

The Effects on Lignin Structure of Overexpression of Ferulate 5-Hydroxylase in Hybrid Poplar^{1[W]}

Jaclyn J. Stewart, Takuya Akiyama, Clint Chapple, John Ralph, and Shawn D. Mansfield*

Department of Wood Science, University of British Columbia, Vancouver, Canada V6T 1Z4 (J.J.S., S.D.M.); Department of Biochemistry and Great Lakes Bioenergy Research Center, University of Wisconsin, Madison, Wisconsin 53706 (T.A., J.R.); and Department of Biochemistry, Purdue University, West Lafayette, Indiana 47907 (C.C.)

Poplar (*Populus tremula* × *alba*) lignins with exceedingly high syringyl monomer levels are produced by overexpression of the ferulate 5-hydroxylase (F5H) gene driven by a cinnamate 4-hydroxylase (C4H) promoter. Compositional data derived from both standard degradative methods and NMR analyses of the entire lignin component (as well as isolated lignin fraction) indicated that the C4H::F5H transgenic's lignin was comprised of as much as 97.5% syringyl units (derived from sinapyl alcohol), the remainder being guaiacyl units (derived from coniferyl alcohol); the syringyl level in the wild-type control was 68%. The resultant transgenic lignins are more linear and display a lower degree of polymerization. Although the crucial β -ether content is similar, the distribution of other interunit linkages in the lignin polymer is markedly different, with higher resinol (β - β) and spirodienone (β -1) contents, but with virtually no phenylcoumarans (β -5, which can only be formed from guaiacyl units). *p*-Hydroxybenzoates, acylating the γ -positions of lignin side chains, were reduced by >50%, suggesting consequent impacts on related pathways. A model depicting the putative structure of the transgenic lignin resulting from the overexpression of F5H is presented. The altered structural features in the transgenic lignin polymer, as revealed here, support the contention that there are significant opportunities to improve biomass utilization by exploiting the malleability of plant lignification processes.

In the continuing search for improved biomass utilization in processes including chemical pulping, natural ruminant digestibility, and biomass conversion to ethanol, considerable attention has focused on improving lignocellulosic feedstocks through genetic engineering. Perturbing plant biomass deposition by misregulating key genes/enzymes integral to major cell wall pathways can provide rich insights into cell wall development and architecture and also create significant opportunities for improved lignocellulosic utilization (Baucher et al., 2003; Boerjan et al., 2003; Ralph et al., 2004). Lignins comprise the second most abundant polymer class in the biosphere; however, their combinatorial biosynthesis renders them among the more complex biomacromolecules synthesized by plants. Alterations in plant cell wall chemistry or ultrastructure, including lignin content or structure, can have a profound effect on chemical or enzymatic degradability. For example, in chemical pulping, lignin structure and content have been shown to significantly impact delignification efficiency (both pulp yield and residual lignin [RL] content) and pulp bleachability (Chang and Sarkanen, 1973; Huntley et al., 2003; Stewart

et al., 2006). Similarly, improvements in fermentable sugar yields have been reported in lignin-engineered *Medicago sativa* (Chen and Dixon, 2007).

In recent years, the genes encoding enzymes specific to the lignin branch of the phenylpropanoid pathway have been cloned and their roles evaluated using a combination of forward and reverse genetics (Franke et al., 2002; Baucher et al., 2003; Boerjan et al., 2003). The down-regulation of genes early in the pathway may limit the overall flux of metabolites to lignin synthesis. In contrast, the genes common to the latter part of the pathway generally affect the distribution of *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) lignin units resulting from the primary monomers (the three monolignols *p*-coumaryl, coniferyl, and sinapyl alcohols). And, in several cases when the biosynthesis of a normal monolignol is severely curtailed, lignins appear to incorporate unique phenolics (e.g. 5-hydroxyconiferyl alcohol, hydroxycinnamaldehydes, and ferulic acid) that have not traditionally been considered lignin monomers (Sederoff et al., 1999; Boerjan et al., 2003; Ralph et al., 2004, 2008a, 2008b).

Ferulate 5-hydroxylase (F5H), also referred to as coniferaldehyde 5-hydroxylase to reflect one of its preferred substrates (Humphreys et al., 1999; Osakabe et al., 1999), is a key enzyme involved in synthesizing the monolignol sinapyl alcohol and, ultimately, S lignin moieties. F5H therefore affects the partitioning between the two major traditional monolignols, coniferyl and sinapyl alcohols, and is fundamental to the evolutionary differences between gymnosperms (with

* Corresponding author; e-mail shawnman@interchange.ubc.ca.

The author responsible for the distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Shawn D. Mansfield (shawnman@interchange.ubc.ca).

^[W] The online version of this article contains Web-only data.

www.plantphysiol.org/cgi/doi/10.1104/pp.109.137059

no S components) and angiosperms (with compositions favoring the S monomers). For example, the *fah1* Arabidopsis (*Arabidopsis thaliana*) mutant, deficient in F5H, has little to no S lignin (Meyer et al., 1998; Marita et al., 1999). Like gymnosperms, it produces G-rich lignins, derived almost exclusively from coniferyl alcohol. In contrast, the overexpression of F5H in the mutant background produces plants displaying substantially higher than normal sinapyl alcohol-derived S units and is consequentially severely depleted in coniferyl alcohol-derived G units. Wet chemical analyses of cell wall lignins estimated S contents of up to about 92% in F5H-up-regulated Arabidopsis (Meyer et al., 1998), up to 84% in tobacco (*Nicotiana tabacum*; Franke et al. 2000), and as high as 93.5% in hybrid poplar (*Populus tremula* × *Populus alba*; Huntley et al., 2003; Li et al., 2003). These engineered cell walls, rich in S units, exceed the highest reported in nature to date (Baucher et al., 1998), with kenaf (*Hibiscus cannabinus*) bast fiber lignin, at 85% S, being among the highest (Ralph, 1996; Morrison et al., 1999). In poplars, the final methylation step, catalyzed by caffeic acid 3-O-methyl transferase (COMT), appears to adequately accommodate the increased flux from coniferaldehyde to 5-hydroxyconiferaldehyde to produce sinapaldehyde and ultimately sinapyl alcohol (Li et al., 2000, 2003). In Arabidopsis, however, evidence suggests that the COMT is not able to keep pace with the increased 5-hydroxyconiferaldehyde generated by F5H overexpression, since the ensuing lignins comprise a significant component derived from 5-hydroxyconiferyl alcohol (Ralph et al., 2001a). Novel 5-hydroxyguaiacyl benzodioxane structures, which result from incorporation of 5-hydroxyconiferyl alcohol into the lignification scheme, were the same as those noted in COMT-deficient plants (Ralph et al., 2001a, 2001b; Marita et al., 2001, 2003).

The availability of woody plant material possessing an extreme S lignin concentration affords unique opportunities to explore the mechanisms and consequences of pushing plants toward compositional limits and has both major fundamental and industrial consequences. Engineering lignin composition to extreme levels provides rare and novel insights into the ramifications on the structure of the lignin component and on other cell wall characteristics. Here, we investigate the nature of the chemical modifications to the lignin polymer in up-regulated cinnamate 4-hydroxylase (C4H)::F5H poplar and propose a model for how such gene perturbation impacts lignin biosynthesis.

RESULTS

The aim of this study was to delineate the consequences of C4H::F5H up-regulation on lignin structure and biosynthesis in hybrid poplar secondary xylem tissue. More detailed descriptions of the lines analyzed herein and other lines with varying degrees of C4H::F5H overexpression are reported elsewhere (Huntley

et al., 2003). Briefly, the transgenic hybrid poplar, with substantially more S units (ranging from 65% in the wild type to as high as 93/95 mol % S units in transgenics, by thioacidolysis/derivatization followed by reductive cleavage [DFRC]), appeared phenotypically normal and displayed no pleiotropic growth effects.

Lignin Abundance and Aromatic Unit (S:G and *p*-Hydroxybenzoate) Distribution

Stems from 2.5-year-old greenhouse-grown trees were harvested from the C4H::F5H-up-regulated and wild-type hybrid poplar lines. Total cell wall lignin contents were estimated by the Klason lignin analysis, while thioacidolysis was used to quantify monomer yield (Table I). Klason analysis indicated that although the total lignin content was similar between the wild-type and transgenic trees (Table I), there is a substantial change in the distribution between acid-soluble and insoluble lignin fractions, with the transgenic trees displaying elevated levels of acid-soluble lignin and reduced levels of acid-insoluble lignin in comparison to the wild-type trees (Table I). Similarly, the transgenic trees also had significantly higher thioacidolysis monomer yields. Thioacidolysis (Rolando et al., 1992) and DFRC (Lu and Ralph, 1997) provided an estimate of monomer contribution, expressed as relative %G (derived from coniferyl alcohol) and %S (from sinapyl alcohol). Both methods release quantifiable monomers from units linked by β -ether bonds (Table II), and as such the measured S:G ratio is a reflection of the units involved only in so-called uncondensed units. However, the quantification of S:G ratios is confounded due to the poor release of the S monomer from β -ether units acylated with *p*-hydroxybenzoates, which are inherently common to poplars. NMR analyses (below) were therefore concurrently used to measure the S:G monomer ratios of the entire lignin (Table II) and to estimate the *p*-hydroxybenzoate contents. Despite the significant increase in the proportion of S units in the lignin, the β -ether content was essentially the same for the two lines.

NMR EVALUATION

Aromatic Region

Changes in the S:G distribution in the lignins are most readily visualized from the aromatic region of NMR spectra, particularly the two-dimensional (2D) ^{13}C - ^1H correlation (HSQC) spectra correlating protons with their attached carbons. Figure 1 shows the impressive differences in the aromatic nature of the lignins in the wild-type poplar (Fig. 1, A and C) compared to the C4H::F5H-overexpressing transgenic poplar (Fig. 1, B and D). As established previously, native wild-type poplar lignin is a S-rich, S-G lignin (Stewart et al., 2006). While both S and G aromatic resonances were readily observed, only trace levels of H units were detected. The lignin from the C4H::F5H-up-regulated

Table I. Lignin content, thioacidolysis yield, S:G monomer ratio, and lignin characteristics of wild-type and C4H::F5H transgenic hybrid poplar wood and isolated lignin fractions ($\pm$ SD)

Fractions are defined in "Materials and Methods." ML, Dioxane:water-soluble milled lignin; EL, enzyme-digested cell wall; RL, residual lignin residue remaining after dioxane:water and enzyme extractions. Asterisk indicates that it was based on lignin basis (wt %). Boldface represents statistically significant ($P = 0.05$) from three biological replicates.

Chemical	Unit	Wild Type	C4H::F5H
Acid-insoluble lignin	%	20.06 (0.32)	16.53 (0.22)
Acid-soluble lignin	%	2.67 (0.13)	5.28 (0.21)
Wood thioacidolysis yield	$\mu\text{mol/g}$	7739 (21)	895 (30)
Wood S:G ratio		1.90	14.17
ML thioacidolysis yield	$\mu\text{mol/g}$	1,120 (38)	1,530 (101)
ML S:G ratio		1.51	12.04
EL thioacidolysis yield	$\mu\text{mol/g}$	2,053 (42)	1,948 (29)
EL S:G ratio		2.24	12.37
RL Lignin thioacidolysis yield	$\mu\text{mol/g}$	848 (28)	846 (33)
RL Lignin S:G ratio		1.20	8.76
Phenolic OH*	(wt %)	6.62 (0.31)	14.1 (0.26)
OMe*	(wt %)	16.3 (0.71)	29.9 (0.88)

poplar (Fig. 1, B and D) is strikingly S rich. Relatively weak, but diagnostic, G correlations remain in a spectrum that is dominated by the S correlations; traces of H units at levels (approximately 0.3%) apparently elevated over the wild-type level (approximately 0.15%) were also detected (data not shown). Volume integration (Table II) allows reasonable quantification of the differences in monomer composition and suggests that the solvent-soluble lignin (ML) consists of 97.5% S, while the enzyme lignin (EL) is composed of 97.4% S following cell wall dissolution (Fig. 2; Table II). This implies that minimal partitioning of lignin structural types has occurred during the fractionation process and that the isolated ML from ball-milled material is representative of that from the whole ball-milled cell wall.

From the aromatic region of the spectra, it is also evident that both the wild-type and transgenic C4H::F5H poplar (Fig. 1) possess *p*-hydroxybenzoate

groups. These free-phenolic units are commonly found acylating the γ -positions of lignin side chains of poplars, willows (*Salix* spp.), and palms (*Palmae*; Ralph et al., 2004). Although the quantitative methodology here has not been firmly established, *p*-hydroxybenzoate levels were measured by 2D NMR volume integration, and comparisons between samples are independent of any quantification response factors. It appears that levels (relative to the total lignin) in the C4H::F5H-up-regulated poplar trees are only half those in the corresponding wild-type trees.

Side Chain Region

The side chain region peripherally reflects the changes in the S:G distribution and is rich in detail regarding the types and distribution of interunit linkage patterns present in the isolated lignin fractions. The wild-type lignin spectrum (Fig. 2, A and C) depicts a typical G/S

Table II. Wild-type and C4H::F5H transgenic hybrid poplar lignin G/S data from NMR, DFRC, and thioacidolysis; NMR-derived *p*-hydroxybenzoate levels; and interunit linkage data

Fractions are defined in "Materials and Methods." ML, Dioxane:water-soluble milled lignin; EL, enzyme-digested cell wall; Ac, acetylated samples. WT is the wild-type control. F5H is the F5H-up-regulated transgenic. PB, *p*-Hydroxybenzoate; **A**, β -O-4 (β -aryl ether); **A3**, β -O-4 (α -keto- β -aryl ether); **B**, β -5 (phenylcoumaran); **C**, β - β (resinol); **S**, β -1 (spirodienone); **X1**, cinnamyl alcohol end group. DFRC and thioacidolysis S/G data are from whole (extracted) wood, not on the isolated lignin fractions.

Sample	%G	%S	S/G ^a	%PB ^b	%A	%A3	%B	%C	%S	%X1	$\Sigma\beta$ -O-4
WT AcML	30.8	69.2	2.24	11.8	86.7	0.7	2.0	7.9	2.1	0.7	87.4
WT AcEL	34.3	65.7	1.91	13.5	86.9	0.8	2.2	6.2	1.6	2.3	87.7
WT DFRC	27.6	72.4	2.63								
WT thio	34.5	65.5	1.90								
F5H AcML	2.5	97.5	38	4.8	83.5	2.6	0.1	9.7	3.3	0.9	86.1
F5H AcEL	2.6	97.4	38	5.3	82.5	3.7	0.1	8.5	3.5	1.8	86.2
F5H DFRC	4.7	95.3	20								
F5H thio	6.6	93.4	14.2								

^aS/G values become unstable at very high-G or high-S values; the %S and %G values are a better indicator in these cases. ^b%PB is the percentage of the *p*-hydroxybenzoate on a lignin (S+G) basis.

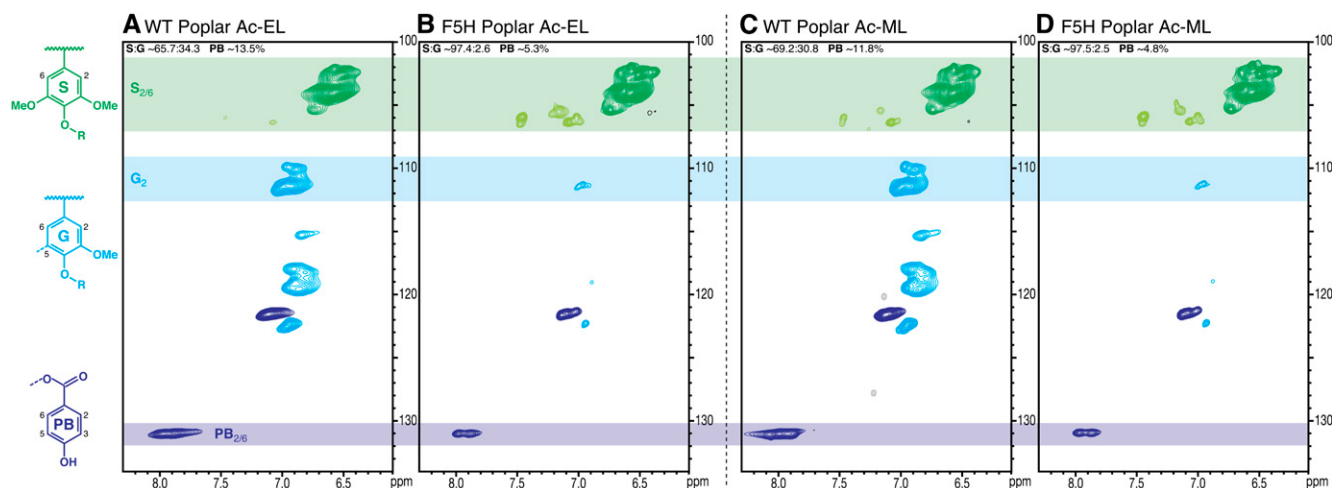


Figure 1. Partial short-range 2D ^{13}C - ^1H (HSQC) correlation NMR spectra (aromatic regions only) of acetylated enzyme lignins (EL) isolated from wild type (WT; A) and the most highly C4H::F5H-up-regulated poplar (B), and acetylated milled wood lignins (ML) isolated from the wild type (C) and the most highly C4H::F5H-up-regulated line (D). Comparison of spectra from the transgenic (B and D) versus the controls (A and C) illustrates the substantive S/G compositional shift. Ac, Acetylated.

lignin containing a small amount of residual polysaccharides (Ralph et al., 1999). The lignin is seen as being rich in β -aryl ether units **A**, with more modest amounts of phenylcoumaran **B**, resinol **C**, and spirodienone **S** units, as is typical for lignins derived from angiosperms. Although α -keto- β -ethers units **A3** are not normally reported, they were apparent in these lignins and are indicative of more highly S-enriched lignins. Spirodienone structures **S**, β -1-coupled units only recently authenticated in lignin spectra (Zhang and Gellerstedt, 2001; Zhang et al., 2006), are readily seen in these poplar samples. Dibenzodioxocins (usually labeled **D**), eight-membered ring structures resulting from radical coupling of a monolignol with a 5-5-coupled end unit (Karhunen et al., 1995), were not apparent. This is not surprising given that S units cannot participate in 5-coupling, and as such, dibenzodioxocins are most prevalent in G-rich lignin fractions (Ralph et al., 1999). Resinols **C**, arising from monomer dimerization, usually from sinapyl alcohol, are especially pronounced in high-S lignins. Cinnamyl alcohol end groups **X1** arising from monomer-monomer coupling (often involving G units) are relatively minor; increasing evidence suggests that the X1 γ correlation peak is contaminated, so we are careful not to overinterpret the significance of any differences in the integral from this component. There is also evidence for arylglycerol units **X7** (structure not shown); the C β -H β correlation is labeled and colored orange.

The lignin from the C4H::F5H-up-regulated trees has a spectrum that has several distinct differences when compared to the parent wild-type trees (Fig. 2, B and D; Table II). First, there are no apparent phenylcoumarans **B**. Second, there are elevated levels of oxidized β -ether units, represented by the α -keto- β -ether units **A3**. It is also apparent from the relative

simplicity of, for example, the β -aryl ether **A** α -proton/carbon-correlation contours (see the expansions in Fig. 2) that the diversity of structures is far less than those observed in corresponding wild-type poplar lignin; the C4H::F5H transgenic lignin clearly has a smaller range in diversity of its side chain linkages and, thus, a much more homogenous structure. This logically explains the high thioacidolysis monomer yields (Table I).

Long-Range Correlations

Long-range correlations (resulting from carbons and protons linked via two to three intervening bonds), from heteronuclear multiple-bond correlation (HMBC) spectra, are valuable in providing information on the types of units (G or S) involved in each linkage type (Hu et al., 1999; Marita et al., 1999; Ralph et al., 1999; Li et al., 2003). The following observations are evident from the wild-type lignin spectrum (Fig. 3A): both β -ether **A** and phenylcoumaran **B** units are derived from both sinapyl and coniferyl alcohol coupling reactions—they are associated with both S and G units; resinol units **C** are essentially all S, deriving from sinapyl alcohol; traces involving G units are observable. The β -ether units **A'** bearing γ -*p*-hydroxybenzoate substituents are all S. Arylglycerol units **X7** are all S.

HMBC correlations from the lignin originating from the C4H::F5H transgenic trees (Fig. 3B) emphasize the extremely S-rich nature of these secondary cell walls. The β -ether units **A** are essentially only S (as is the entire lignin); only traces of G β -ethers can be detected. The lower level of γ -*p*-hydroxybenzoylated β -ether units **A'** remain all S, as do the resinols **C**. Finally, the source of the phenylcoumarans is indistinguishable, as

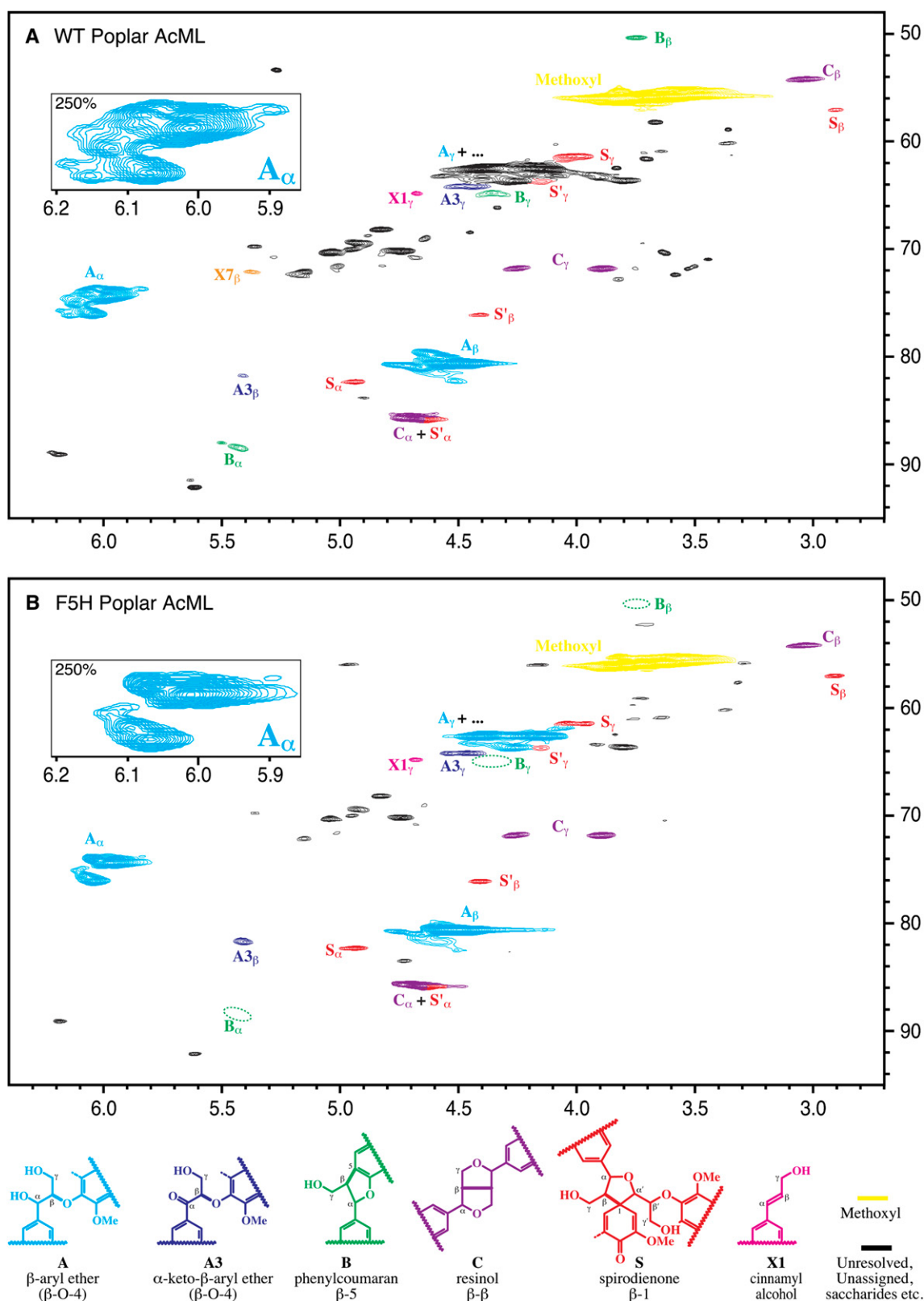


Figure 2. Partial short-range 2D ^{13}C - ^1H (HSQC) NMR spectra (side chain regions) of acetylated milled lignins (ML) isolated from the wild type (WT; A), the most highly C4H::F5H-up-regulated line (B), and EL from the wild type (C) and C4H::F5H-up-regulated poplar (D). Comparison of the spectra indicates the substantial structural differences between the two lignin samples (see text). Volume integrals and semiquantitative data are given in Table II. Interunit type designations A-C, S, X1, and X7 follow

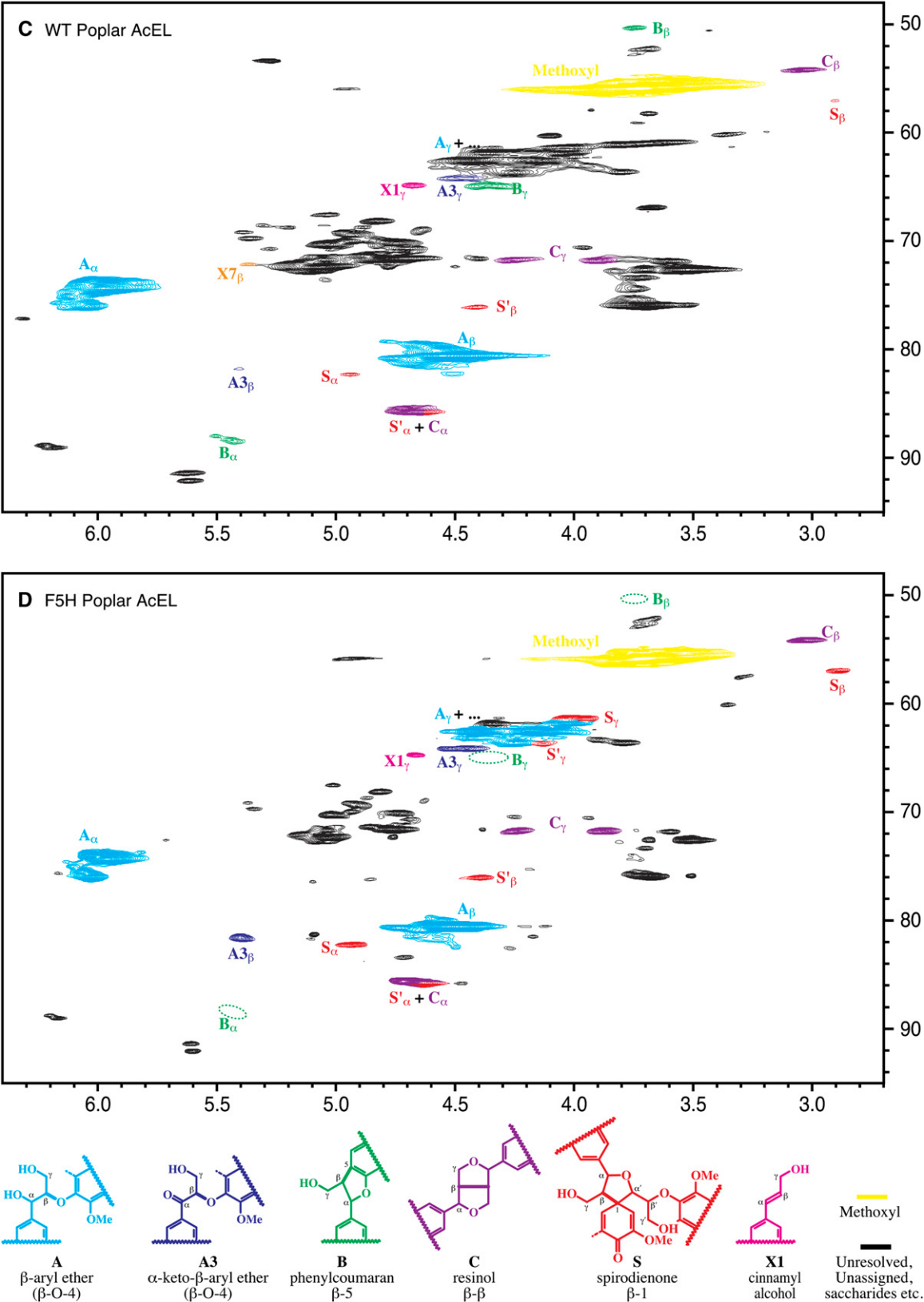


Figure 2. (Continued.) conventions established previously (Ralph et al., 1999). Ac, Acetylated. (Note: A structure for the minor arylglycerol component X7 is not shown.)

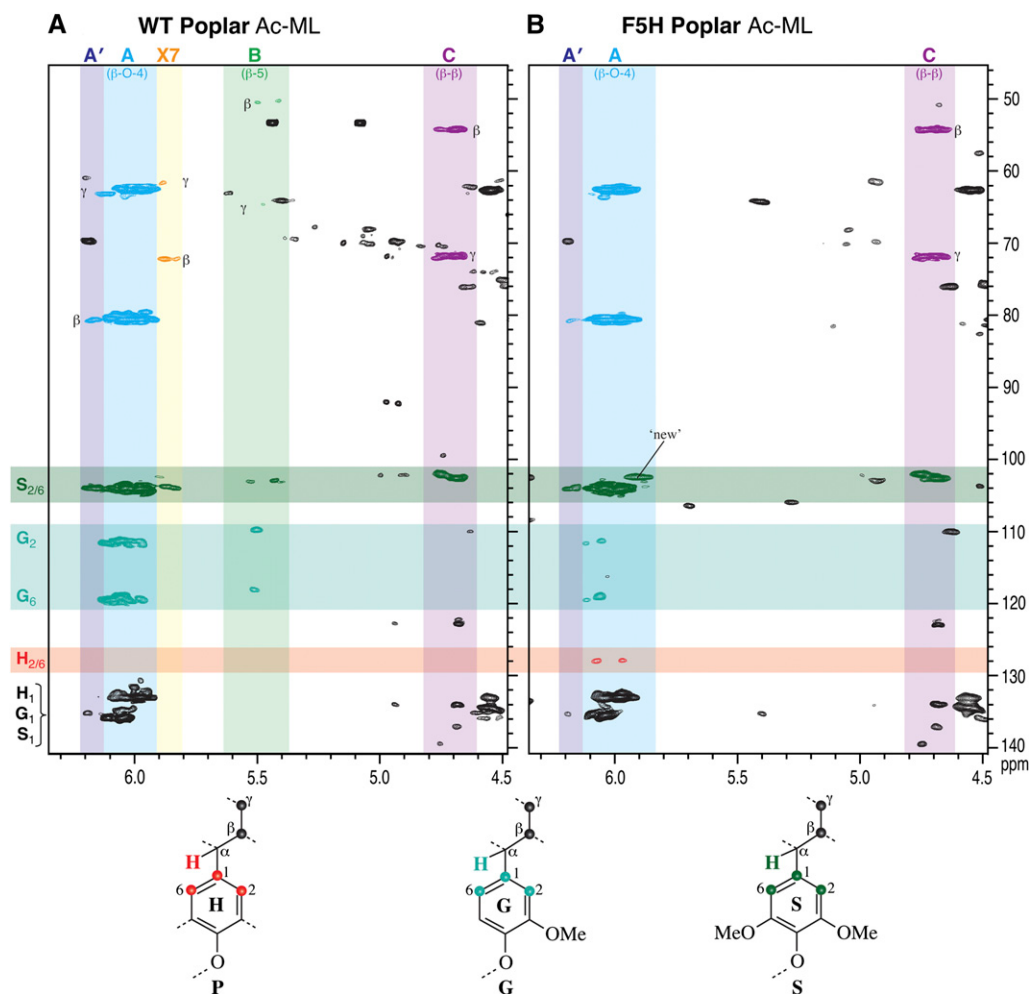


Figure 3. Partial long-range ^{13}C - ^1H (HMBC) spectra (side chain regions) of acetylated milled-wood lignins (ML) isolated from the wild-type (WT) control (A) and the most highly F5H-up-regulated line (B), C4H::F5H. Interunit type designations **A** to **C**, and **X7** (arylglycerol) are the same as in Figure 2; **A'** is the designation used for β -ether units **A** that are γ -acylated by *p*-hydroxybenzoate. Ac, Acetylated.

they are virtually nonexistent, and arylglycerol units **X7** are also not apparent.

Molecular Weight Distribution of Isolated Lignin Fractions

Gel permeation chromatography (GPC) was used to investigate the qualitative M_r distribution of the isolated lignin fractions. GPC traces of the ML fractions demonstrate that the M_r distribution is different between samples (Fig. 4), with the wild type having a higher M_r . The width of the peak for the C4H::F5H ML is broader than that of the wild type, which indicates that the molecular polydispersity is larger in the transgenic sample. In contrast, the EL fractions have almost identical distributions except for a small lower M_r component eluting at approximately 45 min, which is also present, but to a smaller extent, in the wild-type sample (Fig. 4). The most dramatic difference is ob-

served in the RL fractions, where the mass distribution shows that the wild type has a substantially greater portion of its lignin as high M_r polymers (Fig. 4C). The opposite is true for the C4H::F5H RL.

The RL fraction GPC traces are bimodal. The wild-type RL has a higher portion of its lignin at a higher M_r , whereas for the C4H::F5H, a greater portion of its lignin is in the lower M_r fraction. Since the RL makes up the highest proportion of the total lignin by weight, differences in this fraction have a significant impact on the overall composition and structure. Furthermore, the EL fraction has a higher M_r (based on peak retention times) than the milled-wood lignin (ML) for the modified C4H::F5H lignin. This suggests that the walls in the transgenic are more susceptible to the cellulase treatment than the corresponding wild-type trees. The wild-type milled-wood lignin is actually lower in M_r than its EL fraction, indicating that higher M_r lignin is liberated from the carbohydrate matrix

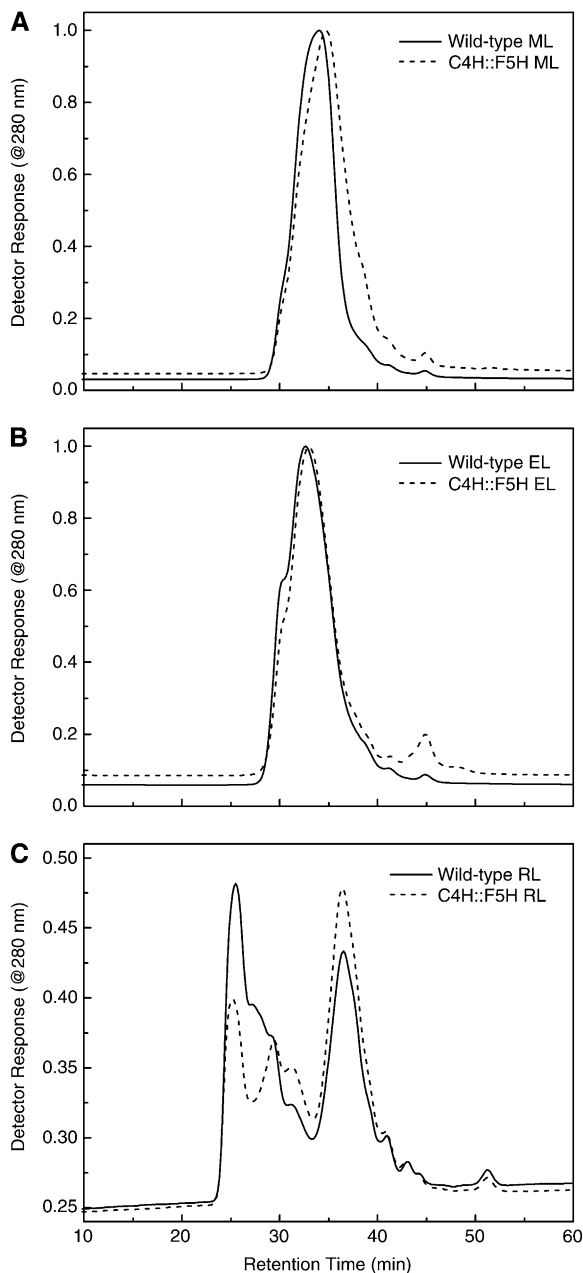


Figure 4. Molecular weight profiles of acetylated milled lignin (ML; A), acetylated EL (B), and acetylated RL (C) from wild-type and C4H::F5H poplar.

during the enzyme treatment. Although it is thought that EL is similar in chemical structure to milled-wood lignin, the thioacidolysis yield (a measure of β -O-4 units) is higher for EL, and the M_r of the EL samples is different than the milled-wood lignin.

These dramatic changes in M_r distribution of the three lignin fractions clearly demonstrate differences in the lignin superstructure resulting from altering the expression of a gene in the lignin biosynthetic pathway. The C4H::F5H modification not only leads to

lignin that has been altered in S:G ratio but has a lower M_r as well. It is conceivable that this large change in lignin degree of polymerization is a result of a restrictive bonding environment created in the S-rich lignin, where the majority of the monomers are substituted at both the 3'- and 5'-positions on the aromatic ring (see lignin model; Fig. 5).

DISCUSSION

The overexpression of F5H in hybrid poplar produces trees with normal growth phenotypes but with substantial compositional and structural differences in their lignins (Huntley et al., 2003). An evaluation of monomer composition by thioacidolysis clearly demonstrated that the transgenic trees are substantially enriched in S units and had a higher overall thioacidolysis yield, while displaying similar levels of total cell wall lignin.

Poplar Plants Incorporating High Levels of Sinapyl Alcohol

The characterization of the F5H-deficient *fah1* mutant of *Arabidopsis* demonstrated that angiosperms can use coniferyl alcohol as a sole precursor to synthesize adequate G-rich lignins reminiscent of those in gymnosperms (Meyer et al., 1998). In addition, complementation of the *fah1* mutant with an F5H-overexpression construct clearly illustrated the fundamental role of this hydroxylase in controlling the partitioning between the two major lignin monomers; such plants had an extremely low G component. It was initially assumed that the S content was compensatorily high, as confirmed by reported S contents of up to 92% (S/G > 11). However, it was not recognized until later that a second monomer, 5-hydroxyconiferyl alcohol, a monomer resulting from incomplete monolignol biosynthesis, was also directed to lignin biosynthesis in those plants (Ralph et al., 2001a). In *Arabidopsis*, this implies that methylation of the direct product of F5H, 5-hydroxyconiferaldehyde, by COMT is unable to keep pace with its production, resulting in the export of 5-hydroxyconiferyl alcohol monomers to the lignifying cell wall and ultimately 5-hydroxyguaiacyl units being incorporated into the lignin macromolecules. Thus, the actual S content was lower than that implied by the initial S:G ratio determined. The strategy in *Arabidopsis* was therefore successful at diverting biosynthesis beyond G units to generate a mixed pool of 5-hydroxyconiferyl and sinapyl alcohol monomers, implying that COMT may be limiting. More recently it has been shown that poplar COMTs are not limiting (Li et al., 2003). Only trace amounts of 5-hydroxyguaiacyl products can be found in this evaluation. The C4H::F5H overexpression herein demonstrates that the lignins are markedly divergent from those in wild-type trees but that essentially normal growth and development patterns are nevertheless retained.

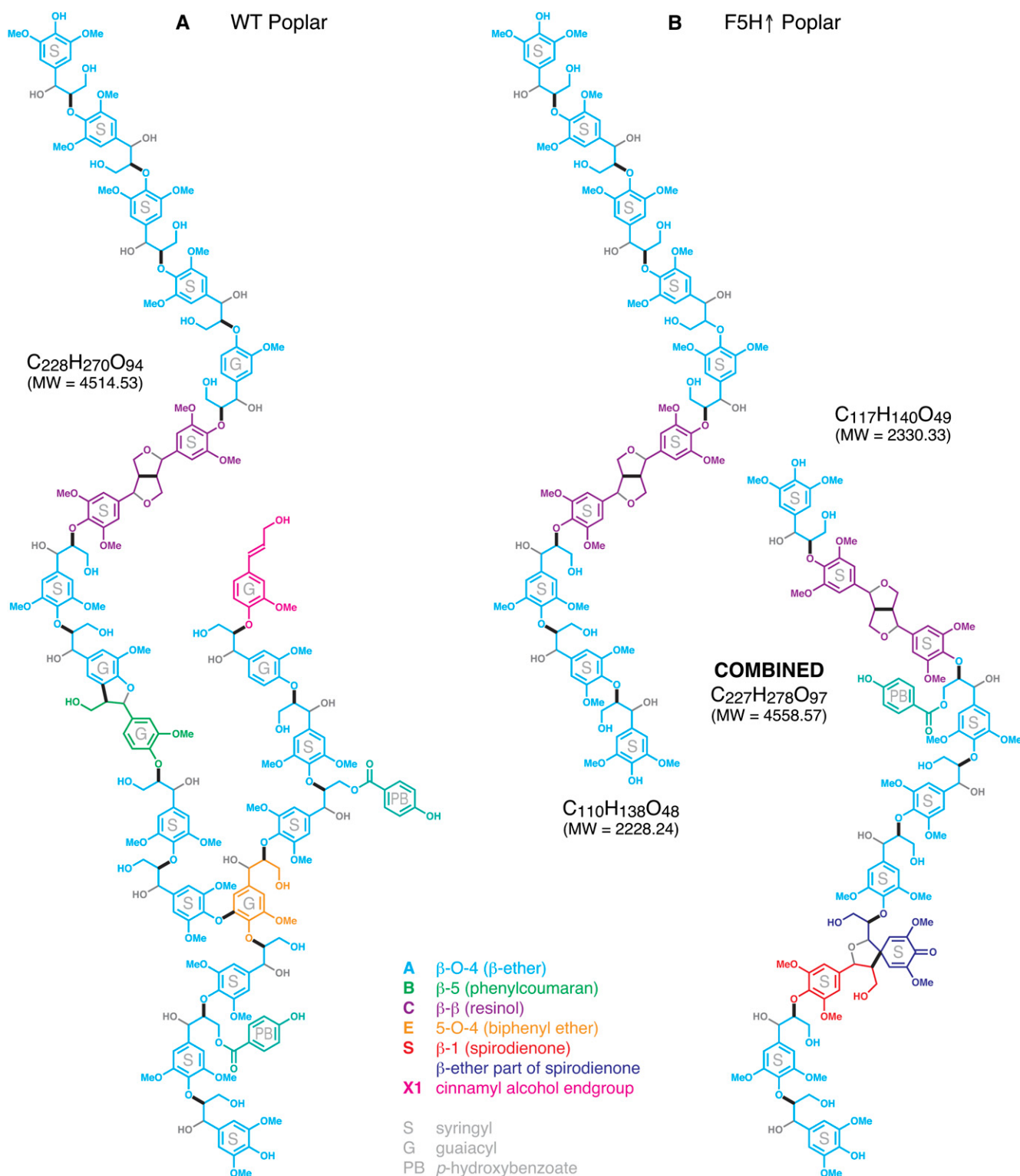


Figure 5. Lignin polymer models for wild-type (WT) poplar lignin (A) and the high-S lignin (B) synthesized from the C4H::F5H-up-regulated transgenic trees, illustrating notably its lower M_n (MW), entirely linear structure with unbranched all-S chains, with a higher resinol **C** and spirodienone **S** content, but with lower *p*-hydroxybenzoate (PB) substitution. The wild-type lignin, with 20 units, is based on prior models (Boerjan et al., 2003; Ralph et al., 2004, 2007) but conforms as closely as possible to the data reported herein. Color coding is uniform across the two models. This way it is readily seen, for example, that β -ether units **A** (cyan) are prevalent in both lignins. For the most part, the individual monomer-derived unit is colored based on the type of unit resulting from the radical coupling step, but structures **E** and **S** represent special cases. In **E**, we colored the unit that is involved in

Lignin Composition and Structure

The anticipated effect of C4H::F5H up-regulation, an enhancement of the relative level of S units in the lignin, is clearly demonstrated in the aromatic profiles revealed in NMR spectra (Fig. 1; Tables I and II). Analysis of two isolated lignin fractions (ML and EL; Fig. 1; Table II) suggests that neither partitioning of S/G units nor of the various interunit linkage types is a serious issue. The C4H::F5H lignin (Fig. 1B) was basically G depleted and strikingly S rich, estimated to be 97.5% S by NMR (Table II). As illustrated in Table II, the NMR-derived S:G ratios are consistent with those ratios derived from wet-chemical degradative methods, methods that have been used for decades for estimating lignin composition. Collectively, these analyses indicate that the C4H::F5H overexpression influences the degree of component homogeneity and, based on the observed extreme S/G ratios, that many chains are all S and therefore linear. In contrast, G units are required in lignin to facilitate branching, as G units are only methoxylated at the 3 position of the aromatic ring, leaving the 5 position free to participate in branching reactions. These extreme S lignins are therefore structurally and physically different than any previously studied (Fig. 5).

The shift in relative concentrations of *p*-hydroxybenzoate is a key new observation. In comparison to the wild-type poplar lignin, the C4H::F5H lignin displayed a >2-fold reduction in *p*-hydroxybenzoate incorporation. These findings were not anticipated, as it has recently been established that *p*-hydroxybenzoate residues in palms and poplars are found acylating S units (J. Ralph, unpublished data). Like their acetate (Lu and Ralph, 2002) and *p*-coumarate (Ralph et al., 1994; Grabber et al., 1996; Lu and Ralph, 1999) analogs, they arise from lignification via preformed monolignol esters (i.e. sinapyl *p*-hydroxybenzoate; Meyermans et al., 2000; Lu et al., 2004; Lu and Ralph, 2005). These *p*-hydroxybenzoate appendages acylating monolignols are essentially exempt from the radical coupling reactions that ultimately give rise to the lignin macromolecule (Hatfield et al., 2008). For this reason, their levels are expressed as a proportion of the S+G units (and not as a proportion of the total aromatics). In previous studies (Jouanin et al., 2000; Meyermans et al., 2000) in which lignin content was observed to be reduced by down-regulation of the genes encoding for 4-hydroxycinnamic acid:CoA

ligase or caffeoyl-CoA O-methyltransferase, the *p*-hydroxybenzoate contents increased in direct proportion; when the amount of lignin was lowered, the proportion of the lignin acylated with *p*-hydroxybenzoate was proportionally higher. Although there was no obvious explanation, the production of the acylated monolignol (sinapyl *p*-hydroxybenzoate) appeared to be uncoupled from the production of the unacylated monomers, the conventional monolignols (coniferyl and sinapyl alcohols). In the C4H::F5H transgenic trees described herein, this does not appear to be the case. Although *p*-hydroxybenzoates appear to be associated with S units in poplar trees, increasing the proportion of S units was not accompanied by a commensurate increase in *p*-hydroxybenzoate levels in the transgenic lignin, as might be expected.

The HSQC spectra of the side chain region are illustrative of the manner in which the monomeric units are assembled into the core lignin polymer. Characterized by a variety of interunit bonds, the most prominent in lignin being labeled A-C, S, and X1 (Fig. 2) to be consistent with standard assignment labeling (Ralph et al., 1999; Boerjan et al., 2003; Ralph et al., 2004). This linkage distribution differs substantially between the wild-type and transgenic C4H::F5H poplar lignins, as is visually apparent in the spectra (Fig. 2) and quantitatively (Table II). Semiquantitative evaluation demonstrates an almost complete disappearance of phenylcoumaran (β -5) units **B**, which are seen only when spectra are viewed at lower contour levels (data not shown, but quantified in Table II). Negligible levels of phenylcoumarans **B** are a logical consequence of lignification with only minor proportions of coniferyl alcohol; only coupling of a monolignol (either coniferyl or sinapyl alcohol) with a G unit can result in β -5-coupled products. Similarly, lignification with diminished coniferyl alcohol proportions is also responsible for the lower levels of cinnamyl alcohol end groups **X1**. Such end groups result from monomer-monomer β -5 or β -O-4 coupling reactions. Dimerization of coniferyl alcohol results in comparable levels of the three dimers from β - β , β -5, or β -O-4 coupling, whereas sinapyl alcohol dimerization predominantly results in syringaresinol, and β - β coupling produces the sole dimer that does not possess cinnamyl alcohol end groups. The reduced levels of phenylcoumarans **B** (from approximately 2% down to approximately 0.1% in the transgenic poplar) and

Figure 5. (Continued.)

both 5-O-4 and β -O-4 coupling in orange; in **S**, the monolignol unit that underwent β -1 coupling is in red, whereas the β -ether unit on the polymer with which it coupled is shown in dark blue to differentiate the resultant spirodienone **S** unit from a normal β -ether unit **A** (cyan). Bold black bonds indicate the bonds formed by radical coupling during lignification; lighter (gray) bonds result from postcoupling internal rearomatization reactions; analogously, atoms added postcoupling (e.g. α -OH groups from nucleophilically added water) are shown in gray. Note that each of these structures represents only one of many isomers (Ralph et al., 2004). These are only models; they do not imply any primary structure or sequencing in the lignins themselves, but depict the general features of the two lignins and attempt to accommodate the main linkage types and their approximate relative frequencies.

cinnamyl alcohol end groups **X1** is compensated by increased relative proportions of resinols **C** and spirodienones **S**. These findings concur with previous observations (Zhang and Gellerstedt, 2001), where it has been noted that high-S lignin plants (e.g. kenaf) are richer in both structures.

Quantification of linkages in the lignins from the transgenic trees appears to indicate a higher proportion of β -ether units **A** (approximately 87% versus approximately 83% of the units quantified). However, a correction for the β -ether units that have become oxidized (i.e. units **A3**; Fig. 2) is required. **S** units are more amenable to such oxidation (Ralph et al., 2004; Bunzel and Ralph, 2006), probably during lignification, so **A3** units are consequently severalfold higher in the transgenic trees (Table II). When both forms of the β -ethers are summed, the levels of β -ether units appear to be similar, approximately 87.5% versus 86%. A similar β -ether level complicates the explanation for the higher thioacidolysis monomer yield. Initially it was assumed that the higher thioacidolysis yields logically arose as a consequence of the high-S lignin polymer being inherently rich in cleavable β -ethers. The NMR data do not support this contention. The polymer derived essentially only from sinapyl alcohol is necessarily more linear; any branching (involving **G** units) will lower the yield. The lower degree of polymerization and the lower *p*-hydroxybenzoate levels are therefore both responsible for increased thioacidolysis monomer yields from the transgenic.

Monolignols invariably couple at their β -positions. Coupling of a monolignol, either coniferyl or sinapyl alcohol, with a **S** unit in the growing polymer can essentially only proceed via β -O-4 coupling (whereas coupling with a **G** unit may additionally be via β -5 coupling). **S**-rich lignins therefore have higher β -ether contents. However, at the extreme **S** levels observed in these C4H::F5H transgenic poplars, smaller effects and less prevalent coupling modes must be taken into account, along with the issue of monomer dimerization. First, the normally neglected β -1 coupling mode plays a significant role. Coupling of a monomer, at its favored β -position, with a preformed β -ether unit, at its 1 position (Zhang and Gellerstedt, 2001; Ralph et al., 2004; Zhang et al., 2006), results in the formation of spirodienones (Fig. 2). As might be anticipated, high levels of β -ethers in high-S lignins relatively favor β -1 coupling. The levels here are tabulated as doubling, from approximately 1.6% in the wild-type EL to approximately 3.5% of the linkages quantified in the C4H::F5H-overexpressing poplar EL.

A second factor is the increased contribution of dimerization to the polymer in high-S lignins. The stability of sinapyl alcohol radicals (compared with their coniferyl alcohol analogs) allows them to be more selective in their coupling and cross-coupling reactions. Thus, whereas dimerization reactions represent only a few percent of the lignin in **G**-only softwood lignins, they comprise a more substantial component in angiosperms. The estimation derives as follows: dimerization pro-

duces one of two NMR-recognizable structural entities, namely, resinols **C** (from β - β -dehydrodimerization) or an unsaturated cinnamyl alcohol side chain **X1** (from β -5- or β -O-4-dehydrodimerization). In gymnosperms, the contents of both are low, totaling just a few percent of the total cell wall lignin. However, in angiosperms, the resinol content is significantly higher. For example, of the units quantified in the ELs herein (Table II), resinols **C** comprise some 6.2% and 8.5% of the linkages in the wild-type and C4H::F5H transgenics, respectively. Resinols **C** plus end groups **X1** comprise approximately 8.5% and 10.3% of all units.

The comparable β -ether levels suggest that the observed pulping performance enhancement of the C4H::F5H-overexpressing poplars (Huntley et al., 2003) is due to factors other than higher β -ether content. While it is recognized that a **S** polymer chain is restricted to linear conformation and can only contain a single resinol unit, it becomes apparent that the high-S lignins, particularly in the extreme transgenics, have a lower average M_r ; a 10% resinol content suggests average linear-chain degrees of polymerization only on the order of 10 monomer units. An evaluation of the M_r distributions supports this contention, showing that the **S**-rich C4H::F5H lignin fractions are of lower M_r (Fig. 4).

Branching of the lignin polymer is a factor in determining both the degree of polymerization and M_r of the polymer. In polymers synthesized from coniferyl and sinapyl alcohol, only two types of branch points, involving 4-O-5- or 5-5-linkages have been observed. Both require at least one **G** unit. The 4-O-5 units are difficult to quantify and may not be as prevalent as previously considered (Ralph, 2005). Even in the control lignin, the amount of 5-5 coupling appears to be extremely low. Dibenzodioxocin structures (the major 5-5 structures, resulting from monolignol addition to 5-5-linked phenolic end groups) are only evident in the HSQC spectra when examined at very low contour levels (data not shown); they are estimated to be only 0.3% of the β -ethers **A** in the control but are readily apparent in other lignins with a similar **G** content (Ralph et al., 2006). And, the extremely low **G** levels in the transgenic lignin assure that none are detectable. Thus, branching is logically very minor in the C4H::F5H transgenic.

Coupling and Cross-Coupling Propensities of Coniferyl and Sinapyl Alcohols in **S**-Rich Environments

Details regarding lignification via enhanced levels of sinapyl alcohol come from analysis of the HMBC spectra. These spectra (Fig. 3) allow a determination of which monomers are involved in forming each type of interunit linkage (Ralph et al., 1999). In the wild-type lignin, coniferyl alcohol is deduced to react producing β -ether **A** and phenylcoumaran **B** units. All of the *p*-hydroxybenzoylated β -ether units **A'** are **S**, suggesting derivation largely/exclusively from sinapyl *p*-hydroxybenzoate. Even when there is a substantial **G** com-

ponent, most of the resinol **C** in angiosperms are derived from sinapyl alcohol. Finally, arylglycerol units **X7** appear to be all **S**.

In the C4H::F5H transgenic lignin, the levels of coniferyl alcohol monomer are clearly so low that finding traces of **G** units in these structures is difficult. However, traces can be seen in the β -ether units **A** but not those from *p*-hydroxybenzoylated β -ethers **A'**, nor resinols **C**, which remain all **S**. **H** units are authenticated in Figure 3B by the correlations observed in the β -ether units **A**. There is a previously unobserved β -ether environment, labeled "new" in Figure 3B. At present, we can only speculate on its origin, but the following seems reasonable and will ultimately be tested. The dilemma is in trying to understand what new pathway is possible in a system with essentially only a single monomer available that is not available, or at least not prevalent, in a system where both monomers are available. In the past, the major oligomers derived from syringaresinol in poplar systems have been shown to be the products of coniferyl alcohol, rather than sinapyl alcohol, addition (Morreel et al., 2004). It has further been noted (Morreel et al., 2004) in synthetic systems that, whereas coniferyl alcohol readily adds to syringaresinol, sinapyl alcohol does not readily add. We suspect, therefore, that in the wild-type trees it is coniferyl alcohol that invariably adds to syringaresinol. Only in the case where there is essentially no coniferyl alcohol available, i.e. in the C4H::F5H transgenic trees, is sinapyl alcohol found adding to syringaresinol during lignification. The so-called new β -ether correlation is therefore conjectured to be due to the normally relatively sparse syringyl-(β -O-4)-syringaresinol moieties. If this ultimately turns out to be correct, a subtle new type of structural change will be implicated, one in which the circumstances (of having no coniferyl alcohol available for coupling) result in the prevalence of a typically minor coupling pathway, i.e. the β -O-4 coupling of sinapyl alcohol with syringaresinol.

Lignin Models

The above details have been incorporated into an illustrative 20-unit model for the lignin polymers in the wild-type versus transgenic samples (Fig. 5).

The main design points include (1) an approximately 2:1 **S**:**G** ratio in the wild type but almost solely **S** units in C4H::F5H lignin and (2) double the resinols **C** in C4H::F5H. It is not possible to have two resinols in a polymer chain (Ralph et al., 2008a). With >10% resinol units **C**, and with no chain branching by **G** units, two decamers are therefore shown for the C4H::F5H lignin. In the wild type, the resinol is underrepresented by the single unit. The resinols are **S**-**S** as most are in hardwoods. (3) One 5-O-4 branchpoint, generally considered to account for up to 5% of units in hardwoods (Adler, 1977) in the wild-type lignin; without **G** units, 5-O-4 units cannot occur (at significant levels) in the C4H::F5H material. (4) Two *p*-hydroxybenzoates (i.e.

about 10%), both on **S** units, in the wild type, with only one in the C4H::F5H. (5) One phenylcoumaran (**B**, β -5 unit) in the wild type and none in the C4H::F5H (since there are no **G** units). (6) Spirodienones (**S**, β -1 units); the level is too low to include one in the wild type, but the elevated levels in the C4H::F5H are reflected by the inclusion of one in the C4H::F5H. (7) One cinnamyl alcohol end group is represented in the wild type; they are usually **G**, although more of them are probably on phenylcoumaran (β -5) units. (8) The β -ether frequency is approximately equal in both models; most of the units arise from β -O-4 coupling reactions. (9) No dibenzodioxocins **D** (data not shown) are depicted; they require **G** units for their formation but are not seen even in the wild-type poplar lignin here. Other details are covered in the caption.

CONCLUSION

Poplar lignins extremely rich in **S** units are produced by C4H::F5H overexpression. In fact, the plants analyzed herein appear to have the highest **S** content on record, with some 97.5% of the lignin deriving from sinapyl alcohol. The **S**/**G** of approximately 38 suggests that the lignin chains are predominantly linear, and this is supported by comparative M_r determination of isolated lignin fractions. As might be anticipated from the substantial compositional shifts documented between the control and the transgenic trees, the structure of the polymer is quite drastically altered. Whereas the β -ether **A** remains predominant, phenylcoumarans **B** become rare (at high **S** levels) and spirodienones **S** increase. An unexpected finding was that β -ether **A** levels were not higher at the extreme **S** levels (compared the native **S** levels common to wild-type trees). The contribution from β - β -units (resinols **C**) was compensatorily higher, suggesting more dehydrodimerization that consequently results in shorter chain lengths in the modified lignin, as confirmed by M_r analysis. The compositional and structural changes in the polymer noted here remain consistent with the existing theory of lignification based on combinatorial radical coupling reactions under simple chemical control.

MATERIALS AND METHODS

Plant Materials

The generation of nine lines of C4H::F5H transgenic hybrid poplar (*Populus tremula* $\times$ *Populus alba*) was previously described (Meyer et al., 1998; Franke et al., 2000; Huntley et al., 2003). Both wild-type and the transgenic lines having the highest **S** content as determined by thioacidolysis (with 93.4% of its releasable monomers being **S**) were selected based on previous studies (Huntley et al., 2003) as the substrates for detailed analysis herein. Stems were harvested from 2.5-year-old poplar wild-type and transgenic trees, manually debarked, and the pith removed by scraping with a sterile razor blade. Lignins were isolated using methods largely described previously (Ralph et al., 2001a).

Cell Wall Preparation, Grinding, and Ball Milling

Stems were ground in a cyclone mill to pass through a 1-mm screen and extensively Soxhlet extracted sequentially with methanol and acetone. The isolated cell walls, 2 g at a time, were ball milled for 3.0 h (in 20 min on/10 min off cycles to avoid excessive sample heating) using a Retsch PM100 ball mill running at 600 rpm with zirconium dioxide vessels (50 mL) containing ZrO₂ ball bearings (3 × 30 mm, 7 × 10 mm).

Crude Cellulase Digestion

The ball-milled walls (5.85 and 9.13 g for wild-type and C4H::F5H transgenic poplar, respectively) were then digested at 30°C with a crude cellulase mixture (Cellulysin; Calbiochem; 30 mg/g of sample, in pH 5.0 acetate buffer, 3 × 48 h, fresh buffer and enzyme each time) leaving the EL fractions containing all of the lignin and residual polysaccharides totaling 1.56 g (26.7% of the original cell wall, wild type) and 1.94 g (21.2%, C4H::F5H) (Table I).

Isolation of Dioxane:Water-Soluble Milled-Wood Lignins

The ball-milled lignocellulosics, 0.502 g (wild-type) and 1.503 g (C4H::F5H), were extracted twice with 96:4 dioxane:water (100 mL) overnight at room temperature. The extracts were pooled and freeze-dried. The lignins were then washed with water and filtered using nylon membrane filters (0.2 μm pore size; Nylaflow; Pall Life Science) to remove water-soluble components (mainly low *M_r* sugars) resulting in the milled-wood lignin (ML) fraction (87.9 and 715 mg, respectively). The residual, insoluble lignin-rich material remaining following dioxane extraction and enzyme treatment is referred to as RL.

Milled-Wood Lignin Acetylation

The milled-wood fraction was acetylated overnight using 1:1 acetic anhydride:pyridine. The solvents/reagents were removed by repeated coevaporation with ethanol at reduced pressure on a rotary evaporator.

EL Dissolution

In order to characterize as much of the lignin fraction as possible, the entire cellulase-digested cell wall fraction was subjected to solubilization in DMSO/*N*-methylimidazole (NMI), a solvent shown to dissolve the whole cell wall fraction of ball-milled woods and other plants (Lu and Ralph, 2003). Dissolution of these poplar samples was essentially complete. The cellulase-digested cell walls, 101 and 122 mg from wild-type and C4H::F5H transgenics, respectively, yielded 105 and 153 mg of acetylated sample.

Cell Wall Analytical Methods

Extractive-free ground wood was prepared by grinding representative samples of wood in a Wiley mill to pass through a 40-mesh sieve, and Soxhlet extracted overnight with hot acetone. The extract free wood was then dried over P₂O₅ and retained for complete chemical analysis.

Klason Lignin

Lignin content was determined using a modified Klason procedure, where extracted ground stem tissue (0.2 g) was treated with 3 mL of 72% H₂SO₄ as per Huntley et al. (2003). The composition of neutral cell wall-associated carbohydrates (Ara, rhamnose, Gal, Glc, Man, and Xyl) was determined using HPLC (Dionex DX-600) equipped with an ion-exchange PA1 (Dionex) column, a pulsed amperometric detector with a gold electrode, and a Spectra AS 3500 auto-injector (Spectra-Physics). Each experiment was run in duplicate.

Thioacidolysis

Thioacidolysis monomer yields and S:G ratios were determined by thioacidolysis (Rolando et al., 1992) using 10 mg samples of extractive-free wood. Tetracosane (2 μL of 0.25 mg/mL in CH₂Cl₂) was used as the internal standard. The silylation reaction proceeded for a minimum of 1.5 h. Gas

chromatographic analyses were performed on an HP 5890 Series II instrument fitted with a 15 m × 0.25 mm DB-5 column (J&W Scientific) and a flame ionization detector. The gas chromatography method used a 2.0-μL injection volume, an initial injector temperature of 250°C, and a detector temperature of 270°C. The initial oven temperature was set to 130°C (held for 3 min) and thereafter ramped at a rate of 3°C/min to 260°C and held for 5 min.

The DFRC Method

Release and quantification of acetylated monolignols by reductive cleavage of β-aryl ethers was performed as described (Lu and Ralph, 1997; Lu and Ralph, 1999).

Methoxyl Analysis

Methoxyl content was determined by a modified Ziesel method (Browning, 1967).

Hydroxyl Content

Aliphatic and phenolic hydroxyl contents were determined via their acetates by proton NMR spectroscopy. *p*-Nitrobenzaldehyde was used as the internal standard (stock solution of 1 mg/mL in CDCl₃). Acetylated milled-wood lignin (Ac-ML) was dissolved into this solution at a concentration of 30 mg/mL of internal standard solution. Standard one-dimensional proton NMR spectra (256 scans) were acquired on a Bruker AV-300 spectrometer. Integrated aliphatic and phenolic acetate proton signals were quantified by comparison to the aldehyde proton signal from the internal standard.

Molecular Weight Determination

Milled-wood lignin, EL, and RL samples were acetylated by dissolving milligram quantities in NMI and dimethylsulfoxide in a ratio of 1:2 and adding acetic anhydride (0.6 × NMI), as per the cell wall dissolution method (Lu and Ralph, 2003). Samples were left for 3 h, precipitated into 6 mM EDTA (pH 8), filtered through a 0.2-μm membrane filter, and dried under vacuum. Samples of 1 mg/mL in tetrahydrofuran were filtered through a 0.45-μm filter. A Dionex Summit HPLC fitted with a Waters Styragel H5 column was used to obtain the *M_r* distribution at 50°C with a 0.5 mL/min flow rate using a photodiode array detector with the wavelength set to 280 nm. The total run time was 60 min.

NMR Spectroscopy

The lignin NMR spectra presented here were acquired on a Bruker Biospin DMX-500 instrument fitted with a sensitive cryogenically cooled 5-mm TXI ¹H/¹³C/¹⁵N gradient probe with inverse geometry (proton coils closest to the sample). Acetylated lignin preparations (60–80 mg) were dissolved in 0.5 mL CDCl₃; the central chloroform solvent peak was used as internal reference (δ_C 77.0, δ_H 7.26 ppm). We used the standard Bruker implementations of the traditional suite of one-dimensional and 2D (gradient-selected, ¹H-detected; e.g. COSY, TOCSY, HSQC, HSQC-TOCSY, and HMBC) NMR experiments for structural elucidation and assignment authentication (Ralph et al., 1999). NMR data for model compounds can be found in the NMR database of lignin and cell wall model compounds (Ralph et al., 2005). Normal HSQC experiments typically had the following parameters: acquired from 8.6 to 2.4 ppm in F₂ (¹H) using 1864 data points (acquisition time of 200 ms), 160 to 40 ppm in F₁ (¹³C) using 512 increments (F₁ acquisition time of 11.3 ms) of 16 or 32 scans with a 1-s interscan delay, total acquisition time of 2 h and 48 min, or 5 h and 34 min; the d₂₄ delay was set to 1.72 ms (approximately 1/4 J). Processing used typical matched Gaussian apodization in F₂ and squared sine-bell in F₁. HMBC experiments had the following parameters: acquired from 8.6 to 2.4 ppm in F₂ (¹H) using 3,000 data points (acquisition time of 329 ms), 180 to 40 ppm in F₁ (¹³C) using 400 increments (F₁ acquisition time of 7.6 ms) of up to 112 scans with a 1-s interscan delay, 80-ms long-range coupling delay, and total acquisition time of up to 18 h. Processing to a final matrix of 2,000 by 1,000 data points used typical matched Gaussian apodization in F₂ (LB –80; GB 0.338) and squared sine-bell in F₁. One level of linear prediction in F₁ (32 coefficients) gave improved F₁ resolution but was not required.

Volume integration of contours in HSQC plots was accomplished using Bruker's TopSpin 1.3 software as described (Ralph et al., 2006). For quanti-

fication of S/G distributions, only the carbon-2 correlations from G units and the carbon-2/6 correlations from S units were used, and the G integrals were logically doubled. No correction factors were necessary. A correction factor was deemed necessary for *p*-hydroxybenzoate quantification. The response factor of 1.5 (for the 3/5-correlation peak) was determined from two model compounds, acetylated sinapyl *p*-hydroxybenzoate, and the γ -*p*-hydroxybenzoate ester of a syringyl- β -O-4-syringyl model (acetylated). For quantification of the interunit linkage types, the following well-resolved contours (see Fig. 2) were integrated: A α , A3 α , B α , C α , S α , and X1 γ . Integral correction factors determined previously (Ralph et al., 2006) were used: A α 1.00, B α 0.71, and C α 1.06; A3 α , S α , and X1 γ were either not, or not reliably, determined and are assumed as 1.00. These values were used to correct the volume integrals to provide the semiquantitative estimates of unit ratios in Table II.

Statistical Analysis

All analysis was carried out using unpaired two-tailed *t* tests at 95% confidence.

Sequence data from this article can be found in the National Center for Biotechnology Information data library under number 829779.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. A, Quantitative ^{13}C NMR spectra of nonacetylated ML. Expanded region is shown in Supplemental Figure S3B. B, Quantitative ^{13}C NMR spectra of nonacetylated ML in the region from 157 to 140 ppm.

Supplemental Figure S2. Quantitative ^{13}C NMR spectrum of acetylated ML.

Supplemental Figure S3. Quantitative ^{13}C NMR spectra of acetylated RL.

Supplemental Figure S4. Lignin substructures.

Supplemental Table S1. Signal assignment in the NMR spectrum of nonacetylated ML. Peaks correspond to substructures in Supplemental Figure S4.

Supplemental Table S2. Signal assignment in the NMR spectrum of acetylated ML. Peaks correspond to substructures in Supplemental Figure S4.

Supplemental Table S3. Lignin substructures determined by NMR. Substructures are shown in Supplemental Figure S4.

ACKNOWLEDGMENTS

The authors gratefully acknowledge funding from the Natural Sciences and Engineering Research Council of Canada Discovery Program held by S.D.M., the Department of Energy Biosciences program (DE-AI02-00ER15067) and the Department of Energy Great Lakes Bioenergy Research Center (DE-FC02-07ER64494) to J.R., and the Office of Science (BER), U.S. Department of Energy (grant no. DE-FG02-06ER64301), to C.C. NMR experiments on the Bruker DMX-500 cryoprobe system were carried out at the National Magnetic Resonance Facility at Madison with support from the National Institutes of Health Biomedical Technology Program (RR02301) and additional equipment funding from the University of Wisconsin, National Science Foundation Academic Infrastructure Program (BIR-9214394), National Institutes of Health Shared Instrumentation Program (RR02781 and RR08438), National Science Foundation Biological Instrumentation Program (DMB-8415048), and the U.S. Department of Agriculture.

Received February 17, 2009; accepted April 20, 2009; published April 22, 2009.

LITERATURE CITED

- Adler E (1977) Lignin chemistry: past, present and future. *Wood Sci Technol* **11**: 169–218
- Baucher M, Halpin C, Petit-Conil M, Boerjan W (2003) Lignin: genetic engineering and impact on pulping. *Crit Rev Biochem Mol Biol* **38**: 305–350
- Baucher M, Monties B, Van Montagu M, Boerjan W (1998) Biosynthesis and genetic engineering of lignin. *Crit Rev Plant Sci* **17**: 125–197
- Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. *Annu Rev Plant Biol* **54**: 519–549
- Browning BL (1967) *Methods of Wood Chemistry*. Wiley-Interscience, New York
- Bunzel M, Ralph J (2006) NMR characterization of lignins isolated from fruit and vegetable insoluble dietary fiber. *J Agric Food Chem* **54**: 8352–8361
- Chang HM, Sarkanen KV (1973) Species variation in lignin: effect of species on the rate of Kraft delignification. *Tappi J* **56**: 132–136
- Chen F, Dixon RA (2007) Lignin modification improves fermentable sugar yields for biofuel production. *Nat Biotechnol* **25**: 759–761
- Franke R, Humphreys JM, Hemm MR, Denault JW, Ruegger MO, Cusumano JC, Chapple C (2002) The Arabidopsis REF8 gene encodes the 3-hydroxylase of phenylpropanoid metabolism. *Plant J* **30**: 33–45
- Franke R, McMichael CM, Meyer K, Shirley AM, Cusumano JC, Chapple C (2000) Modified lignin in tobacco and poplar plants over-expressing the Arabidopsis gene encoding ferulate 5-hydroxylase. *Plant J* **22**: 223–234
- Grabber JH, Quideau S, Ralph J (1996) *p*-Coumaroylated syringyl units in maize lignin; implications for β -ether cleavage by thioacidolysis. *Phytochemistry* **43**: 1189–1194
- Hatfield RD, Ralph J, Grabber JH (2008) A potential role of sinapyl *p*-coumarate as a radical transfer mechanism in grass lignin formation. *Planta* **228**: 919–928
- Hu WJ, Lung J, Harding SA, Popko JL, Ralph J, Stokke DD, Tsai CJ, Chiang VL (1999) Repression of lignin biosynthesis in transgenic trees promotes cellulose accumulation and growth. *Nat Biotechnol* **17**: 808–812
- Humphreys JM, Hemm MR, Chapple C (1999) Ferulate 5-hydroxylase from Arabidopsis is a multifunctional cytochrome P450-dependent monooxygenase catalyzing parallel hydroxylations in phenylpropanoid metabolism. *Proc Natl Acad Sci USA* **96**: 10045–10050
- Huntley SK, Ellis D, Gilbert M, Chapple C, Mansfield SD (2003) Significant increases in pulping efficiency in C4H:F5H-transformed poplars: improved chemical savings and reduced environmental toxins. *J Agric Food Chem* **51**: 6178–6183
- Jouanin L, Goujon T, de Nadaï V, Martin MT, Mila I, Vallet C, Pollet B, Yoshinaga A, Chabbert B, Petit-Conil M, et al (2000) Lignification in transgenic poplars with extremely reduced caffeic acid O-methyltransferase activity. *Plant Physiol* **123**: 1363–1373
- Karhunen P, Rummakko P, Sipilä J, Brunow G, Kilpeläinen I (1995) Dibenzo[dioxocins; a novel type of linkage in softwood lignins. *Tetrahedron Lett* **36**: 169–170
- Li L, Popko JL, Umezawa T, Chiang VL (2000) 5-Hydroxyconiferyl aldehyde modulates enzymatic methylation for syringyl monolignol formation, a new view of monolignol biosynthesis in angiosperms. *J Biol Chem* **275**: 6537–6545
- Li L, Zhou Y, Cheng X, Sun J, Marita JM, Ralph J, Chiang VL (2003) Combinatorial modification of multiple lignin traits in trees through multigene cotransformation. *Proc Natl Acad Sci USA* **100**: 4939–4944
- Lu F, Ralph J (1997) The DFRC method for lignin analysis. Part 1. A new method for *p*-aryl ether cleavage: lignin model studies. *J Agric Food Chem* **45**: 4655–4660
- Lu F, Ralph J (1999) Detection and determination of *p*-coumaroylated units in lignins. *J Agric Food Chem* **47**: 1988–1992
- Lu F, Ralph J (2002) Preliminary evidence for sinapyl acetate as a lignin monomer in kenaf. *Chem Commun (Camb)* **1**: 90–91
- Lu F, Ralph J (2003) Non-degradative dissolution and acetylation of ball-milled plant cell walls: high-resolution solution-state NMR. *Plant J* **35**: 535–544
- Lu F, Ralph J (2005) Novel β - β -structures in natural lignins incorporating acylated monolignols. In *Thirteenth International Symposium on Wood, Fiber, and Pulping Chemistry*. APPITA, Auckland, Australia, pp 233–237
- Lu F, Ralph J, Morreel K, Messens E, Boerjan W (2004) Preparation and relevance of a cross-coupling product between sinapyl alcohol and sinapyl *p*-hydroxybenzoate. *Org Biomol Chem* **2**: 2888–2890
- Marita J, Ralph J, Hatfield RD, Chapple C (1999) NMR characterization of

- lignins in *Arabidopsis* altered in the activity of ferulate-5-hydroxylase. *Proc Natl Acad Sci USA* **96**: 12328–12332
- Marita JM, Ralph J, Hatfield RD, Guo D, Chen F, Dixon RA (2003) Structural and compositional modifications in lignin of transgenic alfalfa down-regulated in caffeic acid 3-O-methyltransferase and caffeoyl coenzyme A 3-O-methyltransferase. *Phytochemistry* **62**: 53–65
- Marita JM, Ralph J, Lapierre C, Jouanin L, Boerjan W (2001) NMR characterization of lignins from transgenic poplars with suppressed caffeic acid O-methyltransferase activity. *J Chem Soc Perkin Trans 1*: 2939–2945
- Meyer K, Shirley AM, Cusumano JC, Bell-Lelong DA, Chapple C (1998) Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in *Arabidopsis*. *Proc Natl Acad Sci USA* **95**: 6619–6623
- Meyermans H, Morreel K, Lapierre C, Pollet B, De Bruyn A, Busson R, Herdewijn P, Devreese B, Van Beeumen J, Marita JM, et al (2000) Modifications in lignin and accumulation of phenolic glucosides in poplar xylem upon down-regulation of caffeoyl-coenzyme A O-methyltransferase, an enzyme involved in lignin biosynthesis. *J Biol Chem* **275**: 36899–36909
- Morreel K, Ralph J, Kim H, Lu F, Goeminne G, Ralph SA, Messens E, Boerjan W (2004) Profiling of oligolignols reveals monolignol coupling conditions in lignifying poplar xylem. *Plant Physiol* **136**: 3537–3549
- Morrison WH, Akin DE, Archibald DD, Dodd RB, Raymer PL (1999) Chemical and instrumental characterization of maturing kenaf core and bast. *Ind Crops Prod* **10**: 21–34
- Osakabe K, Tsao CC, Li L, Popko JL, Umezawa T, Carraway DT, Smeltzer RH, Joshi CP, Chiang VL (1999) Coniferyl aldehyde 5-hydroxylation and methylation direct syringyl lignin biosynthesis in angiosperms. *Proc Natl Acad Sci USA* **96**: 8955–8960
- Ralph J (1996) An unusual lignin from Kenaf. *J Nat Prod* **59**: 341–342
- Ralph J, Akiyama T, Kim H, Lu F, Schatz PF, Marita JM, Ralph SA, Reddy MSS, Chen F, Dixon RA (2006) Effects of coumarate-3-hydroxylase downregulation on lignin structure. *J Biol Chem* **281**: 8843–8853
- Ralph J, Brunow G, Boerjan W (2007) Lignins. In F Rose, K Osborne, eds, *Encyclopedia of Life Sciences*. John Wiley & Sons, Chichester, UK, doi/10.1002/9780470015902.a0020104
- Ralph J, Brunow G, Harris PJ, Dixon RA, Schatz PF, Boerjan W (2008a) Lignification: Are lignins biosynthesized via simple combinatorial chemistry or via proteinaceous control and template replication? In F Daayf, A El Hadrami, L Adam, GM Ballance, eds, *Recent Advances in Polyphenol Research*. Wiley-Blackwell Publishing, Oxford, , pp 36–66
- Ralph J, Hatfield RD, Quideau S, Helm RE, Grabber JH, Jung HJG (1994) Pathway of *p*-coumaric acid incorporation into maize lignin as revealed by NMR. *J Am Chem Soc* **116**: 9448–9456
- Ralph J, Kim H, Lu F, Grabber JH, Leplé JC, Berrio-Sierra J, Mir Derikvand M, Jouanin L, Boerjan W, Lapierre C (2008b) Identification of the structure and origin of a thioacidolysis marker compound for ferulic acid incorporation into angiosperm lignins (and an indicator for cinnamoyl-CoA reductase deficiency). *Plant J* **53**: 368–379
- Ralph J, Lapierre C, Marita J, Kim H, Lu F, Hatfield RD, Ralph SA, Chapple C, Franke R, Hemm MR, et al (2001a) Elucidation of new structures in lignins of CAD- and COMT-deficient plants by NMR. *Phytochemistry* **57**: 993–1003
- Ralph J, Lapierre C, Lu F, Marita JM, Pilate G, Van Doorselaere J, Boerjan W, Jouanin L (2001b) NMR evidence for benzodioxane structures resulting from incorporation of 5-hydroxyconiferyl alcohol into lignins of O-methyl-transferase-deficient poplars. *J Agric Food Chem* **49**: 86–91
- Ralph J, Lundquist K, Brunow G, Lu F, Kim H, Schatz PF, Marita JM, Hatfield RD, Ralph SA, Christensen JH, et al (2004) Lignins: natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. *Phytochem Rev* **3**: 29–60
- Ralph J, Marita JM, Ralph SA, Hatfield RD, Lu F, Ede RM, Peng J, Quideau S, Helm RE, Grabber JH, et al (1999) Solution-state NMR of lignins. In DS Argyropoulos, T Rials, eds, *Advances in Lignocellulosics Characterization*. TAPPI Press, Atlanta, GA, pp 55–108
- Ralph SA (2005) Conundrums regarding 5-O-4-linkages in softwood lignins. In Thirteenth International Symposium on Wood, Fiber, and Pulping Chemistry. APPITA, Auckland, Australia, pp 315–319
- Ralph SA, Landucci LL, Ralph J (2005) NMR database of lignin and cell wall model compounds. <http://ars.usda.gov/Services/docs.htm?docid=10429> (previously <http://www.dfrc.ars.usda.gov/software.html>) (August 1, 2008)
- Rolando C, Monties B, Lapierre C (1992) Thioacidolysis. In CW Dence, SY Lin, eds, *Methods in Lignin Chemistry*. Springer-Verlag, Berlin, pp 334–349
- Sederoff RR, MacKay JJ, Ralph J, Hatfield RD (1999) Unexpected variation in lignin. *Curr Opin Plant Biol* **2**: 145–152
- Stewart JJ, Kadla JF, Mansfield SD (2006) The influence of lignin chemistry and ultrastructure on the pulping efficiency of clonal aspen (*Populus tremuloides* Michx.). *Holzforschung* **60**: 111–122
- Zhang L, Gellerstedt G (2001) NMR observation of a new lignin structure, a spiro-dienone. *Chem Commun (Camb)* **24**: 2744–2745
- Zhang L, Gellerstedt G, Ralph J, Lu F (2006) NMR studies on the occurrence of spirodienone structures in lignins. *J Wood Chem Technol* **26**: 65–79

The Effects on Lignin Structure of Over-Expression of Ferulate 5-Hydroxylase in Hybrid Poplar

Jaclyn J. Stewart¹, Takuya Akiyama², Clint C. Chapple³, John Ralph² and Shawn D. Mansfield¹

¹Department of Wood Science, University of British Columbia, Vancouver, BC, Canada;

²Department of Biochemistry, and Great Lakes Bioenergy Research Center, University of Wisconsin, Madison, Wisconsin, USA.; ³Department of Biochemistry, Purdue University, West Lafayette, Indiana, USA.

Address correspondence to: Shawn D. Mansfield, Department of Wood Science, University of British Columbia, 4030-2424 Main Mall, Vancouver, BC, Canada, V6T 1Z4. Tel. (604) 822-0196; Fax (604) 822-9104; E-mail: shawnman@interchange.ubc.ca

SUPPLEMENTARY MATERIAL

This supplementary data is largely to document the NMR analysis of other lignin fractions and synthetic lignin models that are mentioned in the main paper but are not crucial to it.

1D Nuclear Magnetic Resonance Spectroscopic Experiments derived from Quantitative Data.

Samples of acetylated cell wall material were prepared at a concentration of 150 mg mL⁻¹ in CDCl₃ (Cambridge Isotope Laboratories). Ten minutes of sonication was usually required to dissolve the sample. The sample was filtered through a Kimwipe plugged disposable pipette into a 5 mm NMR tube. All runs were performed at room temperature, approximately 293 K.

ML or Ac-ML was dissolved in d₆-DMSO (Cambridge Isotope Laboratories) at the approximate concentration of 60 mg in 0.25 mL. The sample was allowed to fully dissolve overnight, and stored over desiccant. The next day, the sample was filtered through a Kimwipe into a 5 mm Shigemi NMR microtube (with the magnetic susceptibility matched to DMSO, Shigemi Co.).

For quantitative ¹³C NMR acquisition, 10 mM final concentration of chromium (III) acetylacetonate, Cr(acac)₃ (Aldrich), was added to shorten the T₁ relaxation times of the carbon nuclei to allow signals to be quantified. For ML, a Bruker AM-400 spectrometer (dual ¹³C/¹H probe) was used, while for the RL, a Bruker AV-300 fitted with a QNP probe was used. A 90° ¹³C pulse, a 1.4 s acquisition time and a 1.7 s relaxation delay were the parameters for the inverse gated decoupling (decoupler on during acquisition only). Approximately 20,000 scans were collected for ML samples. The d₆-DMSO peak at 39.5 ppm or the CDCl₃ peak at 77 ppm was used as the reference.

Figure S1a. Quantitative ^{13}C NMR spectra of non-acetylated ML. Expanded region is shown in Figure S3b.

Figure S1b. Quantitative ^{13}C NMR spectra of non-acetylated ML in the region from 157 to 140 ppm.

Figure S2. Quantitative ^{13}C NMR spectrum of acetylated ML.

Figure S3. Quantitative ^{13}C NMR spectra of acetylated RL.

Figure S4. Lignin substructures. Note: these substructures are designated here for the supplementary material and are different from the substructures referred to in the main document and Supplementary figures S1 and S2 (JR, new 2D designations).

Table S1. Signal assignment in the NMR spectrum of non-acetylated ML. Peaks correspond to substructures in Figure S4.

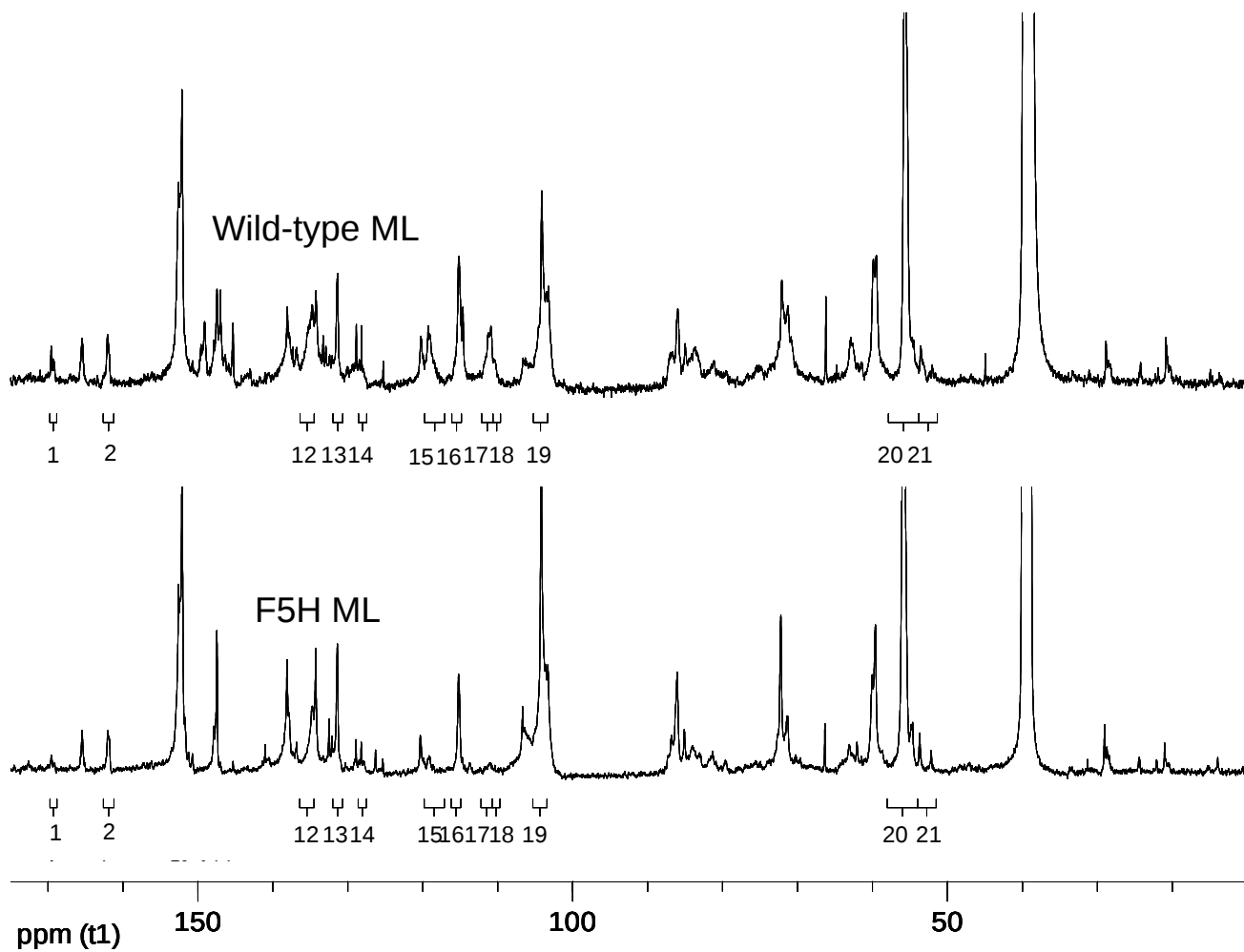
Spectral Region	Chemical Shift Range (ppm)	WT (per Ar)	C4H- F5H (per Ar)	Peak Label
aliphatic COOR	169.7-168.8	0.07	0.04	1
C-4 in H units	162.6-161.2	0.10	0.11	2
C-3 in T _{et} , C3,5 in S _{et} , C- α in L, C-3,6 in I, C-4 in conjugated S, unknown	155-151	1.01	1.19	3
C3/C3' in etherified 5-5' units, C α in Ar-CH=CH-CHO units	155.1-152.4	0.47	0.49	4
C3/C5 in B-ring of 4-O-5' units, C4 in etherified G units with α -C=O	152.4-150.9	0.50	0.65	5
C3 in etherified G	150.2-149.5	0.06	-	6
	149.5-148.5	0.13	0.02	
C3 in nonetherified G units, C4 in etherified G units	148.5-147.2	0.21	-	7
C3 in E, C-4 in conjugated S _{et}	145.6-144.9	0.07	0.02	8
C-3 in E, C-4 in T _{ne} , C-4 in conjugated S, unknown	144.5-142.5	0.07	-	9
C4/C4' in nonetherified 5-5' units, C-3 in B rings of β -5 units				
C4 in nonetherified G units, C4 in B-ring of β -ring of β -5 units	147.3-146.6	0.14	0.02	10
C4/C4' in nonetherified 5-5' units	146.6-145.7	0.09	-	11
C1 in etherified G units	136.4-134.5	0.35	0.29	12
C5/C5' in etherified 5-5' units	132.0-130.7	0.20	0.25	13
C α and C β in Ar-CH=CH-CH ₂ OH	128.6-127.5	0.08	0.06	14
C6 in G units	119.8-117.1	0.25	0.08	15
C5 in G units	116.2-114.9	0.27	0.21	16
C2 in G units	112.2-110.7	0.21	0.04	17
C2 in G, G in stilbene units	110.7-109.7	0.06	0.02	18
C2/6 in S units	105.3-103.4	0.58	1.09	19
OMe, C- β in H, C-1 in I, C- γ in R	58-54	1.50	2.35	20
C- β in E and F	54-51.5	0.16	0.12	21
Clusters				
C _{Ar-O}	160-141	2.02	1.83	
C _{Ar-C}	141-125	1.65	1.69	
C _{Ar-H}	125-103	2.13	2.26	
Alk-O-	90-58	2.84	2.78	
Alk-O-Ar, α -O-Alk	90-77	0.87	0.88	
γ -O-Alk, OH _{sec}	77-65	1.06	0.96	
OH _{prim}	65-58	0.90	0.93	

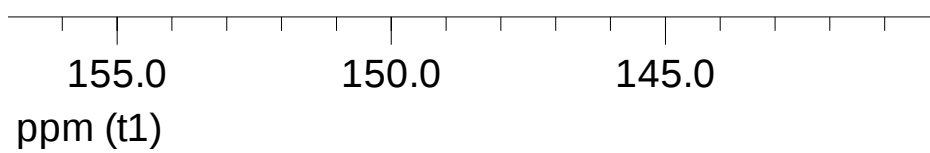
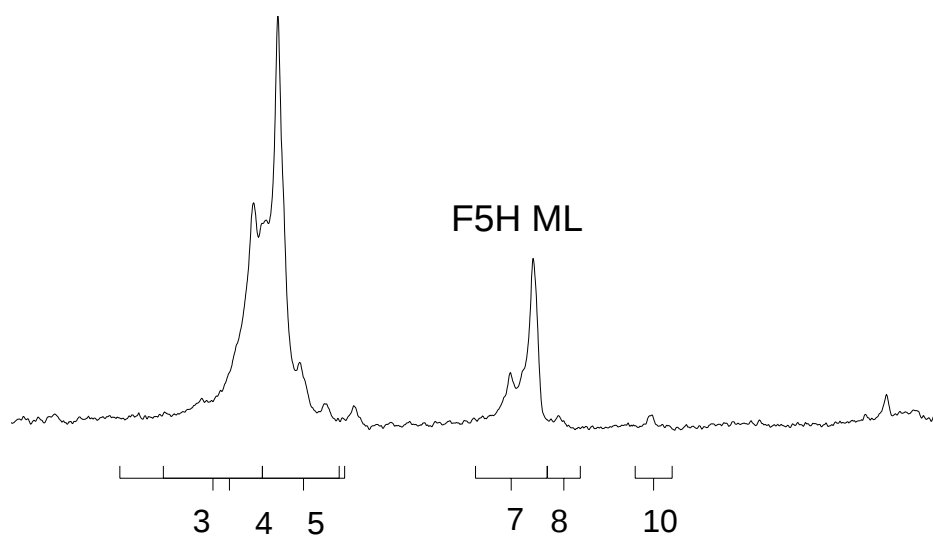
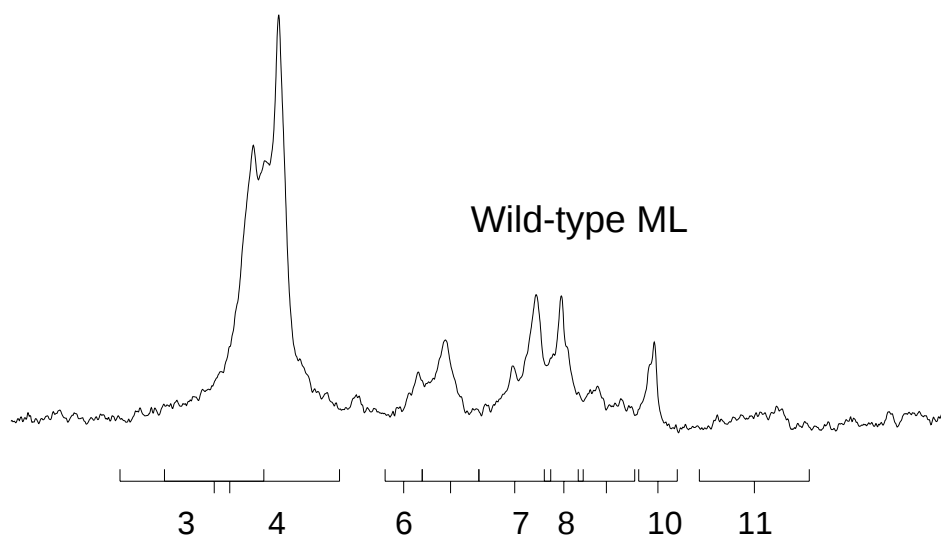
Table S2. Signal assignment in the NMR spectrum of acetylated ML. Peaks correspond to substructures in Figure S4.

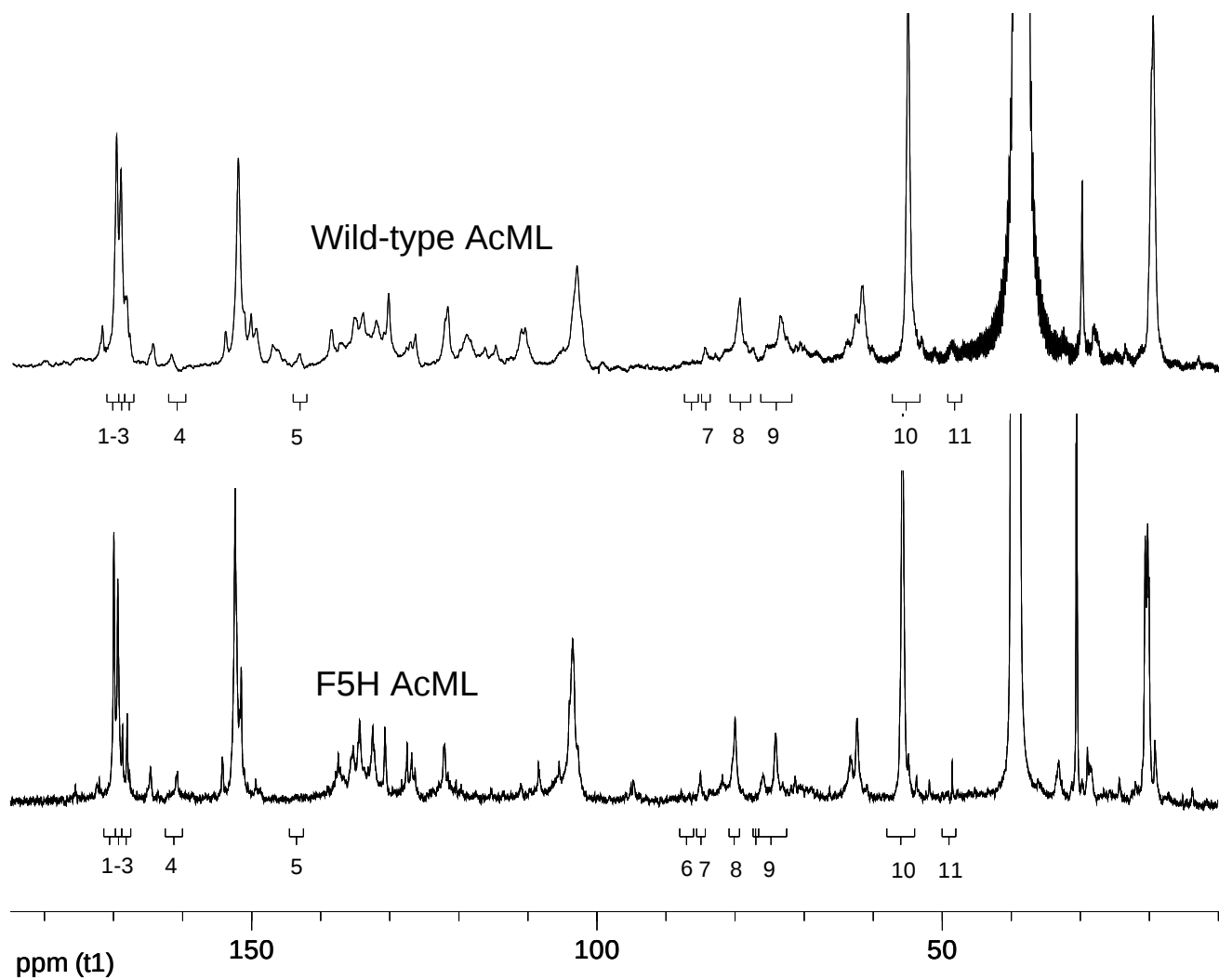
Spectral Region	Chemical Shift Range (ppm)	WT (per Ar)	C4H- F5H (per Ar)	Peak Label
non-conjugated CO (acetone impurity)	207.4-205.1	-	0.33	
primary aliphatic OH	171.4-169.7	0.66	0.45	1
secondary aliphatic OH	169.7-168.8	0.52	0.44	2
phenolic OH, conjugated COOR	168.8-167.5	0.24	0.24	3
all C-3 (except E and H-units), C-5 in S, C- α in L, C-6 in I, C-4 in conjugated CO/COOR etherified and H-units	162-148	1.53	1.70	
	162.5-160	0.04	0.14	4
C-3 in E, C-4 in conjugated S _{et} , unknown	144.5-142.5	0.06	0.06	5
C- α in E	88-86	0.05	0.06	6
C- α in F, G	85.55-84.3	0.08	0.08	7
C- β in G	80.84-79.34	0.30	0.29	8
C- α in A, H, P, carbohydrates	77-72.5	0.55	0.45	9
OMe, C-1, and C- β in I	58-54	1.44	1.27	10
C- β in H, E	50-48	0.17	0.09	11
Clusters				
C _{Ar-H}	125-103	2.07	2.01	
Alk-O-	90-58	2.86	2.30	
Alk-O-Ar, α -O-Alk	90-77	0.91	0.81	
γ -O-Alk, OH _{sec}	77-65	1.06	0.83	
OH _{prim}	65-58	0.88	0.63	

Table S3. Lignin substructures determined by NMR. Substructures are shown in Figure S4.

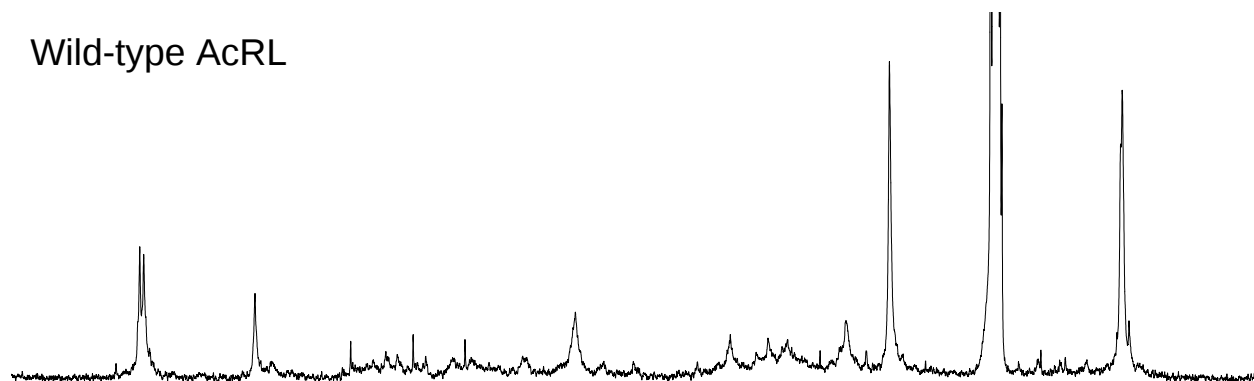
Structure	Calculation if Necessary	WT (per Ar)	C4H-F5H (per Ar)
β -O-4/ α -OH (A)	(77-72.5, ac) – H	0.43	0.42
phenylcoumaran (E)	(88-86, ac)	0.05	0.06
dibenzodioxocin (G)	(86-83, ac) - F	0	0.02
pinoresinol (F)	(54-51) – E	0.11	0.06
β -1 (H)	(50-48, ac) - E	0.12	0.03
OMe	(58-54, na) - H	1.28	2.32
OH _{ph}	(168.6-166, ac) – (168.6-166, na)	0.22	0.24
Alk-O-Ar	A + E + 2xG	0.48	0.48
total etherified	1.00-OH _{ph}	0.78	0.76
γ -O-alk	(77-65)-OH _{sec}	0.54	0.52
maximum α -O-alk	(90-77) – Alk-O-Ar	0.43	0.40
minimum α -O-alk	(90-77) - (1.00-OH _{ph})	0.09	0.12
side chain	(90-52) – OMe + (175-165, na)	3.37	2.90
degree of condensation	theoretical – H units – (125-103, na)	0.17	0



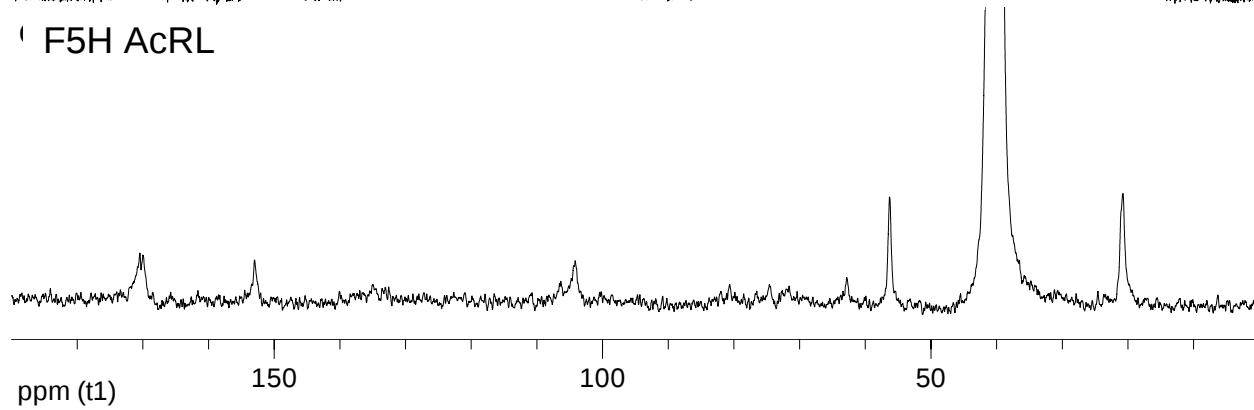


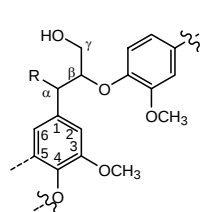


Wild-type AcRL

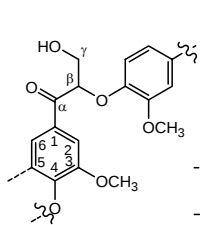


F5H AcRL

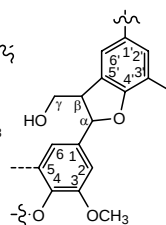




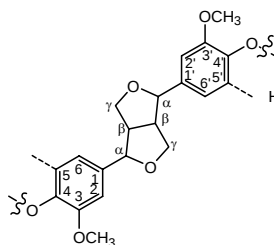
- A. R = OH
B. R = OAr
C. R = OAlk



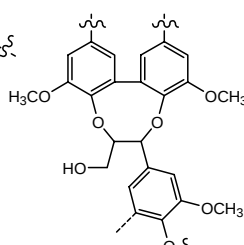
D



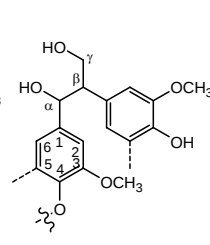
E



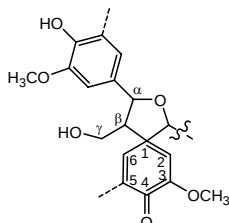
F



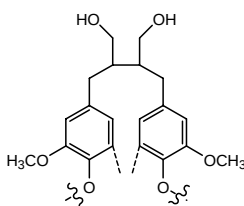
G



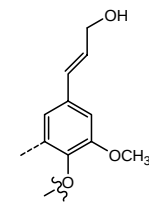
H



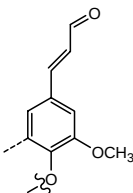
I



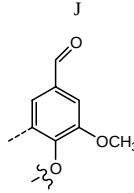
J



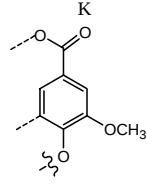
K



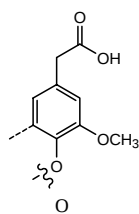
L



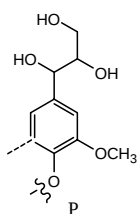
M



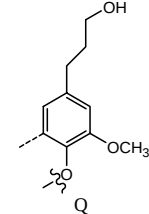
N



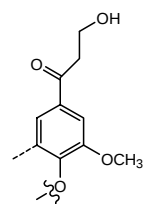
O



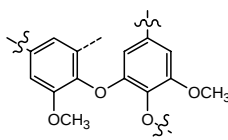
P



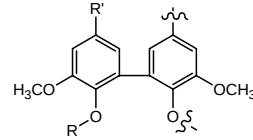
Q



R



S



T

R = H
R' = CHO, CH=CH-CHO, COOH