ELECTRONIC SUPPORTING INFORMATION FOR

Molecular structure and composition elucidation of an industrial humin and its fractions

Sandra Constant,[a,b] Christopher S. Lancefield,[a] Willem Vogelzang,[c] Rajeesh Kumar Pazhavelikkakath Purushothaman,[c] Augustinus E. Frissen,[c] Klaartje Houben,[d] Peter de Peinder,[e] Marc Baldus,[d] Bert M. Weckhuysen,[a] Daan S. van Es*[c] and Pieter C. A. Bruijnincx*[a,f]

[a]Inorganic Chemistry and Catalysis, Institute for Sustainable and Circular Chemistry and Debye Institute for Nanomaterials Science, Faculty of Science, Utrecht University, Universiteitsweg 99, 3584 CG Utrecht, The Netherlands.
[b]Avantium Renewable Polymers B.V., Zekeringstraat 29, 1014 BV Amsterdam, The Netherlands.
[c]Wageningen Food & Biobased Research, Bornse Weilanden 9, 6708 WG Wageningen, The Netherlands. Email: daan.vanes@wur.nl
[d]NMR Spectroscopy Research Group, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands.
[e]VibSpec, Haaftenlaan 28, 4006 XL Tiel, The Netherlands.
[f]Organic Chemistry and Catalysis, Institute for Sustainable and Circular Chemistry, Utrecht University, Faculty of Science, Universiteitsweg 99, 3584 CG Utrecht, The Netherlands.
Email: p.c.a.bruijnincx@uu.nl

Article: DOI:10.1039/d4gc00429a

Number of pages: 85
Number of Figures: 81
Number of Tables: 14

TABLE OF CONTENT

TABLES

FIGURES

A. Materials and methods

I. Materials

a) Humin sample

The industrial humin sample was provided by Avantium, The Netherlands, and produced by conversion of fructose in methanol solvent. The crude industrial humin (CIH) was obtained as a dark coloured syrup.

b) Materials for NMR analysis

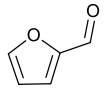
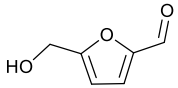
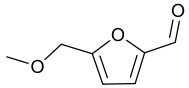
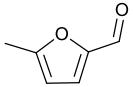
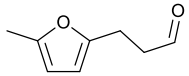
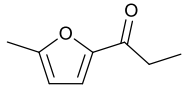
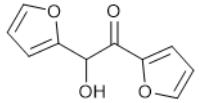
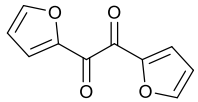
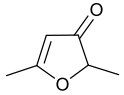
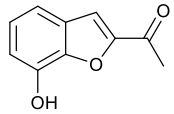
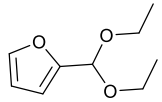
- **^{1}H and ^{13}C NMR analysis**
  Deuterated DMSO-d_6 (99.9%D) solvent was purchased from Cambridge Isotope Labs via Buchem.

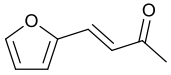
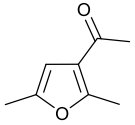
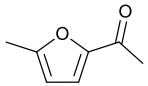
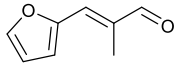
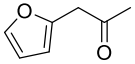
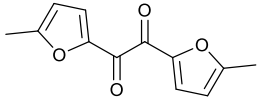
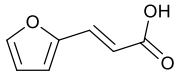
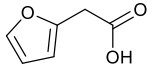
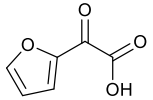
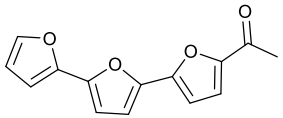
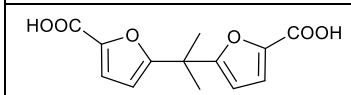
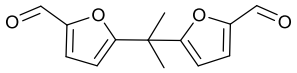
- **^{31}P NMR analysis**
  For ^{31}P NMR, high grade DMF (anhydrous, 99.8%), pyridine (anhydrous, 99.8%), cyclohexanol (Reagent Plus, 99%), chromium(III) acetylacetonate (99.99%), 2-chloro-4,4,5,5-tetramethyl-1,3,2-dioxaphopholane (95%), and $CDCl_3$ (99.8%D, contains 0.03% TMS) were purchased from Sigma- Aldrich.

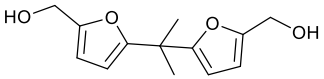
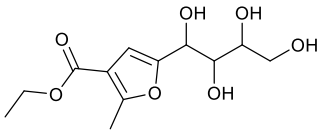
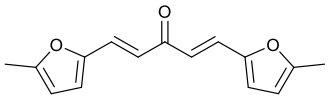
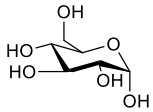
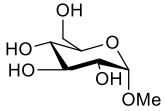
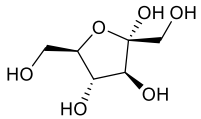
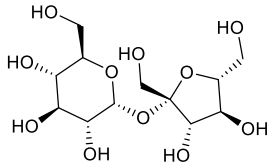
- **Reagents for ^{19}F NMR analysis**
  All reagents for the derivatisation step were obtained commercially: 4-(trifluoromethyl)phenylhydrazine (96%, Sigma Aldrich, stored dry at 4 °C and used in small batches), 1-methyl-4-(trifluoromethyl)benzene (98%, Sigma Aldrich) and chromium(III) acetylacetonate (97%, Acros).

- **Model compounds**
  Details about the origin and purity of the model compounds used in this study are reported in Table S1.

Table S 1. Origin and purity of the model compounds used

Compound	IUPAC Name (Chemdraw generated)	(Trivial) name	CAS	Origin/ purity
	furan-2-carbaldehyde	furfural	98-01-1	Sigma / 99%
	5-(hydroxymethyl)furan-2-carbaldehyde	hydroxymethylfurfural	67-47-0	Sigma / 99%
	5-(methoxymethyl)furan-2-carbaldehyde	methoxymethylfurfural	1917-64-2	Avantium
	5-methylfuran-2-carbaldehyde	5-methylfuran-2-carbaldehyde	620-02-0	Sigma / 99%
	3-(5-methylfuran-2-yl)propanal	3-(5-methyl-2-furyl)propionaldehyde	34756-16-6	TCI / 98%
	1-(5-methylfuran-2-yl)propan-1-one	2-methyl-5-propionylfuran	10599-69-6	TCI
	1,2-di(furan-2-yl)-2-hydroxyethan-1-one	furoin	552-86-3	TCI
	1,2-di(furan-2-yl)ethane-1,2-dione	furil	492-94-4	TCI
	2,5-dimethylfuran-3(2H)-one	2,5-dimethyl-3(2H)-furanone	14400-67-0	TCI
	1-(7-hydroxybenzofuran-2-yl)ethan-1-one	2-acetyl-7-hydroxybenzofuran	40020-87-9	TCI
	2-(diethoxymethyl)furan	2-furaldehyde Diethyl Acetal	13529-27-6	Sigma / 97%

Compound	Chemdraw Name	Usual name	CAS	Origin/ purity
	(E)-4-(furan-2-yl)but-3-en-2-one	4-(2-furyl)-3-buten-2-one	623-15-4	ABCR / 98%
	1-(2,5-dimethylfuran-3-yl)ethan-1-one	3-acetyl-2,5-dimethylfuran	10599-70-9	ABCR / 98%
	1-(5-methylfuran-2-yl)ethan-1-one	2-acetyl-5-methylfuran	1193-79-9	ABCR / 98%
	(E)-3-(furan-2-yl)-2-methylacrylaldehyde	2-methyl-3-furylacrolein	874-66-8	ABCR
	1-(furan-2-yl)propan-2-one	1-(2-furyl)acetone	6975-60-6	ABCR / 99%
	1,2-bis(5-methylfuran-2-yl)ethane-1,2-dione	1,2-bis(5-methylfuran-2-yl)ethane-1,2-dione	51079-12-0	ABCR
	(E)-3-(furan-2-yl)acrylic acid	3-(2-furyl)acrylic acid	539-47-9	Sigma / 99%
	2-(furan-2-yl)acetic acid	2-furanacetic acid	2745-26-8	Sigma / 97%
	2-(furan-2-yl)-2-oxoacetic acid	α-oxo-2-furanacetic acid	1467-70-5	Sigma / >97%
	1-([2,2':5',2''-terfuran]-5-yl)ethan-1-one	1-([2,2':5',2''-terfuran]-5-yl)ethan-1-one RCL L167177	892151-04-1	Sigma / CPR
	5,5'-(propane-2,2-diyl)bis(furan-2-carboxylic acid)	5,5'-(propane-2,2-diyl)bis(furan-2-carboxylic acid)	91903-78-5	WFBR
	5,5'-(propane-2,2-diyl)bis(furan-2-carbaldehyde)	5,5'-(propane-2,2-diyl)bis(furan-2-carbaldehyde)	145659-75-2	WFBR

Compound	Chemdraw Name	Usual name	CAS	Origin/ purity
	(propane-2,2-diylbis(furan-5,2-diyl))dimethanol		1083171-37-2	WFBR
	ethyl 2-methyl-5-(1,2,3,4-tetrahydroxybutyl)furan-3-carboxylate	4-ethoxycarbonyl-2-(D-arabino-tetrahydroxybutyl)-5-methylfuran.	6330-61-6	WFBR
	(1E,4E)-1,5-bis(5-methylfuran-2-yl)penta-1,4-dien-3-one	1,5-bis(5-methyl-2-furyl)-1,4-pentadien-3-one	69239-15-2	Sigma / CPR
		α-D-glucose	492-62-6	Sigma / 96%
		α-methyl glucoside	97-30-3	Acros / 98%
		β-D-fructose	57-48-7	Acros / 99%
		sucrose	57-50-1	Sigma / ACS reagent

II. Crude industrial humin (CIH) purification and fractionation

a) Purification by solvent extraction of CIH

A three-step procedure was developed, aiming at optimal removal of the low molecular weight compounds. The first purification step consisted of washing the syrupy CIH with diethyl ether in small batches to optimise the purification. All residues were then combined and a second purification step was applied by repeated extensive washings with diethyl ether until no more HMF or MMF was detected in the extract by GC-MS analysis. Finally, the final purification step consisted of Soxhlet extraction of the residue with diethyl ether followed by drying under vacuum, giving the PIH sample as a dark brown powder. The experimental details of this three-step process are described below.

Step 1:
The first step of purification by solvent extraction was performed in five different batches using the same batch of crude industrial humin (CIH) supplied by Avantium (Table S2). In a typical extraction step, a round bottom flask equipped with a mechanical stirrer, was charged with the syrupy CIH and diethyl ether. The resulting black suspension was subsequently vigorously stirred for two hours. Next, the black suspension was filtered over a paper filter. The residue was dried in a vacuum oven (~20 mbar, 40 °C) over Sicapent® for 16 h, giving a dark brown powder. The filtrate was evaporated to dryness using a rotary evaporator (45 °C, 8 mbar), giving the diethyl ether extractives in the form of a brown oil (DEE). The mass recovery after solvent extraction (sum of solid and extractives) ranged between 80-92%. The mass loss (13%) can be attributed to the removal of volatiles and water in CIH.

Table S 2. Preparation of PIH from CIH: washing of CIH with diethyl ether: extraction volumes and masses of the different fractions

Batch	**m_{CIH} (g)** Starting mass of CIH to purify	**$v_{diethyl\ ether}$ (mL)** Volume of diethyl ether used for the washing CIH	**m_{PIH} (g, (wt%))** Dried mass of the Purified Humin fraction	**m_{DEE} (g, (wt%))** Dried mass of the Diethyl ether extractives fraction	**m_{total} (g)** Sum of m_{PIH} and m_{DEE}
1	98.9	800	53.2	31.2	84.9
2	25.1	500	15.4	7.7	23.1
3	25.5	500	17.6	5.7	23.3
4	25.0	500	14.5	5.4	19.9
5	25.1	500	16.2	6.1	22.3
Sum	199.6 g		116.9 g (59 wt%)	56.1 g (28 wt%)	173 g (87 wt%)

Step 2:
All residues (~117 g) were combined and repeatedly washed with diethyl ether (20 x 800 mL) until no more HMF or MMF was observed by GC-MS analysis of the liquid phase. The combined ether phases were evaporated to dryness, yielding a dark brown oil (14 g, 12%). The solid residues were dried in a vacuum oven (~20 mbar, 40 °C) over Sicapent® for 16 h, giving a dark brown powder (80 g, 68 wt%). Total mass recovery was 80 wt%.

Step 3:
The final purification step was performed using a 2 L Soxhlet extractor. The dark brown powder (80 g) was placed in a thimble and washed with diethyl ether (3 L) for 24 h. Finally, the powder was dried in a vacuum oven (~20 mbar, 40 °C) over Sicapent® for 16 h, again giving a brown powder (purified industrial humin, PIH) (78 g, 97.5 wt%). Evaporation of the diethyl ether used for washing the sample yielded a brown oil (2 g, 2.5 wt%).

The diethyl ether fractions obtained in steps 1-3 were combined (DEE, 72.1 g) and qualitatively analysed by GC-MS, indicating the presence of hydroxymethylfurfural (HMF), methoxymethylfurfural (MMF) with some minor amounts of methyl levulinate (ML) and furfuraldehyde (Figure S1).

To summarize the mass balances of the purification step: 78 g of PIH (39.1 wt%) and 72.1 g of DEE (36.1 wt%) were obtained from 199.6 g of CIH. The mass loss difference (24.5 wt%) can be attributed to losses occurring during the purification, to the removal of volatiles and water in CIH.

b) Fractionation of PIH into WIPIH and WES

PIH (1 g) was suspended in 75 mL of deionised water in a 100 mL round bottom flask. The mixture was stirred for 24 h at room temperature under magnetic stirring (600 rpm), and subsequently filtered over a paper filter. The aqueous filtrate was then concentrated using a rotary evaporator. Both the aqueous concentrate and the water-insoluble residue were finally dried in a vacuum oven over Sicapent® (~20 mbar, 40 °C) for 16 h, yielding dry powders. The fractions were subsequently weighed to give 55 wt% of a water-soluble fraction (WES) and 45 wt% of water-insoluble purified industrial humin (WIPIH).

III. Acid-Treated Purified Industrial Humin (ATPIH)

PIH (1.03 g) was suspended in 50 mL 0.1 M H_2SO_4 in a Parr 75 mL reactor under nitrogen atmosphere. The mixture was magnetically stirred (800 rpm) for 6 h at 180 °C (autogenous pressure of 5.6 bar). Subsequently, the suspension was allowed to cool down to RT (residual pressure of 0.6 bar) and filtered over a paper filter. This material was subsequently dried in a vacuum oven over Sicapent® (20 mbar, 40 °C) for 16 h, yielding a dry black powder (0.48 g). To remove the residual H_2SO_4, the crude acid treated humin was Soxhlet-extracted with water for 24 h.

IV. Analytical methods

a) GPC

Table S 3. GPC methods

Method	Eluent	Column	Column specification	Standard	Detection method
A	THF	PL-gel Mixed-E columns	Up to 25,000 g mol^{-1}	PS	UV
B	HFIP	PSS PFG	~250 – 2,500,000 g mol^{-1}	PMMA	Refractive Index

Method A (THF):
GPC measurements were performed on a Polymer Labs GPC 50 system, equipped with a series of 3xPL-gel Mixed-E columns were used at 40 °C with a 1 mL/min flow rate of stabilised THF with 250 ppm of 2,5-di-t-butylhydroxytoluene and 1 % v/v of acetic acid using polystyrene standards for calibration and toluene as flow marker. Detection was with an external Knauer UV detector at 280 nm and molecular weight determinations were based on calibration with polystyrene standards (Mn=162, 570, 1060, 1400, 2240, 3690, 4760, 7130, 12 800 and 19 690).
Method B (HFIP):
Relative molecular weights were determined by Gel Permeation Chromatography (GPC) on a Viscotek HP-SEC system, VE-2001 GPC max (pump and auto sampler) equipped with a TDA305 Triple Detector Array, Refractive Index was used as detection system. 2xPSS PFG GPC analytical linear M columns and guard column. Hexafluoroisopropan-2-ol (HFIP, Apollo Scientific Limited) containing 0.02 M potassium trifluoroacetate was used as an eluent with a flow-rate of 0.7 ml min^{-1}. Control measurements were performed with Easy vials PMMA standards (550-2140000 g mol^{-1}) from Agilent Technologies. Prior analysis the samples were dissolved in the eluent using a low profile laboratory reciprocal shaker (IKA HS 501 digital) for 16 h, at RT.

b) Residual sugar analysis

The residual sugars were determined using two analytical methods. The hot water extraction method was used to quantify monomeric reducing sugars or adsorbed sugars. While the chemically bound/incorporated neutral sugars were determined after sulphuric acid hydrolysis. For both methods, the analysis was done by HPAEC with pulsed amperometric detection on a CarboPac PA1 column (Dionex) with a water-sodium hydroxide gradient.

Method for monomeric reducing sugars or adsorbed sugars (hot water extraction):
2 g of samples were extracted with demineralised water at 100 °C for 1 h and subsequently analysed by HPAEC.

Method for chemically bound/incorporated neutral sugars (sulfuric acid hydrolysis):
The acid hydrolysis was performed by reacting 50 mg of humin with 5 mL of a 1 M sulfuric acid solution for 3 hours at 100 °C. The hydrolysates were subsequently analysed by HPAEC.

c) Elemental Analysis

Elemental analyses were performed on a CHNS analyser Model Vario Micro Cube of the brand Elementar.

d) Thermogravimetric Analysis (TGA)

Thermogravimetric analysis was performed on an STA 6000 (Simultaneous Thermal Analyzer) from PerkinElmer Instruments. The samples were heated from 30 to 900 °C at a heating rate of 10 °C min^{-1} under a nitrogen flow of 20 mL min^{-1}.

e) Solid-state NMR

Solid-state nuclear magnetic resonance (ss-NMR) experiments were performed using a Bruker Avance III spectrometer operating at a ^{1}H Larmor frequency of 500 MHz equipped with a 3.2 mm magic angle spinning (MAS) probe at a temperature of 298 °K at a MAS rate of 10 kHz. Referencing of ^{1}H and ^{13}C chemical shifts was done externally to adamantane. All spectra were processed and analysed with Bruker Topspin3.5.

In general, the experimental conditions were as follows: Radio frequency (r.f.) power levels used for ^{1}H decoupling and ^{1}H/^{13}C cross polarisation were set to 94 kHz and 58 kHz/50 kHz with a 70-100 % ramp, respectively. The recycle delay was set to 10 s, and the number of recorded transients varied between 7k and 14k. The dipolar dephasing (DD)-MAS spectrum was recorded using a cross polarisation (CP) pulse program with rotor synchronised dephasing time of 68 µs and refocusing pulse in order to selectively observe the quaternary (and mobile) carbons.
While recording a quantitative direct excitation spectrum was precluded due to long relaxation times, we chose to optimise the CP settings. Having investigated various contact times, 800 µs was used as an optimum for the observation of all the different type of carbons: primary to quaternary (Figure S12).

The ss-NMR-MAS spectra of PIH (Figure S13) were recorded using cross polarisation (CP) in combination with DD methods. The ss-NMR-MAS spectra of WIPIH (Figure 4) and ATPIH (Figure S32 and S33) were recorded using cross polarisation (CP).

f) FT-IR

Fourier transform infrared spectroscopy measurements were carried out at room temperature on a Bruker Tensor 27 instrument. FT-IR data were recorded with a deuterated triglycerine sulphate (DTGS) detector. The samples were recorded using a KBr pellet in transmission mode or using a diamond ATR accessory. The optical resolution of the IR spectra was 4 cm^{-1} and 16 scans were accumulated for each spectrum.

g) SEM

The SEM measurements were carried out on an FEI Helios G3 UC microscope after coating of the humin samples.

h) Solution-state NMR

The NMR spectra of the model compounds were measured on a Varian 400 MHz NMR and a Bruker Avance III 400 MHz spectrometer. The NMR spectra of the humin samples were recorded on a Bruker Avance II 600 MHz spectrometer equipped with a 5 mm CPTCI ^{1}H-^{13}C/^{15}N/^{2}H cryogenic probe. The substances were solubilised in DMSO-d_6 and the spectra referenced against the signal of the residual protio impurity of the solvent (^{1}H 2.50 ppm, ^{13}C 39.52 ppm). The ^{1}H, ^{13}C and HSQC spectra were processed using MestReNova software.

<u>Model compounds:</u>
The 1D ^{1}H NMR measurements of the model compounds were performed using a standard proton pulse program, a relaxation delay of 2 s and 32 acquired scans. The 2D [^{1}H;^{13}C] NMR measurements of the model compounds were carried out using a hsqcetgpsi2 pulse program, a relaxation delay of 1 s and 32 acquired scans.

<u>Humin samples:</u>
The 1D ^{1}H NMR measurements of the humin samples were carried out using a standard proton pulse program, a recycle delay of 1 s and 64 acquired scans. The 1D ^{13}C NMR measurements were carried out using a zgpg pulse program, which uses a 90° pulse and power-gated WALTZ-16 ^{1}H-decoupling, a recycle delay of 2 s and 6144 acquired scans. The 2D [^{1}H;^{13}C] NMR measurements were performed using a hsqcetgpsp pulse program, a relaxation delay of 1 s and 128 acquired scans.

i) ^{31}P NMR

<u>Phospholane derivatisation of the humin samples</u>

A dried solvent mixture composed of pyridine/deuterated chloroform (1.6/1.0 v/v) was prepared (protected from moisture with 3 Å molecular sieves). Stock solutions of relaxation reagent (chromium (III) acetylacetonate, 11.4 mg/mL) and internal standard (cyclohexanol, 19 mg/mL) were prepared in the dry solvent mixture.
Approximately 30 mg of humin sample was mixed with 200 µl of internal standard and 50 µL of relaxation reagent solutions in a GC vial. The components were mixed thoroughly at room temperature for 18h under magnetic stirring to ensure complete solubilisation of the sample. Next,100 µL of derivatisation reagent (2-chloro-4,4,5,5-tetramethyl-1,3,2- dioxaphospholane) was added to this mixture and stirred for at least 10 minutes before the mixture was transferred into a 5-mm-OD NMR tube.

Nuclear Magnetic Resonance Measurements
^{31}P NMR spectra were obtained on a Bruker Avance III 400 MHz spectrometer using a standard phosphorus pulse programme with a relaxation delay of 10 s and 512 acquired scans. Chemical shifts were referenced from the sharp signal arising from the reaction product between residual water and 2-chloro-4,4,5,5-tetramethyl-1,3,2-dioxaphospholane at 132.2 ppm.

j) ^{19}F NMR

Hydrazine derivatisation of the humin samples
Around 150 mg of PIH sample (accurately weighed) and 10 mg of 1-methyl-4-(trifluoromethyl)benzene (internal standard) were dissolved in 500 µL of DMSO-d_6. The mixture was then manually homogenised until there is full dissolution of the polymer. In another vial, 0.8 mmol of 4-(trifluoromethyl)-phenylhydrazine was dissolved in 200 µL of DMSO-d_6 and added dropwise to the solution containing the substrate. The mixture was transferred into a standard NMR tube and placed in an oven at 40 °C for 24 h to ensure complete derivatisation. Just prior to analysis, 100 µL of relaxation agent solution (chromium acetylacetonate, 28 mg/mL in DMSO-d_6, 0.008 mmol) was added to the NMR tube after which the sample was homogenised and analysed immediately.
The ^{19}F method has been slightly modified to adjust the reagent amount with the quantity of WIPIH available: 70 mg of WIPIH, 5 mg of 1-methyl-4-(trifluoromethyl)benzene (internal standard) and 0.5 mmol of 4-(trifluoromethyl)phenylhydrazine were used.

Nuclear Magnetic Resonance Measurements
The NMR spectra were measured on a Varian 400 MHz NMR spectrometer. The ^{19}F NMR measurements were carried out using a standard fluorine 90° pulse program, a relaxation delay of 3 s, and 256 acquired scans between -50 and -70 ppm. The spectra were processed using MestReNova software; a Bernstein polynomial (order 3) baseline correction was applied, and chemical shifts were referenced to the 1-methyl-4-(trifluoromethyl)benzene signal at -60.90 ppm, prior to integration of the peaks.
The commercial 4-(trifluoromethyl)phenylhydrazine reagent contained an impurity (~3%), giving rise to a low-intensity peak in the ^{19}F NMR spectrum at -59.48 ppm. The amount of this impurity was determined for every batch of hydrazine used and corrected for the specific impurity concentration in the batch of hydrazine used.

Calculation
The number of millimoles of carbonyl groups per gram of sample corresponds to the millimoles of hydrazone groups per gram of sample and was calculated from the NMR spectra according to the following equation.

$$\text{mmol carbonyl group/g sample} = \frac{\left[A_C - \left(\frac{A_{H^\circ} \times m_H}{m_{H^\circ}}\right)\right] \times m_{IS}}{A_{IS} \times m_C \times M_{IS}} \times 10^3$$

A_C corresponds to the area of the hydrazone peak; the integrated area was set from −-59.20 to -60.10 ppm. A_{IS} corresponds to the area of the internal standard peak normalised to 1.00. A_{H° represents the area of the peak at -59.48 ppm in the ^{19}F spectrum of a sample of 4-(trifluoromethyl)phenylhydrazine and 1-methyl-4-(trifluoromethyl)benzene with the area of the internal standard peak normalised to 1.00. m_{H° corresponds to the mass of hydrazine of the sample containing 4-(trifluoromethyl)phenylhydrazine and 1-methyl-4-(trifluoromethyl)-benzene (mg). m_{IS} is the mass of the internal standard (mg); m_C is the mass of humin (mg). m_H corresponds to the mass of hydrazine (mg). M_{IS} is the molar mass of the internal standard (160.14 g.mol^{-1}).

B. Additional data on CIH, DEE, PIH, WIPIH and WES

I. Purification and further fractionation of the industrial humin

a) Preparation of PIH from CIH

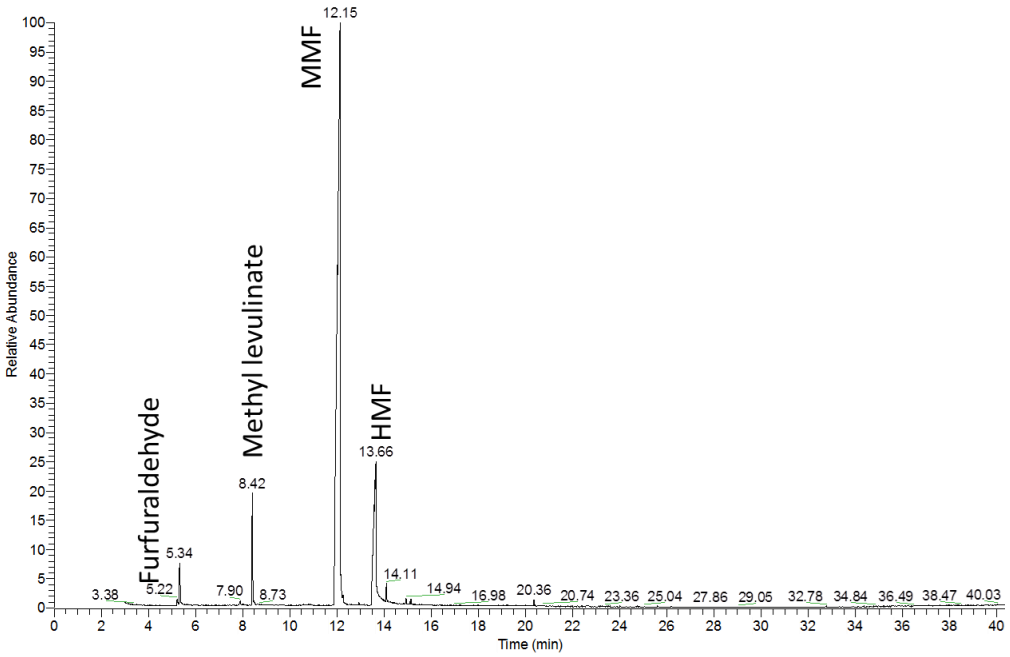


Figure S 1. GC-MS analysis of DEE.

b) Solubility of PIH

The solubility of PIH was determined by stirring 50 mg of sample with 5 mL of solvent for 24 hours.

Table S 4. Solubility of PIH in various organic solvents, by visual appearance.

Solvent	
Acetic acid	+
Dimethyl sulfoxide	+
Diethyl ether	-
Formic acid	+
Ethanol	+/-
HFIP	+
TFA	+
0.5 M NaOH	+
THF	+

II. Residual sugar analysis

Table S 5. Adsorbed residual sugar analysis after direct hot water extraction.

Sample	Carbohydrates (wt%)							Sum
	Arabinose	Xylose	Mannose	Galactose	Glucose	Fructose	Sucrose	(wt%)
CIH	0.00	0.00	0.01	0.00	0.06	0.04	0.37	0.48
PIH	0.00	0.00	0.02	0.00	0.09	0.05	0.36	0.52
WES	0.04	0.00	0.03	0.11	0.33	0.65	0.00	1.16
WIPIH	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table S 6. Sugar analysis after aqueous sulphuric acid hydrolysis.

Sample	Carbohydrates (wt%)							Sum
	Arabinose	Xylose	Mannose	Galactose	Glucose	Fructose	Sucrose	(wt%)
CIH	0.00	0.00	0.17	0.00	0.69	0.46	0.22	1.54
PIH	0.00	0.00	0.21	0.00	0.89	0.59	0.26	1.95
WES	0.04	0.02	0.45	1.54	0.01	0.33	0.00	2.40
WIPIH	0.00	0.00	0.02	0.02	0.00	0.00	0.00	0.05

III. Elemental analysis

Table S 7. Elemental composition of the humin samples.

Humin samples	C (%)	H (%)	O (%)	H/C	O/C
PIH	59	5	37	1.01	0.47
WES	49	5	46	1.21	0.70
WIPIH	63	5	32	0.94	0.38
ATPIH	65	4.5	29.5	0.82	0.34
Fructose-based humin[1]				0.76	0.36
Glucose-based humin[1]				0.79	0.36
Glucose-based HTC[2]	62	4	34	0.77	0.41
Water soluble oligomer from glucose[3]	53.4	6.5	39	1.97	0.98
Glucose-based humin[3]				0.85	0.44

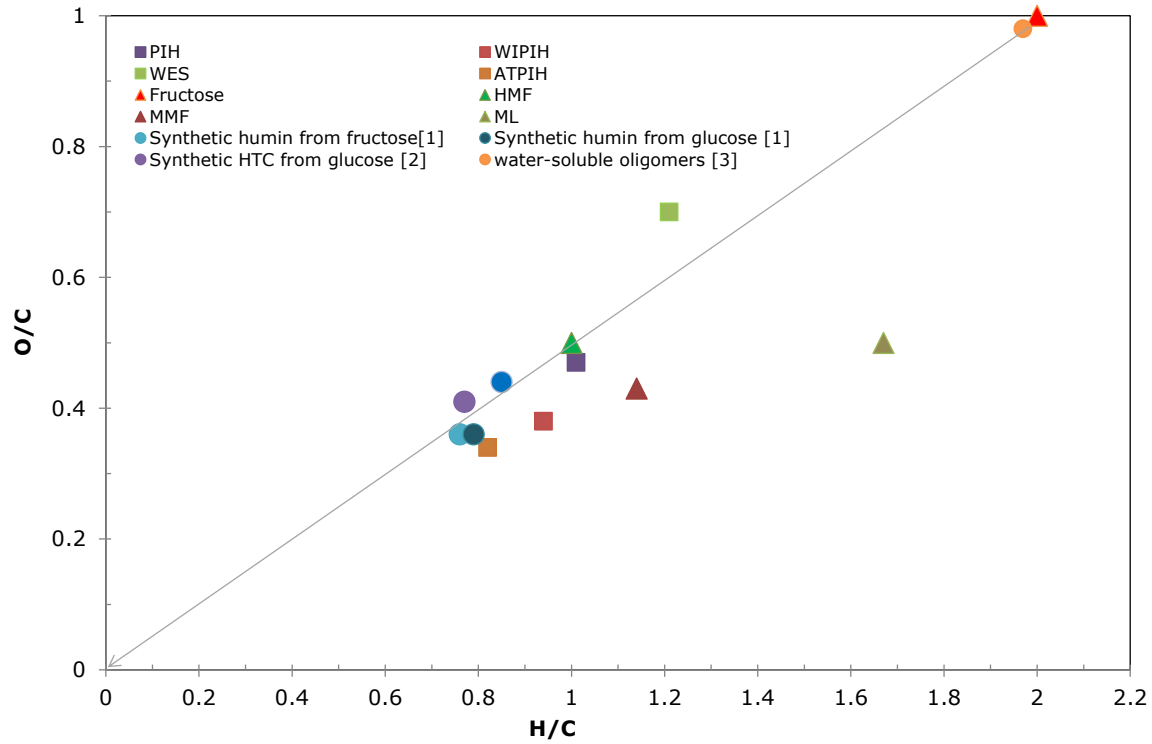


Figure S 2. Van Krevelen plot of the humin samples (PIH, WIPIH, WES and ATPIH), chemicals (fructose, HMF, MMF and ML) and various humin-like materials.[1–3]

IV. Thermogravimetric Analysis (TGA)

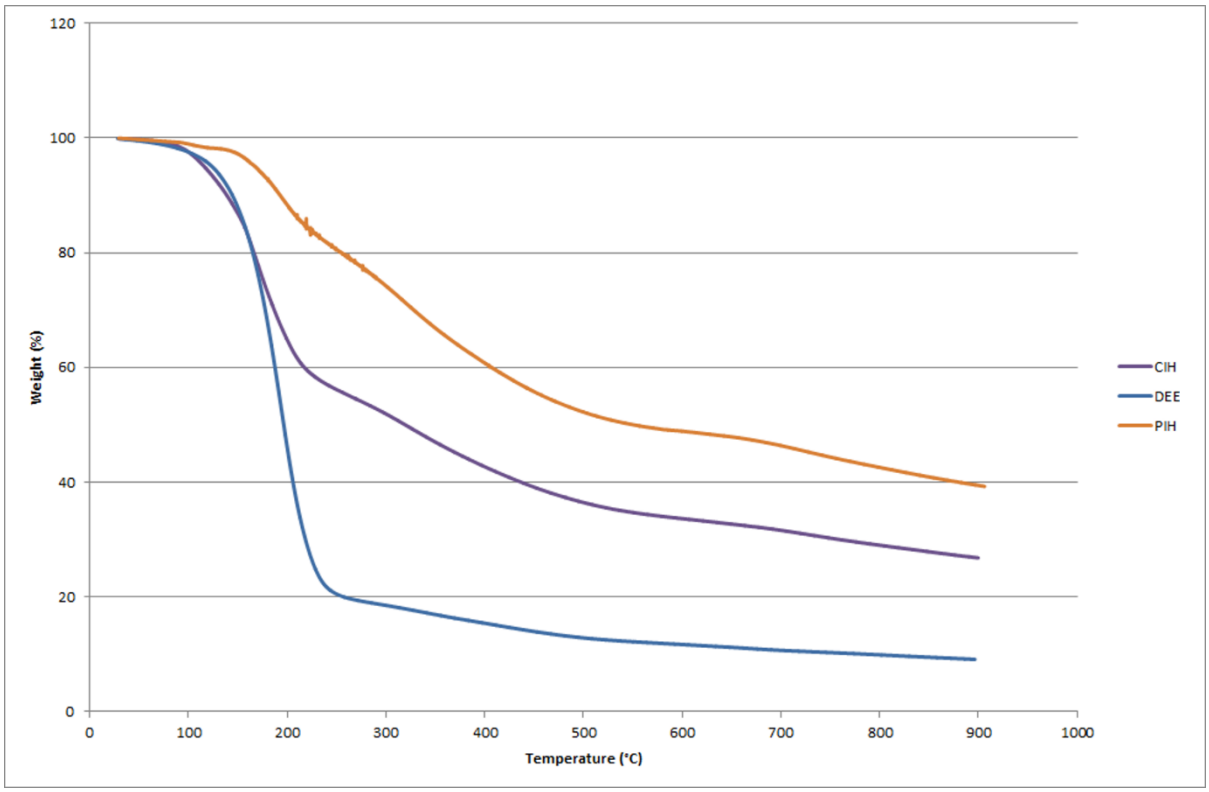


Figure S 3.TGA traces of CIH, DEE and PIH measured under N_2.

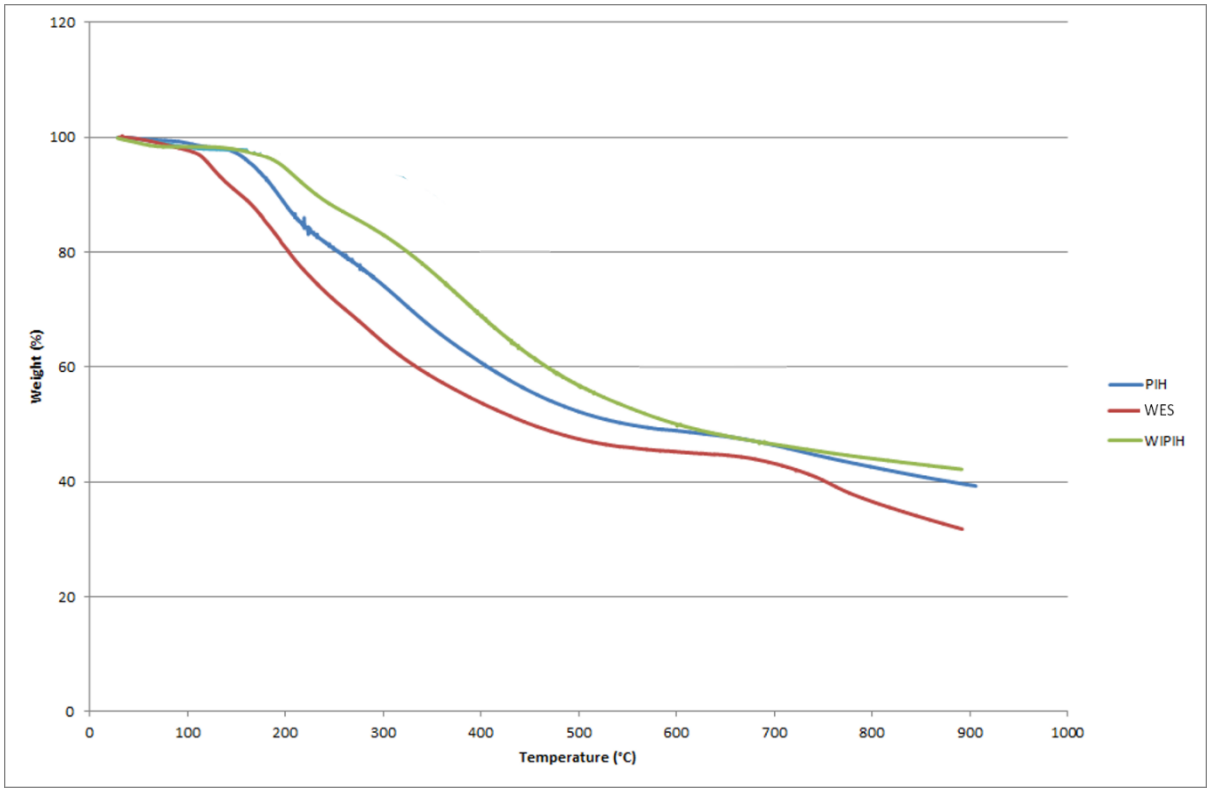


Figure S 4. TGA traces of PIH, WES and WIPIH measured under N_2.

V. SEM images of WIPIH

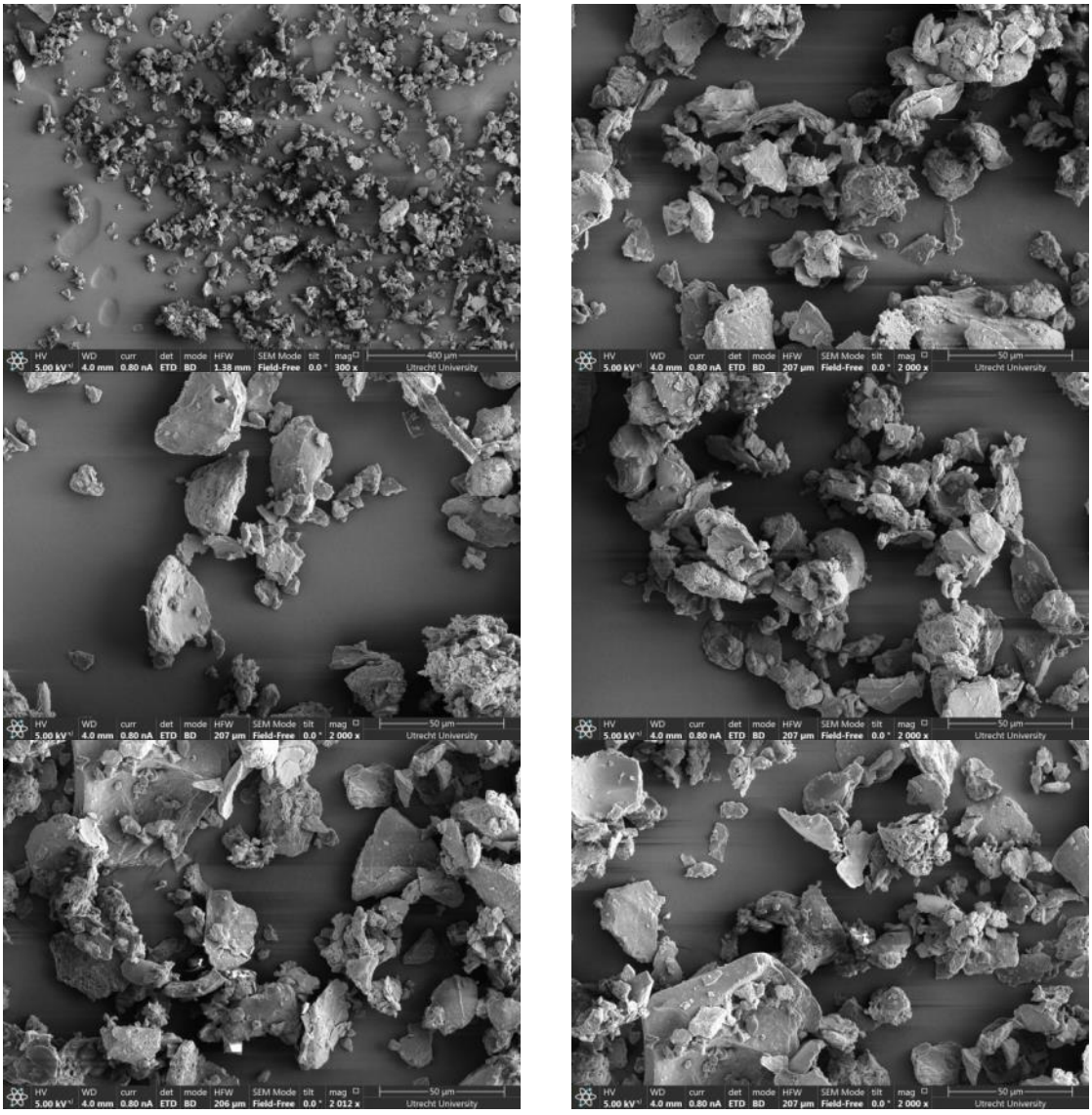


Figure S 5. SEM images of WIPIH.

VI. Gel Permeation Chromatography (GPC)

Table S 8. GPC methods.

Method	Eluent	Column	Column specification	Standard	Detection method
A	THF	PL-gel Mixed-E columns	Up to 25,000 g mol^{-1}	PS	UV
B	HFIP	PSS PFG	~250-2,500,000 g mol^{-1}	PMMA	Refractive Index

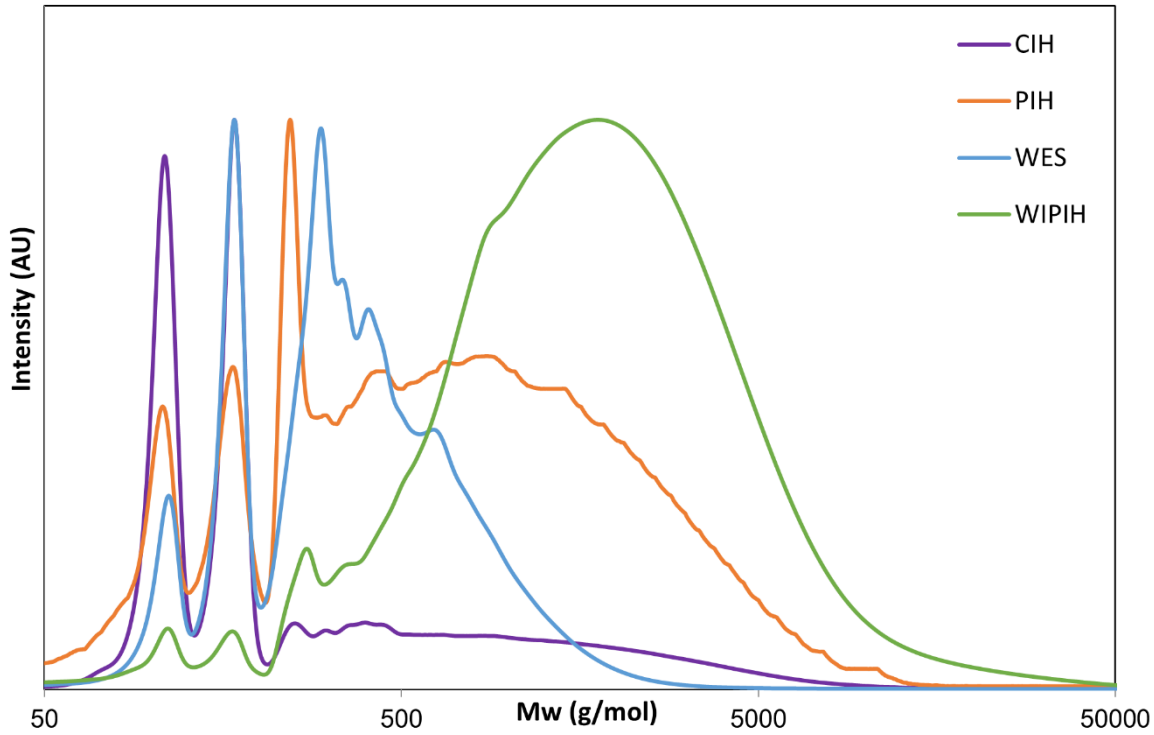


Figure S 6. GPC curves of CIH, PIH, WES and WIPIH measured with method A (duplicate of Figure 3).

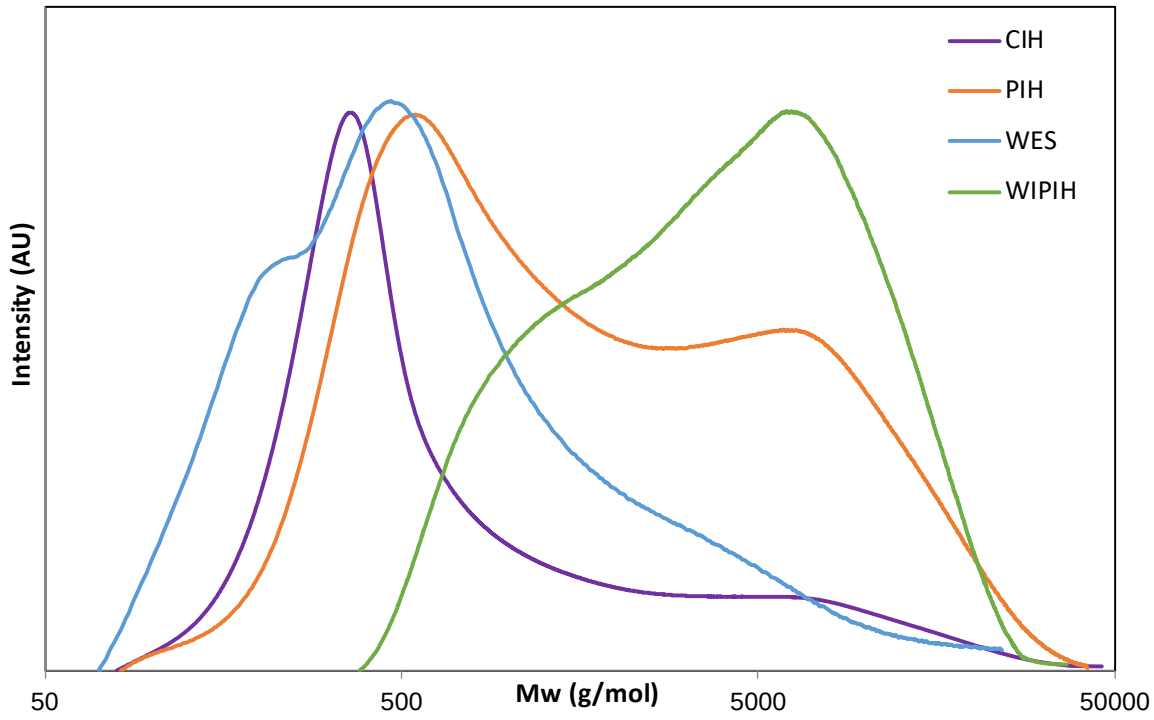


Figure S 7. GPC curves of PIH, WES and WIPIH measured method B.

Table S 9. GPC data of CIH, PIH, WIPIH and WES.

Method	Humin sample	Mn (g/mol)	Mw (g/mol)	PDI
A (THF)	CIH	220	680	3.1
	PIH	400	1100	2.8
	WIPIH	970	2250	2.3
	WES	300	470	1.5
B (HFIP)	CIH	200	950	4.8
	PIH	960	3200	3.3
	WIPIH	840	5900	7.0
	WES	420	1400	3.3

VII. FT-IR

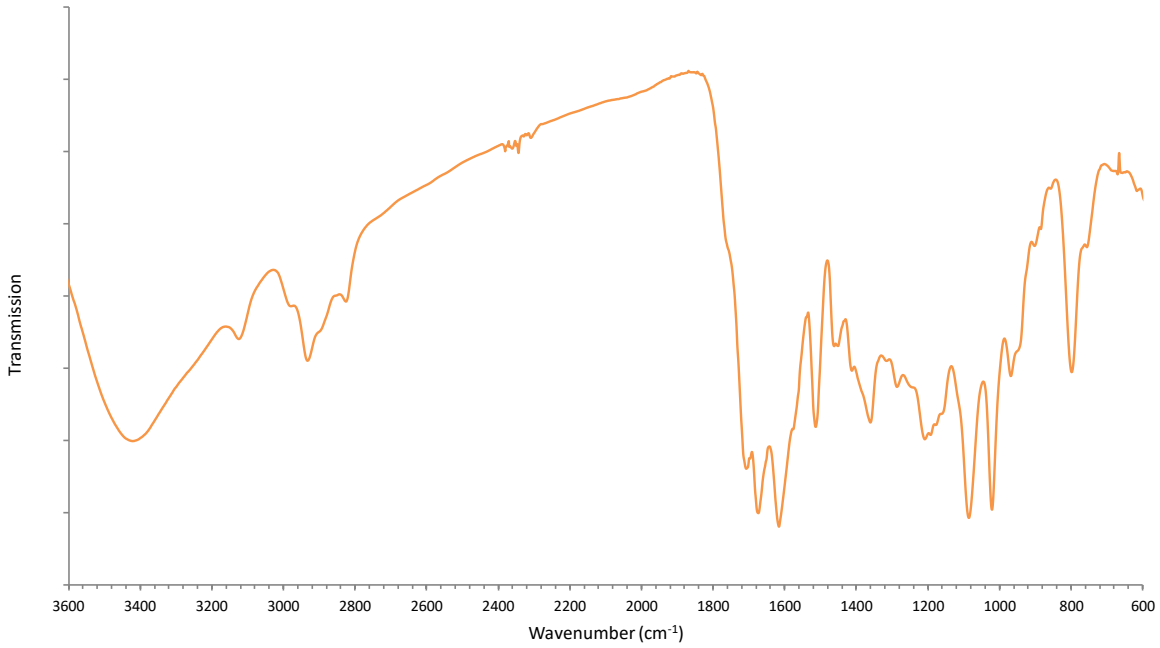


Figure S 8. FT-IR spectrum of WIPIH.

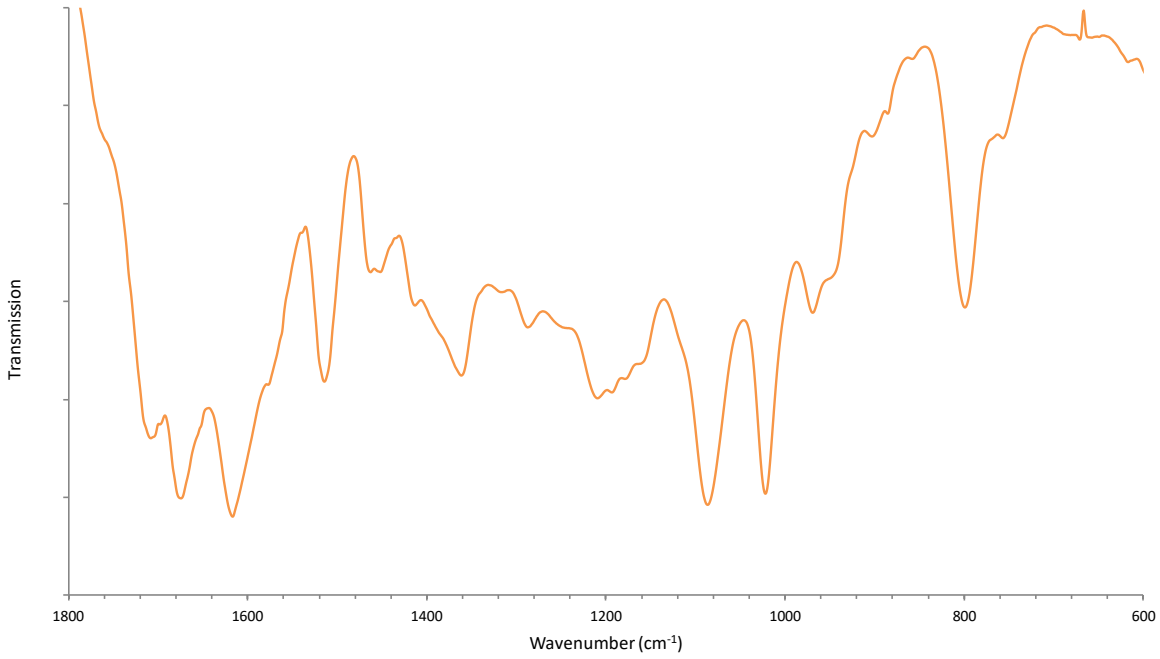


Figure S 9. Zoomed FT-IR spectrum of WIPIH.

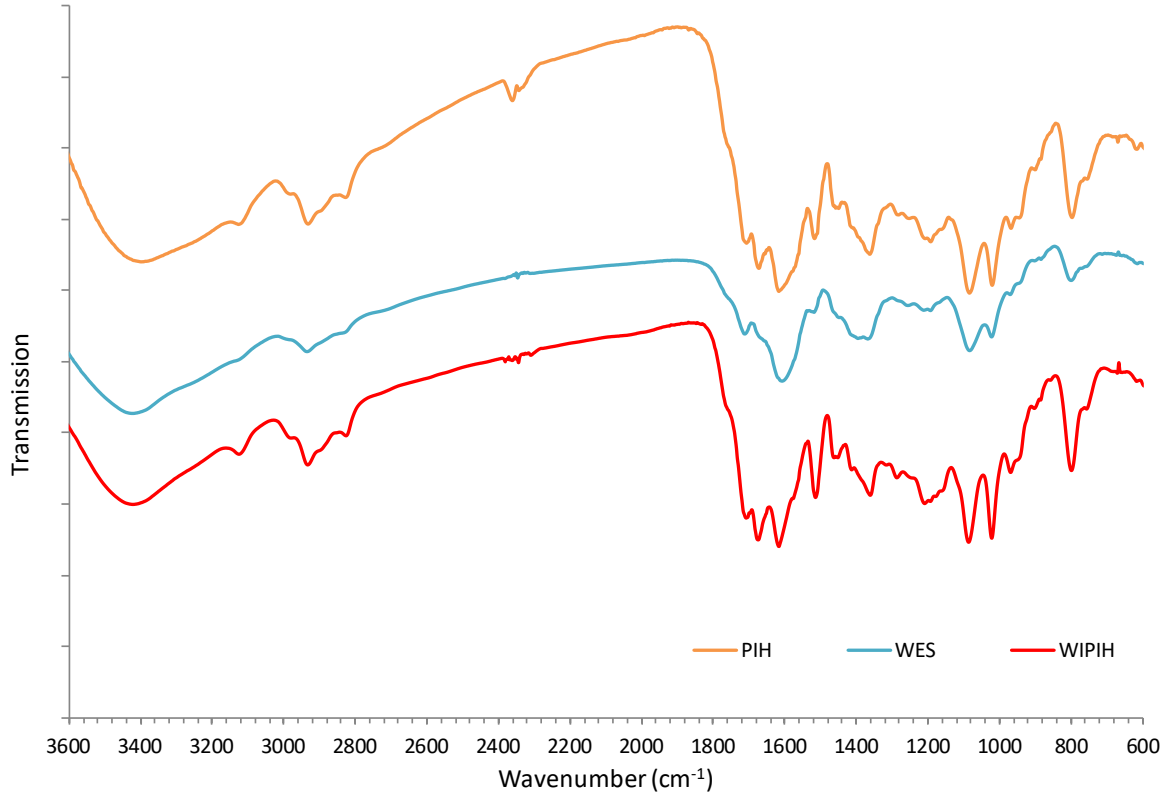


Figure S 10. Comparison of the FT-IR spectra of PIH, WIPIH and WES.

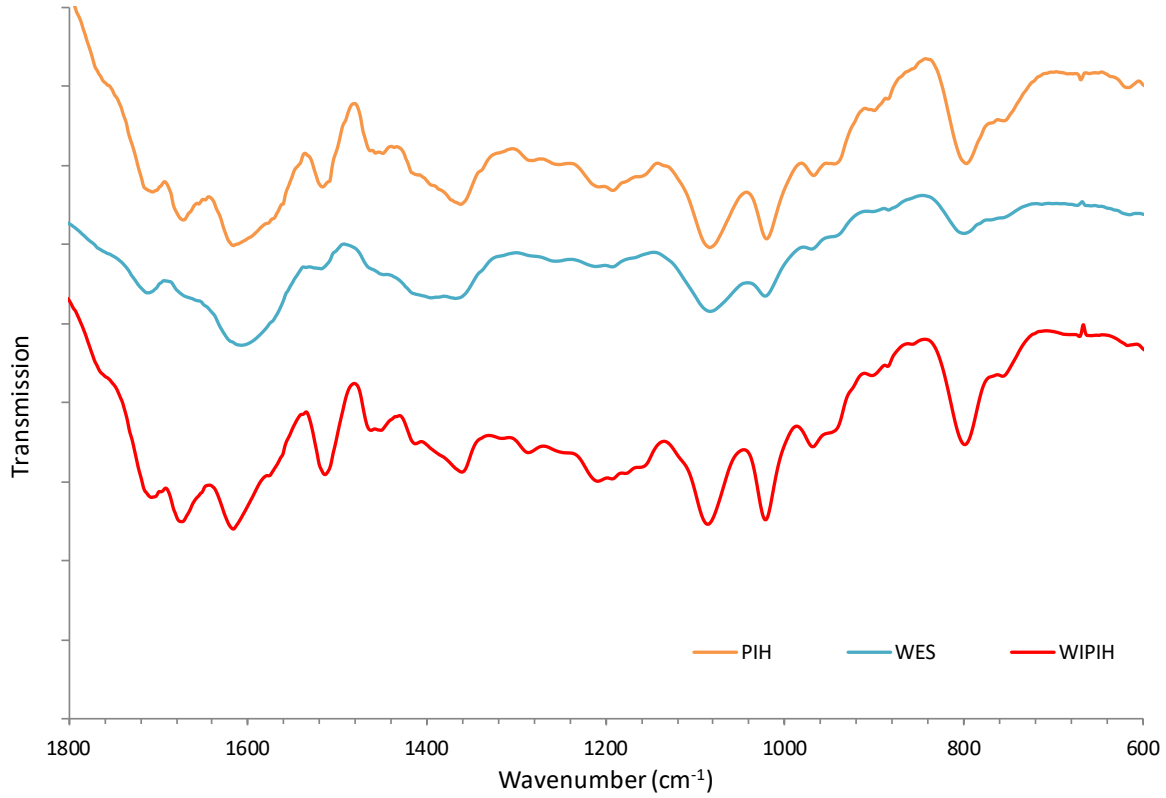


Figure S 11. Zoomed comparison of the FT-IR spectra of PIH, WIPIH and WES.

Table S 10. Assignments of bands observed in the FT-IR spectra.

Wavenumber (cm^{-1})	**Assignments**
3400	O-H stretching of alcohols
3130	C-H of furanic rings
2983	C-H stretching in methyl groups
1745	C=O stretching of esters
1710	C=O stretching of ketones, aldehydes and carboxylic acids
1670	C=O stretch of the HMF-like and MMF-like aldehyde group
1600	ring stretching vibration of furan C=C
1518	ring stretching vibration of furan C=C
1365	Bending of methyl group next to carbonyl
1200	sp^2-O
1020 and 1082	C-O related vibrations in alcohols, sugars, ethers, and acetals
1050	sp^3-C-O in furan rings and O-CH_3
795	C-H out of plane of furan rings
760	C-H out of plane of furan end groups substituted with a C=O in position 2

VIII. Solid-state NMR

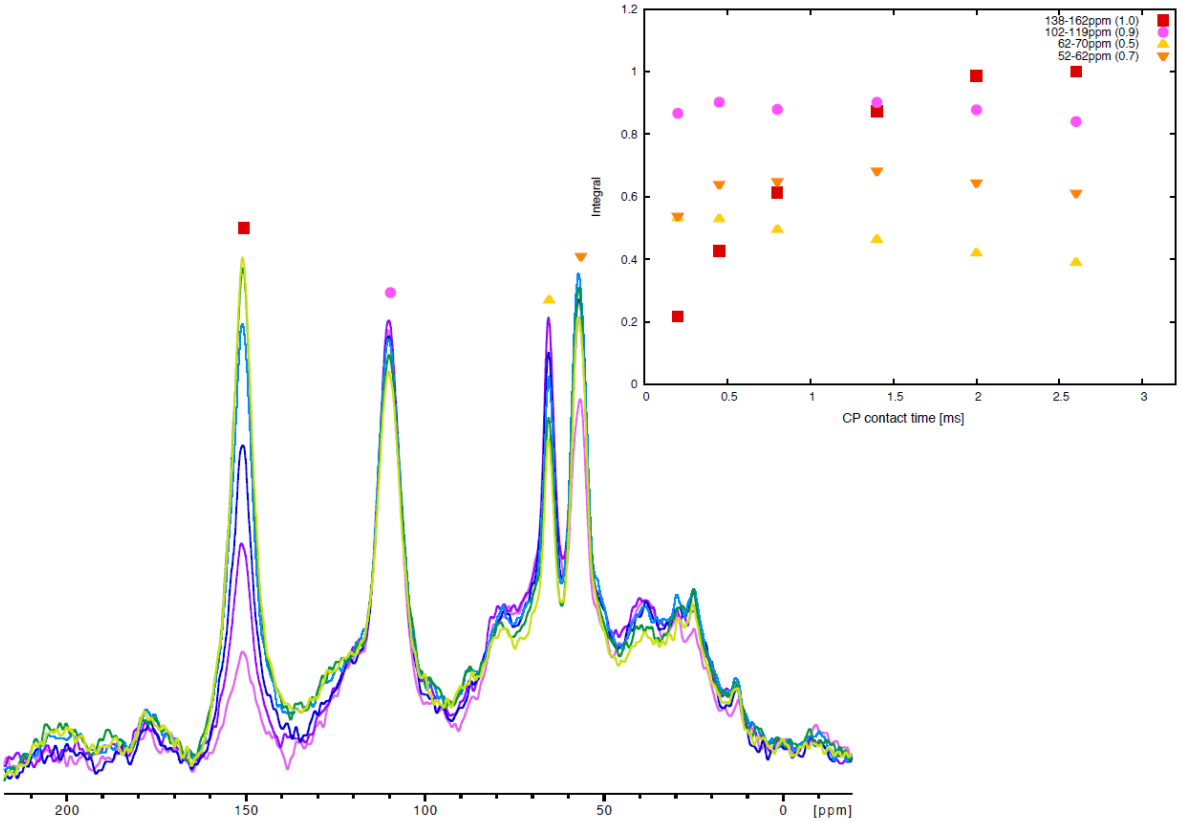


Figure S 12. 1D ^{1}H-^{13}C CP MAS NMR spectra of PIH recorded with different CP contact times.

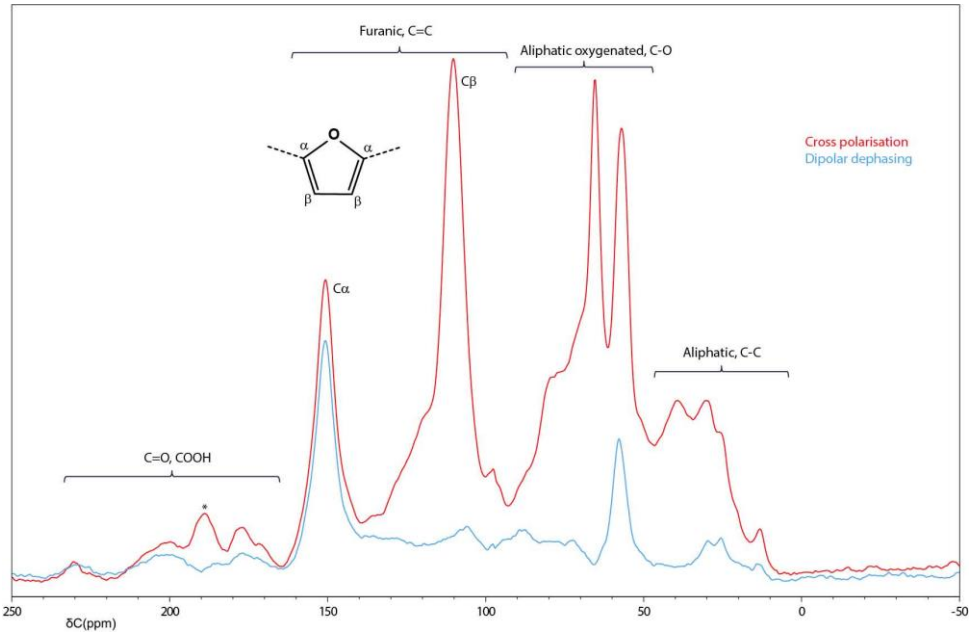


Figure S 13. 1D ss-NMR spectra of PIH recorded by CP and (dipolardephasing) DD experiments.

Table S 11. Assignments of peaks observed on the ss-NMR spectra.

^{13}C Chemical shift (ppm)	**Assignments**
15	aliphatic (-CH_3)
25	aliphatic (-CH_2-)
30	aliphatic (-C-H, C)
39	aliphatic (-C-H, C)
57	methoxy in MMF-derived structures
65	secondary alcohol in HMF/Hydroxyacetyl Furan-derived structures
78	secondary alcohols in carbohydrate derivatives
110	β carbon furan
120	linked β carbon
130	furan conjugated
150	linked α carbon
177	acid, ester
200	ketone, aldehyde

IX. Solution-state NMR

a) ^{1}H NMR

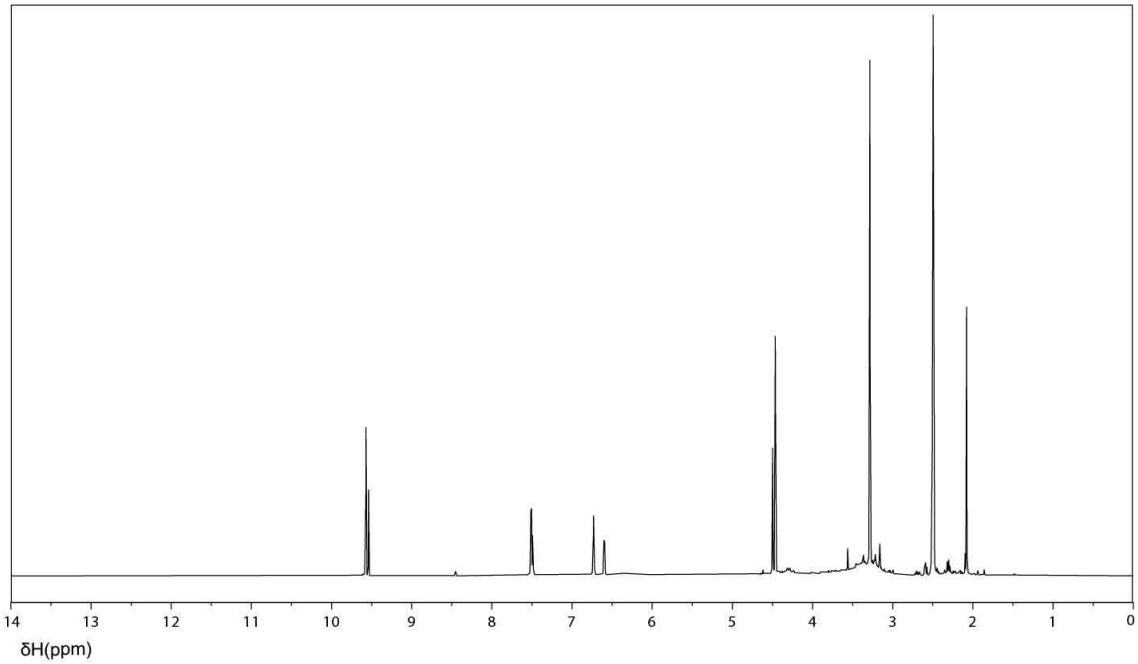


Figure S 14. ^{1}H NMR spectrum of CIH in DMSO-d_6.

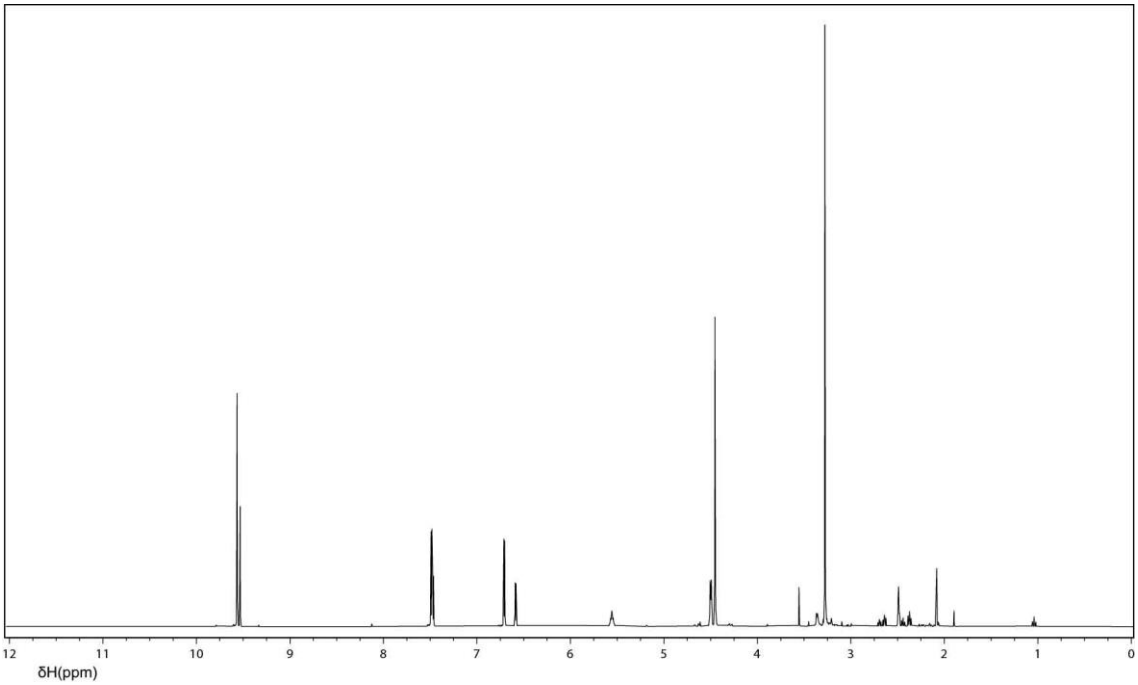


Figure S 15. ^{1}H NMR spectrum of DEE in DMSO-d_6.

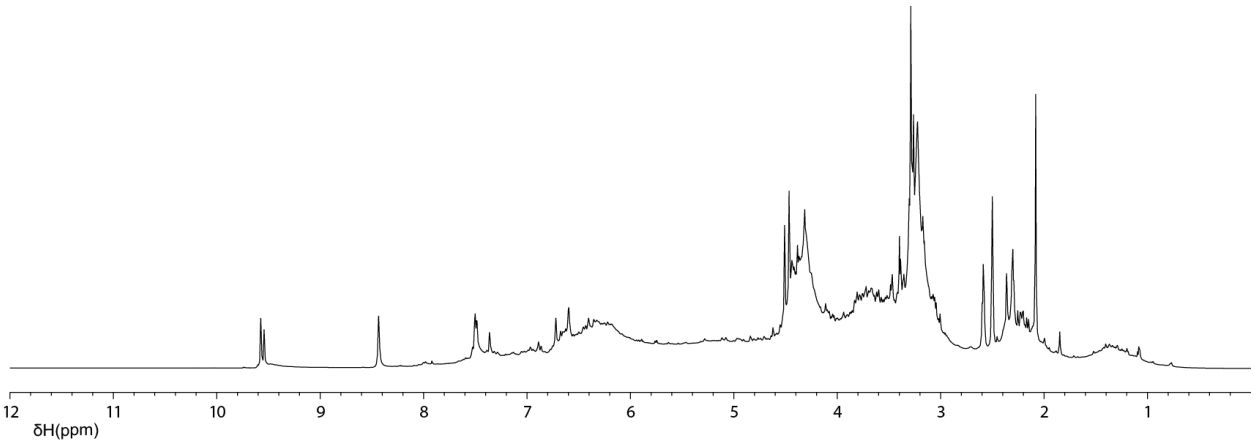


Figure S 16. ^{1}H NMR spectrum of PIH in DMSO-d_6.

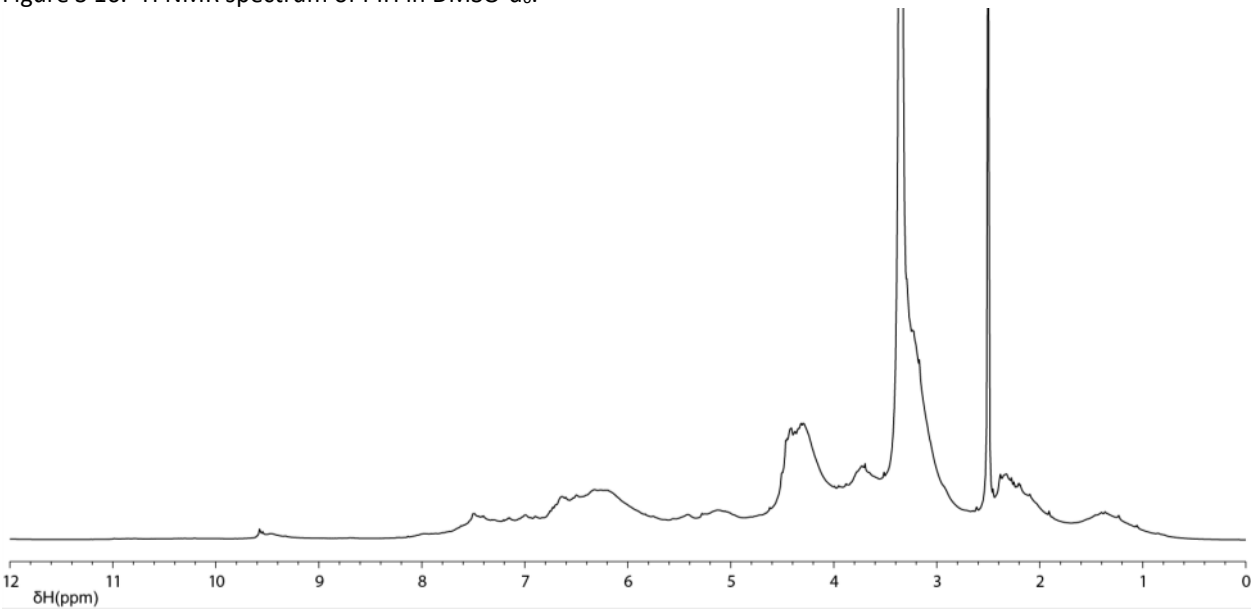


Figure S 17. ^{1}H NMR spectrum of WIPIH in DMSO-d_6.

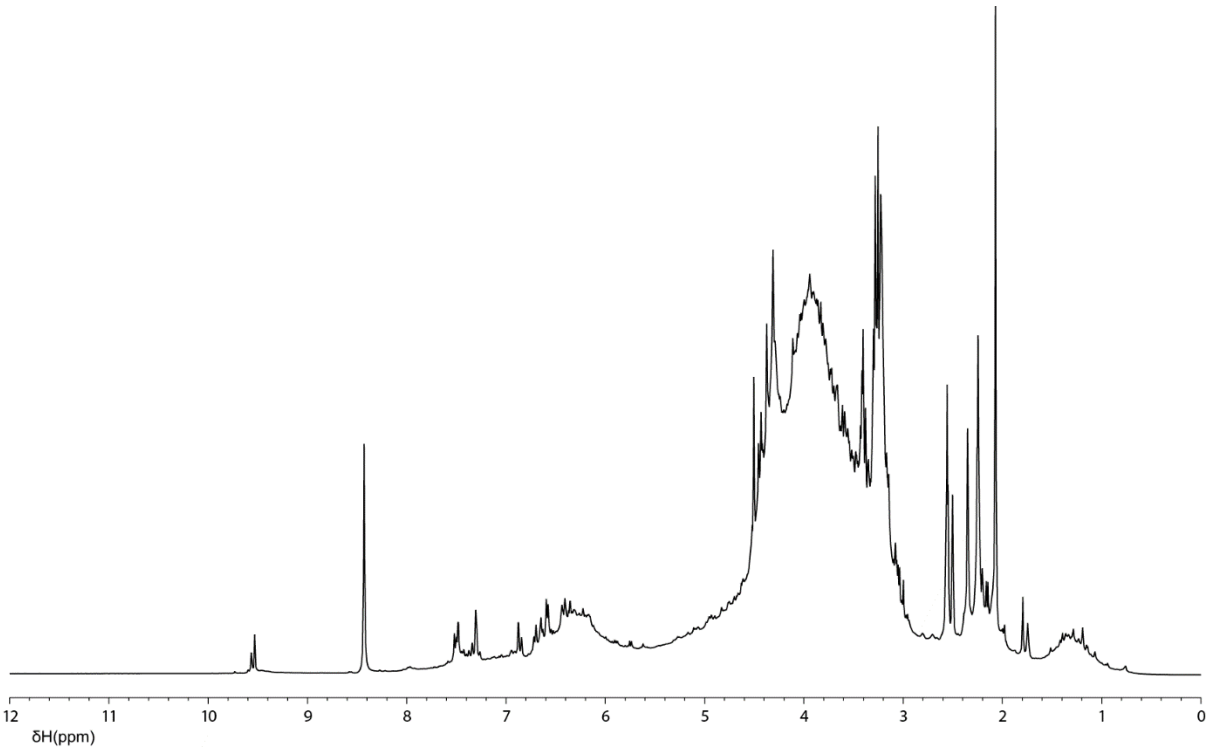


Figure S 18. ^{1}H NMR spectrum of WES in DMSO-d_6.

b) ^{13}C NMR

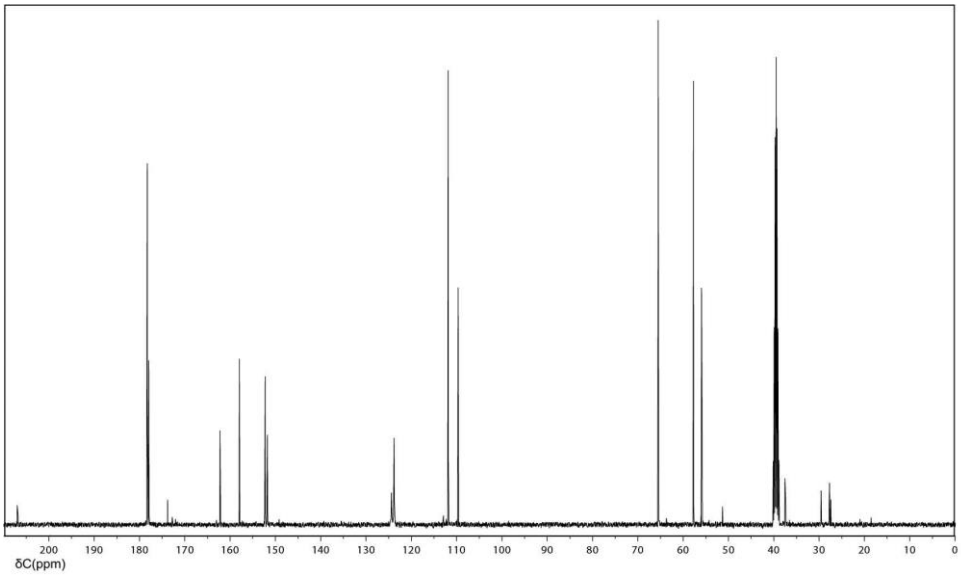


Figure S 19. ^{13}C NMR spectrum of DEE in DMSO-d_6.

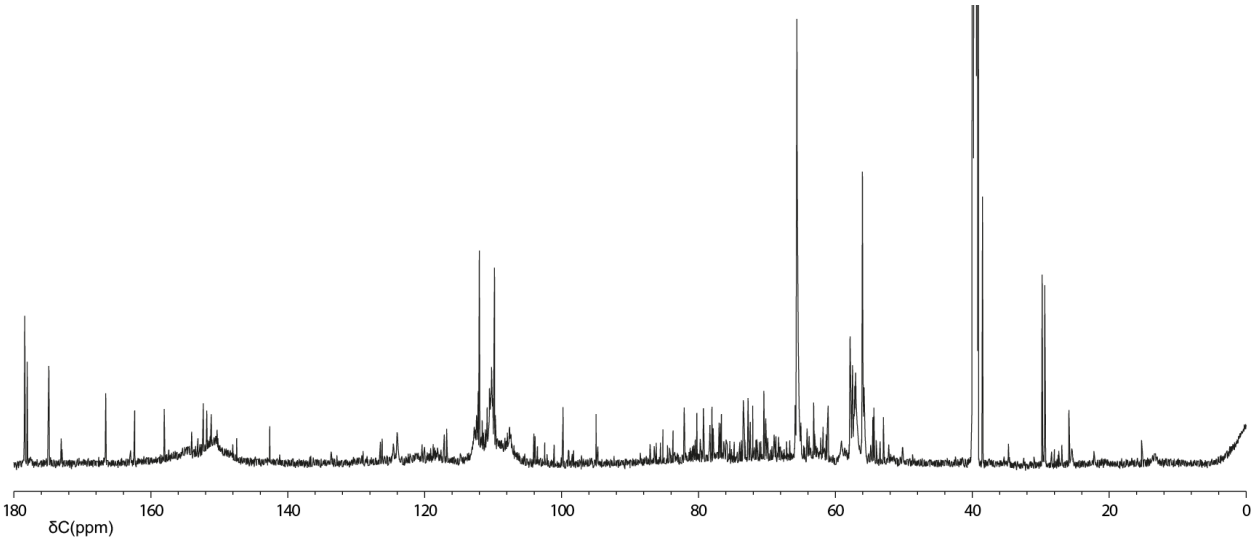


Figure S 20. ^{13}C NMR spectrum of PIH in DMSO-d_6.

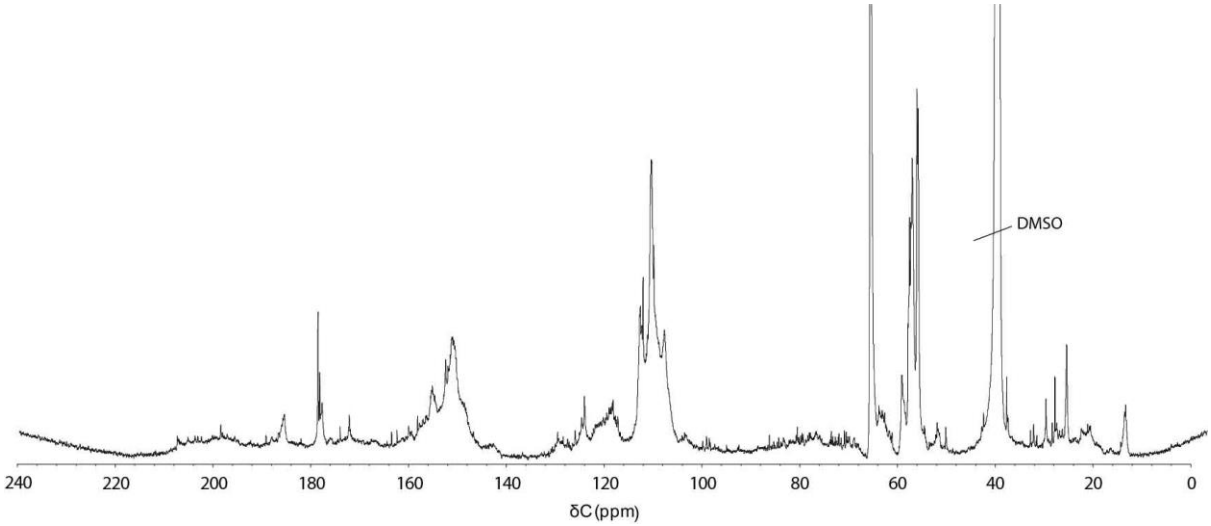


Figure S 21. ^{13}C NMR spectrum of WIPIH in DMSO-d_6.

Table S 12. Assignments of peaks observed on the ^{13}C NMR spectra in DMSO-d_6.

^{13}C Chemical shift (ppm)	**Assignments**
13	methyl
26	methyl groups vicinal to a ketone
163	conjugated carboxylic acid in furoic units
172	conjugated carboxylic acid in allyl units
174	methyl ester in aliphatic units
178	aldehyde in furfural units, carboxylic acid in aliphatic units
186	ketone conjugated on both sides
197	conjugated methyl ketone
205	aliphatic ketones

c) ^{13}C DEPT 135

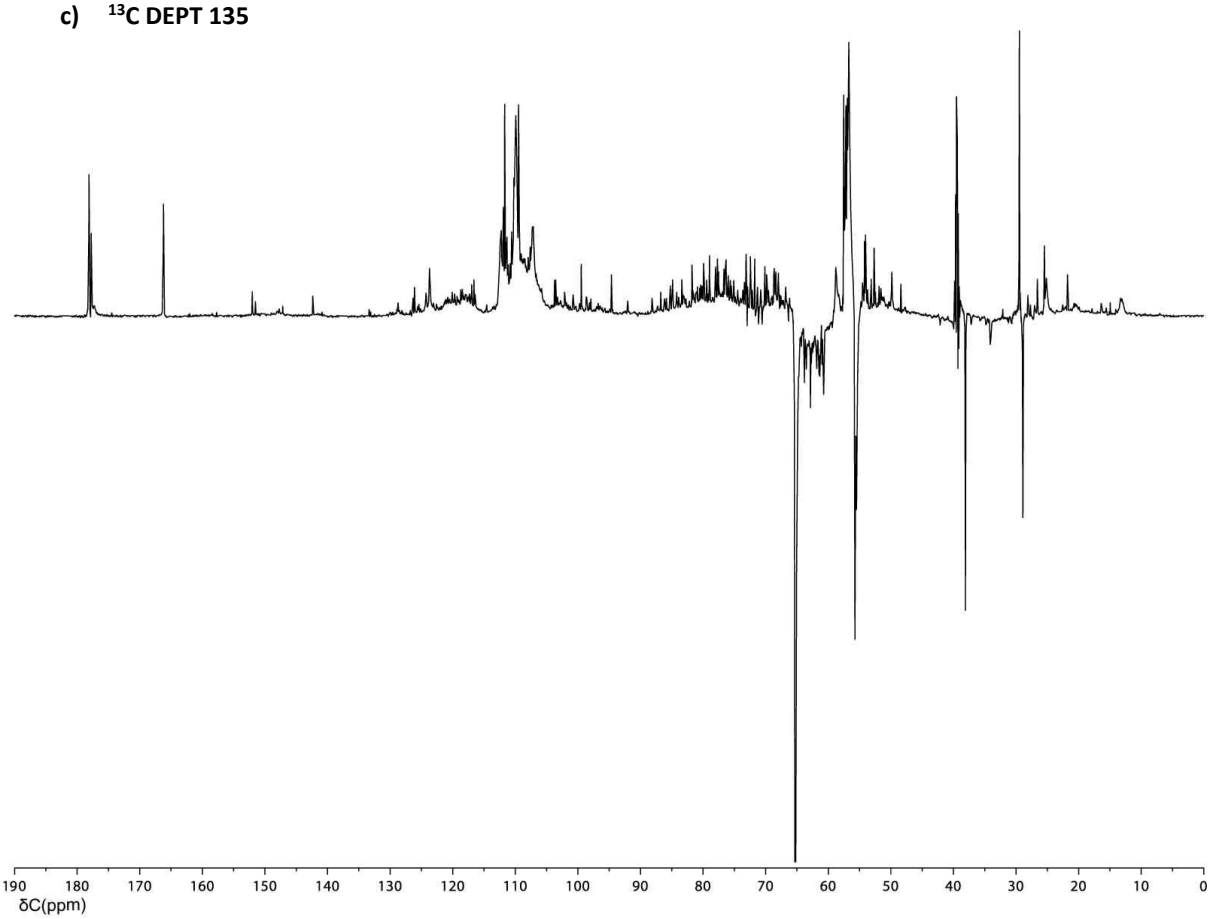


Figure S 22. ^{13}C DEPT NMR spectrum of PIH in DMSO-d_6; negative peaks correspond to CH_2 carbons.

d) HSQC NMR

Table S 13. Assignments of peaks observed on the [^{1}H;^{13}C] HSQC NMR spectra in DMSO-d_6.

^{1}H/^{13}C Chemical shift (ppm)	**Assignments**
1.4/22	methyl next to secondary alcohols (**E**)
2.1/30	methyl next to ketone (**C**,**D**)
2.2/13	methylated furanic units in position 5 (**A**)
2.3/25	methyl next to conjugated ketone (**B**)
2.4/28	CH_2 next to carboxylic acid or methyl ester in LA/ML like structures (**F**)
2.6/37	CH_2 next to ketone (**D**)
3.1/57	methoxy in ethers
3.2/55	methoxy in methylated sugars
3.4/63	CH_2-OH in sugars
3.5/52	methoxy in esters
4.3/65	OCH_2 in MMF-like units (**G**)
4.4/56	OCH_2 in HMF-like units (**H**)
5.2/57.7	OCH_2 in furyl methoxy units (**I**)
6.0/105.5	position 3 and 4 of furanic units connected with aliphatic links (**J**)
6.2/108	position 4 in furanic units conjugated with an allyl ketone (**L**) and furanic units substituted with a carbonyl (**K**)
6.6/112	position 4 in furanic end-groups **M** and **N**
6.9/116	position 3 in furanic units conjugated with an allyl ketone (**L**)
6.9/119	β of in furanic units conjugated with an allyl ketone (**L**)
7.2/114	position 1 in furan benzoquinone-like moieties (**P**)
7.5/129	α of in furanic units conjugated with an allyl ketone (**L**)
7.6/143	position 5 in furanic end-groups **M**
8.0/148	position 5 in furanic end-groups **N**
8.4/167	CH in formic acid
9.5/178	aldehyde in 2,5 substituted furfuraldehyde units (**O**)

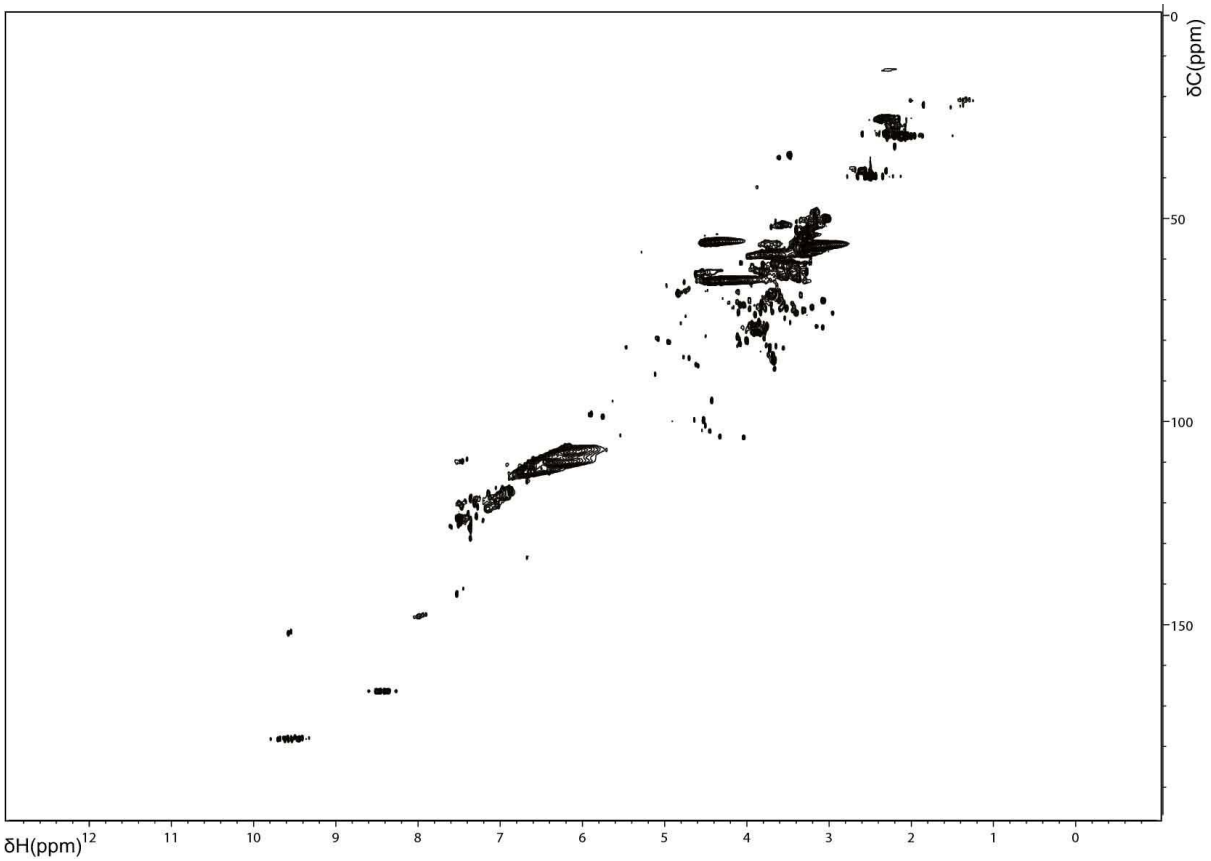


Figure S 23. Full spectrum of the [^{1}H; ^{13}C] HSQC spectrum of PIH in DMSO-d_6.

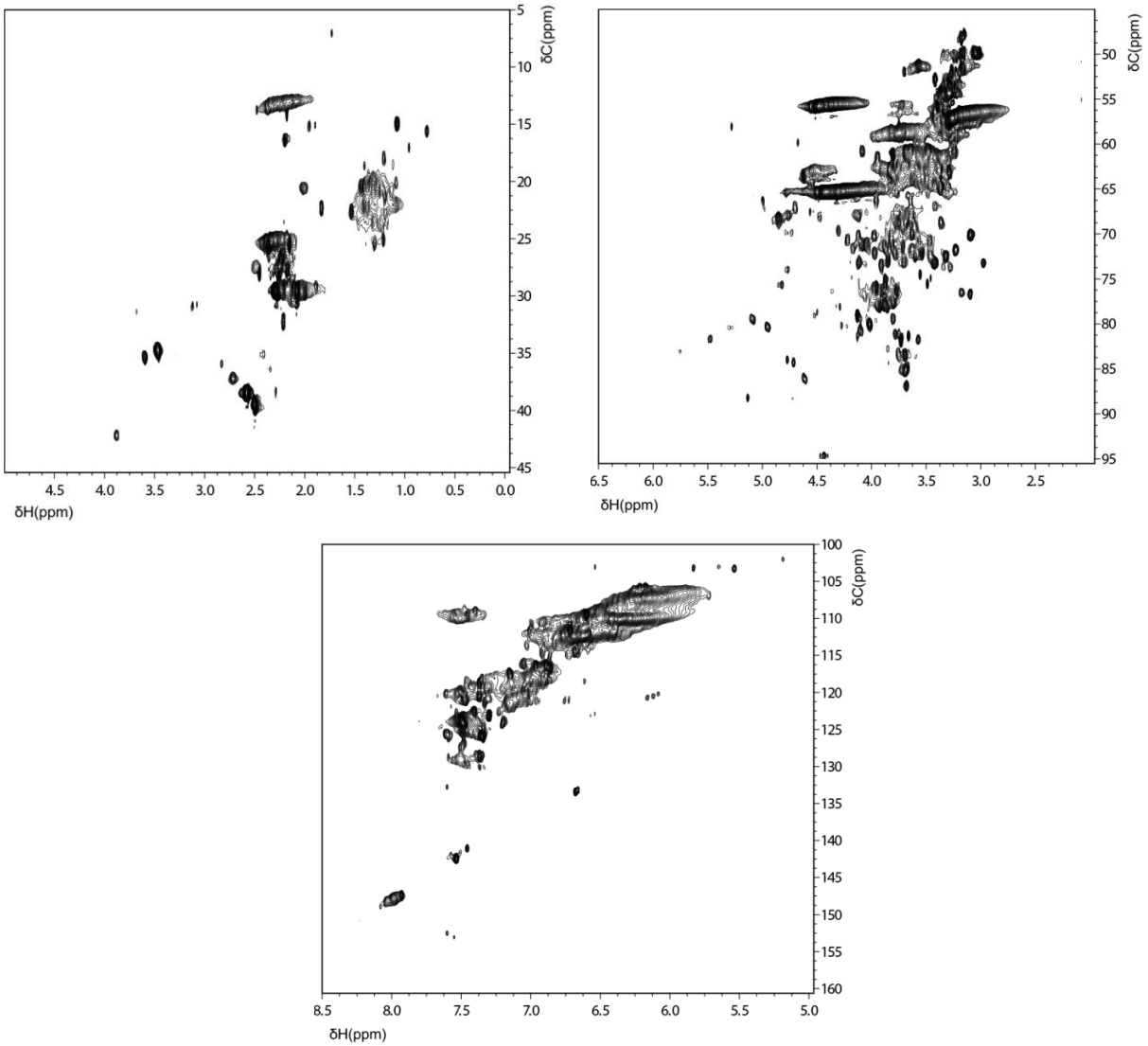


Figure S 24. Spectra of the zoomed regions of the [^{1}H; ^{13}C] HSQC spectrum of PIH in DMSO-d_6.

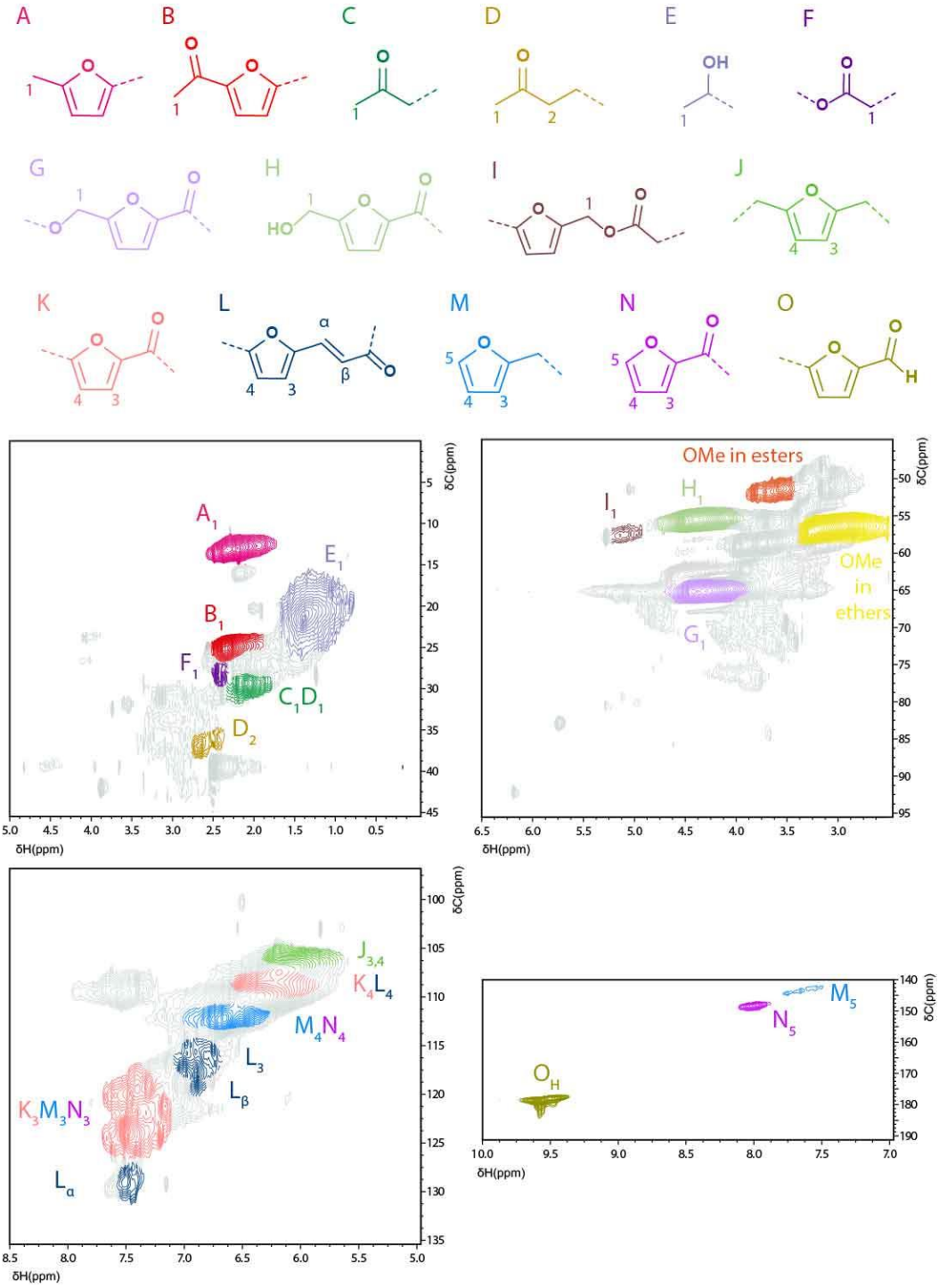


Figure S 25. Assignments of the zoomed regions of the [^{1}H; ^{13}C] HSQC spectrum of WIPIH in DMSO-d_6.

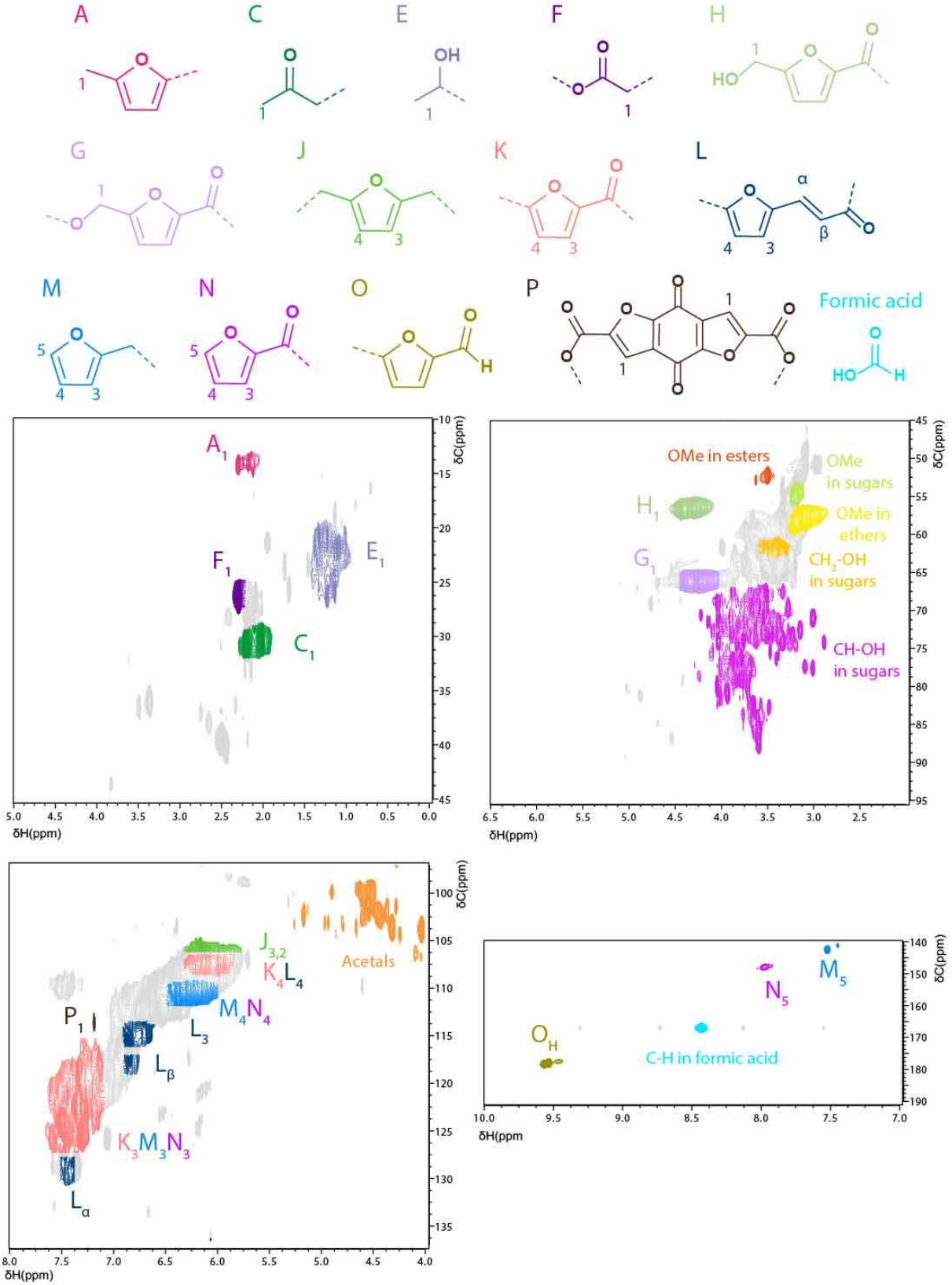
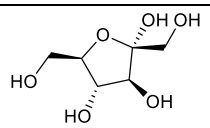
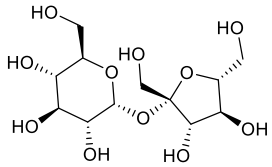
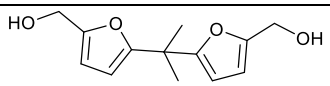
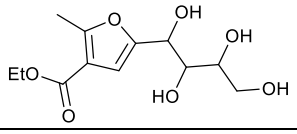
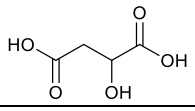
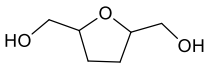
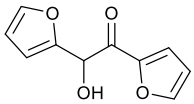
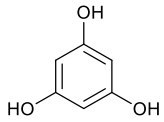
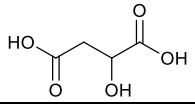
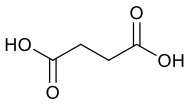
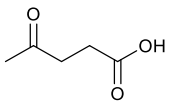
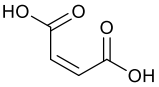
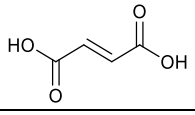
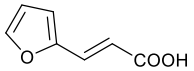
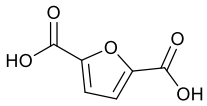
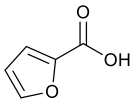
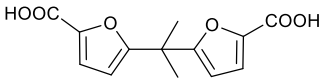
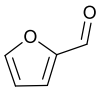
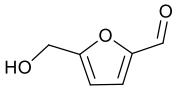


Figure S 26. Assignments of the zoomed regions of the [^{1}H; ^{13}C] HSQC NMR of WES in DMSO-d_6.

e) ^{31}P NMR

Table S14. Library of model compounds derivatised with 2-chloro-4,4,5,5-tetramethyldioxaphospholane

Compound	Structure	δ ^{31}P NMR (ppm)
Aliphatic OH		
D-Fructose		149.09-145.55
D-Sucrose		148.30-145.47
(propane-2,2-diylbis(furan-5,2-diyl))dimethanol		147.66, 147.54
ethyl 2-methyl-5-(1,2,3,4-tetrahydroxybutyl) furan-3-carboxylate		147.50-146.37
Malic acid		147.41
(tetrahydrofuran-2,5-diyl)dimethanol		147.11-146.98
1,2-di(furan-2-yl)-2-hydroxyethan-1-one		146.79,146.49, 143.68
Phenolic OH		
Phloroglucinol		137.25, 136.95, 136.64
Aliphatic carboxylic acid		
Formic acid		136.27
Malic acid		134.72, 134.59
Succinic acid		134.30
Levulinic acid		134.22
Acetic acid		134.15
Unsaturated carboxylic acid		
Maleic acid		134.92
Fumaric acid		134.91
3-(furan-2-yl)acrylic acid		134.94

Compound	**Structure**	**δ ^{31}P NMR (ppm)**
Furanic carboxylic acid		
2,5-FDCA		134.29, 134.70
2-Furoic acid		134.37
5,5'-(propane-2,2-diyl)bis(furan-2-carboxylic acid)		133.91
Aldehyde		
Furfural		149.52,149.38, 146.76
5-(hydroxymethyl)furan-2-carbaldehyde		149.40,149.26, 148.12, 146.67

f) ^{19}F NMR

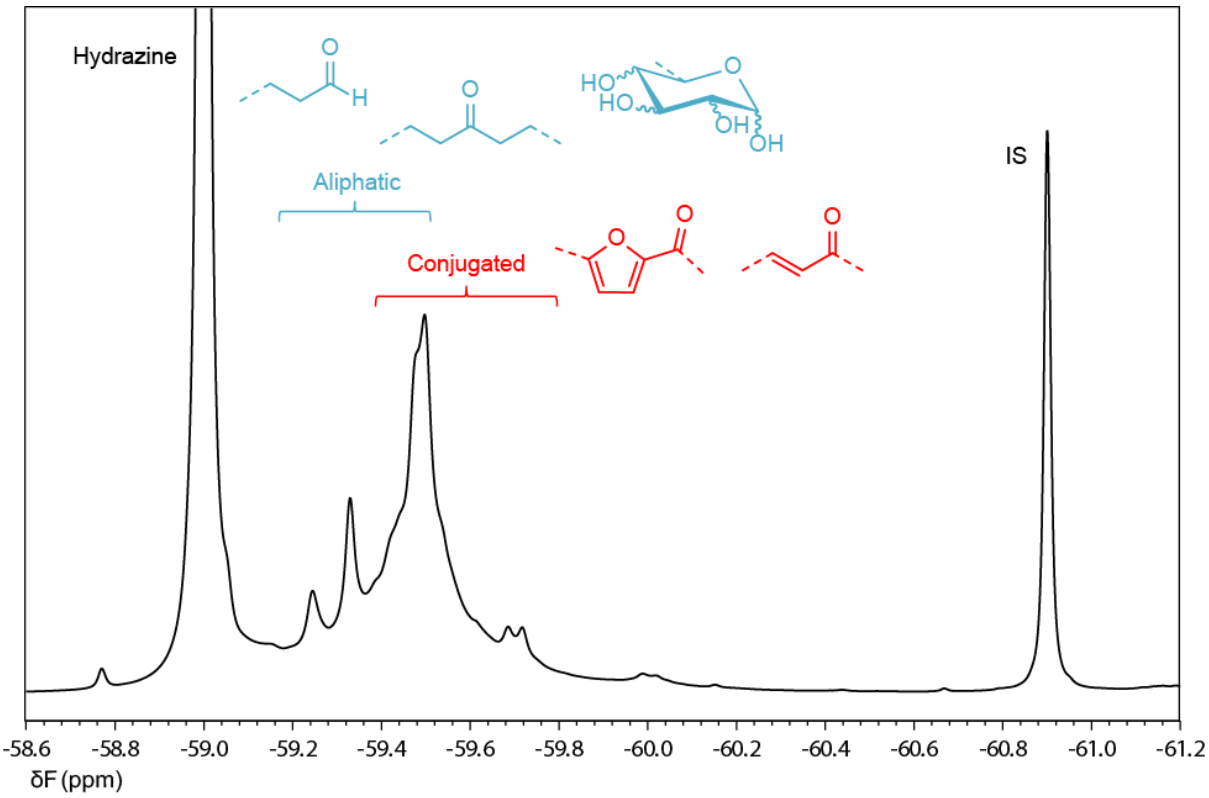


Figure S 27. ^{19}F NMR spectrum of PIH derivatised with 4-(trifluoromethyl)phenylhydrazine.

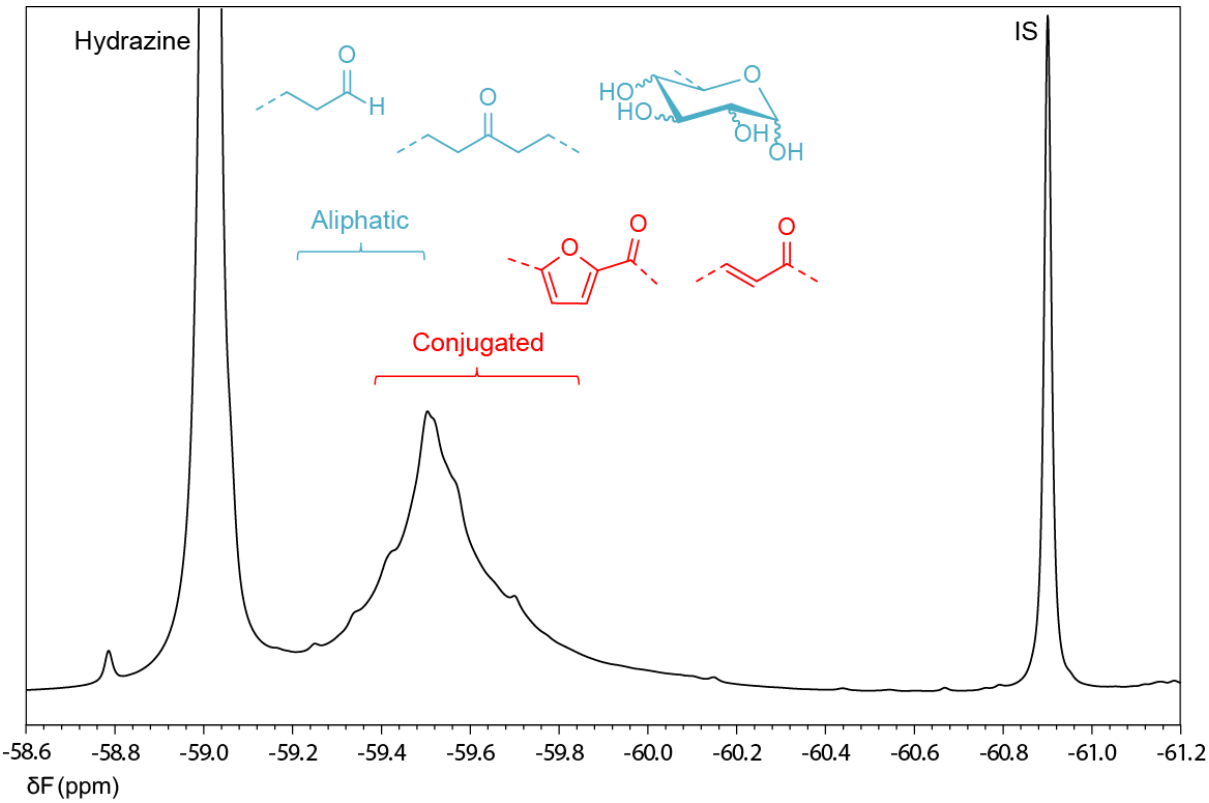


Figure S 28. ^{19}F NMR spectrum of WIPIH derivatised with 4-(trifluoromethyl)phenylhydrazine.

C. Additional data on ATPIH

I. Thermogravimetric Analysis (TGA)

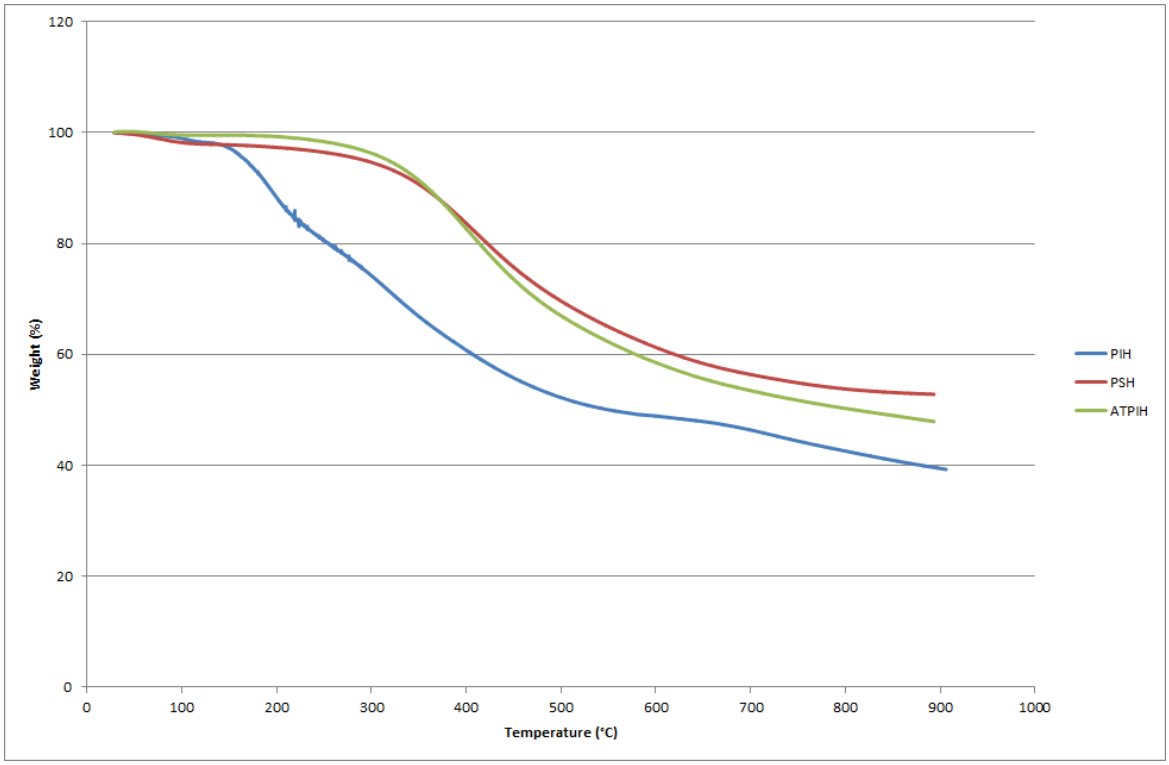


Figure S 29. TGA traces of PIH, PSH and ATPIH measured under N_2.
PSH: purified synthetic humin produced from glucose at 180 °C for 6 h.[1]

II. SEM images of ATPIH

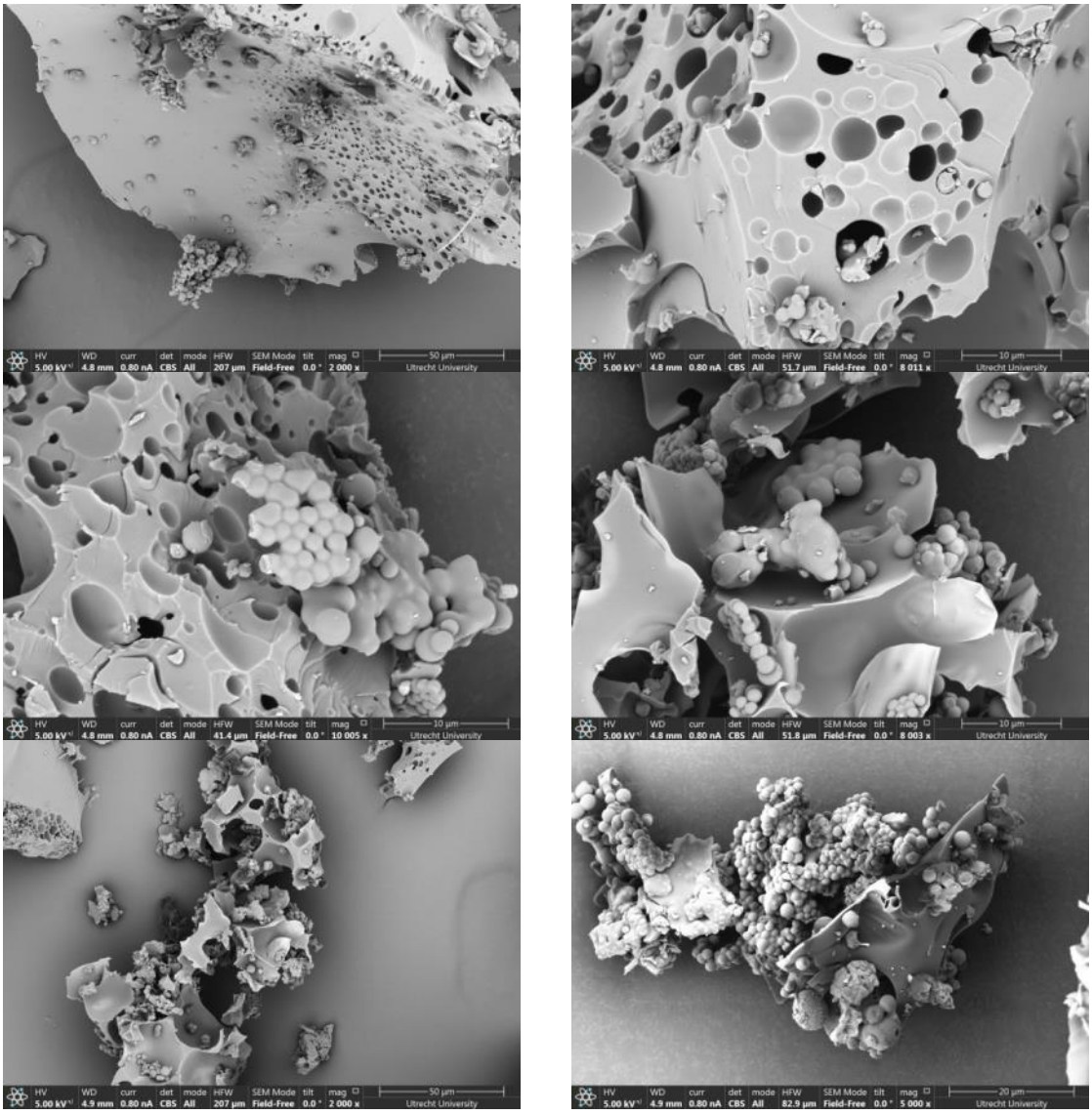


Figure S 30. SEM images of ATPIH.

III. FT-IR spectrum of the ATPIH

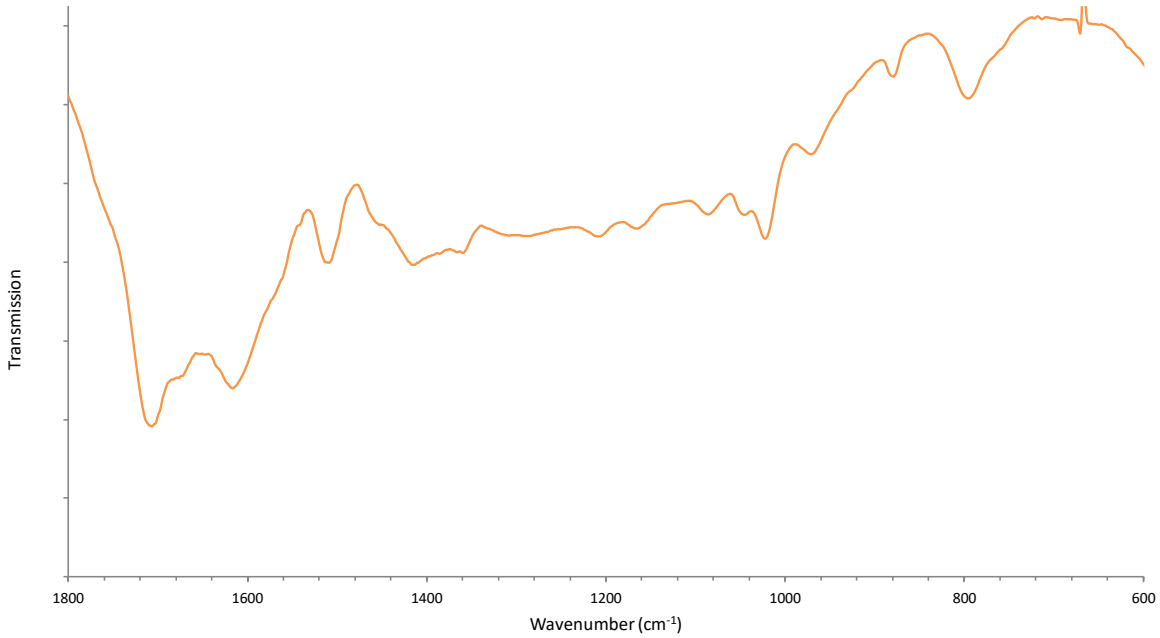


Figure S 31. FT-IR spectrum of ATPIH.

IV. Solid-state NMR spectrum of ATPIH

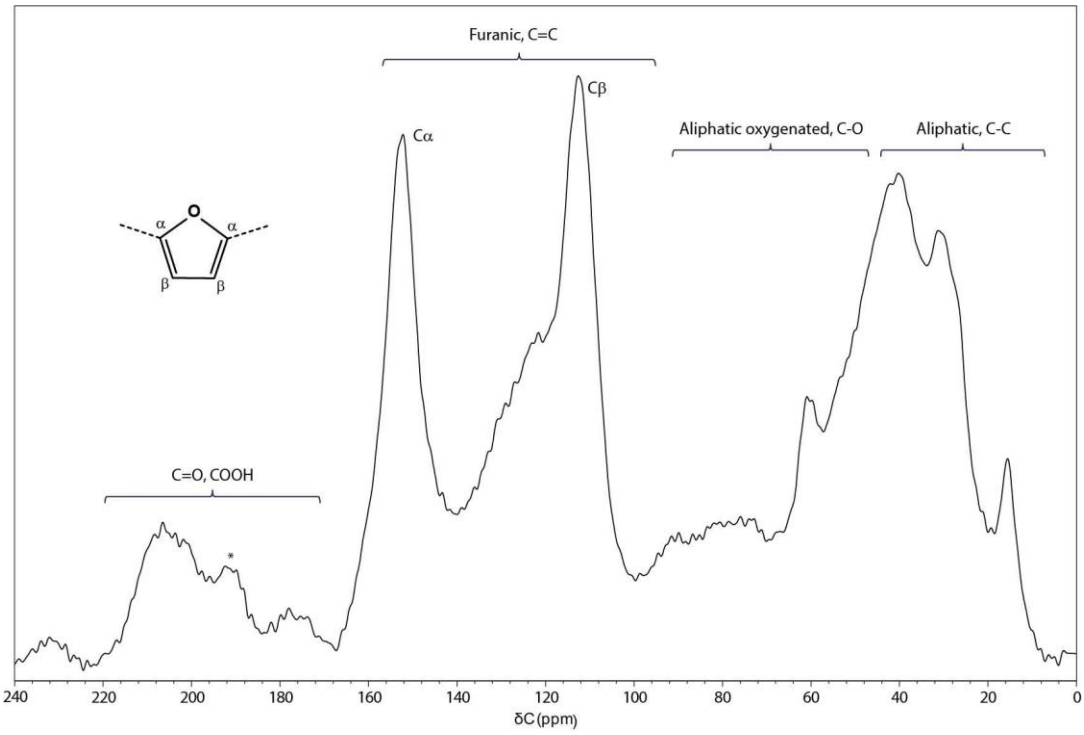


Figure S 32. ^{13}C CP spectrum of ATPIH.

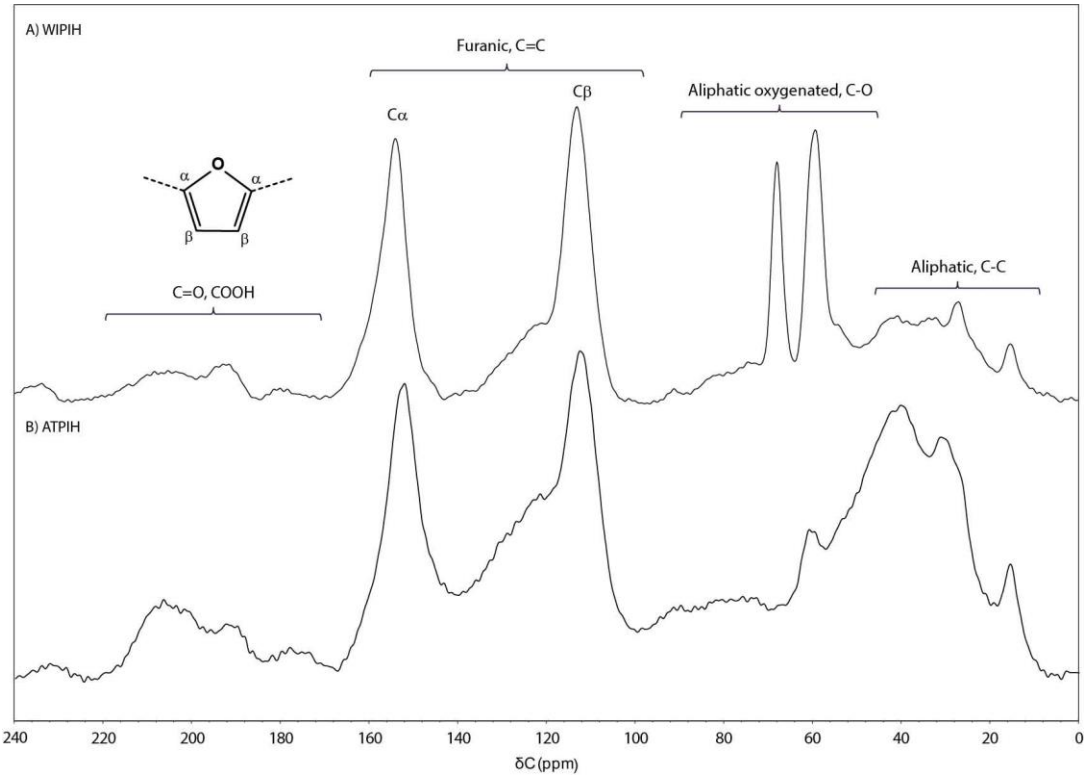
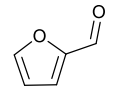
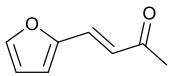
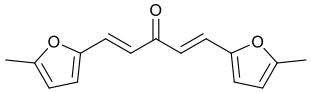
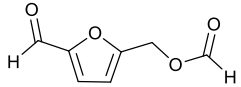
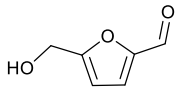
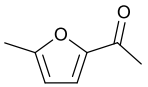
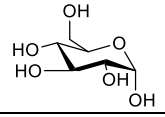
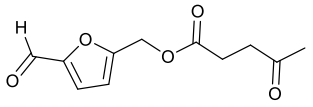
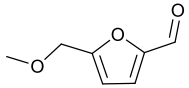
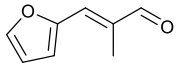
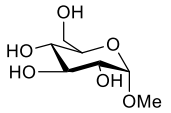
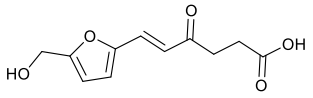
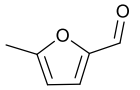
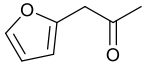
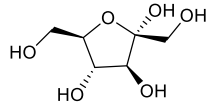
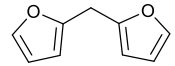
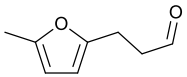
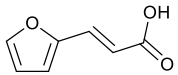
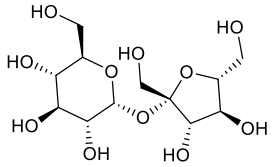
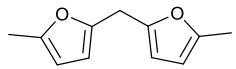
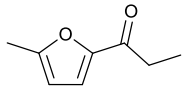
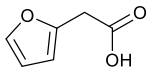
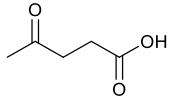
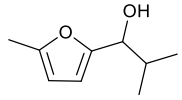
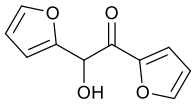
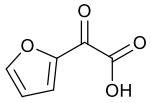
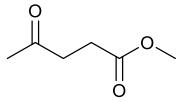
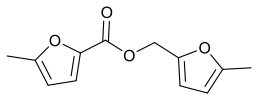
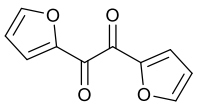
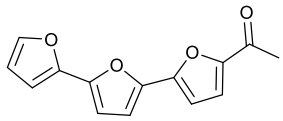
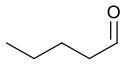
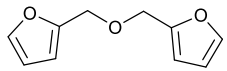
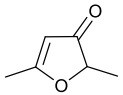
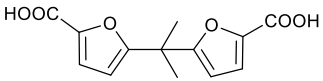
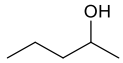
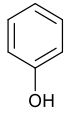
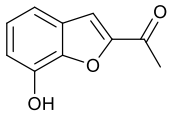
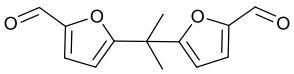
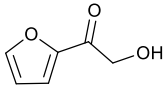
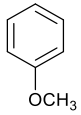
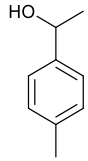
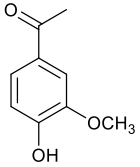
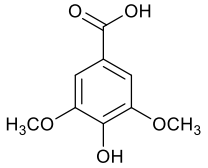
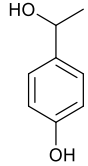
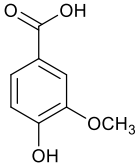
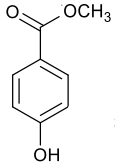
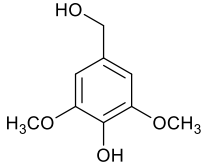
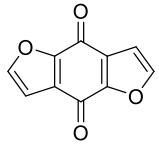
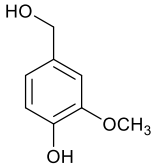
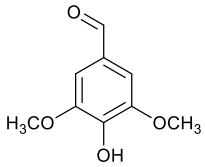
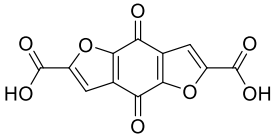
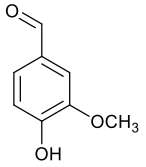
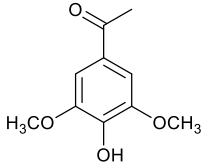


Figure S 33.Overlay of the ^{13}C CP spectra of WIPIH (A) and ATPIH (B).

D. Database of NMR spectra

Compound	Page nb	Compound	Page nb	Compound	Page nb	Compound	Page nb
	50		61		72		78
	51		62		73		78
	52		63		74		79
	53		64		75		79
	54		65		76		79
	55		66		77		79
	56		67		77		80
	57		68		77		80
	58		69		77		80
	59		70		78		80
	60		71		78		

Compound	Page nb	Compound	Page nb	Compound	Page nb
	81		82		83
	81		82		83
	81		82		84
	81		83		84
	82		83		

Furan-2-carbaldehyde

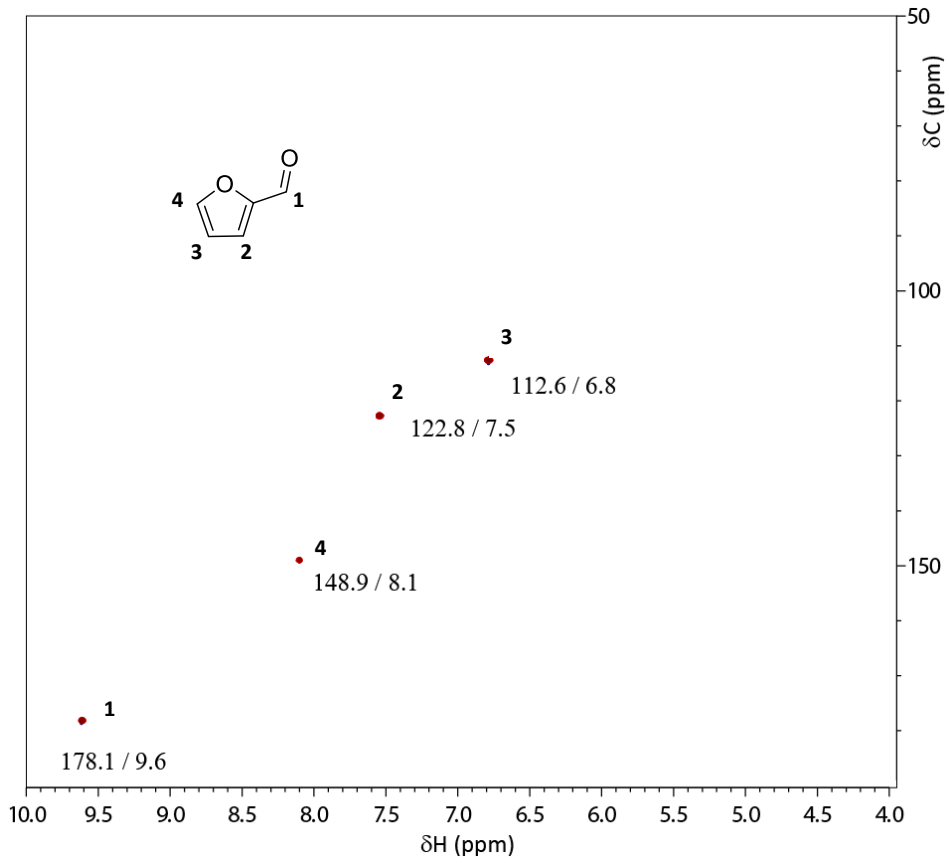


Figure S 34. [^{1}H; ^{13}C] HSQC spectrum of furan-2-carbaldehyde in DMSO-d_6.

5-(Hydroxymethyl)furan-2-carbaldehyde

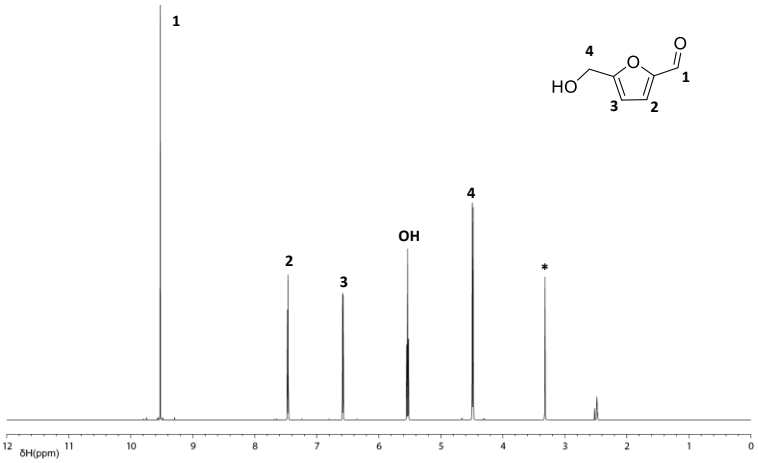


Figure S 35. ^{1}H NMR spectrum of 5-(hydroxymethyl)furan-2-carbaldehyde in DMSO-d$_6$.
* peak due to water contamination in DMSO-d$_6$.

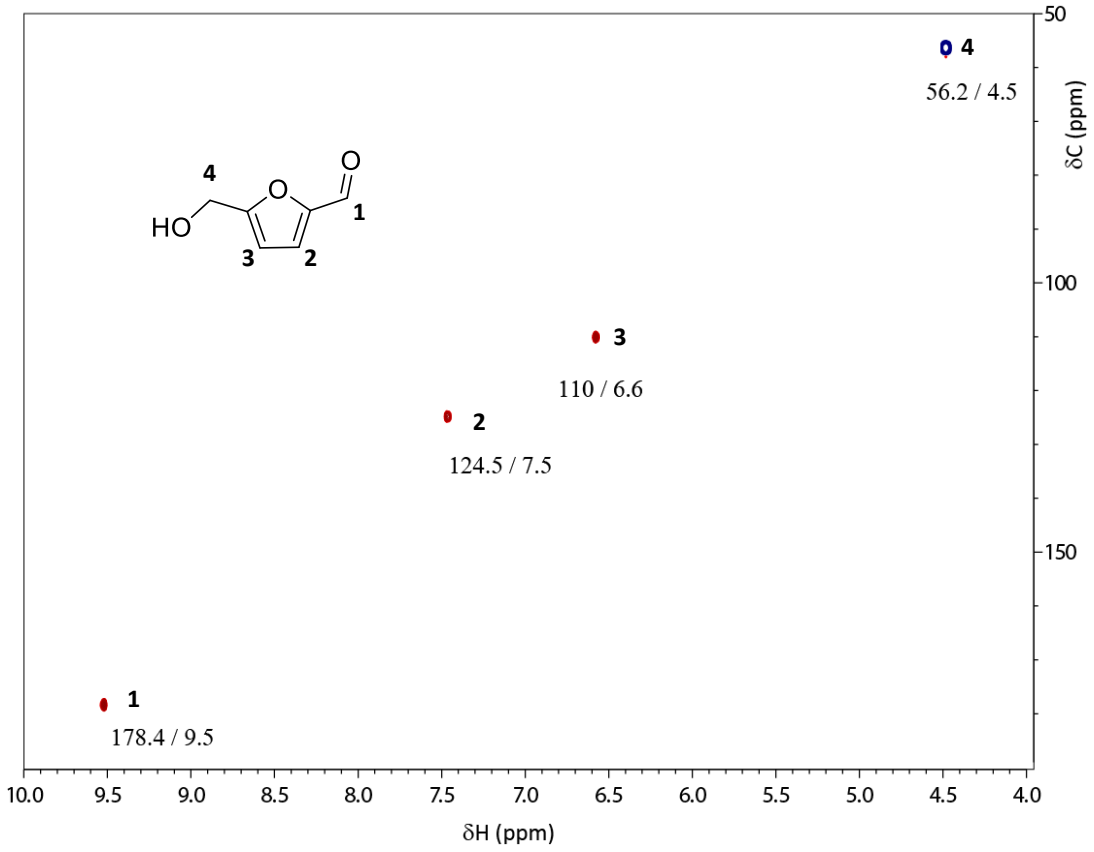


Figure S 36. [^{1}H; ^{13}C] HSQC spectrum of 5-(hydroxymethyl)furan-2-carbaldehyde in DMSO-d$_6$.

5-(Methoxymethyl)furan-2-carbaldehyde

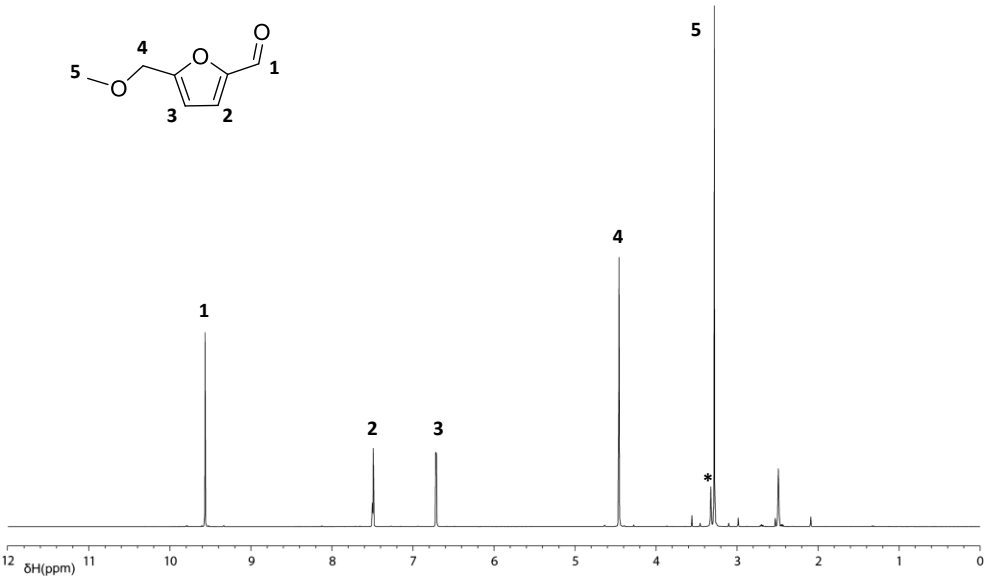


Figure S 37. ^{1}H NMR spectrum of 5-(methoxymethyl)furan-2-carbaldehyde in DMSO-d_6.
* peak due to water contamination in DMSO-d_6.

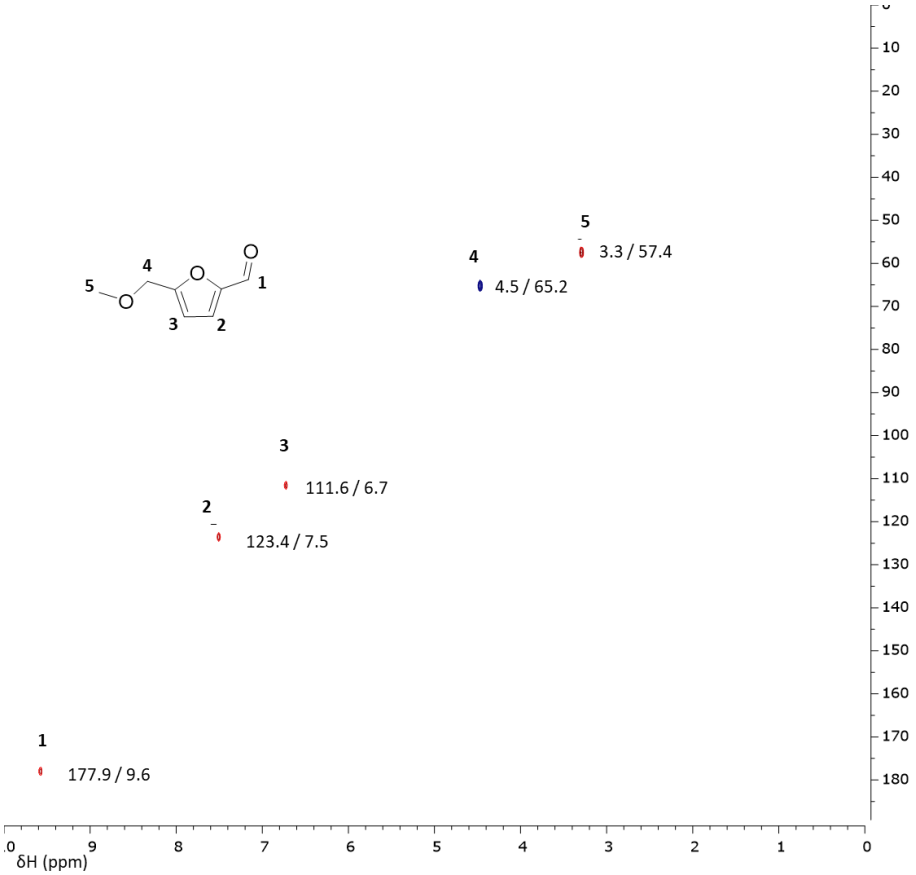


Figure S 38. [^{1}H; ^{13}C] HSQC spectrum of 5-(methoxymethyl)furan-2-carbaldehyde in DMSO-d_6.

5-Methylfuran-2-carbaldehyde

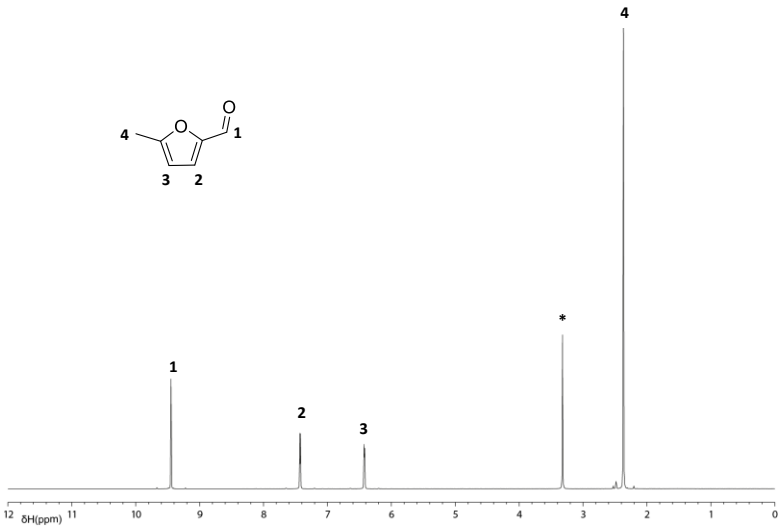


Figure S 39. ^{1}H NMR spectrum of 5-methylfuran-2-carbaldehyde in DMSO-d_6.
* peak due to water contamination in DMSO-d_6.

[^{1}H,^{13}C] HSQC NMR data in $CDCl_3$ taken from reference[4] δ:
1: 9.5/176.8
2: 7.2/124
3: 6.2/109.6
4: 2.4/13.

3-(5-Methylfuran-2-yl)propanal

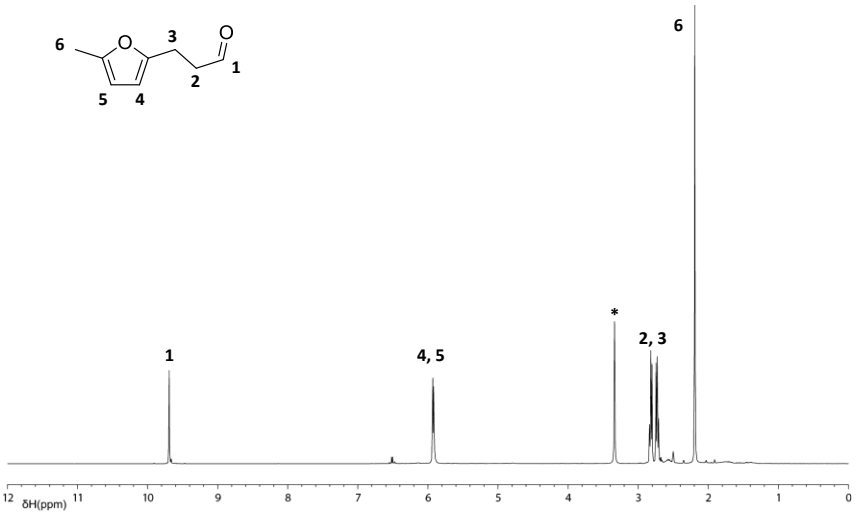


Figure S 40. ^{1}H NMR spectrum of 3-(5-methylfuran-2-yl)propanal in DMSO-d_6.
* peak due to water contamination in DMSO-d_6.

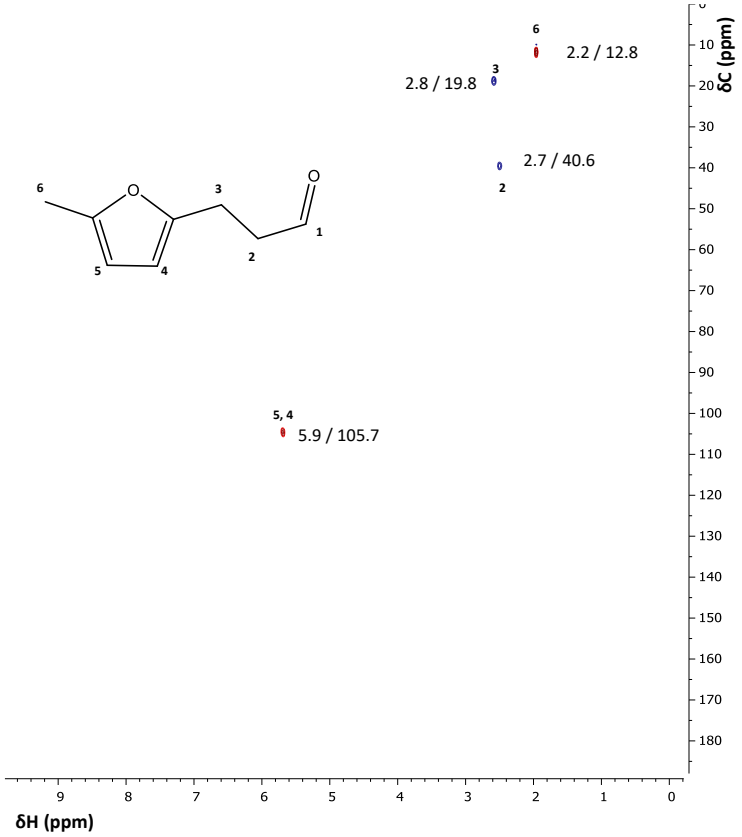


Figure S 41. [^{1}H; ^{13}C] HSQC spectrum of 3-(5-methylfuran-2-yl)propanal in DMSO-d_6.

1-(5-Methylfuran-2-yl)propan-1-one

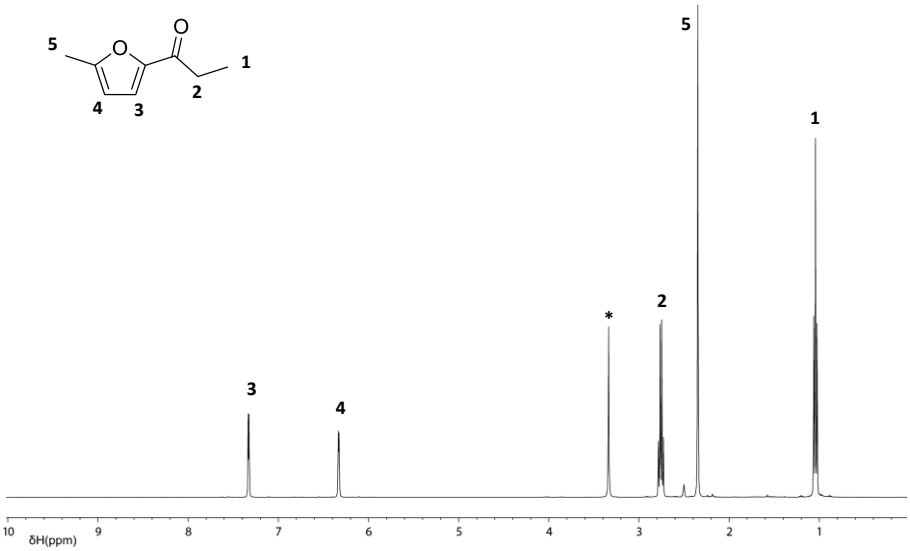


Figure S 42. ^{1}H NMR spectrum of 1-(5-methylfuran-2-yl)propan-1-one in DMSO-d_6.
* peak due to water contamination in DMSO-d_6.

1,2-Di(furan-2-yl)-2-hydroxyethan-1-one

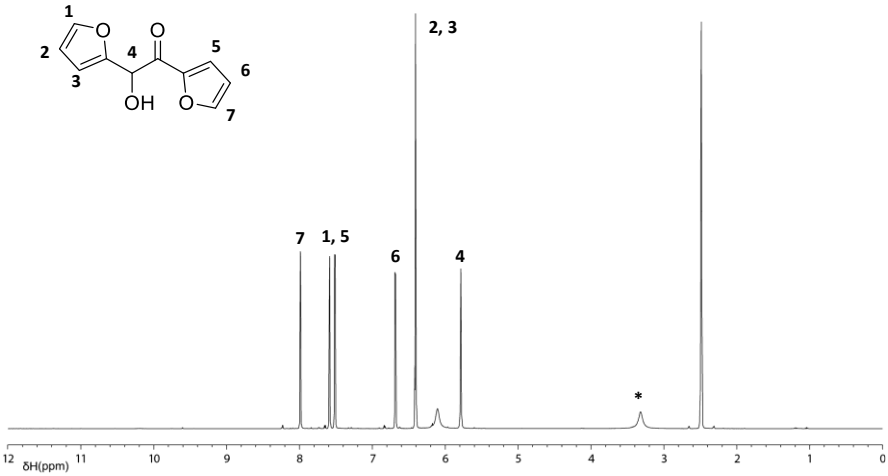


Figure S 43. ^{1}H NMR spectrum of 1,2-di(furan-2-yl)-2-hydroxyethan-1-one in DMSO-d_6.

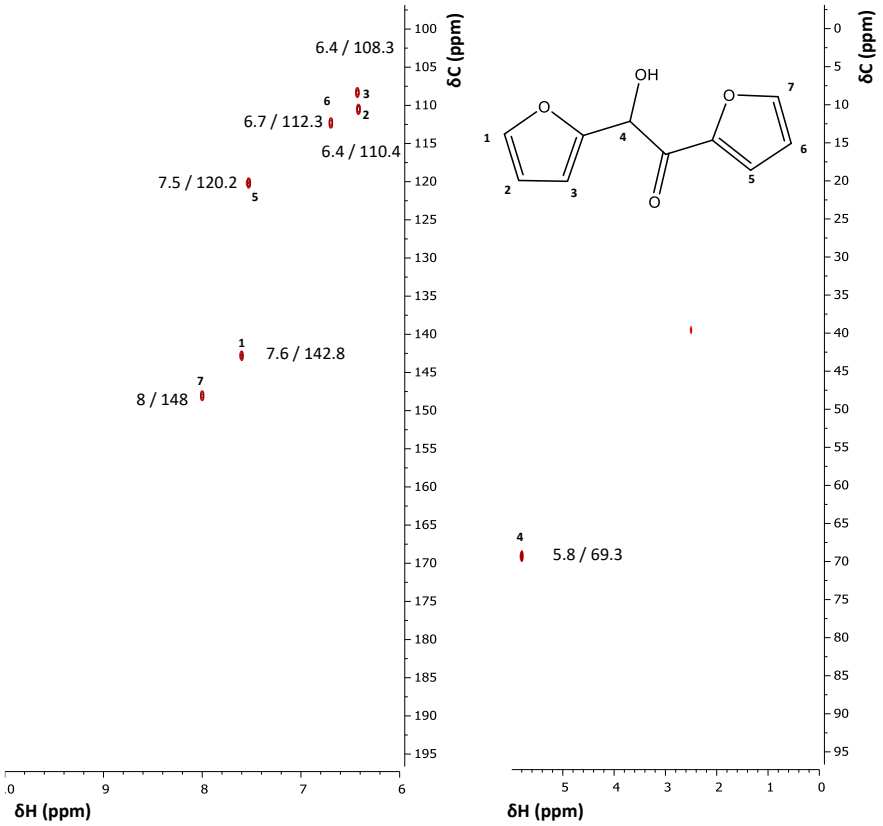


Figure S 44. [^{1}H; ^{13}C] HSQC spectrum of 1,2-di(furan-2-yl)-2-hydroxyethan-1-one in DMSO-d_6.

1,2-Di(furan-2-yl)ethane-1,2-dione

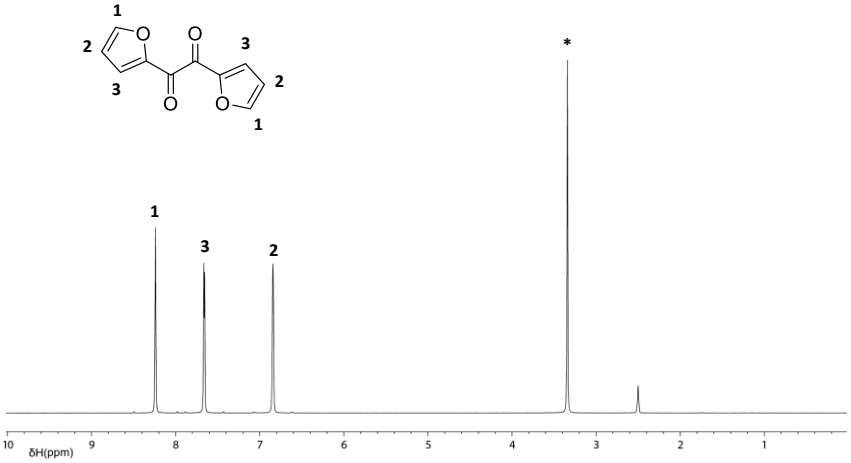


Figure S 45. ^{1}H NMR spectrum of 1,2-di(furan-2-yl)ethane-1,2-dione in DMSO-d_6.
* peak due to water contamination in DMSO-d_6.

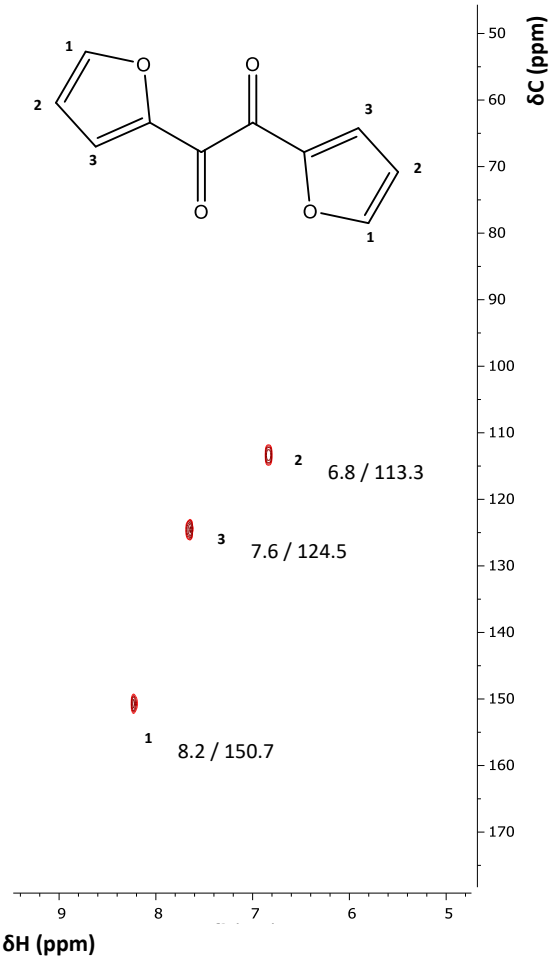


Figure S 46. [^{1}H; ^{13}C] HSQC spectrum of 1,2-di(furan-2-yl)ethane-1,2-dione in DMSO-d_6.

2,5-Dimethylfuran-3(2H)-one

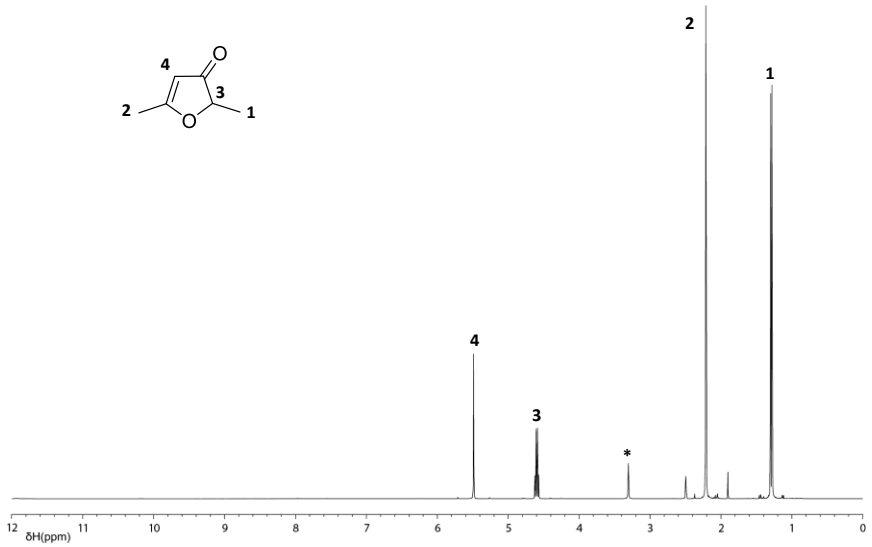


Figure S 47. ^{1}H NMR spectrum of 2,5-dimethylfuran-3(2H)-one in DMSO-d_6.
* peak due to water contamination in DMSO-d_6.

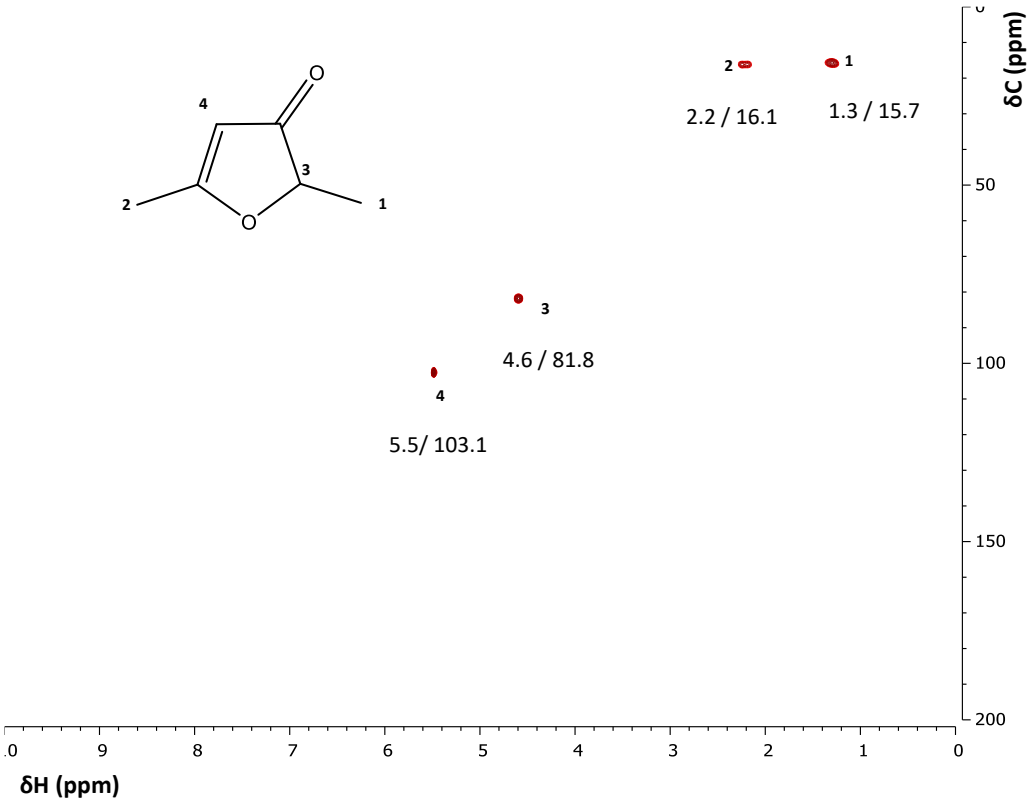


Figure S 48. [^{1}H; ^{13}C] HSQC spectrum of 2,5-dimethylfuran-3(2H)-one in DMSO-d_6.

1-(7-Hydroxybenzofuran-2-yl)ethan-1-one

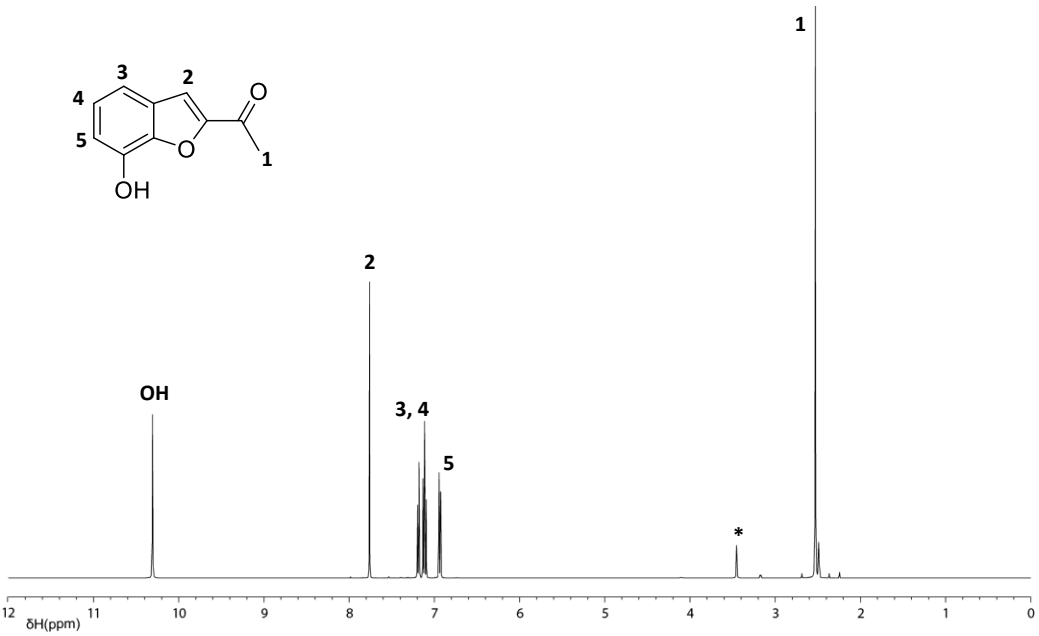


Figure S 49. ^{1}H NMR spectrum of 1-(7-hydroxybenzofuran-2-yl)ethan-1-one in DMSO-d_6.
* peak due to water contamination in DMSO-d_6.

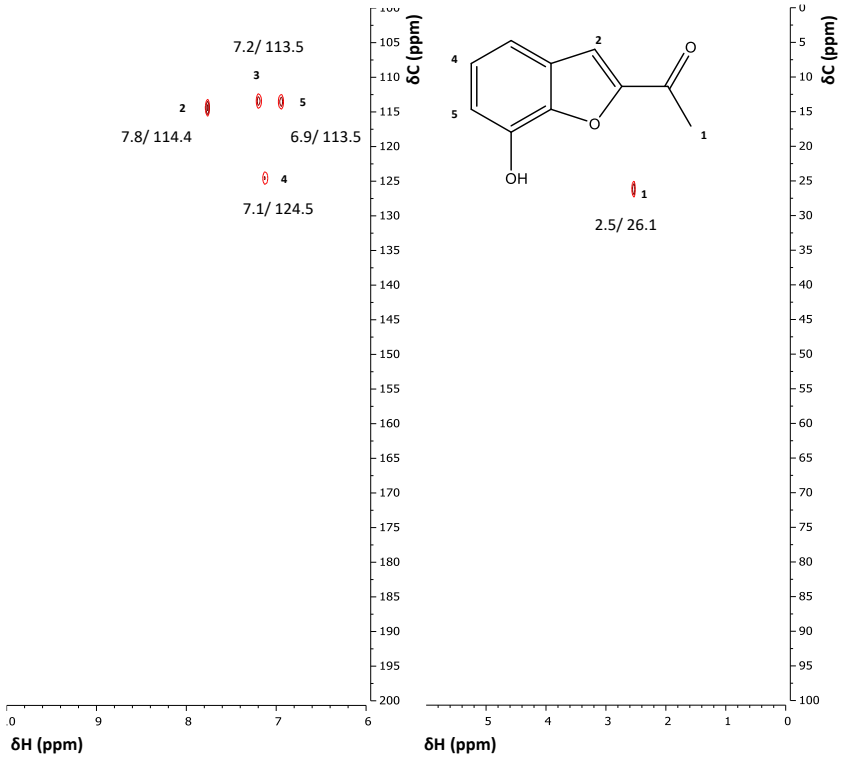


Figure S 50. [^{1}H; ^{13}C] HSQC spectrum of 1-(7-hydroxybenzofuran-2-yl)ethan-1-one in DMSO-d_6.

2-(Diethoxymethyl)furan

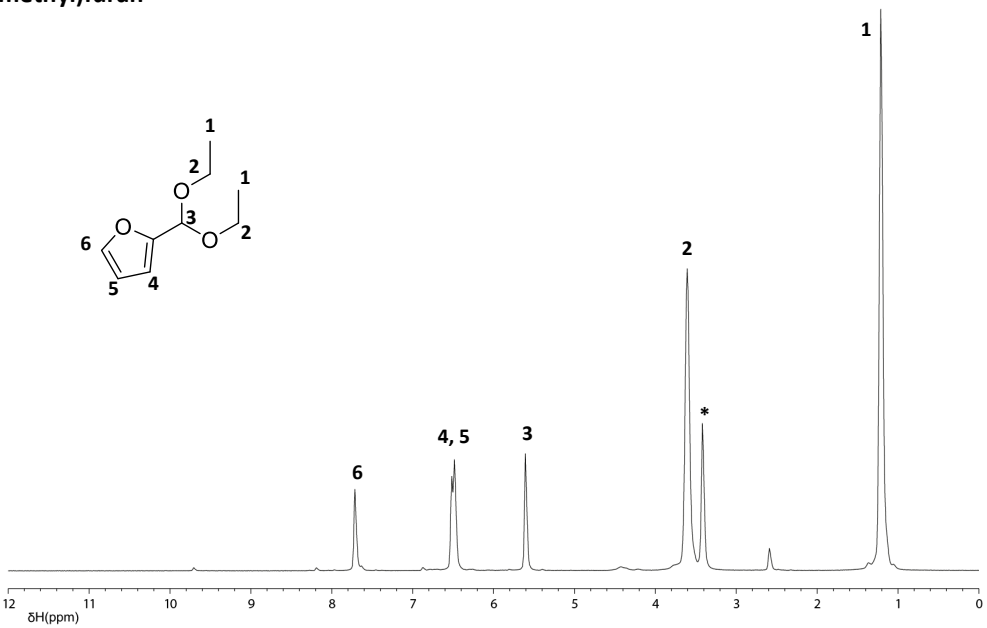


Figure S 51. ^{1}H NMR spectrum of 2-(diethoxymethyl)furan in DMSO-d_6.
* peak due to water contamination in DMSO-d_6.

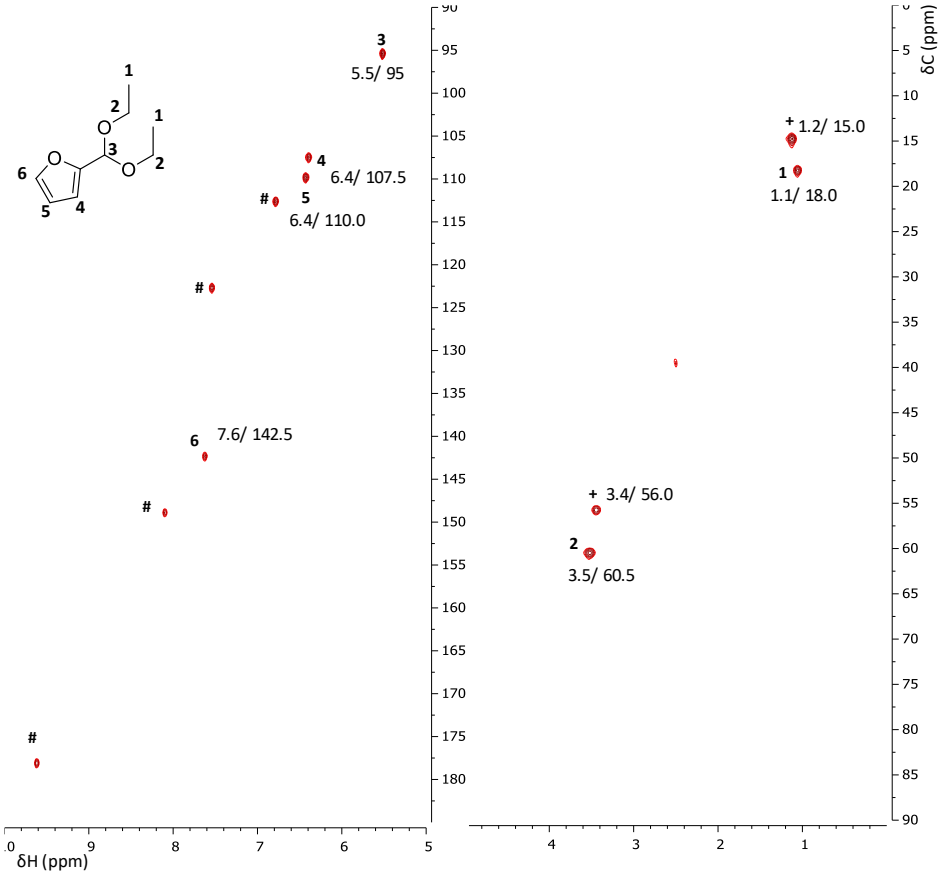


Figure S 52. [^{1}H; ^{13}C] HSQC spectrum of 2-(diethoxymethyl)furan in DMSO-d_6. It also contains traces of ethanol (+) and furfural (#) as impurities

(E)-4-(furan-2-yl)but-3-en-2-one

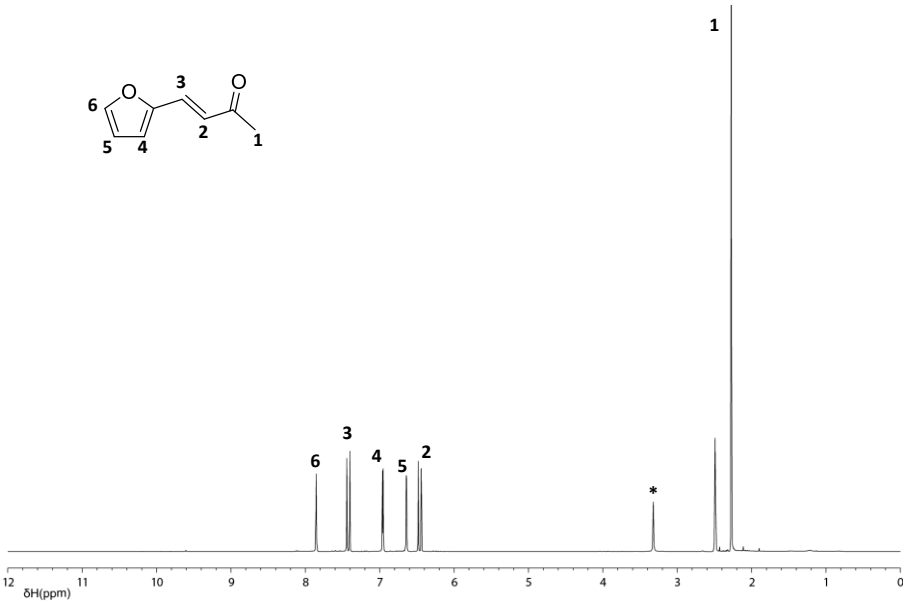


Figure S 53. ^{1}H NMR spectrum of (E)-4-(furan-2-yl)but-3-en-2-one in DMSO-d$_6$.
* peak due to water contamination in DMSO-d$_6$.

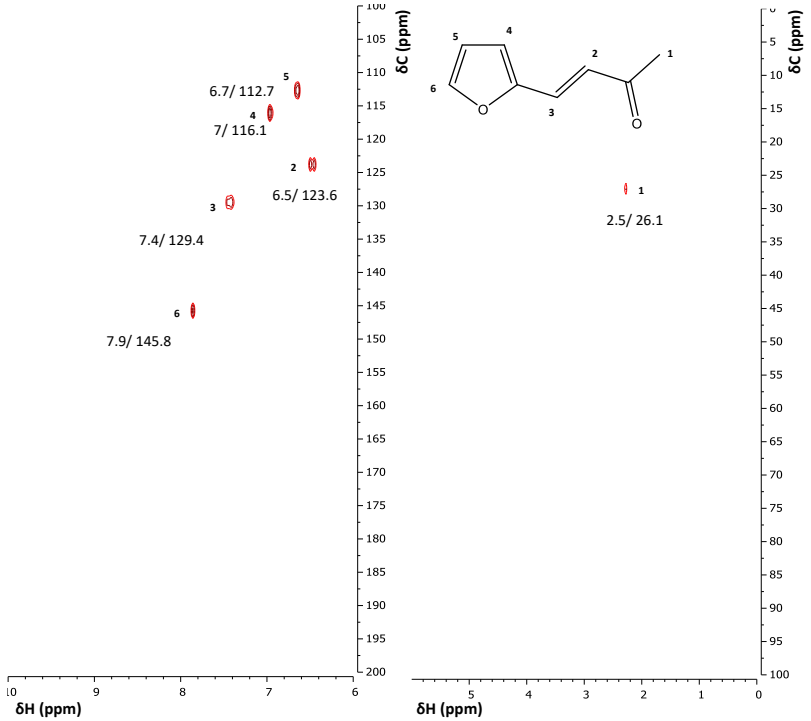


Figure S 54. [^{1}H; ^{13}C] HSQC spectrum of (E)-4-(furan-2-yl)but-3-en-2-one in DMSO-d$_6$.

1-(5-Methylfuran-2-yl)ethan-1-one

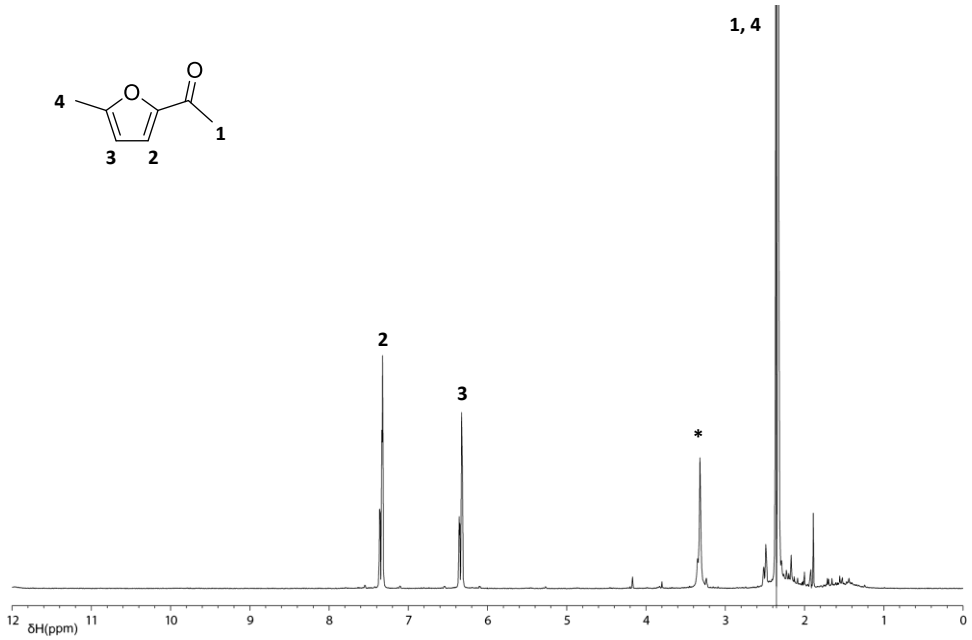


Figure S 55. ^{1}H NMR spectrum of 1-(5-methylfuran-2-yl)ethan-1-one in DMSO-d$_6$.
* peak due to water contamination in DMSO-d$_6$.

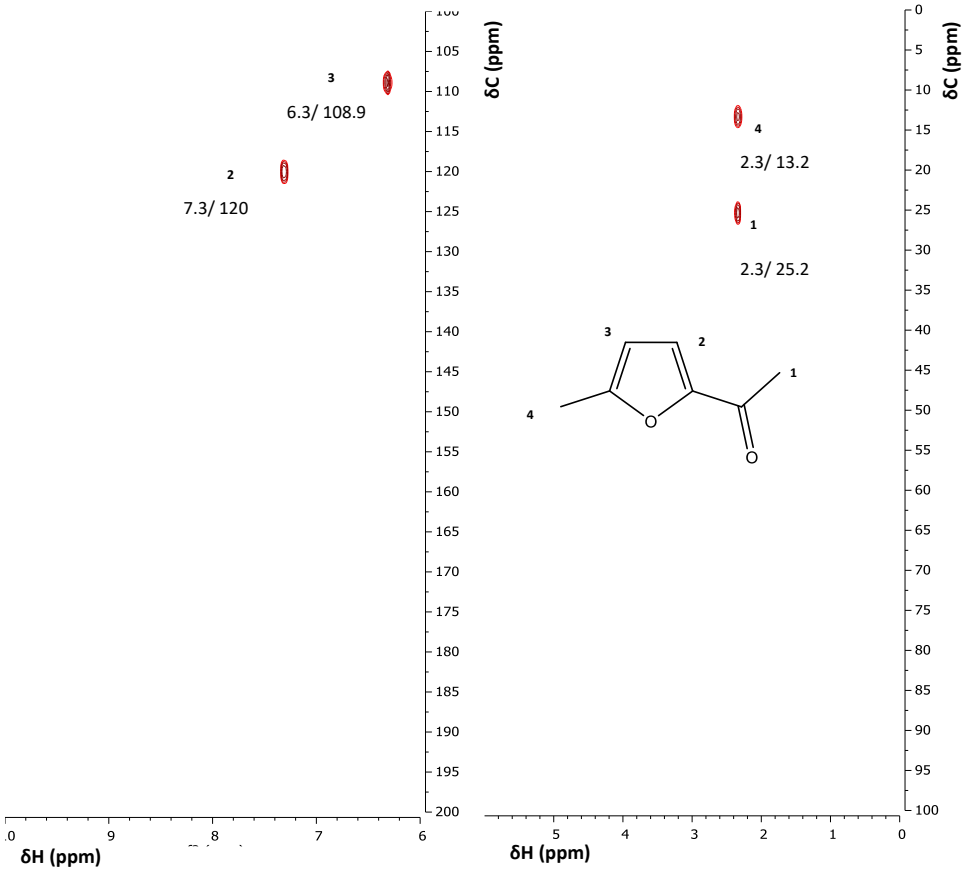


Figure S 56. [^{1}H; ^{13}C] HSQC spectrum of 1-(5-methylfuran-2-yl)ethan-1-one in DMSO-d$_6$.

(E)-3-(furan-2-yl)-2-methylacrylaldehyde

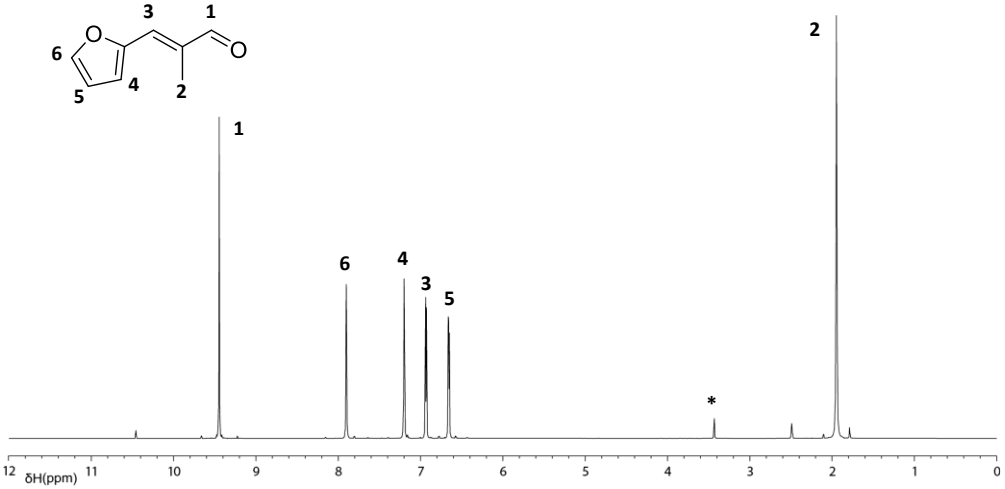


Figure S 57. ^{1}H NMR spectrum of (E)-3-(furan-2-yl)-2-methylacrylaldehyde in DMSO-d_6.

1-(Furan-2-yl)propan-2-one

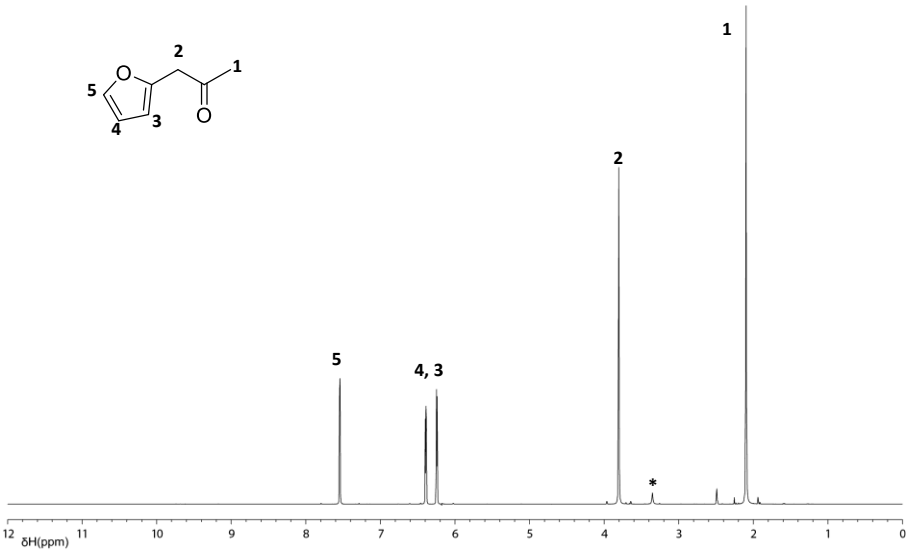


Figure S 58. ^{1}H NMR spectrum of 1-(furan-2-yl)propan-2-one in DMSO-d_6.
* peak due to water contamination in DMSO-d_6.

[^{1}H,^{13}C] HSQC NMR data in $CDCl_3$ taken from reference[4] δ:
1: 2.2/29.1
2: 3.7/43.2
3: 6.2/108.3
4: 6.3/110.7
5: 7.4/142.2

(E)-3-(Furan-2-yl)acrylic acid

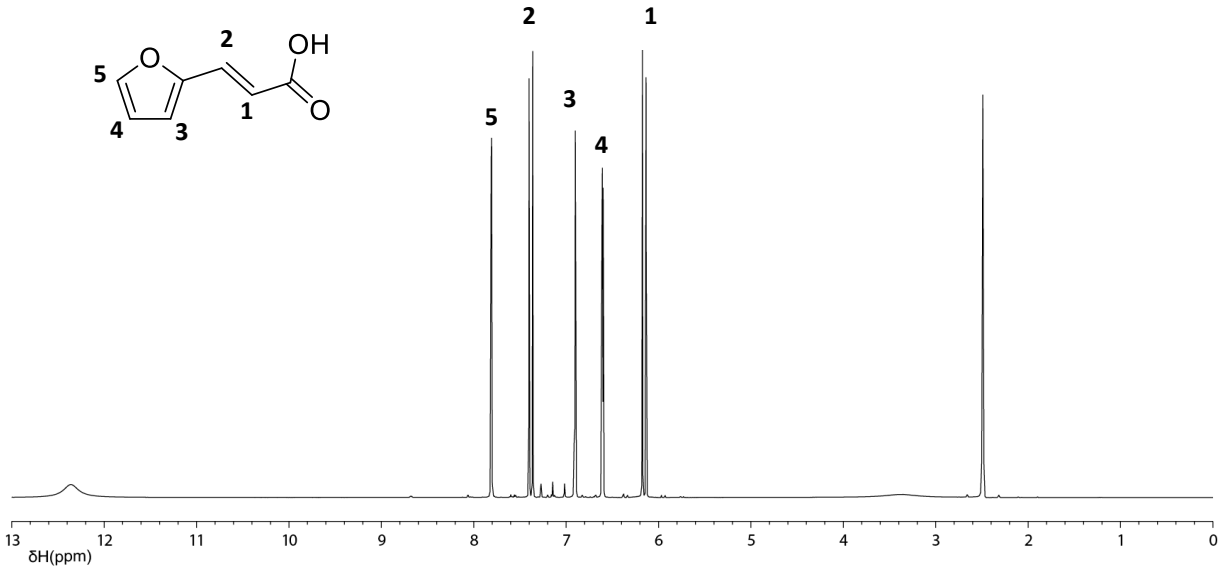


Figure S 59. ^{1}H NMR spectrum of (E)-3-(furan-2-yl)acrylic acid in DMSO-d$_6$.

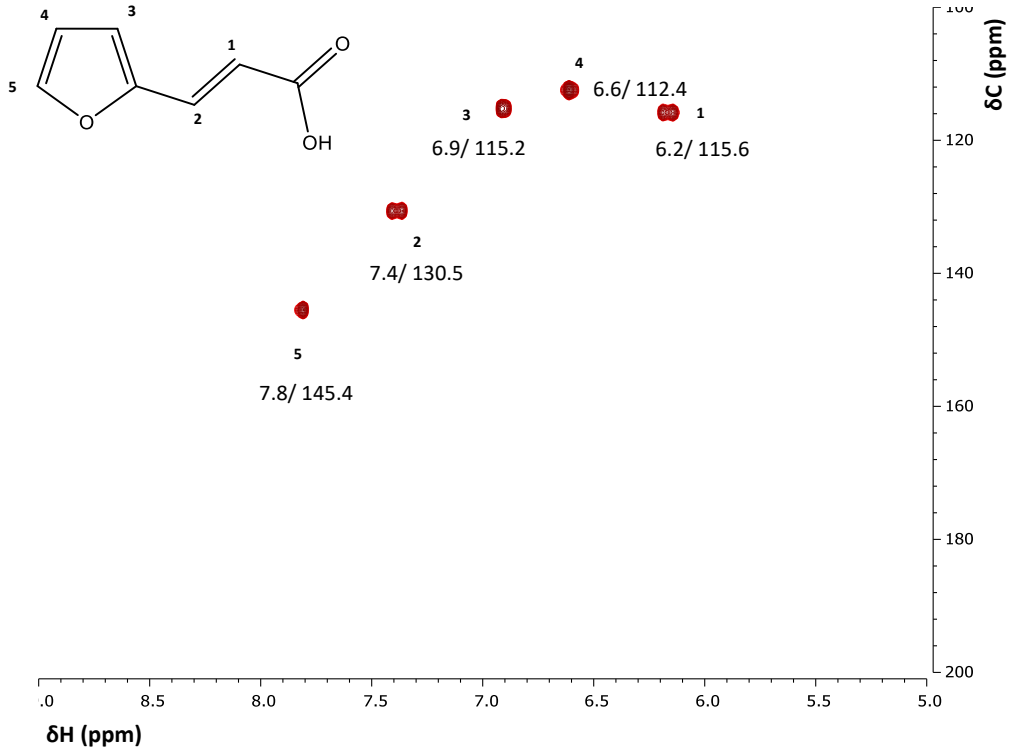


Figure S 60. [^{1}H; ^{13}C] HSQC spectrum of (E)-3-(furan-2-yl)acrylic acid in DMSO-d$_6$.

2-(Furan-2-yl)acetic acid

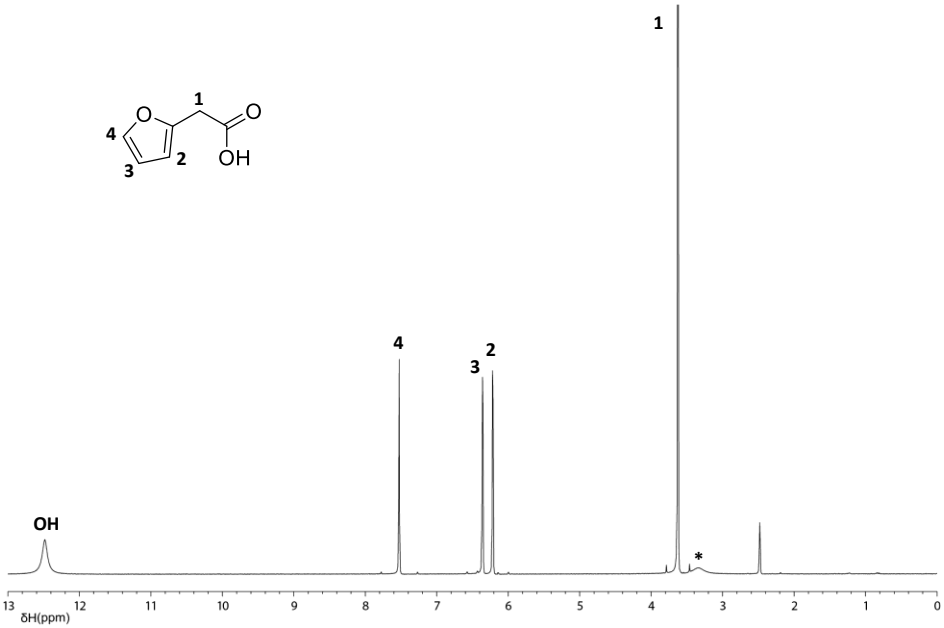


Figure S 61. ^{1}H NMR spectrum of 2-(furan-2-yl)acetic acid in DMSO-d_6.
* peak due to water contamination in DMSO-d_6.

2-(Furan-2-yl)-2-oxoacetic acid

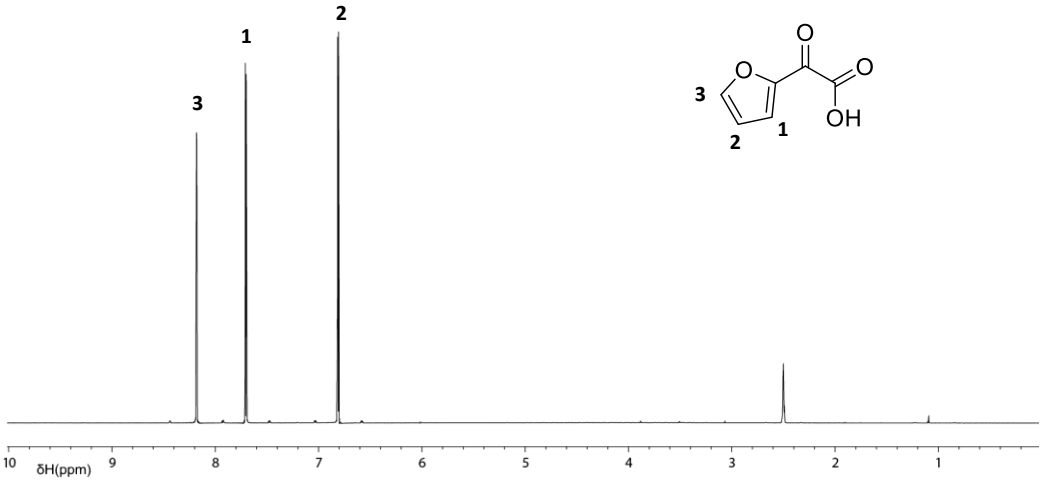


Figure S 62. ^{1}H NMR spectrum of 2-(furan-2-yl)-2-oxoacetic acid in DMSO-d_6.

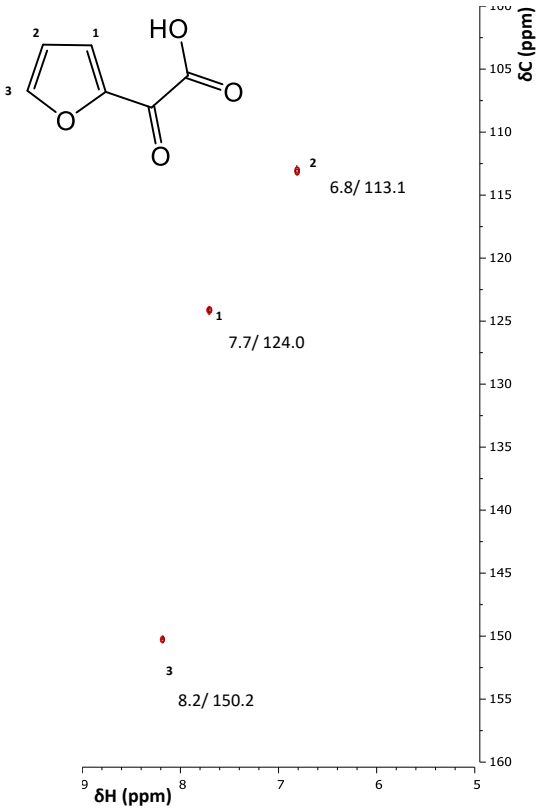


Figure S 63. [^{1}H; ^{13}C] HSQC spectrum of 2-(furan-2-yl)-2-oxoacetic acid in DMSO-d_6.

1-([2,2':5',2''-Terfuran]-5-yl)ethan-1-one

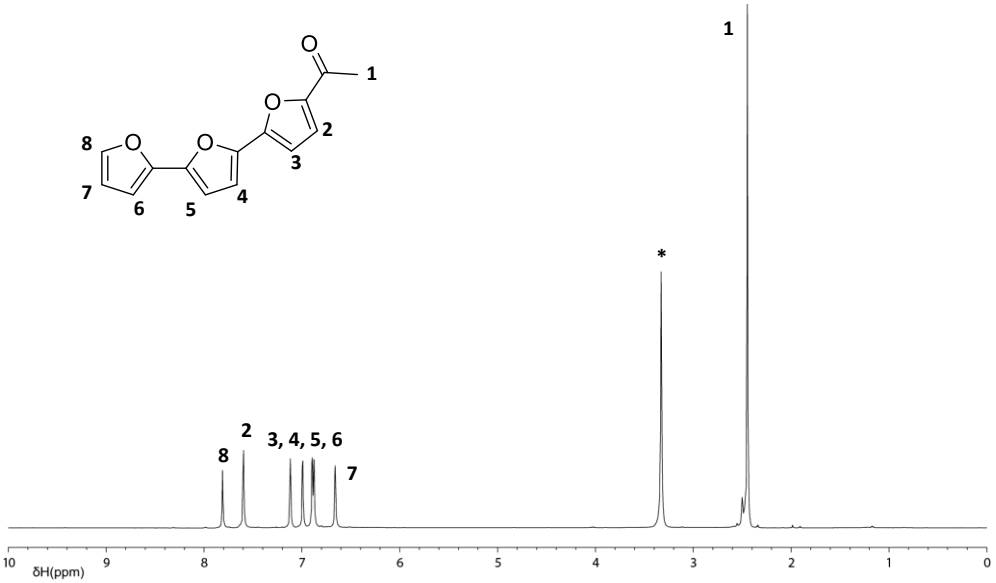


Figure S 64. ^{1}H NMR spectrum of 1-([2,2':5',2''-terfuran]-5-yl)ethan-1-one in DMSO-d_6.
* peak due to water contamination in DMSO-d_6.

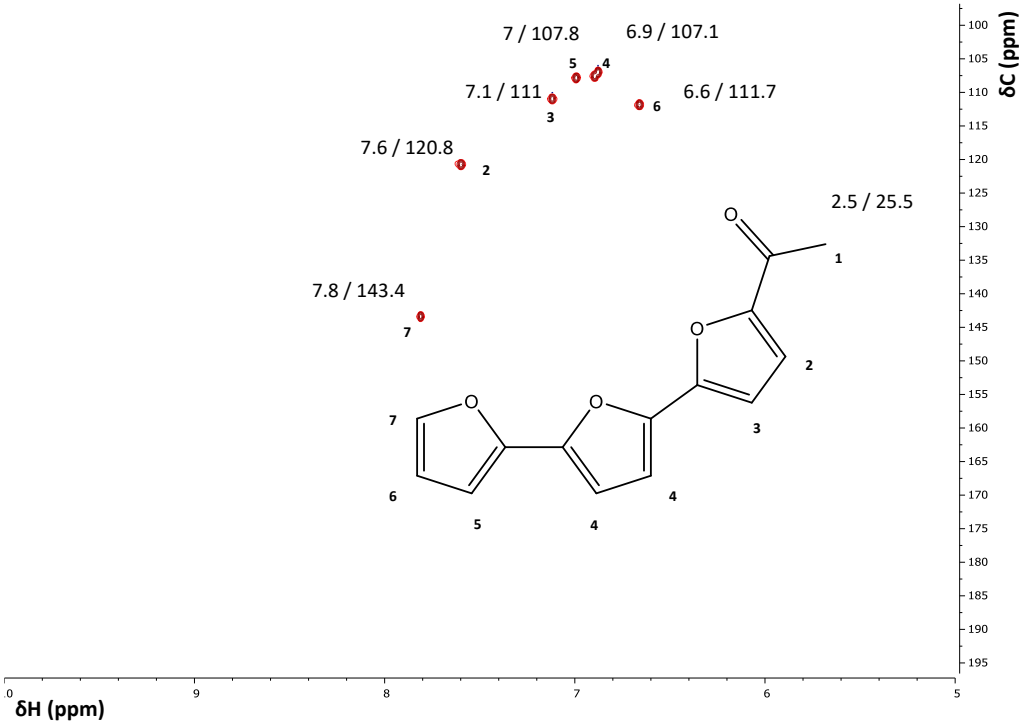


Figure S 65. [^{1}H; ^{13}C] HSQC spectrum of 1-([2,2':5',2''-terfuran]-5-yl)ethan-1-one in DMSO-d_6.

5,5'-(Propane-2,2-diyl)bis(furan-2-carboxylic acid)

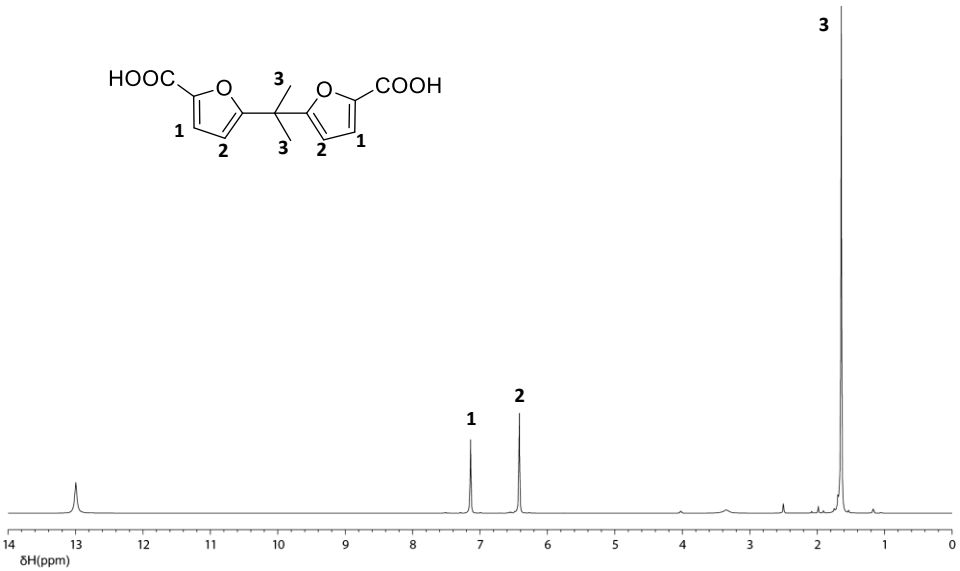


Figure S 66. ^{1}H NMR spectrum of 5,5'-(propane-2,2-diyl)bis(furan-2-carboxylic acid) in DMSO-d_6.

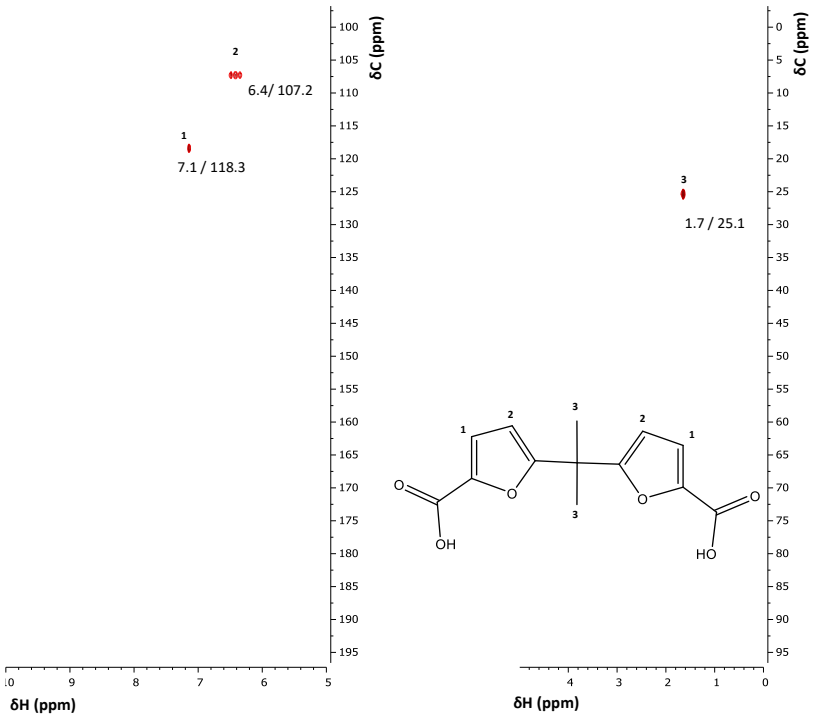


Figure S 67. [^{1}H; ^{13}C] HSQC spectrum of 5,5'-(propane-2,2-diyl)bis(furan-2-carboxylic acid) in DMSO-d_6.

5,5'-(Propane-2,2-diyl)bis(furan-2-carbaldehyde)

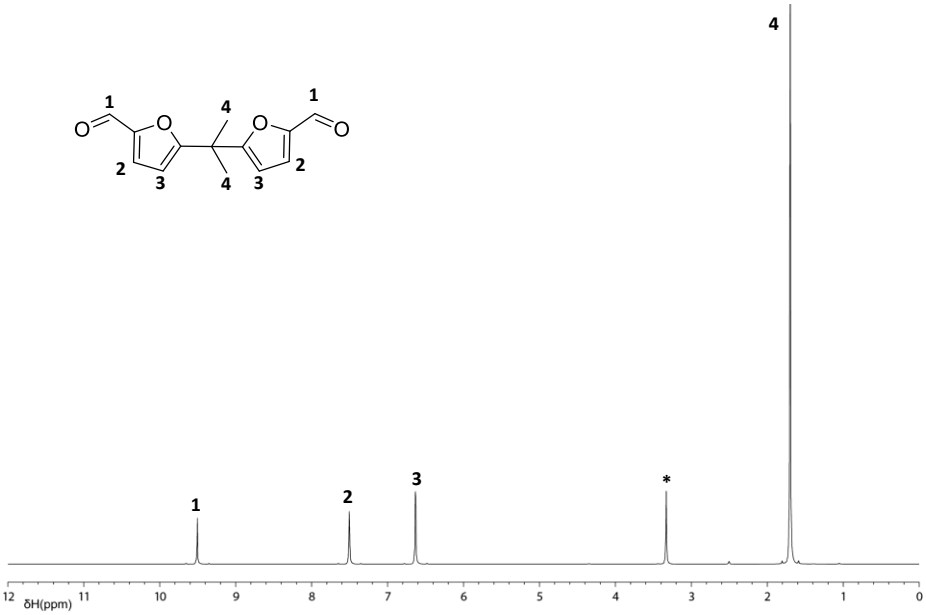


Figure S 68. ^{1}H NMR spectrum of 5,5'-(propane-2,2-diyl)bis(furan-2-carbaldehyde) in DMSO-d_6.
* peak due to water contamination in DMSO-d_6.

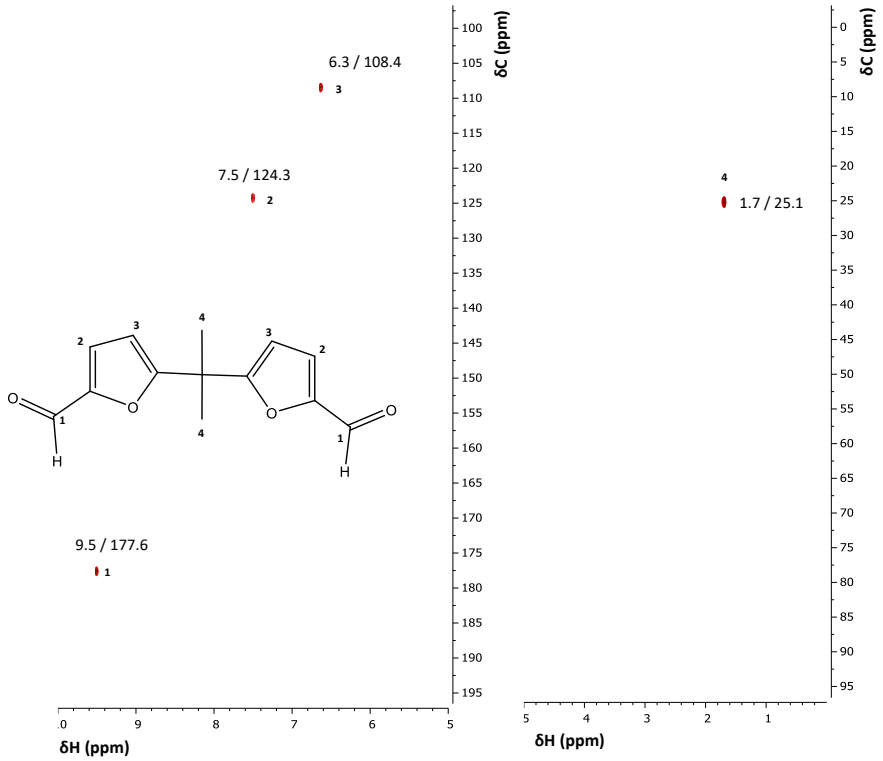


Figure S 69. [^{1}H; ^{13}C] HSQC spectrum of 5,5'-(propane-2,2-diyl)bis(furan-2-carbaldehyde) in DMSO-d_6.

Ethyl 2-methyl-5-(1,2,3,4-tetrahydroxybutyl)furan-3-carboxylate

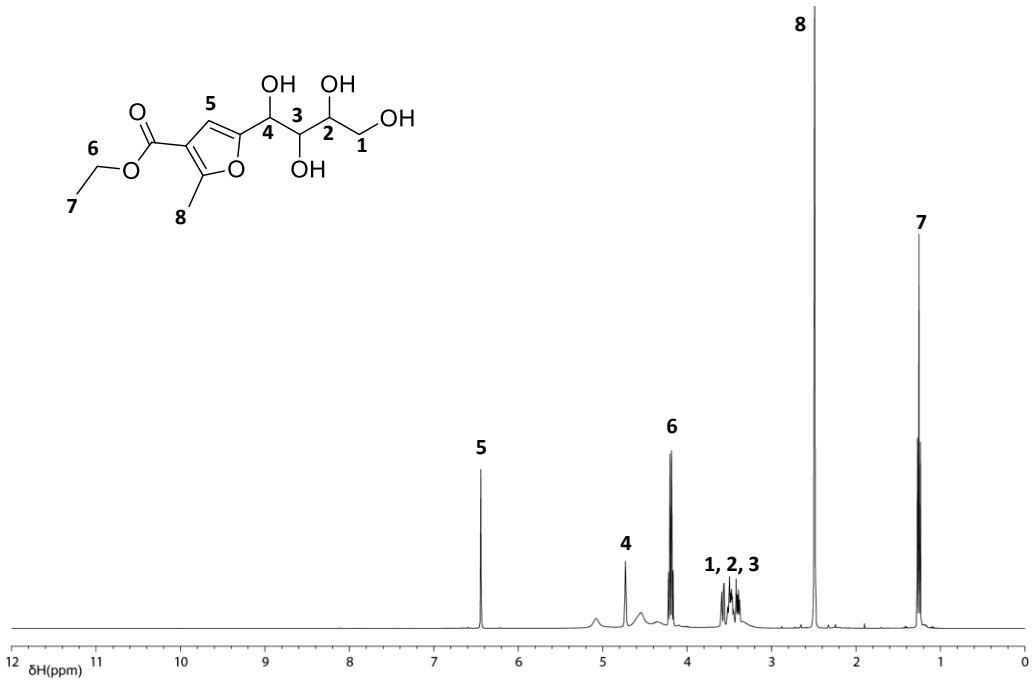


Figure S 70. ^{1}H NMR spectrum of ethyl 2-methyl-5-(1,2,3,4-tetrahydroxybutyl)furan-3-carboxylate in DMSO-d_6.

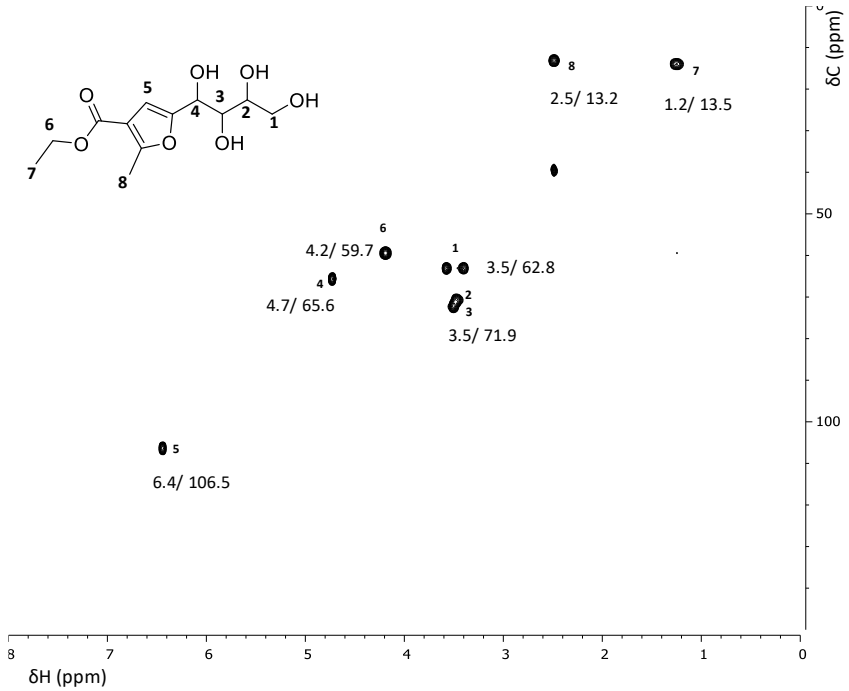


Figure S 71. [^{1}H; ^{13}C] HSQC spectrum of ethyl 2-methyl-5-(1,2,3,4-tetrahydroxybutyl)furan-3-carboxylate in DMSO-d_6.

(1E,4E)-1,5-bis(5-methylfuran-2-yl)penta-1,4-dien-3-one

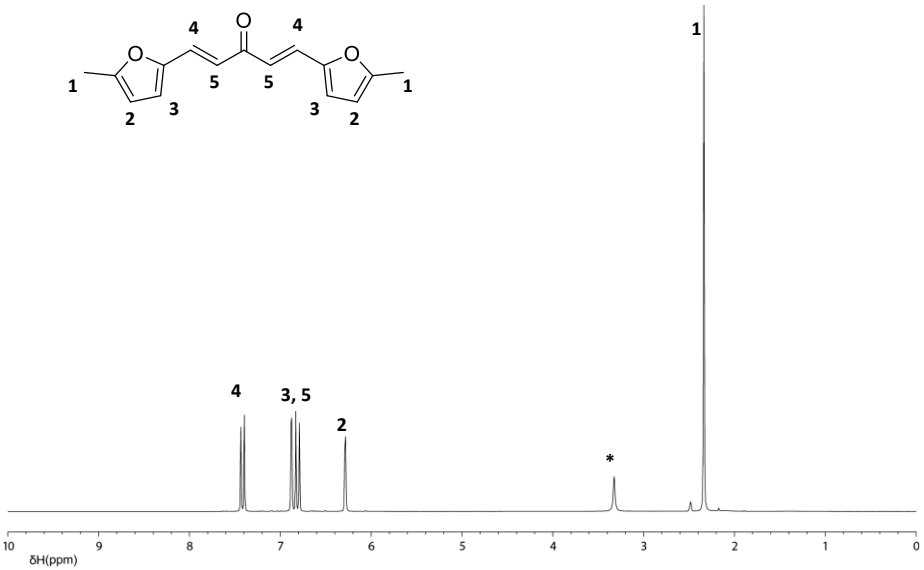


Figure S 72. ^{1}H NMR spectrum of (1E,4E)-1,5-bis(5-methylfuran-2-yl)penta-1,4-dien-3-one in DMSO-d_6.
* peak due to water contamination in DMSO-d_6.

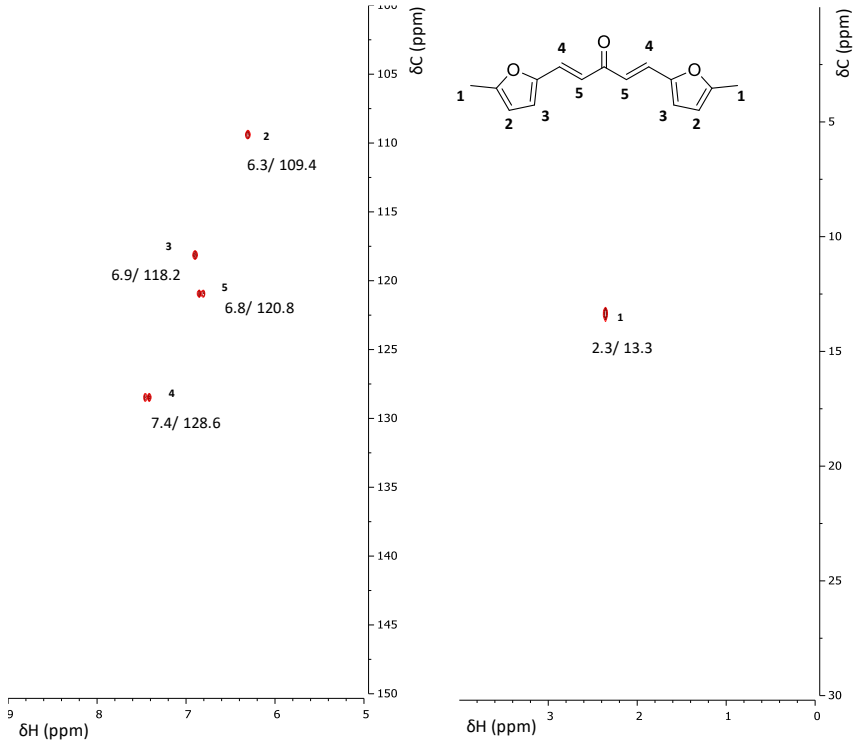


Figure S 73. [^{1}H; ^{13}C] HSQC spectrum of (1E,4E)-1,5-bis(5-methylfuran-2-yl)penta-1,4-dien-3-one in DMSO-d_6.

α-D-Glucose

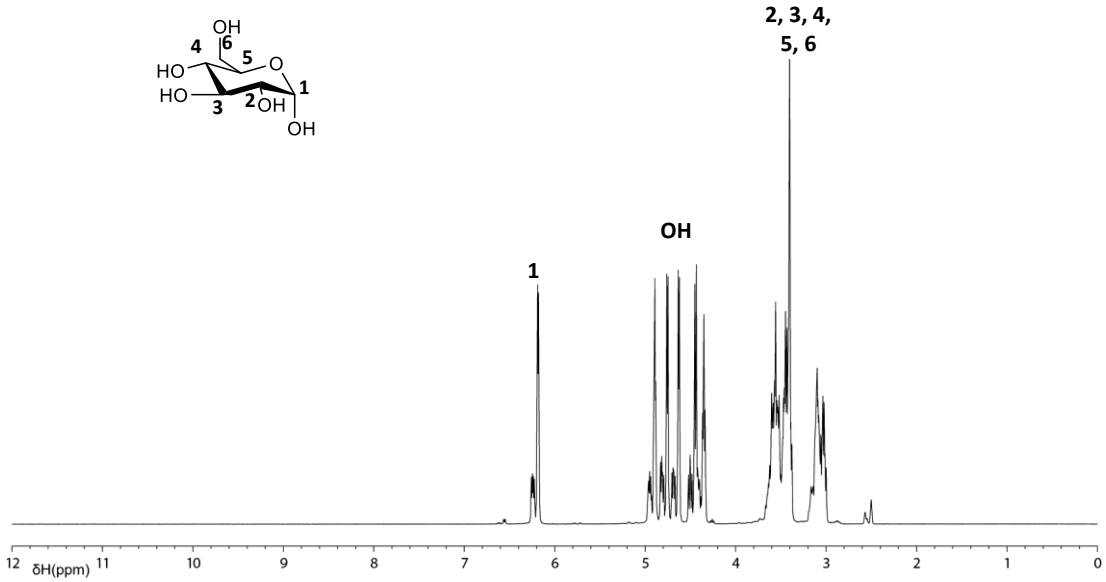


Figure S 74. ^{1}H NMR spectrum of α-D-glucose in DMSO-d$_6$.

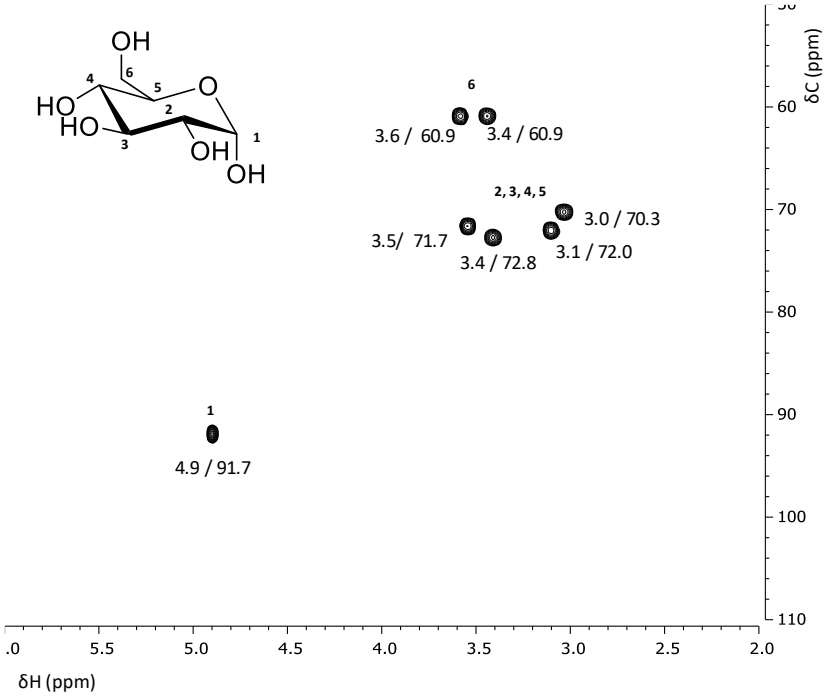


Figure S 75. [^{1}H; ^{13}C] HSQC spectrum of α-D-glucose in DMSO-d$_6$.

α-Methyl glucoside

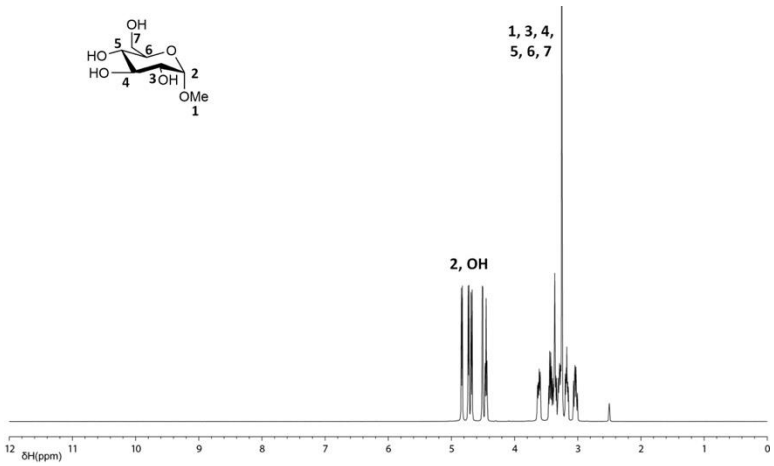


Figure S 76. ^{1}H NMR spectrum of α-methyl glucoside in DMSO-d$_6$.

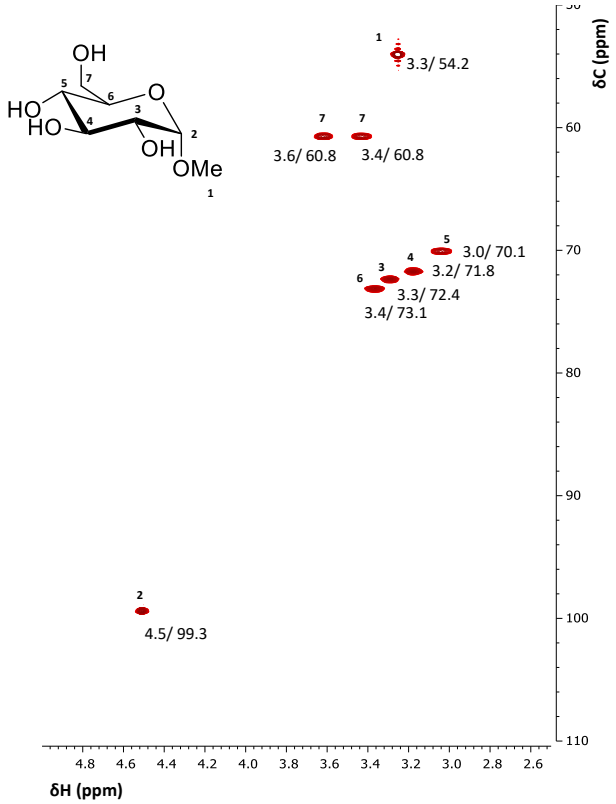


Figure S 77. [^{1}H; ^{13}C] HSQC spectrum of α-methyl glucoside in DMSO-d$_6$.

β-D-Fructose

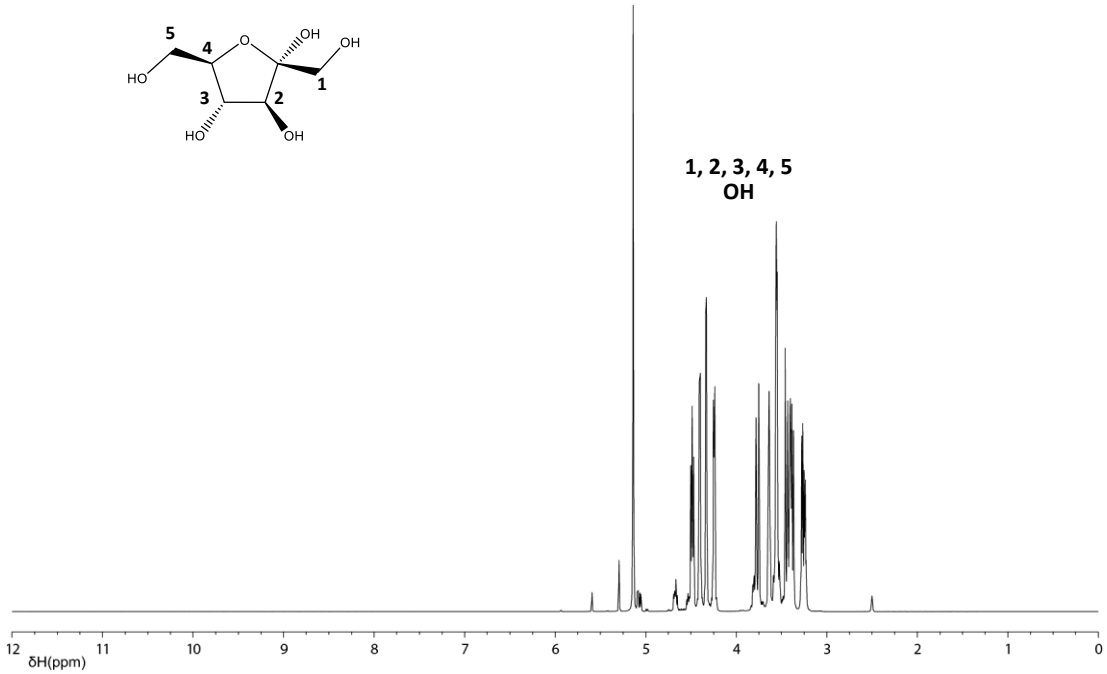


Figure S 78. ^{1}H NMR spectrum of β-D-fructose in DMSO-d$_6$.

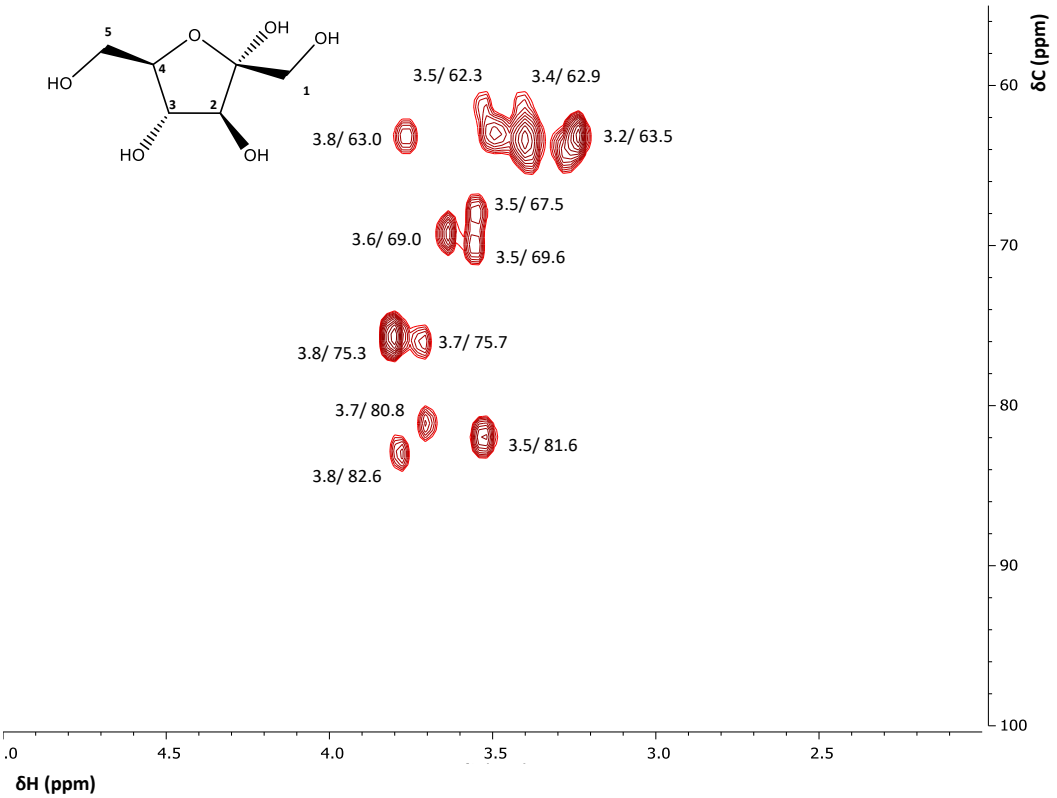


Figure S 79. [^{1}H; ^{13}C] HSQC spectrum of β-D-fructose in DMSO-d$_6$.

Sucrose

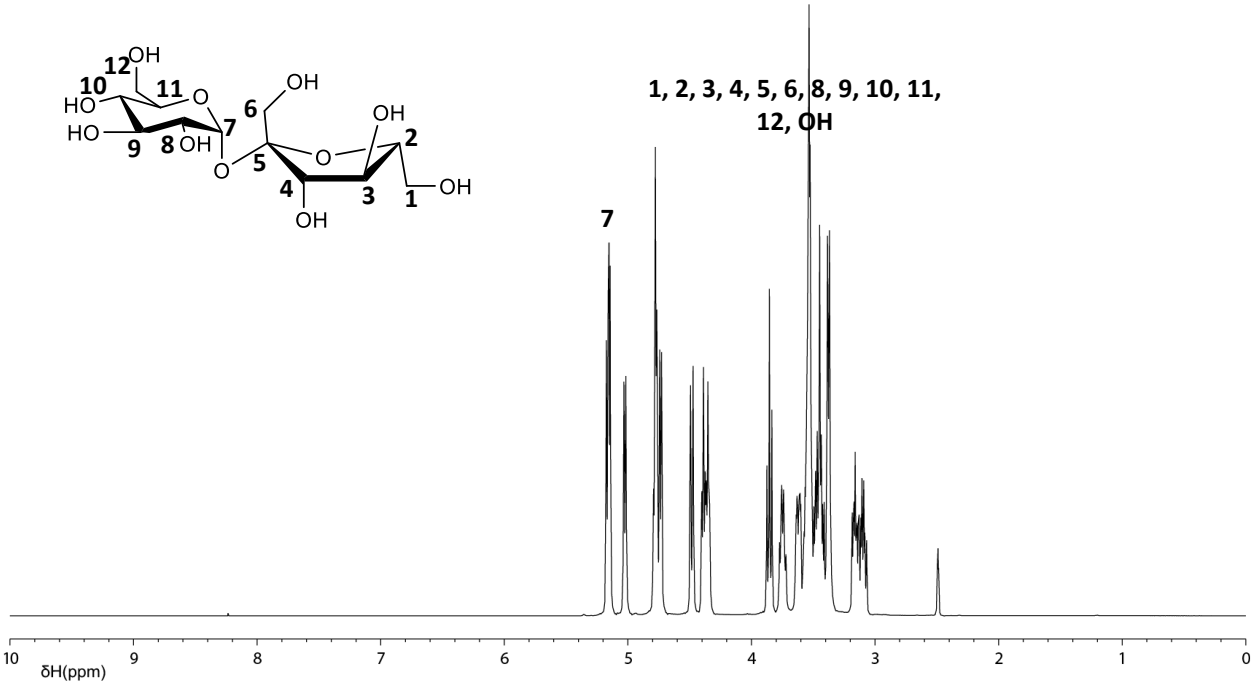


Figure S 80. ^{1}H NMR spectrum of sucrose in DMSO-d$_6$.

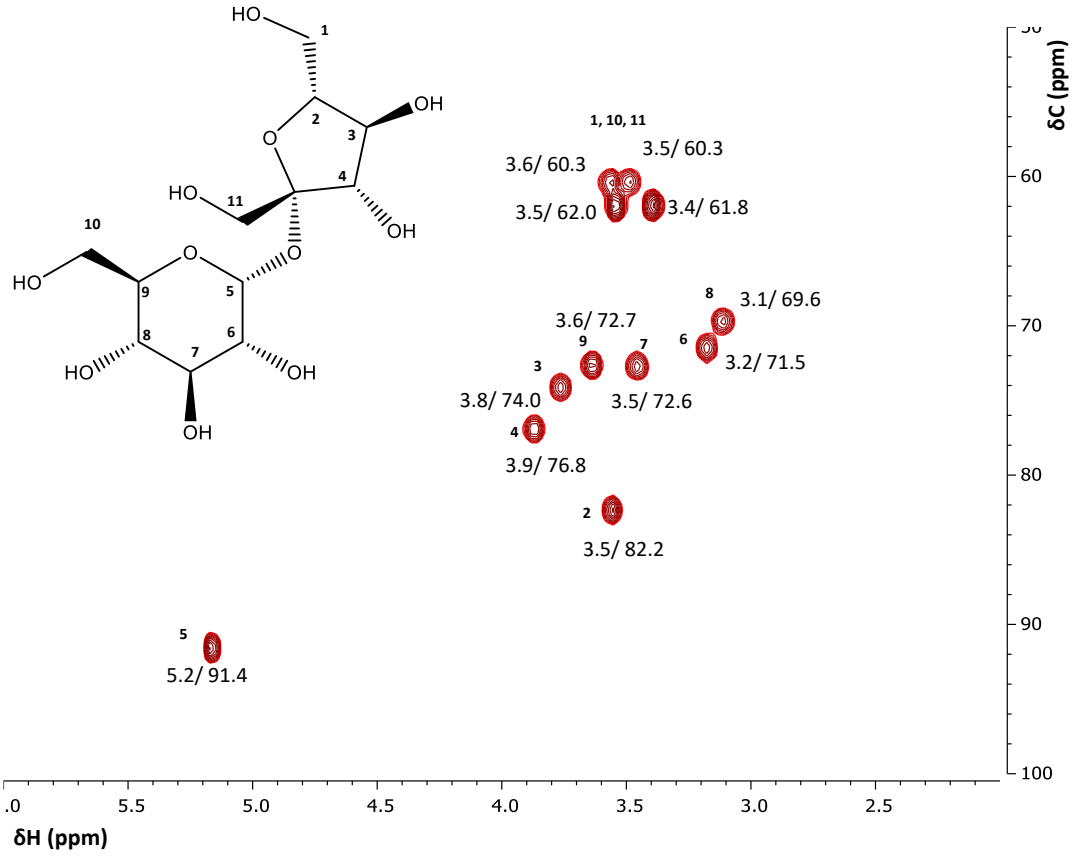
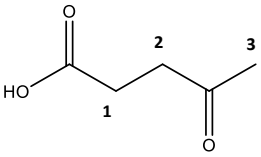


Figure S 81. [^{1}H; ^{13}C] HSQC spectrum of sucrose in DMSO-d$_6$.

Other NMR data found in literature

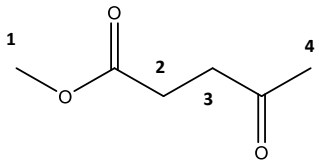
Levulinic acid[5]



^{1}H NMR ($CDCl_3$) δ: 2.17 (s, 3H), 2.59 (t, 2H), 2.73 (t, 2H)
^{13}C NMR ($CDCl_3$) δ: 27.9, 29.9, 37.8, 178.2, 207.0.

[^{1}H,^{13}C] HSQC NMR δ:
1: 2.6/29.9 **2**: 2.7/37.8 **3**: 2.2/27.9

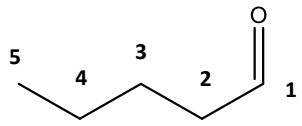
Methyl levulinate[6]



^{1} H NMR (CDCl3) δ: 3.66 (3H, s, OCH3), 2.19 (3H, s, CH3), 2.56 (2H, t, CH2, J5), 2.70 (2H, t, CH2, J 6.3),
^{13}C NMR δ: 27.47 (C-2), 29.39 (CH3), 37.65 (C-3), 51.37 (OCH3), 173.03 (COO)

[^{1}H,^{13}C] HSQC NMR δ:
1: 3.6/51.4 **2**: 2.6/27.5 **3**: 2.7/37.7 **4**: 2.2/29.4

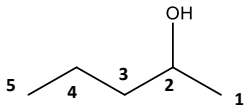
Valeraldehyde[4]



^{1}H NMR ($CDCl_3$) δ: 0.93, 1.35, 1.59, 2.42, 9.76
^{13}C NMR ($CDCl_3$) δ: 13.8, 22.4, 24.3, 43.7, 202.8

[^{1}H,^{13}C] HSQC NMR δ:
1: 9.8/202.8 **2**: 2.4/43.7 **3**: 1.6/24.3 **4**: 1.4/22.4 **5**: 0.9/13.8

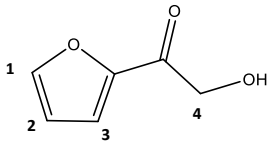
2-Pentanol[4]



^{1}H NMR ($CDCl_3$) δ: 0.93, 1.18, 1.52-1.29, 1.92, 3.79
^{13}C NMR ($CDCl_3$) δ: 14.1, 19.0, 23.4, 41.6, 67,7

[^{1}H,^{13}C] HSQC NMR δ:
1: 1.2/23.4 **2**: 3.8/67.7 **3**: 1.3-1.5/41.6 **4**: 1.3-1.5/19 **5**: 0.9/14.1

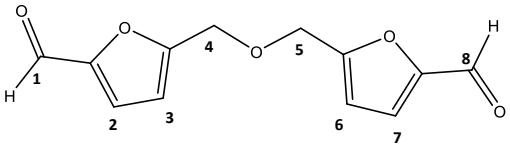
2-(2-Hydroxyacetyl)furan[5]



^{1}H NMR ($CDCl_3$) δ 3.26 (s, 1H), 4.71 (s, 2H), 6.56 (t, 1H), 7.26 (d, 1H), 7.60 (d, 1H)
^{13}C NMR ($CDCl_3$) δ: 65.2, 112.7, 118.02, 147.2, 150.3, 187.8;

[^{1}H,^{13}C] HSQC NMR δ:
1: 7.6/147.2 **2**: 6.6/112.6 **3**: 7.3/118 **4**: 4.7/65.2

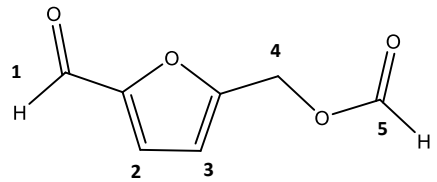
5,5'-(Oxybis(methylene))bis(furan-2-carbaldehyde): OBMF [7]



^{1}H NMR ($CDCl_3$) δ: 9.64(1H, s, CH=O),7.22 (1H, d, J4, =CH), 6.58 (1H, d, J4, =CH), 4.64 (2H, s,O-CH_2);
^{13}C NMR ($CDCl_3$) δ: 177.7, 157.2, 152.7, 121.9, 111.9, 64.6

[^{1}H,^{13}C] HSQC NMR δ:
1, 8: 9.6/177.7 **2, 7**: 7.2/121.9 **3, 6:** 6.6/111.9 **4,5**: 4.6/64.6

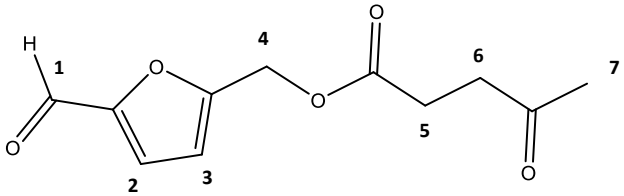
(5-Formylfuran-2-yl)methyl formate: HMF formate[7]



^{1}H NMR ($CDCl_3$) δ: 9.66 (1H, s, CH=O),8.12 (1H, s, O=C-H), 7.23 (1H, d, J4, =CH), 6.64 (1H, d, J4, =CH), 5.24 (2H, s, O-CH_2)
^{13}C NMR ($CDCl_3$) δ: 177.8, 160.1, 154.4, 152.9, 121.6, 112.9, 57.1

[^{1}H,^{13}C] HSQC NMR δ:
1: 9.7/177.8 **2**: 7.2/121.6 **3**: 6.6/112.9 **4**: 5.2/57.1 **5**: 8.1/160.1

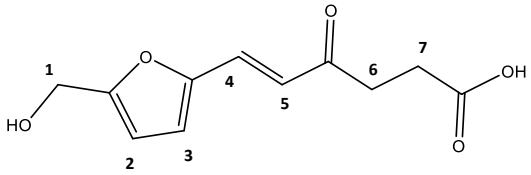
(5-Formylfuran-2-yl)methyl 4-oxopentanoate: HMF levulinate [7]



^{1}H NMR ($CDCl_3$) δ: 9.58 (1H, s, CH=O), 7.18 (1H, d, J4, =CH), 6.56 (1H, d, J4, =CH), 5.10 (2H, s, O-CH_2), 2.74 (2H, t, J6, CH_2), 2.58 (2H, t, J6, CH_2), 2.15 (3H, s, CH_3)
^{13}C NMR ($CDCl_3$) δ: 206.3, 177.7, 172.1, 155.4, 152.8, 121.8, 112.4, 58.0, 37.7, 29.7, 27.7

[^{1}H,^{13}C] HSQC NMR δ:
1: 9.6/177.7 **2**: 7.2/121.8 **3**: 6.6/112.4 **4**: 5.1/58 **5**: 2.6/27.7 **6**: 2.7/37.7
7: 2.1/29.7

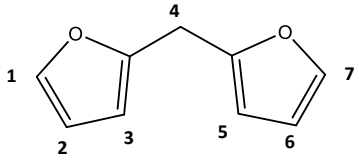
(E)-6-[5-(hydroxymethyl)furan-2-yl]-hex-4-oxo-5-enoic acid[7]



^{1}H NMR (DMSO-d_6) δ: 12.10 (1H, s, O=C-O-H), 7.41 (1H, d, J16, =CH), 6.92 (1H, d, J4, =CH), 6.53 (1H, d, J16, =CH), 6.47 (1H, d, J4, =CH), 5.39 (1H, t, O-H), 4.45 (2H, d, J6, O-CH_2), 2.88 (2H, t, J6, CH_2), 2.48 (2H, t, J6, CH_2)
^{13}C NMR (DMSO-d_6) δ: 198.1, 174.2, 159.2, 150.2, 129.2, 122.8, 118.0, 110.4, 56.3, 35.4, 28.3

[^{1}H,^{13}C] HSQC NMR δ:
1: 4.5/56.3 **2**: 6.5/110.4 **3**: 6.9/118 4: 6.6/129.2 **5**: 7.4/122.8 **6**: 2.9/35.4
7: 2.5/28.3

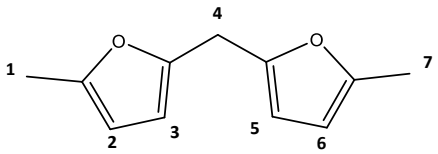
Di-furyl methane[4]



^{1}H NMR ($CDCl_3$) δ: 3.99, 6.08, 6.30, 7.33
^{13}C NMR ($CDCl_3$) δ: 27.4, 106.4, 110.38, 141.6, 151.58

[^{1}H,^{13}C] HSQC NMR δ:
1, 7: 7.3/141.6 **2, 6**: 6.3/110.4 **3, 5**: 6.1/106.4 **4**: 4/27.4

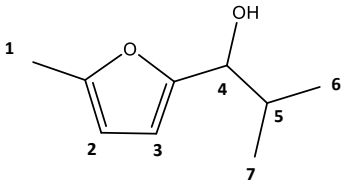
Bis(5-methylfuran-2-yl)methane[8]



^{1}H NMR (CD_3OD) δ: 5.91 (dd, 2H), 5.87 (dd, 2H), 3.83 (s, 2H), 2.21 (s, 6H)
^{13}C NMR ($CDCl_3$) δ: 150.55, 150.03, 106.44, 105.69, 26.71, 12.00

[^{1}H,^{13}C] HSQC NMR δ:
1, 7: 2.2/12 **2, 6**: 5.9/105.7 **3, 5**: 5.9/106.4 **4**: 3.8/26.7

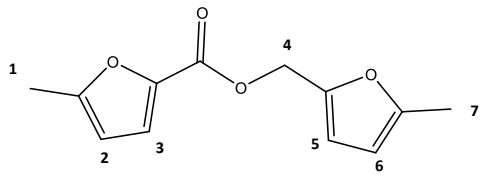
2-Methyl-1-(5-methylfuran-2-yl)propan-1-ol [8]



^{1}H NMR (CD_3OD) δ: 6.08 (d, 1H), 5.91 (d, 1H), 4.19 (d, 1H), 2.25 (s, 3H), 2.10 (t, 1H), 1.01 (d, 3H), 0.82 (d, 3H)
^{13}C NMR (CD_3OD) δ: 154.70, 150.74, 106.85, 105.41, 72.88, 32.90, 17.98, 17.71, 12.09

[^{1}H,^{13}C] HSQC NMR δ:
1: 2.3/12.1 **2**: 5.9/105.4 **3**: 6.1/106.9 **4**: 4.2/72.9 **5**: 2.1/32.9 **6, 7**: 1.0-0.8/17.7

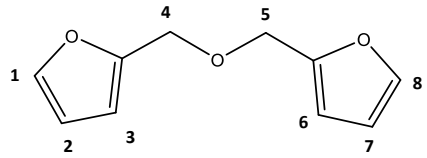
(5-Methylfuran-2-yl)methyl 5-methylfuran-2-carboxylate [8]



^{1}H NMR (DMSO-d_6) δ: 7.22 (d, 1H), 6.46 (d, 1H), 6.32 (d, 1H), 6.08 (d, 1H), 5.18 (s, 2H), 2.34 (s, 3H), 2.26 (s, 3H)
^{13}C NMR (DMSO-d_6) δ: 157.91, 157.88, 153.18, 147.81, 142.42, 120.70, 112.83, 109.42, 107.33, 58.20, 13.97, 13.74

[^{1}H,^{13}C] HSQC NMR δ:
1: 2.3/13.7 **2**: 6.5/112.8 **3**: 7.2/120.7 **4**: 5.2/58.2 **5**: 6.3/109.4 **6**: 6.1/107.3
7: 2.3/13.9

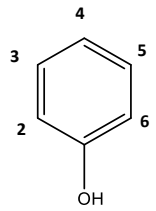
2, 2'-Difurfuryl ether [9]



^{1}H NMR ($CDCl_3$) δ: 4.48 (4H, s, -CH_2-O), 6.34(4H, s, -CH = CH-), 7.42(2H, d, J = 0.9 Hz, C = CH-O)
^{13}C NMR ($CDCl_3$) δ: 63.38, 109.54, 110.19, 142.81, 151.30

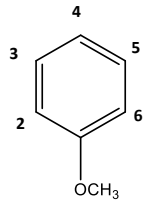
[^{1}H,^{13}C] HSQC NMR δ:
1, 8: 7.4/142.8 **2, 7**: 6.3/110.2 **3, 6:** 6.3/109.5 **4,5**: 4.5/63.4

Phenol [10]



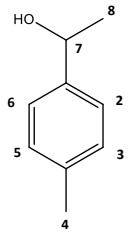
[^{1}H,^{13}C] HSQC NMR δ:
2, 6: 6.8/115.4 **3, 5**: 7.2/129.7 **4**: 6.9/120.9

Anisole [10]



[^{1}H,^{13}C] HSQC NMR δ:
2, 6: 6.9/113.9 **3, 5**: 7.3/129.5 **4**: 6.9/120.6

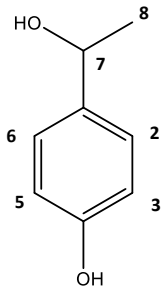
1-(p-Tolyl)ethanol[4]



^{1}H NMR ($CDCl_3$) δ: 1.44, 2.14, 2.32, 4.80, 7.12, 7,21
^{13}C NMR ($CDCl_3$) δ: 21.0, 25.1, 69.9, 125.4, 129.0, 136.8, 143,1

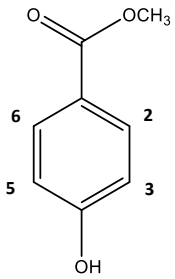
[^{1}H,^{13}C] HSQC NMR δ:
2, 6: 7.2/125.4 **3, 5**: 7.1/129 4: 2.3/21 **7**: 4.8/69.9 **8**: 1.4/25

1-(4-Hydroxyphenyl)ethanol [10]



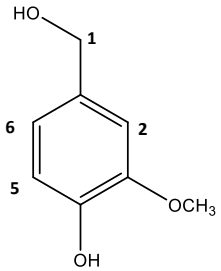
[^{1}H,^{13}C] HSQC NMR δ:
2, 6: 7.3/126.9 **3, 5**: 6.8/115.3 **7**: 4.9/67.7 **8**: 1.5/25.8

Methyl 4-hydroxybenzoate [10]



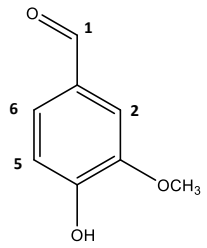
[^{1}H,^{13}C] HSQC NMR δ:
2, 6: 7.9/131.9 **3, 5**: 6.9/115.3

4-Hydroxy-3-methoxybenzyl alcohol [10]



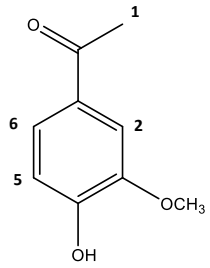
[^{1}H,^{13}C] HSQC NMR δ:
1: 4.6/65.4 **2**: 6.9/108 **5**: 6.8/114.2 **6**: 6.8/118.3

Vanillin [10]



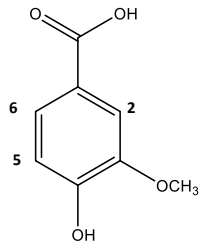
[^{1}H,^{13}C] HSQC NMR δ:
1: 9.8/191.2 **2**: 7.4/109.0 **5**: 7.0/114.6 **6**: 7.4/127.6

Acetovanillone [10]



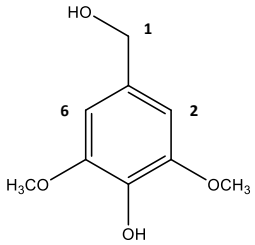
[^{1}H,^{13}C] HSQC NMR δ:
1: 2.6/26.2 **2**: 7.5/109.9 **5**: 6.9/113.9 **6**: 7.5/124.1

Vanillic acid [10]



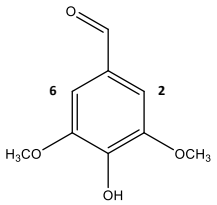
[^{1}H,^{13}C] HSQC NMR δ:
2: 7.6/112.7 **5**: 6.9/115.0 **6**: 7.7/123.4

Syringyl alcohol [10]



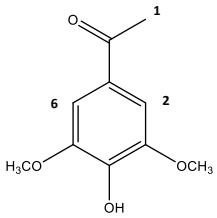
[^{1}H,^{13}C] HSQC NMR δ:
1: 4.6/65.68 **2, 6**: 6.6/103.9

Syringaldehyde [10]



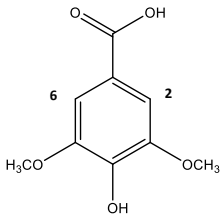
[^{1}H,^{13}C] HSQC NMR δ:
2, 6: 7.1/106.8

Acetosyringone [10]



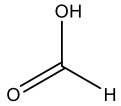
[^{1}H,^{13}C] HSQC NMR δ:
1: 2.6/26.2 **2, 6**: 7.2/105.8

Syringic acid [10]



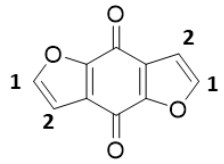
[^{1}H,^{13}C] HSQC NMR δ:
2, 6: 7.4/108.2

Formic acid [4]



[^{1}H,^{13}C] HSQC NMR δ:
1: 8.3/166.2

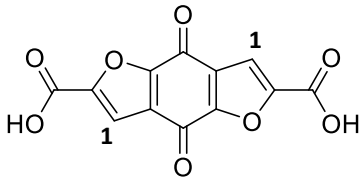
Furo[2,3-f]benzofuran-4,8-dione [11]



^{1}H NMR ($CDCl_3$) δ: d 7.7 (d, J = 1.8 Hz, 2H), 6.9 (d, J = 1.8 Hz, 2H)
^{13}C NMR ($CDCl_3$) δ: 170.9, 152.0, 148.2, 128.4, 108.3

[^{1}H,^{13}C] HSQC NMR δ:
1: 7.7/152.0 **2**: 6.9/108.3

4,8-Dioxofuro[2,3-f]benzofuran-2,6-dicarboxylic acid [11]



^{1}H NMR ($CDCl_3$) δ: 7.24 (s, 2H), 6.30 (s, 2H, COOH)
^{13}C NMR ($CDCl_3$) δ: 170.9, 163.9, 151.9, 151.4, 129.1, 114.0

[^{1}H,^{13}C] HSQC NMR δ:
1: 7.2/114.0

E. References

1 I. van Zandvoort, Y. Wang, C. B. Rasrendra, E. R. H. van Eck, P. C. A. Bruijnincx, H. J. Heeres and B. M. Weckhuysen, *ChemSusChem*, 2013, **6**, 1745–58.

2 N. Baccile, G. Laurent, F. Babonneau, F. Fayon, M.-M. Titirici and M. Antonietti, *The Journal of Physical Chemistry C*, 2009, **113**, 9644–9654.

3 V. Maruani, S. Narayanin-Richenapin, E. Framery and B. Andrioletti, *ACS Sustain Chem Eng*, 2018, **6**, 13487–13493.

4 National Institute of Advanced Industrial Science and Technology, Spectral Database for Organic compounds, https://sdbs.db.aist.go.jp, (accessed 15 January 2024).

5 M. Mascal and E. B. Nikitin, *Angewandte Chemie International Edition*, 2008, **47**, 7924–7926.

6 R. I. Khusnutdinov, A. R. Baiguzina, A. A. Smirnov, R. R. Mukminov and U. M. Dzhemilev, *Russian Journal of Applied Chemistry*, 2007, **80**, 1687–1690.

7 H. Shen, H. Shan and L. Liu, *ChemSusChem*, 2020, **13**, 513–519.

8 S. Qin, T. Li, M. Zhang, H. Liu, X. Yang, N. Rong, J. Jiang, Y. Wang, H. Zhang and W. Yang, *Green Chem.*, 2019, **21**, 6326–6334.

9 S. Yang, Y. Hao, J. Wang, H. Wang, Y. Zheng, H. Tian, Y. Liu and B. Sun, *Sci Rep*, 2017, **7**, 12954.

10 S. A. Ralph and L. L. Landucci, NMR Database of Lignin and Cell Wall Model Compounds, https://www.glbrc.org/databases_and_software/nmrdatabase/NMR_DataBase_2009_Complete.pdf, (accessed 15 January 2024).

11 T. Rosenau, A. Potthast, N. S. Zwirchmayr, H. Hettegger, F. Plasser, T. Hosoya, M. Bacher, K. Krainz and T. Dietz, *Cellulose*, 2017, **24**, 3671–3687.