

# Profiling of Oligolignols Reveals Monolignol Coupling Conditions in Lignifying Poplar Xylem<sup>1[w]</sup>

Kris Morreel, John Ralph, Hoon Kim, Fachuang Lu, Geert Goeminne, Sally Ralph, Eric Messens, and Wout Boerjan\*

Department of Plant Systems Biology, Flanders Interuniversity Institute for Biotechnology, Ghent University, B-9052 Ghent, Belgium (K.M., G.G., E.M., W.B.); United States Dairy Forage Research Center, Agricultural Research Service, United States Department of Agriculture, and Department of Forestry, University of Wisconsin, Madison, Wisconsin 53706 (J.R., H.K., F.L.); and United States Forest Products Laboratory, United States Department of Agriculture Forest Service, Madison, Wisconsin 53705 (S.R.)

Lignin is an aromatic heteropolymer, abundantly present in the walls of secondary thickened cells. Although much research has been devoted to the structure and composition of the polymer to obtain insight into lignin polymerization, the low-molecular weight oligolignol fraction has escaped a detailed characterization. This fraction, in contrast to the rather inaccessible polymer, is a simple and accessible model that reveals details about the coupling of monolignols, an issue that has raised considerable controversy over the past years. We have profiled the methanol-soluble oligolignol fraction of poplar (*Populus* spp.) xylem, a tissue with extensive lignification. Using liquid chromatography-mass spectrometry, chemical synthesis, and nuclear magnetic resonance, we have elucidated the structures of 38 compounds, most of which were dimers, trimers, and tetramers derived from coniferyl alcohol, sinapyl alcohol, their aldehyde analogs, or vanillin. All structures support the recently challenged random chemical coupling hypothesis for lignin polymerization. Importantly, the structures of two oligomers, each containing a  $\gamma$ -*p*-hydroxybenzoylated syringyl unit, strongly suggest that sinapyl *p*-hydroxybenzoate is an authentic precursor for lignin polymerization in poplar.

Lignin is an aromatic heteropolymer that is mainly present in the walls of secondary thickened cells, where it provides strength and impermeability, allowing transport of water and solutes through the vascular system. There is wide interest in understanding the process of lignin biosynthesis and deposition because of its economic relevance; during chemical pulping, lignin needs to be extracted from the wood chips, a process that is expensive and environmentally hazardous. In addition, lignin limits the digestibility of forages. Hence, plant varieties with altered lignin contents may have improved performance as fodder crops or in the production of pulp and paper (Guo et al., 2001; Pilate et al., 2002; Baucher et al., 2003; Boudet et al., 2003).

In dicotyledonous plants, the lignin polymer is made predominantly from the monolignols coniferyl

and sinapyl alcohol (Baucher et al., 1998), whereas the lignin of gymnosperms, on the other hand, lacks sinapyl alcohol. After their synthesis, the lignin monomers are transported to the cell wall where they are polymerized in a combinatorial fashion by free-radical coupling mechanisms, generating a variety of structures within the polymer (Boerjan et al., 2003; Ralph et al., 2004b).

By means of a number of chemical degradation methods, such as derivatization followed by reductive cleavage (Lu and Ralph, 1997), acidolysis (Lundquist, 1992), and thioacidolysis (Rolando et al., 1992), and spectroscopic techniques such as NMR (Ralph et al., 1999a; Lu and Ralph, 2003) and Fourier-transform infrared spectroscopy (Faix, 1986), the nature of the chemical bonds and their relative abundance in the final polymer has been elucidated (Adler, 1977; Brunow et al., 1999; Ralph et al., 2004b). However, during lignin polymerization, a fraction of lower  $M_r$  phenolic compounds is produced that has escaped a detailed characterization, despite the early use of in vitro dehydropolymerization to obtain low- $M_r$  oligomers for characterization (Freudenberg and Neish, 1968). The study of this plant phenolic fraction is important to better understand lignin polymerization and deposition and to answer some pertinent questions about monolignol coupling in vivo.

The first step in lignin polymerization involves the dehydrogenation of the monolignols by oxidative enzymes, such as peroxidases or laccases, with the formation of radicals (Christensen et al., 2000). Resonance

<sup>1</sup> This work was supported by the Fund for Scientific Research-Flanders (G.0040.00N) and by the European Commission program (QLK5-CT-2000-01493 to E.M. and W.B.), in part by the Department of Energy Biosciences program (DE-AI02-00ER15067) and by the U.S. Department of Agriculture-National Research Initiatives (2001-02176 to J.R.), and by the Instituut voor de aanmoediging van Innovatie door Wetenschap en Technologie in Vlaanderen (predoctoral fellowship to K.M.).

\* Corresponding author; e-mail wout.boerjan@psb.ugent.be; fax 32-9-331-3809.

<sup>[w]</sup>The on-line version of this article contains Web-only data.

Article, publication date, and citation information can be found at [www.plantphysiol.org/cgi/doi/10.1104/pp.104.049304](http://www.plantphysiol.org/cgi/doi/10.1104/pp.104.049304).

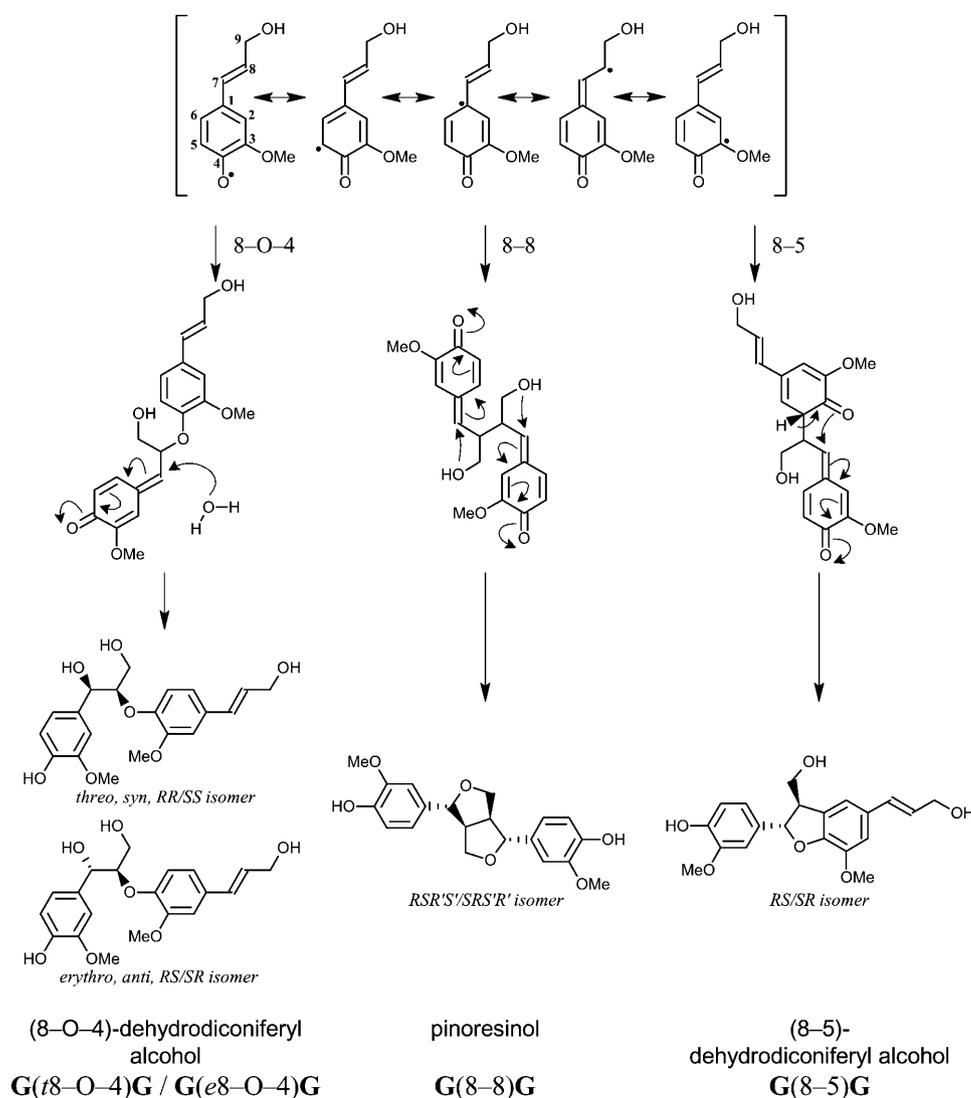
stabilizes the radical with unpaired electron density at the C<sub>1</sub>, C<sub>3</sub>, 4-O, C<sub>5</sub>, and C<sub>8</sub> positions (Fig. 1). According to the conventional random coupling hypothesis, the monomeric radicals couple according to their relative supply and coupling propensities, and these reactions are influenced by the macromolecular environment of the cell wall, finally leading to a racemic polymer (Freudenberg and Neish, 1968; Grabber et al., 1996; Ralph et al., 1999b, 2004b; Syrjänen and Brunow, 2000).

A new class of so-called dirigent proteins that are capable of guiding the stereospecific coupling of two coniferyl alcohol radicals into the lignan (+)-pinosresinol has been described (Davin et al., 1997). Lignans are typically optically active compounds, thought to serve as defense substances in plants and derived from the very same monolignols used to generate the lignin polymer (Lewis and Davin, 1999; Sakakibara et al., 2003; Umezawa, 2004). However, the discovery of dirigent proteins has led to a controversial, but widespread, hypothesis that lignin polymerization is tightly

controlled by protein-mediated coupling reactions (Lewis, 1999; Chen and Sarkanen, 2004). Although only one dirigent protein, catalyzing the formation of (+)-pinosresinol, has been functionally characterized so far (Davin et al., 1997), the large size of the *DIRIGENT* gene families in a variety of species has been used as an argument that the other linkages in lignin are also protein mediated (Lewis, 1999), although no functional proof has supported this hypothesis yet. A dirigent protein has been immunolocalized to the cambial region and the cell wall (Burlat et al., 2001) and a *DIRIGENT* gene is highly expressed in the lignifying zone of poplar (*Populus* spp.; Hertzberg et al., 2001), corroborating an important role for dirigent proteins in these tissues.

As a first step in deepening our understanding of monolignol coupling and polymerization, and in discriminating lignin from lignan biosynthesis, we reasoned that the structures of low-*M<sub>r</sub>* oligolignols should reflect the *in vivo* coupling conditions. Hence, we

**Figure 1.** Dilignol formation. Radical-radical coupling involving the C<sub>8</sub> position of a coniferyl alcohol and the 4-O, C<sub>8</sub>, or C<sub>5</sub> position of another coniferyl alcohol leading to 8-O-4, 8-8, and 8-5 linkages. The quinone methide intermediates are subsequently rearomatized to (8-O-4)-dehydrodiconiferyl alcohol, pinosresinol, and (8-5)-dehydrodiconiferyl alcohol. The R/S nomenclature specifies carbons 7 and 8, namely RS = 7R,8S.



characterized this fraction in poplar xylem, a tissue that is heavily lignified. In transgenic poplars with reduced lignin content, this oligolignol fraction was severely depleted. We identified the structures of 38 phenolic compounds, most of which were dimeric, trimeric, or tetrameric oligolignols derived from cinneryl and sinapyl alcohols and their aldehydes. In addition, the structures of two compounds demonstrate that sinapyl *p*-hydroxybenzoate has to be considered as an authentic lignin precursor in poplar. The structures of all identified compounds are in accordance with the recently challenged combinatorial coupling hypothesis. This is the first study to our knowledge describing the low- $M_r$  oligolignol fraction from lignifying tissue.

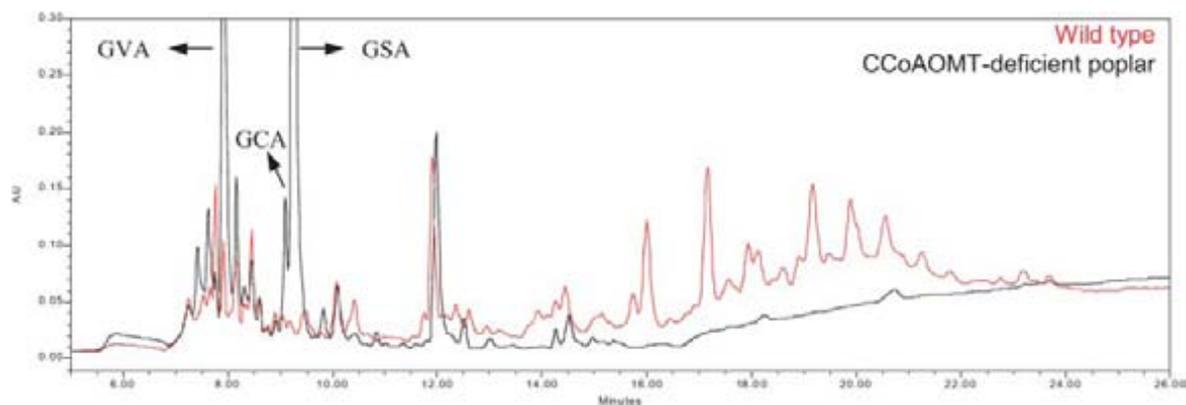
## RESULTS

### Characterization of Oligolignols from Lignifying Poplar Xylem

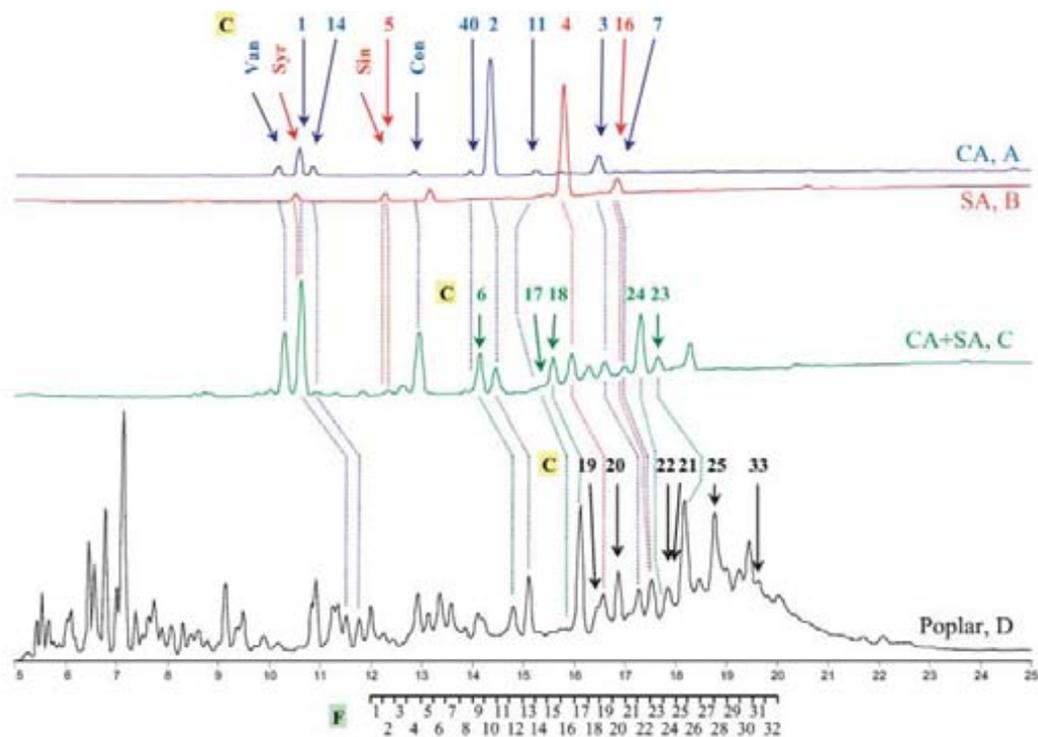
Our aim was to obtain insight into the process of monolignol coupling in the cell wall by characterizing the chemical structures of a large number of low- $M_r$  monolignol-coupling products, and to investigate whether these structures are consistent with a combinatorial coupling process under chemical control. Because monolignol coupling occurs excessively during lignin polymerization, such a low- $M_r$  oligolignol fraction is expected to be present in the walls of lignifying cells. To identify this oligolignol fraction, we profiled the methanol-soluble phenolics present in xylem extracts of wild-type and caffeoyl-CoA *O*-methyltransferase (CCoAOMT)-deficient poplars by HPLC (Fig. 2). Because xylem of the latter plants accumulates less lignin (Meyermans et al., 2000; Zhong et al., 2000), it is an ideal material to identify this fraction, because the oligolignols are expected to be less abundant. Indeed, in the last half of the chromatogram between 11 and 24 min (Fig. 2), a family of compounds abundantly present in wild-type poplar

was barely detectable in HPLC profiles of poplars down-regulated for CCoAOMT, suggesting that their synthesis involved the monolignol biosynthesis pathway. A similar HPLC profile, showing a depletion of peaks in the second half of the chromatogram was also observed for transgenic poplars down-regulated for cinnamoyl-CoA reductase (J.-C. Leplé, K. Morreel, C. Lapierre, K. Ruel, J.-P. Joseleau, G. Goeminne, R. De Rycke, E. Messens, G. Pilate, and W. Boerjan, unpublished data). All these compounds had similar UV/visible (Vis) adsorption spectra with a maximum at approximately 270 nm. Interestingly, although the large fraction of peaks in the last half of the chromatogram was almost absent in the xylem of CCoAOMT-deficient plants, the total peak height (the sum of the heights of all peaks in a chromatogram, divided by the dry weight of xylem tissue) was 2.4-fold higher than that of the wild type. This increase could be attributed primarily to three newly accumulating compounds (the phenolic glucosides of vanillin, caffeic acid, and sinapic acid) that had been identified previously (Meyermans et al., 2000). These three glucosides are thought to be detoxification and/or storage products of the free acids that accumulate as a consequence of a redirection of the flux of caffeoyl-CoA to caffeic acid and further to sinapic acid, rather than to feruloyl-CoA (Meyermans et al., 2000).

To identify the structures of the presumed oligolignols in wild-type poplar, HPLC fractions (Fig. 3, chromatogram D and table) of the complete set of peaks in the second half of the chromatogram were collected and separated on liquid chromatography-mass spectrometry (LC-MS) for further structural elucidation. By mass spectrometry/mass spectrometry (MS/MS) analysis, a tentative structure was proposed. The assigned structure for a number of peaks could be authenticated by spiking and MS/MS analysis of the synthesized compound. Several compounds were trivially assigned by analogy with confirmed peaks, solely from their mass spectral data. As presented in Figure 4, 38 oligolignols were authen-



**Figure 2.** HPLC chromatogram overlay. Chromatograms from xylem extracts of a wild-type poplar (red) and a transgenic CCoAOMT-deficient poplar (black) are shown. GVA,  $O^4$ - $\beta$ -D-glucopyranosyl-vanillic acid; GCA,  $O^3$ - $\beta$ -D-glucopyranosyl-caffeic acid; GSA,  $O^4$ - $\beta$ -D-glucopyranosyl-sinapic acid.



C	M	t <sub>R</sub>	F	Conc		Shorthand name	Trivial name
				m	sem		
1	✓	11.5		31.3	12	G(8-O-4)G	<i>threo</i> -(8-O-4)-dehydrodiconiferyl alcohol
2	✓	15.1		190	70	G(8-5)G	(8-5)-dehydrodiconiferyl alcohol
3	✓	17.3		70.2	26	G(8-8)G	pinoresinol
4	✓	16.6		29.1	7	S(8-8)S	syringaresinol
5	✓		6			S( <i>trans</i> -O-4)S	<i>threo</i> -(8-O-4)-dehydrodisinapyl alcohol
6	✓	14.8		57.3	18	S(8-5)G	simulanol
7	✓	17.5		13.0	4	G(8-5)G'	balanophonin
8			5			G( <i>trans</i> -O-4)S	<i>threo</i> -guaiacylglycerol 8-O-4 sinapyl ether
9			12			S( <i>trans</i> -O-4)S'	
10			15			G( <i>trans</i> -O-4)S' / S( <i>trans</i> -O-4)G'	
11	✓		15			G(8-5)V'	ficusal
12			21			S(8-5)G'	
13			22			G(8-O-4)G'	
14	✓	11.8		21.2	9	G( <i>trans</i> -O-4)G( <i>trans</i> -O-4)G	
15			18			G( <i>trans</i> -O-4)G(8-8)G	
16	✓	17.4		n.d.	n.d.	S( <i>trans</i> -O-4)S(8-8)S	<i>threo</i> -buddlenol D
17	✓	15.8		10.5	3	S( <i>trans</i> -O-4)S(8-5)G	
18	✓	16.1		168	40	G( <i>trans</i> -O-4)S(8-5)G	<i>threo</i> -buddlenol B
19		16.4		21.0	13	SP(8-8)S	
20	✓	16.9		55.3	16	G( <i>trans</i> -O-4)S(8-5)G	<i>erythro</i> -buddlenol B
21		17.8		31.2	3	S( <i>trans</i> -O-4)S(8-5)G'	
22		18.0		n.d.	n.d.	G( <i>trans</i> -O-4)S(8-8)G	<i>threo</i> -buddlenol E
23	✓	18.2		123	15	G( <i>trans</i> -O-4)S(8-5)G'	<i>threo</i> -buddlenol A
24	✓	18.3		n.d.	n.d.	G( <i>trans</i> -O-4)S(8-8)S	<i>threo</i> -buddlenol C
25	✓	18.8		85.5	10	G( <i>trans</i> -O-4)S(8-5)G'	<i>erythro</i> -buddlenol A
26			16			G(8-5)G( <i>trans</i> -O-4)G'	
27			17			G( <i>trans</i> -O-4)G(8-5)V'	
28			20			G( <i>trans</i> -O-4)G(8-5)G'	
29			21			G( <i>trans</i> -O-4)G(8-5)G'	
30			22			S( <i>trans</i> -O-4)S(8-5)G	
31			27			G( <i>trans</i> -O-4)S'(8-8)S / G( <i>trans</i> -O-4)S(8-8)S'	
32			27			G( <i>trans</i> -O-4)S(8-8)S	<i>erythro</i> -buddlenol C
33	✓	19.6		62.2	4	G( <i>trans</i> -O-4)S(8-8)S(8-O-4)G	hedyotisol
34			22			S( <i>trans</i> -O-4)SP(8-8)S / S( <i>trans</i> -O-4)S(8-8)SP	
35			27			G( <i>trans</i> -O-4)G( <i>trans</i> -O-4)S(8-8)S' / G( <i>trans</i> -O-4)G( <i>trans</i> -O-4)S'(8-8)S	
36			22			G( <i>trans</i> -O-4)G(8-8)S(8-O-4)G	
37			25			S(8-O-4)S(8-8)S(8-O-4)G	
38			23			G( <i>trans</i> -O-4)G(8-O-4)S(8-8)S	
39						S( <i>trans</i> -O-4)G(8-5)G	
40	✓					G( <i>trans</i> -O-4)G(8-5)G	

Figure 3. (Legend appears on following page.)

ticated or tentatively identified in this way (see supplemental data, available at [www.plantphysiol.org](http://www.plantphysiol.org), for the MS/MS spectra of all identified compounds and the arguments for the assignment of a particular structure).

The oligomers were mainly composed of units derived from coniferyl alcohol (guaiacyl, **G**) from sinapyl alcohol (syringyl, **S**) and from coniferaldehyde (**G'**), and a few contained units derived from sinapaldehyde (**S'**), vanillin (**V'**), and sinapyl *p*-hydroxybenzoate (**SP**; for nomenclature, see "Materials and Methods"). All possible 8-O-4-, 8-5-, and 8-8-coupled homodimers of **G** and **S** units (Fig. 4), i.e. **G**(*t*8-O-4)**G**, **1**; **G**(8-5)**G**, **2**; **G**(8-8)**G**, **3**; **S**(8-8)**S**, **4**; and **S**(*t*8-O-4)**S**, **5**, were detected. In addition to **G** and **S** units, alternative units, such as **G'**, **S'**, **V'**, and **SP** were found in the heterodimeric fraction, i.e. **S**(8-5)**G**, **6**; **G**(8-5)**G'**, **7**; **G**(*t*8-O-4)**S**, **8**; **S**(8-O-4)**S'**, **9**; **G**(8-O-4)**S'** or **S**(8-O-4)**G'** (compound **10**); **G**(8-5)**V'**, **11**; **S**(8-5)**G'**, **12**; **G**(8-O-4)**G'**, **13**; and **SP**(8-8)**S**, **19**. Homodimers of **G'**, **S'**, **V'**, and **SP** were not detected, presumably because of the low abundance of their precursor monomers. Except for **G**(8-5)**G**(*t*8-O-4)**G'**, **26**, all tri- and tetrameric compounds were composed of a **G** or **S** unit linked by a  $\beta$ -aryl ether bond to a moiety derived from one of the dimers mentioned above (Figs. 3 and 4), or to an **S**(8-8)**G** or **S**(8-8)**S'** moiety, i.e. **G**(*t*8-O-4)**G**(*t*8-O-4)**G**, **14**; **G**(*t*8-O-4)**G**(8-8)**G**, **15**; **S**(*t*8-O-4)**S**(8-8)**S**, **16**; **S**(*t*8-O-4)**S**(8-5)**G**, **17**; **G**(*t*8-O-4)**S**(8-5)**G**, **18**; **G**(*e*8-O-4)**S**(8-5)**G**, **20**; **S**(*t*8-O-4)**S**(8-5)**G'**, **21**; **G**(*t*8-O-4)**S**(8-8)**G**, **22**; **G**(*t*8-O-4)**S**(8-5)**G'**, **23**; **G**(*t*8-O-4)**S**(8-8)**S**, **24**; **G**(*e*8-O-4)**S**(8-5)**G'**, **25**; **G**(*t*8-O-4)**G**(8-5)**V'**, **27**; **G**(*t*8-O-4)**G**(8-5)**G'**, **28**; **G**(*e*8-O-4)**G**(8-5)**G'**, **29**; **G**(*t*8-O-4)**S**(8-8)**S'** or **G**(*t*8-O-4)**S'**(8-8)**S** (compound **31**); **G**(*e*8-O-4)**S**(8-8)**S**, **32**; **G**(8-O-4)**S**(8-8)**S**(8-O-4)**G**, **33**; **S**(*t*8-O-4)**SP**(8-8)**S** or **S**(*t*8-O-4)**S**(8-8)**SP** (compound **34**); **G**(*e*8-O-4)**G**(8-O-4)**S'**(8-8)**S** or **G**(*e*8-O-4)**G**(8-O-4)**S**(8-8)**S'** (compound **35**); **G**(8-O-4)**G**(8-8)**S**(8-O-4)**G**, **36**; **S**(8-O-4)**S**(8-8)**S**(8-O-4)**G**, **37**; and **G**(*t*8-O-4)**G**(8-O-4)**S**(8-8)**S**, **38**. Compound **30**, **S**(8-O-4\*)**S**(8-5)**G**, is likely formed by simple benzylic oxidation of the trimer **S**(8-O-4)**S**(8-5)**G**, **17**.

All detected tetramers were derived from an 8-8-dimeric moiety to which **G** and/or **S** units were attached. The biosynthesis of these compounds is initiated by monomer-monomer coupling, yielding an 8-8-linked dimer with two phenolic groups that are amenable to oxidation by peroxidase/H<sub>2</sub>O<sub>2</sub>, for instance. Hence, further chain extension might be

initiated at either phenol of this dimer, yielding tetramers characterized by an internal 8-8-linked unit or an 8-8-linked end group. Higher order oligomers (pentamers, hexamers, etc.) might be present in poplar xylem as well; some of the corresponding *m/z* values were found by LC-MS analysis, but they were present in minute amounts.

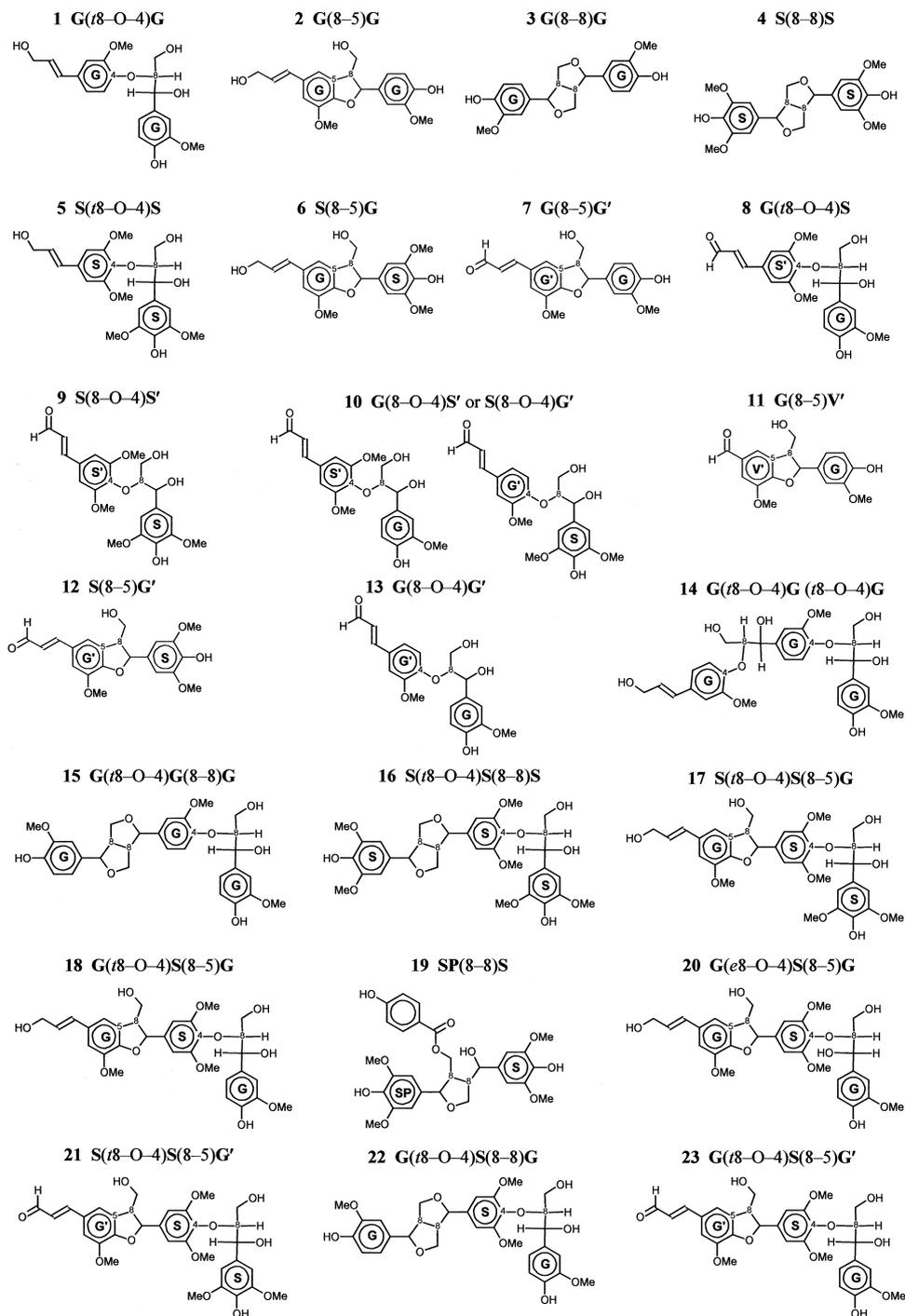
If the production of these oligolignols solely depended on the chemical coupling conditions in the cell wall, their concentrations would be in accordance with the relative supply and cross-coupling propensities of the monomers. Therefore, the concentrations were estimated for the identified oligolignols based on the HPLC chromatograms of the xylem extracts (Fig. 3). Fifteen of the identified compounds were separated sufficiently and abundantly allowing their pseudo-quantification. Together, these 15 oligolignols accounted for approximately 0.05% of the dry weight of xylem tissue. The major detected dilignol was (8-5)-dehydroconiferyl alcohol, **G**(8-5)**G**, **2**, whereas the major trilignols were *threo*-buddlenol B, **G**(*t*8-O-4)**S**(8-5)**G**, **18**, and its corresponding cinnamaldehyde, i.e. *threo*-buddlenol A, **G**(*t*8-O-4)**S**(8-5)**G'**, **23**. The *erythro*-isomers of these trilignols accounted for 25% and 42% of the total amounts (*threo* + *erythro*) of buddlenol B and A, respectively. The only tetralignol that could be quantified was **G**(8-O-4)**S**(8-8)**S**(8-O-4)**G**, **33**. Overall, taking the concentrations into account, the quantified oligolignols were composed mainly of **G** (59%), **S** (31%), and **G'** (10%) units and traces of **V'**, **S'**, and **SP** units and were linked by 8-5 (47%), 8-O-4 (42%), and 8-8 (11%) bonds. No **H** units were detected in any of the coupling products.

### Oligolignol Profiling of Synthetic Mixtures

Our hypothesis is that the oligolignols are derived from phenolic units through oxidation, followed by chemical coupling that is not protein mediated. Thus, synthetic reaction mixtures, prepared by the oxidation of coniferyl alcohol, sinapyl alcohol, or both coniferyl and sinapyl alcohols, resulting in **G**, **S**, or **G** + **S** synthetic oligolignol mixtures, respectively, should reveal the same oligolignol structures as those detected in poplar xylem extracts. These oligolignol mixtures were prepared and separated with the same reversed-phase HPLC method and compared to the oligolignol profiles obtained from the poplar xylem extracts (Fig. 3).

**Figure 3.** Oligolignol profile. HPLC chromatograms of oligomers present in mixtures obtained from the oxidation of coniferyl alcohol (CA; A), sinapyl alcohol (SA; B), and coniferyl and sinapyl alcohols (CA + SA; C), and of a methanol xylem extract from poplar (poplar; D). Oligomers identified in chromatograms A to D are shown in blue, red, green, and black, respectively. Vertical dashed lines indicate the corresponding positions of the identified peaks in the different chromatograms. Isolated HPLC fractions (F) for LC-MS/MS analysis are shown below the retention time. Compound number (C), whether present or not in the **G** + **S** synthetic mixture (M), retention time (*t<sub>R</sub>*), or HPLC fraction where the compound was detected, concentrations (Conc; pmol mg<sup>-1</sup> dry weight of xylem tissue), shorthand, and trivial names of the oligolignols are given below. Retention times are based on the chromatogram of the xylem extract (D). m, mean; n.d., not determined due to coelution; sem, SE of the mean. The presence of **G**(*e*8-O-4)**S**(8-5)**G**, **20**; **G**(*e*8-O-4)**S**(8-5)**G'**, **25**; and **G**(8-O-4)**S**(8-8)**S**(8-O-4)**G**, **33** in the synthetic mixtures was only established by NMR analysis following their purification.

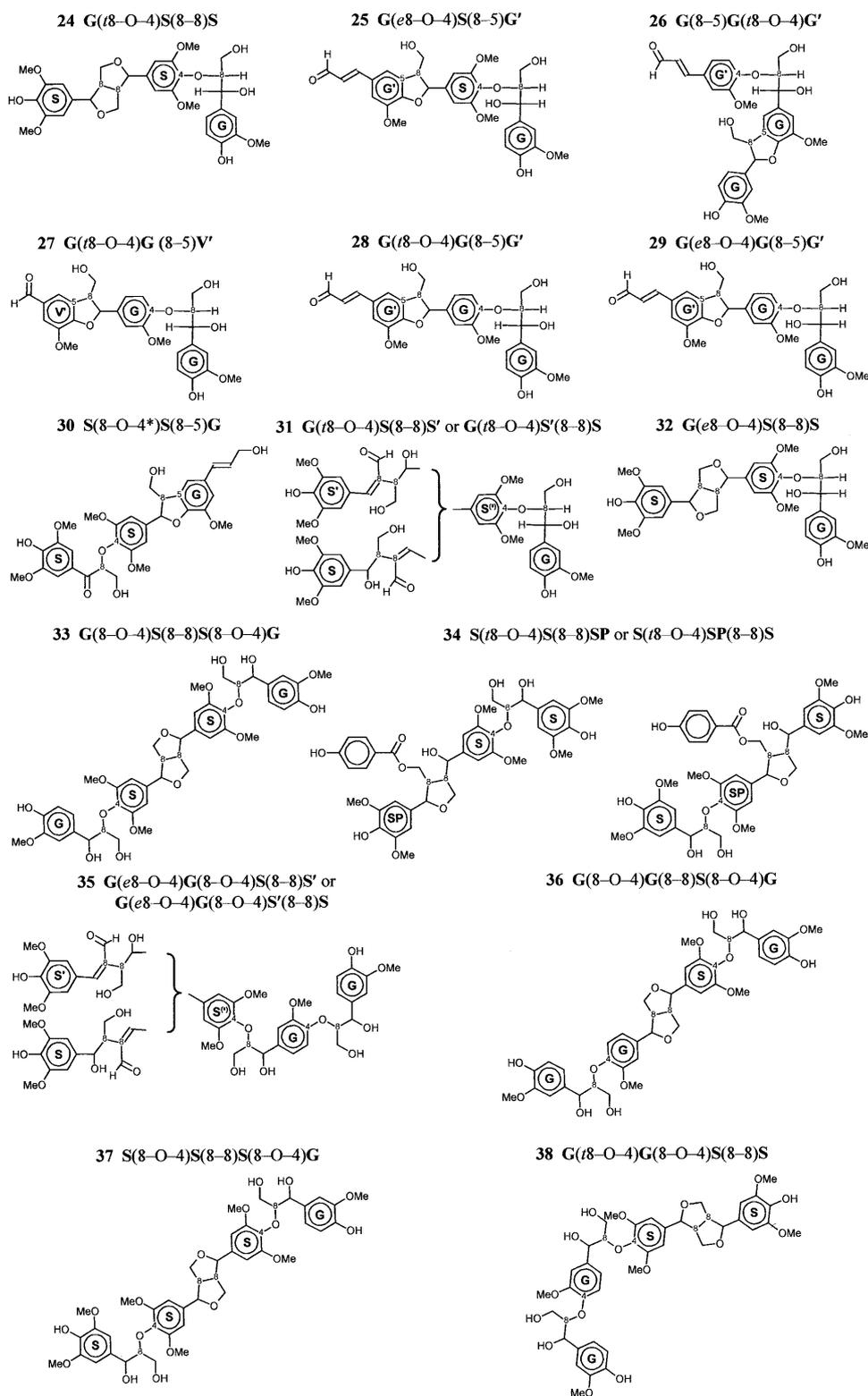
**Figure 4.** Oligolignol structures present in poplar xylem extracts. Molecular structure and shorthand naming inferred from the spectra are shown. (Figure continues on following page.)



The chromatogram of the synthetic **G** oligolignol mixture showed the main types of dimerization products involving the  $C_8$  position, i.e.  $G(t8-O-4)G$ , **1**,  $G(8-5)G$ , **2**, and  $G(8-8)G$ , **3**. In accordance with the *in vivo* situation, two phenylcoumaran dimers were detected for which a **G** unit was connected to a unit derived from coniferaldehyde or vanillin, i.e.  $G(8-5)G'$ , **7** and  $G(8-5)V'$ , **11**. Although the coniferyl alcohol used for the synthetic mixture was virtually pure, both free coniferaldehyde and vanillin were

present as well in the **G** oligolignol mixture based on their MS/MS spectra and the spiking of synthetic products. This indicates that coniferyl alcohol is oxidized to aldehydes under the synthetic conditions. Two trimers were found, namely  $G(t8-O-4)G(t8-O-4)G$ , **14** and  $G(t8-O-4)G(8-5)G$  (compound **40**; Fig. 3), the latter of which was not detected in the xylem extracts. No higher order oligomers were detected. Oligolignol units were mainly 8-5 linked (Fig. 3).

Figure 4. (Continued.)



HPLC analysis of the synthetic **S** oligolignol mixture showed the presence of both **S(8-8)S**, **4** and **S(t8-O-4)S**, **5** dimers, and only one trimer, **S(t8-O-4)S(8-8)S**, **16**. **S(8-8)S**, **4** was the major compound in this synthetic oligolignol mixture (Fig. 3).

By MS/MS analysis and the spiking of standards, two peaks were identified as free sinapaldehyde and syringaldehyde, although the oligolignol mixture was prepared starting from virtually pure sinapyl alcohol.

All compounds identified in the **G** or **S** oligolignol mixtures were also found in the synthetic **G** + **S** oligolignol mixture, which, in addition, contained the **S**(8–5)**G**, **6** dimer, its  $\beta$ -aryl ether-derived trimers, i.e. **S**(*t*8–O–4)**S**(8–5)**G**, **17** and **G**(*t*8–O–4)**S**(8–5)**G**, **18**, the corresponding aldehyde of the latter trimer, **G**(*t*8–O–4)**S**(8–5)**G'**, **23**, and the trimer where a **G** unit is connected via a  $\beta$ -aryl ether to a syringaresinol substructure, **G**(*t*8–O–4)**S**(8–8)**S**, **24**. The latter was the most abundant oligolignol present in the **G** + **S** synthetic mix (Fig. 3).

All identified peaks in the synthetic mixtures were detected in poplar xylem extracts, except for the poorly abundant **G**(*t*8–O–4)**G**(8–5)**G**, **40**, vanillin, and syringaldehyde. Compared to the synthetic mixtures, xylem extracts contained some additional oligolignols, especially tri- and tetralignols (Figs. 3 and 4). The *erythro*-diastereomers of the more abundant xylem oligolignols, such as **G**(*e*8–O–4)**S**(8–5)**G**, **20** and **G**(*e*8–O–4)**S**(8–5)**G'**, **25**, and the tetralignol **G**(8–O–4)**S**(8–8)**S**(8–O–4)**G**, **33**, were clearly detected in the HPLC profiles of xylem extracts, but their presence in the synthetic mixtures was only established by NMR analysis following their purification as *threo/erythro* mixtures; *threo*- $\beta$ -aryl ethers are strongly favored in the borate buffer system used (Landucci et al., 1995).

## DISCUSSION

### The Oligolignol Structures of Poplar Xylem Extracts Are in Agreement with Chemical Coupling Reactions

We have characterized the methanol-soluble oligolignol fraction of poplar xylem to investigate whether their structures are consistent with a chemical coupling process. UV/Vis and MS/MS spectra were used for the initial elucidation of the structure of these compounds. For many of them, the proposed structures were subsequently validated by spiking and MS/MS analysis of synthesized reference compounds. We have authenticated or tentatively identified the structures of 38 compounds, most of which correspond to simple coupling products of monolignols, including dimers, trimers, and tetramers. All structures suggest they correspond with products of radical coupling reactions, and no further modifications invoking enzymatic reactions were evidenced.

The high frequency of  $\beta$ -aryl ether units in trimers and tetramers (all composed of units 8–O–4 linked to an 8–O–4-, 8–5-, or 8–8-coupled dimer) is in agreement with the chemical cross-coupling reactions between a monomer and an oligomer. There are only two possibilities for a hydroxycinnamyl alcohol to couple at its favored C<sub>8</sub> position with a **G** phenolic end group (at its 4–O or C<sub>5</sub> position) and only one for coupling with an **S** phenolic end group (at its 4–O position) (Boerjan et al., 2003; Ralph et al., 2004b).

A survey of the  $\beta$ -aryl ether units in the trimers and tetramers shows that 8–O–4 coupling occurs between coniferyl alcohol and both **G** and **S** units, whereas

sinapyl alcohol forms only 8–O–4 linkages to **S** units; this observation is again in agreement with a non-protein-mediated chemical coupling reaction, where the 8–O–4 radical coupling propensities do not favor a reaction between sinapyl alcohol and a **G** unit because of factors, such as oxidation potential and radical reactivity (Landucci et al., 1992; Syrjänen and Brunow, 1998). To further support this hypothesis, the compound **S**(*t*8–O–4)**G**(8–5)**G**, **39**, which is the **S**-type  $\beta$ -aryl ether of the most abundant dimer in the xylem extract, **G**(8–5)**G**, **2**, was synthesized and searched for by both HPLC analysis and LC-MS analysis of isolated HPLC fractions. This compound, **S**(*t*8–O–4)**G**(8–5)**G**, **39** was found neither in poplar xylem extracts nor in synthetic oligolignol mixtures (see below).

### The Structures of the Oligolignols Are in Agreement with Endwise Polymerization Conditions in the Cell Wall

As is the case for lignification, the oligolignols described here are likely produced by an endwise rather than by a bulk polymerization process. Zulauf-verfahren dehydrogenation polymer (DHP) reactions, which mimic a bulk polymerization process (Freudenberg, 1959), have shown that monolignol radicals prefer to couple with like monolignol radicals rather than to form heterodimers or to cross-couple with dimers or higher oligomers. Hence, the detection of both heterodimers and heterooligomers in the xylem oligolignol fraction supports an endwise polymerization process. Furthermore, bulk polymerization results in oligomer-oligomer couplings, producing 5–5- and 4–O–5-linked structures (Sarkanen, 1971), which are not detected in the xylem oligolignol mixture. On the other hand, endwise polymerization, mimicked by Zutropf-verfahren DHP reactions, results from the gradual supply of monomers to the site of polymerization and represses especially the 8–8-coupling mode (Grabber et al., 1996; Syrjänen and Brunow, 2000). In the xylem oligolignol fraction, only 11% of the linkages were 8–8, a frequency that is in agreement with Zutropf-verfahren DHP reactions. Together, the oligolignol structures in the xylem extract are in agreement with coupling conditions favoring endwise coupling.

### The $\beta$ -Aryl Ether Units in Oligolignols Are Mainly *threo*-Diastereomers

The characterization of the oligolignols present in poplar xylem showed that 14 of the 25 dimeric, trimeric, and terminal tetrameric  $\beta$ -aryl ethers were present only in the *threo* configuration, whereas the  $\beta$ -aryl end group of compound **35**, **G**(*e*8–O–4)**G**(8–O–4)**S**(8–8)**S'** or **G**(*e*8–O–4)**G**(8–O–4)**S'**(8–8)**S**, was present in the *erythro* form. The *threo/erythro* configuration could not be determined from the MS/MS spectra of **S**(8–O–4)**S'**, **9**; **G**(8–O–4)**S'** or **S**(8–O–4)**G'**, **10**; and **G**(8–O–4)**G'**, **13**, and from the MS/MS spectra of the tetralignols with an internal 8–8-linked moiety, i.e. **G**(8–O–4)**S**(8–8)**S**(8–O–4)**G**, **33**; **G**(8–O–4)**G**(8–8)**S**(8–

O-4)G, 36; and S(8-O-4)S(8-8)S(8-O-4)G, 37. Both *threo*- and *erythro*-diastereomers were detected for the remaining 4 of the 25  $\beta$ -aryl ethers, i.e. for the structures G(*t*8-O-4)S(8-8)S, 24 and G(*e*8-O-4)S(8-8)S, 32, and G(*t*8-O-4)G(8-5)G', 28 and G(*e*8-O-4)G(8-5)G', 29, and the more abundant trilignols, i.e. G(*t*8-O-4)S(8-5)G, 18 and G(*e*8-O-4)S(8-5)G, 20, G(*t*8-O-4)S(8-5)G', 23 and G(*e*8-O-4)S(8-5)G', 25. These four compounds could be clearly quantified in the HPLC profile of xylem extracts; the *threo* forms were present for 58% and 75% of the total amount of G(*t*8-O-4)S(8-5)G, 18 and G(*e*8-O-4)S(8-5)G, 20; and G(*t*8-O-4)S(8-5)G', 23 and G(*e*8-O-4)S(8-5)G', 25 with *threo/erythro* ratios of 3:2 and 3:1, respectively. Taken together, the *threo*-diastereomers are clearly the most present among the  $\beta$ -aryl ethers.

Both *threo*- and *erythro*- $\beta$ -aryl ethers are also found in lignin, but, in contrast to the xylem-extracted oligolignols, the *erythro* forms of the  $\beta$ -aryl ether linkages predominate in angiosperm lignins (Brunow et al., 1993; Akiyama et al., 2003). Because gymnosperm lignins, composed of G units, contain approximately equal amounts of *threo* and *erythro* configurations and because the *threo/erythro* ratios correlate inversely with the S:G ratios in dicots (Akiyama et al., 2003), the preponderance of the *erythro* form has been attributed to the presence of  $\beta$ -syringyl ether structures in angiosperm lignins. Both in vivo and in vitro, 8-O-4-guaiacyl ethers and 8-O-4-syringyl ethers in lignin are produced with approximately 50:50 and approximately 25:75 *threo/erythro* ratios, respectively, whereas their equilibrium ratios are nearly equal (Brunow et al., 1993; Ralph et al., 2004b). The reason for the apparent *threo*-isomer predominance in the xylem oligolignol fraction is currently not clear, unless for some reason *erythro*-isomers couple (to higher oligomers) more rapidly.

### Radical Coupling Reactions Accept Alternative Monomers

Besides traditional G and S units, some oligolignols contain alternative units, such as G', S', V', and SP. Importantly, the structures of a few of these oligolignols, namely the trimers, imply that these alternative units arise from the coupling of the corresponding monomers rather than from postcoupling oxidation or derivatization reactions.

For example, cross-coupling of sinapaldehyde appears to result in S(8-8)S', which, after further coupling with sinapyl alcohol, results in compound 31, i.e. G(*t*8-O-4)S(8-8)S' or G(*t*8-O-4)S'(8-8)S and compound 35, i.e. G(*e*8-O-4)G(8-O-4)S'(8-8)S or G(*e*8-O-4)G(8-O-4)S(8-8)S' (Fig. 4). In contrast to the 8-8 coupling of two sinapyl alcohol radicals to S(8-8)S 4, with two tetrahydrofuran rings, no ring structures are formed during the 8-8 coupling of sinapyl alcohol with a cinnamaldehyde. The 8-8 coupling between two sinapyl alcohol radicals forms a bis-quinone methide intermediate. Each quinone methide is rearomatized by internal nucleophilic attack of the 9-OH of the other unit resulting in a resinol unit (Fig. 1).

However, when one of the C<sub>9</sub> positions in the dimer is oxidized or derivatized, it is no longer available to trap the quinone methide of the other unit. In this case, rearomatization of the other unit can only proceed by the nucleophilic attack of, for example, an incoming water molecule and no tetrahydrofuran ring is formed (Lu and Ralph, 2002). The quinone methide derived from the sinapaldehyde unit is rearomatized by the elimination of the C<sub>8</sub> proton, regenerating the enone function, rather than by a nucleophilic attack at C<sub>7</sub>. It should be noted that attempted cross-coupling of coniferaldehyde or sinapaldehyde with normal monolignols did not result in any (G' or S')(8-O-4)(G or S) products; the only cross-product isolated was S(8-O-4)S' (H. Kim, unpublished data). Therefore, structures 31 and 35 remain unauthenticated, but, if correct, indicate that sinapaldehyde is involved directly in the coupling reactions, i.e. that sinapaldehyde is the monomer for this moiety, and not sinapyl alcohol.

The cinnamaldehyde monomers themselves can either be the reaction products of cinnamoyl-CoA reductase, which are transported to the cell wall as aldehydes, or be derived from the precoupling oxidation of the cinnamyl alcohols already present in the cell wall. Because the cinnamaldehyde-derived units are higher in lignins of transgenic plants down-regulated for cinnamyl alcohol dehydrogenase than those in lignins of wild-type plants (Baucher et al., 1996; Kim et al., 2003), these units in the xylem oligolignol fraction are probably made from coniferaldehyde that is synthesized within the cell and transported to the cell wall. However, we cannot exclude that at least part of the G', S', and V' units are derived from the oxidation of the monolignols in the cell wall prior to cross-coupling. For example, although difficult to extrapolate to the in vivo situation, coniferaldehyde, sinapaldehyde, and vanillin were also found in the synthetic mixtures made from coniferyl and sinapyl alcohols, in addition to the coupling products; peroxidase/H<sub>2</sub>O<sub>2</sub> also causes such oxidation.

### Sinapyl *p*-Hydroxybenzoates Are Precursors in Lignin Biosynthesis

More compelling evidence for the incorporation of alternative units is obtained by the identification of compounds 19 and 34, i.e. SP(8-8)S and its sinapyl alcohol coupling product S(*t*8-O-4)SP(8-8)S or S(*t*8-O-4)S(8-8)SP. These results strongly indicate that *p*-hydroxybenzoic acid is esterified by sinapyl alcohol prior to radical cross-coupling, because the product is clearly derived from cross-coupling of sinapyl *p*-hydroxybenzoate with sinapyl alcohol (Lu et al., 2004). Alternatively, SP(8-8)S is formed by ring opening of S(8-8)S, followed by the acylation with *p*-hydroxybenzoic acid. This process would require two enzymatic activities, which is less likely, given that the

radical-radical coupling reactions for lignification occur in the cell wall.

$\gamma$ -Acylated **G** and **S** units have been detected in the lignins of many species. For example, sinapyl acetate is implicated similarly as a monomer in lignification in kenaf bast fibers (Lu and Ralph, 2002). The lignin of grasses is adorned with  $\gamma$ -*p*-coumarate substituents on a variety of lignin units. Lignin structural analyses of poplar, aspen, willow, and palm have shown that  $\gamma$ -*p*-hydroxybenzoate esters are present (Smith, 1955; Nakano et al., 1961; Sun et al., 1999; Meyermans et al., 2000; Landucci and Ralph, 2001). The fact that only **S** units are  $\gamma$ -*p*-hydroxybenzoylated in both poplar lignin and the oligolignol fraction supports the contention that sinapyl *p*-hydroxybenzoate is produced enzymatically and used as an authentic monomer for lignification in poplar.

### Are These Oligomers Destined for Lignin?

Taken together, the dilignols, trilignols, and tetralignols described here are produced by radical endwise condensation reactions and no postcoupling enzymatic reactions seem to be involved because no products were detected resulting from further metabolism of the oligolignols. A pathway in which the oligolignols are used as the main building blocks of lignin is considered to be, at best, a minor one because lignins contain relatively few cinnamyl alcohol end groups, indicating that lignin is mainly produced by the addition of monolignols to the growing polymer and not by the concatenation of preformed oligomers. **G**- and **S**-type  $\beta$ -aryl ether and phenylcoumaran dimers can only add to another monolignol or oligolignol via the free-phenolic function; the unsaturated propenol side chain is blocked from any further reactions because peroxidase oxidizes specifically the phenol function and all couplings are radical-radical reactions. Thus, coupling involving substantial amounts of oligolignols would result in a high proportion of terminal alcohol residues in lignin, which is not observed. For example, it was estimated that over 95% of the lignin units are not derived from dimerization reactions, at least in softwood (Hatfield and Vermerris, 2001).

Of the 38 compounds characterized in this study, 20 have been previously identified from a variety of plant species and tissues, but because of their sporadic identification from various species and tissues, most of these oligolignols have been considered as lignans, compounds with a defensive role in plants. It is the large number of oligolignols identified in this study from one species and from a single tissue with extensive lignification and the nature of their chemical structures that allows us to conclude that these compounds have to be considered as a class of monolignol-coupling products that are formed under the ambient monolignol concentrations and oxidative conditions in the cell wall.

Our data support the recently challenged combinatorial chemical coupling hypothesis of monolignols (Ralph et al., 2004b), but do not exclude that the coupling of certain dilignols may be assisted by dirigent proteins. To investigate this possibility, the oligolignols present in xylem extracts should be purified and analyzed by chiral HPLC to determine their enantiomeric ratios. This will give insight into the elusive role of dirigent proteins in oligolignol synthesis. In another article (Morreel et al., 2004), we show that the relative abundance of these oligolignols is altered dramatically in transgenic poplar down-regulated for caffeic acid *O*-methyltransferase and that novel oligolignols, derived from products of incomplete monolignol biosynthesis, are produced.

## MATERIALS AND METHODS

### Growth Conditions and Plant Material

Wild-type and CCoAOMT down-regulated poplars (*Populus tremula*  $\times$  *P. alba* clone INRA no. 717-1B4; Meyermans et al., 2000) were propagated in vitro on Murashige and Skoog medium (Duchefa, Haarlem, The Netherlands). Rooted plantlets were transferred to the greenhouse (21°C, 60% humidity, 16-h light/8-h dark regime, 40–60  $\mu\text{mol m}^{-2} \text{s}^{-1}$  photosynthetic photon flux) and grown for 3 months until harvest, reaching a height of approximately 1.5 m.

Approximately 300 mg of xylem tissue were harvested from a 10-cm-long, debarked stem (by scraping with a scalpel), cut at 15 cm above ground. After grinding in liquid nitrogen, the tissue was extracted with 15 mL of methanol. The supernatant was subsequently removed and the residue lyophilized and weighed (approximately 70 mg).

### HPLC Analysis

An aliquot (1.0 mL) of the methanol phase was lyophilized and extracted with cyclohexane/water acidified with 0.1% trifluoroacetic acid (1:1; v/v), and separated on HPLC with a Luna C18(2) column (250  $\times$  4.6 mm, 5  $\mu\text{m}$ ; Phenomenex, Torrance, CA), as previously described (Meyermans et al., 2000). A valley-to-valley integration of the chromatogram was applied using the following parameter values: peak width, 15 s and threshold, 17  $\mu\text{V s}^{-1}$ . Using the maximum absorbance value between 230 and 450 nm, quantification was based on the peak height instead of the peak area as the latter method is more sensitive to impurities (Snyder et al., 1997) and standardized to the dry weight. In addition, the HPLC procedure was carried out on sufficient plant material to collect 0.3-mL fractions that were freeze-dried and redissolved in 0.1 mL 1% aqueous triethylammonium acetate for LC-MS/MS.

For NMR analysis, repetitive HPLC separations were used to collect at least 0.1 mg of the compound of interest, followed by a final repurification on the Luna C18(2) column described above.

### LC-MS/MS Analysis

HPLC fractions were injected by means of a SpectraSystem AS1000 autosampler (Thermo Separation Products, Riviera Beach, FL) onto a reversed-phase Luna C18(2) column (150  $\times$  2.1 mm, 3  $\mu\text{m}$ ; Phenomenex). A gradient separation (SpectraSystem P1000XR HPLC pump; Thermo Separation Products) was run from 1% aqueous triethylammonium acetate (solvent A, pH 5) to methanol-acetonitrile (25:75; v/v; 1% triethylammonium acetate; solvent B) using the following conditions: flow 0.25 mL  $\text{min}^{-1}$ , column temperature 40°C, time 0 min, 5% B, time 20 min, 100% B.

A SpectraSystem UV6000LP detector (Thermo Separation Products) measured UV/Vis absorption between 200 and 450 nm with a scan rate of 2 scans/s. Atmospheric pressure chemical ionization, operated in the negative ionization mode, was used as an ion source to couple HPLC with an MS instrument (LCQ Classic; ThermoQuest, San Jose, CA; vaporizer temperature 450°C, capillary temperature 150°C, source current 5  $\mu\text{A}$ , sheath gas flow 21, aux gas flow 3).

During separation, the most abundant ion in each full MS scan was fragmented in the next scan with the dependent MS/MS mode.

Additionally, each fraction was separated on LC-MS/MS under higher acidity buffer conditions. A gradient separation was run from solvent C (1% aqueous acetic acid, pH 2) to solvent D (acetonitrile, 1% acetic acid) under the following conditions: column temperature 40°C, flow 0.3 mL min<sup>-1</sup>, time 0 min, 5% D, time 1 min, 17% D, time 19 min, 77% D, time 20 min, 100% D. The MS conditions were: vaporizer temperature 350°C, capillary temperature 100°C, source current 5  $\mu$ A, sheath gas flow set at 34, aux gas flow set at 4.

## NMR Spectroscopy

Compounds were authenticated by the normal range of 1D and standard 2D (COSY, TOCSY, HSQC, HMBC) experiments on a 360 MHz DRX-360 instrument (Bruker, Karlsruhe, Germany) fitted with a 5-mm <sup>1</sup>H/broadband gradient probe with inverse geometry (proton coils closest to the sample). The solvent was acetone-d<sub>6</sub> unless otherwise noted; the central acetone solvent peak was used as internal reference ( $\delta_c$  29.8,  $\delta_H$  2.04 ppm). NMR data will be deposited in our NMR database of lignin and cell wall model compounds (<http://www.dfrc.ars.usda.gov/software.html>; Ralph et al., 2004a).

## Shorthand Naming Convention for Oligolignols (Dimers, Trimers, and Tetramers)

To describe the oligolignols in a logical and informative manner, the following convention has been adopted. Bold **G** and **S** are used for guaiacyl and syringyl units, to name the units derived from coupling reactions of coniferyl and sinapyl alcohol; bold **SP** for units derived from the incorporation of sinapyl *p*-hydroxybenzoate esters; and **G'**, **S'**, and **V'** for units derived from coniferaldehyde, sinapaldehyde, and vanillin, respectively. The interunit bond formed during the radical coupling reaction is specified in parentheses: (8-O-4), (8-5), or (8-8). For example, **G(8-O-4)S(8-5)G'** results from sinapyl alcohol coupling at its C<sub>8</sub> position with coniferaldehyde at its C<sub>5</sub> position to make dimer **S(8-5)G'**, followed by coupling of this dimer at its phenolic 4-O position with another coniferyl alcohol radical at its favored C<sub>8</sub> position. The descriptor for the trimer is unambiguous because coupling the other way round is not possible because coupling always requires a free-phenolic group on the unit's aromatic ring; for instance, first coupling of coniferyl alcohol at its C<sub>8</sub> position with sinapyl alcohol at its 4-O position could produce the dimer **G(8-O-4)S**, but the specified trimer can no longer result from further coupling to this dimer, because only the **G** unit in the dimer is capable of entering coupling reactions (Ralph et al., 2004b). Whenever it could be determined or tentatively identified, *erythro*- and *threo*-isomers of the (8-O-4) structures are indicated as (*e*8-O-4) and (*t*8-O-4). The shorthand notation (8-O-4\*) indicates a benzylic oxidized (8-O-4)-linked unit, i.e. bearing a 7-oxo group (Fig. 4, compound 30).

## Preparation and HPLC Analysis of Oligolignol Mixtures

**G**, **S**, and **G + S** synthetic oligolignol mixtures were prepared by using Cu(OAc)<sub>2</sub> oxidation of coniferyl alcohol, sinapyl alcohol, or both coniferyl and sinapyl alcohols, respectively, as described previously (Landucci et al., 1995). For example, the **S + G** oligolignols were prepared as follows. A mixture of coniferyl alcohol (480 mg, 2.66 mmol) and sinapyl alcohol (560 mg, 2.66 mmol) was dissolved in 0.05 M borate solution (200 mL, pH 9.2), which was prepared from sodium borate (Na<sub>2</sub>B<sub>4</sub>O<sub>7</sub>·10H<sub>2</sub>O) at 100°C with stirring. A solution of copper acetate (1.06 g, 5.32 mmol) in water (10 mL) was added into the reaction solution. A green suspension formed and was stirred for 30 min at 100°C, after which time the solution color turned to orange. Insoluble copper salts were removed by filtration and the bright yellow solution was collected. The solution was acidified to pH 5 with 0.5 M HCl and extracted with ethyl acetate. The yellow extract was dried over MgSO<sub>4</sub> and concentrated by rotary evaporation. The boric acid was removed as methyl borate by adding MeOH and evaporating three times. A white foamy solid (83%) was obtained. Part of the crude product was used to separate compounds by thin-layer chromatography (TLC; CHCl<sub>3</sub>:MeOH; 10:1) to isolate and identify the structures. The mixtures were also used for HPLC analysis under conditions comparable to those used for the poplar methanol extracts. Coniferyl alcohol (200 mg) or sinapyl alcohol (250 mg) were treated independently in a similar manner to prepare oligolignols consisting of **G** or **S** units only. Higher amounts of some compounds were obtained in Mn(OAc)<sub>3</sub>/pyridine reactions (Landucci et al., 1995).

The following compounds were purified from these preparations and identified by MS and NMR analysis: **G(t8-O-4)G**, **1**; **S(8-5)G**, **6**; **S(8-O-4)S'**, **9**;

**S(t8-O-4)S(8-8)S**, **16**; **S(t8-O-4)S(8-5)G**, **17**; **G(t8-O-4)S(8-5)G**, **18**; **G(e8-O-4)S(8-5)G**, **20**; **G(t8-O-4)S(8-8)G**, **22**; **G(t8-O-4)S(8-5)G'**, **23**; **G(t8-O-4)S(8-8)S**, **24**; **G(e8-O-4)S(8-5)G'**, **25**; **G(8-O-4)S(8-8)S(8-O-4)G**, **33**.

## Chemical Synthesis of Dilignols, Trilignols, and Tetralignols

To authenticate the compounds isolated from the xylem fraction, the required compounds were prepared by the above Cu(OAc)<sub>2</sub> oxidation unless described specifically. Non-HPLC separations were by preparative TLC or by flash chromatography. Where available, the compound number in our NMR database of lignin and cell wall model compounds (<http://www.dfrc.ars.usda.gov/software.html>; Ralph et al., 2004a) is noted.

Compound **1**, **G(t8-O-4)G**, *threo*-1-(4-hydroxy-3-methoxy-phenyl)-2-[4-(3-hydroxy-propenyl)-2-methoxy-phenoxy]-propane-1,3-diol, (8-O-4)-dehydrodiconiferyl alcohol, database number 2013t: prepared from the Cu(OAc)<sub>2</sub> system described above using coniferyl alcohol. Oxidation of coniferyl alcohol by Fe(NH<sub>4</sub>)(SO<sub>4</sub>)<sub>3</sub> gave a mixture of *threo*- and *erythro*-isomers.

Compound **2**, **G(8-5)G**, 4-[3-hydroxymethyl-5-(3-hydroxy-propenyl)-7-methoxy-2,3-dihydrobenzofuran-2-yl]-2-methoxy-phenol, (8-5)-dehydrodiconiferyl alcohol, database number 2004: prepared as described previously (Quideau and Ralph, 1994).

Compound **3**, **G(8-8)G**, 1,4-bis-(4-hydroxy-3-methoxy-phenyl)-tetrahydro-furo[3,4-c]furan], pinosresinol, database number 2020: prepared according to Syrjänen and Brunow (2000).

Compound **4**, **S(8-8)S**, 1,4-bis-(3,5-dimethoxy-4-hydroxy-phenyl)-tetrahydro-furo[3,4-c]furan], syringaresinol, database number 117: synthesized by coupling sinapyl alcohol via CuSO<sub>4</sub>, as described previously (Freudenberg et al., 1958).

Compound **6**, **S(8-5)G**, 4-[3-hydroxymethyl-5-(3-hydroxy-propenyl)-7-methoxy-2,3-dihydrobenzofuran-2-yl]-2,6-dimethoxy-phenol, simulanol, database number 3063: prepared from the Cu(OAc)<sub>2</sub> system described above using coniferyl and sinapyl alcohols.

Compound **7**, **G(8-5)G'**, 3-[2-(4-hydroxy-3-methoxy-phenyl)-3-hydroxymethyl-7-methoxy-2,3-dihydro-benzofuran-5-yl]-propenal, balanophonin, database number 2021, is the cinnamaldehyde analog of compound **G(8-5)G**, **2**. For its synthesis, compound **2** (70 mg, 0.21 mmol) was dissolved in tetrahydrofuran (10 mL) and 2,3-dichloro-5,6-dicyanobenzoquinone (Becker et al., 1980) was added. The reaction mixture was stirred overnight at room temperature. Crude products were dried and compound **7** **G(8-5)G'** was isolated by TLC (chloroform:ethyl acetate; 1:1). Yellow needle crystals (55.4 mg, yield 74%) were obtained from acetone-petroleum ether. NMR data were consistent with those reported previously (Quideau and Ralph, 1994; Sy and Brown, 1999).

Compound **8**, **G(t8-O-4)S**, *threo*-1-(4-hydroxy-3-methoxy-phenyl)-2-[2,6-dimethoxy-phenoxy-4-(3-hydroxy-propenyl)]-propane-1,3-diol, database number 3067: small amounts were reported (Landucci et al., 1995) in the oligomer reactions described above, but were insufficient to isolate here. The peracetate (database number 188) was previously isolated, following acetylation, from Mn(OAc)<sub>3</sub> reactions in pyridine (Landucci et al., 1995). Here, compound **8** was more conveniently isolated from synthetic oligolignol mixtures from Cu(OAc)<sub>2</sub> oxidations of coniferyl plus sinapyl alcohols in acetone at room temperature.

Compound **9**, **S(8-O-4)S'**, (2*E*)-3-[4-[2-hydroxy-1-(hydroxymethyl)-2-(4-hydroxy-3,5-dimethoxyphenyl)-ethoxy]-3,5-dimethoxyphenyl-acrylaldehyde: prepared from the Mn(OAc)<sub>3</sub> oxidation of sinapyl alcohol in pyridine (Landucci et al., 1995).

Compound **11**, **G(8-5)V'**, 2-(4-hydroxy-3-methoxy-phenyl)-3-hydroxymethyl-7-methoxy-2,3-dihydro-benzofuran-5-carbaldehyde, fucusal, database number 3061: prepared in low yields from **G(8-5)G** **2** by oxidation with CrO<sub>3</sub>/montmorillonite K-10 in CH<sub>2</sub>Cl<sub>2</sub> (which also produced low yields of compound **7**; Heravi et al., 1999).

Compound **16**, **S(t8-O-4)S(8-8)S**, 1-(4-hydroxy-3,5-dimethoxy-phenyl)-2-[4-[4-(4-hydroxy-3,5-dimethoxy-phenyl)-tetrahydro-furo[3,4-c]furan-1-yl]-2,6-dimethoxy-phenoxy]-propane-1,3-diol, *threo*-buddlenol D, database number 198: isolated from the oligomers of the above Cu(OAc)<sub>2</sub> oxidation of sinapyl alcohol.

Compound **17**, **S(t8-O-4)S(8-5)G**, 1-(4-hydroxy-3,5-dimethoxy-phenyl)-2-[4-[3-hydroxymethyl-5-(3-hydroxy-propenyl)-7-methoxy-2,3-dihydro-benzofuran-2-yl]-2,6-dimethoxy-phenoxy]-propane-1,3-diol: present in oligolignol mixtures from Cu(OAc)<sub>2</sub> oxidations of coniferyl plus sinapyl alcohols described above.

Compounds **18** and **20**, G(*t*8-O-4)S(8-5)G and G(*e*8-O-4)S(8-5)G, *threo*- and *erythro*-1-(4-hydroxy-3-methoxy-phenyl)-2-[4-[3-hydroxymethyl-5-(3-hydroxy-propenyl)-7-methoxy-2,3-dihydro-benzofuran-2-yl]-2,6-dimethoxy-phenoxy]-propane-1,3-diol, buddlenol B, database number 181: prepared from the Cu(OAc)<sub>2</sub> system described above using coniferyl and sinapyl alcohols (Landucci et al., 1995). NMR showed that the synthesized compound was mainly present in the *threo* form.

Compound **19**, SP(8-8)S, tetrahydro- $\alpha$ 4,2-bis-(4-hydroxy-3,5-dimethoxyphenyl) $\alpha$ 3-O-(4-hydroxybenzoyl)-3,4-furandimethanol, database number 3066: this compound, crucial to establishing SP as a lignin precursor, was prepared by coupling sinapyl alcohol and sinapyl *p*-hydroxybenzoate via peroxidase/H<sub>2</sub>O<sub>2</sub> (approximately 30% yield) as described in detail elsewhere (Lu et al., 2004).

Compound **22**, G(*t*8-O-4)S(8-8)G, 1-(4-hydroxy-3-methoxy-phenyl)-2-[4-[4-(4-hydroxy-3-methoxy-phenyl)-tetrahydro-furo[3,4-c]furan-1-yl]-2,6-dimethoxy-phenoxy]-propane-1,3-diol, buddlenol E, database number 3064: isolated from the synthetic oligolignol mixture from Cu(OAc)<sub>2</sub> oxidation of sinapyl and coniferyl alcohols described above.

Compounds **23** and **25**, G(*t*8-O-4)S(8-5)G' and G(*e*8-O-4)S(8-5)G', *threo*- and *erythro*-3-(2-[4-[2-hydroxy-2-(4-hydroxy-3-methoxy-phenyl)-1-hydroxy-methyl-ethoxy]-3,5-dimethoxy-phenyl]-3-hydroxymethyl-7-methoxy-2,3-dihydro-benzofuran-5-yl)-propanal, buddlenol A, database number 3065: isolated from the synthetic oligolignol mixture from Cu(OAc)<sub>2</sub> oxidation of sinapyl and coniferyl alcohols described above. NMR showed that the synthesized compound was mainly present in the *threo* form.

Compound **24**, G(*t*8-O-4)S(8-8)S, 2-[4-[4-(4-hydroxy-3,5-dimethoxy-phenyl)-tetrahydrofuro[3,4-c]furan-1-yl]-2,6-dimethoxy-phenoxy]-1-(4-hydroxy-3-methoxy-phenyl)-propane-1,3-diol, buddlenol C, database number 183: isolated from the synthetic oligolignol mixture from Cu(OAc)<sub>2</sub> oxidation of sinapyl and coniferyl alcohols described above.

Compound **33**, G(8-O-4)S(8-8)S(8-O-4)G, 2-[4-[4-[2-hydroxy-2-(4-hydroxy-3-methoxy-phenyl)-1-hydroxymethyl-ethoxy]-3,5-dimethoxy-phenyl]-tetrahydro-furo[3,4-c]furan-1-yl]-2,6-dimethoxy-phenoxy]-1-(4-hydroxy-3-methoxy-phenyl)-propane-1,3-diol, hedyotisol, database number 194: isolated from the synthetic oligolignol mixture from Cu(OAc)<sub>2</sub> oxidation of sinapyl and coniferyl alcohols described above. The *threo/erythro* configurations of the two  $\beta$ -aryl ether units were not determined.

Compound **39**, S(*t*8-O-4)G(8-5)G, 1-(4-hydroxy-3,5-dimethoxy-phenyl)-2-[4-[3-hydroxymethyl-5-(3-hydroxy-propenyl)-7-methoxy-2,3-dihydro-benzofuran-2-yl]-2-methoxy-phenoxy]-propane-1,3-diol, was prepared via traditional synthetic  $\beta$ -ether lignin model methods, for instance as in Ralph et al. (1986). Briefly, the coniferyl alcohol dimer G(8-5)G (compound **2**, (8-5)-dehydroconiferyl alcohol) was added to acetate-protected  $\alpha$ -bromo-aceto-syringone, formaldehyde was added (to create the three-carbon side chain) and the benzylic ketone was reduced with NaBH<sub>4</sub> in ethanol:H<sub>2</sub>O (1:1). Completing the deacetylation in pyrrolidine:MeOH (1:1) provided S(*t*8-O-4)G(8-5)G, **39**. NMR (data from 2D, average chemical shift values only, using A(*t*8-O-4)B(8-5)C to unambiguously identify the units in S(*t*8-O-4)G(8-5)G):  $\delta_C/\delta_H$  105.3/6.76 (A2/6), 130.4/6.53 (C7), 128.3/6.23 (C8), 88.0/5.58 (B7), 73.8/4.87 (A7), 87.9/4.24 and 86.2/4.33 (A8), 62.2/4.19 (C9), 64.6/3.88, 3.82 (B9), 61.9/3.8-3.7, 3.51 (A9), 56.5/3.78 and 56.3/3.86 (OMe), 54.8/3.52 (B8).

## Distribution of Materials

Upon request, all novel materials described in this publication will be made available in a timely manner for noncommercial research purposes, subject to the requisite permission from any third-party owners of all or parts of the material. Obtaining any permissions will be the responsibility of the requester.

## ACKNOWLEDGMENT

The authors thank Martine De Cock for help in preparing the manuscript.

Received July 7, 2004; returned for revision September 28, 2004; accepted September 28, 2004.

## LITERATURE CITED

Adler E (1977) Lignin chemistry - past, present and future. *Wood Sci Technol* **11**: 169–218

- Akiyama T, Matsumoto Y, Okuyama T, Meshitsuka G (2003) Ratio of *erythro* and *threo* forms of  $\beta$ -O-4 structures in tension wood lignin. *Phytochemistry* **64**: 1157–1162
- Baucher M, Chabbert B, Pilate G, Van Doorselaere J, Tollier M-T, Petit-Conil M, Cornu D, Monties B, Van Montagu M, Inzé D, et al (1996) Red xylem and higher lignin extractability by down-regulating a cinnamyl alcohol dehydrogenase in poplar. *Plant Physiol* **112**: 1479–1490
- 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
- Becker HD, Bjoerk A, Adler E (1980) Quinone dehydrogenation. Oxidation of benzylic alcohols with 2,3-dichloro-5,6-dicyanobenzoquinone. *J Org Chem* **45**: 1596–1600
- Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. *Annu Rev Plant Biol* **54**: 519–546
- Boudet AM, Kajita S, Grima-Pettenati J, Goffner D (2003) Lignins and lignocellulosics: a better control of synthesis for new and improved uses. *Trends Plant Sci* **8**: 576–581
- Brunow G, Karlsson O, Lundquist K, Sipilä J (1993) On the distribution of the diastereomers of the structural elements in lignins: the steric course of reactions mimicking lignin biosynthesis. *Wood Sci Technol* **27**: 281–286
- Brunow G, Lundquist K, Gellerstedt G (1999) Lignin. In E Sjöström, R Alén, eds, *Analytical Methods in Wood Chemistry, Pulping and Papermaking*. Springer-Verlag, Berlin, pp 77–124
- Burlat V, Kwon M, Davin LB, Lewis NG (2001) Dirigent proteins and dirigent sites in lignifying tissues. *Phytochemistry* **57**: 883–897
- Chen Y-R, Sarkanen S (2004) Macromolecular lignin replication: a mechanistic working hypothesis. *Phytochemistry Rev* (in press)
- Christensen JH, Baucher M, O'Connell AP, Van Montagu M, Boerjan W (2000) Control of lignin biosynthesis. In SM Jain, SC Minocha, eds, *Molecular Biology of Woody Plants, Vol 1 (Forestry Sciences, Vol 64)*. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 227–237
- Davin LB, Wang H-B, Crowell AL, Bedgar DL, Martin DM, Sarkanen S, Lewis NG (1997) Stereoselective bimolecular phenoxy radical coupling by an auxiliary (dirigent) protein without an active center. *Science* **275**: 362–366
- Faix O (1986) Investigation of lignin polymer models (DHP's) by FTIR spectroscopy. *Holzforschung* **40**: 273–280
- Freudenberg K (1959) Biosynthesis and constitution of lignin. *Nature* **183**: 1152–1155
- Freudenberg K, Harkin JM, Reichert M, Fukuzumi T (1958) Die Dehydrierung des Sinapinalkohols. *Chem Ber* **91**: 581–590
- Freudenberg K, Neish AC (1968) *Constitution and Biosynthesis of Lignin*. Springer-Verlag, Berlin
- Grabber JH, Ralph J, Hatfield RD, Quideau S, Kuster T, Pell AN (1996) Dehydrogenation polymer—cell wall complexes as a model for lignified grass walls. *J Agric Food Chem* **44**: 1453–1459
- Guo D, Chen F, Wheeler J, Winder J, Selman S, Peterson M, Dixon RA (2001) Improvement of in-rumen digestibility of alfalfa forage by genetic manipulation of lignin O-methyltransferases. *Transgenic Res* **10**: 457–464
- Hatfield R, Vermerris W (2001) Lignin formation in plants. The dilemma of linkage specificity. *Plant Physiol* **126**: 1351–1357
- Heravi MM, Kiakojoori R, Tabar-Hydar K (1999) Ammonium chlorochromate adsorbed on montmorillonite K-10: a new reagent for the oxidation of alcohols to the corresponding carbonyl compounds under non-aqueous conditions. *Monatsh Chem* **130**: 581–583
- Hertzberg M, Aspeborg H, Schrader J, Andersson A, Erlandsson R, Blomqvist K, Bhalerao R, Uhlén M, Teeri TT, Lundeberg J, et al (2001) A transcriptional roadmap to wood formation. *Proc Natl Acad Sci USA* **98**: 14732–14737
- Kim H, Ralph J, Lu F, Ralph SA, Boudet A-M, MacKay JJ, Sederoff RR, Ito T, Kawai S, Ohashi H, et al (2003) NMR analysis of lignins in CAD-deficient plants. Part 1. Incorporation of hydroxycinnamaldehydes and hydroxybenzaldehydes into lignins. *Org Biomol Chem* **1**: 268–281
- Landucci LL, Deka GC, Roy DN (1992) A <sup>13</sup>C NMR study of milled wood lignins from hybrid *Salix* clones. *Holzforschung* **46**: 505–511
- Landucci LL, Luque S, Ralph J (1995) Reaction of *p*-hydroxycinnamyl

- alcohols with transition metal salts. 2. Preparation of guaiacyl/syringyl di-, tri- and tetralignols. *J Wood Chem Technol* **15**: 493–513
- Landucci LL, Ralph SA** (2001) Reactions of *p*-hydroxycinnamyl alcohols with transition metal salts. IV. Tailored syntheses of  $\beta$ -O-4 trimers. *J Wood Chem Technol* **21**: 31–52
- Lewis NG** (1999) A 20th century roller coaster ride: a short account of lignification. *Curr Opin Plant Biol* **2**: 153–162
- Lewis NG, Davin LB** (1999) Lignans: biosynthesis and function. In U Sankawa, ed, *Polyketides and Other Secondary Metabolites Including Fatty Acids and Their Derivatives, Comprehensive Natural Products Chemistry, Vol 1*. Elsevier Science, Amsterdam, pp 639–712
- Lu F, Ralph J** (1997) DFRC method for lignin analysis. 1. New method for  $\beta$ -aryl ether cleavage: lignin model studies. *J Agric Food Chem* **45**: 4655–4660
- Lu F, Ralph J** (2002) Preliminary evidence for sinapyl acetate as a lignin monomer in kenaf. *Chem Commun* **2002**: 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, 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
- Lundquist K** (1992) Acidolysis. In SY Lin, CW Dence, eds, *Methods in Lignin Chemistry*. Springer-Verlag, Heidelberg, pp 287–300
- 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) Modification 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, Lu F, Goeminne G, Busson R, Herdewijn P, Goeman J, Van der Eycken J, Boerjan W, Messens E** (2004) Phenolic profiling of caffeic acid *O*-methyltransferase-deficient poplar reveals novel benzo-dioxane oligolignols. *Plant Physiol* **136**: (in press)
- Nakano J, Ishizu A, Migata N** (1961) Studies on lignin. XXXII. Ester groups of lignin. *Tappi* **44**: 30–32
- Pilate G, Guiney E, Holt K, Petit-Conil M, Lapierre C, Leplé J-C, Pollet B, Mila I, Webster EA, Marstorp HG, et al** (2002) Field and pulping performances of transgenic trees with altered lignification. *Nat Biotechnol* **20**: 607–612
- Quideau S, Ralph J** (1994) A biomimetic route to lignin model compounds via silver (I) oxide oxidation. 1. Synthesis of dilignols and non-cyclic benzyl aryl ethers. *Holzforschung* **48**: 12–22
- Ralph J, Ede RM, Wilkins AL** (1986) Synthesis of trimeric lignin model compounds composed of  $\beta$ -aryl ether and phenylcoumaran structures. *Holzforschung* **40**: 23–30
- Ralph SA, Landucci LL, Ralph J** (2004a) NMR database of lignin and cell wall model compounds. US Dairy Forage Research Center, US Department of Agriculture, Agricultural Research Service. <http://www.dfrc.ars.usda.gov/software.html> (November 15, 2004)
- Ralph J, Lundquist K, Brunow G, Lu F, Kim H, Schatz PF, Marita JM, Hatfield RD, Ralph SA, Christensen JH, et al** (2004b) Lignins: natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. *Phytochemistry Rev* **3**: (in press)
- Ralph J, Marita JM, Ralph SA, Hatfield RD, Lu F, Ede RM, Peng J, Quideau S, Helm RF, Grabber JH, Kim H, et al** (1999a) Solution-state NMR of lignins. In DS Argyropoulos, ed, *Advances in Lignocellulosic Characterization*. TAPPI Press, Atlanta, pp 55–108
- Ralph J, Peng J, Lu F, Hatfield RD, Helm RF** (1999b) Are lignins optically active? *J Agric Food Chem* **47**: 2991–2996
- Rolando C, Monties B, Lapierre C** (1992) Thioacidolysis. In SY Lin, CW Dence, eds, *Methods in Lignin Chemistry*. Springer-Verlag, Berlin, pp 334–349
- Sakakibara N, Suzuki S, Umezawa T, Shimada M** (2003) Biosynthesis of yatein in *Anthriscus sylvestris*. *Org Biomol Chem* **1**: 2474–2485
- Sarkanen KV** (1971) Precursors and their polymerization. In KV Sarkanen, CH Ludwig, eds, *Lignins, Occurrence, Formation, Structure and Reactions*. Wiley-Interscience, New York, pp 95–163
- Smith DCC** (1955) *p*-Hydroxybenzoate groups in the lignin of aspen (*Populus tremula*). *J Chem Soc* **1955**: 2347–2351
- Snyder LR, Kirkland JJ, Glajch JL** (1997) *Practical HPLC Method Development*, Ed 2. John Wiley & Sons, New York
- Sun RC, Fang JM, Tomkinson J** (1999) Fractional isolation and structural characterization of lignins from oil palm trunk and empty fruit bunch fibres. *J Wood Chem Technol* **19**: 335–356
- Sy L-K, Brown GD** (1999) Conifer aldehyde derivatives from tissue culture of *Artemisia annua* and *Tanacetum parthenium*. *Phytochemistry* **50**: 781–785
- Syrjänen K, Brunow G** (1998) Oxidative cross coupling of *p*-hydroxycinnamic alcohols with dimeric arylglycerol  $\beta$ -aryl ether lignin model compounds. The effect of oxidation potentials. *J Chem Soc [Perkin 1]* **1998**: 3425–3429
- Syrjänen K, Brunow G** (2000) Regioselectivity in lignin biosynthesis. The influence of dimerization and cross-coupling. *J Chem Soc [Perkin 1]* **2000**: 183–187
- Umezawa T** (2004) Diversity of lignan biosynthesis. *Phytochemistry Rev* **3**: (in press)
- Zhong R, Morrison WH III, Himmelsbach DS, Poole FL II, Ye Z-H** (2000) Essential role of caffeoyl coenzyme A *O*-methyltransferase in lignin biosynthesis in woody poplar plants. *Plant Physiol* **124**: 563–577