How to cite: Angew. Chem. Int. Ed. 2024, e202404442
doi.org/10.1002/anie.202404442

Biomass

Quantification of Native Lignin Structural Features with Gel-Phase 2D-HSQC$_0$ Reveals Lignin Structural Changes During Extraction

*Claire L. Bourmaud, Stefania Bertella, Anna Bosch Rico, Steven D. Karlen, John Ralph, and Jeremy S. Luterbacher**

Abstract: Our ability to study and valorize the lignin fraction of biomass is hampered by the fundamental and still unmet challenge of precisely quantifying native lignin's structural features. Here, we developed a rapid elevated-temperature ^{1}H–^{13}C Heteronuclear Single-Quantum Coherence Zero (HSQC$_0$) NMR method that enables this precise quantification of native lignin structural characteristics even with whole plant cell wall (WPCW) NMR spectroscopy, overcoming fast spin relaxation in the gel phase. We also formulated a Gaussian fitting algorithm to perform automatic and reliable spectral integration. By combining HSQC$_0$ measurements with yield measurements following depolymerisation, we can confirm the combinatorial nature of radical coupling reactions during biosynthesis leading to a random sequential organization of linkages within a largely linear lignin chain. Such analyses illustrate how this analytical method can greatly facilitate the study of native lignin structure, which can then be used for fundamental studies or to understand lignin depolymerization methods like reductive catalytic fractionation or aldehyde-assisted fractionation.

[*] C. L. Bourmaud, S. Bertella, A. Bosch Rico, Prof. J. S. Luterbacher
Laboratory of Sustainable and Catalytic Processing,
Institute of Chemical Sciences and Engineering
École Polytechnique Fédérale de Lausanne (EPFL)
CH-1015 Lausanne (Switzerland)
E-mail: jeremy.luterbacher@epfl.ch

S. D. Karlen, J. Ralph
U.S. Department of Energy (DOE) Great Lakes Bioenergy Research Center
Wisconsin Energy Institute
University of Wisconsin
Madison, WI 53726 (USA)

J. Ralph
Department of Biochemistry
University of Wisconsin
Madison, WI 53706 (USA)

Introduction

Lignocellulosic biomass tissues, composed of cellulose (35–50 wt %), hemicelluloses (20–35 wt %), and lignin (10–25 wt %) form the largest fraction of terrestrial biomass and provide a possible renewable alternative to fossil-carbon feedstocks.[1] Amongst its components, lignin has the potential to become a feedstock for producing high-value renewable aromatic chemicals as it is the largest natural source of aromatic functionalities.[2–4] Although some have proposed that lignin synthesis is biologically directed,[5,6] lignin biosynthesis is currently considered to be a combinatorial process in which polymerization is only initiated biologically but then proceeds entirely chemically.[7–9] As a result, lignification *in planta* is theorized to occur from combinatorial radical polymerization of monolignols, the major reaction being the cross-coupling of monolignols with the growing oligomer/polymer in the cell wall in a so-termed endwise (cross-)coupling process.[10] Based on monolignol substitution and oxidation potentials, the formation of the β-ether linkage is largely favoured leading to ratios of given linkages that are controlled by such aforementioned reaction energetics.[11] However, the sequence in which these linkages assemble is largely unknown and has frequently been assumed to follow a random order.[12] The debate surrounding lignin branching is also intricately tied to the identification and characterization of various types of branching points,[13] including biphenyl ether (5-O-4),[14–16] dibenzodioxocin (5-5/4-O-β),[10] or benzyl ethers (alkyl-O-alkyl),[17,18] highlighting the complexity and ongoing exploration of lignin's structural diversity. As a result, lignin is recognized as a heterogeneous and complex racemic phenylpropanoid polymer, perhaps linked to polysaccharide sub-units in lignin carbohydrate complexes (LCC) to form a highly connected and resistant matrix (Figure 1a).[9] LCC characterization is mostly done via solution-state nuclear magnetic resonance (NMR) spectrometry to determine its chemical composition, and solid-state NMR to determine the connectivity between components.[19,20] Pyrolysis gas chromatography mass-spectrometry (GC-MS), which requires no isolation step, can also be used to characterize LCCs, provided that suitable standards are used.[21–23] Five types of lignin-carbohydrate bonds are suggested in the literature, including the phenolic glycoside and the benzyl ether moieties such as those presented in Figure 1a, γ-esters, ferulate or *p*-coumarate esters.[24] The most well-established cross-linking mechanism between hemicelluloses and lignin is that mediated by the radical cross-coupling of lignin monomers with ferulate (and

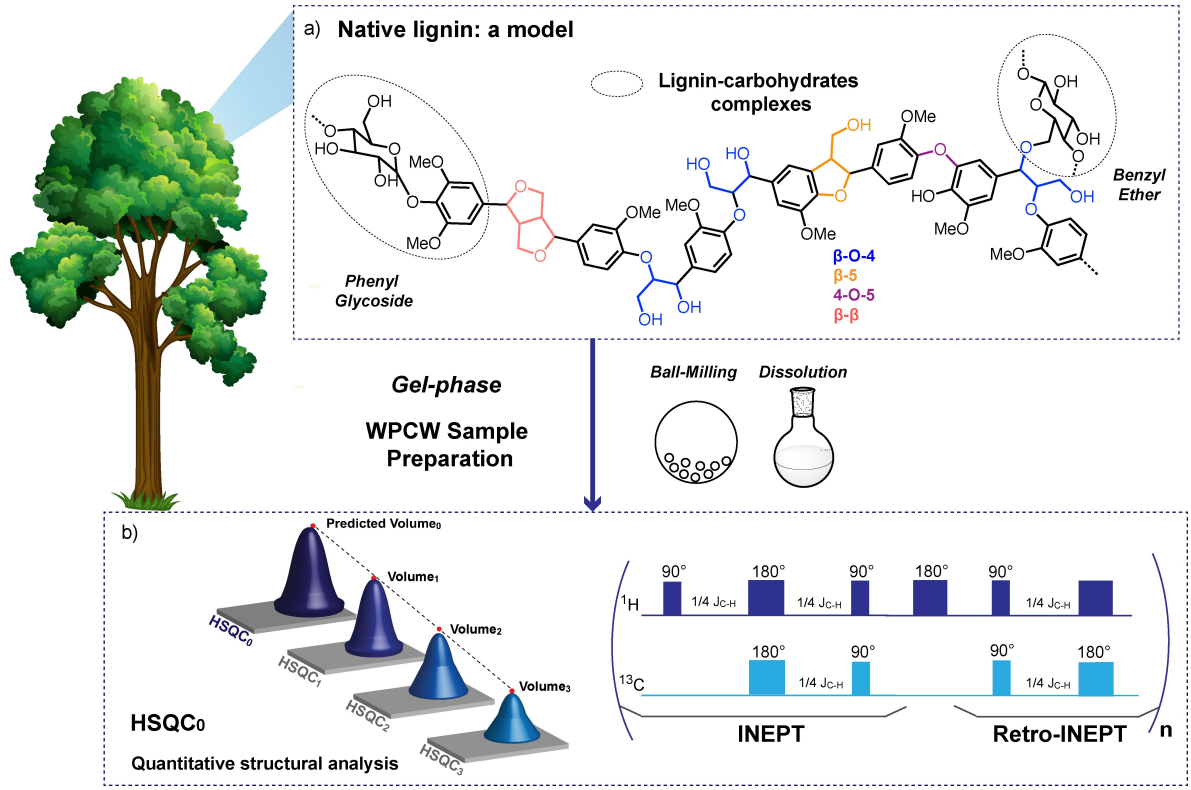


Figure 1. a) A model for native lignin structure within the complex cell wall network comprising lignin-carbohydrates complexes. b) The general principle of WPCW $HCQC_0$ NMR approach and a simplified $HSQC_n$ sequence.

diferulate) esters acylating arabinosyl units in monocot arabinoxylans.[25]

Lignin's complex macromolecular structure involving various ether and carbon-carbon interunit linkages (Figure 1a) makes it difficult to accurately and simultaneously resolve lignin content and chemical composition and structure through traditional spectroscopic methods.[26,27] Most quantitative information on structural components is obtained or inferred from degradative analytical methods that include DFRC,[28,29] nitrobenzene oxidation,[30–32] and thioacidolysis,[33,34] all of which have limitations for determining the composition of native lignins.[13] Total lignin content remains best-determined by the Klason method, a biomass compositional analysis procedure involving hydrolysis of polysaccharides in concentrated and then dilute sulfuric acid and gravimetric determination of the residual insoluble lignin followed by an independent determination of 'acid-soluble lignin' in solution.[35,36] Even though it is frequently used as the reference analysis, the tedious Klason protocol often leads to overestimation of the lignin content, as sugar units may condense during extraction and are then counted as part of the remaining insoluble organic product as "lignin".[37] The actual lignin is also known to substantially condense and degrade during the process which leads to an only semi-accurate determination of its content and gives little to no structural insights into the biomass itself.[38]

More recently, new protocols based on NMR have been developed, allowing for structural analysis of whole-plant-cell-wall (WPCW) materials notably using ^{1}H–^{13}C bidimensional heteronuclear single-quantum coherence (HSQC) NMR sequences.[39] Following controlled milling, biomass dissolution in non-derivatizing deuterated solvents has been extensively studied to adequately solubilize the majority of the lignin to allow solution-state NMR spectra that are essentially devoid of signals from the highly crystalline cellulose. The main challenge of this approach is to find a suitable solvent system, allowing for dissolution and gelation of the wood powder to reach sufficient polymer mobility.[40] Among the solvent systems used, DMSO-d_6/pyridine-d_5 offers better dissolution properties than DMSO-d_6 alone which leads to enhanced gelation of the biomass and highly viscous samples. By disrupting labile intermolecular interactions (e.g., hydrogen bonds), pyridine improves lignin polymer mobility within the gel and decreases the overall viscosity of the solution.[41,42] A general procedure has been extensively detailed and applied to several feedstocks.[43] In these protocols, the biomass particle size appeared to be an essential parameter to maximize the dissolution process while avoiding polymer condensation/destruction during milling. Once optimized, sufficiently intense 2D HSQC spectra could be obtained, allowing for the identification of native lignin structural features before its valorization.[44]

However, these HSQC methods are not quantitative, notably because of traversal spin relaxation (T_2) that characterizes the decay to zero of the *xy* magnetization between pulses and before detection. As the intensity of the signal depends on this relaxation, and because this relaxation rate is resonance-specific (i.e., non-uniform across the

different lignin functionalities), this effect prevents quantification via simple integration of peak volumes. Several alternative sequences have been developed to overcome this relaxation issue.[45–50] These approaches were largely not applicable to whole-cell-wall materials however. They either involve the extraction of small fractions of lignin or neglect the large T_2 variations among cell wall components, especially problematic in viscous WPCW samples. We previously demonstrated that the use of extrapolated time-zero $HSQC_0$ NMR experiments could accurately quantify lignin chemical functionalities in solution (Figure 1b).[49] Relaxation errors were overcome by repeating HSQC blocks one to three times and logarithmically extrapolating signal intensities back to zero time at which no T_2 relaxation had yet occurred. These extrapolated signals are then proportional to C–H bond quantities and independent of resonance-specific decay rates. The resulting peak volumes can also be used to measure real analyte concentration by comparison to the extrapolated volume of an internal standard. Using this method, we confirmed that the level of β-O-4 ether units was a key parameter controlling lignin depolymerization yield by hydrogenolysis.[51,52]

The $HSQC_0$ method corrects for the varying T_2 relaxation rates of a mixture by acquiring three spectra, with increasing acquisition time (between the end of the first ^1H excitation pulse and the beginning of data acquisition) and signal decay, and projecting measured volumes over the three modules back to the beginning of the experiment. It is particularly useful for heterogeneous polymers like lignin, in which potential end-units (*p*-coumaric acids, ferulic acids, LCC…) relax really differently from the polymer core. This decay rate is highly dependent on the molecular weight of the molecule and the viscosity of the NMR sample, both of which significantly accelerate the decrease of the signal-to-noise ratio. In gel-phase samples, the signal attenuation via T_2 relaxation is therefore particularly important with the backbone units in long lignin chains. In fact, when applying the $HSQC_0$ sequence to whole-cell-wall samples, the T_2 has always been too short to acquire usable spectra. Specifically, the last and longest HSQC module in a sequence made of three blocks always had a limited signal-to-noise ratio that prevented successful integration (Figure 2a). These limitations drastically limit known methods to only quantifying and studying isolated soluble lignins that are inherently different than native lignins. The study of native lignins is, however, essential to answer fundamental questions in plant science and better understand lignin extraction methods.

By studying lignin relaxation pathways and magnetization transfer characteristics of several extracted lignins and ball-milled whole-cell-wall samples, we propose an elevated-temperature HSQC method that reduces solution viscosity, decreases relaxation rates, and improves resolution. Coupled with the use of an appropriate internal standard, using improved acquisition parameters significantly enhances the NMR signal. In parallel, a careful control of lignin dissolution after ball-milling also enables us to accurately estimate the lignin content of biomass, offering an alternative to the Klason procedure by quantifying both total lignin content and the level of individual interunit ether linkages. Applying this method to both biomass and extracted lignins, we studied and contrasted the production of aromatic monomers directly from native biomass and from isolated lignins, highlighting important differences in the lignin linkages distribution between these two substrates.

Results and Discussion

Controlling Spin Relaxation at Elevated Temperature

With our previous application of the $HSQC_0$ sequence,[49] spectral noise reached up to 15 % of lignin peak intensity, even with soluble samples, during the third HSQC module

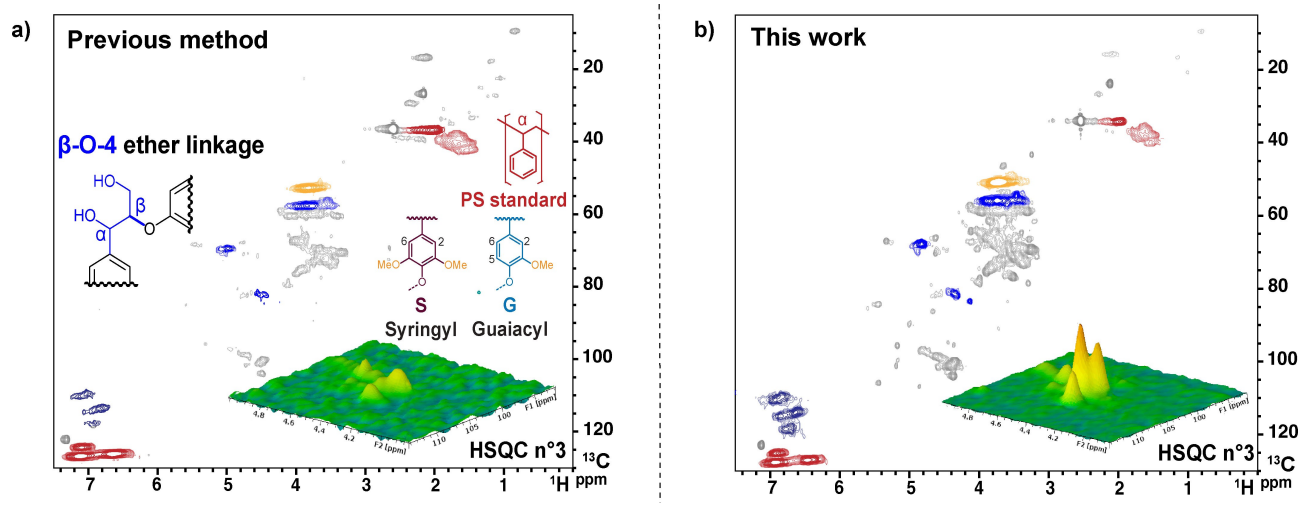


Figure 2. a) 2D pine wood WPCW-$HSQC_3$ contours and mountain-plot sub-spectra highlighting the β-ether signals acquired with the previous conditions[49] employed at 25 °C, which shows an unusable third HSQC module. b) Analogous data acquired with the present work at 55 °C, showing the higher signal-to-noise ratio with controlled T_2 relaxation.

because of spin relaxation. Such signal decay is a source of errors when processing the last spectra because of this low signal-to-noise ratio (Figure S1).

Relaxation was even more important with viscous WPCW solutions and almost no residual signal could be detected in the third HSQC module, rendering an $HSQC_0$ approach impossible (Figure 2a). Short T_2 also impacted the peak definition due to line-broadening.[53] This leads to important overlap and difficult integration of peak volumes in the aliphatic region in which most ether units are encountered (δ_H/δ_C 2.8–5.2/50–110 ppm, Figure 2a). In these instances, integration appeared to be particularly challenging and sensitive to the integration method.

The $HSQC_0$ sequence was thus run at a higher temperature primarily to reduce the solution viscosity and the rotational correlation time, τ_c. The specific impact of temperature on T_2 relaxation was monitored with HSQC spin-echo sequences on extracted lignins. Both ^1H and ^{13}C T_2 were increased by at least 40 % at 55 °C vs 25 °C, resulting in a greater signal intensity and smaller signal width (sigma) in both ^1H and ^{13}C dimensions (Figure 3a–b). In an NMR experiment, the temperature dependence of the spin distribution between high and low energy levels is predicted by the Boltzmann distribution. As the signal-to-noise ratio increases with the population ratio of lower to higher energy spin states, sensitivity is usually better at lower temperatures. However, in viscous samples such as WPCW, decreasing viscosity with increasing temperature strongly reduces signal loss due to spin-spin T_2 relaxation and more than compensates for the signal decrease associated with the population decrease of the lower-energy spin states. We found increases in signal-to-noise ratio from 25 °C to 55–65 °C; at the higher end, a plateau begins to appear (Figure S2). Even if we had used a more heat-resistant probe, increasing temperature further may have led to lignin degradation and may not have helped with signal-to-noise issues; we therefore chose to perform acquisitions at 55 °C. Under these conditions, no lignin degradation or structural modifications were observed during analysis (Figure S3). The improved signal intensity and reduced broadening contributed to better peak definition and easier integration due to a higher signal-to-noise ratio in 2D NMR.

Additionally, the use of high-field NMR and a cryoprobe with what was originally termed "inverse geometry", in which the ^1H coil is located on the inside closest to the sample, was necessary to boost the third HSQC signal's relative intensity, especially in gel-phase samples (Figure S4). Cooling the probe's electronics with liquid helium to 20 K reduces electronic noise and therefore increases the sensitivity by 5-fold. Inverse geometry is particularly useful with indirect experiments like HSQC in which the signal is detected via the ^1H nucleus. These features proved to be essential when working with gel-phase samples in which lignin represents only 15–25 wt % of the cell wall mass that was characterized. Altogether, when probe-temperature elevation was applied to gel-phase samples, the experiment was sufficiently sensitive to maintain intense residual signals during the 3rd HSQC module, allowing for the extraction of quantitative information from native lignin in such WPCW samples (Figure 2b).

Our previous method[49] was run continuously for 60 h, which strongly limited its use as a routine experiment on shared instruments. This long run time was due to the long spin-lattice relaxation T_1 of the internal standard, tetramethylsilane (TMS). T_1 times characterize how fast the magnetization returns to its Boltzmann equilibrium state after a series of NMR pulses. It determines the time required between scans during each acquisition to build the three 2-dimenisonal spectra in an $HSQC_0$ sequence. Acquiring the data at 55 °C significantly shortened this parameter for lignin, allowing the acquisition to be reduced to 15 h and performed overnight. In fact, T_1 does not correlate linearly with τ_c, and the T_1 of polymers in this molecular weight range tends to decrease with temperature (Figure 3c) unlike what is observed for small molecules (Figure S5). Because of this difference between polymers and small molecules, the use of a polymeric internal standard like polystyrene (PS) was necessary to provide a standard with a sufficiently small

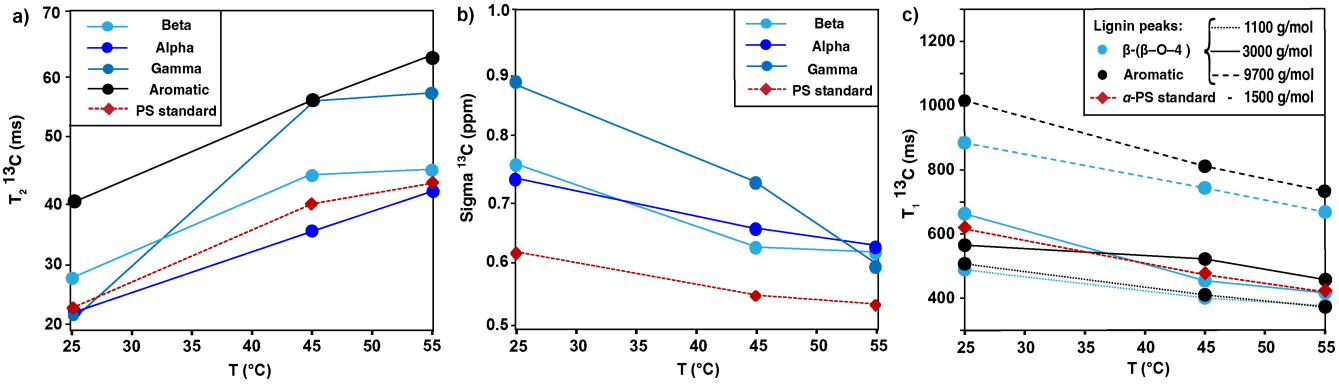


Figure 3. Effect of temperature on the HSQC results. a) Transverse relaxation time (T_2) evolution with temperature for aromatic and β-ether features (α, β, γ). b) Impact of temperature on the HSQC's spectral resolution represented by ^{13}C signal-width reduction. c) Reduction of longitudinal relaxation time (T_1) with temperature for lignin of different average molecular weights compared to a polymeric polystyrene standard's benzylic (α) C–H correlation peak: similar evolutions are observed for lignin's aliphatic α-C–H and β-C–H correlations (blue lines) and aromatic (black lines) groups and for the internal standard (red lines), without a significant impact of molecular weight.

T_1 at the elevated temperature used (Figure S5). Relaxation tests were performed to verify the impact of PS in the solution and we observed that 5–10 mg of 3000 g/mol PS in a 500–1000 μL solution led to similar T_2 constants for both ^1H and ^{13}C relaxations (Table S1, Figure S6).

Given that T_1 relaxation can vary with τ_c, we explored whether lignin polymers of varying molecular weight and various hydroxyl content would see similar relaxation effects with temperature to ensure that the method was robust and consistent across lignin sources.

A sample of propionaldehyde-stabilized birch wood lignin was extracted via Aldehyde-Assisted-Fractionation[55,56] (AAF) (a method used to extract lignin while preserving β-O-4 linkages, as discussed below). This sample was then used to study the effect of molecular weight and polarity in lignin samples. Lignin was fractionated with a solvent gradient to isolate different ranges of molecular weights. To do so, we assumed that higher molecular weight oligomers would be more soluble in a highly polar solvent (Table S2), as was confirmed by gel-permeation chromatography. The analysis of all eleven fractions (from 0 to 40 % acetone/methanol in diethyl ether) confirmed an efficient fractionation with molecular weight ranges varying from 1,100 to 18,700 g/mol, all of which had low polydispersity (PDI) (around 1.5, Table S3, Figure S7). Whereas all fractions reached similar monomer yields after hydrogenolysis (Figure S8), higher aliphatic and carboxylic OH content was monitored within smaller fractions (Figure S9). Relaxation analyses at varying temperatures revealed that the T_1 of all fractions decreased consistently with temperature, even with the lowest molecular weight fraction (Mn = 1100 g/mol) (Figure 3c). This consistent decrease ensured that all lignin samples could be studied at 55 °C, and that this temperature could systematically reduce experimental time.

HSQC sequences are also built around an Insensitive Nuclei Enhanced by Polarization Transfer (INEPT) signal enhancement block, based on scalar coupling for the polarization transfer from ^1H to ^{13}C. To maximize the signal intensity via the INEPT block, HSQC spectra should therefore be acquired at a frequency matching the proton-carbon scalar coupling frequency, $J_{C\text{-}H}$, of the chemical group studied. For example, aliphatic carbons have protons coupled by 145 Hz whereas an aromatic carbon is at around 170 Hz to its attached protons. The use of polystyrene, which has structural similarities to lignin polymers (in both aliphatic and aromatic groups), enables the acquisition of corresponding HSQC signals under the same conditions (SI Section 3.3.10, Figure S10).

Spectral Processing

We found that user choice during manual integration could lead to significant differences in reproducibility (Figure S11a–b), and thus explored the use of an automatic integration method. Chylla et al. had developed a software called Newton that performs 2D deconvolution of NMR peaks based on a fast maximum likelihood reconstruction (FMLR) algorithm.[57,58] This method allows for an accurate fitting of peaks with Lorentzian shapes. However, when applied to gel-phase HSQC experiments, the residual spectrum, i.e., the difference between NMR data and that from peak fitting, revealed that peaks were poorly fitted leading to high quantification variance. We therefore developed an automatic integration algorithm based instead on a spectral deconvolution using Gaussian functions to analyze the ether region and significantly improved the resulting peak fit. This algorithm included the following three steps:

i. *Data conversion*: Each HSQC NMR spectrum was read and converted into three text matrices specifying the lists of ^1H chemical shifts, ^{13}C chemical shifts, and 2D intensities using the *rbnmr* Matlab function developed by Bruker. Alternatively, NMRPipe, a UNIX spectral processing system, can be used to perform this operation.
ii. *Local maxima investigation and selection*: Each peak I[m,n] was identified and defined by its local maximum with m and n, the ^1H and ^{13}C matrix indexes respectively. This initial identification produced a list of peaks along with their maximal positions from which the user is then asked to choose for further individual analysis.

$$\mathrm{I}[\mathrm{m},\mathrm{n}] \text{ if } \begin{pmatrix} I[\mathrm{m},\ \mathrm{n}+1] \textit{ and } I[\ \mathrm{m},\ \mathrm{n}-1] \\ \textit{and } I[\mathrm{m}+1,\ \mathrm{n}] \textit{ and } I[\mathrm{m}-1,\ \mathrm{n}]\ < I\ [\mathrm{m},\mathrm{n}] \end{pmatrix} \quad (1)$$

iii. *Peak modelling*: For each peak, the head of the peak (defined as the peak volume that occurs above 50 % of the maximum intensity) was fitted with a 2D Gaussian function. The use of the peak head was only to avoid interference by the overlapping peak regions at the bottom (Figure S11c). This method is yet to be demonstrated with smaller peaks or more overlapping ones. Peak width (^1H and ^{13}C), height, volume, and quality of the fitting were then extracted and reported into a .csv file. Once the three spectra were analyzed, a hypothetical peak volume at a relaxation time of zero was estimated, using a logarithmic extrapolation of the integrated peak volumes. An average of all β-ether sidechain signals (α, β, γ) was calculated and returned by the algorithm.

The natural shape of NMR peaks in solution is typically Lorentzian, whereas Gaussian shapes are usually used to describe solid or inhomogeneous systems.[60] We obtained better results by fitting Gaussian rather than Lorentzian functions (Figure S12). This effect can be explained by local magnetic field and sample inhomogeneities in addition to the high viscosity of the whole-cell-wall NMR sample. The better Gaussian fit could explain why poor fits were obtained with Lorentzian FMLR in prior studies,[47,57,58] despite its reliability in alternate systems.

Method Validation in Gel-Phase Samples

High-resolution $HSQC_0$ spectra were successfully recorded for whole-cell-wall gel samples allowing for the unprecedented quantitative analysis of native lignin (Figure 2b). These well-resolved two-dimensional spectra contrasted with the corresponding ^1H NMR spectra that had very poor resolution and resembled solid-state NMR spectra of whole cell walls (Figure S13). We suspect that the HSQC sequence acted as a pseudo T_2-filter, suppressing the rapidly relaxing signal of suspended cellulose that can still be detected in ^1H NMR in which we see significant line broadening compared to fully dissolved lignins.

Assessing the reliability of the method was quite challenging as other quantification methods also come with uncertainties and poorly describe native lignin features.[13] We therefore started our analysis by simulating a cell wall model with known component quantities. This model was obtained by mixing cellulose-rich solids isolated via the AAF method,[56] commercial xylan from beechwood, and synthetic lignin polymer model compounds so as to reach a similar viscosity to ball-milled wood samples (as confirmed by their similar T_2 parameters, Table S4). The accuracy of the method was validated as quantification of β-ethers, based on the average of duplicate experiments, was within 5% error based on the mass of polymer introduced (Figure S14).

We also used solid-state NMR to assess the quantification of dissolved lignin. A reference quantitative ^{13}C NMR spectrum was acquired using direct excitation of carbon (*echomas* sequence). This sequence gives us a 1D-^{13}C spectrum of the wood powder and enables the measurement of the total lignin content in the aromatic region by integrating multiple peaks, without resolution of individual functionalities. Lignin peaks were fitted using *dmfit* software,[20,61] and quantification of lignin was obtained by comparison to an internal standard, in mmol per gram of wood. The WPCW method was then applied to the same wood sample after additional milling, and we could quantify the dissolved lignin content using $HSQC_0$. The remaining solid was collected, filtered, and submitted to additional *echomas* analysis to quantify the amount of residual undissolved lignin. A molar balance successfully estimated the initial lignin content with a 3.75% error (dissolved lignin was 58% and residual lignin was 46% of the originally measured lignin) (Table S5). However, to avoid (i) characterizing only a small fraction of the total lignin and (ii) the need for an additional solid-state measurement due to partial dissolution of the lignin, we further developed a biomass dissolution protocol to improve lignin solubilization.

Improving Biomass Dissolution

As efficient dissolution is key to enable sufficient NMR signal during the $HSQC_0$ acquisition and to ensure that we are measuring a representative fraction of the lignin, we further improved the lignin dissolution protocol. Ball-milling decreases particle size and thus increases the surface area to facilitate solvent diffusion and lignin dissolution, but it also can alter the lignin's native structure if too extensive. Procedures for ball-milling depending on the feedstock (hardwood or softwood) and material loading have already been reported.[42,54,62] When reaching a particle size of around 5 nm, alteration and destruction of specific cell wall structures was observed.[63] Our own experience revealed a limited reproducibility of ball-milling and dissolution steps from the literature, which were highly dependent on the loading, starting and final material particle size, as well as equipment available. Therefore, we focused on a need to optimize the procedure based on the apparatus used and the material studied.

Starting from birch wood that had been through a cutting Mill with a 125 µm screen, we performed several ball-milling steps, using 50 mL ZrO_2-lined jars in a Retsch ball-mill and ten 10 mm ZrO_2 ball-bearings, to reach a distribution of particles size and studied the impact of milling time on lignin dissolution and β-O-4 degradation (Figure 4). Although a short milling time preserved lignin β-ether content almost completely, we observed a substantial degradation of such ether units when the sample was submitted to extensive ball-milling (which is necessary for complete lignin dissolution). Though some of this decrease could be due to lignin with fewer ether linkages being harder to dissolve and coming later, the fact that lignin dissolution plateaus while β-O-4 content continues to decrease points to degradation. We concluded that quantifying both the β-O-4 and lignin contents simultaneously and accurately was not possible. Therefore, we propose a two-step protocol that involves milling 200 mg for 90 min (0.45 min/mg, 10 ZrO_2

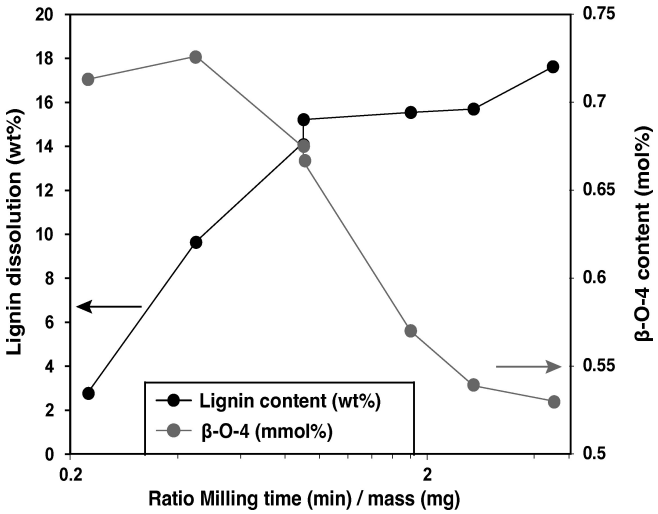


Figure 4. Evolution of lignin dissolution (wt%) (left, black) and β-ether degradation (mol%) (right, grey) based on birch (hardwood) ball-milling time and mass ratio. Limited milling time yields large particles and low lignin dissolution but maintains high β-ether content (left part). Conversely, extensive ball-milling solubilizes more lignin while simultaneously degrading β-ethers. Conditions: 50 mL ZrO_2-lined jar in a Retsch ball-mill, 200 mg of 125 µm-wood (the same batch was used), ten 10 mm ZrO_2 ball-bearings, 450 rpm, 5 min on, 5 min off.

ball bearings, 10 min on, 10 min off) for β-O-4 content analysis and milling 200 mg for 9 h (2.7 min/mg, 10 ZrO_2 ball-bearings, 10 min on, 10 min off) to fully dissolve lignin and quantify its content. The effect of condensation of aromatic subunits with β-ether degradation was neglected during this measurement. This assumption was justified by the fact that even if one assumes that condensation resulting from the loss of a β-ether unit was 100 % complete and selective (which is highly unlikely) on the aromatic ring positions, this would always lead to an underestimation of aromatic content below 10 % when integrating syringyl and G_2 signals (SI. 3.3.13).

Feedstock Analysis

Having developed an optimized procedure, we investigated several well-studied feedstocks[56,64–69] and analyzed their lignin and β-O-4 content (SI. 3.3.14). *p*-Coumarate (*p*CA) and *p*-ferulate (*p*FA) esters are notably found as terminal units in grass lignins; these aromatic units were also integrated and taken into account in the total lignin content. Our data were compared with analyses of lignin using the Klason lignin procedure (Table S6). We reached similar, but usually slightly reduced, lignin contents compared to the Klason Lignin procedure (Table 1). As Klason lignin analyses can overestimate lignin by including sugar degradation products or leftover extractives,[37,70] this slight decrease in lignin content was not surprising and could be indicative of a more accurate result.

We also compared our lignin content analysis to lignin monomer yields following reductive catalytic fractionation (RCF) of this material (Figure 5a, Table 1).[56,64–69,71] RCF involves treating whole biomass (i.e., whole-cell-wall samples) at high temperature in an organic solvent in the presence of a hydrogenolysis catalyst and a hydrogen source (in this case H_2). These conditions lead to the direct conversion of lignin to aromatic monomers and oligomers via hydrogenolysis of β-O-4 linkages.[26] As hydrogenolysis largely outpaces lignin condensation reactions and cleaves nearly all β-O-4 linkages, the yield of monomers can provide direct insights into lignin structure.

Monomer yields are related to the original easily-cleavable ether content by making assumptions about the native lignin structure. We make the distinction between easily cleavable ether linkages such as β-O-4 ethers, more recalcitrant ethers like 4-O-5 ethers, and all C–C linkages that are even more recalcitrant to cleavage because hydrogenolysis (under these conditions) is inactive against them. Assuming a random sequence of β-O-4 ether linkages within the polymer, monomeric units can only be produced via hydrogenolysis when linked to two cleavable ether linkages, or by one at terminal positions.[52] Therefore, monomer yield can be predicted by the β-O-4 content measured using the following correlation, which assumes a linear lignin polymer, where m is the number of aromatic subunits within the polymer chains:

$$\textit{Monomer yield}\ (\textit{mol\%}) = \frac{(\boldsymbol{m}-2)\times(\boldsymbol{\beta}-\boldsymbol{O}-4\ \textit{mol\%})^2}{\boldsymbol{m}} + \frac{2\times(\boldsymbol{\beta}-\boldsymbol{O}-4\ \textit{mol\%})}{\boldsymbol{m}} \quad (2)$$

With a large number of monolignols, m, this equation becomes:

$$\textit{Monomer yield}_{\textit{Random}}(\textit{mol\%}) = (\boldsymbol{\beta}-\boldsymbol{O}-4\ \textit{mol\%})^2 \quad (3)$$

Assuming long linear native lignin chains, and a random distribution of ethers, Equation 3 has frequently been used to qualitatively explain monomer yields measured after RCF. Monophenolic yields around 50 % for hardwoods have been explained by presumed ether contents of 70–80 mol % within these same hardwoods,[4,73] whereas lower monophenolic yields around 20–25 % in softwoods have been explained by ether contents of 40–50 mol % β-O-4 in this substrate. However, an exact quantification of ether linkages has been elusive and so these correlations could not be precisely verified. Here however, when comparing RCF monophenolic yields with our WPCW $HSQC_0$ data (Table 1), we can precisely verify that this model assuming a random order in the linkage sequence does in fact accurately predict RCF yields (dashed line, Figure 6). This accurate yield prediction by the model strongly suggests that lignin can be represented as a linear polymer with a random sequence of interunit linkages, whereas a branched model fails to predict the data (Figure S15). This result can therefore provide additional evidence for a view in which lignin linkages exist in ratios set by reaction energetics but their sequence is random and largely linear, both of which have been controversial conclusions.

Table 1: Characterization of lignin content (total aromatic content, Tables S7–8) and β-O-4 content from representative biomass samples (hardwood, softwood, grass). The β-O-4 content was calculated with the optimized WPCW $HSQC_0$ method.

Feedstock	Klason lignin content (wt%)	$HSQC_0$ lignin content (wt%)	RCF monomer yield (mol%)	β–O–4 content (mol%)
Birch (★)	17.4	17.3 ± 0.51	48.7	72.5 ± 3.6
Poplar (x)	25.0	22.0 ± 0.66	38.7	60.5 ± 3.0
High-syringyl Poplar (●)	12.8	14.3 ± 0.43	64.4	78.9 ± 3.9
Maple (+)	25.0	21.4 ± 0.64	10.1	35.3 ± 1.8
Black Pine (◆)	29.1	28.1 ± 0.84	24.5	48.0 ± 2.4
Corn Cobs (■)	13.7	15.2 ± 0.46	12.4	30.4 ± 1.5
Sorghum (▲)	16.0	12.9 ± 0.39	11.7	29.7 ± 1.5

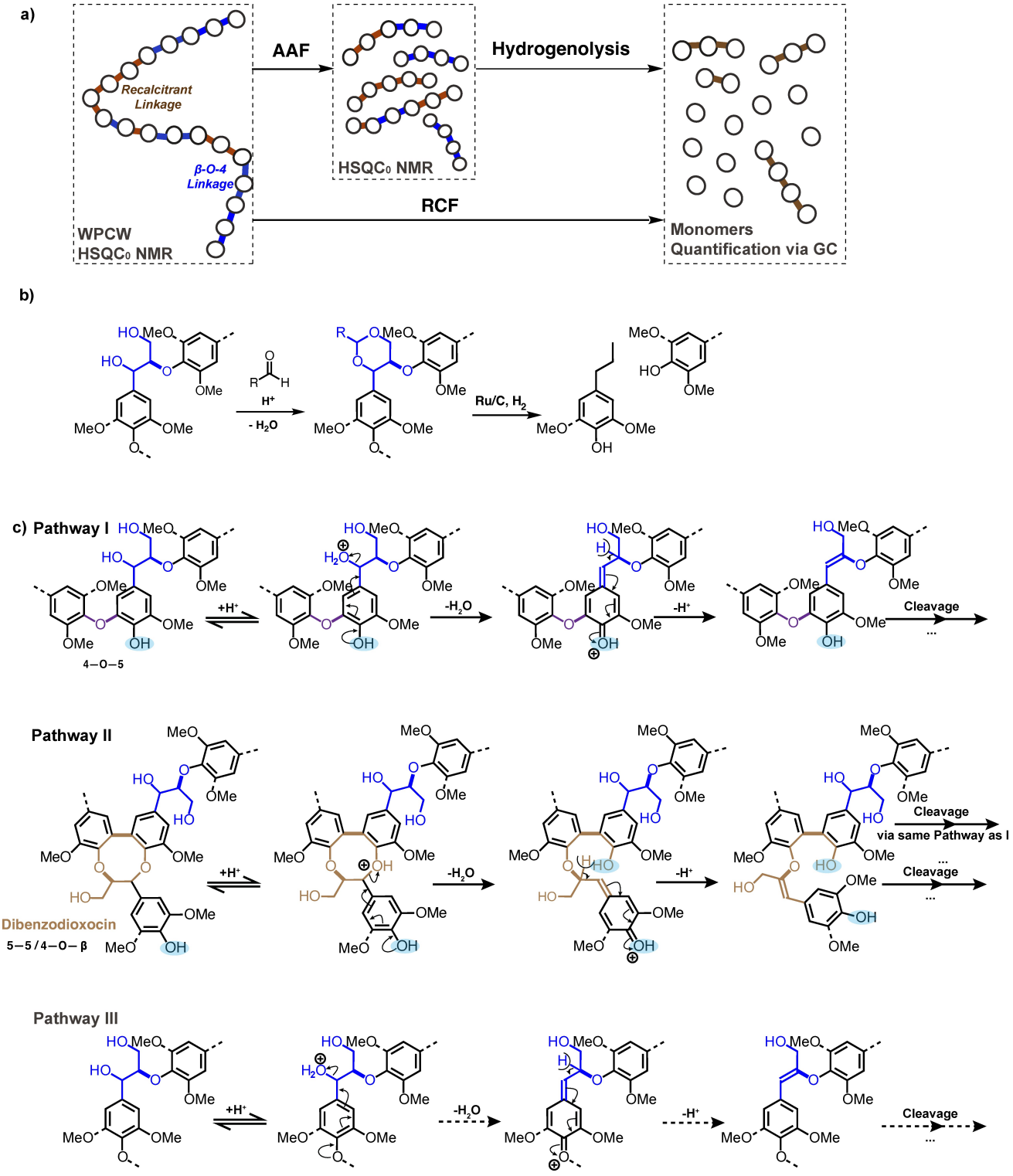


Figure 5. a) Lignin analytical strategy starting from whole cell wall biomass, involving RCF or Aldehyde-Assisted Fractionation (AAF) followed by hydrogenolysis. Possible reorganization of ether linkages from a random to a block sequence may occur during AAF, while still reaching near-theoretical RCF monomer yields after hydrogenolysis. b) Lignin extraction, stabilization with aldehyde, and subsequent hydrogenolysis during AAF. c) Hypothesized preferred β-O-4 cleavage positions under acidic conditions. From top to bottom, diaryl ether (4-O-5), dibenzodioxocins (DBDOX: 5-5, 4-O-β) and non-phenolic β-O-4. Notably, the hydrolysis of non-etherified diaryl ether and DBDOX generates reactive free-phenolic β-ether units (blue shades), that can further be hydrolyzed (or hydrogenolyzed). Starting from non-phenolic β-O-4 linkages, the hydrolysis reaction seems less favored, possible due to limited O-methoxy group electron donor ability (therefore represented with dashed arrows).

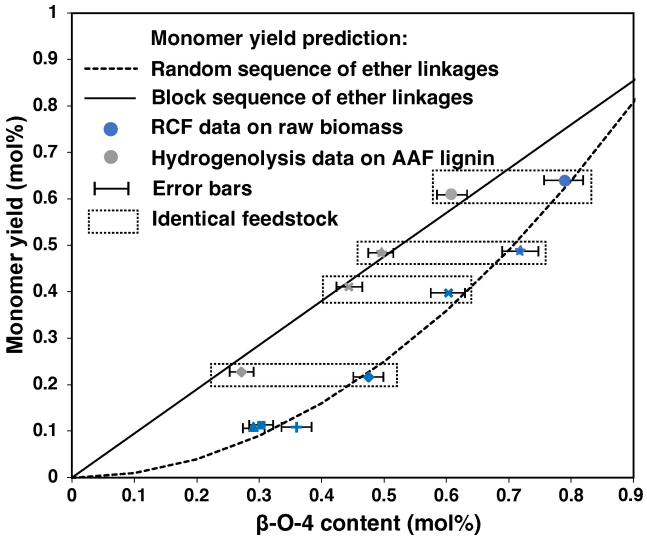


Figure 6. Comparison of experimental and predicted monomer yield (via RCF or hydrogenoylsis) based on native (blue) and AAF extracted lignin (grey) β-O-4 content measured from WPCW using solution-state $HSQC_0$ NMR. Feedstocks studied: (●) High-syringyl Poplar, (★) Birch, (x) Poplar, (■) Corn cobs, (♦) Black Pine, (+) Switchgrass, (▲) Sorghum. Notably, AAF-extracted lignins produce almost theoretical monomer yields after hydrogenolysis compared to RCF yields on raw biomass.

Understanding Lignin Extraction Processes with $HSQC_0$: the Example of Aldehyde-Assisted Fractionation

An alternate model to a random organization of linkages could be one based on the assumption that cleavable β-O-4 linkages and recalcitrant 4-O-5 linkages or C–C linkages are organized in block structures. Specifically, this non-random distribution would feature either β-O-4 linkages or 4-O-5 and C–C linkages clustered together in separate stretches or blocks of polymer. In this case, one β-O-4 bond breaking can generate one monomer after cleavage:

$$\textit{Monomer yield}_{\textit{Block}}(\textit{mol}\%) = \frac{m}{m+1} \times \beta - O - 4 \textit{ content } (\textit{mol}\%) \quad (4)$$

Prior results with $HSQC_0$ on isolated lignin had suggested that acidolysis might result in some such clustering of the isolated lignin oligomers.[49] Our improved analysis method can be used to directly compare lignin within the whole cell wall to lignin extracted via acidolysis. Specifically, we used lignin isolated via AAF for this comparison. This isolation method uses an aldehyde during fractionation in an organic solvent in the presence of an acid catalyst to functionalize and protect β-ether units from condensation (Figure 5b). These conditions lead to removal of the lignin from the plant cells by acidolysis but preserve the lignin's ability to be depolymerized to aromatic monomers at similar yields to RCF on the native biomass, which facilitates direct comparison of the lignin structure.[55] Here we performed AAF with propionaldehyde (Figure 5b) on four feedstocks amongst the ones previously studied with WPCW $HSQC_0$ (Table 1). The β-O-4 content of each extracted and isolated lignin was compared to that from the native lignin using our new $HSQC_0$ procedure. We systematically observed an important decrease of the β-O-4 ether content, which was consistent with cleavage during lignin extraction (Figure 6, blue vs. grey datapoints on the x-axis). However, we are still able to use this extracted lignin to reach monophenolic yields after hydrogenolysis that were equivalent to those from RCF on the native biomass (Figure 6, blue vs. grey data points on the y-axis).

With the lower ether content of the isolated lignin, the random distribution theoretical model (Eq. 3) no longer successfully predicted monophenolic yields (dashed line), suggesting that linkages were no longer randomly distributed. However, as we had previously suggested,[49] the block model (Eq. 4) was successful in predicting monophenolic yields (full line, Figure 6). The successful prediction of the block model for isolated lignin suggests that lignin undergoes a structural/sequential reorganization during AAF (i.e., acid-catalyzed fractionation, Figure 5b). We concluded that a non-random cleavage of β-O-4 linkages occurs during lignin extraction under the acidic AAF conditions. In particular, the recently suggested singular presence of free-phenolic units in the vicinity of certain ether or C–C linkages (diaryl ether, dibenzodioxocins) could explain this preferential cleavage.[10,14,15,74] Indeed, past kinetic studies with model compounds have shown that a free-phenolic group accelerates acidolysis and hydrolysis of ether linkages in lignin,[75–77] due to a benzylic carbocation at the α-position, which is the precursor to cleavage (and lignin condensation[78]), that is better stabilized through resonance (Figure 5c, Pathways I and II). In the case of a non-phenolic β-ether position, this stabilization is less efficient, which makes cleavage slower (Figure 5c, Pathway III), as observed in AAF conditions (Figure S16, Table S9). Therefore, we hypothesize that phenolic β-ether positions that could be found in the vicinity of distinct linkages may act as reactive positions at which cleavage occurs preferentially during lignin extraction. These specific linkages maybe form nearby dibenzodioxocin (5-5/4–O–β) and biphenyl ether (4-O-5) linkages, evidenced to remain free-phenolic (Figure 5c, Pathways I and II) via NMR spectroscopy,[10,14,15,74] or after cleavage of phenyl-glycoside LCCs (Figure 1a) during biomass acidic organosolv treatment.[24] While some controversy exists over phenolic 4-O-5 units as they have not been detected by thioacidolyis,[16] suggesting that there may be a low abundance of β-O-4-linked-4-O-5 units, the observed structural modifications might arise from double hydrolysis of the two phenols arising from the 5-5 units, coupled to LCC hydrolysis. These particular β-ether linkage positions can cleave, leading to soluble oligomeric blocks that are enriched in either ethers or C–C linkages (Figure 5b).

Conclusion

We demonstrated that WPCW $HSQC_0$ NMR can be used as an essentially non-destructive method for quantifying native lignin structural features in biomass. Our results show that

this quantification is relatively simple, reproducible, and less subject to error due to unknown structural features compared to state-of-the-art methods like the Klason protocol. We have used this non-destructive method to unravel unique lignin structural information, accurately and completely quantifying native structural features. Notably, we show strong evidence that lignin ether linkages are randomly organized in the native sequence but appear to be non-randomly distributed following acid-catalyzed isolation methods like AAF. We suspect that preferential β-O-4 cleavage occurs when breaking down the wood 3D structure, and in particular the lignin-polysaccharide complexes. This example illustrates how a precise quantitative structural analysis method like the one presented here can be used to shed new light on longstanding questions in lignin structure and upgrading.

Supporting Information

The authors have cited additional references within the Supporting Information.[79–82]

Acknowledgements

This work was supported by the Swiss National Science foundation through the NCCR Catalysis, a National Center of Competence in Research (Grant 565280) and through Grants 200021_182605 and CRSII5_180258. SDK and JR were funded by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Biological and Environmental Research Program under Award Number DE-SC0018409. The authors thank Y. Lavanchy (EPFL) for his help running gel-permeation chromatography, Dr. A. Bornet (EPFL) for his help and availability running and optimizing NMR experiments. Open Access funding provided by École Polytechnique Fédérale de Lausanne.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

Data will be published in a public repository (zenodo) at the time of publication.

Keywords: biomass · fractionation · lignin · NMR spectroscopy · structure

[1] D. R. Dodds, R. A. Gross, *Science* **2007**, *318*, 1250–1251.
[2] J. Ralph, K. Lundquist, G. Brunow, F. Lu, H. Kim, P. F. Schatz, J. M. Marita, R. D. Hatfield, S. A. Ralph, J. H. Christensen, W. Boerjan, *Phytochem. Rev.* **2004**, *3*, 29–60.
[3] S. Bertella, J. S. Luterbacher, *Trends Chem.* **2020**, *2*, 440–453.
[4] W. Schutyser, T. Renders, S. Van den Bosch, S.-F. Koelewijn, G. T. Beckham, B. F. Sels, *Chem. Soc. Rev.* **2018**, *47*, 852–908.
[5] N. G. Lewis, L. B. Davin, in *Lignin Lignan Biosynth.*, American Chemical Society **1998**, pp. 334–361.
[6] L. B. Davin, N. G. Lewis, *Plant Physiol.* **2000**, *123*, 453–462.
[7] E. Adler, *Wood Sci. Technol.* **1977**, *11*, 169–218.
[8] K. Freudenberg, *Science* **1965**, *148*, 595–600.
[9] R. Vanholme, B. Demedts, K. Morreel, J. Ralph, W. Boerjan, *Plant Physiol.* **2010**, *153*, 895–905.
[10] J. Ralph, C. Lapierre, W. Boerjan, *Curr. Opin. Biotechnol.* **2019**, *56*, 240–249.
[11] K. Syrjänen, G. Brunow, *J. Chem. Soc. Perkin Trans. 1* **1998**, 3425–3430.
[12] R. R. Sederoff, J. J. MacKay, J. Ralph, R. D. Hatfield, *Curr. Opin. Plant Biol.* **1999**, *2*, 145–152.
[13] M. Balakshin, E. A. Capanema, X. Zhu, I. Sulaeva, A. Potthast, T. Rosenau, O. J. Rojas, *Green Chem.* **2020**, 3985–4001.
[14] Y. Li, T. Akiyama, T. Yokoyama, Y. Matsumoto, *Biomacromolecules* **2016**, *17*, 1921–1929.
[15] F. Yue, F. Lu, S. Ralph, J. Ralph, *Biomacromolecules* **2016**, *17*, 1909–1920.
[16] K. Saito, Y. Makimura, H. Nishimura, T. Watanabe, *ChemSusChem* **2021**, *14*, 2554–2563.
[17] X. Zhu, J. Sipilä, A. Potthast, T. Rosenau, M. Balakshin, *J. Agric. Food Chem.* **2023**, *71*, 580–591.
[18] M. Karlsson, J. Romson, T. Elder, Å. Emmer, M. Lawoko, *Biomacromolecules* **2023**, *24*, 2314–2326.
[19] T.-Q. Yuan, S.-N. Sun, F. Xu, R.-C. Sun, *J. Agric. Food Chem.* **2011**, *59*, 10604–10614.
[20] X. Kang, A. Kirui, M. C. Dickwella Widanage, F. Mentink-Vigier, D. J. Cosgrove, T. Wang, *Nat. Commun.* **2019**, *10*, 347.
[21] J. Ralph, R. D. Hatfield, *J. Agric. Food Chem.* **1991**, *39*, 1426–1437.
[22] G. van Erven, R. de Visser, D. W. H. Merkx, W. Strolenberg, P. de Gijsel, H. Gruppen, M. A. Kabel, *Anal. Chem.* **2017**, *89*, 10907–10916.
[23] G. van Erven, R. de Visser, P. de Waard, W. J. H. van Berkel, M. A. Kabel, *ACS Sustainable Chem. Eng.* **2019**, *7*, 20070–20076.
[24] N. Giummarella, Y. Pu, A. J. Ragauskas, M. Lawoko, *Green Chem.* **2019**, *21*, 1573–1595.
[25] J. Ralph, *Phytochem. Rev.* **2010**, *9*, 65–83.
[26] M. M. Abu-Omar, K. Barta, G. T. Beckham, J. S. Luterbacher, J. Ralph, R. Rinaldi, Y. Román-Leshkov, J. S. M. Samec, B. F. Sels, F. Wang, *Energy Environ. Sci.* **2021**, *14*, 262–292.
[27] L. Shuai, M. Talebi Amiri, J. S. Luterbacher, *Curr. Opin. Green Sustain. Chem.* **2016**, *2*, 59–63.
[28] F. Lu, J. Ralph, *J. Agric. Food Chem.* **1997**, *45*, 4655–4660.
[29] F. Lu, J. Ralph, *J. Agric. Food Chem.* **1997**, *45*, 2590–2592.
[30] J. C. Villar, A. Caperos, F. García-Ochoa, *J. Wood Chem. Technol.* **1997**, *17*, 259–285.
[31] C.-L. Chen, in *Methods Lignin Chem.* (Eds.: S. Y. Lin, C. W. Dence), Springer, Berlin, Heidelberg **1992**, pp. 301–321.
[32] M. Yamamura, T. Hattori, S. Suzuki, D. Shibata, T. Umezawa, *Plant Biotechnol.* **2010**, *27*, 305–310.
[33] C. Lapierre, in *Forage Cell Wall Struct. Dig.*, John Wiley & Sons, Ltd **1993**, pp. 133–166.
[34] C. Rolando, B. Monties, C. Lapierre, in *Methods Lignin Chem.*, Springer, Berlin, Heidelberg **1992**, pp. 334–349.
[35] J. Sluiter, A. Sluiter, *Summative Mass Closure: Laboratory Analytical Procedure (LAP) Review and Integration*, National Renewable Energy Laboratory **2011**.
[36] H. G. Jung, D. R. Buxton, R. D. Hatfield, J. Ralph, *Am. Soc. Agron. Inc* **1993**, xv + 794 pp.

[37] M. Bunzel, A. Schüssler, G. Tchetseubu Saha, *J. Agric. Food Chem.* **2011**, *59*, 12506–12513.
[38] G. J. Leary, R. H. Newman, K. R. Morgan, *Holzforschung* **1986**, *40*, 267–272.
[39] J. Ralph, F. Lu, *Org. Biomol. Chem.* **2004**, *2*, 2714.
[40] M. Foston, R. Samuel, J. He, A. J. Ragauskas, *Green Chem.* **2016**, *18*, 608–621.
[41] H. Kim, J. Ralph, *Org. Biomol. Chem.* **2010**, *8*, 576–591.
[42] D. J. Yelle, J. Ralph, C. R. Frihart, *Magn. Reson. Chem.* **2008**, *46*, 508–517.
[43] S. D. Mansfield, H. Kim, F. Lu, J. Ralph, *Nat. Protoc.* **2012**, *7*, 1579–1589.
[44] L. Zhang, G. Gellerstedt, *Magn. Reson. Chem.* **2007**, *45*, 37–45.
[45] S. Heikkinen, M. M. Toikka, P. T. Karhunen, I. A. Kilpeläinen, *J. Am. Chem. Soc.* **2003**, *125*, 4362–4367.
[46] K. Hu, W. M. Westler, J. L. Markley, *J. Am. Chem. Soc.* **2011**, *133*.
[47] K. Hu, J. J. Ellinger, R. A. Chylla, J. L. Markley, *Anal. Chem.* **2011**, *83*, 9352–9360.
[48] M. Sette, R. Wechselberger, C. Crestini, *Chem. Eur. J.* **2011**, *17*, 9529–9535.
[49] M. Talebi Amiri, S. Bertella, Y. M. Questell-Santiago, J. S. Luterbacher, *Chem. Sci.* **2019**, *10*, 8135–8142.
[50] T. Koso, D. Rico del Cerro, S. Heikkinen, T. Nypelö, J. Buffiere, J. E. Perea-Buceta, A. Potthast, T. Rosenau, H. Heikkinen, H. Maaheimo, A. Isogai, I. Kilpeläinen, A. W. T. King, *Cellulose* **2020**, *27*, 7929–7953.
[51] E. M. Anderson, M. L. Stone, R. Katahira, M. Reed, W. Muchero, K. J. Ramirez, G. T. Beckham, Y. Román-Leshkov, *Nat. Commun.* **2019**, *10*, 2033.
[52] T. Phongpreecha, N. C. Hool, R. J. Stoklosa, A. S. Klett, C. E. Foster, A. Bhalla, D. Holmes, M. C. Thies, D. B. Hodge, *Green Chem.* **2017**, *19*, 5131–5143.
[53] J. Kowalewski, L. Mäler, *Nuclear Spin Relaxation in Liquids: Theory, Experiments, and Applications*, CRC Press, Boca Raton **2017**.
[54] H. Kim, J. Ralph, *Org. Biomol. Chem.* **2010**, *8*, 576–591.
[55] L. Shuai, M. T. Amiri, Y. M. Questell-Santiago, F. Héroguel, Y. Li, H. Kim, R. Meilan, C. Chapple, J. Ralph, J. S. Luterbacher, *Science* **2016**, *354*, 329–333.
[56] M. Talebi Amiri, G. R. Dick, Y. M. Questell-Santiago, J. S. Luterbacher, *Nat. Protoc.* **2019**, *14*, 921–954.
[57] R. A. Chylla, B. F. Volkman, J. L. Markley, *J. Biomol. NMR* **1998**, *12*, 277–297.
[58] R. A. Chylla, K. Hu, J. J. Ellinger, J. L. Markley, *Anal. Chem.* **2011**, *83*, 4871–4880.
[59] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, *J. Biomol. NMR* **1995**, *6*.
[60] L. Petrakis, *J. Chem. Educ.* **1967**, *44*, 432.
[61] D. Massiot, F. Fayon, M. Capron, I. King, S. Le Calvé, B. Alonso, J.-O. Durand, B. Bujoli, Z. Gan, G. Hoatson, *Magn. Reson. Chem.* **2002**, *40*, 70–76.
[62] H. Kim, J. Ralph, T. Akiyama, *BioEnergy Res.* **2008**, *1*, 56–66.
[63] M. Hedenström, S. Wiklund-Lindström, T. Öman, F. Lu, L. Gerber, P. Schatz, B. Sundberg, J. Ralph, *Mol. Plant Pathol.* **2009**, *2*, 933–942.
[64] J. Behaghel de Bueren, F. Héroguel, C. Wegmann, G. R. Dick, R. Buser, J. S. Luterbacher, *ACS Sustainable Chem. Eng.* **2020**, *8*, 16737–16745.
[65] W. Lan, J. Rencoret, F. Lu, S. D. Karlen, B. G. Smith, P. J. Harris, J. C. del Río, J. Ralph, *Plant J.* **2016**, *88*, 1046–1057.
[66] F. Lu, S. D. Karlen, M. Regner, H. Kim, S. A. Ralph, R.-C. Sun, K. Kuroda, M. A. Augustin, R. Mawson, H. Sabarez, T. Singh, G. Jimenez-Monteon, S. Zakaria, S. Hill, P. J. Harris, W. Boerjan, C. G. Wilkerson, S. D. Mansfield, J. Ralph, *BioEnergy Res.* **2015**, *8*, 934–952.
[67] J. M. Perez, C. Sener, S. Misra, G. E. Umana, J. Coplien, D. Haak, Y. Li, C. T. Maravelias, S. D. Karlen, J. Ralph, T. J. Donohue, D. R. Noguera, *Green Chem.* **2022**, *24*, 2795–2811.
[68] M. Regner, A. Bartuce, D. Padmakshan, J. Ralph, S. D. Karlen, *ChemSusChem* **2018**, *11*, 1600–1605.
[69] V. I. Timokhin, M. Regner, A. H. Motagamwala, C. Sener, S. D. Karlen, J. A. Dumesic, J. Ralph, *ACS Sustainable Chem. Eng.* **2020**, *8*, 17427–17438.
[70] O. Theander, E. Westerlund, in *Forage Cell Wall Struct. Dig.*, John Wiley & Sons, Ltd **1993**, pp. 83–104.
[71] T. Renders, G. Van den Bossche, T. Vangeel, K. Van Aelst, B. Sels, *Curr. Opin. Biotechnol.* **2019**, *56*, 193–201.
[72] J. H. Jang, A. R. C. Morais, M. Browning, D. G. Brandner, J. K. Kenny, L. M. Stanley, R. M. Happs, A. S. Kovvali, J. I. Cutler, Y. Román-Leshkov, J. R. Bielenberg, G. T. Beckham, *Green Chem.* **2023**, *25*, 3660–3670.
[73] S. Van den Bosch, W. Schutyser, R. Vanholme, T. Driessen, S.-F. Koelewijn, T. Renders, B. De Meester, W. J. J. Huijgen, W. Dehaen, C. M. Courtin, B. Lagrain, W. Boerjan, B. F. Sels, *Energy Environ. Sci.* **2015**, *8*, 1748–1763.
[74] C. Crestini, F. Melone, M. Sette, R. Saladino, *Biomacromolecules* **2011**, *12*, 3928–3935.
[75] M. R. Sturgeon, S. Kim, K. Lawrence, R. S. Paton, S. C. Chmely, M. Nimlos, T. D. Foust, G. T. Beckham, *ACS Sustainable Chem. Eng.* **2014**, *2*, 472–485.
[76] S. Omori, M. Aoyama, A. Sakakibara, *Holzforschung* **1998**, *52*, 391–397.
[77] J. Tsujino, H. Kawamoto, S. Saka, *Wood Sci. Technol.* **2003**, *37*, 299–307.
[78] Y. Pu, F. Hu, F. Huang, A. J. Ragauskas, *BioEnergy Res.* **2015**, *8*, 992–1003.
[79] T. Kishimoto, Y. Uraki, M. Ubukata, *J. Wood Chem. Technol.* **2008**, *28*, 97–105.
[80] T. Kishimoto, Y. Uraki, M. Ubukata, *Org. Biomol. Chem.* **2008**, *6*, 2982.
[81] "luterbachergroup - Overview," can be found under https://github.com/luterbachergroup **n.d**.
[82] S. V. Vassilev, D. Baxter, L. K. Andersen, C. G. Vassileva, *Fuel* **2010**, *89*, 913–933.

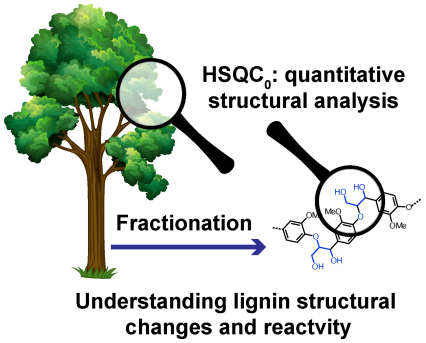
Manuscript received: March 4, 2024
Accepted manuscript online: May 13, 2024
Version of record online: ■■, ■■

Research Articles

Biomass

C. L. Bourmaud, S. Bertella, A. Bosch Rico, S. D. Karlen, J. Ralph, J. S. Luterbacher* **e202404442**

Quantification of Native Lignin Structural Features with Gel-Phase 2D-HSQC$_0$ Reveals Lignin Structural Changes During Extraction



Using a method featuring an extrapolated time-zero Heteronuclear Single-Quantum Coherence Zero (HSQC$_0$) sequence on whole cell walls samples, we can successfully quantify native lignin's structural features. Applying this tool to compare whole cell wall samples to extracted lignin reveals fundamental differences in the apparent linkage distribution within the polymer, which provides insight into lignin's fractionation mechanism.