, 0.75H), 3.63 – 3.59 (m, 0.75H), 3.51 – 3.47 (m, 0.25H), 3.42 – 3.38 (m, 0.25H); **^{13}C NMR** (126

MHz, $CDCl_3$) δ 149.1, 148.9, 148.7, 133.1, 133.1, 119.2, 118.7, 111.1, 109.8, 109.5, 76.2, 75.4, 75.0, 74.8, 63.3, 63.0, 56.0.

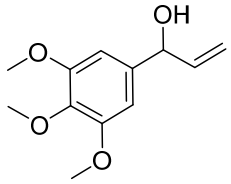
1-(3,4-dimethoxyphenyl)-3-hydroxypropan-2-one (3)



Followed general procedure **C** with **6** (1.00 g; 4.42 mmol) in 2 M HCl: dioxane (4.5 mL: 40 mL) at reflux for 1 hour. Purification on silica (20-40% acetone in petroleum ether) afforded the desired ketone **3** (0.42 g; 2.00 mmol; 45%).

IR (FTIR)ν_{max}: 3462 (b, O-H str), 2924 (b, C-H str), 1716 (C=O, m), 1514 (m) cm^{-1}; **HRMS** (NSI+) *m/z* [M + H^+] calcd. for $C_{11}H_{14}O_4H^+$ 211.0965; found 211.0965; **^{1}H NMR** (500 MHz, DMSO-d_6) δ 6.87 (d, *J* = 8.2 Hz, 1H), 6.79 (d, *J* = 2.0 Hz, 1H), 6.71 (dd, *J* = 8.2, 2.0 Hz, 1H), 4.15 (s, 2H), 3.72 (s, 6H), 3.64 (s, 2H); **^{13}C NMR** (126 MHz, DMSO-d_6) δ 208.6, 148.5, 147.6, 126.9, 121.6, 113.3, 111.8, 67.2, 55.5, 55.4, 44.2.

1-(4-((tert-butyldimethylsilyl)oxy)-3,5-dimethoxyphenyl)prop-2-en-1-ol (S2)

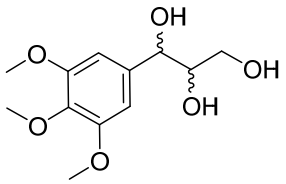


Followed general procedure **A** using: 3,4,5-trimethoxybenzaldehyde (**7**) (3.01 g, 15.4 mmol, 1 eq.); vinylmagnesium bromide solution (1.0 M in THF, 17.0 mL, 17.0 mmol, 1.1 eq.) and THF (75 mL) to yield **S2** as a yellow amorphous solid in quantitative yield.

^{1}H NMR (400 MHz, $CDCl_3$) δ 6.61 (d, *J* = 0.6 Hz, 2H), 6.04 (ddd, *J* = 17.1, 10.3, 6.0 Hz, 1H), 5.38 (dt, *J* = 17.1, 1.4 Hz, 1H), 5.22 (dt, *J* = 10.3, 1.3 Hz, 1H), 3.87 (s, 6H), 3.83 (s, 3H).

Spectroscopic data was consistent with the literature.[S10]

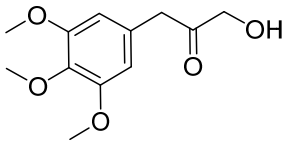
1-(3,4,5-trimethoxyphenyl)propane-1,2,3-triol (8)



Followed general procedure **B** with **S2** (15.4 mmol, 1 eq.)**;** NMO (2.70 g, 23.1 mmol, 1.5 eq.) and OsO_4 (0.2 mL) in THF: water (9:1; 75 mL) and the mixture was left to stir for 16 hours. Purification on silica afforded the desired triol **8** (1.43 g; 5.54 mmol; 36% - over 2 steps) as a colourless thick oil (2.6:1 mixture of diastereoisomers - deduced by relative integrals of peaks at 4.73 ppm and 4.60 ppm in ^{1}H NMR.)

IR (FTIR)ν_{max}: 3394 (b, O-H str), 2940 (b, C-H stretch), 1589 (m) cm^{-1}; **HRMS** (NSI+) *m/z* [M + Na^+] calcd. for $C_{12}H_{18}O_6Na^+$ 281.0996, found 281.0998; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 6.58 – 6.56 (m, 2H), 4.73 (d, *J* = 5.3 Hz, 0.72H), 4.60 (d, *J* = 6.5 Hz, 0.28H), 3.84 – 3.82 (m, 6H), 3.81 – 3.80 (m, 3H), 3.76 – 3.58 (m, 2.72H), 3.53 – 3.48 (m, 0.28H); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 153.4, 153.4, 137.6, 137.5, 136.4, 136.4, 103.6, 103.3, 75.9, 75.1, 74.8, 63.5, 63.2, 61.0, 56.3.

1-hydroxy-3-(3,4,5-trimethoxyphenyl)propan-2-one (4)

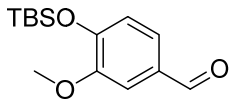


Followed general procedure **C** with **8** (1.01 g; 3.91 mmol) in 2 M HCl: dioxane (2.8 mL: 25 mL) at reflux for 1 hour. Purification on silica (20-40% acetone in petroleum ether) afforded the desired ketone **4** (0.35 g, 1.45 mmol, 37%).

IR (FTIR)ν_{max}: 3387 (b, O-H str), 2924 (b, OMe C-H str), 1720 (C=O, m), 1589 (m) cm^{-1}; **HRMS** (NSI+) *m/z* [M + H^+] calcd. for $C_{12}H_{17}O_5^+$ 241.1071, found 241.1070; **^{1}H NMR** (500 MHz, DMSO-d_6) δ 6.51 (s, 2H), 4.17 (s, 2H), 3.74 (s, 6H), 3.66 (s, 2H), 3.63 (s, 3H); **^{13}C NMR** (126 MHz, DMSO-d_6) δ 208.3, 152.7, 136.1, 130.2, 106.9, 67.3, 60.0, 55.8, 44.8.

Experimental procedures for the synthesis of the G- and S- phenolic Hibbert Ketones

4-((tert-butyldimethylsilyl)oxy)-3-methoxybenzaldehyde (10)

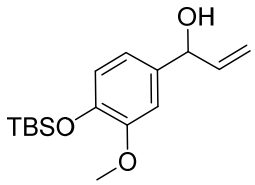


Followed general procedure **D** using **9** (5.00 g, 32.9 mmol, 1 eq.); imidazole (4.47 g, 65.7 mmol, 2 eq.); DMAP (0.20 g, 1.65 mmol, 0.05 eq.); TBS-Cl (5.95 g, 39.5 mmol, 1.2 eq) and DCM (165 mL) to yield **10** in quantitative yield.

^{1}H NMR (500 MHz; $CDCl_3$) δ 9.85 (s, 1H), 7.40 (d, *J* 1.9, 1H), 7.37 (dd, *J* 8.0, 1.9, 1H),6.96 (d, *J* 8.0, 1H), 3.87 (s, 3H), 1.00 (s, 9H), 0.19, (s, 6H).

Spectroscopic data was consistent with the literature.[S11]

1-(4-((tert-butyldimethylsilyl)oxy)-3-methoxyphenyl)prop-2-en-1-ol (S3)

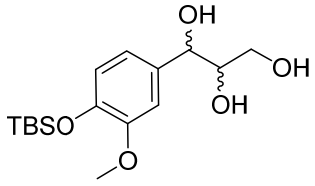


Followed general procedure **A** with **9** (32.9 mmol, 1 eq.) and vinylmagnesium bromide (1.0 M in THF, 39 mL, 1.2 eq.) in THF (150 mL) to give **S3** (9.10 g, 31.0 mmol, 94% - over 2 steps)

^{1}H NMR (500 MHz; $CDCl_3$) δ 6.89 (d, *J* 1.8, 1H), 6.82 (d, *J* 7.9, 1H), 6.80 (dd, *J* 8.2, 1.8, 1H), 6.05 (ddd, *J* 17.1, 10.3, 5.8, 1H), 5.34 (dt, *J* 17.1, 1.5, 1H), 5.19 (dt, *J* 10.3, 1.4, 1H), 5.14 (d, *J* 5.8, 1H), 3.81 (s, 3H), 0.99 (s, 9H), 0.14 (s, 6H).

Spectroscopic data was consistent with the literature.[S12]

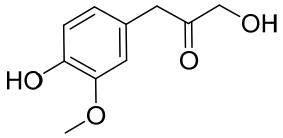
1-(4-((tert-butyldimethylsilyl)oxy)-3-methoxyphenyl)propane-1,2,3-triol (11)

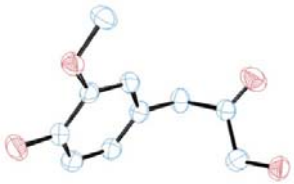


Followed general procedure **B** with **S3** (6.62 g, 22.5 mmol, 1 eq.); NMO (3.95 g, 33.8 mmol, 1.5 eq.) and OsO_4 (1.5 mL) in THF/H_2O (9:1; 115 mL). Purification on silica gel afforded **11** (4.82 g, 14.7 mmol, 65%) (2.1:1 mixture of diastereoisomers - deduced by relative integrals of peaks at 4.77 ppm and 4.61 ppm in ^{1}H NMR.).

IR (FTIR)ν_{max}: 3342 (b, O-H str), 2929 (b, OMe C-H str), 1508 (m), 1278 (s) cm^{-1}; **HRMS** (NSI+) *m/z* [M + Na$^+$] calcd. for $C_{16}H_{28}O_5SiNa^+$ 351.1598; found 351.1593; **^{1}H NMR** (500 MHz, $CDCl_3$) 6.90 – 6.87 (m, 1H), 6.85 – 6.82 (m, 1H), 6.81 – 6.78 (m, 1H), 4.77 (d, *J* = 5.5 Hz, 0.68H), 4.61 (d, *J* = 7.1 Hz, 0.32H), 3.80 – 3.78 (m, 3H), 3.77 – 3.70 (m, 1.68H), 3.68 – 3.65 (m, 0.68H), 3.59 – 3.55 (m, 0.32H), 3.49 – 3.45 (m, 0.32H), 0.99 – 0.98 (m, 9H), 0.15 – 0.13 (m, 6H); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 151.3, 145.0, 133.9, 133.8, 121.0, 119.2, 118.8, 110.4, 110.1, 76.1, 75.8, 75.1, 74.8, 63.4, 63.2, 55.6, 25.8, 18.6, -4.5.

1-hydroxy-3-(4-hydroxy-3-methoxyphenyl)propan-2-one (1)

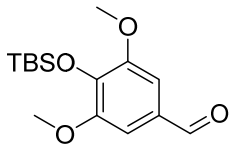




Followed general procedure **C** with **11** (1.28 g; 3.90 mmol) in 2 M HCl: dioxane (4 mL: 35 mL) at reflux for 0.5 hours. Purification on silica (20-40% acetone in petroleum ether) afforded the desired ketone **1** (0.52 g; 2.65 mmol; 68%). Slow evaporation of a sample of **1** from acetone afforded crystals which were sufficient for small molecule X-ray crystallographic analysis: CCDC 1500178.[S13]

IR (FTIR)ν_{max}: 3396 (b, O-H str, OMe C-H str), 2926 (b), 1716 (C=O, m), 1602 (m), 1514 (s), 1269 (s) cm^{-1}; **HRMS** (NSI+) *m/z* [M + Na$^+$] calcd. for $C_{10}H_{12}O_4Na^+$ 219.0628; found 219.0628; **^{1}H NMR** (500 MHz, $CDCl_3$) 6.87 (d, *J* = 8.5 Hz, 1H), 6.72 – 6.68 (m, 2H), 5.62 (s, 1H), 4.29 (s, 2H), 3.88 (s, 3H), 3.65 (s, 2H); **^{13}C NMR** (126 MHz, $CDCl_3$) 207.9, 146.9, 145.2, 124.5, 122.4, 114.8, 111.6, 67.6, 56.1, 45.6.

4-((*tert*-butyldimethylsilyl)oxy)-3,5-dimethoxybenzaldehyde (13)

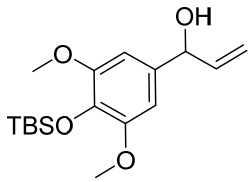


Followed general procedure **D** using: **12** (4.99 g, 27.4 mmol, 1 eq.); imidazole (3.73 g, 54.9 mmol, 2 eq.); DMAP (0.16 g, 1.31 mmol, 0.05 eq.); TBS-Cl (4.96 g, 32.9 mmol, 1.2 eq.) and DCM (135 mL) to yield **13** in quantitative yield.

^{1}H NMR (500 MHz; $CDCl_3$) δ 9.82 (s, 1H), 7.10 (s, 2H), 3.87 (s, 6H), 1.01 (s, 9H), 0.07 (s, 6H).

Spectroscopic data was consistent with the literature.[S14]

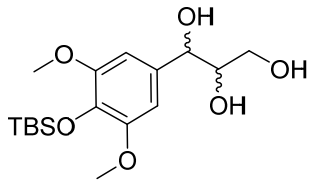
1-(4-((tert-butyldimethylsilyl)oxy)-3,5-dimethoxyphenyl)prop-2-en-1-ol (S4)



Followed general procedure **A** with **13** (27.4 mmol, 1 eq.) and vinylmagnesium bromide (1.0 M in THF, 33 mL, 1.2 eq.) in THF (150 mL). The crude product was purified on silica (5-20% EtOAc in petroleum ether) to afford the desired compound **S4** as a colourless oil (6.39 g, 19.7 mmol, 72% - over 2 steps).

HRMS (ESI) *m/z* $[M+Na]^+$ calcd. For $C_{17}H_{28}O_4SiNa^+$ 347.1655; found 347.1647; **IR** (neat) ν_{max}/cm^{-1}: 3491, 2928, 1591, 1506, 1119; **^{1}H NMR** (500 MHz; $CDCl_3$) δ 6.55 (s, 2H), 6.05 (ddd, *J* = 17.1, 10.3, 6.0, 1H), 5.35 (dt, *J* 17.1, 1.4, 1H), 5.20 (dt, *J* = 10.3, 1.4, 1H), 5.12 (d, *J* = 6.0, 1H), 3.79 (s, 6H), 1.00 (s, 9H), 0.12 (s, 6H); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 151.8, 140.3, 135.2, 134.0, 115.1, 103.5, 75.6, 55.9, 25.9, 18.9, -4.5.

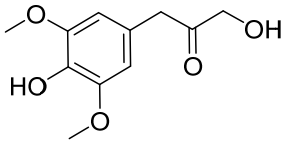
1-(4-((tert-butyldimethylsilyl)oxy)-3,5-dimethoxyphenyl)propane-1,2,3-triol (14)

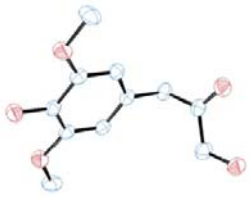


Followed general procedure **B** with **S4** (6.30 g, 19.4 mmol, 1 eq.); NMO (3.86 g, 33.0 mmol, 1.7 eq.) and OsO_4 (1 mL) in THF/H_2O (9:1; 105 mL). Purification on silica afforded **14** (5.13 g, 14.3 mmol, 74%) as a colourless oil (3.3:1 mixture of diastereoisomers - deduced by relative integrals of peaks at 4.74 ppm and 4.57 ppm in ^{1}H NMR.)

IR (FTIR)ν_{max}: 3363 (b, O-H str), 2929 (b, OMe C-H str), 1589 (m), 1508(m) cm^{-1}; **HRMS** (NSI+) *m/z* [M + NH_4^+] calcd. for $C_{17}H_{34}O_5NSi^+$ 376.2150; found 376.2148; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 6.55 – 6.53 (m, 2H), 4.74 (d, *J* = 5.3 Hz, 0.77H), 4.57 (d, *J* = 7.0 Hz, 0.23H), 3.77 (d, *J* = 2.5 Hz, 6.77H), 3.74 – 3.69 (m, 1H), 3.68 – 3.63 (m, 0.77H), 3.59 – 3.54 (m, 0.23H), 3.49 – 3.44 (m, 0.23H), 1.00 – 0.98 (m, 9H), 0.12 – 0.10 (m, 6H); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 134.2, 134.0, 133.0, 132.9, 103.6, 103.2, 76.0, 75.4, 74.8, 63.4, 63.2, 55.9, 25.9, 18.8, -4.5.

1-hydroxy-3-(4-hydroxy-3,5-dimethoxyphenyl)propan-2-one (2)

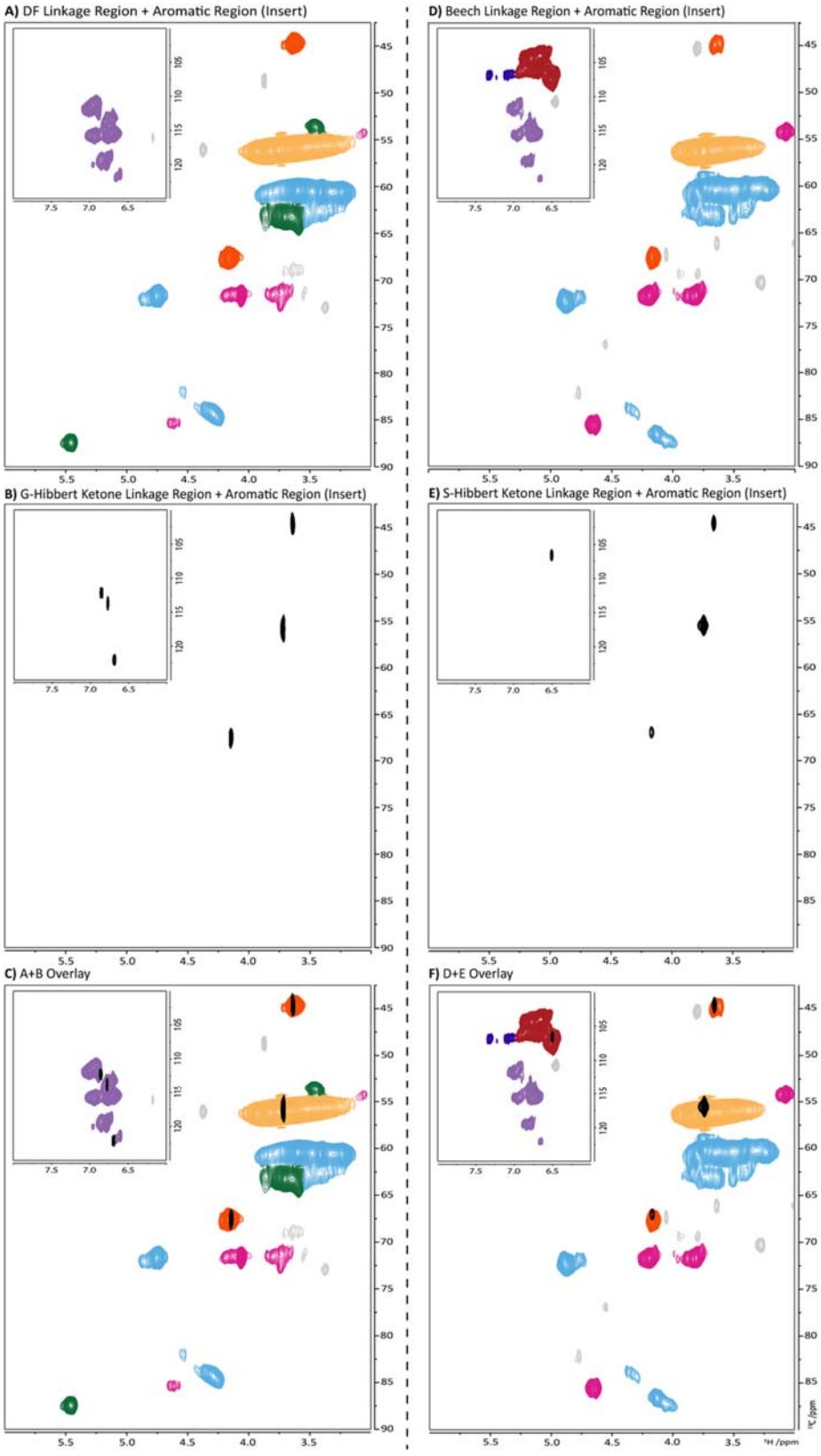




Followed general procedure **C** with **14** (1.40 g; 3.91 mmol) in 2 M HCl: dioxane (3.90 mL: 35 mL) at reflux for 0.5 hours. Purification on silica (20-40% acetone in petroleum ether) afforded the desired ketone **2** (0.42 g; 1.86 mmol; 48%). Slow evaporation of a sample of **2** from acetone afforded crystals which were sufficient for small molecule X-ray crystallographic analysis: CCDC 1500179.[S13]

IR (FTIR)ν_{max}: 3392 (b, O-H str), 2937 (b, OMe C-H str), 1716 (C=O, m), 1602 (m), 1458 (s) cm^{-1}; **HRMS** (NSI+) *m/z* [M + H^+] calcd. for $C_{11}H_{14}O_5H^+$ 227.0914; found 227.0914; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 6.42 (s, 2H), 5.49 (s, 1H), 4.30 (d, *J* = 4.8 Hz, 2H), 3.88 (s, 6H), 3.64 (t, *J* = 0.5 Hz, 2H), 2.99 (t, *J* = 4.8 Hz, 1H).

Assignment of Lignin-Bound Hibbert Ketones in Dioxasolv Lignins



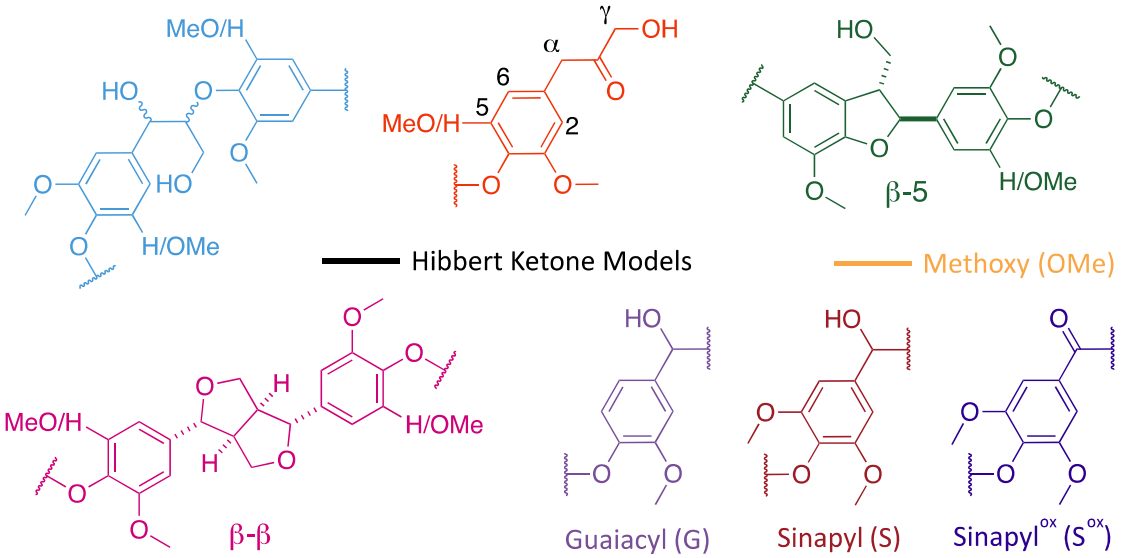


Figure S2: 2D HSQC NMR analysis (700 MHz, d_6-DMSO) of: **A)** DF linkage + aromatic region; **B**) G-Hibbert ketone model **3** linkage and aromatic region **C)** A+B overlay **D)** Beech linkage and aromatic region; **E**) S-Hibbert ketone model **4** linkage and aromatic region. **F**) D+E overlay.

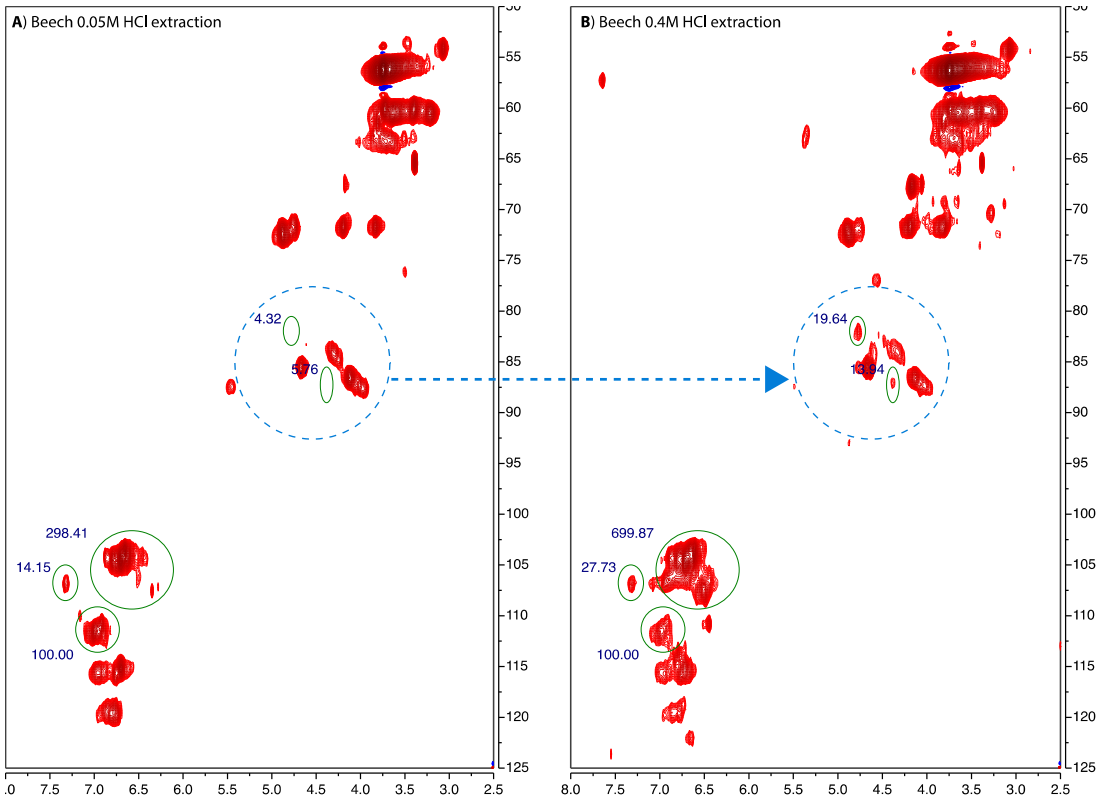


Figure S3: 2D HSQC NMR analysis (700 MHz, d_6-DMSO) of Beech dioxasolv extracted lignin examining *epi*-β–β content. **A**) Beech 0.05M HCl extraction; **B**) Beech 0.4M HCl extraction.

GPC Data

Table S1: GPC data for DF lignins extracted using different acid concentrations. M_n = number average molecular weight; M_w weighted average molecular weight. PDI = polydispersity index (M_w/M_n).

Concentration /M	M_n	M_w	PDI
0.05	1899	5422	2.85
0.2	2193	7832	3.57
0.4	1695	6510	3.84

Table S2: GPC data for Beech lignins extracted using different acid concentrations. M_n = number average molecular weight; M_w weighted average molecular weight. PDI = polydispersity index (M_w/M_n).

Concentration /M	M_n	M_w	PDI
0.05	4537	12053	2.66
0.2	2911	8915	3.06
0.4	1823	5966	3.27

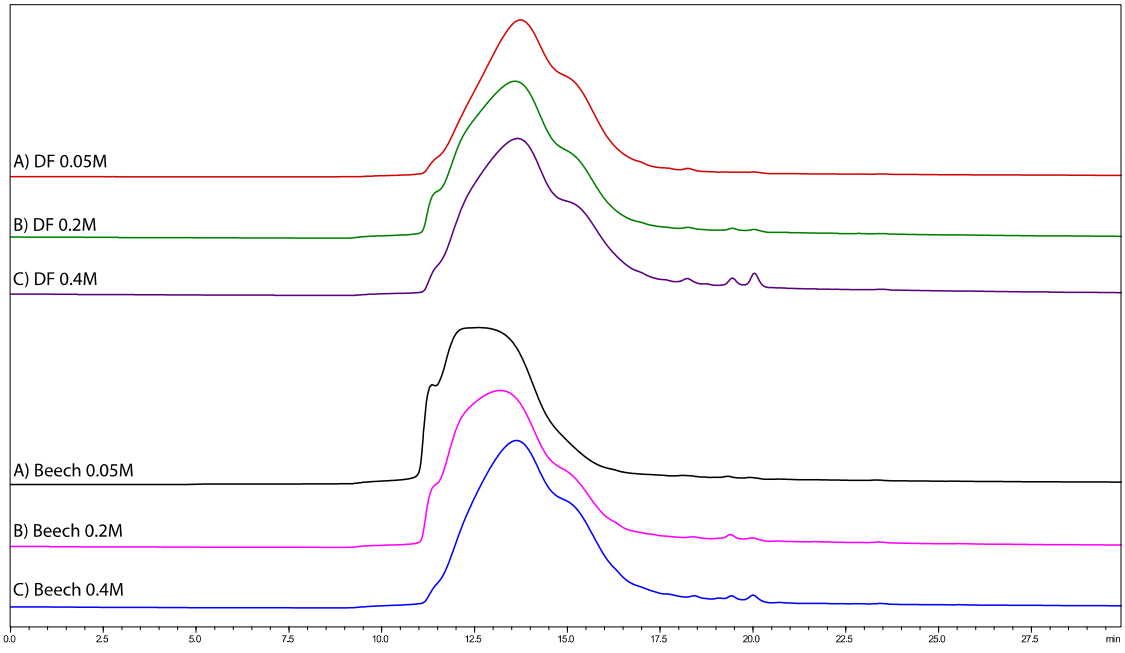


Figure S3a: GPC elution profiles of dioxasolv lignins plotted as UV response (280 nm) over time. This data was analyzed to give the values shown in Table S1-S2. It should be noted that the GPC calibration curve is set using polystyrene standards which are not a true representation of lignin's heterogeneous structure.

Raw Data from Lignin Dioxasolv Extractions

Douglas Fir (Manuscript Table 1): Extraction at each acid concentration was performed on a 10 g scale in triplicate and an example of the integral regions from 2D HSQC NMR used for analysis is provided below (Figure S4). Regions remain consistent throughout DF analysis. 2D HSQC NMR spectrum are phased in f_2. No other adjustments are made to spectra.

Table S3: Integral and error analysis of dioxasolv extractions (0.05M, 0.2M, 0.4M) of Douglas fir sawdust. Number *per* 100 C9 units values used in error analysis to represent the error based on the number of linkages in each lignin sample S.E: standard error. S.D: standard deviation. *Data from Figure S4

Repeat / Conc.		G_2	Aromatics	β-5	β-O-4	β–β	LBHK
1: 0.05M	Integral	100.00	100.00	14.0	37.0	4.7	7.5
2: 0.05M*		100.00	100.00	14.4	33.7	4.5	7.0
3: 0.05M		100.00	100.00	14.0	32.2	5.3	8.1
S.E		-	-	0.13	1.44	0.24	0.33
S.D		-	-	0.23	2.49	0.41	0.57
1: 0.2M	Integral	100.00	100.00	14.1	27.9	5.8	18.7
2: 0.2M		100.00	100.00	14.4	32.2	5.8	16.9
3: 0.2M		100.00	100.00	17.7	24.8	6.6	18.7
S.E		-	-	1.42	2.13	0.25	0.61
S.D		-	-	2.01	3.70	0.44	1.05
1: 0.4M	Integral	100.00	100.00	14.0	25.5	6.8	22.0
2: 0.4M		100.00	100.00	14.2	25.0	6.8	23.0
3: 0.4M		100.00	100.00	17.1	20.6	5.1	23.3
S.E		-	-	1.00	1.55	0.55	0.39
S.D		-	-	1.73	2.68	0.95	0.68

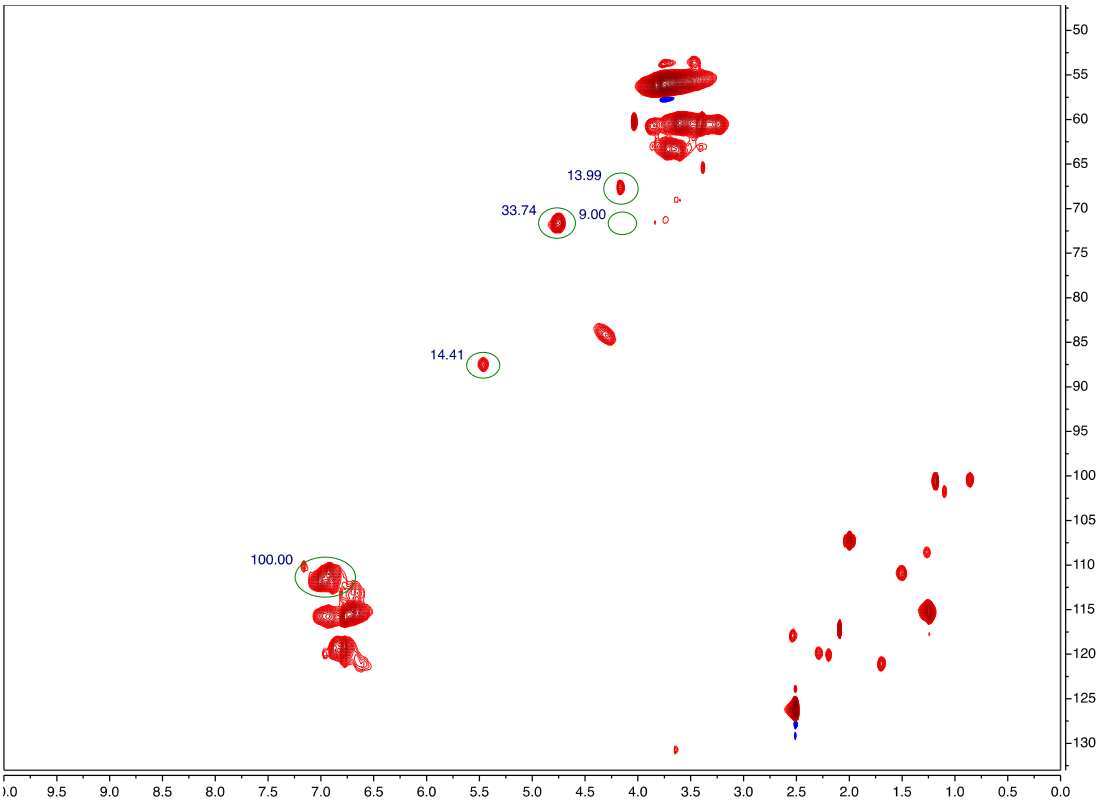


Figure S4: 2D HSQC NMR analysis (700 MHz, d_6-DMSO) of Douglas fir dioxasolv extracted lignin, *c.f.* Table S3, entry **2** for 0.05M concentration. This figure illustrates the integral regions used throughout to keep analysis consistent. Aliphatic region (δ_C 0-47 ppm) folds in the carbon dimension due to acquisition parameters. N.B. S2/6, β–β and HK integrals are halved for values as they correspond to protons each.

Beech (Manuscript Table 1): Extraction at each acid concentration was performed on a 10 g scale in triplicate and an example of the integral regions from 2D HSQC NMR used in analysis is provided below (Figure S5). Regions remain consistent throughout Beech analysis. NMR spectrum are phased in f_2. No other adjustments are made to spectra.

Table S4: Integral and error analysis of dioxasolv extractions (0.05M, 0.2M, 0.4M) of Beech sawdust. S.E: standard error. S.D: standard deviation. Number *per* 100 C9 units values used in error analysis to represent the error based on the number of linkages in each lignin sample. Values are calculated using the average of the aromatic integrals in the following equation: *Data from Figure S5

No. *per* 100 C9 units = (linkage integral/ average aromatic Integral from 3 extractions) *100

Repeat / Conc.		G_2	S2/6	$S2/6^{OX}$	Aromatics	β-5	β-O-4	β–β	HK
1: 0.05M	Integral	100.00	165.6	11.5	**277.1**	13.2	148.0	22.6	8.1
		No. per 100 C9 units				**4.8**	**53.4**	**8.2**	**2.9**
2: 0.05M		100.00	172.0	9.2	**281.2**	12.5	162.9	23.7	8.1
		No. per 100 C9 units				**4.5**	**57.9**	**8.4**	**2.9**
3: 0.05M*		100.00	176.6	11.4	**287.9**	12.4	161.2	24.0	8.6
		No. per 100 C9 units				**4.3**	**56.3**	**8.3**	**3.0**
S.E		-	3.19	0.75	3.17	0.13	2.44	0.08	0.04
S.D		-	5.53	1.30	5.49	0.22	4.23	0.14	0.07
1: 0.2M	Integral	100.00	354.2	23.2	**477.4**	15.1	188.7	49.2	38.5
		No. per 100 C9 units				**3.2**	**39.5**	**10.3**	**8.1**
2: 0.2M		100.00	252.8	14.3	**367.1**	13.4	169.3	36.5	17.4
		No. per 100 C9 units				**3.7**	**46.1**	**10.0**	**4.7**
3: 0.2M		100.00	334.1	24.7	**458.8**	14.0	150.6	52.4	37.8
		No. per 100 C9 units				**3.1**	**32.8**	**11.4**	**8.2**
S.E		-	30.99	3.25	34.09	0.18	3.84	0.44	1.14
S.D		-	53.68	5.62	59.05	0.32	6.65	0.76	1.97
1: 0.4M	Integral	100.00	314.4	25.4	**439.8**	15.1	129.5	47.6	49.3
		No. per 100 C9 units				**3.4**	**29.5**	**10.8**	**11.2**
2: 0.4M		100.00	321.9	22.2	**444.1**	14.3	146.88	46.8	45.3
		No. per 100 C9 units				**3.2**	**33.1**	**10.5**	**10.2**
3: 0.4M		100.00	380.8	12.61	**506.8**	16.2	161.8	48.4	41.0
		No. per 100 C9 units				**3.2**	**31.9**	**9.5**	**8.1**
S.E		-	21.00	1.16	21.64	0.08	1.07	0.39	0.92
S.D		-	36.38	2.02	37.49	0.13	1.85	0.67	1.59

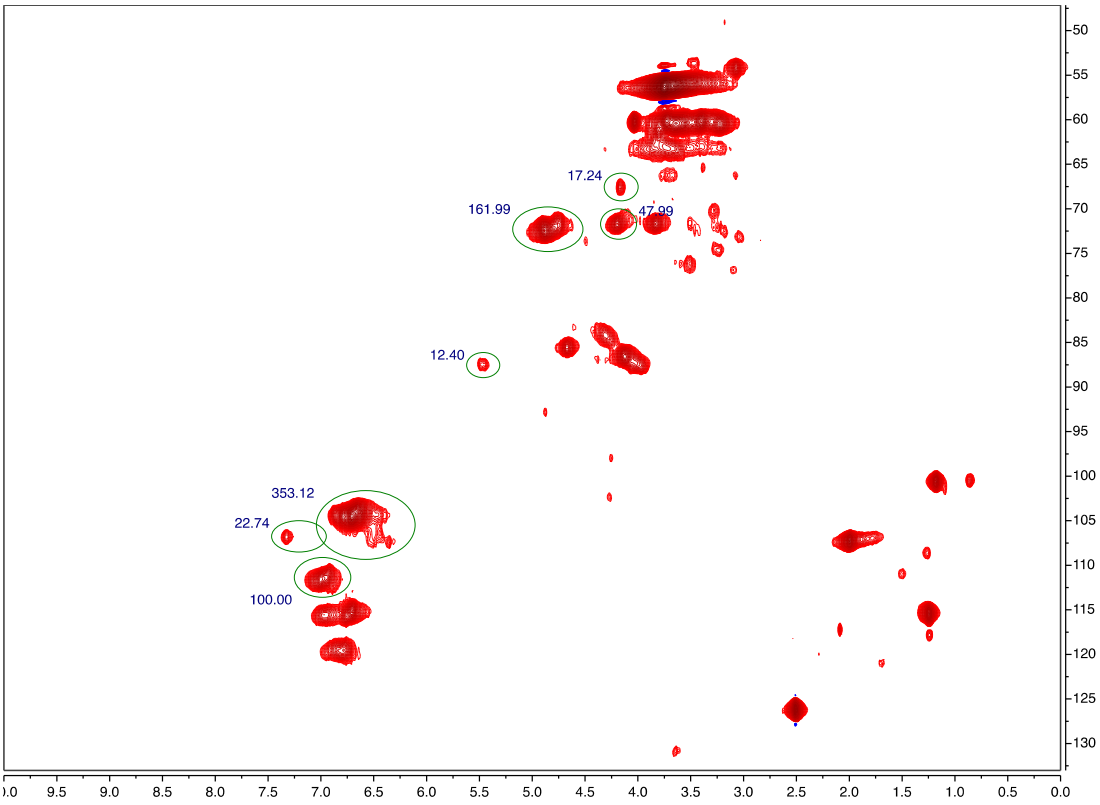


Figure S5: 2D HSQC NMR analysis (700 MHz, d_6-DMSO) of Beech dioxasolv extracted lignin, *c.f.* Table S4, entry **3** for 0.05M concentration. This figure illustrates the integral regions used throughout to keep analysis consistent. Aliphatic region (δ_C 0-47 ppm) folds in the carbon dimension due to acquisition parameters. N.B. S2/6, β–β and HK integrals are halved as they correspond to 2 protons each.

D_1 experiments to eliminate T_1 issues in NMR

In an attempt to eliminate concerns over T_1 (spin lattice relaxation for large molecules such as lignin is thought to be slower, hence T_1 gets bigger) within the 2D HSQC NMRs of lignins, comparison experiments were run comparing the integrals when D_1 (delay time/relaxation delay) = 1s and 15s. As the values were found to be similar, it was assumed that the use of the D_1=1s for all lignin samples would deliver sufficiently quantitative data for the purposes of this study.

Table S5: Integral Data for D_1 = 1s and D_1 = 15s experiments of Douglas fir lignin samples. Integrations relative to aromatic cross-peak (G_2 = 100)

Conc. / M		β-5	β-O-4	β-β	LBHK
0.05	**D_1=1s**	14.14	32.99	5.24	7.71
	D_1=15s	13.93	35.97	4.45	6.20
Difference		**-0.21**	**+2.98**	**-0.79**	**-1.51**
0.2	**D_1=1s**	14.76	25.93	6.91	15.34
	D_1=15s	13.64	26.14	5.40	11.48
Difference		**-1.12**	**+0.21**	**-1.52**	**-3.87**
0.4	**D_1=1s**	13.49	26.53	6.84	20.95
	D_1=15s	12.46	27.73	5.43	16.10
Difference		**-1.03**	**+1.20**	**-1.41**	**-4.85**

Interestingly, the difference data between D_1=1s and D_1=15s for DF shows the following information:

- β-5, β-β and LBHK integrals all decrease as D_1 increases. LBHK integrals reduce more than β-5 and β-β, consistent with end-groups within lignin having local mobility and relaxing differently.
- The β-O-4 integral is the only group increasing suggesting this linkage is under-represented in this lignin analysis.
- All differences were relatively small suggesting that having a D_1=1s or D_1=15s has no major impact on the conclusions drawn in this study.

Table S6: Integral Data for D_1 = 1s and D_1 = 15s experiments of Beech lignin samples. **N.B.** Data is reported only as raw integral data relative to aromatic cross-peaks.

Conc./ M		G	S2/6	S2/6^{ox}	Aromatics	β-5	β-O-4	β–β	LBHK
0.05	**D_1=1s**	100.00	198.97	9.62	308.59	13.50	113.29	30.99	12.20
	D_1=15s	100.00	201.05	7.47	308.52	12.93	121.73	24.08	8.82
	Difference	-	**+2.08**	**-2.15**	**-0.07**	**-0.57**	**+8.44**	**-6.92**	**-3.38**
0.2	**D_1=1s**	100.00	335.48	15.14	450.61	14.01	139.12	51.05	40.21
	D_1=15s	100.00	342.86	13.11	455.97	14.02	148.97	39.09	31.12
	Difference	-	**+7.38**	**-2.03**	**+5.35**	**-0.01**	**+9.85**	**-11.96**	**-9.09**
0.4	**D_1=1s**	100.00	359.76	12.37	472.12	14.92	126.03	49.24	48.29
	D_1=15s	100.00	362.90	11.61	474.51	12.86	135.31	38.09	37.41
	Difference	-	**+3.14**	**-0.76**	**+2.38**	**-2.06**	**+9.28**	**-11.16**	**-10.88**

The difference data between D_1=1s and D_1=15s for Beech shows the following information:

- The S2/6 and S2/6^{ox} increase and decrease respectively but the change is small (approximately 2% at most).
- β-5, β-β and LBHK integrals all decrease. LBHK integrals reduce more than β-5 and β-β, again consistent with end groups within lignin having local mobility and relaxing differently.
- The β-O-4 integral is the only group increasing suggesting this linkage is under-represented in this lignin analysis.
- All differences, besides LBHK are small suggesting that having a D_1=1s or D_1=15s has no major impact on the conclusions drawn in this study.

HMBC Analysis of dioxasolv lignins containing low M_w impurities of HKs 1 & 2

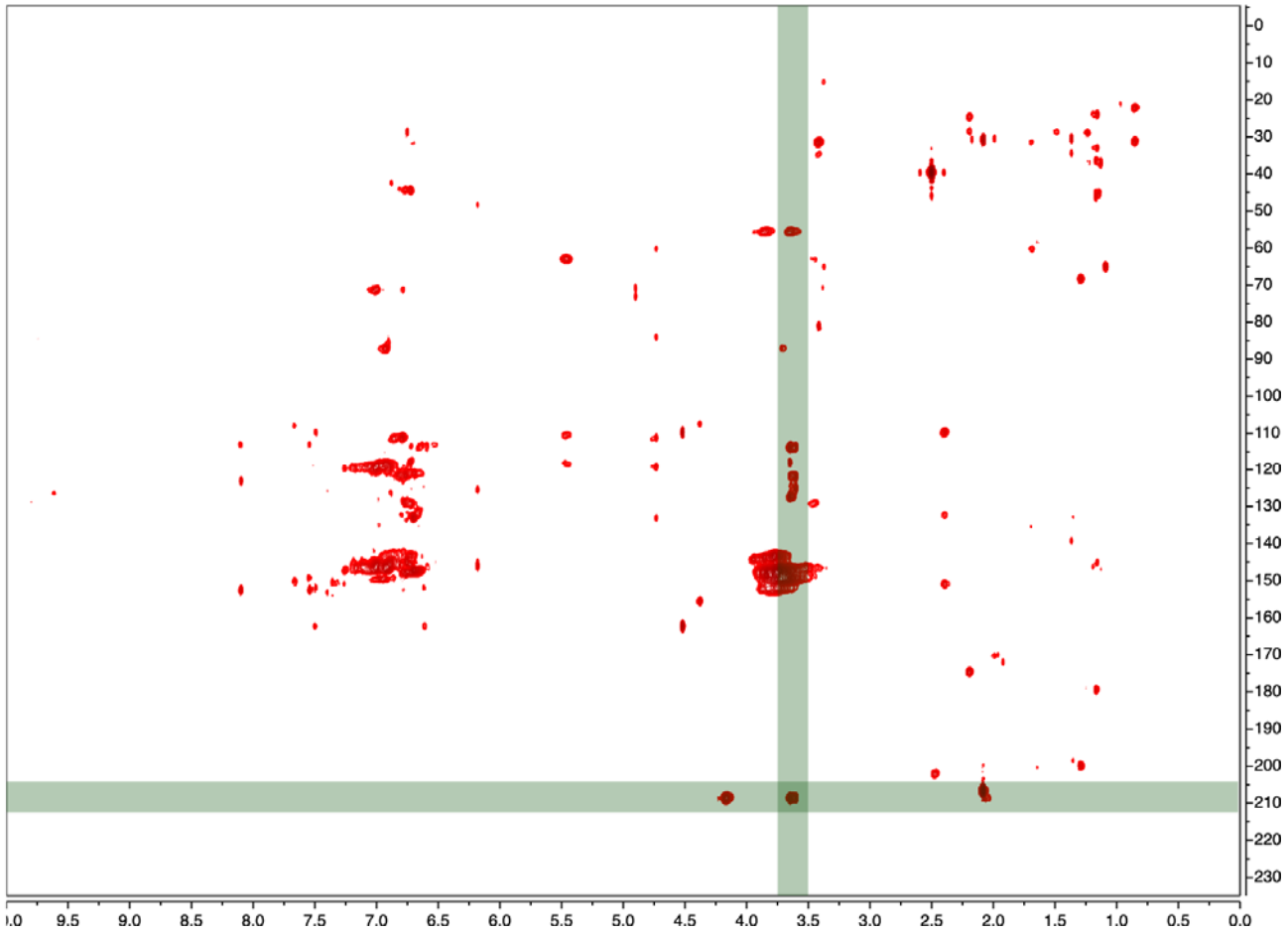


Figure S6: 2D HMBC NMR analysis (700 MHz, d_6-DMSO) of DF lignin (0.4M dioxasolv extracted). Highlighted is the region examined in Figures S8 and S10.

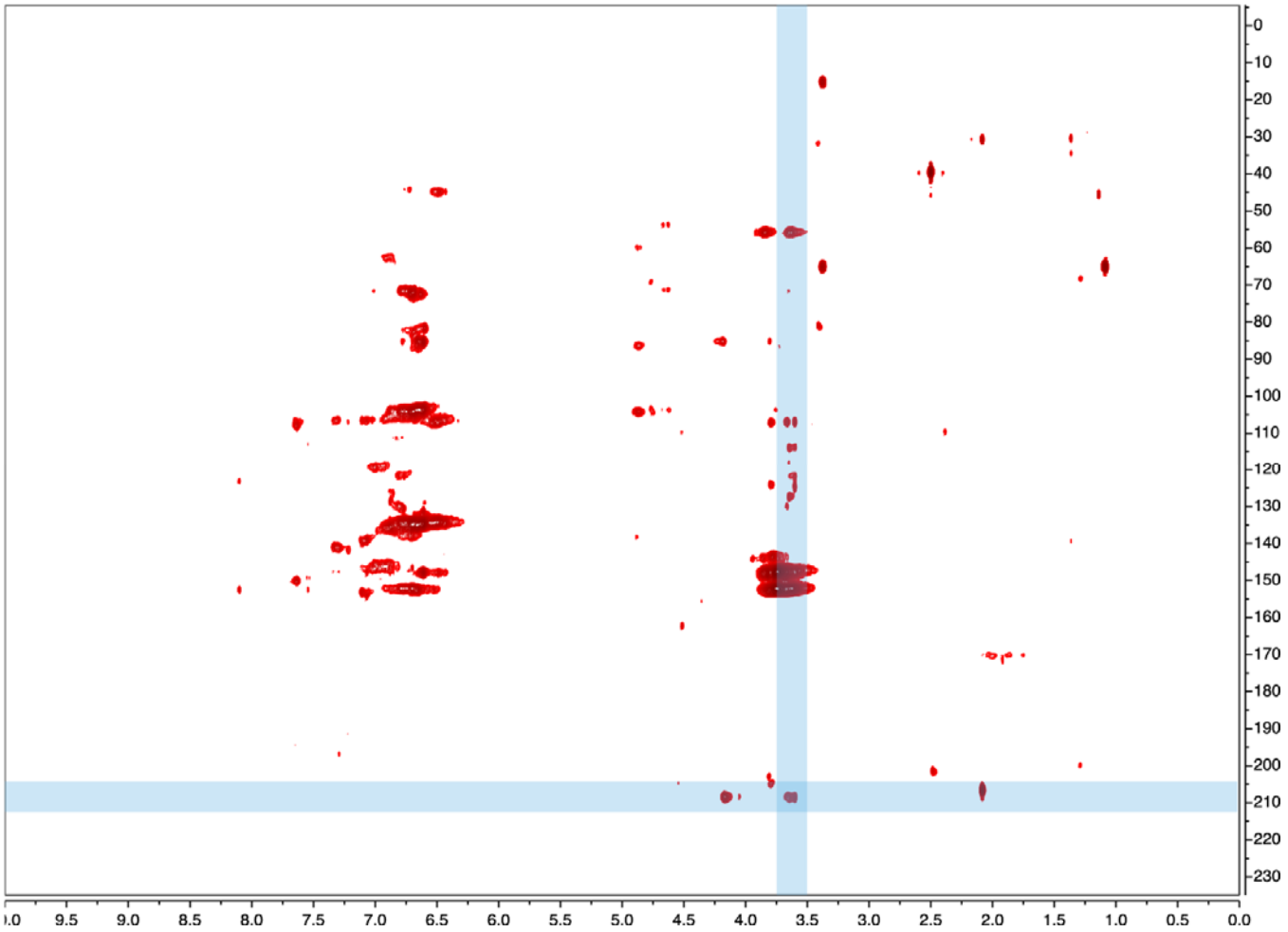


Figure S7: 2D HMBC NMR analysis (700 MHz, d_6-DMSO) of Beech lignin (0.4M dioxasolv extracted). Highlighted is the region examined in Figures S9 and S10.

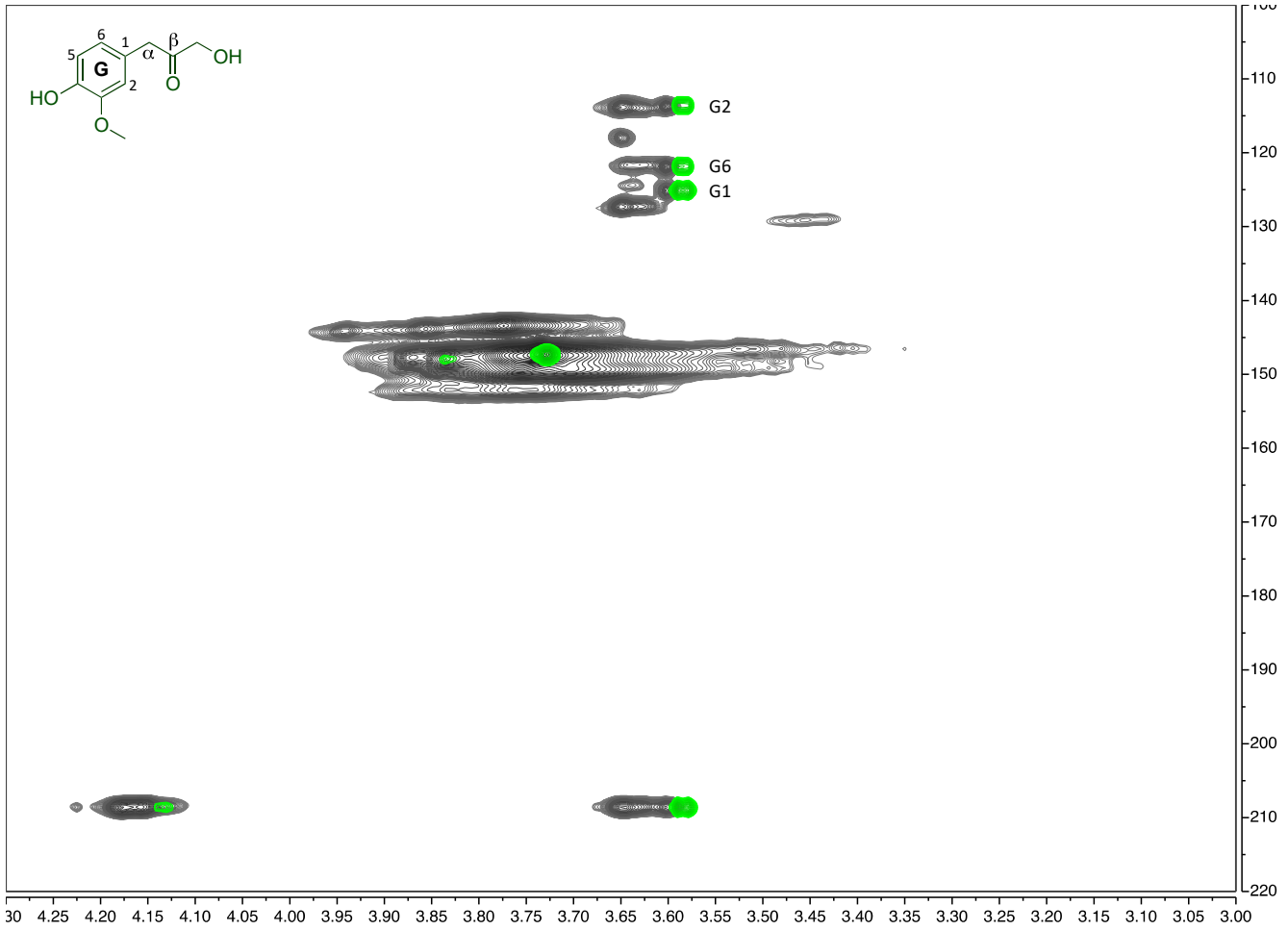


Figure S8: 2D HMBC NMR analysis (700 MHz, d_6-DMSO) of DF lignin (grey) overlaid with G-phenolic Hibbert Ketone model **1** (green). In this DF lignin sample, apparent contamination with **1** was observed leading to further purification being undertaken on these samples. See Figure S11 for before and after purification comparison.

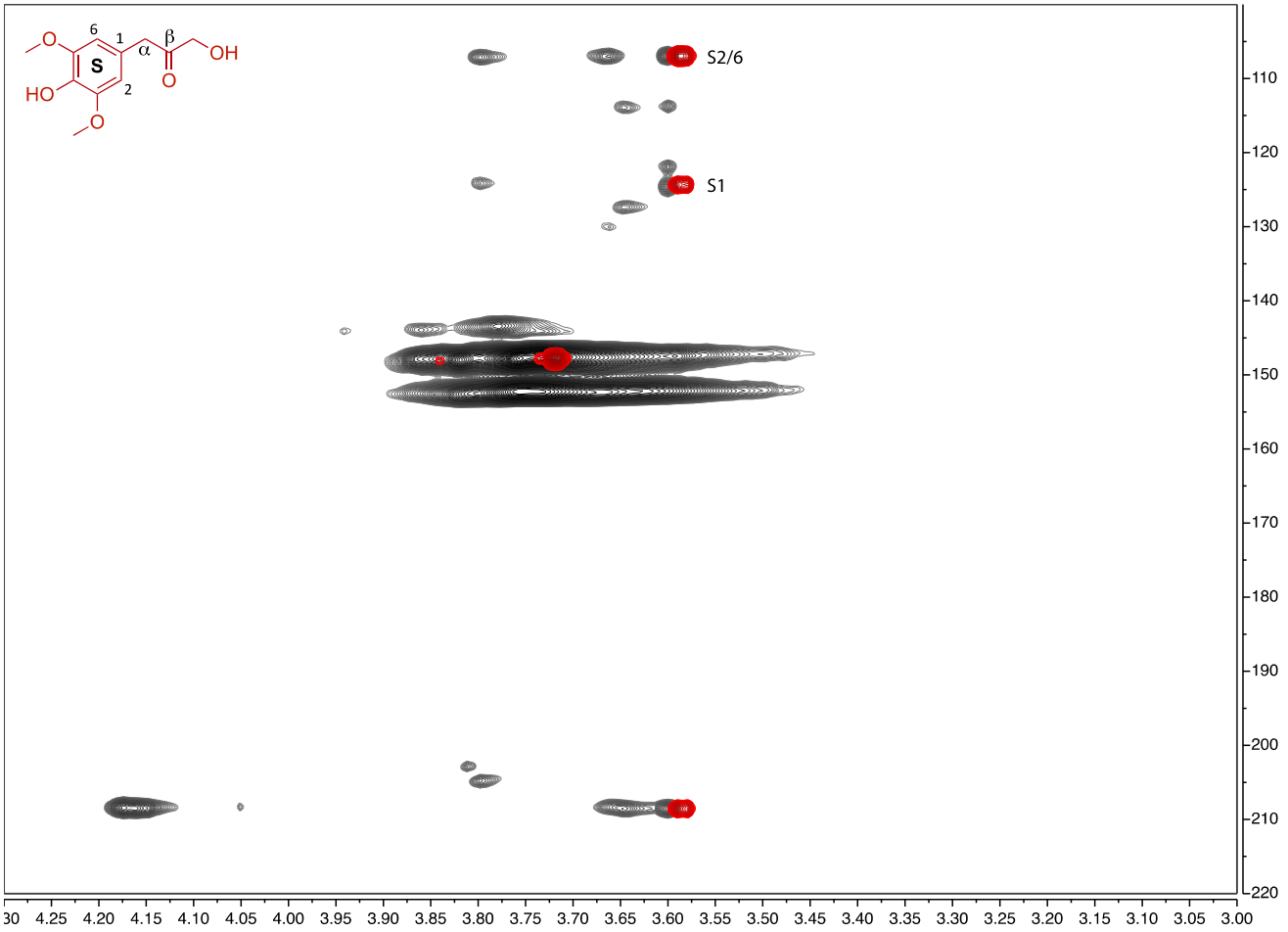


Figure S9: 2D HMBC NMR analysis (700 MHz, d_6-DMSO) of beech lignin (grey) overlaid with S-phenolic Hibbert Ketone model **2** (red). In this beech lignin sample, apparent contamination with **2** (and also **1**, *c.f.* Figure S8) was observed leading to further purification beingundertaken on these samples. See Figure S12 for before and after purification comparison.

HMBC Analysis of dioxasolv lignins (containing low M_w impurities)

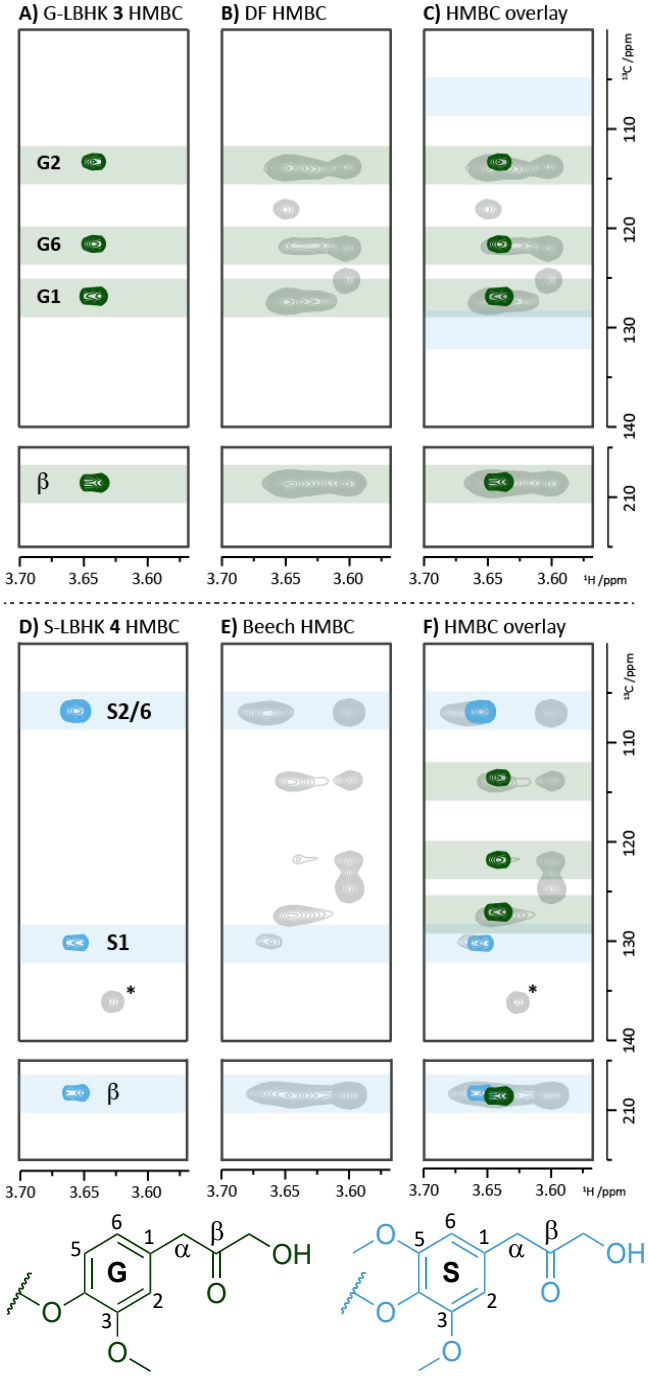


Figure S10: 2D HMBC NMR Analysis (700 MHz, d_6-DMSO) of A) G Hibbert ketone model **3**; B) Douglas fir lignin (0.4M dioxasolv extracted); C) HMBC overlay of A and B; D) S Hibbert ketone model **4**; E) Beech lignin (0.4M dioxasolv extracted); F) HMBC overlay of D and E. *peak at δ_C/δ_H 3.63/136.1 ppm corresponds to C3/5 of S-LBHK model **4** (blue). For carbon numbering, see annotated figures above. For assignment of additional cross-peaks at ^{1}H 3.60 ppm in **B, C, E** and **F,** see ESI Figure S8-S9. Blue bands in 3C indicate the lack of S-LBHK present.

It can be seen from Figures S8-10 that low M_w impurities are not washed out sufficiently during our originally used precipitation stages of dioxasolv extraction. Further Soxhlet extractions/ re-precipitation into ethyl acetate allowed for 2D HMBC data to be acquired containing no contaminants of **1** and **2**, see Manuscript Figure 3 and ESI Figures S11-S12.

Comparison of 2D HMBC Data before and after purification

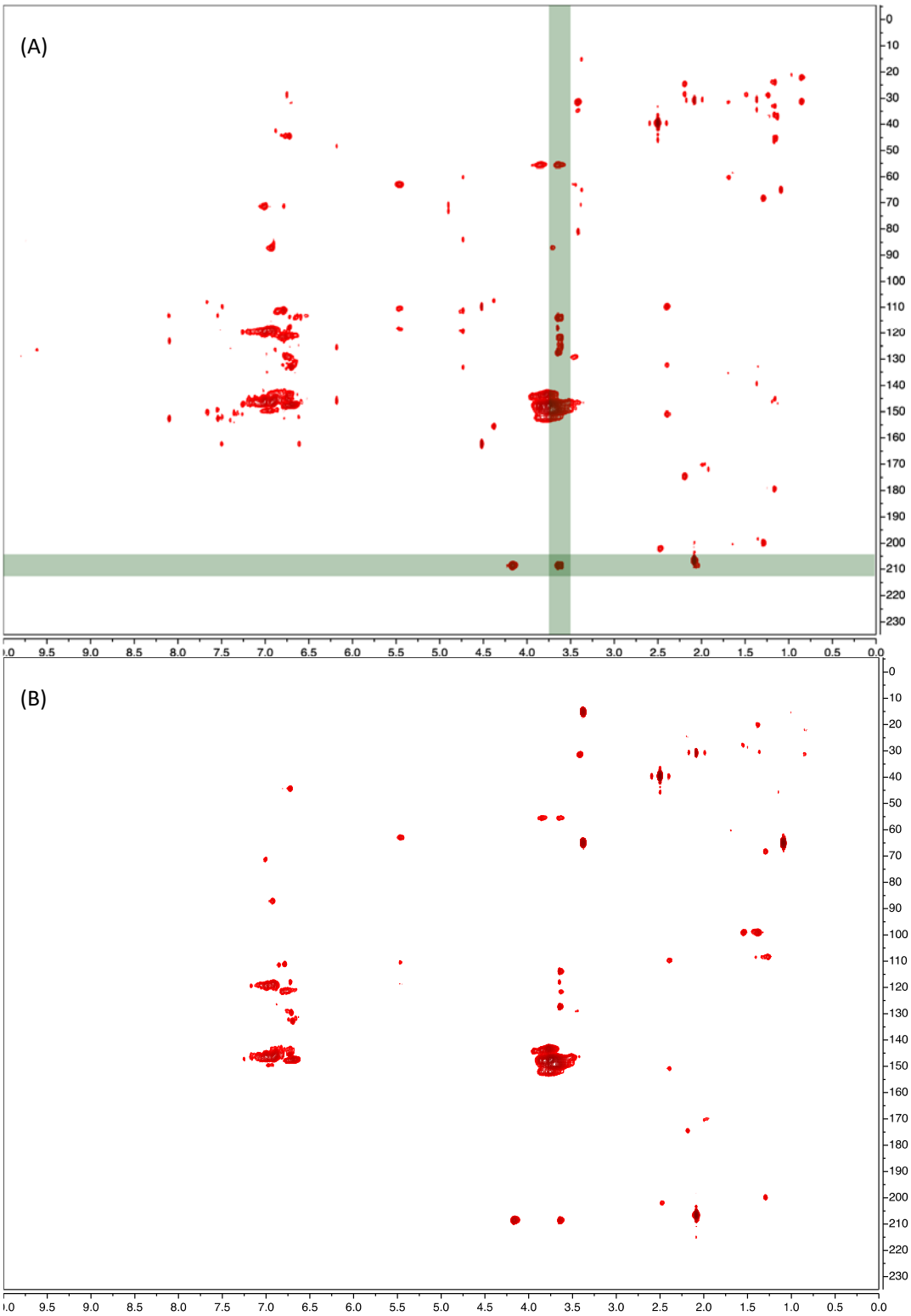


Figure S11: Raw 2D HMBC NMR data (700 MHz, d_6-DMSO) of (**A**) un-purified DF and; (**B**) purified DF lignin.

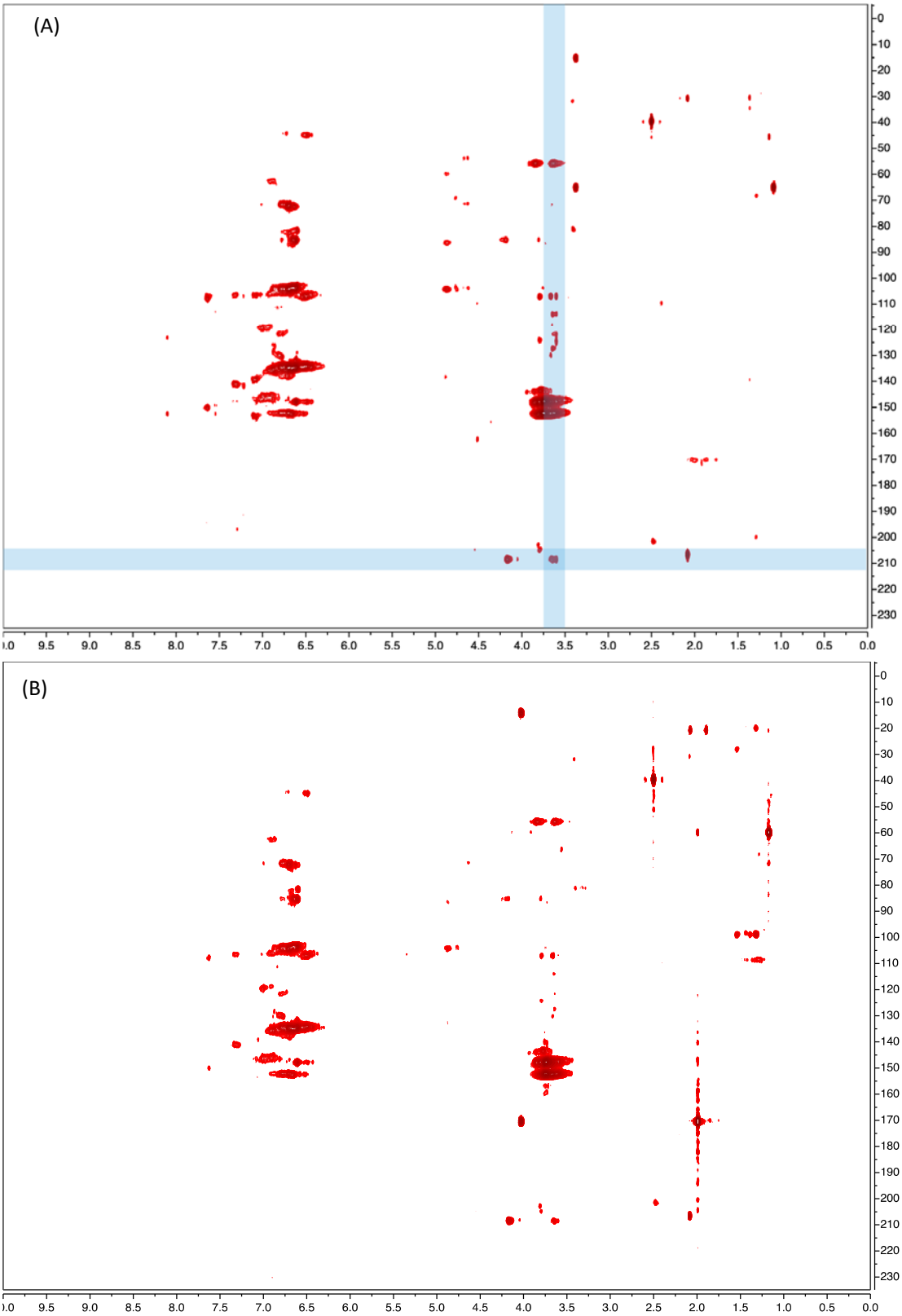


Figure S12: Raw 2D HMBC data (700 MHz, d_6-DMSO) of (A) un-purified beech and; (B) purified beech lignin.

2D HSQC-TOCSY analysis of Hibbert ketones in dioxasolv lignin

A 2D HSQC-TOCSY experiment was used to assess correlations through the 3-carbon spin-system. Because the TOCSY transfer will not pass through aromatics or quaternary carbons (e.g. ketone of the HK), the lignin was first subject to a reduction to generate reduced Douglas fir lignin (DF^{RD}) (Figure S14) *via* the following procedure:

To a stirred solution of DF lignin in THF: H_2O (2:1, 2 mL per 150 mg) was added $NaBH_4$ (20 mg per 150 mg of lignin) and the reaction was stirred overnight at room temperature. The mixture was then concentrated *in vacuo*, suspended in H_2O (~5 ml per 150 mg) and quenched by the addition of NH_4Cl (sat. solution). The mixture was acidified to pH 2 slowly by the addition of 2N HCl causing the lignin to precipitate. The precipitate was collected by filtration, washed with an excess of H_2O and dried in a vacuum desiccator over $CaCl_2$. Recovered weight yields were approx. 80-90%.

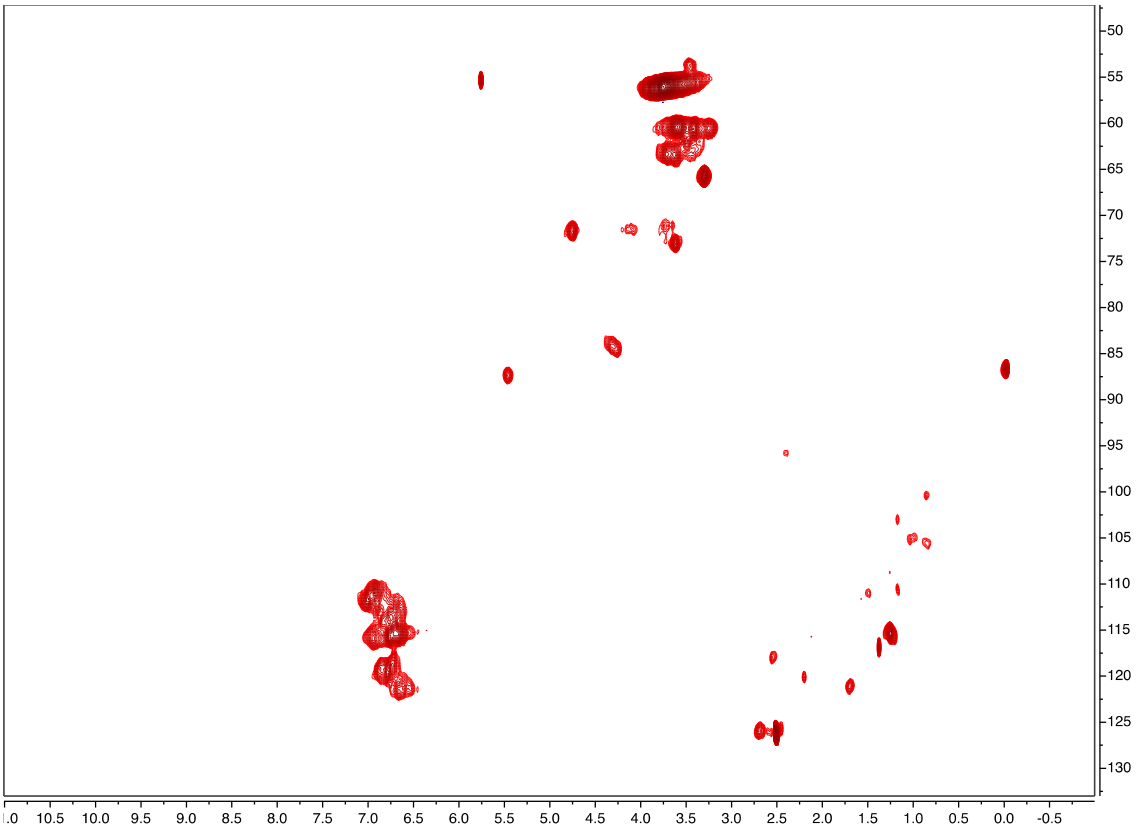


Figure S13a: Full 2D HSQC NMR analysis (700 MHz, d_6-DMSO) of DF^{RD}.

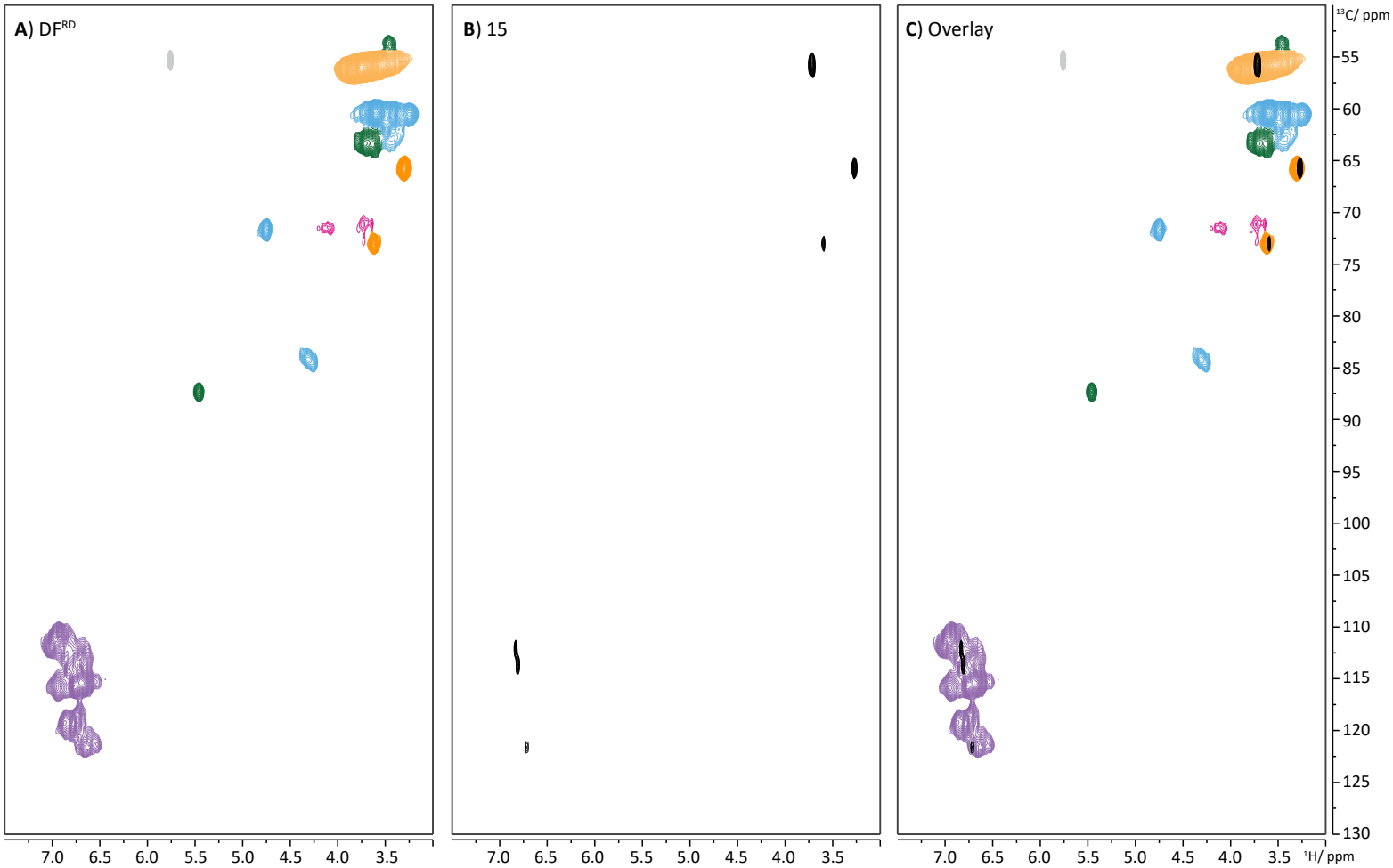
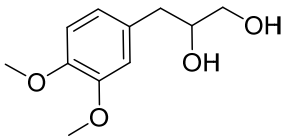


Figure S13b: 2D HSQC NMR analysis (700 MHz, d_6-DMSO) of A) DFRD; B) model compound **15** (see Manuscript Figure 4) and; C) Overlay of A) and B). For assignment of coloured cross-peaks, see Figure S2. Cross-peaks corresponding to the α-protons are located near the DMSO residual solvent peak (ca. 2.5/40 ppm), see Figure S13a.

2D HSQC-TOCSY NMR Data

For more information, see Manuscript Figure 4.

3-(3,4-dimethoxyphenyl)propane-1,2-diol (15)



To a stirred solution of ketone **3** (24 mg; 0.11 mmol, 1 eq.) in MeOH (1.5 mL) at room temperature was added $NaBH_4$ (7 mg, 0.18 mmol, 1.6 eq.) and the reaction was left to stir for 0.5 hours. The reaction was quenched by the addition of water (1.5 mL) before adding DCM (10 mL). The organic phase was washed with water (2 x 10 mL) and brine (2 x 10 mL) before being dried (Na_2SO_4) and concentrating *in vacuo*. Purification on silica gel (20 – 50% EtOAc in petroleum ether) afforded the desired diol **15** (15 mg, 0.07 mmol, 63%).

IR (FTIR)ν_{max}: 3408 (b, O-H str), 2922 (m), 1589 (s), 1516 (s) cm^{-1}; **HRMS** (NSI-) *m/z* [M - H^+] calcd. for $C_{11}H_{15}O_4^-$ 211.0976, found 211.0978; **^{1}H NMR** (700 MHz, DMSO-d_6) δ 6.82 (d, *J* = 8.1 Hz, 1H), 6.81 (d, *J* = 2.0 Hz, 1H), 6.71 (dd, *J* = 8.1, 2.0 Hz, 1H), 4.53 – 4.49 (m, 2H), 3.72 (s, 3H), 3.70 (s, 3H), 3.62 – 3.57 (m, 1H), 3.30 – 3.23 (m, 2H), 2.68 (dd, *J* = 13.7, 5.0 Hz, 1H), 2.46 (dd, *J* = 13.7, 7.6 Hz, 1H); **^{13}C NMR** (176 MHz, DMSO-d_6) δ 148.3, 146.9, 132.1, 121.2, 113.2, 111.7, 72.6, 65.3, 55.5, 55.3, 39.4.

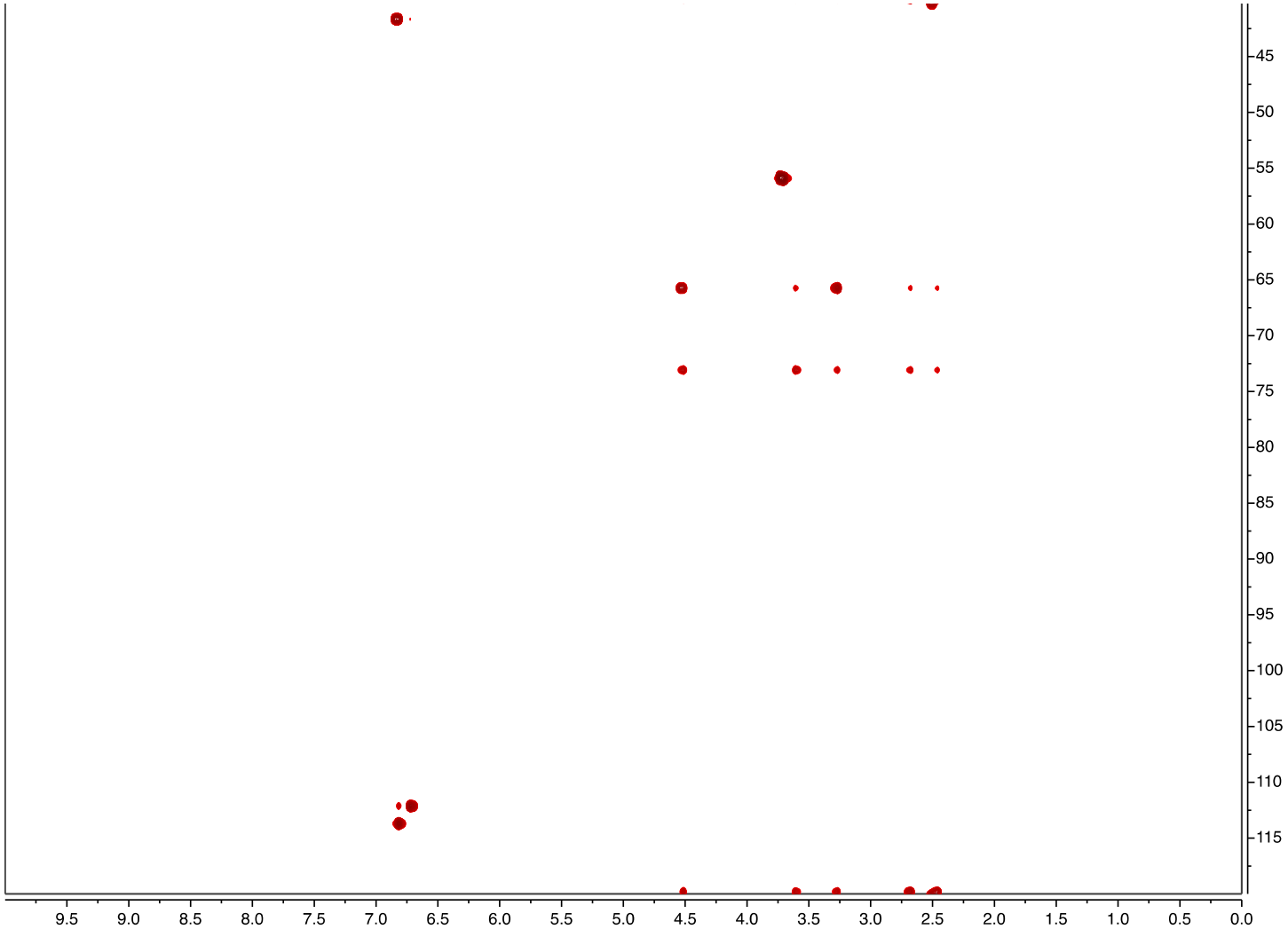


Figure S14: Full 2D HSQC-TOCSY NMR (700 MHz, d_6-DMSO) of **15** (see general procedures for details on NMR acquisition).

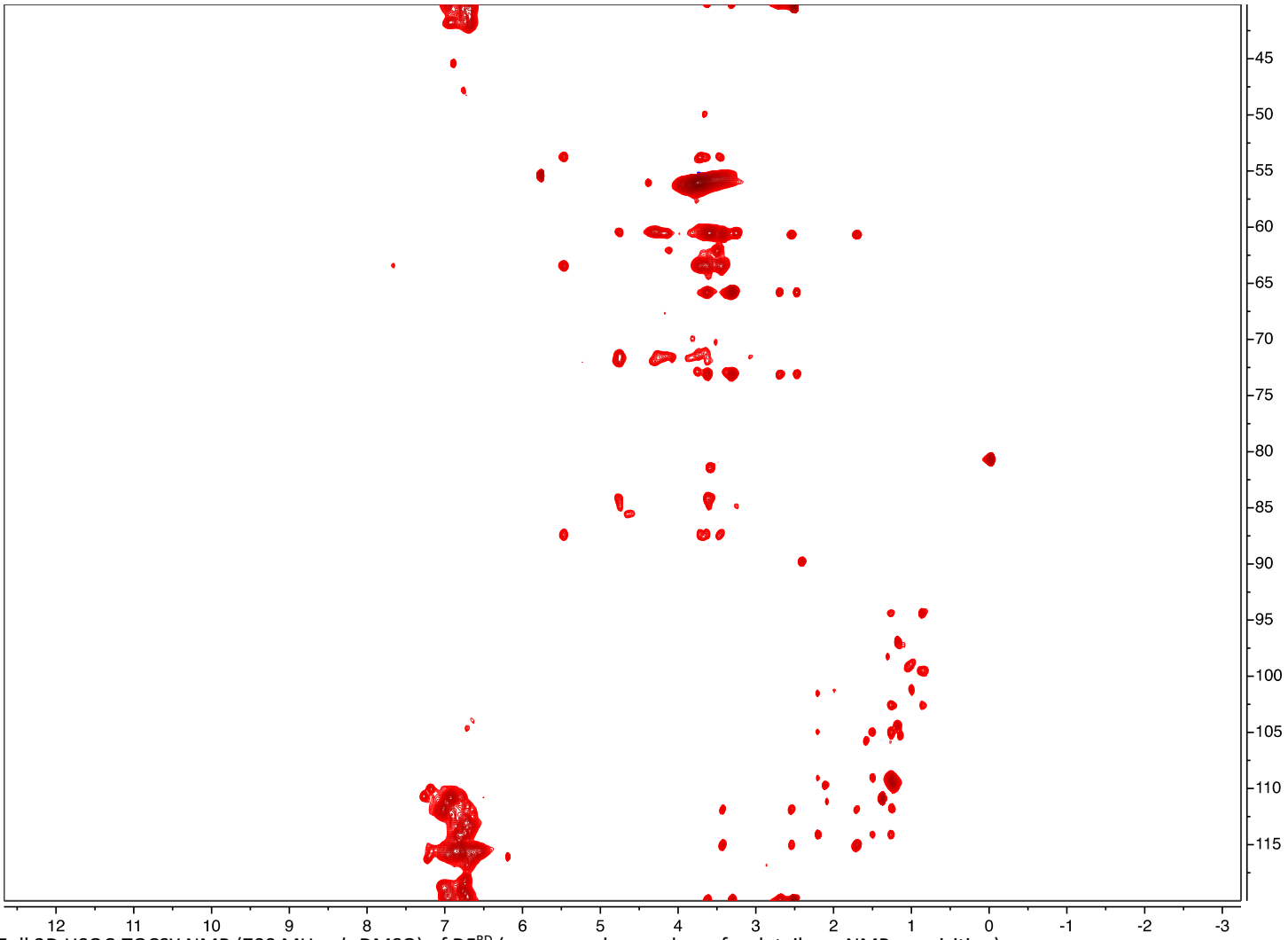
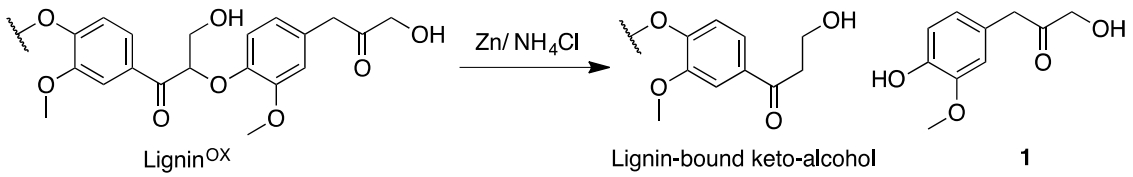


Figure S15: Full 2D HSQC-TOCSY NMR (700 MHz, d_6-DMSO) of DF^{RD} (see general procedures for details on NMR acquisition).

Attempted Isolation of LBHK-derived aromatics by Zinc Reductive Cleavage

Initial attempts to release HKs (e.g. **1** and **2**) or derivatives of HKs from lignin focussed on using a selective benzylic oxidation followed by a reductive cleavage strategy (Scheme S4) as reported by Westwood *et al*.[2] For isolation of HK **1** to be possible, it must not react during the zinc reductive cleavage step. To model this, **3** was subjected to zinc reductive cleavage conditions.



Scheme S4: Proposed reaction of Lignin^{OX} under zinc reductive cleavage conditions. For further detail on reaction conditions, see Reference S2.

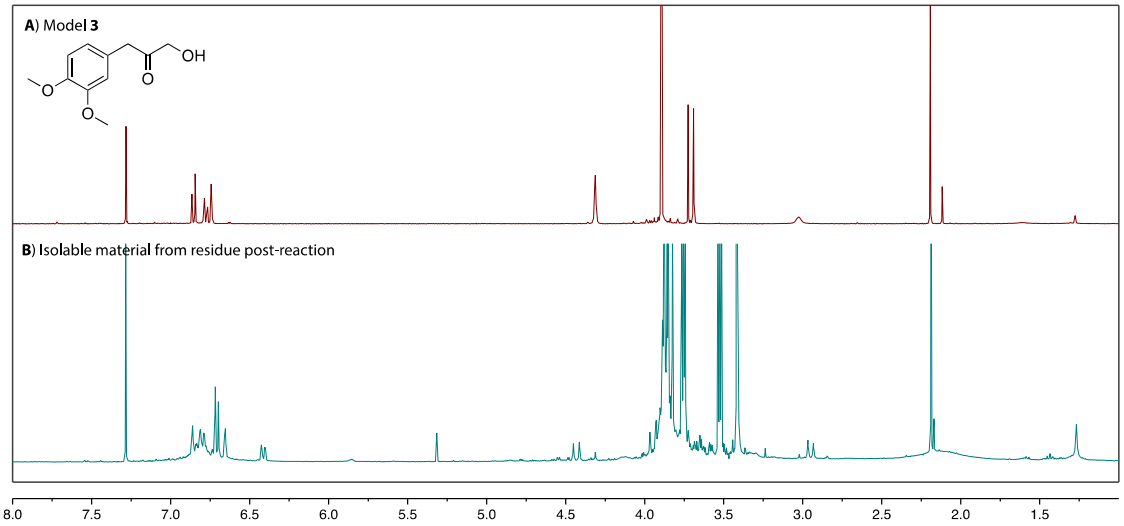
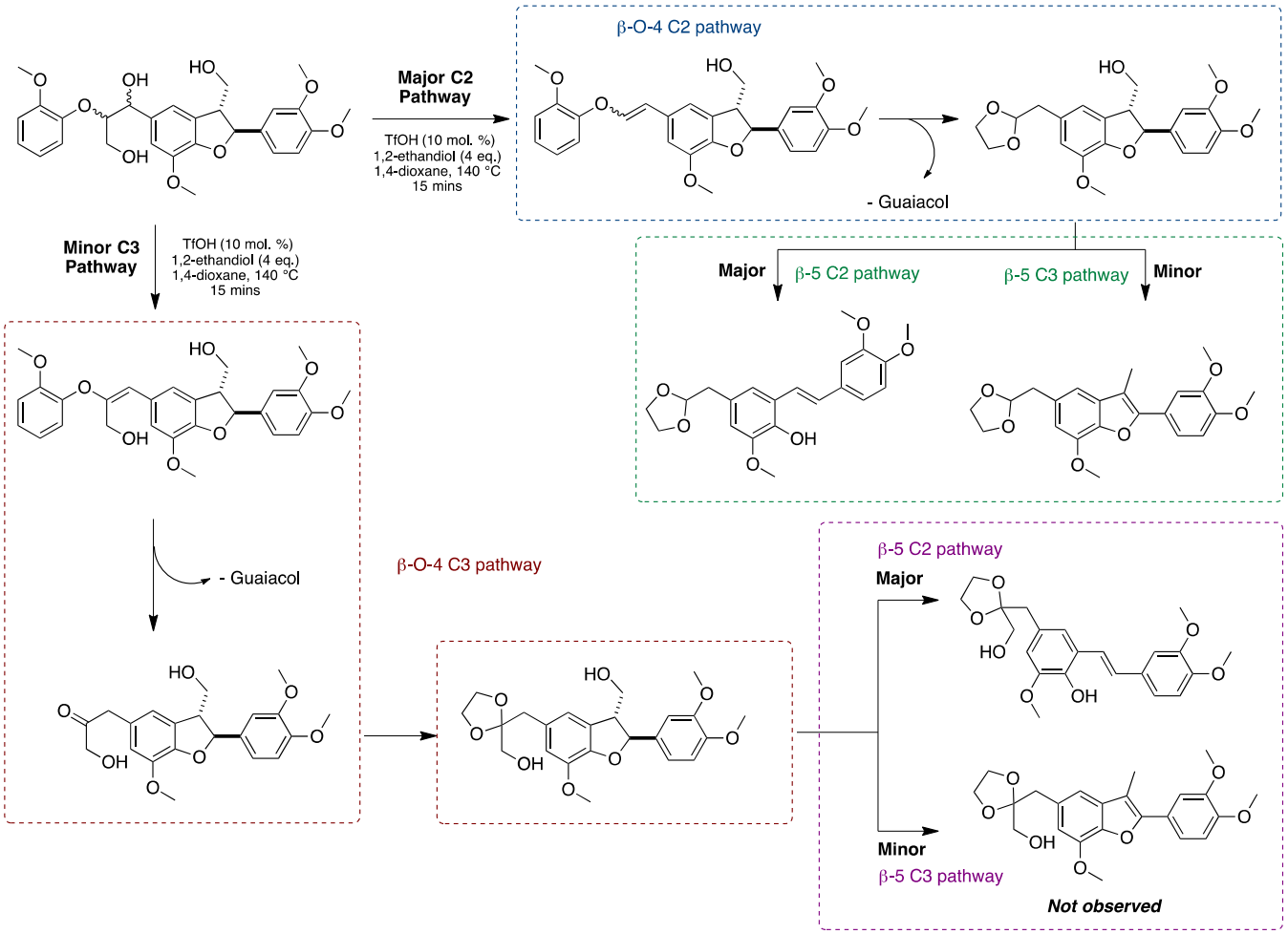


Figure S16: ^{1}H NMR (400 MHz, $CDCl_3$) of **A**) Model **3** and; **B**) isolable material from residue post-zinc reductive cleavage reaction of **3**.

Unfortunately, it was found that **3** underwent extensive degradation when subject to zinc reductive cleavage conditions, forming a black residue and attempts to isolate any products by chromatography failed. Trace amounts of soluble low molecular weight compounds were identified (Figure S16B) post-reaction but the mixture mainly comprised of the reaction solvent (2-methoxyethanol) and no products were obtained in pure form. This approach was abandoned and work instead focussed on the protocols using the metal triflate induced depolymerisation of lignin.[S15]

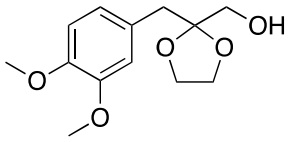


Scheme S5: Proposed C2-major and C3-minor pathways from TfOH/ $M(OTf)_x$ induced depolymerisation of lignin.[S15] In this previous study a peak in the GC-MS analysis of the lignin depolymerisation reaction was assigned as compound **17** (in this manuscript, **P12** in Figure 5 in reference S15). To the best of our knowledge this assignment was carried out based on the obtained mass spectrum. No mention of the formation of compound **18** in this work is made in reference S15.

Metal Triflate Reactions with Model Compounds

Authentic standards of **16**, **17** and **S5** were prepared from **3**, **1** and **2** respectively by reaction with 1,2-ethanediol for comparison with metal triflate crude reaction mixtures. Attempted purification by silica chromatography of **17** and **S5** was found to lead to unknown degradation products.

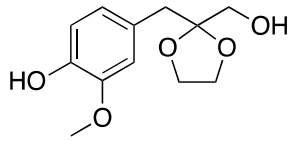
(2-(3,4-dimethoxybenzyl)-1,3-dioxolan-2-yl)methanol (16)



To a stirred solution of ketone **3** (40 mg, 0.19 mmol, 1 eq.) and ethylene glycol (47 mg; 0.76 mmol, 4 eq.) in toluene (2.5 mL) was added camphor-10-sulfonic acid (2 mg; 0.008 mmol, 0.05 eq.). The solution was heated to reflux for 3 hours before cooling and concentrating *in vacuo*. The crude mixture was dissolved in DCM (20 mL) and washed with water (2 x 10 mL) and brine (2 x 10 mL). The organic phase was dried (Na_2SO_4) and concentrated *in* vacuo to afford ketal **16** in quantitative yield as an orange oil. No further purification of **16** was performed.

IR (FTIR)ν_{max}: 3321 (b, O-H str), 2954 (b, C-H str), 1589 (m), 1514 (s), 1463 cm^{-1}. **HRMS** (ESI+) *m/z* [M + Na^+] calcd for $C_{13}H_{18}O_5Na^+$ 277.1052, found 277.1043; **^{1}H NMR** (400 MHz, $CDCl_3$) δ = 6.86 – 6.76 (m, 3H), 3.94 – 3.90 (m, 2H), 3.88 (s, 3H), 3.86 (s, 3H), 3.74 – 3.69 (m, 2H), 3.51 (d, *J*=6.5, 2H), 2.91 (s, 2H), 1.81 (t, *J*=6.7, 1H).**^{13}C NMR** (101 MHz, $CDCl_3$) δ 148.59, 147.91, 128.51, 122.85, 113.91, 110.93, 110.06, 65.86 (2C), 65.68, 56.02, 55.98, 40.75.

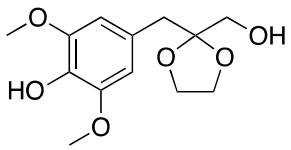
4-((2-(hydroxymethyl)-1,3-dioxolan-2-yl)methyl)-2-methoxyphenol (17)



To a stirred solution of ketone **1** (106 mg, 0.54 mmol, 1 eq.) and ethylene glycol (135 mg; 2.18 mmol, 4 eq.) in toluene (2.5 mL) was added camphor-10-sulfonic acid (6 mg; 0.026 mmol, 0.05 eq.). The solution was heated to reflux for 3 hours before cooling and concentrating *in vacuo*. The crude mixture was dissolved in DCM (25 mL) and washed with water (2 x 20 mL) and brine (2 x 20 mL). The organic phase was dried (Na_2SO_4) and concentrated *in* vacuo to afford ketal **17** in quantitative yield as a brown oil. No further purification of **17** was performed.

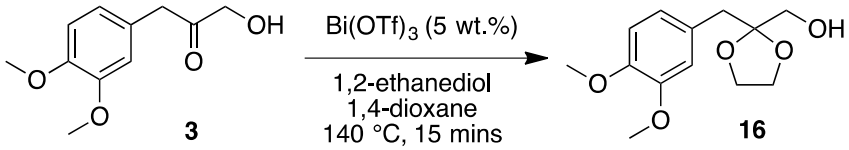
IR (FTIR)ν_{max}: 3342 (b, O-H str), 2934 (b, C-H str), 1601 (m), 1515 (s) cm^{-1} **HRMS** (NSI-) *m/z* [M - H^+] calcd. for $C_{12}H_{15}O_5^-$ 239.0925, found 239.0925; **^{1}H NMR** (500 MHz, $CDCl_3$) δ 6.84 – 6.81 (m, 2H), 6.78 – 6.76 (m, 1H), 5.60 (s, 1H), 3.93 – 3.90 (m, 2H), 3.87 (s, 3H), 3.73 – 3.70 (m, 2H), 3.51 (s, 2H), 2.89 (s, 2H); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 146.2, 144.4, 127.8, 123.5, 114.1, 113.2, 110.1, 65.8 (2C), 65.6, 56.1, 40.8.

4-((2-(hydroxymethyl)-1,3-dioxolan-2-yl)methyl)-2,6-dimethoxyphenol (S5)



To a stirred solution of ketone **2** (50 mg, 0.22 mmol, 1 eq.) and ethylene glycol (60 mg; 8.96 mmol, 4 eq.) in toluene (1.1 mL) was added camphor-10-sulfonic acid (2.5 mg; 0.011 mmol, 0.05 eq.). The solution was heated to reflux for 3 hours before cooling and concentrating *in vacuo*. The crude mixture was dissolved in DCM (15 mL) and washed with water (2 x 10 mL) and brine (2 x 10 mL). The organic phase was dried (Na_2SO_4) and concentrated *in* vacuo to afford acetal **S5** in quantitative yield as a brown oil. No further purification of **S5** was performed.

IR (FTIR)ν_{max}: 3433 (b, O-H str), 2940 (m), 1609 (s), 1516 (m) cm^{-1}; **HRMS** (NSI-) *m/z* [M - H^+] calcd. for $C_{13}H_{17}O_6^-$ 269.1031, found 269.1030; **^{1}H NMR** (500 MHz, $CDCl_3$) δ 6.55 (s, 2H), 5.51 (s, 1H), 3.95 – 3.91 (m, 2H), 3.89 (s, 6H), 3.75 – 3.71 (m, 2H), 3.54 (s, 2H), 2.91 (s, 2H); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 146.7, 133.5, 126.9, 110.0, 107.4, 65.8 (2C), 65.7, 56.4, 41.2.



Scheme S6: Reaction of model **3** under metal triflate catalysed reaction conditions. Reaction conditions: $Bi(OTf)_3$ (5 wt. %), 1,4-dioxane (33 mg/mL), 1,2-ethanediol (1 wt. eqv), 140 °C, 15 mins, sealed tube.

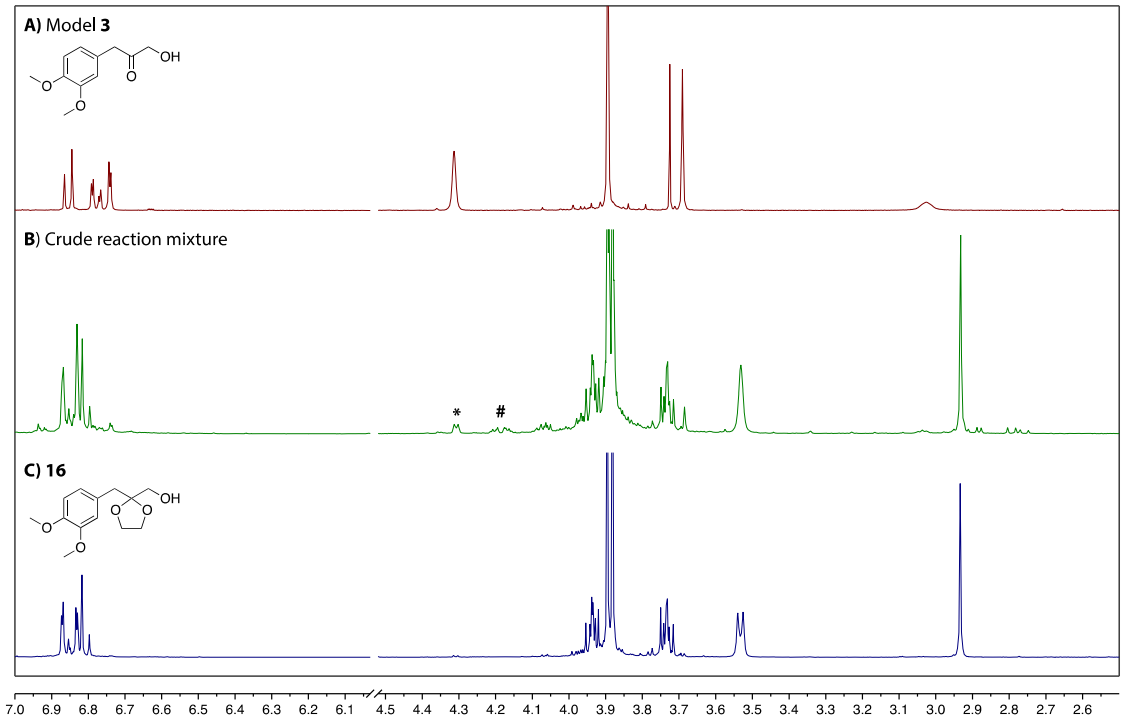
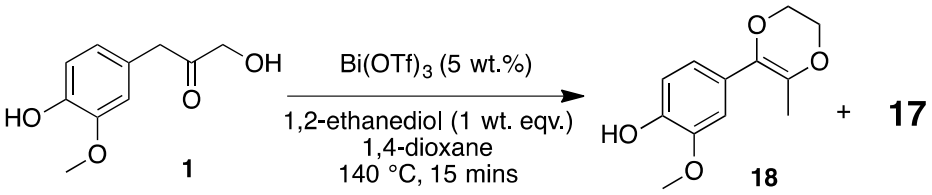


Figure S17: ^{1}H NMR spectra of **A**) Model **3**; **B**) crude reaction mixture from the reaction shown in Scheme S6 and; **C**) **16.** * Trace amounts of **3** were observed under these conditions. # Trace amounts of a by-product were observed but there was insufficient material to enable isolation of this product. Based on the highlighted NMR signal it is tentatively proposed that this by-product corresponds to a non-phenolic version of **18**.



Scheme S7: Reaction of Hibbert ketone **1** under metal triflate catalysed reaction conditions. Reaction conditions: $Bi(OTf)_3$ (5 wt. %), 1,4-dioxane (33 mg/mL), 1,2-ethanediol (1 wt. eqv.), 140 °C, 15 mins, sealed tube.

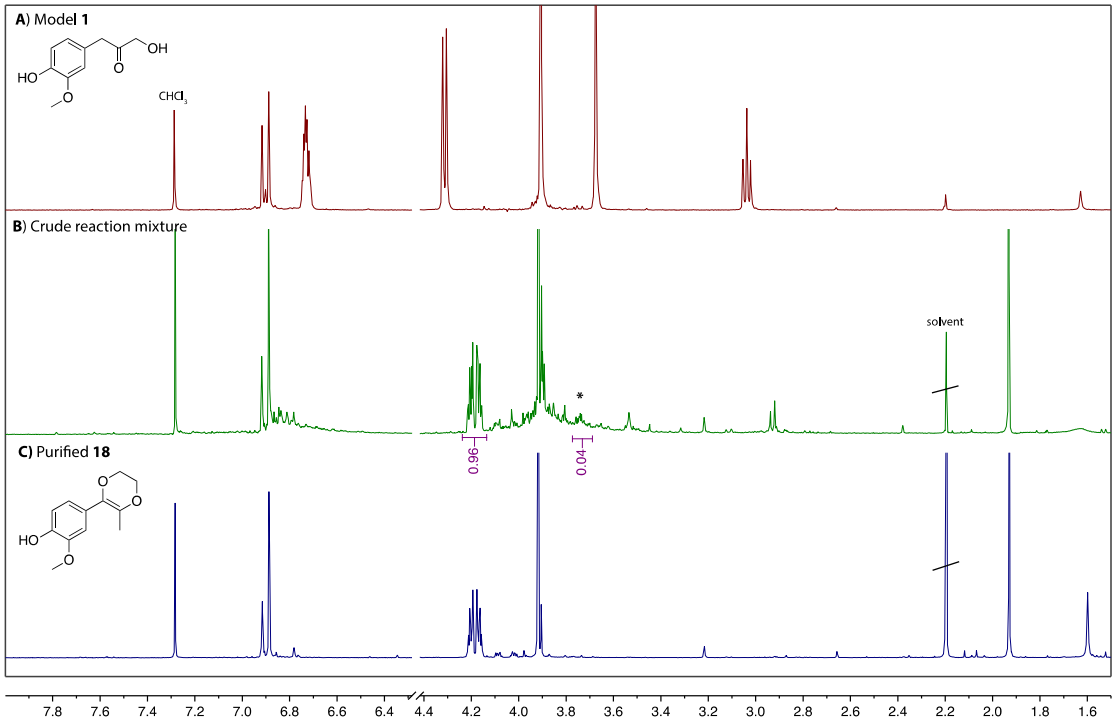
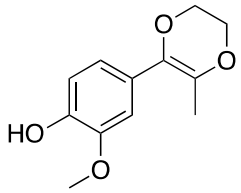


Figure S18: ^1H NMR spectra of **A**) Hibbert ketone **1**; **B**) crude reaction mixture from the reaction shown in Scheme S7 and; C) purified dioxene **18. *** trace amounts of **17** were observed under these reaction conditions.

2-methoxy-4-(3-methyl-5,6-dihydro-1,4-dioxin-2-yl)phenol (18)



IR (FTIR)ν_{max}: 3426 (b, O-H str), 2931 (b, C-H str), 1674 (m), 1512 (s) cm^{-1}; **HRMS** (NSI+) *m/z* [M + H^+] calcd. for $C_{12}H_{15}O_4^+$ 223.0965, found 223.0965; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 6.92 – 6.86 (m, 1H), 6.88 – 6.84 (m, 2H), 5.62 (s, 1H), 4.21 – 4.11 (m, 4H), 3.90 (s, 3H), 1.91 (s, 3H); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 146.2, 145.0, 132.0, 131.0, 127.8, 122.0, 113.9, 110.9, 64.8, 64.8, 56.1, 17.0.

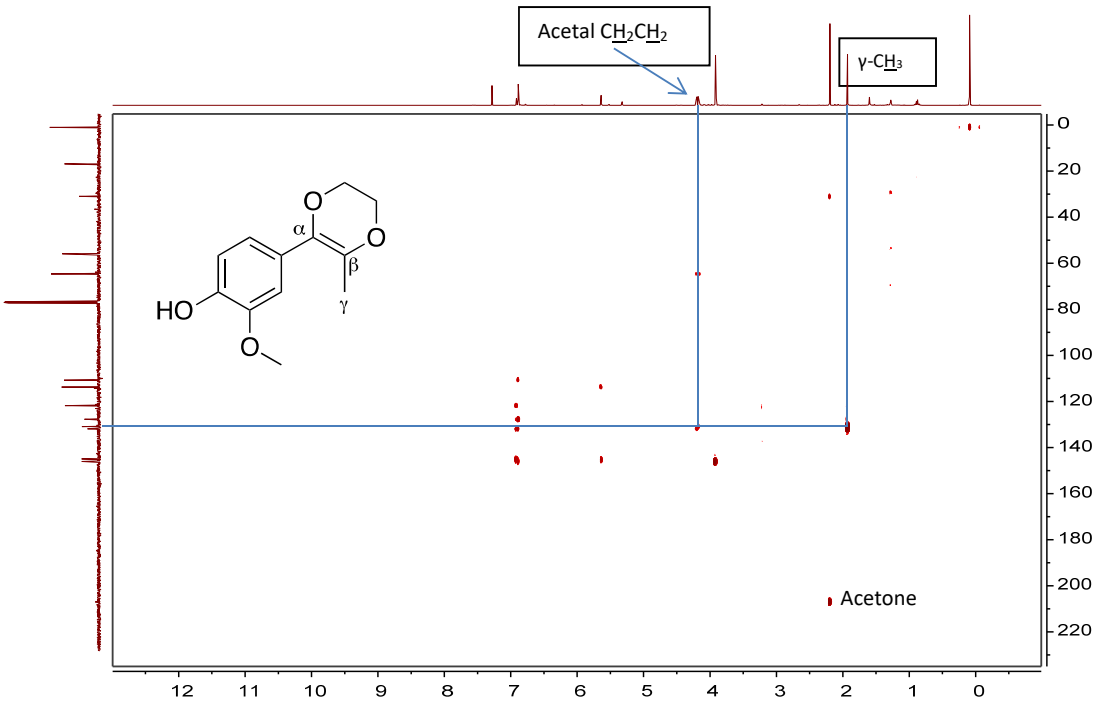
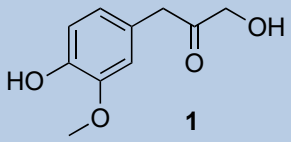
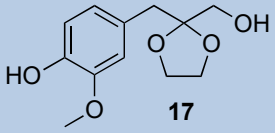
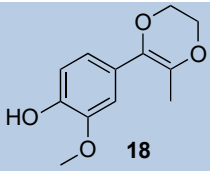


Figure S19: The HMBC obtained for **18**. A correlation was observed between the γ-CH_3 protons and the α- and ß- carbons (131.0/ 132.0 ppm). Strong HMBC correlations were also observed between the acetal protons (2 x CH_2; m at 4.21 – 4.11 ppm) and the α- and ß- carbons.

Metal Triflate Screen

A screen (using otherwise identical reaction conditions) of different metal triflates was undertaken to observe the effect of the Lewis-acid structure on the product distribution of **1**:**17**:**18** (in view of favouring **17**) (See also other work by de Vries *et al*.[S15,S16]). The initial $Bi(OTf)_3$ reaction (Table S7, entry 1) showed predominant formation of dioxene **18**. Similarly, $Hf(OTf)_4$ (Table S7, entry 4) also favoured **18** as expected based on lewis acidity. Acids $Y(OTf)_3$, $Yb(OTf)_3$ and $Er(OTf)_3$ yielded none of dioxene **18** (Table S7, entries 2, 7 and 8), however, unsatisfactory conversion to **17** was observed. It was found that $Zn(OTf)_2$ and $Sc(OTf)_3$ (Table S7, entries 5 and 6 respectively) offered the best conversion to ketal **17**. $Sc(OTf)_3$ was carried through for test reactions on lignin.

Table S7: Metal triflate screen: reactivity of HK **1**. General reaction conditions: $M(OTf)_x$ (5 wt. %), ethylene glycol (1 wt. eqv.), 1,4-dioxane, 140 °C, 15 minutes. All ratios were determined from quantitative ^1H NMR analysis integrating peaks corresponding to **1**, **17** and **18** (see Figure S20). n.d = not determined (or negligible amounts).

Entry	$M(OTf)_x$	Product Distribution		
		 **1**	 **17**	 **18**
1	$Bi(OTf)_3$	n.d	0.15	0.85
2	$Y(OTf)_3$	0.55	0.45	n.d
3	$Cu(OTf)_2$	n.d	0.62	0.38
4	$Hf(OTf)_4$	n.d	0.30	0.70
5	$Zn(OTf)_2$	0.20	0.8	n.d
6	$Sc(OTf)_3$	0.04	0.77	0.19
7	$Yb(OTf)_3$	0.29	0.71	n.d
8	$Er(OTf)_3$	0.47	0.53	n.d

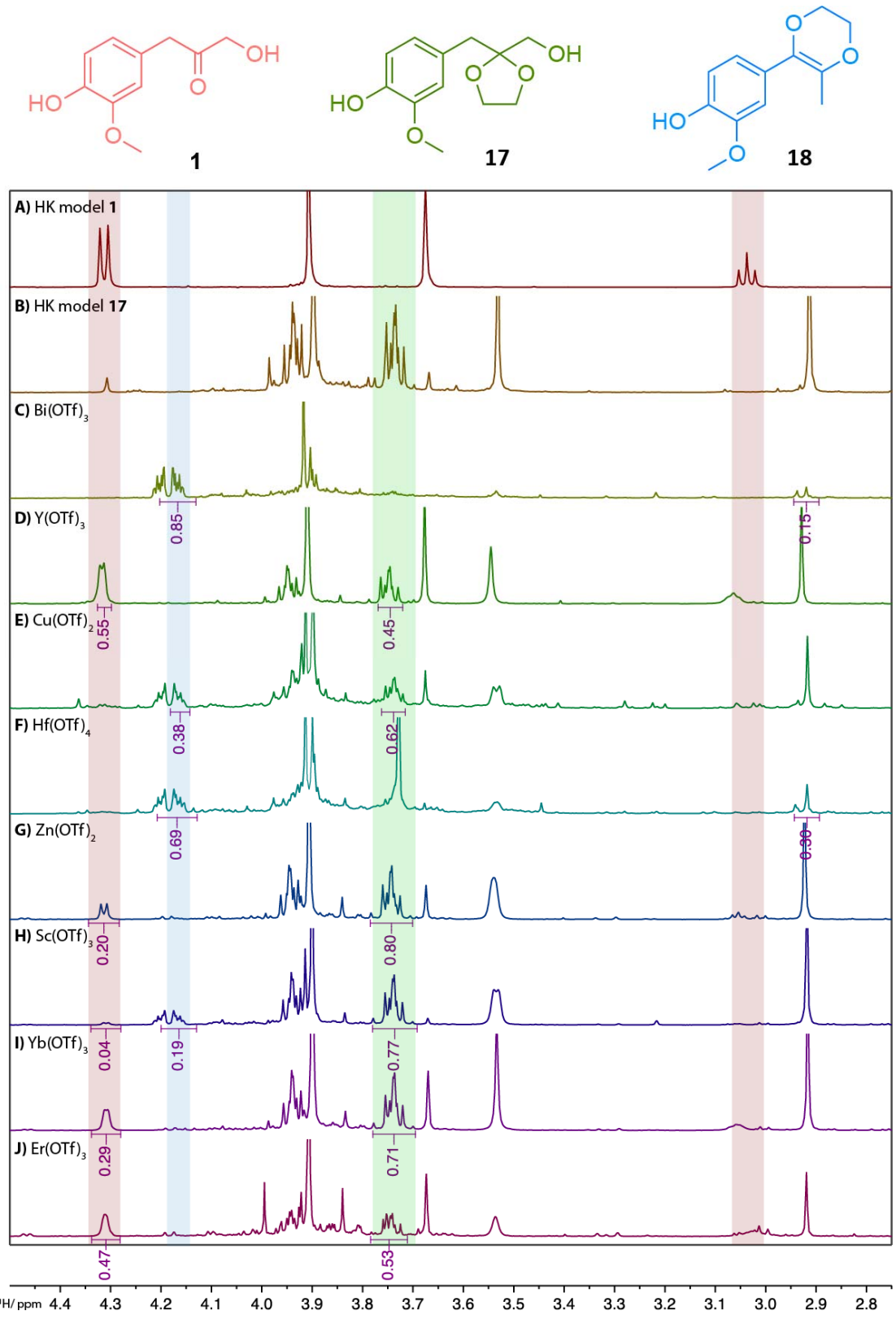
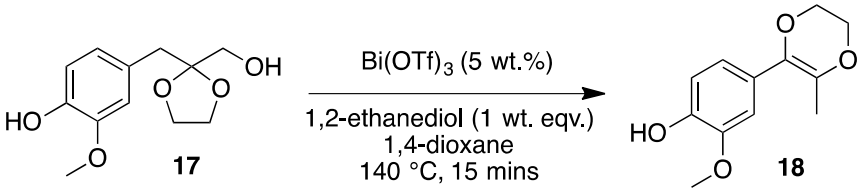


Figure S20: ^{1}H qNMR (in $CDCl_3$) spectra of the crude reaction mixtures obtained on reaction of Hibbert ketone **1** in the metal triflate catalysed reaction (see Table S7 for more details).



Scheme S8: Reactivity of model **17** under metal triflate catalysed acetalization conditions. Reaction conditions: $Bi(OTf)_3$ (5 wt. %), 1,4-dioxane (33 mg/mL), 1,2-ethanediol (1 wt. eqv), 140 °C, 15 mins, sealed tube.

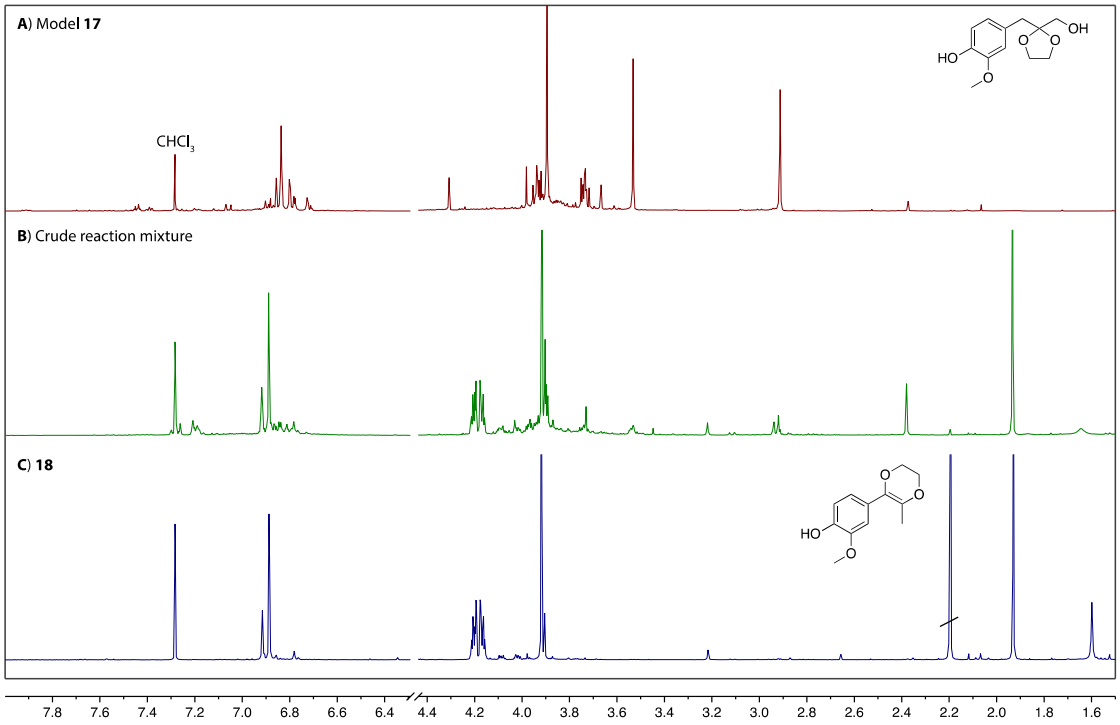


Figure S21: ^1H NMR spectra of A) Model **17**; B) crude reaction mixture from Scheme S8 reaction conditions; C) purified dioxene **18**.

GC-FID and GC-MS Data

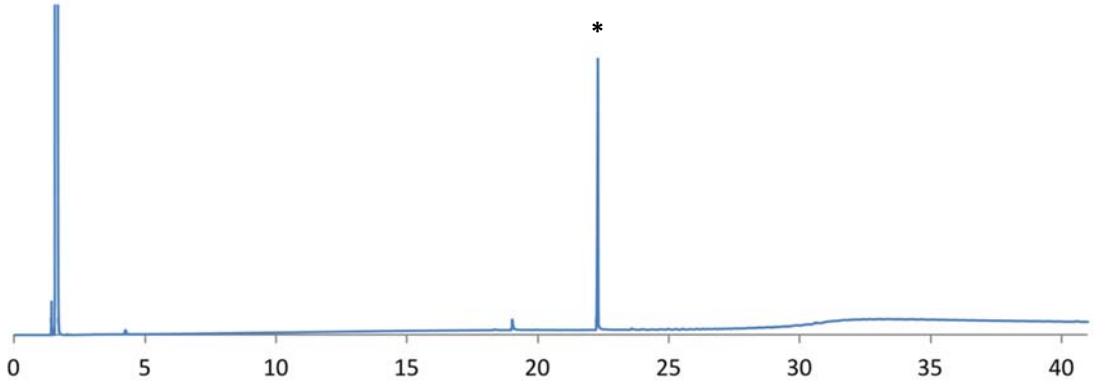


Figure S22: GC-FID trace of G-acetal **19**. ***** Retention time for G-acetal **19**: 22.29 mins.

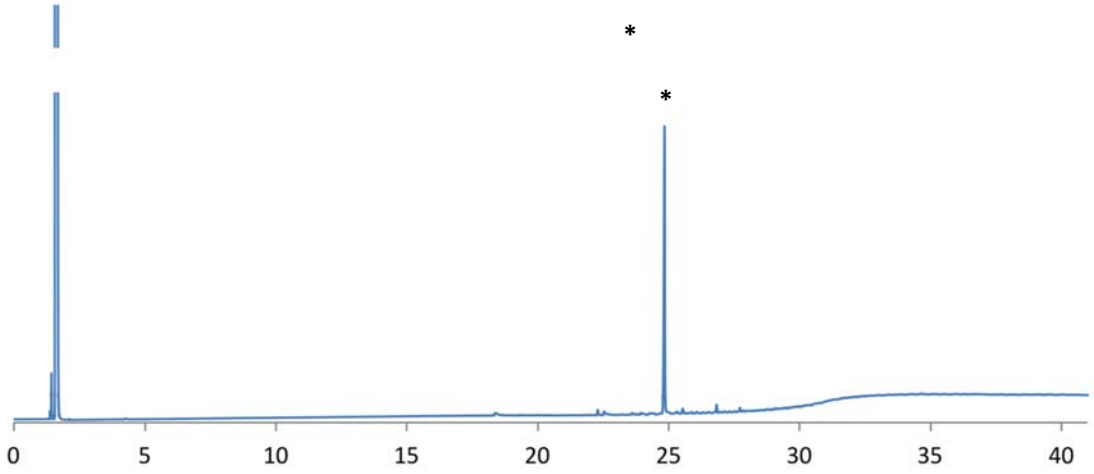


Figure S23: GC-FID trace of semi-purified **18**. ***** Retention time for **18**: 23.60 mins.

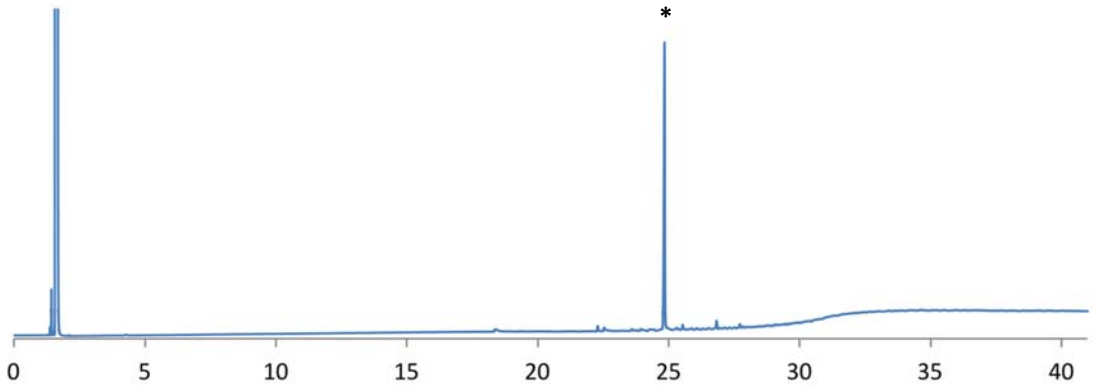


Figure S24: GC-FID trace of **17**. ***** Retention time for **17**: 24.83 mins.

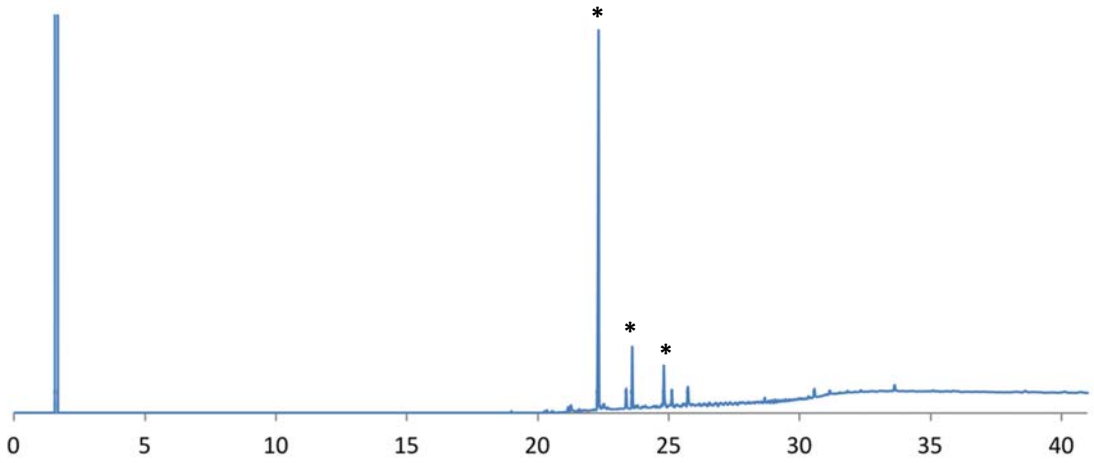


Figure S25: GC-FID trace of crude reaction mixture post DF depolymerisation with $Sc(OTf)_3$. Retention times: 22.32 (**19**), 23.37, 23.60 (**18**), 24.80 (**17**), 25.11, 25.7, 28.66, 30.55, 33.6.

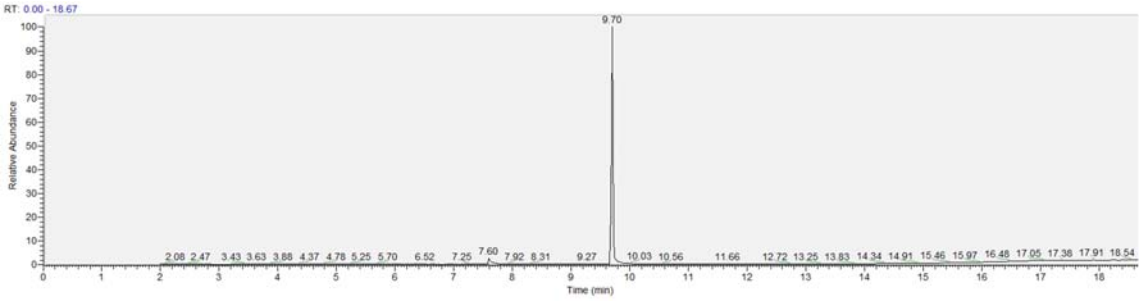
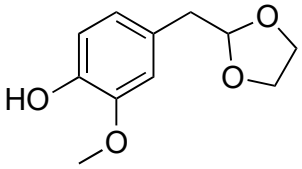


Figure S26: GC-MS Trace of G-acetal **19**

Table S8: Selected GC-MS signals from Figure S26. Expected m/z for compound **19** = 210.09

Entry	r.t. (mins)	*m/z*	Structure
1	7.60	166.10	-
2	9.70	**210**.15	

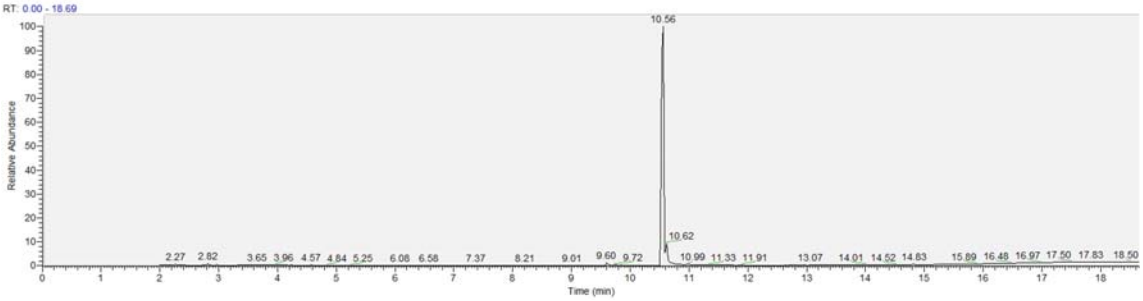
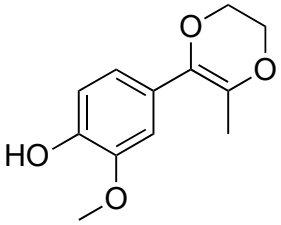


Figure S27: GC-MS trace of **18**

Table S9: Selected GC-MS signals from Figure S27. Expected m/z for compound **18** = 222.09

Entry	r.t. (mins)	*m/z*	Structure
1	9.60	137.09	-
2	10.56	**222**.08	
3	10.62	**222**.08	

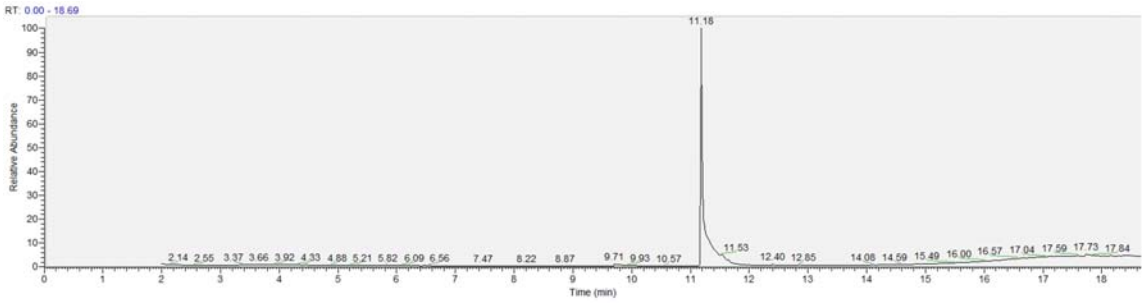
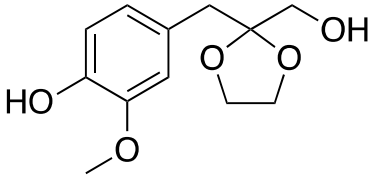


Figure S28: GC-MS trace of **17**

Table S10: Selected GC-MS signals from Figure S28. Expected m/z for compound **17** = 240.10.

Entry	r.t. (mins)	*m/z*	Structure
1	11.18 (broad)	**240**.08	

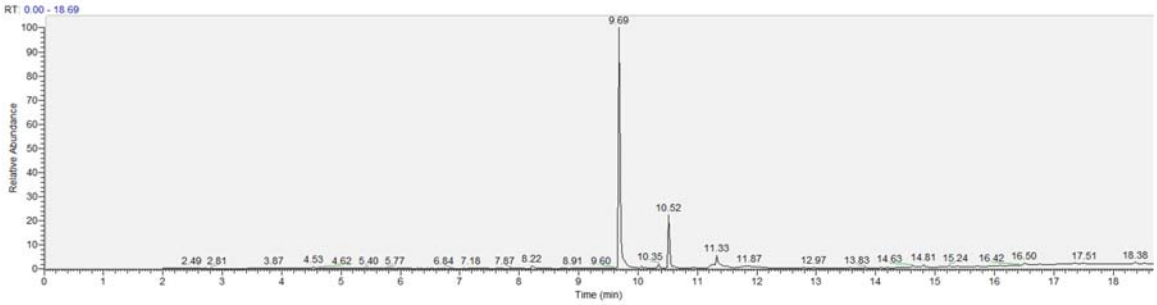
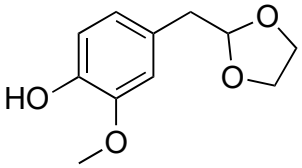
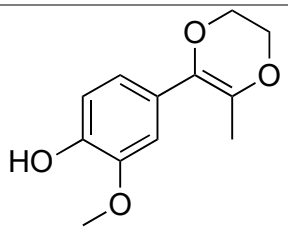
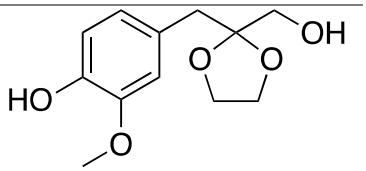


Figure S29: GC-MS trace of toluene-DCM extracts from DF depolymerisation using $Sc(OTf)_3$.

Table S11: Selected GC-MS signals from Figure S29.

Entry	r.t. (mins)	*m/z*	Structure
1	8.22	281.20	-
2	9.69	**210**.18	
3	10.52	**222**.18	
4	11.33 (broad)	**240**.17	

Bibliography

S1 F. Tran, C. S. Lancefield, P. C. J. Kamer, T. Lebl and N. J. Westwood, *Green Chem*., 2015, **17**, 244–249.

S2 C. S. Lancefield, O. S. Ojo, F. Tran and N. J. Westwood, *Angew. Chem. Int. Ed.*, 2015, **54**, 258–262.

S3 L. Mitchell, T. H. Evans and H. Hibbert, *J. Am. Chem. Soc.*, 1944, **66**, 604–607.

S4 L. Mitchell and H. Hibbert, *J. Am. Chem. Soc.,* 1944, **66**, 602–604.

S5 H. E. Fisher, M. Kulka and H. Hibbert, *J. Am. Chem. Soc.*, 1944, **66**, 598–601.

S6 E. West, A. S. MacInnes and H. Hibbert, *J. Am. Chem. Soc.*, 1943, **65**, 1187–1192.

S7 M. Kulka and H. Hibbert, *J. Am. Chem. Soc.*, 1943, **65**, 1180–1185.

S8 J. Zakzeski, P. C. A. Bruijnincx, A. L. Jongerius and B. M. Weckhuysen, *Chem. Rev.*, 2010, **110**, 3552–3599.

S9 J. Štambaský, A. V. Malkov and P. Kočovský, *J. Org. Chem.*, 2008, **73**, 9148–9150

S10 P. R. Reddy, B. Das, *RSC Adv.*, 2014, **4**, 7432-7434.

S11 M. Brasholz, X. Luan and H.-U. Reissig, *Synthesis* (*Stuttg*)., 2005, **20**, 3571–3580.

S12 S. Strych and D. Trauner, *Angew. Chem. Int. Ed.*, 2013, **52**, 9509–9512.

S13 CCDC 1500178 (**1**) and 1500179 (**2**) contain the supplementary crystallographic data for this paper. These data can be obtained free of charge from The Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/data_request/cif

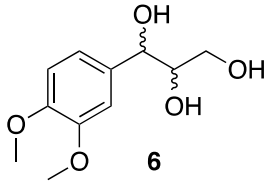
S14 C. Xie, D. Bai, S.-H. Huang, X. Jia and R. Hong, *Asian J. Org. Chem.*, 2014, **3**, 277–280.

S15 C. W. Lahive, P. J. Deuss, C. S. Lancefield, Z. Sun, D. B. Cordes, C. Young, F. Tran, A. M. Z. Slawin, J. G. de Vries, P. C. J. Kamer, N. J. Westwood and K. Barta, *J. Am. Chem. Soc.*, 2016, **138** (28), 8900–8911.

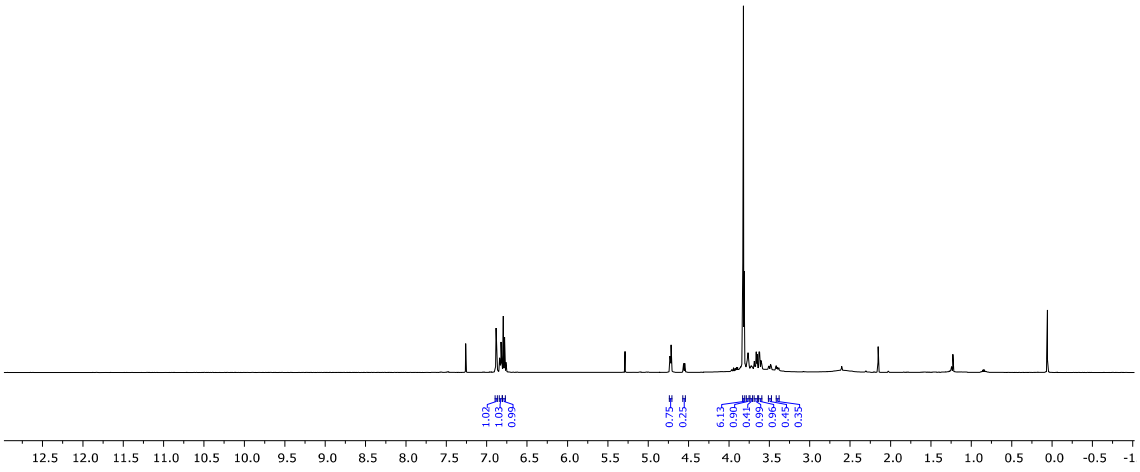
S16 P. Deuss, C. W. Lahive, C. S. Lancefield, N. J Westwood, P. C. J. Kamer, K. Barta, J. G. de Vries, *ChemSusChem.*, 2016, *Manuscript in Press*.

^{1}H and ^{13}C NMR Spectra of novel compounds

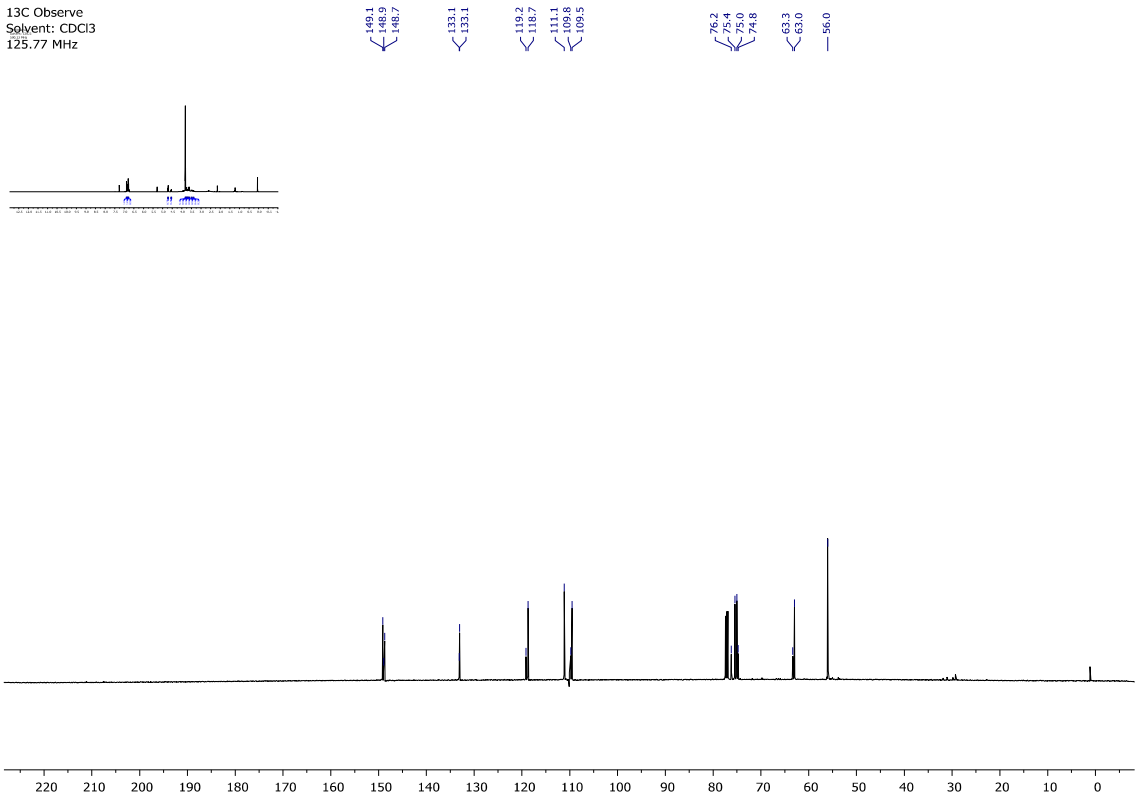
1H Observe
Solvent: CDCl3
500.13 MHz



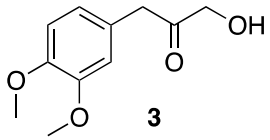
6



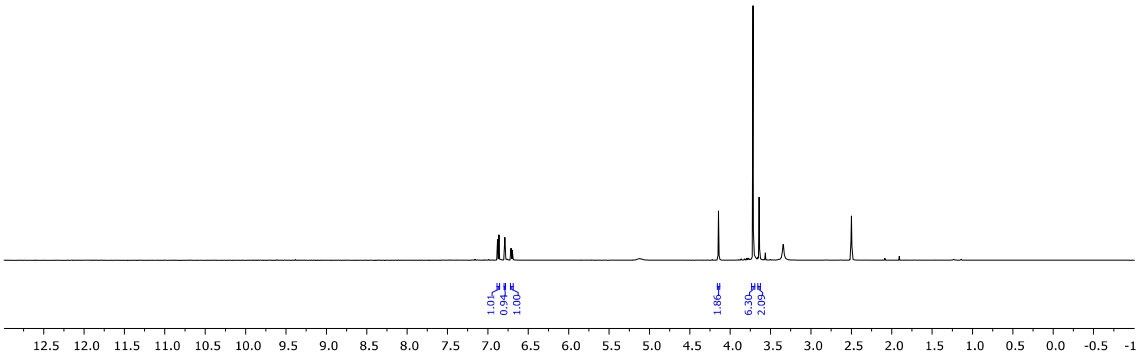
13C Observe
Solvent: CDCl3
125.77 MHz



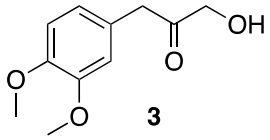
1H Observe
Solvent: DMSO
500.13 MHz



3

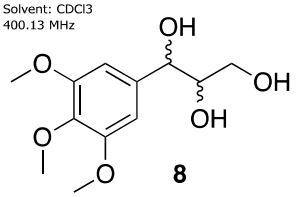


13C Observe
Solvent: DMSO
125.77 MHz



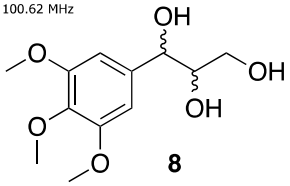
3

1H Observe
Solvent: CDCl3
400.13 MHz



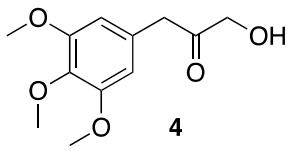
8

13C Observe
Solvent: CDCl3
100.62 MHz

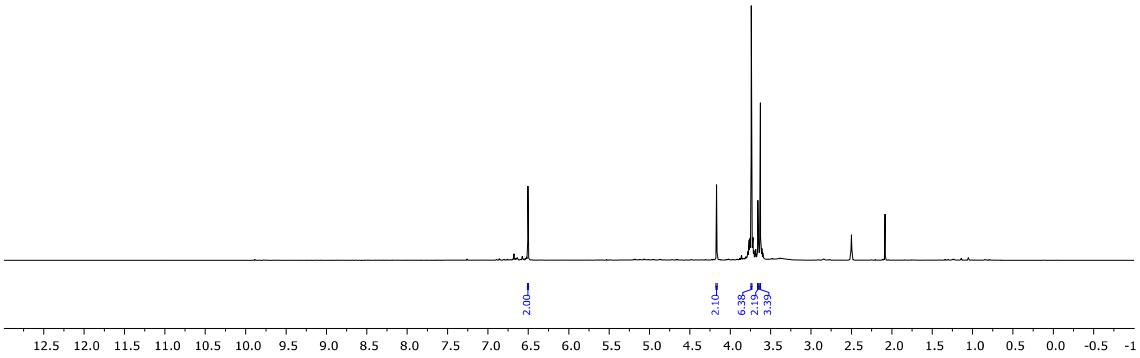


8

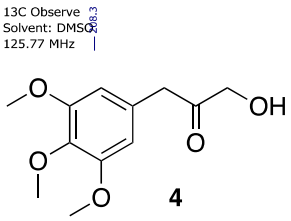
1H Observe
Solvent: DMSO
500.13 MHz



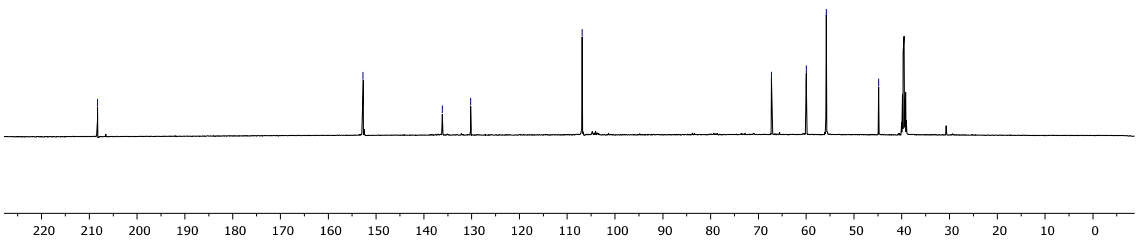
4



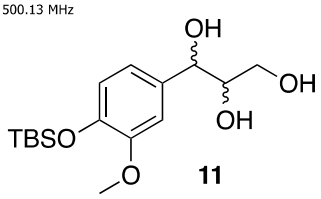
13C Observe
Solvent: DMSO
125.77 MHz



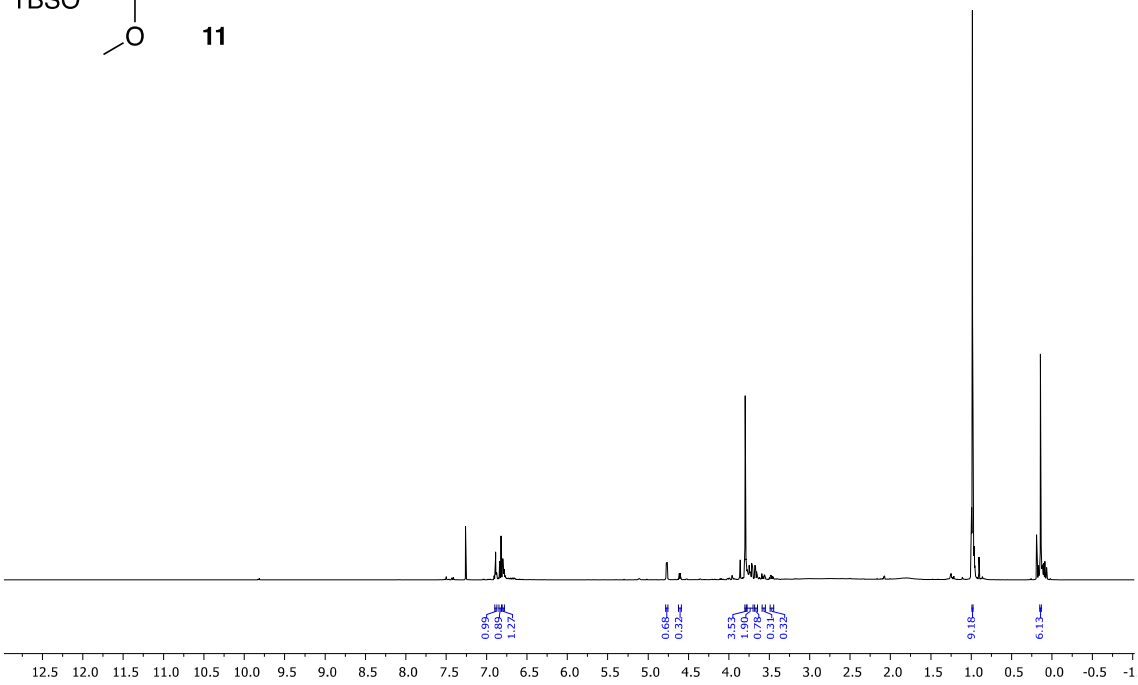
4



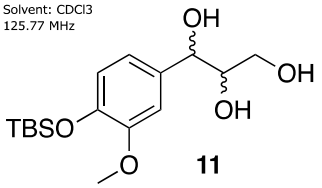
1H Observe
Solvent: CDCl3
500.13 MHz



11

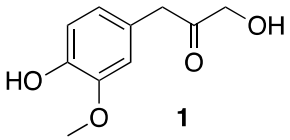


13C Observe
Solvent: CDCl3
125.77 MHz

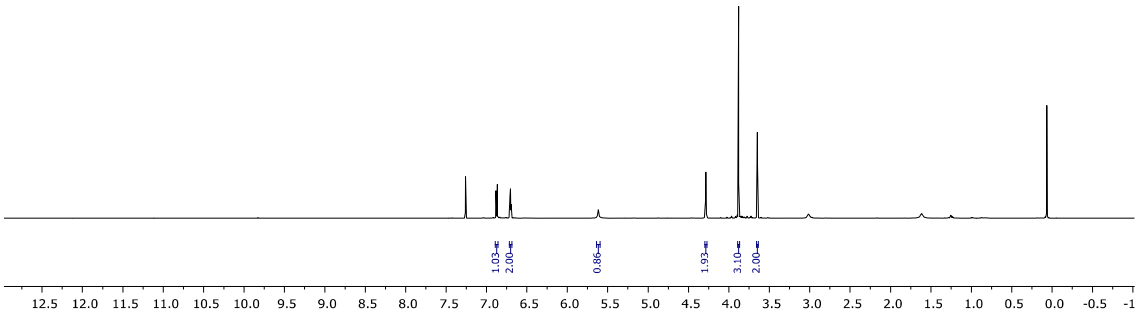


11

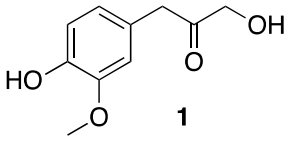
1H Observe
Solvent: CDCl3
499.93 MHz



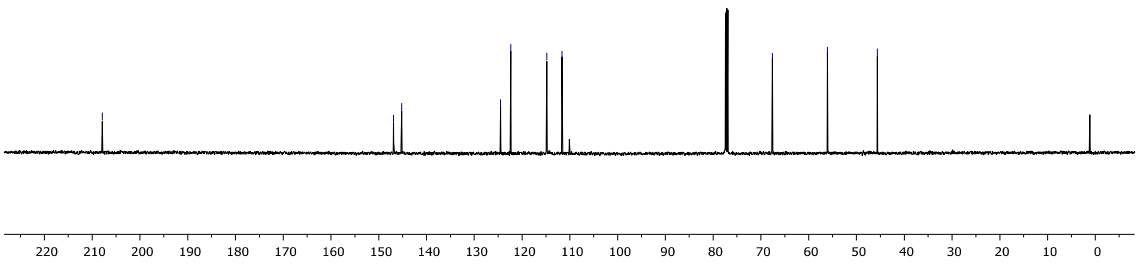
1

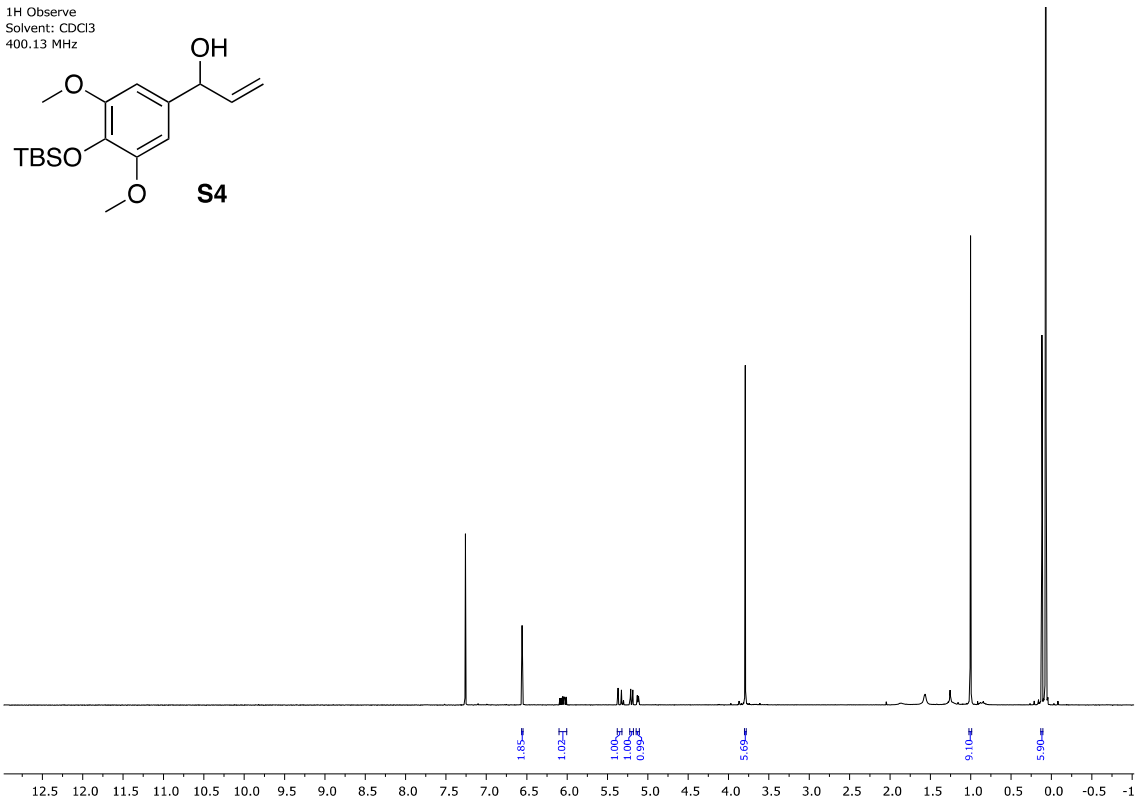


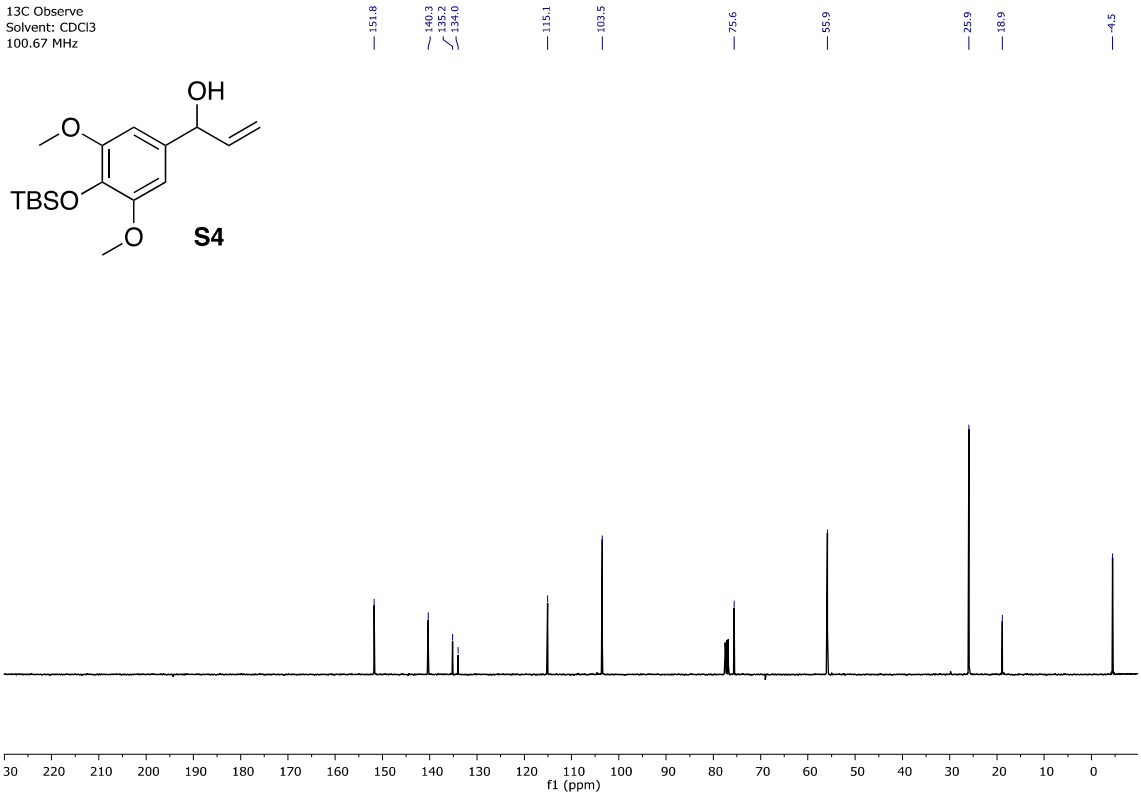
13C Observe
Solvent: CDCl3
125.72 MHz



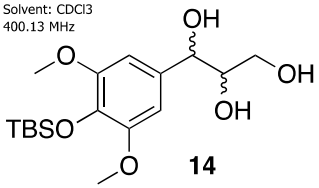
1




1H Observe
Solvent: CDCl3
400.13 MHz
OH
O
TBSO
O
S4
1.85
1.02
1.00
1.00
0.99
5.69
9.10
5.90
12.5 12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5 -1

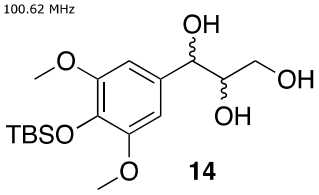

13C Observe
Solvent: CDCl3
100.67 MHz
151.8
140.3
135.2
134.0
115.1
103.5
75.6
55.9
25.9
18.9
-4.5
OH
O
TBSO
O
S4
30 220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

1H Observe
Solvent: CDCl3
400.13 MHz



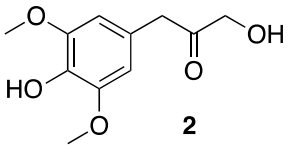
14

13C Observe
Solvent: CDCl3
100.62 MHz

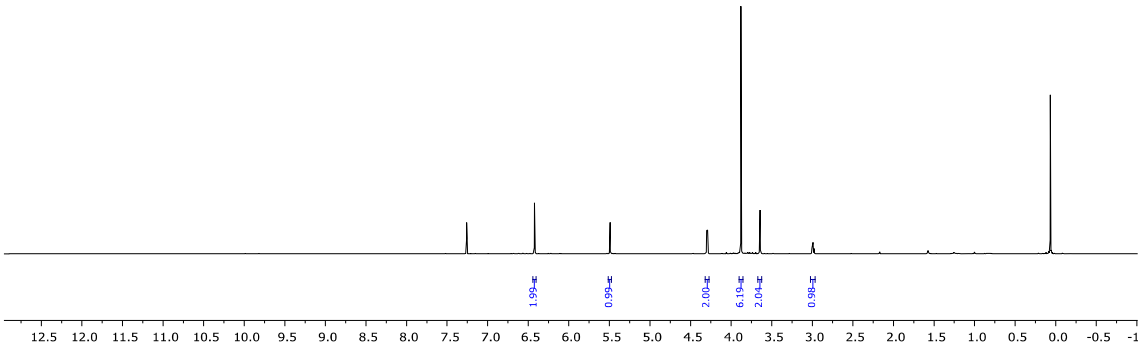


14

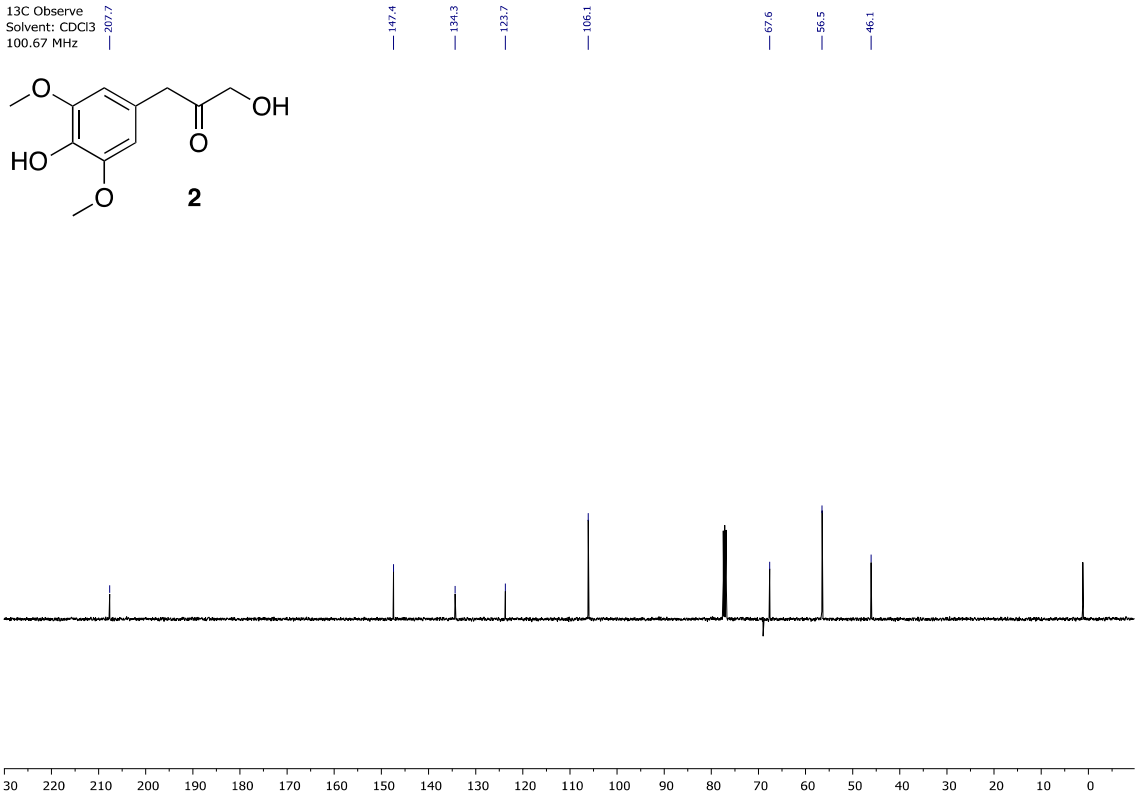
1H Observe
Solvent: CDCl3
400.30 MHz

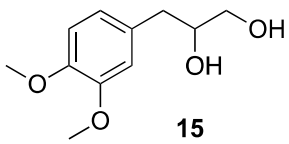


2

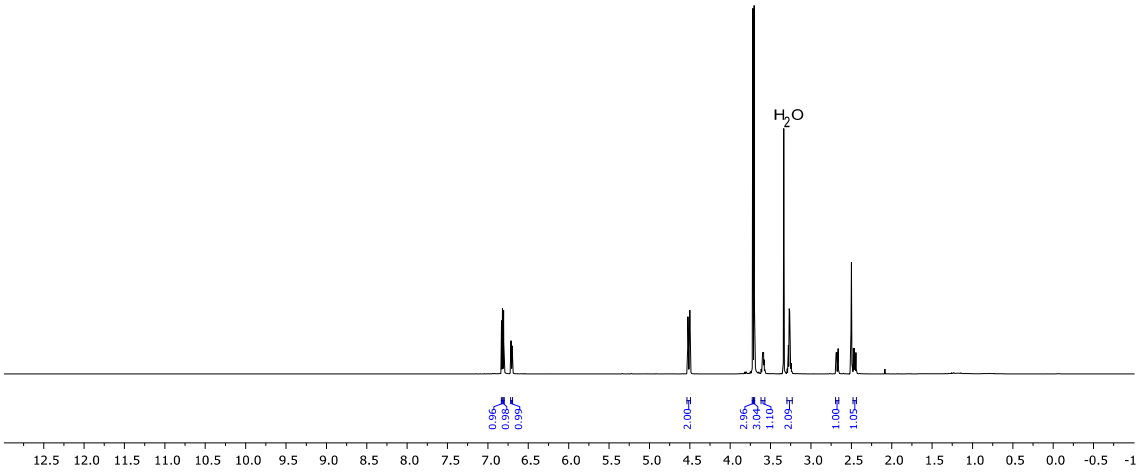


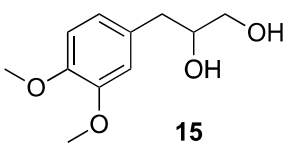
13C Observe
Solvent: CDCl3
100.67 MHz




OH
OH
O
O

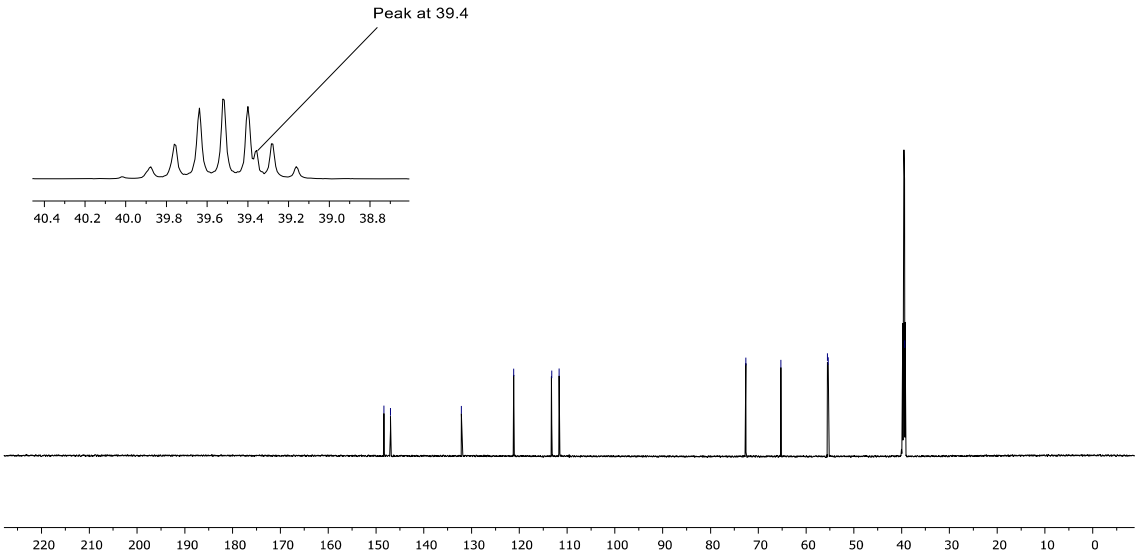
15

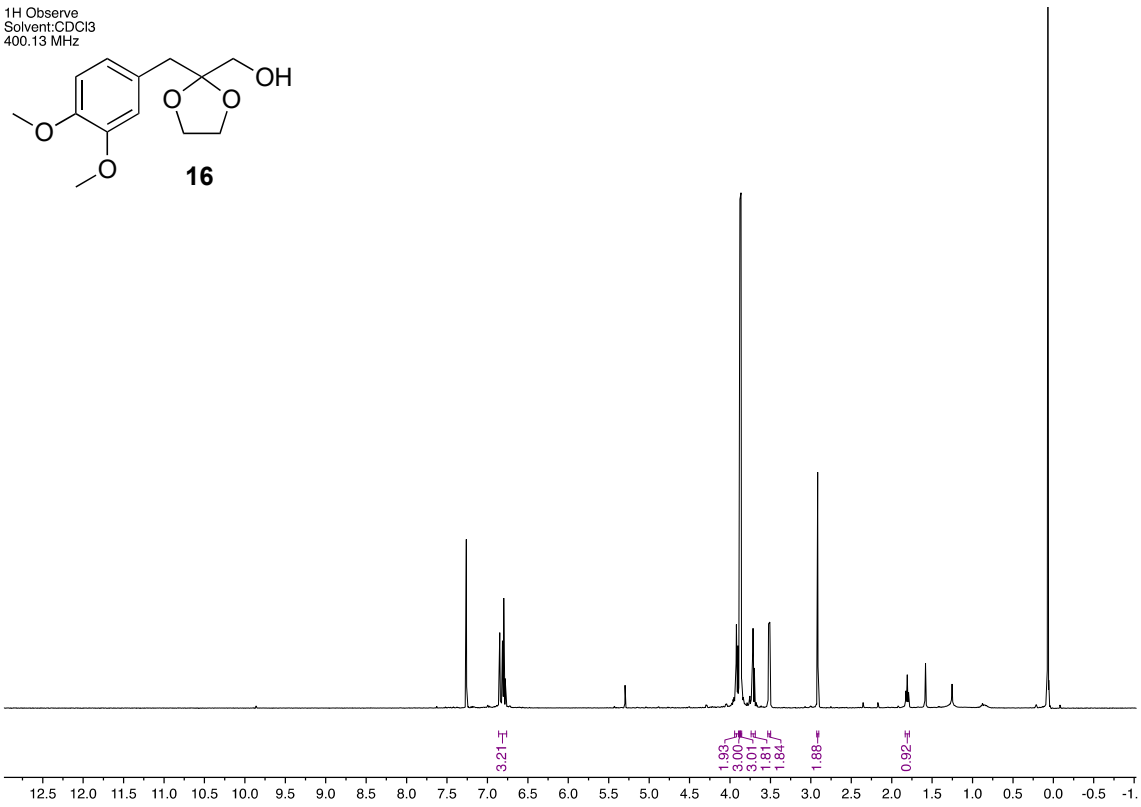

H2O
0.95
0.98
0.99
2.00
2.96
3.04
1.10
2.09
1.00
1.05
12.5 12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5 -1

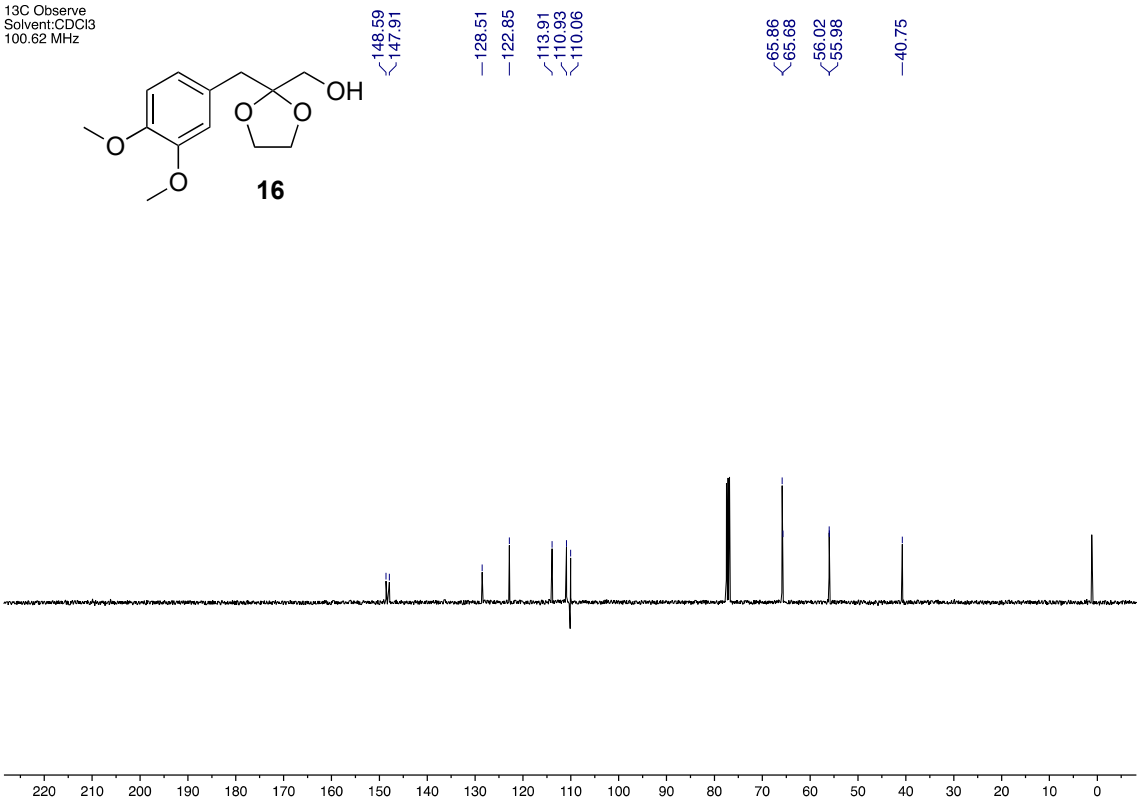

OH
OH
O
O

15

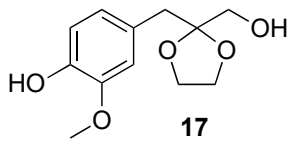

148.3
147.0
132.1
121.2
113.3
111.7
72.6
65.3
55.5
55.4
39.4


Peak at 39.4
40.4 40.2 40.0 39.8 39.6 39.4 39.2 39.0 38.8
220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0

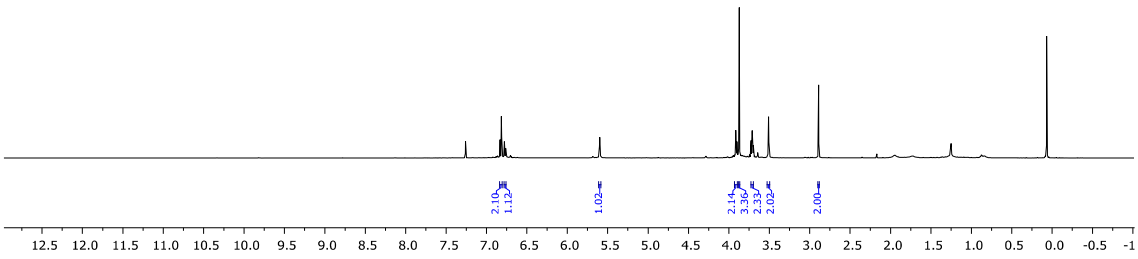

1H Observe
Solvent:CDCl3
400.13 MHz
OH
O
O
O
O
16
3.21
1.93
3.00
3.01
1.81
1.84
1.88
0.92
12.5 12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5


13C Observe
Solvent:CDCl3
100.62 MHz
148.59
147.91
128.51
122.85
113.91
110.93
110.06
65.86
65.68
56.02
55.98
40.75
OH
O
O
O
O
16
220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0

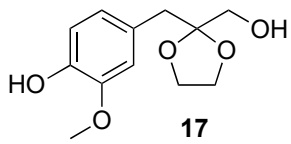
1H Observe
Solvent: CDCl3
499.93 MHz



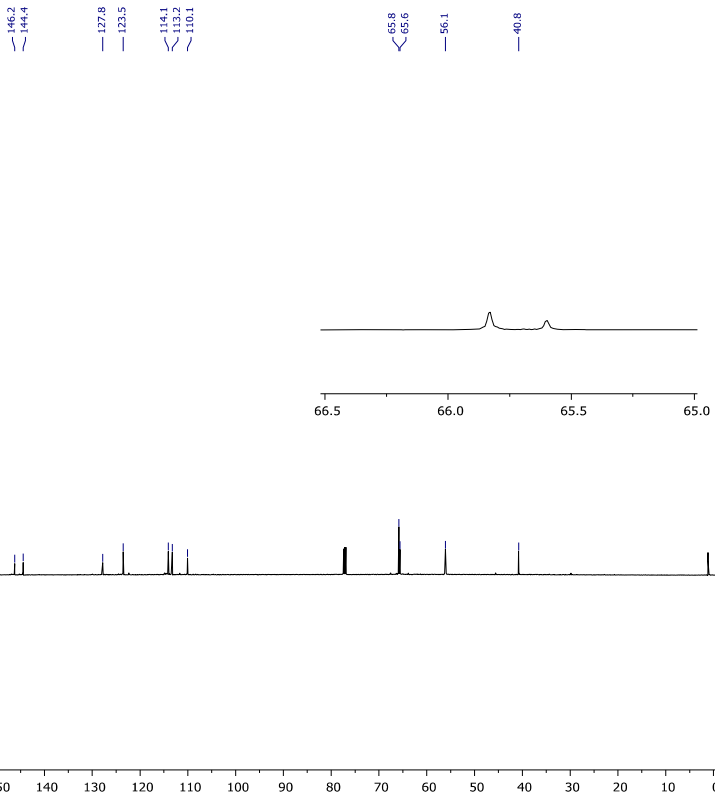
17



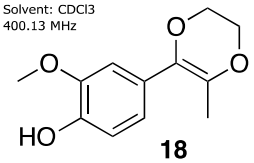
13C Observe
Solvent: CDCl3
125.77 MHz



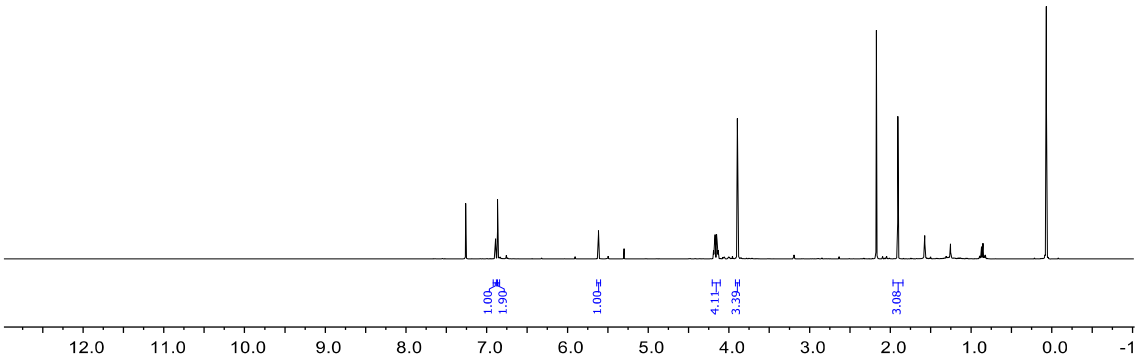
17



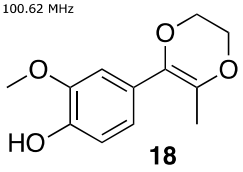
1H Observe
Solvent: CDCl3
400.13 MHz



18

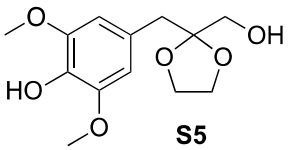


13C Observe
Solvent: CDCl3
100.62 MHz

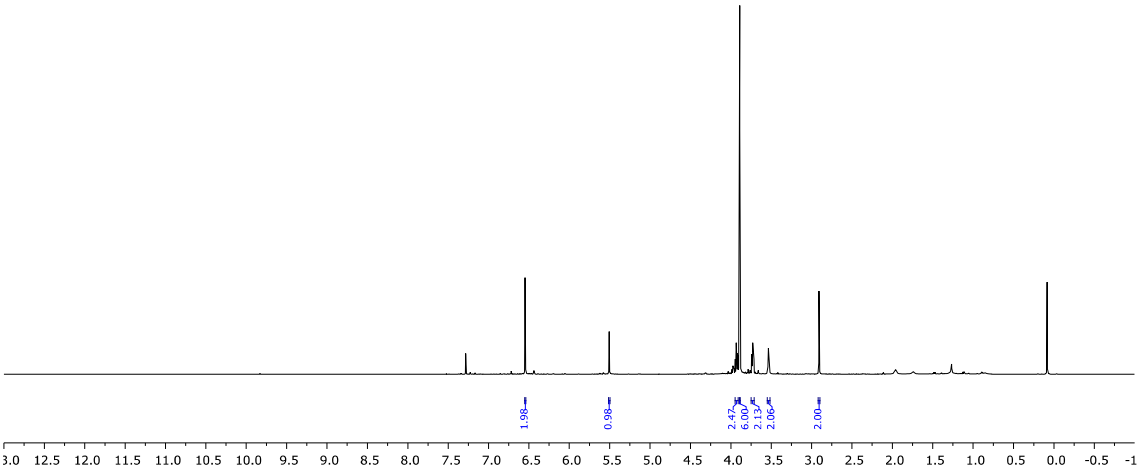


18

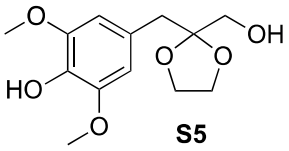
1H Observe
Solvent: CDCl3
500.13 MHz



S5



1H Observe
Solvent: CDCl3
125.77 MHz



S5