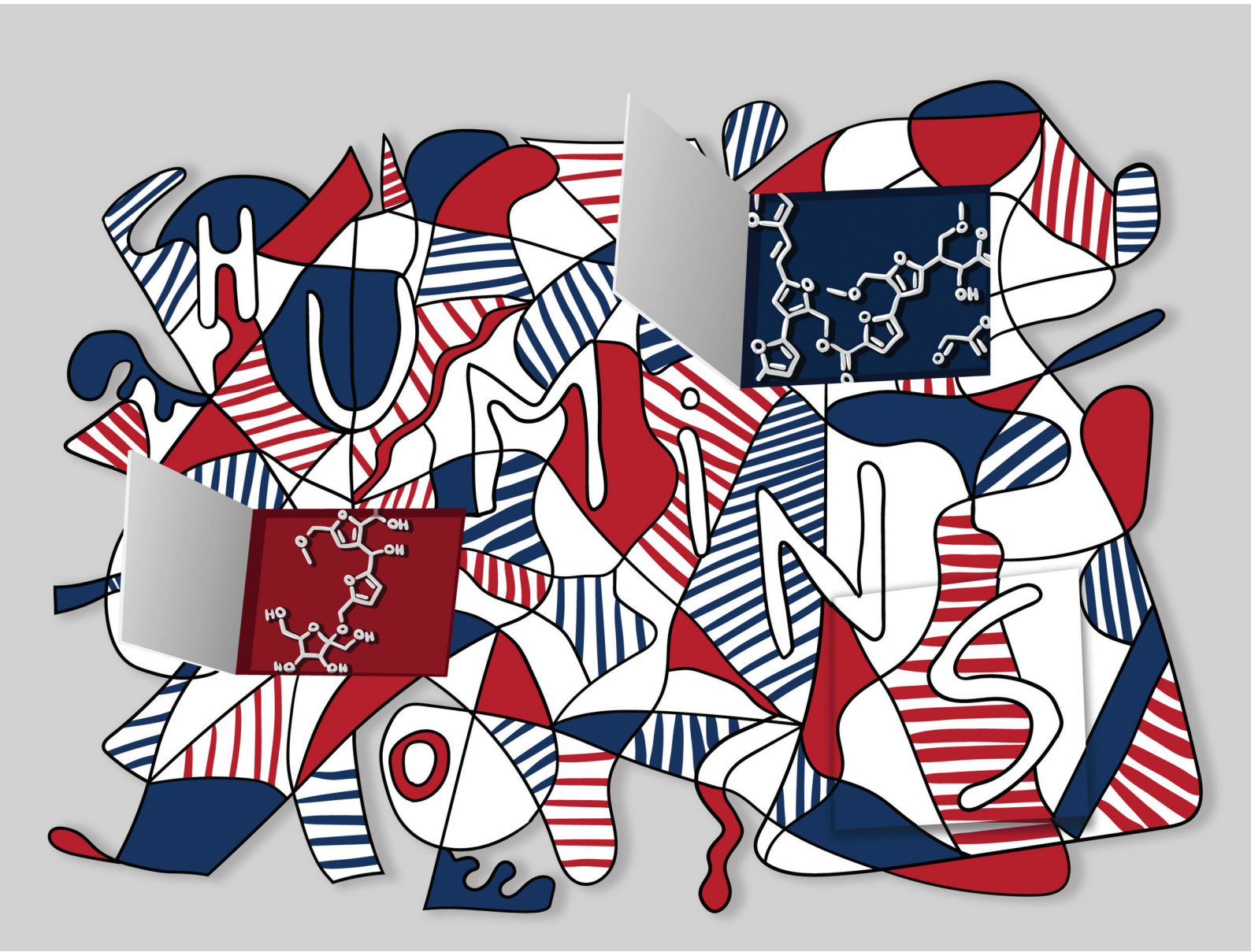
Volume 26
Number 13
7 July 2024
Pages 7429-8040

Green Chemistry

Cutting-edge research for a greener sustainable future

rsc.li/greenchem



ISSN 1463-9262

ROYAL SOCIETY OF CHEMISTRY

PAPER

Daan S. van Es, Pieter C. A. Bruijnincx *et al.*
Molecular structure and composition elucidation of an industrial humin and its fractions

25 YEARS ANNIVERSARY

Green Chemistry

PAPER

View Article Online
View Journal | View Issue

Cite this: *Green Chem.*, 2024, **26**, 7739

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]

Humins, (side-)products of the acid-catalysed dehydration of carbohydrates, will be produced in substantial quantities with the development of industrial biorefining processes. Most structural knowledge about such humins is based on synthetic model humins prepared at lab-scale from typical carbohydrate (-derived) compounds. Here, we report the first extensive characterisation study of an industrial humin. The soluble humin was generated from pilot plant-scale methanolic cyclodehydration of D-fructose to 5-methoxymethyl-2-furfural (MMF), as part of the Avantium YXY® process to produce FDCA. Purification of the industrial humin followed by fractionation allowed isolation of a water-insoluble, high molecular weight fraction (WIPIH) and a water-soluble, low-to-middle molecular weight soluble fraction (WES). Characterisation by elemental analysis, thermogravimetry, IR and NMR spectroscopy and size exclusion chromatography provided a detailed picture of the humin structure in both fractions. Aided by a comprehensive NMR spectral library of furanic model compounds, we identified the main furanic building blocks and inter-unit linkages and propose a structure for this industrial humin sample. The WIPIH and WES fractions were found to be composed of furanic rings interconnected by short aliphatic chains containing a wide range of functionalities including alcohols, ethers, carboxylic acids, esters, aldehydes and ketones. The low level of crosslinking and high functional group content of the industrial humin differ from the more extensively studied, (highly over-)condensed synthetic model humins, towards which they can be considered intermediates. The structural and compositional insights into the nature of an actual industrial humin open up a broad spectrum of valorisation opportunities.

Received 24th January 2024,
Accepted 21st March 2024

DOI: 10.1039/d4gc00429a

rsc.li/greenchem

Introduction

Lignocellulosic biomass feeds are being actively investigated as possible alternatives to fossil feedstocks to produce bio-based materials, fuels and chemicals.[1] Such biomass valorisation efforts need industrial biorefineries in which several integrated biomass processing and upgrading operations are combined. In a typically adopted set-up, the major components of the lignocellulosic biomass are first separated and subsequently upgraded individually. Downstream processing of the carbohydrate fraction, for example, often involves acid-catalysed hydrothermal processes to produce key renewable platform molecules, such as furfural, 5-hydroxymethyl-2-furfural (HMF) and levulinic acid (LA).[2–6] Unfortunately, hydrothermal treatment of sugars under acidic conditions almost inevitably leads to the formation of ill-defined, dark-coloured polymeric side-products called humins.[7,8] These materials are, therefore, anticipated to be produced in considerable amounts in future biorefineries,[9] such as in Avantium's FDCA plant for renewable polymers that is currently being built.[10] To improve the econ-

[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. E-mail: p.c.a.bruijnincx@uu.nl*
[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. E-mail: 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*

†Electronic supplementary information (ESI) available: Material and methods, supplemental experimental, EA, GPC, ss-NMR, SEM, TGA, FT-IR, ^{31}P NMR, ^{19}F NMR, other solution state NMR data and database of NMR spectra. See DOI: https://doi.org/10.1039/d4gc00429a

omic viability of such biorefineries and to adhere to the no-waste axiom of sustainable and circular chemistry, full valorisation of all carbon in the feed is essential. Thus, the formation of these humins either needs to be avoided, or, when partial loss of carbon to such a side stream is unavoidable, valorisation routes need to be found that go beyond low-value solutions such as incineration.[10] For both scenarios, insight into the chemical structure of these humins is essential. Several valorisation approaches have been investigated for humins, aimed at materials applications,[11–22,73] solar cell dyes,[23] hydrogen production[24] and catalytic/pyrolysis conversion to humins oils.[25–28] Humins valorisation is hampered, however, by the fact that the structure of humins is highly complex, highly heterogeneous and not yet unequivocally established. Importantly, and in close analogy with that other recalcitrant, structurally complex biomass feed, *i.e.* lignin, a humin structure will be highly dependent on the applied biomass pretreatment and downstream processing conditions. As with industrial lignins, efforts clearly need to be invested in structure elucidation of industrial humins.

Most structural studies on carbohydrate-derived humins have been performed on synthetic humins prepared at lab scale under conditions that aim to mimic different biorefinery carbohydrate processing conditions. Humins formation has been investigated from various (combinations of) carbohydrate and carbohydrate-derived feedstocks including glucose, fructose, maltose, xylose, and HMF, among many others.[29,30] Contrary to the extensively studied and related hydrothermal carbons (HTC), these synthetic humins are generated in presence of an acid catalyst, classically H_2SO_4 in concentrations up to 1 M. In addition, several reaction media such as water, methanol or DMSO and a broad range of reaction temperatures (4 to 300 °C) have been investigated. These differences in reaction severity, reaction media, residence time, scale and carbohydrate feedstock will obviously affect the chemical structure of the humins, in particular their degree of condensation, and therefore their physical properties such as thermal stability, morphology and solubility. These synthetic humins are often rather insoluble, thus limiting structure characterisation efforts to the use of solid state analytical techniques. The analytical approaches used for the investigation of HTC[31–34] and the structures proposed for these materials obviously served as inspiration for the structural studies on humins.[35]

This way, considerable insights into the structural motifs present in synthetic humins were obtained, *e.g.* by FT-IR spectroscopy and solid state NMR studies. Some of us, for example, have reported on the chemical structure of a series of insoluble, carbohydrate-derived humins based on SEM, FT-IR, NMR and pyro-GC-MS analysis.[29] This multi-technique approach revealed those humins to be composed of furanic rings with aliphatic chains as linkers and containing alcohol, acid, and carbonyl functional groups. Later, furan coupling was further investigated by detailed 1D and 2D solid state NMR experiments on ^{13}C labelled synthetic humins,[36,37] showing that crosslinks between C_α–$C_{aliphatic}$ and C_α–C_α are more abundant than ones between the C_α–C_β and C_β–C_β positions of the furans (Fig. 1). Prior to this, Sumerskii *et al.* reported on humins-like substances, derived from *e.g.* mono-, di-saccharides and furanic compounds,[38] to consist of 60% furan rings and 20% aliphatic fragments based on ^{13}C NMR, FT-IR spectroscopy, and pyro-GC-MS analysis. Based mainly on FT-IR spectroscopy, Lund and co-workers proposed that humins derived from glucose, fructose and HMF were formed *via* aldol addition/condensation polymerisation of 2,5-dioxo-6-hydroxyl-hexanal (DHH) and HMF.[39,40] An idealised structural model was proposed to consist of a conjugated network of C=C bonds with several functional groups such as alcohols, aldehydes, and ketones. The extent of incorporation of furan rings in this structure was observed to be dependent on the starting reactant and the HMF prevailing in the reaction media.[39] Tsilomelekis *et al.*[41] also reported on the molecular structure and growth mechanism of humins derived from HMF under acidic conditions. In this case, humins were found to be partially soluble when DMSO was used as co-solvent for sugar dehydration. A multiple parallel reaction network for the formation and agglomeration of oligomeric humins was proposed to involve nucleophilic attack, condensation, addition, and etherification reactions. Later, the Vlachos group studied the multistage solubilisation of (fructose-derived) humins in various solvents and the structure of the obtained fractions.[42] They concluded that the humin structure was spatially and chemically heterogeneous and consisted of insoluble macromolecules and small soluble species that are weakly associated within the structure. An X-ray scattering study on kinetic growth of humins suggested that humins were formed primarily from HMF rather than fructose.[43] Maruani *et al.* also identified water-soluble oligomers (WSO) of D-glucose during the decomposition of D-glucose into humins.[44] It was suggested that the WSO act as D-glucose "reservoir" that can progressively release D-glucose upon acidic hydrolysis of the glycosidic bonds, but can also be incorporated directly in the humin structure through aldol condensation reactions. Shen *et al.* studied the intermediates resulting from HMF-derived humins formed at low temperatures (4 and 20 °C) up to 170 °C.[45]

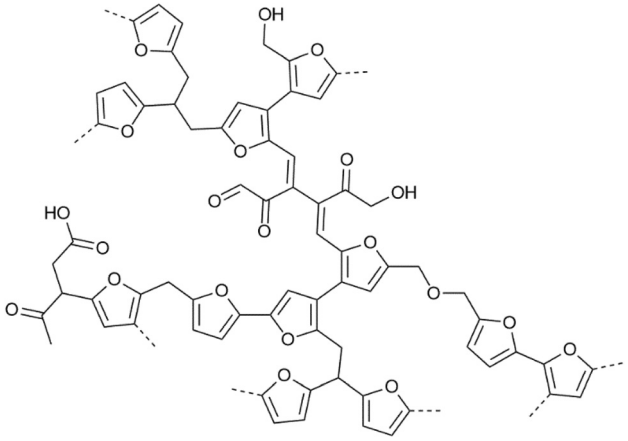


Fig. 1 Model for the molecular structure of a synthetic glucose-derived model humin material.[37]

Several soluble intermediates were identified resulting from aldol condensation between the aldehyde groups of HMF and levulinic acid.

The efforts above thus provided insight into the structure of, essentially, synthetic humin model systems and evidence for the formation of soluble humin-derived oligomers. Here, we describe the structure characterisation of a humin material produced in an industrial process, *i.e.* the humin obtained from pilot plant-scale operation of Avantium YXY® technology, by methanolic cyclodehydration of D-fructose to 5-methoxymethyl-2-furfural (MMF) and related derivatives.[46,47] The humin sample, a dark-coloured, highly viscous syrup, proved to be surprisingly soluble in a number of organic solvents, allowing also for powerful solution state analytical tools to be applied for structure characterisation. As the structural complexity of humins is reminiscent of lignins, the multi-technique analytical approach previously developed for industrial lignins[48] is extended here to the industrial humin. 1D and 2D solution state NMR techniques, for example, provided particularly valuable insights into the overall structure and functional group distribution. For instance, the structure of the industrial humin is found to be composed of furanic rings interconnected by short aliphatic chains decorated by a large range of functionalities. The structure elucidation efforts reported here shed light on the (dis)similarities with the synthetic humins studied previously and will provide input to biorefinery solutions aiming at preventing humins formation. Alternatively, it can support the development of new valorisation approaches for humin materials, this in combination with a lowered footprint for the biorefinery's primary product: FDCA.[49] Indeed, when the humins are used for durable applications, *e.g.* in wood durification,[76] the renewable carbon is stored much longer than if burned directly, creating additional environmental benefits.

Results and discussion

In the following section, the purification and fractionation of the Crude Industrial Humin (CIH) into the Water-Extractable Solubles (WES) and Water-Insoluble Purified Industrial Humin (WIPIH) fractions will be described. These fractions were extensively characterised using elemental analysis, thermogravimetry, FT-IR, size exclusion chromatography and NMR spectroscopy. Based on these analyses, structural models for each fraction are proposed for this industrial humin. Finally, the acid-treated purified industrial humin (ATPIH), discussed at the end of the manuscript, illustrates the relationship between the industrial and synthetic humins.

Purification and fractionation of the crude industrial humin (CIH)

The highly viscous CIH was found to still contain considerable amounts of volatile low molecular weight species. Indeed, gel permeation chromatography (GPC) of CIH showed the sample to have a bimodal distribution and high dispersity, with a relatively large amount of low M_W components (Fig. 3 and Table 1). This was further confirmed by thermogravimetric analysis (TGA), which indicated the sample to consist of ~40 wt% of volatile species measured at 220 °C (Fig. S3†). The thermal behaviour of CIH has been previously extensively studied,[19] revealing, based on headspace GC-MS analysis, that volatile components including furanics are released at ~170 °C. This was confirmed here by GC-MS, ^{1}H NMR spectroscopy and HPLC analyses, which showed that the low molecular weight fraction mainly consisted of HMF, MMF, as well as minor amounts of LA and methyl levulinate (ML) and furfural (Fig. S1 and S14†). The amount of free, monomeric reducing sugars, determined by ion chromatography after a hot water wash of CIH, proved to be very minor (0.2 wt%).

Table 1 Molar masses and carbohydrates content of the CIH, PIH, WES and WIPIH humins determined by GPC-UV using THF as eluent with PL-gel mixed-E columns and polystyrenes as standards

	Molar mass			Carbohydrates (wt%)	
Humin sample	M_n (g mol^{-1})	M_w (g mol^{-1})	PDI	Monomeric reducing sugars	All carbohydrates
CIH	220	680	3.1	0.2	1.1
PIH	400	1100	2.8	0.3	1.4
WES	300	470	1.5	1.1	2.2
WIPIH	970	2250	2.3	0.0	0.1

As the focus of this study was to investigate the chemical structure of the oligomeric/polymeric components of the industrial humins, rigorous purification of the CIH material was required to remove the physically embedded low M_W compounds from the sample. Initial purification attempts by bulk vacuum distillation and bulb-to-bulb short path distillation did effectively remove the volatile fractions, but the prolonged exposure to elevated temperatures resulted in severe sample degradation, giving black insoluble residues reminiscent of the synthetically prepared solid model humins. To preserve the molecular structure of the humin and prevent unwanted condensation reactions, purification by selective solvent extraction was adopted instead. A preliminary solvent screening showed that the targeted low M_w furanic components could be selectively extracted with diethyl ether (see ESI†). After optimisation, a two-step protocol was adopted to ultimately give a dark brown, powdered purified industrial humin fraction (PIH, 39.1 wt%), the diethyl ether extractives (DEE, 36.1 wt%); and a non-isolated fraction (24.5 wt%) which included contributions from losses in purification step, water and volatiles (Fig. 2). The PIH sample was still readily soluble in various polar (a)protic solvents, allowing for the use of solution state analytical tools such as GPC and HPLC (Table S4†). As previously observed in the case of lignins and other (bio)polymers, molar mass determination by GPC depends heavily on the experimental set-up used, including column type and eluent, as well as on the calibration standards and calculation methods applied.[48,50,51] The molar mass distribution of CIH and its fractions was investigated by two complementary GPC measurements,[48] using HFIP and THF as eluents (Table S8†).

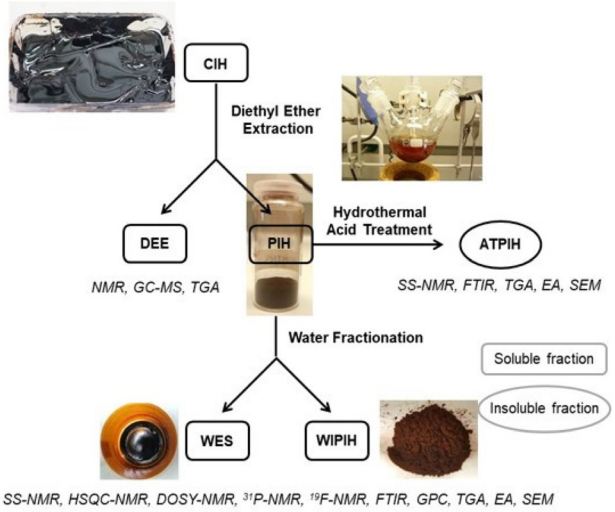


Fig. 2 Flow chart of the preparation of the different fractions of the industrial humin: Crude Industrial Humin (CIH), Diethyl Ether Extractives (DEE), Purified Industrial Humin (PIH), Acid-Treated Purified Industrial Humin (ATPIH), Water-Extractable Solubles (WES) and Water-Insoluble Purified Industrial Humin (WIPIH).

All GPC chromatograms and resulting molar mass values are reported in Table S9 and Fig. S6, S7.† The THF GPC results (Table 1 and Fig. 3) showed that the extraction process was highly effective in removing the furanic monomers from the crude humin. CIH and PIH were found to indeed have different multi-modal distributions. PIH contained a lower proportion of low M_w species than seen for the CIH sample (Fig. 3). The average molar mass of the CIH and PIH were measured to be around 680 and 1100 g mol^{-1} with a dispersity of 3.1 and 2.8 respectively, using polystyrene standards for calibration (Table 1). For comparison, molar masses of soluble macromolecular side-products, formed during the hydro-conversion of glucose and furfural over a NiMo/alumina catalyst, were reported to be in the same range, up to 700 g mol^{-1}, also against polystyrene standards.[52] The GPC data also indicated that highly hydrophilic low M_w components, such as any residual occluded monomeric carbohydrates or other hydrophilic compounds, were not removed by diethyl ether extraction. A second solvent fractionation was therefore necessary to further purify the sample and to avoid any (dominating) contributions of these species to the humin spectra. Conveniently, extraction of PIH with water proved to be most effective in removing the polar, low to middle molecular weight fraction from the polymeric humin sample, giving the water-soluble extractives (WES) and a water-insoluble purified humin fraction (WIPIH). This water-extractable fraction was significant, accounting for 45 wt% of the PIH. This is in clear contrast with synthetic humins, where isolated humins are insoluble in water with very limited extraction options.[29,39] The GPC traces for these two fractions are also reported in Fig. S6 and S7† and molar mass values can be found in Table 1. The GPC chromatogram of WIPIH displayed a broad, intense peak in the high M_w region (M_w 2250, PDI 2.3) with only very little residual intensity in the low M_w region (Fig. 3), showing the efficiency of the second water extraction. GPC data shown in Fig. 3 should be taken with some caution, as the UV detector employed would not detect components such as carbohydrates, as these don't absorb at the applied wavelength. The HFIP GPC results, using a complementary refractive index detector, however, show a similar trend, suggesting that the fractionation indeed efficiently removes lower M_w components from the PIH (Table S9†). The analyses of the total sugars content of the various fractions were performed after an acid hydrolysis procedure similar to the one used for lignin analysis. Sugars were present in the crude industrial humin with 0.2 and 1.1 wt% of reducing and total sugars, respectively (Table 1). The results also showed that after fractionation of the PIH, most sugars are expectedly found in the WES fraction, with only 0.1 wt% of sugars remaining in the WIPIH. TGA showed WIPIH to be more thermally stable than PIH and CIH, with $T_{10\%}$ of 232 °C *vs.* 193 and 148 °C, respectively. The removal of low M_w components in WIPIH thus increased the thermal stability of the humin sample. The extraction steps with diethyl ether followed by the water fractionation thus conveniently allowed for the polymeric fraction of the industrial humin to be isolated and purified. A further characterisation of the WIPIH and WES was then conducted to better understand the polymeric structure of the humin.

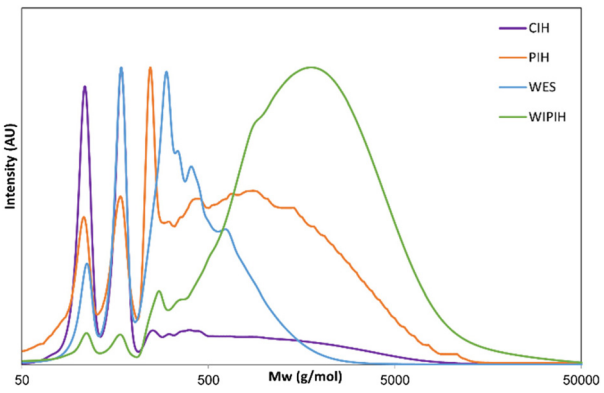


Fig. 3 GPC chromatograms of the different humin fractions measured with the THF method with UV detection.

Characterisation of the water-insoluble fraction of the purified industrial humin (WIPIH)

The WIPIH fraction was soluble in various organic solvents, allowing its structure to be investigated not only by the solid state techniques typically used for synthetic humins (SEM, elemental analysis, solid state NMR, FT-IR, *etc.*), but also with higher-resolution solution state 1D and 2D NMR. First, SEM analysis of WIPIH (Fig. S5†), showed a morphology of agglomerated irregular particles of 10–100 μm, with a few particles appearing as porous materials. This is clearly different from the typically spherical synthetic humins particles.[29,40] Elemental analysis showed that the WIPIH fraction is composed of 63% C, 5% H and 32% O (Table S7†). A Van Krevelen

plot (Fig. S2†) shows the WIPIH fraction to be an intermediate between fructose, HMF, MMF and synthetic humins.[29,31,53] The higher H/C ratio could indeed indicate a significantly lower degree of crosslinking of the humin, probably combined with the presence of methoxy groups due to the use of methanol in the sugar dehydration process.

FT-IR spectroscopy. The FT-IR analysis of the WIPIH fraction provided valuable insights into the general structure and specific functional groups of the humin (Fig. S8–S11 and Table S10†). The furanic skeleton of the humins, previously reported in literature,[29,39,41] was confirmed by the bands observed at 3130, 1600, and 1518 cm^{-1}; the latter was also observed by Pin *et al.* for a very similar humin (at 1515 cm^{-1}) and attributed to the stretching vibration of a furan C=C bond.[12] Furthermore, functional groups such as aldehydes, ketones, carboxylic acids, esters, ethers, acetals and alcohols could be identified in the FT-IR spectrum. The carbonyl band at 1710 cm^{-1} could be ascribed to any type of C=O ketones, aldehydes, and carboxylic acids. The ester band at 1745 cm^{-1} was much less intense than the one associated with ketone and acid functionalities. The band at 1670 cm^{-1} was also previously observed in HMF-derived humins[12,41] and could correspond to the C=O stretch of the HMF-like and MMF-like aldehyde groups. Cheng *et al.* suggested that HMF actively participates in the growth of the humins and HMF-like units could be physisorbed to the core humins.[42] Interestingly, the band at 1200 cm^{-1}, previously ascribed to C–O stretch related vibrations of an sp^2-hybridised OH group such as phenol was also present in the WIPIH sample.[29] The tentative assignment of a vibration at 1020 and 1082 cm^{-1} to acetals was ultimately rejected as no corresponding signals were seen in the HSQC NMR (see below). Finally, evidence of the presence of methyl groups can be observed with the quite intense band at 2983 cm^{-1}. Another strong band related to methyl groups was observed at 1365 cm^{-1} and could be assigned to a CH_3 next to carbonyl functions like in LA/ML units.

Solid state NMR spectroscopy. The solid state cross polarisation (CP) MAS ^{13}C NMR spectrum of the WIPIH fraction is shown in Fig. 4. The CP conditions do not allow signal quantification but lead to significantly reduced experimental measurement times. The CP contact time was optimised to observe all the different types of carbon, from primary to quaternary. For comparison, the CP ^{13}C spectrum of PIH is reported in Fig. S13.† Peak assignments, based on previous studies,[26,31,34,37,54] are reported in Table S11.† Four sharp peaks were observed at 57, 65, 110 and 150 ppm and, when considered together, are thought to indicate HMF/MMF-derived structures. The shoulder observed at 120 ppm could correspond to Cβ next to Cα substituted by a carbonyl group also present in the HMF/MMF-derived structures. Both differences and similarities with previously reported spectra of synthetic humins can be observed for each region of the spectrum. Notably, besides these four peaks that dominate the spectrum, four other peaks are observed in the aliphatic region at 15, 25, 30 and 39 ppm corresponding to saturated primary, secondary and tertiary carbons. These aliphatic

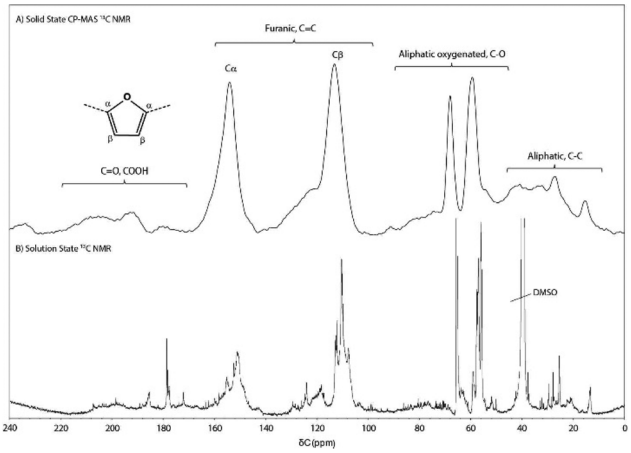


Fig. 4 Solid state CP ^{13}C NMR (A) and solution state ^{13}C NMR (B) spectra of the WIPIH fraction.

signals are quite similar to those observed in the aliphatic spectral regions of previous synthetic humins and HTC.[26,29,31,37,52] In the aliphatic oxygenated region, an intense multicomponent peak is seen, corresponding to various alcohol, and ether contributions. As noted above, the peaks at 57 and 65 ppm are ascribed to methoxy- and secondary alcohol groups in HMF/MMF/hydroxyacetyl furan-derived structures,[55] for example. The broad peak observed at 78 ppm can be assigned to secondary alcohols in carbohydrate derivatives and is clearly less intense in WIPIH compared to the parent PIH (Fig. S13†). This shows that, while some carbohydrate derivatives can be removed, oxygenated aliphatic components constitute a considerable part of the WIPIH structure. In fact, after purification and fractionation, the aliphatic oxygenated region of WIPIH shows a much more intense signal for alcohol and ether functions than the previously studied humins and HTC.[26,29,31,37,52]

The aromatic region of the spectrum is essentially composed of two sharp peaks at 110 and 150 ppm, respectively ascribed to the β-carbon of furan and linked α-carbon atoms. Interestingly, the peak around 130 ppm, previously observed in humins and HTC samples,[26,29,31,37,52] corresponding to conjugated structures, is very weak and ill-defined in the WIPIH spectrum. Again, this suggests a (much) lower degree of condensation of the industrial humin. Finally, in the C=O region, weak signals were observed around 177 and 200 ppm corresponding to carboxylic acid/ester and aldehyde/ketone functions.[37] It should be noted that the aldehyde carbon signal in HMF/MMF(-derivatives), at 178 ppm, would overlap with any carboxylic acid/ester peaks. Unfortunately, the weak intensity signal in this region did not allow further distinction between aldehydes or ketones. More insights on the carbonyl structures present were obtained, however, by solution state ^{19}F NMR spectroscopy after derivatisation as described below.

Solution state NMR spectroscopy. Being soluble in common NMR solvents, WIPIH was further analysed by 1D (^1H, ^{13}C, ^{31}P and ^{19}F after labelling) and [^1H,^{13}C] HSQC solution state NMR

spectroscopy. Very few reports on NMR analysis of soluble humins precursors, and as such on chemical shift assignment for furanic compounds, were available in the literature at the beginning of this study. Wassenberg *et al.*,[56] Tarabanko *et al.*,[57] Maruani *et al.*,[53] Nguyen *et al.*[58] and Shen *et al.*[45] reported solution state 1D NMR analysis of carbohydrate-derived product mixtures and intermediates towards humins. Following an approach similar to lignin structure elucidation efforts, we therefore compiled a database of NMR spectra of compounds representing possible substructures of the humin, both from recorded spectra and data found in literature[45,59–63] (see section D in ESI†). This database then served to assign the 1D and 2D humin NMR spectra.

^{1}H and ^{13}C NMR spectroscopy. The ^{1}H NMR spectrum of WIPIH showed broad signals in the aliphatic, aliphatic oxygenated and aromatic regions (Fig. S17†). As a comparison, the NMR spectra of intermediates isolated by Wassenberg *et al.*[56] were composed of sharper and fewer peaks compared to our humin spectra, showing that the industrial humin structure is more polymeric and complex. Interestingly, the large peak at δ 3.2 ppm confirmed the presence of methoxy groups, in line with the use of methanol in the process in which the humin is formed. The ^{13}C solution state NMR spectrum of WIPIH (Fig. 4) was comparable to the one obtained using solid state NMR. Sharp peaks that can be found on top of the broad ones could be assigned to a number of structural motifs and functionalities. Their presence in the spectrum can be explained by signals resulting from end groups that still can rotate or end groups part of smaller fragments, for example. In the aliphatic region, they are originating from multiple methyl groups; the broad signal at δ 13 ppm points to a large variety of methyl groups in the polymeric structure, *e.g.* methyls in α position of various furanic rings and the signal at δ 26 ppm in turn can be attributed to methyl ketone group. In the aliphatic oxygenated region, most peak intensity is seen in the δ 60–70 ppm region rather than at around δ 80 ppm, suggesting that the humins contain more primary and secondary alcohols than tertiary ones. In the carbonyl and carboxylic acid/ester region, the specific peaks detected can be attributed to, amongst others, methyl esters of aliphatic units, conjugated methyl ketones, aliphatic ketones, ketones conjugated on both sides, aldehydes, and carboxylic acid in furoic and allyl units (Table S12†). Finally, the absence of sharp peaks at chemical shifts higher than δ 208 ppm suggests that long chain aliphatic and cyclic ketones are likely not present in the humin structure.

[^{1}H–^{13}C] HSQC NMR spectroscopy. Even though the ^{13}C spectra gave some insight into the composition and possible linkages in WIPIH, especially the ones involving tertiary and quaternary carbons, more insight could be obtained from 2D HSQC NMR analysis (Fig. 5). Similar to lignins,[48] the HSQC spectrum of the WIPIH was composed of regions of broad signals confirming the heterogeneous polymeric structure of the humin. Aided by the compiled database, we could assign various cross peaks in the spectrum, see Table S13.† It should be noted that these assignments in the 2D spectrum can be biased due to the overlap of some cross peaks. Again, similar to lignin HSQC spectra, four regions of interest are distinguished: the aliphatic side chain region (δ_H/δ_C 0–5/0–45 ppm), aliphatic oxygenated side chain region (δ_H/δ_C 2.5–6.5/45–95 ppm), aromatic (furanic)/unsaturated region (δ_H/δ_C 5–8.5/95–135 ppm), and an end-groups region (δ_H/δ_C 7–10/140–190 ppm). In the aliphatic region, four different cross-correlation peaks corresponding to methyl groups were observed at 2.2/13, 1.4/22, 2.3/25 and 2.1/30 ppm. Interestingly, among these, a strong signal corresponding to methylated furanic units in position 5 (**A**) could result from methylation reaction of the furanic ring in presence of methanol (as deoxygenation is less likely). The very broad peak centred at 1.4/22 ppm could originate from a methyl group next to a secondary alcohol (**C**). Interestingly, the presence of the signal at 2.6/37 ppm, assigned to a CH_2 next to a ketone (**D**), could indicate the insertion of aliphatic units in the humin side chain structure. This assignment can be confirmed by the presence of an intense negative peak observed at 37 ppm on the ^{13}C DEPT NMR spectrum of PIH (Fig. S22†). In addition, a broad peak at 2.4/28 ppm can be assigned to a CH_2 next to a carboxylic acid or methyl ester in LA/ML-like structures (**F**). Shen *et al.* identified a similar structure by NMR in a dimer resulting from aldol condensation of HMF with LA.[45] This structure is also in agreement with a mechanism proposed by Van Zandvoort *et al.*[29] and differs from Lund's hypothesis where reactions with LA are proposed not to occur.[39,40] Regarding furan alkyl carboxylic acid/ester units (furan–CH_2–COOR), no clear evidence of their presence can be found for in the HSQC NMR spectrum; CH_2 between the furanic ring and the carboxylic/ester function should indeed give a signal in the range 3.3–3.6/30–35 ppm. In this region, the signals observed were weak and broad and thus precluded an assignment without further investigations. The presence of ester units can again be confirmed by the cross-correlation peak at 5.2/57.7 ppm[45,60] corresponding to furyl methoxy units (**I**). HMF-esters of formate may not be part of these, however, as no peak was observed at 8.1/160.1 ppm.[45] Similarly, the absence of a cross peak at around 3.8–4.2/27 ppm[60,64] indicates that CH_2's in bis-furylmethane-like units (furan–CH_2–furan) are also likely not part of the structure.

In the aliphatic oxygenated region, several ether- and alcohol-containing substructures could be identified in the structure of WIPIH. The main ether species identified are methoxy groups with a broad signal centred at 3.1/57 ppm, while the broad signal at 3.5/52 ppm corresponds to methoxy esters. MMF-like structures (**G**) were identified by the broad signal centred at 4.3/65 ppm. Shen *et al.* previously identified the 5,5′-oxy(bis-methylene)-2-furaldehyde (OBMF), dimer of HMF, formed through self-etherification,[45] and such structures could indeed contribute to the broad peak seen for these **G** units. Relatedly, a signal corresponding to the analogous HMF-like structures (**H**) was observed at 4.4/56 ppm. The other peaks observed in the aliphatic oxygenated region correspond to aliphatic primary and secondary hydroxyl/ether type structural moieties but could not be more specifically assigned. The broad cross-correlation peak centered at 5.1/65 ppm could, for

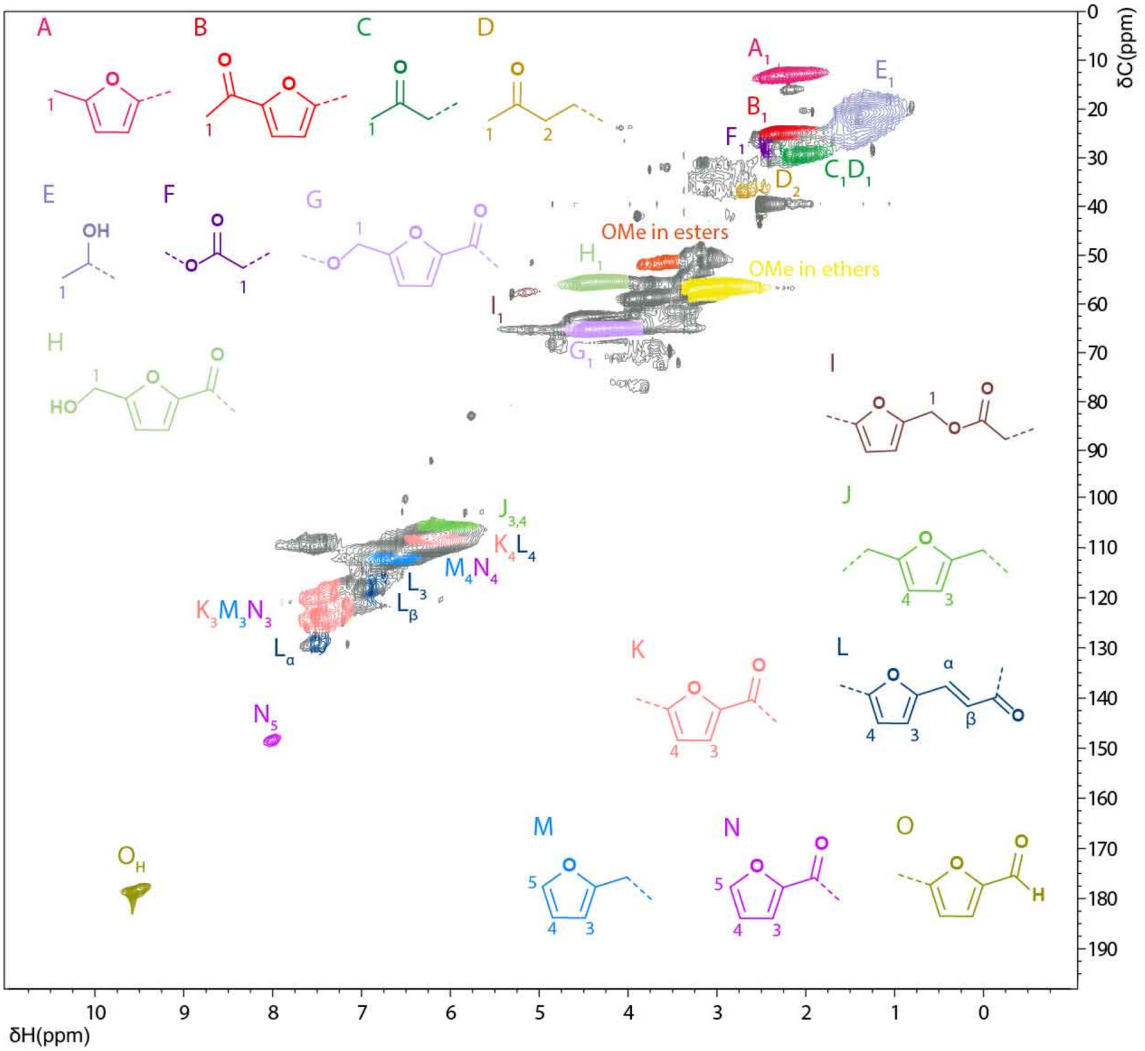


Fig. 5 Solution state [^{1}H; ^{13}C] HSQC NMR spectrum of the WIPIH fraction in DMSO-d_6 and main structural motifs identified.

example, possibly include contributions from 2-hydroxyacetylfuran (HAF)-like structures.[62] Remarkably, no evidence was seen for acetals, often included in several humin structure proposals,[45,65] in the HSQC spectrum of WIPIH (expected in the δ_H/δ_C 4.3–5.5/95–105 ppm region).

In the aromatic region, mainly 2,5-substituted furanics were identified with the furanic units being connected to aliphatic linkers (**J**), conjugated with an allyl ketone (**L**) or substituted with a carbonyl (**K**). Chemical shifts corresponding to furanic end-groups were observed at 7.6/143 and 8.0/148 ppm, assigned here to **M** and **N** type units (Fig. S25†). Interestingly, no furanic end groups conjugated with an alkene, for which the Cβ would be expected at 6.2/115 ppm, were found. Besides furanic end groups, 5-substituted furfuraldehyde units were also observed at 9.5/178 ppm (**O**). It should also be noted that no signal corresponding to aliphatic aldehydes was observed in the humin spectrum (expected at δ_H/δ_C 9.7/202 ppm). Some benzene-derived units can in principle also be expected, *e.g.* as a result of aldol condensation reactions followed by ring closure, which have been previously proposed to occur during humins formation processes.[66] Furthermore, pseudo-lignin, a polymeric material formed from carbohydrate degradation reactions with HMF and furfural as intermediates has been shown to contain 1,3,5-trisubstituted arenes in its structure.[67,68] Unfortunately, our NMR study could not unequivocally confirm their presence or absence. First, cross-correlation peaks corresponding to mono and 1,4-disubstituted phenoxy units were not observed. For example, the positions 2 and 6 of the benzene should give a signal at around 7.2–7.3/125–127 ppm or 7.9/132 ppm, depending on the substituents of the ring.[69] Moreover, no evidence of 1,3,4,5-tetrasubstituted benzene units was found either. However, the cross-correlation peak centred at 7.5/110 ppm may correspond to 1,3,4-trisubstituted benzene units. Further investigations are required to substantiate such assignments.

^{31}P and ^{19}F NMR spectroscopy. Functional group analysis of the WIPIH fraction by ^{31}P and ^{19}F NMR after derivatisation allowed for the hydroxyl, carboxylic acid and carbonyl contents

Table 2 Composition of PIH, WES and WIPIH in terms of functional groups. The carbonyl contents were quantified by ^{19}F NMR whereas the hydroxyl and carboxylic acid contents were determined by ^{31}P NMR both after derivatisation

Content of functional groups (mmol g^{-1} of humin)	PIH	WES	WIPIH
Aliphatic OH	2.75	3.33	1.38
Conjugated OH	1.06	0.63	1.09
Total OH	3.81	3.96	2.47
COOH	0.68	0.98	0.23
COOH (free)	0.18	0.30	0.02
Total carbonyl	2.35	n.d.	1.62

to be determined (Table 2). The quantification of OH and NH groups by derivatisation with 2,2-chloro-4,4,5,5-tetramethyl-1,3,2-dioxaphospholane (CTMP) is a standard characterisation tool for lignin analysis in particular.[70–72] This method was applied on the (fractionated) industrial humin samples and a ^{31}P NMR database similar to the HSQC one was built with furanic and sugar derivatives to assign the regions of the humin spectra (Table S14†). Three distinct chemical shift ranges can be observed correlating with different types of hydroxyl functionalities: 150–144 ppm for aliphatic hydroxyls, 144–138 ppm for aromatic/furanic and conjugated hydroxyls and 137–133 ppm for carboxylic acids. Interestingly, aldehyde functions were also found to react with CTMP and their signals overlap with the aliphatic hydroxyl region between 151 and 146 ppm. Previously, Archipov *et al.* showed that for a series of (lignin-relevant) aldehydes, ^{31}P signals were observed that were shifted downfield compared to aliphatic hydroxyl groups, probably resulting from a formal base-assisted addition of the P–Cl group to the C=O group.[71] This reactivity was also confirmed by P-labelling of the furan-2-carbaldehyde and HMF model compounds (Table S14†). Thus, the ^{31}P NMR measurements probably overestimate the aliphatic OH content due to the presence of aldehydes in the humin.

The ^{31}P NMR spectra of the humin fractions and their compositions in terms of functional groups are respectively shown in Fig. 6 and Table 2. Fractionation of PIH into the WES and WIPIH fractions showed substantial differences in type, distribution, and quantity of OH functional groups contribution over the two samples. The ^{31}P spectrum of the parent PIH was mainly composed of sharp signals in the aliphatic OH and carboxylic acid regions. ^{31}P analysis of the WIPIH fraction provided a spectrum with lower OH intensities and a concomitant lack of resolution. As expected, the OH group content was lower than for WES, indicative of a higher degree of condensation. The aliphatic OHs were the most abundant hydroxyl groups in the WIPIH. A broad signal, not present in WES, is observed from 144–138 ppm, indicating a significant amount of conjugated OH groups. In lignin analysis, signals in this range are attributed to various types of phenolic groups, which are abundantly present in lignins.[48] However, there is no unequivocal evidence for phenol incorporation in the industrial humin. In any case, no free phenolics are formed in the HMF/MMF reaction from which the industrial humin were

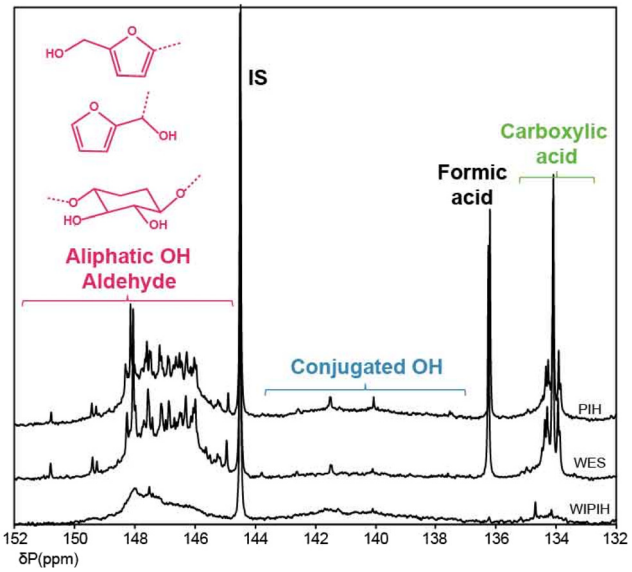


Fig. 6 ^{31}P NMR spectra of derivatised PIH, WES and WIPIH fractions.

obtained. Alternatively, the signals on the ^{31}P spectrum may originate from enols or condensed aromatics which might be formed over time by *e.g.* aldol condensation reactions followed by ring closure.[55,74] Carboxylic acid groups, detected in the 135.5–133.5 ppm range as suggested by 1D NMR, were indeed present, but only in limited amounts (0.23 mmol g^{-1} of humin).

Quantification of the aldehyde and ketone carbonyl functional group contents of the humin was performed by ^{19}F NMR after derivatisation.[75] The ^{19}F NMR spectrum of the WIPIH after derivatisation with 4-(trifluoromethyl)phenylhydrazine and carbonyl quantification are reported in Fig. S28† and Table 2, respectively. The WIPIH spectrum consisted of a multicomponent peak centred at −59.50 ppm with few distinct low intensity peaks at −59.25, −59.33–59.42, −59.57 and −59.71 ppm. This broad peak covered both the conjugated and non-conjugated carbonyls regions. The WIPIH spectrum presented some differences compared with the parent PIH[75] probably due to the removal of water-soluble components by the fractionation step. For instance, the reduced intensity of the sharp peak at −59.50 ppm after fractionation could be explained by the removal of monomeric reducing sugars of PIH reacting with the derivatisation agent.

The quantification of the WIPIH spectrum revealed a total carbonyl group content of about 1.62 mmol per gram of humins, which represents 4.53 wt% of the sample. This would suggest that carbonyls are mainly present as conjugated enones with roughly only 1/5 of the carbonyls being aliphatic.

Characterisation of the water-extractable solubles (WES)

While the focus of this study was on the WIPIH fraction for reasons of application value, the WES fraction was also characterised, although more limitedly. The structure of WES is clearly different from that of WIPIH, having a much lower

average molar mass (470 g mol^{-1}) and different functional group distribution and density. Indeed, the high water-solubility of WES is thought to be due to the high amount of aliphatic OH groups, as indicated by the ^{31}P NMR measurements (1.5× times higher than WIPIH) and the aliphatic oxygenated region of the HSQC spectra (Table 2 and Fig. 7). The sharp signals observed by ^{31}P NMR were indicative of the low molar mass of this fraction as previously noticed by GPC analyses (Fig. 3 and 6). The HSQC spectrum of WES showed sharper and different peaks compared to the WIPIH spectrum, which confirms the presence of low molecular weight components. The most notable differences were observed in the aliphatic region where cross-correlation peaks derived from sugars were predominant. In addition to the free carbohydrates quantified in Table 1, sugar derivatives might then also be covalently incorporated into the structure of WES. Furthermore, acetal peaks, absent in WIPIH, were actually found in WES. This fraction was also found to be rich in carboxylic acids, as evidenced by ^{31}P NMR. Formic acid, a degradation product of HMF, was found by ^{31}P and HSQC NMR. Interestingly, the weak signal observed at δ_H/δ_C 7.2/114.0 ppm could correspond to the CH of the furan ring of a furan benzoquinone-like moiety (**P**), *i.e.* 4,8-dioxofuro[2,3-*f*] benzofuran-2,6-dicarboxylic acid and/or its methoxy form. Such a carboxylic acid compound was previously isolated from degradation products of cellulosic pulping residues by Rosenau *et al.*[62]

Overall, the data thus suggest that the WES fraction consists of much less condensed structures and could be an intermediate to the WIPIH fraction.

Industrial humin structure proposal

Taking together the analyses discussed above, structural models are proposed for each fraction in Fig. 8. These models include identified fragments and inter-unit linkages and thus represent structures that can be expected within the humin. It is important to keep in mind though that, as with technical lignins, no single structure drawing can capture the complexity and heterogeneity of the chemical structure of humins.

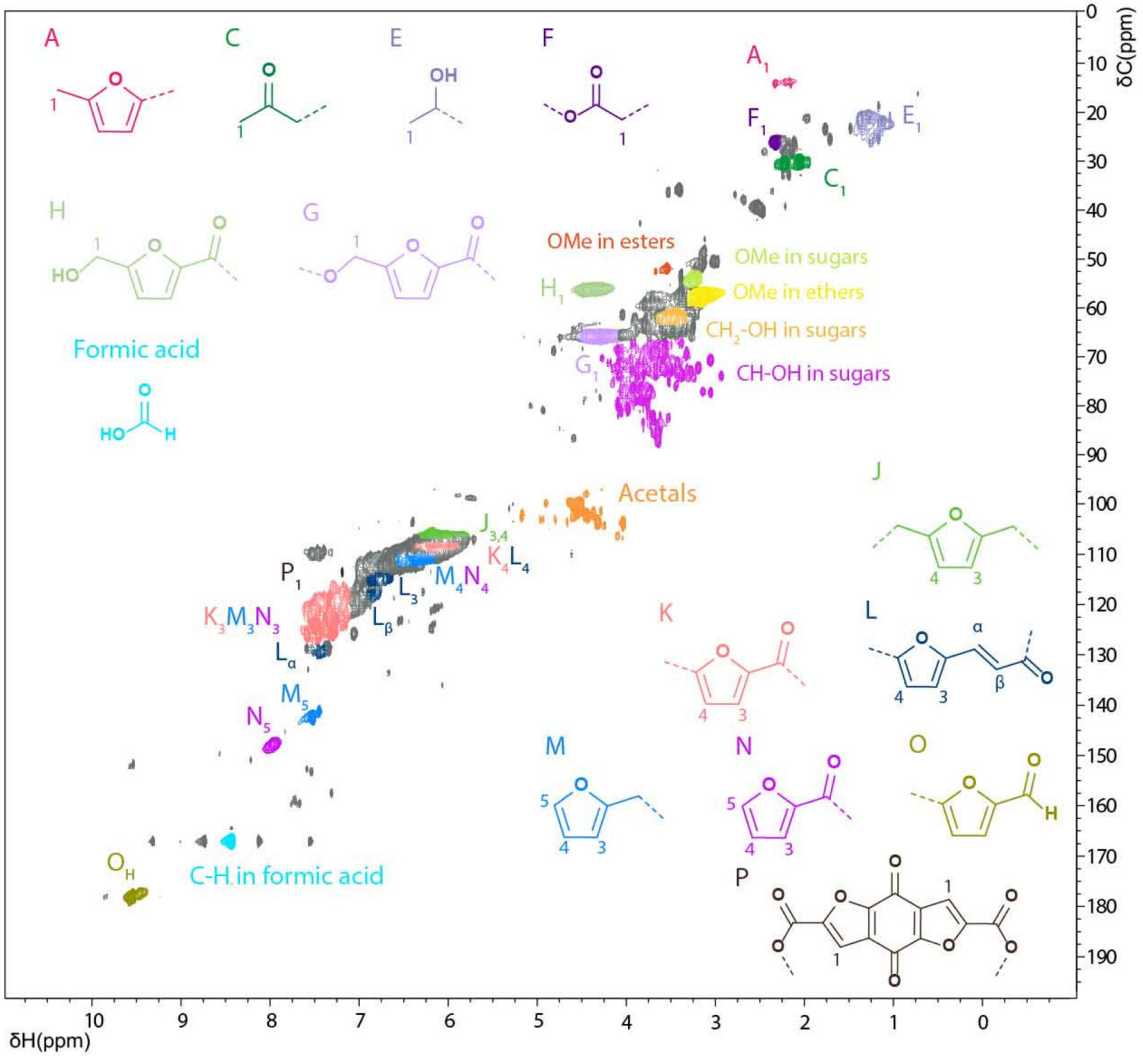


Fig. 7 Solution state [^{1}H; ^{13}C] HSQC NMR spectrum of the WES fraction in DMSO-d_6 and main structural motifs identified.

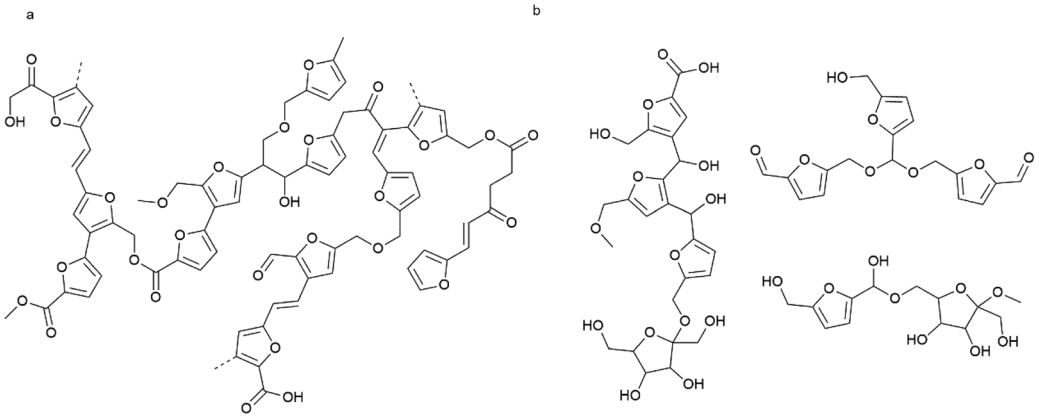


Fig. 8 Revised model for the molecular structure of industrial fructose-derived humin fractions; linkages included in the fragments are based on literature precedence, the previously proposed model[37] and have been refined based on the present work; (a) high molecular weight fraction (WIPIH), (b) low to middle molecular weight fraction (WES).

Further compounding the complexity of structure elucidation, humins might also react like a 'living polymer' with their structure changing over storage time. Shen *et al.* indeed observed self-condensation of HMF in acidic media at room temperature.[45] Furanic units are key constituents of both the WES and WIPIH fractions, but relative abundance and interunit linkages are different. In WIPIH (Fig. 8a), furanic rings are connected with aliphatic linkers *via* ester or ether bonds. Wassenberg *et al.*[56] observed similar linkages for sugar-derived humins obtained from ethanolic reaction mixtures. The linkages originate from aldol condensation, etherification or esterification reactions of the carbonyl, alcohol and carboxylic acid groups present in the humin material and its precursors. No acetal or bisfurylmethane-like units were included in the structure of WIPIH, as we did not find any evidence for them (based on HSQC NMR). This in contrast to WES, where acetals together with ethers and esters seem to constitute the main linkages between the furanic rings. Considering the less condensed, lower M_w structure of WES (470 g mol^{-1}), some oligomers are proposed as components of this fraction (Fig. 8b). Notably, among other fragments, the WES structure is specifically proposed to include residual sugar derivatives, as evidenced by HSQC NMR. Alcohols, ethers, carboxylic acids, esters, aldehydes, and ketones are well represented in both fractions and are contributing to the high solubility of this industrial humin.

The methanolic reaction mixture from which this humin originate leads to the inclusion of methyl, methyl ether and methoxy groups in the structure, which can be considered process-specific features. Methanol incorporation likely inhibited further dehydration/condensation reactions and thus ultimately resulted in a lower crosslinking level in the industrial humin compared to (synthetic) humins formed in aqueous reaction mixtures. Vlachos *et al.* observed a similar trend by suppressing nucleophilic reactions and humins growth with the use of DMSO as solvent.[41]

Compared to previous studies, our proposal is based on a top-down approach *i.e.* combines identified fragments, in contrast to bottom-up approaches where specific mechanistic considerations are leading. Our top-down approach implicitly allows for the complex underlying chemistry and multiple different condensation mechanisms to be accounted for, without excluding others possible mechanisms. In comparison with previous humin structure proposals, the structure proposed here incorporates parts of the structures proposed by van Zandvoort *et al.*[37] (Fig. 1), Shen *et al.*[45] and Wassenberg *et al.*[56] For instance, short aliphatic chains resulting from the incorporation of LA and ML were included together with ether and ester linkages. The short aliphatic chains might also result from DHH incorporation through aldol additions, suggested by Lund *et al.*,[39,40] as conjugated enones were observed by HSQC NMR. Sumerskii *et al.*[38] proposed that humins formation involves nucleophilic additions of a hydroxyl group of HMF with an aldehyde through hemiacetals and acetals. As acetals were only observed in WES, their formation might be the first step of humins genesis, but hydrolysis followed by (irreversible) recondensation through another pathway would result in an expected, gradual disappearance of these linkages upon M_w build up. Another possible mechanism for condensation involves carbocation species, *e.g.* formed from protonated hydroxyl groups, undergoing electrophilic addition with any nucleophilic group to yield oligomers and increasingly condensed humins.[57] Finally, substituted benzene units resulting from aldol condensation reactions followed by ring closure or pseudo-lignin condensations were not included as we could not unequivocally confirm their presence in the industrial humin.

Various valorisation routes have been already explored for humins, from gasification to use in catalyst synthesis, foams, composites, resins, wood durification[76] and self-healing materials.[9,10] The insights gained into the structure of industrial humins can drive the (further) development of these as well as new applications. For instance, the high amount of hydroxyl groups renders the industrial humin a good candidate for applications requiring crosslinking reactions such as

in composite and adhesives applications. In addition, more specific application development of the purified humins or the individual fractions with distinct features and properties can now be envisaged, again similar to lignin fractionation efforts. The WES fraction, soluble in water, for example could be used as replacement of furfuryl alcohol/phenol formaldehyde derivatives during the impregnation of different materials in aqueous environments.[76] The WIPIH fraction, composed of high M_w furanic polymer, could be used as a replacement of existing filler in adhesives or bitumen.

Relation between the industrial and synthetic humins

Given the differences in solubility and chemical fingerprint between the industrial humin studied here and the previously reported synthetic ones, it can be questioned how comparable or related these different humins actually are. Besides the scale of production and process parameters (*i.e.* time, temperature, pressure, catalyst, residence time), one of the main differences in the protocol for the humin formation was the reaction medium. The industrial humin was formed in methanolic solution while the synthetic ones are typically prepared in aqueous media or DMSO. Methanol is expected to contribute to the trapping of reactive intermediate species and hence, together with milder process conditions, to a less condensed humin structure. To study the relationship between the industrial and previously reported synthetic humins, *i.e.* to establish if they are the product of similar chemistry but arrested and isolated at different stages of condensation, the PIH sample was subjected to thermal treatment in aqueous acid at 180 °C, yielding an acid-treated purified humin sample (ATPIH). After treatment, an insoluble black solid was obtained with a spherical microscopic morphology similar to the synthetic humins (Fig. S30†).[29,39] As expected, elemental composition analysis showed that ATPIH to be more condensed than PIH, with O/C and H/C ratios of 0.34 and 0.82 being obtained, respectively. These values are in fact quite similar to those of glucose-derived synthetic humins.[37] The CP-MAS NMR spectrum of ATPIH (Fig. S32†) differed considerably from the WIPIH fraction and parent PIH (Fig. S33 and S13†). Few intense peaks were present in the aliphatic and the aromatic regions, while much less intense signals were observed in the aliphatic oxygenated region compared to the PIH spectrum. Interestingly, a signal around 130 ppm became stronger indicating that the ATPIH consists of condensed structures similar to previous synthetic humins and HTC.[26,29,31,37,52] The FT-IR spectrum (Fig. S31†) also proved very similar to the synthetic humins and HTC, confirming the condensed structure of the ATPIH.[29,40,41] For instance, the intensity of the bands at 1200 and 795 cm^{-1} assigned to the C–O vibrations in alcohols and the C–H out of plane of furan rings, respectively, was strongly decreased after the acid thermal treatment of the PIH. These results indicate that ATPIH is rather similar in structure to the previously studied synthetic humins and thus suggest that the soluble industrial humin can indeed be likely considered as a precursor to the more recalcitrant insoluble humins.

Conclusions

The comprehensive characterisation effort described above provides the first detailed structural insights into an industrial humin. Purification and fractionation reduced the heterogeneity of the sample and provided new insights into the structurally rather different high molecular weight fraction (WIPIH) and the oligomeric fraction (WES). Similar to the efforts in lignin structure elucidation, a multitechnique approach was adopted, with structure identification was aided by a model compound NMR library specifically targeting possible building blocks of the humin structure. Both WIPIH and WES fractions displayed distinctive features in terms of M_w, functionality, structural motifs, and inter-unit linkages. WES oligomers (average molecular weight of 470 g mol^{-1}) were composed of furanic rings linked to aliphatic chains derived from carbohydrate skeletons mainly consisting of hydroxyl, acetal, ether, carbonyl, carboxylic acid and esters functions, in order of abundance. The higher M_w WIPIH fraction (2250 g mol^{-1}) consists of a more condensed structure with furanic rings interconnected with short aliphatic chains *via* ester or ether bonds. The humin structure is proposed to contain methyl, methoxy and methoxy ester groups, but not bismethylfuranes or acetal units. The WES oligomers and WIPIH humin are likely consecutive intermediates to the highly condensed humins previously studied in literature. Indeed, further acid treatment of the industrial sample led to a highly condensed insoluble humin polymer similar to the synthetic humins. Again, analogous to technical lignins, it should be emphasised that any humin structure will heavily depend on the process conditions under which they are generated. Importantly, in this case the presence of methanol during the industrial process is thought to be responsible for the inhibition of more extensive dehydration/condensation reactions, leading to a relatively low level of crosslinking in the humin structure. But, again similar to lignin, the main structural motifs can be generally expected to be present in industrial humin generated in other processes than the one studied here. Further efforts are now needed to refine the structures proposed here and to translate them to other biorefinery humins materials.

Author contributions

Conceptualisation, S.C., C.S.L., W.V., R.K.P., D.V.E., P.C.A.B.; experimentation, S.C., C.S.L., W.V., R.K.P., A.E.F., K.H., P.D.P.; writing – original draft preparation, S.C., D.V.E., P.C.A.B.; writing – review and editing, S.C., C.S.L., R.K.P., A.E.F., K.H., P.D.P., M.B., D.V.E., B.M.W., P.C.A.B.; visualisation, S.C., P.C.A.B.; supervision, D.V.E., P.C.A.B.; funding acquisition, M.B., B.M.W., P.C.A.B. All authors have read and agreed to the published version of the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This project has been performed within the framework of the CatchBio program. The authors gratefully acknowledge the financial support of the Smart Mix Program of the Netherlands Ministry of Economic Affairs and the Netherlands Ministry of Education, Culture and Science. The authors acknowledge contribution of PEFerence project that has received funding from the Bio-based Industries Joint Undertaking (JU) under the European Union's Horizon 2020 research and innovation programme under grant agreement no. 744409. The JU received support from the European Union's Horizon 2020 research and innovation programme and the Bio-based Industries Consortium. We also acknowledge the financial support of uNMR-NL, an NWO-funded National Roadmap Large-Scale Facility for the Netherlands (No. 184.032.207). Dr Hans L. J. Wienk and Prof. Rolf Boelens, are acknowledged for technical support and access to the Utrecht NMR facility. Dr Johann Jastrzebski, Dr Andrei Gurinov, Marjan Versluijs-Helder and Pascal Wijten (Utrecht University) are thanked for technical support with the NMR, SEM and GPC measurements. We also thank Dr Ed de Jong, Dr Gerard van Klink and Tom Claessen (Avantium) for the insightful discussions on the humin material. Finally, we thank Prof. Erik Heeres (Rijksuniversiteit Groningen) for providing the purified synthetic humin sample (PSH).

References

1 H. Stichnothe, G. Bell, H. Jørgensen, I. De Bari, J. Haveren, J. Lindorfer, J. Kepler and E. de Jong, Bio-Based Chemicals A 2020 Update. **https://www.ieabioenergy.com/blog/publications/new-publication-bio-based-chemicals-a-2020-update/**, (accessed 15 January 2024).
2 J. C. Van Der Waal and E. De Jong, in *Industrial Biorenewables: A Practical Viewpoint*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2016, pp. 97–120.
3 R. J. Van Putten, J. C. Van Der Waal, E. De Jong, C. B. Rasrendra, H. J. Heeres and J. G. De Vries, *Chem. Rev.*, 2013, **113**, 1499–1597.
4 A. Mukherjee, M. J. Dumont and V. Raghavan, *Biomass Bioenergy*, 2015, **72**, 143–183.
5 H. Kobayashi and A. Fukuoka, *Green Chem.*, 2013, **15**, 1740–1763.
6 J. Ftouni, A. Muñoz-Murillo, A. Goryachev, J. P. Hofmann, E. J. M. Hensen, L. Lu, C. J. Kiely, P. C. A. Bruijnincx and B. M. Weckhuysen, *ACS Catal.*, 2016, **6**, 5462–5472.
7 M. Mascal and S. Dutta, in *Selective Catalysis for Renewable Feedstocks and Chemicals*, ed. K. M. Nicholas, Springer International Publishing, Norman, USA, 2014, pp. 41–54.
8 P. Bhaumik and P. L. Dhepe, *Catal. Rev. Sci. Eng.*, 2016, **58**, 36–112.
9 E. de Jong, M. Mascal, T. Claessen, P. Tosi, S. Constant and A. Mija, The Origin, Composition, and Applications of Industrial Humins – A Review, to be submitted.
10 E. de Jong, H. (Roy) A. Visser, A. S. Dias, C. Harvey and G.-J. M. Gruter, *Polymers*, 2022, **14**, 943.
11 A. Sangregorio, A. Muralidhara, N. Guigo, G. Marlair, E. de Jong and N. Sbirrazzuoli, *Composites, Part C*, 2021, **4**, 100109.
12 J.-M. Pin, N. Guigo, A. Mija, L. Vincent, N. Sbirrazzuoli, J. C. van der Waal and E. de Jong, *ACS Sustainable Chem. Eng.*, 2014, **2**, 2182–2190.
13 A. Mija, J. C. van der Waal, J. M. Pin, N. Guigo and E. de Jong, *Constr. Build. Mater.*, 2017, **139**, 594–601.
14 R. Dinu and A. Mija, *Green Chem.*, 2019, **21**, 6277–6289.
15 A. Sangregorio, N. Guigo, J. C. van der Waal and N. Sbirrazzuoli, *Compos. Sci. Technol.*, 2019, **171**, 70–77.
16 A. Sangregorio, N. Guigo, E. de Jong and N. Sbirrazzuoli, *Polymers*, 2019, **11**, 1804.
17 P. Tosi, G. P. M. van Klink, A. Celzard, V. Fierro, L. Vincent, E. de Jong and A. Mija, *ChemSusChem*, 2018, **11**, 2797–2809.
18 X. Montané, R. Dinu and A. Mija, *Molecules*, 2019, **24**, 4110.
19 A. Sangregorio, N. Guigo, J. C. van der Waal and N. Sbirrazzuoli, *ChemSusChem*, 2018, **11**, 4246–4255.
20 C. Cantarutti, R. Dinu and A. Mija, *Biomacromolecules*, 2020, **21**, 517–533.
21 P. Tosi, G. P. M. van Klink, C. Hurel, C. Lomenech, A. Celzard, V. Fierro, C. Delgado-Sanchez and A. Mija, *Appl. Mater. Today*, 2020, **20**, 100622.
22 G. P. M. Van Klink, E. de Jong, J. N. Hoogeboom and K. J. Kramer, *EU Pat*, EP3728710A1, 2020.
23 A. C. Mija, J. C. Van der Waal, E. de Jong and G. P. M. Van Klink, *EU Pat*, EP3519480B1, 2017.
24 T. M. C. Hoang, L. Lefferts and K. Seshan, *ChemSusChem*, 2013, **6**, 1651–1658.
25 Y. Wang, S. Agarwal, A. Kloekhorst and H. J. Heeres, *ChemSusChem*, 2016, **9**, 951–961.
26 S. Wang, H. Lin, Y. Zhao, J. Chen and J. Zhou, *J. Anal. Appl. Pyrolysis*, 2016, **118**, 259–266.
27 Y. Wang, S. Agarwal and H. J. Heeres, *ACS Sustainable Chem. Eng.*, 2017, **5**, 469–480.
28 S. Agarwal, D. van Es and H. J. Heeres, *J. Anal. Appl. Pyrolysis*, 2017, **123**, 134–143.
29 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**, 1745–1758.
30 Li Liu, Biomass-Derived Humins, Springer, Singapore, 2023.
31 N. Baccile, G. Laurent, F. Babonneau, F. Fayon, M.-M. Titirici and M. Antonietti, *J. Phys. Chem. C*, 2009, **113**, 9644–9654.
32 M. M. Titirici, M. Antonietti and N. Baccile, *Green Chem.*, 2008, **10**, 1204–1212.
33 M. Sevilla and A. B. Fuertes, *Chem. – Eur. J.*, 2009, **15**, 4195–4203.
34 N. Baccile, C. Falco and M. M. Titirici, *Green Chem.*, 2014, **16**, 4839–4869.
35 S. Liu, Y. Zhu, Y. Liao, H. Wang, Q. Liu, L. Ma and C. Wang, *Appl. Energy Combust. Sci.*, 2022, **10**, 100062.

36 D. Mance, M. Weingarth and M. Baldus, in *Modern Magnetic Resonance*, Springer International Publishing, 2018, pp. 487–503.
37 I. van Zandvoort, E. J. Koers, M. Weingarth, P. C. A. Bruijnincx, M. Baldus and B. M. Weckhuysen, *Green Chem.*, 2015, **17**, 4383–4392.
38 I. V. Sumerskii, S. M. Krutov and M. Ya. Zarubin, *Russ. J. Appl. Chem.*, 2010, **83**, 320–327.
39 S. K. R. Patil, J. Heltzel and C. R. F. Lund, *Energy Fuels*, 2012, **26**, 5281–5293.
40 S. K. R. Patil and C. R. F. Lund, *Energy Fuels*, 2011, **25**, 4745–4755.
41 G. Tsilomelekis, M. J. Orella, Z. Lin, Z. Cheng, W. Zheng, V. Nikolakis and D. G. Vlachos, *Green Chem.*, 2016, **18**, 1983–1993.
42 Z. Cheng, J. L. Everhart, G. Tsilomelekis, V. Nikolakis, B. Saha and D. G. Vlachos, *Green Chem.*, 2018, **20**, 997–1006.
43 Z. Cheng, K. A. Goulas, N. Quiroz Rodriguez, B. Saha and D. G. Vlachos, *Green Chem.*, 2020, **22**, 2301–2309.
44 V. Maruani, S. Narayanin-Richenapin, E. Framery and B. Andrioletti, *ACS Sustainable Chem. Eng.*, 2018, **6**, 13487–13493.
45 H. Shen, H. Shan and L. Liu, *ChemSusChem*, 2020, **13**, 513–519.
46 G. J. M. Gruter and F. Dautzenberg, *EU Pat*, EP1834950A1, 2007.
47 G. J. M. Gruter and F. Dautzenberg, *EU Pat*, EP1834951A1, 2007.
48 S. Constant, H. L. J. Wienk, A. E. Frissen, P. De Peinder, R. Boelens, D. S. Van Es, R. J. H. Grisel, B. M. Weckhuysen, W. J. J. Huijgen, R. J. A. Gosselink and P. C. A. Bruijnincx, *Green Chem.*, 2016, **18**, 2651–2665.
49 A. J. J. E. Eerhart, A. P. C. Faaij and M. K. Patel, *Energy Environ. Sci.*, 2012, **5**, 6407–6422.
50 R. J. A. Gosselink, A. Abächerli, H. Semke, R. Malherbe, P. Käuper, A. Nadif and J. E. G. Van Dam, *Ind. Crops Prod.*, 2004, **19**, 271–281.
51 W. J. J. Huijgen, G. Telysheva, A. Arshanitsa, R. J. A. Gosselink and P. J. de Wild, *Ind. Crops Prod.*, 2014, **59**, 85–95.
52 M. Ozagac, C. Bertino-Ghera, D. Uzio, M. Rivallan, D. Laurenti and C. Geantet, *Biomass Bioenergy*, 2016, **95**, 182–193.
53 V. Maruani, S. Narayanin-Richenapin, E. Framery and B. Andrioletti, *ACS Sustainable Chem. Eng.*, 2018, **6**, 13487–13493.
54 I. Van Zandvoort, E. R. H. Van Eck, P. De Peinder, H. J. Heeres, P. C. A. Bruijnincx and B. M. Weckhuysen, *ACS Sustainable Chem. Eng.*, 2015, **3**, 533–543.
55 R. J. van Putten, J. C. van der Waal, M. Harmse, H. H. van de Bovenkamp, E. de Jong and H. J. Heeres, *ChemSusChem*, 2016, **9**, 1827–1834.
56 A. Wassenberg, T. Esser, M. J. Poller and J. Albert, *Materials*, 2023, **16**, 2864.
57 V. Tarabanko, M. Smirnova, M. Chernyak, A. Kondrasenko and N. Tarabanko, *J. Sib. Fed. Univ. Chem.*, 2015, **8**, 6–18.
58 H. Nguyen, V. Nikolakis and D. G. Vlachos, *ACS Catal.*, 2016, **6**, 1497–1504.
59 S. Yang, Y. Hao, J. Wang, H. Wang, Y. Zheng, H. Tian, Y. Liu and B. Sun, *Sci. Rep.*, 2017, **7**, 12954.
60 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.
61 R. I. Khusnutdinov, A. R. Baiguzina, A. A. Smirnov, R. R. Mukminov and U. M. Dzhemilev, *Russ. J. Appl. Chem.*, 2007, **80**, 1687–1690.
62 M. Mascal and E. B. Nikitin, *Angew. Chem., Int. Ed.*, 2008, **47**, 7924–7926.
63 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.
64 P. Delliere and N. Guigo, *Eur. Polym. J.*, 2023, **187**, 111869.
65 N. Shi, Q. Liu, R. Ju, X. He, Y. Zhang, S. Tang and L. Ma, *ACS Omega*, 2019, **4**, 7330–7343.
66 G. C. A. Luijkx, F. van Rantwijk and H. van Bekkum, *Carbohydr. Res.*, 1993, **242**, 131–139.
67 F. Hu, S. Jung and A. Ragauskas, *Bioresour. Technol.*, 2012, **117**, 7–12.
68 S. D. Shinde, X. Meng, R. Kumar and A. J. Ragauskas, *Green Chem.*, 2018, **20**, 2192–2205.
69 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).
70 D. S. Argyropoulos, H. I. Bolker, C. Heitner and Y. Archipov, *J. Wood Chem. Technol.*, 1991, **11**, 137–157.
71 Y. Archipov, D. S. Argyropoulos, H. Bolker and C. Heitner, *Carbohydr. Res.*, 1991, **220**, 49–61.
72 A. Granata and D. S. Argyropoulos, *J. Agric. Food Chem.*, 1995, **43**, 1538–1544.
73 A. Muralidhara, P. Tosi, A. Mija, N. Sbirrazzuoli, C. Len, V. Engelen, E. de Jong and G. Marlair, *ACS Sustainable Chem. Eng.*, 2018, **6**, 16692–16701.
74 S. Thiyagarajan, H. C. Genuino, J. C. van der Waal, E. de Jong, B. M. Weckhuysen, J. van Haveren, P. C. A. Bruijnincx and D. S. van Es, *Angew. Chem., Int. Ed.*, 2016, **55**, 1368–1371.
75 S. Constant, C. S. Lancefield, B. M. Weckhuysen and P. C. A. Bruijnincx, *ACS Sustainable Chem. Eng.*, 2016, **1**, 965–972.
76 A. Sangregorio, A. Muralidhara, N. Guigo, L. G. Thygesen, G. Marlair, C. Angelici, E. de Jong and N. Sbirrazzuoli, *Green Chem.*, 2020, **22**, 2786–2798.