Regiochemical control of monolignol radical coupling: a new paradigm for lignin and lignan biosynthesis

David R Gang[1], Michael A Costa[1], Masayuki Fujita[1]*, Albena T Dinkova-Kostova[1†], Huai-Bin Wang[1], Vincent Burlat[1], William Martin[1], Simo Sarkanen[2], Laurence B Davin[1] and Norman G Lewis[1]

Background: Although the lignins and lignans, both monolignol-derived coupling products, account for nearly 30% of the organic carbon circulating in the biosphere, the biosynthetic mechanism of their formation has been poorly understood. The prevailing view has been that lignins and lignans are produced by random free-radical polymerization and coupling, respectively. This view is challenged, mechanistically, by the recent discovery of dirigent proteins that precisely determine both the regiochemical and stereoselective outcome of monolignol radical coupling.

Results: To understand further the regulation and control of monolignol coupling, leading to both lignan and lignin formation, we sought to clone the first genes encoding dirigent proteins from several species. The encoding genes, described here, have no sequence homology with any other protein of known function. When expressed in a heterologous system, the recombinant protein was able to confer strict regiochemical and stereochemical control on monolignol free-radical coupling. The expression in plants of dirigent proteins and proposed dirigent protein arrays in developing xylem and in other lignified tissues indicates roles for these proteins in both lignan formation and lignification.

Conclusions: The first understanding of regiochemical and stereochemical control of monolignol coupling in lignan biosynthesis has been established via the participation of a new class of dirigent proteins. Immunological studies have also implicated the involvement of potential corresponding arrays of dirigent protein sites in controlling lignin biopolymer assembly.

Addresses: [1]Institute of Biological Chemistry, Washington State University, Pullman, WA 99164-6340, USA. [2]Department of Wood and Paper Science, University of Minnesota, Saint Paul, MN 55108-6128, USA.

Present addresses: *Faculty of Agriculture, Kagawa University, Kagawa, 761-0795, Japan. †Department of Pharmacology and Molecular Sciences, Johns Hopkins University School of Medicine, Baltimore, MD 21205-2185 USA.

Correspondence: Norman G Lewis
E-mail: lewisn@wsu.edu

Key words: dirigent protein, lignans, lignins, radical coupling

Received: **1 December 1998**
Revisions requested: **4 January 1999**
Revisions received: **12 January 1999**
Accepted: **13 January 1999**

Published: **16 February 1999**

Chemistry & Biology March 1999, **6**:143–151
http://biomednet.com/elecref/1074552100600143

 ISSN 1074-5521

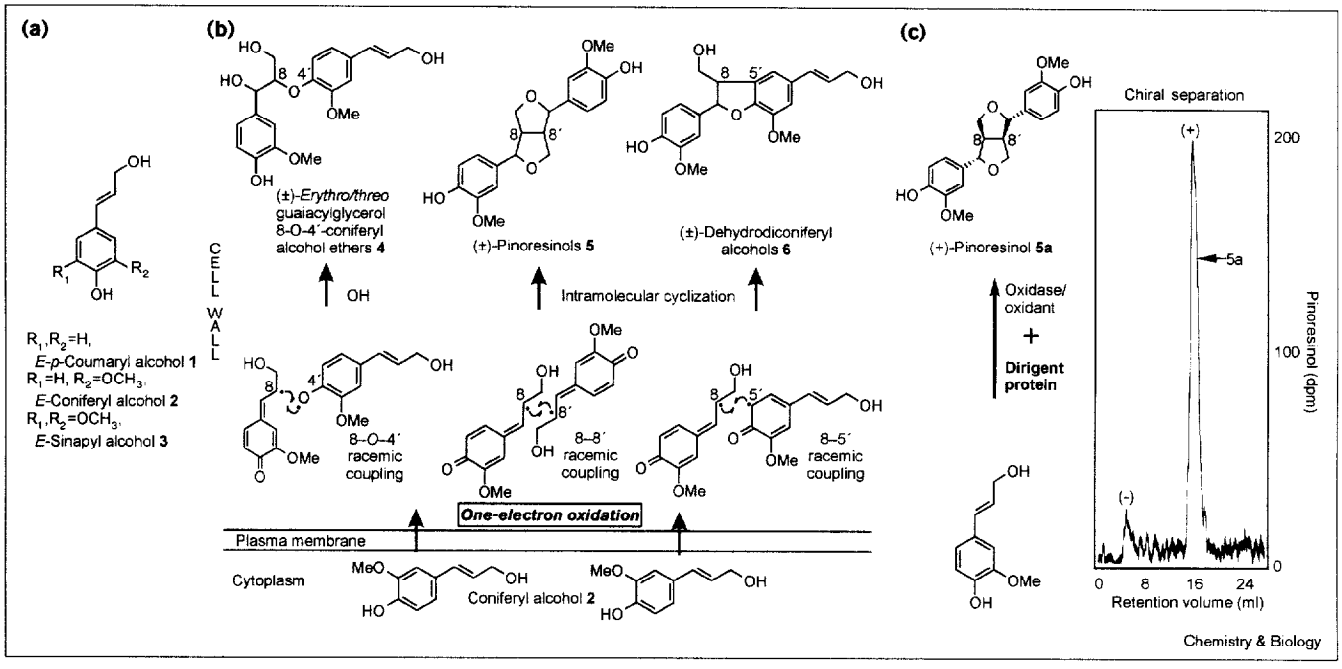
Introduction

The polymeric lignins and (oligomeric) lignans, both monolignol-derived coupling products, are major constituents of vascular plants and account for nearly 30% of the organic carbon circulating in the biosphere. Since the 1950s, however, dehydrogenative monolignol coupling —particularly that affording lignins, nature's second-most abundant biopolymers — has been postulated to be a random process. According to this view, elaborated at some length by Freudenberg [1,2], three monolignols **1**–**3** (Figure 1a) that are all derived from phenylalanine undergo random free-radical coupling *in vivo* through a process that ultimately produces the lignin biopolymers.

The original random-coupling hypothesis can be illustrated using the monolignol *E*-coniferyl alcohol **2** (Figure 1b; *E*, entgegen) as an example. The only ostensible prerequisite for coupling leading to the lignin biopolymer was proposed to be the provision of sufficient oxidative capacity (one-electron oxidation via an oxidase) to generate the free-radical species [1–4]. This assertion, however, was based largely on the observation that random coupling can occur *in vitro* at several atomic centers on the monolignol radical with the major coupling modes resulting in the three racemic 8–*O*–4′, 8–8′ and 8–5′ linked dimers **4**–**6** shown (Figure 1b). For lignin formation *in vivo*, further oxidative coupling of the dimers is required but there was believed to be no requirement for any additional enzyme or protein. The original notion that lignin biopolymers and monolignol dehydropolymerisates (generated by random coupling only) could be identical represents, in fact, the central tenet of the random-coupling hypothesis for macromolecular lignin assembly [1,2]. In stark contrast to any other known biochemical system, lignin formation was therefore envisaged to be satisfactorily reproduced in a test-tube by random encounters between its free-radical precursors.

The random-coupling hypothesis was completely incapable of explaining the formation of optically active lignans, which clearly indicated that some other biochemical mechanism must be operative *in vivo* for controlling both the regiochemistry and stereochemistry of free-radical coupling. Lignans are a group of plant metabolites that are biochemically related to the lignins. They are normally found as

Figure 1



(a) The three monolignols involved in lignin biosynthesis. **(b)** The random-coupling hypothesis for monolignol dehydrodimerization during softwood lignin biosynthesis. (For simplicity, the possible involvement of monolignol glucosides in transport into the cell wall is not included.) Additionally, the numerical structural designation refers to both (+)- and (–)-antipodes. **(c)** Chiral HPLC analysis of [9,9′-^{3}H]-pinoresinol **5** obtained following incubation as in (b), but with recombinant dirigent protein. As can be seen, essentially only (+)-antipode **5a** is formed (**5a** = (+)- antipode and **5b** = (–)-antipode).

monolignol-derived dimers, although higher oligomers can also exist, and their roles in plants appear to be primarily defense related. In a marked departure from what would be anticipated from random coupling, an insoluble *Forsythia suspensa* plant stem residue was discovered that catalyzed the stereoselective dehydrogenative coupling of two molecules of coniferyl alcohol **2** *in vitro* to give a product mixture in which the lignan (+)-pinoresinol **5a** predominated over the racemic lignan products **4**–**6** [5]. This stereoselectivity in phenoxy radical coupling differed from all previous chemical and enzymatic bimolecular phenoxy radical coupling reactions *in vitro*, which had lacked regiospecific and stereospecific control (Figure 1b). A ~78 kDa protein, responsible for the observed stereoselectivity of the phenoxy radical coupling process [5], was purified to apparent homogeneity. The name 'dirigent' protein (from the Latin: *dirigere*, to guide or align) was coined for this new type of protein. Apparently, the protein has no catalytically active (oxidative) center, and serves to bind and orient the monolignol-derived free-radical species, generated by one-electron oxidases, to ensure stereoselectivity in their coupling. A nonspecific oxidase (such as laccase) is therefore required only in the generation of the free-radical species from the monolignols, which, in the absence of the dirigent protein would give random coupling (Figure 1b). Significantly, neither the radicals of *p*-coumaryl **1** nor those of sinapyl **3** alcohols that differ only in the extent of aromatic methoxylation serve as substrates for the (+)-pinoresinol dirigent protein.

Developing flax (*Linum usitatissimum*) seed also appears to contain an analogous dirigent protein that stereospecifically produces only (–)-pinoresinol **5b** from *E*-coniferyl alcohol **2** (J.D. Ford and N.G.L., unpublished observations). This finding, together with the known existence of various optically active lignans linked through atomic centers other than 8–8′, indicates that a family of dirigent proteins exists in which each member stipulates a distinct coupling mode. The purpose of the present investigation was to establish whether or not such a protein family exists and, if it does, to consider what the overall significance of members of this family might be to both lignan and lignin formation.

The present study focused on three goals: obtaining the gene encoding the dirigent protein and establishing its authenticity via functional expression; evaluating its level of uniqueness by sequence homology comparisons and determining possible evolutionary relationships with other polypeptides; and identifying any potential connections between lignan and lignin formation, as both can originate from the same monolignol precursors. As described below,

dirigent protein genes have now been cloned and show no sequence homology to any other protein of known function. When expressed in a heterologous system, they are able to confer regiochemical and stereochemical control on monolignol coupling. In addition, the tissue-specific and subcellular localization of dirigent protein epitopes, as determined by immunolocalization studies, coincides with lignification in developing tissues.

Results and discussion

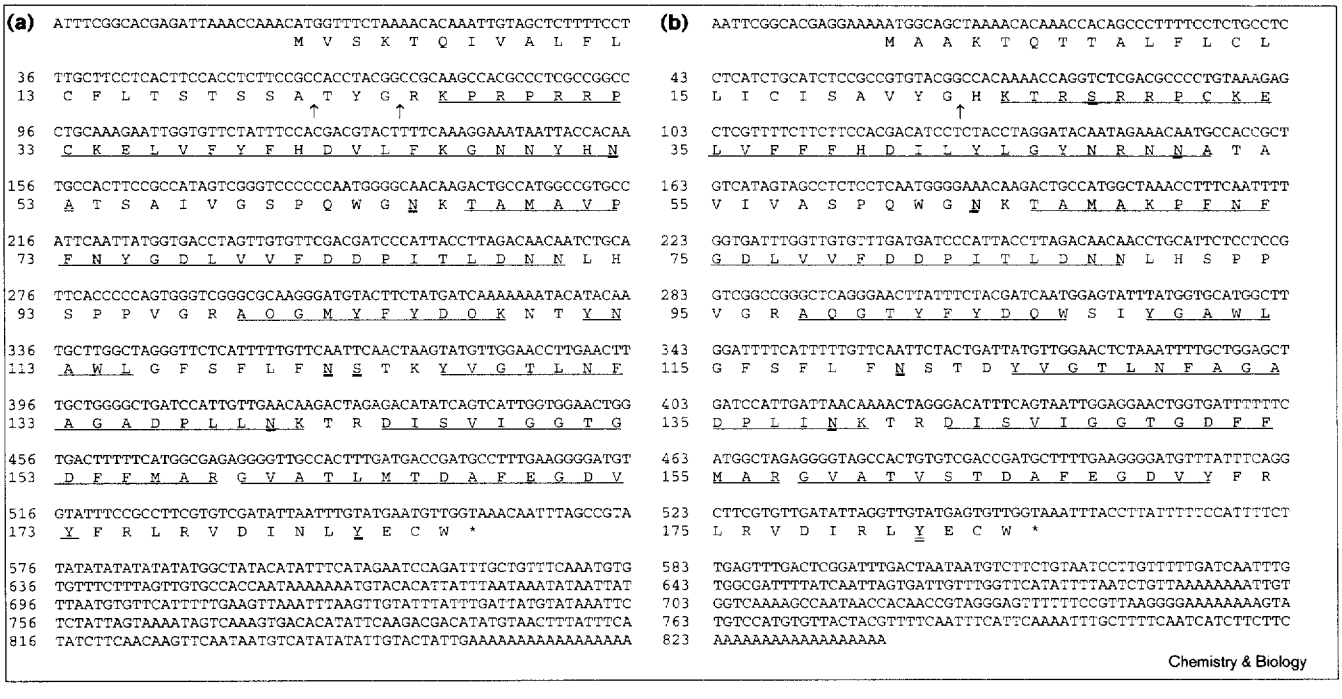
The dirigent protein results from a gene encoding an 18 kDa polypeptide

Analysis of the purified 78 kDa dirigent protein *Forsythia intermedia*, using sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE), indicated the protein has a subunit molecular weight of ~26 kDa [5], suggesting that the native protein is a trimer. From Concanavalin A affinity chromatography and carbohydrate analysis, the dirigent protein appears to be a glycoprotein with a range of pI values between 8 and 8.5 arising from several forms (separated by isoelectric focusing), all of which had identical amino termini [5].

A polymerase chain reaction (PCR) strategy was used to obtain a portion of the cDNA encoding the dirigent protein from a *F. intermedia* stem cDNA library [6], employing degenerate primers designed from the amino terminus and several internal amino-acid microsequences. Using the resulting amplification product (~370 base pairs) of one of these primer pairs as a probe, two distinct but very similar putative pinoresinol synthesizing dirigent protein cDNAs, *psd-Fi1* and *psd-Fi2* (see Figure 2a and 2b, respectively), were obtained. Subsequent cDNA sequence analyses of each identified possible secretory signal peptides, potential *N*-glycosylation sites and serine and tyrosine phosphorylation sites in both proteins [7,8]. Significantly, each cDNA, although encoding all seven of the sequenced peptide fragments from the mature native protein, only encoded a protein of ~18 kDa, after removal of the secretory signal sequences. It was, therefore, tentatively concluded that the differences in protein subunit size (26 versus 18 kDa) were due to glycosylation.

It was instructive to establish whether or not the putative recombinant 18 kDa dirigent protein could be functionally expressed, to verify the authenticity of the cloned cDNAs. A eukaryotic system (*Spodoptera frugiperda*/baculovirus) [9] was chosen for heterologous expression purposes, because it can produce recombinant proteins in the glycosylated form. Accordingly, with *psd-Fi1* under the control of the

Figure 2

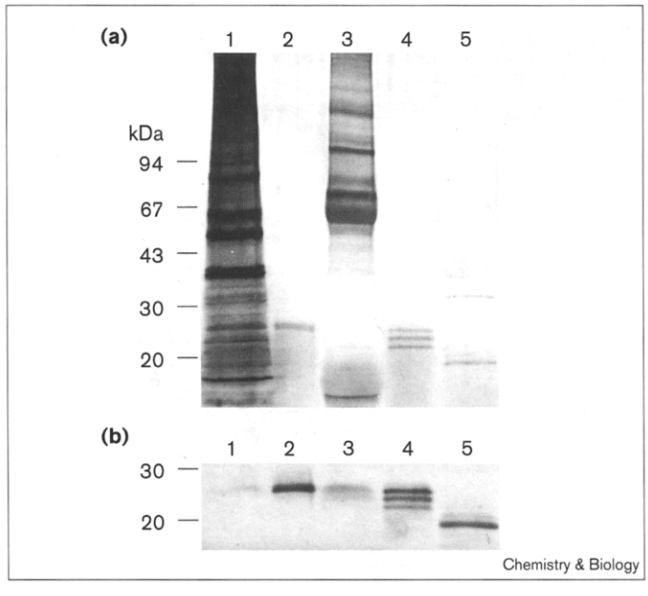


Complete sequence of **(a)** *Forsythia intermedia* dirigent protein cDNA psd-Fi1 and **(b)** psd-Fi2. Peptide sequences determined by amino-terminal and internal microsequencing are underlined. The signal peptide cleavage sites (before Arg25 and His24 in PSD-Fi1 and PSD-Fi2, respectively, in the native proteins, and before Thr21 in the recombinant PSD-Fi1 protein) are indicated by an arrow. Potential *N*-glycosylation sites (Asn52, Asn65, Asn122 and Asn140 in PSD-Fi1; Asn51, Asn64, Asn121 and Asn139 in PSD-Fi2), and serine (Ser123 in PSD-Fi1; Ser28 in PSD-Fi2) and tyrosine (Tyr183 in PSD-Fi1; Tyr182 in PSD-Fi2) phosphorylation sites are indicated by double underlining. The stop codons are indicated by an asterisk.

Autographa californica polyhedrin promoter, log-phase *Spodoptera* Sf9 cells were infected with a high titer AcMNPV-BB/*psd-Fi1* viral stock and SDS–PAGE analyses showed that a protein with the expected ~26 kDa subunit size was secreted into the culture medium (Figure 3a, lanes 3,4).

Unlike its native counterpart, which appears as a single band on SDS–PAGE gels (Figure 3a, lane 2), however, the (partially) purified recombinant PSD-Fi1 dirigent protein expressed in *Spodoptera* Sf9 cells appeared as three major bands on the gel (Figure 3a, lane 4). These proteins had at least ten identical residues at their amino termini, but they ranged in size from ~22 kDa to ~26 kDa, suggesting that PSD-Fi1 could be differentially glycosylated by the insect culture cells. Subsequent western blot analyses (Figure 3b) also revealed that polyclonal antibodies raised against heterologously expressed dirigent protein cross-reacted not only with the multiple recombinant PSD-Fi1 protein bands (expressed in *Spodoptera* Sf9 cells, Figure 3b, lanes 3 and 4), but also with the native *F. intermedia* dirigent protein (Figure 3b, lanes 1 and 2). In the latter case, the cross-reaction appears to be highly specific for the native dirigent protein, because only a single band was observed (~26 kDa) in the total stem protein extract applied to an SDS–PAGE gel (Figure 3a,b, lane 1). Deglycosylation of the purified recombinant protein (the multiple bands observed in Figure 3a, lane 4), followed by SDS–PAGE analysis, gave a single polypeptide band at ~18–19 kDa (Figure 3a, lane 5), in accordance with the size predicted from the cDNA for the nonglycosylated form of the secreted protein. (The two faint bands at 32–40 kDa (Figure 3a, lane 5) are the enzymes used to deglycosylate the heterologously expressed dirigent protein.)

Figure 3



(a) SDS–PAGE and **(b)** western blot analysis of native *F. intermedia* (lanes 1 and 2) and recombinant (*Spodoptera*/baculovirus expressed, lanes 3–5) dirigent protein. Lanes 1 (100 μg) and 3 (25 μg) are crude total cellular protein extracts. Lanes 2 and 4 (0.4 μg each) are purified proteins. Lane 5 contains 0.4 μg of the deglycosylated form of the recombinant protein. A single band is observed in the native protein extract, and at least three bands are observed for the recombinant protein. These are converted to a band of ~18 kDa by deglycosylation (lane 5).

Final proof that the recombinant dirigent protein had been expressed came when the heterologously expressed protein was purified to apparent homogeneity and assayed for functional capability. As can be seen in Figure 1c, the protein demonstrated the ability to confer regioselectivity and stereoselectivity upon bimolecular phenoxy radical coupling of *E*-coniferyl alcohol **2**, with the same substrate specificity as the native protein, provided that oxidative capacity (laccase) was supplied. The glycosylated recombinant dirigent protein from the *Spodoptera*/baculovirus system, therefore, was capable of controlling dehydrogenative monolignol coupling, demonstrating that the *Forsythia* dirigent protein had indeed been cloned.

Interestingly, the amino terminus of the recombinant dirigent protein secreted by the *Spodoptera* Sf9 cells is three amino-acid residues longer than the native (*Forsythia*) protein (Figure 2a), which was exactly as predicted using the sequence analysis software [7,8], indicating differences in signal-sequence cleavage sites between the animal and plant systems.

As plants can contain a series of lignans with differing skeleta, it was of interest to screen for other dirigent protein analogs or homologs in various species known to contain optically active lignans. Two gymnosperms, *Thuja plicata* (western red cedar) and *Tsuga heterophylla* (western hemlock), were chosen first for examination because they have high levels of pinoresinol-derived lignans in their tissues, especially in their heartwoods [10,11]. Using the same method as with *Forsythia* [6], several unique dirigent-protein-like cDNAs were found: two in western hemlock (*psd-Th1* and *psd-Th2*) and eight in western red cedar (*psd-Tp1* to *psd-Tp8*). The presence of dirigent-protein-like genes in other species was also determined using a PCR-guided strategy on the DNA isolated from several plant species [12] (Table 1), yielding sequences of about 360–370 base pairs in length for each species — the size expected from the *psd-Fi1* sequence, assuming no introns. The sequences from two of these species, *Populus tremuloides* and *Fraxinus mandschurica*, were chosen because they are known to possess pinoresinol-derived lignans and represent two divergent angiosperm families, the Salicaceae (Salicales) and Oleaceae (Gentianales), respectively. The cDNAs were subcloned and sequenced, revealing two distinct genes for both *F. mandschurica* (*psd-Fm1* and *psd-Fm2*) and *P. tremuloides* (*psd-Pop1* and *psd-Pop2*).

Table 1

Species evaluated using PCR and Southern blotting for the presence of dirigent protein genes.

Species name	Common name	Family name	Lignans present
Arctium lappa	Burdock	Asteraceae	Yes
Eucommia ulmoides	Eucommia	Eucommiaceae	Yes
Forsythia intermedia	Forsythia	Oleaceae	Yes
Forsythia suspensa	Forsythia	Oleaceae	Yes
Forsythia viridissima	Forsythia	Oleaceae	Yes
Fraxinus mandschurica	Manchurian ash	Oleaceae	Yes
Syringa vulgaris	Lilac	Oleaceae	Yes
Populus tremuloides	Quaking aspen	Salicaceae	Yes
Linum usitatissimum	Flax	Linaceae	Yes
Lupinus alba	White lupine	Fabaceae	?
Medicago sativa	Alfalfa	Fabaceae	?
Nicotiana tabacum	Tobacco	Solanaceae	Yes
Liriodendron tulipifera	Yellow poplar	Magnoliaceae	Yes
Sesamum indicum	Sesame	Pedaliaceae	Yes
Abies grandis	Grand fir	Pinaceae	Yes
Pinus taeda	Loblolly pine	Pinaceae	Yes
Pseudotsuga menziesii	Douglas fir	Pinaceae	Yes
Taxus brevifolia	Pacific yew	Taxaceae	Yes
Thuja plicata	Western red cedar	Cupressaceae	Yes
Tsuga heterophylla	Western hemlock	Cupressaceae	Yes

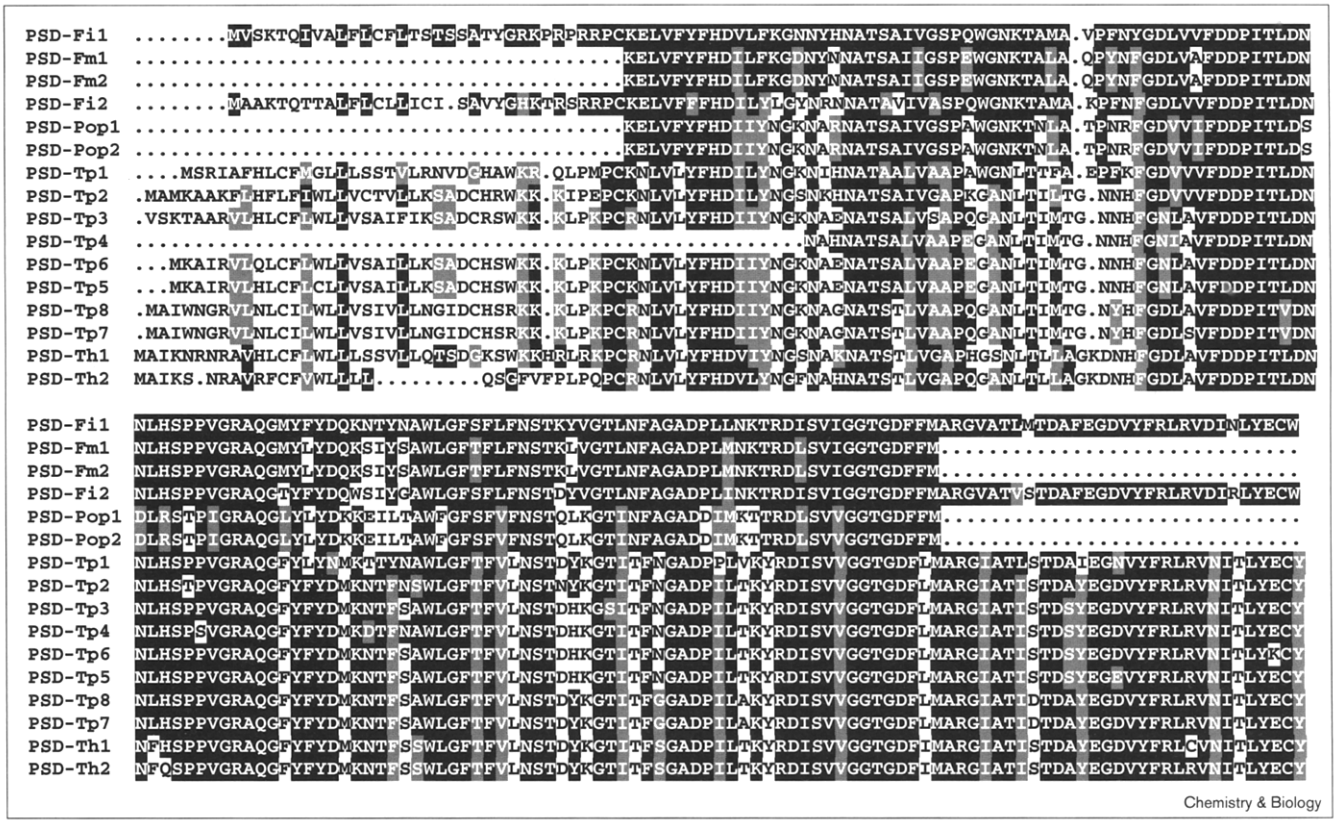
As is evident from Figure 4, the putative dirigent protein clones from these very diverse plant species have a very high degree of similarity to each other. Indeed, similarity/identity of full-length sequences range from 87.6%/81.6% for PSD-Fi2 compared with PSD-Fi1 to 78.0%/62.9% for PSD-Th1 compared with PSD-Fi1. With signal peptides removed, these values increase to 90.1%/84.6% for PSD-Fi2 compared with PSD-Fi1 and 76.5%/66.7% for PSD-Th1 compared with PSD-Fi1. The rest of the dirigent protein clones fall within this range of sequence similarity/identity. Except for the amino termini of the proteins that code for the cleaved signal peptides, most of the amino-acid substitutions between the various dirigent protein clones are conservative (i.e. charged residue for charged residue, hydrophobic for hydrophobic, and so on), indicating that all of these proteins probably perform very similar, if not identical, functions in their respective plants. Interestingly, two pairs of sequences, PSD-Fm1/PSD-Fm2 and PSD-Pop1/PSD-Pop2, showed apparent identity within the pair at the amino-acid level. At the nucleotide level, however, these sequence pairs (alleles) each displayed sequence identities of only about 90%–92%, with all nucleotide changes leading to no change in the amino-acid sequence. Southern blotting of DNA [13] from the species listed in Table 1, using *psd-Fi1* as a probe, verified further the presence of dirigent proteins (or very similar proteins) throughout the plant kingdom because all of these species appear to contain sequences homologous to the *Forsythia* dirigent protein (data not shown). Such conservation, involving very diverse representatives of angiosperms and gymnosperms, indicates that dirigent protein genes might be ubiquitous throughout vascular plants, exactly as proposed for another protein in the lignan biosynthetic pathway, (+)-pinoresinol/(+)-lariciresinol reductase [6,14]. Dirigent proteins therefore appear to represent a very widely distributed, but previously unknown, class of protein.

Comparisons of dirigent protein gene sequence homologies

With the entire *Forsythia* dirigent protein cDNA sequence in hand, a BLAST/BLAST-BEAUTY search [15,16] revealed no homology to any protein of known function, confirming that the 18 kDa proteins comprise a unique class. There was, however, significant identity (77%) to a gene of unknown function — the so-called pea disease resistance response gene 206-d [17] (Genbank accession U11716). Interestingly, 206-d is induced in conjunction with isoflavone reductase, a branch-point enzyme in the isoflavonoid phytoalexin (plant defense) biosynthetic pathway. Lower levels of homology were also observed with portions of three other plant genes, all of which encode proteins of unknown function, namely: a hypothetical protein (Genbank accession SF000657, 54% similarity over 82% of the sequence) from an *Arabidopsis thaliana* random BAC clone; an unknown protein (Genbank accession U32427, 47% similarity over 56% of the sequence) from wheat (*Triticum aestivum*), which, interestingly, is induced by benzothiadiazole — an inducer of systemic acquired resistance in wheat [18]; and a putative 32.7 kDa jasmonate-induced gene (Genbank accession U43497, 50% similarity over 35% of the sequence) from barley (*Hordeum vulgare*).

It can be reasonably concluded, therefore, that the only reported proteins or genes with significant homology to the *F. intermedia* dirigent protein or corresponding cDNA are the dirigent protein homologs that have been identified in other species here, and the plant genes (of unknown

Figure 4



Sequence alignment of the cloned dirigent proteins. Identity, similarity/conservative changes and differences, when compared with the PSD-Fi1 sequence, are indicated by lettering on black, grey and white backgrounds, respectively (for abbreviation descriptions, see text).

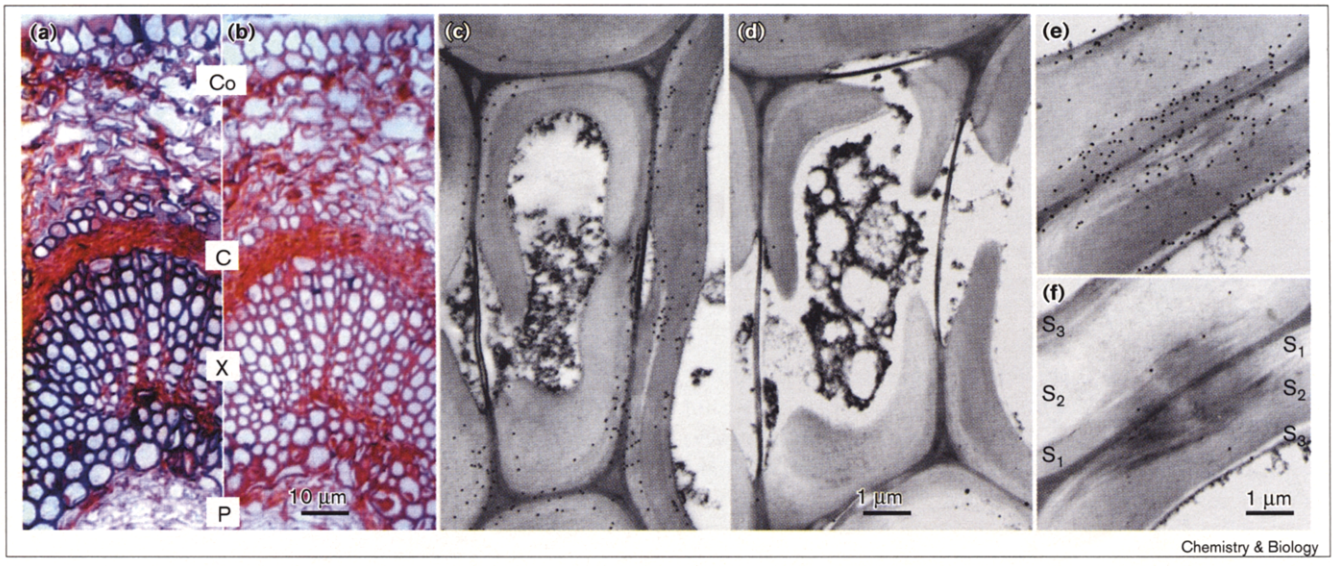
function) from pea, *Arabidopsis*, wheat and barley. As comparisons of sequence homology did not reveal any correlation with other proteins, it is premature to speculate on the origin of these highly unusual proteins and on the physiological function of their progenitors.

Biochemical relationships between lignin and lignan formation *in vivo*

With the discovery of the dirigent proteins, a re-evaluation of the fundamental biochemical basis of *in vivo* control of dehydrogenative monolignol coupling was warranted. Is lignin biopolymer assembly as strictly controlled as the regiospecific and stereospecific coupling of monolignol radicals during lignan formation? A full consideration of the evidence presently available indicates that it would be very surprising indeed if lignification were not under absolute biochemical control.

Indeed, any convincing model for lignification must account for everything that is known about native lignin configurations: the predominance of 8-*O*-4′ linkages *in vivo*; the reputed lack of any optical activity in isolated lignin-derived preparations (and their partially depolymerized fragments); and the characteristic variations in lignin monomer and linkage composition within cell walls among different tissues, organs and cells, leading, in some cases, to the formation of alternative microdomains of distinct lignin types [19]. The question then arises as to how the individual plant cell can determine so precisely both the content of monomer residues and the extent of lignification in regions ostensibly beyond the direct control of the living part of the cell (across the plasma membrane and in the far reaches of the cell wall or middle lamella). In this regard, a very important clue emerged recently concerning the nature of lignification processes: the observation that growing lignin microdomains, at sites far removed from the plasma membrane in the maturing cell wall, expand uniformly with apparent constant density until coalescing. Donaldson [20] reported that the initiation of lignin biosynthesis in *Pinus radiata* (radiata pine) occurs near or adjacent to the original cell plate in the region where the middle lamella between cells ultimately develops. Further initiation sites for lignin biosynthesis were also observed throughout the primary and secondary cell wall regions. It would, therefore, appear that the key to how lignin biosynthesis is actually controlled lies in the vicinity of these initiation sites.

Figure 5



(a,b) Light microscopy and **(c–f)** transmission electron microscopy immuno-gold labeling of *F. intermedia* young stem using the polyclonal antiserum raised against the recombinant dirigent protein. (a,c,e), Labeling with the immune-serum. (b,d,f), Control with the pre-immune serum on same zones of tissues on serial sections. Co, cortex; C, cambium; X, xylem; P, pith cavity; S1, S2, S3, sub-layers of the secondary wall from the outer part (near the middle lamella) to the inner part (near the lumen).

It is, therefore, reasonable to propose that the initiation of lignin macromolecular assembly processes could be under the control of arrays of monolignol or oligolignol radical binding sites (determining both monomer and linkage type), specifically pre-localized within the cell wall during primary-wall formation. Such specific spatial and temporal localization in the developing cell wall has been observed for several other cell-wall proteins [21–23]. Through the same basic mechanism as that viewed responsible for regiospecific and stereospecific monolignol radical coupling during lignan biosynthesis, these dirigent protein arrays would stipulate the primary structure of the progenitorial lignin chain being formed, perhaps via an endwise polymerization process [24]. There is a key conceptual difference, however, between lignin and lignan formation: the dirigent proteins are envisaged to form an array that stipulates macromolecular lignin configuration, unlike the dirigent proteins involved in dimer (lignan) formation. Accordingly, once the progenitorial lignin chain is formed, it would not be able to diffuse away from the array of dirigent sites encoding its primary structure; instead, it would be replicated through a direct template polymerization process by the dehydrogenative incorporation of the corresponding monolignols. Indeed, preliminary evidence for the possibility that macromolecular lignin replication can occur through such a mechanism has been obtained *in vitro* [25].

Supporting experimental evidence for the hypothesis of dirigent protein arrays also comes from immunocytochemical localization results (Figure 5a–f), as well as from tissue-printing, northern analyses and *in situ* hybridization (data not shown). Thus, for immunocytochemical localization [19,26], polyclonal antibodies raised against the *F. intermedia* recombinant dirigent protein were specifically employed to recognize not only dirigent proteins in lignan formation but also the epitopes of presumed dirigent protein arrays involved in lignin biosynthesis. Using light microscopy, tissue-specific localization of dirigent protein (arrays) is clearly observed; only the xylem (ray cells, fibers and vessels) and, to a lesser extent, the fibers surrounding the phloem were immuno labeled (Figure 5a), as compared with the pre-immune serum control (Figure 5b), whereas the nonlignifying cambium, cortex and parenchyma cells bordering the pith cavity were not significantly labeled with the antiserum. In the young stem samples examined, specific labeling was therefore clearly associated only with the lignifying fiber and vessel cell walls, as well as with the ray parenchyma, which are viewed to be involved in lignan deposition (Figure 5a) [27]. Enhanced magnification of the cell walls and better resolution of the labeling were next obtained using transmission electron microscopy (TEM), which showed that the labeling of the xylem cell walls was primarily restricted to the outer part of the wall (S1 sub-layer of the secondary wall), and to a much lesser extent, to the internal S3 sub-layer (Figure 5c,e). The labeling within the outer part of the wall is consistent with the localization of the initial stages of lignin deposition as described earlier by Donaldson [20], and therefore strongly implicates the involvement of dirigent protein sites (perhaps

existing in arrays) in lignin biosynthesis. In contrast, lignan deposition is unlikely to occur at these subcellular sites because post-coupling modification of the initially formed lignans involves cytosolic proteins [6,14].

This working hypothesis and the results obtained thus far, taken together, are consistent with the control of all known aspects of lignification processes *in vivo*. Contiguous arrays of lignol radical binding sites, in different relative orientations, could establish the sequence of inter-unit linkages characterizing particular macromolecular lignin primary structures in the cell-wall domains of each specific cell type. Indeed, such a mechanism could also account for the reputed lack of optical activity in native lignins, if individual dirigent radical coupling sites were to encode mirror images of the same inter-unit linkages, or if template polymerization automatically engenders the formation of enantiomeric structures.

In summary, there is now little doubt that the biochemical control of lignification is implemented at the levels of both monomer deposition and linkage specification, given that differing monomer compositions and frequencies of inter-unit linkages are observed in distinct regions of plant cell walls among the various cell types of different tissues and species. Accordingly, the distinct possibility exists that the genes encoding dirigent proteins and/or arrays have a direct role in this regulatory process. Such control has been established for lignan biosynthesis, and it would be surprising if a corresponding process did not operate in lignification as well, given that dirigent lignol radical binding sites have presumably evolved for the express purpose of preventing random coupling. The enigma of monolignol coupling is therefore at last yielding to biochemical scrutiny, revealing yet again the remarkable level of control that nature exercises over even the most complicated biological processes.

Significance

Monolignol free-radical coupling, the mechanism whereby polymeric lignins and (oligomeric) lignans are produced in plants, is not a random process. Dirigent proteins have been identified, first biochemically and now at the gene level, that are responsible for conferring control over this process. We have identified cDNAs encoding dirigent proteins from several species, both angiosperms and gymnosperms, that have no sequence homology with any other protein of known function and, accordingly, their evolutionary ancestor is unknown. When expressed in a heterologous system the recombinant protein is able to control both the regiochemical and stereochemical outcome of monolignol coupling, leading to the formation of the lignan (+)-pinoresinol. These proteins are produced in the plant in developing vascular tissues, especially in the xylem cell wall, indicating a role in lignification. Furthermore, we propose that the formation of lignins might be under the control of a contiguous array of dirigent sites, either on adjoining protein molecules or along a single polypeptide chain. In this new hypothesis, the dirigent array confers a primary structure on the progenitorial lignin macromolecules, perhaps during an endwise polymerization process, with subsequent replication being effected by a template polymerization mechanism.

Materials and methods

Dirigent protein cDNA cloning strategy

The construction of a *F. intermedia* stem cDNA library has been reported previously [6]. Amino-acid amino terminal and several internal peptide microsequences were obtained from the purified *F. intermedia* dirigent protein according to the methodology described [6]: KPRPXR-XX*KELVFYFX*DILFKGXNYNXA (amino terminal; using single-letter amino-acid code), TA*MAVPFNY*GDL*VVFDDPIT*LDNN (internal sequence 1), YVGTL*NFAGADP*LLXK (internal sequence 2), DISVIGGTGDFFMAR (internal sequence 3), GVATLMTDAFEGDXY (internal sequence 4), AQGMYFYDQK (internal sequence 5), YNAWL (internal sequence 6). Degenerate oligonucleotide primers were designed from regions of the amino-acid sequences in italics: AARGARYTIGTITTYTAYTTY (amino terminal primer PSDNT1); TARTTRAAIGGIACIGCCAT (reverse primer PSDI1R, from internal sequence 1); GTIATIGGRTCRTCRAAIAC (reverse primer PSDI2R, from internal sequence 1); CCATAAAAAAATCICCIGT (reverse primer PSDI3R, from internal sequence 2). These primers were used for PCR amplifications of portions of the dirigent protein cDNA from the cDNA library under the conditions used previously [6]. The amplification probe (~ 370 bp) of the PSDNT1+PSDI3R primer pair was used as a probe to screen the *F. intermedia* cDNA library as described previously [6]. Two clones, psd-Fi1 and psd-Fi2, were obtained and sequenced (Figures 2a,2b).

Heterologous expression of dirigent protein

The full-length 1 kb cDNA, psd-Fi1, containing both the 5′ and 3′ untranslated regions, was excised from the pBlueScript SK(−)-derived (Stratagene) plasmid pPSD-Fi1 using the restriction endonucleases BamHI (Promega) and XhoI (Promega). This 1 kb fragment was directionally subcloned into the baculovirus transfer vector, pBlueBac4 (Invitrogen), producing the 6 kb construct pBB4/PSD-Fi1, which generates a nonfusion dirigent protein with translation being initiated at the dirigent protein cDNA start codon. The construct was then co-transfected with linearized Bac-N-Blue DNA (Invitrogen) into *Spodoptera frugiperda* Sf9 cells through cationic liposome mediated transfection to produce, by means of homologous recombination, the recombinant *Autographa californica* nuclear polyhedrosis viral (AcMNPV) DNA-Bac-N-Blue dirigent protein recombinant virus (AcMNPV-BB/PSD-Fi1); this was purified from plaques according to procedures described by Invitrogen. The final recombinant AcMNPV-BB/PSD-Fi1 contains the *psd-Fi1* gene under the control of the *Autographa californica* polyhedrin promoter as well as the essential sequences needed for replication of the recombinant virus. Next, log phase Sf9 cells were infected with a high titer AcMNPV-BB/PSD-Fi1 viral stock. Maximal heterologous dirigent protein production occurred by 48–70 h post-infection (yield: 1 to 1.5 mg pure dirigent protein/liter, not optimized).

Western blot analysis

Western blots were performed using rabbit polyclonal antibodies raised against recombinant *F. intermedia* dirigent protein. Horseradish peroxidase-conjugated goat anti-rabbit antibodies were used as a secondary marker and binding was visualized using 4-chloro-1-naphthol.

Deglycosylation

Purified recombinant dirigent protein (5 μg) was incubated with NANase II, O-glycosidase and PNGase F (Bio-Rad Laboratories) at 37°C, according to the manufacturer's instructions. After a 2 h incubation, the resulting deglycosylated dirigent protein was analyzed by SDS–PAGE (4–15% gradient gel, Bio-Rad Laboratories) with the proteins visualized by silver staining.

Recombinant dirigent protein purification

A culture (1l) of Sf9 cells infected with AcMNPV-BB/PSD-Fi1 was grown at 27°C in Grace's medium (Invitrogen). After 48 h, the medium was harvested, centrifuged (1200 × g, 20 min, 4°C) and the supernatant concentrated to 100 ml (Diaflo® Ultrafilter YM30 membrane, Amicon). The resulting concentrate was subjected to polyethylene glycol (PEG8000) precipitation (6–12% w/v). Proteins recovered by centrifugation (15,000 × g, 30 min, 4°C) were next resuspended in MES-NaOH buffer (40 ml, pH 5.0) and loaded onto a MonoS column (HR5/5, Pharmacia) for purification of the dirigent protein as described previously [5].

Assay for dirigent protein

Assays were performed as described previously [5] using 30 ng of purified *F. intermedia* laccase as an oxidase.

Genomic PCR reactions

Total genomic DNA (1 μg), isolated from the plant species listed in Table 1 [12], was used in PCR amplifications [6] with degenerate primers designed from two regions of the *Forsythia* (+)-pinoresinol dirigent protein sequences that are also highly conserved with the pea disease response gene 206-d, – residues 34–40 (of PSD-Fi1) for the forward primer (PSDH-NT: CCMTGYAAARAIYTIGTITT) and residues 151–156 for the reverse primer (PSDH-CT: CCAACAYT-CATASAAIIARAT). PCR reactions using both primers with genomic DNA gave bands of about 360–370 bp in length, the size expected based on the psd-Fi1 sequence assuming no introns, for all species in Table 1. Control reactions, of each primer alone with each DNA sample or of both primers and water (no DNA), gave no bands under these conditions. The bands from two of these species, *Fraxinus mandschurica* and *Populus tremuloides*, were gel purified [13] and subcloned into a T-vector plasmid (pT7-Blue T-vector, Novagen, Inc.). Several individual clones were sequenced and two distinct genes were identified for both *F. mandschurica* and *P. tremuloides* (see the Results and Discussion section).

Immunocytochemical localization of dirigent protein in F. intermedia *young stem*

Using the polyclonal antiserum raised against the recombinant dirigent protein, post-embedding indirect immunogold labeling experiments were performed on cross-sections of *F. intermedia* young stems (developing internode). The overall fixation, embedding and labeling protocols were essentially as those described previously [26] except for the following modifications. For the light microscopy experiments, semi-thin sections (800 nm) were carried on plastic rings, and the secondary marker employed was a goat anti-rabbit antibody conjugated to 1 nm gold particles (Amersham, Life Science Inc.), with subsequent silver enhancement performed according to Amersham's instructions. The resulting sections were then transferred to gelatin coated slides and counter-stained with safranin O. For TEM labeling experiments were carried out on ultra-thin sections (80 nm) supported on uncoated 300 mesh nickel grids, in order to have both sides of the sections accessible to antibodies. Under these conditions, the secondary marker employed was protein A conjugated to 10 nm gold particles (Amersham, Life Science Inc.) with subsequent silver enhancement [19] as described previously.

Acknowledgements

The authors thank the United States Department of Energy DE-FG03-97ER20259, the National Science Foundation MCB9631980, the National Aeronautics and Space Administration NAG2-1198, the DOE-NSF-USDA Plant Biochemistry Research and Training Center, the Lewis B. and Dorothy Cullman and G. Thomas Hargrove Center for Land Plant Adaptation Studies and Minnesota Agricultural Experiment Station Project 43-68 (Hatch Funds) for generous support of these studies.

References

1. Freudenberg, K. (1959). Biosynthesis and constitution of lignin. *Nature* **183**, 1152-1155.
2. Freudenberg, K. (1968). The constitution and biosynthesis of lignin. In *Constitution and Biosynthesis of Lignin.* (Freudenberg, K. & Neish, A.C., eds) pp. 47-122, Springer-Verlag, New York, NY.
3. Higuchi, T. (1957). Biochemical studies of lignin formation I. *Physiol. Plant.* **10**, 356-372.
4. Sterjiades, R., Dean, J.F.D., Gamble, G., Himmelsbach, D.S. & Eriksson, K.-E.L. (1993). Extracellular laccases and peroxidases from sycamore maple (*Acer pseudoplatanus*) cell suspension cultures. *Planta* **190**, 75-87.
5. Davin, L.B., *et al.*, &. Lewis, N.G. (1997). Stereoselective bimolecular phenoxy radical coupling by an auxiliary (dirigent) protein without an active center. *Science* **275**, 362-366.
6. Dinkova-Kostova, A.T., Gang, D.R., Davin, L.B., Bedgar, D.L., Chu, A. & Lewis, N.G. (1996). (+)-Pinoresinol/(+)-lariciresinol reductase from *Forsythia intermedia*: protein purification, cDNA cloning, heterologous expression and comparison to isoflavone reductase. *J. Biol. Chem.* **271**, 29473-29482.
7. Appel, R.D., Bairoch, A. & Hochstrasser, D.F. (1994). A new generation of information retrieval tools for biologists: the example of the ExPASy WWW server. *Trends Biochem. Sci.* **19**, 258-260.
8. Program Manual for the Wisconsin Package (1997). Genetics Computer Group, Madison, WI, USA.
9. O'Reilly, D.R., Miller, L.K. & Luckow, V.A. (1994). *Baculovirus Expression Vectors: A Laboratory Manual.* Oxford University Press, New York, N.Y.
10. Swan, E.P., Jiang, K.S. & Gardner, J.A.F. (1969). The lignans of *Thuja plicata* and the sapwood-heartwood transformation. *Phytochemistry* **8**, 345-351.
11. Krahmer, R.L., Hemingway, R.W. & Hillis, W.E. (1970). The cellular distribution of lignans in *Tsuga heterophylla* wood. *Wood. Sci. Technol.* **4**, 122-139.
12. Doyle, J.J. & Doyle, J.L. (1992). Isolation of plant DNA from fresh tissue. *Focus* **12**, 13-15.
13. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1994). *Molecular Cloning: A Laboratory Manual, 3rd ed.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
14. Gang, D.R., Dinkova-Kostova, A.T., Davin, L.B. & Lewis, N.G. (1997). Phylogenetic links in plant defense systems: lignans, isoflavonoids and their reductases. *ACS Symp. Ser.* **658**, 58-89.
15. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410.
16. Worley, K.C., Wiese, B.A. & Smith, R.F. (1995). BEAUTY: An enhanced BLAST-based search tool that integrates multiple biological information resources into sequence similarity search results. *Genome Res.* **5**, 173-184.
17. Fristensky, B., Horovitz, D. & Hadwiger, L.A. (1988). cDNA sequences for pea disease resistance response genes. *Plant Mol. Biol.* **11**, 713-715.
18. Görlach, J., *et al.*, &. Ryals, J. (1996). Benzothiadiazole, a novel class of inducers of systemic acquired resistance, activates gene expression and disease resistance in wheat. *Plant Cell* **8**, 629-643.
19. Burlat, V., Ambert, K., Ruel, K. & Joseleau, J.P. (1997). Relationship between the nature of lignin and the morphology of degradation performed by white-rot fungi. *Plant Physiol. Biochem.* **35**, 645-654.
20. Donaldson, L.A. (1994). Mechanical constraints on lignin deposition during lignification. *Wood Sci. Technol.* **28**, 111-118.
21. Müsel, G., *et al.*, &. Schöpfer, P. (1997). Structure and distribution of lignin in primary and secondary walls of maize coleoptiles analyzed by chemical and immunological probes. *Planta* **201**, 146-159.
22. Ryser, U., *et al.*, &. Keller, B. (1997). Structural cell-wall proteins in protoxylem development: evidence for a repair process mediated by a glycine-rich protein. *Plant J.* **12**, 97-111.
23. Marcus, A., Greenberg, J. & Averyhart-Fullard, V. (1991). Repetitive proline-rich proteins in the extracellular matrix of the plant cell. *Physiol. Plant.* **81**, 273-279.
24. Sarkanen, K.V. (1971). Precursors and their polymerization. In *Lignins –Occurrence, Formation, Structure and Reactions.* (Sarkanen, K.V. & Ludwig, C.H., eds) pp. 95-163, Wiley Interscience, New York.
25. Guan, S.-Y., Mlynár, J. & Sarkanen, S. (1997). Dehydrogenative polymerization of coniferyl alcohol on macromolecular lignin templates. *Phytochemistry* **45**, 911-918.
26. Klauer, S.F., Franceschi, V.R. & Ku, M.S.B. (1991). Protein compositions of mesophyll and paraveinal mesophyll of soybean leaves at various developmental stages. *Plant Physiol.* **97**, 1306-1316.
27. Gang, D.R., Fujita, M., Davin, L.B. & Lewis, N.G. (1998). The abnormal lignins: mapping heartwood formation through the lignan biosynthetic pathway. *ACS Symp. Ser.* **697**, 389-421.