Insights into lignin primary structure and deconstruction from *Arabidopsis thaliana* COMT (caffeic acid *O*-methyl transferase) mutant *Atomt1*†

Syed G. A. Moinuddin,‡ Michaël Jourdes,‡ Dhrubojyoti D. Laskar, Chanyoung Ki, Claudia L. Cardenas, Kye-Won Kim, Dianzhong Zhang, Laurence B. Davin and Norman G. Lewis*

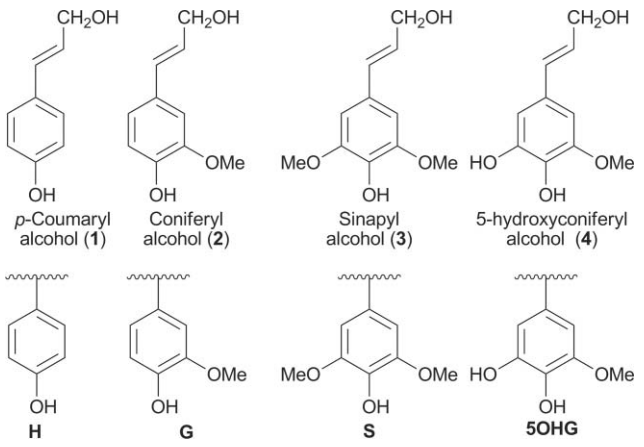
Received 31st March 2010, Accepted 19th May 2010
First published as an Advance Article on the web 22nd July 2010
DOI: 10.1039/c004817h

The *Arabidopsis* mutant *Atomt1* lignin differs from native lignin in wild type plants, in terms of sinapyl (S) alcohol-derived substructures in fiber cell walls being substituted by 5-hydroxyconiferyl alcohol (5OHG)-derived moieties. During programmed lignin assembly, these engender formation of benzodioxane substructures due to intramolecular cyclization of their quinone methides that are transiently formed following 8-*O*-4′ radical-radical coupling. Thioacidolytic cleavage of the 8-*O*-4′ inter-unit linkages in the *Atomt1* mutant, relative to the wild type, indicated that cleavable sinapyl (S) and coniferyl (G) alcohol-derived monomeric moieties were stoichiometrically reduced by a *circa* 2 : 1 ratio. Additionally, lignin degradative analysis resulted in release of a 5OHG–5OHG–G trimer from the *Atomt1* mutant, which then underwent further cleavage. Significantly, the trimeric moiety released provides new insight into lignin primary structure: during polymer assembly, the first 5OHG moiety is linked *via* a C8–*O*–X inter-unit linkage, whereas subsequent addition of monomers apparently involves sequential addition of 5OHG and G moieties to the growing chain in a 2 : 1 overall stoichiometry. This quantification data thus provides further insight into how inter-unit linkage frequencies in native lignins are apparently conserved (or near conserved) during assembly in both instances, as well as providing additional impetus to resolve how the overall question of lignin macromolecular assembly is controlled in terms of both type of monomer addition and primary sequence.

Introduction

The lignin pathway in angiosperms results in formation of the three main monomeric lignin precursors, *p*-coumaryl (**1**), coniferyl (**2**) and sinapyl (**3**) alcohols, which lead to the H (*p*-hydroxyphenyl), G (guaiacyl), and S (syringyl) components of lignins in their vascular cell walls, respectively.[1,2] The S-component is, however, mainly, if not exclusively, deposited in fiber cell walls in plant species such as *Arabidopsis*.[3] It was initially considered that for formation of the S and G precursors, both methylation steps were catalyzed by a single enzyme, the so-called caffeic acid *O*-methyl transferase (COMT).[4,5] This was due to its ability to methylate both caffeic (**5**) and 5-hydroxyferulic (**6**) acids *in vitro* to afford ferulic (**12**) and sinapic (**13**) acids, respectively (Fig. 1). Accordingly, COMT was considered to be substrate versatile, and responsible for producing both G and S lignins derived from coniferyl (**2**) and sinapyl (**3**) alcohols in angiosperms.[4,5]

In subsequent pioneering work from the Legrand/Fritig laboratory, various putative COMT isoforms (I, II and III) were purified from tobacco (*Nicotiana tabacum*)[6] and their corresponding cDNAs cloned.[7,8] In contrast to long held assumptions of COMT being substrate versatile, down-regulation of COMT I in tobacco clearly established that it was instead involved in the second methylation step leading to sinapyl alcohol (**3**) and S-moieties of lignin.[9] This was conclusively demonstrated by massive reduction of S-lignin deposition in these transformants, as shown by thioacidolysis release of monomers **19** and **21** (S/G ratios of ~1 for wild type (WT) and 0.07–0.09 in most COMT I down-regulated plants), with these resulting from presumed 8–*O*–4′ interunit linkage cleavage. There were also no significant differences in estimated Klason lignin contents for WT (~18.4% cell wall residue, CWR) and the COMT down-regulated (~18.3% CWR) lines. Thus, COMT was responsible for the second methylation step and was, therefore, effectively mono-functional in terms of substrate utilized *in vivo*.

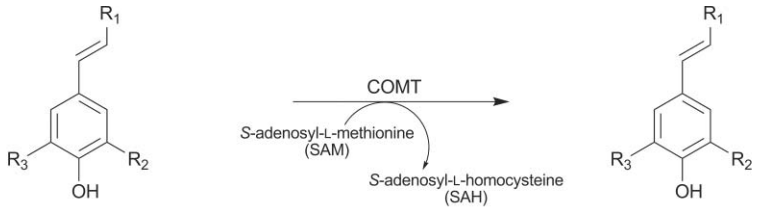


Yet just before this seminal study, reports from the Chiang[10] and Dixon[11] laboratories had concluded that heterologous

Institute of Biological Chemistry, Washington State University, Pullman, WA 99164-6340, USA. E-mail: lewisn@wsu.edu.; Fax: +1 509 335 8206; Tel: +1 509 335 2682

† This work was previously reported in oral presentations at the annual retreat of the BioEnergy Science Center (June 2009, Chattanooga, TN) and at the Phytochemical Society of North America (PSNA) annual meeting (August 2009, Towson, MD).

‡ Both contributed equally to the work, with S. G. A. Moinuddin completing all the synthetic chemistry.

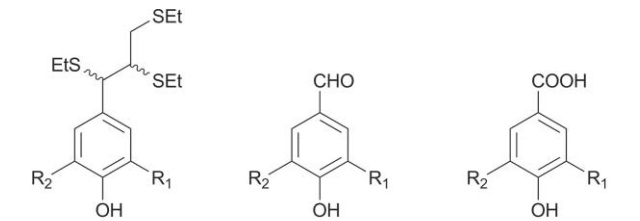


$R_1 = CO_2H$, $R_2 = OH$, $R_3 = H$, caffeic acid (**5**)
$R_1 = CO_2H$, $R_2 = OMe$, $R_3 = OH$, 5-hydroxyferulic acid (**6**)
$R_1 = COSCoA$, $R_2 = OH$, $R_3 = H$, caffeoyl CoA (**7**)
$R_1 = COSCoA$, $R_2 = OMe$, $R_3 = OH$, 5-hydroxyferuloyl CoA (**8**)
$R_1 = CHO$, $R_2 = OH$, $R_3 = H$, caffeyl aldehyde (**9**)
$R_1 = CHO$, $R_2 = OMe$, $R_3 = OH$, 5-hydroxyconiferyl aldehyde (**10**)
$R_1 = CH_2OH$, $R_2 = OH$, $R_3 = H$, caffeyl alcohol (**11**)
$R_1 = CH_2OH$, $R_2 = OMe$, $R_3 = OH$, 5-hydroxyconiferyl alcohol (**4**)

$R_1 = CO_2H$, $R_2 = OMe$, $R_3 = H$, ferulic acid (**12**)
$R_1 = CO_2H$, R_2, $R_3 = OMe$, sinapic acid (**13**)
$R_1 = COSCoA$, $R_2 = OMe$, $R_3 = H$, feruloyl CoA (**14**)
$R_1 = COSCoA$, R_2, $R_3 = OMe$, sinapoyl CoA (**15**)
$R_1 = CHO$, $R_2 = OMe$, $R_3 = H$, coniferyl aldehyde (**16**)
$R_1 = CHO$, R_2, $R_3 = OMe$, sinapyl aldehyde (**17**)
$R_1 = CH_2OH$, $R_2 = OMe$, $R_3 = H$, coniferyl alcohol (**2**)
$R_1 = CH_2OH$, R_2, $R_3 = OMe$, sinapyl alcohol (**3**)

Fig. 1 *In vitro* substrate versatile COMT-catalyzed conversions.

down-regulation of COMT in tobacco using aspen (*Populus tremuloides*) and alfalfa (*Medicago sativa*) COMTs, respectively, had no significant effect on the G/S ratios of the lignins. Furthermore, down-regulation of the alfalfa COMT in tobacco[11] gave transformants whose overall putative lignin contents were reportedly reduced by approximately 2/3 relative to WT, whereas the transformants resulting from down-regulation of aspen COMT in tobacco[10] had no apparent changes in lignin contents. However, determinations in the first of these studies were based on the unreliable thioglycolic acid lignin methods (discussed in Anterola and Lewis),[12] as well as in the use of an alkaline nitrobenzene protocol which can release *p*-hydroxybenzaldehydes (*e.g.*, **22**, **23** and **25**) and *p*-hydroxybenzoic acids (*e.g.*, **26**, **27** and **29**) from various phenylpropanoids, including those from lignins. Both studies had thus incorrectly supported the earlier view that COMT was substrate versatile *in vivo*.



$R_1, R_2 = H$ (**18**)
$R_1 = OMe$, $R_2 = H$ (**19**)
$R_1 = OMe$, $R_2 = OH$ (**20**)
$R_1, R_2 = OMe$ (**21**)

$R_1, R_2 = H$ (**22**)
$R_1 = OMe$, $R_2 = H$ (**23**)
$R_1 = OMe$, $R_2 = OH$ (**24**)
$R_1, R_2 = OMe$ (**25**)

$R_1, R_2 = H$ (**26**)
$R_1 = OMe$, $R_2 = H$ (**27**)
$R_1 = OMe$, $R_2 = OH$ (**28**)
$R_1, R_2 = OMe$ (**29**)

The pioneering work by Atanassova *et al.*,[9] establishing COMT as involved in the second methylation step *in vivo* leading to the S-component of its lignin, was later confirmed using poplar,[13–15] aspen,[16] and alfalfa.[17] Accordingly, confusion over COMT being bi-functional/substrate versatile *in vivo* for a period encompassing nearly two decades had been unambiguously resolved. Interestingly, other *in vitro* studies using heterologously expressed COMTs from tobacco,[18] aspen,[19] and alfalfa[20] in *Escherichia coli* later showed they were also able to methylate caffeic (**5**)/5-hydroxyferulic (**6**) acids, caffeoyl (**7**)/5-hydroxyferuloyl (**8**) CoAs, caffeyl (**9**)/5-hydroxyconiferyl (**10**) aldehydes or caffeyl (**11**)/5-hydroxyconiferyl (**4**) alcohols. This once again illustrates differences that can result between *in vitro* assays and the much more metabolically relevant biochemical processes *in vivo* due to substrate/enzyme compartmentalization in the latter.

Earlier studies, undertaken nearly eight decades ago, had also described various *brown-midrib* (*bm1–bm4*) mutants of maize (*Zea mays*) with putative colored lignins in cell walls of leaf midribs and sclerenchyma cells surrounding the vascular bundles of the stems.[21,22] Of these, the so-called bm_3 mutant had a putative lower ("acid detergent") lignin content as estimated by the potassium permanganate method,[23] as well as lower stalk strength as suggested by 17–26% less crushing strength and 8–14% less stalk-section weight than WT plants.[24] Later, it was reported that when all then known monolignol pathway enzymes were assayed in the bm_3 mutant *in vitro*, apparently only the activity of the COMT purportedly catalyzing methylation of caffeic acid (**5**) was strongly reduced.[25] Additionally, a later study of the bm_3 mutant line by Lapierre *et al.*,[26] employing thioacidolytic treatment of the corresponding stem tissue, resulted in a reportedly reduced level of releasable S-monomers (by a factor of *ca.* 4.4) relative to WT, together with an unidentified, putative, 5-hydroxyguaiacyl (5OHG) derived monomeric component (representing ~7% of total monomeric units released). For the WT line, the recovery of monomeric units **18**, **19** and **21** relative to estimated Klason lignin content was quite low (~20.2%), with an H : G : S ratio = 1 : 9.3 : 16, whereas the yield of estimated monomers **18**, **19** and **21** was ~16.1% in the mutant line, with a putative H : G : S : 5OHG ratio = 1 : 16 : 6.2 : 1.7. The data also suggested a small reduction in estimated lignin contents (Klason lignin: WT = 15.2%, bm_3 = 11.3%), as well as an overall reduction in S-lignin deposition. However, given the variability in the data, the study did not conclusively ascertain that the effect of this mutation was only on S-lignin deposition.

The overall purpose of this investigation was thus to begin to gain urgently needed insight into lignin primary structure using the *Arabidopsis* COMT1 (At5g54160) knockout mutant line *Atomt1*. Accordingly, an *Atomt1* lignin-derived isolate was subjected to non-exhaustive thioacidolytic cleavage in order to begin to precisely identify the substructure(s) resulting from coupling of the putative 5OHG units in the lignin primary structure, such as benzodioxanes (Fig. 2), and thus the effect on overall lignin polymer assembly (lignin amounts/monomeric compositions). This treatment resulted in release of cleavable benzodioxane substructures, a 5OHG–5OHG–G trimer and a 5OHG–G dimer from the *Atomt1* lignin isolate, with the latter presumed generated through cleavage of the corresponding trimer. Significantly, the

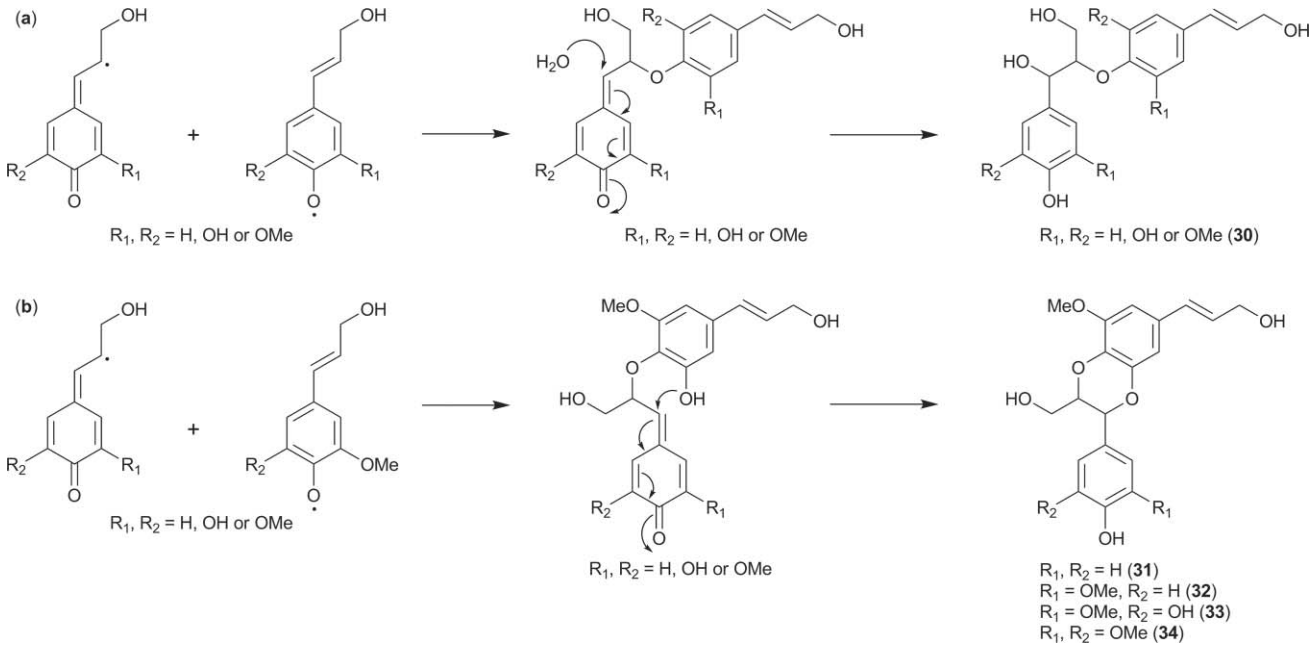


Fig. 2 Potential 8–*O*–4′ (**30**) and benzodioxane (**31–34**) lignin subunits in *Arabidopsis via* (a) H, G and S monolignol coupling and (b) H, G and S monolignol coupling as above, but with 5OHG for formation of substituted benzodioxanes (**31–34**).

G unit placement in the trimer/dimer isolated was either that of a terminal unit in this segment of the lignin biopolymer, or *perhaps* more likely as a 5OHG–5OHG–G repeating unit.

This study thus now clarifies further how the programmed lignification response is predictably modulated when COMT is either down-regulated or "knocked-out", while apparently nearly conserving overall 8–*O*–4′ inter-unit linkage frequency, as well as that of other inter-unit linkages. This limited substrate degeneracy/polymerization process also helps provide further insight into why, from an evolutionary perspective, this mutation was unfavorable throughout the plant kingdom. Additionally, of the 17 putative *COMT* multigene members in *Arabidopsis*, 12 were heterologously expressed in *Escherichia coli* with only one being able to *O*-methylate caffeic (**5**)/5-hydroxyferulic (**6**) acid, caffeyl (**9**)/5-hydroxyconiferyl (**10**) aldehyde and caffeyl (**11**)/5-hydroxyconiferyl (**4**) alcohol *in vitro* (Fig. 1). Taken together, these findings place further constraints on lignification *in vivo* in general, and further underscore the need to develop new methods to fully establish lignin primary structure(s).

Results and discussion

Growth characteristics and quantification of thioacidolytically released lignin-derived monomeric entities

Prior to investigating the nature of the lignin formed in the cell-wall stem tissues of *A. thaliana* WT (ecotype Wassilewskija) and the corresponding *Atomt1* knock-out lines, a comparison of various gross growth characteristics of their developing stems at different time intervals was carried out. Sampling of plant stem tissues was conducted from germination until maturation/senescence, with particular attention being directed to the vascular apparatus. The plots of overall stem lengths and basal stem widths during growth and development until senescence are depicted in Fig. 3. There were no readily visible phenotypical differences

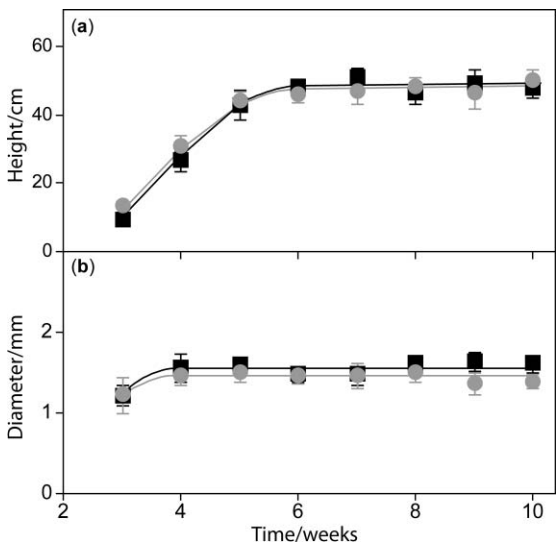


Fig. 3 Plots of growth/development parameters between *Arabidopsis* WT (black squares) and *Atomt1*-knockout (grey circles) lines until senescence; (a) stem lengths and (b) basal stem widths. [All measurements are an average of 20 plants per sampling point].

noted between both lines (WT and *Atomt1*) in terms of overall gross growth/development indices of the stems from rosette leaf development stage (3 weeks) until maturation/senescence (10 weeks).

It was next considered instructive to plot the amounts of cleavable lignin 8–*O*–4′ interunit linkages, in terms of release of their G and S thioacidolysis-derived monomeric moieties **19** and **21**, relative to the putative lignin contents of both lines as a function of growth/development.[1,2,27] For the WT line, the ▽, ◇ and □ values in Fig. 4a depict the monomeric S, G, and G+S amounts released (μmol g^{-1} putative AcBr lignin content) at weekly intervals. By contrast, for the *Atomt1* line, the

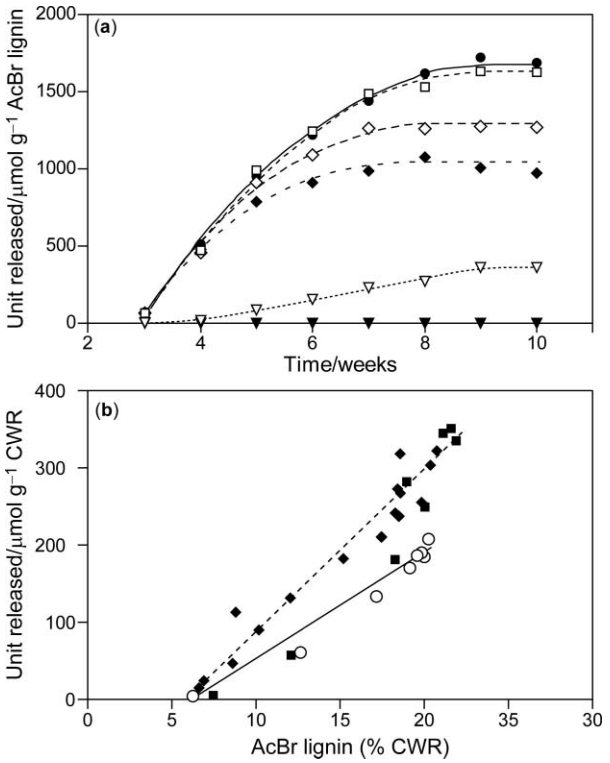


Fig. 4 (**a**) Plots of G and S thioacidolysis monomers (as μmol g^{-1} AcBr lignin) released from both *Arabidopsis* WT and *Atomt1* mutant lines, at different stages of growth and development. WT: ▽, S-derived monomer (**21**); ◇, G-derived monomer (**19**); □, G+S derived monomers (**19**)/(**21**). *Atomt1*: ▼, S-derived monomer (**21**); ◆, G-derived monomer (**19**); ● hypothetical inclusion of S : G monomers in 2 : 1 ratio. (**b**) Correlations of estimates of amounts of total monomeric G+S derivatives releasable by thioacidolysis as a function of estimated AcBr lignin contents; WT = dark symbols, *Atomt1*, open circles.

S-component (▼) was essentially undetectable as anticipated, with the G-amounts released (◆) also being *ca.* 20% lower than that of WT G levels (◇) at maturity. The estimated AcBr lignin contents of stems of both WT and *Atomt1* lines from the earliest stages of stem development/elongation (*circa* 3 weeks) until maturity were also plotted relative to releasable monomeric G and S moieties (Fig. 4b). In the WT line, after *circa* 6% "AcBr lignin" deposition, there is a linear increase in amounts of G and S monomers (**19** and **21**) released until maturation (highest lignin content). For the *Atomt1* line, there is also a linear increase in releasable G monomers (**19**) from early stages of stem development until maturation, with this again increasing proportionally relative to putative lignin level. It was useful at this stage of enquiry to attempt to correlate, for each time point, the differences in monomer release between the WT and *Atomt1* lines. Provisionally, this corresponded to that of the reductions in both S amount released, as well as the reduction in G values between the WT (G) and *Atomt1* (G) lines. For each time point, the reduction in monomer release corresponded to a S : G reduction in a 2 : 1 ratio. When hypothetically factoring in this stoichiometric correction (●, Fig. 4a) for each stage of growth/development, the corresponding plot became essentially superimposable over that of WT (G+S).

These data could thus be *hypothetically* interpreted as follows: (i) 5-hydroxyconiferyl alcohol (**4**) moieties presumed present in the *Atomt1* mutant were, at least in part, covalently linked to adjoining G residues in the benzodioxane substructure(s) of the cell-wall lignin, in the ratio of 2 : 1 (Fig. 4a), and (ii) since overall levels of lignification had not, in fact, been significantly altered in terms of amounts in either line, the overall assembly process had apparently only been modulated due to limited substrate degeneracy during programmed (template) polymerization.[28–30]

Note also for both lines (Fig. 4b), the plot of monomeric moieties released relative to putative AcBr lignin contents again established that these data points do not intersect at the origin of the X and Y axes, as previously noted.[1–3,27,31] In other work, we have demonstrated that some of the putative 'lignin' at the early stages of plant growth and development contain other non-lignin moieties which interfere with these estimations (manuscript in preparation).

NMR spectroscopic analyses of lignin-derived-isolates of WT and *Atomt1* lines

In order to examine the possible formation of benzodioxane substructure **VI** (Fig. 5c) during programmed lignification in the *Atomt1* mutant line, the lignin-enriched isolates from the WT and *Atomt1* lines were next obtained following published procedures.[27] Each was then individually subjected to ^{13}C, 2D heteronuclear multiple quantum coherence (HMQC) and 2D heteronuclear multiple bond coherence (HMBC) NMR spectroscopic analyses. In this regard, analyses of the oxygenated carbon resonances in the aliphatic region (50–90 ppm) confirmed that the lignin-isolates of the WT line (Fig. 5a) contained the five expected lignin substructures **I–V**, namely: 8–*O*–4′ aryl ether (substructure **I**), resinol-like (substructure **II**), phenylcoumaran (substructure **III**), dibenzodioxocin (substructure **IV**) and cinnamyl alcohol end groups (substructure **V**) (Fig. 5c).

The substructures in the WT lignin-enriched isolates (Fig. 5a) were also individually confirmed by analysis of their characteristic one-bond δ_C/δ_H correlations: 60.3/3.19 and 60.3/3.56 (C-9/9-H), 84.1/4.20 (C-8/8-H) and 71.7/4.68 (C-7/7-H) ppm for substructure **I**; 71.8/3.78 and 71.8/4.11 (C-9/9-H, C-9′/9′-H), 54.3/3.01 (C-8/8-H, C-8′/8′-H) and 85.7/4.59 (C-7/7-H, C-7′/7′-H) ppm for substructure **II**; 63.4/3.69 and 63.4/3.45 (C-9/9-H), 53.8/3.38 (C-8/8-H) and 87.6/5.42 (C-7/7-H) ppm, for substructure **III**; 62.3/4.08 and 62.3/3.95 (C-9/9-H), 86.6/4.08 (C-8/8-H) and 84.2/4.69 (C-7/7-H) ppm for substructure **IV**, as well as 63.1/4.13 (C-9/9-H) ppm for cinnamyl alcohol end groups (**V**). Other inter-unit linkages, such as 8–1′, diphenyl and/or diphenyl ether bonds, were not further characterized individually as earlier[27,32] due to overlapping resonances, although they are also present. However, there was no detection of any cross-peaks consistent with substructure **VI** (down to baseline) under the conditions employed for analysis of the WT lignin-isolate.

The carbon–proton correlations in the 2D-HMQC spectra of the lignin-enriched isolates from the *Atomt1* mutant line also established the presence of the five expected lignin substructures **I–V** (Fig. 5b). These were individually detected by their characteristic cross-peak δ_C/δ_H correlations as follows: 59.9/3.12 and 59.9/3.47 (C-9/9-H), 84.2/4.19 (C-8/8-H) and 71.2/4.65 (C-7/7-H) ppm for 8–*O*–4′ aryl ether substructure **I**; 71.5/3.65 and 71.5/4.06

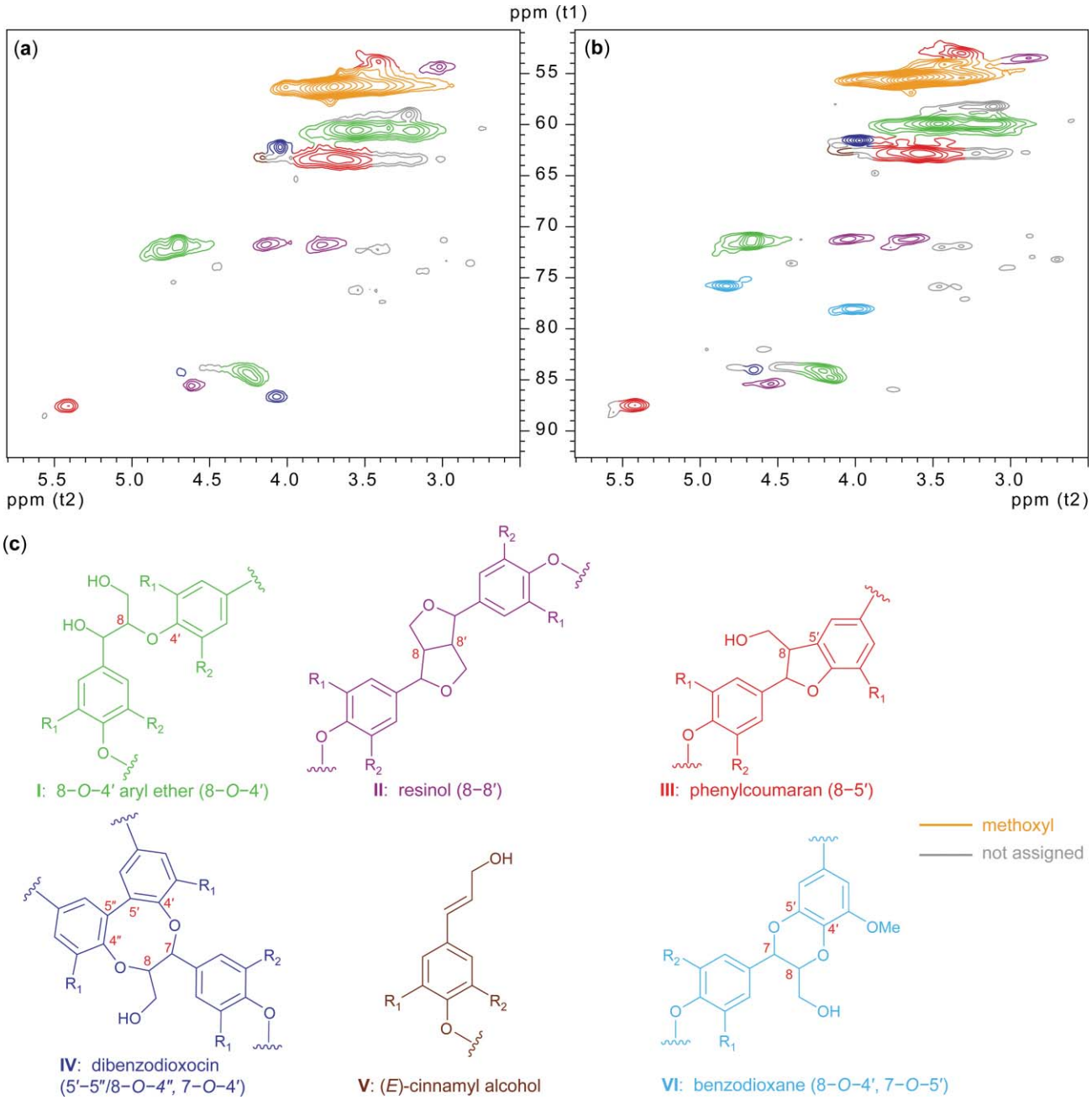


Fig. 5 Partial NMR spectra of lignin-enriched isolates from *Arabidopsis* WT and *Atomt1* lines. 2D HMQC spectra of oxygenated aliphatic region of lignin-enriched isolate from (**a**) WT and (**b**) *Atomt1*. (**c**) Lignin substructures **I–VI** assignable by NMR spectroscopic analyses.

(C-9/9-H, C-9′/9′-H), 53.4/2.84 (C-8/8-H, C-8′/8′-H) and 85.4.6/4.54 (C-7/7-H, C-7′/7′-H) ppm for resinol-like substructure **II**; 62.8/3.68 and 62.8/3.39 (C-9/9-H), 52.9/3.32 (C-8/8-H) and 87.5/5.41 (C-7/7-H) ppm for phenylcoumaran substructure **III**; 61.7/4.03 and 61.7/3.92 (C-9/9-H) and 84.1/4.67 (C-7/7-H) ppm for dibenzodioxocin substructure **IV**, and 62.6/4.08 (C-9/9-H) ppm for cinnamyl alcohol end groups (**V**). In contrast to the WT sample, analysis of the 2D HMQC NMR spectrum of the *Atomt1* mutant lignin-enriched isolate (Fig. 5b) established the presence of an additional substructure that was characterized by two correlations (δ_C/δ_H) at 76.0/4.81 (C-7/7-H) and 78.2/4.0 (C-8/8-H) consistent with benzodioxane substructure **VI** (Fig. 5c). In addition, the observed long-range HMBC correlation from the 5OHG aromatic proton-2 at δ_H ~6.80 ppm to the aromatic ring carbon-4 (δ_C ~133.9 ppm) further extended the correlation from δ_C ~133.9 ppm across to the ether-linked proton-8 (δ_H ~4.1 ppm) (data not shown). This established the presence of a 5OHG aromatic ring in the benzodioxane subunit linked *via* an 8–*O*–4′/7–*O*–5′ bonding pattern to the aliphatic side-chains of both G, as well as 5OHG–G, units in benzodioxane substructure **VI**. This, in turn, suggested that the 5OHG moieties were both homo- and hetero-(G) linked in the lignin of the *Atomt1* line.

Table 1 Provisional relative frequencies (%) of inter-unit linkages by 2D HMQC NMR analyses in lignin isolates of WT (*Arabidopsis* and alfalfa) *versus* the *Atomt1* mutant and alfalfa COMT down-regulated lines

	Arabidopsis		Alfalfa [a]	
Inter-unit linkage	WT	*Atomt1*	WT	COMT down-regulated
8–*O*–4′ (**I**)	79	60	81	44
Resinol (**II**)	7	5	6	3
Phenylcoumaran (**III**)	9	9	8	8
Dibenzodioxocin (**IV**)	2	4	1	4
CA end groups (**V**)	2	3	3	4
Benzodioxane (**VI**)	0	18	0	38

[a] Re-calculated from Marita *et al.*[33]

Conservation of 8–*O*–4′ inter-unit lignin frequencies in WT and COMT knocked-out/down-regulated lines (*Arabidopsis versus* alfalfa).

Currently, neither thioacidolysis nor nitrobenzene oxidation degradative analyses of the lignins in either WT or *Atomt1* lines can account for the bulk of the putative lignins in these plant samples, since recoveries of monomers relative to overall *putative* lignin contents are *circa* 20–40%.[1] Moreover, accurate quantification of benzodioxane substructure (**VI**) in the lignin of the *Atomt1* line is not yet currently achievable using either degradation method (unpublished results). Accordingly, attempts to provisionally estimate the relative overall frequencies of lignin substructures (*i.e.*, **I**, **II**, **III**, **IV**, **V** and **VI**) were next carried out using NMR spectroscopic analyses. These were estimated by measuring the volume integrals in the 2D-HMQC spectra of the lignin-enriched isolates of both WT and *Atomt1* mutant lines, as had been previously reported for alfalfa lignin-derived preparations.[33] In this way, the NMR spectroscopic analysis of the HMQC cone volumes of C-7/7-H correlations at 71.7/4.68 and 76.2/4.87 ppm, that are characteristic for the 8–*O*–4′ (**I**) and benzodioxane (**VI**) substructures, respectively, suggested that substructure **VI** can contribute to ≤25–26% of the total detectable 8–*O*–4′ interunit linkages present in the lignin-enriched isolate of the *Atomt1* line. Estimated inter-unit linkage (substructures **I**, **II**, **III**, **IV**, **V** and **VI**) frequencies in the lignin isolates of WT and *Atomt1* mutant lines are summarized in Table 1.

Based on this approach, the WT *Arabidopsis* line provisionally had ~79% of 8–*O*–4′ (**I**), ~7% of resinol-like (**II**), ~9% of phenylcoumaran (**III**), ~2% of dibenzodioxocin (**IV**), ~2% of cinnamyl alcohol end groups (**V**) and no detectable benzodioxane (**VI**) inter-unit linkages, whereas the *Atomt1* line accounted for ~60% of 8–*O*–4′ (**I**), ~5% of resinol-like (**II**), ~9% of phenylcoumaran (**III**), ~4% of dibenzodioxocin (**IV**), ~3% of cinnamyl alcohol end groups (**V**) and ~18% of benzodioxane (**VI**) inter-unit linkages, respectively. The benzodioxane substructure (**VI**), however, is a chemical variation of the 8–*O*–4′ (**I**) inter-unit linkage. Thus, taken together, the total 8–*O*–4′ linkage levels in the *Atomt1* line are estimated as ~78%, *i.e.* the sum of ~60% 8–*O*–4′ (**I**) and ~18% benzodioxane (**VI**) inter-unit linkages, respectively. This, in turn, suggests that the 8–*O*–4′ linkage frequency in the lignin isolates of both the *Atomt1* mutant and the WT line was conserved. In addition, the relative ratios and amounts of the other inter-unit linkages detected in this way are provisionally also reasonably well conserved using this approach. However, these data now need to be further verified through other means such as by quantitative degradative analyses.

Furthermore, a re-evaluation of the previously estimated levels of comparable inter-unit linkages in alfalfa (*Medicago sativa* L.), as reported in an earlier study of COMT down-regulation,[33] resulted in our interpretation that the overall 8–*O*–4′ inter-unit linkage frequency was also conserved in both WT and COMT alfalfa lines. We deduced, from the previously reported NMR spectroscopic analysis, that the WT line had ~81% of overall inter-unit linkages, whereas the COMT-line was considered to have ~38% of benzodioxane (**VI**) and ~44% of 8–*O*–4′ (**I**) inter-unit linkages.[1] Once again, the combined relative amounts totaled ~82%, indicating that estimated levels of 8–*O*–4′ inter unit linkages were conserved in both alfalfa lines. Yet, by contrast, this same alfalfa lignin inter-unit-linkage frequency data and the presence of the benzodioxane subunits was considered as further evidence for random assembly, *via* a "combinational or combinatorial biochemistry" process.[34,35] These same researchers are committed though to a random assembly process, with lignins having up to 10^{66} possible isomers *via* monolignol-derived polymerization, as well as putatively being able to utilize a wide range of non-monolignol phenolics in cell-wall assembly, *etc.*[35] Additionally, it was also ventured that no two lignin structures are the same within all plant species in the plant kingdom,[35] but with no proof for such an assertion.

By contrast, our own consideration of how lignification occurs is diametrically opposite to that above.[1,2,36] Firstly, the frequency of interunit linkages for both the *Arabidopsis* and alfalfa WT and COMT knockout/mutant appears to be quite similar for all linkage types *within* each plant line, albeit with some small changes (*e.g.* in resinol (**II**) levels from 7 to 5% in *Arabidopsis* and 6 to 3% in alfalfa). Whether such estimated differences are significant, or whether the programmed polymerization template is perturbed because of the presence of benzodioxane substructures, cannot be distinguished at this time. This high level of interunit linkage conservation has occurred even though there are substantial differences in the levels of benzodioxane-linked subunits in both cases (*i.e.* 18 *versus* 38% subunits in *Arabidopsis* and alfalfa, respectively) which could, in part, have also differentially perturbed the proposed template-assisted lignin assembly process.[1,2,28–30] Yet, while the S : G (cleavable monomer) ratio is quite different in both lines at plant maturity, (*i.e.* 0.28 and 0.64 for *Arabidopsis* and alfalfa, respectively), these ratios when divided by the benzodioxane level (as a fraction of 1) give near identical ratios of *circa* 1.6 and 1.7. Secondly, differences noted in S : G ratios between both plant lines presumably also reflect differences in overall numbers and thickness of S-rich fiber cell types in both species. Taken together, there is thus apparently a remarkably unplastic assembly process during lignification in both *Arabidopsis* and alfalfa, albeit with substantial differences in numbers of specific cell types that result in the distinct S : G ratios noted. We consider this indicative of a tightly controlled/conserved assembly process in distinct cell types, as we have already reported in other studies at the cell wall specific level.[3]

Molecular weight distributions of lignin-isolates

The molecular weight distributions of the lignin-enriched isolates from *Atomt1* and WT lines were estimated by gel permeation

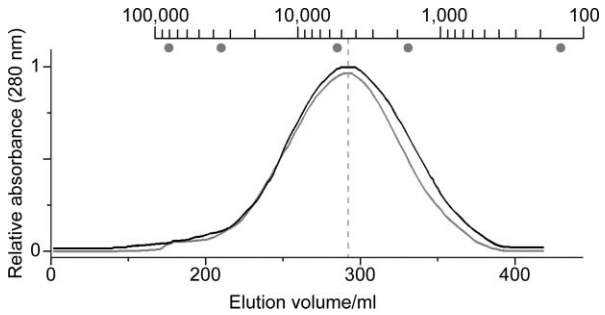
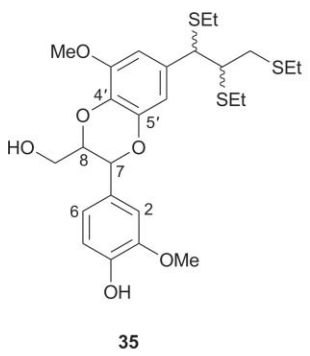


Fig. 6 Molecular weight distribution profiles of the lignin-enriched isolates from *Atomt1* (grey) and WT (black) lines.

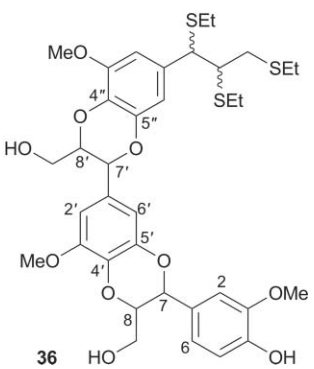
chromatography using a pre-calibrated Sephadex G-100 column[32] eluted with 0.1 M NaOH, in order to minimize lignin association effects.[37] The molecular weight distribution elution profiles of the enriched lignin isolates from *Atomt1* and WT lines were relatively broad, but very similar, with both centered at ~4500 Da (Fig. 6). These values are also in good agreement with previously reported Mw values of lignin-enriched isolates from *A. thaliana* WT.[27,32] The polydispersity index values (Mw/Mn) were also similar, being calculated as ~2.2 and ~2.5 for WT and *Atomt1* lignin-enriched isolates, respectively. However, due to the preparation procedure for the lignin-enriched isolates, the polydispersity index does not reflect the *in situ* molecular weight distributions of the actual lignin polymers but presumably better reflects only the "random" cleavage occurring during mechanical ball milling treatment (~4 days).

Lignin deconstruction and identification of isolable benzodioxane-linked subunits

Isolation of *Atomt1* lignin-derived trimer (36) and dimer (35) thioacidolysis derivatives. From the initial thioacidolysis analyses of intact *Atomt1* stem tissues described above, it was provisionally considered that the benzodioxane substructure (**VI**) was more resistant to cleavage than the 8–*O*–4′ inter-unit linkage, *i.e.* given that monomeric 5OHG moieties were not detected to any significant level in the mutant following thioacidolysis. Therefore, to investigate this question, thioacidolysis was next carried out under conventional conditions[38] using the previously discussed *Atomt1* lignin-derived-isolate harboring benzodioxane-linked substructures. As anticipated, NMR spectroscopic analysis of the resulting crude thioacidolysis reaction mixture verified presence of benzodioxane substructures [4.81 (7-H) and 76.0 (C-7)/4.0 (8-H) and 78.2 (C-8)], *i.e.* they had not been (fully) cleaved under these conditions.



35



36

The crude thioacidolysis reaction mixture was then fractionated using preparative silica gel thin-layer chromatography and reversed-phase HPLC, with the corresponding fractions individually subjected to 1D and 2D HMQC/HMBC NMR spectroscopic analyses to determine which of those contained benzodioxane subunits. Fractions containing the traces of cleaved trimeric **36** and the corresponding dimeric **35** benzodioxane were subjected to mass spectroscopic analyses using LC-APCI-MS. This gave masses in the positive ion mode at *m/z* 672.9 $(M + H–HSEt)^+$, 611.1 $(M + H–2HSEt)^+$ (Fig. 7a) and *m/z* 478.9 $(M + H–HSEt)^+$, 417.1 $(M + H–2HSEt)^+$ (Fig. 7b) indicative of the presence of trimer **36** and dimer **35**, respectively. This was further confirmed by analysis of three-bond long range HMBC correlations of the isolated fractions. Specifically, the trimer structure 5OHG–5OHG–G (**36**) was deduced from two distinct sets of three bond correlations, correlating the benzodioxane carbon-7 at δ_C ~76.2 ppm to the G aromatic ring proton-2 (δ_H ~7.18 ppm) and 6 (δ_H ~6.98 ppm), as well as the 5OHG aromatic ring proton-2 (δ_H ~6.80 ppm) and 6 (δ_H ~6.69 ppm), respectively. The structure of the 5OHG–G dimer (**35**) was also deduced from the correlation of proton-7 (δ_H ~4.87 ppm) of the benzodioxane functionality to carbon-2 (δ_C ~111.9 ppm) and carbon-6 (δ_C ~121.5 ppm) of the guaiacyl ring (G) unit, as well as from extended individual cross-peaks that correlated with aromatic protons-2 (δ_H ~7.12 ppm) and 6 (δ_H ~6.98 ppm) and aromatic carbon-4 (δ_C ~147.9 ppm), respectively. That G and 5OHG in **36**, as well as G in **35**, moieties were individually linked to the 7 and/or 7′ position of the benzodioxane ring was confirmed further from the strong cross-peak correlation of protons-2 and 6 for the respective aromatic ring (G and/or 5OHG) to that of their corresponding carbons-2, 4 and 6 in the respective HMBC spectra. Interestingly, both 1D and 2D NMR spectroscopic analyses of **36** and **35** established that the hydroxyl groups at the 9- and 9′-positions in trimer **36** and the 9-position in dimer **35** were not thioethylated under thioacidolytic reaction conditions. No other related products were detected under the conditions employed.

Confirmation of lignin subunit structures by total synthesis. To further unambiguously confirm the structures of the putative thioacidolytically cleaved lignin-derived trimer **36** and dimer **35**, a synthetic strategy was next developed, the details of which are summarized in the Experimental section. In this regard, various coupling products were synthesized by modification of an Ag_2CO_3 oxidation[39] in *circa* 58–62% yield from coniferyl (**2**) and sinapyl (**3**) alcohols, methyl 5-hydroxyferulate (**37**), and its *tert*-butyldimethylsilyl ether derivative (**38a/b**), and *O*-*tert*-butyldimethylsilyl coniferyl alcohols (**41a/b**) (Scheme 1). This afforded the corresponding dimers 5OHG–G (**32**),[39] 5OHG–S (**34**),[40–43] and 5OHG–5OHG (**43**)[40,43] with the 5OHG–5OHG–G trimer (**46**) being obtained from coupling of methyl ester of the 5OHG–5OHG dimer (**44**) and coniferyl alcohol (**2**) followed by DIBAL reduction. Initially, the Ag_2CO_3 oxidation of methyl 5-hydroxyferulate (**37**) and coniferyl alcohol (**2**) afforded a mixture of the methyl esters of the *threo*/*erythro*-5OHG–G benzodioxane dimer **39** in a 9 : 1 ratio, which was later reduced with DIBAL to form the *threo*/*erythro*-5OHG–G benzodioxane dimer **32** in a similar diastereomeric ratio with 88% yield. The HRMS spectrum of *threo*/*erythro*-**32** gave a (M + Na) at *m/z* 397.1259, consistent with the molecular formula $C_{20}H_{22}O_7$ + Na (calculated mass:

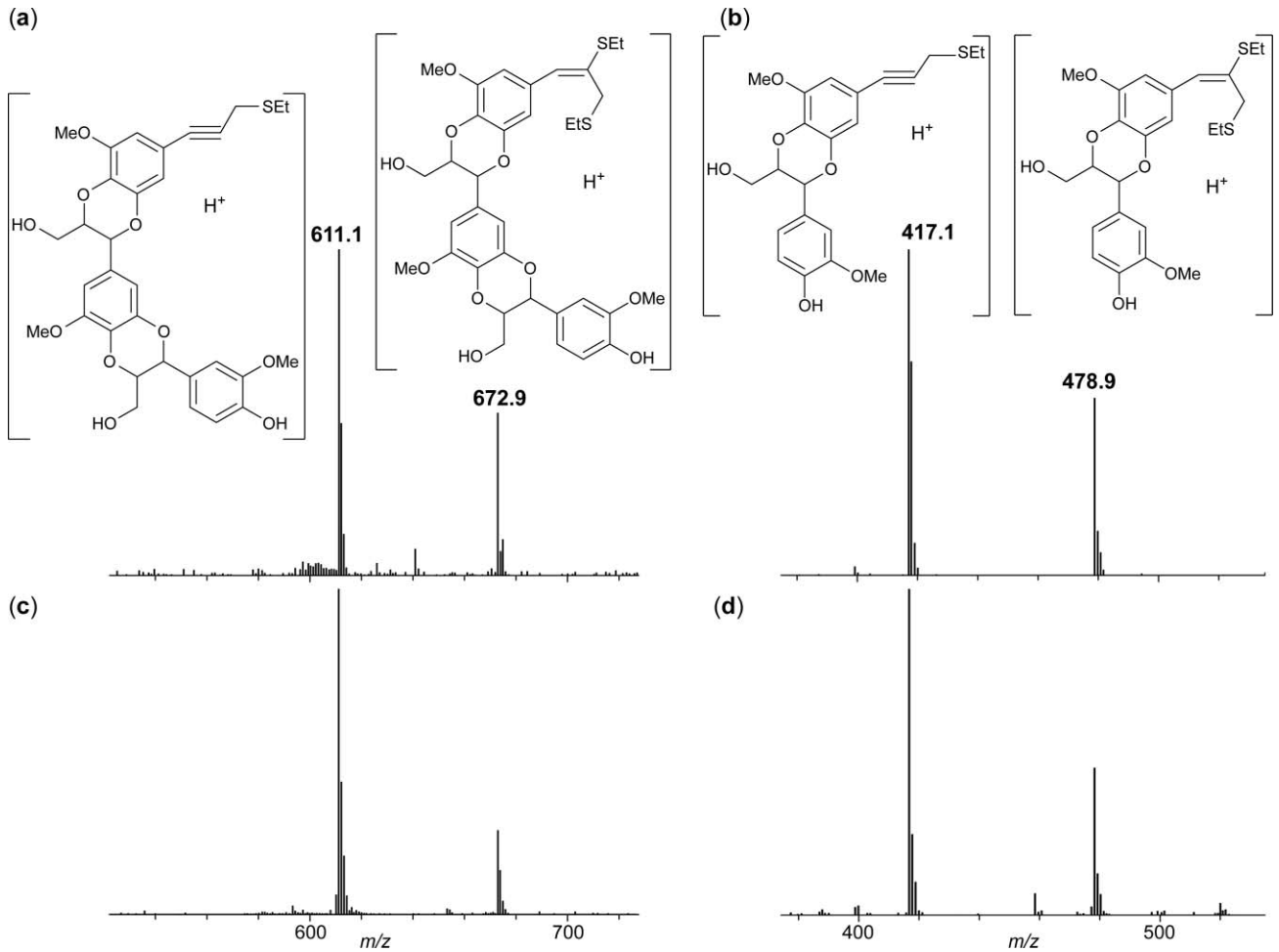
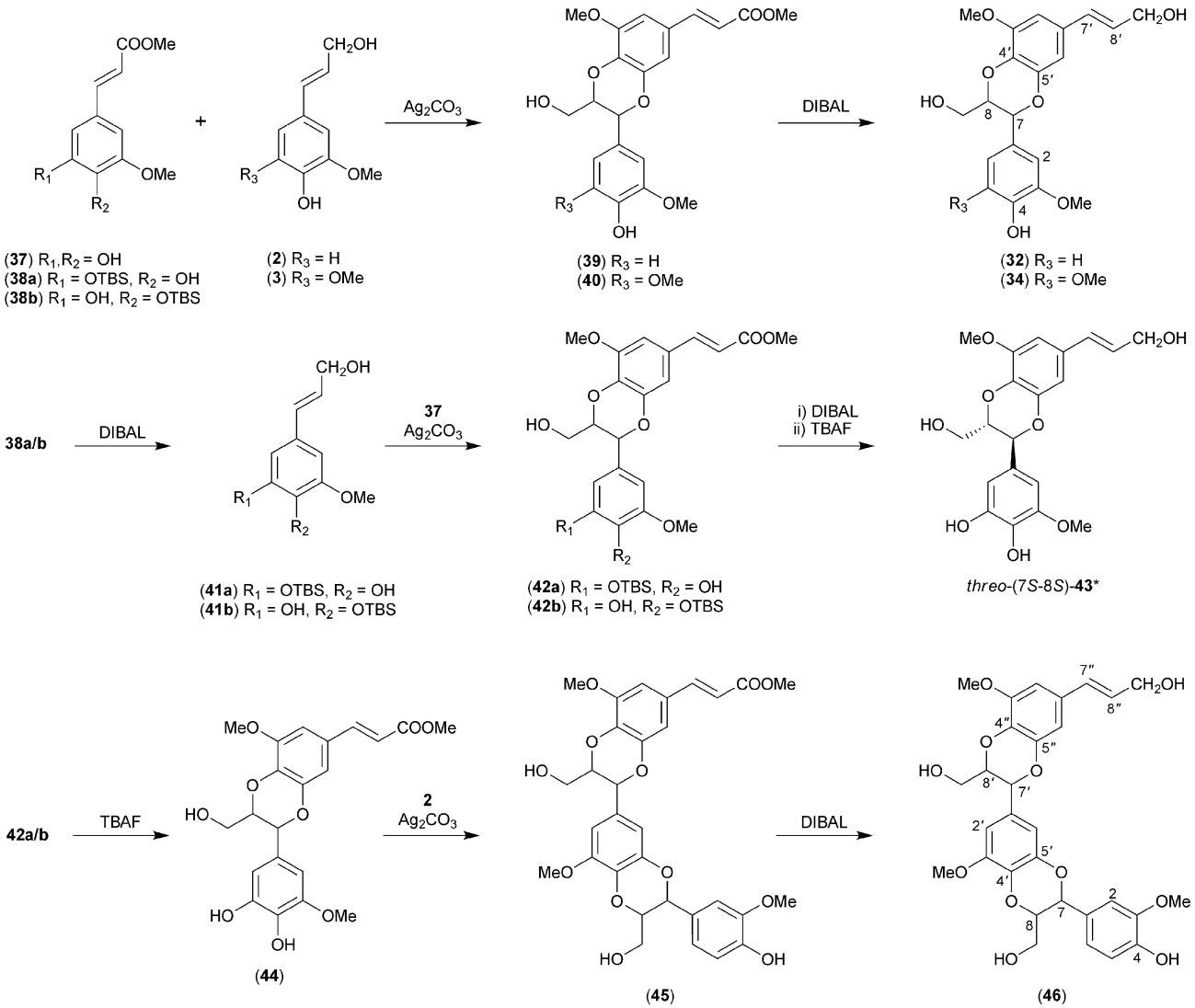


Fig. 7 Mass spectroscopic analyses of *Atomt1* lignin-derived trimer **36** and dimer **35** and their synthetic counterparts. (**a**) Mass spectrum of *Atomt1* lignin-derived trimer **36**; (**b**) mass spectrum of *Atomt1* lignin-derived dimer **35**; (**c**) mass spectrum of synthetic trimer **36**; and (**d**) mass spectrum of synthetic dimer **35**, and their putative charged fragments.

397.1263), and its LC-APCI-MS analysis gave a (M – H) ion at *m/z* 373.0 corresponding to molecular formula $C_{20}H_{22}O_7$. The ^{1}H and ^{13}C NMR spectra of *threo*/*erythro*-**32** were also in close agreement with the reported literature data for nocomtol.[39] Further, the *threo* and *erythro* diastereoisomers of **32** were individually separated by chiral OJ column and their structures were confirmed by 1D and mass spectral analyses (see Instrumentation and Chromatography Materials section). In an analogous manner, the Ag_2CO_3 oxidation of methyl 5-hydroxyferulate (**37**) and sinapyl alcohol (**3**) initially afforded a methyl ester of the *threo*-5OHG–S benzodioxane dimer **40**, which was later reduced with DIBAL to afford the *threo*-5OHG–S benzodioxane dimer **34** in 85% yield. Specifically, the HRMS spectrum of *threo*-**34** gave a (M + Na) ion at *m/z* 427.1400, consistent with a molecular formula of $C_{21}H_{24}O_8Na$ (calculated mass: 427.1369), whereas LC-APCI-MS analysis of **34** gave a (M – H) ion at *m/z* 403.0 corresponding to molecular formula $C_{21}H_{24}O_8$. Additionally, the ^{1}H and ^{13}C NMR spectra of *threo*-**34** were in close agreement with the reported literature data for nitidanin.[40–43]

To obtain the 5OHG–5OHG benzodioxane dimer **43**, three different synthetic protocols were considered. First, a modified Ag_2CO_3 oxidative coupling of 5-hydroxyconiferyl alcohol (**4**) directly gave the desired product **43** in its *threo* form, but in very low yield (~4%). A slightly better yield of product **43** (~10%) was obtained using horseradish peroxidase (HRP), in the presence of H_2O_2, with the product so obtained in a 1 : 9 *erythro*/*threo* ratio. Interestingly, the *erythro*/*threo* product from the HRP/H_2O_2 catalyzed coupling of 5-hydroxyconiferyl alcohol (**4**) could readily be separated into its *threo* and *erythro* diastereoisomers by chiral OJ column chromatography (see Instrumentation and Chromatography Materials section). However, the best overall synthetic procedure involved modified Ag_2CO_3 oxidative coupling of the *tert*-butyldimethylsilyl ether derivatives of 5-hydroxyconiferyl alcohol (**41a**/**b**) with 5-hydroxyferulate methyl ester (**37**) to afford initially the corresponding benzodioxane esters **42a**/**b** which were then converted to **43** by DIBAL reduction and TBAF deprotection (Scheme 1). The overall yield of this conversion from **38a**/**b** was ~90%, with the product **43** obtained in *threo* form.

The HRMS spectrum of *threo*-**43** gave a (M + Na) ion at *m/z* 413.1207, consistent with the molecular formula $C_{20}H_{22}O_8$ + Na (calculated mass: 413.1212), with the LC-APCI-MS also giving a (M – H) ion at *m/z* 388.9. Comparison of the ^{1}H and ^{13}C NMR spectra with 5OHG–G diol **32** and 5OHG–S diol **34** also established the presence of an hydroxyl group in compound



Scheme 1 Synthetic route to dimers **32**, **34**, **43** and trimer **46**. [**threo-(7R-8R)*-**43** is also formed.]

threo-**43** at the 5-position. That is, the ^{1}H NMR spectrum contained two sets of *meta*-coupled signals as doublets at δ 6.7 (*J* 1.5, 2′-H), 6.67 (*J* 1.5, 6′-H) and 6.65 (*J* 1.5, 2-H)/6.59 (*J* 2, 6-H) ppm, along with two methoxyl groups at δ 3.85 (s, 3′-OMe) and 3.83 (s, 3-OMe) ppm, indicative of two 5OHG units coupled together. Furthermore, the ^{1}H NMR spectrum of *threo*-**43** displayed two *trans*-olefinic protons as doublets at δ 6.47 (*J* 16.0, 7′-H), 6.27 (*J* 5.5 and 15.5, 8′-H) ppm, and a doublet at δ 4.20 (*J* 5.0, 9′-H) ppm for an allylic alcohol group connected to one of the two 5OHG units. The *trans*-conformation for *threo*-**43** was also established from analysis of the resonances at δ 4.9 (d, *J* 8.0, 7-H) and 4.0 (m, 8-H) ppm corresponding to two vicinal aliphatic oxymethines as well as two geminal methylene protons as double doublets at δ 3.78 (*J* 2.0 and 12.25, 9-H) and 3.52 (*J* 3.5 and 12.5, 9-H) ppm, respectively. The 5OHG–5OHG diol *threo*-**43** structure was further confirmed by 2D (HMBC, HMQC) NMR spectroscopic analyses. In this regard, the NMR spectra for *erythro*-**43** were also essentially identical to *threo*-**43**, except for the vicinal aliphatic oxymethine proton resonances at δ 5.2 (d, *J* 3.0, 7-H) and 4.46 (m, 8-H) ppm as well as the geminal methylene protons 3.62 (br dd, *J* 5.0, 7.5 and 12.0, 9-H) and 3.49 (dd, *J* 5.0 and 12.0, 9-H) ppm, respectively. Interestingly, this is the first report of the chemical synthesis of 5OHG–5OHG dimer (**43**), although it has been reported to occur as a natural product, simplidin (*threo* configuration), in extracts of *Firmiana simplex*.[43]

In a similar manner, the chemical synthesis of trimer **46** was achieved *via* modified Ag_2CO_3 oxidative coupling of methyl ester of 5OHG–5OHG benzodioxane dimer **44** with coniferyl alcohol (**2**) followed by DIBAL reduction (Scheme 1). The HRMS spectrum of **46** had a (M + Na) ion at *m*/*z* 591.1846, consistent with molecular formula $C_{30}H_{32}O_{11}$ + Na (calculated mass; 591.1842), with the LC-APCI-MS of **46** also giving a (M − H) ion at *m*/*z* 567.0 corresponding to molecular formula $C_{30}H_{32}O_{11}$. Interestingly, its structure closely resembles that of the previously reported benzodioxane/benzodioxane-type sesquineolignans, namely isoamericanol B1-C2,[44] except for the presence of methoxyl groups at the 3, 3′ and 3″- positions in **46**. Analysis of the ^{1}H and ^{13}C NMR spectra of **46** established that it existed as a mixture of diastereoisomers in the ratio of ~1.3 : 1 (*threo*/*erythro*) that could not readily be separated. The ^{1}H NMR spectrum of **46**

also contained three proton signals at δ 7.12 (dd, *J* 2.0 and 4.5, 2-H), 6.97 (d, *J* 8.0, 6-H), 6.88 (d, *J* 8.0, 5-H) ppm for a G-unit and four proton signals at δ 6.78 (br s, 2′-H), 6.71 (dd, *J* 2.0 and 4.5, 2″-H), 6.70 (dd, *J* 2.0 and 4.5, 6″-H) and 6.60 (d, *J* 2.0, 6′-H) ppm corresponding to two 5OHG units. Additionally, the presence of three methoxyl groups at the 3, 3′ and 3″-positions was confirmed from the resonances at δ 3.88 (s, 3-OMe), 3.87 (s, 3′-OMe) and 3.86 (s, 3″-OMe) ppm, respectively. The ^{1}H spectrum included signals attributable to *trans*-olefinic protons at δ 6.47 (d, *J* 16.0, 7″-H), 6.27 (dt, *J* 5.5 and 16.0, 8″-H) ppm and an hydroxy-methyl proton at δ 4.20 (t, *J* 5.0, 9′-H) ppm. In addition to these proton resonances, the ^{1}H NMR spectrum of **46** established the presence of two sets of aliphatic AMXY-type signals at δ 4.99 (d, *J* 8.0, 7-H), 4.97 (d, *J* 7.5, 7′-H), 4.08 (m, 8-H and 8′-H), 3.82 (dd, *J* 4.5 and 12.0, 9-H and 9′-H) and 3.54 (dd, *J* 5.5 and 12.5, 9-H and 9′-H) ppm, respectively. The structure of the 5OHG–5OHG–G benzodioxane trimer **46** was further confirmed by 2D (HMBC, HMQC and HSQC) NMR analyses (data not shown).

Chiral HPLC analysis for 32, 34, 43 and 46. Chiral HPLC was then used to establish conditions for baseline separation of the (+)- and (−)-antipodes of the benzodioxane dimers *threo*/*erythro*-5OHG–G (**32**), *threo*-5OHG–S (**34**), and *threo*-5OHG–5OHG (**43**), whereas the 5OHG–5OHG–G benzodioxane trimer **46** was only partially separable under the conditions employed (data not shown). The chiral HPLC analysis of the three benzodioxane dimers **32**, **34** and **43** established the corresponding (+) and (−)-antipodes were present in an ~1 : 1 ratio, respectively.

Synthesis of lignin-derived thioacidolysis products of 35 and 36

Synthetic dimer **32** and trimer **46** were next subjected to thioacidolysis under conventional conditions,[38,45] to generate the benzodioxane dimer/trimer thioethylated products **35** and **36**, respectively (see Experimental section). In both cases, their structures were unambiguously established using different NMR and mass spectroscopic analyses, with these ultimately demonstrated as being identical to those obtained from the *Atomt1* mutant plant line as indicated below.

Specifically, LC-APCI-MS of synthetic **35** in the positive ion mode gave fragment ion peaks at *m*/*z* 478.9 $(M + H - HSEt)^+$ and 417.1 $(M + H - 2HSEt)^+$ (Fig. 7d) identical to the isolated **35** (Fig. 7b) corresponding to molecular formula of $C_{26}H_{36}O_6S_3$, whereas in the negative ion mode it gave a (M − H) ion at *m*/*z* 538.9. Upon comparison with the 5OHG–G benzodioxane diol **32**, its ^{1}H NMR spectrum indicated the absence of two *trans*-olefinic protons at the 7′, 8′ positions and an allylic hydroxyl group at the 9′ position, with the side-chain also being substituted with three thioethyl moieties (multiplets at δ 2.64–2.33 and 1.25–1.10 ppm), respectively. Again, the hydroxyl group at the 9-position was not thioethylated under the reaction conditions employed, as indicated earlier for the isolated dimer **35** and trimer **36**. The ^{1}H NMR spectrum of **35** contained five aromatic resonances at δ 7.10 (br s, 2-H), 6.96 (d, *J* 8.0, 6-H) and 6.88 (dd, *J* 1.5 and 8.5, 5-H) ppm for the G-unit and δ 6.78 (dd, *J* 2.0 and 7.0, 2′-H), 6.70 (ddd, *J* 1.5 and 6.5, 6′-H) ppm corresponding to the trithioethylated derivative of the 5OHG unit. Presence of a diastereomeric mixture for **35** in the ratio of ~1.7 : 1 (*threo*/*erythro*) was also evident from the resonances at δ 4.98 (d, *J* 7.5, 7-H), 4.97 (d, *J* 8.0, 7-H), 4.08-4.06 (m, 8-H) ppm and 3.79 (br d, *J* 2.5 and 12.5, 9-H), 3.50 (br d, *J* 3.5 and 12.0, 9-H) ppm, respectively.

LC-APCI-MS of synthetic **36** in the positive ion mode (Fig. 7c) gave fragment ions very similar to the isolated **36** (Fig. 7a) at *m*/*z* 672.9 $(M + H - HSEt)^+$ and 611.1 $(M + H - 2HSEt)^+$ corresponding to the molecular formula $C_{36}H_{46}O_{10}S_3$, as well as a (M − H) ion at *m*/*z* 733.1. The ^{1}H NMR spectrum of **36**, upon comparison with the 5OHG–5OHG–G triol **46**, once again demonstrated the absence of two *trans*-olefinic protons at the 7″, 8″ positions and an hydroxyl group at the 9″ position. As for compound **35**, these positions were substituted with three thioethyl moieties as indicated by multiplets at δ 2.63-2.33 and 1.24-1.09 ppm, respectively. Interestingly, the hydroxyl groups at both 9 and 9′-positions were again not thioethylated. The ^{1}H NMR spectrum of **36** also contained seven aromatic resonances at δ 7.12 (dd, *J* 1.5 and 5.5, 2-H), 6.97 (dd, *J* 1.5 and 8.0, 6-H) and 6.86 (dd, *J* 1.5 and 8.0, 5-H) ppm for the G-unit and δ 6.77 (br s, 2′-H and 6′-H), 6.719 (dd, *J* 1.5 and 2.0, 2″-H) and 6.717 (dd, *J* 2.0, 6″-H) ppm corresponding to the trithioethylated derivative of two 5OHG units, respectively. The presence of a diastereomeric mixture of **36** in the ratio of ~1.3 : 1 (*threo*/*erythro*) was also evident from the resonances at δ 5.0 (d, *J* 8.0, 7-H), 4.97 (d, *J* 7.5, 7′-H), 4.08 (m, 8-H, 8′-H) ppm and 3.81 (m, 9-H and 9′-H), 3.54 (m, 9-H and 9′-H) ppm. In this way, the *Atomt1* lignin-derived thioacidolysis products **35** and **36** were fully identified.

Time-course of thioacidolysis of 5OHG–G benzodioxane dimer (32) and 5OHG–5OHG–G benzodioxane trimer (46)

It was next instructive to examine the time course of thioacidolytic conversion of the corresponding lignin model compounds **32** and **46**, and the product distribution and yield (%) at different time points during treatment, taking into account that they contain a 7′–8′ unsaturated side-chain. Of particular interest was whether the 5OHG–5OHG subunit or the 5OHG–G subunit in trimer **46** underwent preferential cleavage, as this could provide useful insight into the nature of the chemistry of the *Atomt1* lignin biopolymer. In this regard, each model compound **32** and **46** was subjected to thioacidolysis for periods ranging from 0 to 5 h duration, with the various products quantified as shown in Fig. 8.

Monothioethylation. Prior to beginning the time course under standard thioacidolysis conditions, the 5OHG–G benzodioxane dimer **32** (3.1 μmol) was subjected to thioacidolysis at room temperature. Within 1 min of reaction time, the bulk of the starting material was converted into mono-substituted derivative **47** (1.11 μmol, ~34%). The LC-APCI-MS of **47** gave a (M − H) ion at *m*/*z* 417.0 corresponding to the molecular formula of $C_{22}H_{26}O_6S$, whereas its ^{1}H NMR spectrum upon comparison with the 5OHG–G benzodioxane diol (**32**) established the presence of *trans*-olefinic protons at 7′, 8′ positions and the absence of an hydroxyl group at the 9′ position which was instead substituted with a thioethyl moiety. The ^{1}H NMR spectrum of **47** also displayed five aromatic resonances at δ 7.11 (d, *J* 2.0, 2-H), 6.96 (dd, *J* 2.0 and 8.0, 6-H) and 6.88 (d, *J* 8.0, 5-H) ppm for the G-unit and δ 6.72 (d, *J* 1.5, 2′-H), 6.60 (d, *J* 1.5, 6′-H) ppm corresponding to the monothioethylated derivative of the 5OHG unit, respectively. The presence of a benzodioxane substructure was evident from resonances in a *trans*-conformation at δ 4.98 (d, *J* 8.0, 7-H) and 4.0 (m, 8-H) ppm, corresponding

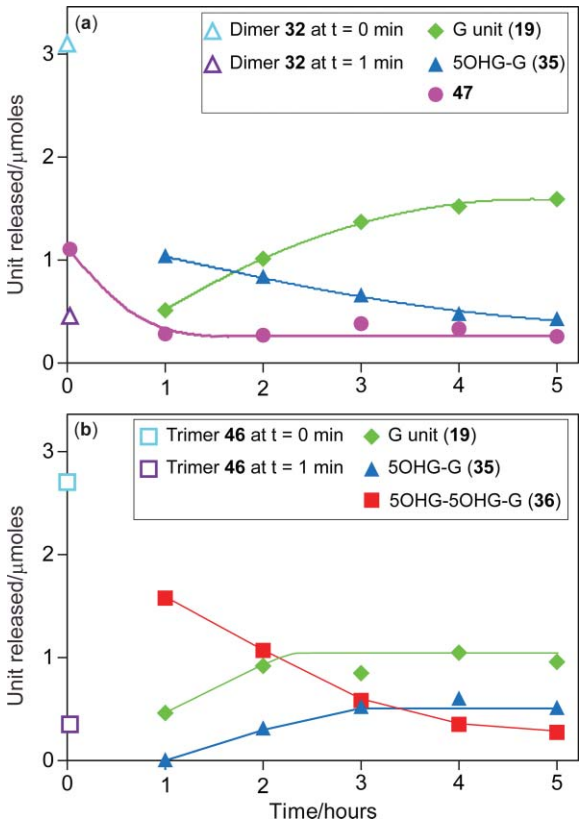


Fig. 8 Thioacidolysis treatment of synthetic (**a**) 5OHG–G benzodioxane dimer **32** and (**b**) 5OHG–5OHG–G benzodioxane trimer **46**.

to two vicinal aliphatic oxymethines and two geminal methylene protons at δ 3.78 and 3.50 (m, 9-H) ppm. Presence of two *trans*-olefinic protons was noted as doublets at δ 6.39 (d, J 15.5, 7′-H), 6.12 (dt, J 7.5 and 16.0, 8′-H) ppm, as well as a doublet at δ 3.30 (d, J 7.5) ppm for the 9′-H proton and a monothioethyl group from the proton resonances at δ 2.49 (q, J 7.5)/1.21 (t, J 6.5 and 7.5) ppm, respectively, for the 10′ and 11′-protons connected to a 5OHG unit. The two methoxyl groups resonated at δ 3.87 (s, 3-OMe) and 3.86 (s, 3′-OMe) ppm.

Thioacidolytic treatment of synthetic 5OHG–G benzodioxane dimer 32. After 1 min at elevated temperature and pressure (*i.e.* conventional thioacidolytic conditions), *circa* 85% of starting dimer **32** was consumed, with concomitant generation of monosubstituted derivative **47** (1.11 μmol, ~34%), based upon direct UPLC analysis and quantification of the reaction mixture. Within 1 h, the mono-substituted product **47** reached basal levels and remained essentially constant thereafter (Fig. 8a). On the other hand, dimer **35** increased in amount to its highest level (1.05 μmol, ~34%) by 1 h, with this then steadily decreasing until reaching a plateau level (0.44 μmol, ~14%) by 5 h. By contrast, formation of the monomeric G-unit (**19**) increased steadily until it also began to attain a plateau level (1.38 μmol, ~44.6%) by 5 h. The corresponding 5OHG thio-derivative **20**, on the other hand, was only detected in trace amounts (data not shown). Thus, taken together, thioacidolysis of dimer **32** resulted in an incomplete conversion of starting material to afford a mixture of products, and in so doing providing some insight into the dynamics which may also occur during lignin polymer thioacidolysis/deconstruction.

Thioacidolytic treatment of synthetic 5OHG–5OHG–G benzodioxane trimer 46. After thioacidolytic treatment of trimer **46** (2.71 μmol) for one minute only, the corresponding monothioethylated product **48** was formed (but not quantified) with only ~28% of trimer **46** remaining (Fig. 8b). Within 1 h, however, trimer **36** was mainly detected (1.58 μmol, 58.5%) together with smaller amounts of the monomeric G-unit **19** (0.45 μmol, ~16.5%) and traces of the 5OHG–G dimer **35**. The trimer **36** levels then began to decrease over a period of 5 h (to ~0.27 μmol, ~9.9%), with the 5OHG–G dimer **35** concomitantly increasing in amounts (0.51 μmol, ~19%). As before, the amounts of the G-monomer **19** increased until reaching a plateau level at 3–4 h (1.04 μmol, ~38.4%). Under the reaction conditions employed, neither the monomer 5OHG (**20**) nor the 5OHG–5OHG benzodioxane dimer **49** were detected. Thus, thioacidolysis of synthetic trimer **46** resulted only in incomplete conversion into products **19** and **35**.

Although thioacidolysis of the corresponding model compounds **32** and **46** currently imposes an analytical limitation in the overall quantification of the tri-substituted derived 5OHG unit (**20**), it interestingly provides evidence for differential (selective) cleavage of the benzodioxane functionality within the 5OHG–5OHG–G benzodioxane trimer **46** subunit, *i.e.* thereby providing evidence for the release of the corresponding 5OHG–G subunit (instead of the 5OHG–5OHG subunit). This result is thus in agreement with the thioacidolysis of the *Atomt1* lignin-derived-isolate, which accounted for isolation of compounds **19** (monomer), **35** (dimer) and **36** (trimer), and could be considered exclusively derived from 5OHG–5OHG–G trimer **46**. Thus, it can be provisionally concluded that a 5OHG–5OHG–G unit corresponds to the benzodioxane substructure (Fig. 9) in the lignin macromolecular assembly of the *Atomt1* mutant line, which upon deconstruction released quantifiable amounts of derived 5OHG–G benzodioxane dimer (**32**) and 5OHG–5OHG–G benzodioxane trimer (**46**), respectively. How this reflects the actual lignin deconstruction is not completely known at this point, *i.e.* given that the side-chain at C7–C8 is not substituted as is mainly the case for lignin *in situ*.

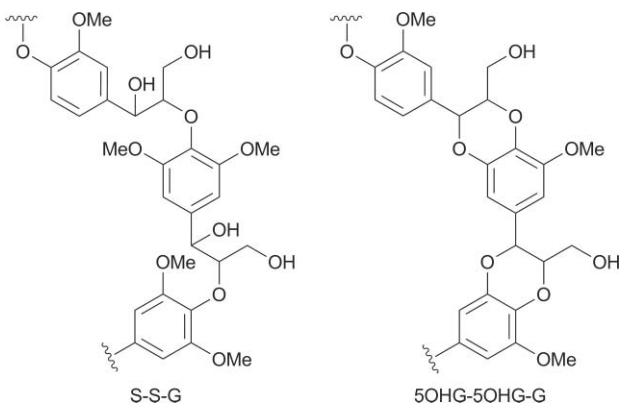


Fig. 9 Putative partial lignin primary structure in fibers from WT (S–S–G) and *Atomt1* (5OHG–5OHG–G) *Arabidopsis* lines.

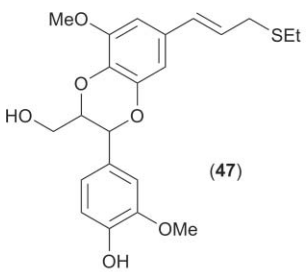
Arabidopsis COMT gene cloning, heterologous expression, purification and kinetic analyses

Analysis of the *Arabidopsis* genome indicated that 17 genes were annotated as putative COMTs.[46] Of these *Arabidopsis* homologues,

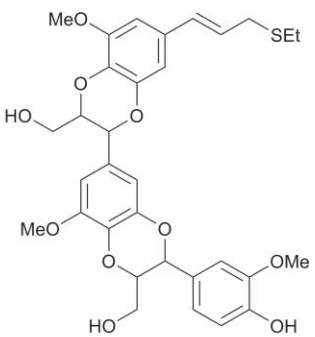
Table 2 Kinetic parameters of AtCOMT1 with various phenylpropanoids and quercetin (**50**)

Substrates	K_m/μM	V_{max}/pkat μg^{-1} protein	k_{cat}/s^{-1}	K_{enz}/M^{-1} s^{-1}
Caffeic acid (**5**)	24.2 ± 3.5	14.6 ± 1.3	0.62	26 000
5-OH Ferulic acid (**6**)	32.0 ± 6.9	30.1 ± 4.3	1.30	40 600
Caffeyl aldehyde (**9**)	19.7 ± 10.8	35.9 ± 10.3	1.55	78 700
5-OH coniferyl aldehyde (**10**)	17.9 ± 9.6	66.2 ± 17.1	2.85	159 000
Caffeyl alcohol (**11**)	51.5 ± 15.8	35.3 ± 5.6	1.52	29 500
5-OH coniferyl alcohol (**4**)	31.6 ± 8.4	51.9 ± 8.2	2.23	70 600
Quercetin (**50**)	23.7 ± 19.3	3.8 ± 2.4	0.16	6800

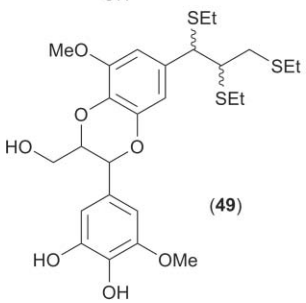
AtCOMT1 has the highest similarity (78.8%) and identity (75.5%) to COMTI from tobacco,[7] with the other homologues showing ~62–42% similarity and 54–33% identity to AtCOMT1. Twelve of these 17 COMTs were cloned (*AtCOMT1*, *3*, *4*, *6*, *8–13*, *15* and *17*), with the corresponding recombinant proteins expressed in *E. coli*, purified to homogeneity by affinity chromatography and biochemically characterized. Only AtCOMT1 displayed COMT activity proper *in vitro*: specifically, it had a very broad substrate specificity for a wide range of potential substrates (Table 2), with caffeyl (**9**) and 5-OH coniferyl (**10**) aldehydes being the preferred substrates, having K_{enz} values of 78 700 and 159 000 M^{-1} s^{-1}, respectively. Interestingly, AtCOMT1 was also previously described as involved in quercetin (**50**) to isorhamnetin (**51**) formation.[47] However, the kinetic data *in vitro* suggest that quercetin (**50**) is a much poorer substrate by at least an order of magnitude (*i.e.* K_{enz} = 6900 M^{-1} s^{-1} for **50** *versus* 159 000 M^{-1} s^{-1} for **10**). The remaining other 11 putative COMTs, all obtained in soluble recombinant protein forms, were unable to methylate any of the potential substrates used.



(**47**)



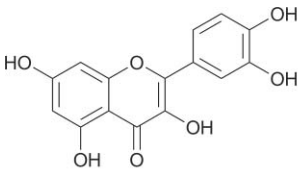
(**48**)



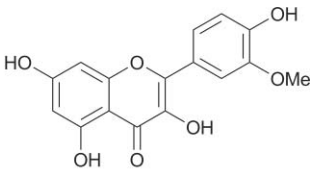
(**49**)

Conclusions

Lignin primary structure in COMT mutated lines produces benzodioxane substructures that are readily cleavable as 5OHG–G (**35**) and 5OHG–5OHG–G (**36**) deconstruction components, albeit where the overall 8–*O*–4′ inter-unit linkage frequency in the *Atomt1* knockout line has apparently been conserved relative to WT. Deconstruction and identification of the released subunits provide further support to lignin macromolecular assembly in *Arabidopsis* being conserved in a highly programmed (template guided) manner, with 5OHG units provisionally being 8–*O*–4′ linked to the primary lignin polymer *via* either one or two 5OHG units capped by a G-unit. It is significant that G-moieties apparently serve as either terminal units in the benzodioxane substructures of the polymer or as end-groups, *i.e.* rather than the previously contemplated/speculated benzodioxane substructures linked n-fold.[33]



Quercetin (**50**)



Isorhamnetin (**51**)

Experimental

Instrumentation and chromatography materials

All solvents and chemicals used were either reagent or HPLC grade and purchased from Sigma-Aldrich (St. Louis, MO, USA), unless otherwise stated. Chemical reactions were carried out under anhydrous conditions using dry freshly distilled solvents in an argon atmosphere. Silica gel thin-layer chromatography (TLC) employed Partisil® PK5F (Whatman, 20 × 20 cm, 1 mm, 150 Å) and AL SIL G/UV$_{254}$ (Whatman, 20 × 20 cm, 0.25 mm) whereas silica gel column chromatography (CC) utilized silica gel 60 (EM Science). UV and a 10% H_2SO_4 spray followed by heating were employed for TLC plate visualization. High-performance liquid chromatography (HPLC) analyses were performed using an Alliance 2690 system (Waters, Milford, MA) equipped with an UV-Vis diode array detector. Reversed-phase separations were carried out with a SymmetryShield RP$_{18}$ column (150 × 3.9 mm, Waters) with detection at 280 nm. Elution conditions at a flow rate of 1 cm^3 min^{-1} were: linear gradients of CH_3CN–H_2O (Optima LC/MS, Fischer Scientific) from 5 : 95 to 40 : 60 in 5 min, then to 80 : 20 in 40 min, to 95 : 5 in 2 min with this composition held for an additional 2 min, to 95 : 5 in 2 min, and finally to 5 : 95 in 2 min with this being held for 5 min. Chiral HPLC separations for determining enantiomeric compositions of diastereoisomers **32**, **34**, **43** and **46** employed a Chiralcel OJ column (250 × 4.6 mm, Chiral Technologies, Inc.) eluted with hexanes–EtOH (7 : 3) at a flow rate of 0.5 cm^3 min^{-1} for **32**/**34**/**43** and 0.3 cm^3 min^{-1} for **46**. Eluant was monitored at 280 nm, as well as with an in-line Advanced Laser Polarimeter detector (PDR-Chiral, Lake Park, Florida, USA, cell volume 56 μl, length 5.17 cm). Optical rotations were measured with a JASCO DIP-181 digital polarimeter at 20 °C (Hg-D line, 546 nm) using a quartz cell (length: 5 cm, volume 1 cm^3), with values given in 10^{-1} deg cm^2 g^{-1}.

NMR spectra were recorded on Varian Mercury-Vx 300, Inova 500 and Varian 600 MHz spectrometers operating at 300.1, 499.85 and 599.69 MHz (^{1}H), and at 75.5, 125.67 and 150.8 MHz (^{13}C), respectively, with *J* values given in Hz. Liquid

chromatography-mass spectrometry with an atmospheric pressure chemical ionization interface system (LC-APCI-MS) was performed using a Finnigan MAT (San Jose, CA, USA) ion trap mass spectrometer equipped with a Finnigan electrospray interface, with liquid chromatographic separations carried out on an ACQUITY Ultra Performance Liquid Chromatograph® (UPLC, Waters) system using a BEH C18 column (Waters, 2.1 × 50 mm) eluted as follows: flow rate of 0.3 cm^3 min^{-1}; linear gradients of CH_3CN-3% (v/v) AcOH in H_2O from 10 : 90 to 45 : 55 in 1.5 min, then to 60 : 40 in 5 min, to 80 : 20 in 2 min with this composition held for an additional 1 min, then to 100 : 0 for 1 min and finally to 10 : 90 in 0.5 min with this being held for 4 min. High resolution mass spectrometric analyses (HRMS) were performed at the University of Mississippi using an Agilent HPLC 1100 series equipped with a diode array detector and mass detector in series (Agilent Technologies, Palo Alto, CA, USA). The mass detector was a time of flight (Model G1969A) equipped with an electrospray ionization interface controlled by Agilent software (Agilent MassHunter Work Station, A.02.01). Data acquisition and processing were done using the AnalystTM QS software (Analyst, QS 1.1). All acquisitions were performed under negative ionization mode with a capillary voltage of 3000 V. Nitrogen was used as the nebulizer gas (30 psi) as well as the drying gas at 11 m^3 min^{-1} at a temperature of 350 °C. The voltage of the PMT, fragmentor and skimmer was set at 850, 175 and 60 V, respectively. Full scan mass spectra were acquired from *m/z* 100–1300.

Plant material and growth parameters

Seeds from the *Atomt1* knockout mutant line, provided by Dr. Lise Jouanin [Institut National de la Recherche Agronomique (INRA), Versailles, France], were first grown on MS medium containing kanamycin (50 μg cm^{-3}). Transformants were next transferred to soil and grown in Washington State University greenhouses as previously described[27] until maturation with the seeds next harvested. These were then subjected to another round of kanamycin screening with seedlings also subjected to GUS screening as described.[48]

Both wild type [ecotype Wassilewskija (Ws)] and the selected *Atomt1* lines were grown. All lines were harvested weekly from 3 weeks until 10 weeks, with growth and development parameters of the main (bolting) stem also measured (*i.e.* length and basal diameters) using twenty randomly chosen plants/harvest sampling point.

Lignin analysis: acetyl bromide, alkaline nitrobenzene oxidation and thioacidolysis estimations

Extractive-free cell wall residues (CWR) were prepared as previously described,[27] with lignin contents next estimated by the acetyl bromide method[38,49] using modified absorptivity values for specific lignin type *i.e.* H, G or S-enriched.[27] Monomeric compositions were estimated using both alkaline nitrobenzene oxidation[50] and thioacidolysis[38,45,51] methods, with all products analyzed using an HP 6890 Series GC System equipped with a RESTEK-5Sil-MS (30 m × 0.25 mm × 0.25 μm) column. Both alkaline nitrobenzene oxidation and thioacidolysis products were identified by comparison/calibration against authentic silylated standards and by mass spectrometric analyses using a HP 5973 MS detector (EI mode, 70 eV).

Chemical syntheses

Coniferyl (2), sinapyl (3) and 5-hydroxyconiferyl (4) alcohols, 5-hydroxyferulic acid (6), caffeyl (9) and 5-hydroxyconiferyl (10) aldehydes, methyl 5-hydroxyferulate (37) and methyl *O*-*tert*-butyldimethylsilyl ferulate (38a/b). These were synthesized as previously described in Kim *et al.*[52]

***O*-*tert*-Butyldimethylsilyl coniferyl alcohols (41a/b).** To an ice-cold solution of **38a/b** (2 g, 5.9 mmol) in dry toluene (40 cm^3) was added 1 M diisobutylaluminium hydride (DIBAL) in toluene (18 cm^3, 17.7 mmol). After stirring the contents for 1 h, excess DIBAL was destroyed by drop-wise addition of MeOH (10 cm^3). The reaction mixture was then extracted with EtOAc (2 × 300 cm^3), washed with brine (1 × 100 cm^3) and dried (Na_2SO_4). Following evaporation of the EtOAc solubles to dryness *in vacuo*, the residue so obtained was subjected to silica gel CC eluted with hexanes–EtOAc (4 : 1) to afford a mixture of *O*-*tert*-butyldimethylsilyl coniferyl alcohols (**41a/b**) which was used without further purification (2.5 : 1 ratio, 1.7 g, 5.4 mmol, 90%). δ_H(500 MHz; acetone-d_6; Me_4Si) 7.45 (1 H, ArOH, **41a**), 7.21 (1 H, ArOH, **41b**), 6.73 (1 H, d, *J* 1.5, 2-H, **41a**), 6.61 (1 H, dd, *J* 1.5 and 4.0, 2-H, **41b**), 6.60 (1 H, dd, *J* 1.5 and 4.0, 6-H, **41b**), 6.58 (1 H, d, *J* 2.0, 6-H, **41a**), 6.45 (1 H, d, *J* 16.0, 7-H, **41a/b**), 6.2 (1 H, dt, *J* 5.5 and 16.0, 8-H, **41a/b**), 4.19 (2 H, br d, *J* 5.0, 9-H, **41a/b**), 3.85 (3 H, s, 3-OMe, **41a**), 3.82 (3 H, s, 3-OMe, **41b**), 1.0 (18 H, s, 2 × 3Me-TBS, **41a/b**), 0.2 (6 H, s, 2Me-TBS, **41a**), 0.17 (6 H, s, 2Me-TBS, **41b**); δ_C(125 MHz; acetone-d_6; Me_4Si) 152.38, 149.52, 149.39, 144.17, 138.17, 133.12, 131.35, 130.40 (C-7, **41b**), 130.33 (C-7, **41a**), 129.17, 129.07, 128.44, 113.16 (C-6, **41a**), 107.9 (C-2, **41b**), 103.9 (C-2, **41a**), 102.62 (C-6, **41b**), 63.37 (C-9, **41b**), 63.30 (C-9, **41a**), 56.44 (C-3-OMe, **41b**), 55.74 (C-3-OMe, **41a**), 26.30 (C-3Me-TBS, **41a**), 26.17 (C-3Me-TBS, **41b**), 19.3, 19.0, –4.12, –4.21; *m/z* (LC-APCI) 309.2 (M – H. $C_{16}H_{25}O_4Si$ requires 309.2) (t_R = 4.3 min).

***threo*/*erythro*-5OHG–G benzodioxane dimer (32).** To a mixture of methyl 5-hydroxyferulate (**37**, 0.5 g, 2.23 mmol) and coniferyl alcohol (**2**, 0.4 g, 2.22 mmol) in benzene–acetone (2 : 1, 60 cm^3) was added Ag_2CO_3 (1.5 g, 5.45 mmol).[39] The contents were then stirred at room temperature under N_2 for about 11 h. After completion of the reaction, as monitored by TLC (hexanes–EtOAc 3 : 7), the Ag_2CO_3 was removed by filtration through a Celite pad with the latter washed with acetone (20 cm^3). The resulting filtrate was evaporated to dryness *in vacuo*, with the residue purified further by silica gel CC using hexanes–EtOAc (3 : 7) as eluant to afford *threo*/*erythro*-5OHG–G benzodioxane methyl ester (~9 : 1 ratio) [(*E*)-methyl 3-(3-(4-hydroxy-3-methoxyphenyl)-2-(hydroxymethyl)-8-methoxy-2,3-dihydrobenzo[*b*][1,4]dioxin-6-yl)acrylate] which was used without further purification (**39**, 0.55 g, 1.36 mmol, 61%).[39,53–55] δ_H(500 MHz; acetone-d_6) 7.8 (1 H, br s, ArOH), 7.55 (1 H, d, *J* 16.0, H-7′), 7.12 (1 H, d, *J* 2.0, 2-H), 6.98 (1 H, dd, *J* 2.0 and 8.0, 6-H), 6.96 (1 H, d, *J* 2.0, 2′-H), 6.89 (1 H, d, *J* 8.5, 5-H), 6.88 (1 H, d, *J* 1.5, 6′-H), 6.43 (1 H, d, *J* 16.0, 8′-H), 5.3 (1 H, d, *J* 2.5, 7-$H_{erythro}$), 5.0 (1 H, d, *J* 8.0, 7-H_{threo}), 4.12 (1 H, m, 8-H), 3.90 (3 H, s, 3-OMe), 3.87 (3 H, s, 3′-OMe), 3.81 (1 H, br d,

J 12.25, 9-H), 3.72 (3 H, s, CO*OMe*), 3.52 (1 H, br d, *J* 12.5, 9-H); δ_C(125 MHz; acetone-d_6) 167.8 (C-9′ *CO*OMe), 150.38 (C-3′), 148.55 (C-3), 148.1 (C-4), 145.6 (C-5′), 145.6 (C-7′), 136.78 (C-4′), 129.0 (C-1′), 127.5 (C-1), 121.68 (C-6), 116.8 (C-8′), 115.8 (C-5), 112.0 (C-2), 111.29 (C-6′), 105.1 (C-2′), 79.7 (C-8), 77.0 (C-7), 61.74 (C-9), 56.5 (C-3-OMe), 56.4 (C-3′-OMe), 51.6 (CO-*OMe*); *m*/*z* (LC-APCI) 401.2 (M − H. $C_{21}H_{21}O_8$ requires 401.1) (t_R = 2.2 min).

To a solution of *threo*/*erythro*-**39** (0.50 g, 1.2 mmol) in toluene (20 cm^3) was added DIBAL (1 M in toluene, 4 cm^3) at 0 °C under N_2 for about 1 h, with the excess DIBAL destroyed by adding MeOH (4 cm^3) dropwise. The resulting suspension was then extracted with excess EtOAc (2 × 200 cm^3), washed with dil. HCl solution (1 × 100 cm^3), brine (1 × 100 cm^3) and dried (Na_2SO_4). The combined EtOAc solubles were then evaporated *in vacuo* with the corresponding residue purified by silica gel CC using EtOAc–acetone (9 : 1) as eluant to afford *threo*/*erythro*-5OHG–G benzodioxane dimer (**32**) in ~9 : 1 ratio (0.396 g, 1.05 mmol, 88%).[39] The ^{1}H and ^{13}C NMR spectra of *threo*/*erythro*-**32** [(*E*)-4-(3-(hydroxymethyl)-7-(3-hydroxyprop-1-enyl)-5-methoxy-2,3-dihydrobenzo[*b*][1,4]dioxin-2-yl)-2-methoxyphenol] were in close agreement with the reported literature data.[39] (**32**) *m*/*z* (LC-APCI) 373.0 (M − H. $C_{20}H_{21}O_7$ requires 373.1), (t_R = 2.05 min); HRMS (EI) 397.1259 (M + Na. $C_{20}H_{22}O_7Na$ requires 397.1263). Chiral HPLC separation of (±)-*threo*/*erythro*-**32** was achieved as described above in the 'Instrumentation and Chromatography Materials' section resulted in the separation of (−)-*threo*-**32** [t_R = 46.4 min; $[\alpha]_D^{20}$ −9.39 (*c* 0.00149 in MeOH)], (+)-*threo*-**32** [t_R = 57.7 min; $[\alpha]_D^{20}$ +8.39 (*c* 0.00143 in MeOH)] and (±)-*erythro*-**32** [t_R = 75.1 min]. The ^{1}H NMR and mass spectra of (+)- and (−)-*threo*-**32** were identical with (±)-*threo*-**32** above.

(±)-*erythro*-**32**: δ_H(500 MHz; acetone-d_6) 7.70 (1 H, br s, ArOH), 7.0 (1 H, d, *J* 2.0, 2-H), 6.88 (1 H, dd, *J* 2.0 and 8.0, 6-H), 6.83 (1 H, d, *J* 8.0, 5-H), 6.72 (1 H, d, *J* 2.0, 2′-H), 6.62 (1 H, d, *J* 1.5, 6′-H), 6.48 (1 H, d, *J* 16.0, 7′-H), 6.28 (1 H, dt, *J* 5.5 and 16.0, 8′-H), 5.26 (1 H, d, *J* 2.5, 7-H$_{erythro}$), 4.48 (1 H, m, 8-H), 4.20 (2 H, dd, *J* 1.0 and 5.5, 9′-H), 3.85 (3 H, s, 3-OMe), 3.80 (3 H, s, 3′-OMe), 3.64 and 3.48 (2 H, m, 9-H); δ_C(150 MHz; $CDCl_3$, Shigemi NMR tube) 149.5 (C-3′), 147.0 (C-3), 145.9 (C-4), 144.0 (C-5′), 131.06 (C-4′), 131.01 (C-1′), 129.9 (C-7′), 127.7 (C-1), 127.6 (C-8′), 119.2 (C-6), 114.8 (C-5), 108.6 (C-2), 108.5 (C-6′), 103.0 (C-2′), 78.0 (C-8), 75.57 (C-7), 63.94 (C-9′), 59.1 (C-9), 56.3 (C-3-OMe), 56.2 (C-3′-OMe); *m*/*z* (LC-APCI) 373.1 (M − H. $C_{20}H_{21}O_7$ requires 373.1) (t_R = 1.95 min).

***threo*-5OHG–S benzodioxane dimer (34).** The *threo*-5OHG–S benzodioxane dimer **34** was prepared exactly as for compound **32** using the Ag_2CO_3 (0.7 g, 2.52 mmol) coupling of methyl 5-hydroxy ferulate (**37**, 0.25 g, 1.11 mmol) and sinapyl alcohol (**3**, 0.2 g, 0.95 mmol) to initially afford the methyl ester of the 5OHG–S benzodioxane dimer (**40**, 0.26 g, 0.6 mmol, 63%). The latter (0.20 g, 0.46 mmol) was then reduced and purified as described above for compound **39** by addition of DIBAL (1 M in toluene, 1.5 cm^3) to afford the *threo*-5OHG–S benzodioxane dimer (**34**, 0.16 g, 0.39 mmol, 85%).

threo-5OHG–S benzodioxane methyl ester [(*E*)-methyl 3-(3-(4-hydroxy-3,5-dimethoxyphenyl)-2-(hydroxymethyl)-8-methoxy-2,3-dihydrobenzo[*b*][1,4]dioxin-6-yl)acrylate] (**40**). δ_H(600 MHz; acetone-d_6 and 3 drops CD_3OD): 7.55 (1 H, d, *J* 16.2, 7′-H), 6.98 (1 H, d, *J* 1.8, 2′-H), 6.88 (1 H, d, *J* 1.8, 6′-H), 6.82 (2 H, s, 2-H and 6-H), 6.43 (1 H, d, *J* 16.2, 8′-H), 4.97 (1 H, d, *J* 7.8, 7-H$_{threo}$), 4.1 (1 H, m, 8-H), 3.91 (6 H, s, 3′-OMe), 3.84 (3 H, s, 3-OMe, 5-OMe), 3.78 (1 H, dd, *J* 2.4 and 12.6, 9-H), 3.72 (3 H, s, COOMe), 3.52 (1 H, dd, *J* 4.2 and 12.6, 9-H); δ_C(150 MHz; acetone-d_6 and 3 drops CD_3OD): 167.9 (C-9′, *CO*OMe), 150.4 (C-3′), 148.8 (C-3 and C-5), 145.64 (C-5′), 145.61 (C-7′), 137.38 (C-4), 136.8 (C-4′), 127.9 (C-1′), 127.6 (C-1), 116.8 (C-8′), 111.3 (C-6′), 106.2 (C-2 and C-6), 105.1 (C-2′), 79.8 (C-8), 77.4 (C-7), 61.6 (C-9), 56.77 (C-3-OMe and C-5-OMe), 56.56 (C-3′-OMe), 51.7 (CO-*OMe*); *m*/*z* (LC-APCI) 431.1 (M − H. $C_{22}H_{23}O_9$ requires 431.1) (t_R = 2.5 min).

The ^{1}H and ^{13}C NMR spectra of (±)-*threo*-5OHG-S benzodioxane diol [(*E*)-4-(3-(hydroxymethyl)-7-(3-hydroxyprop-1-enyl)-5-methoxy-2,3-dihydrobenzo[*b*][1,4]dioxin-2-yl)-2,6-dimethoxyphenol] (**34**) were in close agreement with the reported literature data:[40–43] *m*/*z* (LC-APCI) 403.0 (M − H. $C_{21}H_{23}O_8$ requires 403.1 (t_R = 2.0 min); HRMS (EI) 427.1400 (M + Na. $C_{21}H_{24}O_8$ Na requires 427.1369). Chiral separation: (−)-*threo* **34**: t_R = 77.8 min, $[\alpha]_D^{20}$ −2.96 (*c* 0.00135 in MeOH); (+)-*threo* **34**: t_R = 100.6 min, $[\alpha]_D^{20}$ +2.48 (*c* 0.00161 in MeOH). The ^{1}H NMR and mass spectra of (+)-and (−)-*threo*-**34** were identical with (±)-*threo*-**34**.

5OHG–5OHG benzodioxane dimer (43). Three different synthetic protocols were employed.

Protocol 1. To a solution of methyl 5-hydroxyferulate (**37**, 0.8 g, 3.57 mmol) and *O*-*tert*-butyldimethylsilyl coniferyl alcohols (**41a**/**b**, 1.3 g, 4.19 mmol) in benzene–acetone (2 : 1) was added Ag_2CO_3 (2.7 g, 9.8 mmol) with continuous stirring at room temperature for 15 h under a N_2 atmosphere. After completion of the reaction, the Ag_2CO_3 was removed by filtration through a Celite pad with the latter washed with acetone (40 cm^3). The resulting filtrate was evaporated to dryness *in vacuo*, with the residue **42a**/**b** (1.30 g, 2.4 mmol, 57%) so obtained reduced directly with DIBAL (1 M in toluene, 11 cm^3) at 0 °C under N_2 for about 1 h. The resulting suspension was then extracted with excess EtOAc (2 × 400 cm^3), washed with dil. HCl solution (1 × 150 cm^3), brine (1 × 150 cm^3) and dried (Na_2SO_4). The combined EtOAc solubles were next evaporated *in vacuo* with the corresponding residue so obtained latter deprotected with tetrabutylammonium fluoride (TBAF)[52] (1 M in THF, 5 cm^3) and purified by silica gel CC eluted with EtOAc–acetone (8 : 2), to afford the (±)-*threo*-5OHG–5OHG benzodioxane dimer (**43**, 0.77g, 1.97 mmol, 82%) [(*E*)-5-(3-(hydroxymethyl)-7-(3-hydroxyprop-1-enyl)-5-methoxy-2,3-dihydrobenzo[*b*][1,4]dioxin-2-yl)-3-methoxybenzene-1,2-diol]:[40,43] δ_H(500 MHz; acetone-d_6): 7.67 (2 H, br s, ArOH), 6.7 (1 H, d, *J* 1.5, 2′-H), 6.67 (1 H, d, *J* 1.5, 6′-H), 6.65 (1 H, d, *J* 1.5, 2-H), 6.59 (1 H, d, *J* 2.0, 6-H), 6.47 (1 H, d, *J* 16.0, 7′-H), 6.27 (1 H, dt, *J* 5.5 and 15.5, 8′-H), 4.9 (1 H, d, *J* 8.0, 7-H$_{threo}$), 4.20 (2 H, d, *J* 5.0, 9′-H), 4.0 (1 H, m, 8-H), 3.85 (3 H, s, 3′-OMe), 3.83 (3 H, s, 3-OMe), 3.78 (1 H, dd, *J* 2.0 and 12.25, 9-H), 3.52 (1 H, dd, *J* 3.5 and 12.5, 9-H); δ_C(125 MHz; acetone-d_6) 150.15 (C-3′), 149.1 (C-3), 146.3 (C-5), 145.5 (C-5′), 135.2 (C-4), 134.1 (C-1′), 130.6 (C-7′), 130.1 (C-4′), 129.4 (C-1), 128.8 (C-8′), 109.39 (C-6), 108.7 (C-6′), 103.86 (C-2), 103.56 (C-2′), 79.5 (C-8), 77.2 (C-7), 63.3 (C-9′), 61.8 (C-9), 56.6 (C-3-OMe), 56.3 (C-3′-OMe); *m*/*z* (LC-APCI) 388.9 (M − H. $C_{20}H_{21}O_8$ requires 389.1) (t_R = 2.4 min); HRMS (EI) 413.1207 (M + Na. $C_{20}H_{22}O_8Na$ requires 413.1212). Chiral HPLC analysis (see Instrumentation

and Chromatography Materials section) resulted in the separation of (–)-*threo*-**43** [t_R = 84.2 min; $[\alpha]_D^{20}$ –4.7 (*c* 0.00234 in MeOH)] and (+)-*threo*-**43** [t_R = 115.1 min; $[\alpha]_D^{20}$ +4.63 (*c* 0.00216 in MeOH)]. The ^{1}H NMR and mass spectra of (+)-*threo*-**43** and (–)-*threo*-**43** were identical with (±)-*threo*-**43**.

Protocol 2. To a solution of 5-hydroxyconiferyl alcohol (**4**, 53 mg, 0.27 mmol) in benzene and acetone (2 : 1, 15 cm^3) was added Ag_2CO_3 (90 mg, 0.32 mmol), with the suspension then stirred at room temperature under N_2 for 3 h. The contents were next filtered over a Celite pad, washed with acetone (20 cm^3), and dried *in vacuo*. The residue so obtained was subjected to preparative silica gel TLC (eluant, $CHCl_3$–MeOH, 9 : 1) to afford (±)-*threo*-5OHG–5OHG (**43**) (4.5 mg, 0.011 mmol, 4.1%): The ^{1}H and ^{13}C NMR spectra of (±)-*threo*-**43** were identical to that described in *Protocol 1*; *m*/*z* (LC-APCI) 388.9 (M – H. $C_{20}H_{21}O_8$ requires 389.1) (t_R = 2.4 min). Chiral HPLC analysis; (–)-*threo*-**43**: t_R = 84.0 and (+)-*threo*-**43**: t_R = 114.5 min.

Protocol 3: enzymatic synthesis. To 5-hydroxyconiferyl alcohol (**4**, 55 mg, 0.28 mmol) was added acetone (20 cm^3) and 0.2 M citric acid/phosphate buffer (pH 6.0, 5 cm^3) with the contents cooled to 0 °C.[56] Two drops of 30% H_2O_2 and 7.3 mg of horseradish peroxidase type (II) (1100 units mg^{-1}) in H_2O (3 cm^3) were then added every 15 min, with the mixture stirred at 0 °C for 2 h. The progress of the reaction was monitored by reversed-phase HPLC (see Instrumentation and Chromatography Materials section). The reaction mixture was then warmed to room temperature, extracted with EtOAc (2 × 50 cm^3), washed with brine (1 × 25 cm^3), dried (Na_2SO_4) and evaporated to dryness *in vacuo*. The residue so obtained was subjected to preparative silica gel TLC eluted with $CHCl_3$–MeOH (9 : 1), to afford 5OHG–5OHG benzodioxane diol: (±)-*threo*-**43** (9.8 mg, 0.024 mmol, 8.6%) and (±)-*erythro*-**43** (1.2 mg, 0.003 mmol, 1.1%), respectively.

(±)-*threo*-**43**: the ^{1}H and ^{13}C NMR spectra were identical to that described in *Protocol 1*; *m*/*z* (LC-APCI) 388.9 (M – H. $C_{20}H_{21}O_8$ requires 389.1) (t_R = 2.4 min). Chiral HPLC analysis; (–)-*threo*-**43**: t_R = 84.6 and (+)-*threo*-**43**: 115.1 min.

(±)-*erythro*-**43**: δ_H(500 MHz; acetone-d_6) 7.61 (2 H, br s, ArOH), 6.71 (1 H, d, *J* 1.5, 2′-H), 6.62 (1 H, d, *J* 1.5, 6′-H), 6.62 (1 H, d, *J* 2.0, 2-H), 6.59 (1 H, d, *J* 1.5, 6-H), 6.49 (1 H, d, *J* 16.0, 7′-H), 6.28 (1 H, dt, *J* 5.5 and 16.0, 8′-H), 5.2 (1 H, d, *J* 3.0, 7-H), 4.46 (1 H, m, 8-H), 4.2 (2 H, d, *J* 4.5, 9′-H), 3.85 (3 H, s, 3′-OMe), 3.79 (3 H, s, 3-OMe), 3.62 (1 H, br dd, *J* 5.0, 7.5 and 12.0, 9-H), 3.49 (1 H, dd, *J* 5.0 and 12.0, 9-H); δ_C(125 MHz; acetone-d_6; Me_4Si) 150.4 (C-3′), 148.86 (C-3), 146.32 (C-5), 144.91 (C-5′), 134.6 (C-4), 132.84 (C-1′), 130.9 (C-7′), 130.0 (C-4′), 129.5 (C-1), 128.4 (C-8′), 108.66 (C-6), 108.1 (C-6′), 103.85 (C-2), 103.1 (C-2′), 78.48 (C-8), 76.56 (C-7), 63.26 (C-9′), 59.88 (C-9), 56.47 (C-3-OMe), 56.41 (C-3′-OMe); *m*/*z* (LC-APCI) 388.9 (M – H. $C_{20}H_{21}O_8$ requires 389.1) (t_R = 2.4 min). Chiral HPLC analysis; (±)-*erythro*-**43**: t_R = 83.6 and 90.6 min, respectively.

5OHG–5OHG–G benzodioxane trimer (46). To a solution of **42a/b** (1 g, 1.87 mmol) in dry THF (30 cm^3) was added TBAF (1 M in THF, 4 cm^3),[52] with the contents allowed to stir under N_2 in an ice-bath for 30 min. After completion of the reaction, the reaction mixture was quenched with saturated aqueous NH_4Cl (100 cm^3), with the whole extracted with EtOAc (3 × 100 cm^3). The combined organic fractions were washed with brine (1 × 100 cm^3), dried (Na_2SO_4) and evaporated to dryness *in vacuo*. The residue (~ 0.9 g) was purified by silica gel CC using hexanes–ethyl acetate (2 : 8) as eluant to afford *threo*/*erythro*-5OHG–5OHG benzodioxane methyl ester in 9.5 : 0.5 ratio (**44**, 0.69 g, 1.65 mmol, 88%). Further, to a solution of ester **44** (0.68 g, 1.62 mmol) in benzene–acetone (2 : 1) was added coniferyl alcohol (**2**, 0.3 g, 1.66 mmol) and Ag_2CO_3 (1.14 g, 4.15 mmol). After completion of the reaction, the reaction mixture was purified by silica gel CC using hexanes–EtOAc (1 : 9) as eluant to afford the methyl ester of 5OHG–5OHG–G benzodioxane trimer (**45**, 0.58 g, 0.97 mmol, 58%). A portion of the latter ester **45** (0.3 g, 0.5 mmol) was next reduced with DIBAL (2 cm^3) to afford 5OHG–5OHG–G benzodioxane trimer **46** (0.22 g, 0.40 mmol, 80%).

threo/*erythro*-5OHG–5OHG benzodioxane methyl ester [(*E*)-methyl 3-(3-(3,4-dihydroxy-5-methoxyphenyl)-2-(hydroxymethyl)-8-methoxy-2,3-dihydrobenzo[*b*][1,4]dioxin-6-yl-)-acrylate] (**44**): δ_H(500 MHz; acetone-d_6; Me_4Si) 7.68 (2 H, br s, ArOH), 7.56 (1 H, d, *J* 16.0, 7′-H), 6.98 (1 H, d, *J* 1.5, 2′-H), 6.89 (1 H, d, *J* 2.0, 6′-H), 6.7 (1 H, d, *J* 2.0, 2-H), 6.67 (1 H, d, *J* 2.0, 6-H), 6.44 (1 H, d, *J* 16.0, 8′-H), 5.24 (1 H, d, *J* 2.5, 7-$H_{erythro}$), 4.94 (1 H, d, *J* 8.0, 7-H_{threo}), 4.09 (1 H, m, 8-H), 3.91 (3 H, s, 3′-OMe), 3.84 (3 H, s, 3-OMe), 3.83 (1 H, m, 9-H), 3.73 (3 H, s, COOMe), 3.54 (1 H, br dd, *J* 4.5 and 12.5, 9-H); δ_C(125 MHz; acetone-d_6; Me_4Si) 167.8 (C-9′, *CO*OMe), 150.4 (C-3′), 149.1 (C-3), 146.37 (C-5), 145.6 (C-5′), 145.61 (C-7′), 136.8 (C-4), 135.3 (C-4′), 128.47 (C-1′), 127.5 (C-1), 116.8 (C-8′), 111.3 (C-6), 109.43 (C-6′), 105.15 (C-2), 103.9 (C-2′), 79.8 (C-8), 77.21 (C-7), 61.73 (C-9), 56.63 (C-3-OMe), 56.60 (C-3′-OMe), 51.6 (CO-*OMe*); *m*/*z* (LC-APCI) 416.9 (M – H. $C_{21}H_{21}O_9$ requires 417.1) (t_R = 2.2 min).

5OHG–5OHG–G benzodioxane methyl ester [(*E*)-methyl 3-(3′-(4-hydroxy-3-methoxyphenyl)-2′,3-bis(hydroxymethyl)-5,8′-dimethoxy-2,2′,3,3′-tetrahydro-2,6′-bibenzo[*b*][1,4]dioxin-7-yl)-acrylate] (**45**): δ_H(500 MHz; acetone-d_6; Me_4Si) 7.77 (1 H, br s, ArOH), 7.56 (1 H, d, *J* 16.0, 7″-H), 7.13 (1 H, dd, *J* 1.5 and 4.5, 2-H), 6.98 (1 H, dd, *J* 1.5 and 8.25, 6-H), 6.89 (1 H, dd, *J* 3.5 and 8.0, 5-H), 6.80 (2 H, br d, *J* 1.0, 2′-H and 6′-H), 6.73 (1 H, d, *J* 2.0, 2″-H), 6.72 (1 H, d, *J* 2.0, 6″-H), 6.43 (1 H, dd, *J* 1.5 and 16.0, 8″-H), 5.0 (2 H, d, *J* 8.0, 7-H and 7′-H), 4.13 (2 H, m, 8-H and 8′-H), 3.91 (3 H, s, 3-OMe), 3.88 (3 H, s, 3′-OMe), 3.87 (3 H, s, 3″-OMe), 3.82 (2 H, dd, *J* 3.5 and 12.5, 9-H and 9′-H), 3.72 (3 H, s, COOMe), 3.54 (2 H, dd, *J* 3.0 and 12.5, 9-H and 9′-H); δ_C(125 MHz; acetone-d_6; Me_4Si) 167.7 (C-9″, *CO*OMe), 150.3, 150.10 (d), 148.47 (d), 148.0, 145.48 (d), 145.40 (d), 136.6, 134.82 (d), 129.54 (d), 129.18, 127.5, 121.59 (d), 116.8, 115.74 (d), 111.90 (d), 111.2, 109.9 (d), 105.10 (d), 105.0, 79.53 (d), 79.47 (d), 77.09 (d), 76.85 (d), 61.76, 61.63, 56.45 (br s), 56.34 (d), 51.59 (CO*OMe*); *m*/*z* (LC-APCI) 595.1 (M – H. $C_{31}H_{31}O_{12}$ requires 595.2) (t_R = 2.5 min).

5OHG-5OHG-G benzodioxane triol [(*E*)-(3′-(4-hydroxy-3-methoxyphenyl)-7-(3-hydroxyprop-1-enyl)-5,8′-dimethoxy-2,2′,3,3′-tetrahydro-2,6′-bibenzo[*b*][1,4]dioxine-2′,3diyl-)dimethanol] (**46**):[44] δ_H(500 MHz; acetone-d_6): 7.77 (1 H, br s, ArOH), 7.12 (1 H, dd, *J* 2.0 and 4.5, 2-H), 6.97 (1 H, d, *J* 8.0, 6-H), 6.88 (1 H, d, *J* 8.0, 5-H), 6.78 (1 H, br s, 2′-H), 6.71 (1 H, dd, *J* 2.0 and 4.5, 2″-H), 6.70 (1 H, dd, *J* 2.0 and 4.5, 6″-H), 6.60 (1 H, d, *J* 2.0, 6′-H), 6.47 (1H, d, *J* 16.0, 7″-H), 6.27 (1 H, dt, *J* 5.5 and 16.0, 8″-H), 4.99 (1 H, d, *J* 8.0, 7-H), 4.97 (1 H, d, *J* 7.5, 7′-H), 4.20 (2 H, t, *J* 5.0, 9′-H), 4.08 (2 H, m, 8-H and 8′-H), 3.88 (3 H, s, 3-OMe), 3.87 (3 H, s, 3′-OMe), 3.86 (3 H, s, 3″-OMe), 3.82 (2 H, dd, *J* 4.5 and 12.0, 9-H and 9′-H), 3.54 (2 H, dd, *J* 5.5 and 12.5, 9-H and 9′-H); δ_C(125 MHz;

acetone-d_6): 150.07, 148.48, 148.47, 148.0, 145.43, 145.40, 145.26, 134.73, 134.71, 134.0, 130.6, 130.0, 129.91, 129.88, 129.4, 129.2, 121.60 (d), 115.75, 115.73, 111.91, 111.88, 109.87, 108.6 (br d), 105.07, 104.96, 103.5, 79.47, 79.45, 79.28, 79.24, 77.11, 77.06, 76.89, 76.81, 63.27, 61.77 (d), 56.44, 56.34 (d), 56.28; *m/z* (LC-APCI) 567.0 (M − H. $C_{30}H_{31}O_{11}$ requires 567.2) (t_R = 2.0 min); HRMS (EI) 591.1846 (M + Na. $C_{30}H_{32}O_{11}Na$ requires 591.1842). Chiral HPLC analysis; t_R = 162.6, 192.0 (br) min and 254.2, 292.85 (br) min respectively; the enantiomers were not baseline separable.

***threo* and *erythro* 5OHG–G benzodioxane triethanethiol.** [4-(3-(Hydroxymethyl)-5-methoxy-7-(1,2,3-tris(ethylthio)propyl)-2,3-dihydrobenzo[*b*][1,4]dioxin-2-yl)-2-methoxyphenol] (**35**). Thioacidolysis of compound **32** was performed using the procedure described earlier[38,45] with the following modifications. 5OHG–G benzodioxane dimer **32** (5 × 5 mg each, 0.013 mmol each) was dissolved in dry dioxane (1 cm^3) of which an aliquot (100 mm^3) was added to a 20 cm^3 glass tube fitted with a Teflon-lined screwcap containing the thioacidolysis reagent solution [15 cm^3, 0.2 M $BF_3{\cdot}Et_2O$ in dioxane–ethanethiol (8.75 : 1, v/v)]. The tubes were then heated at 100 °C for 4 h with concomitant magnetic stirring of the solution. After completion of the reaction, the reaction mixture vessel was cooled using ice–H_2O, with the reaction mixture then poured into CH_2Cl_2 (40 cm^3); the reaction tubes were next rinsed with H_2O (2 × 10 cm^3). The pH of the combined aqueous solubles was then adjusted to 3–3.2 with 0.4 M $NaHCO_3$ with the CH_2Cl_2 solubles separated and the aqueous phase re-extracted with CH_2Cl_2 (2 × 25 cm^3). The combined organic solubles were next dried (Na_2SO_4) and concentrated *in vacuo*, with the residue purified by a combination of preparative silica gel TLC eluted with hexanes–EtOAc (7 : 3, 1 : 1 and 3 : 7, respectively) and further purified by reversed-phase HPLC (see instrumentation and chromatography materials section), to afford *threo*/*erythro* G triethanethiol in ∽1.2 : 1 ratio (**19**, 9.1 mg, 0.026 mmol, ~40%) and *threo*/*erythro* 5OHG-G benzodioxane triethanethiol in ∽1.7 : 1 ratio (at C7–C8) (**35**, 4.1 mg, 0.0076 mmol, ~12%), respectively.

threo- and *erythro*-**35**: δ_H(500 MHz; acetone-d_6; Me_4Si) 7.75 (1 H, br s, ArOH), 7.10 (1 H, br s, 2-H), 6.96 (1 H, d, *J* 8.0, 6-H), 6.88 (1 H, dd, *J* 1.5 and 8.5, 5-H), 6.78 (1 H, dd, *J* 2.0 and 7.0, 2′-H), 6.70 (1 H, ddd, *J* 1.5 and 6.5, 6′-H), 4.98 (1 H, d, *J* 7.5, 7-H), 4.97 (1 H, d, *J* 8.0, 7-H), 4.40-4.34 (1 H, m, 7′-H), 4.08-4.06 (1 H, m, 8-H), 3.87 (3 H, s, 3-OMe), 3.83 (3 H, br s, 3′-OMe), 3.79 (1 H, br d, *J* 2.5 and 12.5, 9-H), 3.50 (1 H, br d, *J* 3.5 and 12.0, 9-H), 3.24–3.05 (1 H, m, 8′-H), 3.03-2.71 (2 H, m, 9′-H), 2.64-2.33 (6 H, m, 3 × –SCH_2Me), 1.25-1.10 (9 H, m, 3 × –SCH_2Me); δ_C(125 MHz; acetone-d_6; Me_4Si) 149.68, 149.62, 149.44, 149.40, 148.43 (d), 147.94 (d), 145.0 (d), 144.84 (d), 134.15, 133.71 (d), 133.60, 133.55, 132.4, 132.3, 129.3, 121.55 (d), 115.7, 111.90 (t), 111.38, 111.33, 110.85, 110.80, 110.82 (d), 106.6, 106.1, 79.39, 77.05 (d), 77.03 (d), 61.8, 56.38 (d), 56.33, 53.80 (d), 53.73, 53.6, 53.1, 52.5, 37.49, 37.44, 37.34 (d), 27.3, 27.27, 27.16, 27.12, 26.96, 26.91 (d), 26.34, 26.22 (d), 15.40 (d), 15.33, 15.17 (d), 15.05, 15.02 (d), 14.92 (d); *m/z* (LC-APCI) 538.9 (M − H. $C_{26}H_{35}O_6S_3$ requires 539.2); 417.1 (M + H − 2HSEt)$^+$, 478.9 (M + H − HSEt)$^+$ (t_R = 5.7 min).

***threo* and *erythro* 5OHG–5OHG-G benzodioxane triethanethiol.** [(3′-(4-Hydroxy-3-methoxyphenyl)-5,8′-dimethoxy-7-(1,2,3-tris(ethylthio)propyl)-2,2′,3,3′-tetrahydro-2,6′-bibenzo[*b*][1,4]-dioxine-2′,3-diyl)dimethanol] (**36**). Thioacidolysis of the 5OHG–5OHG–G benzodioxane trimer (**46**) (50 mg, 0.088 mmol) was carried out exactly as described above for **32**, with the crude thioacidolysis reaction products purified by silica gel TLC (eluant, EtOAc–acetone, 8 : 2) and further purified by reversed-phase HPLC to afford the *threo*/*erythro*-mixtures of G triethanethiol (**19**, 9.8 mg, 0.028 mmol, ~32%), 5OHG–5OHG–G benzodioxane triethanethiol (**36**, 3.3 mg, 0.0045 mmol, ~5%) and 5OHG–G benzodioxane triethanethiol (**35**, 6 mg, 0.011 mmol, ~13%), respectively.

threo- and *erythro*-**36**: δ_H(500 MHz; acetone-d_6) 7.75 (1 H, br s, ArOH), 7.12 (1 H, dd, *J* 1.5 and 5.5, 2-H), 6.97 (1 H, dd, *J* 1.5 and 8.0, 6-H), 6.86 (1 H, dd, *J* 1.5 and 8.0, 5-H), 6.77 (2 H, br s, 2′-H and 6′-H), 6.719 (1 H, dd, *J* 1.5 and 2.0, 2″-H), 6.717 (1 H, dd, *J* 2.0, 6″-H), 5.0 (1 H, d, *J* 8.0, 7-H), 4.97 (1 H, d, *J* 7.5, 7′-H), 4.39 - 4.34 (1 H, m, 7″-H), 4.08 (2 H, m, 8-H, 8′-H), 3.877 (3 H, br s, OMe), 3.866 (3 H, br s, OMe), 3.834 (3 H, br s, OMe), 3.81 (2 H, m, 9-H and 9′-H), 3.54 (2 H, m, 9-H and 9′-H), 3.22 - 3.04 (1 H, m, 8″-H), 3.02 - 2.70 (2 H, m, 9″-H), 2.63 - 2.33 (6 H, m, 3 × –SCH_2Me), 1.24 - 1.09 (9 H, m, 3 × –SCH_2Me); δ_C(125 MHz; acetone-d_6) 150.15 (t), 149.51, 149.47, 148.55 (d), 148.07, 145.52 (d), 145.47 (d), 144.77 (d), 144.73 (d), 134.79 (d), 134.30 (d), 133.69 (t), 132.57, 132.50 (d), 129.97, 129.92, 129.27 (d), 121.66 (d), 115.80 (d), 111.98, 111.93 (d), 111.40 (br d), 110.89, 110.85 (d), 109.95 (d), 106.73, 106.26, 105.17, 105.13 (d), 105.03 (d), 79.54 (d), 79.31 (d), 79.26 (d), 77.19, 77.12, 76.96 (d), 76.89 (d), 61.84 (C-9 and C-9′), 56.52, 56.45 (d), 56.41 (d), 53.86 (d), 53.78, 53.70, 53.18 (d), 52.58 (d), 37.53 (d), 37.43, 27.40, 27.35, 27.22 (d), 27.00 (dd), 26.42, 26.30 (d), 15.47, 15.41, 15.25, 15.10 (d), 14.99; *m/z* (LC-APCI) 733.1 (M − H. $C_{36}H_{45}O_{10}S_3$ requires 733.2) (t_R = 5.6 min); 672.9 (M + H − HSEt)$^+$, 611.1 (M + H − 2 HSEt)$^+$.

Time course of thioacidolytic cleavage of 5OHG–G benzodioxane dimer (34) and 5OHG–5OHG–G benzodioxane trimer (46)

The 5OHG–G benzodioxane dimer (**32**) (3.1 μmol) was initially subjected to standard quantitative thioacidolytic conditions [(+)-sesamin as internal standard, $C_{20}H_{18}O_6$, *ca* 75 μg],[45,51] while quenching the reaction at hourly intervals from 0 to 5 h. With the resulting thioacidolysis reaction mixture dissolved in THF (10 cm^3), aliquots (1 mm^3) were then subjected to reversed-phase UPLC analyses (BEH C18 column, Waters, 2.1 × 150 mm) with elution conditions at a flow rate of 0.2 cm^3 min^{-1}: linear gradients of MeOH-3% (v/v) AcOH in H_2O from 40 : 60 to 70 : 30 in 3 min, then to 80 : 20 in 4 min, to 100 : 0 in 2.25 min with this composition held for an additional 0.25 min, and finally to 40 : 60 in 0.5 min this being held for 5 min). The 5OHG–5OHG–G benzodioxane trimer (**46**) (2.7 μmol) was then subjected to standard quantitative thioacidolytic conditions exactly as described above for 5OHG–G benzodioxane dimer (**32**). The various thioacidolysis products from **32** and **46** were identified by comparison to retention times of authentic standards and by mass spectrometric analyses. The chemical characterization of intermediate 5OHG–G benzodioxane monoethanethiol **47**, as obtained from 5OHG–G dimer (**32**) at the early phase thioacidolysis reaction, was established by 1D ^1H and ^{13}C NMR spectroscopy and mass spectral analysis. For quantitative analyses, authentic standards G (**19**), 5OHG (**20**), S (**21**), 5OHG–G (**35**), and 5OHG–5OHG–G (**36**) were used to

calculate the corresponding response factors. All samples were analyzed in duplicate.

5OHG–G benzodioxane monoethanethiol [(*E*)-4-(7-(3-(ethylthio)prop-1-enyl)-3-(hydroxymethyl)-5-methoxy-2,3-dihydrobenzo[*b*][1,4]dioxin-2-yl)-2-methoxyphenol] (**47**): δ_H(500 MHz; acetone-d_6; Me_4Si) 7.77 (1 H, ArOH), 7.11 (1 H, d, *J* 2.0, 2-H), 6.96 (1 H, dd, *J* 2.0 and 8.0, 6-H), 6.88 (1 H, d, *J* 8.0, 5-H), 6.72 (1 H, d, *J* 1.5, 2′-H), 6.60 (1 H, d, *J* 1.5, 6′-H), 6.39 (1 H, d, *J* 15.5, 7′-H), 6.12 (1 H, dt, *J* 7.5 and 16.0, 8′-H), 4.98 (1 H, d, *J* 8.0, 7-H), 4.0 (1 H, m, 8-H), 3.87 (3 H, s, 3-OMe), 3.86 (3 H, s, 3′-OMe), 3.78 and 3.50 (2 H, m, 9-H), 3.30 (2 H, d, *J* 7.5, 9′-H), 2.49 (2 H, q, *J* 7.5, 10′-H), 1.21 (3 H, t, *J* 6.5 and 7.5, 11′-H); δ_C(125 MHz; acetone-d_6; Me_4Si) 150.12 (C-3′), 148.45 (C-3), 148.44 (C-4), 145.44 (C-5′), 134.22 (C-4′), 132.41 (C-7′), 130.2 (C-1′), 129.34 (C-1), 125.38 (C-8′), 121.55 (C-6), 115.7 (C-5), 111.9 (C-2), 108.67 (C-6′), 103.4 (C-2′), 79.47 (C-8), 77.04 (C-7), 61.75 (C-9), 56.32 (C-3-OMe), 56.31 (C-3′-OMe), 34.19 (C-9′), 25.0 (C-10′), 14.97 (C-11′); *m/z* (LC-APCI) 417.0 (M - H. $C_{22}H_{26}O_6S$ requires 417.1) (t_R = 3.7 min).

Isolation and molecular weight distributions of lignin-derived isolates from WT and *Atomt1* lines

Whole stem sections from *Atomt1* mutant and WT lines were subjected to a modified Björkmann[57] lignin isolation procedure as described by Bernard-Vailhé *et al.*[58] The extractive-free CWRs obtained from *Atomt1* mutant (6.95 g) and WT (12.15 g) lines, were individually ground by ball-milling for 96 h at 4 °C and then extracted twice using a mixture of dioxane–water (96 : 4, v/v, 30 cm^3 g^{-1}) for 24 h. The crude lignins thus obtained for *Atomt1* mutant (362 mg) and WT (710 mg) were then individually subjected to various precipitation conditions as described by Jourdes *et al.*[27] to furnish the corresponding lignin-enriched fractions: *Atomt1* mutant (157 mg, 11.6%); λ_{max}(dioxane)/nm 281 (ε/g^{-1} cm^{-1} 15.7 ± 0.11). WT (282 mg, 10.8%); λ_{max}(dioxane)/nm 281 (ε/g^{-1} cm^{-1} 16.6 ± 0.16). Molecular weight distributions of both lignin-enriched isolates were estimated by gel permeation chromatography on Sephadex G-100 column (100 cm × 2.5 cm) eluted with 0.1 M NaOH at 40 cm^3 h^{-1} and precalibrated with sodium polystyrene sulfonates (Mw's: ~1430; 5180; 29 000; 79 000) and coniferyl alcohol (**2**, 180) as previously described.[32]

NMR spectroscopic analyses of lignin-enriched isolates from *Atomt1* mutant and WT lines

Each lignin isolate (~30 mg) of both WT and *Atomt1* lines was individually dissolved in DMSO-d_6 (0.5 cm^3), with the ^{13}C NMR, two-dimensional phase-sensitive gradient-selected, HMQC and HMBC spectra individually recorded using similar acquisition parameters as described earlier.[32]

Thioacidolysis of lignin-enriched isolates harboring trimer (36) and dimer (35) from *Atomt1* mutant

Thioacidolytic cleavage of the *Atomt1* lignin-enriched isolate (90 mg) was carried out as described for **32** and **46**. The crude reaction mixture so obtained (120 mg) was first subjected to silica gel TLC eluted with hexanes–EtOAc (7 : 3), with isolated bands individually subjected to NMR and mass spectroscopic analyses. NMR (1D and 2D) spectra of two fractions Fr-2 (3.5 mg) and Fr-3 (17.7 mg) contained benzodioxane substructures [4.81 (7-H) and 76.0 (C-7)/4.0 (8-H) and 78.2 (C-8)]. These were individually subjected to reversed-phase HPLC separation (see Instrumentation and Chromatography Materials section) with fractions corresponding to dimer **35** (from Fr-2) and trimer **36** (from Fr-3) individually pooled, extracted with EtOAc (40 cm^3), evaporated to dryness *in vacuo* and freeze-dried. The 1D ^1H and ^{13}C NMR spectra of **35** and **36** contained identical chemical shift resonance values to synthetic **35** and **36**, respectively.

5OHG–G benzodioxane triethanethiol (**35**, 0.8 mg); *m/z* (LC-APCI) 478.9 $(M + H - HSEt)^+$ and 417.1 $(M + H - 2HSEt)^+$ (t_R = 5.7 min).

5OHG–5OHG–G benzodioxane triethanethiol (**36**, 3.5 mg); *m/z* (LC-APCI) 672.9 $(M + H - HSEt)^+$ and 611.1 $(M + H - 2HSEt)^+$ (t_R = 5.6 min).

Arabidopsis COMT gene cloning, heterologous expression, purification and kinetic analyses

Twelve of 17 putative COMTs (*AtCOMT1*, *3*, *4*, *6*, *8–13*, *15* and *17*)[46] were isolated from *Arabidopsis* (ecotype: Columbia) as follows: After isolation of total RNA using RNeasy Mini kit (Qiagen), single-strand cDNA was generated with reverse transcriptase II (Invitrogen) following the manufacturer's instructions. The twelve putative COMTs were then individually amplified with PCR; their products were cloned using the TOPO TA cloning kit (Invitrogen) and their sequences were confirmed by individual sequencing.

For heterologous protein expression, each gene was re-amplified with directional nucleotide containing forward primer. The products were then directionally cloned into a pET101/D-TOPO expression vector. After sequence confirmation, each of the vectors was individually transformed into *E. coli* expression cell [BL21 Star (DE3)]. *E. coli* cells were cultured in Luria-Bertani (LB) broth supplemented with carbenicillin (0.1 mg cm^{-3}) at 37 °C. When cell growth reached an OD_{600} value between 0.5 and 0.9, 0.5 mM isopropyl thio-β-D-galactoside (IPTG) was added for induction. Cells were individually harvested by centrifugation at 3,000 × *g* for 20 min and stored at −80 °C until needed. Pellets were thawed, suspended in a 1 : 1 mixture of Tris-HCl (20 mM; pH 7.5, buffer A) and BugBuster™ Protein Extraction reagent. Cellular debris was removed by centrifugation (15 000 × *g*, 30 min), with the resulting pellet resuspended in Buffer A (20 mM Tris-HCl, pH 7.9; 500 mM NaCl; 5 mM imidazole). Each suspension was filtered (0.45 μm) and applied to a POROS 20 MC column (100 mm × 4.6 mm) pre-equilibrated with Buffer A) at a flow rate of 10 cm^3 min^{-1} and eluted as previously described.[59] Aliquots (12 mm^3) of every other fraction were analyzed by SDS-polyacrylamide gel electrophoresis.[60] Fractions containing His-tagged AtCOMTs were pooled and concentrated using Centricon® Plus-80 concentrators, diluted with 20 mM Tris-HCl (pH 7.5) and reconcentrated. Protein concentrations were determined using the Bradford method,[61] with BSA as standard.

Enzyme assays for candidate AtCOMTs were performed as follows: Each assay (300 mm^3) contained Tris-HCl buffer (100 mM, pH 7.5), purified candidate AtCOMT (5 μg), substrates [caffeic (**5**)/5-hydroxyferulic (**6**) acid, caffeyl (**9**)/5-hydroxyconiferyl (**10**) aldehyde and caffeyl (**11**)/5-hydroxyconiferyl (**4**) alcohol, 0.33 mM] and S-adenosylmethionine (0.33 mM). Assays were

initiated by addition of each candidate AtCOMT, incubated for 10 min and stopped by acidification (glacial AcOH, 10 mm^3). Controls included both boiled enzyme extracts and omission of substrates. An aliquot (80 µl) of each assay mixture was then subjected to HPLC analysis.

Note added in proof: A paper (Lu *et al.*, Plant Physiology, doi/10.1104/pp.110.154278) has just appeared on an incompletely COMT-suppressed poplar line and analysis of the cell-wall isolates.

Abbreviations

COMT, caffeic acid *O*-methyltransferase; CWR, cell-wall residue; G, guaiacyl; S, syringyl; 5OHG, 5-hydroxyguaiacyl; WT, wild type; HMQC, heteronuclear multiple quantum coherence; HMBC, heteronuclear multiple bond coherence

Acknowledgements

This work was supported by the Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy, by the BioEnergy Science Center, the U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science, as well as the G. Thomas and Anita Hargrove Center for Plant Genomic Research. Thanks are extended to Prof. Ferreira and Dr Bharathi Avula for carrying out the HRMS analyses and to Dr Lise Jouanin for providing seeds of the *Atomt1* knockout.

Notes and references

1 L. B. Davin, M. Jourdes, A. M. Patten, K.-W. Kim, D. G. Vassão and N. G. Lewis, *Nat. Prod. Rep.*, 2008, **25**, 1015–1090.
2 L. B. Davin, A. M. Patten, M. Jourdes and N. G. Lewis, in *Biomass Recalcitrance. Deconstructing the Plant Cell Wall for Bioenergy*, ed. M. E. Himmel, Blackwell Publishing, Oxford, UK, 2008, pp. 213-305.
3 A. M. Patten, M. Jourdes, C. L. Cardenas, D. D. Laskar, Y. Nakazawa, B.-Y. Chung, V. R. Franceschi, L. B. Davin and N. G. Lewis, *Mol. BioSyst.*, 2010, **6**, 499–515.
4 M. Shimada, *Wood Research*, 1972, **53**, 19–65.
5 M. Shimada, H. Fushiki and T. Higuchi, *Mokuzai Gakkaishi*, 1973, **19**, 13–21.
6 C. Hermann, M. Legrand, P. Geoffroy and B. Fritig, *Arch. Biochem. Biophys.*, 1987, **253**, 367–376.
7 E. Jaeck, F. Martz, V. Stiefel, B. Fritig and M. Legrand, *Mol. Plant-Microbe Interact.*, 1996, **9**, 681–688.
8 L. Pellegrini, P. Geoffroy, B. Fritig and M. Legrand, *Plant Physiol.*, 1993, **103**, 509–517.
9 R. Atanassova, N. Favet, F. Martz, B. Chabbert, M.-T. Tollier, B. Monties, B. Fritig and M. Legrand, *Plant J.*, 1995, **8**, 465–477.
10 U. N. Dwivedi, W. H. Campbell, J. Yu, R. S. S. Datla, R. C. Bugos, V. L. Chiang and G. K. Podila, *Plant Mol. Biol.*, 1994, **26**, 61–71.
11 W. Ni, N. L. Paiva and R. A. Dixon, *Transgenic Res.*, 1994, **3**, 120–126.
12 A. M. Anterola and N. G. Lewis, *Phytochemistry*, 2002, **61**, 221–294.
13 J. Van Doorsselaere, M. Baucher, E. Chognot, B. Chabbert, M.-T. Tollier, M. Petit-Conil, J.-C. Leplé, G. Pilate, D. Cornu, B. Monties, M. Van Montagu, D. Inzé, W. Boerjan and L. Jouanin, *Plant J.*, 1995, **8**, 855–864.
14 C. Lapierre, B. Pollet, M. Petit-Conil, G. Toval, J. Romero, G. Pilate, J.-C. Leplé, W. Boerjan, V. Ferret, V. de Nadaï and L. Jouanin, *Plant Physiol.*, 1999, **119**, 153–163.
15 L. Jouanin, T. Goujon, V. de Nadaï, M.-T. Martin, I. Mila, C. Vallet, B. Pollet, A. Yoshinaga, B. Chabbert, M. Petit-Conil and C. Lapierre, *Plant Physiol.*, 2000, **123**, 1363–1373.
16 C.-J. Tsai, J. L. Popko, M. R. Mielke, W.-J. Hu, G. K. Podila and V. L. Chiang, *Plant Physiol.*, 1998, **117**, 101–112.
17 D. Guo, F. Chen, K. Inoue, J. W. Blount and R. A. Dixon, *Plant Cell*, 2001, **13**, 73–88.
18 S. Maury, P. Geoffroy and M. Legrand, *Plant Physiol.*, 1999, **121**, 215–223.
19 L. Li, J. L. Popko, T. Umezawa and V. L. Chiang, *J. Biol. Chem.*, 2000, **275**, 6537–6545.
20 K. Parvathi, F. Chen, D. Guo, J. W. Blount and R. A. Dixon, *Plant J.*, 2001, **25**, 193–202.
21 R. A. Emerson, G. W. Beadle and A. C. Fraser, *Cornell University Agricultural Experimental Station Memoir*, 1935, 180, 1–83.
22 L. R. Jorgenson, *J. Am. Soc. Agronomy*, 1931, **23**, 549–557.
23 L. D. Muller, R. F. Barnes, L. F. Bauman and V. F. Colenbrander, *Crop Sci.*, 1971, **11**, 413–415.
24 M. S. Zuber, T. R. Colbert and L. F. Bauman, *Z. Pflanzen.*, 1977, **79**, 310–314.
25 C. Grand, P. Parmentier, A. Boudet and A. M. Boudet, *Physiol. Veg.*, 1985, **23**, 905–911.
26 C. Lapierre, M.-T. Tollier and B. Monties, *C.R. Acad. Sci., Ser. III*, 1988, **307**, 723–728.
27 M. Jourdes, C. L. Cardenas, D. D. Laskar, S. G. A. Moinuddin, L. B. Davin and N. G. Lewis, *Phytochemistry*, 2007, **68**, 1932–1956.
28 Y.-r. Chen and S. Sarkanen, *Phytochem. Rev.*, 2003, **2**, 235–255.
29 S.-Y. Guan, J. Mlynár and S. Sarkanen, *Phytochemistry*, 1997, **45**, 911–918.
30 S. Sarkanen, in *Lignin and Lignan Biosynthesis*, ed. N. G. Lewis and S. Sarkanen, ACS Symposium Series, Washington, DC, USA, 1998, pp. 194-208.
31 A. M. Patten, C. L. Cardenas, F. C. Cochrane, D. D. Laskar, D. L. Bedgar, L. B. Davin and N. G. Lewis, *Phytochemistry*, 2005, **66**, 2092–2107.
32 D. D. Laskar, M. Jourdes, A. M. Patten, G. L. Helms, L. B. Davin and N. G. Lewis, *Plant J.*, 2006, **48**, 674–686.
33 J. M. Marita, J. Ralph, R. D. Hatfield, D. Guo, F. Chen and R. A. Dixon, *Phytochemistry*, 2003, **62**, 53–65.
34 J. Ralph, C. Lapierre, J. M. Marita, H. Kim, F. Lu, R. D. Hatfield, S. Ralph, C. Chapple, R. Franke, M. R. Hemm, J. Van Doorsselaere, R. R. Sederoff, D. M. O'Malley, J. T. Scott, J. J. MacKay, N. Yahiaoui, A.-M. Boudet, M. Pean, G. Pilate, L. Jouanin and W. Boerjan, *Phytochemistry*, 2001, **57**, 993–1003.
35 J. Ralph, K. Lundquist, G. Brunow, F. Lu, H. Kim, P. F. Schatz, J. M. Marita, R. D. Hatfield, S. A. Ralph, J. H. Christensen and W. Boerjan, *Phytochem. Rev.*, 2004, **3**, 29–60.
36 L. B. Davin and N. G. Lewis, *Curr. Opin. Biotechnol.*, 2005, **16**, 407–415.
37 S. Dutta, T. M. Garver Jr and S. Sarkanen, *ACS Symp. Ser.*, 1989, **397**, 155–176.
38 K. Blee, J. W. Choi, A. P. O'Connell, S. C. Jupe, W. Schuch, N. G. Lewis and G. P. Bolwell, *Phytochemistry*, 2001, **57**, 1159–1166.
39 K. Morreel, J. Ralph, F. Lu, G. Goeminne, R. Busson, P. Herdewijn, J. L. Goeman, J. Van der Eycken, W. Boerjan and E. Messens, *Plant Physiol.*, 2004, **136**, 4023–4036.
40 T. H. Kim, H. Ito, K. Hayashi, T. Hasegawa, T. Machiguchi and T. Yoshida, *Chem. Pharm. Bull.*, 2005, **53**, 641–644.
41 A. Kuboki, T. Yamamoto, M. Taira, T. Arishige and S. Ohira, *Tetrahedron Lett.*, 2007, **48**, 771–774.
42 C. Ma, H. J. Zhang, G. T. Tan, N. V. Hung, N. M. Cuong, D. D. Soejarto and H. H. S. Fong, *J. Nat. Prod.*, 2006, **69**, 346–350.
43 Y. K. Son, M. H. Lee and Y. N. Han, *Arch. Pharmacal Res.*, 2005, **28**, 34–38.
44 R. Waibel, G. Benirschke, M. Benirschke and H. Achenbach, *Phytochemistry*, 2003, **62**, 805–811.
45 C. Lapierre, B. Monties and C. Rolando, *Holzforschung*, 1986, **40**, 113–118.
46 M. A. Costa, R. E. Collins, A. M. Anterola, F. C. Cochrane, L. B. Davin and N. G. Lewis, *Phytochemistry*, 2003, **64**, 1097–1112.
47 I. Muzac, J. Wang, D. Anzellotti, H. Zhang and R. K. Ibrahim, *Arch. Biochem. Biophys.*, 2000, **375**, 385–388.
48 T. Goujon, R. Sibout, B. Pollet, B. Maba, L. Nussaume, N. Bechtold, F. Lu, J. Ralph, I. Mila, Y. Barrière, C. Lapierre and L. Jouanin, *Plant Mol. Biol.*, 2003, **51**, 973–989.
49 K. Iiyama and A. F. A. Wallis, *J. Sci. Food Agric.*, 1990, **51**, 145–161.
50 K. Iiyama and T. B. T. Lam, *J. Sci. Food Agric.*, 1990, **51**, 481–491.
51 C. Rolando, B. Monties and C. Lapierre, in *Methods in Lignin Chemistry*, ed. S. Y. Lin and C. W. Dence, Springer-Verlag, Berlin, Germany, 1992, pp. 334-349.

52 S.-J. Kim, M.-R. Kim, D. L. Bedgar, S. G. A. Moinuddin, C. L. Cardenas, L. B. Davin, C. Kang and N. G. Lewis, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 1455–1460.
53 W. Gu, X. Jing, X. Pan, A. S. C. Chan and T.-K. Yang, *Tetrahedron Lett.*, 2000, **41**, 6079–6082.
54 X. She, S. Qi, W. Gu and X. Pan, *J. Chem. Res. (S)*, 1998, 436–437.
55 X. She, W. Gu, T. Wu and X. Pan, *J. Chem. Res. (S)*, 1999, 100–101.
56 N. R. Guz and F. R. Stermitz, *J. Nat. Prod.*, 2000, **63**, 1140–1145.
57 A. Björkman, *Nature*, 1954, **174**, 1057–1058.
58 M.-A. Bernard-Vailhé, A. Cornu, D. Robert, M.-P. Maillot and J.-M. Besle, *J. Agric. Food Chem.*, 1996, **44**, 1164–1169 .
59 F. C. Cochrane, L. B. Davin and N. G. Lewis, *Phytochemistry*, 2004, **65**, 1557–1564.
60 U. K. Laemmli, *Nature*, 1970, **227**, 680–685.
61 M. M. Bradford, *Anal. Biochem.*, 1976, **72**, 248–254.

