

# Tricin, a Flavonoid Monomer in Monocot Lignification<sup>1</sup>[OPEN]

Wu Lan, Fachuang Lu\*, Matthew Regner, Yimin Zhu<sup>2</sup>, Jorge Rencoret<sup>3</sup>, Sally A. Ralph, Uzma I. Zakai<sup>4</sup>, Kris Morreel, Wout Boerjan, and John Ralph\*

Department of Energy Great Lakes Bioenergy Research Center, Wisconsin Energy Institute (W.L., F.L., M.R., Y.Z., J.Re., U.I.Z., J.Ra.), Department of Biological System Engineering (W.L., J.Ra.), and Department of Biochemistry (F.L., M.R., J.Ra.), University of Wisconsin, Madison, Wisconsin 53726; United States Forest Service, Forest Products Laboratory, Madison, Wisconsin 53726 (S.A.R.); and Department of Plant Systems Biology, Vlaams Instituut voor Biotechnologie, and Department of Plant Biotechnology and Bioinformatics, Ghent University, B-9052 Ghent, Belgium (K.M., W.B.)

Tricin was recently discovered in lignin preparations from wheat (*Triticum aestivum*) straw and subsequently in all monocot samples examined. To provide proof that triclin is involved in lignification and establish the mechanism by which it incorporates into the lignin polymer, the 4'-O- $\beta$ -coupling products of triclin with the monolignols (*p*-coumaryl, coniferyl, and sinapyl alcohols) were synthesized along with the trimer that would result from its 4'-O- $\beta$ -coupling with sinapyl alcohol and then coniferyl alcohol. Tricin was also found to cross couple with monolignols to form triclin-(4'-O- $\beta$ )-linked dimers in biomimetic oxidations using peroxidase/hydrogen peroxide or silver (I) oxide. Nuclear magnetic resonance characterization of gel permeation chromatography-fractionated acetylated maize (*Zea mays*) lignin revealed that the triclin moieties are found in even the highest molecular weight fractions, either linked to lignin units, demonstrating that triclin is indeed incorporated into the lignin polymer. These findings suggest that triclin is fully compatible with lignification reactions, is an authentic lignin monomer, and, because it can only start a lignin chain, functions as a nucleation site for lignification in monocots. This initiation role helps resolve a long-standing dilemma that monocot lignin chains do not appear to be initiated by monolignol homodehydrodimerization as they are in dicots that have similar syringyl-guaiacyl compositions. The term flavonolignin is recommended for the racemic oligomers and polymers of monolignols that start from triclin (or incorporate other flavonoids) in the cell wall, in analogy with the existing term flavonolignan that is used for the low-molecular mass compounds composed of flavonoid and lignan moieties.

Lignin, a complex phenylpropanoid polymer in the plant cell wall, is predominantly deposited in the cell walls of secondary-thickened cells (Vanholme et al.,

2010). It is synthesized via oxidative radical coupling reactions from three prototypical monolignols, *p*-coumaryl, coniferyl, and sinapyl alcohols, differentiated by their degree of methoxylation *ortho* to the phenolic hydroxyl group. Considered within the context of the entire polymer, the main structural features of lignin can be defined in terms of its *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) units, derived respectively from these three monolignols (Ralph, 2010). Several novel monomers, all deriving from the monolignol biosynthetic pathway, have been found to incorporate into lignin in wild-type and transgenic plants. For example, monolignol acetate, *p*-hydroxybenzoate, and *p*-coumarate ester conjugates have all been shown to incorporate into lignin polymers and are the source of naturally acylated lignins (Ralph et al., 2004; Lu and Ralph, 2008); lignins derived solely from caffeoyl alcohol were found in the seed coats of both monocot and dicot plants (Chen et al., 2012a, 2012b); lignins derived solely from 5-hydroxyconiferyl alcohol were found in a cactus (for example, in a member of the genera *Astrophytum*) seed coat (Chen et al., 2012a); a *Medicago truncatula* transgenic deficient in cinnamyl alcohol dehydrogenase exhibited a lignin that was overwhelmingly derived from hydroxycinnamaldehydes (instead of their usual hydroxycinnamyl alcohol analogs; Zhao et al., 2013); and *iso*-sinapyl alcohol was implicated as a monomer in caffeic acid *O*-methyltransferase down-regulated switchgrass (*Panicum virgatum*; Tschaplinski et al., 2012). These findings imply that plants are quite

<sup>1</sup> This work was supported by the Department of Energy Great Lakes Bioenergy Research Center (grant no. DE-FC02-07ER64494) and the Multidisciplinary Research Partnership Biotechnology for a Sustainable Economy (grant no. 01MRB510W) of Ghent University.

<sup>2</sup> Present address: Department of Chemistry, Pennsylvania State University, 3000 IvySide Park, Altoona College, Altoona, PA 16601.

<sup>3</sup> Present address: Instituto de Recursos Naturales y Agrobiología de Sevilla, Consejo Superior de Investigaciones Científicas, Avenida Reina Mercedes 10, 41012 Seville, Spain.

<sup>4</sup> Present address: Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, 1210 Molecular Biology Building, Ames, IA 50011.

\* Address correspondence to fachuanglu@wisc.edu and jralph@wisc.edu.

W.L. performed most of the experiments and wrote the article; M.R., Y.Z., J.Re., S.A.R., and U.I.Z. performed some of the experiments and provided technical assistance to W.L.; K.M. and W.B. provided materials and mass spectral analysis data; F.L. and J.Ra. developed the project from our (J.Re. and J.Ra.) initial discovery of triclin in wheat lignin; F.L. designed and supervised the experimental plans and helped to write the article; J.Ra. supervised the project, defined the scope, obtained the grant related to this work, and editing the article.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors ([www.plantphysiol.org](http://www.plantphysiol.org)) is: John Ralph (jralph@wisc.edu).

[OPEN] Articles can be viewed without a subscription.

[www.plantphysiol.org/cgi/doi/10.1104/pp.114.253757](http://www.plantphysiol.org/cgi/doi/10.1104/pp.114.253757)

flexible in being able to use a variety of monomers during lignification to form the heterogeneous lignin polymer. Most recently, and as addressed more fully here, the flavonoid tricin has been implicated as a monomer in monocot lignins (del Río et al., 2012). To our knowledge, tricin is the first monomer from outside the monolignol biosynthetic pathway to be implicated in lignification.

Tricin [5,7-dihydroxy-2-(4-hydroxy-3,5-dimethoxyphenyl)-4*H*-chromen-4-one], a member of the flavonoid family, is recognized as a valuable human health compound due to its antioxidant, antiaging, anticancer, and cardioprotective potentials (Ogo et al., 2013). Tricin and its derivatives can be solvent extracted from monocot samples such as wheat (*Triticum aestivum*), oat bran (*Avena sativa*), bamboo (*Leleba oldhami*), sugarcane (*Saccharum officinarum*), and maize (*Zea mays*). Extracted compounds can take the form of tricin itself, 7-*O*-glycosylated tricin, or the flavonolignan in which tricin is 4'-*O*-etherified by putative coupling with coniferyl alcohol (Ju et al., 1998; Bouaziz et al., 2002; Wenzig et al., 2005; Duarte-Almeida et al., 2007; Van Hoyweghen et al., 2010; Nakano et al., 2011; Bottcher et al., 2013; Moheb et al., 2013).

In 2012, we reported, to our knowledge, the first evidence that tricin was incorporated into lignin, as implicated by two previously unassigned correlation peaks at  $\delta_C/\delta_H$  94.1/6.56 and 98.8/6.20 in a heteronuclear single-quantum coherence (HSQC) NMR spectrum from the whole cell wall and an isolated milled wood lignin of (unacetylated) wheat straw (del Río et al., 2012). The same evidence has now been found in the HSQC spectrum of wheat straw lignin isolated via different methods (Yelle et al., 2013; Zeng et al., 2013). Additional studies have verified the presence of tricin in lignin fractions from a variety of monocots, including bamboo (You et al., 2013), coconut coir (*Cocos nucifera*; Rencoret et al., 2013), maize, and others examined in our laboratories. The implication that tricin is the first phenolic from outside the monolignol biosynthetic pathway found to be integrated into the polymer has prompted further study with the aim of identifying and mechanistically delineating the role of tricin in lignin and its biosynthetic incorporation pathway.

Tricin, unlike the monolignols that derive from the shikimate biosynthetic pathway (Sarkanen and Ludwig, 1971), is derived from a combination of the shikimate and acetate/malonate-derived polyketide pathways (Winkel-Shirley, 2001), as shown in Supplemental Figure S1. After *p*-coumaroyl-CoA is synthesized from *p*-coumaric acid by 4-coumarate:CoA ligase, it branches from the monolignol biosynthetic route to be transformed via chalcone synthase and chalcone isomerase into naringenin, the central precursor of most flavonoids. Naringenin is subsequently converted into apigenin by flavone synthase. Further hydroxylation at C-3' and C-5' followed by *O*-methylation furnishes tricin (Koes et al., 1994; Winkel-Shirley, 2001). The incorporation of tricin into lignin, therefore, suggests that an additional biosynthetic pathway, namely the polyketide

pathway, may be associated with cell wall lignification in monocots.

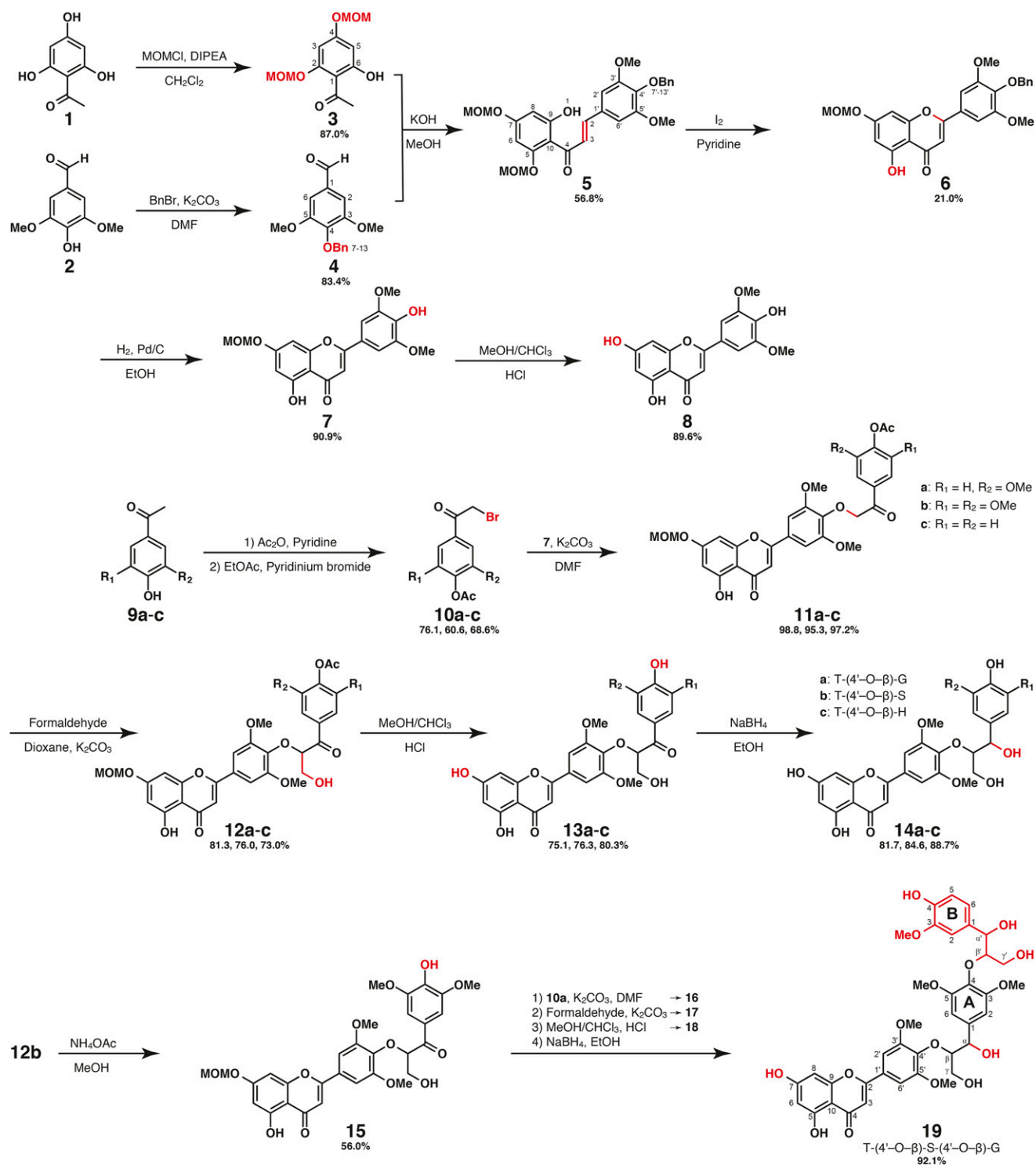
The revelation that tricin is incorporated into the lignin polymer was precipitated by closer study of signals found within the NMR spectra of various monocot samples. Before this discovery, tricin had not been noted in any lignin fractions, and although it is reasonable to anticipate compatibility based on its chemical structure, there is no direct and reliable evidence to date showing that tricin is able to react with monolignols through radical coupling; therefore, the efficiency and selectivity of the coupling reactions between tricin and various monolignols were also unknown. Synthetic model compounds that would facilitate the elucidation of the role of tricin within plant cell walls are desirable as aids to be used in a mechanistic study of flavonolignin generation. (We coin the term flavonolignin to describe the racemic oligomers and polymers of monolignols that start from tricin [or other flavonoids] in the cell wall, in analogy with the existing term flavonolignan that is used for the low-molecular mass compounds composed of flavonoid and lignan moieties that are presumably made in the cytoplasm [Begum et al., 2010; Niculaes et al., 2014; Dima et al., 2015]).

The overall objective of this study is to demonstrate that tricin incorporates into the lignin polymer of monocots, with maize/corn stover as the representative experimental material. To this end, we have synthesized tricin and various model compounds in which tricin is conjugated to monolignols in the manner expected for the lignification process. Next, we verified whether these synthetic compounds could be made from their assumed precursors under the biomimetic radical conditions anticipated for lignification. Subsequently, NMR data generated from these synthetic and biomimetic coupling products were compared with NMR data from native maize stover lignin, including high- $M_n$  fractions. We conclude that tricin is a monomer in monocot lignification and that, because little syringaresinol is found in maize lignin, tricin is functioning as a nucleation site that initiates lignin polymer chains.

## RESULTS AND DISCUSSION

### Synthesis of Tricin-Monolignol Cross-Coupled Oligomers

The synthetic scheme shown in Figure 1 outlines the syntheses of oligomers 14a to 14c and 19 containing a tricin (T) unit linked to arylglyceryl derivatives. 2,4,6-Trihydroxyacetophenone 1 and 4-hydroxy-3,5-dimethoxybenzaldehyde 2 were selected as the starting materials, and the phenolic hydroxyl groups were appropriately protected as the corresponding methoxymethyl (MOM) ether and benzyl ether, respectively (St. Denis et al., 2010). The flavone base structure was synthesized via chalcone 5, which was made via a Claisen-Schmidt condensation between diprotected triol 3 and benzylated syringaldehyde 4 in potassium hydroxide/methanol (MeOH). The initial attempt to synthesize protected tricin proceeded through the cyclization of



**Figure 1.** Scheme for the synthesis of tricrin-monolignol oligomers. Shorthand examples: The dimer 14 arising from the coupling of tricrin (at its 4'-O-position) with coniferyl alcohol/sinapyl alcohol/*p*-coumaryl alcohol (each at its β-position) is denoted in the text as T-(4'-O-β)-G/S/H; analogously, the trimer 19 resulting from the coupling of tricrin with sinapyl alcohol, and then the resultant dimer's coupling with coniferyl alcohol, is denoted as T-(4'-O-β)-S-(4'-O-β)-G.

chalcone 5 to provide the flavanone by first refluxing in ethanol with sodium acetate, followed by several attempts at oxidative dehydrogenation using a variety

of reagents; 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (Shanker et al., 1983), manganese (III) acetate (Singh et al., 2005), and copper (II) acetate were all tried without

success. Successful cyclization of chalcone 5 used a molar equivalent of  $I_2$  in pyridine (Lin et al., 2007). Acetylation of the crude products (not shown in Fig. 1) was performed to facilitate the isolation and purification of the target compound by chromatography. The desired flavonoid product 6 was obtained in approximately 20% yield from compound 5, with unexpected cleavage of the 5-O-MOM ether occurring during the reaction, where the main by-products were the 2,3-dihydro analogs of 6 with and without the 5-O-MOM ether. Removal of the MOM and benzyl groups from 6 would furnish the triclin monomer 8 itself, but selective removal of the benzyl group provided the ideal starting material 7 for synthesizing the required 4'-O-etherified triclin compounds. Additionally, it is worth noting that, although the 5-OH is free, it is strongly H bonded to the carbonyl at C-4 and therefore is essentially protected from most derivatization and coupling reactions.

The synthesis of triclin-(4'-O- $\beta$ )-monolignol dimers [abbreviated as T-(4'-O- $\beta$ )-G/S/H, with the G, S, or H moiety in the dimer depending on whether the formal monolignol involved is coniferyl, sinapyl, or *p*-coumaryl alcohol, respectively] followed traditional  $\beta$ -ether dimer synthetic methods (Kratzl et al., 1959; Ralph et al., 1992), starting with the standard  $S_N2$  reaction between brominated 4-acetoxyacetophenone derivatives 10a to 10c and monoprotected triclin 7, followed by the addition of formaldehyde and final deprotection (removal of the MOM and acetyl groups). After reduction of the carbonyl group at  $C_\alpha$  in 13a to 13c by sodium borohydride in ethanol, each compound 14a to 14c contained two chiral centers and therefore was a mixture of two diastereomers (*anti* [or *erythro*] and *syn* [or *threo*]), as was readily seen in the NMR spectra. The *anti*-diastereomer was the major product according to the Felkin-Anh model (Paddon-Row et al., 1982; Lodge and Heathcock, 1987), with an *anti:syn* ratio of about 80:20, as determined by relative integration of the  $\beta$ - or  $\gamma$ -protons. As noted previously for such  $\beta$ -ethers, the *syn*-isomer has the highest field (lowest  $\delta$ )  $\gamma$ -proton (Ralph and Wilkins, 1985; Ralph and Helm, 1991; Ralph, 1993). The *anti* and *syn* structure assignments were further confirmed by the magnitude of coupling constant  $^3J_{\alpha-H,\beta-H}$ : the doublet with the larger coupling constant of 7 Hz indicated the *anti*-isomer, whereas the doublet with the smaller one (4.8 Hz) corresponded to the *syn*-isomer (Ralph, 1993; Bouaziz et al., 2002). In the obtained HSQC spectra, correlation signals belonging to triclin units C3/H3, C6/H6, and C8/H8 were well dispersed from the lignin signals. The type of monolignol linked to triclin at 4'-OH did not affect the location of these three C-H correlation peaks (i.e. they were invariant for 14a to 14c and 19). All of these correlations, however, shifted in predictable ways upon acetylation. The chemical shifts (before and after acetylation) in various NMR solvents are listed in Table I. Trimer 19, T-(4'-O- $\beta$ )-S-(4-O- $\beta$ )-G, was synthesized from 12b via a similar pathway (Fig. 1). There are a total of eight possible diastereomers ( $2^4 = 16$  optical isomers); liquid

chromatography-mass spectrometry (LC-MS) separated only four components. High-resolution  $^1H$ -NMR of the isomer mixture also readily distinguished four isomers designated as *anti-syn* (*anti* structure in the internal  $\beta$ -O-4-unit and *syn* in the terminal free-phenolic end), *anti-anti*, *syn-anti*, and *syn-syn*, with a ratio of 13:37:37:13.

### Radical Coupling Reactions between Tricin and Monolignols

The key reaction in lignin biosynthesis is the radical coupling of phenolic radicals produced via peroxidases and/or laccases (Dean and Eriksson, 1992). Lignification is generally considered to begin with radical coupling first taking place between monolignols to produce dehydrodimers (henceforth termed dimers), which start the chain that extends by end-wise polymerization with additional monolignols (Freudenberger, 1956; Ralph et al., 2004). Radical reactions between two growing oligomer or polymer chains serve to connect them, thus increasing the polymer size and, as previously thought but recently questioned (Ralph et al., 2008; Crestini et al., 2011), causing the polymer to branch. Therefore, the peroxidase/hydrogen peroxide ( $H_2O_2$ ) system, a two-step one-electron transfer system, is commonly used as a biomimetic system for the preparation of dimeric lignin model compounds or dehydrogenation polymers. Oxidative radical coupling using silver (I) oxide ( $Ag_2O$ ) as the one-electron oxidant is another convenient approach to the synthesis of lignin model compounds (Zanarotti, 1985; Quideau and Ralph, 1994a). Here, we applied both methods to determine whether triclin is capable of reacting with monolignols under radical coupling conditions and to elucidate the nature of the resulting products. Peroxidase-catalyzed reactions were carried out using horseradish peroxidase in acetone:aqueous buffer (2:3, v/v), which contained a larger proportion of acetone than is typically used in this procedure. Furthermore, due to the poor solubility of triclin in this solvent system, a larger overall volume of solvent (500 mL) was needed to completely dissolve approximately 25 mg of triclin along with the monolignols. A reaction time of 2 h, longer than the time used in a prior study (Zhang et al., 2009), was then used to ensure complete reaction at this lower concentration of reactants. Acetone was used as the solvent for the  $Ag_2O$ -catalyzed reaction because it has been shown to give the highest yield of  $\beta$ -O-4 structures in the dimerization of coniferyl alcohol (Quideau and Ralph, 1994b). NMR data acquired from synthetic model compounds 14a to 14c were used for qualitative analysis of the products of triclin-monolignol cross coupling. Additionally, the synthesized model compounds were purified by HPLC and used to produce standard curves for measuring the yields of T-(4'-O- $\beta$ )-G/S/H products resulting from the cross-coupling reactions (Table II).

When triclin and a monolignol were oxidized by peroxidase/ $H_2O_2$ , various coupling products were

**Table I.** Selected diagnostic triclin unit chemical shifts from T-(4'-O-β)-G/S/H dimers 14a to 14c in acetone-d<sub>6</sub>, CDCl<sub>3</sub>, and DMSO-d<sub>6</sub>

| Sample                 | Unacetylated ( $\delta_C/\delta_H$ ) |            |            | Acetylated ( $\delta_C/\delta_H$ ) |             |             |
|------------------------|--------------------------------------|------------|------------|------------------------------------|-------------|-------------|
|                        | C3/H3                                | C6/H6      | C8/H8      | C3/H3                              | C6/H6       | C8/H8       |
| Acetone-d <sub>6</sub> | 106.01/6.82                          | 99.79/6.27 | 94.98/6.56 | 108.67/6.78                        | 110.29/7.49 | 114.80/6.95 |
| CDCl <sub>3</sub>      | 105.83/6.61                          | 99.75/6.33 | 94.34/6.49 | 108.36/6.60                        | 109.08/7.39 | 113.65/6.84 |
| DMSO-d <sub>6</sub>    | 104.87/7.05                          | 98.98/6.21 | 94.37/6.57 | 107.82/7.05                        | 110.18/7.71 | 114.33/7.07 |

formed, as evidenced by NMR spectra of the total crude products (Supplemental Figs. S4 and S5). Generally, most of the triclin remained unreacted, but no monolignol remained in the mixture after reaction, as indicated by the strong signals of C3/H3, C6/H6, and C8/H8 from triclin and the disappearance of C7/H7 and C8/H8 correlations at  $\delta_C/\delta_H$  approximately 130.0/6.50 and approximately 128.0/6.21 from the side chain double bonds of the monolignols. Reaction between coniferyl alcohol and triclin generated a low yield of a cross-coupled product, compound 14a, as demonstrated by the correlation peaks at  $\delta_C/\delta_H$  74.0/5.01, 89.6/4.19, 67.7/3.75, 67.7/3.42 and 73.5/5.02, 89.1/4.35, 61.1/3.89, 61.1/3.52, which correspond to C $\alpha$ /H $\alpha$ , C $\beta$ /H $\beta$ , C $\gamma$ /H $\gamma$ , and C $\gamma$ /H $\gamma$  in the *anti*- and *syn*-diastereomers, respectively. The intensities of the related peaks from each of the isomers were similar, implying that the *anti*:*syn* ratio is close to 1:1. The other products in the mixture were  $\beta$ -5-,  $\beta$ -O-4-, and minor  $\beta$ - $\beta$ -coupled structures resulting from coniferyl alcohol homocoupling. The chemical shifts of  $\alpha$ ,  $\beta$ , and  $\gamma$  in the coniferyl alcohol  $\beta$ -O-4-coupled dimer are quite different from those of compound 14a, allowing these two compounds to be easily distinguished in the products. The formation of compound 14c, T-(4'-O-β)-H, from triclin and *p*-coumaryl alcohol under peroxidase/H<sub>2</sub>O<sub>2</sub> coupling conditions occurred in relatively high yield, with only a small amount of triclin remaining, as was confirmed by comparing the major dimeric product with the synthesized authentic compound. However, compound 14b, T-(4'-O-β)-S, was barely detected in the HSQC spectrum from the triclin-sinapyl alcohol coupling mixture. The major product in this reaction, as in the reaction with sinapyl alcohol alone, was syringaresinol, the  $\beta$ - $\beta$ -homocoupled dimer.

Ag<sub>2</sub>O was also used as an oxidant to invoke cross coupling between triclin and monolignols. HSQC spectra of the total crude products from radical coupling between triclin and the monolignols are shown in Supplemental Figure S5. If monolignols were added in a single shot, no T-(4'-O-β)-monolignol cross-coupling adducts were detected by NMR. In contrast, compounds 14a to 14c were successfully obtained by the slow addition of each respective monolignol into the triclin solution containing Ag<sub>2</sub>O; this sequence is viable because triclin is inert and does not undergo homodimerization under these conditions. Logically, the reactivity of triclin in cross-coupling reactions with a monolignol is lower than that of simple dimerization of the monolignol. Therefore, as seen in the results described earlier, in the presence of equivalent quantities

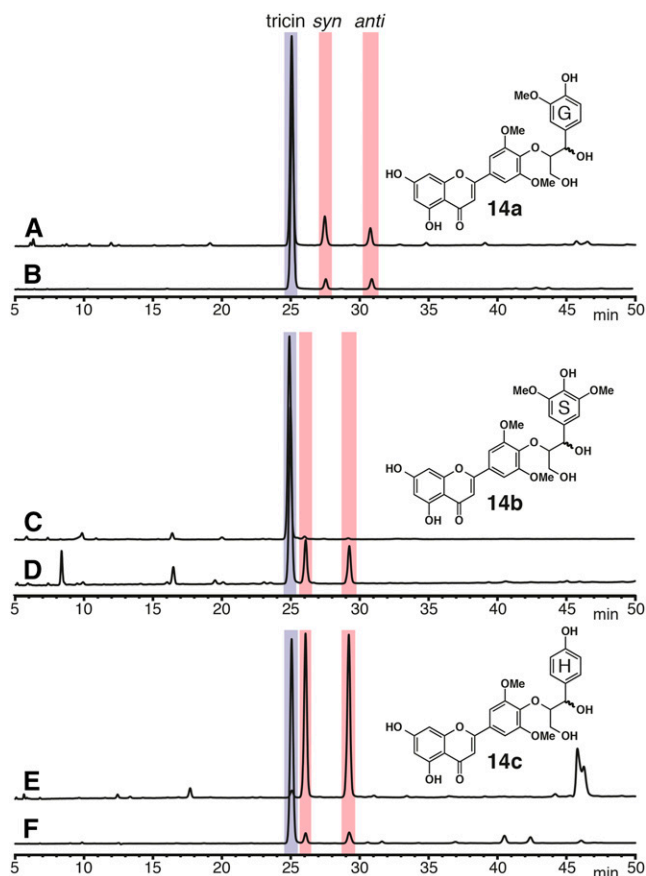
of triclin and a given monolignol, the monolignol will predominantly undergo homocoupling. However, if the concentration of monolignols is limited, the occurrence of cross coupling is enhanced.

The LC-MS profiles of the products 14a to 14c, T-(4'-O-β)-G/S/H, obtained from the triclin-monomolignol cross-coupling reactions, are shown in Figure 2. The peak eluting at 25.3 min with mass-to-charge ratio (*m/z*) of 331 in the positive-ion mode ([M+H]<sup>+</sup>) and 329 in the negative-ion mode ([M-H]<sup>-</sup>) belong to triclin. The *syn*-diastereoisomers of compounds 14a to 14c eluted immediately following triclin at 27.7, 26.4, and 26.3 min, with *m/z* values of 527, 557, and 497 in positive-ion mode and 525, 555, and 495 in negative-ion mode. The later-eluting peaks at 31.1, 29.5, and 29.4 min, with similar peak areas and identical *m/z* values, were obtained from the *anti*-diastereoisomers of compounds 14a to 14c. The yields of cross-coupling products were determined by first combining the peak areas of the peaks from both isomers, the value of which was then used for quantification based on the standard curve. Radical coupling between triclin and monolignols effected by peroxidase/H<sub>2</sub>O<sub>2</sub> produced compounds 14a to 14c in 12.6%, 1.2%, and 51.4% yields; when Ag<sub>2</sub>O was used as the oxidant, the yields were 6.9%, 15.3%, and 10.2%. Previous studies have revealed that the oxidation rate of sinapyl alcohol itself by horseradish peroxidase is quite low (Takahama et al., 1996). Transfer of the radical from *p*-coumarate to this alcohol may aid in the formation of syringyl-rich lignin in some plant species (Grabber, 2005; Hatfield et al., 2008; Ralph, 2010). Without the use of radical transfer agents (or a direct oxidant like Ag<sub>2</sub>O), a much lower yield for 14b than for 14a and 14c would be expected from radical coupling using peroxidase as the oxidant.

Based on the resonance structures that exist for potential triclin phenolic radicals, radical coupling with a monolignol could theoretically take place at the 4'-O-, 3-C-, 8-C-, 7-O-, 6-C-, and 5-O-positions, as shown in Figure 3, indicating the diversity of structures that

**Table II.** Yields of T-(4'-O-β)-G (14a), T-(4'-O-β)-S (14b), and T-(4'-O-β)-H (14c) dimers in the coupling reactions of triclin with coniferyl, sinapyl, and *p*-coumaryl alcohols, respectively, catalyzed by peroxidase/H<sub>2</sub>O<sub>2</sub> and Ag<sub>2</sub>O

| Triclin-Monomolignol Dimer | Peroxidase/H <sub>2</sub> O <sub>2</sub> | Ag <sub>2</sub> O |
|----------------------------|--|-------------------|
|                            | %  |                   |
| T-(4'-O-β)-G (14a)         | 12.6                                     | 6.9               |
| T-(4'-O-β)-S (14b)         | 1.2                                      | 15.3              |
| T-(4'-O-β)-H (14c)         | 51.4                                     | 10.2              |



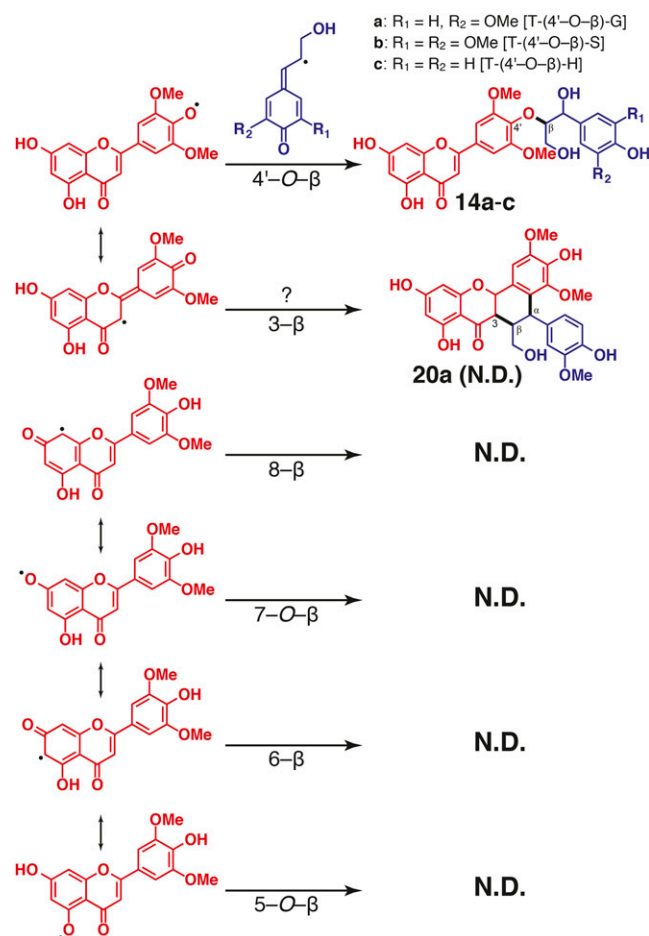
**Figure 2.** HPLC profiles of the products from radical coupling reactions between triclin and monolignols catalyzed by peroxidase/H<sub>2</sub>O<sub>2</sub> (A, C, and E) and Ag<sub>2</sub>O (B, D, and F).

could be formed as a result. However, only the 4'-O-β-coupling products 14 were identified by NMR in the products from peroxidase/H<sub>2</sub>O<sub>2</sub> or Ag<sub>2</sub>O oxidation. In fact, guaiacyl flavonolignan 14b has been isolated from oat (Wenzig et al., 2005), *Hypparrhenia hirta* (Bouaziz et al., 2002), and *Sasa veitchii* (Nakajima et al., 2003). The 4'-O site on triclin can also be glucosylated to form a flavonoid glucoside (with possible further acylation of the Glc by *p*-coumarate); such a compound was isolated from *Acacia nilotica* (Khanam et al., 2011). C-O- and C-C-linked glycosides of triclin at the 5-O-, 7-O-, and 6-C-positions have also been reported (Bouaziz et al., 2002; Duarte-Almeida et al., 2007; Van Hoyweghen et al., 2010; Nakano et al., 2011). These compounds and others have also recently been detected in sugarcane (Bottcher et al., 2013). A dimer (compound 20a in Fig. 3) containing a flavonoid moiety (C<sub>3</sub>) connected to a monolignol-derived unit (C<sub>β</sub>) that has the appearance of deriving from 3-β cross coupling was isolated from oat and *Hydnocarpus wightiana* (Parthasarathy et al., 1979; Wenzig et al., 2005). This type of structure could not be identified among the radical coupling products examined in this study, either by NMR or LC-MS analysis. Furthermore, a reasonable reaction mechanism for the formation of compound 20a

following the coupling of triclin and coniferyl alcohol is not obvious; an earlier intermediate in the biosynthetic pathway may react to form this adduct. No evidence has yet been found to suggest that linkages between monolignols and triclin occur at the 5-O-, 7-O-, 3-C-, 6-C-, and 8-C-positions under biomimetic lignification conditions. Instead, in this study, we prove the incorporation of triclin into lignin via 4'-O-β-coupling with monolignols, as might be anticipated. To determine whether triclin in the lignin polymer has linkages to carbohydrates will require further investigation.

### Evidence for the Incorporation of Tricin into the Lignin Polymer

The objective of this study was to experimentally support the initial claims made regarding the existence of triclin in the milled wood lignin from wheat straw (del Río et al., 2012) by (1) providing the required diagnostic evidence for the presence of triclin in native lignin (i.e. by

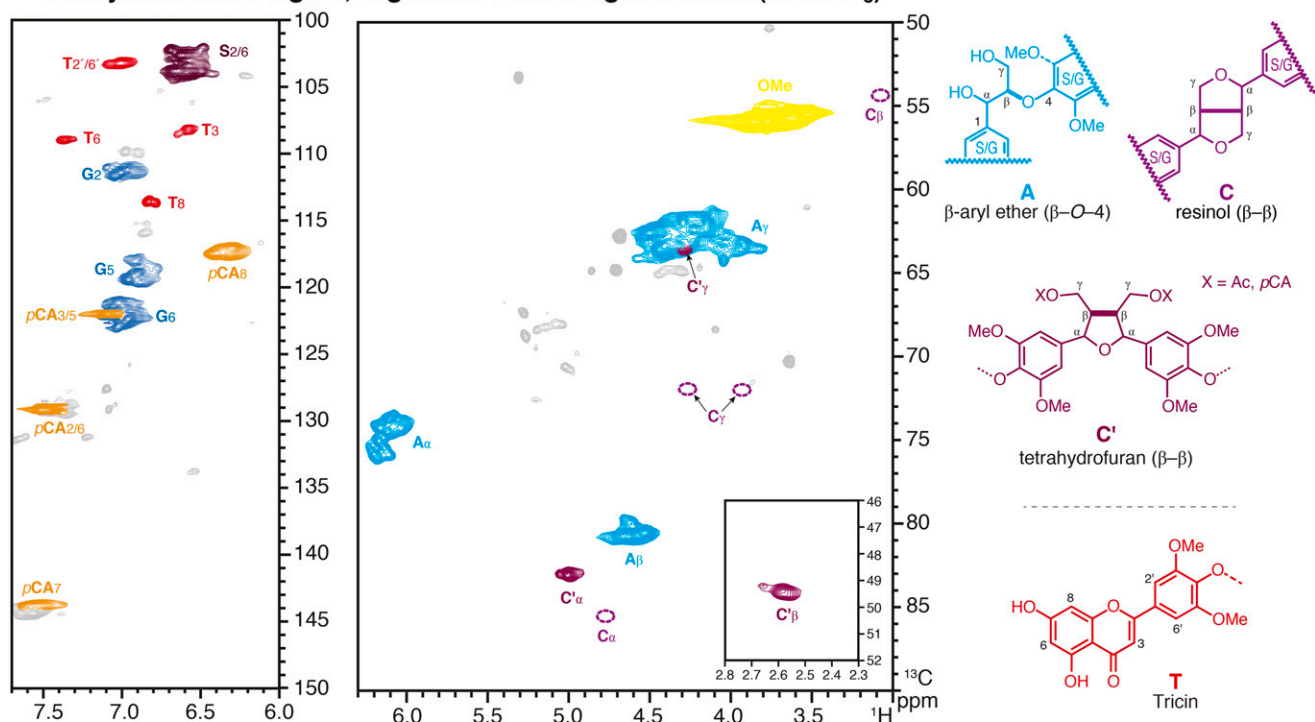


**Figure 3.** Possible radical coupling reactions between triclin and monolignols. N.D. signifies that the product was not detected here. Compound 20a was reported previously (Wenzig et al., 2005), but it was not found in this study, and it is not absolutely clear whether or how it could derive from this coupling reaction (see text).

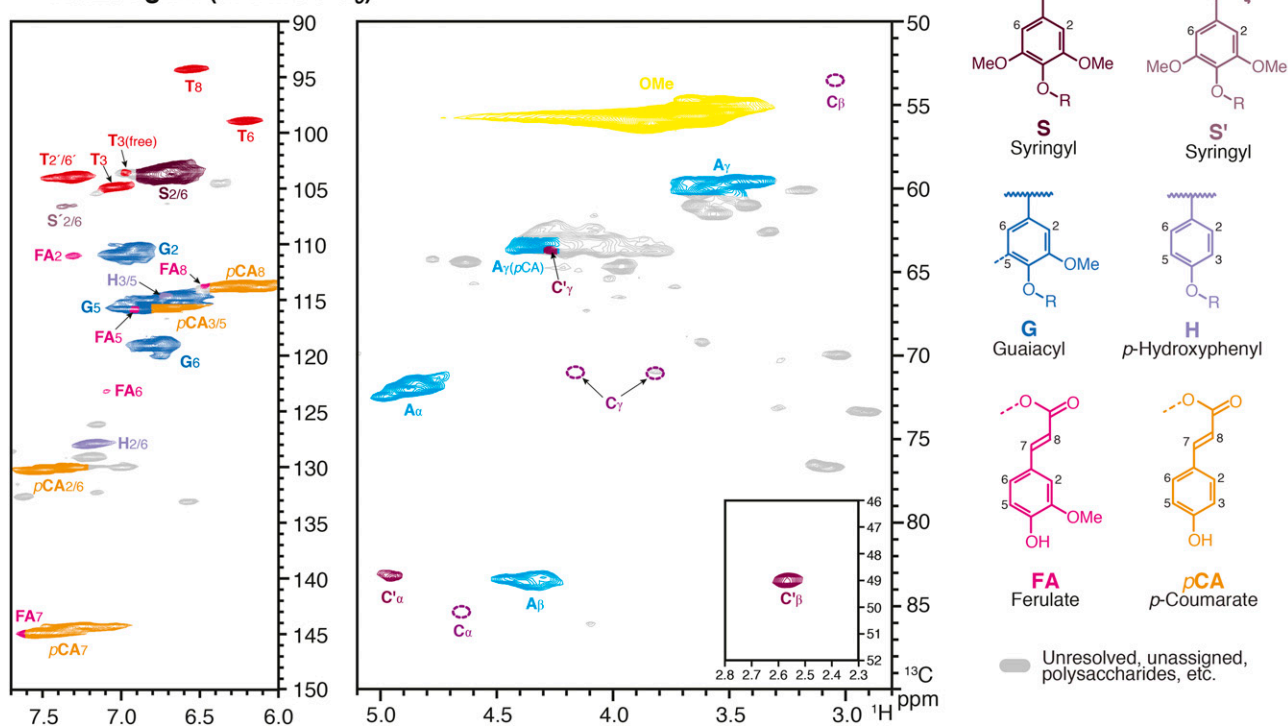
showing that NMR data from synthesized authentic model compounds agree with the data derived from native lignins); (2) confirming the linkage between triclin

and monolignols and demonstrating that such units can be formed by the radical coupling reactions that typify lignification; and (3) showing that triclin is present

### A Acetylated maize lignin, High molecular weight fraction (in $\text{CDCl}_3$ )



### B Maize lignin (in $\text{DMSO}-d_6$ )

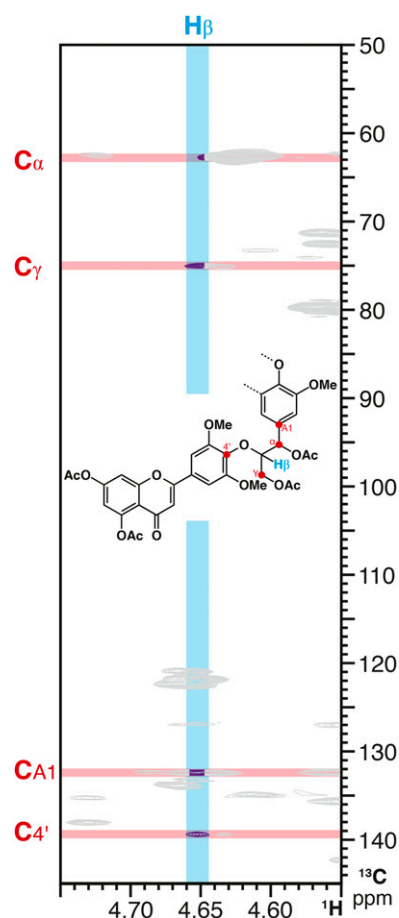


**Figure 4.** HSQC spectra of the highest  $M_n$  fraction of an acetylated maize lignin (in  $\text{CDCl}_3$ ; A) and maize lignin (unacetylated, in  $\text{DMSO}-d_6$ ; B).

(covalently linked) in even the highest  $M_r$  lignin fractions. Thus, NMR data from these synthetic and authenticated model compounds (before and after acetylation) were acquired and used for comparison with tricrin-containing moieties in maize lignin. As noted above, results from biomimetic radical coupling reactions between tricrin and monolignols suggest that tricrin is compatible with lignification. Therefore, lignin preparations extracted by acetic acid pretreatment of maize stover (Pan and Sano, 1999, 2005) were used to confirm the existence of tricrin in the lignin polymer. The chemical shifts of the C-H correlations corresponding to C3/H3, C6/H6, C8/H8, and C2',6'/H2',6' of compounds 14a to 14c and 19 at  $\delta_C/\delta_H$  104.9/7.05, 99.0/6.21, 94.4/6.57, and 104.3/7.31 ppm in dimethyl sulfoxide (DMSO)- $d_6$  for nonacetylated samples or at 108.3/6.60, 109.0/7.39, 113.5/6.84, and 103.4/7.05 ppm in  $CDCl_3$  for acetylated samples were clearly identified in the HSQC spectra of maize lignins in the same solvents (Fig. 4). To provide evidence that tricrin is bonded to lignin units, the HSQC spectrum of compound 8 was compared with those of compounds 14a to 14c, 19, and maize stover lignin. Results showed that the C3/H3, C6/H6, C8/H8, and C2',6'/H2',6' correlations in free tricrin in DMSO- $d_6$  are at  $\delta_C/\delta_H$  103.6/7.05, 98.9/6.30, 94.3/6.63, and 104.4/7.37, differing from those of tricrin connected to monolignols or lignin units via 4'-O-ether bonds (Supplemental Fig. S2). Such differences in chemical shifts were large enough to allow the distinction of free from etherified tricrin, even in the polymeric samples. Heteronuclear multiple-bond correlation (HMBC) experiments provided more direct evidence for covalent bonding between tricrin and the monolignol-derived lignin units in the acetylated maize lignin (Fig. 5; Supplemental Fig. S3). The three-bond correlation between C-4' and H $\beta$  at  $\delta_C/\delta_H$  139.5/4.65 ppm in the HMBC spectrum of acetylated maize lignin sample was validated by comparison with that of model compound 19, T-(4'-O- $\beta$ )-S-(4-O- $\beta$ )-G. This H $\beta$  also correlated with C $\alpha$ , C $\gamma$ , and C $A_1$  at  $\delta_C$  63.9, 76.3, and 132.5 ppm, revealing the 4'-O- $\beta$ -ether bonds between tricrin and lignin units in maize stover lignin. To further elucidate whether tricrin is incorporated into high-molecular mass lignin chains, rather than simply being bonded to monolignols to form dimers or short-chain oligomers, the acetylated maize stover lignin was fractionated via gel permeation chromatography (GPC). Eight fractions were collected, with the first two fractions containing high- $M_r$  components ( $M_w = 5,670$ ,  $M_n = 1,580$  for the first fraction,  $M_w = 2,440$ ,  $M_n = 970$  for the second fraction) accounting for 73% of the sample. Based on NMR characterization, the first four fractions with large to medium  $M_r$  components all contained covalently bonded tricrin. The HSQC spectrum of the highest  $M_r$  fraction is shown in Figure 4A.

### Tricin Initiates Lignin Chains

To date, the accumulated evidence has indicated that tricrin is only incorporated into the polymer (above) in the form of 4'-O- $\beta$ -coupled products 14 and their higher



**Figure 5.** C4'-O-H $\beta$  correlation in the HMBC spectrum of maize stover lignin.

oligomers. We are not stating that tricrin 4'-O-5-coupled units cannot arise from the coupling of a tricrin (radical) with a lignin oligomer (radical), nor that units such as compound 20 will not be found in the lignin. Theoretically, 3-coupled products (like  $\beta$ -coupled products from normal monolignols) are possible, but they have not been shown here in either the biomimetic coupling reactions or in the polymers of the natural samples we examined, notwithstanding the report of compound 20 (Fig. 3) in the literature (Parthasarathy et al., 1979; Wenzig et al., 2005). All of these theoretical products would have very different NMR characteristics from those noted here. Therefore, we deduce that tricrin predominantly incorporates into lignin via 4'-O- $\beta$ -coupling, which by necessity localizes each tricrin unit at one terminus of its lignin chain, and that terminus must be at the starting end of that chain. Tricin, therefore, acts as a nucleation site for lignin chain growth in monocots, a role that has previously been proposed for ferulate on arabinoxylans (Ralph et al., 1995).

### Resolution of a Monocot Lignin Dilemma

The observation that tricrin may be the initiator of many of the lignin chains in the polymer helps to explain an old dilemma arising from many maize lignin



spectra [i.e. that, despite its being an S-G lignin with an S:G ratio similar to those found in many dicots and hardwoods, there is little or no evidence for (syringa)-resinol structures in maize lignin]. That maize lignin has essentially no resinol structures, the correlation positions for which are indicated by the magenta-colored dashed ellipses, can be seen in Figure 4. This means that the polymer chain is not (significantly) started via monolignol dimerization per se (as it is in hardwoods/dicots and softwoods). To our knowledge, the data presented here provide the two-part explanation for the first time. In large measure, the lack of resinols starting polymerization is because the chain is initiated/nucleated by triclin in monocots. Such nucleation behavior has previously been attributed to ferulates on arabinoxylans, and this may also be present here; it is hard to observe or quantify (Ralph et al., 1995). However, careful examination of the NMR spectra reveals that monomer dimerization is in fact occurring, but the monomers in this case are acylated monolignols. Maize lignin is  $\gamma$ -acylated (by *p*-coumarate or acetate; Ralph, 2010). Such acylated monomers cannot cyclize after  $\beta$ - $\beta$ -coupling to give resinol structures (Lu and Ralph, 2008; Ralph, 2010). In fact, the  $\beta$ - $\beta$ -product arising from the coupling of two acylated monolignols, the tetrahydrofuran C' in Figure 4, is readily seen in the side chain region of the HSQC spectra in Figure 4. At some point, it will be intriguing to understand how and why the dimerization reactions are dominated by the acylated monolignols rather than the parent monolignols themselves, but this is not the primary concern here. The observation that syringaresinol is found in single-shot coupling reactions with triclin and sinapyl alcohol, but is less prominent when sinapyl alcohol is slowly added to the triclin solution, further supports the theory that lignification is an end-wise process.

## CONCLUSION

Three dimeric model compounds, the 4'-*O*- $\beta$ -coupling products of triclin with *p*-coumaryl, coniferyl, and sinapyl alcohols, 14a to 14c, and the trimer 19 [T-(4'-*O*- $\beta$ )-S-(4-*O*- $\beta$ )-G] that would result from 4'-*O*- $\beta$ -coupling with sinapyl alcohol and then coniferyl alcohol, were synthesized. Radical coupling between triclin and all three monolignols produced compounds 14a to 14c in yields of 12.6%, 1.2%, and 51.4% under peroxidase-catalyzed oxidation conditions and 6.9%, 15.3%, and 10.2% under Ag<sub>2</sub>O oxidation. By comparison of HSQC and HMBC NMR data from maize stover lignin with those from authentic model compounds, we demonstrated that triclin in maize stover lignin is linked to lignin units via 4'-*O*- $\beta$ -ether bonds. The presence of triclin in high-*M<sub>r</sub>* fractions of GPC-fractionated acylated maize stover lignin was confirmed by NMR analysis. Based on all of these results, it can confidently be asserted that triclin is not only able to couple with monolignols and participate in lignification but that it regularly does so in monocots, where triclin is found

covalently bound into the very lignin polymer itself. In addition, as the triclin that is observed in lignin NMR spectra can only arise from the participation of triclin in the initial coupling reactions with a monolignol, it must be placed at the beginning of a polymer chain, thus acting as an initiator of sorts. The prevalence of triclin in these lignins strongly suggests that triclin units have a role in nucleating the growth of the lignin polymer in monocots. We have also resolved the dilemma of the almost complete absence of syringaresinol units (the dimer that typically starts a lignin chain in pure S-G lignins) in maize and other monocot lignins; the chains are either started by triclin or by dimerization of acylated monolignols that give rise to novel  $\beta$ - $\beta$ -linked dimers that are readily seen in the spectra of the maize lignin here.

Taken together, these findings provide reliable and substantive evidence that triclin is incorporated into maize (and other monocot) lignins via a free radical coupling mechanism and is covalently bound into the lignin polymer (i.e. that triclin should be regarded, in a general sense, as an authentic lignin monomer in monocots). This study, therefore, not only supports the striking observation that monocots routinely incorporate a flavonoid (derived from an entirely different biosynthetic pathway) into their lignins, but it also serves to highlight the remarkable ways in which lignification in monocots differs from the process in other plant classes.

## MATERIALS AND METHODS

### General

All chemicals and solvents used in this study were purchased from commercial sources and used without further purification. Horseradish peroxidase (type II; 180 pyrogallol units mg<sup>-1</sup>) was provided by Sigma.

Maize (*Zea mays*) stover lignin was obtained from acetic acid-pretreated (Pan and Sano, 1999, 2005) Great Lakes Bioenergy Research Center-supplied 2009 standard corn/maize stover. The 2009 corn stover was obtained from corn planted at the Arlington Agricultural Research Station, Wisconsin, on May 5, 2009, and harvested on November 24, 2009. A modified combine to chop whole corn plants and separate the grain from the rest of the plant material was used for harvesting purposes. The chopping height of the plant was approximately 10 inches from the ground. The chopped plant material (except the grain) was then collected into 36- × 21-inch meshed sacks/bags and placed in a dryer at the Arlington Agricultural Research Station, at 50°C for 10 d, before being ground using an 8-inch overhung disintegrator mill with a 5-mm screen (Circ U-flow model 18-7-300; Shutte Buffalo Hammermill). The ground biomass (5-mm particle size) was then collected in plastic bags (each weighing 1.5 kg) and barcoded and labeled via the STARLIMS Laboratory Information Management System. These biomass bags were then stored at the Great Lakes Bioenergy Research Center's temperature-controlled (approximately 25°C) storage facility at the Arlington Agricultural Research Station and were shipped to customers upon request. The material was analyzed by the National Renewable Energy Laboratory ([http://www.nrel.gov/biomass/analytical\\_procedures.html](http://www.nrel.gov/biomass/analytical_procedures.html)) as having 14.3% (w/w) lignin.

Flash chromatography was performed with Biotage snap silica cartridges on an Isolera One instrument (Biotage) using a hexane/ethyl acetate (EtOAc) gradient as the eluent. Preparative thin-layer chromatography (TLC) plates (1- or 2-mm thickness, normal phase) were purchased from Analtech and were run using hexane/EtOAc or MeOH/dichloromethane as the eluent. NMR spectra were recorded on a Bruker Biospin AVANCE 500- or 700-MHz spectrometer fitted with a cryogenically cooled 5-mm TCI (500 MHz) or TXI (700 MHz) gradient probe with inverse geometry (proton coil closest to the sample). Bruker's Topspin 3.1 (Mac) software was used to process spectra. The central solvent peaks were used as internal references ( $\delta_C/\delta_H$ , acetone-*d*<sub>6</sub>,

29.84/2.04;  $\text{CDCl}_3$ , 77.00/7.26; DMSO, 39.50/2.49 ppm). The standard Bruker implementations of one- and two-dimensional (gradient-selected correlation spectroscopy, HSQC, and HMBC) NMR experiments were used for routine structural assignments of newly synthesized compounds. HPLC was conducted with LC-20AD pumps, a SIL-20AC HT autosampler, a CTO-20A column oven, a CBM-20A controller, an FRC-10A fraction collector, and using a Kinetex PFP (250  $\times$  4.6 mm, 5  $\mu\text{m}$ ; Phenomenex) column at 40°C. Detector wavelength ranged from 200 to 400 nm via an SPD-M20A photodiode array detector. The injection volume was 20  $\mu\text{L}$ . Water and MeOH or acetonitrile served as the mobile phase in a gradient mode with a flow rate of 0.7 mL  $\text{min}^{-1}$ . LC-MS analysis was carried out on a Shimadzu LCMS-2020 system using the dual ion source method for ionization, with both positive- and negative-ion spectra being recorded. A Kinetex XB-C18 (250  $\times$  4.6 mm, 5  $\mu\text{m}$ ; Phenomenex) column was used in the liquid chromatography system, using water (solvent A) and acetonitrile (solvent B), each containing 0.1% (v/v) formic acid, as the mobile phase. An aliquot (1  $\mu\text{L}$ ) of the approximately 0.5 mg  $\text{mL}^{-1}$  sample was injected and analyzed at a flow rate of 0.7 mL  $\text{min}^{-1}$  using a gradient method, with 25% (v/v) B at 0 min, 40% (v/v) B at 50 min, 25% (v/v) B at 50.1 min, and 25% (v/v) B at 65 min. GPC fractionation was carried out using gravity feed on a Bio-Rad Bio Bead S-X3 column (96  $\times$  5.1 cm, neutral, porous styrene-divinylbenzene copolymer). About 200 mg of acetylated maize stover lignin was dissolved in dichloromethane and introduced onto the top of the column. The high- $M_n$  exclusion started at about 76 min. A total of eight fractions were collected, out to 180 min.

## Synthesis of Oligomers

### Procedure for Acetylation

The starting material was dissolved in pyridine:acetic anhydride (2:1, v/v) and stirred for 2 h at room temperature. The solution was transferred to a separatory funnel and extracted with EtOAc and washed several times with acidic water to eliminate most of the pyridine. The organic phase was washed with saturated ammonium chloride ( $\text{NH}_4\text{Cl}$ ) solution, dried over anhydrous magnesium sulfate ( $\text{MgSO}_4$ ), filtered, and evaporated under reduced pressure to give the acetylated products. The yield ranged from 92% to 96% (w/w). (All yields here and in the following text are weight-based percentages.) NMR data for all synthetic compounds are provided in Supplemental Text S1.

### Compound 3

A reported method was used for methoxymethylation (St. Denis et al., 2010). Ketone 1 (8.08 g, 43 mmol) was added to dichloromethane (240 mL), and the resulting mixture was cooled to 0°C, to which *N,N*-diisopropylethylamine (24 mL, 120 mmol) was slowly added. After stirring for 20 min, chloromethyl methyl ether (9.20 g, 98.9 mmol) was added dropwise. The mixture was stirred for 20 min at 0°C. Water was added to quench the reaction, and the aqueous layer was separated and washed with chloroform ( $\text{CHCl}_3$ ; 3  $\times$  100 mL). The combined organic layers were washed with saturated  $\text{NH}_4\text{Cl}$  (150 mL), dried over anhydrous  $\text{MgSO}_4$ , filtered, and concentrated to give a dark-colored oil. The crude product was purified by flash chromatography to obtain compound 3 (87% yield) as a clear oil that produced a white solid upon standing for 1 week.

### Compound 4

Syringaldehyde 2 (15 g, 82.3 mmol) was dissolved in *N,N*-dimethylformamide (DMF), to which benzyl bromide (13.41 g, 78.4 mmol) and potassium carbonate ( $\text{K}_2\text{CO}_3$ ; ground to fine powder; 10.84 g, 78.4 mmol) were added. After 12 h, the  $\text{K}_2\text{CO}_3$  was filtered off, and the organic solvent was evaporated under reduced pressure at 70°C. The obtained crude product was dissolved in  $\text{CHCl}_3$  (100 mL) and washed with 1 M NaOH (5  $\times$  100 mL). Then, the separated organic layer was washed with saturated  $\text{NH}_4\text{Cl}$  (150 mL), dried over anhydrous  $\text{MgSO}_4$ , filtered, and the solvent was removed under reduced pressure. Compound 4 (83.4% yield) was obtained as yellowish-white flaky crystals (melting point [m.p.], 61°C–62°C) from crystallization in ethanol.

### Compound 5

Compounds 3 (8.92 g, 42.3 mmol) and 4 (11.51 g, 42.3 mmol) were dissolved in MeOH (600 mL), to which aqueous potassium hydroxide (23.71 g, 423 mmol in 35 mL of water) was added slowly. After 24 h, the solution was

neutralized with 6 M HCl, and the solute was precipitated out as a yellow solid. The MeOH was evaporated at 45°C under reduced pressure. EtOAc (150 mL) and water (150 mL) were added to extract the products, and the aqueous layer was washed with EtOAc (3  $\times$  100 mL). After washing with saturated  $\text{NH}_4\text{Cl}$  solution (150 mL), drying over anhydrous  $\text{MgSO}_4$ , and filtration, the combined organic layers were removed by evaporation, generating compound 5 (56.8% yield) as orange needle-like crystals (m.p., 82°C–83.0°C) after crystallization from ethanol.

### Compound 6

Cyclization of compound 5 was accomplished by dissolving it (5 g, 9.8 mmol) in pyridine with  $\text{I}_2$  (2.48 g, 9.8 mmol) at 120°C for 5 h. Then, the solution was poured into sodium bisulfite solution (5% [w/w], 300 mL). EtOAc (3  $\times$  100 mL) was used to extract the product. The combined EtOAc solutions were dried over anhydrous  $\text{MgSO}_4$ , filtered, and concentrated. The obtained crude products were subjected to acetylation and then TLC purification using hexane and EtOAc (1:1, v/v) as the eluent. Compound 6 (21% yield) was obtained as light-yellow needle crystals (m.p., 140°C–141°C) after crystallization from ethanol.

### Compound 7

Crystalline compound 6 (200 mg, 0.43 mmol) was dissolved in boiling ethanol (200 mL). After cooling to room temperature, palladium on activated carbon (10% [w/w] palladium, 40 mg) was added. The reaction mixture was stirred at room temperature under a hydrogen atmosphere for 45 min. The catalyst powder was filtered using a membrane filter (Teflon; 0.22- $\mu\text{m}$  pore size), and the organic solution was collected, concentrated, and then precipitated into water. The precipitate was lyophilized to give compound 7 (90.9% yield) as a yellow powder.

### Compound 8

Deprotection of the MOM group in 7 was accomplished as follows. Compound 7 (100 mg, 0.27 mmol) was dissolved in MeOH/ $\text{CHCl}_3$  (1:1, v/v, 8 mL), and concentrated HCl (1 mL) was added. After refluxing for 45 min, the solvent was evaporated, and water (100 mL) was added to precipitate the product. The precipitate was collected by filtration and lyophilized to give compound 8 (89.6% yield) as a yellow powder.

### Compounds 10a to 10c

Bromination of commercial compounds 9a to 9c was accomplished via traditional methods as illustrated by the synthesis of compound 10a. Acetylated compound 9a (2.5 g, 12 mmol) and pyridinium tribromide (3.84 g, 12 mmol) were dissolved in EtOAc (100 mL). The mixture was stirred for 2 h at room temperature. Saturated  $\text{NaHCO}_3$  (100 mL) was used to quench the reaction. The EtOAc layer was separated, dried over anhydrous  $\text{MgSO}_4$ , filtered, and evaporated under reduced pressure. Crystallization from ethanol afforded compound 10a (76.1% yield) as white needle crystals (m.p., 91.5°C–92°C). Compounds 10b (60.6% yield, oil) and 10c (68.6% yield, oil) were synthesized similarly.

### Compounds 11a to 11c

The procedure for the preparation of compound 11a is typical. Compound 10a (383.5 mg, 1.34 mmol) and compound 7 (500 mg, 1.34 mmol) were dissolved in DMF (25 mL) with  $\text{K}_2\text{CO}_3$  (184.4 mg, 1.34 mmol). After the reaction was complete (monitored by TLC),  $\text{K}_2\text{CO}_3$  was filtered off, and DMF was evaporated under reduced pressure at 70°C. EtOAc (75 mL) and water (75 mL) were added to the resulting material. The aqueous layer was removed, and the organic layer was washed with water (3  $\times$  50 mL). After drying over anhydrous  $\text{MgSO}_4$ , filtration, and evaporation, compound 11a (98.8% yield) was obtained as a yellow oil. Compounds 11b (95.3% yield, oil) and 11c (97.2% yield, oil) were synthesized analogously.

### Compounds 12a to 12c

A detailed procedure is given using 12a as the example. Compound 11a (600 mg, 1.03 mmol) and formaldehyde solution (37% [w/w], 83.9 mg, 1.03 mmol) were dissolved in 1,4-dioxane (25 mL), to which  $\text{K}_2\text{CO}_3$  (1,427 mg, 10.3  $\mu\text{mol}$ ) was added. The reaction mixture was maintained at 35°C overnight. A second equivalent of formaldehyde solution (37% [w/w], 83.9 mg, 1.03 mmol) was then added. After the reaction was complete (monitored by TLC),  $\text{K}_2\text{CO}_3$  was

filtered off, and the dioxane was evaporated. The pure compound 12a (81.3% yield) was acquired as a light yellow oil after flash chromatographic purification using EtOAc and hexane (1:1, v/v) as the eluent. Compounds 12b (76% yield) and 12c (73% yield) were obtained using the same procedure.

### Compounds 13a to 13c

Deprotection of the MOM group in compounds 12a to 12c was achieved using the same method as described for compound 8. Compounds 13a to 13c were dark yellow powders after freeze drying, with yields of 75.1%, 76.3%, and 80.3%.

### Compounds 14a to 14c

Compound 13a (50 mg, 0.09 mmol) was reduced by stirring with sodium borohydride (18 mg, 0.47 mmol) in ethanol (5 mL) at room temperature. The reaction was monitored by TLC using dichloromethane and MeOH (40:1, v/v) as the eluent. When the reaction was complete, the ethanol was evaporated under reduced pressure at 45°C. The product was dissolved in EtOAc and water (25 mL, 10:1, v/v), and 6 M HCl was added (1 mL). The mixture was stirred at room temperature for 30 min to ensure the cleavage of the borate intermediates. The resulting mixture was washed with 0.5 M NaHCO<sub>3</sub> (3 × 20 mL) and saturated NH<sub>4</sub>Cl solution (1 × 20 mL). The separated EtOAc layer was dried over anhydrous MgSO<sub>4</sub>, filtered, and evaporated to give compound 14a [T-(4'-O-β)-G; 81.7% yield] as a yellow powder. Similar methods were applied for compounds 14b [T-(4'-O-β)-S; 84.6% yield] and 14c [T-(4'-O-β)-H; 88.7% yield].

### Compound 15

The phenolic acetate was selectively removed by ammonium acetate in MeOH. Compound 12b (125 mg, 0.20 mmol) and ammonium acetate (150 mg, 20 mmol) were dissolved in MeOH (20 mL) at 50°C. The solution was stirred overnight, and the MeOH was evaporated at 45°C. EtOAc (25 mL) and water (25 mL) were added to extract the product. The organic layer was washed with water (2 × 25 mL), followed by saturated NH<sub>4</sub>Cl (25 mL). After drying over anhydrous MgSO<sub>4</sub>, filtration, and evaporation, the residue was purified by chromatography using EtOAc and hexane (1:1, v/v) to give compound 15 (56% yield) as a clear yellowish oil.

### Compounds 16 to 19

The T-(4'-O-β)-G-(4-O-β)-S adduct 19 was synthesized by similar methods to those described for compounds 11a to 11c, 12a to 12c, 8, and 14a to 14c. The yields for compounds 16 to 19 were 97%, 56.8%, 90.9%, and 92.1%.

## Radical Coupling Reaction between Tricin and Monolignols

Coniferyl, sinapyl, and *p*-coumaryl alcohols were synthesized from their corresponding acids by ethyl esterification and reduction (Quideau and Ralph, 1992). Using coniferyl alcohol as an example to illustrate the process, ferulic acid (1 g, 5.1 mmol) was dissolved in ethanol (10 mL), to which acetyl chloride (1 mL) was slowly added. After the reaction was complete (monitored by TLC), the organic solvent was removed under reduced pressure. The resulting ethyl ferulate was crystallized from EtOAc/hexane to give crystalline ethyl ferulate (65% yield). Ethyl ferulate (0.5 g, 2.2 mmol) was dissolved in dry tetrahydrofuran (10 mL), and diisobutylaluminum hydride (3.2 g, 22.5 mmol) was slowly added. The reaction was quenched with 1 M HCl solution after 2 h. EtOAc and water were added to extract the product, and the separated EtOAc layer was washed with 1 M HCl solution (3 × 50 mL) and saturated NH<sub>4</sub>Cl solution (1 × 50 mL). After drying over anhydrous MgSO<sub>4</sub>, filtration, and evaporation, the obtained coniferyl alcohol was precipitated in water and then lyophilized to give a white powder (87.8% yield). Sinapyl (73% yield) and *p*-coumaryl (76% yield) alcohols were synthesized analogously. The NMR data of the three obtained monolignols are consistent with those in the NMR database of lignin model compounds (Ralph et al., 2005).

Peroxidase and Ag<sub>2</sub>O were used as oxidative reagents for the radical coupling reactions in this study. The detailed procedure is illustrated using coniferyl alcohol. Tricin (15.7 mg, 47.5 μmol) and coniferyl alcohol (8.6 mg, 47.5 μmol) were dissolved in 200 mL of acetone/phosphate buffer (pH 5, 20 mM, 2:3, v/v). H<sub>2</sub>O<sub>2</sub>-urea complex (4.9 mg, 52.1 μmol) and peroxidase (0.5 mg) were added. The reaction solution was stirred at room temperature for 2 h.

After acidification to pH 3 with 1 M HCl, the reaction mixture was placed in a hood to allow the acetone to evaporate until the radical coupling compounds precipitated, and EtOAc (3 × 100 mL) was added to extract the products. After separation, drying over anhydrous MgSO<sub>4</sub>, filtration, and evaporation, a mixture of coupling products was obtained (86.4% overall yield). The radical coupling reactions of tricin/sinapyl alcohol (87% overall yield) and tricin/*p*-coumaryl alcohol (90% overall yield) were carried out via the same method. The oxidation reaction using Ag<sub>2</sub>O was carried out according to the method used in a previous study (Quideau and Ralph, 1994a) with only slight modifications. Tricin (15.7 mg, 47.5 μmol) was dissolved in acetone (5 mL), and Ag<sub>2</sub>O (16.5 mg, 71.3 μmol) was added. Coniferyl alcohol (8.6 mg, 47.5 μmol) dissolved in acetone (30 mL) was added dropwise via an addition funnel over 5 h, and the resulting mixture was allowed to stir overnight. The reaction was quenched with 1 M HCl (2 mL). The inorganics were filtered off, and the filtrate was collected for evaporation to obtain the crude product mixture (99% overall yield). The same method was applied to the tricin/sinapyl alcohol (99.5% overall yield) and tricin/*p*-coumaryl alcohol (92.5% overall yield) adducts. Yields of the desired tricin-monolignol coupling products 14a to 14c [T-(4'-O-β)-G/S/H] were 12.6%, 1.2%, and 51.4% under peroxidase-catalyzed oxidation conditions and 6.9%, 15.3%, and 10.2% under Ag<sub>2</sub>O oxidation.

## Supplemental Data

The following supplemental materials are available.

**Supplemental Figure S1.** Biosynthetic pathway for tricin.

**Supplemental Figure S2.** Differences in chemical shifts of free tricin versus etherified tricin in maize stover lignin in DMSO-*d*<sub>6</sub>.

**Supplemental Figure S3.** HMBC spectra of tricin in acetylated model compound 19 [acetylated T-(4'-O-β)-S-(4-O-β)-G] and maize stover acetylated lignin.

**Supplemental Figure S4.** Structures of the products of radical coupling reactions.

**Supplemental Figure S5.** HSQC spectra of the product mixtures from cross coupling of tricin with coniferyl alcohol, sinapyl alcohol, and *p*-coumaryl alcohol, catalyzed by peroxidase/H<sub>2</sub>O<sub>2</sub> and Ag<sub>2</sub>O.

**Supplemental Text S1.** <sup>1</sup>H- and <sup>13</sup>C-NMR data for synthetic compounds.

## ACKNOWLEDGMENTS

We thank Dr. Ali Azarpira (University of Wisconsin, Madison) for performing HPLC analyses and Xuejun Pan (University of Wisconsin, Madison) for performing the acidic acid pretreatment of the corn stover.

Received November 14, 2014; accepted February 6, 2015; published February 9, 2015.

## LITERATURE CITED

- Begum SA, Sahai M, Ray AB (2010) Non-conventional lignans: coumarinolignans, flavonolignans, and stilbenolignans. *Fortschritte der Chemie Organischer Naturstoffe* **93**: 1–70
- Botcher A, Cesarino I, Santos AB, Vicentini R, Mayer JL, Vanholme R, Morreel K, Goeminne G, Moura JC, Nobile PM, et al (2013) Lignification in sugarcane: biochemical characterization, gene discovery, and expression analysis in two genotypes contrasting for lignin content. *Plant Physiol* **163**: 1539–1557
- Bouazziz M, Veitch NC, Grayer RJ, Simmonds MSJ, Damak M (2002) Flavonolignans from *Hyparrhenia hirta*. *Phytochemistry* **60**: 515–520
- Chen F, Tobimatsu Y, Havkin-Frenkel D, Dixon RA, Ralph J (2012a) A polymer of caffeoyl alcohol in plant seeds. *Proc Natl Acad Sci USA* **109**: 1772–1777
- Chen F, Tobimatsu Y, Jackson L, Nakashima J, Ralph J, Dixon RA (2012b) Novel seed coat lignins in the Cactaceae: structure, distribution and implications for the evolution of lignin diversity. *Plant J* **73**: 201–211
- Crestini C, Melone F, Sette M, Saladino R (2011) Milled wood lignin: a linear oligomer. *Biomacromolecules* **12**: 3928–3935
- Dean JFD, Eriksson KE (1992) Biotechnological modification of lignin structure and composition in forest trees. *Holzforschung* **46**: 135–147
- del Río JC, Rencoret J, Prinsen P, Martínez ÁT, Ralph J, Gutiérrez A (2012) Structural characterization of wheat straw lignin as revealed by

- analytical pyrolysis, 2D-NMR, and reductive cleavage methods. *J Agric Food Chem* **60**: 5922–5935
- Dima O, Morreel K, Vanholme B, Kim H, Ralph J, Boerjan W** (February 19, 2015) Small glycosylated lignin oligomers are stored in Arabidopsis leaf vacuoles. *Plant Cell* <http://dx.doi.org/10.1105/tpc.114.134643>
- Duarte-Almeida JM, Negri G, Salatino A, de Carvalho JE, Lajolo FM** (2007) Antiproliferative and antioxidant activities of a triclin acylated glycoside from sugarcane (*Saccharum officinarum*) juice. *Phytochemistry* **68**: 1165–1171
- Freudenberg K** (1956) Beiträge zur Erforschung des Lignins. *Angew Chem* **68**: 508–512
- Grabber JH** (2005) How do lignin composition, structure, and cross-linking affect degradability? A review of cell wall model studies. *Crop Sci* **45**: 820–831
- Hatfield R, Ralph J, Grabber JH** (2008) A potential role for sinapyl p-coumarate as a radical transfer mechanism in grass lignin formation. *Planta* **228**: 919–928
- Ju Y, Sacalis JN, Still CC** (1998) Bioactive flavonoids from endophyte-infected blue grass (*Poa ampla*). *J Agric Food Chem* **46**: 3785–3788
- Khanam Z, Adam F, Singh O, Ahmad J** (2011) A novel acylated flavonoid glycoside from the wood of cultivated *Acacia nilotica* (L.) Willd. ex. Delile. *Bioresources* **6**: 2932–2940
- Koes RE, Quattrocchio F, Mol JNM** (1994) The flavonoid biosynthetic pathway in plants: function and evolution. *BioEssays* **16**: 123–132
- Kratzl K, Kisser W, Gatzl J, Silbernagel H** (1959) Der  $\beta$ -Gajacyläther des Guajacylglycerins, sine Umwandlung in Coniferylaldehyd und verschiedene andere Arylpropanderivate. *Monatsh Chem* **90**: 771–782
- Lin AS, Nakagawa-Goto K, Chang FR, Yu D, Morris-Natschke SL, Wu CC, Chen SL, Wu YC, Lee KH** (2007) First total synthesis of protoapigenone and its analogues as potent cytotoxic agents. *J Med Chem* **50**: 3921–3927
- Lodge EP, Heathcock CH** (1987) Steric effects, as well as  $\sigma^*$ -orbital energies, are important in diastereoface differentiation in addition to chiral aldehydes. *J Am Chem Soc* **109**: 3353–3361
- Lu F, Ralph J** (2008) Novel tetrahydrofuran structures derived from  $\beta$ - $\beta$ -coupling reactions involving sinapyl acetate in Kenaf lignins. *Org Biomol Chem* **6**: 3681–3694
- Moheb A, Grondin M, Ibrahim RK, Roy R, Sarhan F** (2013) Winter wheat hull (husk) is a valuable source for triclin, a potential selective cytotoxic agent. *Food Chem* **138**: 931–937
- Nakajima Y, Yun YS, Kunugi A** (2003) Six new flavonolignans from *Sasa veitchii* (Carr.) Rehd. *Tetrahedron* **59**: 8011–8015
- Nakano H, Kawada N, Yoshida M, Ono H, Iwaura R, Tonooka T** (2011) Isolation and identification of flavonoids accumulated in proanthocyanidin-free barley. *J Agric Food Chem* **59**: 9581–9587
- Niculaes C, Morreel K, Kim H, Lu F, McKee LS, Ivens B, Haustraete J, Vanholme B, Rycke RD, Hertzberg M, et al** (2014) Phenylcoumaran benzylic ether reductase prevents accumulation of compounds formed under oxidative conditions in poplar xylem. *Plant Cell* **26**: 3775–3791
- Ogo Y, Ozawa K, Ishimaru T, Murayama T, Takaiwa F** (2013) Transgenic rice seed synthesizing diverse flavonoids at high levels: a new platform for flavonoid production with associated health benefits. *Plant Biotechnol J* **11**: 734–746
- Paddon-Row MN, Rondan NG, Houk KN** (1982) Staggered models for asymmetric induction: attack trajectories and conformations of allylic bonds from ab initio transition structures of addition reactions. *J Am Chem Soc* **104**: 7162–7166
- Pan X, Sano Y** (2005) Fractionation of wheat straw by atmospheric acetic acid process. *Bioresour Technol* **96**: 1256–1263
- Pan XJ, Sano Y** (1999) Acetic acid pulping of wheat straw under atmospheric pressure. *J Wood Sci* **45**: 319–325
- Parthasarathy MR, Ranganathan KR, Sharma DK** (1979)  $^{13}\text{C}$  NMR of flavonolignans from *Hydnocarpus wightiana*. *Phytochemistry* **18**: 506–508
- Quideau S, Ralph J** (1992) Facile large-scale synthesis of coniferyl, sinapyl, and p-coumaryl alcohol. *J Agric Food Chem* **40**: 1108–1110
- Quideau S, Ralph J** (1994a) 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
- Quideau S, Ralph J** (1994b) A biomimetic route to lignin model compounds via silver (I) oxide oxidation. 2. NMR characterization of non-cyclic benzyl aryl ether trimers and tetramers. *Holzforschung* **48**: 124–132
- Ralph J** (1993)  $^1\text{H}$  NMR of acetylated  $\beta$ -ether/ $\beta$ -ether lignin model trimers. *Magn Reson Chem* **31**: 357–363
- Ralph J** (2010) Hydroxycinnamates in lignification. *Phytochem Rev* **9**: 65–83
- Ralph J, Brunow G, Harris PJ, Dixon RA, Schatz PF, Boerjan W** (2008) Lignification: are lignins biosynthesized via simple combinatorial chemistry or via proteinaceous control and template replication? *In* F Daayf, A El Hadrami, L Adam, GM Ballance, eds, *Recent Advances in Polyphenol Research*, Vol 1. Wiley-Blackwell Publishing, Oxford, UK, pp 36–66
- Ralph J, Grabber JH, Hatfield RD** (1995) Lignin-ferulate crosslinks in grasses: active incorporation of ferulate polysaccharide esters into rye-grass lignins. *Carbohydr Res* **275**: 167–178
- Ralph J, Helm RF** (1991) Rapid proton NMR method for determination of *threo:erythro* ratios in lignin model compounds and examination of reduction stereochemistry. *J Agric Food Chem* **39**: 705–709
- Ralph J, Helm RF, Quideau S** (1992) Lignin-feruloyl ester cross-links in grasses. Part 2. Model compound syntheses. *J Chem Soc Perkin Trans 1* 2971–2980
- Ralph J, Lundquist K, Brunow G, Lu F, Kim H, Schatz PF, Marita JM, Hatfield RD, Ralph SA, Christensen JH, et al** (2004) Lignins: natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. *Phytochem Rev* **3**: 29–60
- Ralph J, Wilkins AL** (1985) Rapid NMR method for determination of *threo:erythro* ratios in lignin model compounds. *Holzforschung* **39**: 341–344
- Ralph SA, Landucci LL, Ralph J** (2005) NMR database of lignin and cell wall model compounds. <http://ars.usda.gov/Services/docs.htm?docid=10491> (September 2012)
- Rencoret J, Ralph J, Marques G, Gutiérrez A, Martínez Á, del Río JC** (2013) Structural characterization of lignin isolated from coconut (*Cocos nucifera*) coir fibers. *J Agric Food Chem* **61**: 2434–2445
- Sarkanen KV, Ludwig CH** (1971) Lignins: Occurrence, Formation, Structure and Reactions. Wiley-Interscience, New York
- Shanker CG, Mallaiah BV, Srimannarayana G** (1983) Dehydrogenation of chromanones and flavanones by 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ): a facile method for the synthesis of chromones and flavones. *Synthesis-Stuttgart* **15**: 310–311
- Singh OV, Muthukrishnan M, Raj G** (2005) Manganese(III) acetate mediated oxidation of flavanones: a facile synthesis of flavones. *Synth Commun* **35**: 2723–2728
- St. Denis JD, Gordon JS, Carroll VM, Priefer R** (2010) Novel synthesis of the isoflavone genistein. *Synthesis* **42**: 1590–1592
- Takahama U, Oniki T, Shimokawa H** (1996) A possible mechanism for the oxidation of sinapyl alcohol by peroxidase-dependent reactions in the apoplast: enhancement of the oxidation by hydroxycinnamic acids and components of the apoplast. *Plant Cell Physiol* **37**: 499–504
- Tschaplinski TJ, Standaert RF, Engle NL, Martin MZ, Sangha AK, Parks JM, Smith JC, Samuel R, Jiang N, Pu YQ, et al** (2012) Down-regulation of the caffeic acid O-methyltransferase gene in switchgrass reveals a novel monolignol analog. *Biotechnology for Biofuels* **5**: 71
- Vanholme R, Demedts B, Morreel K, Ralph J, Boerjan W** (2010) Lignin biosynthesis and structure. *Plant Physiol* **153**: 895–905
- Van Hoyweghen L, Karalic I, Van Calenberg S, Deforce D, Heyerick A** (2010) Antioxidant flavone glycosides from the leaves of *Fargesia robusta*. *J Nat Prod* **73**: 1573–1577
- Wenzig E, Kunert O, Ferreira D, Schmid M, Schühly W, Bauer R, Hiermann A** (2005) Flavonolignans from *Avena sativa*. *J Nat Prod* **68**: 289–292
- Winkel-Shirley B** (2001) Flavonoid biosynthesis: a colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol* **126**: 485–493
- Yelle DJ, Kaparaju P, Hunt CG, Hirth K, Kim H, Ralph J, Felby C** (2013) Two-dimensional NMR evidence for cleavage of lignin and xylan substituents in wheat straw through hydrothermal pretreatment and enzymatic hydrolysis. *Bioenergy Research* **6**: 211–221
- You TT, Mao JZ, Yuan TQ, Wen JL, Xu F** (2013) Structural elucidation of the lignins from stems and foliage of *Arundo donax* Linn. *J Agric Food Chem* **61**: 5361–5370
- Zanarotti A** (1985) Synthesis and reactivity of vinyl quinone methides. *J Org Chem* **50**: 941–945
- Zeng J, Helms GL, Gao X, Chen S** (2013) Quantification of wheat straw lignin structure by comprehensive NMR analysis. *J Agric Food Chem* **61**: 10848–10857
- Zhang A, Lu F, Sun R, Ralph J** (2009) Ferulate-coniferyl alcohol cross-coupled products formed by radical coupling reactions. *Planta* **229**: 1099–1108
- Zhao Q, Tobimatsu Y, Zhou R, Pattathil S, Gallego-Giraldo L, Fu C, Jackson LA, Hahn MG, Kim H, Chen F, et al** (2013) Loss of function of *Cinnamyl Alcohol Dehydrogenase 1* leads to unconventional lignin and a temperature-sensitive growth defect in *Medicago truncatula*. *Proc Natl Acad Sci USA* **110**: 13660–13665

# Supporting Information

## Tricin, A Flavonoid Monomer in Monocot Lignification

Wu Lan,<sup>1,2</sup> Fachuang Lu,<sup>1,3,\*</sup> Matthew Regner,<sup>1,3</sup> Yimin Zhu,<sup>1</sup> Jorge Rencoret,<sup>1</sup> Sally A. Ralph,<sup>4</sup> Uzma I. Zakai,<sup>1</sup> Kris Morreel,<sup>5,6</sup> Wout Boerjan,<sup>5,6</sup> and John Ralph.<sup>1,2,3</sup>

<sup>1</sup>DOE Great Lakes Bioenergy Research Center, Wisconsin Energy Institute, University of Wisconsin, Madison, WI, USA

<sup>2</sup>Department of Biological System Engineering, University of Wisconsin, Madison, WI, USA

<sup>3</sup>Department of Biochemistry, University of Wisconsin, Madison, WI, USA

<sup>4</sup>US Forest Service, Forest Products Laboratory, Madison, WI, USA

<sup>5</sup>Department of Plant Systems Biology, VIB, B-9052 Gent, Belgium

<sup>6</sup>Department of Plant Biotechnology and Bioinformatics, Ghent University, B-9052 Gent, Belgium

### Present addresses

Yimin Zhu: Department of Chemistry, Pennsylvania State University, Altoona College, Altoona, PA, USA

Jorge Rencoret: Instituto de Recursos Naturales y Agrobiología de Sevilla, Consejo Superior de Investigaciones Científicas, Seville, Spain

Uzma I. Zakai: Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, Ames, IA, USA

### Corresponding Authors

Prof. John Ralph  
Email: jralph@wisc.edu

Dr. Fachuang Lu  
Email: fachuanglu@wisc.edu

### Table of Contents

- Figure S1.** Biosynthetic pathway for tricrin.
- Figure S2.** Differences in chemical shift of free tricrin vs etherified tricrin in corn stover lignin in DMSO-*d*<sub>6</sub>.
- Figure S3.** HMBC spectra of a) tricrin in acetylated model compound **19** [acetylated T-(4'-O-β)-S-(4-O-β)-G], and b) corn stover lignin.
- Figure S4.** Structures of the products of radical coupling reactions.
- Figure S5.** HSQC spectra of the product mixtures from cross-coupling of tricrin with a) and d) coniferyl alcohol, b) and e) sinapyl alcohol, and c) and f) *p*-coumaryl alcohol, catalyzed by peroxidase/H<sub>2</sub>O<sub>2</sub> (a, b, and c) and Ag<sub>2</sub>O (d, e, and f).
- Data S1.** <sup>1</sup>H and <sup>13</sup>C NMR data for synthetic compounds.

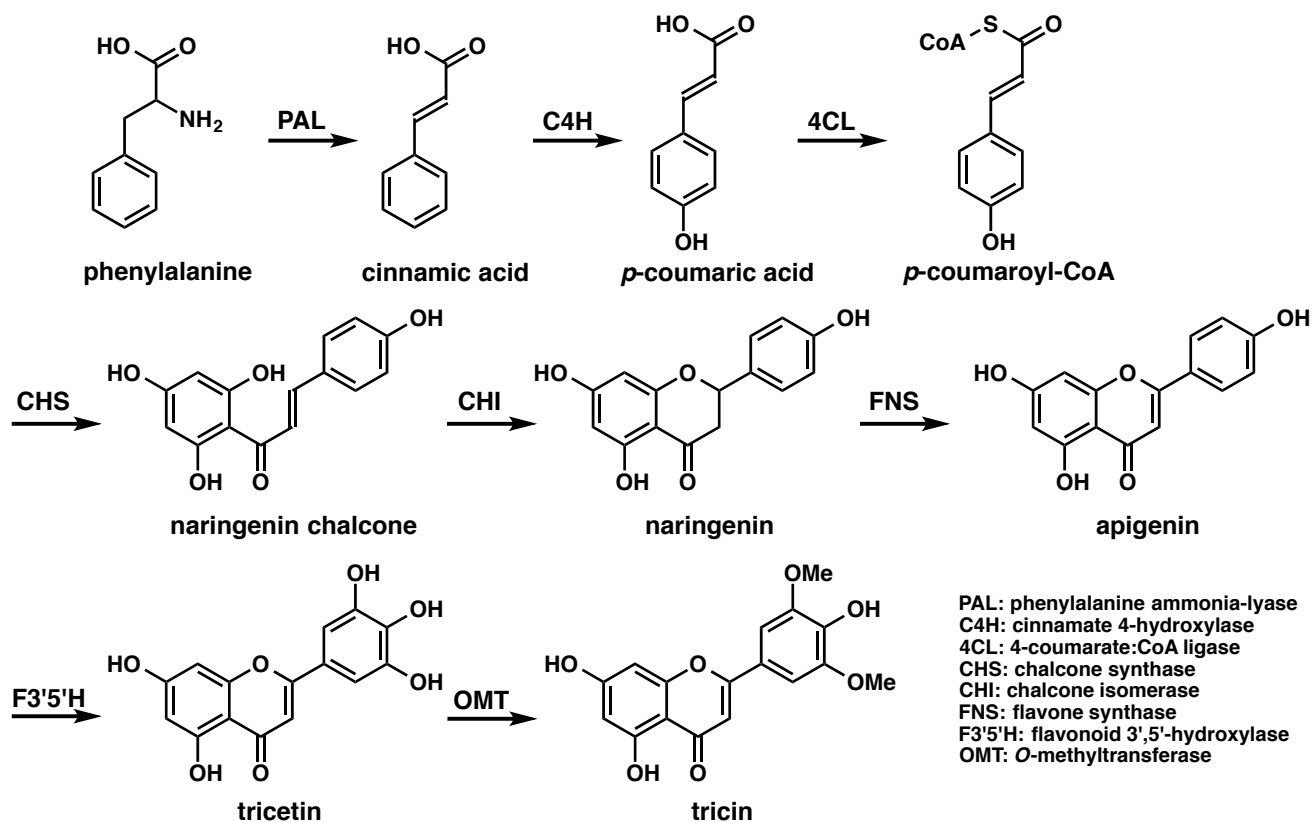


Figure S1. Biosynthetic pathway for tricetin (Morreel *et al.*, 2006; Ogo *et al.*, 2013; Lam *et al.*, 2014).

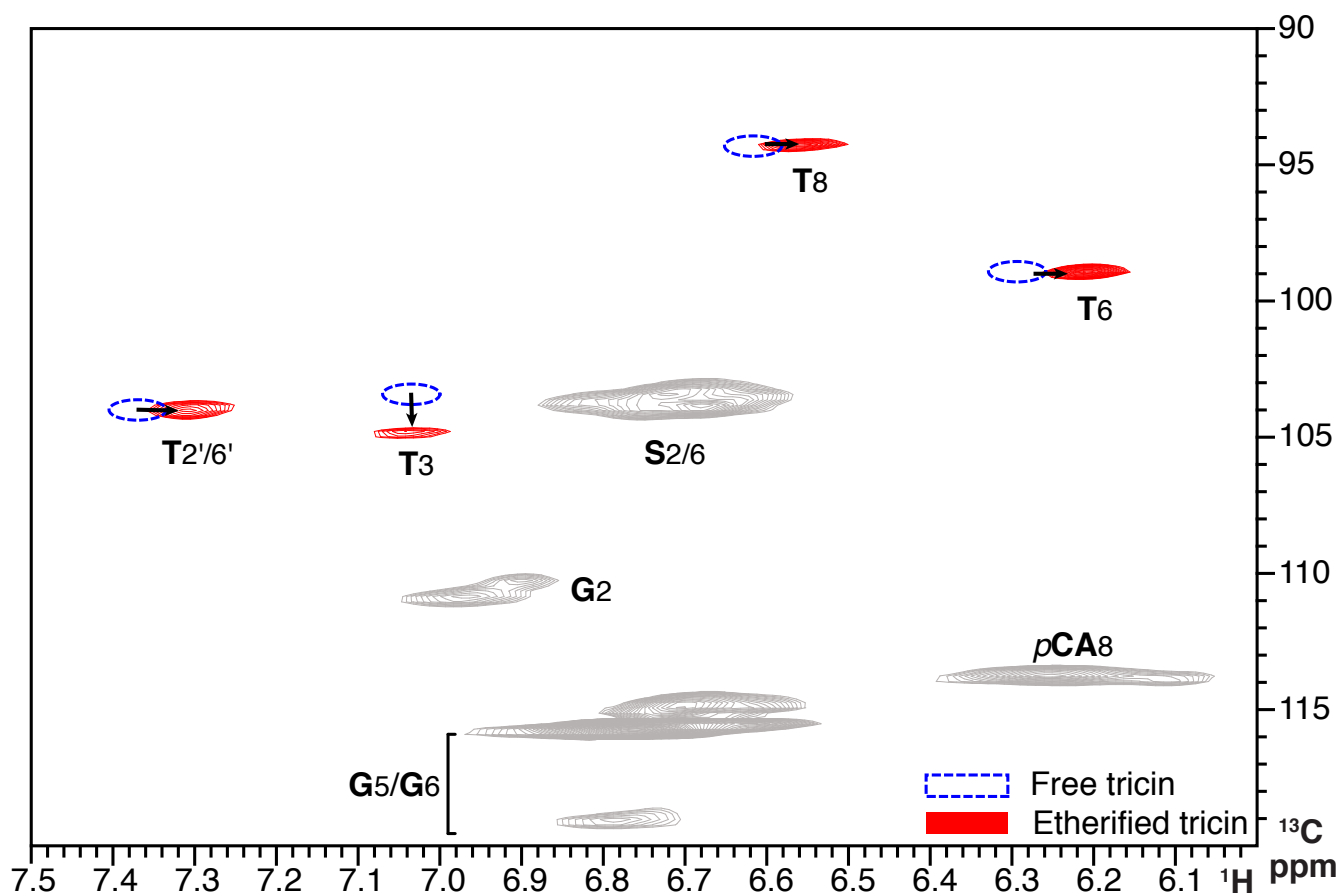
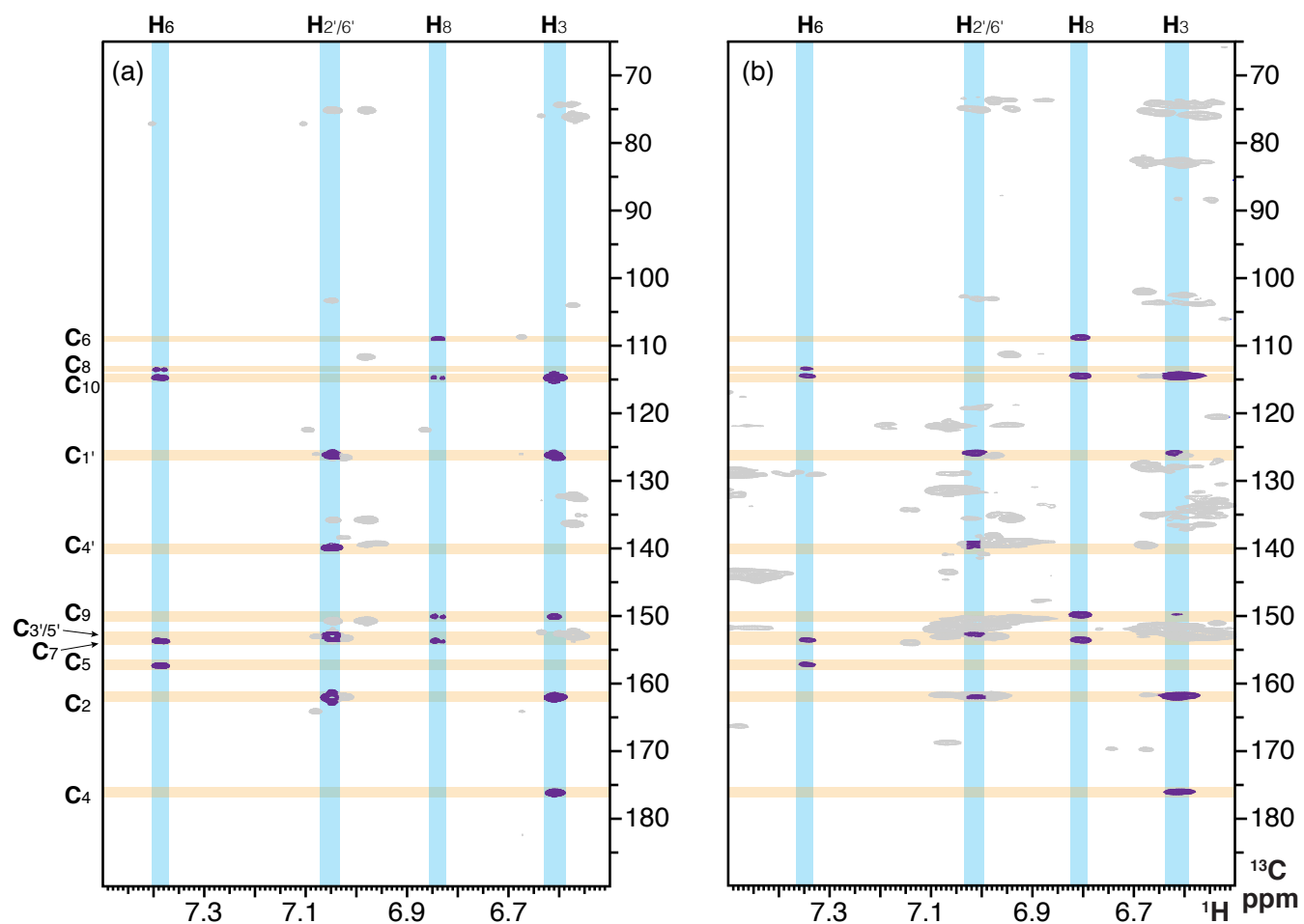
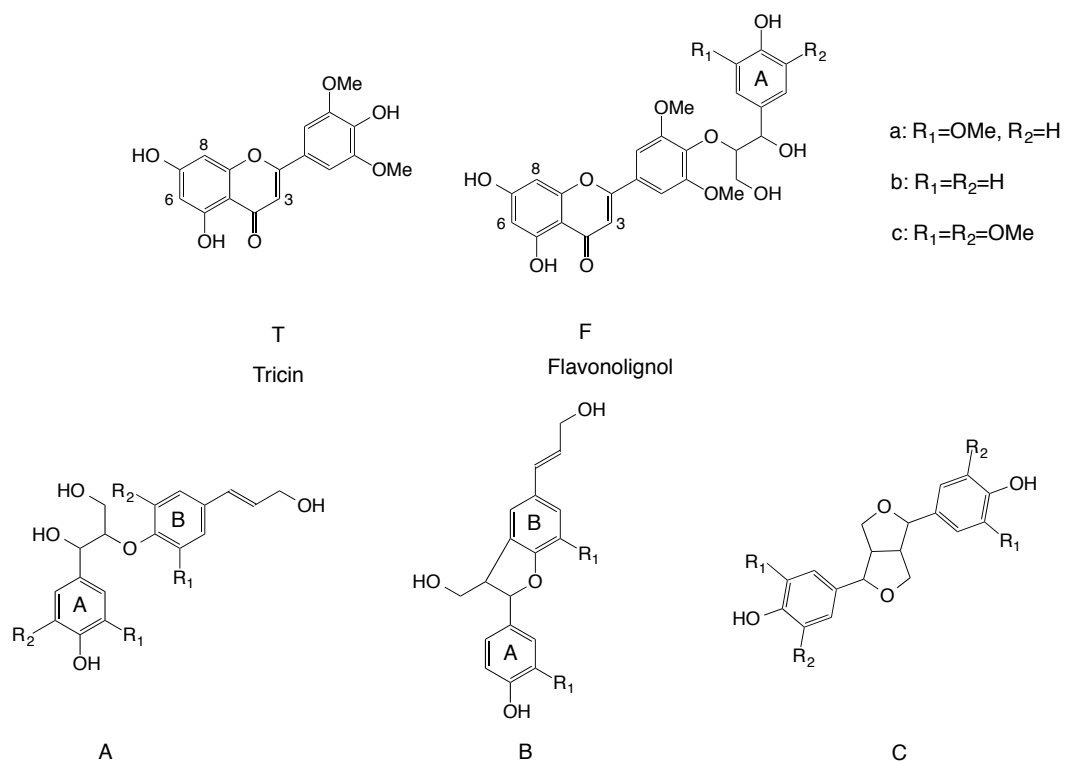


Figure S2. Differences in chemical shift of free tricetin vs etherified tricetin in corn stover lignin in DMSO-*d*<sub>6</sub>.

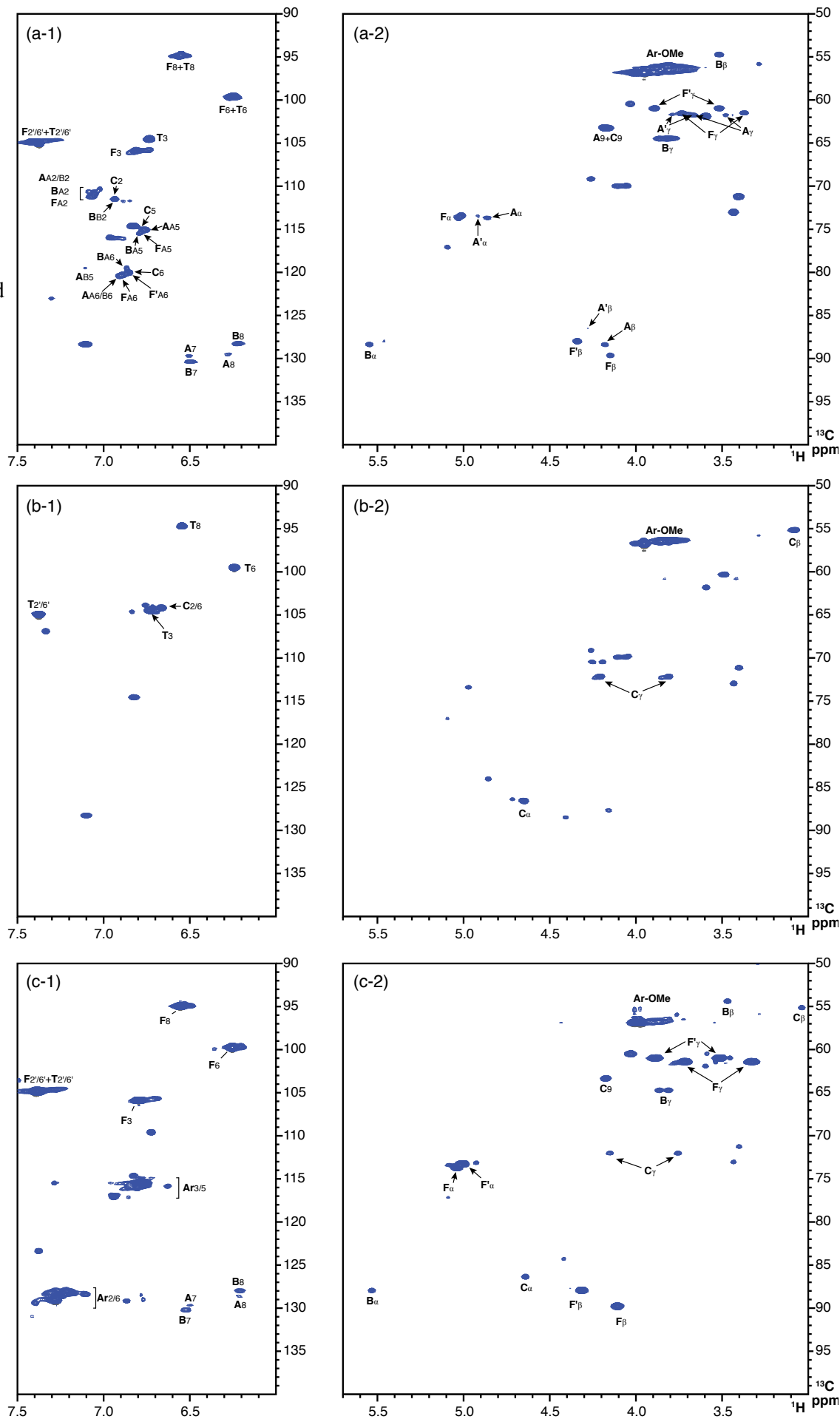


**Figure S3.** HMBC spectra of a) tricrin in acetylated model compound **19** [acetylated T-(4'-O- $\beta$ )-S-(4-O- $\beta$ )-G], and b) corn stover lignin.

**Figure S4.** Structures of the products of radical coupling reactions.

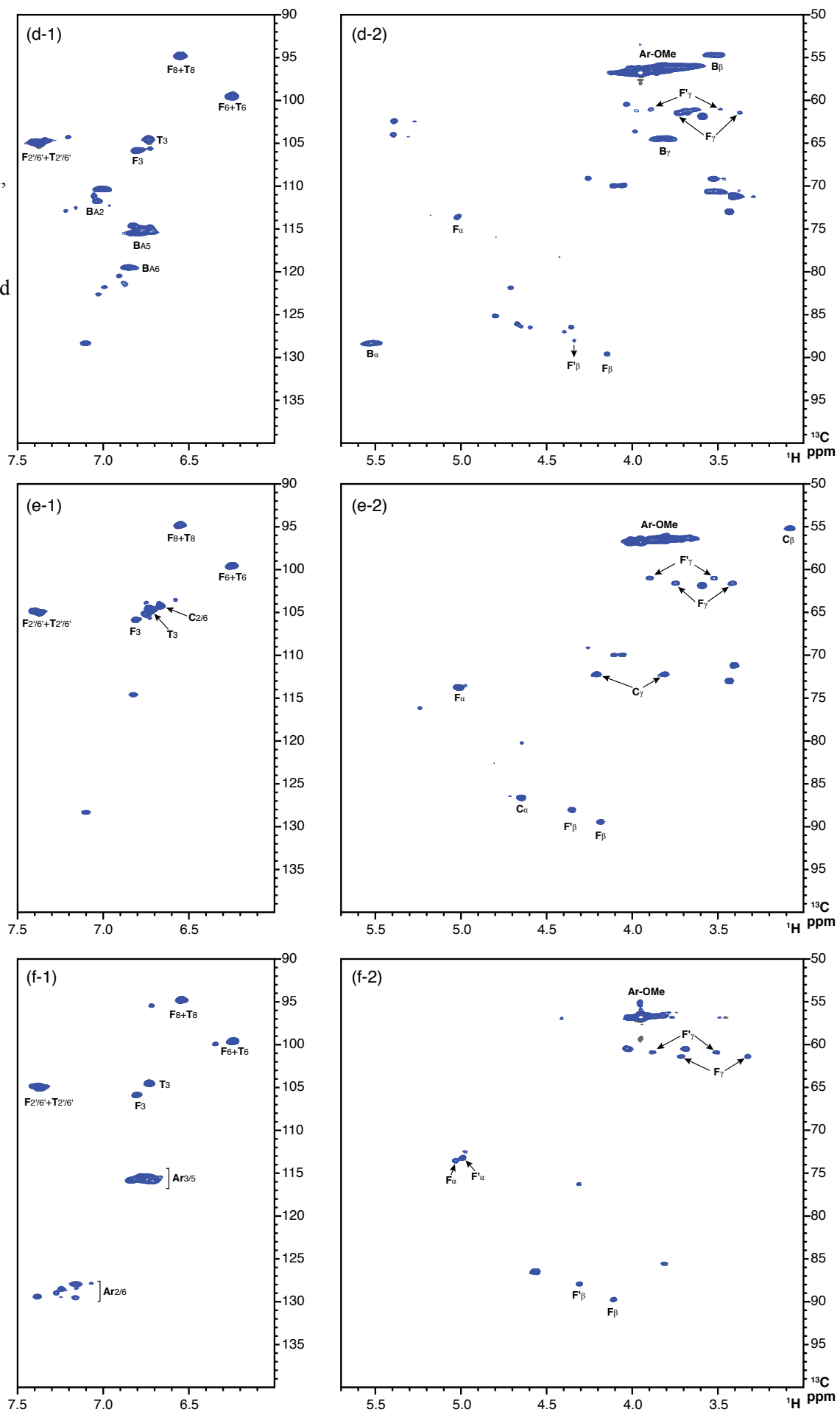


**Figure S5a.**  
 HSQC spectra of the product mixtures from cross-coupling of triclin with a) coniferyl alcohol, b) sinapyl alcohol, and c) *p*-coumaryl alcohol, catalyzed by peroxidase/ $H_2O_2$ .





**Figure S5b.**  
 HSQC spectra of the product mixtures from cross-coupling of triclin with d) coniferyl alcohol, e) sinapyl alcohol, and f) *p*-coumaryl alcohol, catalyzed by Ag<sub>2</sub>O.



**Data S1.** <sup>1</sup>H and <sup>13</sup>C NMR data for synthetic compounds.

All NMR data are for spectra of compounds in acetone-d<sub>6</sub> run on a 700 MHz (176 MHz <sup>13</sup>C) spectrometer unless otherwise stated.

**Compound 3**

<sup>1</sup>H NMR δ 6.28 (1H, d, *J*=2.4 Hz, H3), 6.18 (1H, d, *J*=2.3 Hz, H5), 5.35 (2H, s, 4-OMOM), 5.23 (2H, s, 2-OMOM), 3.51 (3H, s, 4-OMOM), 3.43 (3H, s, 2-OMOM), 2.63 (3H, s, COCH<sub>3</sub>)

<sup>13</sup>C NMR δ 204.17 (COCH<sub>3</sub>), 167.63 (C6), 164.51 (C4), 161.50 (C2), 107.39 (C1), 97.37 (C5), 95.42 (4-OMOM), 94.89 (C3), 94.80 (2-OMOM), 56.91 (4-OMOM), 56.52 (2-OMOM), 33.17 (COCH<sub>3</sub>)

**Compound 4**

<sup>1</sup>H NMR δ 9.89 (1H, s, COH), δ 7.50 (2H, d, *J*=7.4 Hz, H9,13), 7.35 (2H, t, *J*=7.4 Hz, H10, 12), 7.29 (1H, t, *J*=7.4 Hz, OMOM), 7.24 (2H, s, H2,6), 5.09 (2H, s, H7), 3.92 (6H, s, H-OMe); δ<sub>C</sub> 191.73 (COH), 155.02 (C3,5), 143.14 (C4), 138.82(C8), 133.24 (C1), 128.93 (C9,10,12,13), 128.62 (OMOM), 107.38 (C2,6), 75.14 (C7), 56.58 (OMe)

**Compound 5**

<sup>1</sup>H NMR δ 7.97 (1H, d, *J*=15.5 Hz, H3), 7.72 (1H, d, *J*=15.5 Hz, H2), 7.51 (2H, d, *J*=7.5 Hz, H9',13'), 7.35 (2H, t, *J*=7.5 Hz, H10',12'), 7.29 (1H, t, *J*=7.4 Hz, H11'), 7.05 (2H, s, H2',6'), 6.31 (1H, d, *J*=2.1 Hz, H6), 6.25 (1H, d, *J*=2.3 Hz, H8), 5.42 (2H, s, 7-OMOM), 5.25 (2H, s, 5-OMOM), 5.03 (2H, s, H7'), 3.90 (6H, s, OMe), 3.57 (3H, s, 5-OMOM), 3.45 (3H, s, 7-OMOM); δ<sub>C</sub> 193.58 (C4), 168.02 (C9), 164.60 (C7), 161.02 (C5), 154.81 (C3',5'), 143.61 (C2), 140.13 (C4'), 139.13 (C8'), 131.86 (C1'), 128.90 (C9',10',12',13'), 128.52 (C11'), 127.40 (C3), 108.00 (C10), 106.64 (C2',6'), 97.69 (C8), 96.21 (7-OMOM), 95.48 (C6), 94.87 (5-OMOM), 75.18 (C7'), 57.27 (5-OMOM), 56.56 (7-OMOM), 56.47 (OMe)

**Compound 6**

<sup>1</sup>H NMR δ 7.52 (2H, d, *J*=7.6 Hz, H9',13'), 7.41 (2H, s, H2',6'), 7.36 (2H, t, *J*=7.5 Hz, H10',12'), 7.30 (1H, t, *J*=7.4 Hz, H11'), 6.86 (1H, s, H3), 6.79 (1H, d, *J*=2.2 Hz, H8), 6.42 (1H, d, *J*=2.2 Hz, H6), 5.32 (2H, s, OMOM), 5.09 (2H, s, H7'), 3.97 (6H, s, OMe), 3.47 (3H, s, OMOM); δ<sub>C</sub> 183.25(C4), 164.80 (C2), 163.99 (C7), 162.85 (C5), 158.38 (C9), 154.88 (C3', 5'), 141.21 (C4'), 138.86 (C8'), 128.88 (C9',13'), 128.86 (C10',12'), 128.53 (C11'), 127.26 (C1'), 106.54 (C10), 105.92 (C3), 104.88 (C2',6'), 100.26 (C8), 95.25 (C6), 94.92 (OMOM), 75.13 (C7'), 56.73 (OMe), 56.48 (OMOM)

**Compound 7**

<sup>1</sup>H NMR δ 7.41 (2H, s, H2',6'), 6.80 (1H, s, H3), 6.78 (1H, d, *J*=2.1 Hz, H8), 6.40 (1H, d, *J*=2.2 Hz, H6), 5.31 (2H, s, OMOM), 3.96 (6H, s, OMe), 3.47 (3H, s, OMOM); δ<sub>C</sub> 183.23 (C4), 165.42 (C2), 163.90 (C7), 162.92 (C5), 158.41 (C9), 149.12 (C3',5'), 141.19 (C4'), 121.99 (C1'), 106.52 (C10), 105.19 (C2',6'), 104.73 (C3), 100.23 (C6), 95.24 (C8), 94.98 (OMOM), 56.90 (OMe), 56.53 (OMOM)

**Compound 8**

δ<sub>H</sub> 7.39 (2H, s, H2',6'), 6.74 (1H, s, H3), 6.56 (1H, d, *J*=2.1 Hz, H8), 6.25 (1H, d, *J*=2.2 Hz, H6), 3.96 (6H, s, OMe); δ<sub>C</sub> 183.09 (C4), 165.04 (C2), 163.33 (C7), 163.02 (C5), 158.77 (C9), 149.09 (C3',5'), 140.76 (C4'), 122.30 (C1'), 105.34 (C10), 105.12 (C2',6'), 104.64 (C3), 99.62 (C6), 94.86 (C8), 56.89 (OMe)

**Compound 10a**

δ<sub>H</sub> 7.72-7.68 (2H, m, H2,6), 7.24 (1H, d, *J*=7.8 Hz, H5), 4.79 (2H, s, Hβ), 3.91 (3H, s, OMe), 2.27 (3H, s, OAc); δ<sub>C</sub> 191.02 (Cα), 168.61 (OAc), 152.64 (C4), 145.37 (C3), 133.84 (C1), 124.08 (C5), 123.01 (C6), 113.13 (C2), 56.49 (C7), 33.02 (Cβ), 20.47 (OAc)

**Compound 10b**

$\delta_{\text{H}}$  7.39 (2H, s, H2,6), 4.83 (2H, s, H $\beta$ ), 3.89 (6H, s, OMe), 2.26 (3H, s, OAc);  $\delta_{\text{C}}$  190.98 (C $\alpha$ ), 168.09 (OAc), 153.39 (C3,5), 134.12 (C4), 133.06 (C1), 106.40 (C2,6), 56.69 (OMe), 33.22 (C $\beta$ ), 20.15 (OAc).

**Compound 10c**

$\delta_{\text{H}}$  8.10 (2H, d,  $J=8.8$  Hz, H2,6), 7.31 (2H, d,  $J=8.8$  Hz, H3,5), 4.77 (2H, s, H $\beta$ ), 2.29 (3H, s, OAc);  $\delta_{\text{C}}$  190.95 (C $\alpha$ ), 169.25 (OAc), 155.97 (C4), 132.55 (C1), 131.29 (C2,6), 123.10 (C3,5), 32.94 (C $\beta$ ), 20.99 (OAc)

**Compound 11a**

$\delta_{\text{H}}$  7.76 – 7.74 (2H, m, HA2,6), 7.41 (2H, s, H2',6'), 7.23 (1H, d,  $J=8.7$  Hz, HA5), 6.87 (1H, s, H3), 6.78 (1H, d,  $J=2.1$  Hz, H8), 6.42 (1H, d,  $J=2.1$  Hz, H6), 5.33 (2H, s, H $\beta$ ), 5.32 (2H, s, OMOM), 3.93 (6H, s, OMe), 3.91 (3H, s, A-OMe), 3.47 (3H, s, OMOM), 2.27 (3H, s, OAc);  $\delta_{\text{C}}$  194.07 (C $\alpha$ ), 183.23 (C4), 168.60 (OAc), 164.67 (C2), 163.99 (C7), 162.82 (C5), 158.38 (C9), 154.09 (C3',5'), 152.43 (CA3), 144.97 (CA1), 140.63 (C4'), 134.69 (CA4), 127.29 (C1'), 123.84 (CA5), 122.34 (CA6), 112.74 (CA2), 106.54 (C10), 105.99 (C3), 104.99 (C2',6'), 100.27 (C6), 95.27 (C8), 94.92 (OMOM), 75.52 (C $\beta$ ), 56.77 (OMe), 56.49 (OMOM), 56.38 (A-OMe), 20.24 (OAc)

**Compound 11b**

$\delta_{\text{H}}$  7.44 (2H, s, HB2,6), 7.42 (2H, s, HA2,6), 6.87 (1H, s, H3), 6.78 (1H, d,  $J=2.2$  Hz, H8), 6.42 (1H, d,  $J=2.2$  Hz, H6), 5.35 (2H, s, H $\beta$ ), 5.32 (2H, s, OMOM), 3.93 (6H, s, A-OMe), 3.89 (6H, s, B-OMe), 3.47 (3H, s, OMOM), 2.26 (3H, s, OAc);  $\delta_{\text{C}}$  194.03 (C $\alpha$ ), 183.25 (C4), 168.15 (OAc), 164.68 (C2), 164.00 (C7), 162.63 (C5), 158.38 (C9), 154.14 (C3',5'), 153.30 (CA3,5), 140.57 (C4'), 133.88 (CA4), 133.84 (CA1), 127.33 (C1'), 106.55 (C10), 105.99 (C3), 105.89 (CA2,6), 104.98 (C2',6'), 100.28 (C6), 95.27 (C8), 94.92 (OMOM), 75.49 (C $\beta$ ), 56.80 (OMe), 56.64 (A-OMe), 56.49 (OMOM), 20.18 (OAc)

**Compound 11c**

$\delta_{\text{H}}$  8.16 (2H, d,  $J=8.7$  Hz, HA2,6), 7.42 (2H, s, H2',6'), 7.30 (2H, d,  $J=9.0$  Hz, HA3,5), 6.87 (1H, s, H3), 6.79 (1H, d,  $J=2.3$  Hz, H8), 6.42 (1H, d,  $J=2.2$  Hz, H6), 5.32 (2H, s, OMOM), 5.31 (2H, s, H $\beta$ ), 3.92 (6H, s, OMe), 3.47 (3H, s, OMOM), 2.29 (3H, s, OAc);  $\delta_{\text{C}}$  194.21 (C $\alpha$ ), 183.33 (C4), 169.33 (OAc), 164.78 (C2), 164.08 (C7), 162.93 (C5), 158.47 (C9), 155.65 (CA4), 154.16 (C3',5'), 140.74 (C4'), 133.55 (CA1), 130.88 (CA2,6), 127.40 (C1'), 122.91 (CA3,5), 106.55 (C10), 106.08 (C3), 105.08 (C2',6'), 100.36 (C6), 95.35 (C8), 95.00 (OMOM), 75.70 (C $\beta$ ), 56.83 (OMe), 56.56 (OMOM), 21.00 (OAc)

**Compound 12a**

$\delta_{\text{H}}$  7.80 – 7.77 (2H, m, HA2,6), 7.40 (2H, s, H2',6'), 7.21 (1H, d,  $J=8.7$  Hz, HA5), 6.86 (1H, s, H3), 6.76 (1H, d,  $J=2.2$  Hz, H8), 6.41 (1H, d,  $J=2.2$  Hz, H6), 5.45 (1H, t,  $J=5.2$  Hz, H $\beta$ ), 5.30 (2H, s, OMOM), 3.99 – 3.95 (2H, m, H $\gamma$ ), 3.89 (3H, s, A-OMe), 3.86 (6H, s, OMe), 3.46 (3H, s, OMOM), 2.27 (3H, s, OAc);  $\delta_{\text{C}}$  195.77 (C $\alpha$ ), 183.22 (C4), 168.60 (OAc), 164.55 (C2), 163.99 (C7), 162.83 (C5), 158.35 (C9), 153.69 (C3',5'), 152.30 (CA3), 144.77 (CA1), 140.40 (C4'), 135.56 (CA4), 127.32 (C1'), 123.68 (CA5), 122.78 (CA6), 113.19 (CA2), 106.53 (C10), 106.04 (C3), 104.86 (C2',6'), 100.28 (C6), 95.25 (C8), 94.91 (OMOM), 86.52 (C $\beta$ ), 63.70 (C $\gamma$ ), 56.69 (OMe), 56.48 (OMOM), 56.35 (A-OMe), 20.42 (OAc)

**Compound 12b**

$\delta_{\text{H}}$  7.48 (2H, s, HA2,6), 7.42 (2H, s, H2',6'), 6.87 (1H, s, H3), 6.78 (1H, d, H8  $J=2.3$  Hz), 6.42 (1H, d,  $J=2.2$  Hz, H6), 5.52 (1H, t,  $J=5.3$  Hz, H $\beta$ ), 5.31 (2H, s, OMOM), 3.99 – 3.96 (2H, m, H $\gamma$ ), 3.88 (6H, s, A-OMe), 3.87 (6H, s, OMe), 3.47 (3H, s, OMOM), 2.26 (3H, s, OAc);  $\delta_{\text{C}}$  195.68 (C $\alpha$ ), 183.24 (C4), 168.16 (OAc), 164.58 (C2), 164.01 (C7), 162.84 (C5), 158.37 (C9), 153.75 (C3',5'), 153.17 (CA3,5), 140.35 (C4'), 134.82 (CA4), 133.65 (CA1), 127.33 (C1'), 106.55 (C10), 106.35 (CA2,6), 106.05 (C3), 104.87 (C2',6'), 100.29 (C6), 95.27 (C8), 94.91 (OMOM), 85.92 (C $\beta$ ), 63.61 (C $\gamma$ ), 56.74 (OMe), 56.63 (A-OMe), 56.48 (OMOM), 20.18 (OAc)

### Compound 12c

$\delta_{\text{H}}$  8.19 (2H, d,  $J=8.7$  Hz, HA2,6), 7.41 (2H, s, H2',6'), 7.29 (2H, d,  $J=8.8$  Hz, HA3,5), 6.87 (1H, s, H3), 6.78 (1H, d,  $J=2.3$  Hz, H8), 6.41 (1H, d,  $J=2.3$  Hz, H6), 5.36 (1H, t,  $J=4.9$  Hz, H $\beta$ ), 5.31 (2H, s, OMOM), 3.98 – 3.95 (2H, m, H $\gamma$ ), 3.86 (6H, s, OMe), 3.46 (3H, s, OMOM), 2.29 (3H, s, OAc);  $\delta_{\text{C}}$  196.10 (C $\alpha$ ), 183.23 (C4), 169.35 (OAc), 164.61 (C2), 164.04 (C7), 162.62 (C5), 158.46 (C9), 155.50 (CA4), 153.74 (C3',5'), 140.53 (C4'), 134.48 (CA1), 131.29 (CA2,6), 127.43 (C1'), 122.74 (CA3,5), 106.53 (C10), 106.20 (C3), 104.97 (CA2,6), 100.31 (C6), 95.41 (C8), 94.99 (OMOM), 87.16 (C $\beta$ ), 63.68 (C $\gamma$ ), 56.74 (OMe), 56.55 (OMOM), 20.82 (OAc)

### Compound 13a

$\delta_{\text{H}}$  7.73 (1H, dd,  $J=8.3, 2.0$  Hz, HA6), 7.67 (1H, d,  $J=2.0$  Hz, HA2), 7.39 (2H, s, H2',6'), 6.93 (1H, d,  $J=8.3$  Hz, HA5), 6.81 (1H, s, H3), 6.56 (1H, d,  $J=2.2$  Hz, H8), 6.26 (1H, d,  $J=2.2$  Hz, H6), 5.47 (1H, t,  $J=5.4$  Hz, H $\beta$ ), 3.91 (5H, d, H $\gamma$ , A-OMe), 3.87 (6H, s, OMe);  $\delta_{\text{C}}$  194.61 (C $\alpha$ ), 183.02 (C4), 165.01 (C2), 164.27 (C7), 162.95 (C5), 158.80 (C9), 153.71 (C3',5'), 152.45 (CA1), 148.32 (CA3), 140.60 (C4'), 129.08 (CA4), 127.31 (C1'), 124.61 (CA6), 115.37 (CA5), 112.32 (CA2), 105.95 (C3), 105.30 (C10), 104.95 (C2',6'), 99.66 (C6), 94.97 (C8), 86.11 (C $\beta$ ), 63.87 (C $\gamma$ ), 56.78 (OMe), 56.28 (A-OMe)

### Compound 13b

$\delta_{\text{H}}$  7.46 (2H, s, HA2,6), 7.40 (2H, s, H2',6'), 6.81 (1H, s, H3), 6.56 (1H, d,  $J=2.1$  Hz, H8), 6.27 (1H, d,  $J=2.1$  Hz, H6), 5.52 (1H, t,  $J=5.4$  Hz, H $\beta$ ), 3.92 (2H, d,  $J=5.5$  Hz, H $\gamma$ ), 3.88 (6H, s, A-OMe), 3.87 (6H, s, OMe);  $\delta_{\text{C}}$  194.61 (C $\alpha$ ), 183.03 (C4), 164.95 (C2), 164.28 (C7), 162.97 (C5), 158.81 (C9), 153.78 (C3',5'), 148.37 (CA3,5), 142.03 (CA4), 140.54 (C4'), 127.59 (CA1), 127.35 (C1'), 107.52 (CA2,6), 105.97 (C3), 105.34 (C10), 104.93 (C2',6'), 99.66 (C6), 94.97 (C8), 85.71 (C $\beta$ ), 63.79 (C $\gamma$ ), 56.81 (OMe), 56.71 (A-OMe)

### Compound 13c

$\delta_{\text{H}}$  8.05 (2H, d,  $J=8.8$  Hz, HA2,6), 7.39 (2H, s, H2',6'), 6.95 (2H, d,  $J=8.8$  Hz, HA3,5), 6.81 (1H, s, H3), 6.56 (1H, d,  $J=2.2$  Hz, H8), 6.26 (1H, d,  $J=2.2$  Hz, H6), 5.40 (1H, dd,  $J=6.1, 4.6$  Hz, H $\beta$ ), 3.91 (2H, d,  $J=6.1$  Hz, H $\gamma$ ), 3.86 (6H, s, OMe);  $\delta_{\text{C}}$  194.73 (C $\alpha$ ), 183.03 (C4), 164.91 (C2), 164.28 (C7), 162.98 (C5), 162.75 (CA4), 158.81 (C9), 153.73 (C3',5'), 140.64 (C4'), 132.15 (CA2,6), 128.92 (CA1), 127.37 (C1'), 115.92 (CA3,5), 105.99 (C3), 105.35 (C10), 104.94 (C2',6'), 99.65 (C6), 94.97 (C8), 86.64 (C $\beta$ ), 63.84 (C $\gamma$ ), 56.75 (OMe)

### Compound 14a [T-(4'-O- $\beta$ )-G]

ESIMS  $m/z$  [M+H]<sup>+</sup> 527; HRESIMS calculated for C<sub>27</sub>H<sub>26</sub>O<sub>11</sub> [M+H]<sup>+</sup> 527.1548, found 527.1552.

**anti-14a:**  $\delta_{\text{H}}$  7.42 (2H, s, H2',6'), 7.06 (1H, d,  $J=1.9$  Hz, HA2), 6.91 (1H, dd,  $J=8.0, 1.9$  Hz, HA6), 6.82 (1H, s, H3), 6.76 (1H, d,  $J=8.1$  Hz, HA5), 6.56 (1H, d,  $J=2.1$  Hz, H8), 6.27 (1H, d,  $J=2.1$  Hz, H6), 5.03 (1H, d,  $J=7.1$  Hz, H $\alpha$ ), 4.15 (1H, dt,  $J=7.1, 3.5$  Hz, H $\beta$ ), 4.03 (6H, s, OMe), 3.81 (3H, s, A-OMe), 3.74 (1H, dd,  $J=12.2, 3.5$  Hz, H $\gamma$ ), 3.38 (1H, dd,  $J=12.2, 3.4$  Hz, H $\gamma$ );  $\delta_{\text{C}}$  183.12 (C4), 165.01 (C2), 164.29 (C7), 163.31 (C5), 158.80 (C9), 154.30 (C3',5'), 147.91 (CA3), 146.79 (CA4), 140.50 (C4'), 133.58 (CA1), 127.64 (C1'), 120.60 (CA6), 115.20 (CA5), 111.29 (CA2), 106.01 (C3), 105.44 (C10), 104.95 (C2',6'), 99.79 (C6), 94.98 (C8), 89.76 (C $\beta$ ), 73.84 (C $\alpha$ ), 61.62 (C $\gamma$ ), 56.97 (OMe), 56.17 (A-OMe)

**syn-14a:**  $\delta_{\text{H}}$  7.41 (2H, s, H2',6'), 7.06 (1H, d,  $J=1.9$  Hz, HA2), 6.86 (1H, dd,  $J=8.3, 1.6$  Hz, HA6), 6.82 (1H, s, H3), 6.78 (1H, d,  $J=8.1$  Hz, HA5), 6.56 (1H, d,  $J=2.1$  Hz, H8), 6.27 (1H, d,  $J=2.1$  Hz, H6), 5.02 (1H, d,  $J=4.8$  Hz, H $\alpha$ ), 4.35 (1H, m, H $\beta$ ), 4.00 (6H, s, OMe), 3.89 (1H, dd,  $J=12.1, 3.9$  Hz, H $\gamma$ ), 3.83 (3H, s, A-OMe), 3.52 (1H, dd,  $J=12.1, 3.6$  Hz, H $\gamma$ );  $\delta_{\text{C}}$  183.05 (C4), 165.01 (C2), 164.36 (C7), 163.31 (C5), 158.80 (C9), 154.66 (C3',5'), 147.98 (CA3), 146.54 (CA4), 140.06 (C4'), 133.73 (CA1), 127.58 (C1'), 120.12 (CA6), 115.13 (CA5), 111.90 (CA2), 105.98 (C3), 105.41 (C10), 104.99 (C2',6'), 99.75 (C6), 95.05 (C8), 89.14 (C $\beta$ ), 73.54 (C $\alpha$ ), 61.10 (C $\gamma$ ), 56.97 (OMe), 56.17 (A-OMe)

**Compound 14b** [T-(4'-O-β)-S]ESIMS  $m/z$  [M+H]<sup>+</sup> 557; HRESIMS calculated for C<sub>28</sub>H<sub>28</sub>O<sub>12</sub> [M+H]<sup>+</sup> 557.1654, found 557.1660.

**anti-14b**: δ<sub>H</sub> 7.42 (2H, s, H2',6'), 6.82 (1H, s, H3), 6.76 (2H, s, HA2,6), 6.57 (1H, d,  $J=2.1$  Hz, H8), 6.27 (1H, d,  $J=2.2$  Hz, H6), 5.01 (1H, d,  $J=6.8$  Hz, Hα), 4.19 (1H, dt,  $J=7.0, 3.6$  Hz, Hβ), 4.03 (6H, s, OMe), 3.79 (6H, s, A-OMe), 3.75 (1H, d,  $J=12.1, 3.6$  Hz, Hγ), 3.42 (1H, dd,  $J=12.1, 3.5$  Hz, Hγ); δ<sub>C</sub> 183.12 (C4), 165.02 (C2), 164.30 (C7), 163.31 (C5), 158.81 (C9), 154.28 (C3',5'), 148.31 (CA3,5), 140.54 (C4'), 136.12 (CA4), 132.56 (CA1), 127.62 (C1'), 106.01 (C3), 105.43 (C10), 105.27 (C2A,6), 104.96 (C2',6'), 99.79 (C6), 94.98 (C8), 89.60 (Cβ), 73.97 (Cα), 61.73 (Cγ), 56.98 (OMe), 56.56 (A-OMe)

**syn-14b**: δ<sub>H</sub> 7.42 (2H, s, H2',6'), 6.82 (1H, s, H3), 6.73 (2H, s, HA2,6), 6.57 (1H, d,  $J=2.1$  Hz, H8), 6.27 (1H, d,  $J=2.2$  Hz, H6), 5.02 (1H, m, Hα), 4.36 (1H, dt,  $J=4.8, 3.4$  Hz, Hβ), 4.01 (6H, s, OMe), 3.89 (1H, m, Hγ), 3.80 (6H, s, A-OMe), 3.52 (1H, dd,  $J=12.1, 3.3$  Hz, Hγ); δ<sub>C</sub> 183.05 (C4), 165.02 (C2), 164.34 (C7), 163.31 (C5), 158.81 (C9), 154.66 (C3',5'), 148.31 (CA3,5), 140.07 (C4'), 136.12 (CA4), 132.56 (CA1), 127.58 (CA1), 106.09 (C3), 105.40 (C10), 104.96 (C2',6'), 104.85 (CA2,6), 99.79 (C6), 94.98 (C8), 88.28 (Cβ), 73.87 (Cα), 61.01 (Cγ), 56.98 (OMe), 56.56 (A-OMe)

**Compound 14c** [T-(4'-O-β)-H]ESIMS  $m/z$  [M+H]<sup>+</sup> 497; HRESIMS calculated for C<sub>26</sub>H<sub>24</sub>O<sub>10</sub> [M+H]<sup>+</sup> 497.1443, found 497.1436.

**anti-14c**: δ<sub>H</sub> (acetone-*d*<sub>6</sub>, 500 MHz) 7.42 (1H, s, H2',6'), 7.29 (2H, d,  $J=8.4$  Hz, HA2,6), 6.83 (1H, s, H3), 6.78 (2H, d,  $J=8.6$  Hz, HA3,5), 6.57 (1H, d,  $J=2.1$  Hz, H8), 6.26 (1H, d,  $J=2.0$  Hz, H6), 5.04 (1H, dd,  $J=7.3, 3.3$  Hz, Hα), 4.11 (1H, dt,  $J=7.0, 3.4$  Hz, Hβ), 4.02 (6H, s, OMe), 3.72 (1H, ddd,  $J=12.0, 7.2, 3.4$  Hz, Hγ), 3.33 (1H, ddd,  $J=11.8, 7.9, 3.3$  Hz, Hγ); δ<sub>C</sub> (acetone-*d*<sub>6</sub>, 126 MHz) 183.12 (C4), 164.97 (C2), 164.29 (C7), 163.31 (C5), 158.80 (C9), 157.64 (CA4), 154.32 (C3',5'), 140.44 (C4'), 132.99 (CA1), 129.09 (CA2,6), 127.65 (C1'), 115.52 (CA3,5), 106.02 (C3), 105.45 (C10), 104.93 (C2',6'), 99.79 (C6), 94.99 (C8), 89.87 (Cβ), 73.69 (Cα), 61.48 (Cγ), 56.95 (OMe)

**syn-14c**: δ<sub>H</sub> (acetone-*d*<sub>6</sub>, 500 MHz) 7.42 (2H, s, H2',6'), 7.29 (2H, d,  $J=8.5$  Hz, HA2,6), 6.83 (1H, s, H3), 6.78 (2H, d,  $J=8.6$  Hz, HA3,5), 6.57 (1H, d,  $J=2.1$  Hz, H8), 6.26 (1H, d,  $J=2.0$  Hz, H6), 5.01 (1H, d,  $J=4.8$  Hz, Hα), 4.32 (1H, dt,  $J=5.2, 3.4$  Hz, Hβ), 3.98 (6H, s, OMe), 3.89 (1H,  $J=12.0, 3.4$  Hz, Hγ), 3.52 (1H, dd,  $J=12.1, 3.4$  Hz, Hγ); δ<sub>C</sub> (acetone-*d*<sub>6</sub>, 126 MHz) 183.05 (C4), 164.97 (C2), 164.36 (C7), 163.31 (C5), 158.80 (C9), 157.64 (CA4), 154.65 (C3',5'), 139.98 (C4'), 132.99 (CA1), 128.59 (CA2,6), 127.60 (C1'), 115.52 (CA3,5), 106.00 (C3), 105.43 (C10), 104.98 (C2',6'), 99.75 (C6), 95.02 (C8), 88.18 (Cβ), 73.35 (Cα), 60.97 (Cγ), 56.90 (OMe)

**Compound 15**

δ<sub>H</sub> 7.46 (2H, s, HA2,6), 7.42 (2H, s, H2',6'), 6.87 (1H, s, H3), 6.78 (1H, d,  $J=2.3$  Hz, H8), 6.42 (1H, d,  $J=2.2$  Hz, H6), 5.53 (1H, t,  $J=5.4$  Hz, Hβ), 5.31 (2H, s, OMOM), 3.94 – 3.91 (2H, m, Hγ), 3.88 (6H, s, A-OMe), 3.87 (6H, s, OMe), 3.47 (3H, s, OMOM); δ<sub>C</sub> 194.56 (Cα), 183.31 (C4), 164.69 (C2), 164.07 (C7), 162.92 (C5), 158.45 (C9), 153.82 (C3',5'), 148.43 (CA3,5), 142.13 (CA4), 140.70 (C4'), 127.59 (CA1), 127.20 (C1'), 107.52 (CA2,6), 106.62 (C10), 106.06 (C3), 104.99 (C2',6'), 100.36 (C6), 95.34 (C8), 94.99 (OMOM), 85.76 (Cβ), 63.93 (Cγ), 56.83 (OMe), 56.72 (A-OMe), 56.56 (OMOM)

**Compound 16**

δ<sub>H</sub> 7.75 – 7.71 (2H, m, HB2,6), 7.46 (2H, s, HA2,6), 7.42 (2H, s, H2',6'), 7.21 (1H, d,  $J=8.6$  Hz, HB5), 6.87 (1H, s, H3), 6.78 (1H, d,  $J=2.3$  Hz, H8), 6.42 (1H, d,  $J=2.2$  Hz, H6), 5.52 (1H, t,  $J=5.3$  Hz, Hβ), 5.33 (2H, s, Hβ'), 5.31 (2H, s, OMOM), 3.97 – 3.93 (2H, m, Hγ), 3.90 (2H, s, B-OMe), 3.88 (6H, s, OMe), 3.85 (6H, s, A-OMe), 3.47 (3H, s, OMOM), 2.27 (3H, s, OAc); δ<sub>C</sub> 195.40 (Cα), 194.13 (Cα'), 183.31 (C4), 168.67 (OAc), 164.66 (C2), 164.08 (C7), 162.92 (C5), 158.46 (C9), 153.82 (C3',5'), 153.51 (CA3,5), 152.48 (CB3), 145.03 (CB4), 141.76 (CA4), 140.48 (C4'), 134.74 (CB1), 132.21 (CA1), 127.34 (C1'), 123.90 (CB5), 122.38 (CB6), 112.80 (CB2), 107.33 (CA2,6), 106.62 (C10), 106.11 (C3), 104.97 (C2',6'), 100.36 (C6), 95.34 (C8), 94.99 (OMOM), 85.87 (Cβ), 75.50 (Cβ'), 63.75 (Cγ), 56.83 (OMe), 56.67 (A-OMe), 56.56 (OMOM), 56.44 (B-OMe), 20.48 (OAc)

### Compound 17

$\delta_{\text{H}}$  7.79 – 7.75 (2H, m, HB2,6), 7.45 (2H, s, HA2,6), 7.42 (2H, s, H2',6'), 7.20 (1H, dd,  $J=8.7, 3.4$  Hz, HB5), 6.87 (1H, s, H3), 6.78 (1H, d,  $J=2.2$  Hz, H8), 6.42 (1H, d,  $J=2.1$  Hz, H6), 5.51 (1H, t,  $J=5.3$  Hz, H $\beta$ ), 5.45 (1H, t,  $J=5.4$  Hz, H $\beta'$ ), 5.31 (2H, s, OMOM), 3.98 – 3.92 (4H, m, H $\gamma, \gamma'$ ), 3.88 (3H, s, B-OMe), 3.87 (6H, d,  $J=1.7$  Hz, OMe), 3.79 (6H, s, A-OMe), 3.47 (3H, s, OMOM), 2.27 (3H, s, OAc);  $\delta_{\text{C}}$  195.81 (C $\alpha'$ ), 195.36 (C $\alpha$ ), 183.32 (C4), 168.68 (OAc), 164.67 (C2), 164.09 (C7), 162.93 (C5), 158.46 (C9), 153.83 (C3',5'), 153.15 (CA3,5), 152.36 (CB3), 144.83 (CB4), 141.50 (CA4), 140.44 (C4'), 135.62 (CB1), 132.26 (CA1), 127.37 (C1'), 123.74 (CB5), 122.82 (CB6), 113.24 (CB2), 107.21 (CA2,6), 106.63 (C10), 106.12 (C3), 104.96 (C2',6'), 100.37 (C6), 95.35 (C8), 94.99 (OMOM), 86.52 (C $\beta'$ ), 85.88 (C $\beta$ ), 63.79 (C $\gamma'$ ), 63.66 (C $\gamma$ ), 56.83 (OMe), 56.59 (A-OMe), 56.56 (OMOM), 56.42 (B-OMe), 20.49 (OAc)

### Compound 18

$\delta_{\text{H}}$  7.71 (1H, dd,  $J=8.3, 2.0$  Hz, HB6), 7.65 (1H, d,  $J=2.2$  Hz, HB2), 7.46 (2H, s, H2',6'), 7.39 (2H, s, HA2,6), 6.92 (1H, d,  $J=8.3$  Hz, HB5), 6.81 (1H, s, H3), 6.57 (1H, d,  $J=2.2$  Hz, H8), 6.27 (1H, d,  $J=1.6$  Hz, H6), 5.51 (1H, t,  $J=5.2$  Hz, H $\beta$ ), 5.48 (1H, t,  $J=5.3$  Hz, H $\beta'$ ), 3.94 (2H, m, H $\gamma$ ), 3.90 (5H, m, B-OMe, H $\gamma'$ ), 3.88 (6H, s, OMe), 3.79 (6H, s, A-OMe);  $\delta_{\text{C}}$  195.37 (C $\alpha$ ), 194.51 (C $\alpha'$ ), 183.11 (C4), 165.02 (C2), 164.29 (C7), 163.30 (C5), 158.80 (C9), 153.81 (C3',5'), 153.13 (CA3,5), 152.45 (CB4), 148.31 (CB3), 141.79 (CA4), 140.30 (C4'), 132.07 (CB1), 129.11 (CA1), 127.53 (C1'), 124.59 (CB6), 115.39 (CB5), 112.27 (CB2), 107.27 (CA2,6), 106.05 (C3), 105.42 (C10), 104.90 (C2',6'), 99.78 (C6), 94.97 (C8), 86.16 (C $\beta'$ ), 85.92 (C $\beta$ ), 64.02 (C $\gamma'$ ), 63.91 (C $\gamma$ ), 56.81 (OMe), 56.62 (A-OMe), 56.30 (B-OMe)

### Compound 19 [T-(4'-O- $\beta$ )-S-(4-O- $\beta$ )-G]

ESIMS  $m/z$  [M+H]<sup>+</sup> 753; HRESIMS calculated for C<sub>38</sub>H<sub>40</sub>O<sub>16</sub> [M+H]<sup>+</sup> 753.2390, found 753.2396.

**anti-anti & syn-anti:**  $\delta_{\text{H}}$  7.38 (2H, s, H2',6'), 7.02 (1H, d,  $J=1.9$  Hz, HB2), 6.88 – 6.86 (1H, dd,  $J=8.1, 1.9$  Hz, HB6), 6.85 (2H, s, HA2,6), 6.78 (1H, s, H3), 6.74 (1H, d,  $J=8.0$  Hz, HB5), 6.55 (1H, d,  $J=1.9$  Hz, H8), 6.26 (1H, d,  $J=1.8$  Hz, H6), 5.07 (1H, d,  $J=5.9$  Hz, H $\alpha$ ), 4.95 (1H, t,  $J=5.1$  Hz, H $\alpha'$ ), 4.31 (1H, m, H $\beta$ ), 4.01 (6H, s, OMe), 3.87 (6H, s, A-OMe), 3.86 (1H, m, H $\beta'$ ), 3.81 – 3.80 (1H, m, H $\gamma$ ), 3.79 (3H, s, B-OMe), 3.60 (1H, dd,  $J=12.3, 3.5$  Hz, H $\gamma'$ ), 3.50 – 3.45 (1H, m, H $\gamma$ ), 3.25 (1H, dd,  $J=11.9, 3.1$  Hz, H $\gamma'$ );  $\delta_{\text{C}}$  183.04 (C4), 165.28 (C2), 164.14 (C7), 163.27 (C5), 158.78 (C9), 154.20 (C3',5'), 153.43 (CA3,5), 147.87 (CB3), 146.70 (CB4), 140.29 (C4'), 138.69 (CA4), 135.85 (CA1), 133.62 (CB1), 127.66 (C1'), 120.62 (CB6), 115.18 (CB5), 111.32 (CB2), 105.96 (C3), 105.31 (C10), 104.83 (C2',6', CA2,6), 99.87 (C6), 95.00 (C8), 89.64 (C $\beta'$ ), 88.73 (C $\beta$ ), 73.97 (C $\alpha'$ ), 73.64 (C $\alpha$ ), 61.80 (C $\gamma$ ), 61.19 (C $\gamma'$ ), 56.95 (OMe), 56.52 (A-OMe), 56.13 (B-OMe)

**syn-syn & anti-syn:**  $\delta_{\text{H}}$  7.36 (2H, s, H2',6'), 7.01 (1H, d,  $J=1.9$  Hz, HB2), 6.85 (2H, s, HA2,6), 6.78 (2H, m, H3, HB6), 6.74 (1H, d,  $J=8.0$  Hz, HB5), 6.55 (1H, d,  $J=1.9$  Hz, H8), 6.26 (1H, d,  $J=1.8$  Hz, H6), 5.07 (1H, m, H $\alpha$ ), 4.92 (1H, m, H $\alpha'$ ), 4.37 (1H, m, H $\beta$ ), 4.11 (1H, m, H $\beta'$ ), 4.00 (6H, s, OMe), 3.91 (1H, dd,  $J=12.0, 6.1$  Hz, H $\gamma$ ), 3.89 (6H, s, A-OMe), 3.81 (3H, s, B-OMe), 3.79 (1H, m, H $\gamma'$ ), 3.52 (1H, m, H $\gamma$ ), 3.38 (1H, m, H $\gamma'$ );  $\delta_{\text{C}}$  182.98 (C4), 165.26 (C2), 164.23 (C7), 163.27 (C5), 158.78 (C9), 154.56 (C3',5'), 153.70 (CA3,5), 147.91 (CB3), 146.37 (CB4), 139.98 (C4'), 138.78 (CA4), 133.70 (CB1), 135.50 (CA1), 127.66 (C1'), 119.97 (CB6), 115.18 (CB5), 110.78 (CB2), 106.03 (C3), 105.30 (C10), 104.90 (C2',6', CA2,6), 99.86 (C6), 95.08 (C8), 87.85 (C $\beta$ ), 87.74 (C $\beta'$ ), 73.24 (C $\alpha'$ ), 73.64 (C $\alpha$ ), 61.06 (C $\gamma$ ), 60.86 (C $\gamma'$ ), 56.95 (OMe), 56.52 (A-OMe), 56.16 (B-OMe)

### References

- Lam, P.Y., Zhu, F.Y., Chan, W.L., Liu, H.J. and Lo, C. (2014) Cytochrome P450 93G1 is a flavone synthase II that channels flavanones to the biosynthesis of tricin O-linked conjugates in rice. *Plant Physiol.*, **165**, 1315-1327.
- Morreel, K., Goeminne, G., Sterck, L., Ralph, J., Coppeters, W., Storme, V., Breyne, P., Steenackers, M., Georges, M., Messens, E. and Boerjan, W. (2006) Genetical metabolomics of flavonoid biosynthesis in *Populus*: a case study. *Plant J.*, **47**, 224-237.
- Ogo, Y., Ozawa, K., Ishimaru, T., Murayama, T. and Takaiwa, F. (2013) Transgenic rice seed synthesizing diverse flavonoids at high levels: a new platform for flavonoid production with associated health benefits. *Plant Biotechnology Journal*, **11**, 734-746.