STRUCTURAL CHARACTERIZATION OF WHEAT STRAW LIGNIN. EVIDENCE FOR A NOVEL MONOMER IN GRASSES

José C. del Río,[a] Jorge Rencoret,[a] Pepijn Prinsen,[a] Ángel T. Martínez,[b] Ana Gutiérrez,[a] and John Ralph[c]

[a]*IRNAS-CSIC, Seville, Spain;* [b]*CIB-CSIC, Madrid, Spain;* [c]*Univ. Wisconsin, Madison, USA*

ABSTRACT

The structure of the lignin in wheat straw has been investigated by 2D-NMR and DFRC. It is a *p*-hydroxyphenyl-guaiacyl-syringyl lignin (with a H:G:S ratio of 6:64:30) associated with *p*-coumarates and ferulates. The main substructures present are β–*O*–4´-ethers (~75%), followed by phenylcoumarans (~11%), with lower amounts of other units. The lignin is partially acylated (~10%) at the γ-carbon, predominantly with acetates that preferentially acylate guaiacyl (12%) rather than syringyl (1%) units. *p*-Coumarate esters were barely detectable (<1%) on monomer conjugates released by selectively cleaving β-ethers in DFRC, indicating that they might be preferentially involved in condensed or terminal structures. A major new finding is that the flavone tricin is apparently incorporated into the lignins.

Keywords: Wheat straw; Milled Wood lignin; HSQC; *p*-Coumarate; Tricin

INTRODUCTION

Nowadays, increasing attention is being paid to the use of lignocellulosic biomass as a renewable feedstock for the production of second generation bioethanol [1]. Common lignocellulosic feedstocks considered for biofuel production include woods (e.g., poplar or eucalyptus), perennial energy crops (e.g., switchgrass or *Miscanthus* species), and agricultural wastes (e.g., corn stover or cereal straws). Among them, wheat straw has the greatest potential of all agricultural residues because of its wide availability and low cost, with a world production of wheat estimated to be around 680 million tons in 2011 [1]. Wheat straw contains 35–45% cellulose, 20–30% hemicelluloses, and around 15% lignin, which makes it an attractive feedstock to be converted to ethanol and other value-added products. The conversion of lignocellulosic biomass to bioethanol involves saccharification of carbohydrates to fermentable reducing sugars via hydrolysis and then fermentation of these free sugars to ethanol. However, the presence of lignin limits the accessibility of enzymes to cellulose, thus reducing the efficiency of the hydrolysis. Pretreatment of lignocellulosic materials to remove or modify the lignin is therefore needed to enhance the hydrolysis of carbohydrates. The knowledge of the structure of the lignin polymer is important to develop appropriate pretreatment methods for lignin modification and/or removal.

The composition and structure of the lignin in wheat straw has been a matter of study for many years [2,3]. In this paper, a more in-depth and complete characterization of the lignin of wheat straw has been performed by 2D-NMR and DFRC. The knowledge of the composition and structure of wheat straw lignin will help to maximize the exploitation of this important agricultural waste as a feedstock for biofuels and other biorefinery products.

MATERIAL AND METHODS

Wheat straw (*Triticum durum* var. Carioca) was air-dried and the dried samples were milled using a knife mill and successively extracted with acetone in a Soxhlet apparatus for 8 h and hot water. 'Milled-Wood Lignin' (MWL) from wheat straw was obtained by dioxane extraction according to the classical method, from extractive-free wheat straw. The experimental procedure has been explained in detail in previous papers [4,5].

For NMR analysis, **a**round 40 mg of MWL were dissolved in 0.75 mL of DMSO-d_6. NMR spectra were recorded at 25 ºC on a Bruker AVANCE III 500 MHz instrument equipped with cryoprobe. HSQC experiments used Bruker's 'hsqcetgpsisp2.2' pulse program with spectral widths of 5000 Hz (from 10 to 0 ppm) and 20,843 Hz (from 165 to 0 ppm) for the ^{1}H- and ^{13}C-dimensions. The number of collected complex points was 2048 for the ^{1}H-dimension with a recycle delay of 1.5 s. The number of transients was 64, and 256 time increments were always recorded in the ^{13}C-dimension. The $^{1}J_{CH}$ used was 145 Hz. Processing used typical matched Gaussian apodization in the ^{1}H dimension and squared cosine-bell apodization in the ^{13}C dimension. Prior to Fourier transformation, the data matrixes were zero-filled up to 1024 points in the ^{13}C-dimension. The central solvent peak was used as an internal reference (δ_C 39.5; δ_H 2.49). Long range *J*-coupling evolution times of 66 and 80 ms were used in different HMBC acquisition experiments. HSQC correlation peaks were assigned by comparing with the literature [4–8]. A semiquantitative analysis of the volume integrals of the HSQC correlation peaks was performed using Bruker's Topspin 2.1 processing software. In the aliphatic oxygenated region, the relative abundances of side-chains involved in the various inter-unit linkages were estimated from the C_α–H_α correlations to avoid possible interference from homonuclear ^{1}H–^{1}H couplings, except for substructures A_{ox} and I, for which C_β–H_β and C_γ–H_γ correlations were used. In the aromatic/unsaturated region, C_2–H_2 correlations from H, G and S lignin units and from *p*-coumarate and ferulate were used to estimate their relative abundances.

DFRC degradation was performed according to the developed protocol [9]. To assess the presence of naturally acetylated lignin units, the described modification of the standard DFRC method using propionylating instead of acetylating reagents (DFRC´) was used [10].

RESULTS AND DISCUSSION

The structural characteristics of the MWL isolated from wheat straw were analyzed in detail by 2D-NMR and DFRC.

The side-chain (δ_C/δ_H 50-90/2.5-5.8) and the aromatic/unsaturated (δ_C/δ_H 90-155/6.0-8.0) regions of the HSQC NMR spectrum are shown in **Fig. 1**, together with the main substructures found. The aliphatic-oxygenated region of the spectra gave information about the different inter-unit linkages present in the lignin. In this region, correlation peaks from methoxyls and side-chains in β–*O*–4´ substructures (A) were the most prominent. Other substructures were also clearly visible, including signals for phenylcoumarans (B), resinols (C), dibenzodioxocins (D), and spirodienones (F). Minor amounts of α,β-diaryl ether substructures (E) could also be detected. The main correlation peaks in the aromatic/unsaturated region of the HSQC spectra corresponded to the aromatic rings and unsaturated side-chains of the different lignin units and hydroxycinnamates. Signals from *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) units were observed. In addition, prominent signals corresponding to *p*-coumarate (PCA) and ferulate (FA) structures were also observed. Other signals in this region of the spectrum are from the unsaturated side-chains of cinnamyl alcohol end-groups (I) and cinnamaldehyde end-groups (J).

The HSQC spectrum also reveals the presence of characteristic signals corresponding to the C_γ–H_γ correlations of γ-acylated β-*O*-4' (A') and other structures in the range from δ_C/δ_H 63.5/3.83 and ~4.30. Therefore, it is possible to conclude that the lignin of wheat straw is partially acylated at the γ-position of the lignin side-chain, as already observed in other grasses [2,4–6]. An estimate of ~10% for the percentage of γ-acylation of lignin side-chains was calculated by integration of the C_γ-H_γ correlation peaks corresponding to the hydroxylated (A_γ) and acylated (A'_γ) substructures in the HSQC spectrum of the isolated MWL (**Table 1**). The fact that the side-chain of the lignin in wheat straw is partially acylated at the γ-OH, together with the presence of significant amounts of *p*-coumarate moieties (4% with respect to lignin), which are known to acylate the γ-OH of the lignin side-chain in many plants, and particularly in grasses [2,4–7], seems to indicate that *p*-coumarates also acylate the γ-OH in the lignin of wheat straw. However, the HSQC spectrum only indicates that the lignin in wheat straw is partially acylated at the γ-position, but cannot provide information on the nature of the acyl group.

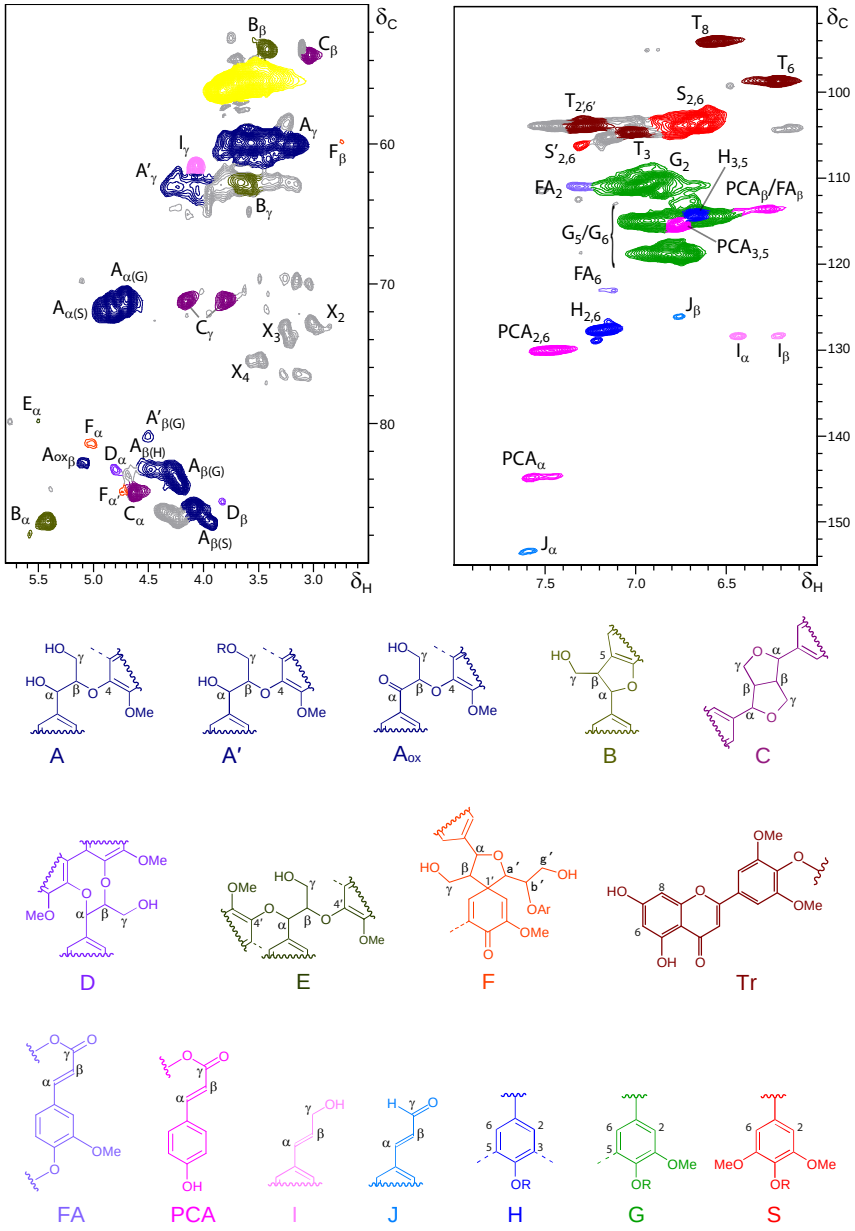

Fig. 1 - Side-chain (δ_C/δ_H 50-90/2.5-5.8) and aromatic/unsaturated (δ_C/δ_H 90-155/5.5-8.0) regions in the 2D HSQC NMR spectra of wheat straw MWL. Main structures present: (A) β–*O*–4´ alkyl-aryl ethers; (A') β–*O*–4´ alkyl-aryl ethers with acylated γ-OH; (A_{ox}) Cα-oxidized β–*O*–4´ structures; (B) phenylcoumarans; (C) resinols; (D) dibenzodioxocins; (E) α,β-diaryl ethers; (F) spirodienones; (I) cinnamyl alcohol end-groups; (J) cinnamyl aldehyde end-groups; (PCA) *p*-coumarates; (FA) ferulates; (H) *p*-hydroxyphenyl units; (G) guaiacyl units; (S) syringyl units.

For this purpose, we performed HMBC experiments that conclusively demonstrate that *p*-coumarate is acylating the γ-position of the lignin side-chains in wheat straw, as also occurs in other grasses [4–7,11]. In addition, the HBMC spectrum also revealed that acetates are also acylating the γ-OH of the lignin side-chain. Acetates have also been previously found acylating the γ–OH in the lignins of many plants, including grasses [4,5,7,10,12,13]. Further confirmation regarding the lignin acylation in wheat, via *p*-coumarate and acetate, was provided by DFRC analyses (as shown below).

The main structural characteristics of wheat straw lignin, estimated from volume integration of contours in the HSQC spectra, are shown in **Table 1**. The data indicated that the structure of the lignins is mostly made up of β–*O*–4′ linkages (accounting for 75% of all the inter-unit linkages), followed by phenylcoumarans (11%) and lower amounts of resinols, dibenzodioxocins, α,β-diaryl ethers, and spirodienones.

Table 1. Structural Characteristics from Integration of ^{13}C-^{1}H Correlation Peaks in the HSQC of the Wheat Straw MWL

	Wheat straw MWL
Lignin inter-unit linkages (%)	
β–*O*–4′ aryl ethers (**A/A'**)	75
α-oxidized β–*O*–4′ aryl ethers ($\mathbf{A_{ox}}$)	2
Phenylcoumarans (**B**)	11
Resinols (**C**)	4
Dibenzodioxocins (**D**)	3
α,β-diaryl ethers (**E**)	2
Spirodienones (**F**)	3
Lignin end-groups[a]	
Cinnamyl alcohol end-groups (**I**)	4
Cinnamaldehyde end-groups (**J**)	4
Lignin side-chain γ-acylation (%)	**10**
Lignin aromatic units[b]	
H (%)	6
G (%)	64
S (%)	30
S/G ratio	0.5
***p*-Hydroxycinnamates**[c]	
p-Coumarates (%)	4
Ferulates (%)	2
p-Coumarates/ferulates ratio	2.0

[a]Expressed as a fraction of the total lignin inter-unit linkage types **A-F**
[b]Molar percentages (H+G+S=100)
[c]*p*-Coumarate and ferulate molar contents as percentages of lignin content

Additional information regarding the acylation of the γ-OH of the lignin side-chain was obtained from DFRC, a degradation method that cleaves α- and β-ether linkages in the lignin polymer leaving γ-esters intact [9,10]. *p*-Coumarate esters were barely detectable (<1%) on monomer conjugates released by selectively cleaving β-ethers in DFRC, indicating that they might be preferentially involved in condensed or terminal structures. In addition, DFRC confirmed the occurrence of native acetylation at the γ-carbon of the lignin side-chain of wheat straw. The analyses indicated that up to 12% of the releasable G-lignin units are acetylated, while only 1% of the total S-lignin units are acetylated.

Identification of the flavone tricin in wheat straw lignin

We also report here the structural elucidation of a component giving unusual correlation peaks in the aromatic regions of HSQC spectra, and provide the first evidence that a flavone is linked to lignin in wheat and other grasses. Two strong and well resolved signals from unknown structures, not previously reported in lignin, were readily observed at δ_C/δ_H 94.1/6.56 and 98.8/6.20 in the HSQC spectrum (**Fig. 1**).

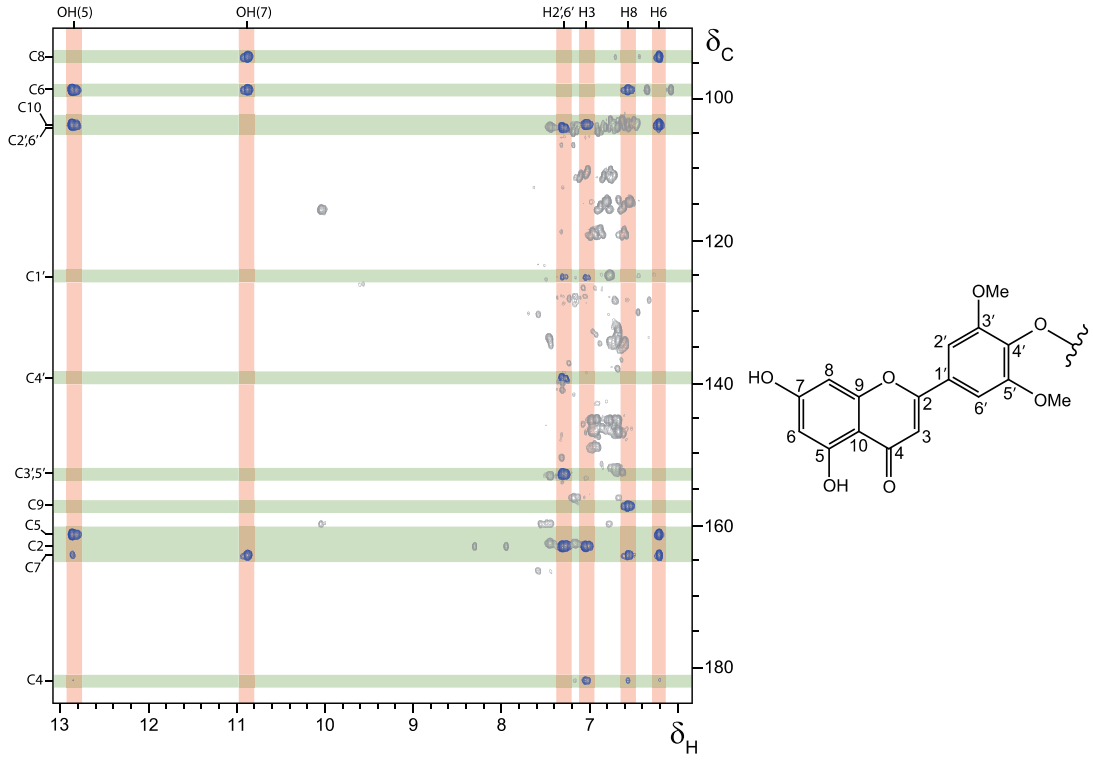

Fig. 2 - Partial HMBC spectrum (δ_C/δ_H 90-185/6.0-13.0) of wheat straw MWL showing the main correlations and the structure of tricin (5,7,4´-trihydroxy-3´,5´-dimethoxyflavone) units in the lignin.

Further valuable information about the nature of this structure was obtained by performing long-range $^{13}C–^{1}H$ correlation (HMBC) experiments (**Fig. 2**). The correlation peaks in the HMBC spectrum clearly indicate that this moiety has a flavone-type structure. Flavones are a class of flavonoids that have the 2-phenylchromen-4-one backbone. From the HMBC data, it was then possible to conclude that the structure of the flavone moiety present in the MWL of wheat straw is tricin (5,7,4´-trihydroxy-3´,5´-dimethoxyflavone, **Fig. 2**); the ^{1}H and ^{13}C shifts match those published for tricin [14].

The signals appearing in the HSQC spectrum at δ_C/δ_H 94.1/6.56 and 98.8/6.20 (**Fig. 1**) thus correspond to the C_8-H_8 and C_6-H_6 correlations, respectively. The HSQC also shows the C_3-H_3 correlation at δ_C/δ_H 104.5/7.04, near the $S_{2,6}$ signal. On the other hand, the phenyl moiety linked at C-2 is of syringyl-type, the correlations for $C_{2'}$-$H_{2'}$ and $C_{6'}$-$H_{6'}$ being also observed in the HSQC spectrum at δ_C/δ_H 103.9/7.31. Tricin has two phenolic hydroxyls at C-5 and C-7 of the chroman-4-one skeleton, with diagnostic phenolic proton chemical shifts that are readily apparent in the HMBC. In addition, the absence of the signals for the phenolic 4´-OH of tricin in the HMBC proton indicates that it is not free. Therefore, incorporation of tricin into the lignin network through 4–O–β ether linkages is indicated. In fact, a signal for the correlation of the tricin C4´ carbon (at 139.5 ppm) and a proton at the β-position of a G-unit at 4.28 ppm was clearly observed in the HMBC spectrum (**Fig. 3**), providing evidence for this incorporation.

Flavones (as all other flavanoids/flavonoids) are metabolic hybrids as they are derived from a combination of the shikimate-derived phenylpropanoid and the acetate/malonate-derived polyketide pathways. Since polyphenols are formed in lignified regions by oxidative coupling, incorporation into the lignin structure is a possibility. In fact, other related benzene diols and triols, such as the flavanols epicatechin, epigallocatechin, or epigallocatechin gallate, although they are not known in actual plant cell walls, have been shown to produce lignin copolymers with normal monolignols [15].

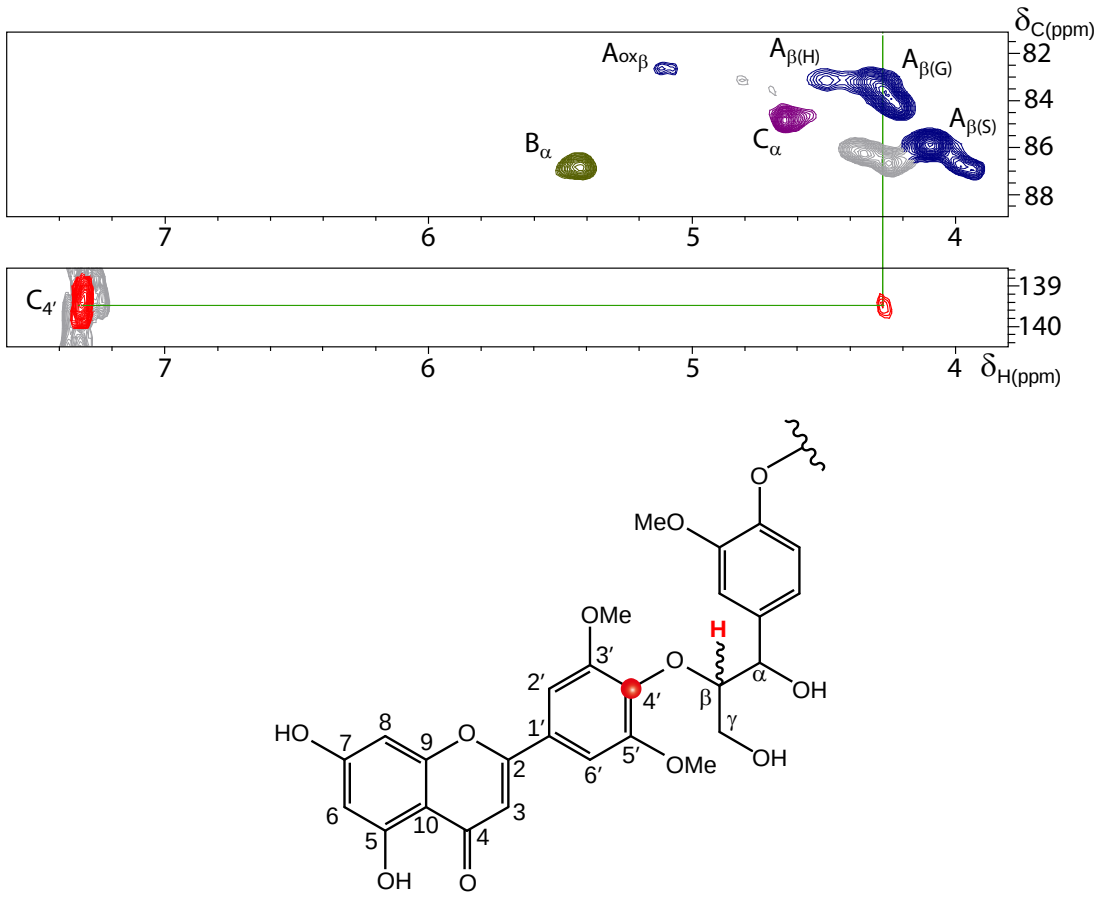

Fig. 3 - Section of the HMBC spectrum of wheat straw MWL showing the correlation for the tricin C4´ carbon (at 139.5 ppm) and the proton at the β-position of a G-unit at 4.28 ppm (bottom). The section of the HSQC spectrum for the Cβ-Hβ correlations of the β–*O*–4´ alkyl-aryl ethers is also shown (top). The structure illustrates the likely incorporation of tricin into the lignin polymer through a 4´–*O*–β ether linkage with a G unit.

The presence in lignins of components from other pathways is of significant interest. It has never been demonstrated, for example, that lignans, dimers and higher oligomers that also arise from radical coupling of monolignols, become incorporated into lignins [16,17]. Lignans are produced under proteinaceous control such that they are always optically active. Lignins are completely racemic [17]; no components excised from lignins have ever been shown to be optically active, so the lignin polymer and the lignan 'extractives' are assumed to be independently produced in time and space. For this reason, this observation and the evidence presented here that a flavone, tricin, is incorporated into lignin, is a new phenomenon with rather profound implications. It implies that the monomer is exported to the cell wall where it undergoes radical coupling reactions with monolignols, or at least with the primary monolignol coniferyl alcohol, and becomes part of the lignin polymer. The occurrence of tricin in the MWL of wheat straw seems to be evidence for this incorporation. At the very least, then, this observation will require further recognition of the malleability of lignification and perhaps another addition to the list of phenolics that must be considered to be 'lignin monomers'. Interestingly, tricin has also been recently found in the lignin from coconut coir [18]. Coconut, like the grasses, belongs to the monocots. Therefore, it is beginning to appear that tricin may be a feature restricted to monocot lignins although its clade range remains to be determined.

CONCLUSIONS

The lignin from wheat straw is a H:G:S lignin, with a strong predominance of G-lignin units (S/G 0.5), and with some amounts of associated *p*-coumarates and ferulates. The main lignin inter-unit linkages are β–*O*–4´ alkyl-aryl ethers, followed by phenylcoumarans and minor amounts of resinols, spirodienones, dibenzodioxocins and α,β-diaryl ethers, together with cinnamyl alcohol and cinnamaldehyde end-groups. 2D-NMR also indicated that this lignin is partially acylated (~10% of all side-chains), and exclusively at the γ-carbon of the side-chain, with acetates and *p*-coumarates. DFRC analyses indicated that acetates preferentially acylate the γ-OH in guaiacyl (12%) rather than in syringyl units (1%). On the other hand, and despite *p*-coumarates' having been found acylating the γ-OH, they were barely detectable as the monolignol conjugates after selectively cleaving β-ethers in lignin in the DFRC method, which seems to indicate that *p*-coumarates must be preferentially involved in structures other than β-ethers. Finally, we present the first evidence that the flavone tricin was found in wheat lignin, etherified by a G-type unit. If it is ultimately shown to have incorporated, in the cell wall, into the lignin by the radical coupling reactions that typify lignification (as it appears), the definition of lignin, and what constitutes a lignin monomer, will need further refinement.

ACKNOWLEDGMENTS

This study has been funded by the Spanish project AGL2011-25379, the CSIC project 201040E075 and the EU-project LIGNODECO (KBBE-244362). Dr. Jorge Rencoret thanks the CSIC for a JAE-DOC contract of the program "Junta para la Ampliación de Estudios" co-financed by Fondo Social Europeo (FSE). John Ralph was funded in part by the DOE Great Lakes Bioenergy Research Center (DOE Office of Science BER DE-FC02-807ER64494).

REFERENCES

[1] Sarkar, N., Ghosh, S.K., Bannerjee, S., Aikat, K., "Bioethanol production from agricultural wastes: An overview". Renewable Energy, 37: 19–27 (2012)

[2] Crestini, C., Argyropoulos, D.S., "Structural analysis of wheat straw lignin by quantitative ^{31}P and 2D NMR spectroscopy. The occurrence of ester bonds and α–*O*–4 substructures". Journal of Agriculture and Food Chemistry, 45: 1212–1219 (1997)

[3] Sun, X.-F., Sun, R.-C., Fowler, P., Baird, M.S., "Extraction and characterization of original lignin and hemicelluloses from wheat straw". Journal of Agriculture and Food Chemistry, 53: 860–870 (2005)

[4] del Río, J.C., Prinsen, P., Rencoret, J., Nieto, L., Jiménez-Barbero, J., Ralph, J., Martínez, A.T., Gutiérrez, A., "Structural characterization of the lignin in the cortex and pith of elephant grass (*Pennisetum purpureum*) stems". Journal of Agriculture and Food Chemistry, 60: 3619–3634 (2012)

[5] del Río, J.C., Rencoret, J., Prinsen, P., Martínez, A.T., Ralph, J., Gutiérrez A., "Structural characterization of wheat straw lignin as revealed by analytical pyrolysis, 2D-NMR, and reductive cleavage methods". Journal of Agricultural and Food Chemistry, 60, 5922–5935 (2012)

[6] Ralph, J., Hatfield, R.D., Quideau, S., Helm, R.F., Grabber, J.H., Jung, H.-J.G., "Pathway of *p*-coumaric acid incorporation into maize lignin as revealed by NMR". Journal of the American Chemical Socity, 116: 9448–9456 (1994)

[7] del Río, J.C., Rencoret, J., Marques, G., Gutiérrez, A., Ibarra, D., Santos, J.I., Jiménez-Barbero, J., Zhang, L., Martínez, A.T., "Highly acylated (acetylated and/or *p*-coumaroylated)

native lignins from diverse herbaceous plants". Journal of Agriculture and Food Chemistry, 56: 9525–9534 (2008)

[8] 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, pp. 137–234 (2010)

[9] Lu, F., Ralph, J., "Derivatization followed by reductive cleavage (DFRC method), a new method for lignin analysis: protocol for analysis of DFRC monomers". Journal of Agriculture and Food Chemistry, 45: 2590–2592 (1997)

[10] Ralph, J., Lu, F., "The DFRC method for lignin analysis. 6. A simple modification for identifying natural acetates in lignin". Journal of Agriculture and Food Chemistry, 46: 4616–4619 (1998)

[11] Ralph, J., "Hydroxycinnamates in lignification". Phytochemistry Reviews, 9: 65–83 (2010)

[12] del Río, J.C., Marques, G., Rencoret, J., Martínez A.T., Gutiérrez, A., "Occurrence of naturally acetylated lignin units". Journal of Agriculture and Food Chemistry, 55: 5461–5468 (2007)

[13] Ralph, J., "An unusual lignin from kenaf". Journal of Natural Products, 59: 341–342 (1996)

[14] Chang, C.L., Wang, G.J., Zhang, L.J., Tsai, W.J., Chen, R.Y., Wu, Y.C., Kuo, Y.H., "Cardiovascular protective flavolignans and flavonoids from *Calamus quiquesetinervius*". Phytochemistry, 71: 271–279 (2010)

[15] Grabber, J.H., Schatz, P.F., Kim, H., Lu, F., Ralph, J., "Identifying new lignin bioengineering targerts: 1. Monolignol-substitute impacts on lignin formation and cell wall fermentability". BMC Plant Biology, 10: 114 (2010)

[16] Ralph, J., Lundquist, K., Brunow, G., Lu, F., Kim, H., Schatz, P.F., Marita, J.M., Hatfield, R. D., Ralph, S.A., Christensen, J.H., Boerjan, W., "Lignins: Natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids". Phytochemistry Reviews, 3: 29–60 (2004)

[17] Ralph, J., Peng, J., Lu, F., Hatfield, R.D., "Are lignins optically active?". Journal of Agriculture and Food Chemistry, 47: 2991–2996 (1999)

[18] Rencoret, J., Ralph, J., Marques, G., Gutiérrez, A., Martínez, A.T., del Río, J.C., "Structural characterization of lignin isolated from coconut (*Cocos nucifera*) coir fibers". Journal of Agriculture and Food Chemistry, 61: 2434−2445 (2013)