THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 13, pp. 8843–8853, March 31, 2006
Printed in the U.S.A.

# Effects of Coumarate 3-Hydroxylase Down-regulation on Lignin Structure*

Received for publication, October 26, 2005, and in revised form, January 5, 2006 Published, JBC Papers in Press, January 17, 2006, DOI 10.1074/jbc.M511598200

**John Ralph[‡§1], Takuya Akiyama[‡], Hoon Kim[‡¶], Fachuang Lu[‡§], Paul F. Schatz[‡], Jane M. Marita[‡], and Sally A. Ralph[‖], M. S. Srinivasa Reddy[**], Fang Chen[**], and Richard A. Dixon[**]**

*From the [‡]United States Dairy Forage Research Center, United States Department of Agriculture-Agricultural Research Service, Madison, Wisconsin 53706, Departments of [§]Forestry and [¶]Horticulture, University of Wisconsin, Madison, Wisconsin 53706, [‖]United States Forest Products Laboratory, United States Department of Agriculture-Forest Service, Madison, Wisconsin 53705, and [**]Plant Biology Division, Samuel Roberts Noble Foundation, Ardmore, Oklahoma 73401*

**Down-regulation of the gene encoding 4-coumarate 3-hydroxylase (C3H) in alfalfa massively but predictably increased the proportion of *p*-hydroxyphenyl (P) units relative to the normally dominant guaiacyl (G) and syringyl (S) units. Stem levels of up to ~65% P (from wild-type levels of ~1%) resulting from down-regulation of C3H were measured by traditional degradative analyses as well as two-dimensional $^{13}C$–$^{1}H$ correlative NMR methods. Such levels put these transgenics well beyond the P:G:S compositional bounds of normal plants; *p*-hydroxyphenyl levels are reported to reach a maximum of 30% in gymnosperm severe compression wood zones but are limited to a few percent in dicots. NMR also revealed structural differences in the interunit linkage distribution that characterizes a lignin polymer. Lower levels of key β-aryl ether units were relatively augmented by higher levels of phenylcoumarans and resinols. The C3H-deficient alfalfa lignins were devoid of β–1 coupling products, highlighting the significant differences in the reaction course for *p*-coumaryl alcohol *versus* the two normally dominant monolignols, coniferyl and sinapyl alcohols. A larger range of dibenzodioxocin structures was evident in conjunction with an approximate doubling of their proportion. The nature of each of the structural units was revealed by long range $^{13}C$–$^{1}H$ correlation experiments. For example, although β-ethers resulted from the coupling of all three monolignols with the growing polymer, phenylcoumarans were formed almost solely from coupling reactions involving *p*-coumaryl alcohol; they resulted from both coniferyl and sinapyl alcohol in the wild-type plants. Such structural differences form a basis for explaining differences in digestibility and pulping performance of C3H-deficient plants.**

The effects on lignin structure of perturbing one crucial step in the monolignol biosynthetic pathway remain to be addressed. Genes encoding all of the enzymes in Fig. 1 have been identified, and the effects of perturbing (by down- and/or up-regulation in transgenic plants or via their knockouts in mutants) all but the *p*-coumarate 3-hydroxylase (C3H)[2]/hydroxycinnamoyl transferase (HCT) steps have been studied in some detail, as reviewed in Refs. 1–3. Down-regulation of some genes, particularly those early in the pathway, may limit the overall flux of metabolites into lignin. In other cases, the distribution of units resulting from the primary monomers (the three monolignols *p*-coumaryl **1a**, coniferyl **1b**, and sinapyl **1c** alcohols, differing in their degree of methoxylation; Fig. 1) may be dramatically altered, sometimes far beyond the limits that have been observed in nature. In some intriguing cases, lignification appears to be able to accommodate phenolics (*e.g.* 5-hydroxyconiferyl alcohol) that are not normally considered to be lignin monomers when the biosynthesis of the normal monolignols is thwarted (1, 3, 4). Such studies are not only providing rich insights into the lignification process, but are also opening up opportunities for improving the utilization of plant cell walls in a range of natural and industrial processes, *e.g.* ruminant digestion (5–13) and chemical pulping (2, 14–20, 22).

Aromatic hydroxylation steps are considered key reactions in plant secondary metabolism, in part due to their irreversibility. Ferulate 5-hydroxylase (F5H), now often called coniferaldehyde 5-hydroxylase (CAld-5H) to more accurately reflect the preferred substrate (23, 24), is the crucial enzyme allowing syringyl (S) lignin production via the monolignol sinapyl alcohol **1c**. An *Arabidopsis* mutant deficient in F5H has no syringyl lignin component (25, 26). Similar to gymnosperms, it produces guaiacyl-rich lignins derived almost exclusively from the monolignol coniferyl alcohol **1b**. Up-regulation of F5H, as might be anticipated, produces plants with lignins having higher syringyl contents and that are relatively depleted in guaiacyl units (G). Analyses of lignins suggest syringyl contents of up to ~92% in F5H-up-regulated *Arabidopsis* (25) and as high as 93% in poplar or aspen (27, 28). Such syringyl levels are comparable with the highest reported in nature (29), with kenaf bast fiber lignin being among the highest at 85–94% (30, 31). In poplar and aspen, the following methylation step via caffeic acid 3-*O*-methyl transferase (COMT), also operating at the aldehyde level and therefore misnamed (32), appears to be able to accommodate the increased flux from coniferaldehyde to 5-hydroxyconiferaldehyde to produce sinapaldehyde and ultimately sinapyl alcohol **1c** (27). In *Arabidopsis*, however, evidence suggests that the COMT is not able to keep pace with the increased 5-hydroxyconiferaldehyde production, because the lignins

* This work was supported in part by the United States Department of Energy Biosciences Programs DE-AI02-00ER15067 (to J. R.) and DE-FG03-97ER20259 (to N. G. Lewis and R. A. D.) and by the Samuel Roberts Noble Foundation and Forage Genetics International (to R. A. D.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

S The online version of this article (available at http://www.jbc.org) contains additional methods and figures.

1 To whom correspondence should be addressed: U.S. Dairy Forage Research Ctr., USDA-Agricultural Research Svc., 1925 Linden Dr. West, Madison WI 53706-1108. Tel.: 608-890-0071; Fax: 608-890-0076; E-mail: jralph@wisc.edu.

2 The abbreviations used are: C3H, *p*-coumarate 3-hydroxylase; HCT, *p*-hydroxycinnamoyl-coenzyme A:quinate or shikimate *p*-hydroxycinnamoyltransferase; WT, wild-type; P, *p*-hydroxyphenyl unit in lignin or *p*-coumaryl alcohol monolignol; G, guaiacyl unit in lignin or coniferyl alcohol monolignol; S, syringyl unit in lignin or sinapyl alcohol monolignol; DFRC, derivatization followed by reductive cleavage; ML, milled lignin; AL, acidolysis lignin; EL, enzyme lignin; HSQC, heteronuclear single quantum correlation; HMBC, heteronuclear multiple bond correlation; COSY, two-dimensional correlated spectroscopy; TOCSY, total correlation spectroscopy; COMT, caffeic acid *O*-methyl transferase; F5H, ferulate 5-hydroxylase; CAld-5H, coniferaldehyde 5-hydroxylase; 5HG, 5-hydroxyguaiacyl; ABSL, acetyl bromide-soluble lignin; ppm, parts/million.

have a significant component derived from 5-hydroxyconiferyl alcohol (33). Novel 5-hydroxyguaiacyl (5HG) benzodioxane structures, which result from incorporation of 5-hydroxyconiferyl alcohol into the lignification scheme, analogously to their normal monolignol counterparts, were the same as those noted in COMT-deficient plants (33–36). The 92% syringyl levels in *Arabidopsis* therefore derive only from S/(G + S); there is no reliable method to measure 5HG levels nor to measure S/(G + 5HG + S) ratios that truly reflect the monomer distribution.

F5H affects the partitioning between the two major traditional monolignols, coniferyl **1b** and sinapyl **1c** alcohols (Fig. 1). Because plants with no syringyl components (*i.e.* softwoods) are well known and because natural plants can have very high syringyl components, perturbing F5H might not be expected to greatly interfere with the requirements of the lignin polymer in the plant. The same is presumably *not* true for C3H, however. Although it is likely that perturbing C3H only affects the monolignol distribution, the monolignol *p*-coumaryl alcohol **1a** does not normally contribute to high levels of *p*-hydroxyphenyl (P) units in normal lignins. Rather, these are minor components, typically just 1–3% (29), in the lignins of both gymnosperms and angiosperms. *p*-Hydroxyphenyl units have long been thought to be significantly higher in grasses (37). They may be, but much of the data has come from incorrectly interpreting the products of degradative methods as deriving from *p*-hydroxyphenyl units in lignin, whereas substantial proportions may derive from *p*-coumarate ester moieties adorning the lignins of grasses (37); such *p*-coumarate moieties are not involved in the backbone of the polymer and should not be confused with lignin monomers (38). Softwood compression wood fractions have the richest *p*-hydroxyphenyl unit content, reportedly ranging as high as 30% (39). Nevertheless, even this is below the levels that might be expected via C3H down-regulation.

Chapple and co-workers (40) have reported on analyses of a C3H-deficient *Arabidopsis ref8* mutant. Degradative methods release *p*-hydroxyphenyl units but no detectable guaiacyl or syringyl components (41). This is consistent with the key role of the hydroxylase on the pathway toward coniferyl and sinapyl alcohols. C3H in *Arabidopsis* is now understood to operate on *p*-coumarate esters of shikimic acid (Fig. 1) or quinic acid (not shown), themselves produced by hydroxycinnamoyl transferases (42). The difficulty in securing sufficient cell wall material from these stunted *ref8* mutants has limited more detailed structural studies of the resultant lignins. Yet, such studies are required, because little is known about the coupling and cross-coupling propensities of *p*-coumaryl alcohol **1a** (in lignification reactions), and therefore little can be predicted about the structure and properties of the lignins that might be expected. More recently, one of our groups has successfully generated transgenic plants of the forage legume alfalfa (*Medicago sativa*) in which C3H levels have been reduced to as low as 5% of the wild-type level (Table 1) in the absence of seriously impaired growth phenotypes (13). A detailed structural analysis of the unusual lignins in these plants is described here.

## MATERIALS AND METHODS

### General

All chemicals were purchased from Aldrich (Milwaukee, WI) unless otherwise noted. UV spectra were recorded on DU-800 Series spectrophotometers (Beckman Coulter, Inc., Fullerton, CA) at wavelengths between 400 and 700 nm.

### Cell Wall Analytical Methods

*Acetyl Bromide-soluble Lignin (ABSL)*—ABSL lignins were determined using essentially the methods previously described (43) but on a smaller scale (2–5 mg of cell wall). However, neither the wild-type (WT) nor the transgenic alfalfa walls were fully soluble in acetyl bromide. Also, no molar extinction coefficient has been reported for *p*-hydroxyphenyl-rich lignins; a value of 17.2 (as determined for WT alfalfa) was used.

*Analytical Thioacidolysis*—Analytical thioacidolysis to release diagnostic lignin monomers was carried out as described previously (44). The products were examined by gas chromatography-mass spectrometry. Gas chromatography (Thermoquest Trace GC 2000) conditions were as follows. A DB1 column (25 m × 0.2 mm, 33 $\mu$m film thickness, J & W Scientific) was used. An initial column temperature of 200 °C, held for 1 min, was first ramped at a rate of 4 °C/min to 248 °C and then ramped at a rate of 30 °C/min to 300 °C and held for 25 min. Inlet temperature was 250 °C. Mass spectrometry (Thermoquest GCQ/Polaris MS) conditions were 220 °C for the ion source temperature and 300 °C for the transfer line temperature.

*Derivatization Followed by Reductive Cleavage (the DFRC Method)*—DFRC release and quantification of acetylated monolignols by reductive cleavage of $\beta$-aryl ethers was performed as described previously (45, 46).

*Neutral Sugars and Total Carbohydrate Levels*—Neutral sugar residues (glucose, xylose, arabinose, mannose, galactose, rhamnose, and fucose) were quantified by gas chromatography as alditol acetate derivatives (47). Approximate lignin levels in isolated ML and AL fractions were determined simply as 100% minus the percent of carbohydrates.

### NMR Spectroscopy

The NMR spectra presented here were acquired on a Bruker Biospin (Rheinstetten, Germany) DMX-750 instrument fitted with a sensitive cryogenically cooled 5-mm TXI $^1H/^{13}C/^{15}N$ gradient probe with inverse geometry (proton coils closest to the sample). Lignin preparations (30–60 mg) were dissolved in 0.5 ml of $CDCl_3$; the central chloroform solvent peak was used as an internal reference ($\delta_C$, 77.0; $\delta_H$, 7.26 ppm). We used the standard Bruker implementations of the traditional suite of one-dimensional and two-dimensional (gradient-selected, $^1$H-detected, *e.g.* COSY, TOCSY, HSQC, HSQC-TOCSY, HMBC) NMR experiments for structural elucidation and assignment authentication. The fully authenticated NMR data for model compounds will be deposited in the "NMR data base of lignin and cell wall model compounds" available via the internet at ars.usda.gov/Services/docs.htm?docid=10491.[3] TOCSY experiments used a 100-ms mixing time; HMBC spectra used an 80- or 100-ms long range coupling delay. Normal HSQC experiments at 750 MHz typically had the following parameters. Spectra were acquired from 8.6–2.4 ppm in $F_2$ ($^1$H) using 1864 data points (acquisition time 200 ms), 160–40 ppm in $F_1$ ($^{13}$C) using 512 increments ($F_1$"acquisition time" 11.3 ms) of 16 or 32 scans with a 1-s interscan delay, with a total acquisition time of 2 h and 48 min or 5 h and 34 min; the $d_{24}$ delay was set to 1.72 ms (~0.25/*J*; *J* is the coupling constant in Hz). The higher resolution (Fig. 3*b*, *inset*) acquired just the side chain region from 6.8–2.4 ppm in $F_2$ ($^1$H) using 1336 data points (acquisition time 200 ms), 110–40 ppm in $F_1$ ($^{13}$C) using 800 increments ($F_1$ acquisition time 30 ms) of eight scans with a 1-s interscan delay and a total acquisition time of 2 h and 12 min. Processing used typical matched Gaussian apodization in $F_2$ and a squared sine-bell in $F_1$. HMBC experiments at 750 MHz (Fig. 4) had the following parameters. Spectra were acquired from 8.6–2.4 ppm in $F_2$ ($^1$H) using 3k data points (acquisition time 329 ms), 180–40 ppm in $F_1$ ($^{13}$C) using 400 increments ($F_1$ acquisition time 7.6 ms) of up to 112 scans with a 1-s interscan delay and a 100-ms long range coupling delay, with a total acquisition time of up to 18 h. Processing to a final matrix of 2k by 1k

[3] S. A. Ralph, L. L. Landucci, and J. Ralph, NMR data base of lignin and cell wall model compounds at ars.usda.gov/Services/docs.htm?docid-10429.

data points used typical matched Gaussian apodization in $F_2$ (LB, −80; GB, 0.338) and a squared sine-bell in $F_1$. One level of linear prediction in $F_1$ (32 coefficients) gave improved $F_1$ resolution but was not required.

Volume integration of contours in HSQC plots was accomplished more conveniently and more accurately than in the past by using Bruker's TopSpin 1.3 software. For quantification of P:G:S ratios, only the carbon-2 correlations from guaiacyl units and the carbon-2/6 correlations from syringyl or *p*-hydroxyphenyl units were used, and the guaiacyl integrals were logically doubled. No correction factors were deemed necessary after noting only slight deviations from 1:1:1 volume integral ratios in a range of model dimers and trimers with mixed P/G/S units. For quantification of the various interunit linkage types, the following well resolved contours (see Fig. 3) were integrated: $\mathbf{A}_\alpha$, $\mathbf{B}_\alpha$, $\mathbf{C}_\alpha$, $\mathbf{D}_\alpha$, $\mathbf{S}_\alpha$, $\mathbf{X1}_\gamma$, and $\mathbf{X7}_\beta$, as well as $\mathbf{B}_\beta$, $\mathbf{C}_\beta$, $\mathbf{D}_\beta$, and $\mathbf{S}_\beta$ as checks. Integral correction factors were determined by acquiring spectra from a range of mixed unit trimers and tetramers. Such models have an exact integral molar ratio of units, usually 1:1, making this approach superior to using mixtures of more simple models. The models chosen were from our collection in the NMR data base[3] (noted as Lib. #) or from a recent study on oligolignol metabolic profiling (49) denoted by just #: acetylated S–(β–O-4)–G–(β–5)–G, compound #39; phenol-methylated acetylated G–(β–O-4)–S–(β–β)–S–(4-O–β)–G, #33; acetylated G–(β–5)–Glycerol, Lib. #262; and an acetylated mixture of a simple GG/G-dibenzodioxocin **3b** (see below) and the β-ether model veratryl-glycerol-β-guaiacyl ether (Lib. #105, acetate #3). The determined relative response factors were: $\mathbf{A}_\alpha$ 1.00, $\mathbf{B}_\alpha$ 0.71, $\mathbf{C}_\alpha$ 1.06, $\mathbf{D}_\alpha$ 0.87, $\mathbf{S}_\alpha$ not determined, $\mathbf{X1}_\gamma$ possibly 2.0 (not reliably determined), and $\mathbf{X7}_\beta$ 0.77. These values were used to correct the volume integrals to provide the semiquantitative estimates of unit ratios in Table 3, but it should be noted that these values can only be used here and should not be considered universal; they are dependent upon the spectrometer and acquisition conditions.

### Plant Materials

Transgenic alfalfa (*M. sativa* cv. *Regen SY*) plants down-regulated in C3H transcripts and corresponding enzyme activity were generated as described elsewhere (13). The C3H-4a and C3H-9a lines mentioned herein had 5 and 20% residual C3H activity compared with the WT control (Table 1).

### Lignin Isolation

Stems (internodes 4–10) were harvested from control (WT) and the most heavily C3H-deficient (C3H-4a) alfalfa lines. Lignins were isolated using methods largely described previously (34). Briefly, alfalfa stems were ground and extensively soxhlet-extracted sequentially with water, methanol, acetone, and chloroform. The isolated cell walls were ball-milled for 2.5 h (in 0.5 h on/0.5 h off cycles to avoid excessive sample heating) using a custom-made ball mill using an offset 1/4-hp Dayton motor running at 1725 revolutions/min with rotating (0.2 Hz) stainless steel vessels (12.2 cm diameter, 11.4 cm high) containing ~3.7 kg of 5-mm stainless steel ball bearings; total weight of jar and bearings was ~6.15 kg. The ball-milled walls (9.90 and 10.10 g for WT and C3H-4a transgenic) were then digested at 30 °C with crude cellulases (Cellulysin, Calbiochem, San Diego, CA 219446 lot number B29887, 40 mg/g of sample, in acetate buffer, pH 5.0, for 3 × 48 h using fresh buffer and enzyme each time) leaving all of the lignin and residual polysaccharides, totaling 2.338 g (23.6% of the original cell wall, WT) and 1.487 g (14.7%, C3H-4a) (Table 1). Soluble lignins were difficult to extract from alfalfa as has been noted previously (36). Extraction of 1.001 g (WT) and 1.003 g (C3H-4a) with 96:4 dioxane:$H_2O$ and then freeze drying, washing with water, and filtration of the product using a 10-kDa membrane ultrafilter (YM10–43 mm; Amicon-Millipore Corp., Bedford, MA) to remove water-soluble components (mainly low molecular weight sugars) gave the soluble lignin (ML) fraction (64 and 98 mg) containing 5.1 and 4.8% polysaccharides. After correction, this represents ~1.43 and 1.36% of the cell wall and 7.6 and 10.8% of the total lignin (Table 1). This ML was acetylated overnight and water/EDTA-washed to remove trace metal contaminants (38, 50), giving the acetylated ML used for NMR. To obtain a further fraction, the residue from the ML extraction was subjected to mild acidolysis (51) to yield fractions AL (see supplemental material) comprising another ~14.6 and 6.6% of the original lignin. Finally, to characterize as much of the lignin fraction as possible, the entire polysaccharidase-digested cell wall fraction was subjected to solubilization in $Me_2SO$/*N*-methylimidazole, a solvent shown to dissolve the whole cell wall fraction of ball-milled woods and other plants (52). Dissolution of these alfalfa samples was, however, incomplete. The polysaccharidase-digested cell walls (330 and 370 mg from WT and transgenic C3H-4a, respectively) yielded 360 and 440 mg of acetylated sample. Chloroform fractionation yielded 200 and 220 mg of soluble fractions, 100 and 130 mg of insoluble residue, but 60 and 80 mg in losses. The soluble fraction in chloroform was then subjected to molecular weight fractionation on Biobeads S-X2 (Bio-Rad, Hercules, CA) yielding 137 mg of high molecular mass fractions in each case; a further 29 mg (WT) and 44 mg (C3H-4a) of lower molecular mass material was also recovered but not examined. The fraction of the polysaccharidase-digested cell wall examined by NMR was therefore 38 (WT) and 31% (C3H-4a).

### Model Compounds and Polymers

*Synthesis of p-Coumaryl Alcohol Dimeric, Trimeric, and Oligomeric Coupling and Cross-coupling Products*—The synthesis and characterization of model dimers and oligomers from *in vitro* radical coupling reactions will be described more fully elsewhere once compounds have been separated, fully identified, and the data have been rigorously interpreted.

*Synthetic Dehydrogenation Polymers*—The synthesis of polymers from *p*-coumaryl alcohol alone or copolymers with all three monolignols, used to authenticate the P-aromatic units in alfalfa lignins, are detailed in the supplemental material.

*Synthesis of a PP/P-Dibenzodioxocin Model Compound*—A PP/P-type dibenzodioxocin **3a** (4-(7-hydroxymethyl-2,11-dimethyl-6,7-dihydro-5,8-dioxadibenzo[*a*,*c*]cycloocten-6-yl)-phenol) was synthesized in low yield by coupling a simple 5–5-dimer with *p*-coumaryl alcohol using Ag(I) oxide (53) (Scheme 1). 5,5′-Dimethyl-biphenyl-2,2′-diol **2a** (0.107 g, 0.5 mmol) (54) was dissolved in acetone (5 ml), and Ag(I) oxide (0.463 g, 2 mmol) was added. *p*-Coumaryl alcohol **1a** (0.150 g, 1 mmol) (55) in acetone (30 ml) was added slowly to the suspension over a period of 4 h. The mixture was stirred overnight and partially concentrated under reduced pressure at 25 °C to ~10 ml. The suspension was diluted with EtOAc and washed with saturated $NH_4Cl$, dried over $Na_2SO_4$, and the solvent evaporated under reduced pressure. The crude products obtained were submitted to preparative TLC eluting with $CHCl_3$-EtOAc (1:1, v/v) to obtain a fraction containing the dibenzodioxocin. Further separation by preparative TLC (eluent:hexane-EtOAc 2:1, multiple elution) provided compound **3a** (8 mg, 4% yield; attempts will be made to improve this yield in further studies on the dibenzodioxocin components of these samples). NMR ($CDCl_3$): $\delta_H/\delta_C$ 4.82/86.7 (α), 4.25/86.1 (β), (4.03, 4.12)/63.9 (γ).

*Synthesis of a GG/G-Dibenzodioxocin Model Compound*—The GG/G-type dibenzodioxocin **3b** (4-(7-hydroxymethyl-4,9-dimethoxy-2,11-dimethyl-6,7-dihydro-5,8-dioxadibenzo[*a*,*c*]cycloocten-6-yl)-2-

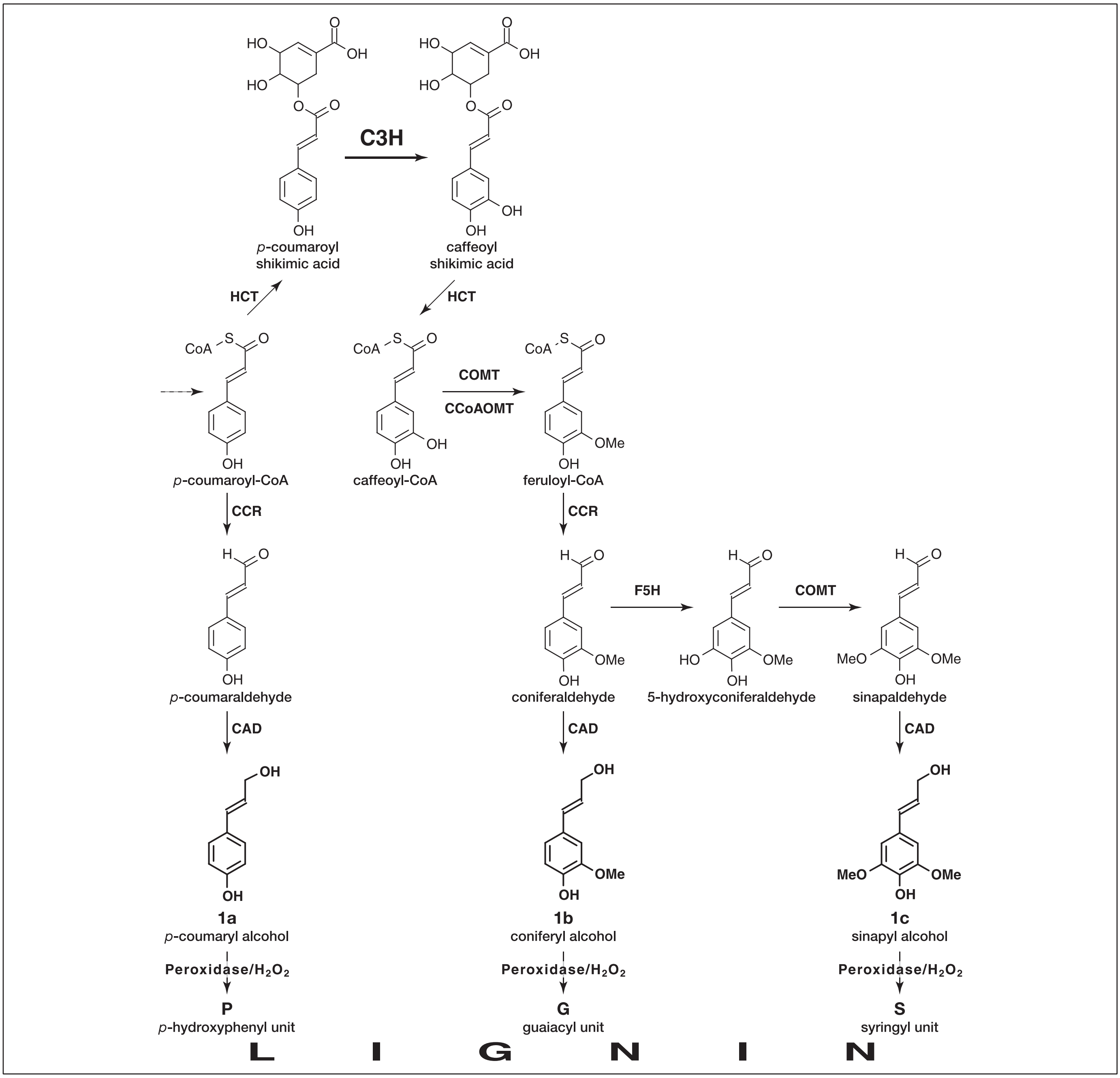



FIGURE 1. **Partial monolignol biosynthetic pathway.** *p*-Coumarate 3-hydroxylase (C3H) is now understood to operate on *p*-coumarate esters of shikimic acid (shown), quinic acid, or possibly others, themselves produced by recently discovered *p*-hydroxycinnamoyl-coenzyme A:quinate shikimate *p*-hydroxycinnamoyltransferases (HCTs) (42). Enzymes on the remainder of the pathway have their standard abbreviations: *CAD*, cinnamyl alcohol dehydrogenase; *CCoAOMT*, caffeoyl-CoA *O*-methyltransferase; *CCR*, cinnamoyl-CoA reductase; *COMT*, caffeic acid *O*-methyltransferase; *F5H*, ferulate 5-hydroxylase.

methoxy-phenol) was prepared using $Mn(OAc)_3$ in pyridine (56) from 3,3′-dimethoxy-5,5′-dimethyl-biphenyl-2.2′-diol **2b** and coniferyl alcohol **1b** (Scheme 1). Yield was 23%. NMR ($CDCl_3$): $\delta_H/\delta_C$ 4.84/84.5 ($\alpha$), 4.13/82.8 ($\beta$), (4.03, 4.50)/64.1 ($\gamma$).

## RESULTS

The aim of the current study was to delineate the consequences of C3H down-regulation for lignin structure. More detailed descriptions of the lines analyzed here and other lines with varying C3H levels are reported elsewhere (13). Briefly, plants with <15% of wild-type C3H activity (demonstrated using *p*-coumaroyl shikimic acid as the substrate and detecting the production of the caffeoyl shikimic acid, Fig. 1) appeared somewhat smaller at flowering than corresponding vector control lines (see Fig. 3, *top left picture*). C3H line 4a, with only 5% residual C3H activity, showed delayed flowering by 10–20 days (Table 1). However, C3H line 9a, with ~20% residual C3H activity, was of normal size and exhibited delayed flowering by only 1–2 days. The bright red coloration throughout the WT stem cross-section vascular tissue following staining with Mäule reagent, a stain normally used to detect S units in lignin (57), was reduced in the C3H lines to a dark brown coloration with more limited distribution between the major xylem cells, consistent with an overall reduction in S-lignin content (13). The C3H-4a line, with the lowest C3H level, was logically chosen for this initial study to delineate

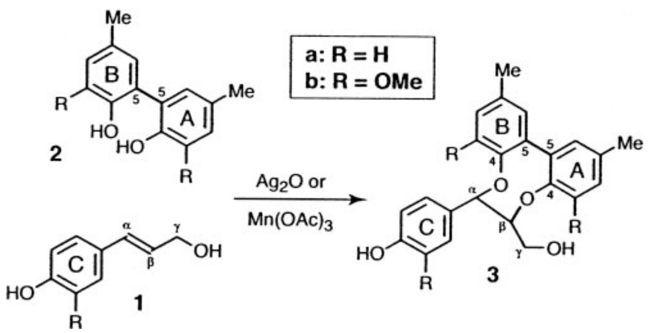
**TABLE 1**

**Basic data on wild-type control and C3H-deficient (C3H-4a, C3H-9a) alfalfa**

Fractions are defined under "Materials and Methods." ML = dioxane: water-soluble milled lignin; AL = acidolysis lignin (from the ML residue); EL = enzyme-digested cell wall; Ac = indicates acetylated samples; ABSL = acetyl bromide-soluble lignin; P = *p*-hydroxyphenyl, G = guaiacyl, S = syringyl, CW = cell wall; thio = thioacidolysis-derived monomers; DFRC = monomers derived from the derivatization followed by reductive cleavage method.

| | WT | C3H-9a | C3H-4a |
|---|---|---|---|
| **C3H activity of WT (%)** | 100 | 20 | 5 |
| Days to early bud | 39 | 41 | 48 |
| Days to bloom | 50 | 51 | 60–70 |
| **Whole Plant** | | | |
| AcBr lignin (% ABSL) | 9.72 | 7.38 | 6.76 |
| Thioacidolysis yield ($\mu$m/g) | 149 | 65 | 54 |
| P (thio) (%) | 2.5 | 28 | 55 |
| G (thio) (%) | 66 | 48 | 30 |
| S (thio) (%) | 31 | 24 | 15 |
| **Stems (%)** | | | |
| P (DFRC) | 3 | | 59 |
| G (DFRC) | 51 | | 22 |
| S (DFRC) | 46 | | 20 |
| **EL (enzyme-digested CW) (%)** | | | |
| CW after digestion | 23.6 | | 14.7 |
| Carbohydrates | 20.0 | | 14.1 |
| Lignin[a] | 80.0 | | 85.9 |
| Soluble (Ac-EL) for NMR | 38 | | 31 |
| **ML (dioxane/water milled lignin) (%)** | | | |
| Carbohydrates | 5.1 | | 4.8 |
| Lignin[a] | 94.9 | | 95.2 |
| ML of CW | 1.43 | | 1.36 |
| ML of total lignin | 7.6 | | 10.8 |
| **AL (acidolysis lignin) (%)** | | | |
| AL of CW | 2.8 | | 0.8 |
| AL of total lignin | 14.6 | | 6.6 |

[a] Lignin calculated as 100% carbohydrates.




SCHEME 1. **Preparation of dibenzodioxocin model compounds 3.** In the text, dibenzodioxocin structures are referred to using *p*-hydroxyphenyl (P), guaiacyl (G) and syringyl (S) descriptors for the A, B, and C rings; *e.g.* the GG/S dibenzodioxocin derives from a sinapyl alcohol (ring C) monomer coupling with a guaiacyl-guaiacyl (GG) 5–5-coupled unit. Bolder bonds (5–5 between rings A and B, $\beta$-*O*-4 between units C and B) are formed during the radical coupling steps. GG/G-dibenzodioxocin, all guaiacyl dibenzodioxcin, *cf.* model compound ***3b***; PP/P-dibenzodioxocin, all *p*-hydroxyphenyl dibenzodioxocin, *cf.* model compound ***3a***.

the most extreme structural effects associated with C3H down-regulation.

*Lignin Levels and Aromatic Unit (P:G:S) Distribution*—Total forage samples (leaf plus stem) from internodes 1–5 were harvested from the C3H-4a-down-regulated and empty vector control lines at the first bud stage. Lignin content was estimated by the ABSL procedure and by total thioacidolysis yield (Table 1). Thioacidolysis also provided estimates of monomer abundance expressed as the relative percent of P (derived from *p*-coumaryl alcohol **1a**), percent of G (from coniferyl alcohol **1b**), and percent of S (from sinapyl alcohol **1c**). ABSL levels of forage samples were significantly reduced in C3H-down-regulated lines. These results were magnified in the corresponding total thioacidolysis yields, and sizeable differences were observed in the thioacidolysis yields of the individual P, G, and S monomers. The reasons become clear from the NMR analyses (below) showing that the $\beta$-ether content is lower in the C3H-deficient plants. Down-regulation of C3H resulted in a massive increase in the proportion of P units in the lignin and a significant decrease in the ratio of G to total units.

**TABLE 2**

**NMR-derived p-hydroxyphenyl: guaiacyl: syringyl (P:G:S) data for stem lignins from control and C3H-deficient plants**

Fractions are defined under "Materials and Methods." C3H-deficient transgenic C3H-4a has 5% residual C3H levels. ML = dioxane: water-soluble milled lignin; AL = acidolysis lignin (from the ML residue); EL = enzyme-digested cell wall; Ac = indicates acetylated samples; Control is the wild type.

| Sample | P | G | S |
|---|---|---|---|
| Control Ac-ML | 0.8 | 58 | 41 |
| Control Ac-AL | 0.7 | 61 | 39 |
| Control Ac-EL | 0.8 | 58 | 41 |
| C3H-4a Ac-ML | 65 | 17 | 18 |
| C3H-4a Ac-AL | 68 | 17 | 15 |
| C3H-4a Ac-EL | 66 | 16 | 18 |

The stem fractions were used for lignin isolation and NMR analysis; some compositional data are given in Table 1. As an independent measure of the wild-type and the C3H-deficient (C3H-4a) lines, the DFRC method gave similar P:G:S ratios as determined for the whole plant by thioacidolysis (Table 1). Both thioacidolysis (44) and the DFRC method (45) release quantifiable monomers from units linked by $\beta$-ether bonds. As such, the measured P:G:S ratio is a reflection of the units involved only in so-called uncondensed units. A measure of the actual P:G:S ratio in soluble lignin fractions was also made by NMR (see below, and Table 2). Although the methodology here has not been firmly established, similarly high P-levels were indicated by two-dimensional NMR volume integration.

*NMR, Aromatic Region*—Changes in the P:G:S distribution in the lignins are most readily visualized from the aromatic region of NMR spectra, particularly the 2D $^{13}$C–$^{1}$H correlation (HSQC) spectra, correlating protons with their attached carbons. Fig. 2 shows the impressive differences in the aromatic nature of the polymers in the wild-type (Fig. 2*a*) *versus* the C3H-deficient (Fig. 2*b*, *C3H-4a*) lignins. As established previously (36), alfalfa lignin is a typical, slightly guaiacyl-rich, syringyl-guaiacyl lignin. Syringyl and guaiacyl aromatic resonances are well separated at 750 MHz but also at lower fields. Traces of the *p*-hydroxyphenyl component are visible in this spectrum (Fig. 2*a*). The lignin from the C3H-deficient alfalfa is strikingly unlike any lignin seen by these investigators. Relatively weak, but diagnostic, syringyl (S) and guaiacyl (G) correlations remain in a spectrum that is overwhelmed by the *p*-hydroxyphenyl (P) correlations. Volume integration (Table 2) allows reasonable quantification of the differences that are plainly visible. The lignins in all three types of fractions analyzed (the solvent-soluble lignin ML, the acidolysis lignin AL on the residue, and the crude enzyme lignin EL following our cell wall dissolution procedure) all showed similar distributions (see Table 2 and supplemental Figs. 2 and 3). As might be anticipated, the enzyme lignin (EL, the best representative of the whole lignin) has a distribution between that of the easily removed solvent-soluble ML and the most rigorously extracted acidolysis lignin (AL) fraction from the ML residue. This suggests that only minimal partitioning of structure types between the fractions has occurred and that the isolated solvent-soluble lignin from ball-milled material is representative of that from the whole ball-milled cell wall.

*NMR, Side Chain Region*—The side chain region only peripherally reflects the changes in the P:G:S distribution but is rich in detail regarding the types and distribution of interunit bonding patterns present in the lignin fraction.

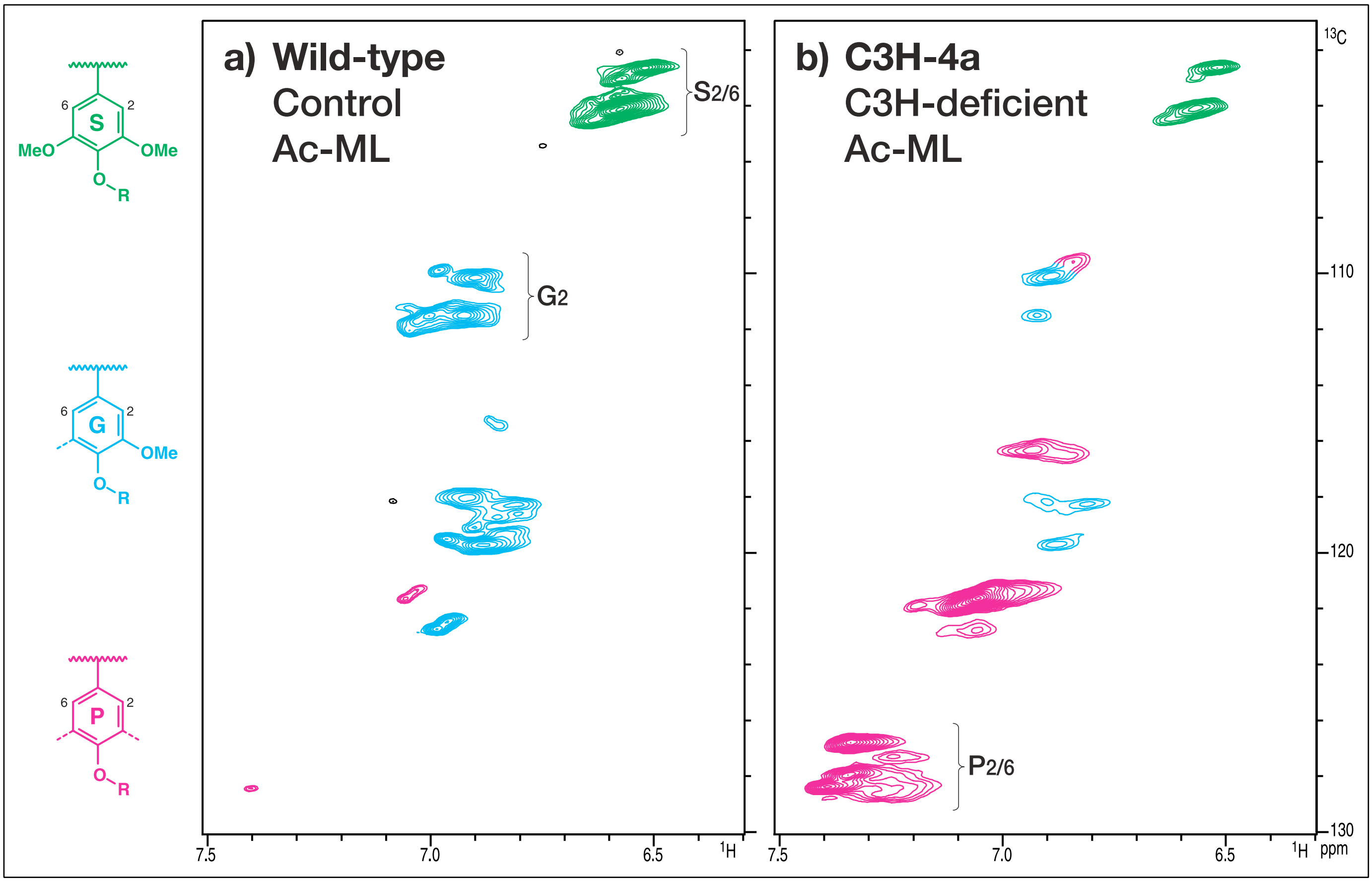



FIGURE 2. **Partial short range $^{13}$C–$^{1}$H (HSQC) correlation spectra (aromatic regions only) of milled lignins (ML) isolated from the wild-type control (*a*) and the most highly C3H-deficient line, C3H-4a (b).** Traces of *p*-hydroxyphenyl (P) units are seen in the typically syringyl/guaiacyl (S/G) lignin in the wild-type alfalfa, whereas P-units entirely dominate the spectrum in the transgenic alfalfa. Semiquantitative volume integrals are given in Table 2. Analogous spectra for the other lignin isolates (acidolysis lignins, AL, and enzyme lignins, EL) are provided in the supplemental material.

The control lignin spectrum (Fig. 3*a*) is typical of a guaiacyl/syringyl lignin containing residual polysaccharides (58). The HSQC spectrum resolves most of the correlations for the various linkage types in the polymer, the exception being in the complex $\gamma$-region, where only the correlations from the phenylcoumarans **B** and the cinnamyl alcohol end groups **X1** are fully resolved. The lignin is seen as being rich in $\beta$-aryl ether units **A**, with more modest amounts of phenylcoumaran **B** and resinol **C** units, as is typical for all lignins. Arylglycerol units **X7**, not normally reported, are identified here; it is suspected that they may arise from $\beta$-ether units during ball milling, but they can also be produced under oxidative coupling reaction conditions (see supplemental material). Their 5HG analogs have recently been documented in COMT-deficient alfalfa lignins (36). Spirodienone structures **S**, $\beta$–1-coupled units only recently authenticated in lignin spectra (59, 60), are readily seen in alfalfa (3). The diagnostic dibenzodioxocins **D** are also relatively newly discovered eight-membered ring structures resulting from radical coupling of a monolignol with a 5–5-coupled end unit (61). Because syringyl units cannot be involved in 5-coupling, dibenzodioxocins have previously been considered to be most prevalent in guaiacyl-rich lignin fractions (58). Finally, the cinnamyl alcohol end groups **X1**, similar to the resinols **C**, arise from monomer-monomer coupling and are therefore relatively minor; the deceptively strong $\mathbf{X1}_{\gamma}$-C/H correlation peak is due to the sharpness caused by the relative invariance of proton and carbon chemical shifts in such structures where the bonding is on the aromatic ring, well distant from the $\gamma$-position.

The C3H-deficient lignin has a spectrum that has several conspicuous differences. In addition to the relative intensity differences (seen more easily from the volume integral data in Table 3) are two notable structural changes. First, there are no apparent spirodienones **S**, even at contour levels closer to the noise level. Second, there are now several types of dibenzodioxocins **D**, with considerable differences in chemical shifts. The *inset* in Fig. 3*b* shows lower contour levels from an experiment run at higher $^{13}$C resolution. Sets of contours labeled with **D** are clearly visible, with only the minor $\mathbf{D}_{\alpha 1}$ and $\mathbf{D}_{\beta 1}$ pair corresponding to the component observed in the wild-type lignin.

*NMR, Long Range Correlations*—Long range correlations (via carbons and protons linked via 2–3 intervening bonds) from HMBC spectra are valuable in providing information on the types of units (P, G, or S) involved in each linkage type (26, 27, 58, 62). This is because the carbon chemical shifts in such units are diagnostically different; syringyl 2/6-carbons are at 102–105 ppm, guaiacyl 2-carbons at ~110–112 ppm, guaiacyl 6-carbons at ~ 118–120 ppm, and *p*-hydroxyphenyl 2/6 carbons at 126–130 ppm. Correlations of $\alpha$-protons in any of the structures to carbons at these disperse frequencies diagnostically determine whether the unit involved is P, G, or S. Although the correlations are not quantitatively relevant, the following are clear from the control lignin spectrum (Fig. 4*a*): (i) both $\beta$-ether **A** and phenylcoumaran **B** units derive from both sinapyl and coniferyl alcohol coupling reactions (they are associated with both syringyl (*S*) and guaiacyl (*G*) units), (ii) resinol units **C** are essentially all syringyl units from sinapyl alcohol, (iii) arylglycerol units **X7** (as seen only at lower contour levels than shown) are largely syringyl, and (iv) unfortunately, useful correlations are not evident for dibenzodioxocins **D** using the parameters of this experiment. Interestingly, $\beta$-ether units **A** also apparently derive from the low levels of *p*-coumaryl alcohol in the normal plant, suggesting that *p*-cou-

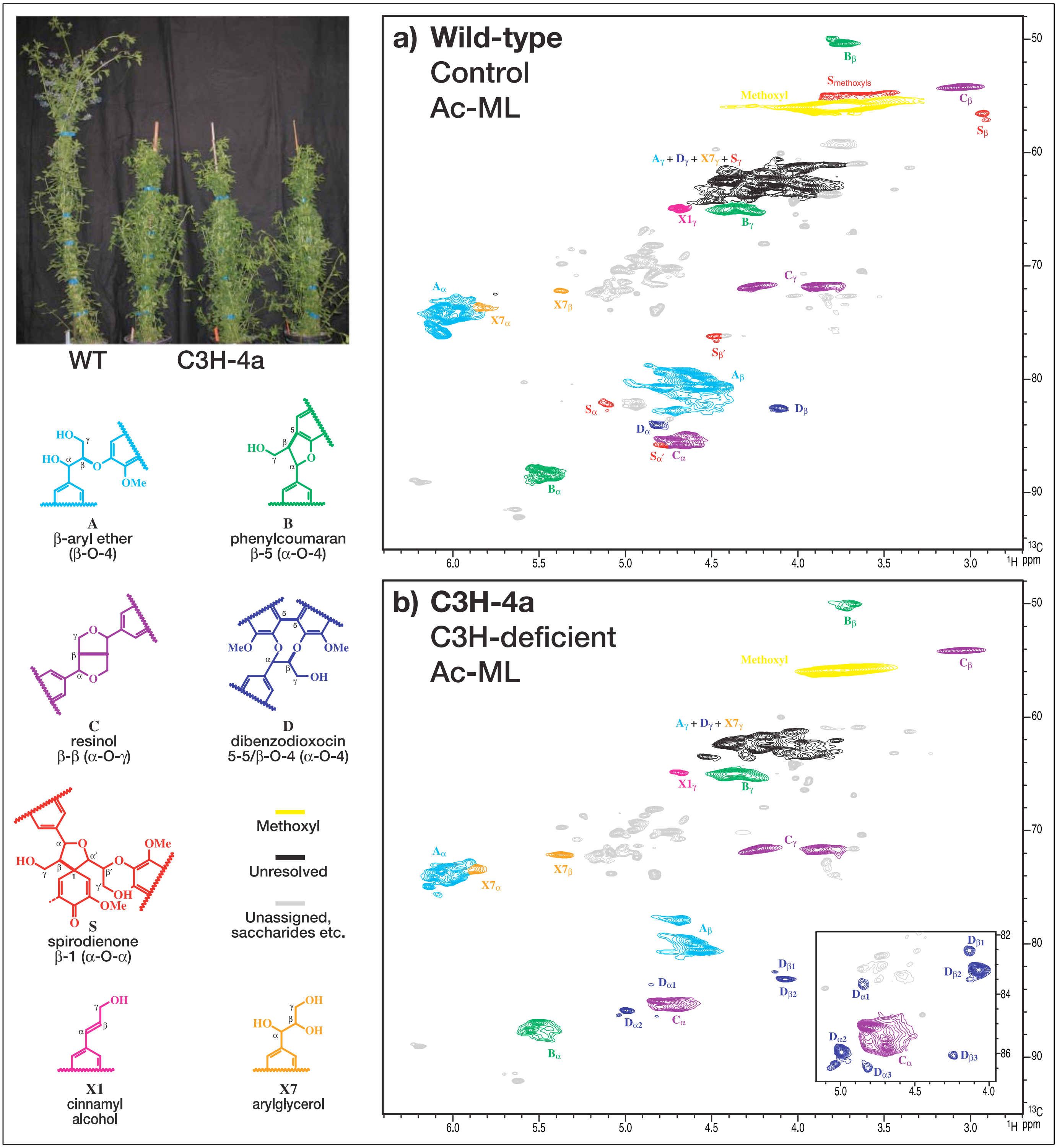



FIGURE 3. **Partial short range $^{13}C$–$^{1}H$ (HSQC) spectra (side chain regions) of milled lignins (ML) isolated from the wild-type control (*a*) and the most highly C3H-deficient line, C3H-4a (*b*).** C3H deficiency and the incorporation of higher levels of *p*-coumaryl alcohol into the lignin produce substantial changes in the distribution of interunit linkage types. The absence of spirodienone S units in the transgenic alfalfa reveals that *p*-coumaryl alcohol does not apparently favor $\beta$–1 cross-coupling reactions. Several types of new dibenzodioxocins **D** are more readily seen at the lower contour levels in the more highly resolved partial spectrum in the *inset*. Note that the contour levels used to display the two spectra were chosen to highlight the structural similarities and differences; with no internally invariant peaks, interpretation of apparent visual quantitative differences needs to be cautious. The *upper left corner photograph* shows WT and C3H-4a transgenic plants at the WT flowering stage; pictures of the C3H-9a transgenic and histochemical staining are provided elsewhere (13). Volume integrals and semiquantitative data are given in Table 3. Analogous spectra for the other lignin isolates (acidolysis lignins (AL) and enzyme lignins (EL)) are provided in the supplemental material. Interunit type designations **A**-**D**, **S**, **X1**, and **X7** follow conventions established previously (1, 3, 58).

maryl alcohol efficiently cross-couples with the dominant guaiacyl and/or syringyl units.

HMBC correlations from the C3H-deficient lignin (Fig. 4*b*) provide the first indications of the coupling and cross-coupling propensities of *p*-coumaryl alcohol and *p*-hydroxyphenyl units (see "Discussion"). The following appear to be evidenced quite clearly: (i) $\beta$-ether units **A** are of all three types, P, G and S, but with relatively low levels of the G- (especially) and S-units; (ii) the phenylcoumarans **B** are almost entirely P; (iii)

**TABLE III**

**NMR-derived interunit linkage data for stem lignins from control and C3H-deficient plants**

Fractions are defined under "Materials and Methods." C3H-deficient transgenic C3H-4a has 5% residual C3H levels. Control is the wild type. ML = dioxane: water-soluble milled lignin; AL = acidolysis lignin (from the ML residue); EL = enzyme-digested cell wall; Ac = indicates acetylated samples; A = $\beta$–$O$–4 ($\beta$-aryl ether); B = $\beta$–5 (phenylcoumaran); C = $\beta$–$\beta$ (resinol); D = dibenzodioxocin; S = $\beta$–1 (spirodienone); X1 = cinnamyl alcohol endgroup; X7 = arylglycerol endgroup.

| Sample | %A | %B | %C | %D | %S | %X1 | %X7 |
|---|---|---|---|---|---|---|---|
| Control Ac-ML | 75 | 9 | 9 | 1.1 | 0.6 | 4.8 | 0.5 |
| Control Ac-AL | 80 | 8 | 7 | 0.6 | 0.2 | 3.7 | 0.4 |
| Control Ac-EL | 77 | 8 | 8 | 0.7 | 0.6 | 4.7 | 0.6 |
| C3H′ Ac-ML | 56 | 18 | 16 | 2.6 | | 2.9 | 4.6 |
| C3H′ Ac-AL | 56 | 16 | 14 | 2.1 | | 5.3 | 6.0 |
| C3H′ Ac-EL | 53 | 17 | 16 | 1.8 | | 5.6 | 6.6 |

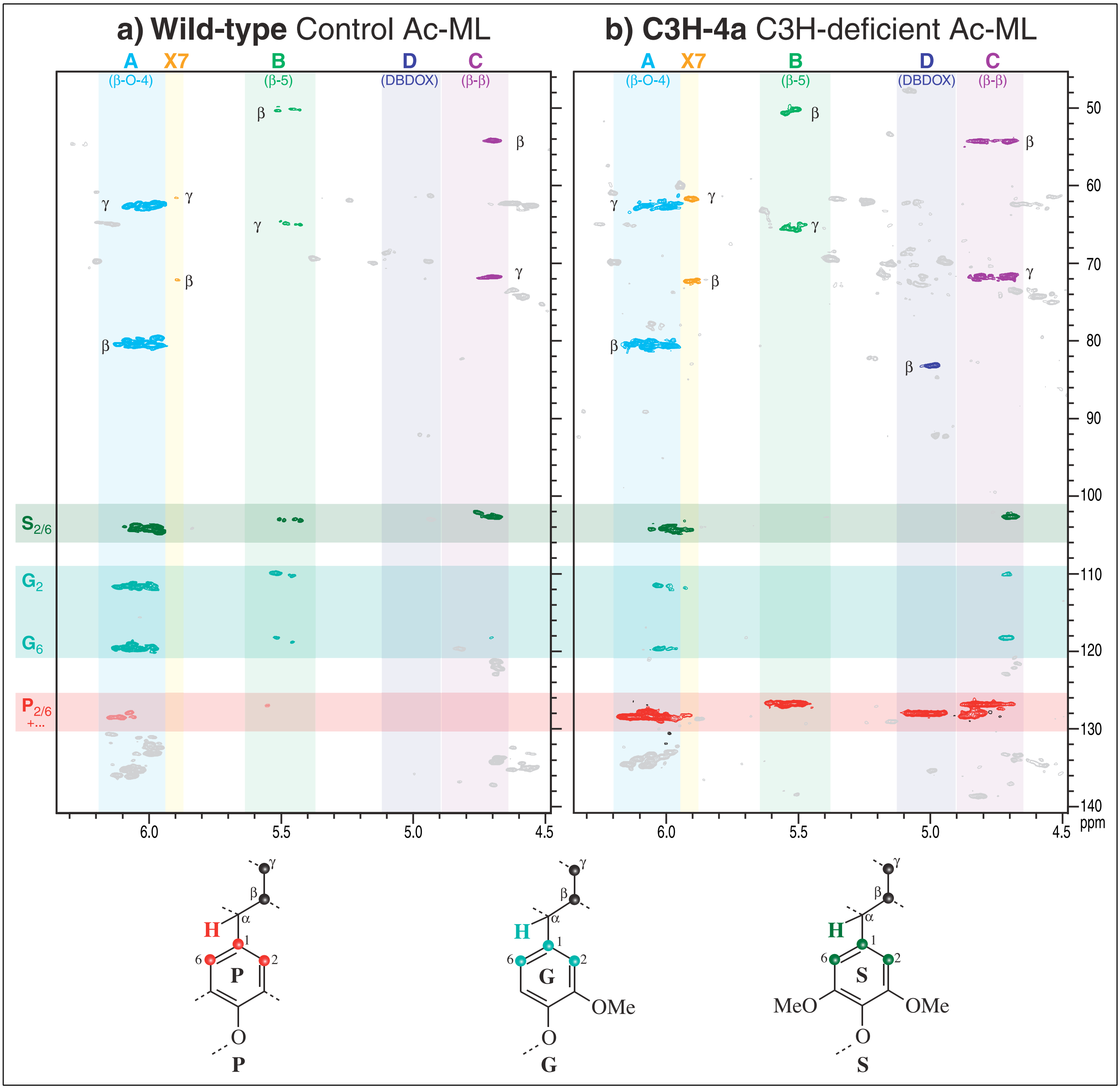



FIGURE 4. **Partial HMBC spectra of ML lignins isolated from the wild-type control (*a*) and the most highly C3H-deficient line, C3H-4a (*b*).** These long range (2–3-bond) $^{13}$C–$^{1}$H correlation spectra allow a determination of the monolignol involved in forming each type of structural unit (see "Results"). The correlations *highlighted* are from $\alpha$-protons to the carbons within three bonds, most diagnostically those to the 2- and 6-positions on the aromatic rings of *p*-hydroxyphenyl (P), guaiacyl (G), and syringyl (S) units.

the dibenzodioxocins **D** now derive from *p*-coumaryl alcohol addition to 5–5-coupled units but, because correlations from G- and S-analogs do not show up in the control, it is not possible to discern whether or not they remain in the C3H-down-regulated material; (iv) resinols **C** are no longer just syringyl (syringyl and guaiacyl units are evident, as well as dominant *p*-hydroxyphenyl units for the resinols that have their $\alpha$-protons displaced to higher chemical shift); (v) arylglycerol units **X7** appear to now derive from all three monolignols but mainly from sinapyl and *p*-coumaryl alcohols.

## DISCUSSION

The NMR spectra reveal both massive and subtle structural differences between the syringyl/guaiacyl lignins in normal wild-type alfalfa *versus* the *p*-hydroxyphenyl-rich lignins in the heavily C3H-down-regulated plants. They provide the first information regarding the incorporation profile for *p*-coumaryl alcohol into (*p*-hydroxyphenyl-rich) copolymer lignins. For the first time in such studies, we have attempted to determine whether significant partitioning of lignin structures has occurred during the lignin fractionation and isolation. We analyzed not only the typical solvent-soluble lignins (ML, ~7.6% and 10.8% of the total lignins, Table 1) but also a further sequential fraction from mild acidolysis (51) (AL, another 14.6% and 6.6% of the total lignins) and, more importantly, a major fraction via dissolution of the enzyme lignin (EL); this latter fraction retains the entire lignin component, but our attempts to analyze the whole lignin component, as we have done for various woody samples (52, 63), were thwarted by ending up with only 38 (WT) and 31% (C3H-4a) of this fraction in the $CDCl_3$-soluble NMR fraction (following EDTA washing and molecular weight fractionation). The difficulty in obtaining soluble lignins from alfalfa has been noted previously (36). However, analysis of all three fractions (ML, AL, and EL; Tables 2–3) suggests that neither partitioning of P/G/S units nor of the various interunit linkage types is a serious issue.

*Alfalfa Plants Incorporating High Levels of p-Coumaryl Alcohol*—A C3H-deficient *Arabidopsis* mutant demonstrated that plants without access to the two primary monolignols, coniferyl and sinapyl alcohols (Fig. 1), could be viable, although with stunted growth (40). The alfalfa plants here demonstrate that low levels of C3H may result in lignins markedly divergent from those in wild-type controls but with the plants exhibiting more normal growth and development patterns, as is discussed in more detail elsewhere (13). Thus, the C3H-4a line, with about 5% residual C3H-activity and about 55–60% P-lignin content (compared with 1–3% in wild-type; Tables 2–3), grows more slowly and flowers later (Table 1, Fig. 3). However, plants with ~20% residual C3H activity compared with wild-type levels, as in line C3H-9a, restores the growth phenotype to more closely resemble the wild-type (13). Nevertheless, the relative *p*-hydroxyphenyl lignin level is still nearly 10-fold more than that found in the wild-type plant (Table 1). These intermediate level plants will eventually be structurally analyzed in more detail.

*Lignin Composition and Structure*—The anticipated effect of C3H deficiency, an enhancement of the relative level of *p*-hydroxyphenyl (P) units in the lignin, is compellingly demonstrated in the aromatic profiles revealed by HSQC NMR spectra (Fig. 2). Wild-type plants have syringyl/guaiacyl lignins with only low levels of P-units. Reduction of C3H depressed the synthesis of coniferyl and sinapyl alcohols, although as noted for all other enzymes in the pathway, not in direct proportion to the enzyme expression level. The most severely down-regulated C3H-4a line (Fig. 2*b*) was G- and S-depleted and strikingly P-rich (~65% P; Table 2). At this time, without the availability of a range of model compound data for *p*-hydroxyphenyl-derived dimers and oligomers, only limited assignments of the correlation peaks in this region can be made. The spectra nevertheless demonstrate the dramatically altered composition of the lignins. Further insight and reasonable quantification of the effects can be gained by volume-integrating the contours in such spectra. Judiciously choosing only the S2/6, G2, and P2/6 contours assures meaningful comparisons, because these ring positions are unsubstituted in lignins; carbons-3 and -5 may be carbon- or oxygen-linked and therefore do not show up in these direct $^{13}C$–$^{1}H$ correlation spectra. It is necessary to double the G2 contribution to allow for the fact that only a single carbon-proton pair (G2) contributes to this integral in guaiacyl units *versus* two pairs (S2/6 or P2/6) in the symmetrical S and P units. The assumption that all S2/6 and G2 units are in the regions indicated is fairly sound, but there is simply not enough model compound data to know whether all P2/6 resonances are in the indicated region. The HSQC integrals were expected to be quantitatively relevant, because they are similar structurally and in an NMR sense. Utilizing dimeric and trimeric model compounds containing mixtures of P, G, and S units showed that essentially no correction is needed and that the "response factors" for each unit type are within a few percent of each other; no corrections are therefore necessary to estimate P:G:S ratios from the volume integrals. As seen in the data of Table 2, the NMR-derived P:G:S ratios are consistent with those ratios derived from degradative methods (Table 1), methods that need not reflect the *actual* ratio of units in the polymer, because they measure the distribution only in the relatively low level of monomers that can be released by cleaving $\beta$-ether bonds.

The high field HSQC spectra of the side chain regions are more revealing regarding the manner in which the monomeric units are assembled. Lignins are characterized by various types of interunit bonds, the most prominent being labeled **A**-**D**, **S**, and **X1** and **X7** (Fig. 3) to retain consistency with standard assignment labeling (1, 3, 58). This linkage-type distribution differs substantially between the wild-type and the C3H-depleted alfalfa lignins. The minor but important spirodienones **S** resulting from $\beta$–1-coupling reactions are absent in the C3H-deficient lignin. The predominant reaction in lignification is the cross-coupling of a monolignol with the growing polymer. Little enough is known about the cross-coupling of coniferyl and sinapyl alcohols with guaiacyl and syringyl lignins (64), and virtually nothing is known about such reactions involving *p*-coumaryl alcohol. It seems logical from the absence of $\beta$–1-coupling products that none of the three monolignols efficiently $\beta$–1-couples with P $\beta$-ether phenolic end units in the lignin, nor does *p*-coumaryl alcohol readily $\beta$–1-couple with G or S $\beta$-ether end units. As has been discovered with other units, this is likely because of a simple chemical incompatibility. For example, angiosperms efficiently incorporate coniferaldehyde because it will cross-couple with syringyl units (33, 65, 66). However, its *in vitro* or *in vivo* inability to cross-couple with guaiacyl units limits its incorporation into gymnosperm lignins (3).

A multitude of dibenzodioxocins appear in the C3H-deficient lignins. Elucidating the exact nature of these will require considerably more work. The **D1** correlations match those in a GG/G-dibenzodioxocin model **3b** (Scheme 1) and in the wild-type lignin, so are logically attributable to dibenzodioxocins **D** formed by coupling of a monolignol with 5–5-linked units derived from the coupling of two guaiacyl units, *i.e.* GG/G-, GG/S- or possibly GG/P-dibenzodioxocins. The data for the dibenzodioxocin peaks labeled **D3** ($\mathbf{D}_{\alpha 3}$ and $\mathbf{D}_{\beta 3}$ for the $\alpha$- and $\beta$-C/H correlations) match those for a synthesized PP/P-dibenzodioxocin model **3a** in which all three aromatic nuclei are *p*-hydroxyphenyl (see "Materials and Methods"). The nature of the large **D2** correlations remains equivocal. We assumed, from the high levels of P-units in this lignin, it would be an all PP/P-dibenzodioxocin, but the data do not match model **3a**. As noted below, from analysis of the HMBC data, it does at least appear to derive from coupling of *p*-coumaryl alcohol to a 5–5-unit. The shifted correlations may be due to 3-substitution on P-units involved in 5–5-coupled structures. Thus PP′/P (where P′ is a 5-linked P-unit) or possibly PG/P dibenzodioxocins seem likely candidates (to be confirmed by future studies).

The other differences in unit-type distribution can be visualized in the spectra but are more readily revealed from data reported in Table 3. Semiquantification used volume-integral correction factors derived from model data, as described under "Materials and Methods." The lower proportion of β-ether units **A** (∼53–56% *versus* ∼75–80% of the units quantified) is clearly a major reason for the lower thioacidolysis yields (on a lignin basis) for the C3H-deficient plants *versus* the wild-type. It also suggests that alkaline-pulping efficiency will be lower, because pulping depends on ether cleavage reactions to depolymerize the lignin and render its fragments soluble in the pulping liquor. However, because lignin-polysaccharide cross-linking can occur via the trapping of intermediate β-ether quinone methides during lignification, reducing the β-ether content may reduce lignin-polysaccharide cross-linking and produce cell walls that are more enzymatically degradable, as demonstrated for the C3H-down-regulated lines via their improved digestibility in ruminant animals (13). Much of the decrease in β-ether **A** levels appears to be due to the two other major units, phenylcoumarans **B** and resinols **C**, each of which nearly doubles in relative proportion. The higher resinol concentration particularly suggests that more monomer-monomer coupling reactions are occurring during the lignification in the P-rich lignins. Although still quite low, the relative dibenzodioxocin **D** level is about double that found in wild-type plants. Quantification of cinnamyl alcohol end groups **X1** was the most variable, but relative levels seem to be similar. Finally, the glycerol structures **X7** are at substantially higher (∼8–11-fold) levels. Because it is not known whether glycerol side chains derive from β-ether units during ball milling or are produced during lignification, we do not speculate on their relevance at this time. However, we have noted (see supplemental material) that oxidative coupling reactions using *p*-coumaryl alcohol produce substantial levels of glycerols (in synthetic polymers that have not been ball-milled), so we are beginning to suspect that they may be, at least in part, authentic units in the native lignins.

*Coupling and Cross-coupling Propensities of p-Coumaryl Alcohol in Lignification*—Details regarding lignification via enhanced levels of *p*-coumaryl alcohol come from analysis of the HMBC spectra. These spectra allow us to determine which monomers are involved in the formation of each interunit linkage type. The observation from Fig. 4 that essentially all phenylcoumarans **B**, for example, are formed by coupling reactions involving *p*-coumaryl alcohol is the kind of information that is required to understand the cross-coupling propensity of this monomer. Unfortunately, the spectra do not reveal whether the coupling was with a guaiacyl or another *p*-hydroxyphenyl unit. Similarly, the change in resinols **C** from being almost entirely derived from sinapyl alcohol in the wild-type to deriving from all three monomers (and *p*-coumaryl alcohol in particular) in the transgenic plants is notable. It is logical that sinapyl alcohol monomers find themselves only rarely able to dimerize (with other sinapyl alcohol monomers) in this sinapyl alcohol-depleted plant, so presumably β–β cross-couple with coniferyl alcohol and possibly *p*-coumaryl alcohol. Studies are required to determine which of these cross-coupling reactions are chemically feasible, *i.e.* if sinapyl alcohol and *p*-coumaryl alcohol, for example, are compatible for radical cross-coupling. Examination of thioacidolysis or DFRC dimers by gas chromatography-mass spectrometry (67) may eventually shed light on the occurrence of mixed resinols. Although dibenzodioxocins **D** do not show useful correlations in the wild-type lignins, it appears that the coupling in the C3H-deficient plant also involves coupling of *p*-coumaryl alcohol with the dibenzodioxocin-precursor 5–5-end unit. The crucial β-aryl ether units derive from all three monolignols, illustrating that *p*-coumaryl alcohol is able to function in the most important end-wise coupling reactions that allow polymer growth. Because *p*-coumaryl alcohol is such a major constituent in these heavily down-regulated plants, the way in which coniferyl and sinapyl alcohols cross-couple with P-units is less clear but should eventually be revealed through studies on alfalfa and other plants with varying levels of down-regulation.

*Comments Regarding the Mechanism of Lignification*—For decades the accepted theory for lignification is one in which monolignols polymerize largely by radical cross-coupling reactions with the growing polymer in a purely chemical reaction, as recently reviewed (3). A challenge (68) hypothesizes that the lignin primary structure could be absolutely dictated by synthesis on arrays of dirigent coupling sites and replicated by template polymerization. As the review (3) notes, the existing theory explains the current facts and readily explains how massive alterations in the monomer profile may drastically affect lignification and the resulting lignin structure but are readily accommodated by the lignification process. Essentially, because it is simply a chemical process, any phenol finding itself in the lignifying zone of the cell wall is capable of entering into the combinatorial free radical-coupling process to the extent allowed by simple chemical concerns, such as structural compatibility, and influenced by typical physical parameters, such as pH, temperature, ionic strength, and the matrix in general. This is the case here for the at least 20-fold-increased levels of the typically minor monolignol *p*-coumaryl alcohol but is also seen in more extreme cases. For example, COMT-deficient plants supplant sinapyl alcohol with the non-traditional monomer 5-hydroxyconiferyl alcohol during lignification (33–36). The dirigent array hypothesis encounters difficulty with monomer substitution. The challenge remaining is for the new hypothesis to explain how a template allows monomer substitution and to provide direct evidence for ordered macrostructures in lignins. The plant materials examined in the present work may provide materials for such studies.

## CONCLUSIONS

Alfalfa lignins strikingly rich in *p*-hydroxyphenyl (P) units are produced by C3H down-regulation. NMR analysis of the lignins suggests that *p*-coumaryl alcohol undergoes coupling and cross-coupling reactions that are, for the most part, analogous to those of the normally dominant monolignols, coniferyl and sinapyl alcohols. The absence of β–1 structures and a considerable shift in the proportions of others demonstrate that the lignification profile is, however, significantly different. Although the total lignin level appears to be lower in C3H-deficient plants, it is apparent that the plants are substituting the more available *p*-coumaryl alcohol for the normally higher levels of coniferyl and sinapyl alcohols; thus P levels that are typically 1–3% of the lignin rise to ∼65% of the lignin in the most heavily down-regulated line. At the extreme, the C3H-deficient *Arabidopsis ref8* mutant (41), which produces lignin at the 100% P level (because it is totally devoid of S and G units), is utilizing the only available monolignol. 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.

*Acknowledgments—We thank Dr. Ron Hatfield for carbohydrate analysis. NMR experiments on the Bruker DMX-750 cryoprobe system were carried out at the National Magnetic Resonance Facility at Madison, WI, with support from the National Institutes of Health (NIH) Biomedical Technology Program (RR02301) and additional equipment funding from the University of Wisconsin, National Science Foundation (NSF) Academic Infrastructure Program (BIR-9214394), NIH Shared Instrumentation Program (RR02781, RR08438), NSF Biological Instrumentation Program (DMB-8415048), and the United States Department of Agriculture.*

## REFERENCES

1. Boerjan, W., Ralph, J., and Baucher, M. (2003) *Ann. Rev. Plant Biol.* **54,** 519–549
2. Baucher, M., Halpin, C., Petit-Conil, M., and Boerjan, W. (2003) *Crit. Rev. Biochem. Mol. Biol.* **38,** 305–350
3. 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., and Boerjan, W. (2004) *Phytochem. Rev.* **3,** 29–60
4. Sederoff, R. R., MacKay, J. J., Ralph, J., and Hatfield, R. D. (1999) *Current Opin. Plant Biol.* **2,** 145–152
5. Krause, D. O., Denman, S. E., Mackie, R. I., Morrison, M., Rae, A. L., Attwood, G. T., and McSweeney, C. S. (2003) *FEMS Microbiol. Rev.* **27,** 663–693
6. Guo, D. G., Chen, F., Wheeler, J., Winder, J., Selman, S., Peterson, M., and Dixon, R. A. (2001) *Transgenic Res.* **10,** 457–464
7. He, X., Hall, M. B., Gallo-Meagher, M., and Smith, R. L. (2003) *Crop Sci.* **43,** 2240–2251
8. Bernard Vailhé, M. A., Besle, J. M., Maillot, M. P., Cornu, A., Halpin, C., and Knight, M. (1998) *J. Sci. Food Agric.* **76,** 505–514
9. Chen, L., Auh, C.-K., Dowling, P., Bell, J., Chen, F., Hopkins, A., Dixon, R. A., and Wang, Z.-Y. (2003) *Plant Biotechnol. J.* **1,** 437–449
10. Barrière, Y., Ralph, J., Méchin, V., Guillaumie, S., Grabber, J. H., Argillier, O., Chabbert, B., and Lapierre, C. (2004) *C. R. Biol.* **327,** 847–860
11. Ralph, J., Guillaume, S., Grabber, J. H., Lapierre, C., and Barrière, Y. (2004) *Comptes Rend. Biologies* **327,** 467–479
12. Jung, H.-J. G., Ni, W., Chapple, C. C. S., and Meyer, K. (1999) *J. Sci. Food Agric.* **79,** 922–928
13. Reddy, M. S. S., Chen, F., Shadle, G. L., Jackson, L., Aljoe, H., and Dixon, R. A. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102,** 16573–16578
14. Boudet, A. M., and Grima-Pettenati, J. (1996) *Mol. Breed.* **2,** 25–39
15. Lapierre, C., Pollet, B., Petit-Conil, M., Toval, G., Romero, J., Pilate, G., Leple, J. C., Boerjan, W., Ferret, V., De Nadai, V., and Jouanin, L. (1999) *Plant Physiol.* **119,** 153–163
16. Chen, C. Y., Baucher, M., Christensen, J. H., and Boerjan, W. (2001) *Euphytica* **118,** 185–195
17. Baucher, M., Christensen, J. H., Meyermans, H., Chen, C. Y., Van Doorsselaere, J., Leplé, J. C., Pilate, G., Petit-Conil, M., Jouanin, L., Chabbert, B., Monties, B., Van Montagu, M., and Boerjan, W. (1998) *Polym. Degrad. Stab.* **59,** 47–52
18. Lapierre, C., Pollet, B., Petit-Conil, M., Pilate, G., Leple, C., Boerjan, W., and Jouanin, L. (2000) *ACS Symposium Series 742, Lignin: Historical, Biological, and Materials Perspectives***,** pp. 145–160
19. Whetten, R., and Sederoff, R. (1991) *For. Ecol. Manage.* **43,** 301–316
20. O'Connell, A., Holt, K., Piquemal, J., Grima-Pettenati, J., Boudet, A., Pollet, B., Lapierre, C., Petit-Conil, M., Schuch, W., and Halpin, C. (2002) *Transgenic Res.* **11,** 495–503
21. Landucci, L. L. (2000) *J. Wood Chem. Technol.* **20,** 243–264
22. Boerjan, W., Meyermans, H., Chen, C., Baucher, M., Van Doorsselaere, J., Morreel, K., Messens, E., Lapierre, C., Pollet, B., Jouanin, L., Leplé, J. C., Ralph, J., Marita, J. M., Guiney, E., Schuch, W., Petit-Conil, M., and Pilate, G. (2001) in *Molecular Breeding of Woody Plants* (Morohoshi, N., and Komamine, A., eds) pp. 187–194, Elsevier Science, Amsterdam
23. Osakabe, K., Tsao, C. C., Li, L., Popko, J. L., Umezawa, T., Carraway, D. T., Smeltzer, R. H., Joshi, C. P., and Chiang, V. L. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 8955–8960
24. Humphreys, J. M., Hemm, M. R., and Chapple, C. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 10045–10050
25. Meyer, K., Shirley, A. M., Cusumano, J. C., Bell-Lelong, D. A., and Chapple, C. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 6619–6623
26. Marita, J., Ralph, J., Hatfield, R. D., and Chapple, C. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 12328–12332
27. Li, L., Zhou, Y., Cheng, X., Sun, J., Marita, J. M., Ralph, J., and Chiang, V. L. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 4939–4944
28. Huntley, S. K., Ellis, D., Gilbert, M., Chapple, C., and Mansfield, S. D. (2003) *J. Agric. Food Chem.* **51,** 6178–6183
29. Baucher, M., Monties, B., Van Montagu, M., and Boerjan, W. (1998) *CRC Crit. Rev. Plant Sci.* **17,** 125–197
30. Ralph, J. (1996) *J. Nat. Prod.* **59,** 341–342
31. Morrison, W. H., Akin, D. E., Archibald, D. D., Dodd, R. B., and Raymer, P. L. (1999) *Industrial Crops and Products* **10,** 21–34
32. Li, L., Popko, J. L., Umezawa, T., and Chiang, V. L. (2000) *J. Biol. Chem.* **275,** 6537–6545
33. Ralph, J., Lapierre, C., Marita, J., Kim, H., Lu, F., Hatfield, R. D., Ralph, S. A., Chapple, C., Franke, R., Hemm, M. R., Van Doorsselaere, J., Sederoff, R. R., O'Malley, D. M., Scott, J. T., MacKay, J. J., Yahiaoui, N., Boudet, A.-M., Pean, M., Pilate, G., Jouanin, L., and Boerjan, W. (2001) *Phytochemistry* **57,** 993–1003
34. Ralph, J., Lapierre, C., Lu, F., Marita, J. M., Pilate, G., Van Doorsselaere, J., Boerjan, W., and Jouanin, L. (2001) *J. Agric. Food Chem.* **49,** 86–91
35. Marita, J. M., Ralph, J., Lapierre, C., Jouanin, L., and Boerjan, W. (2001) *J. Chem. Soc. Perkin Trans.* **1,** 2939–2945
36. Marita, J. M., Ralph, J., Hatfield, R. D., Guo, D., Chen, F., and Dixon, R. A. (2003) *Phytochemistry* **62,** 53–65
37. Sarkanen, K. V., and Hergert, H. L. (1971) in *Lignins. Occurrence, Formation, Structure, and Reactions* (Sarkanen, K. V., and Ludwig, C. H., eds) pp. 43–94, Wiley-Interscience, New York
38. Ralph, J., Hatfield, R. D., Quideau, S., Helm, R. F., Grabber, J. H., and Jung, H.-J. G. (1994) *J. Am. Chem. Soc.* **116,** 9448–9456
39. Timmel, T. E. (1986) *Compression Wood in Gymnosperms*, Springer, Heidelberg, Germany
40. Franke, R., Humphreys, J. M., Hemm, M. R., Denault, J. W., Ruegger, M. O., Cusumano, J. C., and Chapple, C. (2002) *Plant J.* **30,** 33–45
41. Franke, R., Hemm, M. R., Denault, J. W., Ruegger, M. O., Humphreys, J. M., and Chapple, C. (2002) *Plant J.* **30,** 47–59
42. Schoch, G., Goepfert, S., Morant, M., Hehn, A., Meyer, D., Ullmann, P., and Werck-Reichhart, D. (2001) *J. Biol. Chem.* **276,** 36566–36574
43. Fukushima, R. S., and Hatfield, R. D. (2001) *J. Agric. Food Chem.* **49,** 3133–3139
44. Rolando, C., Monties, B., and Lapierre, C. (1992) in *Methods in Lignin Chemistry* (Dence, C. W., and Lin, S. Y., eds) pp. 334–349, Springer-Verlag, Berlin-Heidelberg
45. Lu, F., and Ralph, J. (1997) *J. Agric. Food Chem.* **45,** 4655–4660
46. Lu, F., and Ralph, J. (1999) *J. Agric. Food Chem.* **47,** 1988–1992
47. Hatfield, R. D., and Weimer, P. J. (1995) *J. Sci. Food Agric.* **69,** 185–196
48. Deleted in proof
49. Morreel, K., Ralph, J., Kim, H., Lu, F., Goeminne, G., Ralph, S. A., Messens, E., and Boerjan, W. (2004) *Plant Physiol.* **136,** 3537–3549
50. Ralph, J., Hatfield, R. D., Piquemal, J., Yahiaoui, N., Pean, M., Lapierre, C., and Boudet, A.-M. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 12803–12808
51. Wu, S., and Argyropoulos, D. S. (2003) *J. Pulp Paper Sci.* **29,** 235–240
52. Lu, F., and Ralph, J. (2003) *Plant J.* **35,** 535–544
53. Quideau, S., and Ralph, J. (1994) *Holzforschung* **48,** 12–22
54. Reinhoudt, D. N., Dejong, F., and Vandevondervoort, E. M. (1981) *Tetrahedron* **37,** 1753–1762
55. Quideau, S., and Ralph, J. (1992) *J. Agric. Food Chem.* **40,** 1108–1110
56. Landucci, L. L., Luque, S., and Ralph, S. A. (1995) *J. Wood Chem. Technol.* **15,** 493–513
57. Nakano, J., and Meshitsuka, G. (1992) in *Methods in Lignin Chemistry* (Lin, S. Y., and Dence, C. W., eds) pp. 23–32, Springer-Verlag, Heidelberg, Germany
58. Ralph, J., Marita, J. M., Ralph, S. A., Hatfield, R. D., Lu, F., Ede, R. M., Peng, J., Quideau, S., Helm, R. F., Grabber, J. H., Kim, H., Jimenez-Monteon, G., Zhang, Y., Jung, H.-J. G., Landucci, L. L., MacKay, J. J., Sederoff, R. R., Chapple, C., and Boudet, A. M. (1999) in *Advances in Lignocellulosics Characterization* (Argyropoulos, D. S., and Rials, T., eds) pp. 55–108, TAPPI Press, Atlanta, GA
59. Zhang, L., and Gellerstedt, G. (2001) *Chem. Commun.* **24,** 2744–2745
60. Zhang, L., Gellerstedt, G., Lu, F., and Ralph, J. (2006) *J. Wood Chem. Technol.*, in press
61. Karhunen, P., Rummakko, P., Sipilä, J., Brunow, G., and Kilpeläinen, I. (1995) *Tetrahedron Lett.* **36,** 169–170
62. Hu, W.-J., Lung, J., Harding, S. A., Popko, J. L., Ralph, J., Stokke, D. D., Tsai, C.-J., and Chiang, V. L. (1999) *Nat. Biotechnol.* **17,** 808–812
63. Ralph, J., and Lu, F. (2004) *Org. Biomol. Chem.* **2,** 2714–2715
64. Syrjanen, K., and Brunow, G. (2000) *J. Chem. Soc. Perkin Trans.* **1,** 183–187
65. Kim, H., Ralph, J., Yahiaoui, N., Pean, M., and Boudet, A.-M. (2000) *Org. Lett.* **2,** 2197–2200
66. Kim, H., Ralph, J., Lu, F., Ralph, S. A., Boudet, A.-M., MacKay, J. J., Sederoff, R. R., Ito, T., Kawai, S., Ohashi, H., and Higuchi, T. (2003) *Org. Biomol. Chem.* **1,** 158–281
67. Lapierre, C. (1993) in *Forage Cell Wall Structure and Digestibility* (Jung, H. G., Buxton, D. R., Hatfield, R. D., and Ralph, J., eds) pp. 133–166, American Society of Agronomy-Crop Science Society of America-Soil Science Society of America, Inc., Madison, WI
68. Gang, D. R., Costa, M. A., Fujita, M., Dinkova-Kostova, A. T., Wang, H. B., Burlat, V., Martin, W., Sarkanen, S., Davin, L. B., and Lewis, N. G. (1999) *Chem. Biol.* **6,** 143–151
69. Grabber, J. H., Hatfield, R. D., and Ralph, J. (2003) *J. Agric. Food Chem.* **51,** 4984–4989
70. Kim, H., and Ralph, J. (2005) *J. Agric. Food Chem.* **53,** 3693–3695