A polymer of caffeyl alcohol in plant seeds

Fang Chen[a,b,1], Yuki Tobimatsu[c,1], Daphna Havkin-Frenkel[d], Richard A. Dixon[a,b,2], and John Ralph[c,e,2]

[a]Plant Biology Division, Samuel Roberts Noble Foundation, Ardmore, OK 73401; [b]Department of Energy, BioEnergy Science Center (BESC), Oak Ridge National Laboratory, Oak Ridge, TN 37831; [c]Department of Biochemistry, Enzyme Institute, University of Wisconsin, Madison, WI 53726; [d]Department of Plant Biology and Pathology, Rutgers, State University of New Jersey, New Brunswick, NJ 08901; and [e]Department of Energy, Great Lakes Bioenergy Research Center, and Wisconsin Bioenergy Initiative, Madison, WI 53706

Contributed by Richard A. Dixon, December 20, 2011 (sent for review November 2, 2011)

Lignins are complex phenylpropanoid polymers mostly associated with plant secondary cell walls. Lignins arise primarily via oxidative polymerization of the three monolignols, *p*-coumaryl, coniferyl, and sinapyl alcohols. Of the two hydroxycinnamyl alcohols that represent incompletely methylated biosynthetic products (and are not usually considered to be monolignols), 5-hydroxyconiferyl alcohol is now well established as incorporating into angiosperm lignins, but incorporation of caffeyl alcohol has not been shown. We report here the presence of a homopolymer of caffeyl alcohol in the seed coats of both monocot and dicot plants. This polymer (C-lignin) is deposited to high concentrations in the seed coat during the early stages of seed development in the vanilla orchid (*Vanilla planifolia*), and in several members of the Cactaceae. The lignin in other parts of the Vanilla plant is conventionally biosynthesized from coniferyl and sinapyl alcohols. Some species of cacti contain only C-lignin in their seeds, whereas others contain only classical guaiacyl/syringyl lignin (derived from coniferyl and sinapyl alcohols). NMR spectroscopic analysis revealed that the Vanilla seed-coat polymer was massively comprised of benzodioxane units and was structurally similar to the polymer synthesized in vitro by peroxidase-catalyzed polymerization of caffeyl alcohol. CD spectroscopy did not detect any optical activity in the seed polymer. These data support the contention that the C-lignin polymer is produced in vivo via combinatorial oxidative radical coupling that is under simple chemical control, a mechanism analogous to that theorized for classical lignin biosynthesis.

cell wall polymer | lignin polymerization | lignin structure | whole cell wall NMR

Lignins are abundant phenylpropanoid polymers produced primarily from oxidative polymerization of three 4-hydroxycinnamyl alcohols differing in their degrees of methoxylation (Fig. S1). Lignins occur mostly in vessels, tracheids, and fibrous tissues of vascular plants where they bind, strengthen, and waterproof cell walls to provide mechanical support, enhance water transport, and help ward off pathogens and pests. The biosynthesis and bioengineering of cell wall lignins, and their chemical and mechanical properties, have attracted significant attention because lignin hinders agro-industrial processes, such as chemical pulping of woody crops (1), forage digestion by livestock (2), and conversion of lignocellulosic plant biomass into liquid biofuels (3, 4). In addition, the variability of biosynthesis, and thereby the structures of various lignins, is considered to be closely correlated with the diversity and evolution of land plants (3, 5–12).

During lignin biosynthesis, the monolignol precursors are functionalized by aromatic hydroxylation and *O*-methylation (as well as successive side-chain reductions) to generate monolignols differing in their aromatic substitution patterns (Fig. 1*A* and Fig. S1). Natural lignins are generally composed of *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) units, that are biosynthesized by polymerization of the three primary monolignols, *p*-coumaryl, coniferyl, and sinapyl alcohols, respectively; natural angiosperm lignins have only low levels (<~2%) of H-units. Catechyl (C) and 5-hydroxyl guaiacyl (5-OH-G) units that may derive from polymerization of the corresponding caffeyl and 5-hydroxy coniferyl alcohols (Fig. 1*A* and Fig. S1) are not found in "normal" lignins.

Extensive studies have revealed the essential plasticity of lignin biosynthesis (6, 10, 12–15), and support the concept that lignin polymerization results from a combinatorial radical coupling process that is under simple chemical control (14, 16, 17). Thus, lignin monomer composition is largely determined by monolignol availability and, under certain circumstances, this permits incorporation of the "unusual" C and 5-OH-G monolignols into the polymer. For example, 5-hydroxyconiferyl alcohol participates in lignification in various angiosperm plants in which caffeic acid/5-hydroxyconiferaldehyde *O*-methyltransferase (COMT), the key enzyme for conversion of monolignol precursors from the 5-OH-G to the S aromatic level (18, 19), is down-regulated. The combination of a mutation in the gene encoding COMT with overexpression of ferulate 5-hydroxylase, which catalyzes hydroxylation of G to 5-OH-G aromatic level precursors, generates lignins largely composed of 5-OH-G units in benzodioxane structures (20, 21). Similarly, down-regulation of caffeoyl-CoA *O*-methyltransferase (CCoAOMT), for conversions from C to G aromatic-level precursors, introduces low levels (less than 10%) of C units into cell wall lignins in tracheary-element cultures of the gymnosperm *Pinus radiata* (22). However, down-regulation of CCoAOMT in angiosperm species, such as *Arabidopsis*, alfalfa, poplar, and tobacco, does not result in the incorporation of C units into lignin (23–27), and neither does down-regulation of both monolignol methylation enzymes (28).

Here we report a lignin in the monocotyledonous angiosperm Vanilla orchid (*Vanilla planifolia*) that is naturally biosynthesized from the unusual C monolignol, caffeyl alcohol. Similar polymers are found in the seeds of other vanilla species and several species of cacti (which are dicots). The *V. planifolia* polymer was structurally characterized by various chemical methods, 2D NMR spectroscopic techniques, and gel-permeation chromatography (GPC). All evidence indicates that the C-lignin is formed by combinatorial oxidative radical coupling under simple chemical control, a mechanism analogous to that occurring in classic lignification.

Results

Identification of C-Lignin Signatures in *V. planifolia*. Initial studies on the lignin of mature beans of the vanilla orchid *V. planifolia* (Fig. 1*C*) by thioacidolysis (29, 30) revealed the presence of a small doublet in the gas chromatography (GC)-MS profile at a retention time consistent with the catechyl (C) monomer, α,β,γ-trithioethyl-propylcatechol (Fig. 1*B*). The beans contained black-coated seeds (Fig. 1*F*), and thioacidolysis revealed that the lignin in the isolated seed coats was entirely composed of C units (Fig. 1*H*), with practically no release of α,β,γ-trithioethyl-propylguaiacol, from

Author contributions: F.C., Y.T., D.H.-F., and R.A.D. designed research; F.C. and Y.T. performed research; F.C., Y.T., R.A.D., and J.R. analyzed data; and Y.T., R.A.D., and J.R. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

[1]F.C and Y.T. contributed equally to this work.

[2]To whom correspondence may be addressed. E-mail: radixon@noble.org or jralph@wisc.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1120992109/-/DCSupplemental.

guaiacyl units, nor the syringyl analog (Fig. 1*B*). In contrast, thioacidolysis of the pod residue (after seed isolation), stem, leaf, and aerial root released normal G and S monomers with essentially no C monomers, indicating that the lignins present in these tissues are typical G-rich G/S-lignins (Fig. 1 *I–K*). C-lignin signatures first appeared in the seed coat at around 8 wk after pollination (Fig. S2), when the coat turns from transparent white to brown (Fig. 1 *D–F*), and at least 2–3 mo before the appearance in the pods of the flavor compound vanillin, which is biosynthetically related to lignin (31). The absolute levels of the C-lignin in the vanilla seed are substantially higher than the values estimated from the released thioacidolysis monomers because of the unusual structure of the polymer (see below).

C-Lignin Signatures in Seeds from Other Plant Species. Similar GC-MS traces to that shown for the *V. planifolia* seed coat were obtained from seed coats of two other Vanilla species, *Vanilla pompona* and *Vanilla tahitensis*. However, C units were not detected in the seed coats of *Phalaenopsis* orchid species, nor Asparagus and Agave, two other members of the monocot order Asparagales, to which Vanilla belongs. We did, however, observe strong C-lignin signatures in the seed coats from several species of the family Cactaceae, specifically in members of the genera *Astrophytum*, *Discocactus*, *Frailea*, *Melocatus* (Fig. 1*L*), *Notocactus*, *Uebelmannia*, and *Wigginsia*. Interestingly, all these species possessed black-coated seeds (Fig. 1*G*) that released only C monomers on thioacidolysis, whereas *Mammillaria* (Fig. 1*M*) and *Opuntia* species possessed brown seeds that contained normal G/S-lignin. During this limited survey, no seeds were found that contained both C- and G/S-lignins.

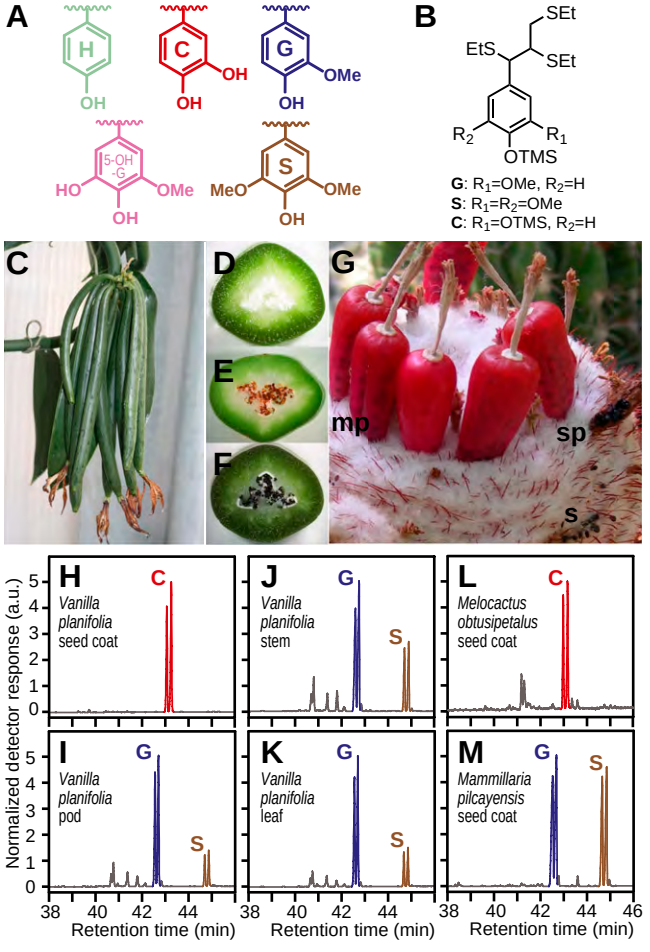

Fig. 1. A C-lignin polymer in plant seeds. (*A*) Aromatic structures of conventional and novel lignin units (see text and Fig. S1). (*B*) Diagnostic monomeric compounds released via thioacidolysis of guaiacyl (G), syringyl (S), and catechyl (C) lignin polymers. (*C*) Mature *V. planifolia* beans. (*D–F*) Cross-sections of *V. planifolia* beans at 6 (*D*), 8 (*E*), and 10 (*F*) wk postpollination. (*G*) Terminal cephalium of *Melocactus obtusipetalus*, showing mature seed pods (mp), a senesced seed pod (sp) revealing contents, and individual released seeds (s). (*H–M*) Partial total-ion chromatograms of thioacidolysis products from *V. planifolia* (*H*) seed coat, (*I*) pod (without seed), (*J*) stem, and (*K*) leaf, and (*L*) *Melocactus obtusipetalus*, and (*M*) *Mammillaria pilcayensis* seed coat.

Wet-Chemical and NMR Characterization of *V. planifolia* Tissues. Separated tissues from mature *V. planifolia* were further characterized by wet-chemical methods and by 2D NMR using direct dissolution/swelling (32, 33). Klason analysis of the seed coat indicated a very high level (~80%) of acid-insoluble lignin polymer (Table S1). Butanol-HCl assay (34) did not detect proanthocyanidins, which are major components in some seed coats (e.g., *Arabidopsis* and Medicago) (35, 36). The majority of the remaining material in the seed coat was crystalline cellulose (16%); very little noncellulosic sugars (2%) were detected. The chemical compositions of the pod remaining after seed isolation and the stem were similar overall (Table S1); these tissues were most rich in cellulose, with modest levels of hemicellulosic and pectic sugars, and Klason lignins.

Two-dimensional gel-state NMR spectra of vanilla tissues were acquired with samples prepared by swelling whole-plant materials, after fine milling, in DMSO-d_6/pyridine-d_5 (4:1, vol/vol) (33) (Fig. 2). The spectra of the pod and stem displayed typical G/S type lignins with an array of hemicellulosic and pectic sugar units (Fig. 2 *B* and *C*; see also Fig. S3 for expanded polysaccharide anomeric regions). However, the overwhelming signals in the spectrum of the seed coat were from the C-lignin polymer (Fig. 2*A*), as evidenced by the striking similarity to the spectrum of a synthetic dehydrogenation polymer (C-DHP) generated by horseradish peroxidase-catalyzed polymerization of caffeyl alcohol (Fig. 2*D*). The aliphatic regions of the spectrum indicated a massive presence of benzodioxane units **V**, for which the α-, β-, and γ-correlations from *trans*-benzodioxane rings V*t*, as well as lower-level contributions from *cis*-benzodioxane rings V*c*, were resolved and readily assigned by comparison with the data from the synthetic model dimers (Fig. S4 *A* and *B*). Conventional lignin aromatic signals were practically absent and, instead, the dominant signals were from C units: compelling confirmation of these assignments could be made via comparison with the C-DHP and model compounds (Fig. S4 *A* and *B*). The benzodioxane polymer in the seed coat is therefore derived from the polymerization, almost exclusively, of caffeyl alcohol (see Fig. 3*D* for the mechanism).

Ball-milled vanilla seed coat was also analyzed by normal solution-state NMR via complete dissolution/acetylation using the DMSO/*N*-methylimidazole (NMI) solvent system (Fig. S5) (32). A massive presence of C-lignins was again firmly established by diagnostic catechyl aromatic signals and benzodioxane units; these signals were predictably shifted following acetylation and could be assigned by comparison with the data from acetylated model dimers and C-DHP (Fig. S4 *C* and *D*). In addition, the signals from cellulosic glucans, which were not significant in the gel-state NMR spectra (Fig. 2) because of the incomplete gelation of crystalline cellulose (33), were then clearly observed, but the signals from hemicellulosic and pectic sugars remain practically absent.

NMR Characterization of Isolated C-Lignin from *V. planifolia* Seed Coat. Representative fractions enriched in the C-lignin polymer were isolated in 24% or 16% yield by cellulase treatment of ball-milled *V. planifolia* seed coat followed by extraction with DMSO or with 96% dioxane:water solution (37–39) (Table S1). Solution-state ^{13}C–^{1}H correlation heteronuclear single quantum coherence (HSQC) spectra of the isolated fractions in DMSO-d_6 indicated successful removal of the polysaccharides and concurrent

enrichment of C-lignin (Fig. S6). The isolated seed-coat lignin was then acetylated so as to be soluble in chloroform-*d*, which facilitates long-range ^{13}C–^{1}H correlation NMR experiments (e.g., heteronuclear multiple bond correlation, HMBC) (40, 41).

The large differences from classical lignins were most readily visualized from the aromatic regions of HSQC spectra of the isolated and acetylated seed-coat lignins (Fig. 3*A*, *Left*). Volume integration of the contour signals confirmed that this lignin is almost exclusively composed of C units; typical G and S lignin aromatics are virtually nonexistent in this polymer. Other correlations (gray) are currently unassigned, but do not seem to arise from caffeyl alcohol because they are not seen in the acetylated C-DHP (Fig. 3*B*).

High-field HSQC spectra of the side-chain (aliphatic) regions resolved most of the correlations for the various linkage types, revealing more clearly the manner in which the monomeric units are assembled (Fig. 3*A*, *Right*). The major contours in the spectrum were almost identical to those of the C-DHP (Fig. 3*B*). Benzodioxanes, resulting from β–O–4-coupling of a monomer with a caffeyl unit, were the dominant units in both the seed-coat lignin and C-DHP, accounting for over 98% of the total identifiable dimeric units. The *trans*/*cis* compositions of the benzodioxane rings (V*t*/V*c*) in the seed polymer and C-DHP were similar (V*t*:V*c* = 97:3 and 96:4, respectively). The normal acyclic β-aryl ether units **I**, which are the predominant linkages in typical natural lignins, were absent in these polymers. Small amounts of phenylcoumaran **II** and resinol **III** units were present in both the seed-coat lignin and C-DHP. HMBC experiments revealed the expected long-range correlations between the side-chain α-protons and the 1-, 2-, and 6-carbons of the catechyl aromatic rings in the polymer, supporting the contention that all of these units derive from caffeyl alcohol (Fig. S7). In addition, at least four unassigned correlations (colored in pink) were observed in the spectrum of the seed polymer (Fig. 3*A*). Because these were also identified in the spectrum of C-DHP (Fig. 3*B*), they presumably represent new structures resulting from radical coupling reactions of caffeyl alcohol. The isolated stem lignin is, in contrast, a G-rich G/S type lignin rich in β-aryl ether units **I**, with more modest amounts of phenylcoumaran **II** and resinol **III** units, and also with more minor amounts of dibenzodioxocin units **IV** (Fig. 3*C*, and Figs. S6 and S7), as is typical for angiosperm stem lignins (40, 41).

Molecular Weights of C-Lignin in *V. planifolia* Seed Coat. Molecular-weight distributions of acetylated samples of ball-milled whole *V. planifolia* seed coat and extracted C-lignins were determined by GPC with UV detection (Fig. 4*A* and Table S2). The molecular-weight profiles showed broad distributions spanning a range of up to 10^5 Da. The number-average degree of polymerization (DP_n, based on the molecular weight of the catechyl benzodioxane internal unit) of the whole seed coat was ∼30. The DP_n of the isolated lignins extracted with DMSO and 96% dioxane:water solution were ∼18 and 13, respectively. It is likely that the insoluble fractions that were left after extractions of isolated seed coat lignins contain C-lignins with higher molecular masses. As expected from the end-unit analysis (Fig. 3 *A* and *B*), the DP_n of the in vitro lignin (C-DHP) was even lower than those of isolated lignins, ∼8. All these values are comparable to literature values for various isolated and synthetic lignins (42–44).

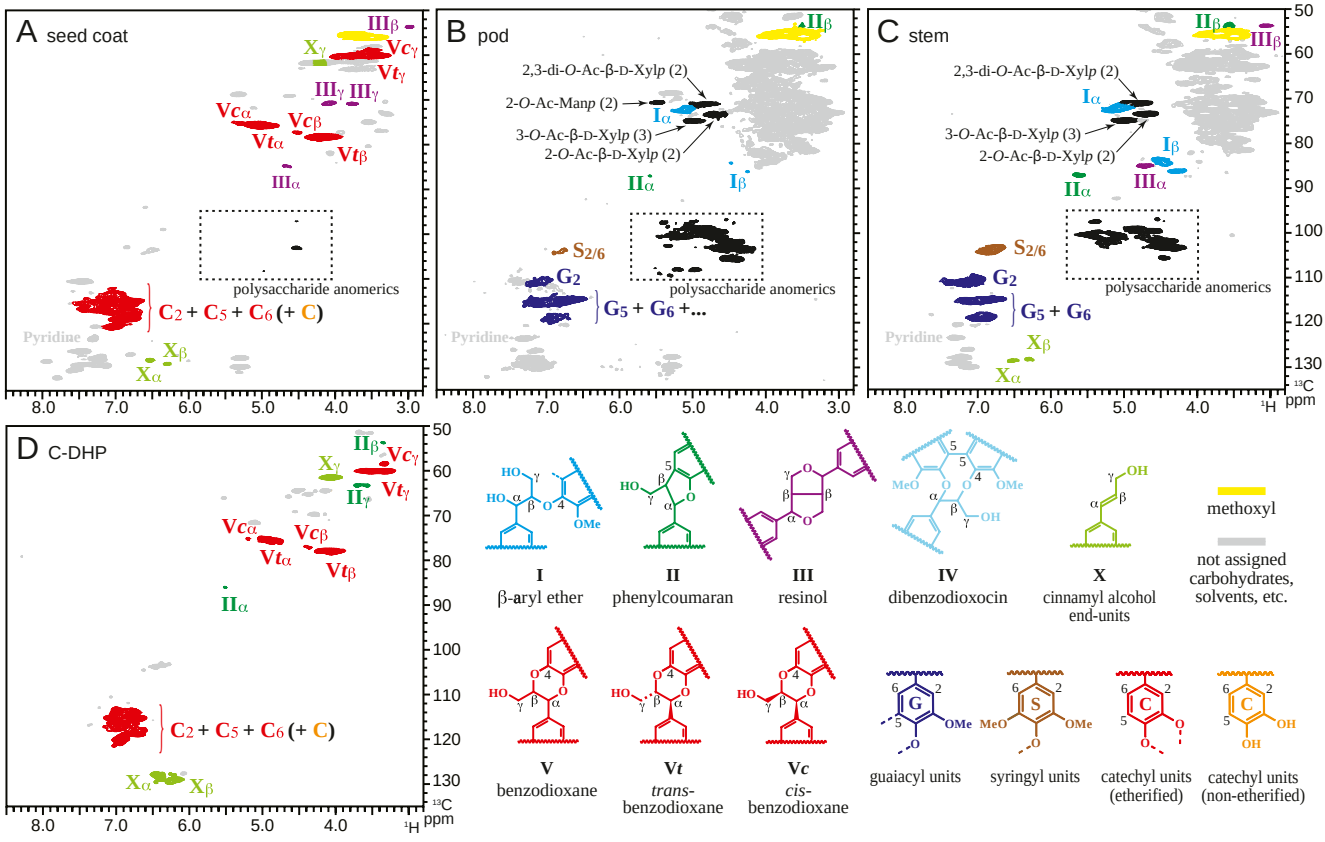

Fig. 2. A 2D NMR characterization of separated *V. planifolia* tissues. (*A*–*C*) Gel-state short-range ^{13}C–^{1}H correlation (HSQC) spectra of whole vanilla tissues, (*A*) seed coat, (*B*) pod, and (*C*) stem in 4:1 dimethyl sulfoxide-d_6/pyridine-d_5. (*D*) Solution-state HSQC spectra of an in vitro polymer (C-DHP), synthesized via peroxidase-catalyzed polymerization of caffeyl alcohol, in dimethyl sulfoxide-d_6.

Optical Activity of C-Lignin in *V. planifolia* Seed Coat. Natural lignins are optically inactive (45, 46). The optical activity of the isolated seed-coat lignin was investigated by CD spectroscopy. In addition, chiral benzodioxane dimers **1a** and **1b** were separated from a racemic mixture of synthetic dimer **1** (Fig. S4) by chiral HPLC and also subjected to CD for comparison (Fig. 4 *B* and *C*). The CD spectrum of the seed-coat lignin indicated no detectable optical activity, whereas the chiral dimer **1a** displayed clear positive Cotton effects in the region of 240–320 nm under the same analytical conditions (Fig. 4*D*). The other enantiomer **1b** had essentially a mirror image spectrum (Fig. 4*C*). Optical activity was readily seen by spiking (46) a preparation of the seed polymer with as little as 5% of chiral dimer **1a** (Fig. 4*D*). Benzodioxane units in the vanilla seed are therefore, within the limits of detection by the current method, optically inactive.

Discussion

Seeds of both monocot and dicot species contain previously unsuspected lignin polymers constructed almost entirely from catechyl (C) units. This C-polymer is a major component of the seed coat of *V. planifolia*, whereas the stem, leaf, and aerial root have only typical angiosperm G/S lignins. Thioacidolysis and 2D-NMR data clearly indicate that the C-polymer is essentially a homopolymer synthesized purely from caffeyl alcohol, and with benzodioxanes as essentially the only intermonomer unit in the polymer. Based on preliminary thioacidolysis data, similar C-lignins are found in the seed coats of certain cactus species. Our data so far suggest that lignins present in seed coat of cactus species are either of the C- or G/S-types, but not both.

It is premature to speculate about the possible distribution of the C-lignin polymer within the plant kingdom. Currently only observed in the Asparagales (Orchidaceae) and Caryophyllales (Cactaceae), it is likely that the polymer has wider distribution, as it is found in both monocots and dicots. The taxonomy of the Cactaceae is constantly under revision, but in a recent analysis *Astrophytum* (C-lignin in seed) and *Mammillaria* (G/S-lignin in seed) are closely related in the same clade (47). This finding suggests that the formation of C-lignin is not an ancient trait, but has occurred recently and probably frequently within the plant kingdom. Thus, the genetic/biochemical mechanisms that allow for the monolignol pathway to be derailed into production of high concentrations of caffeyl alcohol are probably relatively simple.

The C-polymer appears in *V. planifolia* beans at least 2–3 mo before the appearance of the flavor compound vanillin, which is synthesized in hair cells within the pod (48) and is likely derived from an initial phenylpropanoid precursor by side-chain shortening (31). The high concentration (>80% by the Klason lignin method) of C-polymer in the seed coat may imply a lignin-like structural role in addition to a tannin-like role for seed protection.

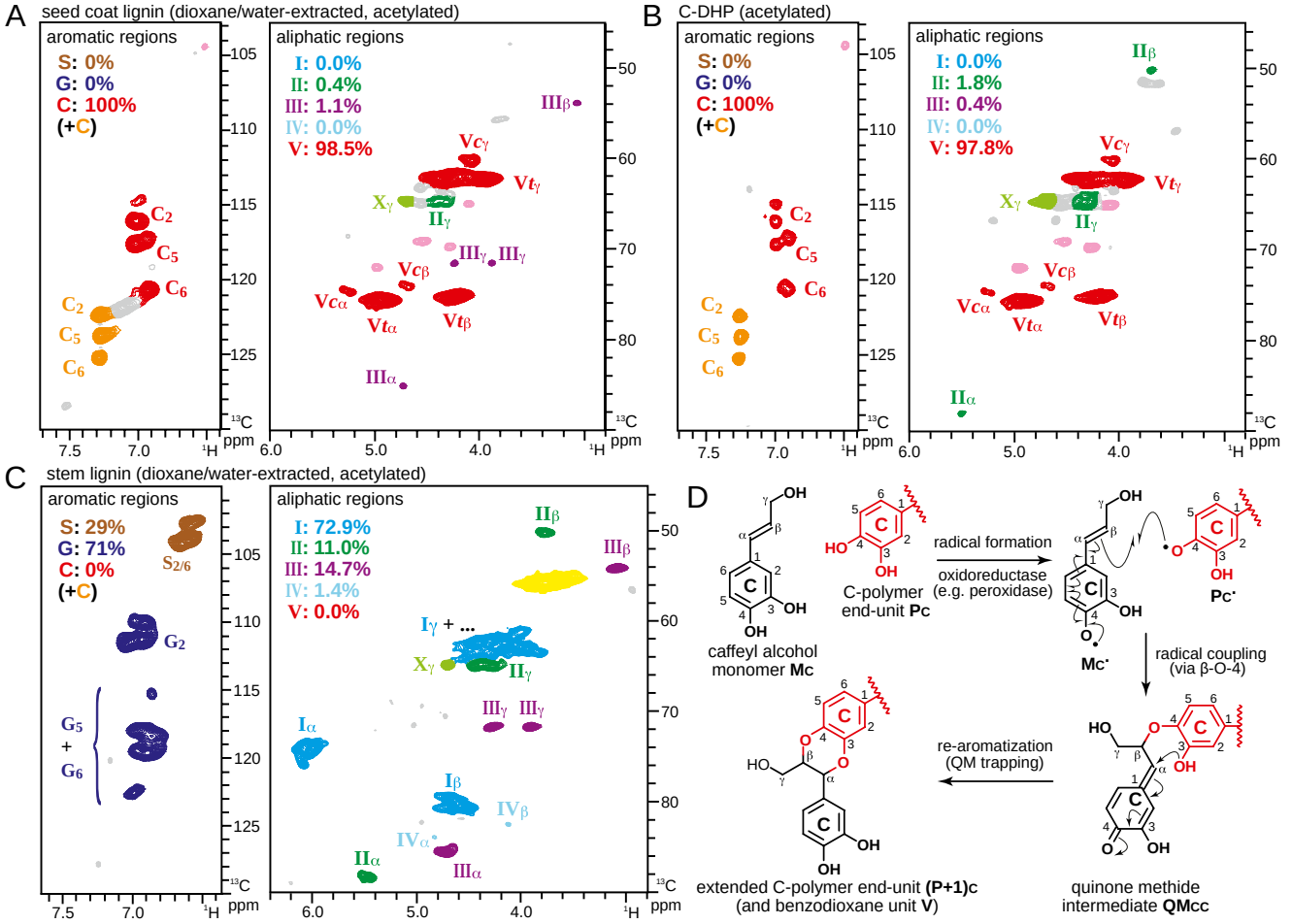

Fig. 3. Short-range ^{13}C–^{1}H correlation (HSQC) spectra of acetylated isolated lignins from *V. planifolia* seed coat (*A*) and stem (*C*) (extracted with dioxane-water, 96:4, vol/vol), and (*B*) an acetylated in vitro polymer (C-DHP), synthesized via peroxidase-catalyzed polymerization of caffeyl alcohol. For abbreviations for signal assignments, see Fig. 2. (*D*) Scheme for generation of benzodioxane units via radical cross-coupling reactions between caffeyl alcohol monomer M_C and catechyl (C)-polymer (P_C) end-units.

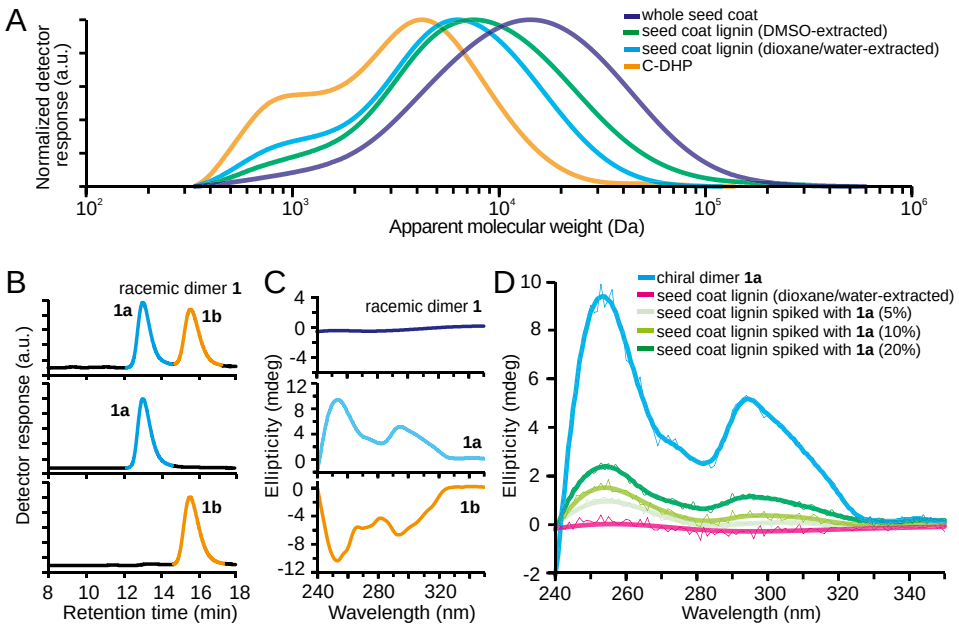

Fig. 4. Size and chirality of *V. planifolia* seed coat lignins. (*A*) GPC elution profiles of acetylated samples of ball-milled whole-seed coat, seed-coat lignins extracted with dimethyl sulfoxide or dioxane-water (96:4, vol/vol), and polymer (C-DHP) synthesized via peroxidase-catalyzed polymerization of caffeyl alcohol. (*B*) Chiral HPLC elution profiles, and (*C*) CD spectra of racemic benzodioxane dimer **1** (Fig. S4), and its separated enantiomers **1a** and **1b**. (*D*) CD spectra of purified phenolic polymer isolated from vanilla seed (seed coat lignin, extracted with dioxane-water, 96:4, vol/vol), chiral benzodioxane dimer **1a**, and seed polymer spiked with 5%, 10%, and 20% (by weight) chiral dimer **1a**; smoothed (thick lines) and raw (thin lines) spectra are presented.

The vanilla seed polymer is strikingly similar to the in vitro polymer synthesized via peroxidase-catalyzed polymerization of caffeyl alcohol. Both polymers are massively composed of benzodioxane units, which are uniquely formed via β–O–4-type radical coupling of the monomer (at its β-position) with a C-polymer end-unit (at its 4–O-position) followed by internal trapping of the quinone methide intermediates (QM) by the *o*-hydroxyl (3-hydroxyl) group in the C unit to form the six-membered ring (Fig. 3*D*). Similar benzodioxanes were recently identified in cell cultures of CCoAOMT-deficient *P. radiata* (22), and analogous benzodioxanes are products of lignification with 5-hydroxyconiferyl alcohol in COMT-deficient angiosperms (20, 21, 26, 49, 50). As the only β–O–4-type units were benzodioxanes, the postcoupling rearomatization of QM seems to be exclusively via the efficient internal trapping by the *o*-hydroxyl group in C units (Fig. 3*D*). The main benzodioxane backbones in the seed polymer are *trans*/*cis*-isomeric mixtures, as in the in vitro polymer and synthetic dimer, suggesting that the stereochemistry of postcoupling rearomatization of the QM is under simple kinetic chemical control. Therefore, it is most plausible that caffeyl alcohol is enzymatically oxidized, presumably by plant oxidoreductases such as peroxidases and laccases initially, but is cross-coupled onto the growing polymer in a chemically controlled fashion, independent of enzymes or other proteins, in the same way as conventional monolignols are during lignin polymerization (14, 16, 17).

In conclusion, the identification of this unique polymer provides compelling evidence for flexibility in the construction of lignin polymers in nature. The mechanisms that allow for formation of caffeyl alcohol in developing vanilla and cactus seeds, and the question of whether such catechyl polymers are much more widespread in nature, remain to be determined. Such studies might contribute to the development of new avenues in lignin bioengineering, and may also provide new insights into the diversity and evolution of land plants.

Materials and Methods

Plant Materials. Mature vanilla beans were provided by Bakto Flavors, separated into the seed and pod (residue left after seed isolation), and processed as described in *SI Materials and Methods*. Vanilla stem material was obtained from vines growing in the greenhouse at the Noble Foundation, Ardmore, OK. Seeds of all cactus species were obtained from flowering plants in the collection of one of the authors (R.A.D.).

Isolations of Vanilla Lignins. Vanilla lignin samples for NMR, GPC, and CD analyses were prepared via methods largely described previously (22, 33, 39), and as further described in *SI Materials and Methods*.

Synthetic Model Dimers. The benzodioxane dimer **1** was synthesized from radical coupling reactions of caffeyl alcohol via silver carbonate (Ag_2CO_3) oxidation, and dimer **2** was via methylation of dimer **1** with methyl iodide; detailed synthetic protocols and complete NMR and MS spectroscopic data are described in *SI Materials and Methods*.

Dehydrogenation Polymer from Caffeyl Alcohol. A dehydrogenation polymer from caffeyl alcohol (C-DHP) was generated via HRP-catalyzed polymerization, as previously described (22, 51, 52), and further outlined in *SI Materials and Methods*.

Chemical Analyses. Determination of Klason lignin, crystalline cellulose, amorphous sugars, protein content, proanthocyanidins and lignin composition (by thioacidolysis) were as described in *SI Materials and Methods*.

NMR Spectroscopy. The NMR methods used were largely described previously (22, 32, 33, 51), and as further described in *SI Materials and Methods*.

GPC. GPC was performed on a Shimadzu LC-20A LC system (Shimadzu) as further described in *SI Materials and Methods*.

Chiral HPLCy. Analytical and preparative chiral HPLC for enantiomeric separation of benzodioxane dimer **1** was performed on a Shimadzu LC-20A LC system as described in *SI Materials and Methods*.

CD Spectroscopy. CD spectra were run on an Model 202SF CD spectrophotometer (Aviv Biomedical) as described in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank Cliff Foster for carbohydrate analysis; Heike Hofstetter for elemental analysis; Darrel McCaslin for assistance with CD spectrometry; Li Shuai, and Sasikumar Elumalai for assistance with Klason lignin analysis; Lisa Jackson for assistance with thioacidolysis analysis; and Hoon Kim for helpful suggestions and assistance with NMR spectroscopy. This work was supported in part by the US Department of Energy's Bioenergy Sciences and Great Lakes Bioenergy Centers, supported by the Office of Biological and Environmental Research in the Department of Energy Office of Science (BER DE-AC05-00OR22725 and DE-FC02-07ER64494, respectively); and by a postdoctoral fellowship from the Japan Society for the Promotion of Science (to Y.T.).

1. Boudet AM, Kajita S, Grima-Pettenati J, Goffner D (2003) Lignins and lignocellulosics: A better control of synthesis for new and improved uses. *Trends Plant Sci* 8:576–581.
2. Guo DG, et al. (2001) Improvement of in-rumen digestibility of alfalfa forage by genetic manipulation of lignin *O*-methyltransferases. *Transgenic Res* 10:457–464.
3. Chen F, Dixon RA (2007) Lignin modification improves fermentable sugar yields for biofuel production. *Nat Biotechnol* 25:759–761.
4. Fu C, et al. (2011) Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. *Proc Natl Acad Sci USA* 108:3803–3808.
5. Weng JK, Chapple C (2010) The origin and evolution of lignin biosynthesis. *New Phytol* 187:273–285.
6. Vanholme R, Demedts B, Morreel K, Ralph J, Boerjan W (2010) Lignin biosynthesis and structure. *Plant Physiol* 153:895–905.
7. Simmons BA, Loqué D, Ralph J (2010) Advances in modifying lignin for enhanced biofuel production. *Curr Opin Plant Biol* 13:313–320.
8. Bonawitz ND, Chapple C (2010) The genetics of lignin biosynthesis: Connecting genotype to phenotype. *Annu Rev Genet* 44:337–363.
9. Weng JK, Li X, Bonawitz ND, Chapple C (2008) Emerging strategies of lignin engineering and degradation for cellulosic biofuel production. *Curr Opin Biotechnol* 19: 166–172.
10. Vanholme R, Morreel K, Ralph J, Boerjan W (2008) Lignin engineering. *Curr Opin Plant Biol* 11:278–285.
11. Reddy MSS, et al. (2005) Targeted down-regulation of cytochrome P450 enzymes for forage quality improvement in alfalfa (*Medicago sativa* L.). *Proc Natl Acad Sci USA* 102:16573–16578.
12. Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. *Annu Rev Plant Biol* 54: 519–546.
13. Ralph J (2006) What makes a good monolignol substitute? *The Science and Lore of the Plant Cell Wall Biosynthesis, Structure and Function*, ed Hayashi T (Universal Publishers; BrownWalker Press, Boca Raton, FL), pp 285–293.
14. Ralph J, et al. (2004) Lignins: Natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. *Phytochem Rev* 3(1):29–60.
15. Sederoff RR, MacKay JJ, Ralph J, Hatfield RD (1999) Unexpected variation in lignin. *Curr Opin Plant Biol* 2:145–152.
16. Ralph J, et al. (2008) Identification of the structure and origin of a thioacidolysis marker compound for ferulic acid incorporation into angiosperm lignins (and an indicator for cinnamoyl CoA reductase deficiency). *Plant J* 53:368–379.
17. Ralph J, et al. (2008) Lignification: Are lignins biosynthesized via simple combinatorial chemistry or via proteinaceous control and template replication? *Recent Advances in Polyphenol Research*, eds Daayf F, El Hadrami A, Adam L, Ballance GM (Wiley-Blackwell Publishing, Oxford, UK), Vol 1, pp 36–66.
18. Li L, Popko JL, Umezawa T, Chiang VL (2000) 5-hydroxyconiferyl aldehyde modulates enzymatic methylation for syringyl monolignol formation, a new view of monolignol biosynthesis in angiosperms. *J Biol Chem* 275:6537–6545.
19. Parvathi K, Chen F, Guo DJ, Blount JW, Dixon RA (2001) Substrate preferences of *O*-methyltransferases in alfalfa suggest new pathways for 3-*O*-methylation of monolignols. *Plant J* 25:193–202.
20. Weng J-K, Mo H, Chapple C (2010) Over-expression of F5H in COMT-deficient *Arabidopsis* leads to enrichment of an unusual lignin and disruption of pollen wall formation. *Plant J* 64:898–911.
21. Vanholme R, et al. (2010) Engineering traditional monolignols out of lignin by concomitant up-regulation of F5H1 and down-regulation of COMT in *Arabidopsis*. *Plant J* 64:885–897.
22. Wagner A, et al. (2011) CCoAOMT suppression modifies lignin composition in *Pinus radiata*. *Plant J* 67:119–129.
23. Zhong RQ, Iii WH, Negrel J, Ye ZH (1998) Dual methylation pathways in lignin biosynthesis. *Plant Cell* 10:2033–2046.
24. Zhong RQ, Morrison WH, 3rd, Himmelsbach DS, Poole FL, 2nd, Ye ZH (2000) Essential role of caffeoyl coenzyme A *O*-methyltransferase in lignin biosynthesis in woody poplar plants. *Plant Physiol* 124:563–578.
25. Meyermans H, et al. (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* 275:36899–36909.
26. Marita JM, et al. (2003) Structural and compositional modifications in lignin of transgenic alfalfa down-regulated in caffeic acid 3-*O*-methyltransferase and caffeoyl coenzyme A 3-*O*-methyltransferase. *Phytochemistry* 62:53–65.
27. Do C-T, et al. (2007) Both caffeoyl Coenzyme A 3-*O*-methyltransferase 1 and caffeic acid *O*-methyltransferase 1 are involved in redundant functions for lignin, flavonoids and sinapoyl malate biosynthesis in *Arabidopsis*. *Planta* 226:1117–1129.
28. Zhou R, et al. (2010) Distinct cinnamoyl CoA reductases involved in parallel routes to lignin in *Medicago truncatula*. *Proc Natl Acad Sci USA* 107:17803–17808.
29. Lapierre C (1993) Application of new methods for the investigation of lignin structure. *Forage Cell Wall Structure and Digestibility*, eds Jung HG, Buxton DR, Hatfield RD, Ralph J (American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, Madison, WI), pp 133–166.
30. Rolando C, Monties B, Lapierre C (1992) Thioacidolysis. *Methods in Lignin Chemistry*, eds Dence CW, Lin SY (Springer-Verlag, Berlin, Heidelberg), pp 334–349.
31. Dixon RA (2011) Vanillin biosynthesis–Not as simple as it seems? *Handbook of Vanilla Science and Technology*, eds Havkin-Frenkel D, Belanger FC (Blackwell Publishing Ltd, Chichester, UK), pp 292–298.
32. Lu F, Ralph J (2003) Non-degradative dissolution and acetylation of ball-milled plant cell walls: High-resolution solution-state NMR. *Plant J* 35:535–544.
33. Kim H, Ralph J (2010) Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO-$d_{(6)}$/pyridine-$d_{(5)}$. *Org Biomol Chem* 8:576–591.
34. Peel GJ, Pang Y, Modolo LV, Dixon RA (2009) The LAP1 MYB transcription factor orchestrates anthocyanidin biosynthesis and glycosylation in Medicago. *Plant J* 59: 136–149.
35. Lepiniec L, et al. (2006) Genetics and biochemistry of seed flavonoids. *Annu Rev Plant Biol* 57:405–430.
36. Pang Y, Peel GJ, Wright E, Wang Z, Dixon RA (2007) Early steps in proanthocyanidin biosynthesis in the model legume *Medicago truncatula*. *Plant Physiol* 145:601–615.
37. Björkman A (1954) Isolation of lignin from finely divided wood with neutral solvents. *Nature* 174:1057–1058.
38. Ralph J, et al. (2006) Effects of coumarate 3-hydroxylase down-regulation on lignin structure. *J Biol Chem* 281:8843–8853.
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* 150:621–635.
40. Ralph J, Landucci LL (2010) NMR of Lignins. *Lignin and Lignans; Advances in Chemistry*, eds Heitner C, Dimmel DR, Schmidt JA (CRC Press, Boca Raton, FL), pp 137–234.
41. Lu F, Ralph J (2011) Solution-state NMR of lignocellulosic biomass. *J Biobased Mater Bio* 5(2):169–180.
42. Faix O, Lange W, Beinhoff O (1980) Molecular weights and molecular weight distributions of milled wood lignins of some wood and *Bambusoideae* species. *Holzforschung* 34(5):174–176.
43. Faix O, Lange W, Salud EC (1981) The use of HPLC for the determination of average molecular weights and molecular weight distributions of milled wood lignins from *Shorea polysperma* (Blco.). *Holzforschung* 35(3):3–9.
44. Tobimatsu Y, Takano T, Kamitakahara H, Nakatsubo F (2008) Studies on the dehydrogenative polymerizations (DHPs) of monolignol beta-glycosides: Part 4. Horseradish peroxidase-catalyzed copolymerization of isoconiferin and isosyringin. *Holzforschung* 62:495–500.
45. Akiyama T, et al. (2000) Proof of the presence of racemic forms of arylglycerol-β-aryl ether structure in lignin: Studies on the stereo structure of lignin by ozonation. *J Wood Sci* 46:414–415.
46. Ralph J, Peng J, Lu F, Hatfield RD, Helm RF (1999) Are lignins optically active? *J Agric Food Chem* 47:2991–2996.
47. Nyffeler R (2002) Phylogenetic relationships in the cactus family (Cactaceae) based on evidence from trnK/ matK and trnL-trnF sequences. *Am J Bot* 89:312–326.
48. Joel DM, et al. (2003) A hairy tissue produces vanillin. *Isr J Plant Sci* 51:157–159.
49. Ralph J, et al. (2001) NMR evidence for benzodioxane structures resulting from incorporation of 5-hydroxyconiferyl alcohol into Lignins of *O*-methyltransferase-deficient poplars. *J Agric Food Chem* 49:86–91.
50. Lu F, et al. (2010) Sequencing around 5-hydroxyconiferyl alcohol-derived units in caffeic acid *O*-methyltransferase-deficient poplar lignins. *Plant Physiol* 153:569–579.
51. Tobimatsu Y, Davidson CL, Grabber JH, Ralph J (2011) Fluorescence-tagged monolignols: Synthesis, and application to studying in vitro lignification. *Biomacromolecules* 12:1752–1761.
52. Tobimatsu Y, Takano T, Kamitakahara H, Nakatsubo F (2008) Studies on the dehydrogenative polymerizations of monolignol β-glycosides. Part 3: Horseradish peroxidase-catalyzed polymerizations of triandrin and isosyringin. *J Wood Chem Technol* 28:69–83.