BioMACROMOLECULES

Article

Characterization and elimination of undesirable protein residues in plant cell wall materials for enhancing lignin analysis by solution-state NMR

Hoon Kim, Dharshana Padmakshan, Yanding Li, Jorge Rencoret, Ronald D. Hatfield, and John Ralph

Biomacromolecules, **Just Accepted Manuscript** • DOI: 10.1021/acs.biomac.7b01223 • Publication Date (Web): 24 Oct 2017

ACS Publications

Biomacromolecules is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Characterization and elimination of undesirable protein residues in plant cell wall materials for enhancing lignin analysis by solution-state NMR

Hoon Kim,[*,†,‡] *Dharshana Padmakshan,*[‡] *Yanding Li,*[‡,§] *Jorge Rencoret,*[#] *Ronald D. Hatfield,*[*,¶] *and John Ralph*[*,†,‡,§]

[†]Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53706, U.S.A.

[‡]Department of Energy Great Lakes Bioenergy Research Center, Wisconsin Energy Institute, Madison, WI 53726, U.S.A.

[§]Department of Biological System Engineering, University of Wisconsin-Madison, Madison, WI 53726, U.S.A.

[#]Instituto de Recursos Naturales y Agrobiología de Sevilla (IRNAS), CSIC, 41012, Seville, Spain

[¶]USDA-ARS Dairy Forage Research Center, 1925 Linden Drive West, Madison, Wisconsin 53706, U.S.A.

Keywords: lignins, proteins, amino acids, plant cell wall, Klason lignin, 2D NMR

ABSTRACT

Protein polymers exist in every plant cell wall preparation, and they interfere with lignin characterization and quantification. Here we report the structural characterization of the residual protein peaks in 2D NMR spectra in corn cob and kenaf samples, and note that aromatic amino acids are ubiquitous and evident in spectra from various other plants and tissues. The aromatic correlations from amino acid residues were identified and assigned as phenylalanine and tyrosine. Phenylalanine's 3/5 correlation peak is superimposed on the peak from typical lignin *p*-hydroxyphenyl (H-unit) structures, causing an overestimation of the H units. Protein contamination also occurs when using cellulases to prepare enzyme lignins from virtually protein-free wood samples. We used a protease to remove the protein residues from the ball-milled cell walls, and we were able to reveal H-unit structures in lignins more clearly in the 2D NMR spectra, providing a better basis for their estimation.

Introduction

Plant cell walls are complex systems comprising a number of components, such as cellulose, hemicelluloses, lignin, structural proteins, enzymes, suberin, and other extraneous components such as water and waxes, depending on the species and tissues. Various of the components are also acylated by various acids, including acetate, ferulate, *p*-coumarate, and *p*-hydroxybenzoate. Compositional analysis and structural characterization is a crucial element in cell wall and lignin-related research, including in biomass and lignocellulosic utilization studies.

One of the most important analytical methods developed for the analysis of lignin and biomass structural characterization over the last several years is the 2D gel-NMR method for whole-cell-wall (WCW) profiling.[1,2] It not only provides detailed information on the original lignocellulosic structure without requiring an isolation/separation process, but also detects components that have been only minimally damaged during the solvent extraction process and required ball-milling steps. In particular, the aromatic region of a WCW 2D HSQC gel-NMR spectrum is virtually identical to that obtained from an isolated lignin (Figure 1). The S, G, and H units originating from typical monolignols, 4-hydroxycinnamyl alcohols (sinapyl, coniferyl, and *p*-coumaryl alcohols), are major components of many plants and tissues. Semi-quantitative evaluation using the 2D NMR method does not provide the absolute levels of the S, G, and H units in lignins, but delivers exceptional diagnostic information on the relative proportions of the aromatics and has the advantage of addressing the entire lignin fraction. Its profiling capability is not limited to the traditionally defined lignin units (S, G, and H); various lignin (or lignin-related) structures from different species can also be evaluated. The grasses (Poaceae family, and the commelinid monocots in particular) have abundant hydroxycinnamates, ferulate (*trans*-4-hydroxy-3-methoxycinnamate, FA) and *p*-coumarate (*trans*-4-hydroxycinnamate, *p*CA) associated with both lignin and hemicellulosic components.[3,4] Poplar, willow, and palm trees have *p*-hydroxybenzoates (*p*BA) associated with their lignins.[2,5,6] The newly discovered tricin [5,7-dihydroxy-2-(4-hydroxy-3,5-dimethoxyphenyl)-4H-chromen-4-one] in grasses is an unusual lignin component that comes from a non-monolignol biosynthetic pathway.[7,8] Most recently, hydroxystilbenes such as piceatannol and resveratrol have also been discovered in the lignins in palm endocarp tissues.[9]

Regarding lignin aromatic compositions, the H units generated from *p*-coumaryl alcohol are typically present at only low levels in many plant species,[10] but can be easily recognized and estimated using analytical assays including NMR methods, although they are frequently overestimated by methods that produce H monomers from components (such as *p*CA and *p*BA) that are not related to H-lignin units — we would like to make it clear that "H-unit" in this paper indicates the lignin component derived strictly from *p*-coumaryl alcohol, as it should always be. High levels of actual H units are naturally found in conifer compression wood and cell-wall middle lamella zones.[11,12] The *p*-coumaryl alcohol used to be considered a candidate for producing highly condensed lignin structures, as the H units have more potential radical-coupling sites, at positions of 3 and 5 of the aromatic ring, than G units from coniferyl alcohol or S units from sinapyl alcohol. However, the most abundant H units in lignins are uncondensed β-ether structures rather than carbon-carbon-bonded condensed structures.[12-14]

Because studies relating to H units are common in transgenic and mutant plant studies, it is important to accurately measure and identify authentic H units. Lange et al. showed 20-fold higher H units as a stress response in spruce cell wall cultures compared to spruce wood lignin.[15] Downregulation of the 4-coumarate 3-hydroxylase (C3H) in various plants significantly increased the H-unit levels over those in WT, 65% vs. ~1% in alfalfa,[10] and 21% vs. 0.2%,[16] and 31% vs. 0.3% in poplar.[17] The Arabidopsis *ref8* mutant is C3H-deficient and has essentially only H-lignin but has collapsed vessels and is particularly stunted;[18] the high-H lignin is retained (95% vs. ~1.9%) but agronomic characteristics are largely rescued by also downregulating mediator genes.[19] Hydroxycinnamoyl-CoA: shikimate hydroxycinnamoyl transferase (HCT) silencing in tobacco (*Nicotiana tabacum*) and Arabidopsis showed a decrease in syringyl units and an increase in H units, with lower lignin contents;[20] downregulating the same gene in radiata

pine, in which the lignin is normally dominated by G units, elevated H levels from ~0.4% to 31%.[21] Recently collaborators found a new enzyme, caffeoyl shikimate esterase (CSE), from *Arabidopsis thaliana*.[22] The CSE-deficient mutant resulted in an increase in lignin H-unit levels of over 30-fold, and a *Medicago truncatula* loss-of-function mutant was H-lignin rich as revealed by both analytical thioacidolysis and NMR,[23] demonstrating CSE's key role in the pathway in some plants.

Such compositional and structural studies could be easily thwarted and the accurate determination of the target components cannot be reliably undertaken if unexpected or unknown components appear in plant cell wall samples. We found previously unassigned (or unauthenticated) peaks around the typical H-unit 2/6 correlation at δ_C/δ_H 127.88/7.21, which is the correlation peak used for the H-unit quantification, in 2D HSQC NMR data of several plant samples including kenaf bast fiber and 1-month-old eucalypt wood.[1,24] We, as others have, inaccurately assigned the additional peaks as part of lignins' H units until we discovered high-volume peaks with unique patterns in the aromatic area of corn leaf and sheath (and sometimes in stem) tissues (Figures 1B–D). Later, the components responsible for these correlation peaks were recognized as being ubiquitous in many grasses, especially in leaf tissues and in young plants.

This finding led us to investigate the structures of the contentious peaks, from which a number of questions regarding the possibilities were raised: 1) Are the unknown peaks associated with the H unit or H-like polymer structures, as the chemical shifts are close to those from *p*-coumaryl alcohol and *p*-coumarate? 2) Do the peaks represent more than one component? 3) Are they in fact aromatic structures as the NMR chemical shifts suggest? 4) Are they different kinds of cell wall components, i.e., not lignin-related structures, or are they not part of the cell wall? If so,

how do we remove them to get reliable spectra of just the authentic cell wall, and an accurate H-unit estimate? 5) Can lignin quantification from the various analytical methods be affected?

In order to answer these questions, we examined whole cell walls (CWs) and lignins primarily by the 2D NMR method.[1,2] Corn cob material was principally used in this research because the unknown peak content is similar to that in corn sheath and other samples, and because the cobs have been considered to be useful biomass wastes in previous studies.[25] As has historically been the case, examining model compounds for possible components was crucial to revealing the nature of the structures. To confirm the structural and compositional identification results, we used common analytical assays for lignin such as the derivatization followed by reductive cleavage (DFRC) method,[26,27] nitrobenzene oxidation (NBO),[28,29] and Klason lignin. As with the related analytical thioacidolysis, the DFRC method can underestimate the H-unit level in the cell wall because it strictly cleaves the ether linkages of lignin, but the released monomers nevertheless confirm the presence of authentic H units in lignin to support 2D NMR results. The NBO method is known to overestimate H units and Klason lignin overestimates total lignins in herbaceous plants.[30,31] In this study, we reveal the specific structural information on the contentious aromatic peaks in 2D HSQC NMR data from corn cob and kenaf bast fiber. Additionally, evidence of overestimating H unit levels from such plants by NMR and NBO methods, and overestimation of Klason lignin, is demonstrated. These results will be beneficial for identifying “true H units” in lignin structural studies, especially for grasses and young plants.

Experimental Section

General

All chemicals, enzymes, and solvents were purchased from Sigma-Aldrich (Milwaukee, WI, USA) unless otherwise noted. Shimadzu GC-MS (GCMS-QP2010 Plus) on a fused-silica column [Restek, Rxi-5Sil MS Column, 30 m, 0.25 mm ID, 0.25 μm film thickness (d_f)] was used for nitrobenzene oxidation (NBO) and Derivatization Followed by Reductive Cleavage (DFRC) analyses. A Shimadzu UV spectrophotometer (UV-1800) was used to estimate acid-soluble lignins. A Thermo Scientific Sorvall Biofuge Primo centrifuge was used for solvent extraction of plant cell walls and in the preparation of milled wood lignins (MWL), enzyme lignins (EL), and cellulolytic enzyme lignins (CEL).

Plant materials and cell wall (CW) preparation

Dried plant materials, corn cob, aspen wood, and kenaf bast fiber, were pre-ground for 30 s in a Retsch MM400 mixer mill at 30 Hz, using stainless steel vessels (50 mL) containing stainless steel ball bearings (1 × 20 mm). The pre-ground material was extracted with distilled water (ultrasonication, 1 h, three times) and 80% ethanol (ultrasonication, 1 h, three times). The resulting isolated cell walls were then dried and ball-milled using a Fritsch planetary micro mill PULVERISETTE 7 (Germany) at 600 rpm, using zirconium dioxide (ZrO_2) vessels (20 mL) containing ZrO_2 ball bearings (10 mm × 10). Each sample (600 mg) was ground for 6 h 40 min (interval: 10 min., break: 5 min, repeated 27 times). The same scale of grinding was repeated until the desired amount of ball-milled sample was collected.

Isolation of milled wood lignin (MWL)

The MWL was isolated from extractive-free corn cob cell walls as stated in previous publications.[32-34] Finely ball-milled CWs (2 g) were extracted with dioxane/water (96:4, v/v) using 40 mL of solvent each time for 24 h with stirring, and the solution was centrifuged and the solvent-solubles collected by decantation. The extraction was repeated three times. The collected supernatant was combined and evaporated on a rotary evaporator at 50 °C. The crude MWL was obtained (263.6 mg, 13.2%), and then dissolved in acetic acid/water (9:1, v/v). The lignin was then precipitated in cold water and centrifuged to collect the MWL (38 mg, 1.9%), and was not further treated with any other solvents.

Isolation of enzyme lignin (EL) and cellulolytic enzyme lignin (CEL)

The extractive-free ball-milled corn cob cell walls were treated with a crude cellulases (Cellulysin, EC 3.2.1.4, activity >10000 units/g, Calbiochem) from *Trichoderma viride* to prepare the EL. The cell walls (2 g) were suspended in acetate buffer (pH 5), and 100 mg of Cellulysin was added. The reaction mixture was incubated and shaken on a rotary incubator shaker at 35 °C for 48 h. The insoluble residue was collected by centrifugation (8000 rpm, 30 min), and the enzyme treatment process was repeated three times. The collected lignin was sonicated and washed with deionized (DI) water three times after the enzyme treatments, and lyophilized to provide the so-called 'enzyme lignin' (EL, 306.1 mg, 15.3%). Aspen EL was prepared as above, and then extracted with dioxane:water (96:4, v/v) to produce its cellulolytic enzyme lignin (CEL).[33]

Polymerization of *p*-coumaryl alcohol and dimerization/polymerization of methyl-*p*-coumarate

Methyl *p*-coumarate and *p*-coumaryl alcohol were prepared from *p*-coumaric acid based on previous methods.[35] To prepare methyl *p*-coumarate, *p*-coumaric acid (4-hydroxycinnamic acid, 50 g, 0.31 mol) was dissolved in methanol (400 mL) and stirred at room temperature. Acetyl chloride (40 mL) was carefully added to the reaction solution, which was then stirred overnight at room temperature. The solvent (and HCl) was removed via rotary evaporation. A small amount of MeOH was added and evaporated several times to ensure the removal of HCl. Purplish-white crystals (41 g, 0.23 mol, 74%) formed after drying. ^{1}H NMR (acetone-d_6): δ 3.71 (3H, s, γ-OMe), 6.34 (1H, d, J = 16.0, β), 6.89 (2H, J = 8.76, m, 3 & 5), 7.53 (2H, J = 8.76, m, 2 & 6), 7.60 (1H, d, J = 16.0 Hz, α), 8.84 (1H, br s, Ph-OH); ^{13}C NMR (acetone-d_6): δ 51.47 (γ-OMe), 115.31 (8), 116.66 (3 & 5), 126.98 (1), 130.89 (2 & 6), 145.34 (7), 160.53 (4), 167.83 (9).

For *p*-coumaryl alcohol synthesis, the methyl *p*-coumarate (30 g, 0.17 mol) was dissolved in THF (500 mL) and then $LiAlH_4$ (12.9 g, 2 eq.) was added at room temperature. The reaction solution was stirred for 6 h, and checked by TLC for reaction completion. The reaction flask was placed in an ice-water bath and satd. aq. ammonium chloride (NH_4Cl) was slowly added. The product was extracted from the gray-colored suspension with EtOAc (3 × 200 mL) and the solution evaporated until half remained. Anhydrous magnesium sulfate ($MgSO_4$) was added to dry the solution, and the solution was filtered through a silica bed in a sintered glass filter. Solid products were obtained after evaporation, and white crystals (24.4 g, 0.16 mol, 96%) were obtained in EtOAc/petroleum ether after recrystallization. ^{1}H NMR (acetone-d_6): δ 4.20 (2H, bd, J = 5.2 Hz, γ), 6.21 (1H, dt, J = 15.8, 5.5 Hz, β), 6.51 (1H, bd, J = 15.9 Hz, α), 6.80 (2H, m, 3,

5), 7.31 (2H, m, 2, 6); ^{13}C NMR (acetone-d_6): δ 63.47 (γ), 116.19 (3, 5), 127.67 (β), 128.33 (2, 6), 129.73 (1), 130.29 (α), 157.78 (4).

For polymerization to the synthetic H-DHP (dehydrogenation polymer), the *p*-coumaryl alcohol (1.2 g, 7.99 mmol) was dissolved in acetone:water (1:10, v/v, 100 mL). Horseradish peroxidase (6.63 mg; EC 1.11.1.7, 181 purpurogallin units per mg solid, type II) was added directly to the reaction solution with stirring. Excess hydrogen peroxide (30%, 1.09 mL) was added at once into the reaction solution with stirring. The color immediately changed to dark red. The reaction solution was stirred for 15 h at room temperature. The crude products were extracted with EtOAc in order to remove low molecular weight products, and the insoluble residue was collected by filtration through a 0.8 μm nylon membrane filter. The polymer was washed with DI water to yield a dark brown material (243.4 mg, 20.3%).

Methyl *p*-coumarate was reacted under the same condition as for the *p*-coumaryl alcohol polymerization but was intentionally reacted for only 20 min in order to obtain low molecular mass products. The EtOAc-soluble fraction was collected and washed with satd. aq. NH_4Cl and water, then dried over anhydrous $MgSO_4$, filtered, and the solvent evaporated.

Acetylation of whole cell wall and model compounds.

For NMR analysis of acetylated cell wall (Ac-CW), the finely ball-milled cell walls were completely dissolved in DMSO/*N*-methylimidazole (2/1, v/v) followed by acetic anhydride addition.[36] The acetylated cell walls were precipitated into cold water and then collected by filtration through a 0.8 μm nylon membrane filter. Model compound acetylation was by reaction

in 1:1 pyridine/acetic anhydride. The reaction mixture was stirred at room temperature for 24 h. The product solution was dried on a rotatory evaporator at below 50 °C, co-evaporating with additional EtOH and water several times to remove the pyridine and acetic acid.

NMR analysis

The ball-milled CWs were directly prepared and transferred into NMR tubes for gel-state NMR experiments as previously described.[2,37] The NMR spectra were acquired on a Bruker Biospin (Billerica, MA) AVANCE-III 700 MHz spectrometer equipped with a 5 mm QCI $^{1}H/^{31}P/^{13}C/^{15}N$ cryoprobe with inverse geometry (proton coils closest to the sample). The central DMSO solvent peak was used as the internal reference (δ_C 39.5, δ_H 2.49 ppm). The $^{1}H–^{13}C$ correlation experiment was an adiabatic heteronuclear single-quantum coherence (HSQC) experiment (Bruker standard pulse sequence 'hsqcetgpsisp.2'; phase-sensitive gradient-edited-2D HSQC using adiabatic pulses for inversion and refocusing).[38] HSQC experiments were carried out using the following parameters: acquired from 11.5 to −0.5 ppm in F2 (^{1}H) with 1682 data points (acquisition time 100 ms), 215 to −5 ppm in F1 (^{13}C) with 620 increments (F1 acquisition time 8 ms) of 56 scans with a 500 ms interscan delay; the d_{24} delay was set to 0.86 ms (1/8J, J = 145 Hz). The total acquisition time was 6 h.

The isolated lignins and acetylated CW (Ac-CW) samples were measured by using the following parameters: acquired from 11.5 to −0.5 ppm in F2 (^{1}H) with 3366 data points (acquisition time 200 ms), 215 to −5 ppm in F1 (^{13}C) with 620 increments (F1 acquisition time 8 ms) of 32 scans with a 1 s interscan delay; the d_{24} delay was set to 0.86 ms (1/8J, J = 145 Hz). The total acquisition time for a sample was 6 h 45 min.

Processing in either case used typical matched Gaussian apodization (GB = 0.001, LB = −0.5) in F2 and squared cosine-bell and one level of linear prediction (32 coefficients) in F1. Volume-integration of contours in HSQC plots used Bruker's TopSpin 3.5pl6 (Mac version) software on spectra generated without the linear prediction. The integrals after the contour level adjustment for optimal peak separation were used for relative comparisons, but they do not represent absolute quantification as more mobile endgroups, in particular, are over-represented;[37] it is for this reason that data is presented on an S+G+H= 100% basis; *p*CA and tricin units are invariably terminal and can therefore be overestimated, and ferulate units may also be.

Proteinase K treatment and removing protein residues

The solvent-extracted ball-milled cell walls (1.5 g scale for each sample) in phosphate buffer (pH 8) were shaken at 37 °C for 2 days with Proteinase K (EC 3.4.21.64, activity 3–15 unit/mg, Sigma-Aldrich) from *Tritirachium album* to remove protein components.[39,40] Proteinase K has little substrate specificity. It degrades many proteins, but has some preference for cleavage at hydrophobic or aromatic amino acids.[41] The predominant site of cleavage is the peptide bond adjacent to the carboxyl group of aliphatic and aromatic amino acids.[39] The enzyme treatment was repeated three times and 100 mg of protease was used each time. The insoluble residues were collected after centrifugation, and the process was repeated three times. Corn cob, corn cob EL, aspen, and kenaf bast fiber cell walls were all treated in the same way producing residues in 61.6%, 77.0%, 70.1% and 63.4% yield, respectively (Table 4). Each supernatant was saved and examined by NMR after being lyophilized.

Nitrogen content and protein determination

Total nitrogen (N) contamination of cell wall residues and lignin samples was determined using an Elemental Combustion System (Costech Analytical Technologies Model 4010, Valencia, CA 91355). Samples (approximately 10 mg) were accurately weighed in tin combustion cups using a microbalance. After complete combustion, total N was measured as N_2 gas. Crude protein level in each sample was estimated based on the conversion factor of 6.25.[42,43] All analyses were run in duplicate (N = 2) to estimate residual protein.

Derivatization Followed by Reductive Cleavage (DFRC) method

The DFRC method was performed according to previous publications.[26,27,44] The extract-free cell wall samples (20 mg) were stirred in a 2 dram vial fitted with a polytetrafluoroethylene (PTFE) pressure-release cap, in acetyl bromide:acetic acid (1:4 v/v, 4 mL) at 50 °C for 2.5 h. The solvent was removed on a SpeedVac (Thermo Scientific SPD131DDA, 50 °C, 35 min, 1.0 torr, 35 torr/min). The reaction products were suspended in absolute ethanol (0.5 mL), and the ethanol was then removed on the SpeedVac (50 °C, 15 min, 6.0 torr, 35 torr/min). The residues were suspended in dioxane:acetic acid:water (5:4:1 v/v/v, 5 mL), and nano-powdered zinc (40-60 nm average particle size, Sigma-Aldrich, 150 mg) was added. The vial was sonicated to ensure that the solids were suspended, and then the solution was vigorously stirred in the dark at room temperature for 16–20 h. The mixture was quantitatively transferred with dichloromethane (DCM, 3 × 2 mL) to a separatory funnel charged with saturated ammonium chloride (10 mL) and deuterated internal standards for each compound, as described.[45] The reaction products were extracted with DCM (4 × 10 mL) and dried over $MgSO_4$. The DCM solution was filtered and

evaporated on a rotovap (water bath at <50 °C). The products were then acetylated for 16 h in the dark using a solution of pyridine and acetic anhydride (1:1 v/v, 5 mL) and dried on a rotovap. To remove most of the polysaccharide-derived products, the crude DFRC product was loaded onto an SPE cartridge (Supelco Supelclean LC-Si SPE tube, 3 mL, P/N: 505048) using DCM (3 × 1 mL). The products were eluted with hexanes:ethyl acetate (1:1, v/v, 8 mL), then the combined solvents were evaporated on a rotovap. The products were dissolved in DCM (GC-MS grade) and transferred into a GC-MS vial to make a final sample volume of 1 mL (or with a 300 µL insert, for a final sample volume of 200 µL when the sample amount was small). A 1 µL aliquot of the solution was injected at a split ratio of 20:1 into the GC-MS (Shimadzu GC-2010 Plus with mass spectrometer: GCMS-QP2010 Plus). The GC-MS analyses were performed under the following conditions: the oven was heated from 150 °C (held for 1 min), then ramped at 10 °C/min to 300 °C, and held for 29 min at the same temperature. The injector was set at 250 °C, and the transfer line was kept at 300 °C. Helium was used as the carrier gas at a linear velocity of 45 cm/sec. The DFRC compounds were identified on the basis of their mass spectra and relative retention times and quantified using synthesized poly-deuterated internal standards for each compound, as published previously.[45]

Nitrobenzene oxidation (NBO)

Nitrobenzene oxidation (NBO) was performed according to a previous paper.[46] The ball-milled CWs (60 mg) were treated with nitrobenzene (400 µL) and 2 M NaOH (7 mL) in steel Parr reactors (10 mL). The reactors were submerged in an oil bath at 170 ºC for 2 h with occasional shaking. The reactors were then cooled in ice-water. An internal standard, ethyl vanillin (EV, 3

μmol/mL in 2 M NaOH), was freshly prepared and added to each reaction mixture. The reaction mixtures were transferred to separatory funnels, and the excess nitrobenzene extracted with DCM (3 × 15 mL). The aqueous layer was then acidified to approximately pH 4 with 2 M HCl and extracted with DCM (3 × 50 mL) and diethyl ether (1 × 50 mL). The DCM and diethyl ether fractions were combined and washed with DI water (20 mL), dried over anhydrous $MgSO_4$ and filtered. The filtrate was collected in 100 mL pear-shaped flask and carefully dried on a rotary evaporator. For silylation, the reaction products were transferred into GC vials using pyridine (3 × 300 μL), and 100 μL of *N,O*-bis(trimethylsilyl)trifluoroacetamide (BSTFA) was added. The vials were heated at 50 ºC for 30 min. The derivatized products were analyzed with GC-MS (Shimadzu GC-2010 Plus with mass spectrometer: GCMS-QP2010 Plus) at a split ratio of 20:1. The initial column temperature was 150 ºC (held for 1 min) and raised at 10 ºC/min to 300 ºC (held for 29 min). The total running time was 45 min. The products, *p*-hydroxybenzaldehyde **Ha**, *p*-hydroxybenzoic acid **HA**, vanillin **Va**, vanillic acid **VA**, syringaldehyde **Sa**, and syringic acid **SA**, were quantified on the basis of the calibration curve of standards. Tyrosine (10.7 mg) was also examined under the same NBO treatment conditions and the product analyzed in the same way (for **Ha** and **HA** products).

Determination of the Klason lignin (KL) and acid-soluble lignin (ASL)

Klason lignin contents were measured by the NREL protocol (LAP: Laboratory Analytical Procedure).[47] Extractive-free ball-milled CWs (two replicates of 100 mg) were weighed into Pyrex bottles (Corning autoclave bottle, Max. 140 °C), and 4 mL of 72% sulfuric acid was added to each sample. The samples were kept at room temperature for 1 h and stirred every 10–15 min

using a glass rod. The samples were diluted to a 4% acid concentration by adding 109.9 mL DI (deionized) water. The diluted samples were autoclaved for 1.5 h at 121 °C. After cooling to room temperature, the hydrolysate was filtered through Gooch-type crucibles (fine porosity) to collect the Klason lignin. The first filtrates (approximately 50 mL) were also collected and set aside for the ASL (acid-soluble lignin) determination, and later for the sugar analysis. DI water was used to quantitatively transfer all remaining solids out of the serum bottle into the filtering crucible. The solids were rinsed with fresh DI water until the pH became neutral. The Klason lignins in the crucibles were kept at 105 °C and dried for 24 h to achieve constant weight, measured to the nearest 0.1 mg.

Acid-soluble lignins (ASLs) were measured from the saved filtrates. The NREL protocol describes the measurement using UV-vis spectroscopy (205 nm), and we used the extinction coefficient (absorptivity) value of 110 L/g-cm from the NREL LAP-004 to calculate the amount of ASL in the hydrolysate.[48]

Results and Discussion

Identification of aromatic protein residues, and structural elucidation of amino acid residues by NMR

The 2D HSQC whole-cell-wall (WCW) NMR experiments provide remarkably detailed information on lignin structures. Figure 1 shows significant H-unit levels (1.3%) in pine lignin (Figure 1A), whereas only a trace (0.4%) can be detected just above the noise level in aspen lignin (Figure 1B). Although low levels of H units are detectable among the thioacidolysis

method,[49-51] they are not always reported.[52] The 'unknown' peaks in the corn samples (Figures 1C–E and Figures 2A, 2F) were initially suspected, due to their chemical shifts, to be polymeric structures of *p*-coumaryl alcohol or *p*-coumarate (*p*CA). However, the chemical shifts and patterns from various models did not match the distinctive arrangement of the commonly observed four correlation peaks centered at approximately δ_C/δ_H 130.04/7.08, 129.06/7.25, 127.88/7.21, and 126.12/7.13 ppm. *p*-Coumaryl alcohol monomer 2/6 and 3/5 correlation peaks appear at δ_C/δ_H 127.30/7.26 and 115.31/6.79, but its polymer product peaks (in synthetic lignins) are spread widely at δ_C/δ_H 129.48–126.13/7.32–7.21 and 117.27–115.15/7.08–6.86 (Figure 2C). Moreover, there are extra peaks at around δ_C/δ_H 122.52/7.43 and 108.80/6.80 in the *in vitro* polymer (Figure 2C), but no such peaks can be found in the cell wall (CW) NMR spectra. As synthetic lignins usually represent all of the structures possible in lignins, just not in the same proportions,[53] it was difficult to imagine that the *p*-coumaryl alcohol polymers could represent the unique pattern of the four peaks from the CWs (Figures 1C–E, Figures 2A and 2F). The same rationale applied to the *p*-coumarate (*p*CA) components. The *p*CA monomer shows 2/6 and 3/5 correlation peaks at 130.14/7.55 and 115.77/6.87 ppm (Figure 2D). They are in agreement with the existing *p*CA peaks identified in the CW spectra (e.g., Figs 2A, 2F) but not with the unknown peaks. Although the NMR spectrum of the polymeric methyl *p*-coumarate product (Figure 2E) shows peaks at δ_C/δ_H 132.11–126.00/7.69–7.06 and 116.13–113.88/7.01–6.44, the intense and complicated peaks are hard reconcile with the clean shape of the new peaks, especially for the 2/6-position correlation peaks. However, there is a possibility that some of the radically coupled *p*CA structures might appear at one of the same positions (δ_C/δ_H 127.88/7.21).

After gaining unconvincing results from the radical polymerization of *p*-coumaryl alcohol or *p*CA, aromatic protein residues became strong candidates because of the species and tissue

specificity of the new components. Cell-wall structural proteins exist in all plants. The aromatic amino acid residues tyrosine, phenylalanine, and tryptophan, are commonly found in such proteins.[54-56] Moreover, recent studies have revealed that protein contamination is invariably found in lignin extraction or estimation,[57,58] raising issues for the contentions of many previous studies, as will be discussed below. Amino acids with aromatic side-chains were carefully examined by NMR. We found the aromatic peaks from tyrosine and phenylalanine correspond precisely to the other peaks in this region (Figure 2G–J). Those peaks are actually from a combination of the two different amino acid residues. The three peaks at δ_C/δ_H 129.06/7.25, 127.88/7.21, and 126.12/7.13 ppm belong to phenylalanine C/H 2/6, 3/5, and 4 correlations. The other peak at the δ_C/δ_H 130.28/7.11 ppm belongs to tyrosine 2/6, and the tyrosine 3/5 correlation at δ_C/δ_H 115.11/6.74 ppm is superimposed on a peak from normal G-units (belonging to a mixture of C/H 5 and 6 correlations). Tryptophan can only be found at a trace level in the CWs. Tyramine, a natural decarboxylation product from tyrosine known to be especially prevalent in *Solanacea* lignins,[59] was also examined, but it is unlikely a part of the unique peaks here because its 2/6 proton chemical shift was slightly off from the tyrosine peak (Figure 2J) and, although it can't strictly be ruled out as contributing to the highest carbon chemical shift peak in the corn cob samples (Figures 2A and 2F) here, evidence for tyramine ferulates, easily seen in thioacidolysis products, is not seen in non-solanaceous plants.

As discussed in many previous papers, the protein residues are difficult to remove.[57,58,60] Efforts to wash proteins from cell walls with extensive solvent extractions have not been successful (data not shown). In order to confirm whether the aromatic protein residues can be removed during the lignin isolation process, we prepared so-called milled wood lignin (MWL), also often termed Björkman lignin,[61] and enzyme lignin (EL) from the corn cob. The final yield of MWL

was 1.9% of the CW and 13.6% of the original Klason lignin content [10.2% of the total lignin – Klason (acid-insoluble) lignin, KL, plus acid-soluble lignin, ASL). After the MWL was precipitated and collected from cold water, and due to the low yield, the lignin was not further purified as has sometimes been done by treating with other solvents such as 1,2-dichloroethane/ethanol, or diethyl ether and precipitation into hexane.[32] The yield of EL was 15.3% of the CW and 109.3 % of the Klason lignin (82.3% of the total lignin, KL + ASL). The high yield of EL implies that there are additional non-lignin components in the prepared EL. Both the prepared corn cob lignins, EL (Figure 3B) and MWL (Figure 3C), were examined by NMR to ascertain the degree of protein residue removal. The lignins are significantly cleaned up, but still showed visible protein residues, implying that complete protein removal is difficult.

An interesting discovery was the cellulase contamination of aspen cellulolytic enzyme lignin (CEL) (Figure 3E). This is one line of evidence supporting the protein residue structural assignments from this research. Typically, aspen CW and lignin samples show only a trace of H units, and no protein peaks in this region of the spectrum (Figure 3D, Table 1) and often show no detectable H units;[52] however, protein residue peaks abruptly appeared after the cellulase treatment used in the EL and CEL (cellulolytic enzyme lignin)[33] preparation. We therefore realized that we had long overlooked the appearance of these unique peaks in EL. To confirm the enzyme contamination, we examined the cellulase and other enzymes commonly used for lignin and cell wall studies by NMR. All the enzymes we tested, cellulase (Cellulysin from Calbiochem), protease (proteinase K from Sigma-Aldrich), and peroxidase (peroxidase type II from Sigma-Aldrich), have aromatic structures but, as expected, have different relative levels of the various amino acid residues (Figure 4). Cellulysin contains all three aromatic amino acid residues, tyrosine (Tyr), phenylalanine (Phe), and tryptophan (Trp). Proteinase K has tyrosine

and phenylalanine residues. Peroxidase II contains only phenylalanine residues, which often also appears in DHPs when the DHP is prepared with peroxidase (data not shown). A possible explanation for the observation of amino acid residues in non-protein-rich lignins or DHPs is that the enzyme residues are trapped in, and/or contaminate, them during enzyme treatment of the cell wall. The possibility of a strong affinity between protein and lignin has already been studied,[62,63] and it is clear that enzyme contamination during sample preparation causes problems in estimating accurate H-unit content by various of the analytical methods, including NMR.

Protease treatment reveals that phenylalanine residues result in H-unit overestimation by NMR experiments

To verify that the amino acid residue peaks we identified are truly protein-derived and to also attempt to reveal the authentic H units in the NMR data, we treated the cell walls with proteinase K, as has been suggested previously for accurate lignin measurement.[31] In this study, corn cob cell walls were used as the primary samples. Kenaf and aspen, which only have traces of H units in their lignins, were also tested for the effectiveness of the protease treatment. A trial with pre-ground (but not finely ball-milled) cell walls was unsuccessful and most aromatic amino acid residues remained after the treatment. Next, cell walls were finely ball-milled for effective treatment under the same conditions as for the lignin sample preparation. As we contended, most of the aromatic amino acid residues were successfully removed from the cell wall after the protease treatment, and the real H-unit 2/6 correlation appeared in the corn cob and kenaf samples at the chemical shift at which one of the phenylalanine peaks appears (δ_C/δ_H 127.88/7.21 ppm, Phe-3/5) (Figures 3 A, B, and F). Proteinase K has tyrosine and phenylalanine residues in

its structure but, unlike in the aspen CEL, the contamination did not occur to any significant extent after the treatment of corn cobs. In the corn cob sample, an unexpected small peak appeared near the H-unit peak in the WCW spectrum after treatment with proteinase K (Figure 3A), but it remains an unknown component; this NMR peak does not appear in the EL (Figure 3B) or MWL (Figure 3C) samples.

From NMR integrals, the "H-unit" level in corn cobs not treated with proteinase K was 15.1% (on a total lignin S+G+H aromatics basis). After protease treatment, this was found to be considerably lower, at 6.5% (Table 1). The corn cob EL, which originally had 8.6% H units by its integral, was decreased to 4.5%; its MWL had 6.9% H units. The H-unit level in untreated kenaf was 4.0%. Most of the protein residues were removed and only a trace level of H units (0.4%) remained after treatment. We therefore consider that H-unit estimation by NMR is much more reasonable after the samples have undergone proteinase K treatment.

During the protease treatment, each supernatant was collected and lyophilized for NMR experiments. A substantial quantity of polysaccharides was extracted and found in the buffer solutions. We suspect this is the main cause of the somewhat unexpected mass loss of 30–38% after protease treatment. Unfortunately, the buffer solutions of kenaf and aspen contained a noticeable amount of water-soluble lignins, whereas the corn cob had only trace amounts of lignin based on the NMR data. The loss of lignin is apparent from the protease treatment process of the aspen CW sample, but we contend that protease treatment is not necessary for aspen because it contains insignificant levels of aromatic amino acid residues (or proteins) in general. The lignin loss result may indicate and support the existence of cell wall structural proteins and lignin-protein complexes.[56,63] To provide additional evidence for tyrosine residues, the corn cob cell wall was acetylated (to produce the Ac-CW) and its spectrum compared with that of the

acetylated tyrosine monomer. The 2/6 and 3/5 correlation peaks of free-phenolic tyrosine at δ_C/δ_H 130.28/7.11 and 115.11/6.74 ppm moved to δ_C/δ_H 129.92/7.29 and 121.33/7.04 ppm after the acetylation (Figure 3G). These tyrosine residues therefore exist in the CW preparation in free-phenolic form. Logically, the chemical shifts of phenylalanine residues were not changed after acetylation as they have no phenolic hydroxyl group to be acetylated. Even though tyrosine is more structurally related to the lignin H units, the real interference in the identification and estimation of actual H units by NMR was the presence of phenylalanine that, as noted above, has one overlapping correlation peak (Figure 2H and Figure 3).

The DFRC method identifies actual H units, whereas tyrosine partially contributes to high H-unit estimates by NBO

Degradative lignin analysis assays, DFRC (Derivatization followed by Reductive Cleavage) and NBO (nitrobenzene oxidation), were performed and the results were compared with the decreased H-unit values from 2D HSQC NMR. The DFRC method strictly cleaves β–*O*–4-linked structures and produces the (acetylated) monolignols from which the polymer is derived as its primary monomeric products.[26,27] Although the DFRC method may underestimate the total amounts of H, G and S units in the lignins, the existence and relative percentage of the H units can be determined by GC-MS or GC-FID. Corn cob showed 2.4% of the detectable total lignin content as H units before protease treatment and 2.6% after treatment (Figure 5A, Table 2). The values are slightly lower than typical thioacidolysis estimates from previous studies (2.1–4%).[64] The results of both methods support the low H-unit levels determined from the NMR data from corn samples that have been protease treated. Aspen (or poplar) is not normally reported to

release (or is reported only release extremely low levels of) H units by thioacidolysis,[52,65,66] but we did detect a low level by the DFRC method (and presumably could by thioacidolysis if sufficient sensitivity was used). Aspen CW has H-unit levels of 0.5% (protease-untreated) and 0.9% (protease-treated). Kenaf CW showed traces of H units (0.2%) for both untreated and protease-treated samples by the DFRC method and the results are consistent with previous results.[67] It is not possible to determine which values of the analytical methods are indeed closer to the actual H-unit contents of the lignins as all methods release only a fraction of the lignin as monomers; nevertheless, the range of values between 2D HSQC NMR semi-quantification (of EL, MWL and protease-treated CW) and the DFRC method (or thioacidolysis) seems reasonable. Most importantly, the presence of H units in plant cell walls was confirmed here by the DFRC method.

We also used the NBO method to measure monomers released from CWs. The results showed high levels of H monomers (100.55−107.15 μmol/g, Table 3), as expected based on previous results; corn cob H units from the NBO method were previously estimated at 31.9%.[68] Corn lignin has *p*CA (*p*-coumarate) structures that become a major contributor to the H monomers obtained during NBO method treatment.[30,69] As a result, much higher (and erroneous) apparent H-unit values are inevitable when compared to analytical assays that are diagnostic for lignin structures (such as DFRC and thioacidolysis). Besides the *p*CA component, we inferred that the tyrosine residues also contributed to the measure of H units. Tyrosine (10.7 and 10.6 mg for two duplicates) was subjected to NBO and showed about 10% of conversion to *p*-hydroxybenzaldehyde (9.1% and 11.6% respectively), a key NBO "H-lignin" monomer; however, there was no notable change in the H-unit values between the corn cob CW (31.6%) and the protease-treated sample (31.8%) based on our measurements (Figure 5B & Table 3).

Obviously, therefore, and as shown in a previous study in which the H-type structures of the NBO products are significantly decreased after alkaline treatment,[69] the high H-unit values in corn cob result primarily from the *p*CA rather than tyrosine residues. Aspen lignin also has *p*BA (*p*-hydroxybenzoate) structures that likewise produce H-monomers by NBO. As with the corn cob, there was no observable change between untreated (8.8%) and protease-treated (8.4%) aspen CW samples. Kenaf bast CW showed a relatively high level of H units (7.3%) by the NBO method that is thought to originate mainly from tyrosine residues because the kenaf lignin lacks *p*CA and *p*BA, unlike corn cob and aspen lignins, and obviously has a low-H lignin by NMR (Figure 3F). The H units were decreased by 27.4% after the protease treatment, but the treated CW still produced substantial amounts of H units (5.3%) by NBO. This result is in contrast to the DFRC result of low (or trace) levels of authentic H units after the elimination of proteins. As seen in Figure 3F, traces of protein residues can still be detected by NMR even though most of the aromatic amino acid residues were removed. Furthermore, reduction of NBO-determined H-unit levels after the protease treatment was rather insignificant, as 47% of proteins remained after the treatment (Table 4).

Proteins in cell walls are a major source of Klason lignin overestimation

To support the implication that the newly assigned aromatic peaks are indeed from protein residues in cell walls, we inspected the extent to which the proteinase K treatment eliminated protein residues (Table 4). The protein contents were calculated using the nitrogen-to-protein conversion factor (N × 6.25).[42,43] The protein content of corn cob was measured to be 4.26% in the previous paper,[70] and we measured it to be 2.60% here. The protein levels were clearly

decreased after the protease treatment (1.57%), but about 60% of the proteins remained in the corn cob cell wall.

The proteins in cell walls are not only a problem for the detection of H units but also for the estimation of total lignin content. Klason lignin is the most commonly accepted method for measuring total lignin, and has been recommended as the best method even in protein-containing forages.[71] However, during the processing with sulfuric acid, the protein residues can be condensed onto lignin, as was recognized in the attempt to provide an improved method for forages.[72] We measured Klason lignin and ASL (acid-soluble lignin) and cumulatively acquired the total lignin (KL + ASL) contents of our samples (Table 5). Corn cob CW lignin was measured as 14.6% (KL: 9.3%, ASL: 5.3%) before the protease treatment and 17.7% (KL: 14.4%, ASL: 3.3%) after treatment. The protein content is similar to that previously reported (17.2%).[73] Because of the initial removal of proteins, the protease-treated CW had a higher lignin content than untreated CW when the same amounts of cell walls were used for the determination. The corrected corn cob lignin value based on the protease treatment became considerably lower at 10.9% (74.7% of the original lignin content). Analogous results can be found in the aspen and kenaf samples. Some of the water-soluble lignin may be lost during protease treatment as noted above, but the corrected total lignin content (10.9%) in the corn cob may still be overestimated (due to residual protein).

It has been a century-long discussion that protein is one of the major impediments for the accurate determination of Klason lignin (and ASL) in herbaceous plants and grasses.[60,72,74] Rencoret et al.[58] measured corrected lignin content by estimating the nitrogen content of Klason lignin from Brewer's spent grain. However, true lignin content measurement can be more complicated. Bunzel et al.[57] rigorously examined the Klason lignin compositions of plant-based

food products including corn bran and concluded that only 2.3% real lignin exists in the cell walls even though they were initially characterized as having 9.9% Klason lignin. The estimated "final" lignin content was much lower than the simple corrected lignin content based on the amount of protein.

Conclusions

Through this study we have identified and characterized aromatic amino acid residues in 2D NMR data from plant cell wall samples, and provided clear evidence to support the characterization. The most important finding in this study is the protein residue peaks in NMR data, and the precise peak assignment to the various aromatic amino acid units allowing researchers to make these designations on their spectra. One of phenylalanine's peaks interferes with the identification and estimation of the lignin H units, and tyrosine residues contribute in part to the overestimation of H units in the nitrobenzene oxidation (NBO) method. These findings reaffirm that the NBO method cannot be relied on to accurately estimate H-unit levels in lignin, whereas DFRC and thioacidolysis, because of their discrete ether-cleaving chemistry and their release of monomers retaining the full lignin propyl sidechain, can provide an 'absolutely diagnostic' confirmation of compositional estimates to aid the NMR analysis, albeit still limited in that the measure is solely on the releasable monomers and not on the whole lignin. Proteins are abundant in certain plant species such as herbaceous plants and grasses, in various tissues such as leaves and sheaths, and particularly in young plants. Proteinase K treatment effectively removed most of the aromatic protein residues from samples. In this study, it was possible to estimate the realistic H-unit content by removing undesirable protein residues, and it was also a

viable method to enhance the lignin analysis by 2D NMR on protein-rich plant materials, as shown here with the corn cob and kenaf bast fiber cell walls. The evidence from this study lead us to firmly conclude that aromatic protein residues have been incorrectly assigned as lignin H units in 2D NMR spectra, and the contents of H units and of lignin itself in such samples and their sourced plants have been substantially overestimated in previous studies. Taken together, our results reaffirm the need to authenticate peak assignments in NMR data and to fully understand the nature of the cell wall and the extraneous components in a given sample.

A protease treatment step can alleviate some of the issues associated with ambiguously assigning NMR peaks, but it is not without limitations. Protease treatment cannot remove all protein from a sample, particularly from cell wall samples, and of course proteases are themselves proteins that can exacerbate aspects of the protein contamination problem. Some lignins were also found to be "lost" to the buffer solution, along with proteins and polysaccharides, during protease treatment.

Protein residues, and other components such as ash, fat, and cutin/suberin, remain the main factors in the overestimation of lignin by the Klason lignin analysis, which is currently the most common method for measuring total lignin. There are many uncertainties associated with the current measurement of lignin content, especially in protein-rich plants. Further research with more focus on developing suitable lignin estimation methods is therefore rather desperately required. Researchers should be able to utilize the data and the newly assigned peaks presented here to both improve the characterization of the plant material and to more accurately reflect the composition of the lignin, particularly with regard to the H-lignin component.

AUTHOR INFORMATION

Corresponding Author

Hoon Kim

Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53706, U.S.A. and Department of Energy Great Lakes Bioenergy Research Center, Wisconsin Energy Institute, Madison, WI 53726, U.S.A.

*E-mail: hoonkim@wisc.edu

ORCID: 0000-0001-7425-7464

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding Sources

This work was supported by grants from the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02-07ER64494).

ACKNOWLEDGMENT

We are grateful for Jane Marita for her various helpful discussions on this project. We also thank Steven Karlen for GC/MS experimental support and Kris Niemann (USDFRC) for assistance with nitrogen analysis.

ABBREVIATIONS

2D NMR, Two-dimensional nuclear magnetic resonance spectroscopy; H-unit, *p*-hydroxyphenyl; WCW, whole-cell-wall (e.g., for WCW NMR); HSQC, (^{1}H–^{13}C) heteronuclear single-quantum coherence; S, syringyl unit; G, guaiacyl unit; H, *p*-hydroxyphenyl unit; FA, ferulate (*trans*-4-hydroxy-3-methoxycinnamate); *p*CA, *p*-coumarate (*trans*-4-hydroxycinnamate); *p*BA, *p*-hydroxybenzoate; C3H, 4-coumarate 3-hydroxylase; HCT, hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl transferase; CSE, caffeoyl shikimate esterase; DFRC, derivatization followed by reductive cleavage; NBO, nitrobenzene oxidation; GC-MS, gas chromatography–mass spectrometry; GC-FID, gas chromatography–flame ionization detection; CW, cell wall; MWL, milled wood lignin or Björkman lignin; EL, enzyme lignin; CEL, cellulolytic enzyme lignin; DI, deionized (water); THF, tetrahydrofuran; $LiAlH_4$, lithium aluminum hydride; DHP, (synthetic) dehydrogenation polymer; DCM, dichloromethane; KL, Klason lignin (acid-insoluble lignin); ASL, acid-soluble lignin; Tyr, tyrosine; Phe, phenylalanine; Trp, tryptophan; PK, proteinase K.

REFERENCES

(1) Kim, H.; Ralph, J.; Akiyama, T. *Bioenergy Res.* **2008**, *1*, 56-66.

(2) Kim, H.; Ralph, J. *Org. Biomol. Chem.* **2010**, *8*, 576-591.

(3) Shimada, M.; Fukuzuka, T.; Higuchi, T. *TAPPI J.* **1971**, *54*, 72-78.

(4) Hatfield, R. D.; Marita, J. M. *Phytochem. Rev.* **2010**, *9*, 35-45.

(5) Lu, S.; Li, Q.; Wie, H.; Chang, M.-J.; Tanulaya-Anukit, S.; Kim, H.; Liu, J.; Song, J.; Sun, Y.-H.; Yuan, L.; Yeh, T.-F.; Ralph, J.; Sederoff, R. R.; Chiang, V. L. *Pro. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 10848-10853.

(6) Lu, F.; Karlen, S. D.; Regner, M.; Kim, H.; Ralph, S. A.; Sun, R.-c.; Kuroda, K.-i.; Augustin, M. A.; Mawson, R.; Sabarez, H.; Singh, T.; Jimenez-Monteon, G.; Hill, S.; Harris, P. J.; Boerjan, W.; Mansfield, S. D.; Ralph, J. *Bioenergy Res.* **2015**, *8*, 934-952.

(7) del Río, J. C.; Rencoret, J.; Prinsen, P.; Martínez, Á. T.; Ralph, J.; Gutiérrez, A. *J. Agric. Food Chem.* **2012**, *60*, 5922-5935.

(8) Lan, W.; Morreel, K.; Lu, F.; Rencoret, J.; del Rio, J. C.; Voorend, W.; Vermerris, W.; Boerjan, W.; Ralph, J. *Plant Physiol.* **2016**, *171*, 810-820.

(9) del Río, J. C.; Rencoret, J.; Gutiiérrez, A.; Kim, H.; Ralph, J. *Plant Physiol.* **2017**, *174*, 2072-2082.

(10) Ralph, J.; Akiyama, T.; Kim, H.; Lu, F.; Schatz, P. F.; Marita, J. M.; Ralph, S. A.; Reddy, M. S. S.; Chen, F.; Dixon, R. A. *J. Biol. Chem.* **2006**, *281*, 8843-8853.

(11) Fengel, D.; Wegener, G., Lignin. In *Wood: Chemistry, Ultrastructure, Reactions*, Walter de Gruyter: Berlin, New York, 1989; pp 132-181.

(12) Önnerud, H. *Holzforschung* **2003**, *57*, 377-384.

(13) Saito, K.; Fukushima, K. *J. Wood Sci.* **2005**, *51*, 246-251.

(14) Nakatsubo, F. *Wood Research* **1981**, *67*, 59-118.

(15) Lange, B. M.; Lapierre, C.; Sandermann, H., Jr. *Plant Physiol.* **1995**, *108*, 1277-87.

(16) Coleman, H. D.; Park, J.-Y.; Nair, R.; Chapple, C.; Mansfield, S. D. *Pro. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 4501-4506.

(17) Ralph, J.; Akiyama, T.; Coleman, H. D.; Mansfield, S. D. *Bioenergy Res.* **2012**, *5*, 1009-1019.

(18) Franke, R.; Humphreys, J. M.; Hemm, M. R.; Denault, J. W.; Ruegger, M. O.; Cusumano, J. C.; Chapple, C. *Plant J.* **2002**, *30*, 33-45.

(19) Bonawitz, N. D.; Kim, J. I.; Tobimatsu, Y.; Ciesielski, P. N.; Anderson, N. A.; Ximenes, E.; Maeda, J.; Ralph, J.; Donohoe, B. S.; Ladisch, M.; Chapple, C. *Nature* **2014**, *509*, 376-380.

(20) Hoffmann, L.; Besseau, S.; Geoffroy, P.; Ritzenthaler, C.; Meyer, D.; Lapierre, C.; Pollet, B.; Legrand, M. *Plant Cell* **2004**, *16*, 1446-1465.

(21) Wagner, A.; Ralph, J.; Akiyama, T.; Flint, H.; Phillips, L.; Torr, K. M.; Nanayakkara, B.; Te Kiri, L. *Pro. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 11856-11861.

(22) Vanholme, R.; Cesarino, I.; Rataj, K.; Xiao, Y.; Sundin, L.; Goeminne, G.; Kim, H.; Cross, J.; Morreel, K.; Araujo, P.; Welsh, L.; Haustraete, J.; McClellan, C.; Vanholme, B.; Ralph, J.; Simpson, G.; Halpin, C.; Boerjan, W. *Science* **2013**, *341*, 1103-1106.

(23) Ha, C. M.; Escamilla-Trevino, L.; Serrani-Yarce, J. C.; Kim, H.; Ralph, J.; Chen, F.; Dixon, R. A. *Plant J.* **2016**, *86*, 363-375.

(24) Rencoret, J.; Gutiérrez, A.; Nieto, L.; Jiménez-Barbero, J.; Faulds, C. B.; Kim, H.; Ralph, J.; Martínez, Á. T.; del Río, J. C. *Plant Physiol.* **2011**, *155*, 667-682.

(25) Li, J.; Jiang, Z.; Hu, L.; Hu, C. *ChemSusChem* **2014**, *7*, 2482–2488.

(26) Lu, F.; Ralph, J. *J. Agric. Food Chem.* **1997**, *45*, 2590-2592.

(27) Lu, F.; Ralph, J. *J. Agric. Food Chem.* **1997**, *45*, 4655-4660.

(28) Yamamura, M.; Hattori, T.; Suzuki, S.; Shibata, D.; Umezawa, T. *Plant Biotechnol-Nar* **2010**, *27*, 305-310.

(29) Pew, J. C. *J. Am. Chem. Soc.* **1955**, *77*, 2831-2833.

(30) Min, D. Y.; Jameel, H.; Chang, H. M.; Lucia, L.; Wanga, Z. G.; Jina, Y. C. *RSC Adv.* **2014**, *4*, 10845–10850.

(31) Hatfield, R.; Fukushima, R. S. *Crop Sci.* **2005**, *45*, 832-839.

(32) Björkman, A. *Sven. Papperstidn.* **1956**, *59*, 477-485.

(33) Chang, H.-M.; Cowling, E. B.; Brown, W.; Adler, E.; Miksche, G. *Holzforschung* **1975**, *29*, 153-159.

(34) Rencoret, J.; Marques, G.; Gutierrez, A.; Nieto, L.; Jimenez-Barbero, J.; Martinez, A. T.; del Rio, J. C. *Ind. Crops. Prod.* **2009**, *30*, 137-143.

(35) Quideau, S.; Ralph, J. *J. Agric. Food Chem.* **1992**, *40*, 1108-1110.

(36) Lu, F.; Ralph, J. *Plant J.* **2003**, *35*, 535-544.

(37) Mansfield, S. D.; Kim, H.; Lu, F.; Ralph, J. *Nat. Protoc.* **2012**, *7*, 1579-1589.

(38) Kupče, E.; Freeman, R. *J. Magn. Reson.* **2007**, *187*, 258-265.

(39) Ebeling, W.; Hennrich, N.; Klockow, M.; Metz, H.; Orth, H. D.; Lang, H. *Eur. J. Biochem.* **1974**, *47*, 91-97.

(40) Gross-Bellard, M.; Oudet, P.; Chambon, P. *Eur. J. Biochem.* **1973**, *36*, 32-38.

(41) Betzel, C.; PAL , G. P.; Saenger, W. *Eur. J. Biochem.* **1988**, *178*, 155-171.

(42) Jones, D. B. *U.S. Department of Agriculture Circular* **1931**, *183*, 1-22.

(43) Sluiter, J. B.; Ruiz, R. O.; Scarlata, C. J.; Sluiter, A. D.; Templeton, D. W. *J. Agric. Food Chem.* **2010**, *58*, 9043-9053.

(44) Karlen, S. D.; Zhang, C.; Peck, M. L.; Smith, R. A.; Padmakshan, D.; Helmich, K. E.; Free, H. C. A.; Lee, S.; Smith, B. G.; Lu, F.; Sedbrook, J. C.; Sibout, R.; Grabber, J. H.; Runge, T. M.; Mysore, K. S.; Harris, P. J.; Bartley, L. E.; Ralph, J. *Sci. Adv.* **2016**, *2*, e1600393: 1-9.

(45) Karlen, S. D.; Smith, R. A.; Kim, H.; Padmakshan, D.; Bartuce, A.; Mobley, J. K.; Free, H. C. A.; Smith, B. G.; Harris, P. J.; Ralph, J. *Plant Physiol.* **2017**, in press (accepted 9/6/2017; DOI: 10.1104/pp.17.01172).

(46) Tamai, A.; Goto, H.; Akiyama, T.; Matsumoto, Y. *Holzforschung* **2015**, *69*, 951-958.

(47) Sluiter, A.; Hames, B.; Ruiz, R.; Scarlata, C.; Sluiter, J.; Templeton, D.; Crocker, D. *Determination of structural carbohydrates and lignin in biomass*; National Renewable Energy Laboratory: Golden, CO, 2012.

(48) Ehrman, T. *Determination of acid-soluble lignin in biomass*; National Renewable Energy Laboratory: Golden, CO, 1996.

(49) del Río, J. C.; Rencoret, J.; Marques, G.; Li, J. B.; Gellerstedt, G.; Jimenez-Barbero, J.; Martinez, A. T.; Gutierrez, A. *J. Agric. Food Chem.* **2009**, *57*, 10271-10281.

(50) del Rio, J. C.; Rencoret, J.; Gutierrez, A.; Nieto, L.; Jimenez-Barbero, J.; Martinez, A. T. *J. Agric. Food Chem.* **2011**, *59*, 11088-11099.

(51) Rencoret, J.; Ralph, J.; Marques, G.; Gutiérrez, A.; Martínez, Á. T.; del Rio, J. C. *J. Agric. Food Chem.* **2013**, *61*, 2434-2445.

(52) Lapierre, C.; Pollet, B.; Rolando, C. *Res. Chem. Intermed.* **1995**, *21*, 397-412.

(53) Ralph, J.; Landucci, L. L., NMR of lignins. In *Lignin and Lignans; Advances in Chemistry*, Heitner, C.; Dimmel, D. R.; Schmidt, J. A., Eds. CRC Press (Taylor & Francis Group): Boca Raton, FL, 2010; pp 137-234.

(54) Cassab, G. I. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1998**, *49*, 281–309.

(55) Laidlaw, R. A.; Smith, G. A. *Holzforschung* **1965**, *19*, 129–134.

(56) Bao, W.; O'Malley, D. M.; Sederoff, R. R. *Pro. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 6604-6608.

(57) Bunzel, M.; Schussler, A.; Saha, G. T. *J. Agric. Food Chem.* **2011**, *59*, 12506-12513.

(58) Rencoret, J.; Prinsen, P.; Gutierrez, A.; Martinez, A. T.; del Río, J. C. *J. Agric. Food Chem.* **2015**, *63*, 603-613.

(59) Ralph, J. *Phytochem. Rev.* **2010**, *9*, 65-83.

(60) Lai, Y. Z.; Sarkanen, K. V., Isolation and structural studies. In *Lignins, Occurrence, Formation, Structure and Reactions*, Sarkanen, K. V.; Ludwig, C. H., Eds. Wiley-Interscience: New York, 1971; Vol. 1, pp 165-240.

(61) Björkman, A. *Nature* **1954**, *174*, 1057-1058.

(62) van Sumere, C. F.; Albrecht, J.; Dedonder, A.; de Pooter, H.; Pe, I., Plant proteins and phenolics. In *The Chemistry and Biochemistry of Plant Proteins*, Harborne, J. B.; van Sumere, C. F., Eds. Academic Press: New York, 1975; p 211.

(63) Whitmore, F. W. *Plant Sci. Lett.* **1978**, *13*, 241-245.

(64) Lapierre, C., Determining lignin structure by chemical degradations. In *Lignin and lignans : Advances in chemistry*, Heitner, C.; Dimmel, D. R.; Schmidt, J. A., Eds. CRC Press, Taylor & Francis Group: New York, 2010; pp 11-48.

(65) Cai, Y.; Zhang, K.; Kim, H.; Hou, G.; Yang, H.; Zhang, X.; Feng, H.; Miller, L.; Ralph, J.; Liu, C.-J. *Nat. Commun.* **2016**, *7*, 11989: 1-14.

(66) Lapierre, C.; Jouin, D.; Monties, B. *Phytochemistry* **1989**, *28*, 1401-1403.

(67) Lu, F.; Ralph, J. *J. Agric. Food Chem.* **1998**, *46*, 547-552.

(68) Shi, Z. *Journal of Northeast Forestry University* **1991**, *2*, 98-107.

(69) Oliveira, L.; Evtuguin, D. V.; Cordeiro, N.; Silvestre, A. J.; Silva, A. M.; Torres, I. C. *J. Agric. Food Chem.* **2006**, *54*, 2598-2605.

(70) Pointner, M.; Kuttner, P.; Obrlik, T.; Jäger, A.; Kahr, H. *Agronomy Research* **2014**, *12*, 391–396.

(71) Hatfield, R. D.; Jung, H. G.; Ralph, J.; Buxton, D. R.; Weimer, P. J. *J. Sci. Food and Agric.* **1994**, *65*, 51-58.

(72) Van Soest, P. J. *J. Anim. Sci.* **1967**, *26*, 119-128.

(73) Jiang, Z.; He, T.; Li, J.; Hu, C. *Green Chem.* **2014**, *16*, 4257–4265.

(74) Christiernin, M.; Ohlsson, A. B.; Berglund, T.; Henriksson, G. *Plant Physiol. Biochem.* **2005**, *43*, 777-785.

TABLES

Table 1. Lignin and associated CW aromatics from 2D HSQC NMR

	H	G	S	*p*BA	*p*CA	FA	Tricin	S/G	H/(S+G)
Corn cob-CW	15.1	39.1	45.8	-	121.1	88.4	10.7	1.2	0.178
Corn cob-CW-PK	6.5	35.4	58.1	-	113.0	51.2	13.2	1.6	0.070

Corn cob-EL	8.6	44.8	46.6	-	172.0	32.7	14.7	1.0	0.094
Corn cob-EL-PK	4.5	47.0	48.5	-	169.6	27.2	11.7	1.0	0.047
Corn cob-MWL	6.9	45.4	47.7	-	123.8	24.1	17.5	1.0	0.074
Kenaf-CW	4.0	14.7	81.3	-	-	-	-	5.5	0.042
Kenaf-CW-PK	0.4	9.3	90.3	-	-	-	-	9.7	0.004
Aspen-CW	0.4	36.3	63.4	5.7	-	-	-	1.8	0.004
Aspen-CW-PK	0.3	29.9	69.8	4.3	-	-	-	2.3	0.003

Relative volume-integral percentages (%), except S/G and H/(S+G), of H, G and S lignin units and other aromatics, *p*BA, *p*CA, FA, and tricin, from whole-cell-wall (WCW) 2D HSQC gel-NMR experiments. The *p*BA, *p*CA, FA, and tricin components were estimated based on the total lignin values (H+G+S). CW, (whole) cell wall; EL, enzyme lignin; MWL, milled wood lignin (or Björkman lignin); PK, treated with proteinase K.

Table 2. DFRC analysis of control and protease-treated cell walls

			Without SA-*p*CA & SA-*p*BA		With SA-*p*CA & SA-*p*BA	
	H	G	S	Total (H+G+S)	S	Total (H+G+S)
Corn cob	1.41 ± 0.03	36.29 ± 1.51	8.98 ± 0.55	46.67 ± 0.99	20.50 ± 1.35	58.20 ± 0.19
Corn cob-PK	1.66 ± 0.13	35.83 ± 2.52	10.75 ± 1.03	48.24 ± 3.68	25.58 ± 1.81	63.07 ± 4.46

Kenaf	0.3 ± 0.00	44.30 ± 1.67	106.97 ± 0.78	151.56 ± 0.89	106.97 ± 0.78	151.56 ± 0.89
Kenaf-PK	0.3 ± 0.06	48.64 ± 2.73	117.17 ± 2.60	166.11 ± 5.39	117.17 ± 2.60	166.11 ± 5.39
Aspen	1.23 ± 0.00	102.91 ± 0.74	135.47 ± 0.02	239.61 ± 0.76	151.82 ± 0.70	255.95 ± 1.44
Aspen-PK	2.51 ± 0.01	110.34 ± 2.69	140.36 ± 5.67	253.21 ± 8.34	158.51 ± 6.25	271.36 ± 8.95

Released DFRC monomers (H, G, and S) from extractives-free cell walls from control and protease-treated samples. The data represent the means (μmol/g) ± SEM (standard error of the mean) from technical duplicates (N=2). "Without SA-*p*CA & SA-*p*BA" represent the G and S values *without* including the conjugate structures, SA-*p*CA (corn cob) and SA-*p*BA (aspen). "With SA-*p*CA & SA-*p*BA" represents the values *including* the conjugates. PK: treated with proteinase K.

Table 3. NBO analysis of control and protease-treated cell walls

	H	G	S	Total (H+G+S)
Corn cob	100.55 ± 1.28	107.98 ± 0.94	109.22 ± 0.68	317.75 ± 2.90
Corn cob-PK	107.15 ± 8.35	89.91 ± 2.56	139.53 ± 8.57	336.59 ± 2.34

Kenaf	16.23 ± 1.36	29.86 ± 1.19	175.16 ± 1.36	221.26 ± 1.54
Kenaf-PK	13.83 ± 1.09	37.00 ± 1.18	210.94 ± 5.19	261.77 ± 2.92
Aspen	43.48 ± 1.94	137.95 ± 5.78	316.15 ± 8.33	497.58 ± 12.17
Aspen-PK	45.70 ± 1.08	151.01 ± 1.74	349.09 ± 5.13	545.81 ± 7.95

Yield (μmol/g) of released NBO monomers (designated as H, G, and S to conventionally indicate their origin, but actually representing total: H, *p*-hydroxybenzaldehyde + *p*-hydroxybenzoic acid; G, vanillin + vanillic acid; S, syringaldehyde + syringic acid) from extractive-free cell walls of control and protease-treated samples. The data represent the means ± SEM from technical duplicates (N=2). H-unit values are significantly higher compared to results in Table 2 from the DFRC method. Corn cob samples in particular, but also Aspen, have extremely high H-unit values. PK: treated with proteinase K.

Table 4. Proteinase K treatment yields and protein estimation of control and treated cell walls

	Relative yields following protease treatment	Protein amounts
Corn cob	100	2.60 ± 0.03
Corn cob-PK	61.6	1.57 ± 0.03 (**60.4**)

Corn cob-EL	100	9.16 ± 0.01
Corn cob-EL-PK	77.0	5.42 ± 0.02 (**59.2**)
Kenaf	100	3.19 ± 0.00
Kenaf-PK	63.4	1.50 ± 0.00 (**47.0**)
Aspen	100	0.58 ± 0.03
Aspen-PK	70.1	1.51 ± 0.01 (**260.3**)

Relative percentages (%) of the obtained cell walls after Proteinase K treatment. The protein percentages were estimated based on nitrogen determination (N × 6.25). The data represent the means ± SEM from technical duplicates (N=2). The bolded values in the parentheses in the protein amounts column indicate the percentages of obtained proteins after proteinase K treatment compare to the original protein amounts in the cell walls. PK: treated with proteinase K.

Table 5. Klason lignin (KL) and acid-soluble lignin (ASL) from control and protease-treated cell walls

	KL	ASL	Total Lignin (KL + ASL)

	Uncorrected	**Corrected**	Uncorrected	**Corrected**	Uncorrected	**Corrected**
Corn cob-CW	9.3 ± 0.2	(9.3 ± 0.2)	5.3 ± 0.1	(5.3 ± 0.1)	14.6 ± 0.3	(14.6 ± 0.3)
Corn cob-CW-PK	14.4 ± 0.5	8.9 ± 0.3	3.3 ± 0.0	2.1 ± 0.0	17.7 ± 0.5	10.9 ± 0.3 (**74.7**)
Kenaf-CW	5.9 ± 0.2	(5.9 ± 0.2)	4.2 ± 0.0	(4.2 ± 0.0)	10.1 ± 0.2	(10.1 ± 0.2)
Kenaf-CW-PK	6.7 ± 0.4	4.3 ± 0.2	3.8 ± 0.0	2.4 ± 0.0	10.5 ± 0.4	6.7 ± 0.2 (**66.3**)
Aspen-CW	14.0 ± 0.1	(14.0 ± 0.1)	4.6 ± 0.1	(4.6 ± 0.1)	18.6 ± 0.1	(18.6 ± 0.1)
Aspen-CW-PK	16.5 ± 0.1	11.6 ± 0.1	4.5 ± 0.0	3.1 ± 0.0	20.9 ± 0.1	14.7 ± 0.1 (**79.0**)

Relative average percentages (%) of obtained Klason lignin (KL) and acid-soluble lignin (ASL) from extractives-free cell walls of control and protease-treated samples. The data represent the means (%) ± SEM from technical duplicates (N=2). Determinations used the same weight (~100 mg) for both untreated and treated samples. The corrected KLs and ASLs of the protease-treated samples were calculated based on yields following enzyme treatments. The bolded values in parentheses in the total lignin column indicate the percentage of obtained lignins from the original lignins after proteinase K treatment. PK: treated with proteinase K.

FIGURES

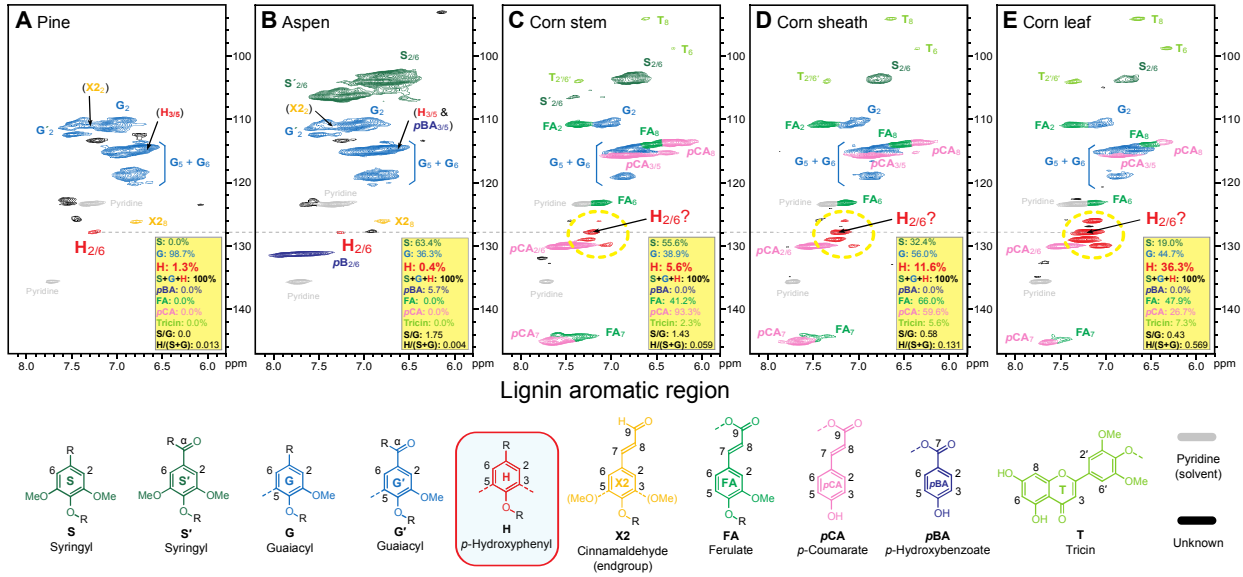

Figure 1. Lignin aromatic regions of 2D HSQC WCW NMR spectra (DMSO-d_6:pyridine-d_5, 4:1, v/v) of ball-milled cell walls (CWs) from pine, aspen, corn stem, corn sheath, and corn leaf. Pine **A** and aspen **B** show typical H units along with G and S units. Corn stem **C**, sheath **D**, and leaf **E** show the contentious peaks (yellow circles) near the H units and *p*CA peaks. The 2/6 position of H units in pine (**A**) and aspen (**B**) appears at δ_C/δ_H 127.88/7.21, and one of the contentious peaks in the corn samples (**C**, **D**, and **E**) shares the same chemical shifts. Much higher H-unit estimation is incurred for corn samples due to 'contamination' of this peak (see text).

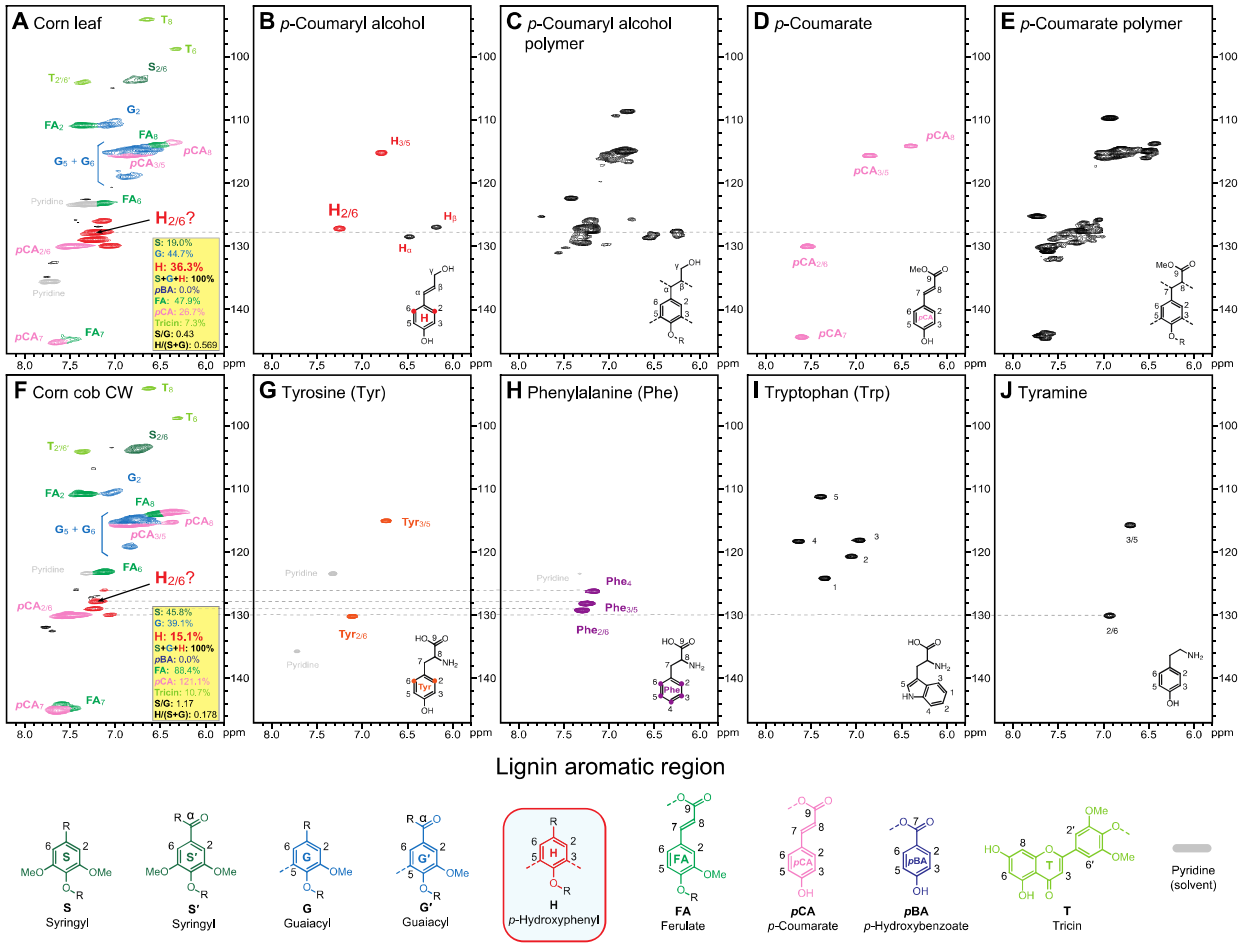

Figure 2. Corn cell wall aromatic regions of 2D HSQC NMR spectra (DMSO-d_6:pyridine-d_5, 4:1, v/v) and lignin and amino acid models for the protein residue peak assignment. The unknown aromatic peaks from corn leaf **A** and corn cob **F** were compared with lignin and amino acid models: *p*-coumaryl alcohol **B**, *p*-coumarate **D**, and the corresponding polymers **C** & **E** were not confirmed as contributing to the contentious peaks, but tyrosine **G** and phenylalanine **H** were confirmed as the peaks in corn samples; tryptophan and tyramine were not evident.

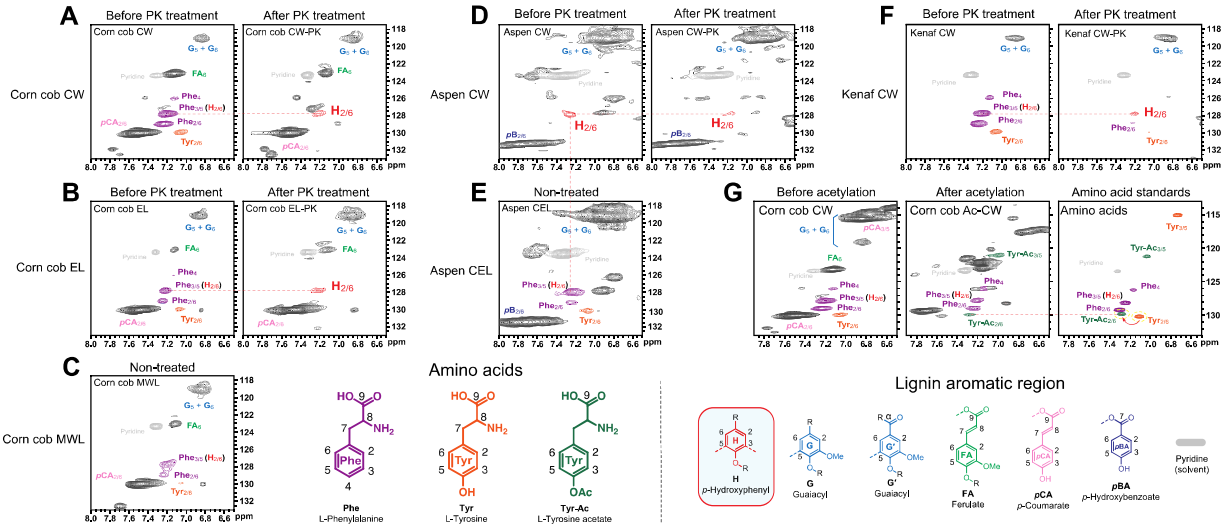

Figure 3. Partial aromatic regions of 2D HSQC NMR spectra (DMSO-d_6:pyridine-d_5, 4:1, v/v) of ball-milled whole cell walls or lignins, and the protease-treated cell walls or lignins. **A**, corn cob CW; **B**, corn cob EL; **C**, corn cob MWL; **D**, aspen CW; **E**, aspen CEL; **F**, kenaf CW; **G**, acetylated corn cob CW (Ac-CW) and amino acids. After protease treatment, most of the aromatic amino acids had been removed, revealing the H units at δ_C/δ_H 127.88/7.21. Prepared ELs showed protein residue contamination after cellulase enzyme treatment (**B**, left spectrum), and produced a clean H-unit peak in the NMR data after treatment with protease (**B**, right spectrum). The MWL had protein residues remaining after the preparation (**C**). Acetylation of corn cob cell wall provided supporting evidence for the NMR peak assignments for tyrosine and phenylalanine residues (**G**).

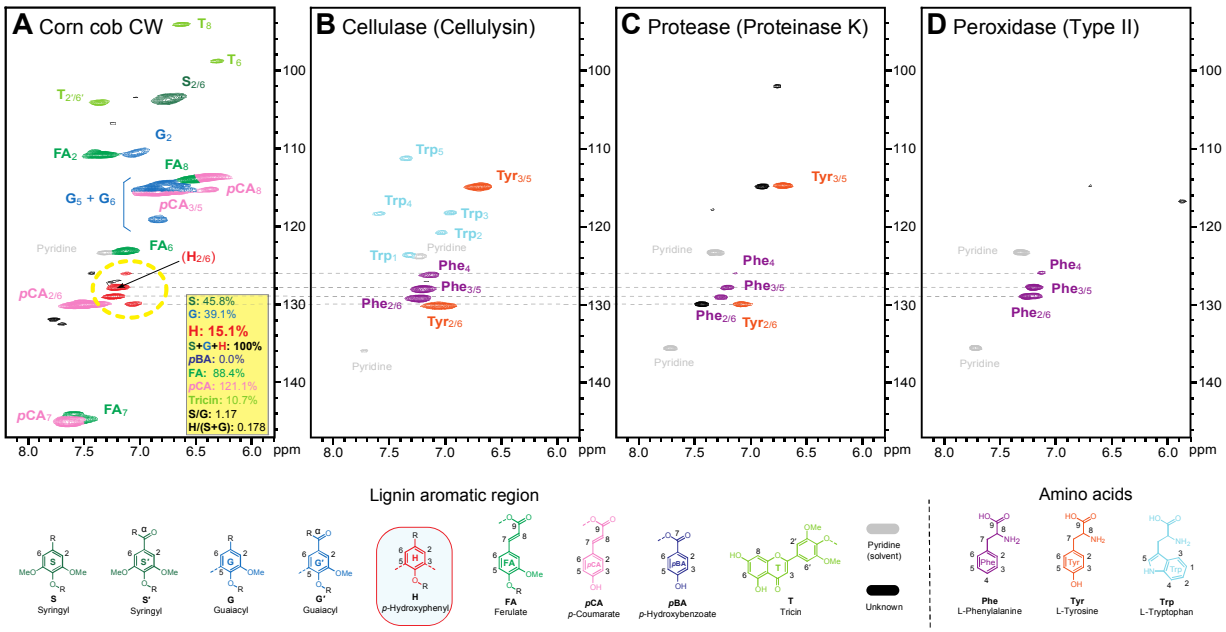

Figure 4. Aromatic regions of 2D HSQC NMR spectra (DMSO-d_6:pyridine-d_5, 4:1, v/v) of corn cob cell wall and commonly used enzymes in lignin-related studies. **A**, corn cob CW has the newly assigned protein residue peaks; **B**, cellulase (Cellulysin) has all three aromatic amino acid residues: phenylalanine (Phe), tyrosine (Tyr), and tryptophan (Trp); **C**, protease (proteinase K) has two aromatic amino acid residues: phenylalanine and tyrosine; **D**, peroxidase (type II) has only phenylalanine.

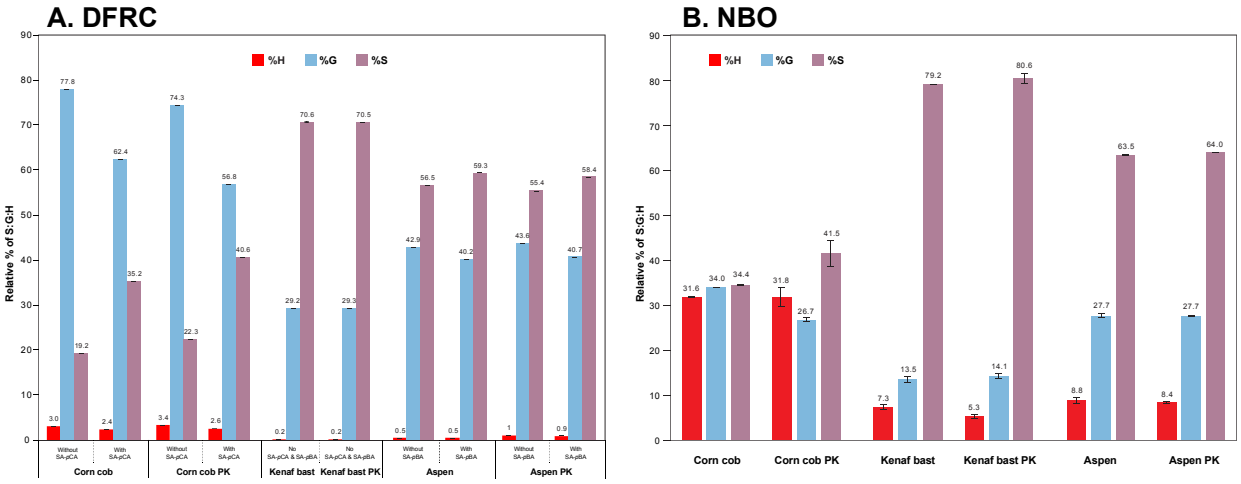

Figure 5. DFRC and NBO analysis of control cell walls and protease-treated cell walls. **A**, Relative average percentages (%) of released DFRC monomers (H, G, and S); and **B**, NBO monomers from extractives-free cell walls of control and protease-treated samples. The data represent the means ± SEM from technical duplicates (N=2). "Without SA-*p*CA & SA-*p*BA" represent the G and S values *without* including the conjugate structures, SA-*p*CA (corn cob) and SA-*p*BA (aspen). "With SA-*p*CA & SA-*p*BA" represent the G and S values *including* the conjugates. PK: proteinase-K-treated. Corn cob shows noticeably elevated S values when *p*CA and *p*BA conjugates are included, whereas the H-unit values were slightly lower. The NBO method released ~100-fold more H monomers from corn cobs compared to the DFRC method.

Table of Contents (TOC)

Protein amino acid residues and *p*-hydroxyphenyl (H) units in cell wall lignins have been characterized by solution-state 2D gel-NMR.