RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

Comprehensive structural analysis of biorefinery lignins with a quantitative ^{13}C NMR approach

Mikhail Yu. Balakshin,* Ewellyn A. Capanema

Renmatix Inc., 660 Allendale Rd., King of Prussia, PA 19406, USA

*Corresponding Author: mikhail.balakshin@renmatix.com

Abstract

The advance in analytical methodology is critical for the progress in the biorefinery and lignin commercialization. This paper reports a comprehensive approach (more than 30 common structural characteristics along with moieties specific for various lignin types) for the analysis of biorefinery lignins with quantitative ^{13}C NMR spectroscopy, which has been demonstrated to be significantly different from the analysis of native lignins. Experimentals required for high precision NMR spectra are highlighted. The statistic data allowed for evaluating the accuracy in the quantification of different lignin units for the first time. The analysis of various lignins originated from the key biorefinery processes of different types of biomass clearly demonstrated that, in general, lignin degradation was always accompanied with decreases in aliphatic OH (primary and especially secondary ones), oxygenated aliphatic moieties, specifically β–O-4 units, and increasing amounts of phenolic OH, COOR, saturated aliphatic moieties and the degree of condensation. However, the differences in the quantity of different functionalities between the lignins investigated were very significant. Hardwood steam explosion lignins were the less degraded ones whereas aspen kraft lignin underwent the most severe structural modification. Finally, this report presents a comprehensive database on the structure of reference biorefinery lignins that is of primary importance for their commercialization.

Key words: biorefinery lignins, kraft lignins, lignin structural analysis, organosolv lignins, quantitative ^{13}C NMR.

1 Introduction

Lignin is the second most abundant natural polymer on Earth, contributing up to approximately 30% of the weight of lignocellulosic biomass. Lignin's chemical structure suggests that it can play a central role as a new chemical feedstock, particularly in the formation of polymeric materials and aromatic chemicals replacing petroleum-derived products with renewable green products, providing sustainability and decreasing greenhouse gas (GHG) emissions.[1] Therefore, lignin offers a significant opportunity for enhancing the economics of lignocellulosic biorefineries based on traditional pulping technologies (such as kraft, soda and sulfite pulping) as well as emerging biorefinery processes. The amount of biofuel necessary to meet the Department of Energy's (DOE) targeted reduction in gasoline usage would result in ~225 million tons of lignin as a co-product in the USA only.[1] As the result, exploring high value end uses of lignin such as adhesives, carbon fibers, thermoplastics, etc. is booming. Lignin commercialization is impossible without strong methodology for comprehensive lignin analysis, including development of advanced analytical techniques as well as validation and standardization of the new and traditional methods.[1-2]

Lignin is a heterogeneous aromatic polymer comprising of C9-units of the p-hydroxy phenyl (H-), guaiacyl (G-) and syringyl (S)-types attached to each other with different types of C-O and C-C linkages[3] (Fig. 1). The most popular method in lignin analysis nowadays is a combination of semi-quantitative 2D Heteronuclear Single Quantum Coherence (HSQC) NMR and ^{31}P NMR.[4] The 2D NMR provides great advantage in separation of signals of a very complex lignin macromolecule.[5] It is very successful in comprehensive analysis of various native lignins when lignin structure can be well described by the quantification of about 10 structural units.[5-6] 2D

NMR was also very beneficial in identification and quantification of newly formed substructures of technical lignins (lignins obtained during thermo- and/or chemical processing of plant biomass).[2,7] However, the structure of technical lignins is extremely heterogeneous; HSQC NMR spectra of technical lignins acquired in a most advanced NMR machine (a Bruker 950 MHz spectrometer equipped with a CryoProbe™) provide with a few hundreds signals.[8] However, most of them are of very low intensity and close to the detection limit. Identification and quantification of all these moieties on a structural level is a tremendous task, but unlikely has a practical rational. Direct application of the HSQC approach used for native lignin to technical ones may cover as low as only 2.5% of lignin structural units.[4] The analysis of newly formed moieties in technical lignins, in a very semi-quantitative mode, increases the amount of quantified units to about 20-30%,[2] but still leaving majority of lignin uncharacterized. Due to such heterogeneity of technical lignins, it makes more sense to characterize them on the functional level quantifying specific lignin functionalities as groups rather than attempting to compute each individual lignin substructures. Apparently, 2D NMR is not capable to provide this analysis. ^{31}P NMR was rather successful in the analysis of low molecular aromatics (such as bio-oil)[9] but, when used for polymeric lignins, provides with only 4-5 values of different OH moieties.[10] Although OH groups are among major lignin functionalities, this information is clearly insufficient to describe the whole lignin structure.

In contrast, ^{13}C NMR covers all lignin functionalities[11] and can overcome the current gap in the structural information. Unfortunately, the high potential of quantitative ^{13}C NMR in lignin analysis is not utilized sufficiently. ^{13}C NMR allows for identification and discussion of more than 80 different signals in lignin spectra on a qualitative base,[5,12] which is an order of magnitude higher than that allowed by ^{1}H and ^{31}P NMR methods. A vast database on the chemical shifts of various lignin model compounds has been generated.[11-13] However, most of recent publications on quantitative ^{13}C NMR analysis for technical lignins report structural data only for a few lignin moieties, that is much less than even the original report.[11] Thus, a common approach strongly

underutilizes the potential of the ^{13}C NMR method. Partially, it can be explained by challenges associated with a very complex nature of spectra and significant signals overlap.

Recently we reported an advanced NMR methodology allowing for a comprehensive, very reliable and reasonably fast characterization of softwood and hardwood milled wood lignins (MWL).[14,15] Relatively simple structures of lignin investigated allowed for establishing a validation baseline and correlation between different resonances in the spectra of MWLs. A similar approach should be applied for the analysis of technical lignins. However, as technical lignins are significantly modified and much more heterogeneous as compared to native lignins, their spectra look considerably different from spectra of native lignins and modification of the quantification algorithm is needed.

Thus, the objectives of this research were to develop a comprehensive method for analysis of softwood (SW), hardwood (HW) and non-wood technical lignins and to compare then lignins derived from key biorefinery technologies. Special attention was made to address some important details of the experimental protocol required for accurate and reproducible spectra acquisition and processing which were not described in the literature. For this purpose we used well known lignin standards based on the US Department of Energy (DOE) and International Lignin Institute (ILI) selection of most realistic biorefinery processes.[16]

2 Experimental

The following lignin preparations were used: Alcell (HW organosolv), Indulin, Curan (both SW kraft), organosolv Doulas Fir (OS-DF), aspen kraft lignin (AKL), soda bagasse lignin (SBL), Sucrolin (acid hydrolysis bagasse lignin), steam explosion aspen and poplar lignins (SEAL and SEPL). An aspen and pine milled wood lignins (AMWL and PMWL) were used for comparison. The detailed lignins histories were reported elsewhere [10,16,17] and summarized in Supplemental.

The NMR spectra were recorded on a Bruker AVANCE 500 MHz spectrometer at 300 K using a dedicated ^{13}C NMR probe. 190 – 210 mg of lignin was dissolved in 0.55 ml of DMSO-*d6*

contained a relaxation reagent, chromium (III) acetylacetonate (0.016M), and an internal standard (IS), trioxane (IS:Lignin ratio was 1:10, w/w). Inverse gate detection and a 90^{o} pulse width were used for the quantitative ^{13}C NMR acquisitions. T1 experiment was run to ensure acquisition conditions requested for the quantitative NMR measurement, especially for IS. To ensure an accurate baseline, the spectra were recorded in the interval of 240-(-40) ppm., About 20,000 scans were collected.

The spectra were Fourier transformed, phased, calibrated and the baseline was manually corrected by using a polynomial function. The correction of baseline was done using the following approximate interval ranges to be adjusted to zero: (220-215ppm)-(185-182ppm)-(97-94ppm)-(5-(-20) ppm). No other regions were forced to 0.

The aromatic region (about 100-163 ppm) in a ^{13}C NMR spectrum was integrated, and this integral set to a value of 600. Subsequent integration of the regions of interest in this spectrum would now be in the units of "per 100 Ar." (Note: the exact chemical shift value is determined based on the local minima in the spectra and can be slightly variables from the numbers below.)

The calculation of the quantity of specific groups in mmol/g lignin was done as follows:

For non-acetylated lignins:

$$X\ (\text{mmol/g lignin}) = I_X * m_{IS} / (30 m_{Lig} * I_{IS}) * 1000$$

For acetylated lignins (recalculated per original non-acetylated lignin):

$$X\ (\text{mmol/g lignin}) = I_X * m_{IS} / (30 m_{Lig} * I_{IS} - 42 * I_{Ac} * m_{IS}) * 1000$$

Where X is the amount of the specific moiety; I_X, I_{IS} and I_{Ac} are the resonance values of the specific moiety, the internal standard and total Ac groups (corresponding to total OH), correspondingly; m_{Lig} and m_{IS} are the masses of the lignin and internal standard; 30 is the equivalent mass of the IS (M=90 g mol^{-1} with three equivalent carbons resonating at about 92 ppm) and 42 is the increment in the mass of lignin after acetylation of each OH group.

Alternatively, the values in mmol/g lignin could be obtained from the original numbers in units/100Ar using the "molecular weight" of an average lignin monomeric unit (M) (Table 1) as:

$X\ (mmol/g\ lignin) = X\ (units/100Ar)/M_{Ar}*10$

Where

$M_{Ar} = 0.3m_{Lig} * I_{IS}/ m_{IS}$ for non-acetylated lignins and

$M_{Ar} = 0.3m_{Lig} * I_{IS}/ m_{IS} - 0.42\ I_{Ac}$ for acetylated lignin.

Molecular weights were determined by size exclusion chromatography (SEC) performed on an Agilent 1260 ultra HPLC, equipped with refractive index and ultraviolet (280 nm) detectors using 0.1 M NaOH at the flow rate of 0.5 mL/min as the mobile phase.[8] The column set employed three sulfonated polystyrene-divinylbenzene PSS MCX columns (a pre-column, a 1000 Å column, and a 100 000 Å column, Polymer Standards). Six different polystyrene standards ranging from 890 g/mol to 65,400 g/mol were used for calibration.

3 Results and discussion

3.1 Aromatic Ring (Ar) and Internal Standard (IS) as quantification references

Importantly, the ^{13}C NMR method with IS[18] allows for two simultaneous modes of direct evaluation of lignin moieties, in units/100 Ar (aka mol%) and in mmol/g lignin,[10,15] in contrast to other methods on structural lignin analysis. The former is very useful in understanding the fundamentals of lignin structures and different reaction mechanisms of lignin transformation. Another advantage of this approach is independence on contaminants in lignin samples, such as carbohydrate, extractives and ash in contrast to the mmol/g mode which is directly dependent on the lignin purity. The mmol/g data is more practical for industrial lignin applications. In addition, it also allows for correlation of the ^{13}C NMR results with those obtained by wet chemistry and ^{31}P NMR methods.[10,15] Therefore, we use the calibration per 100Ar to obtain the main data and the calculations using IS for additional information.

A serious drawback of the original ^{13}C NMR protocol with internal standard was a very long experimental time (as compared to a regular ^{13}C NMR of lignin) to assure complete relaxation of IS.[18] Recently, we optimized the procedure by decreasing the experimental time by 4-fold[10] making this method much more affordable for a routine use.

The calculations of different lignin functionalities are summarized in Table 1.

3.2 Method accuracy

3.2.1 Experimental factors affecting the accuracy of the method

In addition to well defined requirements for quantitative ^{13}C NMR, i.e. 90^0 pulse, suppression of NOE, and complete relaxation of all nuclei by optimization of the pulse delay (pulse delay > $5T_{1max.}$),[11] there are other experimental issues (specifically for lignins), which strongly affect the accuracy of the spectra acquisition and processing and therefore the final quantification values, but are not sufficiently described in the literature. Very importantly, the acquisition parameters should be optimized to obtain a best possible raw baseline. From our experience, the most important parameter affecting the baseline shape is the pre-scan delay (DE). Examples of a good and a bad raw baseline are shown in Fig. S1. In particular, the use of CryoProbe™ technology allows a spectrum with very high resolution and good signal-to-noise (S/N) ratio within only 1 hour, but very unfortunately the raw baseline is very complex and biased that makes accurate quantification extremely difficult.

The correction of the raw baseline should be done as described in Experimental, preferably with one step through the whole spectrum. Point correction is very ambiguous for complex lignin samples and should be avoided as well as automatic baseline correction.

Another important factor strongly affecting the accuracy is an appropriate S/N ratio. To evaluate the S/N ratio quantitatively, we use the ratio between the signal resonance at 163-98 ppm to the noise level at 0-(-10) ppm. Our experience shows that the S/N ratio calculated by this way should be above 200. To achieve it, a few issues should be considered. First, a direct detection NMR

probe should be used, preferably a dual $^{1}H/^{13}C$ probe. Broad-band multinuclear probes can be also used, although their sensitivity is somewhat lower than that of a "dedicated" probe. Second, a rather high concentration of lignin in an NMR solvent should be used for the quantitative ^{13}C NMR experiment; the concentration of technical lignins should be about 350-400 mg per 1 ml of DMSO, that is significantly higher than we used for milled wood lignins (MWL) earlier.[14] Finally, a sufficient number of scans (NS) should be collected. For a routine 500 MHz Bruker NMR spectrometer (without CryoProbe™), NS should be at least 18,000 – 20,000. Importantly, if the lignin concentration is low, good S/N ratio cannot be achieved even with large NS (usually, no significant improvement is observed after acquiring more than 25,000 scans). Following these recommendations allows for accurate and reproducible experimental data.

Furthermore, although the most accurate quantification is directly from the resonance values without any calculations, the latter often provides additional valuable information. However, it is of primary importance to choose an appropriate calculation way to minimize error from data manipulations or at least clearly realize when the data are semi-quantitative. For example, calculations of 5-5' structures (Fig. 1, Structure **T**) in softwood MWLs[14] is rather semi-quantitative whereas calculations of β–1 moieties in HW MWLs appears to be very inaccurate[15] and should be avoided. On the other hand, certain calculation and correction suggested in Table 1 are accurate enough and allows for reproducible data (Table 2).

3.2.2 Reproducibility

Although the quantitative ^{13}C NMR method is used in lignin analysis for more than 30 years, information of its accuracy and reproducibility is very limited.[5] It could be explained by long experimental time (more than 70h) required in traditional quantitative ^{13}C NMR analysis of lignin,[11] which makes replicate statistics difficult. The modified protocol[13,14] dramatically reduces the experimental time and thus makes statistical evaluation more affordable.

Replicate experiments (including sample preparation, NMR acquisition and processing) were performed for selected lignin samples. Importantly, the Alcell and Indulin lignin samples obtained from different sources showed very similar results (within the same lignin type), the within samples deviation did not exceed the deviation between the replicates of exactly the same sample (Table S1). Therefore, the samples circulated inside the lignin community are very similar, and there is little batch-to-batch deviation. Further, we treated the data from these different samples of the same lignins as an average, for Alcell and Indulin lignin, correspondingly.

The reproducibility for 3 types of lignin spectra was examined: native lignins (AMWLa), technical lignins of good solubility and good spectra resolution (Alcell, Indulin) and technical lignins of lower solubility, resulting in lower signal-to-noise (S/N) ratio and resolution (SEAL). The S/N ratios were 280, 200, 180 and 70 for AMWLa, Alcell, Indulin and SEAL samples, correspondingly.

Overall, the standard deviation (StDev) was rather similar for the native and technical lignins (excluding SEAL) (Table 2). However, the relative deviation, RSD, was dependent on the amount of specific moieties and in certain cases was better (lower) for technical lignins of good resolution (Table 2). For example, the RSD values for non-conjugated CO, phenolic OH groups and DC were lower for technical lignins due to higher amounts of these functionalities. In contract, the accuracy in the quantification of aliphatic OH (primary and secondary ones), β–O-4 units and other oxygenated aliphatic moieties was lower for the technical lignins due to degradation of these moieties during processing and their lower amount in the technical lignins as compared to MWLs. Generally, the accuracy in the quantification of technical lignins can be graded as follows (as examples):

1. Highly accurate quantification (RSD <3%): OMe, Total OH, Aliphatic (total) OH, Phenolic OH, S, G, S/G ratio, ArH, and Oxygenated Aliphatic moieties.

2. Moderate accuracy (RSD of 3-10%): Primary and Secondary aliphatic OH, 5-substituted and 5-free PhenOH, total β–O-4, COOR, CO and EtO-groups, Degree of Condensation (DC), H-units (in grass-originated lignins), Alk-O and Saturated Aliphatic moieties.

3. Semi-quantitative (RSD >10%): β–β and β–5 moieties (of low amounts), H-units (in wood-originated lignins), Alk-O-Alk, Degree of demethylation.

Importantly, the accuracy is different for quantification of different moieties, and is usually higher for the majority of important lignin moieties than earlier assumed accuracy of 5-10% for all lignin structures[11] or lower for certain minor structures or specific calculated values. Surprisingly, the accuracy of minor moieties (β–β, β–5, β–1) was similar or even lower in a quantitative HSQC method[19] in spite of much better signal separation in the 2D NMR indicating that the quantitative ^{13}C NMR methodology is not inferior in this respect. The information on the accuracy of the quantification of specific lignin moieties is of primary importance for adequate discussion of the structural information and comparison of different lignins.

Definitely, the low S/N ratio and resolution in the spectra of SEAL resulted in lower accuracy (about twice lower as average) in the quantification of most lignin moieties, especially those of lower intensity (Table 2). Similar results are expected if insufficient numbers of scans are acquired for ^{13}C NMR spectra.

Similarly to the MWL data,[15] there is a rather good correlation in the values for certain resonances in the spectra of non-acetylated and acetylated lignins confirming the absence of lignin fractionation or side reaction during the acetylation by the selected protocol.[14] Therefore, we used the average values for these clusters, when appropriate, for better accuracy (Table 1).

3.3 Differences in the quantification of various structural moieties between native and technical lignins.

The quantification of different structural units in spectra of native lignins has been comprehensively described earlier.[11,14,15] Therefore, the current discussion will be focused on differences in the analysis of technical lignins as compared to the prior works[11,14,15,20]

3.3.1 Aromatic ring

S and G units and protonated aromatic carbons (Ar-H) were estimated from the spectra of the acetylated lignins (Table 1) due to better signals resolution in the area of 125-98 ppm as compared to the non-acetylated spectra (Fig. 2). Importantly, the S/G ratio evaluated from the non-acetylated lignins was about 20% higher than that obtained from the spectra of the corresponding acetylated lignins.

The resonance at 163-148 ppm in the spectra of acetylated lignins embodies G_3 and $S_{3,5}$ as well as H_4 carbons.[15] Then, $[I_{163\text{-}148ac} - \text{H-units}] = S_{3,5} + G_3$. The sum $(S_{3,5} + G_3)$ and the sum $(S_{2,6} + G_2)$ (Fig. 3A-2) show good correlation for MWLs and technical lignins (with the average ratio value of 0.99), which also indicates that the contribution of minor moieties into these resonances is insignificant.

The sum of H+G+S is very close to 100% for the MWLs in agreement with our previous publications.[14,15] This is not the case for technical lignins (Fig. 3A-1). There is a tendency in increasing the misbalance (H+G+S < 100) with higher lignin degradation during the processing. This can be possibly due to formation or/and accumulation of condensed structures (at G_2 and $S_{2,6}$) and demethylated moieties, which resonate at a lower field. It is also important to stress out that the use of an assumption (G+S =100%) for presentation of 2D NMR data as per Ar (or C9-unit)[19] is fine for native lignins, but less accurate for technical lignins and results in overestimation of all values by 10-25% (see Table 3 for G+S amounts).

3.3.2 Oxygenated aliphatic moieties

The amounts of β-β and β-5 units are usually estimated from the peak of C_β at 54-53 ppm in the spectra of non-Ac lignins.[11,20] An alternative way is quantify β-5 and β-β moieties (Fig. 1,

Structures E and F) from their C_α resonances in the spectra of Ac-lignins from the resonance at 88-86 ppm and 86-84 ppm, correspondingly (Fig. 2D, Table 1), when the signals of β-O-4 units are shifted upfield. Although these approaches correlate well for native lignins[14,15] (Fig. 3B-4,5), the amounts of β-β and β-5 quantified from the peak at 54-53 ppm are significantly overestimated (up to 4 times) due to incomplete resolution of the baseline in this region of the spectra of technical lignins (Fig. 2B, D). Therefore, only the resonance at 88-84 ppm should be used for technical lignins. Although it also embody dibenzodioxocin (DBDO) moieties (Fig. 1, Structure U), their amount is very low in technical lignins.[7]

β–O-4 moieties are one of the most important types of lignin structures. In the native lignins, they were quantified from the spectra of non-acetylated lignins by subtracting the resonance at 54-53 ppm (β-β+β-5) from the resonance at 90-82.5 ppm.[15] However, this approach was modified for technical lignins due to significant overestimation of β-β+β-5 amounts from the resonance at 54-53 ppm in their spectra as discussed earlier. The signals of C–β in various β–O-4 moieties, such as those with α–OH, α-O-Alk, α–Ar and DBDO (Fig. 1, Structures A-D, U), are located at 87-82.5 ppm in the spectra of non-Ac lignins and shifted upfield after acetylation.[13] Thus, the total amount of β–O-4 moieties was calculated by subtracting the resonance at 90-82.5 ppm in the spectra of Ac lignins from that in the spectra of the corresponding non-acetylated lignins (Table 1).

The evaluation of β–O-4 moieties from their C-γ signals at 59-61 ppm[11,20] appeared to be very inaccurate for technical lignins (Fig. 3B-6) due to the contribution of other primary alcohol moieties and probably quaternary aliphatic carbons. Their contribution increase during lignin processing, and the more degraded the lignin, the more erroneous this approach will be (Fig. 3B-6).

In contrast to native lignins, only total β–O-4 structures can be quantified from the spectra of technical lignins. The calculations used to determine the amount of β–O-4/α–OH (Fig. 1, structure A), the main type of β–O-4 moieties, from the resonance at ca 77-71 ppm in MWL[15] cannot be used for technical lignins due to significant contribution signals of lignin degradation products into this area.[2,7]

The sum of various OH groups is in good correlation with the total Ac group signals at 22-18 ppm (Fig. 3A-3) indicating once more the quantitative nature of the spectra and reliability of the quantification algorithm. A reasonable correlation has been observed[14,15] between the amount of primary OH groups (OH_{pr}) and the resonance at 65-58ppm for MWL (Fig.3B-7). However, this is not the case for the technical lignins investigated; in most cases, the resonance at 65-58 ppm is significantly higher than the amount of OH_{pr} (Fig. 3B-7). For Alcell lignin, this can be explained by the contribution of ethyl groups (O-$\underline{C}H_2$-CH_3) into the resonance at 65-58 ppm. For other lignins of high degradation, such as AKL, SBL and Sucroline, the contribution of quaternary aliphatic carbons in the signal at 65-58 ppm could be speculated.

3.3.3 Correction for carbohydrate content

Technical lignins can contain significant amount of carbohydrates, for example SEAL and AKL. Carbohydrate signals overlap with signals of lignins in oxygenated aliphatic and aliphatic OH groups regions. This is not a problem when lignin is used 'as-is' because these carbohydrate will be a part of lignin product. Nevertheless, if it is important to understand the structure of "true lignin", correction for sugar content can be made. The C-1 signals of carbohydrates are shifted upfield after acetylation and can be separated from the lignin signals and quantified from the resonance at 102-98 ppm (Table 1). The corrections of the specific functionalities and the data for the "true" lignin for SEAL and AKL are shown in Table 4. It is important to keep in mind that the accuracy of the corrected values will decrease with increasing amount of sugars in lignin preparations.

3.3.4 Ether moieties

It has been suggested to use the ratio of resonance at 155-151 ppm to that at 150-145ppm in the spectra of non-Ac lignin to evaluate the ratio between etherified to non-etherified lignin units.[20] We believe this approach is not optimal because the latter embodies resonances of G_3 and $S_{3,5}$ of non-etherifies and G_4 of etherified and non-etherified lignin units[12,13] and, therefore, it is too complex. We suggest that the resonance at 155-151 ppm alone is much simpler for interpretation and for approximate evaluation of the amounts of S etherified moieties in spectra of HW lignins (Table 3) considering $\mathbf{R}_{et}$ and $\mathbf{T}_{et}$ moieties (Fig. 1) as minor. In contrast, the resonance at 155-151 ppm embodies predominantly 5-5' etherified lignin units (G_3) in the spectra of softwood lignins. Noteworthy, the resonance at 155-151 ppm correlates well with the amounts of Alk-O and β–O-4 moieties, but, as expected, the lines for SW and HW and grass lignins are different (Fig. 3C).

3.3.5 Saturated aliphatic moieties

Saturated aliphatic moieties were quantified from the interval of about 54-0 ppm. Their estimation is partially interfered by strong NMR solvent (DMSO) resonance centered at 39.5 ppm. To avoid solvent interference in the quantification of saturated aliphatic moieties resonance, we run ^{13}C NMR for acetylated Alcell lignin in $CDCl_3$. The spectrum showed (Fig. 2E) that some aliphatic lignin moieties resonate in the area of 45-35 ppm and will be obscured by the solvent signal when DMSO is used. The value is 43/100 Ar or about 25% of the total saturated aliphatic resonance value (without the acetyl peak centered at ca 20 ppm). However, the resonances at 54-45 and 35-0 ppm are the same in the spectra run in DMSO and $CDCl_3$ meaning that the DMSO signal does not strongly inflate these ranges.

Although the use of $CDCl_3$ as solvent allows for characterization of saturated aliphatic moieties, its resonance at ca 78 ppm strongly obscures important resonances in the spectra (β–O-4 units, oxygenated aliphatic moieties) and therefore we would not recommend this solvent for a routine NMR analysis of lignins. In addition, we observed somewhat lower spectral resolution of lignin

signals in $CDCl_3$ as compared to that in DMSO. For a comprehensive NMR analysis, spectra acquired both in $CDCl_3$ and DMSO can be used. For a simplified protocol, DMSO is still the best solvent (considering also the highest lignin solubility in it, especially for non-Ac lignins). The sum of the resonances at 54-45 ppm and 35-0 ppm can be used for comparative evaluation of saturated aliphatic moieties on a routine base then.

Significant amounts of saturated aliphatic moieties are produced during lignin degradation in pulping and other biorefinery processes (Table 3, Fig. 2). Obviously, they are ones of the major functional units in technical lignins, but have not been discussed sufficiently enough. A part of the saturated aliphatic moieties could come from lipophilic extractives linked to lignin physically or more likely chemically as they could not be completely removed by extraction with a solvent of low polarity. Another part of these moieties should come from lignin degradation and re-arrangement during technical process, such as formation of Hibbert's ketons under acidic conditions.[2]

3.3.6 The size of lignin side chain

Eventually, it is of interest to evaluate the length of lignin side chain in technical lignins. It is well known that the side chain degrades during various technical processing. Usually, significant shortage of the side chain is expected based on strong degradation of the oxygenated aliphatic moieties. However, we have to keep in mind that this process is accompanied by a strong increase in the amount of saturated aliphatic moieties. In addition, the amounts of aliphatic carbonyl and carboxyl/esters increase significantly as well. Therefore, the calculated size of the side chains in the technical lignins (Table 3) is not much lower than that in the native ones in spite of dramatic transformation in the lignin structure. This is also consistent with the calculated "molecular mass" values of an averaged lignin monomeric unit (M_{Ar}) for each lignin type, which just slightly decrease in the technical lignins as compared to native MWL lignins (Table 3).

3.3.7 Demethylation

The percentage of methoxy groups eliminated during lignin processing i.e. the degree of demethylation (or/and demethoxylation) was determined semi-quantitatively comparing the number expected from the normalized S/G/H ratio and the actual amounts of OMe groups (Table 1). The demethylation reaction is well known for kraft pulping, but has not been described well for acidic treatments. However, the degree of demethylation is the most significant in the organosolv lignins, especially for the HW Alcell lignin and also substantial for the SE lignins. In addition, it is believed that no demethylation occurs during soda pulping, but our results show that it is noticeable. The reaction mechanisms of lignin demethylation/demethoxylation in soda pulping and during acidic lignin processing are not well understood. It can be speculated that oxidative demethylation/demethoxylation might takes place as certain amount of oxygen is present in wood chips and pulping solution. The solubility of oxygen in organic solvents is of an order of magnitude higher than that in aqueous solutions and this might explain the high degree of demethoxylation for the organosolv lignins (Alcell and OS-DF).

3.3.8 Ethoxyl moieties

Ethoxyl moieties (OEt) are typical for ethanol based organosolv lignins. They were detected earlier with 2D NMR[7,17] and can be quantified with ^{13}C NMR at 16.5-13.0 ppm.[17] They can be tentatively differentiated by ether and ester types centered at ca 15.3 and 14.1 ppm correspondingly (Fig. 3, Table 5). It should be mentioned that some other saturated aliphatic moieties might also contribute to this resonance and therefore the values reported show the highest limit for EtO- groups. However, organosolv lignins have very little amounts of extractives as evident from their HSQC spectra,[7] and therefore these values should correspond to EtO-groups predominantly.

3.3.9 Special features in the analysis of non-wood lignins

The chemical structures of non-wood (grass) lignins are significantly more complex than wood lignins due to the presence of H-units and conjugated acid derivatives (cinnamic acids

399 derivatives), such as ferrulates (FA) and coumarates (CA). Therefore, significant adjustments to
400 the calculation algorithm are required.

401 It should be stated that CA and p-hydroxy benzoic acid (PHBA) derivatives are usually not
402 considered as lignin structural units (C9-units) from the classical biosynthetical point of view.[3]
403 However, from the point of lignin utilization, the biosynthetic origin of different lignin elements
404 is much less important than the properties on the lignin product itself. Therefore, we include all
405 aromatic lignin subunits (including CA and PHBA) into consideration ("per Ar") and they were
406 also included into the G- (FA) and H-unit (CA, PHBA) types.

407 Although the presence of flavonoids, specifically tricin, was demonstrated in native grass
408 lignins,[6] they were not detected in the technical lignins (SBL and Sucrolin). The amounts of the
409 conjugated COOR moieties were determined from the resonance at about 168 – 165 ppm (Fig. 4,
410 Table 6). HSQC 2D NMR spectra showed that the amounts of PHBA were very little in BSL and
411 Sucroline, and therefore all resonance of the conjugated COOR could be assigned to CA and FA
412 derivatives. A correction factor should be used for grass lignins containing significant amounts of
413 conjugated COOR, which contribute to the resonance at 163-102 ppm with 2 olefinic carbons, as:
414 $I_{163\text{-}102}$= (600+2*COORcon.)
415 In addition to the conjugated COOR, grass lignins contained significant amounts of H-units of
416 different types (including CA). Differentiation of various H-units with ^{13}C NMR has been
417 described earlier[14,15,20] and summarized in Table 6. Noteworthy, in addition to CA structures, the
418 grass technical lignins contained significant amounts of non-conjugated H-moieties (Fig. 4, Table
419 6). However, flavonoids (different from tricin) could also contribute to this resonance[22] and their
420 presence could not be excluded.
421 It is not possible to differentiate directly between COOR signals of the CA and FA types using
422 ^{13}C NMR.[13] It was suggested[21] to quantify CA from the amount of the conjugated H-moieties and
423 deduct the amount of FA by difference between the total $COOR_{conj.}$ and the obtained value for

CA. We used this assumption in our current calculations (Table 6), but it would require further confirmation. In addition, it was possible to separate CA/FA acid moieties from the corresponding esters in the ^{13}C NMR spectra by the resonances at ca 168 ppm and 166 ppm correspondingly (Fig. 4, Table 6).[13]

Furthermore, calculation of DC in grass lignins also requires some corrections from the way used for wood lignins (Table 1). As one olephinic carbon contributes into the resonance of ArH (at ca 115 ppm), the corrected value of ArH would be

$Ar\text{-}H = (I_{125\text{-}102})_{ac} - (CA+FA)_{\%}$, where $(CA+FA)_{\%}$ (%) = (CA + FA)/(S + G + H)*100

The Degree of Condensation (DC) is calculated then as following:

$DC = (200 + \mathbf{G}_{\%}) - [(I_{125\text{-}102})_{ac} - (CA+FA)_{\%}]$

3.4 Comparison of different types of technical lignins

Certain general tendencies can be easily observed when compare technical lignins *versus* native ones (Table 3; see Fig. S2 for main lignin structural functional groups in mmol/g). Lignin degradation during various technical processing of different biomass is always accompanied with a decrease in aliphatic OH (primary and especially secondary ones), oxygenated aliphatic moieties in general and β–O-4 units specificallyand an increase in phenolic OH, COOR, saturated aliphatic moieties, DC and demethylation.

At the same time, there were significant structural differences in the lignins investigated. The SE lignins (SEAL and SEPL) were the less degraded while AKL underwent the most severe degradation. A specific characteristic of the acid-derived lignins (organosolv ones and Sucrolin) was high amounts of CO groups, likely Hibbert's ketons. Also, a typical feature of the ethanol-based organosolv lignins was the presence of EtO- groups (Fig. 3, Table 5). Noteworthy, Alcell lignin contains higher amounts of EtO-groups than the DF-OS lignin, specifically those tentatively assigned to the ether types. It could be due to the biomass used or/and in the process

conditions (OS-DF lignin was produced with addition of catalytic amount of sulfuric acid while the Alcell process was auto-catalyzed).

High amounts of COOR groups, specifically conjugated ones, in Sucrolin and SBL lignins (Table 3, 6), were specie related. Noteworthy, significant portions of cinnamic acids survived the treatments under very severe conditions (soda pulping and acid hydrolysis). Interestingly, SBL contains predominantly FA, whereas CA moieties dominated in Sucrolin. Furthermore, Sucrolin lignin had a higher portion of esters than SBL indicating their stronger degradation under alkaline conditions than under acidic conditions. In addition, Sucrolin lignin had much higher amounts of H-units, both of conjugated and non-conjugated types, than SBL (Table 6).

3.5 Molecular weight

Molecular weight analysis of the lignins was performed by SEC method using 0.1 M NaOH as a mobile phase. As significant deviation between different methods for SEC analysis of lignins have been documented,[23] these results (Table 7) should be considered rather relative. However, they are useful for comparison of different lignins analyzed under the same conditions.

The results show that, similarly to the NMR data, the differences between different batches of Alcell and Indulin lignins were subtle. Hardwood Alcell and AKL lignins have the lowest molecular weights and dispersity (D); softwood kraft lignins have higher molecular weight and D values. The steam explosion lignins (SEPL and SEAL) show the highest D, in agreement with the previous studies.[23]

4 Conclusions

The current research provides with an advanced methodology for the quantification of different structural elements in various technical lignins by ^{13}C NMR; more than 30 lignin structural characteristics can be obtained on a routine base along with data on functionalities specific for

certain types of lignins (such as various EtO-groups, H-units and conjugated COOR moieties). The accuracy in the quantification of different moieties has been determined. Different ways for the quantification of specific lignin subunits have been discussed and most reliable approaches have been selected.

A database on 9 technical lignins originated from most common biorefinery processes has been generated with the suggested approach. Lignin degradation during various technical processing of different biomass species is always accompanied with a decrease in aliphatic OH (primary and especially secondary ones), β−O-4 and total oxygenated aliphatic moieties, and an increase in phenolic OH, COOR, saturated aliphatic moieties and in the degree of condensation. Significant structural differences in the lignins investigated can originate from the process conditions and/or can be specie related.

Acknowledgements

The authors are very thankful to Mr. Abaecherli (ILI), Dr. A. Berlin (former UBC, currently at Novozymes), Dr. C.-L. Chen (NCSU) and Prof. Dr. O. Faix for kind donation of lignin samples, Dr. H. Gracz (NCSU) for the NMR experiment and Mr. N. Voznyak (Renmatix Inc.) for the SEC analysis.

References

1. W. Glasser, In: *Lignin: Historical, Biological, and Material Perspectives.* W. Glasser, R. Northey and T. Schultz. Eds., ACS, Washington, DC, 2000; pp. 216-238; J. E. Holladay, J. F. White, J.J. Bozell, and D. Johnson. *Top Value-Added Chemicals from Biomass - Volume II—Results of Screening for Potential Candidates from Biorefinery Lignin* . PNNL-16983, Pacific Northwest National Laboratory, Richland, WA 2007.

2. A. Berlin and M. Balakshin, In *Bioenergy Research: Advances and Application;* V.K. Gupta, C.P. Kubicek, J. Saddler, F. Xu and M.Tuohy, Eds. Elsevier: Amsterdam, 2014; pp 315 - 336.
3. C.-L. Chen, In *Wood Structure and Composition;* M. Lewis, I.S. Goldstein, Eds.; Marcel Dekker Inc., New York, 1991; 183-261; Ralph, J.; Lundquist, K.; Brunow, G.; Lu, F.; Kim, H.; Schatz, P.F.; Marita, J.M.; Hatfield, R.; Ralph, S.A.; Christensen, J.H.; Boerjan, W. *Phytochem. Rev.* 2004, 3, 29-60.
4. S. Yang, J.-L. Wen, T.-Q. Yuan and R.-C. Sun, *RSC Adv*., 2014, **4**, 57996-58004.
5. J. Ralph and L.L. Landucci, In *Lignin and Lignans: Advances in Chemistry;* C. Heitner, D.R. Dimmel and J.A. Schmidt, Eds. CRC Press: Boca Raton, 2010; pp.137-244.
6. J.C. del Río, J. Rencoret, P. Prinsen, A.T. Martinez, J. Ralph and A. Gutiérrez, *J. Agric. Food Chem*. 2012, **60**, 5922-5935.
7. E.A. Capanema, M.Yu. Balakshin, C.-L. Chen, J.S. Gratzl and H. Gracz, *Holzforschung* 2001, **54**, 302-308; M.Yu. Balakshin, E.A. Capanema, C.-L. Chen and H.S. Gracz, *J. Agric. Food Chem*. 2003, **51**, 6116-6127 ; T.M. Liitia, S.L. Maunu, B. Hortling, M. Toikka and I. Kilpelainen, *J. Agric Food Chem.*, 2003, **51**, 2136–2143.
8. E.A. Capanema and M.Yu. Balakshin U.S. Patent Application Publication 2014/0275501.
9. Y. Pu, S. Cao and A.J. Ragauskas, *Energy & Environmental Science*. 2011, **4**, 3154-3166.
10. M. Yu. Balakshin and E.A. Capanema, *J. Wood Chem. Techn.* 2015, **35(3)**, 220 – 237.
11. C.-L. Chen and D. Robert, In *Methods in Enzymology*; Wood, W.A., Kellogg, S.T. Eds. Academic Press Inc., New York. 1988; Vol. 161B, pp 137-174; D.V. Evtuguin, C. Pascoal Neto, A.M.S. Silva, P.M. Dominges, F.M.L. Amado, D. Robert and O. Faix, *J. Agric. Food Chem.* 2000,**149**, 4252 - 4261.
12. K.P. Kringstad and R. Mörck, *Holzforschung* 1983, **37**, 237-244.

13. H.H. Nimz, D. Robert, O. Faix and M. Nemr, *Holzforschung* 1981, **35**, 16–26; S.A. Ralph, J. Ralph and L.L. Landucci, NMR database of lignin and cell wall model compounds. Available at URL https://www.glbrc.org/databases_and_software/nmrdatabase/

14. E.A. Capanema, M.Yu. Balakshin and J.F. Kadla, *J. Agric. Food Chem*. 2004, **52**, 1850-1860; M.Yu. Balakshin, E.A. Capanema, B. Goldfarb, J. Frampton and J. Kadla, *Holzforschung* 2005, **59**, 488-496.

15. M.Yu. Balakshin, E.A. Capanema, R.B. Santos, H.-m. Chang and H. Jameel, *Holzforschung* 2015, **69**, DOI 10.1515/hf-2014-0328.

16. T. Milne, H. Chum, F. Agblevor and D.K. Johnson, *Biomass & Bioenergy* 1992, **2**, 341-366; O. Faix, D.S. Argyropoulos, D. Robert and V. Neirinck, *Holzforschung*. 1994, **48**, 387-394. International Lignin Institute Online: http://www.ili-lignin.com/

17. A. Berlin, M. Balakshin, N. Gilkes, J. Kadla, V. Maximenko, S. Kubo and J. Saddler, *J. Biotechn*. 2006, **125(2)**, 198-209.

18. P. Xia, L. Akim and D. S. Argyropoulos, *J. Agric. Food Chem*. 2001, **49**, 3573-3578.

19. M. Sette, R. Wechselberger, C. Crestini. *Chem. – Eur. J*. 2011, **17**, 9529 – 9535.

20. A.P. Duarte, D. Robert and D. Lachenal, *Holzforschung* 2000, **54**, 365–372.

21. L. Oliveira, D.V. Evtuguin, N. Cordeiro, A.J.D. Silvestre and A.M.S. Silva, In *Characterization of Lignocellulosics*; Hu, T., Ed.; Blackwell; Oxford, UK, 2008, pp. 171 – 188.

22. C-L. Chang, G-J. Wang, L.-J. Zhang, W.-J. Tsai, R.-Y. Chen, Y.-C. Wu and Y.-H. Kuo, *Phytochemistry* 2010, **71**, 271-279.

23. S. Baumberger, A. Abaecherli, M. Fasching, G. Gellerstedt, R. Gosselink, B. Hortling, J. Li, B. Saake and de E. Jong, *Holzforschung* 2007, **61**, 459-468.

Table 1 Quantification of various moieties in technical lignins by ^{13}C NMR.

No	Structures	Quantification	Minor moieties
1	β–O-4 total	$(I_{90\text{-}82.5})_{na} - (I_{90\text{-}82.5})_{ac}$	
2	Pino-/syringylresinol (**F_α**)*	$(I_{86\text{-}84})_{ac}$	**U**
3	Phenylcoumarane (**E_α**)	$(I_{88\text{-}86})_{ac}$	
4	Sugars (C_1)	$(I_{102\text{-}98})_{ac}$	except reducing end units
5	**$S_{2,6}$**	$(I_{110\text{-}102})_{ac}$ /2	**$R_{2,6}$**
6	**G_2**	$(I_{113\text{-}110})_{ac}$	except **$R_{2,6}$**
7	**H_4**	$(I_{163\text{-}156})_{na}$	
8	Degree of Condensation	$(200 + \mathbf{G}_{\%}) - (I_{125\text{-}102})_{ac}$	except **$M_{G\text{-}6}$**
9	OMe	$[(I_{58\text{-}54})_{na} + (I_{58\text{-}54})_{ac}]/2$	
10	Non-conjugated CO	$[(I_{215\text{-}200})_{na} + (I_{215\text{-}200})_{ac}]/2$	
11	Conjugated CO	$[(I_{200\text{-}185})_{na} + (I_{200\text{-}185})_{ac}]/2$	
12	Non-conjugated COOR	$(I_{178\text{-}168.5})_{na}$	
13	Conjugated COOR	$(I_{168.5\text{-}165})_{na}$	
14	Total OH	A. ($\underline{\boldsymbol{C}}H_3COO$-) $(I_{23\text{-}18})_{ac} - (I_{23\text{-}18})_{na}$ B. ($CH_3\underline{\boldsymbol{C}}OO$-) $(I_{172\text{-}166})_{ac} - (I_{172\text{-}166})_{na}$	Sugars
15	Primary aliphatic OH	$(I_{172\text{-}169.7})_{ac} - (I_{172\text{-}169.7})_{na.}$	C_6 in hexoses
16	Secondary aliphatic OH	$(I_{169.7\text{-}168.7})_{ac} - (I_{169.7\text{-}168.7})_{na.}$	$C_{2,3,}$ in sugars
17	5-free Phenolic OH	$(I_{168.7\text{-}168.3})_{ac} - (I_{168.7\text{-}168.3})_{na.}$	
18	5-subst. Phenolic OH	$(I_{168.3\text{-}166})_{ac} - (I_{168.3\text{-}166})_{na.}$	
19	Ar-H	$(I_{125\text{-}102})_{ac}$	except **$H_{2,6}$** and **$M_{G\text{-}6}$**
20	Oxygenated Aliphatic	$[(I_{90\text{-}58})_{na} + (I_{90\text{-}58})_{ac}]/2$	Sugars (except C_1), Aliphatic quaternary C
21	Saturated Aliphatic	$(I_{54\text{-}0})_{na}$ (in $CDCl_3$) $(I_{54\text{-}45})_{na} + (I_{35\text{-}0})_{na}$ (in DMSO-d_6)	Resonance ($I_{45\text{-}35}$) is missing
22	EtO-	$[(I_{16.5\text{-}13.0})_{na} + (I_{16.5\text{-}13.0})_{ac}]/2$	Extractives
23	Alkyl-O-Alkyl	Oxygen. Aliph. – OH_{Aliph}	Sugars
24	Side chain length	CO+COOR + Oxygen.Aliph. + Sat.Aliph.	Resonance ($I_{45\text{-}35}$) is missing (in DMSO-d_6)
25	Demethylation degree	$100 - OMe/(2\mathbf{S}_{\%} + \mathbf{G}_{\%})x100$	
	Clusters, ppm	**Major Moieties**	
26	161-148ac	**H_4**, **$S_{3,5}$**, **G_3**	**R_5**, **$G_{et.conj.\text{-}4}$**, **L_α**
27	156-151na	**($S_{3,5}$**, **$R_{3,5}$**, **T_3)**$_{et.}$	**L_α**, **$G_{et.conj.\text{-}4}$**
28	150-149na	**$G_{et\text{-}3}$** non-condensed	
29	148-144.5ac	**$G_{et\text{-}4}$**	except **R** and **$G_{conj.}$**
30	90-78	Alk-O-Ar, α–O-Alk	
31	78-65	γ–O-Alk, $OH_{sec,}$	Sugars
32	65-58	$OH_{prim,}$	Sugars, Aliphatic quaternary C
33	M_{Ar}	$[(0.3m_{Lig}x(I_{95\text{-}90})_{na}/ m_{IS} + (0.3m_{Lig}x(I_{95\text{-}90})_{ac}/m_{IS} - 0.42xOH_{Total})]/2$	

Note: the exact chemical shift values are determined by the local minima in the spectra and can be slightly different from the numbers listed in Table 1; ($I_{xx\text{-}yy}$) corresponds to the resonance value in the interval (xx-yy) ppm of the ^{13}C NMR spectra; abbreviations "na" and "ac" are used for data obtained from spectra of non-acetylated and acetylated lignins, correspondingly. Structures **E**, **F**, **U**, **R**, **M_G**, **R**, **L**, **S**, **G**, **H** correspond to those in Fig. 1.
*The number of C_9-units involved in resinol structures; as the structure is symmetric, the number of resinol structures is ½ of the C_9-units involved.

556 **Table 2** Deviation in NMR analysis of AMWL, Alcell, Indulin and SEAL lignins.

Moieties/ Integration range (ppm)	StDev (per 100Ar)				RSD (%)			
	AMWL	Alcell	Indulin	SEAL	AMWL	Alcell	Indulin	SEAL
Non-conjugated CO	1.0	1.6	1.1	4.9	32.1	10.7	15.9	61.0
Conjugated CO	1.0	1.2	0.9	4.4	7.9	8.3	10.8	36.7
Total CO	2.0	1.8	2.0	9.2	12.5	6.2	13.0	46.1
Non-conj. COOR	0.6	1.0	0.8	2.8	8.0	6.0	5.1	21.8
Conjugated COOR	0.6	0.5	0.5	1.5	12.0	12.0	23.6	37.0
Total COOR	1.2	1.3	1.1	3.1	9.5	6.0	6.6	18.0
Primary OH	1.1	1.3	0.8	3.5	1.5	6.7	2.5	10.1
Secondary OH	1.0	0.6	0.6	0.3	1.6	4.6	3.5	1.9
Total aliphatic OH	1.3	0.6	1.4	3.2	1.0	1.9	2.9	6.3
5-free PhOH	0.7	0.7	NR	1.3	7.3	3.8	NR	8.8
5-subst. PhOH	1.0	2.7	NR	1.3	8.0	5.2	NR	3.4
Total PhOH	0.5	1.0	1.0	2.6	2.4	1.5	1.5	4.7
Total OH	1.8	1.3	1.3	3.8	1.2	1.3	1.1	3.6
OMe	2.0	1.0	1.1	5.0	1.2	1.0	1.3	3.8
$S_{2,6}$	1.0	0.7	NA	2.1	0.8	0.9	NA	2.1
G_2	0.6	1.4	1.4	0.7	1.9	4.0	1.6	2.0
H-units	0.5	0.6	1.0	0.7	10.0	8.8	13.0	9.4
Ar-H	1.6	2.9	1.4	0.7	0.7	1.4	0.6	0.3
DC, %	0.5	1.1	1.4	0.7	4.5	2.4	2.1	2.0
S/G ratio	0.03	0.06	NA	0.07	1.2	5.0	NA	5.0
β-5	0.5	0.5	0.5	0.8	22.9	15.0	13.0	28.3
β–β	0.5	0.5	0.2	0.8	6.0	15.0	5.5	12.9
β-O-4	1.1	0.5	0.7	2.8	2.1	7.0	9.4	12.9
90-78 ppm	1.0	1.9	1.5	5.8	1.2	8.1	5.2	12.0
78-65 ppm	1.5	2.2	1.0	3.2	1.9	9.0	3.2	15.5
65-58 ppm	1.5	1.4	0.9	6.0	2.0	4.0	2.8	15.3
Oxygenated Aliph.	3.5	2.5	1.7	13.7	1.5	3.0	1.9	12.6
Saturated Aliph.	4.0	11.5	7.0	10.3	7.1	7.7	6.5	8.8
Side chain length	1.0	11.5	7.5	56.8	0.2	4.0	2.9	18.8
M_{Ar}	5.0	5.5	6.1	ND	2.3	3.1	3.5	ND

557 NR - non-resolved signals (therefore, not indegrated); NA – non applicable; ND – not determined
558 (unknown real lignin:IS ratio due to incomplete lignin dissolution)

Table 3 Amounts of various lignin moieties (per 100Ar).

Moieties/range	Alcell	OS-DF	Indulin	Curan	AKL	SEPL	SEAL	Sucrolin	SBL	AMWL	PMWL
Total CO	29	22	15	16	21	23	20	30	19	16	20
Non-conj. CO	15	8	7	7	11	11	8	17	7	3	3
Conj. CO	14	14	8	9	10	12	12	13	11	13	17
Total COOR	21	5	17	21	28	22	17	38	37	13	6
Non-conj. COOR	17	4	15	17	25	18	13	27	27	8	4
Conj. COOR	4	1	2	4	3	4	4	11	10	5	2
Total OH	103	110	115	120	107	124	130	92	96	156	140
Aliph. OH	33	34	49	51	31	61	75	43	37	134	107
OHpr	19	26	31	35	17	33	40	19	17	72	67
OHsec	14	8	18	16	14	28	35	24	20	62	40
Phenolic OH	70	76	66	69	76	63	55	49	59	22	33
PhOH 5-free	18				18	16	15			10	
PhOH 5-subst.	52				58	47	40			12	
S-units	42	NA	NA	NA	46	49	50	25	25	66	NA
G-units	36	104	92	86	35	30	36	47	49	31	99
H-units	7	8	8	5	3	10	7	35	20	5	4
S:G-ratio	1.18	NA	NA	NA	1.31	1.63	1.39	0.53	0.51	2.11	NA
OMe	103	78	81	82	120	126	132	81	92	164	97
% Demethylation	27	16	12	13	21	12	10	10	14	(-1)	(-1)
ArH	202	225	234	218	199	201	208	207	213	221	253
DC, %	44	75	66	82	44	37	34	58	53	11	43
β-5	3	3	4	2	2	2	3	1	1	2	10
β–β**	3	4	4	3	5	4	6	2	1	8	4
β-O-4	7	4	7	5	1	17	22	4	2	52	42
163-148 ppm	130	98	90	92	124	147	145	132	113	174	108
155-151ppm	37	15	13	16	14	61	69	28	25	124	33
90-78 ppm	23	21	29	20	22	40	48	16	14	80	76
78-65 ppm	24	23	33	28	42	46	55	20	16	81	66
65-58 ppm	35	31	31	26	29	42	49	28	22	76	72
Oxygen. Aliph.	82	75	93	74	93	128	152	64	52	237	214
Saturated Aliph.	149	96	109	100	145	116	117	161	140	56	32
Side chain length	281	198	233	211	269*	289	270*	293	248	322	272
Alkyl ethers	50	42	44	23	54*	68	61*	21	15	103	107
Sugars	<1	<1	~1	~1	4	<1	8	<1	~1	<1	~ 1
M_{Ar}	178	164	173	180	201	194		203	195	218	180

*corrected for sugar content; **The number of C_9-units involved in resinol structures; as the structure is symmetric, the number of resinol structures is ½ of the C_9-units involved.

Table 4 Correction for sugar content (per 100 Ar)

Corrected values	Calculations	AKL	SEAL
$OH_{pr\text{-}cor}$	OH_{pr} – Sugars x $Hex_{\%}$/100	15	36
$OH_{sec\text{-}cor}$	OH_{sec} – 2Sugars	6	19
$OH_{Aliph\text{-}cor}$	$OH_{pr\text{-}cor}$ + $OH_{sec\text{-}cor}$	21	55
$OH_{Total\text{-}cor}$	$OH_{Aliph\text{-}cor}$ + OH_{Ph}	97	110
$I_{(78\text{-}65)cor}$	$I_{(78\text{-}65)}$ – Sugars x ($3Xyl_{\%}$ + 4 $Hex_{\%}$)/100	28	27
$I_{(65\text{-}58)cor}$	$I_{(65\text{-}58)}$ – Sugars	25	41
Oxygenated $Aliph._{cor}$	$I_{(90\text{-}78)}$ + $I_{(78\text{-}65)cor}$ + $I_{(65\text{-}58)cor}$	75	116

$Xyl_{\%}$ and $Hex_{\%}$ are the percentages of xylan and hexoses, correspondingly, in the total sugar content, determined by a wet chemistry method.

Table 5 Quantification of various ethoxyl groups in organosolv lignins

Moieties	Integration range (ppm)	Quantities (per 100Ar)	
		Alcell	OS-DF
Ester EtO-	16.5-14.8	6±0.3	6±0.2
Ether EtO-	14.8-13.0	8±0.4	3±0.2
Total EtO-	16.5-13.0	14±0.5	9±0.3

Table 6 Amounts of various H-units and conjugated COOR moieties in grass technical lignins (per 100Ar)

Moieties	Integral range, ppm*	Sucrolin	BSL
Total H	163-156	31	14
Conjugated H	163-158.5	9	3
Non-conjugated H	158.5-156	22	11
Total conjugated COOR	168.5-165	11	9
Conjugated acids	168.5-167.2	5	7
Conjugated esters	167.2-165	6	2
CA COOR	161-158.5	9	3
FA COOR	Total COOR - CA	2	6

*In the spectra of non-acetylated lignins

Table 7 Molecular weight of biorefinery lignins (Da)

Lignin	Mp	Mn	Mw	Mz	D
Alcell-1	1684	757	2002	5166	2.64
Alcell-2	1691	786	2063	7560	2.62
Alcell-3	1710	825	2121	6114	2.57
OS-DF	1965	911	3428	16480	3.76
Indulin-1	2116	1030	4443	16354	4.31
Indulin-2	2183	1177	5539	24339	4.71
Curan	2409	1358	6839	26995	5.04
AKL	1382	708	1722	6367	2.43
SEPL	1866	1025	6801	46990	6.64
SEAL	1815	985	5246	28518	5.33
Sucrolin	1962	1110	5107	25642	4.60
BSL	1844	1023	3423	12110	3.35
PMWL	3276	1139	3708	8917	3.26

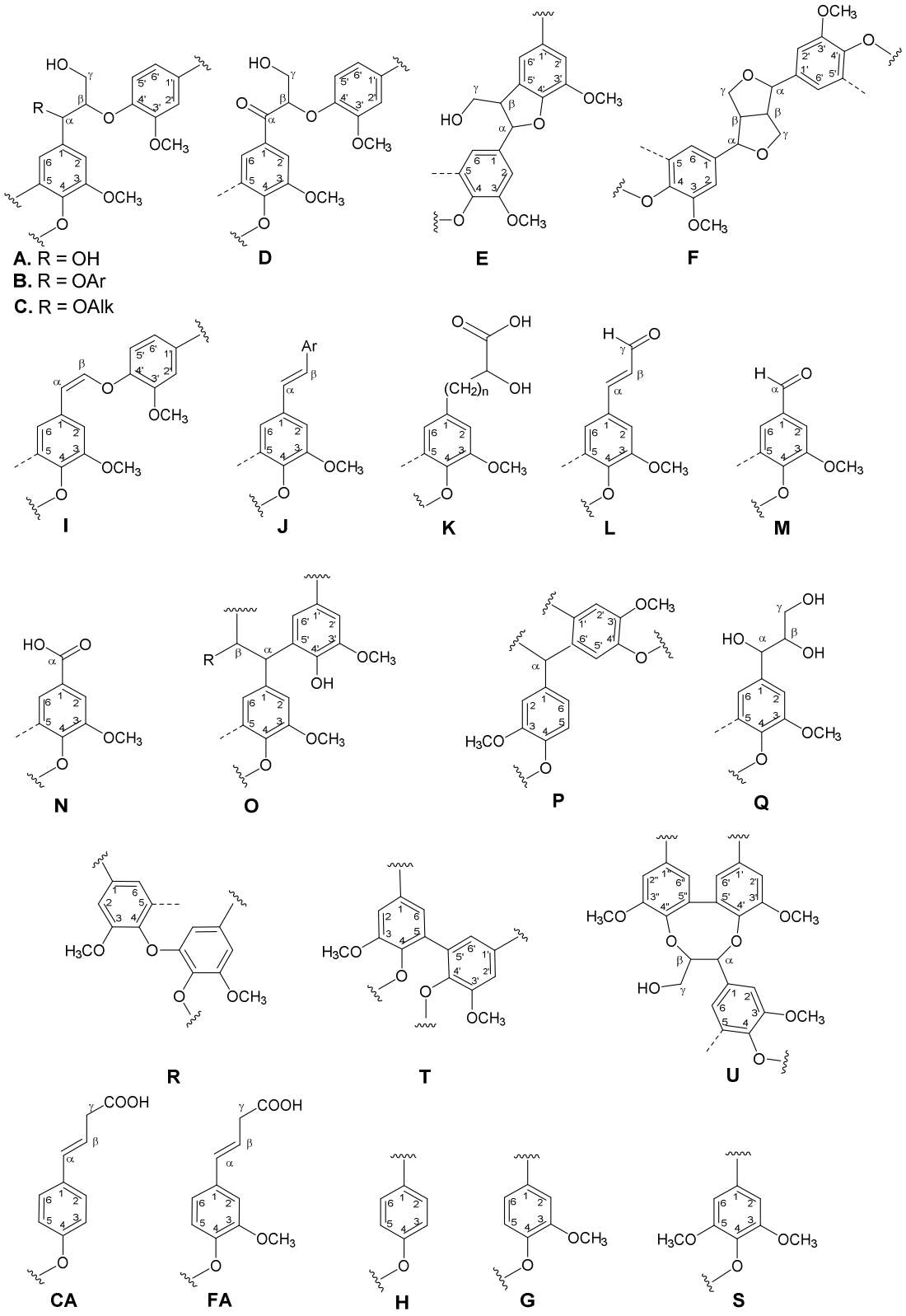
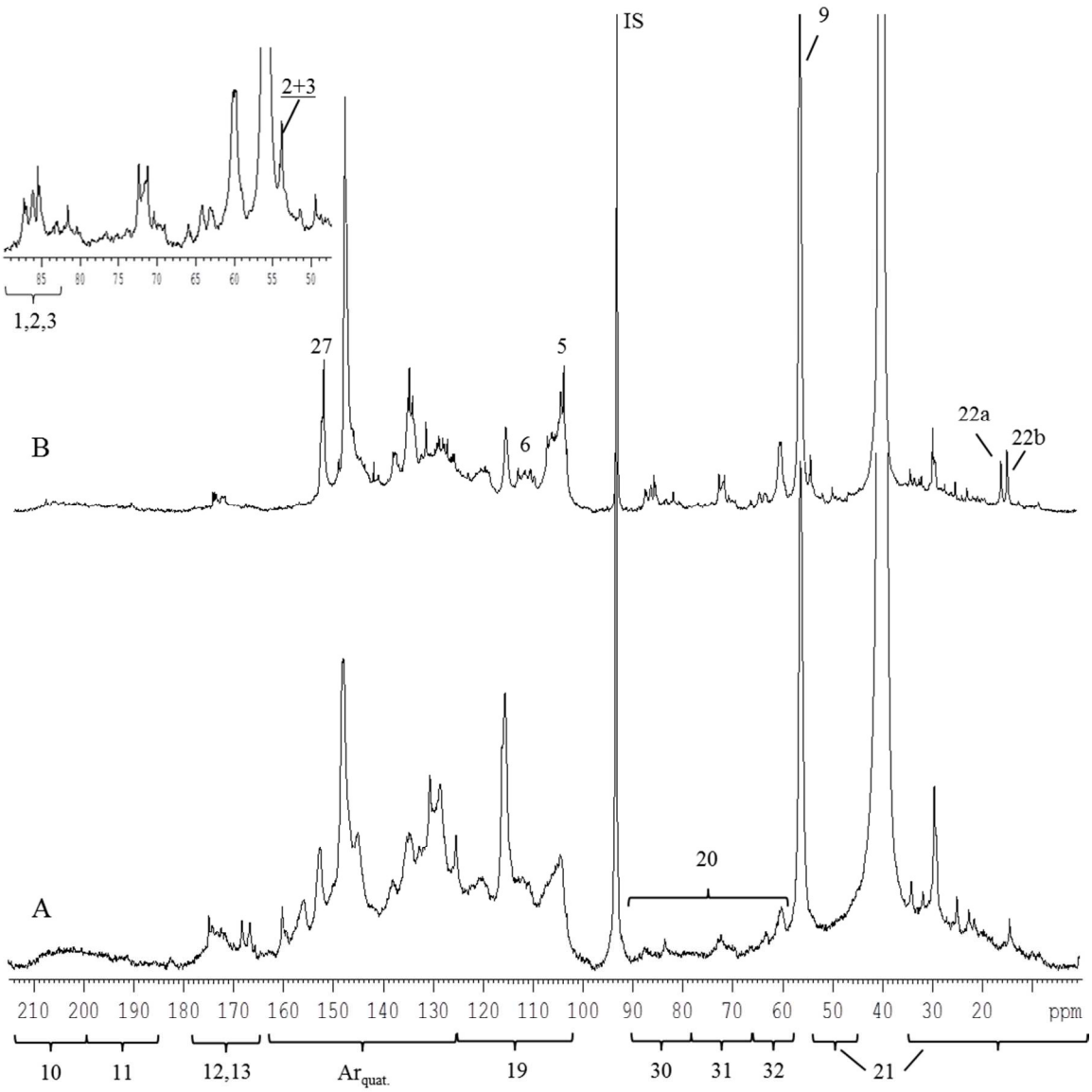


Fig. 1 Substructures detected in technical lignins.[2,7]



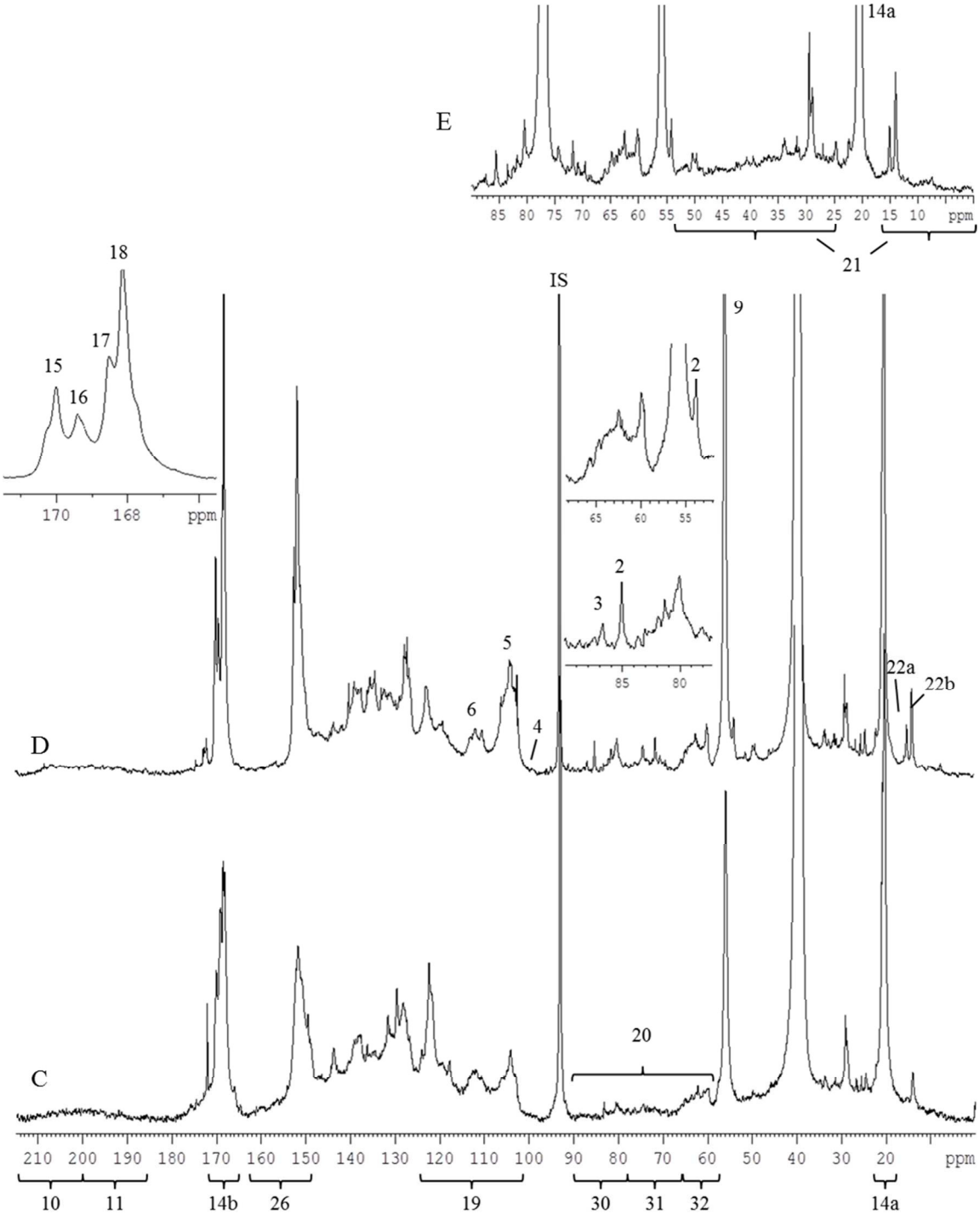


Fig. 2 ^{13}C NMR spectra of non-acetylated Sucrolin lignin (A), non-acetylated Alcell lignin (B), acetylated Sucrolin lignin (C), acetylated Alcell lignin acquired in DMSO-d6 (D) as well as acetylated Alcell lignin (expanded aliphatic area) acquired in $CDCl_3$ (E). Numbers correspond to those in Tables 1 and 5.

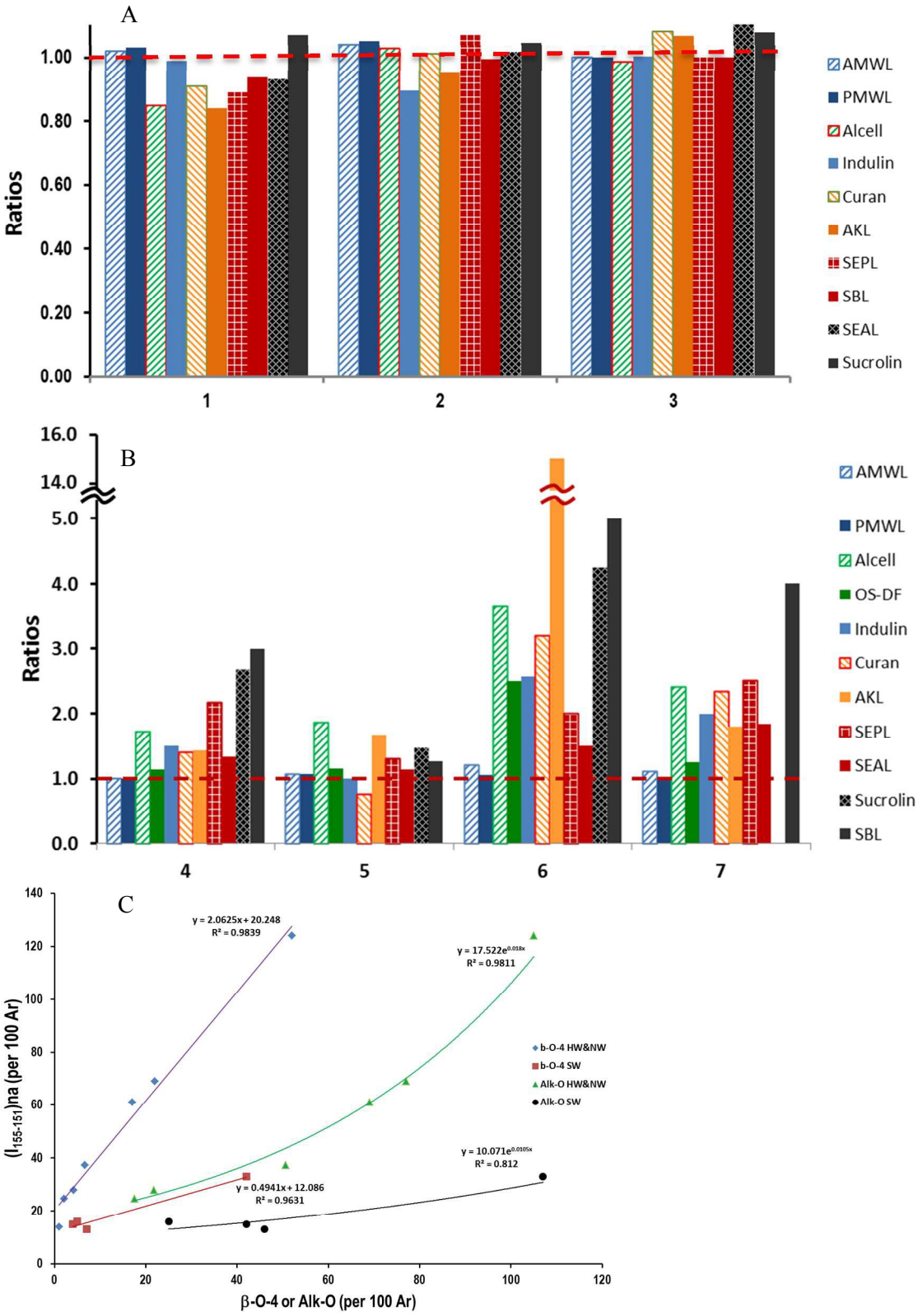


Fig. 3 Various correlations in the quantitative ^{13}C NMR of technical lignins:
A. 1. (S+G+H)/100; 2. ($S_{2,6}$+G_2) / ($S_{3,5}$+G_3) 3. Total OH content measured by way A and B (Table 1) (through the $\underline{C}H_3$-CO and CH_3-$\underline{C}O$ signals, correspondingly)

B. Ratios between 4. the (β–β + β–5) content measured from the resonance at 54-53 ppm in the spectra of non-acetylated lignins to the (β–β + β–5) content measured from the resonance at 88-84 ppm in the spectra of acetylated lignins [(I_{54-53})na/(I_{88-84})ac]; 5. the β–β content measured from the resonance at 54-53 ppm and 86-84 ppm in the spectra of acetylated lignins [(I_{54-53})ac/(I_{86-84})ac]; 6. Resonance at 62-58 ppm in the spectra of non-acetylated lignins to the β–O-4 content [(I_{62-58})na/β–O-4]; 7. Resonance at 67-58 ppm to the amount of primary OH groups [$(I_{67-58})/OH_{pr.}$];
C. correlations between the amount of β–O-4 and alkyl ether moieties (Alk-O) and the resonance at 155-151 ppm in the spectra of non-acetylated lignins.

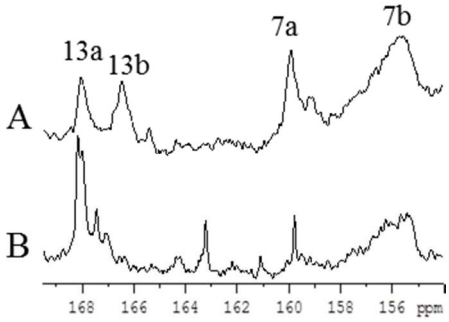


Fig. 4 Expanded regions of conjugated COOR and H-units in the spectra of Sucrolin (A) and SBL (B). Numbers correspond to those in Table 6.