Effect of pretreatment on the formation of 5-chloromethyl furfural derived from sugarcane bagasse

Joshua Howard[1], Darryn W. Rackemann[1], Zhanying Zhang[1], Lalehvash Moghaddam[1], John P. Bartley[2], William O.S. Doherty[1,*]

1. Centre for Tropical Crops and Biocommodities, Queensland University of Technology, Brisbane, Australia
2. School of Chemistry, Physics and Mechanical Engineering, Queensland University of Technology, Brisbane

*Corresponding author.

Postal address: GPO Box 2432, 2 George St, Brisbane, QLD 4001, Australia

Tel: +61 7 31381245; Fax: +61 7 3138 4132

Email: w.doherty@qut.edu.au

Abstract

Chloromethylfurfural (CMF), a valuable intermediate for the production of chemicals and fuel, can be derived in high yields from the cellulose component of biomass. This study examined the effect of sugar cane bagasse components and biomass architecture on CMF/bio-oil yield using a HCl/dichloroethane biphasic system. The type of pretreatment affected bio-oil yield, as the CMF yield increased with increasing glucan content. CMF yield reached 81.9% with bagasse pretreated by acidified aqueous ionic liquid, which had a glucan content of 81.6%. The lignin content of the biomass was found to significantly reduce CMF yield, which was only 62.3% with acid-catalysed steam exploded sample having a lignin content of 29.6%. The change of CMF yield may be associated with fibre surface changes as a result of pretreatment. The hemicellulose content also impacted negatively on CMF yield. Storage of the bio-oil in chlorinated solvents prevented CMF degradation.

Keywords: Chloromethylfurfural, bio-oil, biphasic system, pretreatment, cellulose, stability

Introduction

In recent years, producing fuels and chemicals from lignocellulosic biomass has received significant research interest. Compared to fossil-based resources such as crude oil and coal, lignocellulosic biomass is sustainable and atmospheric CO_2 via photosynthesis is consumed in its production. This makes it a more environmentally friendly resource.

Generally, there are two processes used to produce fuels from lignocellulosic biomass: one is a biochemical process, in which the biomass is converted to fermentable sugars via saccharification, and subsequently the sugars are converted to fuels such as ethanol and butanol. However, this process is time-consuming because of the long fermentation times that take days. [1] The other approach to produce fuel (such as, bio-oils or hydrocarbons) is through a thermo-chemical process. [2, 3] In this approach, the carbohydrate content of the biomass can be converted to furanics, such as 5-hydroxymethyfurfural (HMF) and furfural [4, 5] which are high-energy organic compounds, which can subsequently be converted to fuel. Other furanics such as 5-chloromethylfurfural (CMF), [6, 7] 5-bromochlorofurfural BMF [8] and ethoxymethylfurfural (EMF) [9] can also be produced from biomass in very high yields. These chemicals are also very useful platform chemicals, apart from being a good resource for subsequent fuel production.

Previous studies on CMF production via biphasic systems have principally focused on the optimization of solvents and processes with different carbohydrate materials including glucose, sucrose, cellulose, corn stover, wood, cotton, etc. [10-12] Mascal and Nikitin reported that CMF yields from lignocellulosic biomass such as corn stover were lower than those from microcrystalline cellulose, glucose and sucrose. [10] On the other hand, Gao *et al.* [11] observed significantly lower CMF yields were obtained for sucrose, glucose and cellulose when compared

to yields obtained from Kraft *Eucalyptus* pulp and *Eucalyptus* wood. There were also differences in the CMF yields between *Eucalyptus* pulp and wood. On the basis of these results, it would be constructive to evaluate the effect the biomass components and biomass architecture on CMF yield. As a consequence, untreated sugarcane bagasse and pretreated sugarcane bagasse samples having different proportions of glucan, xylan and lignin, and structural differences, were evaluated for CMF yield using the biphasic system described by Mascal and Nikitin. [10] The solvents used to pretreat bagasse were NaOH, H_2SO_4, and the ionic liquid, 1-butyl-3-methylimidazolium methylsulfonate (IL, $BMIMCH_3SO_3$,). Two bagasse samples (NaOH-bagasse and IL-bagasse) were prepared in the laboratory, and the other two samples, NaOH treated steam exploded bagasse (SSE-bagasse) and sulfuric acid treated steam exploded bagasse (ASE-bagasse) were produced in a pilot plant having a steam explosion facility.

It is known that CMF darkens on storage indicating that it degrades with time. To monitor its stability commercial CMF, crude and purified bio-oils produced in the present study were stored in a number of solvents and characterized using proton nuclear magnetic resonance (^{1}H-NMR). The information will provide way to best store the oil. The solid residue remaining after acid hydrolysis of the biomass was characterized by solid state NMR, ^{31}P-NMR, Mannich reactivity and elemental analysis in order to identify potential applications. This is because the residue (rich in lignin) constitutes a large proportion of the total biomass, and finding a use for it may improve the economics to produce CMF from lignocellulosic biomass.

Experimental

Chemicals

Furfural, HMF, D-(+)xylose, D-(+)glucose, D-(+)arabinose, n-butanol, and 1,2-dichloroethane (DCE) paraformaldehyde, diethyl amine, dioxane, pyridine, chromium acetylacetonate, cyclohexanol, 2-chloro-4,4,5,5-tetramethyl-1,3,2-dioxaphospholane (TMDP), dimethyl sulfoxide-d_6 (DMSO-d_6), chloroform ($CHCl_3$), and deuterated chloroform ($CDCl_3$), were of analytical grades, while NaOH and $MgSO_4 \cdot 7H_2O$ were reagent grades (Sigma-Aldrich Castle Hill, NSW, Australia). Concentrated HCl (32 wt%), H_2SO_4 (98 wt%) and CH_3COOH (32 wt%) were obtained as reagent grades from Merck (Kilsyth, VIC Australia). Deuterium oxide (D_2O) (99.9 atom% D), $BMIMCH_3SO_3$) (>95%) were obtained from Sigma-Aldrich (Castle Hill, NSW, Australia).

Untreated bagasse

Sugarcane bagasse was collected from Racecourse Sugar Mill (Mackay Sugar Limited) in Mackay, Australia. The bagasse was dried to constant weight at 45 °C. One portion of the dried bagasse was ground by a cutter grinder (Retsch SM100, Retsch GmBH, Germany) and passed through a 2.0 mm aperture screen and the other two portions of the dried bagasse were ground and passed through 0.5 mm and 0.2 mm aperture screens respectively. Therefore, the biomass compositions of these three bagasse samples were the same. The ground bagasse samples were used for CMF production.

NaOH-bagasse

Whole (*i.e.*, unmilled) bagasse was passed through a sieve having an aperture size of 1.0 cm to remove the pith. Depithed bagasse (1 kg dry weight) was delignified with 1.0 M NaOH solution and processed according to the procedure described previously. [13] The pretreated bagasse was air-dried for CMF production. A small portion of air-dried biomass was dried to a constant weight at 100 °C for determination of moisture.

IL-bagasse

Depithed and milled bagasse with an aperture range of 0.25 – 0.5 mm was pretreated with aqueous 1-butyl-3-methylimidazolium methylsulfonate solution containing 20% water (w/w). The detailed pretreatment and post-pretreatment procedures were described previously. [14] The pretreated bagasse was air-dried for CMF production. A small portion of air-dried biomass was dried to a constant weight at 100 °C for determination of moisture.

ASE-bagasse and SSE-bagasse

Sulfuric acid-steam explosion treated sample (ASE-bagasse) and NaOH-steam explosion treated sample (SSE-bagasse) at the pilot-scale in the Mackay Renewable Biocommodities Pilot Plant, Mackay, Queensland, Australia using a two-stage pretreatment reactor designed and constructed by Andritz Inc (Glen Falls, NY, USA). The pretreatment reactor consisted of a first-stage, horizontal hydrolysis reactor (150 L) and a second-stage, vertical reactor (69 L) which performs the steam explosion facility. Sugarcane bagasse (20 kg) was used for each pretreatment experiment. Sulfuric acid steam-explosion was achieved with 3% (wt/dry fibre wt) H_2SO_4 at 170 °C for 15 min, followed by steam impregnation at 185 °C for 5 min and steam-explosion (explosion pressure = 2 MPa). Sodium hydroxide steam-explosion was achieved with 15.5% (wt/dry fibre wt) NaOH at 170 °C for 30 min, followed by steam impregnation at 150 °C for 5 min and steam-explosion (explosion pressure = 2 MPa). Pretreated bagasse samples were washed with distilled water (4 × 1 L) and air-dried. The air-dried biomass was used for CMF production. A small portion of air-dried biomass was dried to a constant weight at 100 °C for determination of moisture.

Biomass compositional analysis

The composition of untreated and treated bagasse samples including the amounts of sugars present and amount of acid insoluble residue, were determined by the average of the two duplicate tests based on the standardized National Renewable Energy Laboratory (NREL) method. [15] The morphology of the bagasse and the pretreated bagasse samples (gold coated) was examined using a FEI Quanta 200 Environmental scanning electron microscope, SEM (Hillsboro, OR, USA), at an accelerating voltage range of 5 – 30 kV. Prior to analysis by SEM, photographs of the samples were taken with an Olympus BX41 System Light equipped with an Olympus Digital Camera (Melville, NY, USA).

CMF preparation

The method used for CMF preparation is similar to that described by Mascal and Nikitin. [10] A known amount of biomass was added to a 150 mL glass pressure tube containing 35 mL of concentrated HCl and 70 mL of dichloroehane (DCE). The tube was sealed and heated to a required temperature (80 °C, 90 °C or 100 °C) with vigorous stirring. After 1 h, the reactor was cooled to room temperature (24 ºC) and the organic layer separated. A further 70 mL of fresh DCE was added to the aqueous layer and stirred for 5 min then separated. Another batch of 70 mL of fresh DCE was added to the aqueous layer and heated for 1 h. After cooling the organic layer it was then separated and the aqueous phase mixed with fresh DCE for 5 min then separated. A final batch of 70 mL of fresh DCE was added and a third processing cycle carried out. The solid residue was collected after filtration of both the aqueous and organic phases, and was washed with distilled water (5 × 50 mL) to obtain a neutral filtrate. It was dried to constant weight at 45 ºC and then stored in a sealed container at room temperature for further analysis.

All DCE extracts were mixed and dried over anhydrous $MgSO_4$ and the solvent was evaporated off under reduced pressure at ~40 °C. The product obtained after removal of DCE was called "crude bio-oil" and was dried under vacuum and then weighed. The bio-oil was passed through a silica gel (~2 g) column and the oil was eluted with dichloromethane (~50 mL) and solvent removed by evaporation. The purified oil was dried under vacuum to constant weight at 30 °C.

The stability of CMF was examined for commercial CMF (Excel Asia Enterprises Ltd, China), crude and purified bio-oil. The commercial CMF was dissolved in DMSO, $CDCl_3$ and D_2O and each solution stored in a desiccator containing silica gel and kept under vacuum. Product stability was monitored using ^{1}H-NMR. The stability of the crude and purified bio-oil was monitored in a similar way but was dissolved in D_2O and $CDCl_3$.

Proton nuclear magnetic resonance (^{1}H-NMR) for product analysis

CMF was identified and quantified by proton nuclear magnetic resonance, ^{1}H-NMR (Bruker Avance 400 MHz NMR spectrometer). Prior to NMR analysis, a known amount (~150 mg) of crude bio-oil or purified bio-oil was dissolved in measured volume (10 mL) of $CDCl_3$ and a weighed amount of dimethlysulfone (~50 mg) as internal standard (δ = 3.7 ppm). The proton peaks due to CMF are δ = 9.62 ppm (s, 1H), 7.25 ppm (d, 1H), 6.58 ppm (d, 1H) and 4.60 ppm (s, 2H). Spectra were normalised to the aldehyde peak (9.2 to 9.8 ppm depending on solvent).

Fourier transform infrared spectroscopy (FTIR) for product analysis

FTIR was used to confirm CMF in the bio-crude and purified material. Infra-red (IR) spectra were collected using a Nicolet 870 Nexus FTIR system including a Continuum IR microscope equipped with a liquid-nitrogen-cooled MCT detector, and an Attenuated Total Reflectance (ATR) objective incorporating a Si internal reflection element (Nicolet Instrument Corp. Madison, WI). The contact

area with the sample was circular with an approximate diameter of 100 μm. Spectra were collected in the spectral range 4000 to 650 cm^{-1}, using 128 scans and 4 cm^{-1} resolution. The IR spectra were typical of CMF with characteristic peaks correlating to an aldehyde group (~1710, 2850 cm^{-1}), an ether group (~1090 cm^{-1}), carbon/carbon double bond (~1460 cm^{-1}) and an organic chloride group (~680 cm^{-1}).

Solid-state NMR for solid residue analysis

The macromolecular structure of the solid residue after CMF production was studied using ^{13}C-cross-polarization, magic-angle-spinning (CP-MAS) solid-state probe mounted on Inova 400 Varian NMR spectrometer (Agilent, US) operated at 100 MHz. Magic angle spinning was conducted at 13 kHz, a recycle time of 2 s, an acquisition time of 33 ms, over 4,000 scans.

^{31}P-NMR for solid residue analysis

^{31}P-NMR analysis of solid residues after CMF production was conducted according to the procedure described previously. [13] The concentrations of the different hydroxyl groups were calculated based on the internal standard of cyclohexanol (chemical shift, 144.5–144.0 ppm).

Mannich reaction for solid residue analysis

The Mannich reaction is used to provide information on the degree of substitution associated with the C-3 and C-5 aromatic positions of lignin. [16, 17] The detailed procedure for Mannich reaction was same as that described previously. [13] The final solid after Mannich reaction was subjected to elemental analysis. The nitrogen composition determined by elemental analysis was used to calculate the number of free C-3 and C-5 positions.

Elemental analysis of solid residue

Elemental analysis was performed on the residues using a 'Carlo Erba' Elemental Analyser (Model NA1500, UK) instrument and method according to ASTM D 5373. Samples were first dried to remove moisture prior to analysis. Solid samples recovered after CMF production were weighed into a tin capsule that is flash burned in the presence of pure oxygen (excess) and helium carrier gas. Gas chromatographic methods are used to compare to calibrated standards for analysis of carbon, hydrogen, nitrogen, and sulfur. Oxygen was obtained by difference. The higher heating value (HHV) of the sample was calculated based on the following: [18]

$$HHV\ \left(\frac{MJ}{kg}\right) = -1.3675 + 0.3137 \times Carbon + 0.7009 \times Hydrogen + 0.0318 \times Oxygen$$

Results and discussion

Compositional analysis of untreated and treated bagasse

Results of the compositional analysis of untreated and treated bagasse are shown in Table 1. There is significant increase in the glucan content of NaOH-bagasse and IL-bagasse compared to untreated bagasse due to the removal of lignin, ash and extractives. IL pretreatment resulted in the greatest removal of lignin and also removed significant amounts of xylan. The modest increase in the glucan content for ASE-bagasse relative to bagasse is associated with removal of xylan and extractives, as the proportion of lignin is higher in the sample. The NaOH-bagasse and SSE-bagasse treatments mainly removed lignin, hence a higher proportion of glucan was present compared to the untreated bagasse. The significant differences in the proportions of xylan and lignin contents among the samples are related as to whether pretreatment was performed either in alkali or acidic condition. At very high pH, delignification is predominant, whereas at very low pH, xylan removal is greatest. The high ash content in the untreated bagasse sample is because

the sample was directly obtained from the sugar factory, and soil typically accounts for 1% of the wet mass of sugarcane billets delivered to the factory.

Scanning electron microscopy was used to examine the microscopic structural differences among the samples (Figure 1a-e). The bagasse sample contained fibre bundles (Figures 1a1 and 1a2), while acid treatment (ASE-bagasse) resulted in fibre disintegration (with lengths from 10 μm > 100 μm, Figures 2a1 and 2a2). The IL-bagasse sample mainly contained defibrillated fibre strands (Figure 1c1) and the fibre surface was relatively clean and smooth because of removal of lignin (Figuer 1c2). The NaOH-bagasse sample (Figures 1d1 and 1d2) contained strands of longer defibrillated fibres (with lengths > 200 μm) compared to IL-bagasse sample. The morphological properties of SSE-bagasse (Figure 1e1) were similar to NaOH-bagasse (Figure 1d1). However, at higher magnification (Figure 1e2), micro-cracks can be observed on fibre due to steam explosion. The widths of the defibrillated fibres of IL, NaOH and SSE were in the similar range, ~10 – 30 μm.

Effect of processing conditions on CMF yield

The effect of reaction temperature on the conversion of untreated bagasse to CMF is shown in Table 2 (entries 1-3). Hydrolysis carried out at 90 °C resulted in the highest CMF yield though this was not significant. Bredihhin *et al.* [19] found the optimum temperature to be 65 °C, below this temperature the reaction was slow, and above this temperature the yield of 5-bromomethylfurfural (a furanic similar to CMF) was slightly lower for glucose, cellulose and aspen with a biomass loading of 1%. Similar bio-oil results was obtained with <2 mm and <0.5 mm fractions, though slightly lower yield was obtained with the smallest particle size fraction (which also retains a larger proportion of ash from the whole bagasse). The insignificant differences in the results are due to the very strong acidic

system used, nullifying any mass transfer limitations caused by particle size differences. The difference in biomass loading from 0.5 to 1.5 wt% on the CMF yield was not significant.

The reaction rate at 90 °C was likely to be higher than that at 80 °C since the rate-limiting isomerization of glucose to fructose (formed in situ) has relatively high activation energy. [20] As such, further experiments were conducted at 90 °C with bagasse (<2 mm particle size) at higher feed loading (5 % and 10%), in order to quantify the various furanics present in the bio-oil (Table 3). Maximum CMF and furfural yields are obtained at 1%; thereafter the yields gradually dropped. The decreased yield of CMF is due to increased degradation to the by-products HMF (^{1}H-NMR, δ = 9.75 ppm, 6.34 ppm, 4.64 ppm), levulinic acid (LA) (^{1}H-NMR, δ = 2.51 ppm, 2.35 ppm, 2.17 ppm) and 2-hydroxyacetylfuran (HAF) (^{1}H-NMR, δ = 7.60 ppm, 7.26 ppm, 6.56). These results are consistent with a previous study which showed that increasing biomass loading from 1% to 10% caused 5−10% decrease in CMF yield with different biomass substrates. [10] Low yields of furfural (from the hemicellulose component of bagasse) were achieved (<40 mol%) and is similar to the 40% yields from corn stover achieved by Mascal and Nikitin. [21] The low furfural yield highlights either low reaction selectivity for C5 sugars or reflects the instability of furfural under acidic reaction conditions, whereby furfural degrades to polymers and solid material (humins). The yield of furfural reduced by ~13% at 10% bagasse loading. There was also an increase in the solid residue content with increasing biomass loading.

CMF production from various treated bagasse samples showed that yield was in the order IL-bagasse > SSE-bagasse ~NaOH-bagasse> untreated bagasse >> ASE-bagasse (Figure 2). The IL-bagasse with the highest glucan (*i.e.*, hexose sugars) content and the lowest lignin and ash contents gave the highest yield. This is not unexpected, as CMF conversion is via hexose sugar hydrolysis, and the IL

pretreatment process led to the highest increase in the proportion of cellulose due to the removal of the highest total non-cellulose components (Table 1 and Table 2).

Figure 2 shows that the highest CMF yield of 81.9% was achieved with IL-bagasse, followed by 78.2% with SSE-bagasse, 77.2% with soda (NaOH)-bagasse, 73.5% with untreated bagasse and 62.3% with ASE bagasse, which corresponded to lignin contents of 6.9%, 12.3%, 9.8%, 21.5% and 29.6% respectively in original bagasse samples (Table 2). The results in Figure 2 indicate that lignin content has a negative effect on CMF yield. Figure 2 also indicates high glucan content and low xylan content have positive effects on CMF yield. However, CMF yield was the lowest in spite of its lowest xylan content and a moderate level of glucan possibly because of the highest lignin content. The trend of higher yield with increasing cellulose content was also demonstrated for pure microcrystalline cellulose (83.5% CMF) and corn stover containing 33.9% cellulose (80.2% CMF) with 1% (w/v) substrate loading by Mascal and Nikitin. [10] However, at a much higher substrate loading of 10% (w/v), pure cellulose gave a significantly higher CMF yield than corn stover (78.2% vs 70.4%). [10] This may simply be due to lack of sufficient contact between the biomass and the surrounding acid (*i.e*,. cellulose accessibility) for the corn stover biomass.

The ASE-bagasse sample gave the lowest yield of CMF. This biomass has the highest proportion of lignin but the lowest xylan content. In terms of composition, the main significant differences between ASE-bagasse and SSE-bagasse are the ash and xylan contents (Table 1). It is likely that ash will not influence CMF formation (given the use of concentrated acid), while xylan may because of its reactive nature under acid conditions and propensity of furfural to polymerise with other products and reactants. As such, it should be expected that the CMF yield would be lower for the SSE-bagasse sample because of its significantly higher xylan content. As this is not the case there are likely to be

other reasons for the differences in the result. The difference in the sizes of the fibres appears not to play a role in CMF yield. As shown in Figure 1, scanning electron micrographs reveal differences in ultra-structures. ASE-bagasse was brown, indicating the predominance of lignin on the outer surface of the biomass. This is an indication of lignin redistribution from its original location from the fibre matrix would have occurred to a significant extent compared to the SSE pretreatment or the other pretreatments. Selig *et al.* [22] reported the deposition of lignin droplets on the biomass after dilute acid pretreatment of maize stems. This phenomenon would have likely occurred with ASE-bagasse, and as such ready access to the glucan component of the biomass by the concentrated acid may have been physically blocked. It is also probable that during the reaction process, acid soluble lignin species, which will be highest in the ASE-bagasse acid system, will react with glucose released during hydrolysis, reducing the amount available for conversion into CMF. As such, ultra-structure differences, where is a physical barrier involving lignin, clearly impact on CMF yield.

Bio-oil stability

In an industrial process, the bio-oil is likely to be stored prior to further processing, and so its stability is of vital importance. Commercial CMF was analysed using ^{1}H-NMR after storage in various solvents to examine CMF stability (Figure 3). In $CDCl_3$, CMF was stable at the end of the 14 days of examination. In DMSO, after 7 days peaks in the ^{1}H-NMR spectra appeared at δ = 9.5 ppm, 7.4 ppm, 6.6 ppm, 4.5 ppm and 3.8 ppm associated with HMF (~15 wt%). DMSO is hydrophilic and absorbs moisture, so it is expected that the small amount of water present will hydrolyse CMF to HMF. Additional degradation products were formed from CMF stored in D_2O over the 14 day period. Peaks at 8.0 and 4.7 ppm indicated the presence of HAF or the CMF analogue, chloroketone; 2-chloro-1-(furan-2-yl) ethanone (CFE). As some of the other peaks linked to HAF or CFE [23] were not detected, it is assumed that these peaks may have been swamped by the

CMF peaks. Levulinic acid was also detected in the CMF stored in D_2O and the peaks associated with it dropped over time. So, with commercial CMF, it must be stored in a moisture-free environment or in a chlorinated solvent like chloroform.

Figure 4 presents the ^{1}H-NMR spectra of crude bio-oil stored in various conditions. The spectrum obtained with $CDCl_3$ remains unchanged even after 1 week, and is similar to the fresh crude bio-oil. The spectra of the crude bio-oil stored neat at 20 °C in a desiccator (under reduced pressure) for 24 h and that in DCE after 1 week, show prominent peaks associated with LA formation (δ = 2.51 ppm, 2.35 ppm, 2.17 ppm). The broad singlet at ~1.3 ppm could be due to aliphatic extractives from bagasse, although polymeric degradation products are possible. [24]

The purification procedure which is expected to remove soluble polymeric material produced ^{1}H-NMR spectra with sharper peaks (*c.f.* Figures 4 and 5). Surprisingly, LA is present in the purified bio-oil at a noticeably higher proportion than the crude bio-oil. The peaks < 2 ppm also increased in intensities in the neat sample (Figure 5). Two possibly explanations for this is a relative increase in aliphatic impurities due to relative decrease in CMF content as result of conversion to LA and/or that these peaks are due to CMF degradation products. The question is why there is more CMF break down in the purified bio-oil than the crude bio-oil that contains more impurities. The reason for this is unknown. However, the relative stability of the crude bio-oil may be related to more acidic environment.

Functional groups of solid residue

The solid content after acid hydrolysis accounted for over 45% of the total biomass (on dry basis). As this amount is significant, detailed characterization of the solid residue was carried out to determine its value as a by-product. The function groups present in the solid residue was

investigated by ATR-FTIR (Supplementary Figure S1). The wide peak in the range 2979-3662 cm^{-1} is attributed to O-H stretching vibrations, [25] the peak 2940 cm^{-1} is due to C-H stretching, and the peak at 2892 cm^{-1} is due to C-H stretching vibrations of the methoxy group [25]. These peaks are broader in the spectrum of the solid residue than those of bagasse suggesting modification of these groups through condensation. The peak at 1730 cm^{-1} is associated with conjugated aldehyde or carboxylic acid carbonyl group, [25, 26] and is slightly more prominent in the residue. The residue contains peaks of higher intensities at 1602 cm^{-1} and 1510 cm^{-1} due to furanic ring stretching [25-27] and at 1035 cm^{-1} (C-O stretching or ring deformation), [26] as well at 1360-1390 cm^{-1} and 1280 cm^{-1} (C-O stretching and ring vibrations). [26, 28] This suggests that condensation of furan species has occurred and may explain the low yield of furfural achieved from the hemicellulose content of the bagasse. The peaks at 1462 cm^{-1} and 1421 cm^{-1} is assigned to methoxy groups in lignin, [29] and are of higher intensity in the solid residue relative to bagasse. The solid residue is 45-47% of the starting material, and if it is assumed all lignin in the starting material is transferred to the residue than ~50% of residue comprises of lignin. This explains the extensive presence of lignin structural features present in the residue. However, a large portion of the lignin structure has been modified and/or condensed into humic structures reducing the solubility of the residue to ~15-20% in 0.1 M NaOH solution.

The ^{13}C CP-MAS NMR spectrum of the residue (Supplementary Figure S2), and the assignment of the different regions of the molecular substructures were based on the information obtained in the literature. [30-32] The two main prominent peaks at δ = 100 ppm and 130 ppm are associated with the presence of aromatic compounds. The big shoulder at δ = 85 ppm may be related to C-α,β,γ, substructure, and slight hump at 65 ppm is the methoxyl substituent. The peaks at δ = 157 ppm to 200 ppm are carbonyl substituents, while the peak at 220 ppm belongs to keto groups. As

spectrum profile and the peaks of the phenolic (δ = 157 ppm) and the methoxyl (δ = 65 ppm) substituents are small, it is inferred that the hydrolysis residue is dissimilar from lignin and the lignin has been modified by the concentrated acid process. [33]

Figure 6 shows the ^{31}P-NMR of the solid residue from hydrolysis of untreated bagasse and bagasse soda lignin obtained by acid precipitation and drying of the black liquor produced during NaOH pretreatment. The spectrum of the solid residue is distinctly different from that of soda lignin and contains very few peaks. This may be an indication of a highly polymerized and condensed material. It was observed that only about 20% of the solid residue was soluble in the work-up procedure for the analysis, and so this proportion is what is revealed in the spectrum. The spectrum, however, contains sharp peaks at δ = 149.5 ppm and 146.5 ppm associated with aliphatic hydroxyl groups, a doublet at δ = 136 ppm associated with the carboxylic acid group, [34, 35] and unknown peaks at 129 ppm, 132 ppm and 132.5 ppm. The sharpness of the peak indicates low molecular weight phenolic species.

Elemental analysis and Mannich reactivity of solid residue

Mannich reactivity is an organic synthesis method that is used to study the chemical reactivity of lignin. [36] The elemental analysis of the solid residue (before and after treatment) is presented in Table 4. The increase in nitrogen content indicates presence of C3 and/ C5 active sites on the phenolics present in the solid residue. This amount is far lower than the values of 2.24% and 2.49% obtained for bagasse soda lignin and bagasse IL lignin respectively, from our previous work. [13] The results may therefore indicate that the modified solid residue obtained from the Mannich reaction will not be as suitable for the production of surfactant chemicals and polycationic materials as bagasse and IL-bagasse solid residue. [37] However, the results show that the residue

has a higher calorific heating value (20.3 MJ/kg) than untreated bagasse (18.3 MJ/kg), and so it can be used in combustion boilers to produce energy. The sulfur content is low and so should not be of major concern in these type of boilers.

Conclusion

The results indicated that although pretreatment of bagasse improved CMF yield, this was not significant to warrant its use prior to acid hydrolysis. In fact, the type of pretreatment could significantly reduce CMF yield. Pretreatments that results in lignin re-distribution and possibly other surface changes appear to affect CMF prodution. However, as pretreatment results in fractionation of the main components, converting the hemicellulose and/or the lignin components to value-added products will enhance biomass conversion processes. This is because the use of a biphasic system involving concentrated acid for CMF destroys the hemicellulose component of the biomass and renders the lignin component highly condensed. As such, a fractionation/pretreatment process that separates out the three main lignocellulosic components will allow each component to be treated separately and therefore improve the economics of CMF production. The removal of hemicellulose and lignin reduces the amount and type of impurities that ends up in the crude bio-oil produced, thereby simplifying the purification process and hence will reduce the cost of CMF production.

The present study has also highlighted the instability of CMF. CMF was shown to be fairly stable in chlorinated solvents, but began to break down when stored neat as a bio-oil as it is highly reactive to moisture.

Acknowledgement

The authors thank Sugar Research Alliance for the PhD Scholarship support to Joshua Howard. The authors also appreciate the technical support from Dr Mark Wellard (School of Chemistry, Physics and Mechanical Engineering, Queensland University of Technology) for NMR analysis.

References

1. D. Humbird, R. Davis, L. Tao, C. Kinchin, D. Hsu and A. Aden, *Process design and economics for biochemical conversion of lignocellulosic biomass to ethanol*, National Renewable Energy Laboratory Golden, Colorado, 2011.
2. E. L. Kunkes, D. A. Simonetti, R. M. West, J. C. Serrano-Ruiz, C. A. Gartner and J. A. Dumesic, *Science*, 2008, 322, 417-421.
3. T. R. Carlson, T. R. Vispute and G. W. Huber, *Chemsuschem*, 2008, 1, 397-400.
4. B. Saha and M. M. Abu-Omar, *Green Chem*, 2014, 16, 24-38.
5. C. M. Cai, T. Y. Zhang, R. Kumar and C. E. Wyman, *J Chem Technol Biot*, 2014, 89, 2-10.
6. M. Mascal and E. B. Nikitin, *Angewandte Chemie International Edition*, 2008, 47, 7924-7926.
7. M. Mascal and E. B. Nikitin, *Green Chemistry*, 2010, 12, 370-373.
8. N. Kumari, J. K. Olesen, C. M. Pedersen and M. Bols, *European Journal of Organic Chemistry*, 2011, 2011, 1266-1270.
9. Y. Yang, M. M. Abu-Omar and C. Hu, *Applied Energy*, 2012, 99, 80-84.
10. M. Mascal and E. B. Nikitin, *ChemSusChem*, 2009, 2, 859-861.
11. W. H. Gao, Y. Q. Li, Z. Y. Xiang, K. F. Chen, R. D. Yang and D. S. Argyropoulos, *Molecules*, 2013, 18, 7675-7685.

12. S. W. Breeden, J. H. Clark, T. J. Farmer, D. J. Macquarrie, J. S. Meimoun, Y. Nonne and J. E. S. J. Reid, *Green Chem*, 2013, 15, 72-75.
13. L. Moghaddam, Z. Zhang, R. M. Wellard, J. P. Bartley, I. M. O'Hara and W. O. S. Doherty, *Biomass and Bioenergy*, 2014, 70, 498-512.
14. Z. Y. Zhang, I. M. O'Hara and W. O. S. Doherty, *Green Chemistry*, 2013, 15, 431-438.
15. A. Sluiter, *Determination of Structural Carbohydrates and Lignin in Biomass: Laboratory Analytical Procedure (Lap): Issue Date, 4/25/2008*, National Renewable Energy Laboratory, Golden, Colo. :, 2011.
16. X.-J. Pan and Y. Sano, *Journal of Wood Science*, 1999, 45, 319-325.
17. N.-E. E. Mansouri and J. Salvadó, *Industrial Crops and Products*, 2006, 24, 8-16.
18. C. Sheng and J. L. T. Azevedo, *Biomass and Bioenergy*, 2005, 28, 499-507.
19. A. Bredihhin, U. Mäeorg and L. Vares, *Carbohydrate Research*, 2013, 375, 63-67.
20. C. Usuki, Y. Kimura and S. Adachi, *Food Sci. Technolo. Res.*, 1981, 13, 205-209.
21. M. Mascal and E. B. Nikitin, *ChemSusChem*, 2009, 2, 423-426.
22. M. J. Selig, S. Viamajala, S. R. Decker, M. P. Tucker, M. E. Himmel and T. B. Vinzant, *Biotechnology Progress*, 2007, 23, 1333-1339.
23. M. Brasholz, K. von Kanel, C. H. Hornung, S. Saubern and J. Tsanaktsidis, *Green Chemistry*, 2011, 13, 1114-1117.
24. G. Knothe and J. A. Kenar, *European Journal of Lipid Science and Technology*, 2004, 106, 88-96.
25. K. Bilba and A. Ouensanga, *Journal of Analytical and Applied Pyrolysis*, 1996, 38, 61-73.
26. S. K. R. Patil and C. R. F. Lund, *Energy and Fuels*, 2011, 25, 4745-4755.
27. N. Shi, Q. Liu, Q. Zhang, T. Wang and L. Ma, *Green Chemistry*, 2013, 15, 1967-1974.

28. I. V. Sumerskii, S. M. Krutov and M. Y. Zarubin, *Russian Journal of Applied Chemistry*, 2010, 83, 320-327.
29. N. Labbe, T. G. Rials, S. S. Kelley, Z. M. Cheng, J. Y. Kim and Y. Li, *Wood Science and Technology*, 2005, 39, 61-U19.
30. L. G. Akim, S. M. Shevchenko and M. Y. Zarubin, *Wood Sci.Technol.*, 1993, 27, 241-248.
31. R. K. Sharma, J. B. Wooten, V. L. Baliga, X. Lin, W. Geoffrey Chan and M. R. Hajaligol, *Fuel*, 2004, 83, 1469-1482.
32. H. Wikberg and S. Liisa Maunu, *Carbohydrate Polymers*, 2004, 58, 461-466.
33. F. Liang, Y. Song, C. Huang, J. Zhang and B. Chen, *Catalysis Communications*, 2013, 40, 93-97.
34. P. Sannigrahi, A. J. Ragauskas and S. J. Miller, *Energy & Fuels*, 2010, 24, 683-689.
35. Y. Pu, S. Cao and A. J. Ragauskas, *Energy & Environmental Science*, 2011, 4, 3154-3166.
36. X. Du, J. Li and M. E. Lindström, *Industrial Crops and Products*, 2014, 52, 729-735.
37. H. P. S. A. Khalil, M. M. Marliana and T. Alshammari, *Bioresources*, 2011, 6, 5206-5223.

Legends of Tables and Figures

Tables

Figures

Figure 5 ^{1}H-NMR of purified bio-oil (a) freshly prepared, (b) stored for 24 h, and (c) stored in $CDCl_3$ for 1 week.

Figure 6 ^{31}P-NMR spectrum of bagasse hydrolysis residue (bottom) and soda lignin (top)

Supplementary Figures

Figure S1 FTIR of (a) bagasse and (b) solid residue

Figure S2 ^{13}C CP-MAS NMR spectrum of bagasse hydrolysis residue

Table 1

Bagasse type	Glucan (wt%)	Xylan (wt%)	Arabinan (wt%)	Lignin (wt%)	Ash (wt%)	Extractives (wt%)
Untreated bagasse	43.0	17.4	1.7	21.5	9.4	8.2
NaOH-bagasse	66.3	21.8	1.5	9.8	2.0	ND
IL-bagasse	81.6	10.3	<0.1	6.9	0.8	ND
ASE-bagasse	58.6	3.6	<0.1	29.6	8.2	ND
SSE-bagasse	58.5	16.7	<0.1	12.3	15.1	ND

Table 2

Entry	Loading, wt%	Temperature, °C	Particle size less than, mm	Bio-oil yield*, % conversion based on C6 sugar content
1	1.0	80	0.5	80.0
2	1.0	100	0.5	76.1
3	1.0	90	0.2	80.3
4	1.0	90	0.5	80.9
5	1.0	90	2.0	83.3
6	0.5	90	0.5	81.4
7	1.5	90	0.5	81.2

* The errors on bio-oil yields were within ±3%.

Table 3

Sample	CMF (%)	Furfural (%)	HMF (%)	LA (%)	HAF (%)	Solids (%)
Bagasse-1%	74.1	38.4	n/a	n/a	n/a	45.1
Bagasse-5%	72.3	34.2	3.9	2.3	3.0	45.7
Bagasse-10%	69.8	33.5	3.7	2.6	3.3	47.9

Table 4

Solid residue	N wt%	C wt%	H wt%	S wt%	O wt%
Before	0.01	52.92	5.34	0.05	41.73
After	1.44	42.36	5.23	0.00	50.97

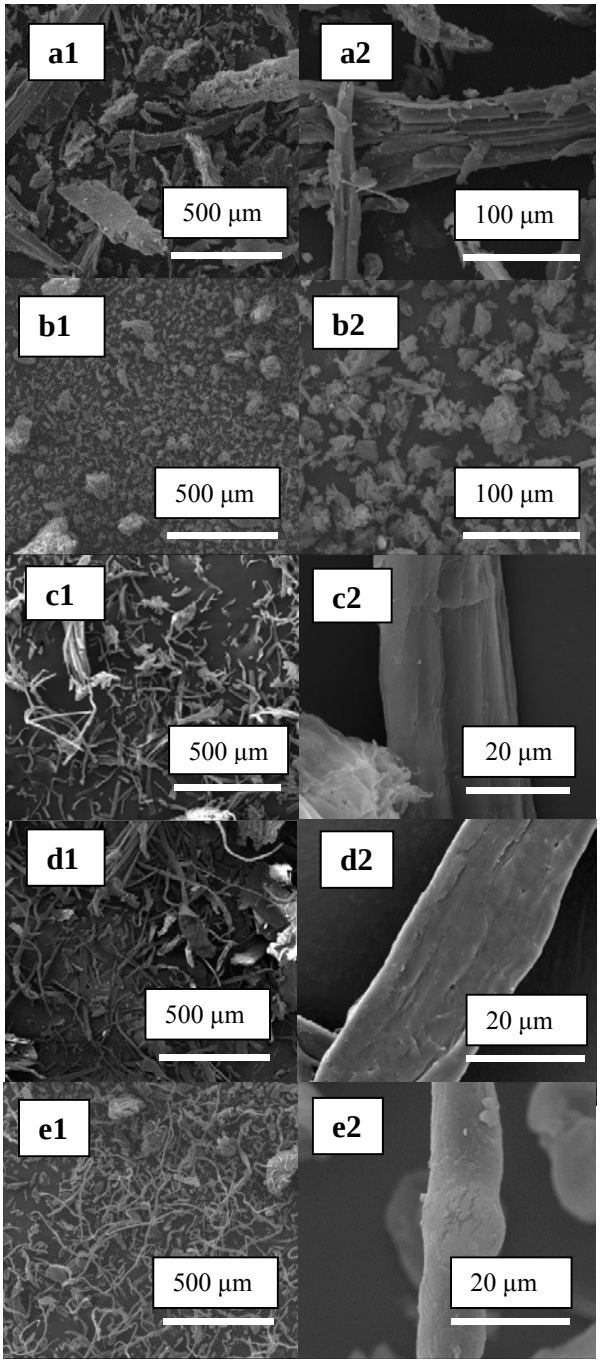
a1
500 µm
a2
100 µm
b1
500 µm
b2
100 µm
c1
500 µm
c2
20 µm
d1
500 µm
d2
20 µm
e1
500 µm
e2
20 µm

Figure 1

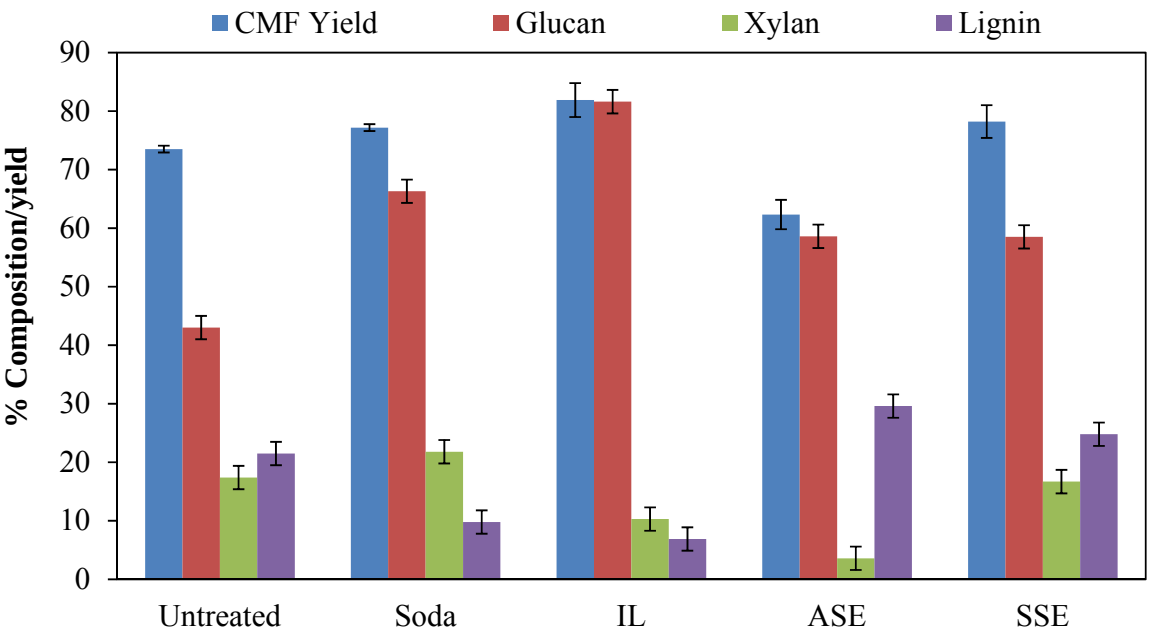
CMF Yield
Glucan
Xylan
Lignin
% Composition/yield
90
80
70
60
50
40
30
20
10
0
Untreated
Soda
IL
ASE
SSE

Figure 2

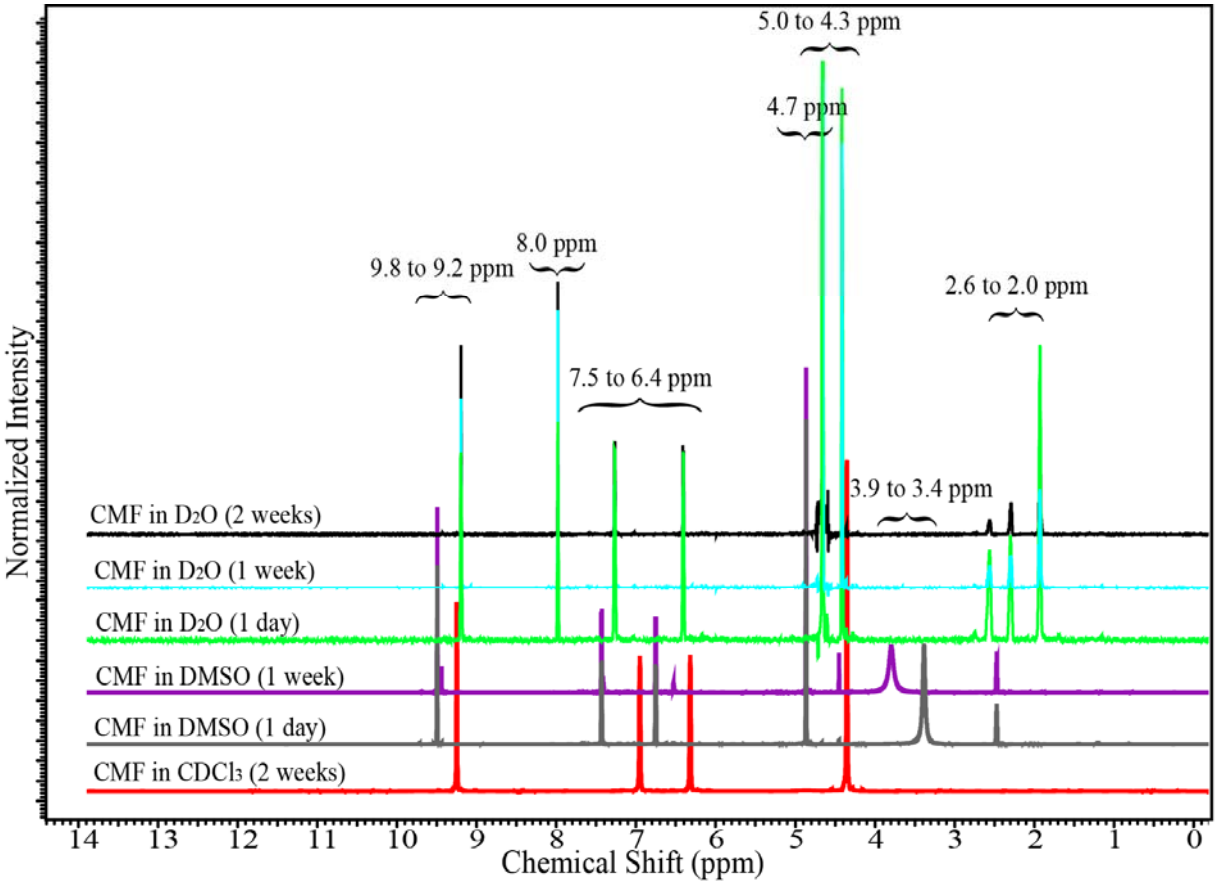
5.0 to 4.3 ppm
4.7 ppm
8.0 ppm
9.8 to 9.2 ppm
2.6 to 2.0 ppm
7.5 to 6.4 ppm
3.9 to 3.4 ppm
CMF in D2O (2 weeks)
CMF in D2O (1 week)
CMF in D2O (1 day)
CMF in DMSO (1 week)
CMF in DMSO (1 day)
CMF in CDCl3 (2 weeks)
Normalized Intensity
14 13 12 11 10 9 8 7 6 5 4 3 2 1 0
Chemical Shift (ppm)

Figure 3

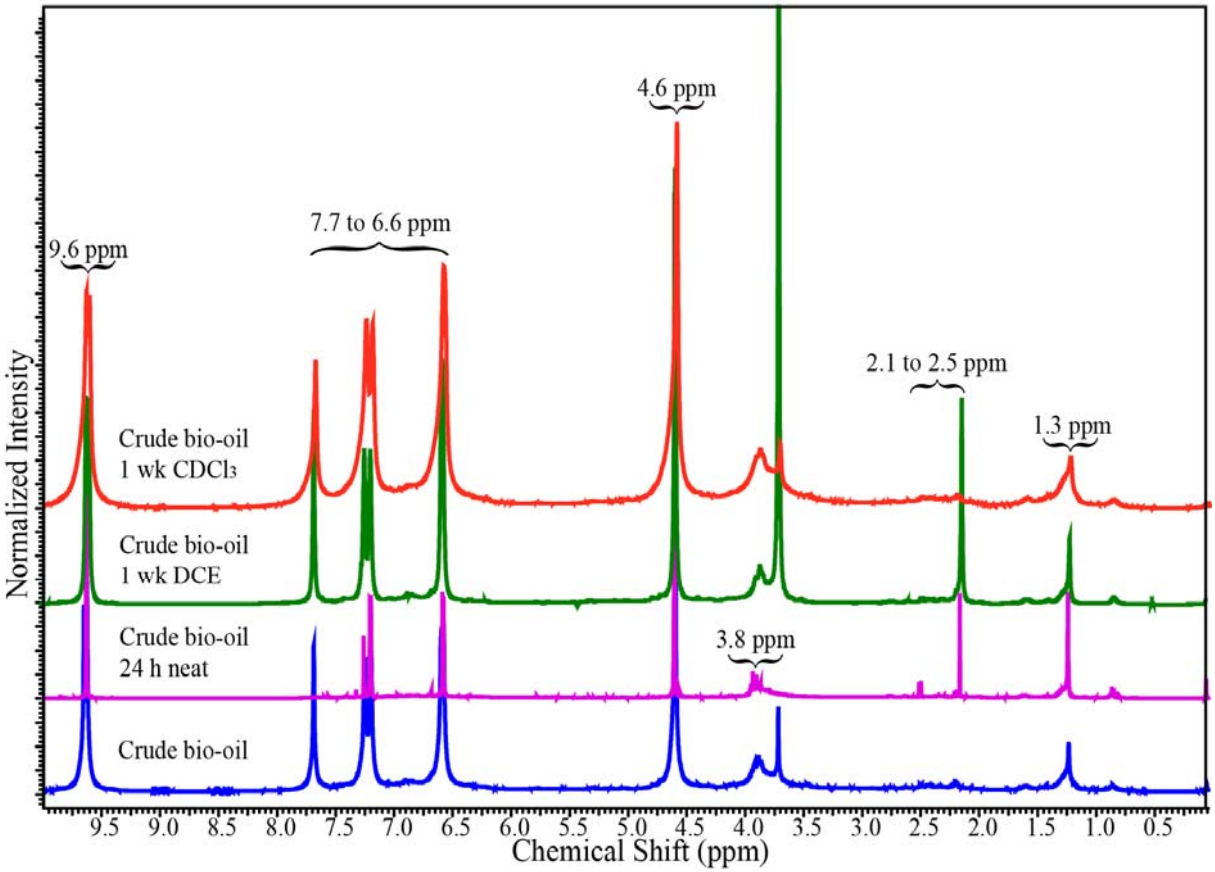
9.6 ppm
7.7 to 6.6 ppm
4.6 ppm
2.1 to 2.5 ppm
1.3 ppm
3.8 ppm
Crude bio-oil
1 wk CDCl3
Crude bio-oil
1 wk DCE
Crude bio-oil
24 h neat
Crude bio-oil
Normalized Intensity
9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5
Chemical Shift (ppm)

Figure 4

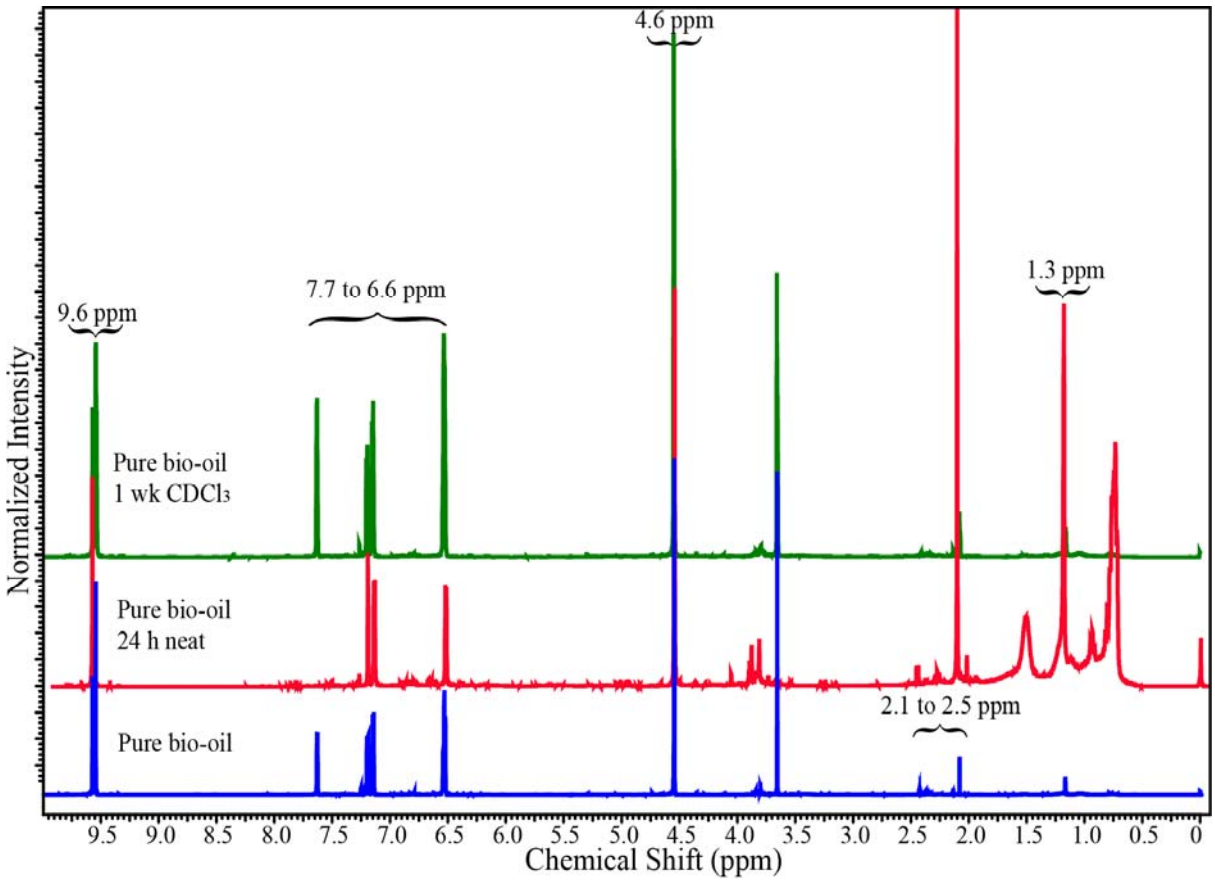
4.6 ppm
1.3 ppm
7.7 to 6.6 ppm
9.6 ppm
Normalized Intensity
Pure bio-oil
1 wk CDCl3
Pure bio-oil
24 h neat
2.1 to 2.5 ppm
Pure bio-oil
9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0
Chemical Shift (ppm)

Figure 5

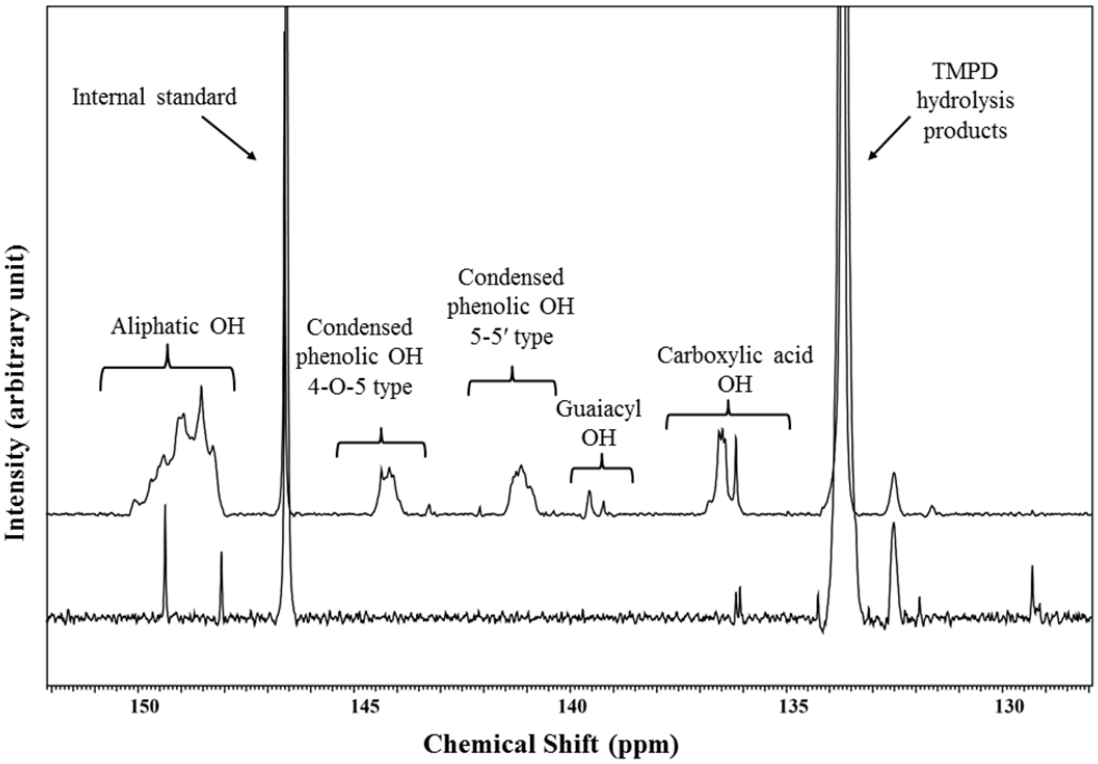
Internal standard
TMPD hydrolysis products
Aliphatic OH
Condensed phenolic OH 4-O-5 type
Condensed phenolic OH 5-5′ type
Guaiacyl OH
Carboxylic acid OH
Intensity (arbitrary unit)
150
145
140
135
130
Chemical Shift (ppm)

Figure 6

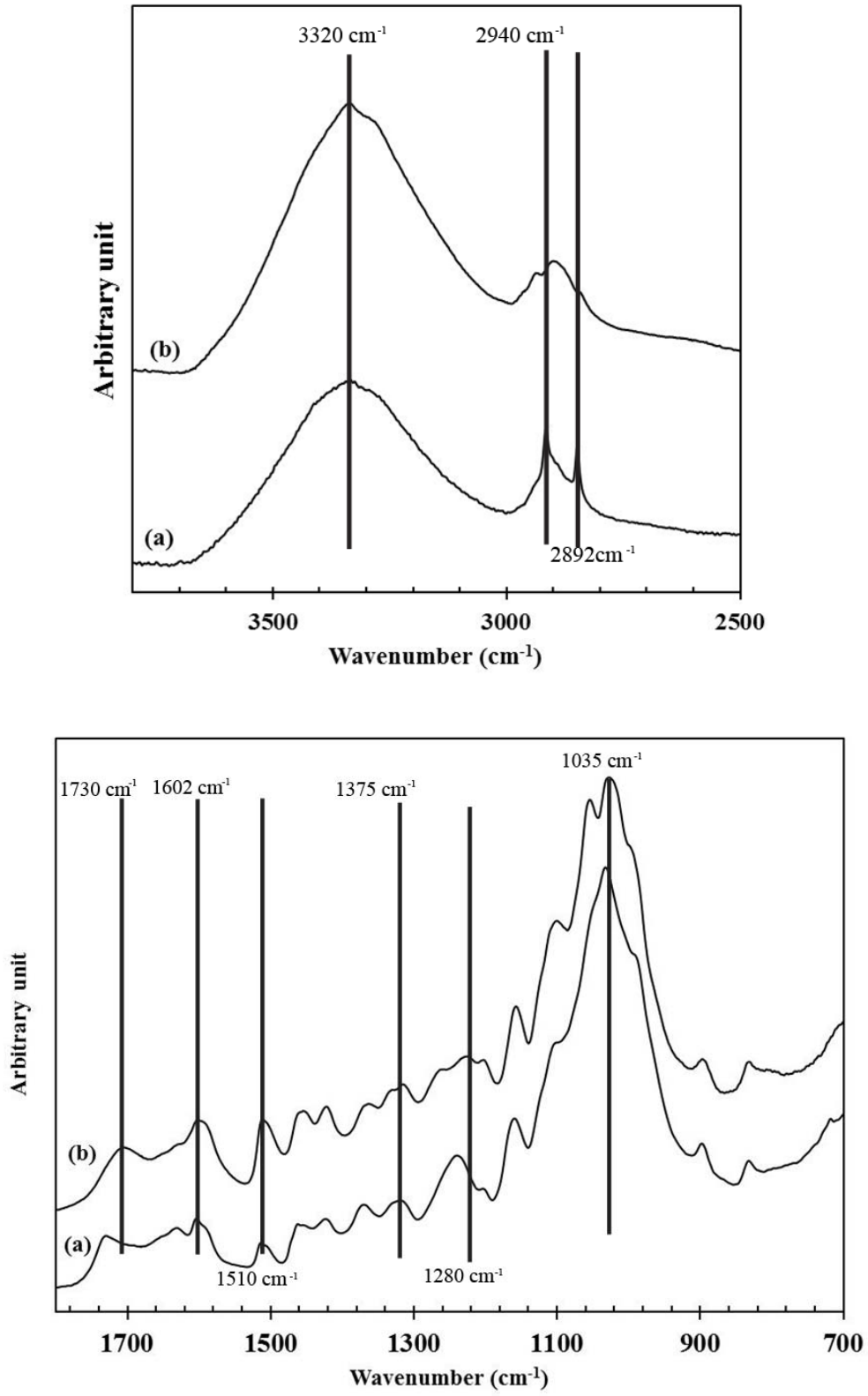
3320 cm-1
2940 cm-1
Arbitrary unit
(b)
(a)
2892cm-1
3500
3000
2500
Wavenumber (cm-1)
1730 cm-1
1602 cm-1
1375 cm-1
1035 cm-1
Arbitrary unit
(b)
(a)
1510 cm-1
1280 cm-1
1700
1500
1300
1100
900
700
Wavenumber (cm-1)

Figure S1

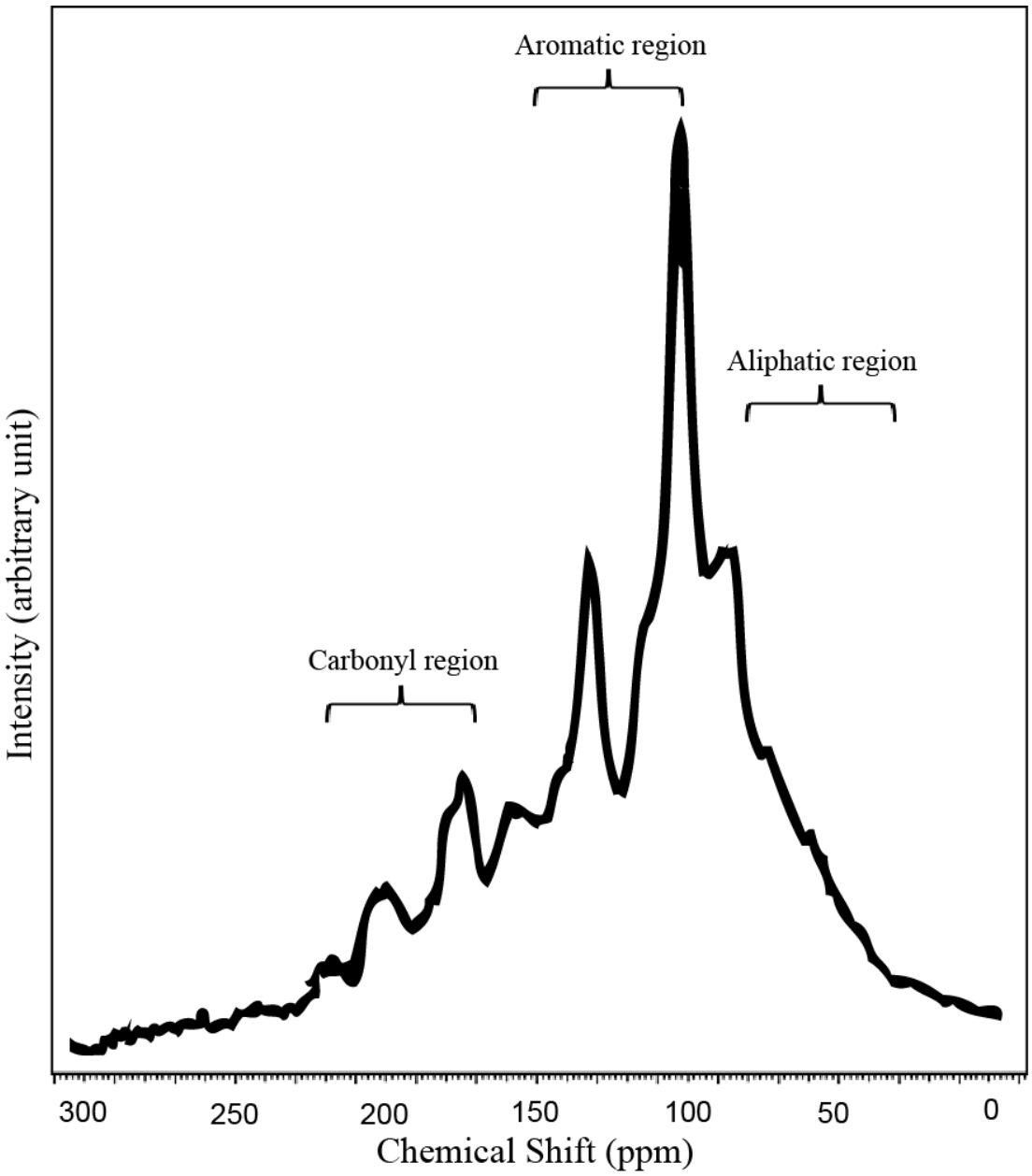
Aromatic region
Aliphatic region
Carbonyl region
Intensity (arbitrary unit)
300
250
200
150
100
50
0
Chemical Shift (ppm)

Figure S2