Unexpected variation in lignin
Ronald R Sederoff*, John J MacKay†, John Ralph‡ and Ronald D Hatfield§

Recent studies on mutant and transgenic plants indicate that lignification may be far more flexible than previously realized. Pines with a mutation affecting the biosynthesis of the major lignin precursor, coniferyl alcohol, show a high level of an unusual subunit, dihydroconiferyl alcohol. These results argue in favor of an increased potential for genetic modification of lignin and indicate that our knowledge of the biosynthesis of lignin is far from complete.

Introduction
Lignification is a process essential to the nature and evolution of vascular plants that is still poorly understood, even though it has been studied for more than a century. Lignin is unusual compared to other abundant natural polymers due to the low degree of order and the high degree of heterogeneity in its structure [1,2]. It is also unusual as a plant polymer in that there are no plant enzymes for its degradation. Unexpected variation in lignin subunit composition has been found recently, particularly in an unusual mutation affecting the wood of loblolly pine [3,4] and also in mutants and genetically engineered variants of a diverse range of higher plants [4–6]. These results have significant implications regarding the traditional definition of lignin, and highlight the need for a better understanding of the lignin precursor biosynthetic pathway. We believe that the observed variation in composition and structure of lignin is still best explained by variation of the monolignol precursors and their abundance in the lignifying zone. The plasticity in lignin composition reveals new potential that extends beyond the traditional monolignol pathway for modification of the polymer by genetic engineering.

Our arguments for a greater level of plasticity in lignin, through variation in precursor composition, have recently been challenged [7–9]. Here, we review and extend our results to support our structural findings and present our conclusions that these structures represent normal lignification with unusual precursors, consistent with a traditional paradigm for lignification. Our view is in sharp contrast with a recent model of lignin biosynthesis requiring template dependent stereospecific control of lignin polymerization [9,10]. Structural information has long been used to guide the search for underlying mechanisms for important biological processes, and the biosynthesis of lignin is no exception. The combination of current methods of structural chemistry, biochemistry, cell biology and genetics should continue to elucidate the nature and origin of the lignin polymer.

Lignin is conventionally defined as a complex hydrophobic network of phenylpropanoid units derived from the oxidative polymerization of one or more of three types of hydroxycinnamyl alcohol precursors [11–16]. These alcohols give rise to p-hydroxyphenyl, guaiacyl and syringyl subunits in lignin (Figure 1). The precursors are themselves derived from phenylalanine by deamination, followed by hydroxylation of the aromatic ring, methylation, and the reduction of the terminal acidic group to an alcohol. These alcohols have long been thought to be the direct precursors for lignin (monolignols). The lignin precursors can radically couple at several sites with each other, or, more frequently, with the growing lignin oligomer, to produce a complex polymer with a variety of intermolecular linkages [15,17–19]. At least 20 types of linkages have been described, and it is likely that many more are present in low proportions. Here, we focus on a mutation in the last step of the precursor pathway: the formation of the monolignol coniferyl alcohol from coniferaldehyde. This step is catalyzed by the enzyme cinnamyl alcohol dehydrogenase (CAD; EC 1.1.1.195) encoded by a single gene in loblolly pine [20,21].

Discovery of a CAD-deficient pine mutant
The discovery of a recessive mutant allele of the cad gene, cad-n1, in loblolly pine has permitted the study of pines with severe deficiencies of CAD enzyme [3,4]. The secondary xylem (wood) in cad-n1 homozygous seedlings acquires a brown color, very distinct from the nearly white color of wild-type pine wood (Figure 2). CAD deficiency causes dramatic changes in the accumulation and nature of soluble phenolics; it also alters the structure of the lignin polymer that is deposited in the cell wall. The color and many of the changes in wood chemistry are similar to those observed in transgenic plants and in brown midrib mutants with suppressed CAD activity [22–24]. The incorporation of one compound, however, dihydroconiferyl alcohol (DHCA) — a major lignin precursor in pines with the cad-n1 mutation [4] — has not yet been noted in other plants with decreased CAD activity. These plants need to be re-examined using newer NMR and degradative methods that reveal sidechain
**Figure 1**  
Lignin precursor biosynthesis. The general monolignol precursor biosynthetic pathway is outlined from 4-coumarate to five lignin precursors. Only three enzymes are shown: 4CL (hydroxycinnamyl-CoA ligase), CCR (hydroxycinnamoyl-CoA:NADPH oxidoreductase) and CAD (cinnamyl alcohol dehydrogenase). For details of the biosynthetic pathway and variation in precursor biosynthesis, the reader is referred to [6,74,75].

Enzymology and molecular characterization of CAD in loblolly pine

CAD from differentiating xylem has been extensively characterized in loblolly pine [20]. The native protein dimer (82,000Da) was purified to homogeneity. The enzyme has a high affinity for coniferaldehyde (Km = 1.7 µm) and a much lower affinity for sinapaldehyde. It is the only form of CAD detected in this tissue, which is predominantly composed of differentiating tracheids active in lignification. The wild-type cad gene has been characterized by cDNA cloning, genomic cloning and Southern blot analyses [20,21]. At least five different allozyme variants are present in loblolly pine, which segregate as expected for alleles of a single structural gene. The gene encoding CAD has been located on a loblolly pine genetic map [3].

Characterization of the cad-n1 allele and inheritance of the mutant phenotype

The cad-n1 allele was originally discovered in loblolly pine (genotype clone 7–56) through allozyme studies of the cad gene [3]. These studies made use of the haploid maternally derived genome of the megagametophyte in pine seeds. Inheritance experiments using haploid seed tissue showed that the CAD deficiency was due to a recessive mutant allele of the cad gene. By further comparing the allozyme phenotypes of haploid and diploid tissues in individual seeds from diverse crosses, we established that the mutation behaves as a recessive mutation with no trans-acting effects, and which maps to, or very close to, the cad locus [3]. Investigation of the inheritance of the cad-n1 allozyme phenotype using diploid xylem tissues was also consistent with the above observations and indicates that CAD deficiency was not restricted to seed tissues. In the xylem of cad-n1 homozygotes, the level of enzyme activity is approximately 1% of wild-type. The steady-state transcript level indicates a reduction of CAD mRNA in the mutant, sufficient to account for the low level of enzyme activity [3]. Genetic association of the brown wood phenotype with
the $\text{cad-n1}$ homozygous genotype is unambiguous. Although localized brown wood color in pine may be induced by stress, a genetic basis for the brown color of the $\text{cad-n1}$ mutant is substantial. The brown wood (mutant) phenotype segregates in a 3:1 ratio of wild-type:mutant when two heterozygous $\text{cad-n1}$ trees are crossed, following Mendelian expectations for a single recessive gene. Among the segregating offspring, the mutant phenotype is found only in seedlings that inherit two copies of the $\text{cad-n1}$ gene [3]. We have obtained homozygous $\text{cad-n1}$ (mutant) loblolly pine from several crosses between heterozygous trees that demonstrate inheritance over three generations. The homozygous trees are still too young for crossing. The genetic basis of the brown wood color is further confirmed by its distinct, uniform and constant appearance throughout the wood, whether plants were grown in the greenhouse, in a nursery environment, or in field plantings (2 and 12 years of age).

**Accumulation of the substrate of CAD and the $\text{cad-n1}$ mutant phenotype**

Chemical analyses of the wood from $\text{cad-n1}$ homozygotes all indicate a dramatic increase in the levels of coniferaldehyde, the predominant substrate of CAD enzyme in pine, and in structures derived from coniferaldehyde [3,4,25]. Simple extraction of ground wood with alcohol or other solvents revealed a several fold increase in soluble phenolics and a large increase in coniferaldehyde and vanillin (4-hydroxy-3-methoxybenzaldehyde) [3,4]. The resultant lignin polymer contains coniferaldehyde and vanillin at levels significantly higher than the wild-type. This inference is based on results from isolated lignin samples analyzed by NMR, FTIR and UV spectroscopy [3,4,26], and from extractive free wood samples by microphotospectrometry, FTIR, PyrMS and thioacidolysis ([3,27,28]; C Lapierre, unpublished data). The development of a brown or red-brown color in lignified tissues of plants that have suppressed CAD activity is well documented in both transgenics [22,23] and mutants [24]. In addition, in vitro synthesis of lignin oligomers (DHPs) starting with a mixture of coniferaldehyde and coniferyl alcohol generates dark red product, as opposed to a nearly colorless product when the aldehyde is absent or at low levels [29,30].

**Discovery of dihydroconiferyl alcohol (DHCA) subunits in $\text{cad-n1}$ homozygotes**

The lignin from $\text{cad-n1}$ homozygotes shows a large increase of unexpected subunits derived from DHCA [4]. The identity and abundance of these subunits were first determined in an isolated lignin preparation by diagnostic NMR experiments (Figure 3). These findings were confirmed by direct chemical analysis of the isolated lignin using the DFRC method [31,32] and analysis of extractive free wood samples by DFRG, PyrMS [28] and thioacidolysis (C Lapierre, unpublished data). Greatly elevated levels of DHCA have been observed in wood of several $\text{cad-n1}$ homozygous seedlings and trees obtained from several crosses as outlined above; DHCA accumulation was not detected in any of the wild-type seedlings. The association between this phenotype and the $\text{cad-n1}$ homozygous genotype is strong and is not dependent on environmental or biotic stress. DHCA is a known minor component of lignins [25,33–35]. It is, therefore, not a new product in the mutant, but one found at highly elevated levels.

**Dihydroconiferyl-alcohol-derived subunits are components of lignin**

Combined chemical degradation and NMR provide unambiguous evidence that DHCA is a *bona fide* and abundant component of the lignin polymer [25]. The non-extractable lignin fractions can be dissolved in acetyl bromide, where NMR again confirms the presence of DHCA [25]. Similar results in support of this claim are obtained whether the analyzed material is milled wood lignin, total wood, or fully solubilized residual lignin from the mutant pine. Comparison with synthetic lignins and model compounds indicates that approximately 50% of the subunits are in 5–5-coupled structures (Figure 3) [4]. The presence of monomeric DHCA in solvent extracts and its predominance in 5–5-linked structures argues strongly against the suggestion [8,36] that the DHCA components are the result of a modification of lignin following coupling of coniferyl alcohol. Detailed examination of dimers from degradation of normal pine lignin by the DFRC method provides evidence that the coniferyl alcohol monomer is essentially not involved in 5–5-coupling reactions ([37], J Peng, J Ralph, F Lu, unpublished data). It has also been suggested that DHCA may be a dioxane:water extractable oligomerized

![Figure 2](image_url)

Wood samples from the $\text{cad-n1}$ mutant (right) and normal pine (left) are shown. Trees are 12 years old. For details on the biological material, the reader is referred to [3] and for lignin analysis to [4,5].
lignan artifact (lignans are nonstructural dimeric phenolic metabolites) [7,8]. Degradative cleavage evidence (DFRC and thioacidolysis) indicates, however, that DHCA is present in all lignin fractions. In addition, the isolated lignin of a CAD deficient tree had a weight-average molecular weight of ~17,000 and did not contain a low molecular weight fraction (J MacKay and Dimmel, unpublished data; F Lu and J Ralph, unpublished data). The abundance of DHCA, however, is unlikely to be attributable to contaminating lignans, which would be much lower in molecular weight.

In fact, since lignans in *Pinus taeda* appear to be optically active [38], the inability to detect optical activity in the mutant lignin (J Ralph, J Peng, F Lu, RD Hatfield; unpublished data) may be a sufficient counter-argument.

The origin of DHCA in the mutant pine and in other normal softwoods is unknown. An NADPH-dependent enzyme activity is reported to reduce the cinnamyl alcohol double bond in β-5 dilignol dehydrodiconiferyl alcohol to produce dihydrodehydrodiconiferyl alcohol [8]. This activity is an interesting candidate for involvement in the conversion of coniferaldehyde to dihydroconiferyl alcohol and should be carefully tested.

**Composition and content of lignin in other mutants and transgenic plants**

Mutant or transgenic plants with genetic deficiencies affecting enzyme activity in the lignin biosynthetic pathway often have novel lignin structures or modified lignin composition, suggesting a high level of metabolic plasticity in lignin precursors [4–6,15,16,39]. A mutation in the enzyme ferulate-5-hydroxylase results in lignin without a syringyl component, whereas an overexpressing transgenic variant produces a lignin that is almost entirely composed of syringyl units [40–42]. Minor components of normal lignins can become more significant when other key enzymes are depleted. Naturally occurring mutants (*e.g.*, the brown-midrib (*bm3*) mutants of maize and sorghum) and transgenic plants deficient in *O*-methyl transferase (OMT) contain significant amounts of units derived from 5-hydroxyconiferyl alcohol [43–47]. Tobacco downregulated in cinnamoyl-CoA reductase (CCR) shows a striking increase in tyramine ferulate [19], a logical sink for the anticipated build-up of feruloxy-CoA. Tyramine ferulate is also seen as a stress response product in tobacco [48,49]. Coniferaldehyde becomes more significant in the CAD deficient pine mutant, radically coupling with aldehydes or with lignin monomers/oligomers [4]. Similarly, sinapaldehyde becomes a major component of antisense-CAD tobacco transgenics [19].

**Variability of lignin composition indicates a high level of metabolic plasticity based upon precursor supply**

The ability of plants to adapt to diverse and large changes in the precursor supply indicates that there is considerable metabolic plasticity in the assembly of the lignin polymer. Lignin must encompass a wider array of phenolic structures, with its composition and structure primarily guided by the precursor supply in the lignifying zone within plant tissues. The precursor supply varies among plant taxa [50]; it also varies among cell types and within the cell wall, thus resulting in macro and micro heterogeneity of lignin itself [2,39]. For example, the wood formed in gymnosperms on the underside of a leaning stem is known as compression wood. The increase in *p*-hydroxyphenyl subunits in lignin from compression wood is well-established [51]. The precursor supply is also affected by genetic lesions or variants that create additional variation in lignin structure as discussed here [4,16,39,40].
Other nontraditional subunits are found in lignins

Lignin is often defined from the compositional standpoint as a polymer derived from three hydroxycinnamyl alcohols (the monolignols: \(p\)-coumaryl, coniferyl and sinapyl alcohols) by a dehydrogenative polymerization involving radical coupling. This definition, however, has long been recognized as too narrow [52]. Many plants have lignins containing significant levels of other unusual components (Figure 4), and it is likely that no plant contains lignins that are solely derived from the three ‘primary’ precursors. For example, all lignins contain aldehyde groups [53,54] – it is this feature that provides the diagnostic lignin staining reaction with acid phloroglucinol [55]. Evidence from mutants and genetic variants where aldehydes accumulate strongly supports the view that aldehydes are incorporated as precursors, because, in these variants, more aldehydes are found in the lignin [3–5,22,23].

Many lignins are biosynthesized by incorporating esterified monolignols into the lignification scheme. Thus, grasses utilize \(p\)-coumarates [56–60], hardwoods and some dicots such as kenaf utilize acetates [61,62], and some plants, notably bamboo, aspen and willow, use \(p\)-hydroxybenzoates as ‘monomers’ for lignification [57,58,63]. Ferulates and diferulates are found intimately incorporated into all grass and some dicot lignins, where they are equal partners in the free-radical polymerization process and may even be nucleation sites for lignification [5,64,65]. Amides may also be incorporated; although it is a wounding response product, tyramine ferulate is found in various lignins, for example, in tobacco [19,48,49]. A general definition of lignin must include more than the traditional three hydroxycinnamyl alcohols, or the phenolic polymers in many plants serving the structure and function of lignin, for example, in grasses, might not be considered lignin. Lignin components do appear to be derived from phenylpropanoids as a general class; this classification has been used frequently [11,13,66].

Redefining lignin

It is often difficult to distinguish cell wall lignin from other polymerized infusions [67]: for example, normal lignification versus wound-response lignin. Definitions of lignin that are both specific enough and general enough are difficult to find. For example, a recent book on ‘Lignin and Lignan Biosynthesis’ does not provide such a definition even in the first chapter, entitled “Lignin and Lignan Biosynthesis: Distinctions and Reconciliations” [36]. Attempts to define lignin in terms of its function within the plant are also not clear-cut. Lignin is often defined for specific applications: for example, it is often regarded as little more than the (non-extractable) phenolic component that must be fragmented or degraded to produce pulp and paper. We prefer to consider lignin as an operational term for a diverse class of naturally occurring phenolic polymers that need to be more precisely defined in each case, on the basis of the source and

Figure 4

Non-monolignol ‘monomers’ that are known or thought to be intimately incorporated into lignins. Sites for radical coupling are indicated with an arrow; a dashed arrow indicates other possible sites (coupling behavior not known).
the method of purification or analysis. Until there is better agreement on how lignin should be defined, it will not be surprising if there is continued controversy about its origin, composition, properties and biosynthesis.

Order and randomness in lignin and lignification

The issue of randomness or the degree of order in lignin formation also remains controversial. Researchers occasionally speculate about the possibility of crystallinity [47]. Studies on simple lignin trimers led one of us to describe lignin as a stereochemical nightmare with crystallinity in the traditional sense being astronomically improbable [65]. That is not to say that other aspects of order in the polymer, such as the alignment of aromatic rings indicated by Raman spectroscopy [68], are impossible. Randomness has been used to imply that the process appeared to have no direct enzymatic control. Obviously, the polymerization of lignin is not completely random, only a subset of the possible linkages are found, and the linkage types are not equally probable. In fact, prior descriptions as ‘random’ presumably never implied a totally random distribution of coupling products; clearly, the coupling of two monomers must be weighted by the propensity for each type of coupling. Lignin formation, however, has little to do with monomer coupling — lignification, unlike lignan formation, almost entirely involves coupling of a single monomer to a growing oligomer. The idea that the coupling would depend on the types of units involved, their concentrations, the matrix and so on are all issues that can determine what type of coupling arises at any step.

Order in the lignin polymer could result from processes of self-assembly, or alternatively, specific finely orchestrated steps might be involved in the assembly. Lewis has argued that “It is inconceivable that lignin formation would be left to the vagaries of such a wide range of enzymes or be realized in a haphazard manner” [10]. A more general opposing view of lignin is succinctly stated by Denton [69], who argued that “Biological defense is well served by loosely ordered chemistry. Where many randomly linked products are needed, sloppy pathways are economical.” It may be that the lack of stereospecificity is required for lignin function. The order inherent in the properties of small molecules may be highly relevant in biological systems and not require direct stereospecific enzymatic control.

Lewis et al. claimed that we have argued for a new and unsupported paradigm for lignin polymerization [8,36]. This is clearly not the case. Rather, we believe that the properties of abnormal lignins are simply due to changes in the relative abundance of precursors likely to be found in normal lignins. For example, the lignin in Arabidopsis can be converted from a guaiacyl-syringyl lignin to a predominantly guaiacyl lignin or to an almost exclusively syringyl lignin by modification of the level of expression of the gene for ferulate-5-hydroxylase [42,70]. Much of the variation in lignin, whether due to genetic or environmental causes, can be readily explained by a change in the relative abundance of the precursors delivered to the lignifying zone. All that is required beyond delivery of precursor is that the mechanism for polymerization be a general one.

Our results challenge the need for the new paradigm proposed by Lewis et al. [9,10,36]. Recent isolation of a ‘dirigent’ protein from Forsythia which facilitates coniferyl alcohol radial coupling to produce the lignan pinoresinol in a regio- and stereoselective manner has led to a proposed new mechanism for lignin biosynthesis because of the similarity of the phenolic precursors. According to their model, lignins form from template arrays of dirigent proteins and are synthesized with absolute structural control [9]. With the possibility that the synthesized lignin chain then structurally dictates the next chain by a template-polymerization process [71], the model resembles the mechanism of biosynthesis of more highly ordered biological polymers such as cellulose or proteins. This idea is intriguing, but highly speculative and currently devoid of evidence.

Exquisitely synthesized polymers should produce an array of discrete products following degradation by such procedures as ether cleaving reactions (acidolysis, thioacidolysis, the DFRC method and high temperature base). Instead, such degradative methods produce a continuous array of oligomers with no members of the series obviously missing. We have not been able to detect any hint of optical activity in various isolated lignins nor in degradation products which retain the optical centers produced in the coupling step (J Ralph, J Peng, F Lu, RD Hatfield, unpublished data). The new paradigm proposal cites two possible explanations for the “perceived lack of optical activity of lignins” [9]. One explanation is ‘two distinct types of proteins each encoding formation of complementary chains that effectively cancel out any measurable optical activity’. This idea requires that the plant would go to the energetically extreme measure of creating an optically active lignin polymer only to carefully negate that structural feature via a complementary set of proteins for which an additional (complementary) biochemical pathway must also be supported. This effort is to produce two complementary lignin polymers when each has identical physical properties, identical to those of the racemic mixture — some (presumably defense-related) reason would have to be envisaged, yet, producing a variety of structures/stereochemistries is an asset in defense. The template argument (that “complementary mirror images form via template replication”) has less serious detractions, although no evidence of any structural replication ability has yet appeared. As lignins are found intimately associated with hemicelluloses, it is not clear how discontinuities in the alignment of one lignin chain might contribute to excess optical activity (of the incompletely replicated section). We look forward to cogent arguments, rationales and diagnostic experimental evidence for the complex issues involved in the proposed new lignification paradigm. Until experimental evidence is provided, the extension of the lignan dirigent protein observations to lignin, as a paradigm for lignification, is without substance.
We have not seen any experimental data that require a precisely controlled synthesis of structurally-defined lignin, that is to say, data that cannot be supported by simply recognizing that the plant does exquisitely, temporally and spatially control the supply of monolignols, oxidizing enzymes and oxidative species. The matrix environment of the polymerization may alter the interunit linkage composition, as has been well demonstrated in synthetic lignification experiments [72,73]. The regulated differences in lignification in various cells and various regions of the cell, in wounding or in stress, demand little more than changes in monomer supply.

Conclusions
Recent genetic studies have shown that manipulating specific lignin-biosynthetic-pathway genes produces profound alterations in the phenolic components of plants. Whether the polyphenolic components produced by radical coupling reactions should be called lignin is little more than semantics. Although the ‘lignins’ in mutant and transgenic plants may appear to be strikingly different from ‘normal lignins,’ findings indicate that they represent merely broad compositional shifts. All of the novel units that have been found to date appear to be minor units in normal lignins. The recognition that such minor units can incorporate into lignin provides significantly expanded opportunities for engineering the composition and consequent properties of lignin.

Acknowledgements
This work has been partially supported by grants from the USDA-NRI 97-35/03-4796, the US Department of Energy DE-FG05-92ER20085, the NIH GM45344-06 and the NCSU Forest Biotechnology Industrial Research Consortium.

References


