RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

Efficient recovery and structural characterization of lignin from cotton stalk based on a biorefinery process using γ-valerolactone/water system

Miao Wu,[a] Jia-Kun Liu,[a] Zhong-Ya Yan,[a] Bo Wang,[a] Xue-Ming Zhang,[*a] Feng Xu[a] and Run-Cang Sun[ab]

In order to understand the integrated reaction behaviors for biomass pretreatment using the γ-valerolactone (GVL)/water system, the recovery and structural changes of lignin component from cotton stalk obtained under different ratios of GVL to water were investigated. Structural elucidation of these lignin samples was performed by Fourier transform infrared spectroscopy (FT-IR), high-performance anion-exchange chromatography (HPAEC), gel permeation chromatography (GPC), 2D heteronuclear single quantum coherence spectroscopy nuclear magnetic resonance (2D HSQC NMR), and derivatization followed by reductive cleavage (DFRC). The results showed that the separated lignin fractions possessed higher yields and purities than milled wood lignin (MWL). From the results of molecular weight, DFRC and 2D NMR analysis, it was also found that lignin component from cotton stalk was G-S type unit analogous to hardwood, and remarkable degradation and repolymerization occurred on lignin in this acid system, which led to more condensed and lower molecular weight lignin than MWL. Particularly, under the condition of GVL/H_2O 80/20, the cleavages of aryl-ether dramatically happened, resulting in the least amount

[a]*Beijing Key Laboratory of Lignocellulosic Chemistry, Beijing Forestry University, Beijing 100083, China. E-mail: xm_zhang@bjfu.edu.cn; Tel/Fax: +86 10 62336189*

[b]*State Key Laboratory of Pulp and Paper Engineering, South China University of Technology, Guangzhou 510640, China*

of *β*-*O*-4′ linkages. Considering the concept of biorefinery, the one-pot acid GVL/H_2O system with 80/20 GVL/H_2O could be an attractive method for simple and efficient recovery of lignin and sugars simultaneously from agricultural wastes.

Introduction

Due to the increasing depletion of fossil resources and contamination of environment, biomass, the only practical renewable resource of carbon, has attracted more attention for the production of fuels and chemicals.[1] Nowadays, many governments have also established legislation and set targets to achieve the sufficient utilization of renewable resources in future.[2,3] Actually, the reserve of biomass is immense, it is reported that more than 370 million dry tons of forestry biomass and 1 billion dry tons of agricultural biomass are available in the United States every year.[2] However, among them, much of the biomass with the potential to become a valuable feedstock is always considered as the waste. As a consequence, the emphasis will be put on the recycling of forest and agricultural waste in biorefinery. In consideration of the considerable researches on corn stalk, rice straw and wheat straw, the relatively less popular biomass, cotton stalk, is chosen in this study.

As one of the three main components in biomass (along with cellulose and hemicellulose), lignin is a complex and heterogeneous biomacromolecule, consisting of phenylpropane units and various functional groups. The primary monomers of lignin are guaiacyl (G), syringyl (S), and *p*-hydroxyphenyl (H), which depend on the species of biomass and link together through C-O bonds of *α*- and *β*-aryl alkyl ethers

and C-C linkages.[4,5] On account of the high degree of aromaticity, lignin represents a benign source of fuels, chemicals and materials. For example, it has been used for the production of aromatics, agrochemicals and high-performance materials, such as carbon fibers, activated carbons and polyurethanes.[6-9] However, among the three main components, lignin is deemed to largely render biomass intractable, thus, in order to maximum the utilization of all-components in biomass, it is necessary to seek an efficient technique for the removal and recovery of lignin, which can facilitate easier access to the carbohydrates and the production of valuable side-product streams based on lignin.[10]

Nowadays, many methods have been developed for lignin isolation within the biorefinery concept, including acid, alkali, organosolv, ionic liquid, hydrothermal pretreatment and so forth. Among these pretreatment technologies, dilute acid pretreatment has been widely studied because it is effective and inexpensive, typically employing 0.5-3.0% H_2SO_4 in a batch reaction system or an acid level lower than 0.1% in a flow-through acid pretreatment at temperatures of 140-220^{o}C for a certain time.[11] Whereas, traditional acid pretreatments always focus on the preliminary removal of hemicelluloses, and lignin component requires a subsequent step to be recovered, such as the most common alkaline extraction.[12,13] Recently, a noteworthy flow-through process has been established to produce concentrated streams of C5 and C6 sugars directly from the cellulose and hemicellulose fractions of intact lignocellulosic biomass in GVL-H_2O solvent mixtures with the addition of low concentration of sulfuric acid.[14] Based on various previous reports and advantages of

this renewable solvent GVL,[15-17] it is conjectured that this method can be referred as a new pretreatment in the shape of a brief batch reaction at a fixed temperature to simultaneously achieve the high conversion efficiency of carbohydrates and the recovery of lignin due to the synergistic effect of acid water and GVL organic solvent. In our previous study, a one-pot γ-valerolactone/water system pretreatment containing a very low dosage of acid was put forward, which achieved an excellent enzymatic saccharification efficiency, resulting in the total recovery of over 92.6% of glucose, and simultaneously, an easy acquisition of the lignin component without tedious steps.[18] While the structure features and changes of the obtained lignin fraction during the process was still unaware and deserved to be explored in order to further integrallty explain the reaction course and promote the complete utilization of biomass.

Various methods of structural characterization of lignin have been investigated in recent years, mainly consisting of wet chemical techniques, thermochemical conversion, chromatography, vibrational spectroscopy, nuclear magnetic resonance, electronic spectroscopy, atomic force and electron microscopy. Thereinto, chromatographic methods are often employed in conjunction with wet chemical methodology. For instance, derivitization followed by reductive cleavage (DFRC), a technique developed by Lu and Ralph to break α and β-ether linkages, leaving intact γ-esters, was used with GC-FID and GC-MS to study different lignin model compounds and actual lignin fractions.[19] This technique has been reported to be more advantageous compared to alkaline nitrobenzene oxidation and thioacidolysis, due to

the less stringent reaction conditions, more simplified procedure and better molecular ion peak resolution in MS.[20] Moreover, nuclear magnetic resonance has also been extensively investigated to qualitatively and quantitatively analyze structural units and linkages of lignin component. Incipiently, one-dimensional ^{1}H- and ^{13}C-NMR were mainly used for lignin characterization, nevertheless, the problem of overlapped signals always exists. With the rapid advances in NMR technology, two-dimensional heteronuclear single quantum coherence (2D-HSQC) NMR became a powerful tool for lignin identification, which provides richer and less unambiguous information.[21]

Therefore, the purpose of the present work was to discover the integrated reaction process of cotton stalk in one-pot GVL/H_2O reaction system containing a low concentration of acid under relatively moderate pretreatment conditions through comprehensive identifications for lignin. In order to investigate the structural changes that may occurred in lignin, the isolated lignin fractions obtained under different ratios of GVL to water were comparatively characterized by Fourier transform infrared spectroscopy (FT-IR), high-performance anion exchange chromatography (HPAEC), gel permeation chromatography (GPC), 2D heteronuclear single quantum coherence nuclear magnetic resonance (2D HSQC NMR) spectroscopy, and derivatization followed by reductive cleavage (DFRC).

Experimental

Materials

Cotton stalk (*Gossypium hirsutum*) was harvested in an agricultural field in Xinjiang

Province in China. More details about the preliminary preparation for samples and the compositions of feedstock have been described in a previous literature.[22] All chemicals used were analytical grade, and sugar reference materials were purchased from Sigma-Aldrich Company (Beijing).

Recovery of lignin component

The lignin component was obtained by the procedures described in the previous literature.[18] Cotton stalk powder and GVL-H_2O mixed solution were added into a 50 mL batch reactor with magnetic stirring (model SLM-50, Sen Long Instruments Company, Beijing, China) at a solid to liquor ratio of 1:15 g/mL with addition of a trace amount of H_2SO_4 (10 mM, ~0.1%). The treatment condition was controlled at 170 ^oC under autogeneous pressure for 1 h, and the variable was the ratio of GVL to water in GVL/H_2O mixture, which was set at 90:10, 80:20, 70:30 and 60:40, respectively. The obtained reaction liquor was filtered to get the filtrate and solid residues. Then, a certain amount of sodium chloride was added into the filtrates, and the resulting solutions were repeatedly shaken and sonicated in a sonication bath until no solids were visible. At the same time, a diphasic state was formed. The mixed solutions were then centrifuged at 8000 rpm for 5 min, and the upper layer (organic-phase) was acquired. And then the lignin dissolved in the organic phase was recovered through precipitation in the acid water, which was denominated as L_1, L_2, L_3 and L_4 corresponding to the ratios of GVL to water with 90:10, 80:20, 70:30 and 60:40, respectively. Fig. 1 shows the schematic illustration of the recovery of lignin

component in the one-pot GVL/H_2O pretreatment.

Structure characterizations of the lignin fractions

FT-IR spectra of lignin samples were obtained on a Thermo Scientific Nicolet iN10 FT-IR Microscope (Thermo Nicolet Corporation, Madison, WI) equipped with a liquid nitrogen cooled MCT detector. Each spectrum was recorded in the range from 4000 to 650 cm^{-1} at 4 cm^{-1} resolution and 128 scans per sample.

Carbohydrate analysis of the lignin fractions was conducted through a high-performance anion-exchange chromatography (HPAEC) system (Dionex ICS3000, USA) with a pulsed amperometric detector and anion exchange Carbopac PA-1 column (4×250 mm). The determination of content of remaining sugars associated with lignin was achieved by hydrolysis with dilute sulfuric acid according to the standard method of the National Renewable Energy Laboratory. Calibration was performed with a standard solution of L-rhamnose, L-arabinose, L-glucose, L-galactose, D-mannose, and D-xylose.

The weight-average (M_w) and number-average (M_n) molecular weights of the lignin samples were determined by gel permeation chromatography (GPC, Agilent 1200, USA) with an ultraviolet detector (UV) at 280 nm on a PL-gel 10 mm Mixed-B 7.5 mm i.d. column, calibrated with polystyrene standards. 4 mg of lignin sample was dissolved in 2 mL of tetrahydrofuran (THF), filtered and then 20 μL solutions were injected by automatic sampler. The column was operated at ambient temperature and a flow rate of 0.5 mL/min was maintained.

2D heteronuclear single quantum coherence nuclear magnetic resonance (2D HSQC NMR) spectra were acquired on a Bruker Avance 400 MHz spectrometer fitted with a 5 mm gradient probe with inverse geometry (proton coils closest to the sample). 40 mg of lignin sample was dissolved in 0.5 mL of DMSO-d_6, and the central solvent peak at δ_C/δ_H 39.5/2.49 ppm was used as an internal reference. The standard Bruker implementations of one- and two-dimensional (gradient-selected, ^{1}H-detected HSQC) NMR experiments were used for structural characterization and assignment authentication.[23] A semi-quantitative analysis of the HSQC cross-signal intensities was also performed and the procedures were identical to those reported in the literature.[24]

Derivatization followed by reductive cleavage (DFRC) analysis was carried out according to the classic method.[19] The isolated lignin fractions (2 mg) were dissolved in 4 mL AcOH and 1 mL AcBr solution, and reacted at 50 ^{o}C for 1 h. After removal of solvent by rotary evaporation, the residue was then dissolved in 2 mL of dioxane/acetic acid/water (5:4:1, v/v/v) solution. Subsequently, zinc dust (50 mg) was added to the solution with stirring and kept for 30 min at room temperature. After that, the internal standard tetracosane (4,4'-Ethylidenebisphenol $C_{14}H_{14}O_2$, M_w=214.26) was added in the solution, which was extracted by the mixture with CH_2Cl_2 and saturated NH_4Cl. The pH of the aqueous phase was adjusted to less than 3 by adding 3% HCl, and then the organic layer was separated. The water phase was extracted twice more with CH_2Cl_2 (2×5 mL). The combined CH_2Cl_2 fractions were dried over $MgSO_4$ and the filtrate was evaporated under reduced pressure. The residue was

acetylated with 1.5 mL dichloromethane containing 0.5 mL acetic anhydride and 0.5 mL pyridine for 50 min. All volatile components were removed completely by coevaporation with isometric ethanol for several times under reduced pressure. The acetylated product dissolved with chromatographically pure CH_2Cl_2 was identified by GC-MS (Agilent 7890A/5978, USA) with a 30 mm×0.25 mm×0.25 μm capillary column (HP-5).

Results and discussion

FT-IR spectra

As it has been expected, in the GVL/H_2O system with low concentration of H_2SO_4, the acid water at high temperature could react with carbohydrates in biomass to degrade hemicelluloses and cellulose into the aqueous solution, and simultaneously, the organic solvent GVL would extract lignin fraction, initially the whole was a homogeneous solution. However, it has been discovered that the system would become biphasic if the aqueous phase contained specific solutes, such as salts and sugars. Moreover, it also has been confirmed that most of the carbohydrates could be recovered from the water phase after the formation of biphasic state. Thus, with the addition of NaCl, the biphasic system appeared, that is the dissolved carbohydrates in the water phase and the lignin in the GVL phase. The former has been detected in our previous report,[18] in order to further verify this conjectured process, the solid residue obtained by precipitation in the acid water, which was considered to be lignin, was characterized by Fourier transform infrared spectroscopy (FT-IR) to preliminarily

prove if it was lignin fraction and what was roughly in it. Fig. 2 illustrates the FT-IR spectra of these four isolated fractions according to the different ratios of GVL to water from cotton stalk, and peaks were identified by comparing their wave numbers with literature data.[25,26]

From this figure, it was obviously observed that typical characteristic peaks ascribed to lignin appeared in all these four samples at 1593, 1508, and 1420 cm^{-1}, corresponding to aromatic skeletal vibrations and at 1458 cm^{-1} attributed to the C-H deformation combined with aromatic ring vibration, which indicated the isolated fraction was indeed lignin component. Apart from these remarkable signals, other absorption peaks were also assigned. The absorption at 3397 cm^{-1} was mainly due to O-H stretching vibration of alcohol hydroxyl and phenolic hydroxyl groups in lignin fractions and residual sugars, and the band at 2935 cm^{-1} was ascribed to C-H stretching vibration in CH_2 groups. The peak at 1734 cm^{-1} on behalf of C=O stretch in unconjugated C=O groups reduced, while that at 1651 cm^{-1} arising from stretching of conjugated C=O groups strengthened with the increase of water in these four fractions, which demonstrated that more conjugated C=O groups were produced with the more water in this process. Other absorptions, such as 1324 cm^{-1} (syringyl and condensed guaiacyl units), 1269 cm^{-1} (guaiacyl units), 1123 cm^{-1} (unmistakable sign of G-S lignin), and 1030 cm^{-1} (aromatic C-H in-plane deformation vibrations) were all presented. Moreover 1219 cm^{-1} originated from the C-C, C-O, and C=O stretching (G condensed > G etherified) also appeared, and the intensity of this band in L_2 was the maximum, implying it occupied more condensed G units than other samples. On the

whole, as can be seen from the spectra, the peaks and the absorption intensity of these four samples were rather similar, indicating an analogous and initial structure of these lignin fractions.

Yields and sugar analysis

The extraction efficiencies of lignin from cotton stalk in this GVL/H_2O system are presented in Table 1. The results showed that the obtained lignin fractions represented the yields of 35.7, 64.7, 52.2 and 53.3% (%, Klason lignin) under different ratios of GVL to water in this reaction system, respectively, which significantly exceeded that of the corresponding MWL. Especially, the yield of lignin isolated under the 80/20 GVL/H_2O condition could reach 64.7%, suggesting a relatively prominent high yield in comparison with lignin extracted from other common pretreatments. This favorable result might be ascribed to the synergistic effect of acid water and relative amounts of GVL in this system. Firstly, the acid hydrolysis mainly happened on carbohydrates to remove hemicelluloses similar to a traditional acid pretreatment, and simultaneously released lignin fragments that could be extracted with organic GVL. As the ratio of GVL to water decreased from 90/10 to 80/20, the presence of more water promoted the permeation of acid into the cell wall matrix structure, thus resulting in more severe degradation of carbohydrates as reported in our previously published work [18], and therefore, more amounts of lignin component was liberated and dissolved in GVL, which would simultaneously contribute the degradation of carbohydrates as lignin fills the spaces in the cell wall between cellulose and hemicellulose just like glue.

However, with the further increase of water, the degradation rate of carbohydrates was reduced due to the weaker acid water in this system, and thus the yield of lignin component also decreased accordingly. In general, this was a dynamic equilibrium process, in which the condition GVL/H_2O 80/20 was proved to be the best condition. Therefore, it could be concluded that the 80/20 GVL/H_2O condition in this acid system is effective on recovery of lignin and corresponding high efficient production of sugars.

To verify the purity of the isolated lignin fractions, the associated polysaccharides were measured by HPAEC, and the results are shown in Table 1. As can be obviously observed, all of the samples contained a relative low amount of neutral sugars contents (1.76-6.34%) in this GVL/H_2O pretreatment, which were a little lower than that of MWL (6.74%). Xylose (1.32-5.27%) was the major monosaccharide followed by glucose (0.26-0.90%), suggesting that the associated polysaccharides in these lignin fractions originated from xylan of hemicelluloses. This phenomenon was attributed to that cellulose was mainly left in the pretreated solid residues and most of hemicelluloses were removed in the hydrolysates according to the data of our previous report,[18] thus residual hemicelluloses connected with lignin were detected. Meanwhile trace amounts of rhamnose, arabinose and galactose were also observed. In addition, the content of neutral sugars detected in L_2 was remarkably low compared with other samples, which indicated polysaccharides were most severely destroyed and removed under in this condition in accordance with previous results.[18]

Molecular weight analysis

Changes in molecular weights of these lignin samples could indirectly provide valuable insights into fragmentation and recondensation reactions of lignin during the GVL/H_2O process. Therefore, Table 2 displays the weight-average *(*M_w), number-average (M_n) molecular weights and the polydispersity (M_w/M_n) of these obtained lignin preparations, which were measured by the gel permeation chromatographic to determine the effect of this system. As shown in the table, the molecular weights of the isolated lignin fractions based on this pretreatment ranged from 614 to 899 g/mol, which was prominently low as compared to the corresponding milled mood lignin. This result suggested that lignin underwent severe depolymerization in this acid system. Another remarkable finding was that although the associated polysaccharides in L_2 were the least, the molecular weight of that was the biggest, and those of L_4 were the completely reverse, which implied that strong recondensation reactions obviously occurred in L_2 combined with the latter NMR analysis. These phenomena were owing to the degradation in the acid water at high temperature and dissolution in GVL of lignin similar to the result of a previous pretreatment,[27] which was reported that the MWL isolated after hydrothermal pretreatment was more condensed and had a lower molecular weight than MWL isolated from untreated material. Moreover, all the lignin fractions exhibited relatively narrow molecular weight distribution (M_w/M_n, 1.45-2.82%) while L_4 exhibited a relatively high polydispersity, indicating a larger amount of low molecular weight species was present as compared to other samples. The relatively high molecular

weight from lignin samples L_1, L_2 and L_3 might be due to the much more serious condensation with the formation of C-C linkages, which was demonstrated by the subsequent data in Table 5 that a smaller amount of C-C linkages (resinols and phenylcoumarans) was detected in lignin sample L_4.

DFRC and GC-MS analysis

DFRC method was also chosen to investigate the lignin structure through selective degradation of *β*-aryl ether structures on account of its more advantages than former degradative methods. In general, the DFRC involves three steps: (i) solubilization of lignin by bromination and acetylation with AcBr; (ii) reductive ether linkage cleavage; and (iii) acetylation (Fig. 3). Finally, the primary monomers derived from DFRC degradation of lignin are essentially 4-acetoxycinnamyl acetate (*p*-coumaryl peracetate, P), 4-acetoxy-3-methoxycinnamyl acetate (coniferyl peracetate, G), and 4-acetoxy-3,5-dimethoxycinnamylacetate (sinapyl peracetate, S) (Figs. 3); trans-isomers predominate.[22]

The degradation monomers from these isolated lignin fractions subjected to the DFRC method were detected by GC-MS. It was obviously observed that the monomers released from the DFRC method were mainly composed of G- and S-type monomers in each lignin samples without the presence of P-type monomers, which further confirmed the lignin component from cotton stalk was G-S type structure. Simultaneously, the predominance of G- over S-type compounds was also found in all lignin samples except L_1, implying the least *β*-aryl ether type of G units in L_1. In

addition, the content of primary degradation monomers G- and S-type monomers was calculated and the results are shown in Table 3. As can be seen from the data, the lignin fraction L_2 was more severely condensed than other samples, while the condensation of L_4 was the weakest, negatively correlated with the amounts of monomers released from these lignin fractions, in accordance with results obtained from molecular weights and NMR analysis.

2D HSQC NMR spectra

To further investigate structural changes of lignin during this process, 2D NMR analysis was performed for all lignin fractions. The representative 2D HSQC NMR spectra and the corresponding main substructures of lignin fractions are shown in Figs. 4, 5 and 6. The main cross-signals assigned according to the previous literatures[28-31] are listed in Table 4.

In the side-chain region of the lignin (δ_C/δ_H 50-90/2.5-6.0 ppm), cross-peaks of different interunit linkages were identified, such as *β*-aryl-ether (*β*-*O*-4′, A), resinol (*β*-*β*′, B), phenylcoumaran (*β*-5′, C), and cinnamyl alcohol end-groups (F), and it can also be evidently observed that, the cross-signals of methoxy groups (-OCH_3, δ_C/δ_H 55.7/3.75 ppm) and *β*-*O*-4′ substructures (A) were the most prominent, C_α and C_γ positions of which were presented at δ_C/δ_H 71.6/4.81 and 59.9/3.28-3.61 ppm. C_β positions of *β*-*O*-4′ in G and S type lignins were also detected at δ_C/δ_H 83.7/4.30 and 85.9/4.11 ppm, respectively. In certain lignin samples, a small but clear signal at δ_C/δ_H 64.7/3.98 ppm assigned to C_γ-H_γ correlations in *γ*-oxidized lignin units (A′/A″) was

also identified. In addition, obvious signals for resinol structures (β-β' linkages, B) were observed in the spectra, with their C-H correlations for α-, β- and double γ-C positions at δ_C/δ_H 84.9/4.68, 53.5/3.01 and 70.9/3.77 and 4.15 ppm, respectively. Phenylcoumaran substructures (β-5′ linkages, C) were also recognized by the signals of C_α-H_α, C_β-H_β and C_γ-H_γ correlations at δ_C/δ_H 87.0/5.45, 53.0/3.47 and 62.8/3.65, respectively. Moreover, weak cinnamyl alcohol end-groups (F) (δ_C/δ_H 61.5/4.08) were also visible in the spectra. Simultaneously, minor amounts of xylan moieties (X) were observed in the side chain region, and the intensity of these signals was the strongest in L_4 and weakest in L_2, which indicating the associated polysaccharides in these lignin fractions were the least in L_2 and most L_4, in agreement with the aforementioned sugar analysis. Prominently, an evident signal at δ_C/δ_H 76.5/4.60 ppm was detected in each sample, which was ascribed to the residual GVL solvent referred to the pertinent literature.[32]

In the aromatic region (δ_C/δ_H 95-130/6.0-7.8 ppm), the main cross-signals from the aromatic rings of guaiacyl (G) and syringyl (S) were clearly observed in the spectra, suggesting that the lignin is G-S type lignin consistent with the previous literature.[27] Specifically, the normal S units showed a prominent signal for $C_{2,6}$-$H_{2,6}$ correlation at δ_C/δ_H 103.5/6.01, while that of the C_α-oxidized structure of syringyl units (S′) was identified at δ_C/δ_H 106.0/7.07. The G units exhibited different correlations for C_2-H_2 (δ_C/δ_H 110.9/6.90), C_5-H_5 (δ_C/δ_H 115.0/6.73), and C_6-H_6 (δ_C/δ_H 118.7/6.78), respectively. Meanwhile, the signal assigned to C_2-H_2 correlation in C_α-oxidized G units (G′) was present in each spectrum at δ_C/δ_H 109.0/7.15 ppm. It can

also be seen that all lignin samples showed different degrees of condensation in the spectra.

Semiquantitative 2D HSQC spectra analysis

According to the semiquantitative methods in the literature,[33] the relative amounts of interunit linkages and S/G molar ratios present in all lignin samples were calculated as shown in Table 5. As shown from the quantitative NMR data, the main lignin sub-structure was *β*-*O*-4′ aryl ether (58.2-72.0% of total side chains), followed by low amounts of phenylcoumaran (*β*-5′) and resinol (*β*-*β*′). In addition, by comparing the proportions of *β*-*O*-4′ in all lignin fractions, it was found that the decreased order of *β*-*O*-4′ linkages was $L_2>L_1>L_3>L_4$, which suggested the different degrees of degradation. The highest amount of *β*-*O*-4′ linkages in lignin sample L_4 (Table 5) was possibly attributed to the less serious degradation due to the relatively weak treatment conditions, which was consistent with the above results from DFRC. However in contrast, the amounts of carbon-carbon linkages (C-C, *β*-*β*′ and *β*-5′) were elevated as the above order, which may be attributed to more condensed lignin structures. Combined with the results of molecular weight (M_w: $L_2>L_1>L_3>L_4$), it could be deduced that this process took place severe condensation accompanied by degradation. Undoubtedly, it was discovered that the strongest depolymerization and repolymerization reactions simultaneously happened in L_2 fraction consistent with each of the aforementioned analysis. Besides *β*-*O*-4′ aryl ether linkage cleavage and carbon-carbon (C-C) linkage degradation or condensation, the change of S/G ratio

was also prominent structural alterations observed after treatment. From the obtained data, the S/G ratios of these four lignin samples suggested the G units dominated in the isolated lignin, and those were similar to that of the corresponding MWL,[22] implying the removing order of G and S units in the conditions given was almost the same with MWL.

Conclusion

The effect of one-pot GVL/H_2O system pretreatment on fractionation of lignin component and its structural changes was comparatively investigated. During this process, lignin could be obtained with no need of additional step after acid treatment on account of the presence of GVL, which represented an advantage over other acid pretreatments. In addition, lignin fractions separated from this system showed relative high yields and purities. As shown from results of molecular weight, DFRC and 2D HSQC NMR analysis, the isolated lignin fractions mainly consisted of noticeable amounts of guaiacyl followed by syringyl units, and remarkable depolymerization and repolymerization simultaneously occurred in lignin, especially outstanding in the sample obtained under GVL/H_2O 80/20. Therefore, in order to realize the biorefinery process, the acid GVL/H_2O system with the condition of GVL/H_2O 80/20 exhibited a simple and efficient method to simultaneously extract lignin components and sugars from cotton stalk for further transformation and application.

Acknowledgements

We are grateful for the financial support of this research from the Fundamental Research Funds for the Central Universities (BLYJ201622), Natural Science Foundation of China (No. 31470606), Program for New Century Excellent Talents in University (NCET-12-0782), China National Funds for Distinguished Young Scientists (31225005) and China Ministry of Education (No. 111).

References

1 Z. Strassberger, P. Prinsen, F. V. D. Klis, D. S. V. Es, S. Tanase and G. Rothenberg, *Green Chem.*, 2015, **17**, 325–334.

2 A. Pinkert, D. F. Goeke, K. N. Marsh and S. S. Pang, *Green Chem.*, 2011, **13**, 3124–3136.

3 Z. Strassberger, S. Tanase and G. Rothenberg, *RSC Adv.*, 2014, **4**, 25310–25318.

4 G. Chatel and R. D. Rogers, *ACS Sustain. Chem. Eng.*, 2014, **2**, 322–339.

5 H. Lange, S. Decina and C. Crestini, *Eur. Polym. J.*, 2013, **49**, 1151–1173.

6 Q. Bu, H. Lei, L. Wang, Y. Wei, L. Zhu and X. Zhang, *Bioresour. Technol.*, 2014, **162**, 142–147.

7 P. Carrott and M. Ribeiro Carrott, *Bioresour. Technol.*, 2007, **98**, 2301–2312.

8 J. Kadla, S. Kubo, R. Venditti, R. Gilbert, A. Compere and W. Griffith, *Carbon*, 2002, **40**, 2913–2920.

9 H. Chung and N. R. Washburn, *ACS Appl. Mater. Inter.*, 2012, **4**, 2840–2846.

10 E. C. Achinivu, R. M. Howard, G. Li, H. Gracz and W. A. Henderson, *Green Chem.*, 2014, **16**, 1114–1119.

11 N. Mosier, C. Wyman, B. Dale, R. Elander, Y. Y. Lee, M. Holtzapple and M. Ladisch, *Bioresour. Technol.*, 2005, **96**, 673–686.

12 K. Wang, H. Y. Yang, X. Yao, F. Xu and R. C. Sun, *Bioresour. Technol.*, 2012, **116**, 99–106.

13 L. Camesasca, M. B. Ramírez, M. Guigou, M. D. Ferrari and C. Lareo, *Biomass Bioenerg.*, 2015, **74**, 193–201.

14 J. S. Luterbacher, J. M. Rand, D. M. Alonso, J. Han, J. T. Youngquist, C. T. Maravelias, B. F. Pfleger and J. A. Dumesic, *Science*, 2014, **343**, 2772–2780.

15 L. Qi, Y. F. Mui, S. W. Lo, M. Y. Lui, G. R. Akien and I. T. Horváth, *Acs Catal.*, 2014, **4**, 1470–1477.

16 I. T. Horváth, H. Mehdi, V. Fábos, L. Boda and L. T. Mika, *Green Chem.*, 2008, **10**, 238–242.

17 S. G. Wettstein, D. M. Alonso, Y. Chong and J. A. Dumesic, *Energy Environ. Sci.*, 2012, **5**, 8199–8203.

18 M. Wu, Z. Y. Yan, X. M. Zhang, F. Xu and R. C. Sun, *Bioresour. Technol.*, 2016, **200**, 23-28.

19 F. Lu and J. Ralph, *J. Agric. Food Chem.*, 1997, **45**, 2590–2592.

20 J. S. Lupoi, S. Singh, R. Parthasarathi, B. A. Simmons and R. J. Henry, *Renew. Sust. Energ. Rev.*, 2015, **49**, 871–906.

21 J. L. Wen, S. L. Sun, B. L. Xue and R. C. Sun, *Materials*, 2013, **6**, 359–391.

22 M. Wu, D. Zhao, J. Pang, X. Zhang, M. Li, F. Xu and R. Sun, *Ind. Crop. Prod.*, 2015, **66**, 123–130.

23 J. L. Wen, T. Q. Yuan, S. L. Sun, F. Xu and R. C. Sun, *Green Chem.*, 2014, **16**, 181–190.

24 S. M. Kang, L. P. Xiao, L. Y. Meng, X. M. Zhang and R. C. Sun, *Int. J. Mol. Sci.*, 2012, **13**, 15209–15226.

25 O. Faix, *Holzforschung*, 1991, **45**, 21–28.

26 M. Wu, J. Pang, X. Zhang and R. Sun, *Int. J. Polym. Sci.*, 2014, 1–10.

27 L. P. Xiao, Z. J. Shi, F. Xu and R. C. Sun, *Holzforschung*, 2012, **66**, 295–302.

28 J. Ralph, R. D. Hatfield, S. Quideau, R. F. Helm, J. H. Grabber and H. J. G. Jung, *J. Am. Chem. Soc.*, 1994, **116**, 9448–9456.

29 J. Rencoret, G. Marques, A. Gutierrez, D. Ibarra, J. Li, G. Gellerstedt, J. I. Santos, J. Jimenez-Barbero, A. T. Martinez and J. C. del Rio, *Holzforschung*, 2008, **62**, 514–526.

30 J. C. Del Rio, J. Rencoret, G. Marques, A. Gutierrez, D. Ibarra, J. I. Santos, J. Jimenez-Barbero, L. M. Zhang and A. T. Martinez, *J. Agric. Food Chem.*, 2008, **56**, 9525–9534.

31 J. C. Del Rio, J. Rencoret, A. Gutierrez, L. Nieto, J. Jimenez-Barbero and A. T. Martinez, *J. Agric. Food Chem.*, 2011, **59**, 11088–11099.

32 A. S. Amarasekara, B. Wiredu and D. N. Edwards, *Biomass Bioenerg.*, 2015, **72**, 39–42.

33 J. Rencoret, G. Marques, A. Gutierrez, L. Nieto, J. Jimenez-Barbero, A. T. Martinez and J. C. del Rio, *Ind. Crop. Prod.*, 2009, **30**, 137–143.

Table 1 Yields (%, Klason lignin) and contents of neutral sugars (relative % dry weight, w/w) of lignin fractions separated from cotton stalk based on the GVL/H_2O system

	Lignin fractions [a]				
	L_1	L_2	L_3	L_4	MWL
Yield (%)	35.7	64.7	52.2	53.3	19.9
Sugars (%)					
Rhamnose	ND [b]	0.1	ND	ND	0.09
Arabinose	0.10	0.01	0.11	0.15	0.09
Galactose	0.15	0.07	0.21	0.28	0.09
Glucose	0.90	0.26	0.48	0.64	2.04
Xylose	3.40	1.32	4.20	5.27	4.43
Total	4.55	1.76	5.00	6.34	6.74

[a] Represents the lignin fractions extracted with GVL to water at ratios of 90:10, 80:20, 70:30 and 60:40, respectively.

[b] ND=not detected.

Table 2 Weight-average (M_w), number-average (M_n) molecular weights and polydispersity (M_w/M_n) of the lignin fractions

	Lignin fractions [a]				
	L_1	L_2	L_3	L_4	MWL
M_w	840	870	730	610	1570
M_n	560	620	490	220	770
M_w/M_n	1.52	1.45	1.53	2.82	2.03

[a] Corresponding to the lignin fractions in Table 1.

Table 3 Yields of the main G and S monomers recovered from DFRC degradation of the isolated lignin fractions

Lignin fractions [a]	DFRC yield (% wt)	molar yield (μmol/g)	rel distribution G:S
L_1	66.42	2364	0.79:1
L_2	19.92	721	1.48:1
L_3	33.60	1214	1.38:1
L_4	74.27	2669	1.13:1

[a] Corresponding to the lignin fractions in Table 1.

Table 4 Assignments of ^{13}C-^{1}H correlation signals in the HSQC NMR spectrum of the obtained lignin fractions from cotton stalk

Lables	δ_C/δ_H(ppm)	Assignment
C_β	53.0/3.47	C_β-H_β in phenylcoumaran substructures (C)
B_β	53.5/3.01	C_β-H_β in β-β' (resinol) substructures (B)
Me*O*	55.7/3.75	C-H in methoxyls
A_γ	59.9/ 3.28-3.61	C_γ-H_γ in β-*O*-4′ substructures (A)
C_γ	62.8/3.65	C_γ-H_γ in phenylcoumaran substructures (C)
F_γ	61.5/4.08	C_γ–H_γ in cinnamyl alcohol end-groups (F)
A'_γ/A''_γ	64.7/3.98	C_γ-H_γ in γ-acylated/γ-*p*-coumaroylated β-*O*-4′ substructures (A′/A″)
B_γ	70.9/3.77 and 4.15	C_γ-H_γ in β-β' resinol substructures (B)
A_α	71.6/4.81	C_α-H_α in β-*O*-4′ substructures linked to an S units (A)
$A_{\beta(G)}$	83.7/4.30	C_β-H_β in β-*O*-4′ substructures linked to G and H units (A)
B_α	84.9/4.68	C_α-H_α in β-β' (resinol) substructures (B)
$A_{\beta(S)}$	85.9/4.11	C_β-H_β in β-*O*-4′ substructures linked to S units (A)
C_α	87.0/5.45	C_α-H_α in phenylcoumaran substructures (C)
$S_{2,6}$	103.5/6.01	$C_{2,6}$-$H_{2,6}$ in S units (S)
$S'_{2,6}$	106.0/7.07	$C_{2,6}$-$H_{2,6}$ in C_α-oxidized S units (S′)
G_2	110.9/6.90	C_2-H_2 in G units (G)
G'_2	109.0/7.15	C_2-H_2 in C_α-oxidized G units (G′)
G_5	115.0/6.73	C_5-H_5 in G units (G)
G_6	118.7/6.78	C_6-H_6 in G units (G)

Table 5 Structural characteristics (inter-unit linkages and S/G ration) from different lignin fractions by integration of ^{13}C-^{1}H correlation signals in the HSQC spectra

Linkage relative abundance (%)	L_1	L_2	L_3	L_4
β-*O*-4′ aryl ethers (A)	62.3	58.2	65.7	72.0
resinols (*β*-*β*′, B)	18.0	17.9	16.3	14.2
phenylcoumarans (*β*-5′, C)	19.7	23.9	18.1	13.8
S/G ratio	0.39	0.44	0.41	0.44

Figure Captions

Fig. 1 Schematic illustration of the one-pot GVL/H_2O pretreatment.

Fig. 2 FT-IR spectra of the isolated lignin fractions from cotton stalk (L_1: GVL/H_2O 90/10; L_2: GVL/H_2O 80/20; L_3: GVL/H_2O 70/30; L_4: GVL/H_2O 60/40).

Fig. 3 Selective ether cleavage in lignin by DFRC method.

Fig. 4 Side chain (δ_C/δ_H 45-90/2.6-6.0 ppm) regions in the 2D HSQC NMR spectra of these four lignin fractions (L_1: GVL/H_2O 90/10; L_2: GVL/H_2O 80/20; L_3: GVL/H_2O 70/30; L_4: GVL/H_2O 60/40).

Fig. 5 Aromatic (δ_C/δ_H 95-130/6.0-7.8 ppm) regions in the 2D HSQC NMR spectra of these four lignin fractions (L_1: GVL/H_2O 90/10; L_2: GVL/H_2O 80/20; L_3: GVL/H_2O 70/30; L_4: GVL/H_2O 60/40).

Fig. 6 Main substructures present in the isolated lignin from cotton stalk: (A) *β-O*-4′ alkyl-aryl ethers; (A′) *β-O*-4′ linkages with acetylated at *γ*-carbon; (A″) *γ-p*-coumaroylated *β-O*-4′ linkages; (B) resinol; (C) phenylcoumaran; (F) cinnamyl alcohol end-groups; (G) guaiacyl unit; (G′) oxidized guaiacyl units with a carbonyl group at C_α; (S) syringyl unit; (S′) oxidized syringyl units with a carbonyl group at C_α; (X) xylose.

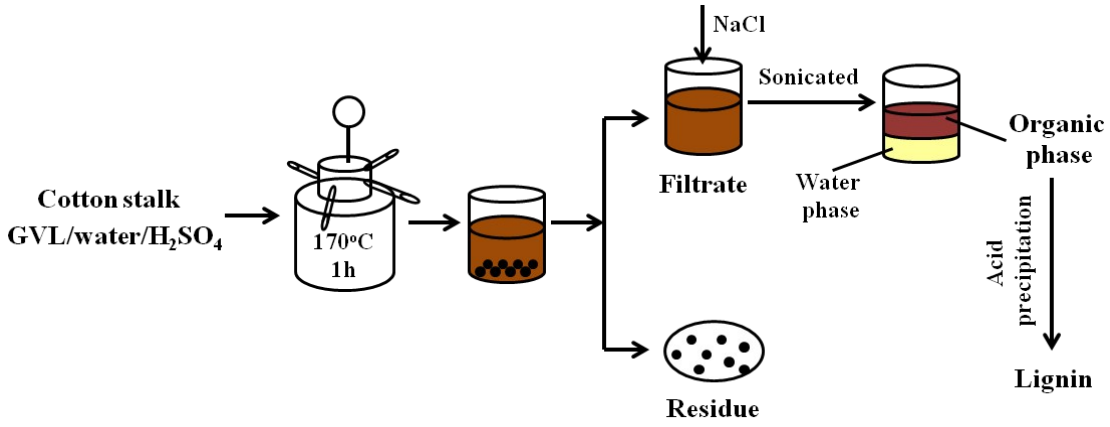

Fig. 1

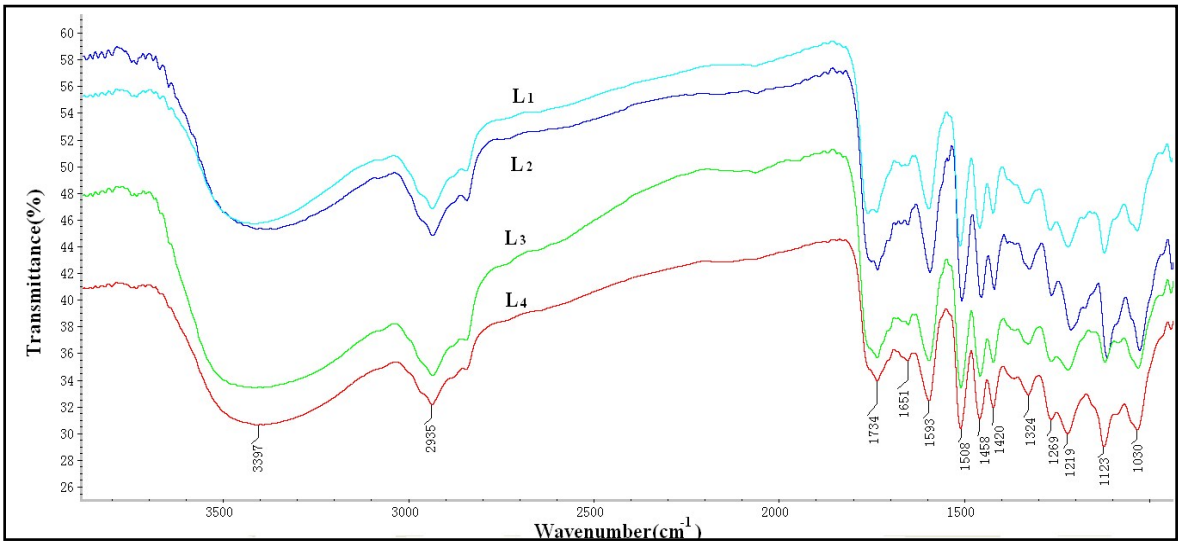

Fig. 2

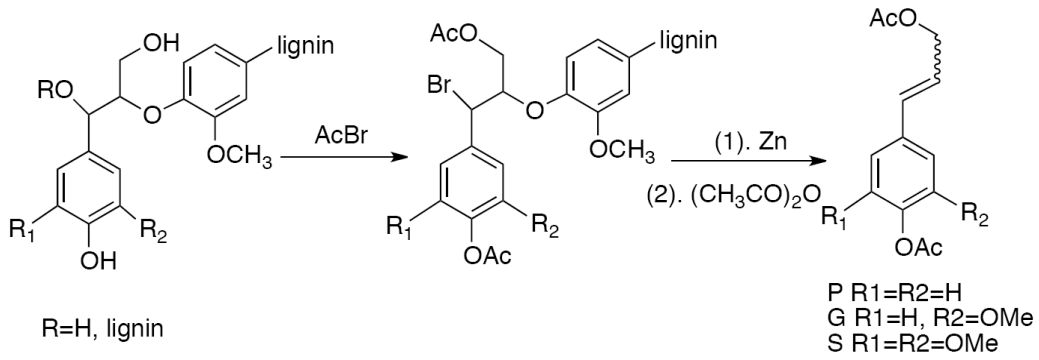

Fig. 3

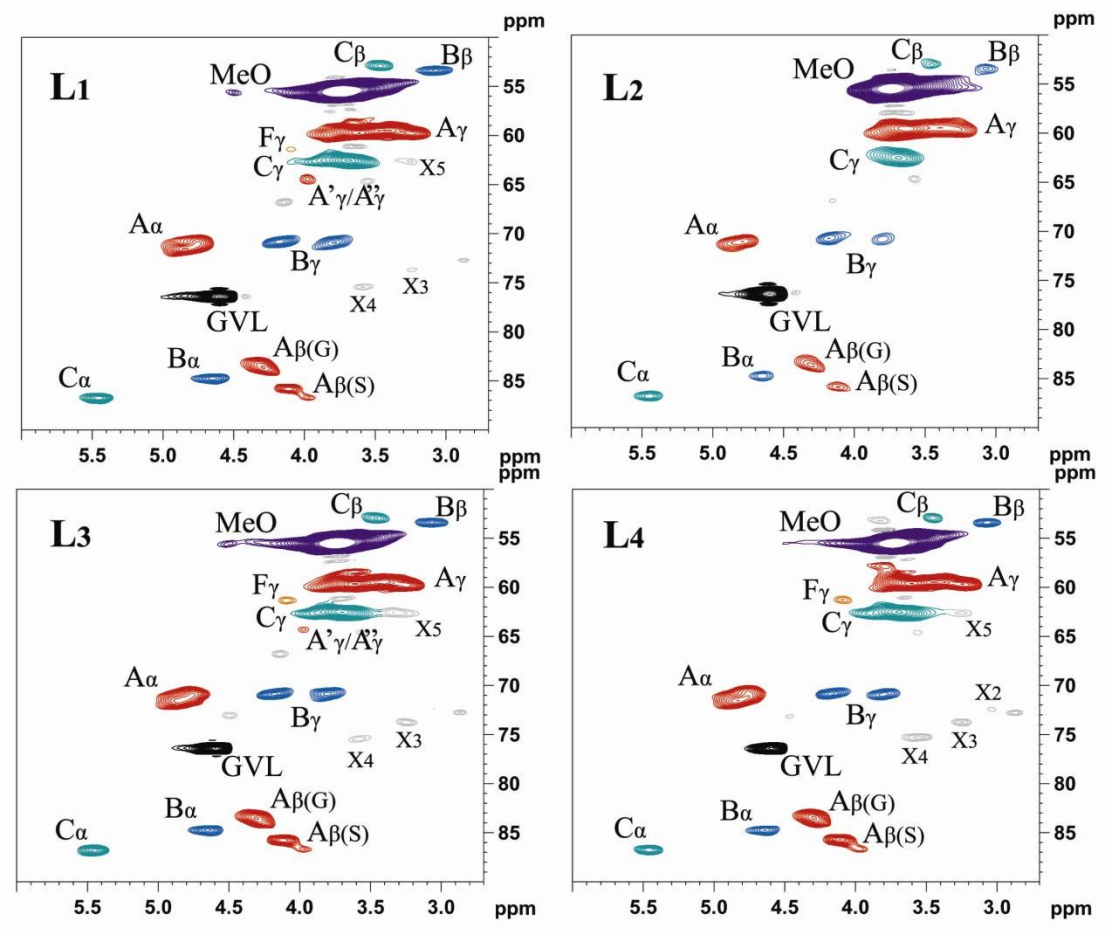

Fig. 4

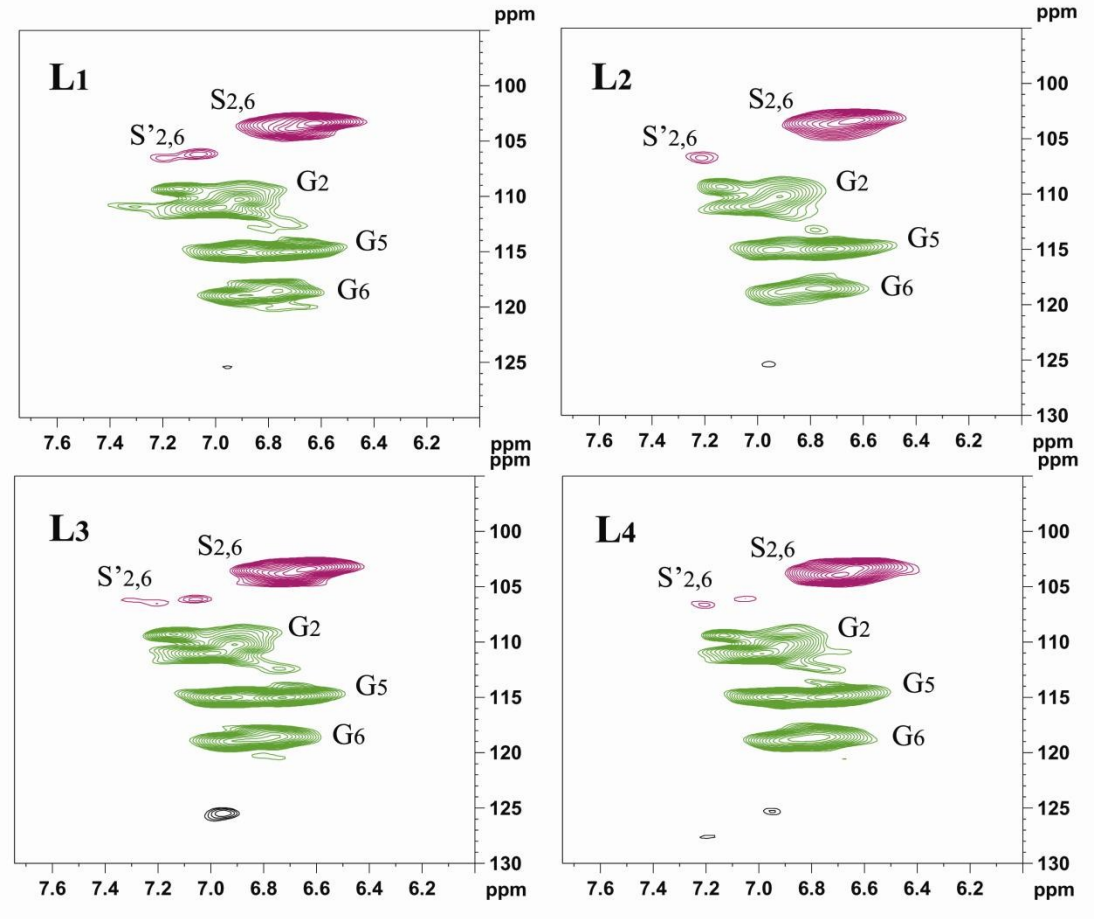

Fig. 5

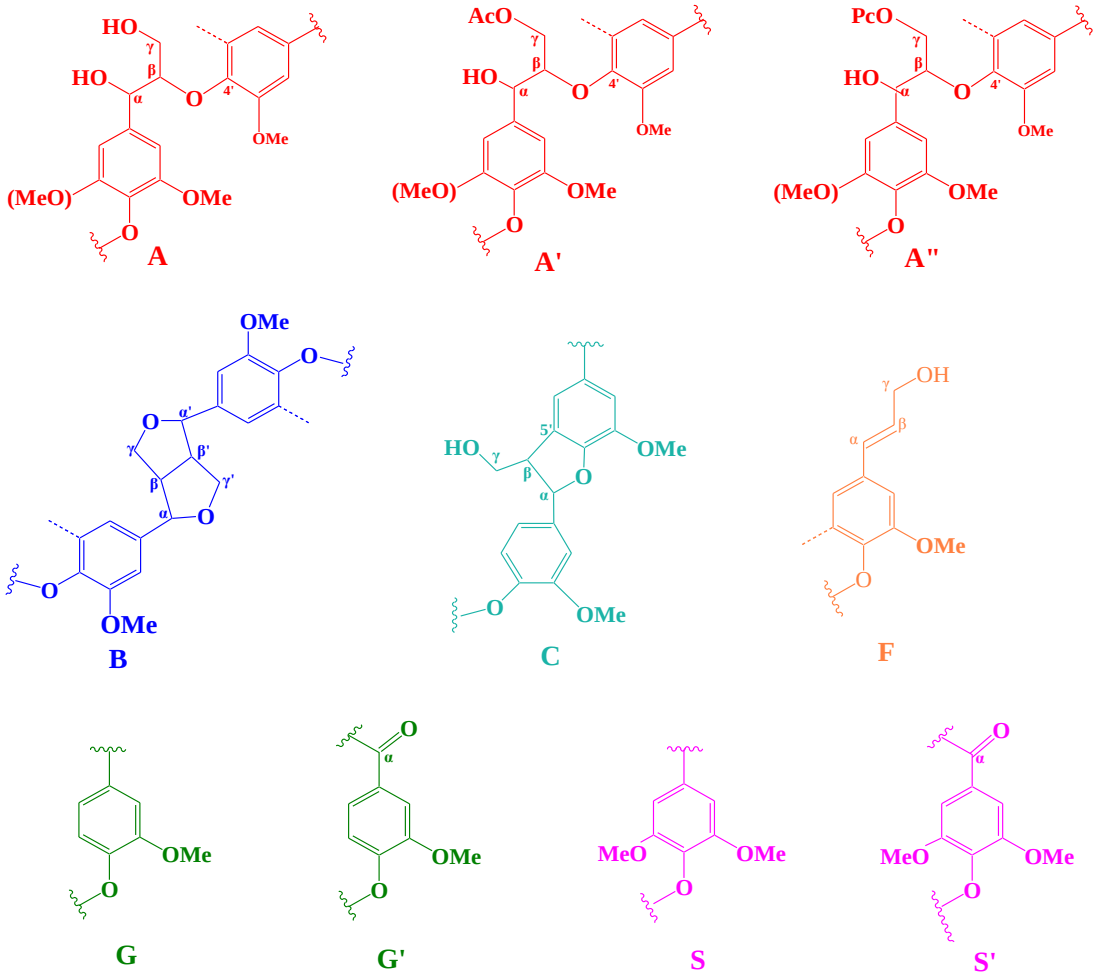

Fig. 6

Graphical Abstract

Lignin component (35.7-64.7%) was recovered by the pretreatment using γ-valerolactone/water system under different ratios of GVL to water.