Bioenerg. Res. (2015) 8:934–952
DOI 10.1007/s12155-015-9583-4

Naturally *p*-Hydroxybenzoylated Lignins in Palms

Fachuang Lu · Steven D. Karlen · Matt Regner · Hoon Kim · Sally A. Ralph · Run-Cang Sun · Ken-ichi Kuroda · Mary Ann Augustin · Raymond Mawson · Henry Sabarez · Tanoj Singh · Gerardo Jimenez-Monteon · Sarani Zakaria · Stefan Hill · Philip J. Harris · Wout Boerjan · Curtis G. Wilkerson · Shawn D. Mansfield · John Ralph

Published online: 27 February 2015

Abstract The industrial production of palm oil concurrently generates a substantial amount of empty fruit bunch (EFB) fibers that could be used as a feedstock in a lignocellulose-based biorefinery. Lignin byproducts generated by this process may offer opportunities for the isolation of value-added products, such as *p*-hydroxybenzoate (*p*Bz), to help offset operating costs. Analysis of the EFB lignin by nuclear magnetic resonance (NMR) spectroscopy clearly revealed the presence of bound acetate and *p*Bz, with saponification revealing that 1.1 wt% of the EFB was *p*Bz; with a lignin content of 22.7 %, 4.8 % of the lignin is *p*Bz that can be obtained as a pure component for use as a chemical feedstock. Analysis of EFB lignin by NMR and derivatization followed by reductive cleavage (DFRC) showed that *p*Bz selectively

F. Lu · S. D. Karlen · M. Regner · H. Kim · J. Ralph
The Department of Energy's Great Lakes Bioenergy Research Center, the Wisconsin Energy Institute, University of Wisconsin, Madison, WI 53726, USA

F. Lu · M. Regner · H. Kim · J. Ralph
Department of Biochemistry, University of Wisconsin, Madison, WI 53706, USA

S. A. Ralph
US Forest Products Laboratory, USDA-Forest Service, Madison, WI 53726, USA

R.-C. Sun
Beijing Key Laboratory of Lignocellulosic Chemistry, Beijing Forestry University, Beijing 100083, China

K.-i. Kuroda
Department of Forest and Forest Products Sciences, Faculty of Agriculture, Kyushu University, Fukuoka 812-8581, Japan

M. A. Augustin · R. Mawson · H. Sabarez · T. Singh
CSIRO Food, Nutrition and Bioproducts Flagship, Werribee, VIC 3030, Australia

G. Jimenez-Monteon
USA-ARS Dairy Forage Research Center, Madison, WI 53706, USA

S. Zakaria
Bioresources and Biorefinery Laboratory, Universiti Kebangsaan Malaysia, Selangor 43600, Malaysia

S. Hill
Scion, Rotorua 3046, New Zealand

P. J. Harris
School of Biological Sciences, The University of Auckland, Auckland, New Zealand

W. Boerjan
Department of Plant Systems Biology, VIB, B-9052 Gent, Belgium

W. Boerjan
Department of Plant Biotechnology and Bioinformatics, Ghent University, B-9052 Gent, Belgium

C. G. Wilkerson
Departments of Plant Biology, and Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824, USA

C. G. Wilkerson
The Department of Energy's Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, MI 48824, USA

S. D. Mansfield
Department of Wood Science, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

J. Ralph (✉)
Department of Biochemistry and the D.O.E. Great Lakes Bioenergy Research Center, Wisconsin Energy Institute, University of Wisconsin, Madison, 1552 University Ave., Madison, WI 53726-4084, USA
e-mail: jralph@wisc.edu

acylates the γ-hydroxyl group of S units. This selectivity suggests that *p*Bz, analogously with acetate in kenaf, *p*-coumarate in grasses, and ferulate in a transgenic poplar augmented with a feruloyl-CoA monolignol transferase (FMT), is incorporated into the growing lignin chain via its γ-*p*-hydroxybenzoylated monolignol conjugate. Involvement of such conjugates in palm lignification is proven by the observation of novel *p*-hydroxybenzoylated non-resinol β–β-coupled units in the lignins. Together, the data implicate the existence of *p*-hydroxybenzoyl-CoA:monolignol transferases that are involved in lignification in the various willows (*Salix* spp.), poplars and aspen (*Populus* spp., family Salicaceae), and palms (family Arecaceae) that have *p*-hydroxybenzoylated lignins. Even without enhancing the levels by breeding or genetic engineering, current palm oil EFB 'wastes' should be able to generate a sizeable stream of *p*-hydroxybenzoic acid that offers opportunities for the development of value-added products derived from the oil palm industry.

Keywords Lignin acylation · Transferase · NMR · DFRC method · Poplar · *p*-Hydroxybenzoic acid · Monolignol

Introduction

In order for a lignocellulose biorefinery to function as an economically viable enterprise, it must be developed to yield value-added coproducts in the process [1, 2]. Coupling the source of biomass to the waste stream of an associated commercial process is one way to create value in the overall system. The production of palm oil generates a substantial amount of empty fruit bunch (EFB) fibers that can be readily processed into sugars for biofuels production and a mixture of unfermentable components including lignin. The resulting lignin stream could then offer new opportunities for the development and isolation of value-added products derived from the oil palm industry [3]. For example, in 2013 Malaysia, the second largest oil palm producer after Indonesia, produced over 19 million tons of palm oil on a total plantation area of over 5 million ha, which also generated an estimated 70–80 million tons of biomass, mostly from oil palm EFB [4]. This biomass, which is a high-quality lignocellulosic fiber, has not been fully utilized commercially and is a largely untapped resource, although research in the biorefinery area has attempted to provide high value-added products via liquefaction [5], solvolysis [6] and pyrolysis [7].

Lignin, a stochastically generated biopolymer found in plants, forms by means of radical coupling reactions, primarily between a monolignol and the growing lignin polymer in an endwise process [8, 9]. In angiosperms, coniferyl and sinapyl alcohols are the predominant monomers and, once incorporated into the lignin backbone, produce guaiacyl (G) and syringyl (S) units in the resulting polymer. A monolignol couples overwhelmingly at its β-position to the growing polymer at available 4–O- (G or S) or 5- (G only) positions. Branching is introduced by 5–O–4- or 5–5-coupling reactions between preformed oligomers, with the latter occurring only between G units and the former involving at least one G unit. Evidence continues to accumulate, particularly from studies of the lignin biosynthetic pathway using mutant and transgenic plants, that lignins can incorporate monomers beyond the traditional three monolignols. As reviewed previously [8–11], these include (but are not limited to) the novel caffeyl and 5-hydroxyconiferyl alcohols from incomplete methylation in monolignol biosynthesis (OMT-deficient plants); the immediate hydroxycinnamaldehyde monolignol precursors, coniferaldehyde and sinapaldehyde, and the hydroxy-benzaldehydes (vanillin and syringaldehyde) derived from them from incomplete reduction (CAD deficiency); and products such as dihydroconiferyl alcohol and its derived guaiacylpropane-1,3-diol in wild-type and mutant softwoods.

Beyond the monolignol biosynthetic pathway, various monolignol conjugates have also been shown to participate in lignification [12–14]. Acylated lignins of four types have been encountered in various plants to date. Angiosperm trees have long been thought to have low-level lignin acetylation [15], whereas kenaf bast fiber lignins and lignins in various other plants are extensively acetylated [16–21]. Grasses produce the enzyme *p*-coumaroyl-CoA:monolignol transferase (PMT) [22, 23], which produces γ-*p*-coumaroylated monolignols that are incorporated into the growing lignin polymer, particularly at advanced maturity [24–26]. Recently, monolignol ferulate conjugates were introduced into poplar trees through the introduction of a gene expressing feruloyl-CoA:monolignol transferase (FMT) [27], further demonstrating the plasticity of lignification and the role of monolignol conjugates. A few hardwoods have partially *p*-hydroxybenzoylated lignins, notably the willows (*Salix spp.*) (family Salicaceae) [28], the poplars and aspens (*Populus spp.*) (family Salicaceae) [29, 24, 30–32], and *Aralia cordata* (family Araliaceae) [33]. In addition, the palms, which are a large family (Arecaceae) of economically important monocotyledonous plants comprising some 2400 species, also have partially *p*-hydroxybenzoylated lignins [34–37]. Several low molecular mass cross-coupling products γ-acylated with *p*-hydroxybenzoate were isolated from actively lignifying poplar xylem [29, 32]. Enhanced interest in *p*-hydroxybenzoates has surfaced in conjunction with reduced lignification in aspen and poplar following down- or upregulation of genes in the monolignol pathway [38–41].

Relatively little is known about the nature of *p*-hydroxybenzoylated lignins or how they arise. From oil palm trunk, nitrobenzene oxidation produced high levels of *p*-hydroxybenzoic acid (9.8–14.9 %) [37], which was also the main product of alkaline hydrolysis of parenchyma milled wood lignin (MWL) [31]. Small quantities of vanillic and syringic

acids were also found. The authors concluded that the phenolic acids acylated the side-chain alcohol groups of lignin subunits. High levels of *p*-hydroxybenzoic acid were also released from oil palm lignins [42–46, 34]. The ^{13}C-NMR chemical shifts in *Salix* (willow) MWL samples showed good concordance with expected shifts for free-phenolic γ-*p*-hydroxybenzoate esters based on model compound studies [28]. Rather than suspecting that the lignin side-chains were acylated after the polymer had already formed, *p*-hydroxybenzoyl monolignol conjugates were proposed as the precursors to these structures in early research [47]; monolignol acylation was further conjectured [25] and eventually proven by the observation of novel β–β-coupling products in lignins that could only have arisen from pre-acylated monolignols [29, 14, 48, 32]. Recent studies have delivered the transferases *p*-coumaroyl-CoA:monolignol transferase (PMT) in grasses responsible for lignin *p*-coumarate esters via *p*-coumaroylation of monolignols [22, 23, 49, 50] and now an exotic feruloyl-CoA:monolignol transferase (FMT) that feruloylates monolignols prior to lignification [27]; the putative *p*-hydroxybenzoyl-CoA and acetyl-CoA monolignol transferases have not yet been identified.

The research presented in this paper seeks to answer several key questions regarding the intricacies of acylation of lignins by *p*-hydroxybenzoate, specifically: whether *p*-hydroxybenzoates exist as free-phenolic appendages on lignins, akin to the *p*-coumarates; whether they, like *p*-coumarates [51, 52, 23, 49] and acetates [18] in most plants, primarily acylate S units; whether the various lignin substructures carry these units and thus implicate the incorporation of preformed monolignol γ-*p*-hydroxybenzoate conjugates into the lignification process; and whether diagnostic β–β-coupled units can be found to provide proof that lignification utilizes preformed monolignol *p*-hydroxybenzoate conjugates, thus, also establishing that monolignol *p*-hydroxybenzoate conjugates are lignin precursors. Although a survey across all palms is beyond the scope of this work, a sago palm lignin previously examined by analytical pyrolysis [35] and prior studies on coconut (*Cocos nucifera*) coir fibers [36] are used to suggest a degree of generality in palms. Together, such studies will help to elucidate the roles of putative genes and enzymes associated with the required transferase activity, knowledge that may clarify the poorly understood reasons for widespread lignin acylation in the plant kingdom. At the same time, it will also shed light on future paths that may be worth exploring for value-added products making lignocellulosic biofuels a more viable option; depending on the level of *p*-hydroxybenzoates found in a given biomass material, and the ease with which it can be cleaved and recovered from the polymer, *p*-hydroxybenzoic acid or its derivatives may be accessible value-added products [34, 31]. A simple net search shows that *p*-hydroxybenzoic acid currently sells for ~$US4,000/ton and has uses: as a source of esters that have antibacterial and antifungal properties, as an intermediate in pesticides and antiseptics, and for preparation of co-polyesters (including liquid crystal polymers).

Materials and Methods

General

Commercial chemicals, including solvents, were of reagent grade or better and used without further purification. Analytical thin-layer chromatography (TLC) was carried out on EM Science TLC plates pre-coated with silica gel 60F_{254}; TLC visualization was by UV. Purification by flash column chromatography was accomplished on silica gel (230–450 mesh). 4-Acetoxybenzoyl chloride ***p*Bz$_{Cl}$** was prepared according to a previous procedure [53]. Nuclear magnetic resonance (NMR) data and assignments for all model compounds are from well-described 1D and 2D NMR experiments (on a Bruker Biospin Avance 500 MHz spectrometer). High-resolution mass spectrometric (HRMS) data were acquired using electrospray ionization (ESI) on a Waters (Micromass) LCT ESI/TOF MS instrument. UV–vis analysis performed in 1-cm quartz cuvettes at λ=280 nm on a Shimadzu UV-1800.

Plant Materials

Palm Samples The African oil palm (*Elaeis guineensis*) fractions used in this study came from two sources. The first sample, consisting of oil palm fronds and empty fruit bunch (EFB) fibers, was supplied by the Forest Research Institute of Malaysia and is the same sample used in a previous study by Sun et al. [45] and will be denoted here as EFB-1. The second sample of oil palm EFB fibers was provided by the Ladang Tai Tak Sdn. Bhd. Palm oil mill in Kota Tinggi, Johor, Malaysia, and will be denoted here as EFB-2. These EFB fibers were steam-treated (45 psi, 130 °C, 70 min) to remove the fruitlets and then shredded into fibers (4–6 in in length). Various data on both oil palm EFB samples are presented in Table 1. Isolated sago palm (*Metroxylon sagu*) lignin was the material previously described in a pyrolysis study [35].

Sample Preparation and Lignin Isolation

All fibrous oil palm samples were dried at 50 °C in an oven for at least 72 h before being ground. Two-step grinding was performed in a Wiley mill using a 2-mm followed by a 1-mm mesh. Removal of extractives from the samples was carried out by successive extractions with water, ethanol, acetone, and chloroform. Extractions were accomplished by refluxing the material in Soxhlet equipment for at least 8 h with each solvent. Prior to analysis, extracted samples, frequently referred to as whole-cell-wall preparations, were dried

Table 1 Analytical data on the two sources of oil palm EFB material and isolated MWL

	EFB-1 (Sun)	EFB-2 (Ladang Tai Tak)
AcBr lignin (wt% WCW)	22.7±0.4 %	18.3±0.2 %
Saponification		
p-Hydroxybenzoic acid (wt% WCW)	1.09±0.06 %	1.06±0.04 %
p-Hydroxybenzoic acid (wt% AcBr lignin, calc.)	4.8 %	5.8 %
p-Hydroxybenzoic acid (wt% MWL)	4.2±0.1 %	–
DFRC		
Coniferyl alcohol (μg/mg AcBr lignin)	40±2	45±5
Sinapyl alcohol (μg/mg AcBr lignin)	88±6	108±12
Coniferyl *p*-hydroxybenzoate (μg/mg AcBr lignin)	2.0±0.1	3.0±0.4
Sinapyl *p*-hydroxybenzoate (μg/mg AcBr lignin)	59±4	85±2
Total monolignols detected (wt% of AcBr lignin)	16.7 %	20.9 %
p-Hydroxybenzoic acid (wt% WCW)	0.6 %	0.7 %
p-Hydroxybenzoic acid (wt% AcBr lignin, calc.)	2.6 %	3.7 %
SA:CA (mole ratio)	73:27	75:25
SA-*p*Bz:CA-*p*Bz (mole ratio)	96:4	96:4
(SA+CA):(SA-*p*Bz+CA-*p*Bz) (mole ratio)	77:23	74:26

DFRC values are raw data, uncorrected for recovery. Other abbreviations are in the text

WCW whole-cell-wall material, *MWL* milled wood (Björkman) lignin

for 2 weeks in a desiccator under vacuum using P_2O_5 as drying agent, which was replaced regularly. The dried materials were ball-milled using a stainless steel vibratory ball mill for 1.5 h (with 0.5 h on, 0.5 h off, to minimize heating). The ball-milled materials (50 g) were submitted to enzymatic digestion with crude cellulases (Cellulysin, Calbiochem) for polysaccharide degradation [25, 54, 55]. As with aspen and poplar, the enzymatic treatment was particularly efficient at removing polysaccharides, leaving just 22.2 and 20.9 % of the cell wall from the EFB and frond samples, respectively. Samples of the enzyme lignins obtained (10 g) were then extracted with dioxane and water (96:4) in complete darkness at room temperature for 2 days [56]. The soluble fractions were obtained by filtration and lyophilized overnight. The isolated lignins were washed extensively with water and 6 mM EDTA aqueous solution (pH 8). This removed as much of the remaining low-molecular-weight saccharides and metal ions as possible [25]. The washed Björkman lignins were lyophilized again overnight and stored. The yields of isolated EFB lignin used in a previous study by Sun et al. [45] were 19.9 % of the original cell wall material. As seen in Table 1, the AcBr lignin contents, measured with ε_{280}=17.9, as previously described [57, 58], of the EFB-1 fibers were 22.7±0.4 %, giving a MWL yield on a lignin basis of ~88 %. The *p*Bz content of the MWL was 4.2±0.1 %. For the EFB-2 sample, the AcBr lignin content was 18.3±0.2 %.

Isolated lignins were acetylated using acetic anhydride and pyridine as previously described [59]. Solutions of the acetylated lignins in EtOAc were washed with 6 mM EDTA to more rigorously remove metal ion contaminants following methods described previously [25].

NMR of Lignins

Spectra have been run over many years on a variety of Bruker instruments, including DP 360 and Avance 400, 500 and 600 MHz instruments, using standard experiments and conditions [13]. The 3D total correlation spectroscopy-heteronuclear single-quantum coherence (TOCSY-HSQC) experiment was run as described previously [60, 55].

Derivatization Followed by Reductive Cleavage (DFRC) Method Adapted for the Analysis of Monolignol *p*-Hydroxybenzoate Conjugates

The derivatization followed by reductive cleavage (DFRC) method [61, 62] was adapted for HPLC analysis, as the monolignol *p*-hydroxybenzoate conjugates were found to be incompatible with GC-MS analysis. The modified procedure is detailed below.

In a 2-dram vial equipped with a polytetrafluoroethylene (PTFE) pressure-release cap and containing a stir bar was suspended each extract-free EFB whole-cell-wall sample (50 mg) in a 20 % solution of acetyl bromide in acetic acid (5 mL). The suspension was heated at 50 °C and allowed to stir for 3 h. The solvent was removed on a SpeedVac (Thermo Scientific SPD131DDA, 50 °C, 35 min, 1.0 Torr, 40 Torr/min, RC lamp on). The crude film was suspended in absolute ethanol (1 mL), and the ethanol was then removed on the SpeedVac (50 °C, 15 min, 6.0 Torr, 30 Torr/min, RC lamp on). To the residue was added a mixture of dioxane:acetic acid:water (5/4/1 *v*/*v*, 5 mL) and nano-powder zinc (250 mg). The sample was sonicated to ensure that the biomass film was dissolved and the zinc was in suspension. The reaction was stirred in the sealed vial in the dark at room temperature for 16–20 h. Additional nano-powder zinc was added to the vial as required to maintain a fine suspension. The reaction was quenched with a mixture of dichloromethane (DCM, 6 mL), saturated ammonium chloride (15 mL), and internal standard (ISTD, diethyl 5-5′-diferulate diacetate, DEDF, 54.0 μg). After isolating the organic fraction using a separatory funnel, the aqueous phase was extracted with DCM (3×10 mL). The combined organic fractions were dried over anhydrous sodium sulfate, filtered, and the solvent evaporated in vacuo. The free hydroxyl groups were acetylated by adding pyridine/acetic anhydride (1/1 *v*/*v*, 5 mL) and allowing the mixture to sit in the

dark for 16 h, after which the solvent was removed to give an oily film.

To remove most of the polysaccharide-derived products, the crude DFRC product was dissolved in ethyl acetate (EtOAc, 0.15 mL), diluted with hexanes (0.15 mL), and loaded onto an SPE cartridge (Supelco Supelclean LC-Si SPE tube, 3 mL, P/N: 505048). The products were eluted with hexanes:ethyl acetate (1:1, 8 mL), and the solvent was removed on a rotary evaporator. The product was dissolved in ethyl acetate and filtered through a 0.2-μm PTFE filter into a high-performance liquid chromatography (HPLC) vial. Analysis of the purified products was performed on an HPLC–PDA-ESI-MS (Shimadzu LC-MS 2020) with an amino stationary-phase column (Phemonenex Luna NH_2, 100 Å, 250 mm×4.60 mm×5 μm). The mobile phase was hexanes and EtOAc with 0.1 % triethylamine, running at 0.7 mL/min, holding at 30 % EtOAc for 1 min then ramping at 1.2 %/min up to 95 % EtOAc. The column eluent was split 80:20 between a PDA detector (265–400 nm) and an ESI-MS detector operating in positive-ion mode. External synthetic standards of the predominantly *trans*-isomers were used for determination of the retention time, calibration of the PDA response, and authentication of each compound by MS, Table 2. The DFRC data is also summarized in Table 1.

Syntheses of γ-*p*-Hydroxybenzoylated β-Ether Model Compounds

As we determined by NMR and DFRC that *p*-hydroxybenzoates acylated essentially only S units, the prime model compounds of interest were the etherified and the free-phenolic S β-ether γ-*p*-hydroxybenzoate models **7** and their peracetates **8**. These model compounds were prepared following the synthetic route shown in Scheme 1.

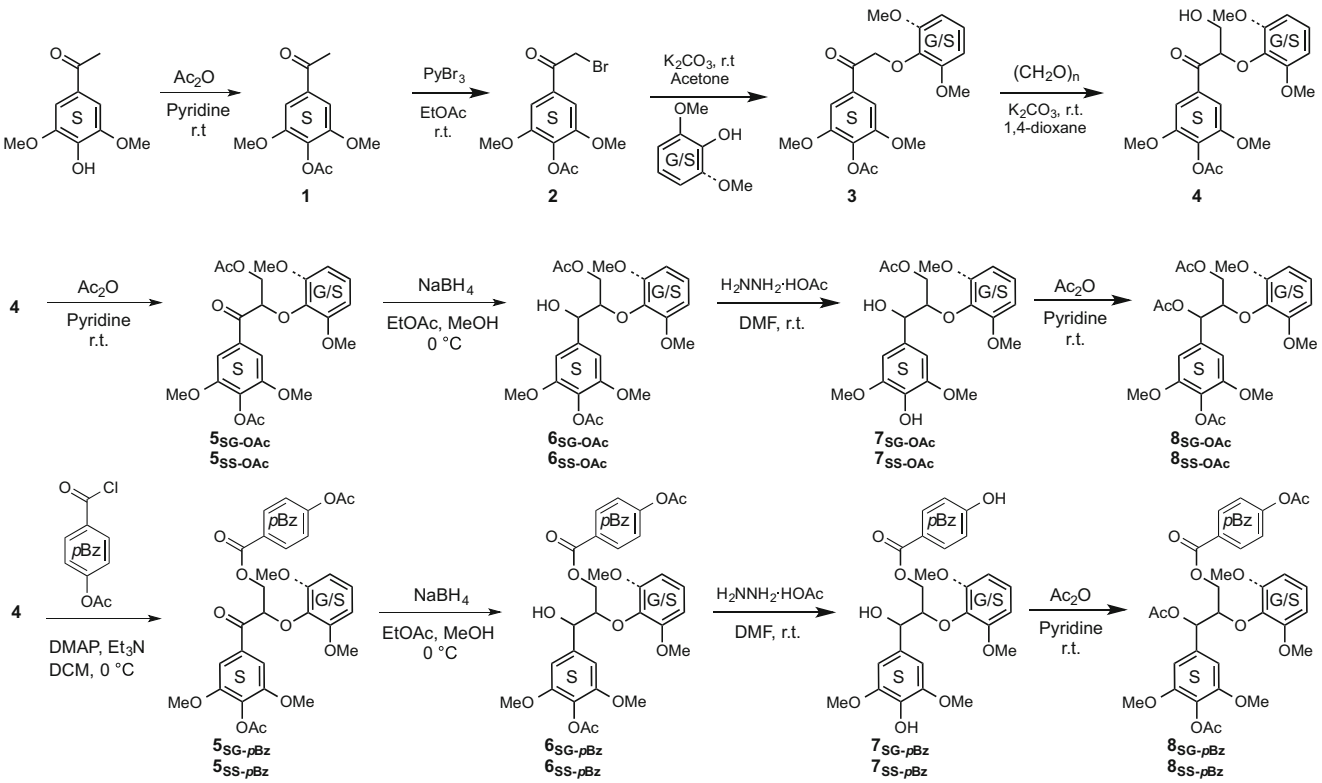
Preparation of 4-Acetoxy-3,5-dimethoxyacetophenone **1** To a solution of the starting acetophenone (3,5-dimethoxy-4-hydroxy-acetophenone, 51.0 mmol, 10.0 g) in pyridine (30 mL) was added Ac_2O (30 mL). The solution was allowed to stir overnight at room temperature. Upon confirmation by TLC that the starting material had been fully consumed, the reaction mixture was diluted with EtOAc (300 mL) and washed successively with 1 M HCl (3×150 mL), sat. $NaHCO_3$ (1×150 mL), and sat. NaCl (1×100 mL). The organic layer was dried over Na_2SO_4 and concentrated in vacuo. Residual AcOH was removed by azeotropic distillation with toluene (3×50 mL), followed by removal of the toluene with MeOH co-distillation (2×50 mL). Recrystallization from 20 % EtOAc in Et_2O afforded **1** (11.0 g, 91 %) as a pale yellow crystalline solid: ^{1}H NMR ($CDCl_3$) δ 7.23 (s, 2/6), 3.89 (s, 6H, OMe), 2.61 (s, 3H, β), 2.36 (s, 3H, OAc); ^{13}C NMR ($CDCl_3$) δ 197.0 (α), 168.4 (OAc), 152.4 (3/5), 135.3 (1), 132.9 (4), 105.2 (2/6), 56.5 (OMe), 26.7 (β), 20.6 (OAc); HRMS (ESI) *m/z* calcd. for $[M+NH_4]^+$=256.1180, found 256.1180.

Bromination of compound **1** *(preparation of compound* **2***)* Compound **1** (46.0 mmol, 11.0 g) was dissolved in EtOAc (460 mL), and the reaction solution was cooled to 0 °C. Pyridinium tribromide (55.4 mmol, 20.0 g) was charged into the reaction flask incrementally over 10 min, and the reaction was allowed to stir for 2 h, with the temperature gradually warming from 0 °C to room temperature (23 °C). If any starting material remained detectable by TLC at that point, additional portions of pyridinium tribromide (in increments of approximately 3 mmol, or 1 g) were added until TLC analysis showed full conversion. The reaction was then quenched with a solution of sat. $NaHCO_3$ (400 mL) and, after separation, the aqueous layer was extracted with DCM (1×100 mL). The combined organic layers were then washed with 1 M HCl

Table 2 HPLC-PDI-ESI-MS elution parameters and observed mass-ion-ratios for parent or fragment ions from synthetic standards CA, SA, CA-*p*Bz, SA-*p*Bz, and internal recovery standard, DEDF (see Fig. 4 for structures of the peracetates of these compounds)

Compound name	Retention time (min)	Formula	MW	m/z (*Rel. Int.*)	Fragment	(*Rel. Int.*)	Fragment
Coniferyl alcohol diacetate (CA)	*cis*: 8.80 *trans*: 9.90	$C_{14}H_{16}O_5$	264.1	306 (100)	$[M\text{-}OAc+Et_3N]^+$	205 (36)	$[M\text{-}AcOH]^+$
Sinapyl alcohol diacetate (SA)	*cis*: 10.65 *trans*: 12.04	$C_{15}H_{18}O_6$	294.11	336 (100)	$[M\text{-}OAc+Et_3N]^+$	235 (60)	$[M\text{-}AcOH]^+$
Coniferyl *p*-hydroxybenzoate diacetate (CA-*p*Bz)	*cis*: 13.60 *trans*: 14.96	$C_{21}H_{20}O_7$	384.12	306 (100)	$[M\text{-}pBz+Et_3N]^+$	486 (7)	$[M+Et_3N+H]^+$
Sinapyl *p*-hydroxybenzoate diacetate (SA-*p*Bz)	*cis*: 16.00 *trans*: 17.76	$C_{22}H_{22}O_8$	414.13	336 (100)	$[M\text{-}pBz+Et_3N]^+$	516 (35)	$[M+Et_3N+H]^+$
Diethyl 5,5′-diferulate diacetate (DEDF)[a]	14.13	$C_{29}H_{30}O_{10}$	526.18	527 (100)	$[M+H]^+$	351 (45)	$[Unknown]^+$

[a] DEDF was used as a control internal standard, spiked at 54.0 μg during the aqueous workup of the zinc reduction step. m/z columns list the observed mass of the most intense parent or fragment ions (usually loss of the γ-acyl group and Et_3N adduct)



Scheme 1 Synthetic scheme for model compounds. [Note that, in the NMR assignments in the Materials and Methods section, compounds **7** and **8** have rings labeled as *A*, the lower S ring; *B*, the middle S/G ring; and *C*, the *p*Bz moiety]

(2×200 mL) and sat. NaCl (1×100 mL), dried over Na_2SO_4, and concentrated under reduced pressure. After crystallization from 30 % EtOAc/hexanes, **2** (9.2 g, 63 %) was obtained as a pale yellow crystalline solid: ^{1}H NMR ($CDCl_3$) δ 7.25 (s, 1H, 2/6$_A$), 4.42 (s, 2H, β), 3.89 (s, 6H, OMe$_A$), 2.36 (s, 3H, OAc$_A$); ^{13}C NMR ($CDCl_3$) δ 190.4 (α), 168.3 (OAc), 152.6 (3/5), 133.6 (1), 132.0 (4), 105.9 (2/6), 56.5 (OMe), 30.6 (β), 20.6 (OAc); HRMS (ESI) *m*/*z* calcd. for $[M+NH_4]^+$= 334.0285, found 334.0289.

Etherification of bromide **2** *(illustrated for preparation of compound* $\mathbf{3_{SG}}$*)* Upon dissolution of compound **2** (10.0 mmol, 3.20 g) in acetone (100 mL), K_2CO_3 (40.0 mmol, 5.5 g) and NaI (0.1 mmol, 15.0 mg) were charged into the reaction flask. After adding guaiacol (11.0 mmol, 1.4 g), the reaction mixture was heated to reflux and allowed to stir for 3 h. At that point, TLC showed complete conversion and after the reaction mixture was allowed to cool to room temperature, the acetone was removed in vacuo. The solids were then reconstituted in EtOAc (200 mL) and H_2O (100 mL), and the aqueous layer removed. The organic layer was washed with 1 M HCl (1× 100 mL) and sat. NaCl (1×75 mL), dried over Na_2SO_4, and concentrated in vacuo to a viscous oil which, upon allowing to stand in the freezer overnight, crystallized. Recrystallization with 20 % EtOAc/hexanes afforded $\mathbf{3_{SG}}$ (2.74 g, 76 %) as a white solid: ^{1}H NMR ($CDCl_3$) δ 7.00–6.96 (m, 1H, 1$_B$), 6.92 (dd, J=8.1, 1.1 Hz, 1H, 2$_B$), 6.89–6.83 (m, 2H, 5$_B$/6$_B$), 5.28 (s, 2H, β), 3.88 (s, 9H, OMe), 2.36 (s, 3H, OAc$_A$); ^{13}C NMR ($CDCl_3$) δ 193.9 (α), 168.3 (OAc$_A$), 152.5 (3$_A$/5$_A$), 149.8 (3$_B$), 147.4 (4$_B$), 133.4 (4$_A$), 132.6 (1$_A$), 122.8 (1$_B$), 121.0 (6$_B$), 114.8 (5$_B$), 112.2 (2$_B$), 105.3 (2$_A$/6$_A$), 72.6 (β), 56.5 (OMe$_A$), 55.9 (OMe$_B$), 20.6 (OAc$_A$); HRMS (ESI) *m*/*z* calcd. for $[M+H]^+$=361.1282, found 361.1286.

Compound $\mathbf{3_{SS}}$ (2.75 g, 71 %), white solid: ^{1}H NMR ($CDCl_3$) δ 7.38 (s, 2H, 2$_A$/6$_A$), 7.03 (t, J=8.4 Hz, 1H, 4$_B$), 6.59 (d, J=8.4 Hz, 2H, 3$_B$/5$_B$), 5.14 (s, 2H, β), 3.88 (s, 6H, OMe$_A$), 3.81 (s, 6H, OMe$_B$), 2.36 (s, 3H, OAc$_A$); ^{13}C NMR ($CDCl_3$) δ 194.1 (α), 168.4 (OAc$_A$), 153.4 (3$_B$/5$_B$), 152.4 (3$_A$/5$_A$), 136.4 (4$_B$), 133.2 (4$_A$), 133.1 (1$_A$), 124.4 (1$_B$), 105.5 (2$_A$/6$_A$), 105.3 (2$_B$/6$_B$), 75.6 (β), 56.5 (OMe$_A$), 56.2 (OMe$_B$), 20.6 (OAc$_A$); HRMS (ESI) *m*/*z* calcd. for $[M+Na]^+$=413.1207, found 413.1199.

Procedure for hydroxymethylation of **3** *(illustrated for preparation of compound* $\mathbf{4_{SG}}$*)* The improved method using dioxane as solvent as described previously [63] was used. Ether $\mathbf{3_{SG}}$ (7.60 mmol, 2.74 g) was dissolved in 1,4-dioxane (76 mL), and K_2CO_3 (30.4 mmol, 4.20 g) was charged into the reaction flask. After adding formaldehyde (9.1 mmol, 0.7 mL, 37 wt% in H_2O) to the flask, the reaction mixture was allowed to stir for approximately 4 h at room temperature, until the starting material could no longer be detected by TLC. The

organics were concentrated in vacuo, with all residual solids then being reconstituted in EtOAc (200 mL) and H_2O (100 mL). The layers were separated, and the aqueous layer was extracted with EtOAc (1×100 mL). After combining the organic layers, they were washed with 1 M HCl (1×150 mL) and sat. NaCl (1×100 mL), dried over Na_2SO_4, and evaporated under reduced pressure. Flash column chromatography (100 g SiO_2) using gradient elution from 10 to 75 % EtOAc/hexane over 30 column volumes afforded **4_{SG}** (2.28 g, 77 %) as a light yellow brittle foam: ^{1}H NMR ($CDCl_3$) δ 7.37 (s, 2H, $2_A/6_A$), 7.02 (td, J=7.8, 1.2 Hz, 1H, 1_B), 6.92 (dd, J=8.1, 1.2 Hz, 1H, 2_B), 6.91 (dd, J=8.0, 1.5 Hz, 1H, 5_B), 6.85 (td, J=7.7, 1.2 Hz, 1H, 6_B), 5.36 (dd, J=5.8, 4.8 Hz, 1H, β), 4.09 (s, 2H, γ), 3.84 (s, 9H, $OMe_{A/B}$), 2.94 (s, 1H, OH), 2.35 (s, 3H, OAc_A); ^{13}C NMR ($CDCl_3$) δ 195.8 (α), 168.3 (OAc_A), 152.5 ($3_A/5_A$), 150.6 (3_B), 146.8 (4_B), 133.5 (4_A), 132.9 (1_A), 124.0 (1_B), 121.4 (6_B), 118.6 (5_B), 112.4 (2_B), 105.9 ($2_A/6_A$), 84.8 (β), 63.6 (γ), 56.5 (OMe_A), 55.9 (OMe_B), 20.6 (OAc_A); HRMS (ESI) *m/z* calcd. for $[M+NH_4]^+$=408.1653, found 408.1659.

Compound **4_{SS}** (2.25 g, 77 %), sticky white foam: ^{1}H NMR ($CDCl_3$) δ 7.39 (s, 2H, $2_A/6_A$), 7.04 (t, J=8.4 Hz, 1H, 1_B), 6.59 (d, J=8.4 Hz, 2H, $2_B/6_B$), 5.09 (dd, J=7.4, 3.2 Hz, 1H, β), 4.03 (ddd, J=11.9, 7.4, 4.5 Hz, 1H, γ), 3.94 (dd, J=9.0, 4.4 Hz, 1H, OH), 3.86 (s, 6H, OMe_A), 3.82 (dd, J=9.0, 3.1 Hz, 1H, γ), 3.75 (s, 6H, OMe_B), 2.36 (s, 3H, OAc_A); ^{13}C NMR ($CDCl_3$) δ 195.6 (α), 168.3 (OAc_A), 152.9 ($3_B/5_B$), 152.3 (3_A), 136.4 (4_B), 133.6 (4_A), 133.1 (1_A), 124.6 (1_B), 105.9 ($2_A/6_A$), 105.3 ($2_B/6_B$), 87.5 (β), 63.4 (γ), 56.5 (OMe_A), 56.1 (OMe_B), 20.6 (OAc_A); HRMS (ESI) *m/z* calcd. for $[M+H]^+$=421.1494, found 421.1494.

Procedure for 4-acetoxybenzoylation of the γ-OH in the β–O–4-linked dimers **4** *(illustrated for preparation of compound* **$5_{SG\text{-}pBz}$***)* Alcohol **4_{SG}** (1.3 mmol, 508 mg) was dissolved in dichloromethane (13 mL), and the resulting solution was cooled to 0 °C. DMAP (0.33 mmol, 40 mg) and Et_3N (3.9 mmol, 0.54 mL) were charged into the reaction mixture, followed by 4-acetoxybenzoyl chloride (*p*BzCl; 1.6 mmol, 318 mg). The reaction was allowed to warm to room temperature gradually with continued stirring for approximately 2 h. Once TLC showed no detectable starting material remaining, the mixture was poured into 1 M HCl (50 mL). After separation, the aqueous layer was extracted with DCM (1×30 mL), and the combined organics were washed with sat. NaCl (1×50 mL). They were then dried over Na_2SO_4 and concentrated in vacuo. Purification by flash column chromatography (50 g SiO_2, gradient elution from 5 to 50 % EtOAc/hexane) gave ester **$5_{SG\text{-}pBz}$** (547 mg, 76 %) as a white foam: ^{1}H NMR ($CDCl_3$) δ 8.00 (d, J=8.8 Hz, 2H, $2_C/6_C$), 7.56 (s, 2H, $2_A/6_A$), 7.15 (d, J=8.8 Hz, 2H, $3_C/5_C$), 7.00–6.95 (m, 2H, $1_B/5_B$), 6.87–6.80 (m, 2H, $2_B/6_B$), 5.74 (dd, J=7.3, 3.6 Hz, 1H, β), 4.99 (dd, J=12.1, 3.7 Hz, 1H, γ), 4.68 (dd, J=11.9, 7.3 Hz, γ), 3.86 (OMe_A), 3.67 (OMe_B), 2.36 (OAc_A), 2.32 (OAc_C); ^{13}C NMR ($CDCl_3$) δ 194.6 (α), 169.0 (OAc_C), 168.3 (OAc_A), 165.8 (9), 154.7 (4_C), 152.5 ($3_A/5_A$), 150.5 (3_B), 146.8 (4_B), 133.5 (4_A), 132.8 (1_A), 131.5 ($2_C/6_C$), 127.2 (1_C), 123.8 (1_B), 121.8 ($3_C/5_C$), 121.2 (6_B), 118.6 (5_B), 112.7 (2_B), 106.1 ($2_A/6_A$), 81.0 (β), 65.3 (γ), 56.5 (OMe_A), 55.7 (OMe_B), 21.3 (OAc_C), 20.6 (OAc_A); HRMS (ESI) *m/z* calcd. for $[M+NH_4]^+$=570.1970, found 570.1949.

Compound **$5_{SS\text{-}pBz}$** (381 mg, 69 %), white foam: ^{1}H NMR ($CDCl_3$) δ 7.90 (d, J=8.8 Hz, 2H, $2_C/6_C$), 7.57 (s, 2H, $2_A/6_A$), 7.12 (d, J=8.8 Hz, 2H, $3_C/5_C$), 6.97 (t, J=8.4 Hz, 1H, 4_B), 6.51 (d, J=8.4 Hz, 2H, $3_B/5_B$), 5.58 (dd, J=6.6, 4.1 Hz, 1H, β), 4.87 (dd, J=11.9, 4.0 Hz, 1H, γ), 4.72 (dd, J=11.9, 6.5 Hz, 1H, γ), 3.84 (s, 6H, OMe_A), 3.64 (s, 6H, OMe_B), 2.36 (s, 3H, OAc_A), 2.31 (s, 3H, OAc_C); ^{13}C NMR ($CDCl_3$) δ 195.0 (α), 169.0 (OAc_C), 168.3 (OAc_A), 165.6 (9), 154.5 (4_C), 153.1 ($3_B/5_B$), 152.2 ($3_A/5_A$), 135.7 (4_B), 133.6 (1_A), 133.2 (4_A), 131.4 ($2_C/6_C$), 127.4 (1_C), 124.4 (1_B), 121.7 ($3_C/5_C$), 106.5 ($2_A/6_A$), 105.1 ($2_B/6_B$), 82.2 (β), 65.1 (γ), 56.5 (OMe_A), 55.9 (OMe_B), 21.3 (OAc_C), 20.6 (OAc_A); HRMS (ESI) *m/z* calcd. for $[M+NH_4]^+$=600.2076, found 600.2092.

Procedure for acetylation of the γ-OH in the β–O–4-linked dimers **4** *(illustrated for preparation of compound* **$5_{SG\text{-}OAc}$***)* - Alcohol **4_{SG}** (1.0 mmol, 400 mg) was dissolved in equal parts pyridine and Ac_2O (2 mL/2 mL) and allowed to stir at room temperature overnight. The mixture was diluted with EtOAc (100 mL) and washed successively with 1 M HCl (3×50 mL), sat. $NaHCO_3$ (2×50 mL), and sat. NaCl (1×50 mL). Residual AcOH was removed by azeotropic distillation with toluene (3×25 mL), followed by removal of the toluene by co-distillation with MeOH (3×25 mL). Ester **$5_{SG\text{-}OAc}$** (407 mg, 94 %) was recovered as a white foam and used without further purification in the next step: ^{1}H NMR ($CDCl_3$) δ 7.49 ($2_A/6_A$), 6.99 (ddd, J=7.8, 7.8, 1.3 Hz, 1H, 1_B), 6.93 (dd, J=8.0, 1.6 Hz, 1H, 5_B), 6.88 (dd, J=8.2, 1.4 Hz, 1H, 2_B), 6.84 (ddd, J=7.7, 7.7, 1.5 Hz, 1H, 6_B), 5.64 (dd, J=7.5, 3.6 Hz, 1H, β), 4.73 (dd, J=12.0, 3.6 Hz, 1H, γ), 4.46 (dd, J=12.0, 7.5 Hz, 1H, γ), 3.87 (s, 6H, OMe_A), 3.76 (s, 3H, OMe_B), 2.35 (OAc_A), 2.07 (OAc_γ); ^{13}C NMR ($CDCl_3$) δ 194.5 (α), 171.2 (OAc_γ), 168.3 (OAc_A), 152.5 (OMe_A), 150.3 (OMe_B), 146.8 (4_B), 133.4 (4_A), 132.7 (1_A), 123.7 (1_B), 121.2 (6_B), 118.1 (5_B), 112.7 (2_B), 105.9 ($2_A/6_A$), 80.6 (β), 64.7 (γ), 56.5 (OMe_A), 55.9 (OMe_B), 21.0 (OAc_γ), 20.6 (OAc_A); HRMS(ESI) *m/z* calcd. for $[M+NH_4]^+$=450.1759, found 450.1740.

Compound **$5_{SS\text{-}OAc}$** (254 mg, 92 %), white foam: ^{1}H NMR ($CDCl_3$) δ 7.49 (s, 2H, $2_A/6_A$), 6.99 (t, J=8.4 Hz, 1H, 1_B), 6.54 (d, J=8.4 Hz, 2H, $2_B/6_B$), 5.46 (J=6.4, 4.6 Hz, 1H, β), 4.61 (dd, J=11.7, 4.6 Hz, 1H, γ), 4.53 (dd, J=11.7, 6.4 Hz, 1H, γ), 3.85 (s, 6H, OMe_A), 3.72 (s, 6H, OMe_B), 2.35 (OAc_A), 1.97 (OAc_γ); ^{13}C NMR ($CDCl_3$) δ 194.8 (α), 171.0 (OAc_γ), 168.3 (OAc_A), 153.1 (OMe_B), 152.2 (OMe_A), 135.9 (4_B), 133.6

(1_A), 133.1 (4_A), 124.4 (1_B), 106.3 ($2_A/6_A$), 105.3 ($2_B/6_B$), 81.7 (β), 64.4 (γ), 56.4 (OMe_A), 56.0 (OMe_B), 20.9 (OAc_γ), 20.6 (OAc_A); HRMS(ESI) *m/z* calcd. for $[M+NH_4]^+$= 480.1865, found 480.1884.

Typical procedure for reduction of α-ketones **5** *(illustrated for preparation of the compound* $\mathbf{6_{SG\text{-}pBz}}$*)* Ester $\mathbf{5_{SG\text{-}pBz}}$ (0.43 mmol, 250 mg) was dissolved in EtOAc (4.4 mL) and MeOH (0.4 mL), and the reaction mixture was cooled to 0 °C. $NaBH_4$ (0.86 mmol, 33 mg) was charged to the flask and allowed to stir continuously for approximately 1 h. After confirming complete consumption of the starting material by TLC, the organics were diluted with EtOAc (25 mL), washed with 1 M HCl (1×30 mL) and sat. NaCl (1×30 mL), dried over Na_2SO_4, and concentrated in vacuo. The resulting clear, viscous oil $\mathbf{6_{SG\text{-}pBz}}$ was used crude in the final deprotection, with an assumed 100 % yield; products $\mathbf{6_{SG\text{-}pBz}}$, $\mathbf{6_{SS\text{-}pBz}}$, $\mathbf{6_{SG\text{-}OAc}}$, and $\mathbf{6_{SS\text{-}OAc}}$ were formed and used in the same way.

Typical procedure for deprotection of phenolic acetates (illustrated for preparation of the compound $\mathbf{7_{SG\text{-}pBz}}$*)* Acetates were removed essentially as in the recently described method [53]. Benzyl alcohol $\mathbf{6_{SG\text{-}pBz}}$ (0.32 mmol, 175 mg) was dissolved in DMF (1.3 mL). Hydrazine acetate (1.1 mmol, 103 mg) was charged to the flask, and the reaction was allowed to stir at room temperature for 1 to 3 h. The solution was diluted with EtOAc (50 mL) and washed with H_2O (5× 15 mL) and NaCl (1×10 mL). After drying over Na_2SO_4, the organics were evaporated under reduced pressure. The resulting crude oil was purified by flash column chromatography (10 g SiO_2, gradient elution using either 10 to 70 % EtOAc/hexane or 1 to 10 % MeOH in DCM) to afford $\mathbf{7_{SG\text{-}pBz}}$ (69 mg, 46 %) as a brittle white foam and a 70:30 mixture of *erythro-* (*anti-*) and *threo-* (*syn-*) isomers: ^{1}H NMR ($CDCl_3$) δ 7.85 (d, J=8.7 Hz, 0.7H, $2_{C,t}/6_{C,t}$), 7.81 (d, J= 8.7 Hz, 2H, $2_{C,e}/6_{C,e}$), 7.18 (dd, J=7.9, 1.5 Hz, 0.3H, $5_{B,t}$), 7.10–7.04 (m, 2.4H, $5_{B,e}$, $2_{B,t}$, $1_{B,e}$), 6.96–6.89 (m, 2.7H, $2_{B,e}/6_{B,t}$, $6_{B,e}$, $1_{B,t}$), 6.83 (d, J=8.8 Hz, 0.8H, $3_{C,t}/5_{C,t}$), 6.81 (d, J= 8.7 Hz, 2H, $3_{C,e}/5_{C,e}$), 6.66 (s, 2H, $2_{A,e}/6_{A,e}$), 6.63 (s, 0.8H, $2_{A,t}/6_{A,t}$), 5.55 (s, 0.3H, $ArOH_t$), 5.48 (s, 1.3H, $ArOH_{e/t}$), 5.46 (s, 1H, $ArOH_e$), 4.99–4.92 (m, 1.4H, α_t, α_e), 4.60–4.54 (m, 2H, γ_e, β_e), 4.46 (dd, J=12.1, 3.5 Hz, 0.3H, γ_t), 4.42–4.36 (m, 1H, γ_e), 4.32 (quint, J=4.1 Hz, 0.4H, β_t), 4.28–4.23 (m, 0.7H, γ_t, αOH_t), 3.89 (s, 1H, $OMe_{B,t}$), 3.87 (s, 3H, $OMe_{B,e}$), 3.86 (s, 6H, $OMe_{A,e}$), 3.83 (d, J=3.8 Hz, 1H, αOH_e), 3.80 (s, 2H, $OMe_{a,t}$); ^{13}C NMR ($CDCl_3$) δ 166.45 ($7_{C,e}$), 166.11 ($7_{C,t}$), 160.59 ($4_{C,t}$), 160.44 ($4_{C,e}$), 151.55 ($3_{B,e}$), 150.93 ($3_{B,t}$), 147.75 ($4_{B,t}$), 147.21 ($3_{A,t}/5_{A,t}$), 147.14 ($3_{A,e}/5_{A,e}$), 147.01 ($4_{B,e}$), 134.78 ($4_{A,t}$), 134.21 ($4_{A,e}$), 132.06 ($2_{C,t}/6_{C,t}$, $2_{C,e}/6_{C,e}$), 130.19 ($1_{A,t}$), 130.14 ($1_{A,e}$), 124.31 ($1_{B,t}$), 124.30 ($1_{B,e}$), 122.07 ($1_{C,e}$), 121.89 ($1_{C,t}$), 121.66 ($B_{6,e}$), 121.61 ($B_{6,t}$), 120.75 ($5_{B,e}$), 120.56 ($5_{B,t}$), 115.42 ($3_{C,t}/5_{C,t}$), 115.32 ($3_{C,e}/5_{C,e}$), 112.34 ($2_{B,t}$), 112.32 ($2_{B,e}$), 103.87 ($2_{A,t}/6_{A,t}$), 103.09 ($2_{A,e}/6_{A,e}$), 85.98 (β_t), 84.36 (β_e), 75.07 (α_t), 72.60 (α_e), 63.44 (γ_t), 62.95 (γ_e), 56.41 ($OMe_{A,e}$), 56.33 ($OMe_{A,t}$), 55.96 ($OMe_{B,e}$), 55.91 ($OMe_{B,t}$); HRMS (ESI) *m/z* calcd. for $[M+NH_4]^+$=488.1916, found 488.1928.

Compound $\mathbf{7_{SS\text{-}pBz}}$ (201 mg, 93 %), brittle white foam: ^{1}H NMR ($CDCl_3$) δ 7.86 (d, J=8.6 Hz, 2H, $2_{C,t}/6_{C,t}$), 7.83 (d, J= 8.6 Hz, 1.4H, $2_{C,e}/6_{C,e}$), 7.05 (t, J=8.8 Hz, 0.8H, $1_{B,e}$), 7.04 (t, J=8.8 Hz, 1H, $1_{B,t}$), 6.82 (d, J=8.7 Hz, 2H, $3_{C,t}/5_{C,t}$), 6.79 (d, J=8.8 Hz, 1.5H, $3_{C,e}/5_{C,e}$), 6.63–6.58 (m, 7H, $2_{A,e}/6_{A,e}$, $2_{A,t}/6_{A,t}$, $2_{B/e}/6_{B,e}$, $2_{B,t}/6_{B,t}$), 5.55 (s, 1.3H, $ArOH_{e/t}$), 5.15 (d, J= 8.1 Hz, 1H, α_t), 5.06 (s, 1H, αOH_t), 4.93 (s, 1H, α_e), 4.67 (dd, J=13.0, 3.9 Hz, 1H, γ_t), 4.61–4.55 (m, 1.3H, β_e, γ_e), 4.52 (d, J=2.7 Hz, 0.6H, αOH_e), 4.42–4.36 (m, 0.6H, γ_e), 4.16–4.12 (m, 2H, β_t, γ_t), 3.82 (s, 6H, $OMe_{B,t}$), 3.81 (s, 4.7H, $OMe_{A,e}$), 3.78 (s, 4.2H, $OMe_{B,e}$), 3.70 (s, 6H, $OMe_{A,t}$); ^{13}C NMR ($CDCl_3$) δ 166.59 ($7_{C,e}$), 166.31 ($7_{C,t}$), 160.85 ($4_{C,t}$), 160.60 ($4_{C,e}$), 153.76 ($3_{B,e}/5_{B,e}$), 152.96 ($3_{B,t}/5_{B,t}$), 147.13 ($3_{A,e}/5_{A,e}$), 147.20 ($3_{A,t}/5_{A,t}$), 136.63 ($4_{B,t}$), 134.71 ($4_{B,e}$), 134.59 ($4_{A,t}$), 133.98 ($4_{A,e}$), 132.01 ($2_{C,e}/6_{C,e}$), 131.99 ($2_{C,t}/6_{C,t}$), 130.34 ($1_{A,t}$), 129.78 ($1_{A,e}$), 124.59 ($1_{B,e}$), 124.51 ($1_{B,t}$), 122.24 ($1_{C,e}$), 121.85 ($1_{C,t}$), 115.39 ($3_{C,t}/5_{C,t}$), 115.22 ($3_{C,e}/5_{C,e}$), 105.33 ($2_{B,e}/6_{B,e}$), 105.30 ($2_{B,t}/6_{B,t}$), 103.80 ($2_{A,t}/6_{A,t}$), 102.83 ($2_{A,e}/6_{A,e}$), 87.30 (β_t), 83.42 (β_e), 75.18 (α_t), 72.18 (α_e), 64.38 (γ_t), 62.87 (γ_e), 56.41 ($OMe_{A,e}$), 56.22 ($OMe_{A,t}$), 56.19 ($OMe_{B,e}$), 56.11 ($OMe_{B,t}$); HRMS(ESI) calcd. for *m/z* $[M+NH_4]^+$= 518.2021, found 518.2025.

Compound $\mathbf{7_{SG\text{-}OAc}}$ (75 mg, 84 %), white foam: ^{1}H NMR ($CDCl_3$) δ 7.15 (dd, J=7.8, 1.3 Hz, 0.3H, $5_{B,t}$), 7.09–7.03 (m, 2.4H, $1_{B,e}$, $2_{B,t}$, $5_{B,e}$), 6.96–6.91 (m, 2.8H, $1_{B,t}$, $2_{B,e}$, $6_{B,e}$, $6_{B,t}$), 6.62 (s, 2H, $2_{A,e}/6_{A,e}$), 6.60 (s, 0.7H, $2_{A,t}/6_{A,t}$), 5.55 (s, 0.3H, $ArOH_t$), 5.52 (s, 1H, $ArOH_e$), 4.86 (s, 1H, α_e), 4.82 (d, J= 8.0 Hz, 0.3H, α_t), 4.44–4.36 (m, 2H, β_e, γ_e), 4.31–4.24 (m, 0.5H, αOH_t, γ_t), 4.24–4.18 (m, 0.5H, β_t), 4.15–4.09 (m, 1H, γ_e), 4.05 (dd, J=11.9, 5.5 Hz, 0.3H, γ_t), 3.91 (s, 1H, $OMe_{B,t}$), 3.88 (s, 3H, $OMe_{B,e}$), 3.87 (s, 6H, $OMe_{A,e}$), 3.86 (s, 3H, $OMe_{A,t}$), 3.77 (s, 1H, αOH_e), 2.04 (s, 1.5H, γOAc_t), 2.01 (s, 3H, γOAc_e); ^{13}C NMR ($CDCl_3$) δ 171.32 (γOAc_e), 170.72 (γOAc_t), 151.69 ($3_{B,e}$), 150.92 ($3_{B,t}$), 148.14 ($4_{B,t}$), 147.18 ($3_{A,t}/5_{A,t}$), 147.11 ($3_{A,e}/5_{A,e}$), 147.00 ($4_{B,e}$), 134.77 ($4_{A,t}$), 134.17 ($4_{A,e}$), 130.35 ($1_{A,t}$), 130.10 ($1_{A,e}$), 124.31 ($1_{B,e}$), 124.24 ($1_{B,t}$), 121.59 ($6_{B,e}$), 121.57 ($6_{B,t}$), 120.85 ($5_{B,e}$), 120.55 ($5_{B,t}$), 112.27 ($2_{B,e}$), 112.23 ($2_{B,t}$), 103.82 ($2_{A,t}/6_{A,t}$), 102.92 ($2_{A,e}/6_{A,e}$), 86.32 (β_t), 84.58 (β_e), 74.74 (α_t), 72.21 (α_e), 63.46 (γ_t), 62.90 (γ_e), 56.42 ($OMe_{A,t}$, $OMe_{A,e}$), 55.96 ($OMe_{B,e}$), 55.92 ($OMe_{B,t}$), 21.19 (γOAc_e), 20.94 (γOAc_t); HRMS (ESI) *m/z* calcd. for $[M+NH_4]^+$=410.1810, found 410.1817.

Compound $\mathbf{7_{SS\text{-}OAc}}$ (143 mg, 79 %), white foam: ^{1}H NMR ($CDCl_3$) δ 7.06 (t, J=8.4 Hz, 1H, $1_{B,e}$), 7.02 (t, J=8.4 Hz, 1H, $1_{B,t}$), 6.63 (d, J=8.4 Hz, 2H, $2_{B,e}/6_{B,e}$), 6.59 (d, J=8.4 Hz, 2H, $2_{B,t}/6_{B,t}$), 6.57 (s, 2H, $2_{A,t}/6_{A,t}$), 6.57 (s, 2H, $2_{A,e}/6_{A,e}$), 5.56 (s, 1H, $ArOH_t$), 5.53 (s, 1H, $ArOH_e$), 4.99 (dd, J=8.5, 1.8 Hz, 1H, α_t), 4.82 (t, J=3.0 Hz, 1H, α_e), 4.80 (d, J=1.9 Hz, 1H,

αOH_t), 4.54–4.48 (m, 2H, β_e, γ_t), 4.37 (dd, J=12.0, 8.1 Hz, 1H, γ_e), 4.33 (d, J=2.9 Hz, 1H, αOH_e), 4.14 (dd, J=12.0, 3.5 Hz, 1H, γ_e), 4.00 (dt, J=8.4, 3.1 Hz, 1H, β_t), 3.90 (dd, J=12.0, 2.9 Hz, 1H, γ_t), 3.87 (s, 6H, $OMe_{A,e}$), 3.85 (s, 12H, $OMe_{B,e}$, $OMe_{B,t}$), 3.83 (s, 6H, $OMe_{A,t}$), 2.05 (s, 3H, OAc_t), 1.99 (s, 3H, OAc_e); ^{13}C NMR ($CDCl_3$) δ 171.03 (γOAc_e), 170.78 (γOAc_t), 153.79 ($3_{B,e}/5_{B,e}$), 152.86 ($3_{B,t}/5_{B,t}$), 147.06 ($3_{A,e}/5_{A,e}$), 147.02 ($3_{A,t}/5_{A,t}$), 136.70 ($4_{B,t}$), 134.52 ($4_{B,e}$), 134.51 ($4_{A,t}$), 133.87 ($4_{A,e}$), 130.67 ($1_{A,t}$), 129.74 ($1_{A,e}$), 124.55 ($1_{B,e}$), 124.36 ($1_{B,t}$), 105.46 ($2_{B,e}/6_{B,e}$), 105.20 ($2_{B,t}/6_{B,t}$), 103.72 ($2_{A,t}/6_{A,t}$), 102.57 ($2_{A,e}/6_{A,e}$), 86.87 (β_t), 82.99 (β_e), 74.57 (α_t), 71.61 (α_e), 64.07 (γ_t), 62.64 (γ_e), 56.41 ($OMe_{A,e}$), 56.31 ($OMe_{A,t}$), 56.30 ($OMe_{B,e}$), 56.07 ($OMe_{B,t}$), 20.99 (γOAc_e), 20.90 (γOAc_t); HRMS (ESI) *m/z* calcd. for $[M+NH_4]^+$=440.1916, found 440.1914.

Typical procedure for peracetylation (illustrated for preparation of compound $\mathbf{8_{SG\text{-}pBz}}$*)* Compound $\mathbf{7_{SG\text{-}pBz}}$ (10.0 mg, 20 μmol) was charged into a small amber vial, followed by addition of a solution containing equal parts pyridine and acetic anhydride (500 μL each). The resulting solution was allowed to stir overnight, at which point, all organic solvents were removed in vacuo to yield $\mathbf{8_{SG\text{-}pBz}}$ as a clear, colorless oil: ^1H NMR ($CDCl_3$) δ 7.98–7.94 (m, 2.7H, $2_{C,e}/6_{C,e}$, $2_{C,t}/6_{C,t}$), 7.16–7.13 (m, 2.7H, $3_{C,e}/5_{C,e}$, $3_{C,t}/5_{C,t}$), 7.02–6.96 (m, 1.7H, $1_{B,e}$, $2_{B,t}$, $5_{B,t}$), 6.89–6.81 (m, 3.8H, $1_{B,t}$, $2_{B,e}$, $5_{B,e}$, $6_{B,e}$, $6_{B,t}$), 6.72 (s, 2H, $2_{A,e}/6_{A,e}$), 6.69 (s, 0.7H, $2_{A,t}/6_{A,t}$), 6.18 (d, J=6.2 Hz, 0.3H, α_t), 6.13 (d, J=5.6 Hz, 1H, α_e), 4.84–4.80 (m, 1H, β_e), 4.76–4.72 (m, 0.3H, β_t), 4.60 (d, J=5.0 Hz, 2H, γ_e), 4.50 (dd, J=11.9, 4.2 Hz, 0.3H, γ_t), 4.37 (dd, J=11.9, 6.0 Hz, 0.3H, γ_t), 3.79 (s, 6H, $OMe_{A,e}$), 3.76 (s, 2H, $OMe_{A,t}$), 3.74 (s, 1H, $OMe_{B,t}$), 3.72 (s, 3H, $OMe_{B,e}$), 2.33 (s, 3H, $OAc_{A,t}$, $OAc_{C,t}$), 2.32 (s, 5H, $OAc_{A,e}$, $OAc_{C,e}$), 2.11 (s, 3H, αOAc_e), 2.09 (s, 1H, αOAc_t); ^{13}C NMR ($CDCl_3$) δ 169.91 (αOAc_t), 169.72 (αOAc_e), 169.00 ($OAc_{A,e}/OAc_{C,e}$), 168.98 ($OAc_{A,t}/OAc_{C,t}$), 168.75 ($OAc_{A,e}/OAc_{C,e}$), 168.70 ($OAc_{A,t}/OAc_{C,t}$), 165.63 ($7_{C,e}$), 165.47 ($7_{C,t}$), 154.59 ($4_{C,t}$), 154.56 ($4_{C,e}$), 152.33 ($3_{A,t}/5_{A,t}$), 152.16 ($3_{A,e}/5_{A,e}$), 151.25 ($3_{B,e}$), 150.99 ($3_{B,t}$), 148.05 ($4_{B,t}$), 147.30 ($4_{B,e}$), 135.08 ($1_{A,e}$), 134.89 ($1_{A,t}$), 131.45 ($2_{C,e}/6_{C,e}$), 131.40 ($2_{C,t}/6_{C,t}$), 128.87 ($4_{A,t}$), 128.76 ($4_{A,e}$), 127.38 ($1_{C,e}$), 127.31 ($1_{C,t}$), 123.85 ($1_{B,e}$), 123.58 ($1_{B,t}$), 121.83 ($3_{C,t}/5_{C,t}$), 121.79 ($3_{C,e}/5_{C,e}$), 121.13 ($6_{B,e}$, $6_{B,t}$), 119.88 ($5_{B,e}$), 119.13 ($5_{B,t}$), 112.62 ($2_{B,e}$), 112.51 ($2_{B,t}$), 104.48 ($2_{A,e}/6_{A,e}$), 104.16 ($2_{A,t}/6_{A,t}$), 80.80 (β_t), 80.56 (β_e), 75.08 (α_t), 74.47 (α_e), 64.07 (γ_t), 63.70 (γ_e), 56.32 ($OMe_{A,e}$, $OMe_{A,t}$), 55.80 ($OMe_{B,e}$, $OMe_{B,t}$), 21.32 ($OAc_{A,e}/OAc_{C,e}$), 21.28 (αOAc_t), 21.24 (αOAc_e), 20.64 ($OAc_{A,t}/OAc_{C/t}$).

Compound $\mathbf{8_{SS\text{-}pBz}}$ (colorless oil): ^1H NMR (acetone-d_6) δ 7.98 (d, J=8.8 Hz, 2H, $2_{C,t}/6_{C,t}$), 7.85 (d, J=8.8 Hz, 1.8H, $2_{C,e}/6_{C,e}$), 7.27 (d, J=8.8 Hz, 2H, $3_{C,t}/5_{C,t}$), 7.21 (d, J=8.8 Hz, 1.8H, $3_{C,e}/5_{C,e}$), 7.00 (t, J=8.4 Hz, 1H, $1_{B,t}$), 6.99 (t, J=8.4 Hz, 1H, $1_{B,e}$), 6.86 (s, 2H, $2_{A,t}/6_{A,t}$), 6.84 (s, 2H, $2_{A,e}/6_{A,e}$), 6.65 (d, J=8.4 Hz, 2H, $2_{B,t}/6_{B,t}$), 6.63 (d, J=8.4 Hz, 2H, $2_{B,e}/6_{B,e}$), 6.24 (d, J=6.5 Hz, 1H, α_t), 6.20 (d, J=4.4 Hz, 0.9H, α_e), 4.89–4.85 (m, 0.8H, β_e), 4.80–4.75 (m, 1H, β_t), 4.64 (dd, J=11.9, 6.3 Hz, 0.8H, γ_e), 4.52–4.47 (m, 2H, γ_e, γ_t), 4.23 (dd, J=11.9, 5.0 Hz, 1H, γ_t), 3.80 (s, 6H, $OMe_{A,e}$), 3.74 (s, 7H, $OMe_{B,t}$), 3.73 (s, 7H, $OMe_{B,e}$), 3.72 (s, 6H, $OMe_{A,t}$), 2.30 (s, 3H, OAc), 2.29 (s, 3H, OAc), 2.22 (s, 3H, OAc), 2.21 (s, 3H, OAc), 2.16 (s, 3H, αOAc_e), 1.99 (s, 3H, αOAc_t); ^{13}C NMR (acetone-d_6) δ 170.04 (αOAc_e), 169.95 (αOAc_t), 169.29 (OAc), 169.28 (OAc), 168.53 (OAc), 168.47 (OAc), 165.75 ($7_{C,e}$), 165.67 ($7_{C,t}$), 155.69 ($4_{C,t}$), 155.59 ($4_{C,e}$), 154.29 ($3_{B,e}/5_{B,e}$), 154.21 ($3_{B,t}/5_{B,t}$), 153.10 ($3_{A,e}/5_{A,e}$), 153.07 ($3_{A,t}/5_{A,t}$), 137.73 ($4_{B,e}$), 137.60 ($4_{B,t}$), 136.87 ($1_{A,e}$), 136.66 ($1_{A,t}$), 131.79 ($1_{C,t}$), 131.75 ($1_{C,e}$), 128.37 ($4_{A,e}$), 128.34 ($4_{A,t}$), 124.75 ($1_{B,e}$), 124.56 ($1_{B,t}$), 122.91 ($2_{C,t}/6_{C,t}$), 122.74 ($2_{C,e}/6_{C,e}$), 106.18 ($2_{B,t}/6_{B,t}$), 106.03 ($2_{B,e}/6_{B,e}$), 104.84 ($2_{A,t}/6_{A,t}$), 104.45 ($2_{A,e}/6_{A,e}$), 81.80 (β_e), 81.69 (β_t), 76.69 (α_t), 75.60 (α_e), 65.19 (γ_t), 64.22 (γ_e), 56.45 ($OMe_{A,e}$), 56.41 ($OMe_{B,t}$), 56.29 ($OMe_{B,e}$), 56.21 ($OMe_{A,t}$), 21.00 (OAc), 20.99 (OAc), 20.27 (OAc), 20.25 (OAc).

Compound $\mathbf{8_{SG\text{-}OAc}}$ (colorless oil): ^1H NMR ($CDCl_3$) δ 7.02–6.94 (m, 1.7H, $1_{B,e}$, $2_{B,t}$, $5_{B,t}$), 6.91–6.80 (m, 3.8H, $1_{B,t}$, $2_{B,e}$, $5_{B,e}$, $6_{B,t}$, $6_{B,e}$), 6.68 (s, 2H, $2_{A,e}/6_{A,e}$), 6.67 (s, 0.8H, $2_{A,t}/6_{A,t}$), 6.09 (d, J=6.3 Hz, 0.3H, α_t), 6.05 (d, J=5.7 Hz, 1H, α_e), 4.69–4.64 (m, 1H, β_e), 4.64–4.59 (m, 0.3H, β_t), 4.46 (dd, J=12.0, 5.7 Hz, 1H, γ_e), 4.32 (dd, J=11.8, 4.5 Hz, 0.3H, γ_t), 4.29 (dd, J=11.9, 3.8 Hz, 1H, γ_e), 4.09 (dd, J=11.8, 5.7 Hz, 0.3H, γ_t), 3.81 (s, 1H, $OMe_{B,t}$), 3.81 (s, 2H, $OMe_{A,t}$), 3.79 (s, 6H, $OMe_{A,e}$), 3.77 (s, 3H, $OMe_{B,e}$), 2.32 (s, 4H, $OAc_{A,e}$, $OAc_{A,t}$), 2.10 (s, 3H, αOAc_e), 2.07 (s, 1.1H, αOAc_t), 2.04 (s, 3H, γOAc_e), 2.01 (s, 1.1H, γOAc_t); ^{13}C NMR ($CDCl_3$) δ 171.04 (γOAc_e), 170.78 (γOAc_t), 169.90 (αOAc_t), 169.69 (αOAc_e), 168.75 ($OAc_{A,e}$), 168.72 ($OAc_{A,t}$), 152.28 ($3_{A,t}/5_{A,t}$), 152.10 ($3_{A,e}/5_{A,e}$), 151.16 ($3_{B,e}$), 150.88 ($3_{B,t}$), 148.07 ($4_{B,t}$), 147.28 ($4_{B,e}$), 135.02 ($1_{A,e}$), 134.95 ($1_{A,t}$), 128.85 ($4_{A,t}$), 128.72 ($4_{A,e}$), 123.76 ($1_{B,e}$), 123.46 ($1_{B,t}$), 121.10 ($6_{B,e}$, $6_{B,t}$), 119.52 ($5_{B,e}$), 118.76 ($5_{B,t}$), 112.59 ($2_{B,e}$), 112.50 ($2_{B,t}$), 104.47 ($2_{A,e}/6_{A,e}$), 104.16 ($2_{A,t}/6_{A,t}$), 80.46 (β_t), 80.36 (β_e), 74.89 (α_t), 74.08 (α_e), 63.23 (γ_t), 62.84 (γ_e), 56.35 ($OMe_{A,t}$), 56.31 ($OMe_{A,e}$), 55.89 ($OMe_{B,t}$), 55.87 ($OMe_{B,e}$), 21.24 (αOAc_t), 21.21 (αOAc_e), 20.97 (γOAc_e), 20.91 (γOAc_t), 20.64 ($OAc_{A,e}$, $OAc_{A,t}$).

Compound $\mathbf{8_{SS\text{-}OAc}}$ (colorless oil): ^1H NMR (acetone-d_6) δ 7.00 (t, J=8.4 Hz, 2H, $1_{B,e}$, $1_{B,t}$), 6.81 (s, 2H, $2_{A,t}/6_{A,t}$), 6.78 (s, 2H, $2_{A,e}/6_{A,e}$), 6.67 (d, J=8.3 Hz, 2H, $2_{B,t}/6_{B,t}$), 6.66 (d, J=8.3 Hz, 2H, $2_{B,e}/6_{B,e}$), 6.10 (d, J=7.0 Hz, 1H, α_t), 6.07 (d, J=4.2 Hz, 1H, α_e), 4.72 (dt, J=6.1, 4.2 Hz, 1H, β_e), 4.59 (dt, J=7.0, 4.2 Hz, 1H, β_t), 4.42 (dd, J=11.8, 6.1 Hz, 1H, γ_e), 4.28 (dd, J=11.9, 4.1 Hz, 1H, γ_t), 4.18 (dd, J=11.8, 4.1 Hz, 1H, γ_e), 3.91 (dd, J=11.9, 4.4 Hz, 1H, γ_t), 3.81 (s, 5H, $OMe_{B,e}$), 3.81 (s, 5H, $OMe_{B,t}$), 3.80 (s, 6H, $OMe_{A,e/t}$), 3.79 (s, 6H, $OMe_{A,e/t}$), 2.21 (s, 6H, $OAc_{A,t}$, $OAc_{A,e}$), 2.14 (s, 3H, αOAc_e), 1.98 (s, 3H, γOAc_t), 1.94 (s, 3H, αOAc_t), 1.86 (s, 3H,

γOAc_e); ^{13}C NMR (acetone-d_6) δ 170.70 (γOAc_e, γOAc_t), 169.98 (αOAc_e), 169.92 (αOAc_t), 168.53 ($OAc_{A,e}$), 168.50 ($OAc_{A,t}$), 154.30 ($3_{B,e}/5_{B,e}$), 154.11 ($3_{B,t}/5_{B,t}$), 153.04 ($3_{A,t}/5_{A,t}$), 153.01 ($3_{A,e}/5_{A,e}$), 137.80 ($4_{B,t}$), 136.81 ($1_{A,t}$), 136.72 ($1_{A,e}$), 136.53 ($4_{B,e}$), 129.44 ($4_{A,t}$), 129.20 ($4_{A,e}$), 124.74 ($1_{B,e}$), 124.45 ($1_{B,t}$), 106.16 ($2_{B,t}/6_{B,t}$), 106.08 ($2_{B,e}/6_{B,e}$), 104.78 ($2_{A,t}/6_{A,t}$), 104.38 ($2_{A,e}/6_{A,e}$), 81.64 (β_t), 81.46 (β_e), 76.75 (α_t), 75.36 (α_e), 64.24 (γ_t), 63.31 (γ_e), 56.49 (OMe), 56.44 (OMe), 56.31 (OMe), 56.29 (OMe), 20.98 (αOAc_t), 20.94 (αOAc_e), 20.69 (γOAc_t), 20.62 (γOAc_e), 20.26 ($OAc_{A,e}$, $OAc_{A,t}$).

Note that compounds $\mathbf{8_{SG\text{-}OAc}}$ and $\mathbf{8_{SS\text{-}OAc}}$, if they are the desired final products, are obviously directly available simply by reducing and acetylating compounds **4** as previously reported [64].

Syntheses of γ-*p*-Hydroxybenzoylated β–β-Unit Model Compounds

Peroxidase-H_2O_2 oxidative coupling reactions of sinapyl alcohol or/and acylated sinapyl alcohols were performed according to previously described procedures [16, 32, 14].

Syntheses of model compounds **9**–**11** Model compounds for **9a**, **10a**, and **11a** (Fig. 5) in which the 4–O-etherification is simply via a 4–O-methyl group were synthesized from sinapyl 4-acetoxybenzoate and sinapyl alcohol by peroxidase-H_2O_2 oxidation, followed by separation of the individual products, methylation (MeI-K_2CO_3, acetone), and acetylation (Ac_2O-Py). Analogous model compounds for **9b**, **10b**, and **11b** were obtained in the similar way starting from sinapyl acetate and sinapyl alcohol as previously described [14].

Results and discussion

The following sections address the various issues posed in the introduction regarding lignin *p*-hydroxybenzoylation in palms.

The Nature of *p*-Hydroxybenzoates and Acylation Regiochemistry

Determining whether *p*-hydroxybenzoates are in fact free-phenolic entities acylating the γ-positions of lignin side-chains is most readily deduced from solution-state NMR of isolated lignins, analogously to similar determinations for *p*-coumarates [25]. Björkman milled wood lignins (MWLs) were prepared from the oil palm (*Elaeis guineensis*) EFB fibers, fronds, and trunk fractions; the EFBs had already been shown to be particularly rich in *p*-hydroxybenzoates [34]. Simple saponification of EFB whole-cell-wall (WCW) samples indicated that the EFB-1 and EFB-2 samples contained 1.09 and 1.05 wt% *p*-hydroxybenzoic acid or, based on AcBr lignin contents of 22.7 and 18.3 %, ~4.8 and 5.8 % of the lignins (Table 1). NMR of the underivatized (i.e., un-acetylated) isolated lignin was made difficult due to the high levels of iron in this sample, as previously reported [65]; unfortunately, we were not able to sufficiently effect its removal by complexation with EDTA. Nevertheless, ^{13}C NMR (Fig. 1) readily revealed the presence of *p*-hydroxybenzoates, assumed at this point to be on lignins. It is worth noting that these signals have sometimes been misinterpreted as arising from *p*-hydroxyphenyl lignin units or *p*-coumarates [43]. Acetylation of the isolated lignin facilitated removal of the metal contaminants. The observed changes in chemical shifts of the aromatic carbons following acetylation indicated that *p*-hydroxybenzoates were present overwhelmingly in their free-phenolic form on the lignin, Fig. 1c–d. The sago palm lignin, used in a previous study in one of our labs [35], had similar features, Fig. 1a; in fact the *p*-hydroxybenzoate level was highest on this particular lignin fraction. Also striking, and not observed in previous studies, was the high level of natural acetylation of the EFB lignin, Fig. 1b–c. The presence of natural acetates on lignins is frequently missed for a variety of reasons, including that lignins are routinely peracetylated to improve NMR properties, extraction conditions may hydrolyze the acetates [35], and lignin purification steps involving acetic acid can introduce artifactual acetates. Circumventing some of these issues, solid-state NMR spectra of various palm fractions indicate the presence of acetates, Fig. 1b. Furthermore, acetates are clearly observable in spectra from other published reports, where they have typically been attributed to polysaccharides [66, 42]; historically, the contribution from lignin may have therefore been overlooked.

The gradient-enhanced 2D HSQC NMR experiment provides a highly detailed view of the molecular structure of plant cell walls with minimal processing of the sample [67]. The fast relaxation caused by the metals lead to extremely poor spectra, Fig. 2a–b, but the chemical shift data are consistent with acylation of the lignin. The *p*-hydroxybenzoate (*p*Bz) groups are revealed in the aromatic regions of Fig. 2b, along with the G and predominant S aromatics. The side-chain region, Fig. 2a, offers good evidence that lignin γ-hydroxyls, and not the alternative α-hydroxyls, are acylated; acylated β-ether α-H/C correlations, for example, would be seen at ~6.2/76 ppm. After peracetylation of the lignins, followed by successive extraction into chloroform and washing with EDTA in order to remove metal contaminants, higher-resolution spectra, Fig. 2c–d, are obtained. At this point, the nature of the acylation (i.e., γ- vs. α-regiochemistry) is masked, and distinguishing acylation patterns indicative of *p*-hydroxybenzoates from those of acetates is less direct; this issue is somewhat ameliorated by the considerably higher resolution of the spectra. The shifts in the *p*-hydroxybenzoate HSQC correlations (especially $pBz_{3/5}$) following acetylation (Fig. 2d) concur with earlier observations regarding

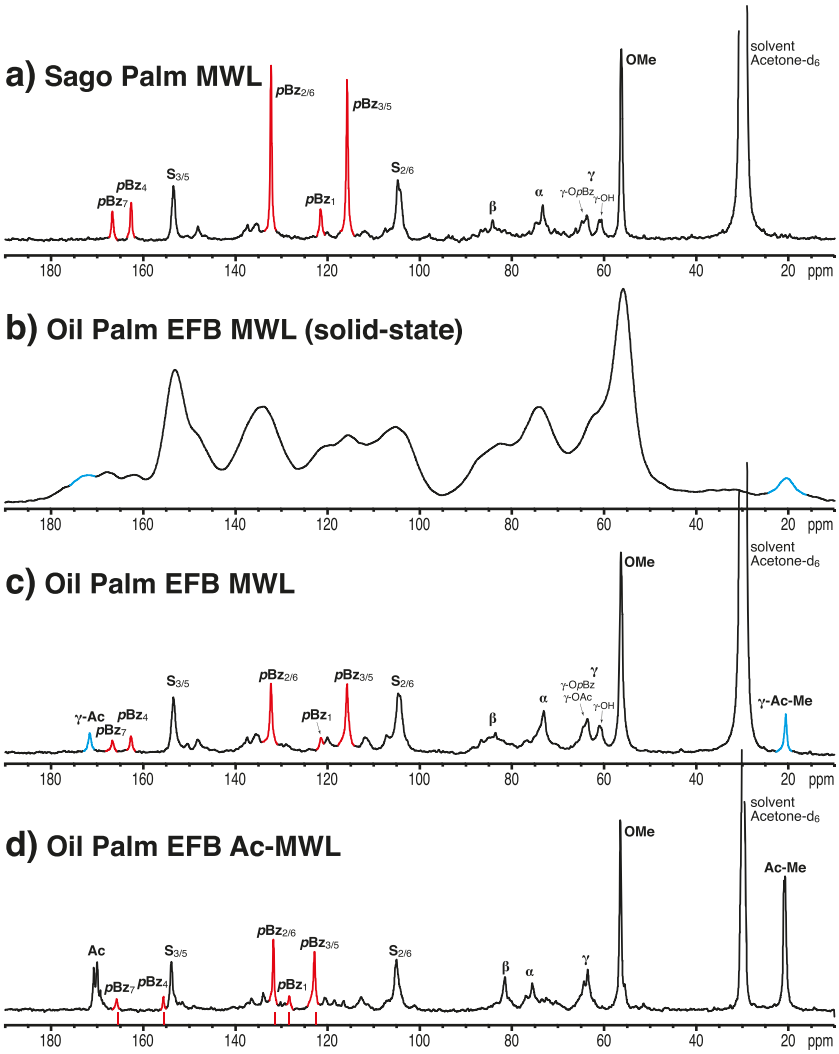


Fig. 1 1D ^{13}C NMR spectra of various palm lignins. **a** Solution-state spectrum from a sago palm MWL isolated as described previously [35]. **b** Solid-state (200 MHz) spectrum and **c** solution-state spectrum of MWL isolated from the oil palm empty fruit bunches (*EFB*) previously characterized [34, 43]. **d** Solution-state spectrum of acetylated EFB isolated lignin (Ac-MWL). The solution-state spectra were run at 400 MHz (100 MHz ^{13}C); the acetylated MWL is in acetone-d_6, underivatized MWLs in 9:1 acetone-d_6:D_2O. *Red-colored peaks* labeled *p*Bz are for carbons in the *p*-hydroxybenzoyl moiety. The *red lines* between plot **d** and its axis denote the chemical shifts from model compound $\mathbf{8}_{\mathbf{SS}\text{-}p\mathbf{Bz}}$. *S* syringyl unit. The side-chain regions are quite disperse but generally labeled α, β, and γ. γ-O*p*Bz signifies *p*-hydroxybenzoylated γ-alcohols, γ-OAc is (naturally) acetylated γ-alcohols. *Cyan-colored peaks* in **b**–**c** indicate the natural γ-acetylation of these lignins; no acetates appear in the sago palm lignin in **a** presumably due to the isolation method [35]. Spectra of lignins isolated from other parts of the oil palm plant (trunk and fronds, not shown) are almost indistinguishable from those shown in **c**–**d**

the 1D ^{13}C spectra, namely, that *p*-hydroxybenzoates on lignins exist in their free-phenolic form and are not etherified; were they etherified, the *p*Bz group would not become acetylated and their correlations would be essentially unchanged from those in the unacetylated lignins (Fig. 2b).

p-Hydroxybenzoates are therefore confirmed to be free-phenolic pendant entities acylating the γ-hydroxyls of lignin side-chains. Additional proof is established below.

Which Lignin Units are *p*-Hydroxybenzoylated?

Next, we needed to establish which units in lignins were acylated by *p*-hydroxybenzoates. Consistent with observations on *p*-coumarates and acetates, were *p*-hydroxybenzoates largely on S units, and were they on a range of lignin unit types? In this case, NMR of the acetylated lignin derivatives could answer both of these questions as, fortuitously, the α-protons in acetylated γ-*p*-hydroxybenzoylated β-ether units **B** were rather well resolved from the normal acetylated β-ether units **A** (Fig. 2c).

The 1D proton projection on the HSQC spectrum in Fig. 2c indicates how well the α-protons in the major β-ether units are resolved by proton NMR. The γ-*p*-hydroxybenzoylated β-ether units **B** have α-proton/α-carbon correlations at 6.20/76.8 ppm (*threo*) and 6.19/75.7 ppm (*erythro*), whereas those from the normal acetylated units **A** are at 6.06/76.8 ppm

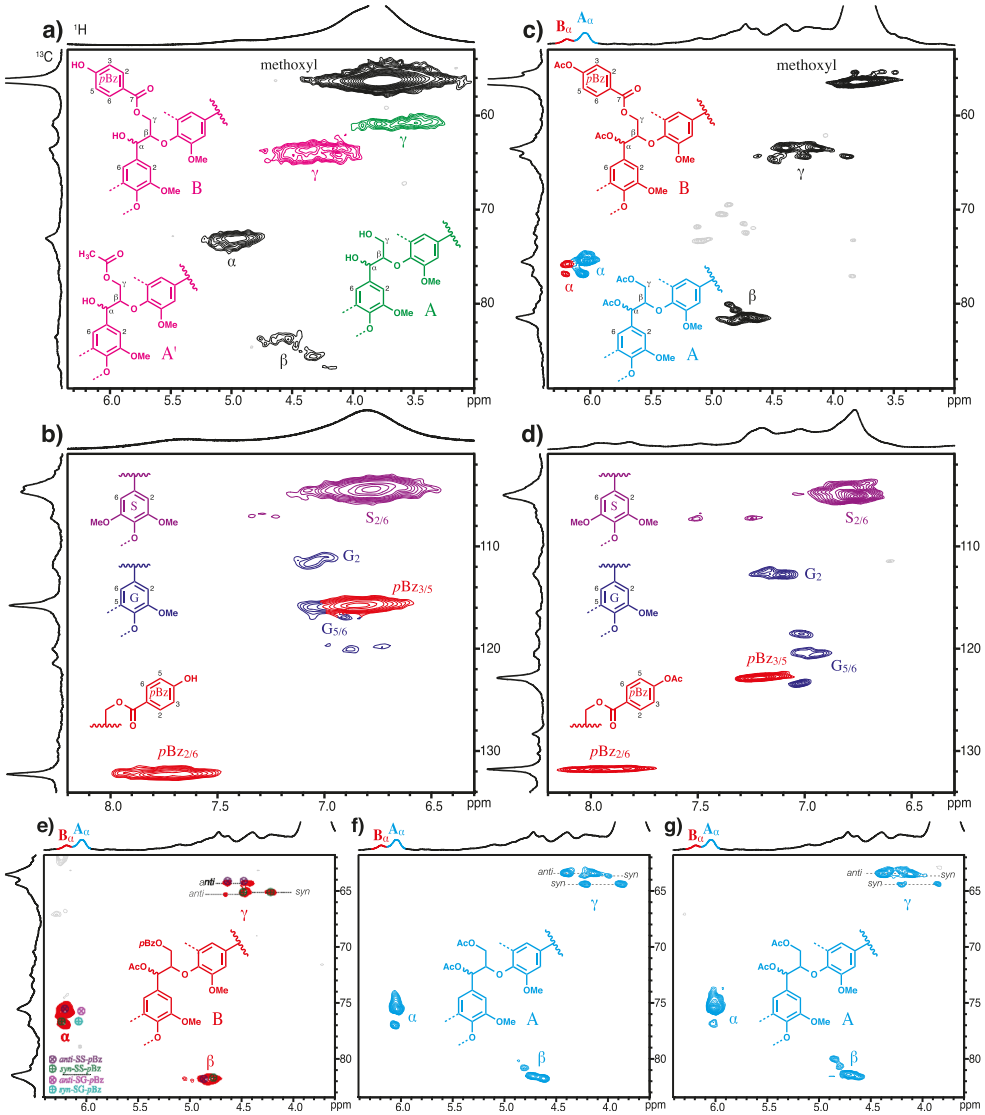


Fig. 2 2D HSQC spectra, and 2D HSQC (F2-F3) planes from 3D TOCSY-HSQC spectra, of oil palm EFB lignins. **a**–**b** The unacetylated MWL (360 MHz); **a** side-chain region, **b** aromatic region. The unacetylated material contained magnetic material, presumably iron compounds, that could not be effectively washed out; the spectra are therefore extremely broad, but **a** still clearly shows the degree to which γ-hydroxyls are acylated, either by acetate or *p*-hydroxybenzoate. **c**–**d** The EDTA-washed acetylated EFB MWL (500 MHz cryoprobe instrument); **c** side-chain region, **d** aromatic region. These much sharper spectra resolve the α-protons (and the corresponding correlations) of γ-*p*-hydroxybenzoylated **B** from γ-acetylated **A** units. **e**–**g** HSQC slices (F2–F3 dimension) from a 3D TOCSY-HSQC experiment on acetylated EFB lignin over a small proton chemical shift range in the F1 dimension: **e** 6.29–6.19 ppm (α-protons in γ-*p*-acetoxybenzoylated β-ether units **B**), **f** 6.13–6.07 ppm (α-protons in normal β-ether units **A**, both *syn*- and *anti*-isomers), and **g** 6.04–5.98 (α-protons in normal β-ether units **A**, mainly *anti*-isomers). The *colored symbols* on the *red contours* in **e** are shifts from model compounds **8**; clearly, only the SS-*p*Bz moieties are present at significant levels in the lignin (with chemical shifts for the α-H/C correlations matching well with those from compounds *syn*- and *anti*-SS-*p*Bz but not with those from the SG-*p*Bz compounds; the β and γ data are not shown for the latter as they are less diagnostic). Units **A** are normal (acetylated) β-ether units, **B** are γ-*p*-acetoxybenzoylated β-ether units. *G* guaiacyl, *S* syringyl, *pBz p*-hydroxybenzoyl

(*threo*) and 6.02/75.4 ppm (*erythro*). Initially, the α-^{1}H–^{13}C correlation contours were not sufficiently resolved, but running higher-resolution spectra restricted to acquiring just the side-chain region resulted in sufficient resolution to readily distinguish the various correlations (Fig. 2c).

It is apparent from the aromatic region of HSQC spectra that the EFB lignin was S-rich, Fig. 2b, d. An S:G ratio of 83:17 can be derived from the data in Fig. 2d by volume integration of the $S_{2/6}$ and G_2 aromatic proton/carbon correlations. In support, the side-chain region (Fig. 2c) revealed very low levels of phenylcoumaran (β–5) and dibenzodioxocin (5–5/β–O–4) units that must involve G units—their contours appear below the levels plotted.

Distinguishing whether *p*-hydroxybenzoates are attached to S or G units requires long-range ^{1}H–^{13}C correlation experiments [13]. The heteronuclear multiple bond correlation

(HMBC) experiment can correlate a proton with carbons up to three bonds away. The simplest probe is via the resolved α-protons in β-ether units (in the acetylated lignins), which should correlate with carbons γ (3-bond), β (2-bond), 1 (2-bond), and 2 and 6 (3-bond). The feature of this experiment that makes distinction of S- from G-units so elegant is that the identical 2/6 carbons in symmetrical S units resonate at about 105 ppm, whereas carbons 2 and 6 are distinctly different in the unsymmetrical G-units, resonating at 114 and 120 ppm. As shown in Fig. 3a, it is readily evident that the β-ethers in normal units **A** are both S (dark blue) and G (cyan), whereas it is only possible to detect γ-*p*-hydroxybenzoylated S units **B** (red). The prevalence of *p*-hydroxybenzoates on S over G units may be either a reflection of the control in the transferase-catalyzed acylation reaction or simply the timing of the conjugation and the levels of the two monolignols present at that time, i.e., that *p*-hydroxybenzoylation of sinapyl alcohol is either significantly favored over coniferyl alcohol, or that sinapyl alcohol is more prevalent at the time of acylation.

The 3-bond correlations in the HMBC data between the γ-protons and the carbonyl (Fig. 3b) also provide compelling proof that the *p*-hydroxybenzoates acylate lignin γ-positions. The HMBC experiment can also be used to identify the variety of units that are acylated. This is achieved here by examining the γ-protons correlating with the carbonyl carbon of the *p*-hydroxybenzoate moiety (Fig. 3b). However, determining the range of products is not as easy as it was for *p*-coumarates, as there are insufficient model compounds, and the dispersion between the unit types is limited. For example, it is currently difficult to prove that *p*-hydroxybenzoates acylate the phenylcoumaran (β–5) units or hydroxycinnamyl alcohol endgroups; each of those unit types is present only at low levels in this S-rich lignin.

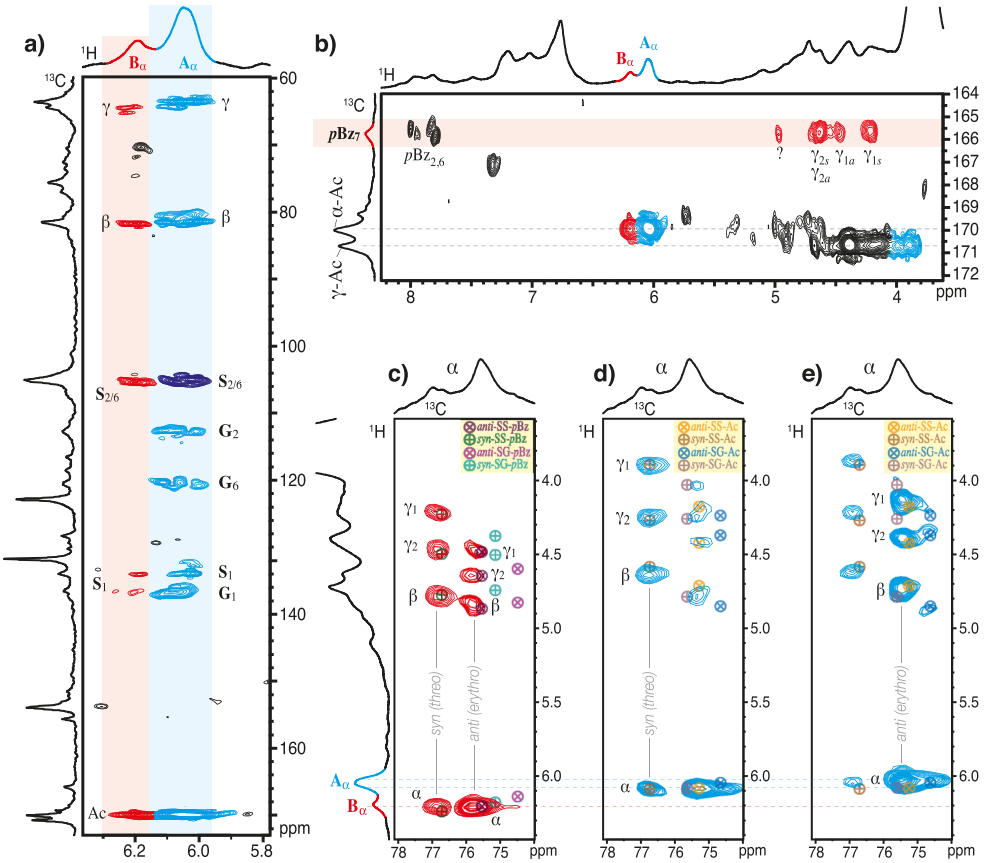


Fig. 3 2D HMBC spectra, and 2D HSQC-TOCSY (F1–F2) planes from 3D TOCSY-HSQC spectra, of oil palm EFB acetylated lignin (Ac-MWL). **a** HMBC section showing long-range ^{1}H–^{13}C correlations to α-protons in β-ether units, elegantly revealing their S and/or G nature. Normal β-ether units **A** are clearly both S and G in nature; γ-*p*-hydroxybenzoylated β-ether units **B** are essentially S-only. **b** HMBC section showing long-range correlations to the *p*-acetoxybenzoyl and acetate carbonyl carbons proving acylation of lignin γ-hydroxyls. **c**–**e** HSQC-TOCSY slices (F1–F2 dimension) from the 3D TOCSY-HSQC experiment on acetylated EFB lignin over a small proton chemical shift range in the F3 dimension: **c** 6.29–6.19 ppm (α-protons in γ-*p*-acetoxybenzoylated β-ether units **B** showing both *syn*- and *anti*-isomers of these units, **d** 6.13–6.07 ppm (α-protons in normal β-ether units **A**, mainly *syn*-isomers), **e** 6.04–5.98 (α-protons in normal β-ether units **A**, mainly *anti*-isomers). The *colored symbols* on the *red* and *blue contours* in **c**–**e** are shifts from model compounds **8**; again, it is clear that data from SS-*p*Bz moieties match the model data well, whereas those of the SG analogs are not evidenced; for the acetates, there is evidence for the SG-Ac components as well as, obviously in this high-S lignin, the dominant SS-Ac analogs

The dispersion that is afforded to the α-protons suggests that it should be possible for a 3D TOCSY-HSQC experiment to delineate whether *p*-hydroxybenzoates acylate both *syn*- (*threo*-) and *anti*- (*erythro*-) isomers of the β-ether units. This is because one can take 2D HSQC planes at given proton frequencies and exploit the α-proton differences between the γ-*p*-hydroxybenzoylated **B** and normal **A** units noted above. Similarly, one can take 2D HSQC-TOCSY planes through the α-proton region. Such 2D HSQC slices are shown in Fig. 2e–g, and HSQC-TOCSY slices are shown in Fig. 3c–e, as described in the captions. It is revealed that there are *syn*- and *anti*-isomers of the normal β-ether units (Figs. 2f–g and 3d–e), with *anti*-isomers predominating because of the S-rich nature of the lignin, as already well established [68–72]. It is equally clear that *p*-hydroxybenzoates acylate both of the β-ether isomers (Figs. 2c and 3c). A comparison of the α-H/C correlation peak data for the lignin versus the overlaid data from the various model compounds (Fig. 2e) indicates that, in addition to the *p*Bz's being almost entirely on S units as noted above, the data for the two SS-*p*Bz isomers match very well, but those for the β–O–4-G analogs SG-*p*Bz do not. Similarly, the data match well for the same SS compounds and not (especially for the β-H/C correlations) for the SG analogs (Fig. 3c). As in the case made early on for *p*-coumarates acylating grass lignins [25], the presence of *p*-hydroxybenzoates indiscriminately on both isomers of β-ether units suggests that they arise by lignification with preformed monolignol *p*-hydroxybenzoate conjugates.

Determination of Regiochemistry and Quantification by DFRC

Proving γ-regiochemistry (although not its exclusivity) and the S versus G nature of the units involved with *p*-hydroxybenzoates is elegantly made by the DFRC (derivatization followed by reductive cleavage) assay [61, 62], Fig. 4a. This is possible as DFRC advantageously cleaves lignin β-ethers but leaves esters fully intact [62, 51, 18, 23, 27]. Unfortunately, the acetylated coniferyl and sinapyl *p*-hydroxybenzoate products anticipated from this method do not survive GC analysis conditions, so the products were analyzed by HPLC-ESI-MS, Fig. 4b. Quantification of the DFRC-released diacetates of coniferyl alcohol (CA), sinapyl alcohol (SA), coniferyl *p*-hydroxybenzoate (CA-*p*Bz), and sinapyl *p*-hydroxybenzoate (SA-*p*Bz) accounted for 0.9, 2.0, 0.1, and 1.3 wt% of the EFB-1 sample (Table 1). The monolignol component, uncorrected for recovery and losses, accounted for some 16.7 % of the lignin. The DFRC-released total S to G monomer ratio was 73:27, whereas the ratio of SA-*p*Bz to CA-*p*Bz was 96:4, confirming the NMR observations that the *p*-hydroxybenzoates are primarily located on S units. Similar data were obtained for the EFB-2 sample (Table 1).

Proof that Lignin *p*-Hydroxybenzoates Arise via Lignification with Preformed Monolignol Conjugates

All of the aforementioned data strongly suggest, but strictly do not prove, that the *p*-hydroxybenzoate groups inherent on these palm lignins arise from lignification incorporating preformed monolignol *p*-hydroxybenzoate conjugates. As we have described previously for acetates (e.g., on kenaf lignins), there is a way to provide compelling proof [62, 14]. Given that post-coupling rearomatization of β–β-coupled quinone methide lignification intermediates is the one reaction that is perturbed by the presence of an acylated γ-OH, observing novel acylated β–β-coupling products in the lignin provides convincing evidence that acylated monolignols are involved in lignification [14].

As seen in Fig. 5, evidence for all three types of products **9a**–**11a** indicating the involvement of sinapyl *p*-hydroxybenzoate monomers in lignification is provided by observing sets of signature H/C correlations in HSQC spectra, correlations that well match the data from synthesized model compounds for all of the structures **9a**–**11a**. These units include both those arising from the cross-coupling of sinapyl *p*-hydroxybenzoate with sinapyl alcohol (**9a**, **10a**) as well as the homo-dimerization of sinapyl *p*-hydroxybenzoate (**11a**). These spectra also provide evidence (with correlations for structures **9b**–**11b**) that sinapyl acetate is a monomer conjugate in this palm EFB lignin, as was previously observed in kenaf [14].

Evidence from metabolite profiling of poplar wood for novel β–β-coupled heterodimers between sinapyl alcohol and sinapyl *p*-hydroxybenzoate [29, 32] suggests that it is safe to contend that all lignin *p*-hydroxybenzoates, across the plant taxa that have them, derive from lignification using monolignol *p*-hydroxybenzoate conjugates. The evidence is therefore sufficiently compelling that currently unknown transferases are responsible for producing monolignol *p*-hydroxybenzoate conjugates that are used as 'monomers' for lignification in palms (as well as in willows, poplars, and aspen). The levels of *p*-hydroxybenzoates on lignins should be manipulable by normal methods of selection and breeding or by misregulating such genes, once found. If *p*-hydroxy benzoic acid or its derived products are sufficiently valuable as commodities and coproducts generated from a biomass conversion process, a long-term strategy might be to upregulate the *p*-hydroxybenzoate levels on lignin in target plants. As determined here, only some 23 % (Table 1) of lignin monomers are *p*-hydroxybenzoylated in palm EFBs, and the level is likely to be lower in poplar and willow, so there appears to be significant potential for its increase. Other factors being equal, for example, if all of the monolignols destined for lignification were *p*-hydroxybenzoylated, the amount of *p*-hydroxybenzoic acid available would be approximately 40 wt% of the lignin or roughly 8 wt% of the EFB fiber.

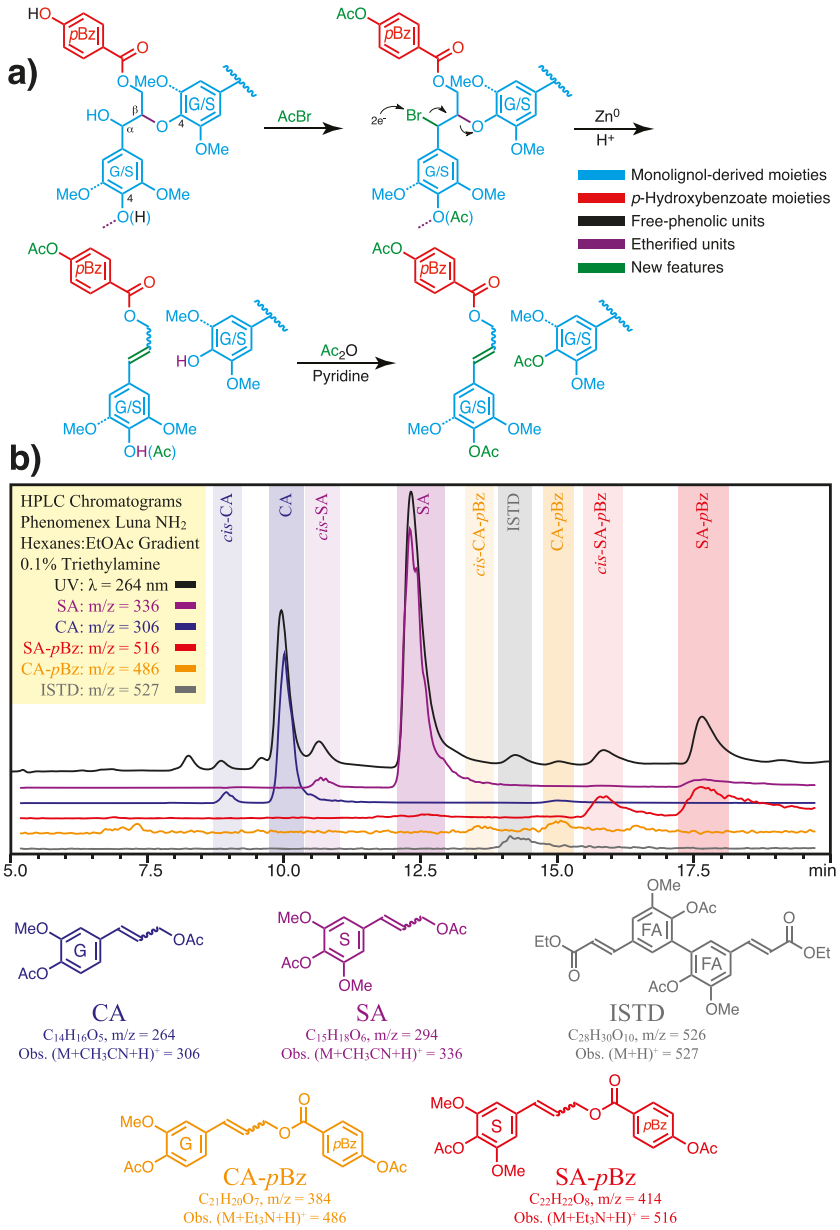


Fig. 4 DFRC proof of γ-*p*-hydroxybenzoylation. **a** A schematic representation of the steps in the DFRC analysis highlighting the bonds that are cleaved and the new functional groups that are formed. **b** The HPLC chromatogram (λ= 264 nm) of the DFRC product of oil palm EFB overlaid (*colored traces*) with extracted-ESI-MS-ion chromatograms for the base-peak masses corresponding to the various compounds of interest. The color-band-highlighted peaks correspond to DFRC products coniferyl alcohol (CA [M+CH_3CN+H]$^+$, *m/z*=306, *blue*), sinapyl alcohol (SA [M+CH_3CN+H]$^+$, *m/z*=336, *purple*), coniferyl *p*-hydroxybenzoate (CA-*p*Bz, [M+Et_3N]$^+$, *m/z*=486, *orange*), sinapyl *p*-hydroxybenzoate (SA-*p*Bz, [M+Et_3N]$^+$, *m/z*=516, *red*), and internal recovery standard diethyl 5,5′-diferulate (DEDF [M+H]$^+$, *m/z*=527, *gray*) all as their acetylated derivatives (as shown); the same compound names are used for the unacetylated compounds in Table 2 and elsewhere. The retention time, UV–vis spectrum, and mass spectrum of each compound were authenticated using external synthetic standards

Conclusions

Our data provide compelling evidence that, as for acetates on kenaf (and many other plants) lignins and *p*-coumarates on grass lignins, *p*-hydroxybenzoates acylate the γ-hydroxyl of lignin side-chains and are found mainly on S units. Furthermore, like their *p*-coumarate analogs, the *p*-hydroxybenzoate moieties are not subjected to significant levels of radical coupling, or other etherification, during lignification and therefore remain overwhelmingly as free-phenolic pendant units on lignin unit side-chains and can therefore be easily cleaved off to deliver a single compound in quite pure form. As has been demonstrated for *p*-coumarates, and as there is no a priori reason that *p*-hydroxybenzoates can not undergo radical coupling, the logical reason is that such units are preferentially subject to radical transfer reactions to the G and S monomer and polymer phenolics present in the radical-limited lignification process [12, 73]. The data presented here also provide evidence that, analogously to the *p*-coumarates and acetates found acylating lignin units, *p*-hydroxybenzoate units arise from lignification using the preformed monolignol *p*-hydroxybenzoate conjugates and, at least in these palm lignins, are almost entirely derived from sinapyl *p*-hydroxybenzoate. It is now logical to contend that any of the acylation products seen in any of the native lignins is via acylation at the monomer level and that post-polymerization (or post-coupling) acylation reactions are unlikely. We therefore

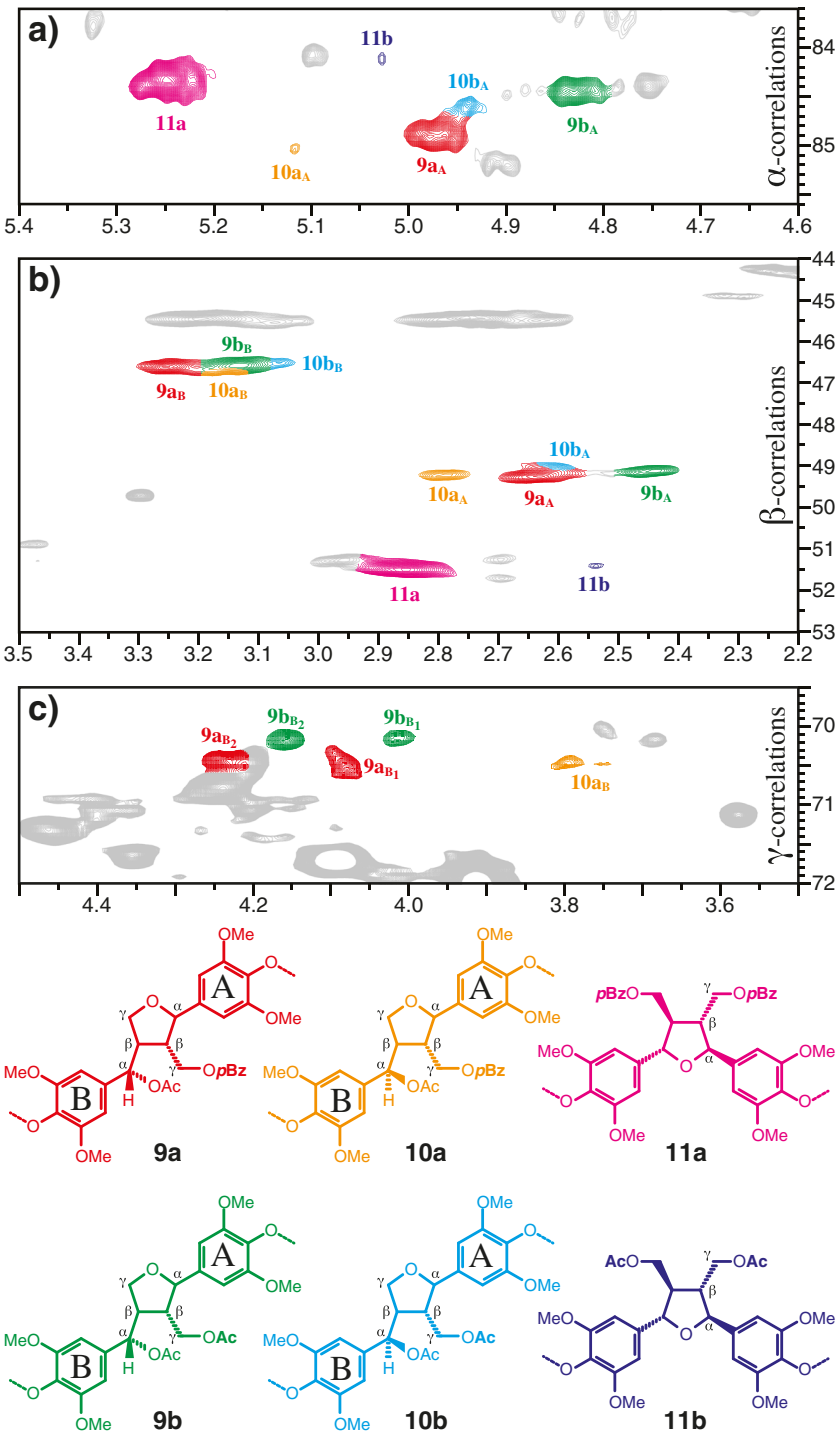


Fig. 5 Proof that preformed monolignol *p*-hydroxybenzoates (and acetates) are responsible for the lignin *p*-hydroxybenzoylation (and acetylation). Diagnostic HSQC correlation contour peaks matching those from model compounds are evidenced in the **a** α-, **b** β-, and **c** γ-regions of spectra for all three types of novel β–β-coupling products (**9a**, **10a**, **11a**) that would arise only from radical coupling reactions involving sinapyl *p*-hydroxybenzoate [and for analogous products (**9b**, **10b**, **11b**) from sinapyl acetate]. See text for explanation. Model compounds were of the 4–O-methylated structures shown; assignments of lignin correlation peaks (colored the same as the structures below) were made from the model data. Products in the lignin may be 4–O-etherified during lignification or may be 4–O-Ac here if the moiety in the lignin remained free-phenolic. Labeling is necessarily complex; $\mathbf{9a_B}$ for example is the α-, β-, or γ-H/C correlation from the diagnostic B unit of compound **9a**; similarly, $\mathbf{9a_{B2}}$ in Fig. 5c (γ-correlations) is simply the second (lowest field, highest ppm) γ-H/C correlation from the B unit of compound **9a**

expect that *p*-hydroxybenzoyl-CoA:monolignol transferases are involved in lignification in the various willows (*Salix* spp.), poplars and aspen (*Populus* spp.), and palms (family Arecaceae) that have *p*-hydroxybenzoylated lignins, and that identification of the genes involved will allow *p*-hydroxybenzoate levels on lignin to be manipulated as recently demonstrated for *p*-coumarates [23]. To date, however, the in planta pathway to *p*-hydroxybenzoate (or its CoA derivative) remains ambiguous. If *p*-hydroxybenzoic acid becomes a viable coproduct from cellulosic bioenergy lignin streams, it could conceivably be upregulated or even introduced into other biomass plants. Right now, the current palm oil empty fruit bunch 'wastes' should be able to generate a sizeable stream of relatively clean *p*-

hydroxybenzoic acid; when available at the multiton level, industrial applications are expected to increase beyond those already in use today.

Acknowledgments We gratefully acknowledge partial funding through the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02-07ER64494), DOE Energy Biosciences program (#DE-AI02-00ER15067, #DE-FG02-03ER15442), USDA-CSREES National Research Initiatives (Improved Utilization of Wood and Wood Fiber #2001-02176), and, for the reported ESI-MS data, the purchase of the Waters LCT® in 2000 that was partially funded by NSF Award #CHE9974839 to the University of Wisconsin Department of Chemistry. WB acknowledges the Multidisciplinary Research Partnership (01MRB510W) 'Biotechnology for a Sustainable Economy.'

References

1. Somerville C, Youngs H, Taylor C, Davis SC, Long SP (2010) Feedstocks for lignocellulosic biofuels. Science 329(5993):790–792. doi:10.1126/Science.1189268
2. Ragauskas AJ, Williams CK, Davison BH, Britovsek G, Cairney J, Eckert CA, Frederick WJ, Hallett JP, Leak DJ, Liotta CL, Mielenz JR, Murphy R, Templer R, Tschaplinski T (2006) The path forward for biofuels and biomaterials. Science 311(5760):484–489. doi:10.1126/Science.1114736
3. Sumathi S, Chai SP, Mohamed AR (2008) Utilization of oil palm as a source of renewable energy in Malaysia. Renew Sust Energ Rev 12(9):2404–2421
4. Distribution of oil palm planted area by state and sector in 2013 (Jan–June). (2014). http://www.kppk.gov.my/statistik_komoditi/palmoil.html.
5. Rasidi R, Zakaria S, Chia CH, Boehm R, Laborie M-P (2014) Physico-mechanical properties of resol phenolic adhesives derived from liquefaction of oil palm empty fruit bunch fibres. Ind Crop Prod 62:119–124
6. Zakaria S, Liew TK, Chia CH, Pua FL, Pin FS, Roslan R, Amran UA, Potthast A, Rosenau T (2013) Characterization of Fe_2O_3/FeOOH catalyzed solvolytic liquefaction of oil palm empty fruit bunch (EFB) products. Bioremediation Biodegradation:1–7. doi:10.4172/2155-6199.S4-001
7. Nor Roslam Wan Isahak W, Hisham MWM, Yarmo MA, Yun Hin T-y (2012) A review on bio-oil production from biomass by using pyrolysis method. Renew Sust Energ Rev 16(8):5910–5923
8. Ralph J, Lundquist K, Brunow G, Lu F, Kim H, Schatz PF, Marita JM, Hatfield RD, Ralph SA, Christensen JH, Boerjan W (2004) Lignins: natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. Phytochem Revs 3(1):29–60
9. Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Annu Rev Plant Biol 54:519–549
10. Vanholme R, Morreel K, Darrah C, Oyarce P, Grabber JH, Ralph J, Boerjan W (2012) Metabolic engineering of novel lignin in biomass crops. New Phytol 196(4):978–1000. doi:10.1111/j.1469-8137.2012.04337.x
11. Ralph J (2007) Perturbing Lignification. In: Entwistle K, Harris PJ, Walker J (eds) The compromised wood workshop 2007. Wood Technology Research Centre, University of Canterbury, New Zealand, pp 85–112
12. Ralph J (2010) Hydroxycinnamates in lignification. Phytochem Revs 9(1):65–83. doi:10.1007/s11101-009-9141-9
13. Ralph J, Landucci LL (2010) NMR of Lignins. In: Heitner C, Dimmel DR, Schmidt JA (eds) Lignin and lignans: advances in chemistry. CRC Press (Taylor & Francis Group), Boca Raton, pp 137–234. doi:10.1201/EBK1574444865
14. Lu F, Ralph J (2008) Novel tetrahydrofuran structures derived from β–β-coupling reactions involving sinapyl acetate in Kenaf lignins. Org Biomol Chem 6(20):3681–3694. doi:10.1039/B809464K
15. Sarkanen KV, Chang H-M, Allan GG (1967) Species variation in lignins. III. Hardwood lignins. Tappi 50(12):587–590
16. Lu F, Ralph J (2002) Preliminary evidence for sinapyl acetate as a lignin monomer in kenaf. Chem Commun 1:90–91
17. Ralph J (1996) An unusual lignin from kenaf. J Nat Prod 59(4):341–342
18. Ralph J, Lu F (1998) The DFRC method for lignin analysis. Part 6. A modified method to determine acetate regiochemistry on native and isolated lignins. J Agric Food Chem 46(11):4616–4619
19. del Rio JC, Rencoret J, Marques G, Gutierrez A, Ibarra D, Santos JI, Jimenez-Barbero J, Zhang LM, Martinez AT (2008) Highly acylated (acetylated and/or *p*-coumaroylated) native lignins from diverse herbaceous plants. J Agric Food Chem 56(20):9525–9534. doi:10.1021/Jf800806h
20. del Rio JC, Marques G, Rencoret J, Martinez AT, Gutierrez A (2007) Occurrence of naturally acetylated lignin units. J Agric Food Chem 55(14):5461–5468
21. Martinez AT, Rencoret J, Marques G, Gutierrez A, Ibarra D, Jimenez-Barbero J, del Rio JC (2008) Monolignol acylation and lignin structure in some nonwoody plants: a 2D NMR study. Phytochemistry 69(16):2831–2843. doi:10.1016/J.Phytochem.2008.09.005
22. Withers S, Lu F, Kim H, Zhu Y, Ralph J, Wilkerson CG (2012) Identification of a grass-specific enzyme that acylates monolignols with *p*-coumarate. J Biol Chem 287(11):8347–8355. doi:10.1074/jbc.M111.284497
23. Petrik D, Karlen SD, Cass C, Padmakshan D, Lu F, Liu S, Le Bris P, Antelme S, Santoro N, Wilkerson CG, Sibout R, Lapierre C, Ralph J, Sedbrook JC (2014) *p*-Coumaroyl-CoA:monolignol transferase (PMT) acts specifically in the lignin biosynthetic pathway in *Brachypodium distachyon*. Plant J 77(5):713–726. doi:10.1111/tpj.12420
24. Monties B, Lapierre C (1981) Donnés récentes sur l'hétérogénéite de la lignine. Physiologie Végétale 19(3):327–348
25. Ralph J, Hatfield RD, Quideau S, Helm RF, Grabber JH, Jung H-JG (1994) Pathway of *p*-coumaric acid incorporation into maize lignin as revealed by NMR. J Am Chem Soc 116(21):9448–9456
26. Crestini C, Argyropoulos DS (1997) Structural analysis of wheat straw lignin by quantitative ^{31}P and 2D NMR spectroscopy. The occurrence of ester bonds and α-O-4 substructures. J Agric Food Chem 45(4):1212–1219
27. Wilkerson CG, Mansfield SD, Lu F, Withers S, Park J, Karlen SD, Gonzales-Vigil E, Padmakshan D, Unda F, Rencoret J, Ralph J (2014) Monolignol ferulate transferase introduces chemically labile linkages into the lignin backbone. Science 344(6179):90–93. doi:10.1126/science.1250161
28. Landucci LL, Deka GC, Roy DN (1992) A ^{13}C NMR study of milled wood lignins from hybrid *Salix* clones. Holzforschung 46(6):505–511
29. Morreel K, Ralph J, Kim H, Lu F, Goeminne G, Ralph SA, Messens E, Boerjan W (2004) Profiling of oligolignols reveals monolignol coupling conditions in lignifying poplar xylem. Plant Physiol 136(3):3537–3549
30. Venverloo CJ (1971) Lignin of *Populus nigra* and some other *Salicaceae*. Holzforschung 25(1):18–24
31. Tomimura Y (1992) Chemical characteristics and utilization of oil palm trunks. Japan Agric Res Q 25(4):283–288

32. Lu F, Ralph J, Morreel K, Messens E, Boerjan W (2004) Preparation and relevance of a cross-coupling product between sinapyl alcohol and sinapyl *p*-hydroxybenzoate. Org Biomol Chem 2:2888–2890
33. Hibino T, Shibata D, Ito T, Tsuchiya D, Higuchi T, Pollet B, Lapierre C (1994) Chemical properties of lignin from *Aralia cordata*. Phytochemistry 37(2):445–448
34. Sun RC, Fang JM, Tomkinson J (1999) Fractional isolation and structural characterization of lignins from oil palm trunk and empty fruit bunch fibres. J Wood Chem Technol 19(4):335–356
35. Kuroda K-i, Ozawa T, Ueno T (2001) Characterization of sago palm (*Metroxylon sagu* Rottb.) lignin by analytical pyrolysis. J Agric Food Chem 49:1840–1847
36. Rencoret J, Ralph J, Marques G, Gutiérrez A, Martínez ÁT, del Rio JC (2013) Structural characterization of the lignin from coconut (*Cocos nucifera*) coir fibers. J Agric Food Chem 61(10):2434–2445. doi:10.1021/jf304686x
37. Tomimura Y (1992) Chemical characteristics of oil palm trunk. Bull For For Prod Res Inst 0(362):133–142
38. Jouanin L, Goujon T, de Nadaï V, Martin M-T, Mila I, Vallet C, Pollet B, Yoshinaga A, Chabbert B, Petit-Conil M, Lapierre C (2000) Lignification in transgenic poplars with extremely reduced caffeic acid *O*-methyltransferase activity. Plant Physiol 123(4):1363–1373
39. Stewart JJ, Akiyama T, Chapple CCS, Ralph J, Mansfield SD (2009) The effects on lignin structure of overexpression of ferulate 5-hydroxylase in hybrid poplar. Plant Physiol 150(2):621–635. doi:10.1104/Pp.109.137059
40. Hedenström M, Wiklund-Lindstrom S, Öman T, Lu F, Gerbner L, Schatz PF, Sundberg B, Ralph J (2009) Identification of lignin and polysaccharide modifications in *Populus* pood by chemometric analysis of 2D NMR spectra from dissolved cell walls. Mol Plant 2(5):933–942. doi:10.1093/Mp/Ssp047
41. Meyermans H, Morreel K, Lapierre C, Pollet B, De Bruyn A, Busson R, Herdewijn P, Devreese B, Van Beeumen J, Marita JM, Ralph J, Chen C, Burggraeve B, Van Montagu M, Messens E, Boerjan W (2000) Modifications in lignin and accumulation of phenolic glucosides in poplar xylem upon down-regulation of caffeoyl-coenzyme A *O*-methyltransferase, an enzyme involved in lignin biosynthesis. J Biol Chem 275(47):36899–36909
42. Jarvis MC (1994) Solid-state NMR study of leaf cell walls of oil palm. Phytochemistry 35(2):485–487
43. Sun RC, Mott L, Bolton J (1998) Isolation and fractional characterization of ball-milled and enzyme lignins from oil palm trunk. J Agric Food Chem 46(2):718–723
44. Sun RC, Mott L, Bolton J (1998) Fractional and structural characterization of ball milled and enzyme lignins from oil palm empty fruit bunch fiber. Wood Fiber Sci 30(3):301–311
45. Sun RC, Fang JM, Tomkinson J, Bolton J (1999) Physicochemical and structural characterization of alkali soluble lignins from oil palm trunk and empty fruit-bunch fibers. J Agric Food Chem 47(7):2930–2936
46. Sun RC, Fang JM, Mott L, Bolton J (1999) Fractional isolation and characterization of polysaccharides from oil palm trunk and empty fruit bunch fibres. Holzforschung 53(3):253–260
47. Nakamura Y, Higuchi T (1978) Ester linkage of *p*-coumaric acid in bamboo lignin. II. Syntheses of coniferyl p-hydroxybenzoate and coniferyl p-coumarate as possible precursors of aromatic acid esters in lignin. Cellul Chem Technol 12(2):199–208
48. Lu F, Ralph J (2005) Novel β–β-structures in natural lignins incorporating acylated monolignols. In: Thirteenth International Symposium on Wood, Fiber, and Pulping Chemistry, Auckland, New Zealand. APPITA, Australia, pp 233–237
49. Marita JM, Hatfield RD, Rancour DM, Frost KE (2014) Identification and suppression of the *p*-coumaroyl CoA: hydroxycinnamyl alcohol transferase in *Zea mays* L. Plant J 78(5):850–864
50. Hatfield RD, Marita JM, Frost K (2008) Characterization of *p*-coumarate accumulation, *p*-coumaroyl transferase, and cell wall changes during the development of corn stems. J Sci Food Agric 88:2529–2537
51. Lu F, Ralph J (1999) Detection and determination of *p*-coumaroylated units in lignins. J Agric Food Chem 47(5):1988–1992
52. Grabber JH, Quideau S, Ralph J (1996) *p*-Coumaroylated syringyl units in maize lignin: implications for β-ether cleavage by thioacidolysis. Phytochemistry 43(6):1189–1194
53. Zhu Y, Regner M, Lu F, Kim H, Mohammadi A, Pearson TJ, Ralph J (2013) Preparation of monolignol γ-acetate, γ-*p*-hydroxycinnamate, and γ-*p*-hydroxybenzoate conjugates: selective deacylation of phenolic acetates with hydrazine acetate. RSC Adv 3(44):21964–21971. doi:10.1039/C3RA42818D
54. Chang H-M, Cowling EB, Brown W, Adler E, Miksche G (1975) Comparative studies on cellulolytic enzyme lignin and milled wood lignin of sweetgum and spruce. Holzforschung 29(5):153–159
55. Marita JM, Ralph J, Lapierre C, Jouanin L, Boerjan W (2001) NMR characterization of lignins from transgenic poplars with suppressed caffeic acid *O*-methyltransferase activity. J Chem Soc Perkin Trans 1(22):2939–2945
56. Björkman A (1954) Isolation of lignin from finely divided wood with neutral solvents. Nature 174:1057–1058
57. Hatfield RD, Grabber JH, Ralph J, Brei K (1999) Using the acetyl bromide assay to determine lignin concentrations in herbaceous plants: some cautionary notes. J Agric Food Chem 47(2):628–632
58. Fukushima RS, Hatfield RD (2001) Extraction and isolation of lignin for utilization as a standard to determine lignin concentration using the acetyl bromide spectrophotometric method. J Agric Food Chem 49(7):3133–3139
59. Marita J, Ralph J, Hatfield RD, Chapple C (1999) NMR characterization of lignins in *Arabidopsis* altered in the activity of ferulate-5-hydroxylase. Proc Natl Acad Sci 96(22):12328–12332
60. Ralph J, Lu F (2004) Cryoprobe 3D NMR of acetylated ball-milled pine cell walls. Org Biomol Chem 2(19):2714–2715
61. Lu F, Ralph J (1997) Derivatization followed by reductive cleavage (DFRC method), a new method for lignin analysis: protocol for analysis of DFRC monomers. J Agric Food Chem 45(7):2590–2592
62. Lu F, Ralph J (2014) The DFRC (Derivatization Followed by Reductive Cleavage) method and its applications for lignin characterization. In: Lu F (ed) Lignin: structural analysis, applications in biomaterials, and ecological significance. Nova Science Publishers, Inc, Hauppauge, New York, pp 27–65
63. Helm RF, Ralph J (1992) Lignin–hydroxycinnamyl model compounds related to forage cell wall structure. 1. Ether-linked structures. J Agric Food Chem 40(11):2167–2175
64. Kratzl K, Kisser W, Gratzl J, Silbernagel H (1959) Der β-Giajacyläther des Guajacylglycerins, sine Umwandlung in Coniferylaldehyd und verschiedene andere Arylpropanderivate. Monatsch Chem 90(6):771–782
65. Wanasuria S, Steyobudi H, Mayun IB, Suprihatno B (1999) Iron deficiency of oil palm in Sumatra. Better Crops Int 13(1):33–35
66. Gallacher J, Snape CE, Hassan K, Jarvis MC (1994) Solid-state ^{13}C NMR study of palm trunk cell walls. J Sci Food Agric 64(4):487–491
67. Kim H, Ralph J, Akiyama T (2008) Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO-d_6. BioEnergy Res 1(1):56–66. doi:10.1007/s12155-008-9004-z
68. Akiyama T, Sugimoto T, Matsumoto Y, Meshitsuka G (2002) Erythro/threo ratio of β-O-4 structures as an important structural characteristic of lignin. I: improvement of ozonation method for the quantitative analysis of lignin side-chain structure. J Wood Sci 48(3):210–215
69. Akiyama T, Goto H, Nawawi DS, Syafii W, Matsumoto Y, Meshitsuka G (2005) *Erythro*/*threo* ratio of β-O-4-structures as an important structural characteristic of lignin. Part 4: variation in the

erythro/*threo* ratio in softwood and hardwood lignins and its relation to syringyl/guaiacyl ratio. Holzforschung 59(3):276–281

70. Bardet M, Robert D, Lundquist K, von Unge S (1998) Distribution of *erythro* and *threo* forms of different types of β-O-4 structures in aspen lignin by carbon-13 NMR using the 2D INADEQUATE experiment. Magn Reson Chem 36(8): 597–600
71. Lundquist K, Stomberg R, Von Unge S (1987) Stereochemical assignment of the threo and erythro forms of 2-(2,6-dimethoxyphenoxy)-1-(3,4-dimethoxyphenyl)-1,3-propanediol from X-ray analyses of the synthetic intermediates (Z)-2-(2,6-dimethoxyphenoxy)-3-(3,4-dimethoxyphenyl)-2-propenoic acid and threo-2-(2,6-dimethoxyphenoxy)-3-(3,4-dimethoxyphenyl)-3-hydroxypropanoic acid. Acta Chem Scand Ser B Org Chem Biochem 41(7):499–510
72. Akiyama T, Magara K, Meshitsuka G, Lundquist K, Matsumoto Y (2015) Absolute configuration of β- and α-asymmetric carbons within β-O-4-structures in hardwood lignin. J Wood Chem Technol 35(1):8–16. doi:10.1080/02773813.2014.892992
73. Hatfield RD, Ralph J, Grabber JH (2008) A potential role of sinapyl *p*-coumarate as a radical transfer mechanism in grass lignin formation. Planta 228:919–928. doi:10.1007/s00425-008-0791-4