ABSTRACT

WANG, JACK PENG-YU. Towards a Systems Approach for Lignin Biosynthesis in *Populus trichocarpa*: Functional Redundancy of the Two 5-Hydroxylases and Kinetic Modeling of the Monolignol Biosynthetic Pathway. (Under the direction of Prof. Vincent Chiang and Ronald Sederoff).

Lignin is one of the most abundant component of biomass in the terrestrial biosphere, and constitutes an important role in the adaptation of plants to the various biotic and abiotic stresses of terrestrialization. Lignin is also the primary recalcitrant in the utilization of cell wall polysaccharides for pulp and energy. A systems approach in the study of lignin biosynthesis would increase our knowledge of the assembly of plant cell walls, and at the same time, help to improve the conversion of lignocellulosic biomass.

In this study, we investigated the functional redundancy of the two 5-hydroxylases in monolignol biosynthesis of *Populus trichocarpa*, through LC-MS/MS based protein quantification and metabolic flux analysis of the recombinantly expressed P450 monooxygenases.

We then performed the biochemical characterization of major monolignol pathway enzymes and deduced their substrate specificities and kinetic parameters. Using this information, a kinetic model of the monolignol biosynthetic pathway of *P. trichocarpa* was assembled. Simulated perturbation studies and sensitivity analyses were performed to investigate the functional roles of numerous monolignol pathway enzymes. Both confirmation and inconsistencies between the model outputs and our knowledge of monolignol biosynthesis were observed. The results presented in this study will be the foundation for discovering novel components and mechanisms that govern lignin biosynthesis, content and composition.

Towards a Systems Approach for Lignin Biosynthesis in *Populus trichocarpa*: Functional Redundancy of the Two 5-Hydroxylases and Kinetic Modeling of the Monolignol Biosynthetic Pathway

by
Jack Peng-Yu Wang

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Forestry and Environmental Resources

Raleigh, North Carolina

2012

APPROVED BY:

_____	_____
Vincent L. Chiang Co-Chair of Advisory Committee	Ronald Sederoff Co-Chair of Advisory Committee
_____	_____
Cranos Williams	Joel Ducoste
_____	
Hou-min Chang	

BIOGRAPHY

My name is Jack, and I like lignin. It is my goal to research and learn the mechanisms underlying how lignin is biosynthesized.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 2

Chapter 3

Chapter 4

LIST OF FIGURES

Chapter 2

Chapter 3

Chapter 4

Chapter 1

Introduction

The Lignified Plant Cell Wall

The lignified cell walls of plants are the predominant biomass in most of our ecosystems, and affect many aspects of our daily lives. They are the source of our wood, pulp and paper, our linen, and some chemical feedstocks. The awareness of the lignified plant cell wall as a renewable resource and as a carbon sink in the global carbon cycle, has driven scientific inquiry into its structure and biosynthesis.

Plant cell walls are of intricate entities, composed of complex yet coordinated structures. Crucial aspects of the biology of plants and their colonization of terrestrial environments have resulted from the evolution and diversification of the lignified cell wall. The lignified secondary walls that surround plant cells are hydrophobic entities that are intimately associated with many areas of plant biology, from growth and development, breeding potential and agricultural utility to their responses to abiotic and biotic stress.

The plant cell walls are specialized organelles affecting critical physiological and morphological functions. Lignocellulose is the major component of the secondary wall. Lignin biosynthesis is initiated when cells cease elongation and adopt functional specialization. Lignocellulose is composed mostly of cellulose and hemicelluloses, impregnated with lignin, and minor quantities of proteins. In the secondary wall of woody

angiosperm, cellulose can account for 50% of the mass of the secondary wall, while hemicellulose and lignin occupy 20% and 25% respectively. During the formation of the secondary wall, cellulose microfibrils are deposited onto the outer surface of the plasma membrane in organized parallel sheets and bundles, cross-linked by a matrix of hemicellulose polymers. The crystallinity of the cellulose microfibers has been associated with many of the secondary walls physical properties including shape, growth and mechanical strength.

Owing to the exceptional mechanical, physical, chemical properties and abundant biomass of wood, research into the biosynthesis of lignocelluloses has been promoted for more than half a century. The biosynthesis of the lignified plant cell wall is a complex and dynamic process, requiring the coordination of well over 2000 different genes. The structure of lignocellulosics can be subdivided to 3 major components, cellulose, hemicellulose and lignin.

Cellulose microfibrils are composed of linear β-1, 4 linked glucan chains. The extent of hydrogen bonding inter-molecularly and intra-molecularly dictates the level of lateral aggregation and crystallinity. Studies of cellulose biosynthesis have been extensive, yet the underlying mechanisms remain largely unknown. *In vitro* reconstruction of cellulose synthase activities have not been accomplished to date. The difficulty arose from the complex interactions within a distinctive populations of proteins localized on the plasma membrane. These rosette terminal protein complexes (RTCs) are hexameric in structure, and synthesize cellulose through linear polymerization of UDP-D-glucose precursors from a steroid primer. 17 putative cellulose synthases were annotated from the genome of *P.*

trichocarpa, of which two classes of RTCs with distinctive localization have been characterized.

Hemicelluloses in the xylem of woody angiosperms are composed mainly of xylan and glucomannan, in approximate proportions of 85% and 15% respectively. Knowledge of the biosynthesis of hemicellulose is limited. In the genome of *P. trichocarpa*, 30 cellulose synthase like (*Csl*) genes have been identified and annotated. These genes are categorized into subfamilies *A, B, C, D, E* and *G* according to protein sequence similarity. The *CslA* subfamily catalyzes the synthesis of glucomannan, and the *CslC* subfamily has been associated with the synthesis of the xyloglucan back bone. The functions of *Csl B, D, E* & *G* remain to be elucidated .

Providing rigidity and stability to the secondary wall is a major function of lignin. Lignin is a complex phenolic polymer, and unique in the absence of order in its primary structure. The diversity in lignin structures is fundamental to the walls stability and resistance to decay or degredation. Lignin deposition usually occupies the later stages of secondary cell differentiation. The hydrophobic lignin polymer surrounds the cellulose microfibrils and hemicellulose matrices forming an insoluble wall surface, providing mechanical support and facilitating vertical water transport.

Lignin Structure and Composition

After cellulose, lignin is the most abundant organic polymer in the terrestrial biosphere. Lignin is a complex aromatic phenolic polymer that imparts rigidity, mechanical strength and resistance of plants to decay and diseases. Deposition of lignin occurs mainly in terminally differentiating cell walls of xylem tissues, providing them with the ability to withstand the force of gravity, mechanical stress, and the negative pressure generated by transpiration. The lignin polymer is highly resistant to both mechanical disruption and enzymatic degradation. This recalcitrance has driven scientific inquiry into its structure and biosynthesis, as the presence of lignin in animal feed and forage greatly reduces its nutritional quality, and removal of lignin from wood pulp in the process of paper production is costly and requires the use of harsh polluting chemicals. Lignin also limits the yield and increases the cost of generating fermentable sugars from bulk plant biomass for subsequent conversion to ethanol and other biofuels. The elucidation of lignin biosynthesis, structure, regulation, and function is therefore not only essential for an understanding of plant biology and evolution, but also has major implications for potential bioengineering strategies to improve feed, fiber and bioenergy crops.

Lignin is polymerized from three major subunits, (*p*-hydroxyphenyl, guaiacyl and syringyl monolignols), that are non-methoxylated, monomethoxylated (C-3) and dimethoxylated (C-3, C-5) respectively. The proportion of monolignol subunits in lignin differs significantly between angiosperms and gymnosperms. The lignin of gymnosperms is composed mostly of guaiacyl subunits with small amounts of *p*-hydroxyphenyl subunits. In contrast, the lignin of

woody dicotyledonous angiosperms consists of both guaiacyl and syringyl subunits. Monocotyledonous lignin is a mixture of all three. Owing to differences in the chemical properties of the three monomers, their relative abundance in a particular lignin polymer can significantly affect its overall structure and chemical properties.

Phenolic Glucosides

A significant fraction of phenolic compounds are found in gymnosperms and some angiosperms to be glucosylated on the phenolic hydroxyl group to form 4-O-β-D-glucosides, namely coniferin and syringin. (Freudenberg & Harkin, 1963, Terazawa et al., 1984). These monolignol glucosides are found in the vacuoles of differentiating xylem cells of conifers, and the root and leaf of angiosperms, leading to the hypothesis that the role of these glucosides are for monolignol storage rather than as a principle intermediate in the monolignol biosynthesis pathway. These glucosides have not yet been detected in the lignifying zones of plants, further supporting their role as a storage pool for monolignols. On the contrary, the role of monolignol glucosides as an indispensable intermediate in monolignol biosynthesis has also been suggested by the radiotracer experiments in poplar (Fukushima & Terashima, 1990) and rice (He & Terashima, 1989). When radio-labeled monolignol glucosides are administered to a growing stem of a plant, they are effectively incorporated into protolignin in specific morphological regions of cell walls by biochemically controlled mechanisms. These observations, coupled with the widespread

occurrence of coniferyl alcohol-glucosyl transferase suggests a general participation of glucosides in the biosynthesis of lignin.

Evolution of Syringyl Lignin in Angiosperms

Angiosperms represent one of the greatest terrestrial radiations. Their rapid diversification and ecological dominance during the late Cretaceous suggest adaptive advantage over other vascular plants (Soltis et al. 2008). A morphological feature restricted primarily to angiosperms is the innovation of syringyl lignin. This recent adaptation is retained in nearly all angiosperms, and in many species the syringyl moiety has become the most abundant monolignol unit in lignin. The characteristic abundance of syringyl monolignol in different species, together with the evidence of a differentially regulated temporal and spatial distribution of lignin components can both be taken as evidence that the incorporation of syringyl sub-units in lignin provides an evolutionary advantage. Syringyl lignin differs from guaiacyl by the presence of a 5' methoxyl group, which reduces the available coupling sites thus lowering the amount of crosslinking in lignin. The more linear syringyl lignin may allow for a more flexible and elastic lignin, reducing the susceptibility of trees to lodging. It is then likely that the metabolic mechanisms regulating the biosynthesis of syringyl monolignol are under a selection pressure, and increasing the syringyl components in lignin will positively impact the fitness of the species.

Lignification

Lignification entails monolignol biosynthesis, transportation to the cell wall, and polymerization. Biosynthesis of monolignols occurs in the cytoplasm. Ten enzyme families mediate the conversion of phenylalanine to monolignols. These lignin precursors are exported through cellular membranes to the cell wall where they are oxidized and polymerized to lignin.

Lignification displays considerable plasticity in incorporating different monomeric precursors. Besides the S, G and H monolignols, under certain circumstances other phenolic monomers are also incorporated into lignin. For example, in cinnamyl-alcohol dehydrogenase deficient mutants of pine and transgenic tobacco, significant amounts of hydroxycinnamaldehydes were incorporated into lignin. Similarly, transgenics down-regulated in caffeic acid O-methyltransferase resulted in the incorporation of 5-hydroxylated guaiacyl subunits in lignin. Many studies also note that monolignols are deposited differentially in discrete regions of particular tissues or cells. The cell walls of vessels in angiosperms are composed mostly of guaiacyl lignin, whereas the interfascicular fibers are enriched in syringyl lignin. Moreover, lignin in the middle lamella is enriched in *p*-hydroxyphenyl lignin moieties, while syringyl and guaiacyl lignin are deposited mostly in the secondary cell wall. These data suggests that the biosynthesis and deposition of lignin monomers into cell wall is a highly organized, regulated process, and that active transportation mechanisms might selectively permit the deposition of the particular monolignols.

Biosynthesis of Monolignols

Lignin, together with other secondary metabolites such as isoprenoids, tannins, flavonoids, lignans and alkaloids are products of the plant phenylpropanoid metabolism. As with all phenylpropanoids, the synthesis of lignin diverges from primary metabolism with the deamination of phenylalanine to form cinnamic acid. A series of hydroxylation and O-methylation reactions modify the aromatic ring of cinnamic acid, and its side chains are activated from acids to coenzyme-A esters, and subsequently reduced to aldehydes and alcohols, resulting in the production of the monolignols.

Phenylalanine ammonia-lyases (PALs)

Studies with radioactive tracers have demonstrated that monolignols are biosynthesized from the phenylpropanoid cinnamic acid. In Poaceae, tyrosine as well as phenylalanine is a precursor for the biosynthesis of cinnamic acid. However, in dicotyledons, the deamination of phenylalanine by PAL is the principle source of cinnamic acid for lignin synthesis. Before PALs were discovered, cinnamic acid was thought to have been synthesized from the dehydration of phenyllactic acid from phenylpyruvic acid. However, the purification and characterization of PAL from *Hordeum vulgare* showed that PAL could directly catalyze the deamination of phenylalanine to cinnamic acid. Radiolabelled tracer studies using celery petioles and sycamore cells also demonstrated the biosynthesis of cinnamic acid from phenylalanine. L-[U-^{3}H] phenylalanine, fed to isolated xylem of celery petioles, was found

to be incorporated into lignin as is L-[U-3H] cinnamic acid when fed to isolated sycamore cells (Rubery, P.H, 1968).

As the enzyme that catalyzes the first reaction in the phenylpropanoid pathway, PAL transcription is heavily regulated spatially and temporally in response to various factors involved in plant defense reactions and growth. Products of the phenylpropanoid pathway are used for a wide range of functions, and the transcriptional regulation of enzyme levels in the phenylpropanoid pathway reflects the diverse controls exerted on the pathway. Flavonoid pigments, lignin, anti-microbial phenolics, UV protectants and cell wall associated phenolics are all derived from the phenylpropanoid pathway, and are synthesized in response to normal developmental cues, pathogen attack, and mechanical and UV stress. Transcriptional regulation by plant development and environmental cues extends into individual members of the PAL family. Differential regulation of PAL genes were first observed in *Phaseolus vulgaris L*. All three *PvPAL*s are expressed at high levels in roots, however only *PvPAL1* and *PvPAL2* are expressed in shoots. *PvPAL3* is not affected by illumination, whereas *PvPAL1* and *PvPAL2* are activated under illumination. Similar observations were documented in *Populus tremuloides*, *Populus trichocarpa* and *Arabidopsis*.

PAL functions at a metabolically important position, it was therefore postulated that PAL is involved in a multienzyme complex that includes downstream enzymes, to facilitate efficient channeling of carbon into the phenylpropanoid pathway (Hirazdina et al. 1978, Wagner and Hrazdina, 1984). Through differential centrifugation, sedimentation of microsomes, gradient

separation of membrane fractions, protein blot analysis of ER, immunolocalization and confocal microscopy, PALs from different plant species have been reported associated with ER membranes where the downstream P450s are localized. However, with the exception of the polyamine metabolon involved in the biosynthesis of spermidine, multienzyme complexes in secondary metabolism involving soluble proteins and P450s are elusive and contain less stable interactions than those that have been reported in primary metabolism. Conclusive demonstration of their existence has been difficult to obtain and additional evidence is still required.

Transgenics perturbed in the expression levels of PALs have been generated in numerous plant species. Severe inhibition of PAL generally leads to decreased levels of lignin and other phenylpropanoids, and an increase in the syringyl to guaiacyl ratio in lignin. Phenotypes such as stunted growth, delayed root formation and altered leaves were also documented in plants with severe PAL down-regulation. However, the high degree of inhibition required to impact lignin deposition suggests that PAL activities are frequently present in excess. This is likely due to its essentiality, as multiple copies of PAL are present in most plant species. Studies of the quantitative relationship between *PAL* expression levels and lignin accumulation in transgenic tobacco revealed that PAL activities becomes the dominant rate-limiting-step in lignin biosynthesis only when the expression levels are 3- to 4-fold below wildtype (Bate et al. 1994).

Cinnamate 4-hydroxylases (C4Hs)

The second step in the phenylpropanoid pathway is the hydroxylation of cinnamic acid to *p*-coumaric acid. The reaction is catalyzed by cinnamate 4-hydroxylase, a member of the superfamily of Cyt P450 heme-thiolate proteins. P450s are ubiquitous enzymes that catalyze the activation of molecular oxygen and the insertion of one of its atoms into physiological and artificial substrates (Porter and Coon, 1991). P450s catalyze the oxidation of a remarkably broad range of endogenous and exogenous chemicals in all organisms. In plants, they play important roles in biosynthetic pathways, including those of sterols, isoprenoids, alkaloids, oxygenated fatty acids, and phenylpropanoids (Bolwell et al., 1994; Durst and O'Keefe, 1995; Schuler, 1996). They are also involved in the metabolism and sometimes in the activation of many herbicides, insecticides, and other xenobiotics.

A characteristic common to many P450s from all organisms is their inducibility by exogenous stimuli. The activity of C4H is increased by light (Benveniste et al. 1978), wounding (Tanaka et al. 1974, Benveniste et al, 1977), infection (Fujita et al. 1982, Maule and Ride 1983, Kochs and Grisebach, 1989) and chemical treatments (Reichhart et al, 1980, Adele et al, 1981). The mechanisms of C4H induction is primarily from gene activation. (Batart et al. 1997).

Transgenic plants perturbed in the expression level of C4H have been generated for tobacco, alfalfa, tomato, *Arabidopsis* and hybrid aspen. Reduction of C4H activity generally leads to decreased levels of lignin and several different classes of phenylpropanoid products, such as sinapoylmalate and flavonoids. The observed changes in lignin composition in response to

lowered C4H are not consistent. Transgenic tobacco and alfalfa exhibited a lowered S/G ratio when C4H is down-regulated, whereas in transgenic tomato and *Arabidopsis*, S/G was greatly increased. C4H perturbed transgenic hybrid aspen did not have an altered S/G ratio. Pleiotropic phenotypes such as dwarfism, male sterility and the development of swellings at branch junctions were also documented in transgenic plants down regulated in C4H.

The idea that C4H functions as an ER localized protein anchor for a large multi-protein complex was proposed decades ago (Czichi and Kindl 1977, Hrazdina and Wagner 1985). Numerous studies have shown that C4H is co-localized and co-regulated with PAL, suggesting that they may be directly involved in a multi enzyme complex, facilitating the channeling of metabolites in the phenylpropanoid pathway (Deshpande et al. 1993, Koopmann et al. 1999, Rasmussen and Dixon. 1999, Sato et al. 2003). However, radiotracer studies using reconstitution of PAL and C4H in yeast demonstrated that PAL and C4H do not require channeling of metabolites through a protein complex (Ro and Douglas. 2004). Through the co expression of multiple plant P450s in yeast, Chen et al. 2011 recently showed that C4H in *Populus trichocarpa* forms a multiple enzyme complex with another plant P450, a coumarate 3-hydroxylase. The formation of C3H/C4H protein complex greatly increased some of the monooxygenation activities of the P450 enzymes.

4-coumarate:coenzyme A ligase (4CLs)

The 4-coumarate:coenzyme A ligase family of enzymes catalyzes the conversion of hydroxycinnamic acids to CoA thioesters. Phenylpropanoid CoA thioesters are precursors to many products of the general phenylpropanoid pathway, including benzoic acid, tannins, lignin, flavonoids and lignans.

4CL isolated from various plant species are generally found to be able to metabolize a number of hydroxycinnamic acid derivatives as substrates, in particular, *p*-coumaric acid, caffeic acid and ferulic acid to p-coumaroyl-CoA, caffeoyl-CoA and feruloyl-CoA respectively (Gross and Zenk, 1974; Knobloch and Hahlbrock, 1975, 1977; Grand et al., 1983; Lozoya et al., 1988; Voo et al., 1995; Lee and Douglas, 1996). Multiple copies of 4CLs are present in the genome of most plant species. It has been proposed, that the different isoforms of 4CL may possess different patterns of substrate preference, and the characterization of 4CL isoforms from some plant species supports this (Knobloch and Hahlbrock, 1975; Ranjeva et al., 1976; Wallis and Rhodes, 1977; Grand et al., 1983). Such diversity could potentially enable particular 4CL isoforms to differenciately supply specific CoA thioesters to various metabolic pathways. However, in other plants in which they have been purified, no evidence towards catalytically distinct 4CL isoforms were observed (Lozoya et al., 1988; Voo et al., 1995).

Transgenic plants perturbed in the expression of 4CLs have been generated in tobacco (Kajita et al. 1996), *Arabidopsis* (Lee et al. 1997), aspen (Hu et al. 1999), *radiata* pine (Wagner et al. 2009), hybrid poplar (Kitin et al. 2010, Voelker et al. 2010), hybrid white poplar (Voelker et al. 2011) and switchgrass ((Xu et al. 2011). Reduction in the activity of 4CL has a significant impact on both lignin content and composition. Lignin quantity is reduced in all 4CL transgenic plants. The reduction in lignin maybe compensated by an increase in the cellulose content (Hu et al. 1999, Li et al. 2003). Soluble and wall bound phenolics *p*-coumaric acid, ferulic acid and sinapic acid accumulate in transgenic plants with severe 4CL reduction (Kajita et al. 1997, Hu et al. 1999). Changes in lignin composition are inconsistent between different plant species downregulated in 4CL expression. Aspen, tobacco, *Arabidopsis* and switchgrass exhibited an increase in the S/G ratio of lignin when 4CL is reduced. The increase in S/G is mainly due to the reduction in G-lignin deposition. The S/G ratio of transgenic hybrid popar (tremula × Populus alba) however, decreased when 4CL expression was down-regulated. The roles of functionally distinct 4CL isoforms in directing the metabolic flux for the biosynthesis of different phenolics remains to be studied further.

Cinnamate 3-hydroxylases (C3Hs)

Cinnamate 3-hydroxylases are one of the three plant P450s required in the phenylpropanoid pathway, and catalyze the hydroxylation at the 3-position of the ring. C3Hs were one of the last of the monolignol genes to be cloned and characterized. This is in part due to the detergent instability, low abundance and the membrane-bound nature of the plant P450s.

Many researchers have attempted to purify and characterize C3Hs. A considerable amount of disagreements concerning the nature of C3H was published in the past 40 years. Putative C3Hs have been documented by various researchers as either an ascorbate-, NADPH- or flavin-dependent, mixed function oxidase (Boniwell and Butt, 1986; Kojima and Takeuchi, 1989; Stafford and Dresler, 1972; Vaughan and Butt, 1970), a plastidic enzyme that uses plastoquinone or ferredoxin as an electron donor (Bartlett et al., 1972), or most commonly, a phenolase that also oxidizes dihydroxyphenols to their corresponding orthoquinones (Stafford and Dresler, 1972).

Through a bioinformatics approach focused on the phylogeny of plant P450s, finally, the first gene encoding C3H (CYP98A3) was cloned and characterized from the genome of Arabidopsis (Schoch et al. 2001). Less than a year later, Arabidopsis mutant ref8 carrying a mutation in the CYP98A3 was characterized, and confirmed the involvement of CYP98A3 in the 3-hydroxylation of phenylpropanoids (Franke et al. 2002). Functional characterization of C3Hs revealed that neither p-coumaric acid or its glucose or CoA esters, p-coumaraldehyde, or p-coumaryl alcohol were efficiently metabolized. However, the shikimate ester of *p*-coumaric acid was very actively 3-hydroxylated to the corresponding caffeic acid conjugate. In vitro characterization of type II methyltransferases from various plant species suggests that shikimates are not substrates of CCoAOMTs. This indicates that caffeoyl shikimic acid must be converted back to a CoA ester for downstream methylation.

Transgenic plants with downregulation of C3H have been generated and characterized from alfalfa (Reddy et al. 2005) and hybrid poplar (Coleman et al. 2008). Reductions in C3H

activity is generally associated with reductions in lignin content. Changes in lignin composition were also observed. Lignin of C3H transgenic plants have significant reductions in guaiacyl subunits, while the level of syringyl lignin is maintained. The reduction in guaiacyl lignin is compensated by an increase in the level of incorporated p-hydroxyphenyl units. Pleiotropic phenotypes associated with the downregulation of C3Hs include developmental defects and a lowered resistance to fungal attacks (Franke et al. 2002).

Hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl transferases (HCTs)

The cloning and characterization of hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl transferases stemmed from the discovery of p-coumaroyl shikimate as a principle intermediate in the monolignol biosynthesis pathway (Schoch et al. 2001). The shikimate transferase activities have been documented in cell suspension cultures of parsley and carrot (Heller and Kuhnl 1985, Kuhnl et al. 1987), however it was not evident at the time that these activities were associated with monolignol pathway. HCT was recently purified from tobacco (Hoffmann et al. 2003). Its characterization showed that p-coumaroyl-CoA and caffeoyl-CoA are efficiently converted to the corresponding shikimate esters, and to some degree, the catalysis is reversible.

Characterization of HCT transgenic Arabidopsis and tobacco plants, provided unequivocal proof of the involvement of HCT in lignin biosynthesis. HCT downregulated plants exhibited significant changes in plant development, lignin content, and susceptibility of the cell walls to enzymatic degradation. Immunolocalization of HCT in wildtypes and HCT transgenic

plants demonstrated a correlation between the disappearance HCT in vascular cells and the impact on lignin structure. Transgenic tobacco with perturbed HCT displayed surprisingly higher autofluorescence, suggesting that the enrichment of p-hydroxyphenyl units in lignin may alter its spectral properties. This phenomenon could be linked to the fact that H-lignin units are essentially terminal lignin units with free phenolic groups (Lange et al., 1995). The decrease in HCT protein accumulation in xylem tissues also strictly correlated with a decrease in lignin S unit content.

O-methyltransferases (COMTs and CCoAOMTs)

Caffeic acid O-methyltransferase (COMT) and caffeoyl coenzyme A O-methyltransferase (CCoAOMT) are two very well characterized enzymes in plant secondary metabolism. These two structurally distinct families of methyltransferases catalyzes the 3- and 5- methoxylation of monolignol precursors in the biosynthesis of guaiacyl and syringyl monolignol (Dixon and Reddy 2003; Davin et al. 2008). The hydroxylation and methoxylation states of the 3- and 5-positions of lignin subunits affect the physical and chemical properties of polymeric lignin and consequentially the digestibility and forage quality of plant tissues (reviewed in Boerjan et al., 2003; Raes et al., 2003).

CCoAOMTs are categorized as type II methyltransferases. The S-adenosyl-L-methionine (SAM) and cation-dependent CCoAOMT is considered specific for the methoxylation of CoA thioesters, which are key intermediates in the biosynthesis of monolignols.

COMTs are type I methyltransferases. They are SAM dependent multifunctional methyltransferases that are less specific in their substrate binding compared to type II methyltransferases and can methoxylate both the 3- and 5-hydroxyl group of various phenylpropanoids (Li et al. 2000; Humphreys and Chapple 2002; Boerjan et al. 2003; Davin et al. 2008).

O-methylation is a critical step in the biosynthesis of lignin. Double knockouts in Arabidopsis targeting both COMT and CCoAOMT result in a severe dwarf phenotype without the development of reproductive organs in a variety of plants (Do et al. 2007). Transgenics perturbed in the expression of O-methyltransferases have been generated in a number of plant species including tobacco, quaking aspen, hybrid poplar, alfalfa, maize and white leadtree (Dwivedi et al. 1994, Tsai et al. 1998, Meyermans et al. 2000, Guo et al. 2001, Piquemal et al. 2002, Rastogi and Dwivedi. 2006). Reduction of COMT activities generally results in mild decreases in lignin content, severe reduction in syringyl content in lignin and the incorporation of 5-hydroxy guaiacyl units into lignin. The down-regulation of CCoAOMTs generally result in a more severe decrease in lignin content with an increased S/G ratio. The reduced lignin is accompanied by the accumulation of phenolic glucosides.

Lastly, even though a reduction in the total lignin content and changes in lignin composition are observed in single knock-down transgenic of various plant species, both enzymes to some extent are functionally complementary (Guo et al. 2001; Chen et al. 2006; Day et al. 2009).

Coniferaldehyde 5-hydroxylase (CAld5H)

Coniferaldehyde 5-hydroxylases define a group of heme-containing cytochrome P450-dependent monooxygenases that catalyzes the first reaction in the divergent branch from guaiacyl monolignols to the syringyl specific pathway. Unlike C4H and numerous other plant P450s, the isolation of CAld5H by protein purification based strategies proved difficult. CAld5Hs constitute a small fraction of total P450s. Their detergent instability and activities that are difficult to detect in plant extracts meant that the isolation of 5-hydroxylases had to be through alternative strategies. The first gene encoding a CAld5H was cloned by mapping the mutation present in the *Arabidopsis fah1* mutant, which was isolated in a forward genetic screen for plants lacking sinapoylmalate, a major phenolic secondary metabolite in *Arabidopsis*. In mutant screens, *fah1* mutants are readily identified by their characteristic red fluorescence when observed under UV light. Sinapoylmalate is a fluorescent secondary metabolite that is accumulated in the leaf epidermis. Wildtype *Arabidopsis* exhibits pale blue fluorescence under UV, while *fah1* mutants lacking in sinapoylmalate appear dark red due to the fluorescence of chlorophyll. Chemical analysis of the *fah1* mutant revealed that it produces lignin that is completely devoid of syringyl moiety.

For decades, the 5-hydroxylations of phenolics toward sinapyl alcohol biosynthesis were thought to occur at the level of the hydroxycinnamic acids (Grand 1984). However, sinapic acid as a principle intermediate in the biosynthesis of syringyl lignin has not been demonstrated in *planta*. *In vitro* characterization of CoA ligases from various angiosperm species cannot activate sinapic acid into its CoA thioester derivative for syringyl lignin biosynthesis (Gross et al. 1975, Kutsuki and Higuchi 1982, Hu et al. 1998). More recently,

biochemical characterization of recombinant CAld5Hs from poplar and *Arabidopsis* revealed that the syringyl pathway stems from hydroxylation at the levels of the free aldehydes and alcohols (Osakabe et al. 1999; Humphreys et al. 1999).

The targeted knockdown and overexpression of CAld5H have been documented in many plant species, which result in changes in the composition of lignin ranging from almost no syringyl subunits in lignin to nearly one hundred percent syringyl lignin (Meyer et al. 1998, Ruegger et al. 1999, Franke et al. 2000, Huntley et al. 2003, Li et al. 2003, Reddy et al. 2005, Stewart et al. 2009). These transgenic plants are virtually indistinguishable from wildtype plants in terms of their morphology, demonstrating that the lignin polymer can undergo substantial alterations in composition, without affecting the plant phenotype.

Recalcitrance to saccharification is a major limitation in the utilization of lignocellulosic biomass for pulp and energy. Accessibility of plant cell wall polysaccharides to chemical and enzymatic digestion is largely limited by the presence of lignin. The physical and chemical properties of lignin is greatly affected by its composition. Transgenic plants with increased syringyl subunits in lignin have been shown to be more efficient for enzymatic hydrolysis and chemical pretreatment.

Cinnamoyl-CoA reductases (CCRs) and Cinnamoyl alcohol dehydrogenases (CADs)

The reduction of hydroxycinnamyl CoA thioesters to hydroxycinnamyl aldehydes and then to monolignols are catalyzed by cinnamoyl-CoA reductases and cinnamyl alcohol dehydrogenases respectively. Reduction of phenolic CoAs constitute the first committed step

in monolignol biosynthesis from the general phenylpropanoid pathway. Both CCR and CAD were identified from the amino acid sequences of the purified proteins. Recent studies suggest that monolignol biosynthesis proceeds via the reduction of feruloyl-CoA to coniferaldehyde and the reduction of either coniferaldehyde to coniferyl alcohol or, following phenolic ring modification, sinapaldehyde to sinapyl alcohol.

Transgenic tobacco down regulated in CCR showed a reduced lignin deposition and dwarfism, associated with a collapse of the xylem (Chanbannes et al. 2001). Similar phenotypes were reported in *Arabidopsis* mutants with nonfunctional CCR genes (Goujin et al. 2003). Biochemical characterization of these transgenics indicate that, CCR deficiency produces lignin that contains increased levels of ferulic acid, demonstrating that atypical intermediates in the monolignol biosynthesis pathway can be incorporated into the lignin polymer when pathway fluxes are perturbed.

The antisense knockdown of CAD in tobacco resulted in plants with lignin that was relatively comparable in quantity to the control plants, but which contained a substantial level of coniferaldehyde (Chanbannes et al. 2001). Similar observations were reported in the pine mutant *cad-n1* (Moller et al. 2006), the maize *bm1* mutant (Halpin et al. 1998), and in *Arabidopsis* mutants (Sibout et al. 2003). In contrast to CCR-deficient plants, transgenics with decreased CAD activities are not dwarfed, suggesting that hydroxycinnamyl aldehydes can, to some extent functionally substitute for the canonical monolignols.

Monolignol transport and polymerization

Monolignols are biosynthesized in the cytoplasm. However, the deposition of lignin takes place in the middle lamella or the secondary cell wall. Lignin precursors must be transported across the plasma and vacuolar membranes to the lignifying zone where they are oxidized and polymerized.

The underlying molecular mechanisms governing this transportation process remains unclear. Recent studies on *Arabidopsis* and poplar, using *in vitro* uptake assays demonstrated that the monolignols are transported to the cell wall through an ATP-dependent primary-transport process, involving ATP-binding cassette-like transporters (Miao and Liu, 2010).

Lignin polymerization denotes the process by which monolignols and/or oligolignols are polymerized via oxidative coupling. Three successive steps are involved in the generation of the lignin polymer. The monolignols are first radicalized by hydrogen peroxide and peroxidase. The monomeric radicals then couple to each other, or to the radicalized phenolic groups on the lignin polymer, in accordance to their relative supply and coupling propensities.

Two models to the polymerization of monolignols have been proposed recently. A combinatorial model suggests that the monolignols in the extracellular space polymerize with one another in the absence of proteinaceous control, subject only to the supply of monolignols and the usual physical parameters governing chemical reactions. In contrast, another model proposes a direct oxidation of the growing polymer by peroxidase. The mechanism requires that peroxidases directly oxidize both lignin macromolecules and

monolignols, and that a phenolic radical on the polylignol be coupled to a monolignol radical. The proteins responsible for such control have been proposed to contain specific radical binding sites similar to those found in characterized dirigent proteins, and dictates the order of monolignols and intersubunit linkages. In line with the current discussion of lignin mutants and transgenics, however, it is important to point out that the genetic data fall clearly in line with the combinatorial model.

Deposition of lignin is always preceded by the deposition of polysaccharides. Early deposition of lignin in the cell corners and middle lamella occurs after pectic substances have been deposited. However, the predominant lignin deposition does not occur until after cellulose microfibrils have been deposited in mature secondary cell walls.

References

Structural Alterations of Lignins in Transgenic Poplars with Depressed Cinnamyl Alcohol Dehydrogenase or Caffeic Acid O-Methyltransferase Activity Have an Opposite Impact on the Efficiency of Industrial Kraft Pulping ;2012.

Achnine L, EB Blancaflor, S Rasmussen, RA Dixon 2004 Colocalization of L-phenylalanine ammonia-lyase and cinnamate 4-hydroxylase for metabolic channeling in phenylpropanoid biosynthesis Plant Cell ;16:3098-3109.

Amrhein N, G Frank, G Lemm, HB Luhmann 1983 Inhibition of lignin formation by L-alpha-aminooxy-beta-phenylpropionic acid, an inhibitor of phenylalanine ammonia-lyase Eur.J.Cell Biol. ;29:139-144.

Bagal UR, JH Leebens-Mack, WW Lorenz, JF Dean 2012 The phenylalanine ammonia lyase (PAL) gene family shows a gymnosperm-specific lineage BMC Genomics ;13 Suppl 3:S1.

Bassard J, J Mutterer, F Duval, D Werck-Reichhart 2012; 2011 A novel method for monitoring the localization of cytochromes P450 and other endoplasmic reticulum membrane associated proteins: a tool for investigating the formation of metabolons FEBS Journal ;279:1576 <last_page> 1583.

Batard Y, M Schalk, MA Pierrel, A Zimmerlin, F Durst, D Werck-Reichhart 1997 Regulation of the Cinnamate 4-Hydroxylase (CYP73A1) in Jerusalem Artichoke Tubers in Response to Wounding and Chemical Treatments Plant Physiol. ;113:951-959.

Bate NJ, J Orr, W Ni, A Meromi, T Nadler-Hassar, PW Doerner, RA Dixon, CJ Lamb, Y Elkind 1994 Quantitative relationship between phenylalanine ammonia-lyase levels and phenylpropanoid accumulation in transgenic tobacco identifies a rate-determining step in natural product synthesis Proc.Natl.Acad.Sci.U.S.A. ;91:7608-7612.

Bell-Lelong DA, JC Cusumano, K Meyer, C Chapple 1997 Cinnamate-4-hydroxylase expression in Arabidopsis. Regulation in response to development and the environment Plant Physiol. ;113:729-738.

Betz C, TG McCollum, RT Mayer 2001 Differential expression of two cinnamate 4-hydroxylase genes in 'Valencia' orange (Citrus sinensis Osbeck) Plant Mol.Biol. ;46:741-748.

Bevan M, D Shufflebottom, K Edwards, R Jefferson, W Schuch 1989 Tissue- and cell-specific activity of a phenylalanine ammonia-lyase promoter in transgenic plants EMBO J. ;8:1899-1906.

Bjurhager I, AM Olsson, B Zhang, L Gerber, M Kumar, LA Berglund, I Burgert, B Sundberg, L Salmen 2010 Ultrastructure and mechanical properties of populus wood with reduced lignin content caused by transgenic down-regulation of cinnamate 4-hydroxylase Biomacromolecules ;11:2359-2365.

Blee K, JW Choi, AP O'Connell, SC Jupe, W Schuch, NG Lewis, GP Bolwell 2001 Antisense and sense expression of cDNA coding for CYP73A15, a class II cinnamate 4-hydroxylase, leads to a delayed and reduced production of lignin in tobacco Phytochemistry ;57:1159-1166.

Blount JW, KL Korth, SA Masoud, S Rasmussen, C Lamb, RA Dixon 2000 Altering expression of cinnamic acid 4-hydroxylase in transgenic plants provides evidence for a feedback loop at the entry point into the phenylpropanoid pathway Plant Physiol. ;122:107-116.

Bubna GA, RB Lima, DY Zanardo, WD Dos Santos, MD Ferrarese, O Ferrarese-Filho 2011 Exogenous caffeic acid inhibits the growth and enhances the lignification of the roots of soybean (Glycine max) J.Plant Physiol. .

C P Vance,A M Nambudiri,G H Towers 1973 Cinnamate-4-hydroxylase in Polyporus hispidus. Canadian journal of biochemistry ;51:731-734.

Chalutz E 1973 Ethylene-induced Phenylalanine Ammonia-Lyase Activity in Carrot Roots Plant Physiol. ;51:1033 <last_page> 1036.

Chen F, MS Srinivasa Reddy, S Temple, L Jackson, G Shadle, RA Dixon 2006 Multi-site genetic modulation of monolignol biosynthesis suggests new routes for formation of syringyl lignin and wall-bound ferulic acid in alfalfa (Medicago sativa L.) Plant J. ;48:113-124.

Chen HC, Q Li, CM Shuford, J Liu, DC Muddiman, RR Sederoff, VL Chiang 2011 Membrane protein complexes catalyze both 4- and 3-hydroxylation of cinnamic acid derivatives in monolignol biosynthesis Proc.Natl.Acad.Sci.U.S.A. ;108:21253-21258.

Cheng CK-, HV Marsh 1968 Gibberellic Acid-Promoted Lignification and Phenylalanine Ammonia-lyase Activity in a Dwarf Pea (Pisum sativum) Plant Physiol. ;43:1755 <last_page> 1759.

Coleman HD, JY Park, R Nair, C Chapple, SD Mansfield 2008 RNAi-mediated suppression of p-coumaroyl-CoA 3'-hydroxylase in hybrid poplar impacts lignin deposition and soluble secondary metabolism Proc.Natl.Acad.Sci.U.S.A. ;105:4501-4506.

Czichi U, H Kindl 1977 Phenylalanine ammonia lyase and cinnamic acid hydroxylases as assembled consecutive enzymes on microsomal membranes of cucumber cotyledons: Cooperation and subcellular distribution Planta ;134:133 <last_page> 143.

Dalkin K, R Edwards, B Edington, RA Dixon 1990 Stress Responses in Alfalfa (Medicago sativa L.): I. Induction of Phenylpropanoid Biosynthesis and Hydrolytic Enzymes in Elicitor-Treated Cell Suspension Cultures Plant Physiol. ;92:440-446.

Deshpande AS, KK Surendranathan, PM Nair 1993 The phenyl propanoid pathway enzymes in Solanum tuberosum exist as a multienzyme complex Indian J.Biochem.Biophys. ;30:36-41.

Dwivedi UN, WH Campbell, J Yu, RS Datla, RC Bugos, VL Chiang, GK Podila 1994 Modification of lignin biosynthesis in transgenic Nicotiana through expression of an antisense O-methyltransferase gene from Populus Plant Mol.Biol. ;26:61-71.

Ehlting J, D Buttner, Q Wang, CJ Douglas, IE Somssich, E Kombrink 1999 Three 4-coumarate:coenzyme A ligases in Arabidopsis thaliana represent two evolutionarily divergent classes in angiosperms Plant J. ;19:9-20.

Emiliani G, M Fondi, R Fani, S Gribaldo 2009 A horizontal gene transfer at the origin of phenylpropanoid metabolism: a key adaptation of plants to land Biol.Direct ;4:7.

Fahrendorf T, RA Dixon 1993 Stress responses in alfalfa (Medicago sativa L.). XVIII: Molecular cloning and expression of the elicitor-inducible cinnamic acid 4-hydroxylase cytochrome P450 Arch.Biochem.Biophys. ;305:509-515.

Frank MR, JM Deyneka, MA Schuler 1996 Cloning of wound-induced cytochrome P450 monooxygenases expressed in pea Plant Physiol. ;110:1035-1046.

Franke R, MR Hemm, JW Denault, MO Ruegger, JM Humphreys, C Chapple 2002a Changes in secondary metabolism and deposition of an unusual lignin in the ref8 mutant of Arabidopsis Plant J. ;30:47-59.

Franke R, JM Humphreys, MR Hemm, JW Denault, MO Ruegger, JC Cusumano, C Chapple 2002b The Arabidopsis REF8 gene encodes the 3-hydroxylase of phenylpropanoid metabolism Plant J. ;30:33-45.

Franke R, CM McMichael, K Meyer, AM Shirley, JC Cusumano, C Chapple 2000 Modified lignin in tobacco and poplar plants over-expressing the Arabidopsis gene encoding ferulate 5-hydroxylase Plant J. ;22:223-234.

Gabriac B, D Werck-Reichhart, H Teutsch, F Durst 1991 Purification and immunocharacterization of a plant cytochrome P450: the cinnamic acid 4-hydroxylase Arch.Biochem.Biophys. ;288:302-309.

Grand C 1984 Ferulic acid 5-hydroxylase: a new cytochrome P-450-dependent enzyme from higher plant microsomes involved in lignin synthesis FEBS Lett. ;169:7 <last_page> 11.

Guerra D, AJ Anderson, FB Salisbury 1985 Reduced phenylalanine ammonia-lyase and tyrosine ammonia-lyase activities and lignin synthesis in wheat grown under low pressure sodium lamps Plant Physiol. ;78:126-130.

Guo D, F Chen, K Inoue, JW Blount, RA Dixon 2001 Downregulation of caffeic acid 3-O-methyltransferase and caffeoyl CoA 3-O-methyltransferase in transgenic alfalfa. impacts on lignin structure and implications for the biosynthesis of G and S lignin Plant Cell ;13:73-88.

Hamada K, T Nishida, K Yamauchi, K Fukushima, R Kondo, Y Tsutsumi 2004 4-Coumarate:coenzyme A ligase in black locust (Robinia pseudoacacia) catalyses the conversion of sinapate to sinapoyl-CoA J.Plant Res. ;117:303-310.

Harding SA, J Leshkevich, VL Chiang, CJ Tsai 2002 Differential substrate inhibition couples kinetically distinct 4-coumarate:coenzyme a ligases with spatially distinct metabolic roles in quaking aspen Plant Physiol. ;128:428-438.

Hoffmann L, S Besseau, P Geoffroy, C Ritzenthaler, D Meyer, C Lapierre, B Pollet, M Legrand 2004 Silencing of hydroxycinnamoyl-coenzyme A shikimate/quinate hydroxycinnamoyltransferase affects phenylpropanoid biosynthesis Plant Cell ;16:1446-1465.

Hotze M, G Schroder, J Schroder 1995 Cinnamate 4-hydroxylase from Catharanthus roseus, and a strategy for the functional expression of plant cytochrome P450 proteins as translational fusions with P450 reductase in Escherichia coli FEBS Lett. ;374:345-350.

Hrazdina G, GJ Wagner 1985 Metabolic pathways as enzyme complexes: evidence for the synthesis of phenylpropanoids and flavonoids on membrane associated enzyme complexes Arch.Biochem.Biophys. ;237:88-100.

Hu GS, JM Jia, YJ Hur, YS Chung, JH Lee, DJ Yun, WS Chung, GH Yi, TH Kim, DH Kim 2011 Molecular characterization of phenylalanine ammonia lyase gene from Cistanche deserticola Mol.Biol.Rep. ;38:3741-3750.

Hu WJ, SA Harding, J Lung, JL Popko, J Ralph, DD Stokke, CJ Tsai, VL Chiang 1999 Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees Nat.Biotechnol. ;17:808-812.

Hu WJ, A Kawaoka, CJ Tsai, J Lung, K Osakabe, H Ebinuma, VL Chiang 1998 Compartmentalized expression of two structurally and functionally distinct 4-coumarate:CoA ligase genes in aspen (Populus tremuloides) Proc.Natl.Acad.Sci.U.S.A. ;95:5407-5412.

Hu Y, Y Gai, L Yin, X Wang, C Feng, L Feng, D Li, XN Jiang, DC Wang 2010 Crystal structures of a Populus tomentosa 4-coumarate:CoA ligase shed light on its enzymatic mechanisms Plant Cell ;22:3093-3104.

Hubner S, M Hehmann, S Schreiner, S Martens, R Lukacin, U Matern 2003 Functional expression of cinnamate 4-hydroxylase from Ammi majus L Phytochemistry ;64:445-452.

Humphreys JM, MR Hemm, C Chapple 1999 New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase Proc.Natl.Acad.Sci.U.S.A. ;96:10045-10050.

Huntley SK, D Ellis, M Gilbert, C Chapple, SD Mansfield 2003 Significant increases in pulping efficiency in C4H-F5H-transformed poplars: improved chemical savings and reduced environmental toxins J.Agric.Food Chem. ;51:6178-6183.

Hyodo H, SF Yang 1971 Ethylene-enhanced Synthesis of Phenylalanine Ammonia-Lyase in Pea Seedlings Plant Physiol. ;47:765-770.

Jouanin L, T Goujon, V de Nadai, MT Martin, I Mila, C Vallet, B Pollet, A Yoshinaga, B Chabbert, M Petit-Conil, C Lapierre 2000 Lignification in transgenic poplars with extremely reduced caffeic acid O-methyltransferase activity Plant Physiol. ;123:1363-1374.

Kajita S, S Hishiyama, Y Tomimura, Y Katayama, S Omori 1997 Structural Characterization of Modified Lignin in Transgenic Tobacco Plants in Which the Activity of 4-Coumarate:Coenzyme A Ligase Is Depressed Plant Physiol. ;114:871-879.

Kajita S, Y Katayama, S Omori 1996 Alterations in the biosynthesis of lignin in transgenic plants with chimeric genes for 4-coumarate: coenzyme A ligase Plant Cell Physiol. ;37:957-965.

Kao YY, SA Harding, CJ Tsai 2002 Differential expression of two distinct phenylalanine ammonia-lyase genes in condensed tannin-accumulating and lignifying cells of quaking aspen Plant Physiol. ;130:796-807.

Kawai S, A Mori, T Shiokawa, S Kajita, Y Katayama, N Morohoshi 1996 Isolation and analysis of cinnamic acid 4-hydroxylase homologous genes from a hybrid aspen, Populus kitakamiensis Biosci.Biotechnol.Biochem. ;60:1586-1597.

Kim YJ, DG Kim, SH Lee, I Lee 2006 Wound-induced expression of the ferulate 5-hydroxylase gene in Camptotheca acuminata Biochim.Biophys.Acta ;1760:182-190.

Kitin P, SL Voelker, FC Meinzer, H Beeckman, SH Strauss, B Lachenbruch 2010 Tyloses and phenolic deposits in xylem vessels impede water transport in low-lignin transgenic poplars: a study by cryo-fluorescence microscopy Plant Physiol. ;154:887-898.

Kochs G, D Werck-Reichhart, H Grisebach 1992 Further characterization of cytochrome P450 involved in phytoalexin synthesis in soybean: cytochrome P450 cinnamate 4-hydroxylase and 3,9-dihydroxypterocarpan 6a-hydroxylase Arch.Biochem.Biophys.

;293:187-194.

Koopmann E, E Logemann, K Hahlbrock 1999 Regulation and functional expression of cinnamate 4-hydroxylase from parsley Plant Physiol. ;119:49-56.

KOUKOL J, EE CONN 1961 The metabolism of aromatic compounds in higher plants. IV. Purification and properties of the phenylalanine deaminase of Hordeum vulgare J.Biol.Chem. ;236:2692-2698.

Lee D, M Ellard, LA Wanner, KR Davis, CJ Douglas 1995 The Arabidopsis thaliana 4-coumarate:CoA ligase (4CL) gene: stress and developmentally regulated expression and nucleotide sequence of its cDNA Plant Mol.Biol. ;28:871-884.

Lee D, K Meyer, C Chapple, CJ Douglas 1997 Antisense suppression of 4-coumarate:coenzyme A ligase activity in Arabidopsis leads to altered lignin subunit composition Plant Cell ;9:1985-1998.

Li L, Y Zhou, X Cheng, J Sun, JM Marita, J Ralph, VL Chiang 2003 Combinatorial modification of multiple lignin traits in trees through multigene cotransformation Proc.Natl.Acad.Sci.U.S.A. ;100:4939-4944.

Liang XW, M Dron, CL Cramer, RA Dixon, CJ Lamb 1989 Differential regulation of phenylalanine ammonia-lyase genes during plant development and by environmental cues J.Biol.Chem. ;264:14486-14492.

Liu S, Y Hu, X Wang, L Han, S Song, H Cheng, Z Lin 2009 Isolation and characterization of a gene encoding cinnamate 4-hydroxylase from Parthenocissus henryana Mol.Biol.Rep. ;36:1605-1610.

Lu H, YL Zhao, XN Jiang 2004 Stable and specific expression of 4-coumarate:coenzyme A ligase gene (4CL1) driven by the xylem-specific Pto4CL1 promoter in the transgenic tobacco Biotechnol.Lett. ;26:1147-1152.

Lu S, Y Zhou, L Li, VL Chiang 2006 Distinct roles of cinnamate 4-hydroxylase genes in Populus Plant Cell Physiol. ;47:905-914.

Luderitz T, G Schatz, H Grisebach 1982 Enzymic synthesis of lignin precursors. Purification and properties of 4-coumarate:CoA ligase from cambial sap of spruce (Picea abies L.) Eur.J.Biochem. ;123:583-586.

Maher EA, NJ Bate, W Ni, Y Elkind, RA Dixon, CJ Lamb 1994 Increased disease susceptibility of transgenic tobacco plants with suppressed levels of preformed phenylpropanoid products Proc.Natl.Acad.Sci.U.S.A. ;91:7802-7806.

Marita JM, J Ralph, RD Hatfield, C Chapple 1999 NMR characterization of lignins in Arabidopsis altered in the activity of ferulate 5-hydroxylase Proc.Natl.Acad.Sci.U.S.A. ;96:12328-12332.

Meyer K, A Shirley, JC Cusumano, DA Bell-Lelong, C Chapple 1998 Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America ;95:6619-6623.

Meyer K, JC Cusumano, C Somerville, CC Chapple 1996 Ferulate-5-hydroxylase from Arabidopsis thaliana defines a new family of cytochrome P450-dependent monooxygenases Proc.Natl.Acad.Sci.U.S.A. ;93:6869-6874.

Meyermans H, K Morreel, C Lapierre, B Pollet, A De Bruyn, R Busson, P Herdewijn, B Devreese, J Van Beeumen, JM Marita, J Ralph, C Chen, B Burggraeve, M Van Montagu, E Messens, W Boerjan 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.

Millar DJ, M Long, G Donovan, PD Fraser, AM Boudet, S Danoun, PM Bramley, GP Bolwell 2007 Introduction of sense constructs of cinnamate 4-hydroxylase (CYP73A24) in transgenic tomato plants shows opposite effects on flux into stem lignin and fruit flavonoids Phytochemistry ;68:1497-1509.

Mizutani M, E Ward, J DiMaio, D Ohta, J Ryals, R Sato 1993 Molecular cloning and sequencing of a cDNA encoding mung bean cytochrome P450 (P450C4H) possessing cinnamate 4-hydroxylase activity Biochem.Biophys.Res.Commun. ;190:875-880.

Nedelkina S, SC Jupe, KA Blee, M Schalk, D Werck-Reichhart, GP Bolwell 1999 Novel characteristics and regulation of a divergent cinnamate 4-hydroxylase (CYP73A15) from French bean: engineering expression in yeast Plant Mol.Biol. ;39:1079-1090.

Nozzolillo C, KB Paul, C Godin 1971 The Fate of L-Phenylalanine Fed to Germinating Pea Seeds, Pisum sativum (L.) var. Alaska, during Imbibition Plant Physiol. ;47:119 <last_page> 123.

Olsen KM, US Lea, R Slimestad, M Verheul, C Lillo 2008 Differential expression of four Arabidopsis PAL genes; PAL1 and PAL2 have functional specialization in abiotic environmental-triggered flavonoid synthesis J.Plant Physiol. ;165:1491-1499.

Osakabe K, CC Tsao, L Li, JL Popko, T Umezawa, DT Carraway, RH Smeltzer, CP Joshi, VL Chiang 1999 Coniferyl aldehyde 5-hydroxylation and methylation direct syringyl lignin biosynthesis in angiosperms Proc.Natl.Acad.Sci.U.S.A. ;96:8955-8960.

Pendharkar MB, PM Nair 1993 Studies on cinnamic acid-4-hydroxylase from wounded potato tissue: isolation and characterization of cinnamic acid-4-hydroxylase activation factor Indian J.Biochem.Biophys. ;30:89-97.

Pincon G, S Maury, L Hoffmann, P Geoffroy, C Lapierre, B Pollet, M Legrand 2001 Repression of O-methyltransferase genes in transgenic tobacco affects lignin synthesis and plant growth Phytochemistry ;57:1167-1176.

Piquemal J, S Chamayou, I Nadaud, M Beckert, Y Barriere, I Mila, C Lapierre, J Rigau, P Puigdomenech, A Jauneau, C Digonnet, AM Boudet, D Goffner, M Pichon 2002 Down-regulation of caffeic acid o-methyltransferase in maize revisited using a transgenic approach Plant Physiol. ;130:1675-1685.

Ralph J, T Akiyama, H Kim, F Lu, PF Schatz, JM Marita, SA Ralph, MS Reddy, F Chen, RA Dixon 2006 Effects of coumarate 3-hydroxylase down-regulation on lignin structure J.Biol.Chem. ;281:8843-8853.

Rasmussen S, RA Dixon 1999 Transgene-mediated and elicitor-induced perturbation of metabolic channeling at the entry point into the phenylpropanoid pathway Plant Cell ;11:1537-1552.

Rastogi S, UN Dwivedi 2006 Down-regulation of lignin biosynthesis in transgenic Leucaena leucocephala harboring O-methyltransferase gene Biotechnol.Prog. ;22:609-616.

Reddy MS, F Chen, G Shadle, L Jackson, H Aljoe, RA Dixon 2005a Targeted down-regulation of cytochrome P450 enzymes for forage quality improvement in alfalfa (Medicago sativa L.) Proc.Natl.Acad.Sci.U.S.A. ;102:16573-16578.

Reddy MS, F Chen, G Shadle, L Jackson, H Aljoe, RA Dixon 2005b Targeted down-regulation of cytochrome P450 enzymes for forage quality improvement in alfalfa (Medicago sativa L.) Proc.Natl.Acad.Sci.U.S.A. ;102:16573-16578.

Reddy MS, F Chen, G Shadle, L Jackson, H Aljoe, RA Dixon 2005c Targeted down-regulation of cytochrome P450 enzymes for forage quality improvement in alfalfa (Medicago sativa L.) Proc.Natl.Acad.Sci.U.S.A. ;102:16573-16578.

Reichhart D, A Simon, F Durst, JM Mathews, PR Ortiz de Montellano 1982 Autocatalytic inactivation of plant cytochrome P-450 enzymes: selective inactivation of cinnamic acid 4-hydroxylase from Helianthus tuberosus by 1-aminobenzotriazole Arch.Biochem.Biophys. ;216:522-529.

Ro DK, CJ Douglas 2004a Reconstitution of the entry point of plant phenylpropanoid metabolism in yeast (Saccharomyces cerevisiae): implications for control of metabolic flux into the phenylpropanoid pathway J.Biol.Chem. ;279:2600-2607.

Ro DK, CJ Douglas 2004b Reconstitution of the entry point of plant phenylpropanoid metabolism in yeast (Saccharomyces cerevisiae): implications for control of metabolic flux into the phenylpropanoid pathway J.Biol.Chem. ;279:2600-2607.

Ro DK, N Mah, BE Ellis, CJ Douglas 2001 Functional characterization and subcellular localization of poplar (Populus trichocarpa x Populus deltoides) cinnamate 4-hydroxylase Plant Physiol. ;126:317-329.

Rodgers MW, A Zimmerlin, D Werck-Reichhart, GP Bolwell 1993 Microsomally associated heme proteins from French bean: characterization of the cytochrome P450 cinnamate-4-hydroxylase and two peroxidases Arch.Biochem.Biophys. ;304:74-80.

Rubery PH, DH Northcote 1968 Site of phenylalanine ammonia--lyase activity and synthesis of lignin during xylem differentiation Nature ;219:1230-1234.

Ruegger M, K Meyer, JC Cusumano, C Chapple 1999 Regulation of Ferulate-5-Hydroxylase Expression in Arabidopsis in the Context of Sinapate Ester Biosynthesis. Plant Physiology ;119:101-110.

Sato T, K Takabe, M Fujita 2004 Immunolocalization of phenylalanine ammonia-lyase and cinnamate-4-hydroxylase in differentiating xylem of poplar C.R.Biol. ;327:827-836.

Saunders JA, EE Conn, CH Lin, M Shimada 1977 Localization of Cinnamic Acid 4-Monooxygenase and the Membrane-bound Enzyme System for Dhurrin Biosynthesis in Sorghum Seedlings Plant Physiol. ;60:629-634.

Schalk M, Y Batard, A Seyer, S Nedelkina, F Durst, D Werck-Reichhart 1997 Design of fluorescent substrates and potent inhibitors of CYP73As, P450s that catalyze 4-hydroxylation of cinnamic acid in higher plants Biochemistry ;36:15253-15261.

Schalk M, F Cabello-Hurtado, MA Pierrel, R Atanossova, P Saindrenan, D Werck-Reichhart 1998 Piperonylic acid, a selective, mechanism-based inactivator of the trans-cinnamate 4-hydroxylase: A new tool to control the flux of metabolites in the phenylpropanoid pathway Plant Physiol. ;118:209-218.

Schilmiller AL, J Stout, JK Weng, J Humphreys, MO Ruegger, C Chapple 2009 Mutations in the cinnamate 4-hydroxylase gene impact metabolism, growth and development in Arabidopsis Plant J. ;60:771-782.

Sewalt V, W Ni, JW Blount, HG Jung, SA Masoud, PA Howles, C Lamb, RA Dixon 1997a Reduced Lignin Content and Altered Lignin Composition in Transgenic Tobacco Down-Regulated in Expression of L-Phenylalanine Ammonia-Lyase or Cinnamate 4-Hydroxylase Plant Physiol. ;115:41-50.

Sewalt V, W Ni, JW Blount, HG Jung, SA Masoud, PA Howles, C Lamb, RA Dixon 1997b Reduced Lignin Content and Altered Lignin Composition in Transgenic Tobacco Down-Regulated in Expression of L-Phenylalanine Ammonia-Lyase or Cinnamate 4-Hydroxylase Plant Physiol. ;115:41-50.

Song J, Z Wang 2011 RNAi-mediated suppression of the phenylalanine ammonia-lyase gene in Salvia miltiorrhiza causes abnormal phenotypes and a reduction in rosmarinic acid biosynthesis J.Plant Res. ;124:183-192.

Stewart JJ, T Akiyama, C Chapple, J Ralph, SD Mansfield 2009 The effects on lignin structure of overexpression of ferulate 5-hydroxylase in hybrid poplar Plant Physiol. ;150:621-635.

Teutsch HG, MP Hasenfratz, A Lesot, C Stoltz, JM Garnier, JM Jeltsch, F Durst, D Werck-Reichhart 1993 Isolation and sequence of a cDNA encoding the Jerusalem artichoke cinnamate 4-hydroxylase, a major plant cytochrome P450 involved in the general phenylpropanoid pathway Proc.Natl.Acad.Sci.U.S.A. ;90:4102-4106.

Tsai CJ, JL Popko, MR Mielke, WJ Hu, GK Podila, VL Chiang 1998 Suppression of O-methyltransferase gene by homologous sense transgene in quaking aspen causes red-brown wood phenotypes Plant Physiol. ;117:101-112.

Urban P, D Werck-Reichhart, HG Teutsch, F Durst, S Regnier, M Kazmaier, D Pompon 1994 Characterization of recombinant plant cinnamate 4-hydroxylase produced in yeast. Kinetic and spectral properties of the major plant P450 of the phenylpropanoid pathway Eur.J.Biochem. ;222:843-850.

Vance CP, RT Sherwood 1976 Regulation of lignin formation in reed canarygrass in relation to disease resistance Plant Physiol. ;57:915-919.

Voelker SL, B Lachenbruch, FC Meinzer, M Jourdes, C Ki, AM Patten, LB Davin, NG Lewis, GA Tuskan, L Gunter, SR Decker, MJ Selig, R Sykes, ME Himmel, P Kitin, O Shevchenko, SH Strauss 2010 Antisense down-regulation of 4CL expression alters lignification, tree growth, and saccharification potential of field-grown poplar Plant Physiol. ;154:874-886.

Voelker SL, B Lachenbruch, FC Meinzer, SH Strauss 2011 Reduced wood stiffness and strength, and altered stem form, in young antisense 4CL transgenic poplars with reduced lignin contents New Phytol. ;189:1096-1109.

VOELKER SL, B LACHENBRUCH, FC MEINZER, P KITIN, SH STRAUSS 2011; 2011 Transgenic poplars with reduced lignin show impaired xylem conductivity, growth efficiency and survival Plant, Cell Environ. ;34:655 <last_page> 668.

Voo KS, RW Whetten, DM O'Malley, RR Sederoff 1995 4-coumarate:coenzyme a ligase from loblolly pine xylem. Isolation, characterization, and complementary DNA cloning Plant Physiol. ;108:85-97.

Wagner A, L Donaldson, H Kim, L Phillips, H Flint, D Steward, K Torr, G Koch, U Schmitt, J Ralph 2009 Suppression of 4-coumarate-CoA ligase in the coniferous gymnosperm Pinus radiata Plant Physiol. ;149:370-383.

Weng JK, X Li, J Stout, C Chapple 2008 Independent origins of syringyl lignin in vascular plants Proc.Natl.Acad.Sci.U.S.A. ;105:7887-7892.

Weng JK, H Mo, C Chapple 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.

Werck-Reichhart D, Y Batard, G Kochs, A Lesot, F Durst 1993 Monospecific polyclonal antibodies directed against purified cinnamate 4-hydroxylase from Helianthus tuberosus. Immunopurification, immunoquantitation, and interspecies cross-reactivity Plant Physiol. ;102:1291-1298.

Xu B, LL Escamilla-Treviño, N Sathitsuksanoh, Z Shen, H Shen, Y Percival Zhang, RA Dixon, B Zhao 2011; 2011 Silencing of 4-coumarate:coenzyme A ligase in switchgrass leads to reduced lignin content and improved fermentable sugar yields for biofuel production New Phytol. ;192:611 <last_page> 625.

Yamauchi K, S Yasuda, K Hamada, Y Tsutsumi, K Fukushima 2003 Multiform biosynthetic pathway of syringyl lignin in angiosperms Planta ;216:496-501.

Yang J, F Chen, O Yu, RN Beachy 2011 Controlled silencing of 4-coumarate:CoA ligase alters lignocellulose composition without affecting stem growth Plant Physiol.Biochem. ;49:103-109.

Zhang XH, VL Chiang 1997 Molecular cloning of 4-coumarate:coenzyme A ligase in loblolly pine and the roles of this enzyme in the biosynthesis of lignin in compression wood Plant Physiol. ;113:65-74.

Zhong R, WH III, J Negrel, ZH Ye 1998 Dual methylation pathways in lignin biosynthesis Plant Cell ;10:2033-2046.

Chapter 2

Functional Redundancy of the Two 5-Hydroxylases in Monolignol Biosynthesis of *Populus trichocarpa*: LC-MS/MS Based Protein Quantification and Metabolic Flux Analysis

Published in Planta (2012) 10.1007/s00425-012-1663-5

Jack P. Wang, Christopher M. Shuford, Quanzi Li, Jina Song, Ying-Chung Lin, Ying-Hsuan Sun, Hsi-Chuan Chen, Cranos M. Williams, David C. Muddiman, Ronald R. Sederoff, Vincent L. Chiang

North Carolina State University, Raleigh, NC 27695

Summary

Flowering plants have syringyl and guaiacyl subunits in lignin in contrast to the guaiacyl lignin in gymnosperms. The biosynthesis of syringyl subunits is initiated by coniferaldehyde 5-hydroxylase (CAld5H). In Populus trichocarpa there are two closely related CAld5H enzymes (PtrCAld5H1 and PtrCAld5H2) associated with lignin biosynthesis during wood formation. We used yeast recombinant PtrCAld5H1 and PtrCAld5H2 proteins to carry out Michaelis-Menten and inhibition kinetics with LC-MS/MS based absolute protein quantification. CAld5H, a monooxygenase, requires a cytochrome P450 reductase (CPR) as an electron donor. We cloned and expressed three P. trichocarpa CPRs in yeast; and show that all are active with both CAld5Hs. The kinetic analysis shows both CAld5Hs have essentially the same biochemical functions. When both CAld5Hs are coexpressed in the same yeast membranes, the resulting enzyme activities are additive suggesting functional redundancy and independence of these two enzymes. Simulated reaction flux based on Michaelis-Menten kinetics and inhibition kinetics confirmed the redundancy and independence. Subcellular localization of both CAld5Hs as sGFP fusion proteins expressed in P. trichocarpa differentiating xylem protoplasts indicate that they are endoplasmic reticulum (ER) resident proteins. These results imply that during wood formation 5-hydroxylation in monolignol biosynthesis of P. trichocarpa requires the combined metabolic flux of these two CAld5Hs to maintain adequate biosynthesis of syringyl lignin. The combination of genetic analysis, absolute protein quantitation based enzyme kinetics, homologous CPR specificity, SNP characterization and ER localization provides a more

rigorous basis for a comprehensive systems understanding of 5-hydroxylation in lignin biosynthesis.

Introduction

Lignin is a heterogeneous phenolic polymer that imparts rigidity, strength and decay resistance to plants (Sarkanen and Ludwig 1971; Whetten and Sederoff 1995; Boerjan et al. 2003; Rogers et al. 2004, Ralph 2004, 2007). Lignin is derived from the coupling of phenylpropanoid monolignols, *p*-coumaryl, coniferyl and sinapyl alcohols to the respective *p*-hydroxyphenyl (H), guaiacyl (G) and sinapyl (G) subunits in lignin (Sarkanen and Ludwig 1971; Higuchi 1997). In angiosperms lignin is composed primarily of S and G subunits with low levels of H and other intermediates. Lignin in gymnosperms consists almost exclusively of G subunits (Sarkanen and Ludwig 1971).

Identification of lignin biosynthetic enzymes and their metabolites arose from tracer studies (Brown 1961; Brown and Neish 1955; Brown and Neish 1956; Wright et al. 1958; Brown and Neish 1959; Higuchi and Brown 1963; Higuchi and Brown 1963). Ten enzyme families in a metabolic grid regulate the biosynthesis of monolignols from L-phenylalanine through a principal pathway (Fig. 1) (Higuchi 1997; Shi et al. 2010). The regulatory mechanisms that underlie the evolutionary divergence of syringyl subunit biosynthesis from the guaiacyl-substituted phenolics are of substantial interest because of its potential role in the evolution of flowering plants. The 5-hydroxylation of guaiacyl intermediates for syringyl monolignol biosynthesis was long thought to occur through the 5-hydroxylation of ferulic acid (Grand

1984). More recent kinetic studies of P450 dependent monooxygenase activity revealed that the syringyl pathway stems from hydroxylation at the levels of the free aldehydes and alcohols (Osakabe et al. 1999; Humphreys et al. 1999). The enzyme activity is more appropriately described as coniferaldehyde-5-hydroxylase (Osakabe et al. 1999).

P. trichocarpa (Nisqually-1) is the woody plant with the most complete set of genomic and transcriptomic sequences and where lignin gene models are established (Tuskan et al. 2006; Shi et al. 2010). Through analysis of tissue specificity and transcript abundance, three *PtrCAld5H* genes were identified from the 45,555 gene models (JGI ID POPTR_0005s11950; POPTR_0007s13720; POPTR_0009s12640). *PtrCAld5H1* and *PtrCAld5H2* (91.4% protein sequence identity) are expressed specifically and abundantly in xylem (Shi et al, 2010), and they are located on two different linkage groups (JGI v2.2) and presumed to have arisen from the salicoid-specific genome duplication 65 Mya (Tuskan et al. 2006). *PtrCAld5H3* has very low transcript abundance in xylem, leaf, root and shoot tissues in *P. trichocarpa* under normal growth conditions. Nisqually-1 is known to have substantial heterozygosity (Tuskan et al. 2006). PtrCAld5Hs are homologous to the previously characterized aspen CAld5H (Osakabe et al. 1999). It is important to distinguish different gene models and potential allelic variation such as single nucleotide polymorphisms (SNPs).

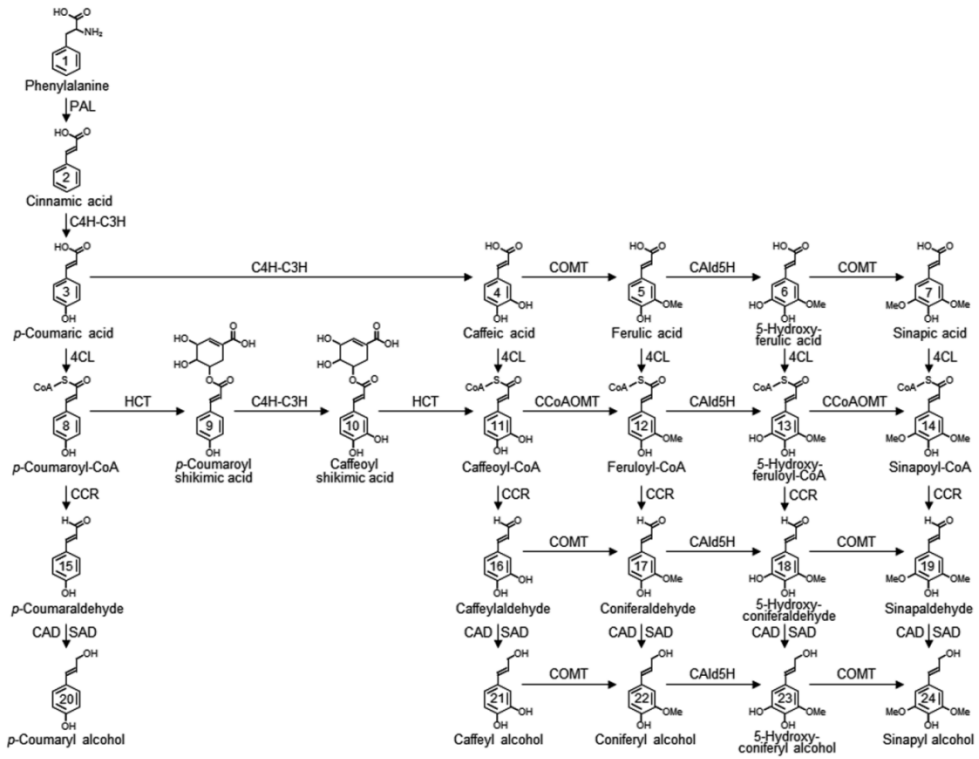

Figure 1 - The proposed monolignol biosynthetic pathway. Ten enzyme families (Phenylalanine ammonia-lyase (PAL), cinnamic acid 4-hydroxylase (C4H), 4-coumarate 3-hydroxylase (C3H), 4-coumarate: coenzyme A ligase (4CL), *p*-hydroxycinnamoyl-CoA:quinate shikimate *p*-hydroxycinnamoyltransferase (HCT), caffeoyl-CoA *O*-methyltransferase (CCoAOMT), coniferyl aldehyde 5-hydroxylase (CAld5H), O-methyltransferase (COMT), cinnamoyl-CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD)) mediate the formation of hydroxycinnamyl alcohols from phenylalanine

The extent to which these two PtrCAld5Hs have specific or redundant biochemical functions had not been determined. Enzyme activity analysis of 5-hydroxylases requires a cytochrome P450 reductase (CPR) as an electron donor (Osakabe et al. 1999). Previous studies of poplar 5-hydroxylases have used CPR from *Arabidopsis* (ATR1) (Osakabe et al. 1999). Whether heterologous combinations of CAld5Hs and CPRs provide a good representation of their homologous systems had not been established. Rigorous understanding of metabolic flux in monolignol biosynthesis requires both kinetic analysis and absolute protein quantification. Current methods of mass spectrometry enable quantitative estimation of specific absolute protein contents using stable isotope labeled surrogate peptides for protein cleavage and isotope dilution mass spectrometry (PC-IDMS) (Shuford et al. Submitted).

Here, we report the functional analysis of both PtrCAld5H1 and PtrCAld5H2 with a homologous CPR to evaluate the extent of specificity or functional redundancy. We have carried out a metabolic flux simulation based on this information to validate our results and provide a basis for further mathematical analysis of monolignol biosynthesis.

Materials and Methods

Plant materials and growth conditions

Clonal propagules of *P. trichocarpa* (Nisqually-1*)* were maintained in a greenhouse under 16-h light / 8-h dark photoperiod in soil mix consisting 1:1 potting mix to peat moss. Stem differentiating xylem (SDX) was collected from 7 months old trees as previously described (Shi et al. 2010), and used immediately for protein, RNA and protoplast isolation.

Vector construction for subcellular localization, cloning and expression of PtrCPRs and SNP variants of PtrCAld5Hs

PtrCAld5H1 and *PtrCAld5H2* full length coding regions were previously cloned (Shi et al. 2010). For the assembly of vectors for expressing recombinant PtrCAld5H1 and PtrCAld5H2 individually in yeast, primer sets PtrCAld5H1-F1/R1 and PtrCAld5H2-F1/R1 (Table 1) were used to amplify the coding regions of *PtrCAld5H1* and *PtrCAld5H2* respectively using PfuUltra high fidelity DNA polymerases (Stratagene). The PCR fragments were gel purified, cloned into the *pGEMT-easy* vector (Promega) and sequence confirmed. The *PtrCAld5H1* and *PtrCAld5H2* coding regions were then excised using *Not*I/*Sal*I and cloned into the *Not*I/*Sal*I site of the *pYeDP60M* (Chen et al. 2011) vector to yield *pYeDP60M-PtrCAld5H1* and *pYeDP60M-PtrCAld5H2*. For the co-expression of PtrCAld5H1 and PtrCAld5H2, the fragment *GAL10-CYC1P-PtrCAld5H1-CPR1T* from *pYeDP60M-PtrCAld5H1* was amplified using the primer set GCPC-F/R (Table 1), partially digested with *Nco*I and cloned into the *NcoI* site of *PtrCAld5H2-pYeDP60M*. For the

subcellular localization of PtrCAld5H1 and PtrCAld5H2 in protoplasts, coding regions were amplified using PtrCAld5H1-F2/R2 and PtrCAld5H2-F2/R2 (Table 1) and cloned into the *pGEMT-easy* vector. After sequence confirmation, the coding regions for *PtrCAld5H1* and *PtrCAld5H2* were excised using *Xba*I/*Sal*I and cloned into the *XbaI/SalI* site of the expression vector *pUC19-35S-sGFP* (Chen et al. 2011) to form *pUC19-35S-PtrCAld5H1:sGFP* and *pUC19-35S-PtrCAld5H2:sGFP*. The *sGFP* sequence was fused to the C-terminus of the 5-hydroxylase coding region to visualize the expression of PtrCAld5H1:sGFP and PtrCAld5H2:sGFP fusion proteins in protoplast cells.

To construct vectors for the expression of *P. trichocarpa* cytochrome P450 reductases in yeast, the full length cDNA of *PtrCPR1*, *PtrCPR2*, and *PtrCPR3* were amplified from SDX using primer sets, CPR1-F/CPR1-R, CPR2-F/CPR2-R, and CPR3-F/CPR3-R (Table 1), respectively, and cloned into *pGEMT*-easy vectors for sequencing. The cDNA fragments were excised from *pGEMT-easy* vectors, and inserted into *pYeDP110* (Urban et al. 1997) at *BamHI*, *KpnI*, and *SmaI*, respectively, resulting in the expression vectors *pYeDP110-PtrCPR1*, *pYeDP110-PtrCPR2*, and *pYeDP110-PtrCPR3*. These three expression vectors were linearized with *Not*I, purified with a QIAquick PCR purification kit (Qiagen), and used for yeast transformation.

Table 1 - Primers used for vector construction

Primer name	**Sequence**
PtrCAld5H1-F1	AGCGGCCGCTTAATAATGGATTCTCTTCTCCAATCT
PtrCAld5H1-R1	A GTCGACTTAAAATGGACAGACCACGCGCTTTCTCGG
PtrCAld5H2-F1	AGCGGCCGCTTAATAATGGATTCTCTTGTCCAATC
PtrCAld5H2-R1	AGTCGACTTAGAGAGGGCATAGCACACGCTTGCT
GCPC-F	ACCATGGAGGGCACAGTTAAC
GCPC-R	ACCATGGGGTGTGCCCGTAAAAGGTTCT
PtrCAld5H1-F2	ACTGTCTAGACATCTGCATCCATGGATTCTCTTCTCCAATC
PtrCAld5H1-R2	ACTGGTCGACAAATGGACAGACCACGCGCT
PtrCAld5H2-F2	ACTGTCTAGATAAACCATCCATGGATTCTCTTGTCCAATC
PtrCAld5H2-R2	ACTGGTCGACGAGAGGGCATAGCACACGCT
CPR1-F	CTCTAGAGGATCCATGAGTTCAGGTGGTTCAAATT
CPR1-SF1	GGCTTTAGCTGAAGAGGTCAA
CPR2-F	AGGTACCATGCAATCATCAAGCAGCTCG
CPR2-R	AGGTACCTTACCATACATCACGCAGATACCT
CPR3-F	CCCGGGATGGAGTCATCAAGCAGCTCG
CPR3-R	CCCGGGCTTCATGGCCTACACGAACG

P. trichocarpa xylem protoplast preparation, transfection and subcellular localization of the two PtrCAld5Hs

The bark was peeled from stem segments of *P. trichocarpa*, and xylem was carefully sliced into ~1mm thick strips and submerged in a freshly prepared digestive solution consisting 20mM MES (pH5.7), 0.5M mannitol, 20mM KCl, 1.5% (w/v) cellulase RS, 0.4% Macerozyme RS, 10mM $CaCl_2$, and 0.1% (v/v) BSA. The digestive mixture was covered in foil, and placed under vacuum for 30 minutes, and held at room temperature for 2 hours. The mixture was gently swirled for 1 minute, and then filtered through a 75µm nylon mesh, and washed twice after centrifugation at 200 x g for 3 minutes by discarding the supernatant and resuspending the protoplasts in W5 solution (5mM glucose, 2mM MES, pH 5.7, 154mM NaCl, 125mM $CaCl_2$, 5mM KCl) (Yoo et al. 2007). The isolated protoplasts were co-transfected with an mCherry-ER marker and either *pUC19-35S-PtrCAld5H1-sGFP or pUC19-35S-PtrCAld5H2-sGFP* using PEG (Yoo et al. 2007). Plasmid DNA for protoplast transfection was prepared with the QIAGEN MAXI-prep Kit, and the transfected protoplasts were observed at 16 hours post transfection using a confocal laser scanning microscope LSM 5 PASCAL (Carl Zeiss).

Yeast strains and recombinant protein expression

The expression vectors pYeDP110-PtrCPR1, pYeDP110-PtrCPR2, and pYeDP110-PtrCPR3 were separately transformed into Saccharomyces cerevisiae W(N) (WB303-1B) (MATa;ade2-1; his3-11, -15; leu2-3, -112; ura3-1; can^R; cyr+) using a LiAc/PEG method (Gietz et al. 2007) to generate WPT1, WPT2 and WPT3 respectively. The transformed yeast

were selected on SGI medium (Ralston et al. 2001), cultured and collected (Ralston et al. 2001). Genomic DNA was isolated using a Blood & Cell Culture DNA Mini Kit (Qiagen) and used for PCR to verify the positive transformants.

WPT1, WPT2 and WPT3 cells were transformed with PtrCAld5H1-pYeDP60, PtrCAld5H2-pYeDP60 and PtrCAld5H1+PtrCAld5H2-pYeDP60 as described by Gietz and Schiestl (2007). The transformed yeasts were cultured and induced for recombinant protein expression (Jiang and Morgan 2004) and the microsomal fractions were collected as described by Ralston et al. (2001) for enzymology and quantification.

Quantification of PtrCAld5H1 and PtrCAld5H2 by protein cleavage isotope dilution mass spectrometry (PC-IDMS)

All recombinant hydroxylases were quantified by PC-IDMS (Barr et al. 1996; Barnidge et al. 2004). The stable isotope-labeled (SIL) peptide standards for PtrCAld5H1 and PtrCAld5H2 were FLEPGVPDFK (aa 426-435) and FMKPGVPDFK (aa 427-436), respectively (Shuford et al. Submitted). The peptides were synthesized by the Mayo Clinic Proteomic Research Center (Rochester, MN) and contained [$^{13}C_5$,^{15}N]-valine as the heavy-labeled amino acid. Each sample was denatured in 4 M urea containing 50 mM dithiothreitol (56 °C, 30 min) and subsequently alkylated in 200 mM iodoacetamide (37 °C, 1 hr). Two hundred µg of yeast microsomal proteins were added to Amicon 10 kDa molecular weight cut off (MWCO) filters (Millipore), concentrated, and buffer exchanged 3 times with trypsin digestion buffer (50 mM Tris-HCl, 2 M urea, 10 mM $CaCl_2$, pH 8.0). Each centrifugation step was for 15

min at 14,000 × *g* at 20 °C. Subsequently, 40 µg of trypsin was added to the MWCO-filters in 90 µL of digestion buffer, immediately followed by 2 pmoles of each SIL peptide in 10 µL of 0.001% zwittergent 3-16 (Calbiochem). SIL peptide quantities were confirmed spectrophotometrically (Scopes 1974). The digestions proceeded for 8 hours at 37 °C, then were quenched by the addition of 100 µL of 1% formic acid containing 0.001% zwittergent 3-16. The peptides were then eluted from the MWCO-filters via centrifugation. The small volume retained in each filter was then diluted with 400 µL of the quenching solution and likewise eluted from the filter units. Finally, each sample was diluted to ~1 mL with quenching solution and analyzed by liquid chromatography coupled with selected reaction monitoring (LC-SRM). The entire digestion process was repeated in duplicate for each sample.

LC-SRM analysis of each sample was carried out using a nanoLC-2D system equipped with an AS1 autosampler and cHiPCL-Nanoflex system (Eksigent). Online desalting and nanoLC separations were carried out on 5 µL of sample using Nano cHiPLC Trap (200 µm × 0.5 mm) and analytical (75 µm × 15 cm) columns packed with ChromXP C18-CL (5 µm, 200 Å) particles. Peptides were loaded/desalted in a 5 % mobile phase B at 1.5 µL/min and subsequently eluted at 400 nL/min by ramping from 5% to 38.5% mobile phase B over the course of 20 min. Mobile phases A and B were comprised of water/acetonitrile/formic acid (98/2/0.2 and 2/98/0.2, respectively). The column eluent was ionized using a 20 µm I.D. PicoTip emitter and ESI potential of 1400 V. Scheduled SRM was performed on TSQ Vantage triple stage quadrupole mass spectrometer (Thermo Scientific) using a retention time window of 2.5 min and SRM duty cycle time of 1.5 sec. Other instrument parameters

were Q1 Peak Width (Full width at half maximum) of 0.7, collision gas pressure of 1.5 mTorr, and Chrom Filter Peak Width of 30 sec. The transitions utilized for each natural and SIL peptide can be found in Table 2. Each digest was analyzed in triplicate and the natural:SIL ratio (measured by total peak area) was utilized to quantify the amount of protein in each sample. All LC-SRM data processing was performed using Skyline v1.1.0.2905 (Maclean et al. 2010).

Table 2 - SRM Transitions for natural (NAT) and stable isotope labeled (SIL) peptides

Protein & Peptide	Precursor Ion	Product Ion	Precursor Ion *m/z*		Product Ion *m/z*		CE (Volts)
			NAT	SIL	NAT	SIL	
PtrCAld5H1	$[M+2H^+]^{2+}$	y_7^+	574.803	577.810	759.404	765.417	18
FLEPG<u>V</u>PDFK		y_4^+			506.261	506.261	
		b_3^+			390.202	390.202	
PtrCAld5H2	$[M+3H^+]^{3+}$	y_4^+	389.207	391.212	506.261	506.261	14
FMKPG<u>V</u>PDFK		y_2^+			294.181	294.181	
		y_4^{2+}			253.634	253.634	
		b_6^+			660.354	666.368	

CE, Collision Energy

5-Hydroxylation assays and kinetic analysis

5-Hydroxylation assays were conducted as described by Osakabe et al. (1999) with modifications. Assay buffer (50µM sodium citrate, pH 5.4, 1mM NADPH, 10mM glucose-6-phosphate, 2U glucose-6-phosphate dehydrogenase) was added to varying concentrations of ferulic acid (4µM to 500µM), coniferaldehyde (0.2µM to 10µM), coniferyl alcohol (0.2µM to 20µM) or a mixture of these substrates in a final volume of 75µl. Assays were initiated by the addition of 25ul of either yeast microsomes expressing recombinant 5-hydroxylase or *P. trichocarpa* SDX microsomes and were incubated at 30°C for 5 minutes before being terminated by the addition of 50µl of termination buffer (4% v/v trichloroacetic acid, 50% v/v acetonitrile). The reaction mixtures were clarified by centrifugation at 20,000 x *g* for 60 minutes and were analyzed through an Agilent HPLC 1200 system. The substrates and products were separated by an Agilent ZORBAX SB-C3 5µm, 4.6 x 150mm column, using a gradient method (Solvent A, 10mM formic acid in water; Solvent B, 10mM formic acid in acetonitrile; 10% to 15% B in 10 minutes, 25% to 30% B in 15 minutes; 30% to 100% B in 3 minutes; flow rate 1ml/min) and quantified through an Agilent Diode-Array Detector SL based on authentic compounds (ferulic acid and 5-hydroxyferulic acid, λ_{max} = 325nm; coniferaldehyde and 5-hydroxyconiferaldehyde, λ_{max} = 335nm; coniferyl alcohol and 5-hydroxyconiferyl alcohol, λ_{max} = 260nm).

To determe the enzyme kinetic constants V_{max} and K_m, Lineweaver-Burk (double reciprocal) plots were used to derive linear regression of the Michaelis-Menten equation (Equation 1). For evaluating inhibition kinetics, different concentrations (0.2 ~ 6 µM) of coniferaldehyde were added to kinetic assays of ferulic acid and coniferyl alcohol. Reaction rates were

analyzed by Lineweaver-Burke plots to determine the modes of inhibition. If addition of inhibitor increases the K_m, the inhibition is competitive. Whereas if V_{max} decreases, the inhibition is non-competitive, and if both K_m and V_{max} decreases proportionally, the mode of inhibition is uncompetitive. Mixed inhibition is when multiple modes of inhibition (competitive, non-competitive, uncompetitive) is present, more commonly in combination of competitive and uncompetitive. To derive the inhibition constants K_{IC} (competitive) and K_{IU} (uncompetitive), we first obtained K'_m and V'_{max}, the apparent K_m and V_{max} of the inhibited enzyme at each inhibitor concentration, and plotted K'_m/V'_{max} against [I] and $1/V'_{max}$ against [I] to derive K_{IC} and K_{IU} respectively according to Equation 2 and Equation 3 (based on the Michaelis-Menten equation).

Equation 1 Michaelis-Menten equation

$$v = \frac{V_{max} \cdot [S]}{[S] + K_m}$$

Equation 2 Competitive inhibition equation

$$v = \frac{V_{max} \cdot [S]}{[S] + K_m \cdot (1 + \frac{[I]}{K_{IC}})}$$

Equation 3 Un-competitive inhibition equation

$$v = \frac{V_{\max} \cdot [S]}{[S] \cdot (1 + \frac{[I]}{K_{IU}}) + K_m}$$

Equation 4 Reaction flux for single enzyme in the presence of multiple substrates

$$v = \frac{V_{\max} \cdot [S]}{[S] \cdot (1 + \frac{[I]}{K_{IU}}) + K_m \cdot (1 + \frac{[I]}{K_{IC}})}$$

Equation 5 Reaction flux for multiple enzymes in the presence of a single substrate

$$v = \frac{V_{\max(PtrCAld5H1)} \cdot [S]}{[S] + K_{m(PtrCAld5H1)}} + \frac{V_{\max(PtrCAld5H2)} \cdot [S]}{[S] + K_{m(PtrCAld5H2)}}$$

Equation 6 Reaction flux for multiple enzymes in the presence of multiple substrates

$$v = \frac{V_{\max(PtrCAld5H1)} \cdot [S]}{[S] \cdot (1 + \frac{[I]}{K_{IU}}) + K_{m(PtrCAld5H1)} \cdot (1 + \frac{[I]}{K_{IC}})} + \frac{V_{\max(PtrCAld5H2)} \cdot [S]}{[S] \cdot (1 + \frac{[I]}{K_{IU}}) + K_{m(PtrCAld5H2)} \cdot (1 + \frac{[I]}{K_{IC}})}$$

Results and Discussion

P. trichocarpa SDX microsomal proteins mediate 5-hydroxylation of ferulic acid, coniferaldehyde and coniferyl alcohol

To investigate the 5-hydroxylation pathway in monolignol biosynthesis of *P. trichocarpa*, we first evaluated the 5-hydroxylation activities in the soluble and microsomal portions of *P. trichocarpa* SDX using coniferaldehyde as substrate. The conversion of coniferaldehyde to 5-hydroxyconiferaldehyde was 7.4-fold higher in the microsomal fraction, as expected. Using SDX microsomes, we assayed the 5-hydroxylation activities of individual substrates, ferulic acid, coniferaldehyde and coniferyl alcohol as well as the activities with all three substrates. 5-Hydroxylation activities were detected for all three substrates when assayed individually. However, when the substrates were present together in the reaction, the activity of ferulic acid and coniferyl alcohol 5-hydroxylation is greatly inhibited, while coniferaldehyde 5-hydroxylation is maintained (Fig. 2). These enzyme activities from SDX microsomes indicate that the two identified 5-hydroxylases are likely to be ER resident proteins. In addition, PtrCAld5H1 and PtrCAld5H2 annotated polypeptide sequences indicated that the N-terminal regions of both peptides resemble a hydrophobic trans-membrane like domain, also suggesting that they are membrane bound proteins. Therefore we conducted subcellular localization of these 5-hydroxylases.

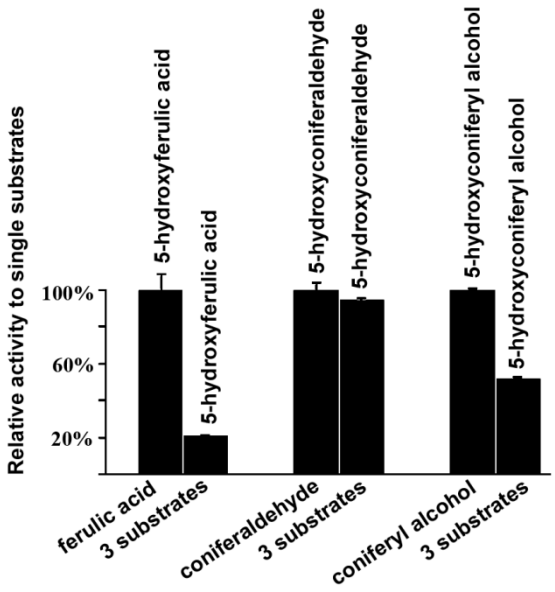

Figure 2 - Xylem microsomes assayed for 5-hydroxylation activity with ferulic acid, coniferaldehyde and coniferyl alcohol. The error bars represent the standard error of three technical replicates

PtrCAld5H1 and PtrCAld5H2 are localized on the endoplasmic reticulum (ER) of P. trichocarpa SDX cells

We fused enhanced green fluorescence protein (*sGFP*) to the C-terminus of full-length *PtrCAld5H1* and *PtrCAld5H2* in *pUC19* based constructs driven by the CaMV 35S promoter. The same construct containing only the sGFP was included as a control. We co-transfected the fusion constructs and an mCherry ER marker plasmid (ABRC stock no. CD3-959) into *P. trichocarpa* SDX protoplasts via PEG-mediated transfection. Using confocal microscopy, strong fluorescence was observed for both PtrCAld5H1 and PtrCAld5H2 samples that follows a distinctive reticulating pattern resembling fluorescence emitted from the ER marker. The fluorescence pattern of the sGFP only control was uniformly distributed in the protoplasts and is distinctively different from the fluorescence pattern of the ER marker (Fig. 3). These results suggest that PtrCAld5H1 and PtrCAld5H2 are likely ER localized proteins in the SDX of *P. trichocarpa*.

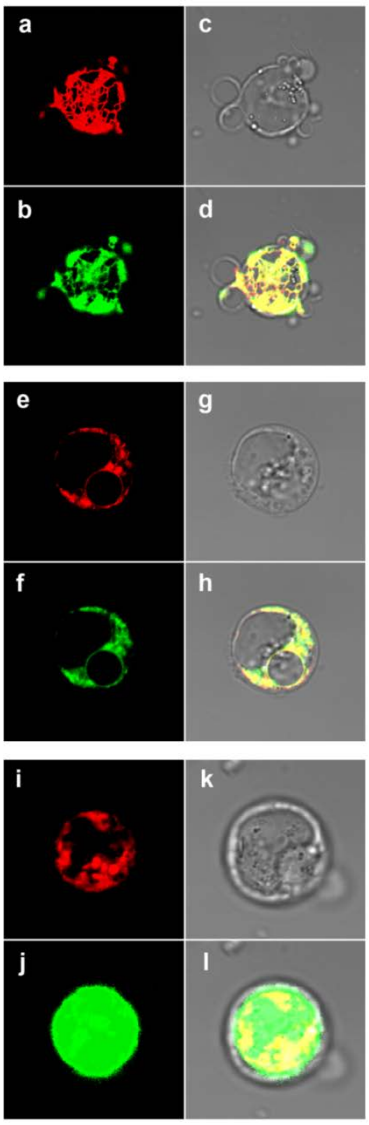

Figure 3 - Sub-cellular localization in xylem protoplasts of *P. trichocarpa* (a) ER marker:mCherry (b) PtrCAld5H1:sGFP (c) White light (d) Merged (e) ER marker:mCherry (f) PtrCAld5H2:sGFP (g) White light (h) Merged (i) ER marker:mCherry (j) sGFP only control (k) White light (l) Merged

Cloning and coexpression of P. trichocarpa CPR and CAld5Hs in S. cerevisiae

Previous studies of recombinant aspen CAld5H kinetics (Osakabe et al. 1999) were based on a yeast expression system co-expressing *Arabidopsis* CPR (ATR1). However, the efficiency of such heterologous P450/CPR combinations cannot be reliably evaluated. To better understand the kinetic properties of recombinant PtrCAld5H1 and PtrCAld5H2, we engineered *S. cerevisiae* to co-express *P. trichocarpa* CPRs so that both the P450s and CPRs are homologous. Four CPR gene models were identified in the genome of *P. trichocarpa* (JGI v2.2), *PtrCPR1* (POPTR_0005s19710.1), *PtrCPR2* (POPTR_0006s18280.1), *PtrCPR3* (POPTR_0018s09980.1), and *PtrCPR4* (POPTR_0002s10680.1). All four *PtrCPRs* are expressed in SDX based on our microarray data (Not shown), with the highest expression by *PtrCPR2* and lowest expression by *PtrCPR4*. The gene model of *PtrCPR4* is truncated. Our attempt to amplify the coding region and 5'UTR of *PtrCRP4* by 5'RACE was not successful, thus *PtrCPR4* was not studied further. The *P. trichocarpa* genome database shows four alternative gene models for *PtrCPR1* with different lengths. Specific primers were designed to distinguish the four models. RT-PCR showed that *PtrCPR1* cDNA was amplified only from the primer pair specific to the longest gene model. The full length cDNAs of *PtrCPR1*, *PtrCRP2* and *PtrCPR3* were cloned and sequence confirmed. Vectors *pYeDP110-PtrCPR1*, *pYeDP110-PtrCPR1*, and *pYeDP110-PtrCPR3* were prepared and used to generate recombinant yeast strains expressing individual PtrCPRs (WPT1, WPT2 and WPT3 expressing PtrCPR1, PtrCPR2 and PtrCPR3 respectively). PCR using yeast genomic DNA verified that the yeast *CPR* had been replaced by *PtrCPRs*.

Functional specificity comparisons of PtrCAld5H1 and PtrCAld5H2 with PtrCPRs in vitro showed that both proteins are coniferaldehyde 5-hydroxylases

PtrCAld5H1 and PtrCAld5H2 recombinant proteins were expressed individually in WPT1, WPT2 and WPT3 and assayed for the 5-hydroxylation of coniferaldehyde. Both hydroxylases are able to utilize all three PtrCPRs as electron donors, however coexpression with PtrCPR3 resulted in the highest hydroxylation activities (Fig. 4). We therefore used PtrCPR3 for all subsequent kinetic analyses. Further work is needed to characterize the specificity of the different CPRs in vivo. Yeast microsomal proteins containing either recombinant PtrCAld5H1 or PtrCAld5H2 were assayed with substrates ferulic acid, coniferaldehyde and coniferyl alcohol individually and in mixtures of all three. Consistent with the observations using microsomal fractions from SDX of *P. trichocarpa*, when ferulic acid, coniferyl alcohol and coniferaldehyde were assayed individually, all three substrates were 5-hydroxylated by PtrCAld5H1 and PtrCAld5H2. However, in the presence of all substrates, ferulic acid and coniferyl alcohol 5-hydroxylation is greatly inhibited, while coniferaldehyde 5-hydroxylation is unchanged (Fig. 5). These results show that the recombinant hydroxylases retain the same substrate specificity and inhibition as in the plant extracts. The consistency between the substrate specificity and inhibition patterns of recombinant PtrCAld5Hs and plant extracts suggest that PtrCAld5H1 and PtrCAld5H2 are likely the predominant 5-hydroxylases in monolignol biosynthesis of *P. trichocarpa*. Several additional phenylpropanoid intermediates (cinnamic acid, *p*-coumaric acid, *p*-coumaraldehyde, *p*-coumaryl alcohol, *p*-coumaroyl shikimic acid) were tested as potential substrates for recombinant PtrCAld5H1 and PtrCAld5H2. However, for both hydroxylases,

5-hydroxylation activities were detected only for ferulic acid, coniferaldehyde and coniferyl alcohol.

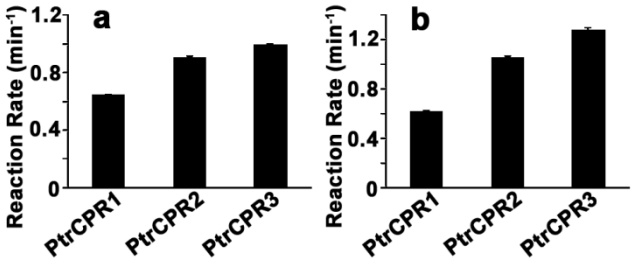

Figure 4 - 5-Hydroxylation rates of PtrCAld5H1 (a) and PtrCAld5H2 (b) with different PtrCPRs. The error bars represent the standard error of three technical replicates

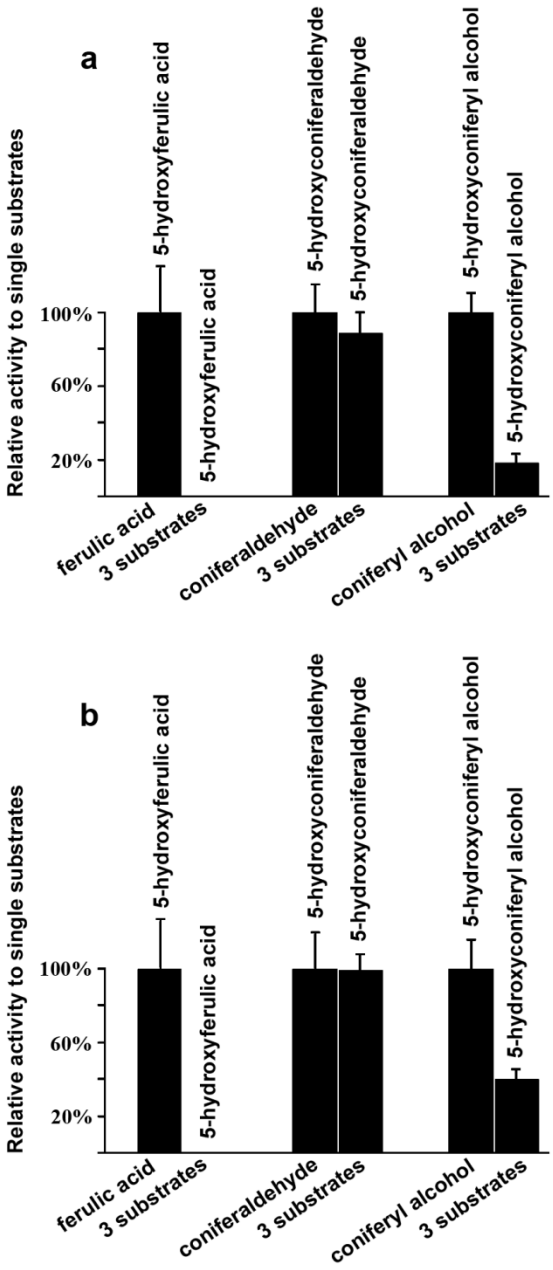

Figure 5 - Recombinant PtrCAld5H1 (a) and PtrCAld5H2 (b) assayed for 5-hydroxylation activities with ferulic acid, coniferaldehyde and coniferyl alcohol. The error bars represent the standard error of three technical replicates

Single-nucleotide polymorphisms (SNPs) in PtrCAld5H1 and PtrCAld5H2 do not alter enzyme activities in vitro

To evaluate the significance of SNPs in regulating 5-hydroxylation, we characterized all variants of *PtrCAld5H1* and *PtrCAld5H2* SNPs in Nisqually-1 by assaying for their activities using coniferaldehyde. SNP variants were identified by full length sequencing of *PtrCAld5H1* and *PtrCAld5H2* coding regions from multiple clones isolated from *P. trichocarpa* SDX. A single C/A SNP was identified for *PtrCAld5H1* that results in an amino acid change from a Proline to a Threonine (Table 3). Three SNPs were identified in the coding region of *PtrCAld5H2*, two were degenerate, and one resulted in an amino acid change from a Methionine to a Leucine (Table 3). The variants were expressed in WPT3 as previously described and assayed using coniferaldehyde as substrate. The 5-hydroxylation activities of all SNP variants of PtrCAld5H1 and PtrCAld5H2 are essentially identical (Fig. 6). PtrCAld5H1 (Proline variant) and PtrCAld5H2 (Methionine variant) were used for all subsequent kinetic analyses (Table 3).

Table 3 - SNP variation in *PtrCAld5H1* and *PtrCAld5H2* in *P. trichocarpa* (Nisqually-1)

	PtrCAld5H1
Proline variant	TTTTGATAAACTA **C** CATTCTTCAAA
Threonine variant	TTTTGATAAACTA **A** CATTCTTCAAA
	PtrCAld5H2
Methionine variant	CTC **T** TACATGAAACATC **C** GAA … GTTT **A** TGAA
Leucine variant	CTC **C** TACATGAAACATC **T** GAA … GTTT **T** TGAA

Underlined nucleotides reflect the SNP variations

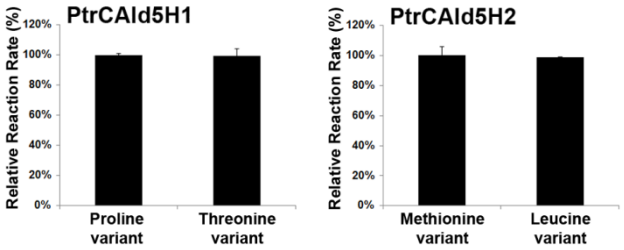

Figure 6 - The 5-hydroxylation activities of PtrCAld5H1 and PtrCAld5H2 SNP variants with coniferaldehyde as the substrate. The error bars represent the standard error of three technical replicates

LC-MS/MS based absolute quantification of recombinant PtrCAld5H1 and PtrCAld5H2

To define enzyme kinetic constants requires knowledge of the absolute enzyme quantities. Previously, quantification of recombinant CAld5Hs for kinetic analyses were based either on mg of total microsomal protein or total P450s derived from the reduced-CO difference spectrum (Osakabe et al. 1999). However, recombinant CAld5Hs constitute only a fraction of total microsomal protein. Such quantification is inadequate for determining precise enzyme catalytic rates for metabolic flux analysis. PC-IDMS can provide specific and sensitive estimates of protein quantities from complex mixtures (Barnidge et al. 2004; Barr et al. 1996; Dass et al. 1991; Gerber et al. 2003). In this method, a protein's concentration is determined by quantifying one of its unique peptides that has been produced by targeted proteolysis. The proteolytic peptide is quantified by comparing its analytical signal (*e.g.*, peak area) to that of its stable isotope-labeled (SIL) analogue, which is spiked into the sample at a known concentration to serve as an internal standard. Because the native/natural (NAT) peptide and its SIL standard possess the same physio-chemical properties (excepting molecular weight), the relative intensity of their signals is equivalent to their relative concentrations.

To provide quantification by PC-IDMS, the targeted peptide must be unique to the protein of interest. In the case of PtrCAld5H1 and PtrCAld5H2, the tryptic peptides (FLEPGVPDFK for PtrCAld5H1and FMKPGVPDFK for PtrCAld5H2) used for quantification (Table 2) were compared against all *S. cerevisiae* peptides annotated in the NCBI protein database *via* Skyline(MacLean et al. 2010) and were found to be unique. We utilized LC-SRM to ensure

specific detection of the NAT and SIL peptides. SRM is a targeted tandem MS (MS/MS) detection mode, which provides a high-degree of detection specificity based on the unique combination of intact peptide ion mass and its fragment ions masses (Lange et al. 2008). To increase confidence in the detected peptides identity, we monitored multiple fragment ions by SRM for each NAT and SIL peptide (Fig. 7, Table 2) and also compared the relative contribution of each fragment ion to the total peak area for each measurement (Kushnir et al. 2005; Sauvage et al. 2008). In all cases, the information afforded by SRM detection was used in combination with retention time (*i.e.*, co-elution of the NAT/SIL pairs) to confirm the identity of the detected peptides and ensure specific quantification of the hydroxylases.

Given that variation between digestions can lead to variation in the protein quantities measured by PC-IDMS (Barnidge et al. 2004), we elected to perform all digests in duplicate and utilize the average for the estimates (Table 4). Good reproducibility was obtained between the two digests for each sample. In all cases, the coefficients of variation (CV) were less than 7%.

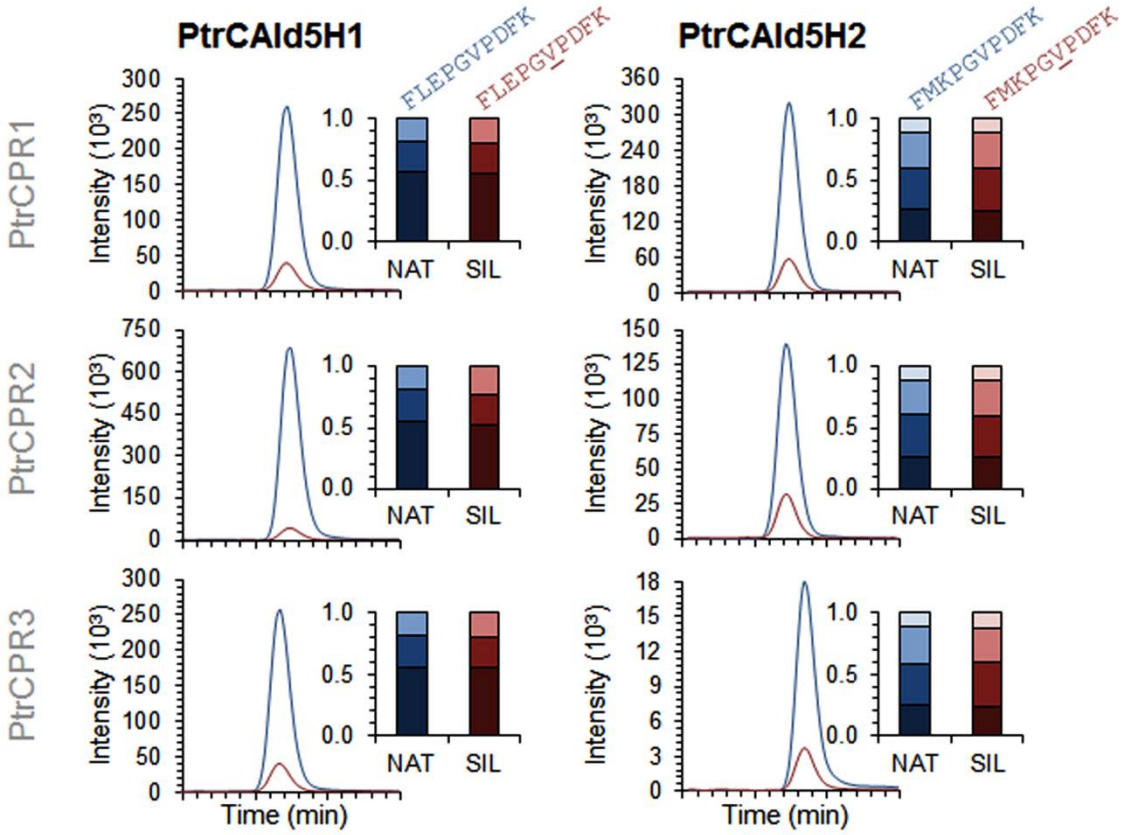

Figure 7 - Shown are representative SRM chromatograms for the target tryptic peptides and the corresponding SIL peptide standards of CAld5H1 (left) and CAld5H2 (right) when expressed recombinantly in combination with the three *P. trichocarpa* CPRs. The SRM traces are a summation of all specific SRM transitions monitored for the specific peptide. The inset bar graphs show the relative contribution of each transition (*i.e.*, each product ion) to the summed measurements of each peptide, which should be constant for a given peptide if the transitions are specific to that peptide species.

Table 4 - Quantification of recombinant PtrCAld5H1 and PtrCAld5H2

		PtrCAld5H1 – WPT1	PtrCAld5H1 – WPT2	PtrCAld5H1 – WPT3	PtrCAld5H2 – WPT1	PtrCAld5H2 – WPT2	PtrCAld5H2 – WPT3
Digest 1	Rep. 1	71.8	160	58.9	62.1	42.9	55.4
	Rep. 2	71.3	161	59.7	65.3	44.1	54.4
	Rep. 3	69.8	156	62.0	64.3	42.3	53.0
	Mean	71.0	159	60.2	63.9	43.1	54.3
Digest 2	Rep. 1	66.0	151	67.0	58.1	43.9	49.3
	Rep. 2	66.6	155	65.6	56.6	44.0	49.6
	Rep. 3	68.8	155	66.0	59.3	44.5	50.5
	Mean	67.1	153	66.2	58.0	44.2	49.8
	Mean	69.1	156	63.2	61.0	43.6	52.1
	SD	2.7	4	4.2	4.2	0.7	3.1
	CV	3.9%	2.7%	6.7%	6.9%	1.7%	6.0%
	$CI_{95\%}$	±3.8	±6	±5.9	±5.8	±1	±4.4

*All units are in fmoles per microgram of microsomal protein (fmol/µg)

Michaelis-Menten kinetics and inhibition kinetics of recombinant PtrCAld5H1 and PtrCAld5H2

We carried out Michaelis-Menten kinetic assays of PtrCAld5H1 and PtrCAld5H2 using ferulic acid, coniferaldehyde and coniferyl alcohol as substrates (Table 5). Consistent with our previous findings (Osakabe 1999), the specificity constants (k_{cat}/K_m) confirmed that for both enzymes, coniferaldehyde and coniferyl alcohol are the preferred substrates, whereas ferulic acid 5-hydroxylation is not likely to be a major pathway in monolignol biosynthesis. These results suggest that, these two 5-hydroxylases have near indistinguishable catalytic activities and are functionally redundant. To understand the above observed inhibition of 5-hydroxylation in xylem microsomes (Fig. 2) we next tested inhibition kinetics of PtrCAld5H1 and PtrCAld5H2 reactions using recombinant proteins. In addition, inhibition kinetics are important for simulating the metabolic flux distributions when the 5-hydroxylases are present in a mixture of substrates.

Table 5 Michaelis-Menten kinetics of PtrCAld5H1 and PtrCAld5H2

	PtrCAld5H1		
	Coniferaldehyde	Coniferyl alcohol	Ferulic acid
K_m	0.0544 ± 0.0014	0.0686 ± 0.0016	19.3 ± 1.4
k_{cat}	1.23 ± 0.02	2.12 ± 0.03	$7.7x10^{-4} \pm 1.3x10^{-5}$

	PtrCAld5H2		
	Coniferaldehyde	Coniferyl alcohol	Ferulic acid
K_m	0.0453 ± 0.0004	0.0526 ± 0.0044	5.2 ± 0.2
k_{cat}	1.008 ± 0.002	1.72 ± 0.05	$1.1x10^{-3} \pm 1.6x10^{-5}$

* K_m is in μM and k_{cat} is in min^{-1}

For PtrCAld5H1, coniferaldehyde has a mixed mode of inhibition for both ferulic acid and coniferyl alcohol 5-hydroxylation, diagnosed by an increased K_m and decreased V_{max}. The K_{IC} by conferaldehyde on ferulic acid 5-hydroxylation is 0.4 ± 0.03 µM, ~50 fold lower than the K_m of ferulic acid (19.3 ± 1.4 µM), indicating strong inhibition of ferulic acid 5-hydroxylation (high affinity of enzyme-inhibitor complex) (Fig. 8). The K_{IC} of coniferaldehyde inhibiting the conversion of coniferyl alcohol is 0.8 ± 0.03 µM, which is greater than the K_m of coniferyl alcohol 5-hydroxylation (0.0686 ± 0.0016 µM) (Fig. 8), indicating a weak inhibition. These results are consistent with reactions using recombinant PtrCAld5H1 (Fig. 5), when ferulic acid, coniferyl alcohol and coniferaldehyde are present in equimolar concentrations, conversion of ferulic acid is drastically inhibited whereas the inhibition of coniferyl alcohol 5-hydroxylation is only partial.

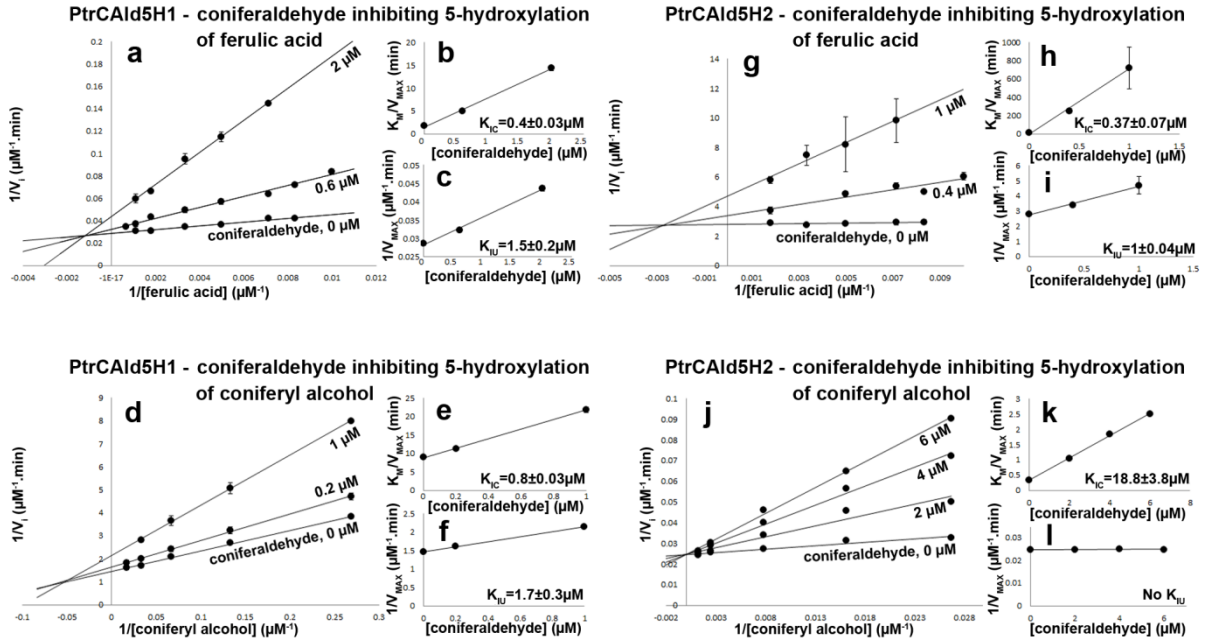

Figure 8 - Inhibition kinetics of PtrCAld5H1 and PtrCAld5H2 (a) Double reciprocal plots showing coniferaldehyde inhibiting 5-hydroxylation of ferulic acid for PtrCAld5H1 (b) K_m/V_{max} versus [coniferaldehyde] for calculating K_{IC} (c) $1/V_{max}$ versus [coniferaldehyde] for calculating K_{IU} (d) Double reciprocal plots showing coniferaldehyde inhibiting 5-hydroxylation of coniferyl alcohol for PtrCAld5H1 (e) K_m/V_{max} versus [coniferaldehyde] for calculating $K_{IC.}$ (f) $1/V_{max}$ versus [coniferaldehyde] for calculating K_{IU} (g) Double reciprocal plots showing coniferaldehyde inhibiting 5-hydroxylation of ferulic acid for PtrCAld5H2 (h) K_m/V_{max} versus [coniferaldehyde] for calculating K_{IC} (i) $1/V_{max}$ versus [coniferaldehyde] for calculating K_{IU} (j) Double reciprocal plots showing coniferaldehyde inhibiting 5-hydroxylation of coniferyl alcohol for PtrCAld5H2 (k) K_m/V_{max} versus [coniferaldehyde] for calculating K_{IC} (l) $1/V_{max}$ versus [coniferaldehyde] for calculating K_{IU}. The error bars represent the standard error of three technical replicates

PtrCAld5H2 mediated conversion of ferulic acid was also effectively inhibited by coniferaldehyde, through a mixed mode inhibition with a K_{IC} value of 0.37 ± 0.07 μM and a K_{IU} value of 1.0 ± 0.04 μM (Fig. 8). Inhibition of coniferyl alcohol 5-hydroxylation by coniferaldehyde was purely competitive, as the $1/V'_{max}$ against [I] plot showed a horizontal line, indicating that V'_{max} does not change with [I] (Fig. 8). The K_{IC} value is much greater than the K_m, suggesting that coniferyl alcohol is less effectively inhibited by coniferaldehyde. These data are consistent with reactions using recombinant PtrCAld5H2. Conversion of ferulic acid is completely inhibited when all three substrates are present, whereas coniferyl alcohol 5-hydroxylation is only mildly inhibited.

Kinetic studies of the recombinant PtrCAld5Hs showed that they are redundant in substrate specificity. Analysis of inhibition kinetics further revealed similarities in their activities, as both hydroxylases effectively inhibit the conversion of ferulic acid.

Simulation of metabolic flux with single and mixed substrates and experiments validate functional redundancy and additivity of PtrCAld5H1 and PtrCAld5H2

Through analysis of single substrate kinetics and inhibition kinetics, we showed that when the recombinant PtrCAld5Hs are characterized individually, both enzymes exhibited very similar catalytic properties. To elucidate whether the 5-hydroxylases are truly functional redundant, it is important to determine when both enzymes are present, whether they react independently or cooperatively. If the 5-hydroxylases react independently of each other, the total activities when both enzymes are present will be equal to the sum of the activities of

PtrCAld5H1 and PtrCAld5H2. If they react cooperatively, the total 5-hydroxylation activity will be modified. To evaluate functional independence, we first derived metabolic flux simulations using the substrate and inhibition kinetic constants (Table 5, Fig. 8) to predict the flux distribution when the two PtrCAld5Hs are coexpressed.

The reaction flux of individual 5-hydroxylases for each substrate is represented by the Michaelis-Menten rate equation (Equation 1). In the case of multiple substrates, the apparent K'_m and V'_{max} is substituted into Equation 1 according to the inhibition constants K_{IC} and K_{IU} (Equation 2 and Equation 3) to give Equation 4, the reaction flux for individual PtrCAld5Hs when multiple substrates are present (Fig. 9). If PtrCAld5H1 and PtrCAld5H2 are functionally independent, the total 5-hydroxylation activity will be the sum of the activities of PtrCAld5H1 and PtrCAld5h2 (additive) for single substrates (Equation 5) (Fig. 9) and mixed substrates (Equation 6) (Fig. 9).

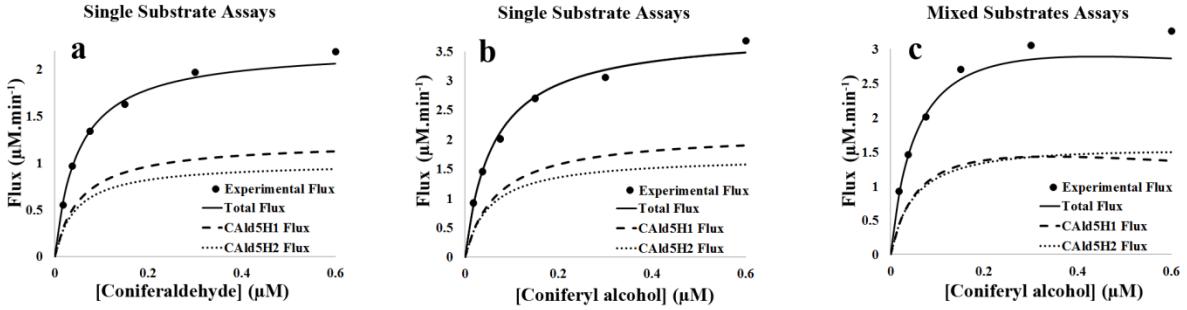

Figure 9 - Simulation and experimental validation of metabolic fluxes with single and mixed substrates for PtrCAld5H1, PtrCAld5H2 and the additive flux of both enzymes. The dashed lines and dotted lines correspond to the simulated flux for PtrCAld5H1and PtrCAld5H2 hydroxylations respectively. The solid lines represent the simulated additive flux of PtrCAld5H1 and PtrCAld5H2. The round dots are the experimental data confirming simulation (a) The simulation of coniferaldehyde 5-hydroxylation fluxes with coniferaldehyde as a single substrate (b) The simulation of coniferyl alcohol 5-hydroxylation fluxes with coniferyl alcohol as a single substrate (c) The simulation of coniferyl alcohol 5-hydroxylation fluxes for mixed substrates (coniferaldehyde and coniferyl alcohol)

To validate the simulation and confirm whether the two 5-hydroxylases act independently or cooperatively, we coexpressed PtrCAld5H1 and PtrCAld5H2 in WPT3. We assayed the coexpressed recombinant proteins with the individual substrates (ferulic acid, coniferyl alcohol and coniferaldehyde) and mixed substrates. Consistant with simulation, the experimentally determined reaction fluxes were additive supporting functional independence for both individual substrate and mixed substrate conditions (Fig. 9).

Functional annotation of putative 5-hydroxylases in plants

PtrCAld5Hs belong to the CYP84A family of P450s, defined by a putative ferulate-5-hydroxylase (F5H) from *Arabidopsis*. We screened available plant genomes in Phytozome v8.0 (*www.phytozome.net*) for *CYP84A* annotations. Single putative *F5H*s were annotated in the genomes of *Morchella esculenta, Medicago truncatula, Proteus vulgaris, Cucumis sativus, Prunus persica, Capsella rubella, Brassica rapa, Thellungiella halophila, Camellia sinensis, Citrus clementina, Mimulus guttatus, Sorghum bicolor, Setaria italica, Oryza sativa, Brachypodium distachyon, Selaginella moellendorffii.* As well as *P. trichocarpa*, multiple putative 5-hydroxylases were annotated in the genomes of *Ricinus communis, Linum usitatissimum, Glycine max, A. thaliana, Arabidopsis lyrata, Carica papaya, Eucalyptus grandis, Vitis vinifera, Aquilegia coerulea and Zea mays.* The *Arabidopsis* genome contains two 5-hydroxylases (*AtF5H1* and *AtF5H2*). The biochemical function of AtF5H1 has been tested (Humphreys et al. 1999). However, AtF5H2 failed to compensate for the loss of AtF5H1 in the *fah1* mutant, whether AtF5H2 is involved in monolignol

biosynthesis remains unknown (Raes et al. 2003). The biochemical functions of the remaining CYP84As in the plant species mentioned above have not yet been determined.

Conclusions

We evaluated the various properties of the paralogous PtrCAld5Hs. Through reactions using xylem microsomes and recombinant PtrCAld5Hs, we showed that substrate specificities and inhibitions are conserved. PtrCAld5H1 and PtrCAld5H2 are likely to be the two predominant 5-hydroxylases in monolignol biosynthesis of P. trichocarpa. We then showed that the enzymatic activities of PtrCAld5Hs are not altered by SNP variants in *P. trichocarpa* Nisqually-1. When expressed as sGFP fusion proteins in SDX protoplasts of P. trichocarpa, the PtrCAld5Hs co-localized with an ER marker, indicating that they are both ER resident proteins. And lastly, through simulation and experimental validation, we demonstrated that PtrCAld5H1 and PtrCAld5H2 are functionally redundant. This functional redundancy implies that the biosynthesis of sinapyl alcohol in *P. trichocarpa* is directly regulated by the combined activities of the two PtrCAld5Hs (Fig. 9), and the suppression of either one's activity will result in a reduction in the overall flux of 5-hydroxylation, lowering the syringyl subunit biosynthesis in lignin. To test this, we are generating transgenic *P. trichocarpa* with gene specific knock-down RNAi constructs targeting individual *PtrCAld5H1* and *PtrCAld5H2* transcripts as well as the down-regulation of both hydroxylases.

The combination of genetic analysis, absolute protein quantitation based enzyme kinetics, homologous CPR specificity, SNP characterization and ER localization provided the most intensive characterization of 5-hydroxylation activities in monolignol biosynthesis. This work provides a rigorous basis for further studies on the developments of molecular and mathematical models for comprehensive systems understanding of lignin biosynthesis.

REFERENCES

Barnidge DR, Goodmanson MK, Klee GG, Muddiman DC (2004) Absolute quantification of the model biomarker prostate-specific antigen in serum by LC-Ms/MS using protein cleavage and isotope dilution mass spectrometry J. Proteome Res 3:644-652

Barr JR, Maggio VL, Patterson Jr DG, Cooper GR, Henderson LO, Turner WE, Smith SJ, Hannon WH, Needham LL, Sampson EJ (1996) Isotope dilution--mass spectrometric quantification of specific proteins: model application with apolipoprotein A-I. Clin Chem 42:1676-1682

Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Annu Rev Plant Biol 54:519-546

Brown SA (1961) Chemistry of Lignification: Biochemical research on lignins is yielding clues to the structure and formation of these complex polymers. Science 134:305-313

Brown SA, Neish AC (1956) Studies of lignin biosynthesis using isotopic carbon. V. Comparative studies on different plant species. Can J Biochem Physiol 34:769-778

Brown SA, Neish AC (1955) Studies of lignin biosynthesis using isotopic carbon. IV. Formation from some aromatic monomers. Can J Biochem Physiol 33:948-962

Brown SA, Neish AC (1959) Studies of lignin biosynthesis by use of isotopic carbon. VIII. isolation of radioactive hydrogenolysis products of lignin. J Am Chem Soc 81:2419-2424

Brown SA, Neish AC (1955) Shikimic Acid as a Precursor in Lignin Biosynthesis. Nature 175:688-689

Chen HC, Shuford CM, Liu J, Muddiman DC, Sederoff RR, Chiang VL (2011) Membrane protein complexes catalyze both 4- and 3-hydroxylation of cinnamic acid derivatives in monolignol biosynthesis. Proc Natl Acad Sci 108:21253-21258

Cocking EC (1960) A Method for the Isolation of Plant Protoplasts and Vacuoles. Nature 187:962-963

Wright D, Brown SA, NEISH AC (1958) Studies of lignin biosynthesis using isotopic carbon. VI. Formation of the side chain of the phenylpropane monomer. Can J Biochem Physiol 36:1037-1045

Gietz RD, Schiestl RH (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG. Method Nat Protoc 2:31-34

Grand C (1984) Ferulic acid 5-hydroxylase: a new cytochrome P-450-dependent enzyme from higher plant microsomes involved in lignin synthesis. FEBS Lett 169:7-11

Higuchi T, Brown SA (1963) Studies of lignin biosynthesis using isotopic carbon. XII. The biosynthesis and metabolism of spinapic acid. Can J Biochem Physiol 41:613-620

Higuchi T, Brown SA (1963) Studies of lignin biosynthesis using isotopic carbon: XIII. The phenylpropanoid system in lignification. Biochemistry and Cell Biology 41:621- 628

Higuchi T (1997) Biochemistry and Molecular Biology of Wood. Springer, New York, pp 131-233

Humphreys JM, Hemm MR, Chapple C (1999) New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. Proc Natl Acad Sci 96:10045-10050

Jiang H, Morgan JA (2004) Optimization of an in vivo plant P450 monooxygenase system inSaccharomyces cerevisiae. Biotechnol Bioeng 85:130-137

Lynch M (2000) The Evolutionary Fate and Consequences of Duplicate Genes. Science 290:1151-1155

Maclean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, Maccoss MJ (2010) Skyline: An Open Source Document Editor for Creating and Analyzing Targeted Proteomics Experiments. Bioinformatics 26:966-968

Nowak MA, Boerlijst MC, Cooke J, Smith JM (1997) Evolution of genetic redundancy. Nature 388:167-171

Osakabe K, Tsao CC, Li L, Popko JL, Umezawa T, Carraway DT, Smeltzer RH, Joshi CP, Chiang VL (1999) Coniferyl aldehyde 5-hydroxylation and methylation direct syringyl lignin biosynthesis in angiosperms. Proc Natl Acad Sci 96:8955-8960

Ralph J (2004) Lignins: natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. Phytochem Rev 3:29-60

Ralph J, Brunow G, Boerjan W (2007) Lignins. In Encyclopedia of Life Sciences. John Wiley Sons, Chichester, UK

Ralston L, Kwon ST, Schoenbeck M, Ralston J, Schenk DJ, Coates RM, Chappell J (2001) Cloning, heterologous expression, and functional characterization of 5-epi-aristolochene-1,3-dihydroxylase from tobacco (Nicotiana tabacum). Arch Biochem Biophys 393:222-235

Rogers LA, Campbell MM (2004) The genetic control of lignin deposition during plant growth and development. New Phytol 164:17-30

Sarkanen KV, Ludwig CH (1971) Lignins: Occurance. Formation, Structure and Reactions. John Wiley and Sons, New York

Scopes RK (1974) Measurement of protein by spectrophotometry at 205 nm. Anal Biochem 59:277-282

Shi R, Sun YH, Li Q, Heber S, Sederoff R, Chiang VL (2009) Towards a systems approach for lignin biosynthesis in Populus trichocarpa: Transcript abundance and specificity of the monolignol biosynthetic genes. Plant Cell Physiol 51:144-163

Shuford CM, Li Q, Sun YH, Chen HC, Wang JP, Shi R, Sederoff RR, Chiang VL, Muddiman DC (2012) Comprehensive Quantification of Monolignol-pathway Enzymes in *Populus trichocarpa* by Protein Cleavage Isotope Dilution Mass Spectrometry. Submitted

Song J, Lu S, Chen ZZ, Lourenco R, Chiang VL (2006) Genetic transformation of Populus trichocarpa genotype Nisqually-1: a functional genomic tool for woody plants. Plant Cell Physiol 47:1582-1589

Tuskan GA et al. (2006) The genome of black cottonwood, Populus trichocarpa (Torr. & Gray) Science 313:1596-1604

Urban P, Mignotte C, Kazmaier M, Delorme F, Pompon D (1997) Cloning, Yeast Expression, and Characterization of the Coupling of Two Distantly Related Arabidopsis thaliana ADPH-Cytochrome P450 Reductases with P450 CYP73A5. J Biol Chem 272:19176-19186

Whetten R, Sederoff R (1995) Lignin biosynthesis. Plant Cell 7:1001-1003

Yoo SD, Cho YH, Sheen J (2007) Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. Nat Protoc 2:1565-1572.

Chapter 3

Kinetic Modeling of the Monolignol Biosynthesis Pathway in *Populus trichocarpa*: Steady-state, Perturbations and Sensitivity Studies

Manuscript to be submitted to the journal PLoS Computational Biology

Jack Peng-Yu Wang, Punith Naik, Liu Jie, Hsi-Chuan Chen, Rui Shi, Christopher M. Shuford, Quanzi Li, Kevin Lin, Joel Ducoste, David C. Muddiman, Ronald Sederoff, Vincent Chiang

North Carolina State University, Raleigh, NC 27695

Introduction

Lignin is an abundant aromatic polymer that is present as an entangled covalent matrix with cellulose and hemi-celluloses in the thickened secondary cell walls of vascular plants (Sarkanen and Ludwig 1971; Whetten and Sederoff 1995; Boerjan et al. 2003; Rogers et al. 2004, Ralph 2004, 2007). Due to its heterogeneity, lignin is the primary recalcitrant in the utilization of lignocellulosic biomass for biofuel and paper production (Sarkanen 1976, Chiang 2002, Ragauskas et al. 2006, Chen and Dixon 2007). For delignification, lignin is often removed in an alkaline solution at elevated temperature and pressure. This energy intensive step enables the cellulosic fibers to be accessible to subsequent bleaching or enzyme hydrolysis, and accounts for the high cost of pulp and biofuel production from lignocellulosic biomass.

In Angiosperms such as *Populus trichocarpa*, lignin is polymerized from phenylpropanoid monomers consisting mostly of coniferyl and sinapyl alcohols, which are polymerized to guaiacyl and syringyl units in lignin respectively. Research towards comprehending the biosynthesis of these monomeric hydroxycinnamyl alcohols has been extensive and elaborate (Brown 1961; Brown and Neish 1955; Brown and Neish 1956; Wright et al. 1958; Brown and Neish 1959; Higuchi and Brown 1963; Higuchi and Brown 1963). Ten enzyme families in a metabolic grid convert phenylalanine to monolignols (Higuchi 1997, Shi et al. 2010).

The first step of the phenylpropanoid pathway, the deamination of phenylalanine to cinnamic acid is catalyzed by phenylalanine ammonia-lyase (PAL). Cinnamic acid is then hydroxylated to 4-coumaric acid by cinnamate 4-hydroxylase (C4H), followed by CoA

thioester formation by 4-coumarate:coenzyme A ligase. A co-operative interaction of coumarate 3-hydroxylase (C3H) and hydroxycinnamoyl-CoA shikimate hydroxycinnamoyl transferase (HCT), 3-hydroxylates 4-coumaroyl-CoA to caffeoyl-CoA. Caffeoyl coenzyme A O-methyltransferase (CCoAOMT) methylates caffeoyl-CoA at the 3' position to give feruloyl-CoA, which is reduced to coniferaldehyde by cinnamoyl-CoA reductase (CCR). Coniferaldehyde 5-hydroxylase and caffeate O-methyltransferase hydroxylate and methylate coniferaldehyde sequentially at the 5' position to give sinapaldehyde. The hydroxycinnamaldehydes are reduced to hydroxycinnamyl alcohols by cinnamyl alcohol dehydrogenase (CAD). The resulting supply and ratio of the hydroxycinnamyl alcohols are the controlling factors that determine lignin composition and structure (Frederik et al. 2010).

The evaluation of the ratio and metabolic flux of hydroxycinnamyl alcohols for lignification necessitate knowledge of the flux dynamics in the monolignol biosynthesis pathway. However, the complexity of the phenylpropanoid/monolignol pathway and their regulations renders an intuitive analysis of its metabolism a difficult task (Schallau and Junker 2010). Mathematical modeling approaches are an appropriate method to rationalize and reconstruct these complex networks. Various modeling strategies have been developed. Basic constraint optimization based models such as elementary mode analysis, minimal metabolic behaviors and flux balance analysis have been widely applied for the computation of stationary fluxes in plant metabolic networks including monolignol biosynthesis (Lee et al. 2010, 2011). The advantages of these basic models lies in the simplicity of their assembly because

knowledge of the stoichiometry alone is sufficient. However, as a result of this, the output of these models are to a large degree hypothetical, as the simulation relies heavily on plausible but hardly provable biological principles and constraints that are thought to govern flux distributions. Consequently, these models' ability to illustrate the biological relevance underlying regulatory mechanisms is largely limited to comparing mass-action based kinetic models (Hoppe et al. 2007). Alternatively, kinetic modeling is assembled upon a complete set of differential equations that defines fully the mechanisms of every enzymatic reaction in a biological pathway. Hence it is the most detailed and complex mathematical description of a metabolic network and constitutes a significant branch in the growing field of systems biology (Schallau and Junker 2010).

Methods

Plant materials and growth conditions

Clonal propagules of *P. trichocarpa* (Nisqually-1*)* were maintained in a greenhouse under 16-h light / 8-h dark photoperiod in a soil mix consisting of 1:1 potting mix to peat moss. Stem differentiating xylem (SDX) was collected from 7 months old trees as previously described (Shi et al. 2010) and used for LC-MS/MS based protein quantification.

Quantification of monolignol biosynthesis proteins from SDX of *P. trichocarpa* by protein cleavage isotope dilution mass spectrometry (PC-IDMS)

Equal weights of SDX from 3 *P. trichocarpa* trees were collected and pooled. This pooled sample was analyzed for the quantification of all 21 monolignol proteins as described in Shuford et al. 2012.

Recombinant proteins production for kinetic studies

Full length coding regions of all *P. trichocarpa* putative monolignol biosynthetic pathway proteins were previously cloned and sequenced (Shi et al. 2010). Recombinant proteins were expressed and purified as described by Shuford et al. (2012).

Chemical and biochemical synthesis of monolignol precursors for enzymatic reactions

Phenylalanine, cinnamic acid, *p*-coumaric acid, caffeic acid, ferulic acid, sinapic acid, coniferaldehyde, coniferyl alcohol, sinapaldehyde and sinapyl alcohol were purchased from Sigma Aldrich (St. Louis MO, USA). For all compounds excepting the alcohols, product ion spectra were acquired for the $[M+H^+]^{1+}$ molecular ion of each synthetic standard using a TSQ Quantum triple-stage quadruple mass spectrometer (Thermo Scientific, San Jose, CA). For the alcohols, product ion spectra were acquired for the $[M–OH^-]^{1+}$ molecular ion. Collision energy of 10 eV and 1.5 mtorr of argon were used in acquiring all spectra, except for the acyl-CoA derivatives which utilized collision energy of 30 eV.

Synthesis of 5-hydroxyferulic acid:

To a stirred mixture of 5-hydroxyvanillin (50 mg, 1 mmol) and malonic acid (75 mg), were added successively pyridine (1 ml), piperidine (5 μl), and aniline (10 μl), and the resulting solution was stirred at 25 °C for 1 week. After acidification with 1 N HCl to pH 2, the resulting mixture was extracted with 10 ml of ethyl acetate, then purified over a silica column, yielding 63 mg 5-hydroxyferulic acid. MS *m/z* (%) 211 ($[M+H^+]^+$, 15), 193 (100), 169 (1), 160 (18), 133 (8).

p-Coumaraldehyde, *p*-coumaryl alcohol, caffeylaldehyde, caffeyl alcohol, 5-hydroxyconiferaldehyde, and 5-hydroxyconiferyl alcohol were chemically synthesized in our lab as described (Chen et al, 2001). Briefly, 4-acetoxy-benzaldehyde (1 mmol) and (1, 3-dioxolan-2-yl-methyl)-triphenylphosphosphonium bromide (1 mmol) were dissolved in CH_2Cl_2 (20 ml). Solid K_2CO_3 (1 mmol) and 18-crown-6 (0.01 mmol) were added. The reaction mixture was kept at room temperature for 6 h, and filtered to separate the organic phase from the solid phase. Aqueous HCl (10%, 10 ml) was added to the organic portion and stirred the mixture at 25 °C for 4 h. Then the mixture was diluted with 10 ml H_2O and extracted three times with CH_2Cl_2. The combined organic layers were washed with saturated $NaHCO_3$ and saturated aqueous NaCl solutions successively, dried over $MgSO_4$ under vacuum. A silica gel column was used for purification. The 4-acetoxycoumaraldehyde (0.1 g) was dissolved in 20 ml of 0.2 M KOH in 95% EtOH, and stirred for 2 h under N_2. The solvent was removed under vacuum, and the mixture diluted with 10 ml H_2O and extracted with ethyl acetate (10 ml ×3). The combined organic layers were dried over Na_2SO_4 and evaporated to give *p*-coumaraldehyde (61 mg, 41% yield). *p*-Coumaryl alcohol was

synthesized from 4-acetoxycoumaraldehyde according to the procedure described by Daubresse et al. (1994). Briefly, 4-acetoxycoumaraldehyde (0.1 g) was dissolved in MeOH along with KH_2PO_4. $NaBH_4$ (50 mg) was slowly added at 0 °C. Cold water was added and MeOH was removed under vacuum and the mixture was extracted with CH_2Cl_2, washed with H_2O, dried and evaporated to give a 4-acetyloxy-2-propen-1-ol (90 mg). 4-Acetyloxy-2-propen-1-ol (50 mg) was dissolved in 20 ml of 0.2 M KOH in 95% EtOH, and stirred for 2 h under N2. The solvent was removed under vacuum, and the mixture diluted with 10 ml H_2O and extracted with ethyl acetate (20 ml ×3). The combined organic layers were dried over Na_2SO_4 and purified by silica gel column to give *p*-coumaryl alcohol (62 mg, 41% yield). Caffeoyl aldehyde and alcohol, 5-hydroxyconiferaldehyde, and 5-hydroxyconiferyl alcohol were synthesized using the same procedure with overall yield of 40–53%. p-Coumaraldehyde: MS m/z (%) 149 ([M+H+]+, 100), 131 (32), 121 (7), 107 (3), 103 (7), 93 (1), 91 (1), 77 (1), 55 (4). Caffeylaldehyde: MS m/z (%) 165 ([M+H+]+, 45), 147 (100), 119 (14), 91 (6), 55 (1). 5-Hydroxyconiferaldehyde: MS m/z (%) 195 ([M+H+]+, 85), 177 (100), 167 (3), 163 (11), 149 (6), 145 (2), 135 (1), 131 (5), 121 (1), 117 (1), 107 (1), 106 (1), 103 (5), 55 (1). *p*-Coumaryl alcohol: MS m/z (%) 133 [M–OH–]1+, 100), 115 (1), 105 (21), 103 (2), 79 (2), 77 (1). Caffeyl alcohol: MS m/z (%) 149 [M–OH–]1+, 100), 131 (96), 103 (31), 77 (2), 73 (1). 5-Hydroxyconiferyl alcohol: MS *m/z* (%) 179 $[M–OH^-]^{1+}$, 70), 164 (1), 161 (4), 147 (100), 133 (2), 119 (9), 91 (4).

p-Coumaroyl-CoA, caffeoyl-CoA and feruloyl-CoA were enzymatically synthesized from each acid (Beuerle and Pichersky 2002). Purified *P. trichocarpa* 4-coumarate: CoA ligase-3 (Ptr4CL3) (Shuford et al. 2012) recombinant protein from *Escherichia coli* was used to

biochemically synthesize *p*-coumaroyl-CoA and caffeoyl-CoA. Briefly, 6 mg acid, 4 mg coenzyme A hydrate (CoA), and 14 mg ATP were dissolved in a total volume of 40 ml of 50 mM Tris-HCl (pH 7.5) buffer containing 2.5 mM $MgCl_2$. 0.3 Milligrams of purified protein was added to the mixture to start the reaction. After 30 min at 37 °C, 1.6 g ammonium acetate was added to stop the reaction. The resulting mixture was purified by using an SPE cartridge (Chromabond C_{18} ec, Macherey-Nagel). Yields of 1.7 mg *p*-coumaroyl-CoA, 2.3 mg caffeoyl-CoA and 1.5 mg feruloyl-CoA were obtained, which represented 36%, 48% and 36% yield respectively, based on CoA used in the reaction. The purity and identity of all synthesized products were confirmed by tandem MS. Product ion spectra were acquired directly for the $[M+H^+]^+$ molecular ion of each compound on a TSQ Quantum Triple Quadrupole mass spectrometer (Thermo Scientific) at a collision energy of 10 eV and 1.5 mtorr of argon. *p*-Coumaroyl-CoA: MS *m/z* (%) 914 ($[M+H^+]^+$, 56.5), 768 (8.7), 505, (9.9), 428 (40.0), 407 (100), 341 (6.3), 305 (21.1), 261 (10.9). Caffeoyl-CoA: MS *m/z* (%) 930 ($[M+H^+]^+$, 60.2), 768 (20.5), 521 (8.8), 428 (43.3), 423 (100), 410 (7.0), 341 (11.5), 321 (20.6), 261 (14.8). Feruloyl-CoA: MS *m/z* (%) 944 ($[M+H^+]^+$, 56.0), 768 (7.0), 535 (11.7), 437 (100), 428 (48.8), 410 (7.4), 335 (19.4), 261 (13.3), 177 (8.3). 5-Hydroxyferuloyl-CoA: MS *m/z* (%) 960 ($[M+H^+]^+$, 5), 768 (3), 551 (3), 453 (100), 428 (41), 410 (3), 351 (19), 341 (8), 261 (19), 193 (8). Sinapoyl-CoA: MS *m/z* (%) 974 ($[M+H^+]^+$, 7), 768 (5), 467 (100), 428 (40), 365 (21), 341 (6), 261 (17), 207 (14).

Enzymatic synthesis of *p*-coumaroyl shikimic acid and caffeoyl shikimic acid used 6 mg *p*-coumaroyl-CoA, or 6 mg caffeoyl-CoA and 2 mg shikimic acid in a volume of 20 ml of potassium phosphate buffer (pH 7). The reaction was started by the addition of purified

PtrHCT6 (Shuford et al. 2012) (0.5 mg). After incubating for 20 min at 30 °C, the product was extracted three times with 20 ml of ethyl acetate. The organic layer was recovered, then dried over Na_2SO_4 and evaporated to give 1.5 mg *p*-coumaroyl shikimic acid (65% yield) or 1.2 mg caffeoyl shikimic acid (54% yield). *p*-Coumaroyl shikimic acid: MS *m/z* (%) 321 ($[M+H^+]^+$, 18.2), 303 (3.5), 165 (5.6), 147 (100), 139 (1.6), 119 (2.3) . Caffeoyl shikimic acid: MS *m/z* (%) 337 ($[M+H^+]^+$, 20.2), 319 (3.0), 181 (4.1), 163 (100), 145 (2.8), 145 (2.8), 139 (1.4), 117 (1.0).

Kinetic analysis of *P. trichocarpa* monolignol pathway enzymes

Kinetic assays for all recombinant proteins were carried out using the optimal assay conditions described by Liu et al. 2012. To ensure that accurate kinetic parameters were obtained, the recombinant proteins were assayed immediately post expression and purification to minimize loss of activity due to protein degradation.

The substrates and products of the enzyme assays were separated by an Agilent ZORBAX SB-C3 5µm, 4.6 x 150mm column. Analysis of hydroxycinnaic acids, hydroxycinnamyl aldehydes and hydroxycinnamyl alcohols were done using a gradient method (Solvent A, 10mM formic acid in water; Solvent B, 10mM formic acid in acetonitrile; 10% to 20% B in 10 minutes, 20% to 100% B in 5 minutes; flow rate 1ml/min). Analysis of reactions involving hydroxycinnamyl CoAs were done using a gradient method (Solvent A, 5mM ammonium acetate, pH 5.6 in water; Solvent B, 2/97.8/0.2 water/acetonitrile/acetic acid; 8% to 10% B in 3 minutes, 10% to 30% B in 5 minutes, 30% to 100% in 5 minutes; flow rate

1ml/min). The metabolites were quantified through an Agilent Diode-Array Detector SL based on authentic compounds.

To determine the optimal recombinant protein concentration for initial-rate kinetic analysis, we made serial dilutions of the recombinant proteins and the substrates. The lowest substrate concentrations were assayed against different concentrations of the recombinant proteins. Graphs of percentage substrate consumption against the recombinant protein concentrations were plotted. These plots ensured that the observed activities follow Michaelis-like rate behaviors (Linear dependence of reaction rate on the concentration of active enzyme) and met the criterion for pseudo-first-order approximations (<10% [S] consumption).

To determine the rate parameters K_m, V_{max}, k_{cat} and k_{cat}/K_m that define the initial-rate behavior of the enzymes, Lineweaver-Burk (double reciprocal) plots were used to derive linear regressions of the Michaelis-Menten equation (Equation 1). The K_m, or the Michaelis constant provides information about the enzyme's affinity for its substrate. V_{max} and k_{cat} describes the maximal velocity or catalytic turnover number of the enzyme, where $k_{cat}=V_{max}/[E_T]$. k_{cat}/K_m is the "Specificity constant" indicates the efficiency of substrate conversion by an enzyme.

For evaluating inhibition kinetics, different concentrations of inhibitors were added to the kinetic assays. Reaction rates were analyzed by Lineweaver-Burke plots to determine the modes of inhibition. If addition of inhibitor increases the K_m, the inhibition is competitive. Whereas if V_{max} decreases, the inhibition is non-competitive, and if both K_m and V_{max} decreases proportionally, the mode of inhibition is uncompetitive. Mixed inhibition is when

multiple modes of inhibition (competitive, non-competitive, and uncompetitive) are present, more commonly in combination of competitive and uncompetitive. To derive the inhibition constants K_{IC} (competitive) and K_{IU} (uncompetitive), we first obtained K'_m and V'_{max}, the apparent K_m and V_{max} of the inhibited enzyme at each inhibitor concentration, and plotted K'_m/V'_{max} against [I] and $1/V'_{max}$ against [I] to derive K_{IC} and K_{IU} respectively according to Equation 2 and Equation 3 (based on the Michaelis-Menten equation).

Equation 1 Reciprocal of the Michaelis-Menten equation

$$\frac{1}{v} = \frac{K_m}{V_{max}} \cdot \frac{1}{[S]} + \frac{1}{V_{max}}$$

Equation 2 Competitive inhibition equation

$$v = \frac{V_{max} \cdot [S]}{[S] + K_m \cdot (1 + \frac{[I]}{K_{IC}})}$$

Equation 3 Un-competitive inhibition equation

$$v = \frac{V_{\max} \cdot [S]}{[S] \cdot (1 + \frac{[I]}{K_{IU}}) + K_m}$$

Modeling Approach

Progress in cell and molecular biology has resulted in identification of the many control mechanisms that affect cell function. The most important step in developing a predictive model to describe a biological process is experimentally obtaining the parameters needed to develop a model. Because of the complexity associated with biological processes, obtaining these parameters is difficult and in most cases impossible. Several models have been developed, which try to model cell mechanisms using the stoichiometry of the network (Klipp and Liebermiester, 2006). One such method is the flux balance approach, which is an optimization method in which constraints were placed on the flux and define an objective function, which in some cases is maximization of biomass. Although these models provide valuable information about the mechanisms of interest, those models would not be accurate because of the great many underlying assumptions used to simplify the model owing to insufficient experimental data. In this paper we try to develop a kinetic model that uses the law of mass action and Michaelis-Menten kinetics to understand the process of lignin biosynthesis. This kinetic model is the most detailed to date and provides a mathematical description of the complex monolignol biosynthetic process.

Kinetic models are described in a continuous deterministic form using the rate equations for the concentration of metabolites and a conservation of mass approach (Klipp and Liebermiester, 2006). A metabolic network can be easily translated into mathematical form, where the rate of change of substrate concentration of interest $S_{i,}$ as a function of time is given by the rate at which the enzymes form the metabolite minus the rate at which the enzymes use this particular substrate to form some other metabolite (Schallau and Junker 2010). As shown in the equation below V_{in} and V_{out} are the rate of formation and degradation (or conversion) respectively.

$$\frac{dS_i}{dt} = V_{in} - V_{out}$$

The rate of formation and degradation of the metabolite can be represented by using the classical Michaelis Menten kinetics (Figure 1). This method is followed for all the metabolites in the network. This process results in a system of ordinary differential equations (ODE's), with each equation explaining the rate of change of concentration of the substrate of interest. The systems of ordinary equations were then solved by using MATLAB© R2011a and ODE solver ODE23tb which is usually used for moderately stiff functions (Shampine 1994). On solving the system of equations the steady state concentration for all the metabolites are obtained and the flux resulting from the metabolites are subsequently calculated. Here steady state means that there is no observed change in the concentration of the metabolites over time. The next step in the simulation process is to

understand the effect of perturbation of enzymes on the steady state flux and eventually on the lignin structure.

The table below shows how the differential equations for the rate change of various metabolites were set up. V_0 is the input flux and V_{38} and V_{40} are the exit fluxes.

To model the lignin biosynthesis pathway in *P. trichocarpa,* a constraint was set to fix the ratio of S to G at 2 because that is the experimentally determined value. The other constraint was that the input flux V_0 is equal to the sum of exit fluxes V_{38} and V_{40}. The analysis of the network was carried out by knocking down the individual enzyme families with respect to their wildtype to understand the effect of down regulation on the lignin structure.

Table 1 – List of ordinary differential equations corresponding to the 24 metabolites that interact in the lignin biosynthesis pathway.

Substrate	Mass Balance Equation
Phenyalanine	$\frac{dS1}{dt} = V_o - V_1$
Cinnamic Acid	$\frac{dS2}{dt} = V_1 - V_2$
p-Coumaric Acid	$\frac{dS3}{dt} = V_2 - V_3 - V_7$
Caffeic Acid	$\frac{dS4}{dt} = V_3 - V_4 - V_8$
Ferulic Acid	$\frac{dS5}{dt} = V_4 - V_5 - V_9$
5-Hydroxy Ferulic Acid	$\frac{dS6}{dt} = V_5 - V_6 - V_{10}$
p-Coumaryl-CoA	$\frac{dS8}{dt} = V_7 - V_{12} - V_{18}$
p-Coumaryl Shikimic Acid	$\frac{dS9}{dt} = V_{12} - V_{13}$
Caffeoyl Shikimic Acid	$\frac{dS10}{dt} = V_{13} - V_{14}$
Caffeoyl-CoA	$\frac{dS11}{dt} = V_8 + V_{14} - V_{15} - V_{19}$
Feruloyl CoA	$\frac{dS12}{dt} = V_9 + V_{15} - V_{16} - V_{20}$
p-Coumaraldehyde	$\frac{dS15}{dt} = V_{18} - V_{23} - V_{27}$
Caffeylaldehyde	$\frac{dS16}{dt} = V_{19} + V_{23} - V_{24} - V_{28}$
Coniferaldehyde	$\frac{dS17}{dt} = V_{20} + V_{24} - V_{25} - V_{29}$
5-Hydroxy-Coniferaldehyde	$\frac{dS18}{dt} = V_{21} + V_{25} - V_{26} - V_{30}$
Sinapaldehyde	$\frac{dS19}{dt} = V_{22} + V_{26} - V_{31}$
p-Coumaryl Alcohol	$\frac{dS20}{dt} = V_{27} - V_{32} - V_{36}$
Caffeyl Alcohol	$\frac{dS21}{dt} = V_{28} - V_{33}$
Coniferyl Alcohol	$\frac{dS22}{dt} = V_{29} + V_{33} - V_{34} - V_{38}$
5-Hydroxy-Coniferyl-Alcohol	$\frac{dS23}{dt} = V_{30} + V_{34} - V_{35}$
Sinapyl Alcohol	$\frac{dS24}{dt} = V_{31} + V_{35} - V_{40}$

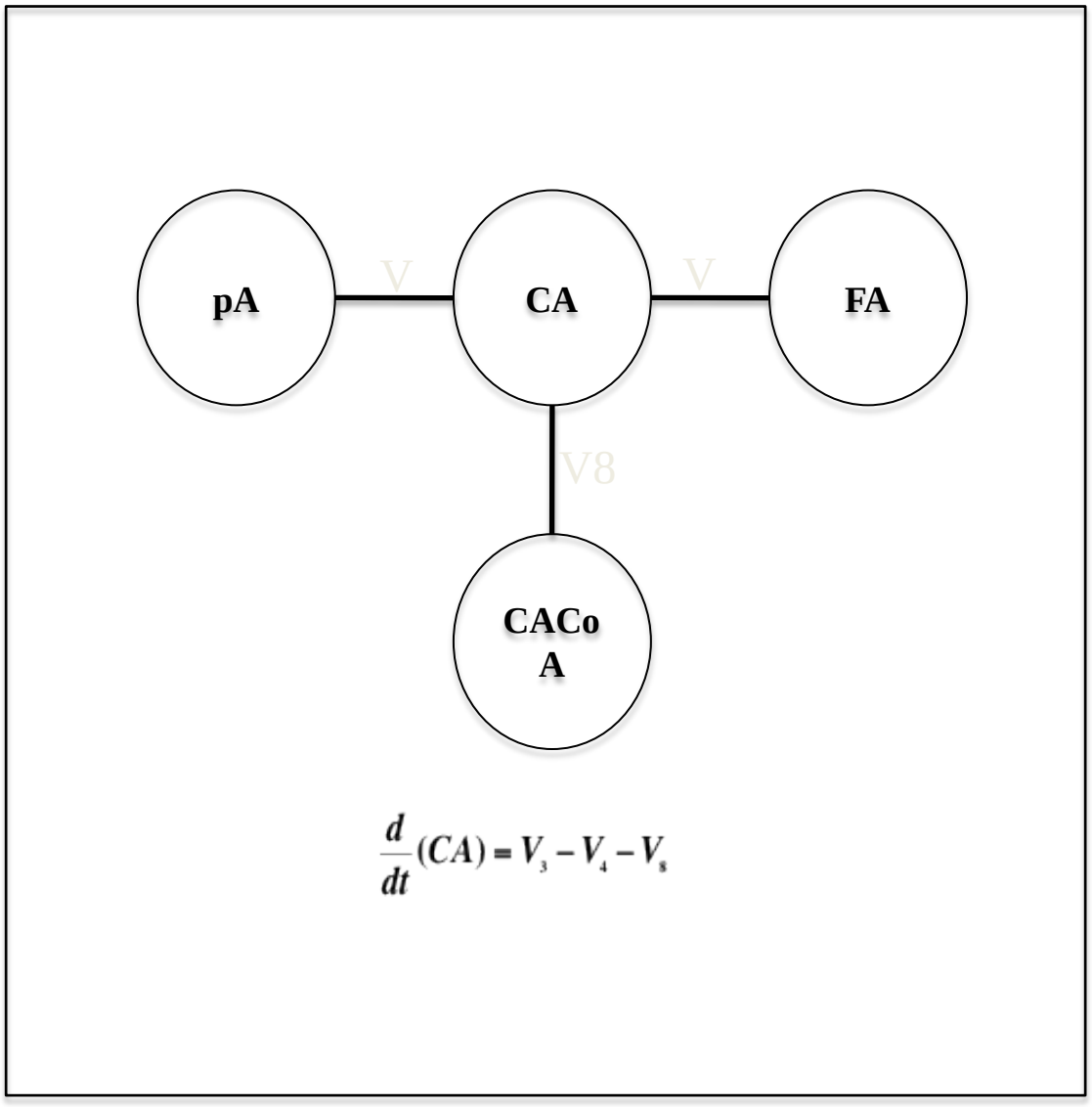

Figure 1 - Simple Representation of a small section of the network to which mass action law is applied. V_3 is the flux that defines the rate of formation of Caffeic Acid from p-Coumaric Acid. V_4 and V_8 define the rate of conversion at which Caffeic Acid is converted to Ferulic Acid and Caffeoyl CoA. All the fluxes are estimated from Michaelis Menten kinetics.

Sensitivity Analysis

Differential equation models depend on the parameters that are determined experimentally; these parameters are usually subject to some kind of error mainly due to the measurement process. Several methods have been developed to explain how the errors associated with these parameters affect the output of the model. The model developed to explain the monolignol biosynthesis pathway is primarily a function of all the activation and inhibition kinetics. Sensitivity analysis of these parameters is not only critical to validate the model, but also to guide future research efforts (McKay et al. 1979). Sensitivity analysis are carried out for a number of reasons, few of which include: (1) which parameters significantly affect the output; (2) which parameters are insignificant, that may be excluded from the final model; (3) which parameters explains the most variability in the output. Sensitivity analysis is rarely done on biological data, but as suggested (Hamby 1995), all biological models must be tested for the parameter sensitivity. The sensitivity analysis in this paper was carried out using the methodology outlined in (Hamby 1995).

There are two kinds of sensitivity analysis that can be carried out, (1) Local Sensitivity Analysis, (2) Global Sensitivity Analysis. In Local sensitivity analysis, one parameter is varied at a time and the change in the output is recorded. This method works well for a model with fewer parameters, but when the model gets complex because of the increased number of parameters, global sensitivity analysis is preferred (Hamby 1995). Global sensitivity analysis is usually carried out by defining a maximum and minimum value for the parameters and

assuming a probability distribution function for each of the parameters. The most commonly used distributions are normal, uniform and triangular. Once the distribution is selected the parameters are randomly selected over the range and the output is calculated. To ensure the parameter values are sampled randomly, a Monte Carlo sampling method is usually employed. We have used Latin Hypercube Sampling (LHS), which is a stratified Monte Carlo sampling (McKay et al. 1979). Latin hypercube sampling selects n different values from each of k parameters. The range of each variable is then divided such that there are n nonoverlapping intervals with equal probability (Iman and Helton 1988). Figure 2 below shows how Latin hypercube sampling is carried out for a sample size of 5 (Iman and Conover 1982).

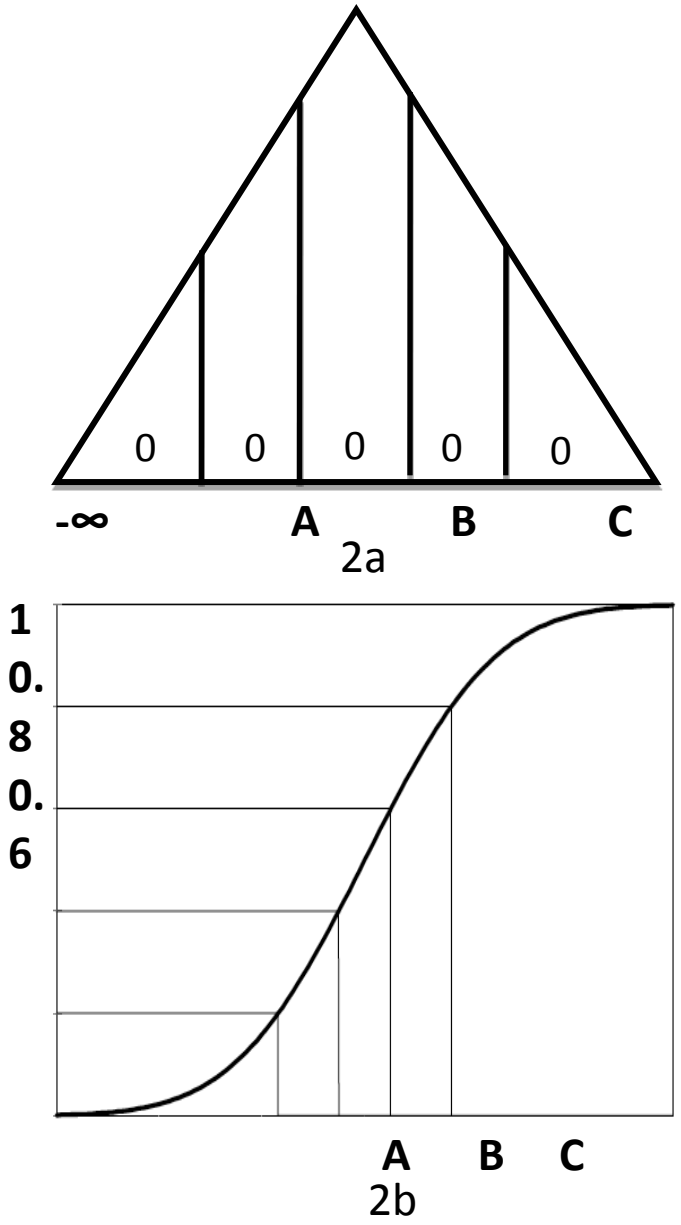

Figure 2 – 2a. Triangular probability distribution curve for a variable with upper and lower bounds divided into 5 equal intervals. 2b. Cumulative distribution curve for a variable with 5 equal intervals (Iman et al. 1981).

The next step in the sensitivity analysis is to understand how the changes in the parameters within the model affect the model output. This can be carried out by calculating the correlation between the parameters and the output of interest. The most commonly used approach is to calculate the Partial Rank Correlation Coefficient for the parameter and the output of interest. PRCC provides a measure of the linear relationship between the parameters and model output, it varies from -1 to 1, where 1 means there is a strong influence of the parameter of interest on the output and value of -1 denotes a strong negative influence of the parameter on the output. For the PRCC method, the output corresponding to the different values of the parameters are tabulated. The input matrix is the matrix of the parameters whose size is (n x k), where n is the number of samples for each parameter and k is the number of parameters, the output matrix is the matrix of size (n x 1) which is a function of the inputs. Because most of the biological models are monotonic because of the number of independent variables and parameters associated with the model, they need to be rank transformed so that they are linear (Iman and Conover 1982). The PRCC between the output variable y and the input variable x is obtained from the use of a sequence of regression models shown in (1). The partial correlation between x and y is defined as the correlation coefficient between $\boldsymbol{y}-\hat{\boldsymbol{y}}$ and $\boldsymbol{x}-\hat{\boldsymbol{x}}_j$ as shown in (2).

$$\hat{x}_j = c_o + \sum_{p \neq j} c_p x_p$$
$$\hat{y} = b_o + \sum_{p \neq j} b_p x_p \qquad (1)$$

Once $\hat{x}_j$ and $\hat{y}$ are obtained, the residuals $x - \hat{x}_j$ and $y - \hat{y}$ were obtained and the partial rank correlation coefficient is calculated using the equation (2), where $\bar{x}$ and $\bar{y}$ are the respective sample means.

$$prcc(x_j, y) = \frac{cov(x_j, y)}{\sqrt{var(x_j)\,var(y)}} = \frac{\sum_{i=1}^{N}(x_{ij} - \bar{x})(y - \bar{y})}{\left[\sum_{i=1}^{N}(x_{ij} - \bar{x})^2\right]^{1/2}\left[\sum_{i=1}^{N}(y - \bar{y})^2\right]^{1/2}}$$
$$j = 1, 2, \ldots k \qquad (2)$$

Results and Discussion

Enzyme initial-rate behavior and substrate specificities

We assayed the first-order kinetic parameters of *P. trichocarpa* monolignol biosynthetic recombinant proteins *in vitro*. 45 sets of kinetic assays were performed in this study (Table 2). Under laboratory conditions, all initial-rate reactions followed Michaelis-Menten kinetics.

Table 2 – Michaelis-Menten kinetic parameters of the 24 *P. trichocarpa* monolignol biosynthetic enzymes

Enzyme	Substrate	KM (µM)	kcat (1/min)	kcat/KM (1/µM.min)
PtrPAL1	Phe	32.7	88.1	2.7
PtrPAL2	Phe	21.4	63.9	2.99
PtrPAL3	Phe	25.8	93.5	3.63
PtrPAL4	Phe	23.3	71.7	3.07
PtrPAL5	Phe	22.7	75.8	3.34
PtrC4H1	CiA	7.05	6.72	0.95
PtrC4H2	CiA	40.7	21.4	0.53
Ptr4CL3	pA	14.9	39.3	2.64
	CA	15.1	25.3	1.68
	FA	70.2	50.3	0.72
	5HFA	13.6	3.84	0.28
Ptr4CL5	CA	45	85.4	1.9
	pA	148	107	0.72
	FA	134	68.4	0.51
	5HFA	193	37.3	0.19
	SA	508	20.8	0.04
PtrC3H3	pACoSA	4.96	57.8	11.65
	pA	227	67.9	0.3
PtrHCT1	pACoA	39	138	3.54
	CaSA	569	4.28	0.01
PtrHCT6	pACoA	25	489	19.53
	CaSA	780	31.7	0.04
PtrCCoAOMT1	CCoA	29.8	75.8	2.55
PtrCCoAOMT2	CCoA	19.1	131	6.87
PtrCCoAOMT3	CCoA	23.1	45	1.95
PtrCOMT2	CAFAld	0.141	4.52	32.04
	5HCAld	0.232	2.6	11.19
	CAFAlc	0.748	6.17	8.25
	5HCAlc	2.46	15.4	6.25
	5HFA	9.93	5.55	0.56
	CA	163	3	0.02
PtrCAld5H1	CAlc	0.0686	2.12	30.93
	CAld	0.0544	1.23	22.59
	FA	17.97	0.0007	0
PtrCAld5H2	CAlc	0.0526	1.72	32.65
	CAld	0.0453	1.01	22.25
	FA	5.24	0.001	0
PtrCCR2	FACoA	44.9	112	2.5
	pACoA	260	38.7	0.15
	CACoA	581	78	0.13
PtrCAD1	CAld	0.68	30.5	44.85
	pAld	3	45.9	15.3
	SAld	2.4	26.9	11.21
PtrCAD2	SAld	35.4	119	3.36
	CAld	70.1	110	1.57

The first step in the phenylpropanoid pathway of *P. trichocarpa* involves the deamination of phenylalanine to cinnamic acid catalyzed by five PtrPALs (Shi et al. 2010). Initial-rate kinetics revealed essentially identical catalytic activities between the PtrPAL family members, suggesting functional redundancy in their roles in the biosynthesis of phenylpropanoids. Multiple PAL genes with similar activities are found in most plants (Hanson and Havir 1981).

PtrC4H1 and PtrC4H2 are homologous proteins that share a 96.4% peptide sequence identity, and mediate the conversion of cinnamic acid to *p*-coumaric acid. The kinetic characterization of the two proteins revealed that both PtrC4H1 and PtrC4H2 have similar specificity constants (k_{cat}/K_m). However, their rate constants are very different. PtrC4H1 has both a low K_m and a low k_{cat}, whereas the K_m and k_{cat} of PtrC4H2 are both significantly higher. When the physiological cinnamic acid concentration is low, PtrC4H1 would be the predominant 4' hydroxylase, whereas when cinnamic acid accumulates, PtrC4H2 would become the predominant 4' hydroxylase. Together, PtrC4H1 and PtrC4H2 are able to efficiently metabolize a broad range of cinnamic acid concentrations.

Two coenzyme-A ligases (Ptr4CL3 and Ptr4CL5) mediate the conversion of hydroxycinnamic acids to hydroxycinnamyl-CoAs. These two proteins have distinct catalytic specificities for the hydroxycinnamic acids (Chen et al. Manuscript in preparation). Ptr4CL3 ligates CoA to *p*-coumaric acid most efficiently, but has a stong affinity for both *p*-coumaric acid and caffeic acid. Ptr4CL5 on the other hand can only effectively bind to caffeic acid.

During the biosynthesis of guaiacyl and syringyl monolignols, hydroxylation of the phenolic ring at the 3-position was thought to occur at the level of the hydroxycinnamic acids. With the discovery of genes encoding C3Hs we learnt that the catalytic efficiency for 3' hydroxylation is significantly higher for *p*-coumaroyl shikimic acid than it is for *p*-coumaric acid (Schoch et al. 2001, Franke et al, 2002). Through characterizing the kinetic properties of PtrC3H3, a *P. trichocarpa* C3H that is expressed specifically and abundantly in SDX (Shi et al. 2010), we validated the specificity of the 3' hydroxylation step in the monolignol biosynthesis of trees and showed that PtrC3H3 has a ~40 fold higher catalytic efficiency for *p*-coumaroyl shikimic acid compared to *p*-coumaric acid (Chen et al. 2011).

To support the involvement of shikimate esters in the biosynthesis of monolignols in *P. trichocarpa*, we evaluated the kinetic properties of PtrHCT1 and PtrHCT6. Both shikimate transferases can convert *p*-coumaroyl-CoA into *p*-coumaroyl shikimic acid, with PtrHCT6 having a significantly higher catalytic efficiency. However, the efficiency of both enzymes for the the formation of caffeoyl-CoA is unexpectately low. For HCT to govern the biosynthesis of caffeoyl-CoA, a principle substrate for CCoAOMTs, caffeoyl-shikimic acid must accumulate to high concentrations *in vivo*. In contrast, past studies have shown that caffeoyl-shikimic acid does not accumulate. These conflicting results challenge our current understanding of the involvement of shikimate esters in the biosynthesis of monolignols, and suggest that alternative mechanisms may co-regulate 3' hydroxylation of the monolignol intermediates.

The O-methylation of the monolignol precursors is governed by two families of enzymes, COMT and CCoAOMT. In *P. trichocarpa*, three PtrCCoAOMTs regulate the 3' methylation of caffeoyl-CoA to feruloyl-CoA (Shi et al. 2010). The kinetic rates of the three PtrCCoAOMTs indicate that all three enzymes can effectively bind to caffeoyl-CoA, but that PtrCCoAOMT2 has a significantly higher catalytic efficiency.

COMTs from a number of different plant species exhibit broad specificity towards the monolignol precursors (Bugos et al. 1992, Inoue et al, 2000, Dixon et al. 2001, Parvathi et al. 2001). Structural analysis of alfalfa COMT suggests that the binding pocket of COMTs is positionally selective for the methylation of hydroxyl groups at the 3- and 5-hydroxyl positions around the phenyl ring, and that, substrate specificity is in the order of aldehydes > alcohols > acids (Zubieta et al. 2002). Our kinetic analysis of recombinant *P. trichocarpa* COMT2 resembled to the catalytic efficiencies and substrate specificities of COMTs from other plant species. PtrCOMT2 methylates the 3- and 5-hydroxyl groups of hydroxycinnamyl aldehydes most efficiently, followed by hydroxycinnamyl alcohols. PtrCOMT2 have a comparatively weak affinity for hydroxycinnamic acids. These results, coupled with the observation that PtrCOMT2 is the most abundant protein in the monolignol biosynthesis pathway of *P. trichocarpa* (Shuford et al. 2012), explains the absence of detectable 5-hydroxylated guaiacyl precursors in the metabolome of *P. trichocarpa* SDX.

PtrCAld5H1 and PtrCAld5H2 are functionally redundant proteins that mediate the first step in the redirection of metabolic flux from the guaiacyl pathway towards syringyl monolignol biosynthesis in *P. trichocarpa* (Wang et al. 2012). The K_m values of PtrCAld5Hs for

substrates coniferaldehyde and coniferyl alcohol are the lowest (Highest substrate affinity) in the 45 kinetic assays performed in this study. This high affinity for the guaiacyl monolignol precursors could explain the high S/G ratio observed in lignin of *P. trichocarpa*.

Reduction of hydroxycinnamyl-CoAs to aldehydes constitutes the first committed step towards the biosynthesis of monolignols from the general phenylpropanoid pathway. In *P. trichocarpa*, a single PtrCCR2 is expressed specifically and abundantly in the SDX (Shi et al. 2010). PtrCCR2 has a strong substrate preference towards the reduction of feruloyl-CoA into coniferaldehyde. Comparatively, the binding affinity of PtrCCR2 to *p*-coumaryl-CoA and caffeoyl-CoA are significantly weaker. This substrate specificity is likely to be accountable for the minimal deposition of *p*-hydroxyphenyl lignin or caffeyl lignin observed in *P. trichocarpa*. These results are consistent with the *in vitro* characterization of CCRs from SDX of different plant species (Pichon et al. 1998, Li et al. 2005, Baltas et al. 2005, Ma 2007, Escamilla-Trevino et al. 2010, Zhou et al. 2010, Tu et al. 2010, Tamasloukht et al, 2011).

Two cinnamyl alcohol dehydrogenases (PtrCAD1 and PtrCAD2) mediate the last step in the monolignol biosynthesis pathway of *P. trichocarpa*, the reduction of hydroxycinnamyl aldehydes into alcohols (Shi et al. 2010). Historically, a single coniferyl alcohol specific CAD (PtrCAD1) was thought to catalyze the biosynthesis of both guaiacyl and syringyl monolignols. Recently, a novel SAD (PtrCAD2) with substrate preference for sinapyl alcohol was identified (Li et al. 2001). We evaluated the kinetics and specificities of PtrCAD1 and PtrCAD2 toward the hydroxycinnamyl aldehydes. PtrCAD1 has highest

substrate specificity for coniferaldehyde, ~40 fold that of the specificity for the reduction of sinapaldehyde and *p*-coumaraldehyde. PtrCAD2 has a relatively weak reductase activity compared to PtrCAD1, but has higher substrate specificity for sinapaldehyde.

Inhibition

We assayed the inhibition kinetic parameters, K_{IC}, K_{IU} and K_{IS} of *P. trichocarpa* monolignol biosynthesis recombinant proteins *in vitro,* against an array of potential feed-back inhibitors. A total of 54 sets of inhibition kinetic assays were performed (Table 3).

Deamination of phenylalanine is regulated by product inhibition (Sato et al. 1981). We assayed the inhibition kinetics of cinnamic acid on the activities of PtrPAL1~PtrPAL5. All PtrPALs are strongly inhibited by cinnamic acid in a competitive mode, with the K_{IC}s much lower than the corresponding K_ms. Therefore cinnamic acid is not only a key intermediate, but also an important regulator of the phenylpropanoid pathway.

Table 3 – Inhibition kinetic parameters of the 24 *P. trichocarpa* monolignol enzymes

Enzyme	Substrate	Inhibitor	K_{IC}	K_{IU}	K_{IS}
PtrPAL1	Phe	CiA	9.77		
PtrPAL2	Phe	CiA	6.18		
PtrPAL3	Phe	CiA	7.11		
PtrPAL4	Phe	CiA	5.1		
PtrPAL5	Phe	CiA	5.6		
Ptr4CL3	pA	CA	14.5		
		FA	45.6		
		5HFA	38.4		
		SA	91.5		
	CA	pA	43.5		
		FA	98.5		
		5HFA	107		
	FA	pA	12.8		
		CA	11.1		
		5HFA	9.62		
		SA	173		
Ptr4CL5	pA	CA	7.06	47.4	
		FA	293	178	
		5HFA	48.8	37.5	
		SA	632	238	
	CA	pA	123	78.1	
		CA			56
		FA	519	372	
		5HFA	45.4	307	
		SA	2842	157	
	FA	pA	27.9		
		CA	18.5	15	
		5HFA	186	18.71	
		SA	18049		
PtrCOMT2	CA	CAFAld	2.08	7.57	
		CAFAlc	1.63	12.1	
		5HFA	38.6	245	
		5HCAld	0.2	2.7	
		5HCAlc	3.8	23.2	
	CAFAld	CA	175		
		CAFAld			14.8
		5HFA	19.8		
		5HCAld	0.554	1.45	
		5HCAlc	0.796		
	CAFAlc	CAFAlc			26.6
		CAFAld	5.65	0.734	
		5HCAld	0.129	0.542	
		5HCAlc		3.14	
	5HCAld	CAFAld	0.287		
		CAFAlc	2.5	14.37	
		5HCAld			1.81

A complex regulatory network governs the CoA ligation of hydroxycinnamic acids in *P. trichocarpa* (Chen et al. manuscript in preparation). Despite a high peptide sequence, Ptr4CL3 and Ptr4CL5 are regulated by distinct modes of metabolic control. Whereas the catalysis by Ptr4CL3 is regulated solely through a competitive mode, activities of Ptr4CL5 exhibited much more complex substrate / inhibitor binding characteristics involving three modes of inhibition, competitive, uncompetitive and substrate.

The methylation of phenolic intermediates by COMT is an important step in the biosynthesis of monolignols. Enzyme inhibition kinetics were determind to identify inhibitors of PtrCOMT2 reactions and their modes of inhibitions. Caffeic acid, 5-hydroxyferulic acid, caffeylaldehyde, 5-hydroxyconiferaldehyde, caffeyl alcohol and 5-hydroxyconiferyl alcohols were tested for inhibition. Both the 3- and 5-methylation of hydroxycinnamyl aldehydes and alcohols are regulated by substrate inhibition, with the 5-hydroxylated phenolics being more severely regulated. The 3-methylation of caffeic acid is strongly inhibited both competitively and uncompetitively by the remaining 5 substrates of PtrCOMT2. The 3-methylation of caffeylaldehyde is strongly inhibited by 5-hydroxyconiferaldehyde and 5-hydroxyconiferyl alcohol, and weakly inhibited by caffeic acid and 5-hydroxyferulic acid. Caffeyl alcohol 3-methylation is strongly inhibited by caffeylaldehyde and 5-hydroxyconiferaldehyde, and weakly inhibited by 5-hydroxyconiferyl alcohol. 5-hydroxyconiferaldehyde is strongly inhibited competitively by caffeylaldehyde, and weakly inhibited by caffeyl alcohol and 5-hydroxyconiferyl alcohol. These elaborate relationships between the phenolic compounds and the catalytic specificities of PtrCOMT2 suggest that the methylation steps in monolignol biosynthesis are under strict metabolic control.

Ferulic acid and coniferyl alcohol 5-hydroxylation are both inhibited by the 5-hydroxylation of coniferaldehyde in *P. trichocarpa* (Wang et al. 2012). While ferulic acid 5-hydroxylation is strongly inhibited both competitively and uncompetitively, the inhibition of coniferyl alcohol 5-hydroxylation is mild in comparison. Both coniferaldehyde and coniferyl alcohol may be viable intermediates for 5-hydroxylation by PtrCAld5H1 and PtrCAld5H2 in *P. trichocarpa*.

PtrCAD1 is the predominant enzyme responsible for the reduction of cinnamaldehydes into monolignols. We determined the extent and modes of inhibition for coniferaldehyde and sinapaldehyde on the PtrCAD1 mediated biosynthesis of coniferyl and sinapyl alcohols. The reduction of sinapaldehyde is strongly inhibited both competitively and uncompetitively by coniferaldehyde, whereas the coniferaldehyde reduction by sinapaldehyde is inhibited to a lesser extent. These results provided further evidence for the specificity of PtrCAD1 for the biosynthesis of guaiacyl monolignols.

Simulated perturbation studies: Regulatory roles of individual enzyme families in the monolignol biosynthetic pathway

Targeted reverse genetic analyses of monolignol biosynthesis genes are comprehensive. However, the mutant phenotypes are not always consistent with our knowledge of the monolignol pathway. These paradoxical observations suggest that the current model of monolignol biosynthesis is an oversimplification. Additional enzymes and regulatory mechanisms may be necessary in the biosynthesis of monolignols.

It is to our interest, in the kinetic modeling of monolignol biosynthesis to identify and quantify pathway components and regulatory mechanisms in the lignin pathway. To do this, we simulated and systematically perturbed individual pathway enzyme families in the kinetic model and evaluated the consequent flux patterns. Contradictions and inconsistencies in the model outputs and our knowledge of monolignol biosynthesis will provide a foundation for discovering novel components and mechanisms that govern lignin content and composition.

Perturbation of PAL

Transgenics perturbed in the expression levels of PALs have been generated in numerous plant species including *P. trichocarpa*. Severe inhibition of PAL generally leads to decreased levels of lignin and other phenylpropanoids, and an increase in the syringyl to guaiacyl ratio in lignin. Phenotypes such as stunted growth, delayed root formation and altered leaves were also documented in plants with severe PAL down-regulation. However,

the degree of inhibition required to impact lignin deposition suggests that PAL activities are present in excess. This is likely due to its essentiality, as multiple copies of PAL are present in most plant species. Studies of the quantitative relationships between *PAL* expression levels and lignin accumulation in transgenic tobacco revealed that PAL activity becomes the rate-limiting-step in lignin biosynthesis only when the expression levels are 3- to 4- fold below wildtype (Bate et al. 1994).

We investigated the functional roles of PtrPALs in our kinetic model. Concentrations of the PAL family proteins were reduced incrementally from 100% to 0.1% of wildtype (Figure 3), and the subsequent effects on the simulated pathway flux patterns, lignin content and S/G ratios were evaluated (Figures 4, 5). Many of the characteristics of the PAL transgenics are captured in our simulation. Detectable changes in the pathway flux and S/G ratio are observed only when PAL concentrations are reduced to <1% of wildtype. PAL activities are not likely to be acting at capacity for the biosynthesis of monolignols.

The observation that severe down-regulation of PALs result in reduced lignin and a slight increase in S/G ratio was first made in transgenic tobacco (Sewart et al. 1997). It has been postulated that metabolic channeling of multiprotein compexes involving PALs and downstream proteins are responsible for the inhibition of G-lignin or activation of S-lignin biosynthesis. However, demonstration of the existence of a multi-protein complex has been difficult to obtain and additional evidence is still required. PAL catalyzes reaction that initiates the biosynthesis of all phenylpropanoid products. Deciphering the molecular mechanisms underlying the relationships between PAL activities and lignin S/G ratio is not

possible without comprehensive knowledge of the dynamics of the monolignol biosynthetic pathway.

Our kinetic simulations show an increase in S/G ratio when PAL activities are severely reduced. Close examination of the model outputs revealed an intricate relationship between the flux entering the monolignol pathway and the biosynthesis of S- and G-lignin. When PAL activity is reduced, the lowered flux is preferentially diverged into the syringyl monolignols at the levels of hydroxycinnamyl aldehydes and alcohols. The preference for syringyl monolignol biosynthesis is in large part due to the kinetic properties of the 5-hydroxylases. The substrate binding capacity of CAld5Hs for the guaiacyl monolignol precursors coniferaldehyde and coniferyl alcohol is the strongest of all enzyme/substrate combinations in monolignol biosynthesis. Therefore, the production of syringyl lignin may be driven by the availability of CAld5Hs rather than by substrate. When the overall monolignol flux is reduced, the synthesis of G-lignin is more severely affected.

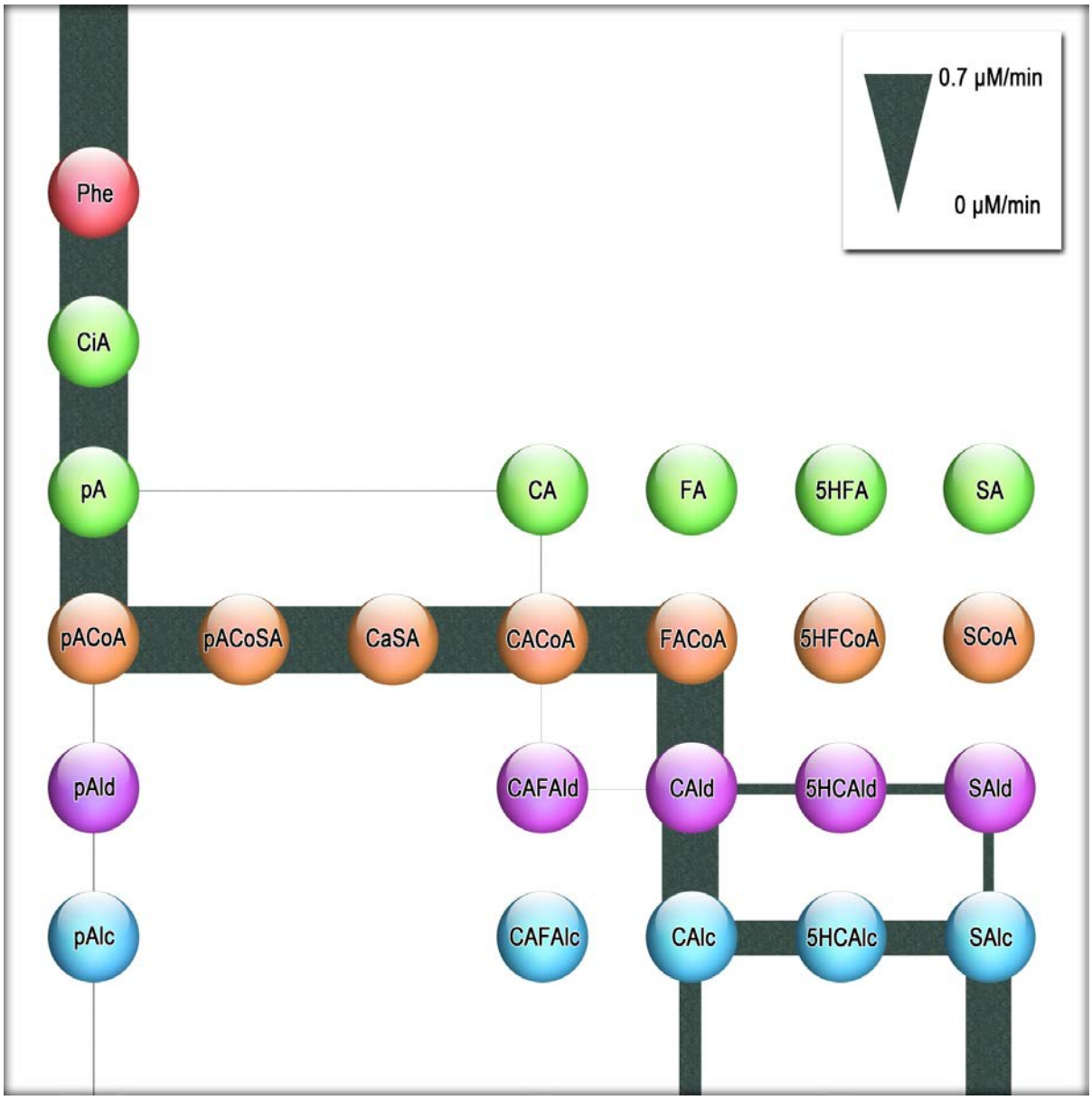

Figure 3 – Simulated monolignol biosynthesis pathway flux pattern of wildtype *P. trichocarpa*

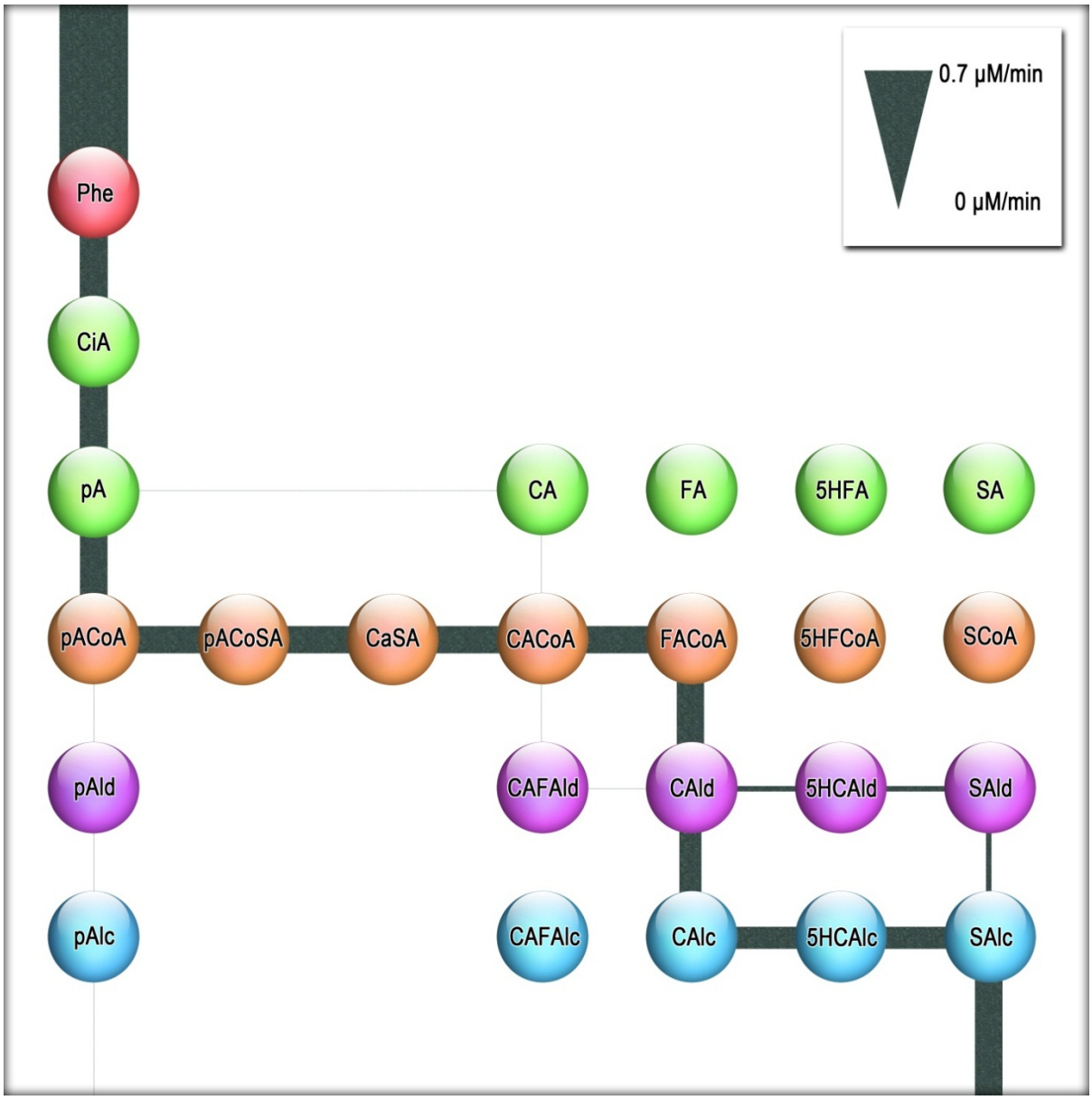

Figure 4 – Flux pattern for 99% knockdown of PAL activity

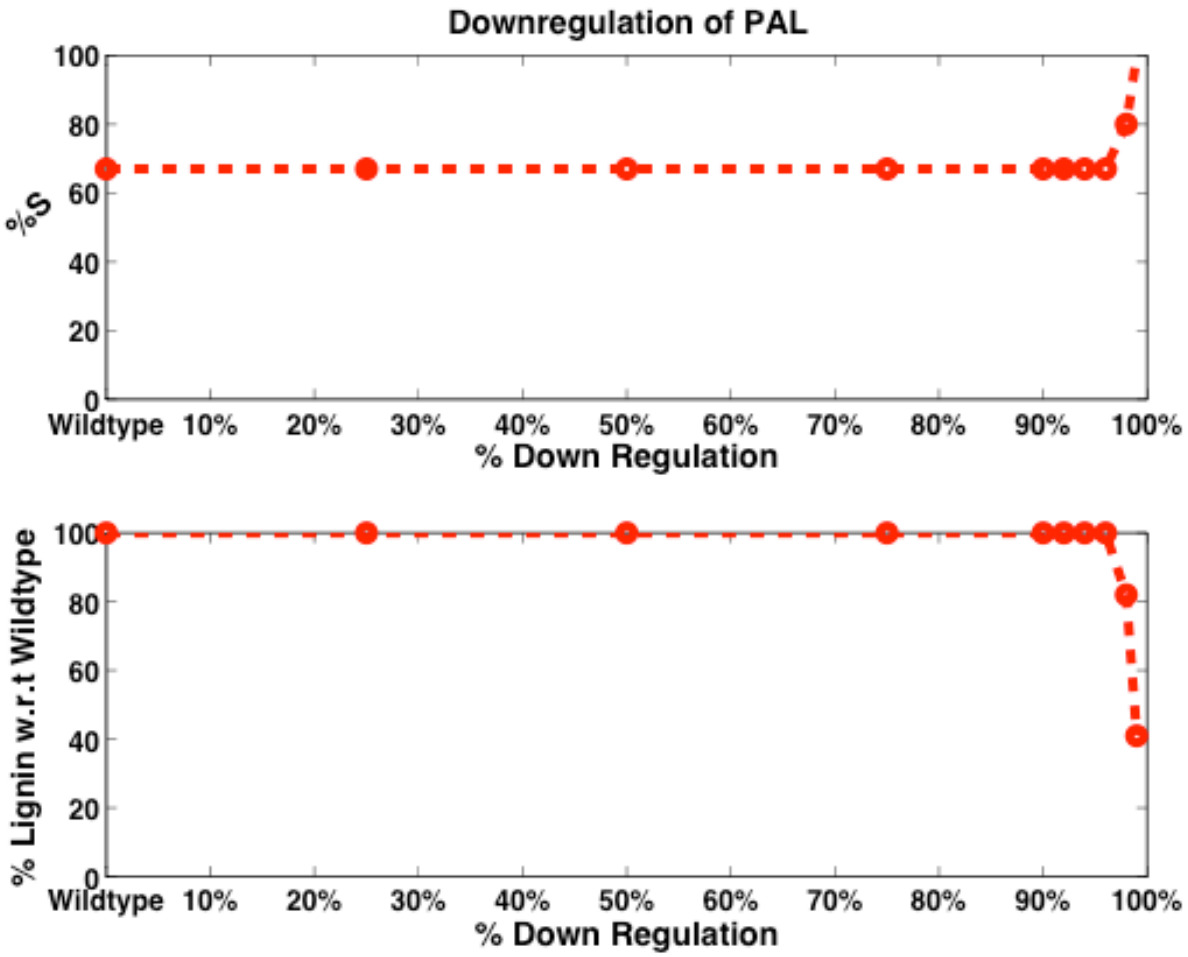

Figure 5 – The effects of PAL perturbations on simulated syringyl moiety abundance in lignin and changes in lignin content.

Perturbation of C4H

C4H transgenic plants have been generated for tobacco (Sewalt et al. 1997), alfalfa (Reddy et al. 2005), tomato (Millar et al. 2007), *Arabidopsis* (Schilmiller et al. 2009) and hybrid aspen (Bjurhager et al. 2010). Reduction in the activity of C4H generally leads to decreased levels of lignin and several different classes of phenylpropanoid products such as sinapoylmalate and flavonoids. The observed changes in lignin composition in response to reduction in C4H are not consistent. Transgenic tobacco and alfalfa exhibited a lowered lignin S/G ratio when C4H is down-regulated, whereas in lignin of transgenic tomato and *Arabidopsis*, S/G is greatly increased. The lignin of transgenic hybrid aspen perturbed in C4H was not altered in S/G ratio.

C4H transgenic plants provided evidence for a feedback loop at the entry point into the phenylpropanoid pathway. Common to all transgenic plant species perturbed in the expression of C4H is the reduction of PAL activity. Cinnamic acid is a feedback modulator of the phenylpropanoid pathway through PAL. Altering the expression of C4H activity leads to the accumulation of cinnamic acid, which inhibits the activity of PAL, and lowers the flux into the phenylpropanoid pathway.

We investigated the functional roles of PtrC4Hs in our kinetic model. Concentrations of the C4H family proteins were reduced incrementally from 100% to 1% of wildtype, and the subsequent effects on the simulated pathway flux patterns, lignin contents and S/G ratios were evaluated (Figure 6, 7). The results of our simulations show that, similar to PAL, the activity of C4H for monolignol biosynthesis in *P. trichocarpa* is present is excess. Changes

in the content and composition of lignin are only detected when C4H concentration is <12% of wildtype.

Our simulations suggest that down-regulating C4H in *P. trichocarpa* results in a drastic increase in S/G, and that the mechanism is the same as for PAL perturbations. However, the recent generation of C4H transgenic hybrid aspen showed that S/G ratios in poplars are not affected by the down-regulation of C4H (Bjurhager et al. 2010). The inconsistency between our simulations and transgenic plants suggest that an alternative mechanism may be present in trees for maintaining homeostasis in lignin composition.

Evidence for the feedback inhibition of PAL activity by reducing C4H were portrayed in our simulations. The extent of feedback inhibition was quantified and evaluated. The model suggests that the inhibitory effect on the entry point into the phenylpropanoid pathway is mild comparing to the extent of C4H downregulation. ~ 10% reduction in PAL activity was observed when C4H is reduced to 1% of wildtype.

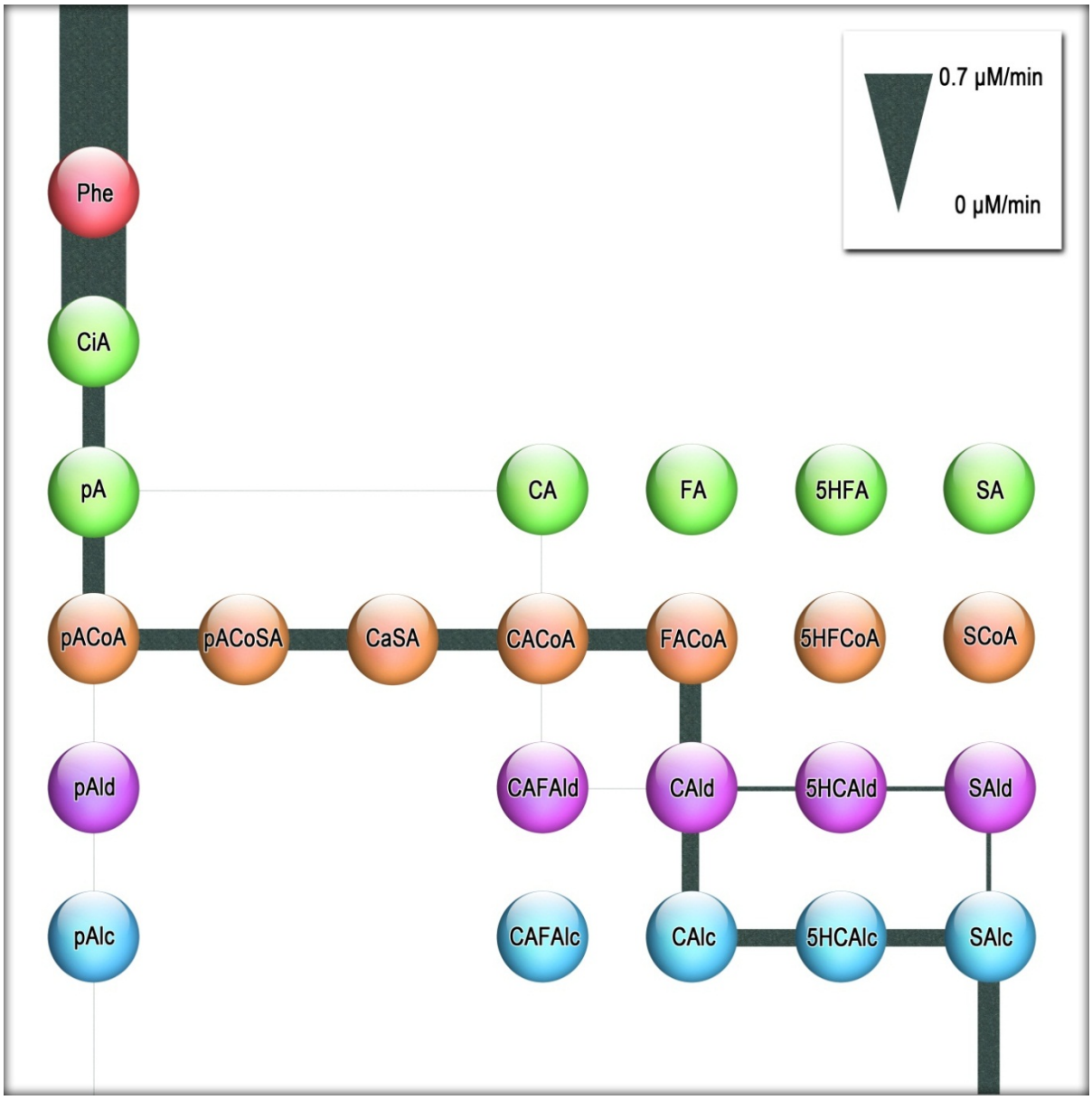

Figure 6 – Flux pattern for 96% knockdown of C4H activities

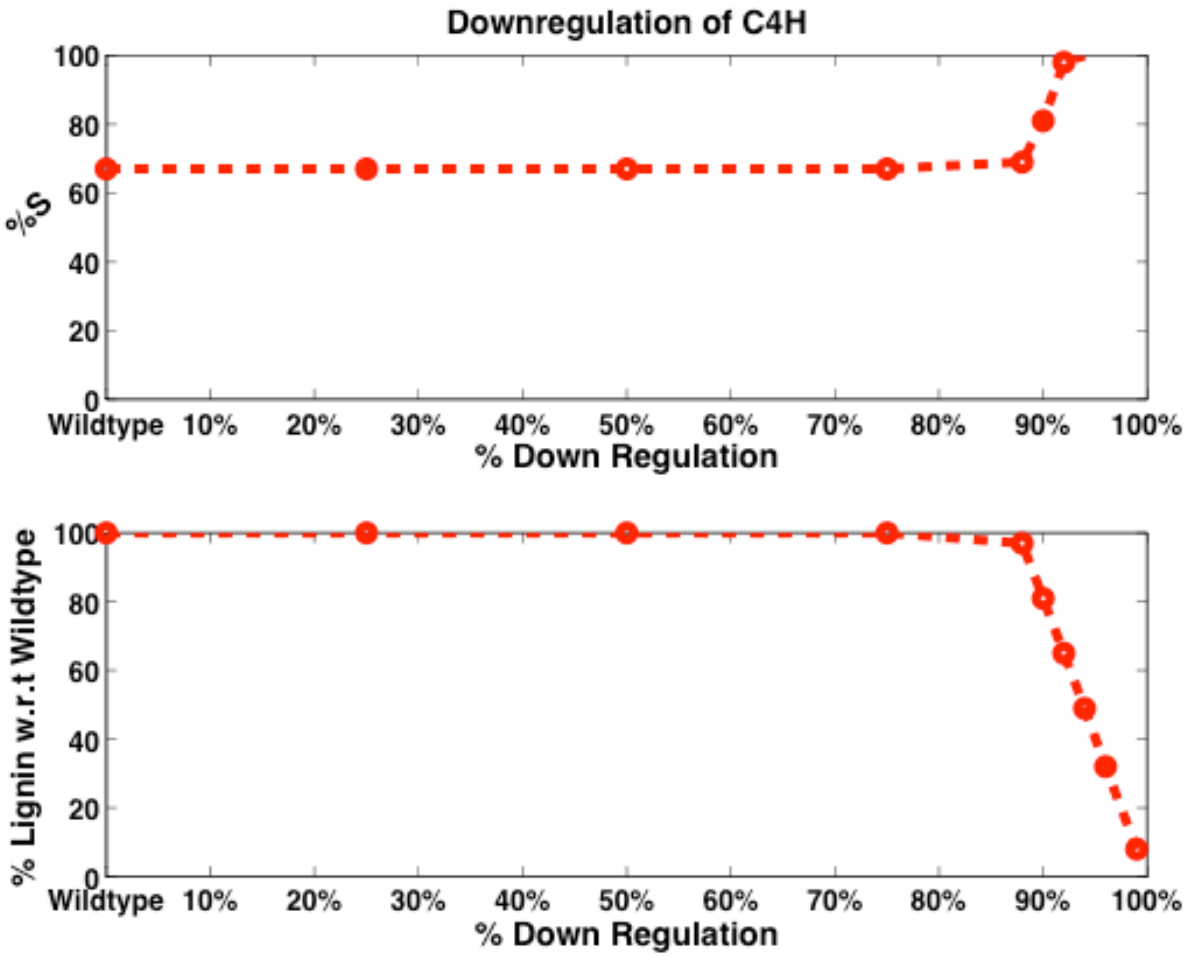

Figure 7 – The effects of C4H perturbations on simulated syringyl moiety in lignin and changes in lignin content.

Perturbation of 4CL

Transgenic plants perturbed in the expression of 4CLs have been generated in tobacco (Kajita et al. 1996), Arabidopsis (Lee et al. 1997), aspen (Hu et al. 1999), *radiata* pine (Wagner et al. 2009), hybrid poplar (Kitin et al. 2010, Voelker et al. 2010), hybrid white poplar (Voelker et al. 2011) and switchgrass ((Xu et al. 2011). Reductions in the activities of 4CL have significant impacts on both lignin content and composition. Lignin quantity is reduced in all 4CL transgenic plants. The reduction in lignin is compensated by an increase in the cellulose content (Hu et al. 1999, Li et al. 2003). Soluble and wall bound phenolics *p*-coumaric acid, ferulic acid and sinapic acid accumulate in transgenic plants with severe 4CL reduction (Kajita et al. 1997, Hu et al. 1999). Changes in lignin composition are inconsistent between different plant species downregulated in 4CL expression. Aspen, tobacco, Arabidopsis and switchgrass exhibited an increase in the S/G ratio of lignin when 4CL is reduced. The increase in S/G is mainly due to the reduction in G-lignin deposition. The S/G ratio of transgenic hybrid popar (*Populus tremula* × *P. alba*) however, decreased when 4CL expression is down-regulated. The roles of 4CL isoforms in directing the metabolic flux for the biosynthesis of different phenolic pathways remains unclear.

We simulated the roles of *P. trichocarpa* 4CLs in the regulation of lignin composition, content and monolignol pathway flux pattern. Simulations were run using reduced concentrations of the Ptr4CL family proteins from 100% to 1% of wildtype in the kinetic model (Figure 9, 10). When the total 4CL activity is lowered, the principle metabolite flux through the 3-hydroxylation and subsequent 3-methylation steps in the biosynthesis of monolignols shifts from the pathway involving phenolic shikimic acids and CoA thioesters to

the level of the hydroxycinnamyl acids. The coenzyme A ligation step then proceeds through the metabolism of ferulic acid instead of *p*-coumaric acid and caffeic acid. These results provide an explanation to the accumulation of hydroxycinnamic acids in 4CL transgenic plants.

Sensitivity analysis

We carried out the sensitivity analysis for the lignin biosynthesis model, assuming a triangular distribution for all the parameters, with the mode being the nominal value for each parameter measured experimentally and the lower and upper bounds were set to ±50% of the nominal values. There were a total of 161 parameters in the model, and the sensitivity analysis was carried out with n=2000, which means that there were 2000 values sampled for each parameter, within the specified range.

The objective of this study is to identify the parameters that influence the steady state concentration of coniferyl alcohol and sinapyl alcohol.

The table 4 shows the list of parameters along with the PRCC values that have a positive effect on the steady state concentration of coniferyl alcohol. The complete list of parameters along with the corresponding PRCC values is shown in the appendix. From the table it is evident that the most sensitive parameters are the ones, which are in the vicinity of coniferyl alcohol in the pathway.

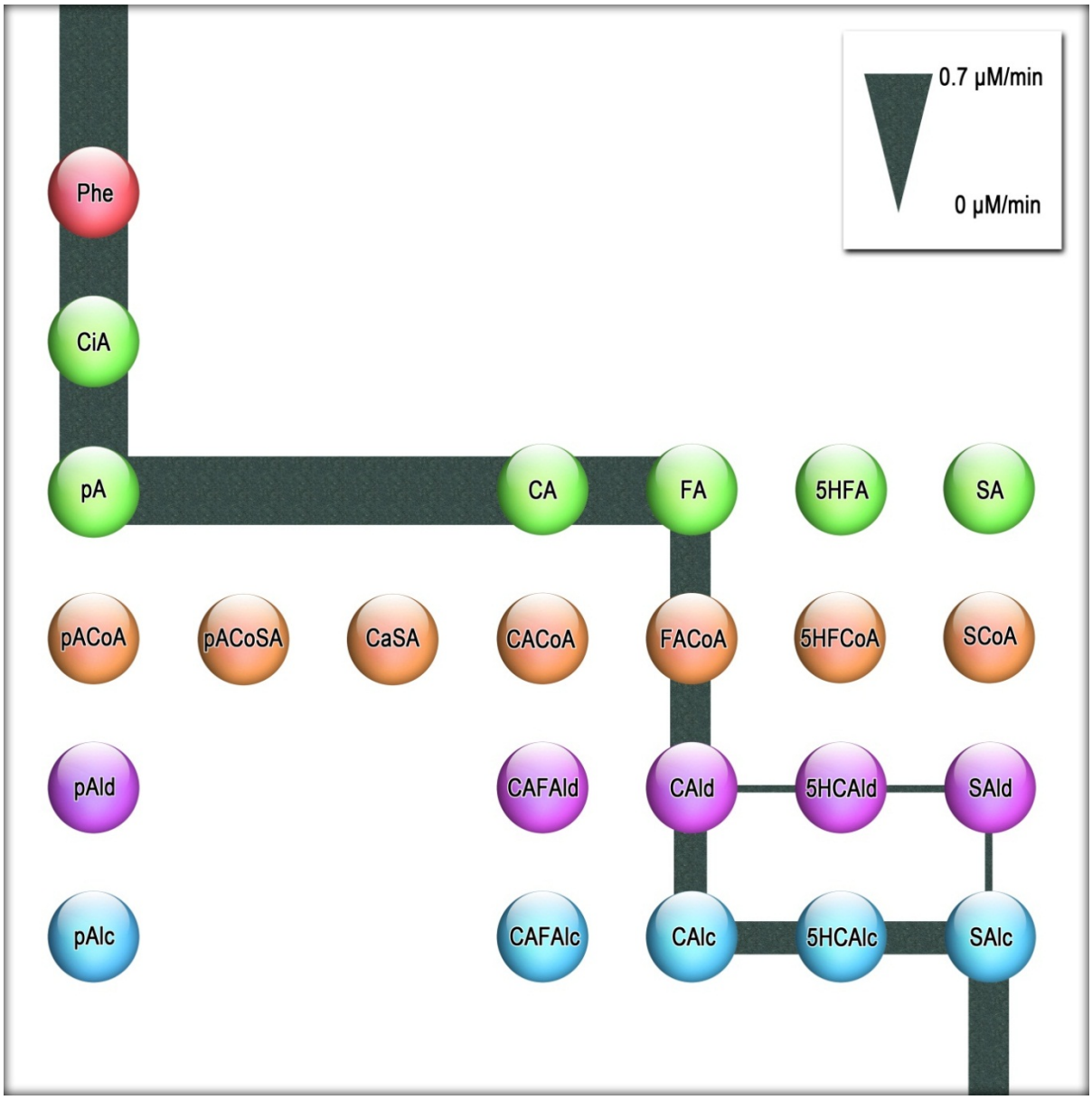

Figure 9 – Flux pattern for 99% knockdown of 4CL activities

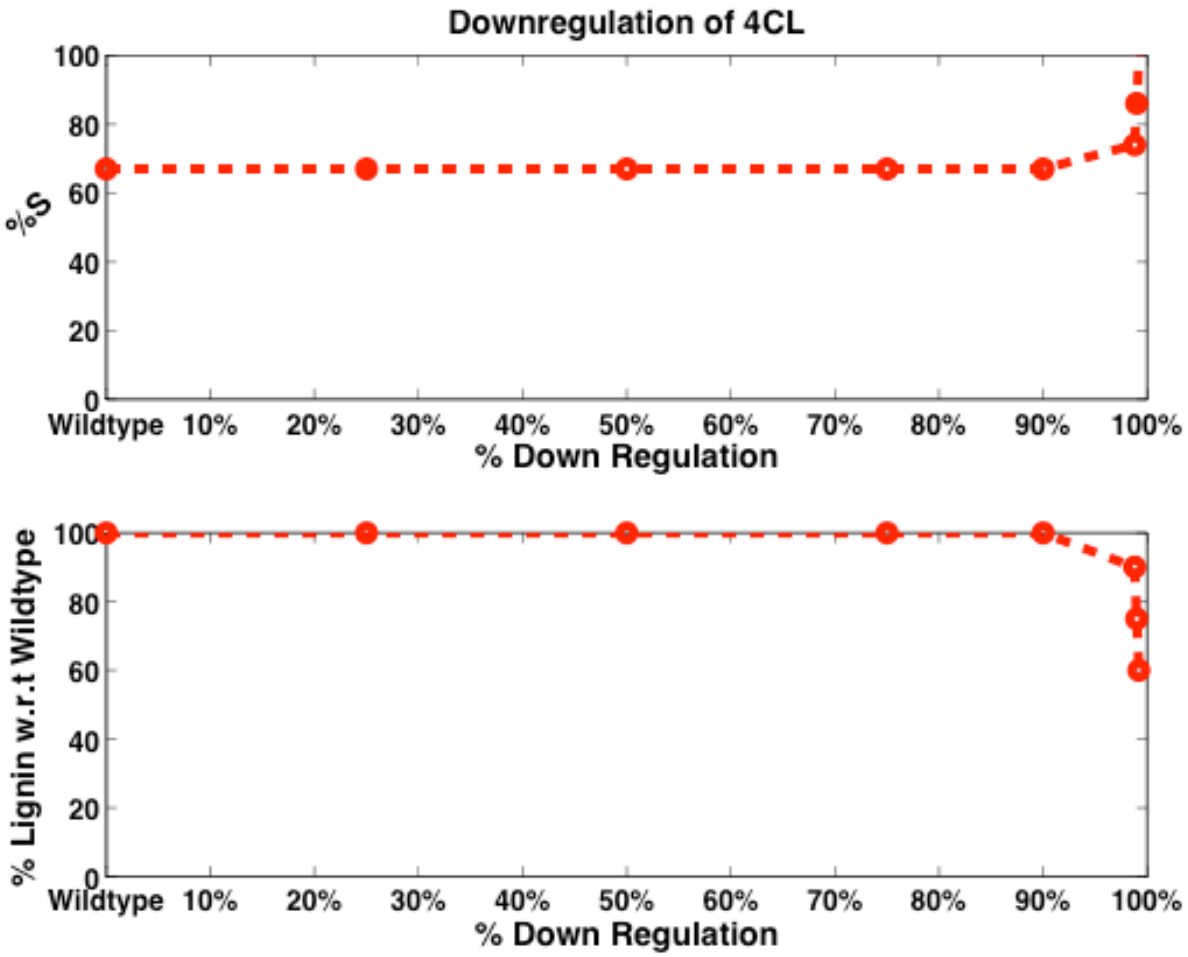

Figure 10 – The effects of 4CL perturbations on simulated syringyl moiety in lignin and changes in lignin content.

Table 4 - List of parameters and the corresponding partial rank correlation coefficient (PRCC) values that have a significant effect on the steady state concentration of coniferyl alcohol.

Parameter	PRCC(Coniferyl Alcohol)
kmCAld5H2CAlc	0.7777
kmCAld5H1CAlc	0.7277
kcatPAL4Phe	0.712
kcatPAL2Phe	0.6988
kisCOMTCAlc	0.476
kmCald5H1CAld	0.2409
kmCald5H2CAld	0.2322
kcatPAL3Phe	0.1625
kcatPAL1Phe	0.1297
kicCOMTCAFAld5HCAld	0.0642
kicCOMT5HCAld	0.045
kicCOMTCAFAldCA	0.0384
kiuCOMTCA5HCald	0.0339
km4CL5SA	0.0332
kic4CL5pA5HFA	0.0329

Parameters K_m associated with the enzyme PtrCAld5Hs for the substrate coniferyl alcohol is related to the equation of flux that takes away coniferyl alcohol, since km is in the denominator it has a positive effect on coniferyl alcohol. The second set of parameters that are most sensitive are the kcat associated with PAL4 with phenylalanine as the substrate, this is because the flux V1 is the one that is at the starting point of the pathway and hence higher the values of kcat the higher is the flux V1, which results in the increased formation of coniferyl alcohol. The scatter plot in figure 11, shows that there is a strong correlation between the parameter kmCAld5H2Calc. The y-axis in the plot denotes the residuals of the steady state concentration of coniferyl alcohol, which is the difference between the steady

state concentrations obtained by the regression model, with and without kmCAld5H2Calc in the model. Similarly, the x-axis denotes the residuals obtained by the regression model, with and without kmCAld5H2Calc.

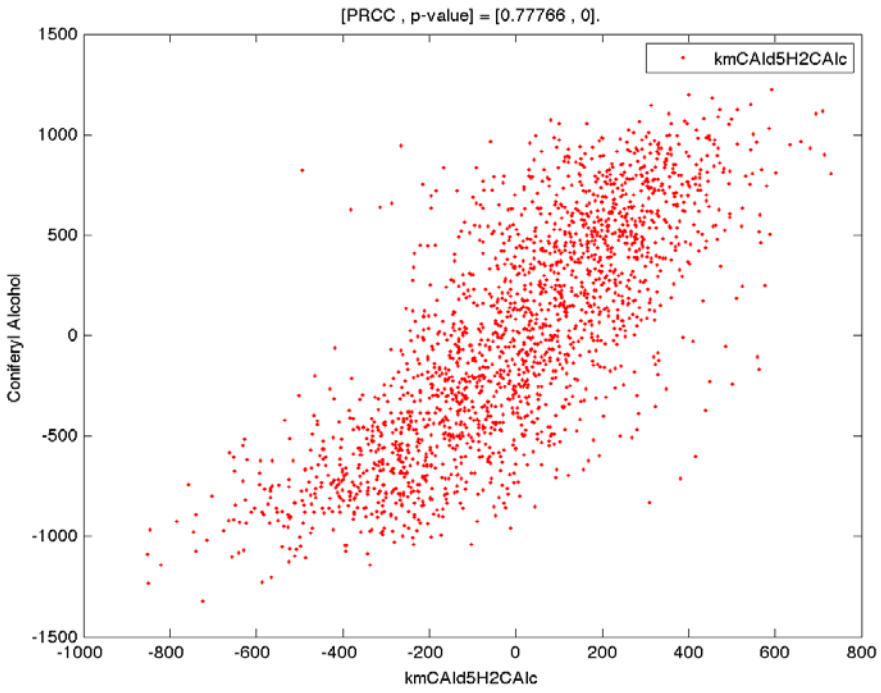

Figure 11 - The partial correlation plot of the parameter kmCAld5H2CAlc and coniferyl alcohol concentration

The table 5 shows the list of parameters that have a negative effect on the steady state concentration of coniferyl alcohol. Coniferyl alcohol concentration is most sensitive to kcatCAld5H2Calc; this can be explained due to the fact that CAld5H2 uses coniferyl alcohol as the substrate to form 5-hydroxy coniferyl alcohol. Since the flux equation, defining the formation of 5-hydroxy coniferyl alcohol is proportional to kcatCAld5H2Calc, the steady

state concentration of coniferyl alcohol decreases as the value of kcatCAld5H2CAlc increases. The partial correlation plot shown in Figure 12, shows the relationship between the concentration of coniferyl alcohol and kcatCAld5H2Calc.

Table 5 - List of Parameters and the corresponding PRCC values that tend to have a negative effect on the steady state concentration of coniferyl alcohol.

Parameter	PRCC(Coniferyl Alcohol)
kcatCAld5H2CAlc	-0.769
kmPAL4Phe	-0.733
kcatCAld5H1CAlc	-0.727
kmPAL2Phe	-0.688
kiuCOMTCAFAld	-0.473
kcatCald5H2CAld	-0.237
kcatCald5H1CAld	-0.221
kmPAL3Phe	-0.173
kmPAL1Phe	-0.147
kcatCCRpACoA	-0.08
kiuCAD2SAldCAld	-0.052
kic4CL5FApA	-0.05
km4CL3pA	-0.048
kcatCCoAoMT1CaCoA	-0.048
kcatCCoAoMT3CaCoA	-0.046
kiuCOMT5HCAld	-0.039
kiuCOMTCACaFald	-0.039
kmHCT6pCCoA	-0.036
kmCCoAoMT1CaCoA	-0.036
kic4CL3FApA	-0.034
km4CL5FA	-0.034

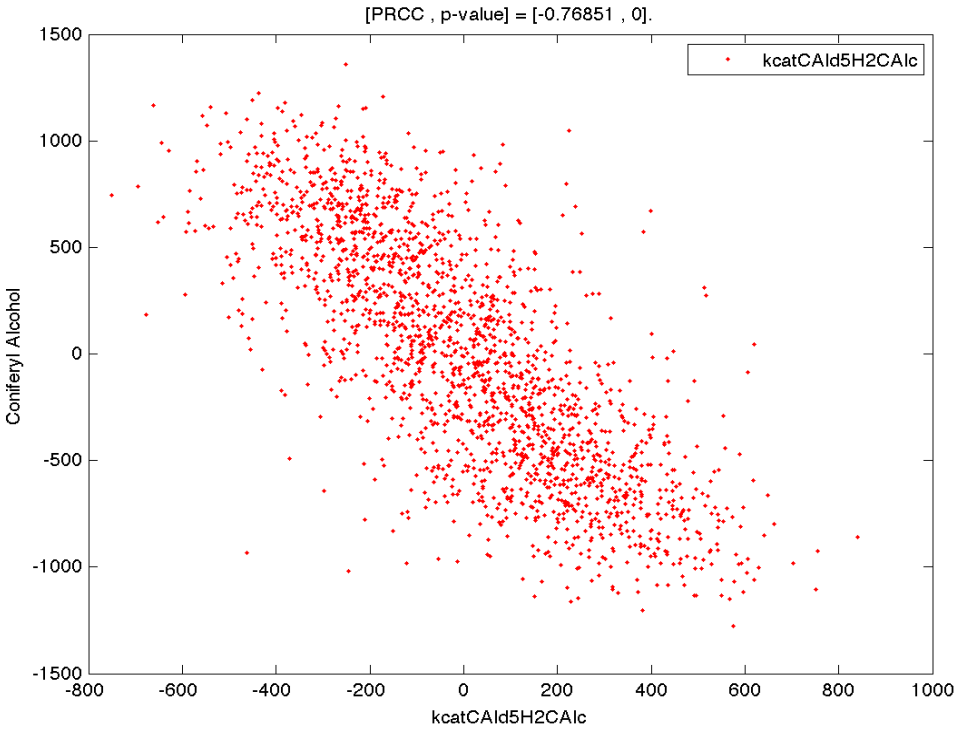

Figure 12 - The partial correlation plot of the parameter kcatCAld5H2CAlc and coniferyl alcohol concentration

Similarly, table 6 shows the sensitivity of sinapyl alcohol to the different parameters. The parameters that have the most positive affect on the steady state concentration is the kcat associated with the enzyme PAL, which converts phenylalanine to cinnamic acid. The effect of the parameter kcatPAL4Phe on sinapyl alcohol is shown in Figure 13. Similarly, Table 7 shows the list of parameters that have a negative effect on the steady state concentration of sinapyl alcohol. The partial correlation plot between kmPAL4Phe is shown in Figure 14.

Table 6 - List of parameters to which sinapyl alcohol is sensitive

Parameter	PRCC(Sinapyl Alcohol)
kcatPAL4Phe	0.842
kcatPAL2Phe	0.8163
kcatPAL3Phe	0.2233
kcatPAL1Phe	0.2209
kmCCRpACoA	0.0945
kcatHCT6pCCoA	0.0443
kic4CL5CA5HFA	0.041
kicCOMTCAFAldCA	0.0346
kicCOMTCAFAld	0.0332
kmHCT1CaSA	0.0329
kiuCOMT5HCAldCAFAlc	0.0324

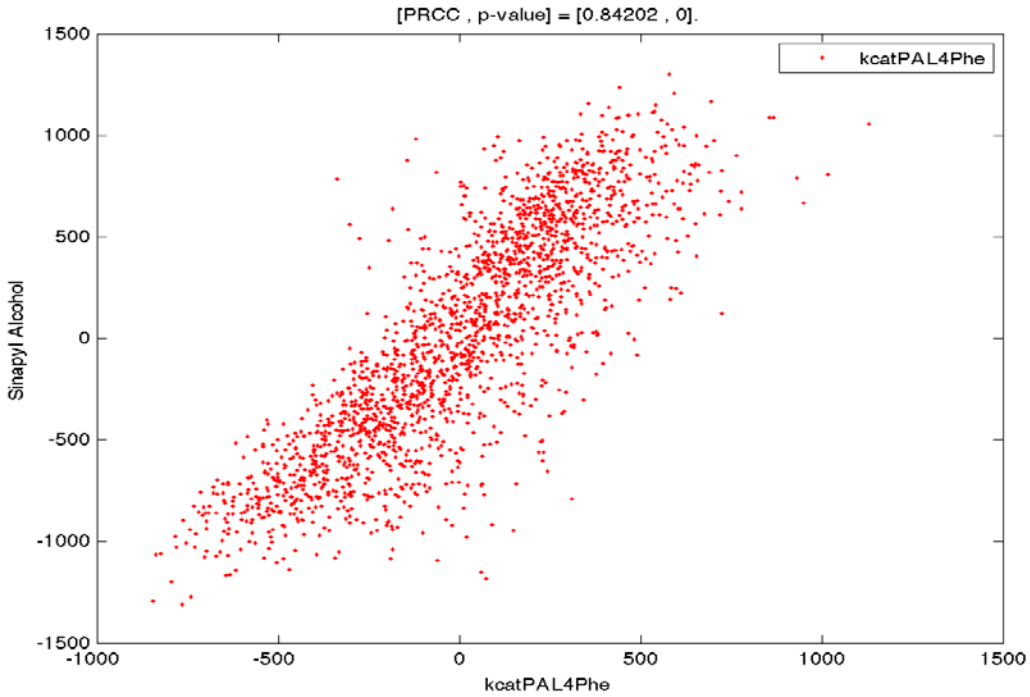

Figure 13 - The partial correlation plot between the parameter kcatPAL4Phe and sinapyl alcohol concentration

Table 7 - List of parameters that have a positive effect the steady state concentration of sinapyl alcohol.

Parameter	PRCC(Sinapyl Alcohol)
kmPAL4Phe	-0.853
kmPAL2Phe	-0.824
kmPAL1Phe	-0.256
kmPAL3Phe	-0.233
kmHCT1pCCoA	-0.087
kmHCT6pCCoA	-0.078
kcatCCRpACoA	-0.07
kmHCT6CaSA	-0.07
kiuCAD2SAldCAld	-0.067
kic4CL5FApA	-0.064
kmCCRFACoA	-0.053
kcatCOMT5HCAld	-0.05
kic4CL5FASA	-0.049
kic4CL3FApA	-0.046
kic4CL5FA5HFA	-0.039
kiu4CL5pASA	-0.039
kcatC3H3pCoSA	-0.038
kcatCald5h1FA	-0.037
kiuCOMTCA5HCald	-0.037
km4CL5pA	-0.033
kicCOMT5HCAld5HCAlc	-0.033
kiuCOMTCAFAld	-0.033

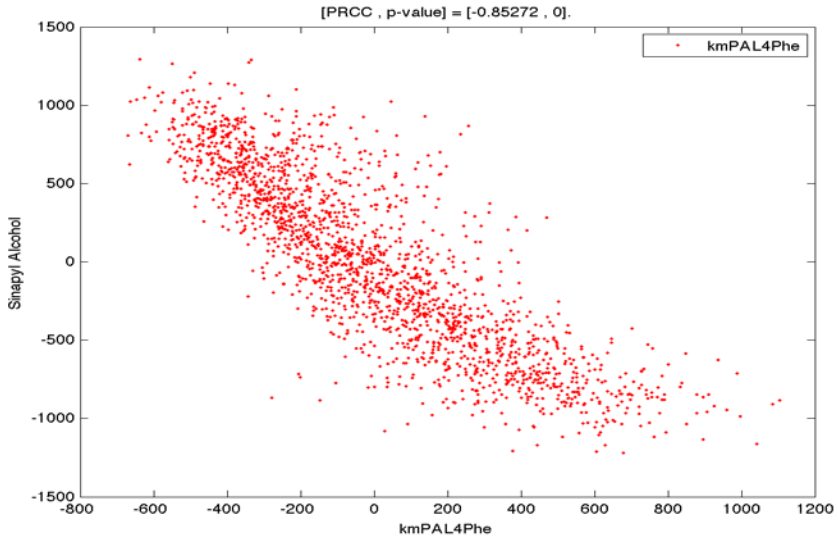

Figure 14 - The partial correlation plot of the parameter kmPAl4Phe and sinapyl alcohol concentration

References

F Chen, P Kota, J W Blount, R A Dixon (2001) Chemical syntheses of caffeoyl and 5-OH coniferyl aldehydes and alcohols and determination of lignin *O*-methyltransferase activities in dicot and monocot species. Phytochemistry, 58: 1035-1042.

Daubresse, N., Francesch, C., Mhamdi F., Rolando, C (1994). A mild synthesis of coumaryl, coniferyl, sinapyl aldehydes and alcohol. Synthesis, 369–371.

Beuerle T, Pichersky E (2002) Enzymatic synthesis and purification of aromatic coenzyme A ester. Anal Biochem 302: 305-312.

Shuford CM, Li QZ, Sun YH, Chen HC, Wang J, Shi R, Sederoff RR, Chiang VL, Muddiman DC (2012) Comprehensive quantification of monolignol-pathway enzymes in Populus trichocarpa by protein cleavage isotope dilution mass spectrometry. J Proteome Res in press.

Structural Alterations of Lignins in Transgenic Poplars with Depressed Cinnamyl Alcohol Dehydrogenase or Caffeic Acid O-Methyltransferase Activity Have an Opposite Impact on the Efficiency of Industrial Kraft Pulping ;2012.

Achnine L, EB Blancaflor, S Rasmussen, RA Dixon 2004 Colocalization of L-phenylalanine ammonia-lyase and cinnamate 4-hydroxylase for metabolic channeling in phenylpropanoid biosynthesis Plant Cell ;16:3098-3109.

Bate NJ, J Orr, W Ni, A Meromi, T Nadler-Hassar, PW Doerner, RA Dixon, CJ Lamb, Y Elkind 1994 Quantitative relationship between phenylalanine ammonia-lyase levels and phenylpropanoid accumulation in transgenic tobacco identifies a rate-determining step in natural product synthesis Proc.Natl.Acad.Sci.U.S.A. ;91:7608-7612.

Bevan M, D Shufflebottom, K Edwards, R Jefferson, W Schuch 1989 Tissue- and cell-specific activity of a phenylalanine ammonia-lyase promoter in transgenic plants EMBO J. ;8:1899-1906.

Bjurhager I, AM Olsson, B Zhang, L Gerber, M Kumar, LA Berglund, I Burgert, B Sundberg, L Salmen 2010 Ultrastructure and mechanical properties of populus wood with reduced lignin content caused by transgenic down-regulation of cinnamate 4-hydroxylase Biomacromolecules ;11:2359-2365.

Blee K, JW Choi, AP O'Connell, SC Jupe, W Schuch, NG Lewis, GP Bolwell 2001 Antisense and sense expression of cDNA coding for CYP73A15, a class II cinnamate 4-hydroxylase, leads to a delayed and reduced production of lignin in tobacco Phytochemistry ;57:1159-1166.

Blount JW, KL Korth, SA Masoud, S Rasmussen, C Lamb, RA Dixon 2000 Altering expression of cinnamic acid 4-hydroxylase in transgenic plants provides evidence for a feedback loop at the entry point into the phenylpropanoid pathway Plant Physiol. ;122:107-116.

Chen F, MS Srinivasa Reddy, S Temple, L Jackson, G Shadle, RA Dixon 2006 Multi-site genetic modulation of monolignol biosynthesis suggests new routes for formation of syringyl lignin and wall-bound ferulic acid in alfalfa (Medicago sativa L.) Plant J. ;48:113-124.

Coleman HD, JY Park, R Nair, C Chapple, SD Mansfield 2008 RNAi-mediated suppression of p-coumaroyl-CoA 3'-hydroxylase in hybrid poplar impacts lignin deposition and soluble secondary metabolism Proc.Natl.Acad.Sci.U.S.A. ;105:4501-4506.

Dwivedi UN, WH Campbell, J Yu, RS Datla, RC Bugos, VL Chiang, GK Podila 1994 Modification of lignin biosynthesis in transgenic Nicotiana through expression of an antisense O-methyltransferase gene from Populus Plant Mol.Biol. ;26:61-71.

Franke R, MR Hemm, JW Denault, MO Ruegger, JM Humphreys, C Chapple 2002a Changes in secondary metabolism and deposition of an unusual lignin in the ref8 mutant of Arabidopsis Plant J. ;30:47-59.

Franke R, JM Humphreys, MR Hemm, JW Denault, MO Ruegger, JC Cusumano, C Chapple 2002b The Arabidopsis REF8 gene encodes the 3-hydroxylase of phenylpropanoid metabolism Plant J. ;30:33-45.

Franke R, CM McMichael, K Meyer, AM Shirley, JC Cusumano, C Chapple 2000 Modified lignin in tobacco and poplar plants over-expressing the Arabidopsis gene encoding ferulate 5-hydroxylase Plant J. ;22:223-234.

Guo D, F Chen, K Inoue, JW Blount, RA Dixon 2001 Downregulation of caffeic acid 3-O-methyltransferase and caffeoyl CoA 3-O-methyltransferase in transgenic alfalfa. impacts on lignin structure and implications for the biosynthesis of G and S lignin Plant Cell ;13:73-88.

Hoffmann L, S Besseau, P Geoffroy, C Ritzenthaler, D Meyer, C Lapierre, B Pollet, M Legrand 2004 Silencing of hydroxycinnamoyl-coenzyme A shikimate/quinate hydroxycinnamoyltransferase affects phenylpropanoid biosynthesis Plant Cell ;16:1446-1465.

Hu WJ, SA Harding, J Lung, JL Popko, J Ralph, DD Stokke, CJ Tsai, VL Chiang 1999 Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees Nat.Biotechnol. ;17:808-812.

Huntley SK, D Ellis, M Gilbert, C Chapple, SD Mansfield 2003 Significant increases in pulping efficiency in C4H-F5H-transformed poplars: improved chemical savings and reduced environmental toxins J.Agric.Food Chem. ;51:6178-6183.

Jouanin L, T Goujon, V de Nadai, MT Martin, I Mila, C Vallet, B Pollet, A Yoshinaga, B Chabbert, M Petit-Conil, C Lapierre 2000 Lignification in transgenic poplars with extremely reduced caffeic acid O-methyltransferase activity Plant Physiol. ;123:1363-1374.

Kajita S, S Hishiyama, Y Tomimura, Y Katayama, S Omori 1997 Structural Characterization of Modified Lignin in Transgenic Tobacco Plants in Which the Activity of 4-Coumarate:Coenzyme A Ligase Is Depressed Plant Physiol. ;114:871-879.

Kajita S, Y Katayama, S Omori 1996 Alterations in the biosynthesis of lignin in transgenic plants with chimeric genes for 4-coumarate: coenzyme A ligase Plant Cell Physiol. ;37:957-965.

Kitin P, SL Voelker, FC Meinzer, H Beeckman, SH Strauss, B Lachenbruch 2010 Tyloses and phenolic deposits in xylem vessels impede water transport in low-lignin transgenic poplars: a study by cryo-fluorescence microscopy Plant Physiol. ;154:887-898.

Lee D, K Meyer, C Chapple, CJ Douglas 1997 Antisense suppression of 4-coumarate:coenzyme A ligase activity in Arabidopsis leads to altered lignin subunit composition Plant Cell ;9:1985-1998.

Li L, Y Zhou, X Cheng, J Sun, JM Marita, J Ralph, VL Chiang 2003 Combinatorial modification of multiple lignin traits in trees through multigene cotransformation Proc.Natl.Acad.Sci.U.S.A. ;100:4939-4944.

Maher EA, NJ Bate, W Ni, Y Elkind, RA Dixon, CJ Lamb 1994 Increased disease susceptibility of transgenic tobacco plants with suppressed levels of preformed phenylpropanoid products Proc.Natl.Acad.Sci.U.S.A. ;91:7802-7806.

Meyer K, A Shirley, JC Cusumano, DA Bell-Lelong, C Chapple 1998 Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America ;95:6619-6623.

Meyer K, JC Cusumano, C Somerville, CC Chapple 1996 Ferulate-5-hydroxylase from Arabidopsis thaliana defines a new family of cytochrome P450-dependent monooxygenases Proc.Natl.Acad.Sci.U.S.A. ;93:6869-6874.

Meyermans H, K Morreel, C Lapierre, B Pollet, A De Bruyn, R Busson, P Herdewijn, B Devreese, J Van Beeumen, JM Marita, J Ralph, C Chen, B Burggraeve, M Van Montagu, E Messens, W Boerjan 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.

Millar DJ, M Long, G Donovan, PD Fraser, AM Boudet, S Danoun, PM Bramley, GP Bolwell 2007 Introduction of sense constructs of cinnamate 4-hydroxylase (CYP73A24) in

transgenic tomato plants shows opposite effects on flux into stem lignin and fruit flavonoids Phytochemistry ;68:1497-1509.

Pincon G, S Maury, L Hoffmann, P Geoffroy, C Lapierre, B Pollet, M Legrand 2001 Repression of O-methyltransferase genes in transgenic tobacco affects lignin synthesis and plant growth Phytochemistry ;57:1167-1176.

Piquemal J, S Chamayou, I Nadaud, M Beckert, Y Barriere, I Mila, C Lapierre, J Rigau, P Puigdomenech, A Jauneau, C Digonnet, AM Boudet, D Goffner, M Pichon 2002 Down-regulation of caffeic acid o-methyltransferase in maize revisited using a transgenic approach Plant Physiol. ;130:1675-1685.

Ralph J, T Akiyama, H Kim, F Lu, PF Schatz, JM Marita, SA Ralph, MS Reddy, F Chen, RA Dixon 2006 Effects of coumarate 3-hydroxylase down-regulation on lignin structure J.Biol.Chem. ;281:8843-8853.

Rastogi S, UN Dwivedi 2006 Down-regulation of lignin biosynthesis in transgenic Leucaena leucocephala harboring O-methyltransferase gene Biotechnol.Prog. ;22:609-616.

Reddy MS, F Chen, G Shadle, L Jackson, H Aljoe, RA Dixon 2005a Targeted down-regulation of cytochrome P450 enzymes for forage quality improvement in alfalfa (Medicago sativa L.) Proc.Natl.Acad.Sci.U.S.A. ;102:16573-16578.

Reddy MS, F Chen, G Shadle, L Jackson, H Aljoe, RA Dixon 2005b Targeted down-regulation of cytochrome P450 enzymes for forage quality improvement in alfalfa (Medicago sativa L.) Proc.Natl.Acad.Sci.U.S.A. ;102:16573-16578.

Reddy MS, F Chen, G Shadle, L Jackson, H Aljoe, RA Dixon 2005c Targeted down-regulation of cytochrome P450 enzymes for forage quality improvement in alfalfa (Medicago sativa L.) Proc.Natl.Acad.Sci.U.S.A. ;102:16573-16578.

Ruegger M, K Meyer, JC Cusumano, C Chapple 1999 Regulation of Ferulate-5-Hydroxylase Expression in Arabidopsis in the Context of Sinapate Ester Biosynthesis. Plant Physiology ;119:101-110.

Schilmiller AL, J Stout, JK Weng, J Humphreys, MO Ruegger, C Chapple 2009 Mutations in the cinnamate 4-hydroxylase gene impact metabolism, growth and development in Arabidopsis Plant J. ;60:771-782.

Sewalt V, W Ni, JW Blount, HG Jung, SA Masoud, PA Howles, C Lamb, RA Dixon 1997a Reduced Lignin Content and Altered Lignin Composition in Transgenic Tobacco Down-Regulated in Expression of L-Phenylalanine Ammonia-Lyase or Cinnamate 4-Hydroxylase Plant Physiol. ;115:41-50.

Sewalt V, W Ni, JW Blount, HG Jung, SA Masoud, PA Howles, C Lamb, RA Dixon 1997b Reduced Lignin Content and Altered Lignin Composition in Transgenic Tobacco Down-Regulated in Expression of L-Phenylalanine Ammonia-Lyase or Cinnamate 4-Hydroxylase Plant Physiol. ;115:41-50.

Song J, Z Wang 2011 RNAi-mediated suppression of the phenylalanine ammonia-lyase gene in Salvia miltiorrhiza causes abnormal phenotypes and a reduction in rosmarinic acid biosynthesis J.Plant Res. ;124:183-192.

Stewart JJ, T Akiyama, C Chapple, J Ralph, SD Mansfield 2009 The effects on lignin structure of overexpression of ferulate 5-hydroxylase in hybrid poplar Plant Physiol. ;150:621-635.

Tsai CJ, JL Popko, MR Mielke, WJ Hu, GK Podila, VL Chiang 1998 Suppression of O-methyltransferase gene by homologous sense transgene in quaking aspen causes red-brown wood phenotypes Plant Physiol. ;117:101-112.

Voelker SL, B Lachenbruch, FC Meinzer, M Jourdes, C Ki, AM Patten, LB Davin, NG Lewis, GA Tuskan, L Gunter, SR Decker, MJ Selig, R Sykes, ME Himmel, P Kitin, O Shevchenko, SH Strauss 2010 Antisense down-regulation of 4CL expression alters lignification, tree growth, and saccharification potential of field-grown poplar Plant Physiol. ;154:874-886.

Voelker SL, B Lachenbruch, FC Meinzer, SH Strauss 2011 Reduced wood stiffness and strength, and altered stem form, in young antisense 4CL transgenic poplars with reduced lignin contents New Phytol. ;189:1096-1109.

VOELKER SL, B LACHENBRUCH, FC MEINZER, P KITIN, SH STRAUSS 2011; 2011 Transgenic poplars with reduced lignin show impaired xylem conductivity, growth efficiency and survival Plant, Cell Environ. ;34:655 <last_page> 668.

Wagner A, L Donaldson, H Kim, L Phillips, H Flint, D Steward, K Torr, G Koch, U Schmitt, J Ralph 2009 Suppression of 4-coumarate-CoA ligase in the coniferous gymnosperm Pinus radiata Plant Physiol. ;149:370-383.

Weng JK, H Mo, C Chapple 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.

Xu B, LL Escamilla-Treviño, N Sathitsuksanoh, Z Shen, H Shen, Y Percival Zhang, RA Dixon, B Zhao 2011; 2011 Silencing of 4-coumarate:coenzyme A ligase in switchgrass leads to reduced lignin content and improved fermentable sugar yields for biofuel production New Phytol. ;192:611 <last_page> 625.

Yang J, F Chen, O Yu, RN Beachy 2011 Controlled silencing of 4-coumarate:CoA ligase alters lignocellulose composition without affecting stem growth Plant Physiol.Biochem. ;49:103-109.

Zhong R, WH III, J Negrel, ZH Ye 1998 Dual methylation pathways in lignin biosynthesis Plant Cell ;10:2033-2046.

Chapter 4

Supporting Research

A novel O-methyl transferase-like gene with a drastic ectopic expression in response to tension wood formation in *Populus trichocarpa*

The following work was reprinted with permission from authors.

Published in Cellulose Chemistry & Technology (2007) 41:521-528

Ting-Feng Yeh, Jack Peng-Yu Wang, Rui Shi, Ying-Hsuan Sun, Vincent L. Chiang

North Carolina State University, Raleigh, NC 27695

Summary

Much of our knowledge on *S*-adenosyl-L-methionine-dependent *O*-methyltransferase (OMT) functions in lignin biosynthesis stems from the pioneering studies developed by Professor Takayoshi Higuchi. Caffeate OMT (COMT) mediating syringyl (S) monolignol biosynthesis is a key function evolved in angiosperms after gymnosperm speciation. The emergence of this function in angiosperms may be associated with abiotic stress adaptation, because angiosperm trees develop specialized, S-lignin enriched "tension wood" for counteracting mechanical stimuli. Although our knowledge on lignin is substantial, many aspects of S monolignol biosynthesis in tension wood are not yet quantified or sufficiently understood. In the *Populus trichocarpa* genome, 24 gene models encoding putative OMTs (PtcOMTs) were identified. Real-time PCR (RT-PCR) quantification of tissue-specific transcript abundances of these 24 genes revealed that 3 of them (*PtcCOMT1*, *2* and *3*) are preferentially expressed in the stem xylem of *P. trichocarpa*. The expression of the three xylem-specific *PtcCOMT*s is, however, only slightly up-regulated during tension wood formation. In contrast, the expression of a normally leaf-specific *OMT* becomes rapidly up-regulated by ~31-fold in the developing xylem of tension wood. The results suggest that this OMT, Ptc-TW-OMT, together with PtcCOMT1, 2 and 3, may be associated with the increased S-lignin formation in tension wood.

Introduction

Our current knowledge on monolignol biosynthesis stems largely from the classical tracer studies[1-8] performed by Brown, Neish and Higuchi in the 1950s, describing key metabolites and enzymes. It was also Higuchi *et al*.[9] who discovered, more than 40 years ago, a caffeate *O*-methyltransferase (COMT, EC 2.1.1.68) function involved in the lignin biosynthesis in bamboo. Subsequently, Higuchi led a series of systematic studies on OMT functions in monolignol biosynthesis and made the first revelation of distinct, species-specific bifunctional and mono-functional COMTs.[10-22] In angiosperms, COMTs are bi-functional and capable of methylating hydroxycinnamic acids, leading to the biosynthesis of both coniferyl alcohol and sinapyl alcohol, guaiacyl (G) and syringyl (S) monolignols, respectively. The COMTs in gymnosperms are monofunctional, mediating primarily the formation of G monolignol. These pioneering studies provided the basic knowledge for the first cloning of a gene encoding the bi-functional COMT from aspen (*Populus tremuloides*),[23] which also led to the discovery of an OMT function in catalyzing the methylation of 5- hydroxyconiferyl aldehyde,[24,25] an alternative OMT function previously suggested in the OMT studies of Higuchi.[26,27]

The emergence of bi-functional OMTs has been considered as an event in the angiosperm evolution adaptation to a longterm growth advantage.[13] A unique adaptation is the formation of “tension wood (TW)”, a specialized wood developed in the angiosperm tree stems or branches, to counteract mechanical stress/load.[28,29] TW synthesizes more cellulose and a lignin enriched with an S monolignol constituent.[30- 33] Because TW can be easily induced in a

tree stem or branch by bending, it therefore provides a convenient and rich source for a more profound understanding of genes and, particularly, of the genetic networks associated with the *OMT*s that regulate S monolignol biosynthesis. Indeed, the expression profile of monolignol genes during the development of TW has been investigated[34-36] by microarrays based primarily on the information of expressed sequence tags (ESTs). However, no consensus is still reached as to the interpretation of the transcripts identified exclusively on the basis of ESTs, as some genes involved in a particular process may not have been identified and would not be present on the arrays.[37] Meaningful gene expression studies begin with the knowledge of all genes of interest. Thus, tree species lacking full genome sequence are inadequate for such investigation.

The present study focuses on the expression response of all possible *COMT*s homologues to TW formation. The model system employed is *P. trichocarpa* Nisqually-1, the genotype whose full genome has been sequenced (http://genome.jgipsf.org/Poptr1_1/Poptr1_1.home.html). A systematic, genome-wide analysis of these genes was carried out, as to their phylogenetic relationship and quantitative transcript abundances in various tissues, to identify those most likely associated with monolignol biosynthesis and those with TW formation.

Materials and Methods

Sequence analysis of *COMT* family members

PtcCOMT homologues were identifiedfrom the currently available 45,555 *P. trichocarpa* gene models (http://genome.jgi-psf.org/Poptr1_1/Poptr1_1.home.html), based on the 16 annotated Arabidopsis (*Arabidopsis thaliana*) COMT (AtCOMT) protein sequences, with values below 1E-05, using the BLASTX program.[38] The redicted protein sequences of the BLASTX-identified gene models were further screened for completeness by multiple sequence alignment based on the Arabidopsis protein sequences. The *AtCOMT* family members were obtained from the GenBank, on the basis of The Arabidopsis Information Resource (TAIR; http://www.arabidopsis.org). Phylogenetic trees and their associated bootstrap values were analyzed with the ClustalW multiple sequence alignment program.[39] The neighbour-joining trees were created based on the distance matrices derived from the results of multiple sequence alignments, using the default settings.

Gene specific primer design

Based on the identified *PtcCOMT* gene models, PCR primers were designed so that the sequences of each set of primers would match perfectly with the target *PtcCOMT* sequence but differ in at least three nucleotides from the sequences of all the other family members (Table 1). Each designed primer set was expected to amplify a PCR product from a specific exon with a size of approximately 100 bp in length.

Table 1 - Unique primer pairs for each individual gene

Gene ID	Forward primer (5’ to 3’)	Reverse primer (5’ to 3’)
COMT834247	TCTTGAAGAATTGCTATGACGCCT	GAATGCACTCAACAAGTATCACCTTG
COMT824484	AGCACAATCGTCTCCAAGTACCCT	AACATTCTCCACACCAGGGAAAGC
COMT731466	CAGTTGCCCACAAAGAACCCAGAT	CATCCACAGAGCAACCAAGCACAT
COMT552360	TCCAGAGGATGGGAAGGTAATTGTT	CTTTCCCTCCTGGGTTTTGAGAC
COMT799151	TCCAGAGGATGGGAAGGTAATTGTA	CTTTCCCTCCTGGGTTTTGAGTT
COMT806142	CCAGTATGCTATGCAACTTTCAAGT	TATCTGAAGCTGAGAGCAAGGCATCT
COMT836247	GATGCTGAAGGCAATGTCAGGTGT	TGCACTCCCACCTACATCAACCAA
COMT811844	ACCAGCCCATTACCCTTCAACAGT	AAGAAACCGGAATGCACCAGTACG
COMT758840	ACCCTTCCACAGTTGGTTTCAGCA	ACACCAGAATGCACCAGTATACGC
COMT555739	TGGTACCGGTTCACTTTCCAGGAT	ATGACTTGTGGAAGCTCAAGCACC
COMT586060	ACATAATCCACAACCATGGCAGGC	GATGCGCATGAGCCGGTATATGCA
COMT780751	ACTATGGTGGCCATGAACCAATGC	AACCCAACCCTTCAAACACCTCCT
COMT829076	CCTACCGATCCACCCATCAAAAG	AAGATTTTGCTGGGAAAAGAAACCC
COMT740577	ATTGCTTGAGCACTTGGTTCCG	TTGGCCAGCGTACTCCCAAAG
COMT669875[a]	TGACTGGAGTGATGAAGACTGCGT	ATCACCTTCCCAGCTTTCTGCTGT
COMT588324[a]	TGACTGGAGTGATGAAGACTGCGT	ATCACCTTCCCAGCTTTCTGCTGT
COMT588323[a]	TGACTGGAGTGATGAAGACTGCGT	ATCACCTTCCCAGCTTTCTGCTGT
COMT828771	TCAGGCAGTTCCTCCAGCAGAA	ACTCATAATTGCTTGTCTGCATCGTC
COMT676887	CAAAATCTTAACCAAGAAGCCTATTCT	TCATACTCAAGGGGTTATCCTTGAGG
COMT582798	CTAGCAAGCTGGTTTTAAGCCAATGC	CTGCGATGCCCTTGGCCATG
COMT280232	TTGTTGTCCTCGACCTCTTCTGGT	TGACTCATCGAAGGAGGATTTCAGCG
COMT597596[b]	TGGAGGAGAAACACTCTTTGGGCT	GCTGCTAGGAACCATGAACTGTGT
COMT571762[b]	TGGAGGAGAAACACTCTTTGGGCT	GCTGCTAGGAACCATGAACTGTGT
COMT8277072	ACTGTCGGAAAGCAATGCCTGAGA	TTGGCCATTTCCTTCTGGTTGCAG

Tissue collection and mechanically-induced tension wood formation

Fully expended leaf (without midrib), stem-developing phloem, stem-developing xylem, and young shoot tip (one to three internodes) were harvested from 7-month-old *P. trichocarpa* Nisqually-1 trees grown in our greenhouse and stored immediately in liquid nitrogen for tissue specific expression analysis.

Twenty-five 7-month-old clonal propagules of *P. trichocarpa* were bent into an arch of 45º from vertical for 7 weeks to allow complete TW development. Differentiating xylem tissues from lower quarter of these bent stems were collected at 1, 3, 7, 21, 35, and 49 days by scraping the xylem tissue directly into liquid nitrogen with razor blades and stored in liquid nitrogen. NW developing xylem tissues were collected from trees without bending treatment (time 0).

RNA isolation and real-time RT-PCR analysis of gene expression

Total RNAs were isolated from the above described tissues, treated with RNase-free DNase I, then purified by a RNeasy plant RNA isolation kit, the RNA quality being examined by UV spectrogram scan and gel electrophoresis, as in previous studies.[40] Total RNA (200 ng) was reversely transcribed, by using TagMan reverse transcription reagents (Applied Biosystems, Roche). Real-time PCR was conducted with an Applied Biosystems 7900HT Sequence Detection System. For each reaction, a 25-μL mixture contained the first cDNA strand (equivalent to 1- 5 ng of total RNA), 5 pmol each of the forward and reverse primer,

and 12.5 μL 2 × SYBR green PCR master mix. The amplification program was as follows: 95 $_{o}$C for 10 min, then 40 cycles of 95$_{o}$C for 15 sec and 60$_{o}$C for 1 min, after which a thermal denaturing cycle was added, to determine the dissociation curve of the PCR products for checking the amplification specificity. A formula for absolute quantification of the transcript copy numbers per unit mass of total RNA was derived, similarly to previous publication.[40] For relative quantification, each reaction was repeated at least three times and the transcript level was normalized to that of 18s rRNA. Such normalized values allowed the comparison of the expression levels of different genes.

Results and Discussion

Tissue-specific expression patterns of *P. trichocarpa COMT* genes

24 full protein sequences were identified in the *P. trichocarpa* genome (http://genome.jgi-sf.org/Poptr1_1/Poptr1_1home.html), homologous to the annotated Arabidopsis COMT proteins. Multiple sequence alignment analysis shows that these 24 *PtcCOMT* genes share 20% to 100% protein sequence homology with each other (Table 2). The three different COMT588324, COMT669875 and COMT588323 genes encode an identical protein, suggesting gene duplication events for redundant functions. Similarly, COMT571762 and COMT 597596 also share identical protein sequences.

To identify the xylem specific *COMT* genes, the expression level of all possible full length *COMT* family genes in different tissues of *P. trichocarpa* were profiled, with real-time RT-

PCR. The tissue-specific expression analysis clearly indicated that COMT834247, COMT824484 and COMT827707 are most xylem-specific, suggesting that they encode monolignol biosynthesis COMTs (named as *PtcCOMT1*, *PtcCOMT2* and *PtcCOMT3*, respectively). Among them, *PtcCOMT1* exhibits the highest expression. *PtcCOMT1* is actually the most homologous one (99% homology) to a *P. tremuloides PtOMT1* that had been characterized in detail as to its biochemical function.[23,24] It also shares a 99% protein sequence homology with a hybrid poplar (*P. trichocarpa* × *P. deltoides*) *OMT*, also previously characterized to encode a xylemspecific COMT.[41] The functions of *PtcCOMT2* and *PtcCOMT3* are unknown, although the *PtcCOMT3* shares a moderate (53%) sequence homology with a xylemspecific loblolly pine (*Pinus taeda*) AEOMT,[42] which exhibits both COMT and caffeoyl-CoA 3-*O*-methyltransferase (CCoAOMT, EC 2.1.1.104) activities. It will be interesting to test the functional correlations of *PtcCOMT2* and *PtcCOMT3* with those of *PtcCOMT1* in regulating the methylation functions involved in monolignol biosynthesis.

For the first time, a highly shoot tipspecific *COMT* gene, the COMT597596 (or COMT571762), was also identified. Being highly specific in internodes one to three, this *COMT* is likely involved in monolignol biosynthesis during the stem primary growth, where only guaiacyl lignin is formed.[25,43] This proposition of the *COMT* gene function still needs to be verified. A few *COMT*s, more abundant in leaf tissues, were also found out (Fig. 1). They may be more related to methylation functions for leaf metabolites, such as phytoalexins and flavonoids.[44]

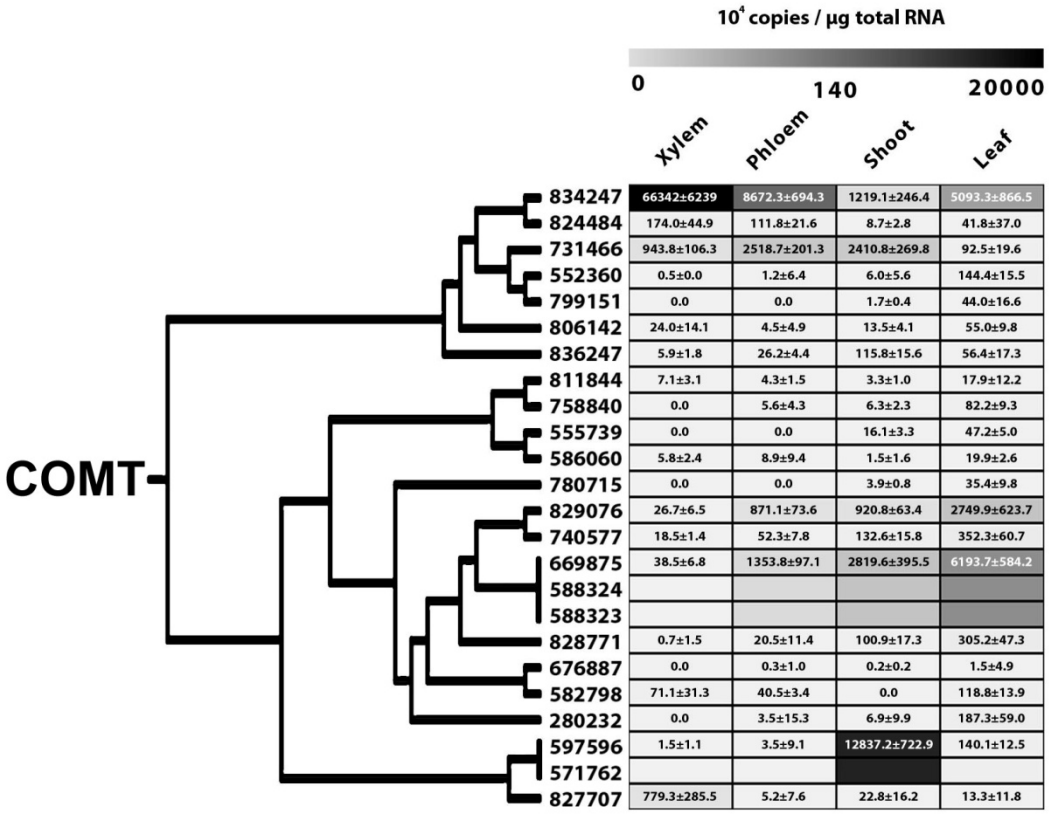

Figure 1 - Dendrogram of 24 COMT family genes of P. trichocarpa and their transcript abundance in different tissues (mean±SD, n=3).

Table 2 - Sequence homology of 24 identified full length COMT protein of *P. trichocarpa*

Gene ID	552360	799151	731446	806142	824484	834247	836247	571762	597596	827707	555739	586060	758840	811844	780715	582798	676887	588323	588324	669875	740577	829076	828771
280232	23	25	23	20	23	23	23	27	27	32	40	39	39	40	46	70	70	60	70	70	70	69	67
552360		97	85	51	54	54	36	34	34	36	28	28	27	27	26	28	28	34	30	30	30	30	29
799151			85	52	55	56	37	34	34	37	30	30	30	29	28	30	30	35	32	32	31	32	31
731466				52	55	57	36	33	33	35	28	28	27	27	28	29	29	34	30	30	29	30	29
806142					52	52	36	31	31	32	27	26	27	28	24	26	26	30	26	26	26	27	26
824484						90	37	35	35	37	29	29	30	32	28	30	30	36	30	30	30	31	30
834247							37	34	34	36	30	28	29	30	28	29	30	36	30	30	30	31	30
836247								31	31	32	25	26	27	27	25	27	27	29	28	28	27	27	26
571762									100	67	34	34	34	35	34	35	36	37	35	35	34	35	34
597596										67	34	34	34	35	34	35	36	37	35	35	34	35	34
827707											34	34	37	38	37	39	38	38	38	38	37	38	36
555739												89	55	52	49	51	52	52	51	51	52	53	52
586060													55	53	50	52	52	50	51	51	51	52	52
758840														88	50	50	50	48	49	49	48	49	50
811844															51	51	50	48	50	50	48	50	50
780715																61	60	63	61	61	60	61	60
582798																	97	83	87	87	86	87	86
676887																		82	86	86	86	88	86
588323																			100	100	89	90	88
588324																				100	91	92	90
669875																					91	92	90
740577																						95	88
829076																							90

Profiles of the expression responses of *P. trichocarpa COMT* genes to TW formation

In a subsequent step, the expression response of the *COMT* genes to TW formation in the stem developing xylem was examined. The TW formation was induced by bending. Because transcripts of COMT588324, COMT669875 and COMT588323 are identical and therefore cannot be distinguished, their expression responses were not analyzed. For the same reason, no examination of the expression responses of COMT571762 and COMT 597596 was possible. Thus, the 19 *COMT* genes were tested. Generally, the expression of 18 of the tested *COMT* genes was affected to various extents by the bending stress. The expression of COMT676887 was found nonresponsive to tension stress. For the sake of clarity, these 18 genes were grouped for further discussion.

Figure 2 shows the profile of the expression of the normally xylem-specific *COMT*s, *PtcCOMT1*, *PtcCOMT2* and *PtcCOMT3*, in TW, relatively to their expression in NW. While the expression of *PtcCOMT1* and *PtcCOMT3* was slightly reduced in response to the tension stress, the *PtcCOMT2* expression under stress was conspicuously up-regulated ~3-fold. Upregulation of this *PtcCOMT2* gene expression occurred during the first three days of bending, being maintained essentially throughout the entire bending period (7 weeks).

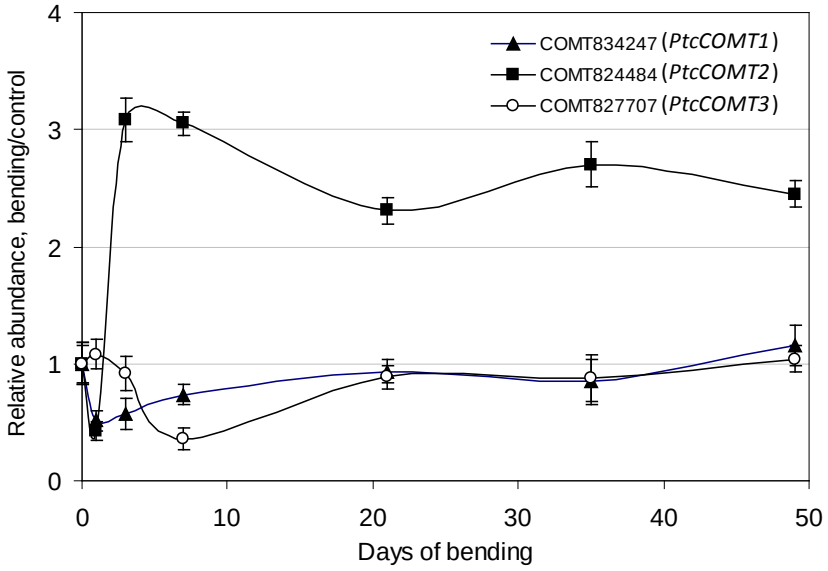

Figure 2 – Relative abundances of xylem-specific COMT genes of *P. trichocarpa* under mechanical stresses (mean±SD, n=3).

The *PtcCOMT2* may therefore be a key candidate gene responsible for the commonly believed increase in the lignin S/G ratio in TW, while the other two xylem-specific *COMT* genes may serve to maintain a basal level of the S/G ratio. It is also likely that a compensatory interplay of these gene functions exists for a coordinated upregulation of the lignin S/G ratio.

Interestingly, a group of normally nonxylem specific genes exhibited elevated xylem-expression in response to the tension stress (Fig. 3). They all first exhibited a rapid but minor up-regulated expression response to tension stress and then reached a peak expression induction after a 3-week bending. The expression of two other more leafspecific *COMT*

genes (COMT811844 and COMT555739) also increased to a peak value after a 3-week bending, although the increase was already obvious after a 1-week bending, as shown in Figure 4. These two genes, showing similar expression responses to the bending stress, are phylogentically related (Fig. 1). Figure 4 also shows another group of phylogentically-related *COMT* genes (COMT731466, COMT806142 and COMT836247) whose expression was found slightly suppressed most of the times throughout the entire bending.

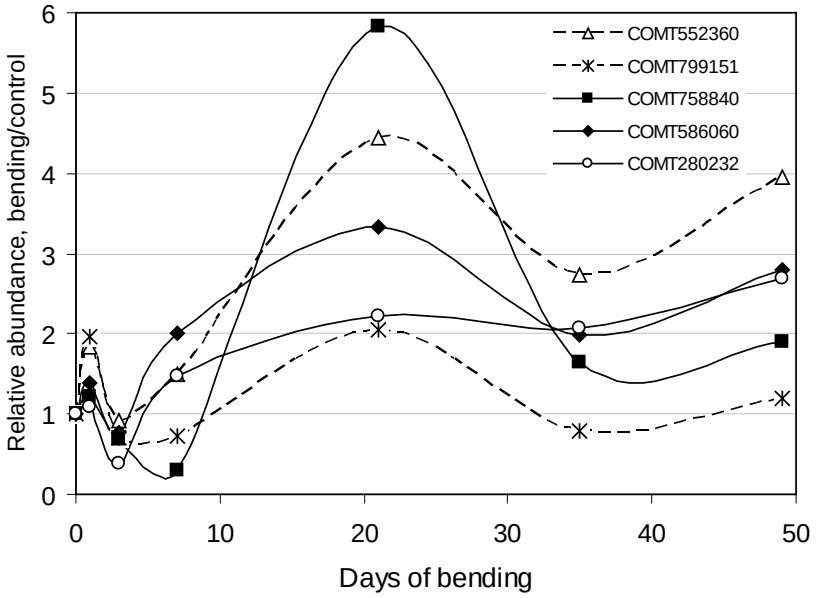

Figure 3 - Relative abundances of COMT genes of *P. trichocarpa* under mechanical stresses, group 1.

The exception is that a rapid expression increase was observed for COMT731466 before the expression suppression. The initial spike in the expression response to tension stress was also observed for a group of more leaf-specific *COMT* genes (Fig. 5). The most striking observation was the remarkable ~31- fold rapid increase (Fig. 5, 24-hr bending) in the expression in xylem of a *COMT* gene (COMT829076) that is normally abundant in leaves but deficient in xylem (Fig. 1).

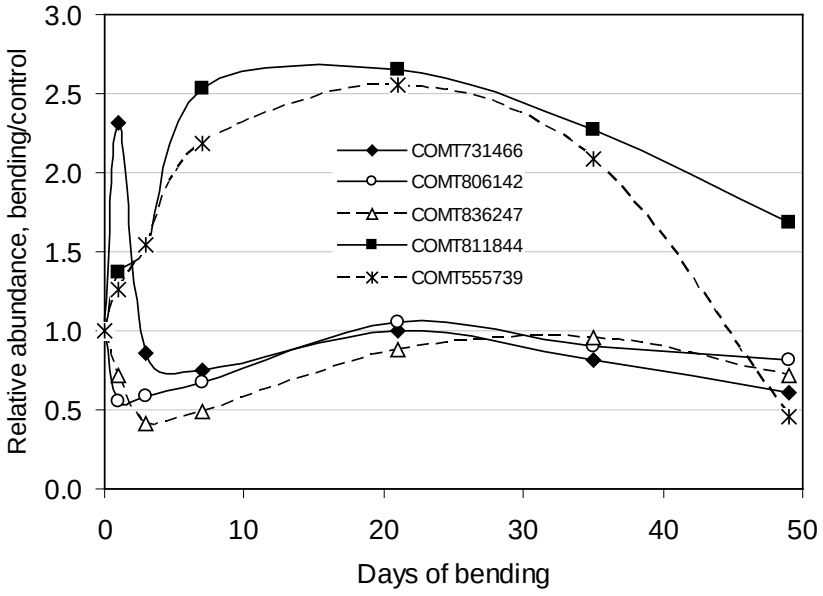

Figure 4 - Relative abundances of COMT genes of *P. trichocarpa* under mechanical stresses, group 2 and group 3.

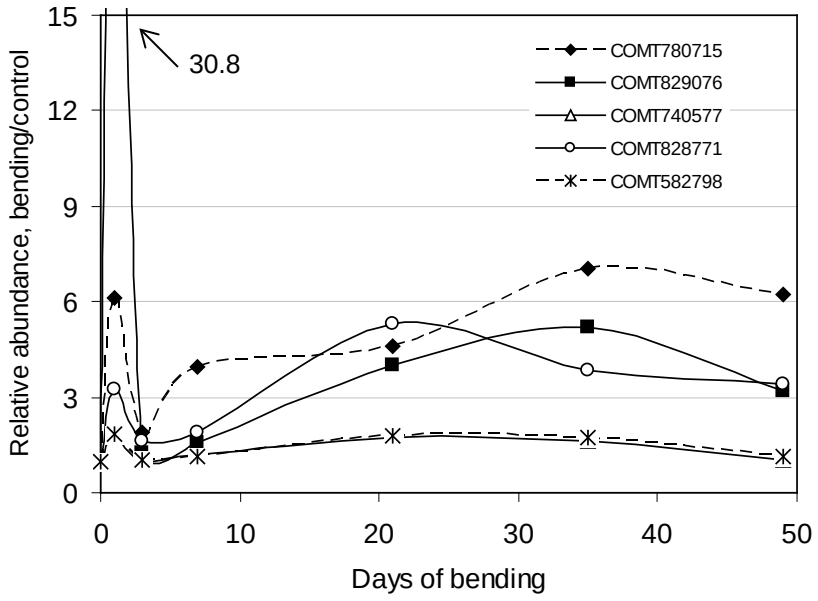

Figure 5 - Relative abundances of COMT genes of *P. trichocarpa* under mechanical stresses, group 4.

The tension stress-induced ectopic expression of this gene, named *Ptc-TWCOMT*, is a novel observation. Although such an ectopic expression was the most pronounced in *Ptc-TW-COMT* (COMT 829076), it also occurred to other tested *COMT* genes, in either early (Fig. 6) or later stages of the stress treatment (Figs. 3-4), which might suggest that certain common *cis*-regulatory elements in these *COMT* genes are receptive to the bending stress signals for gene expression activation and that their interplays may coordinate the onoff switches for gene activation.

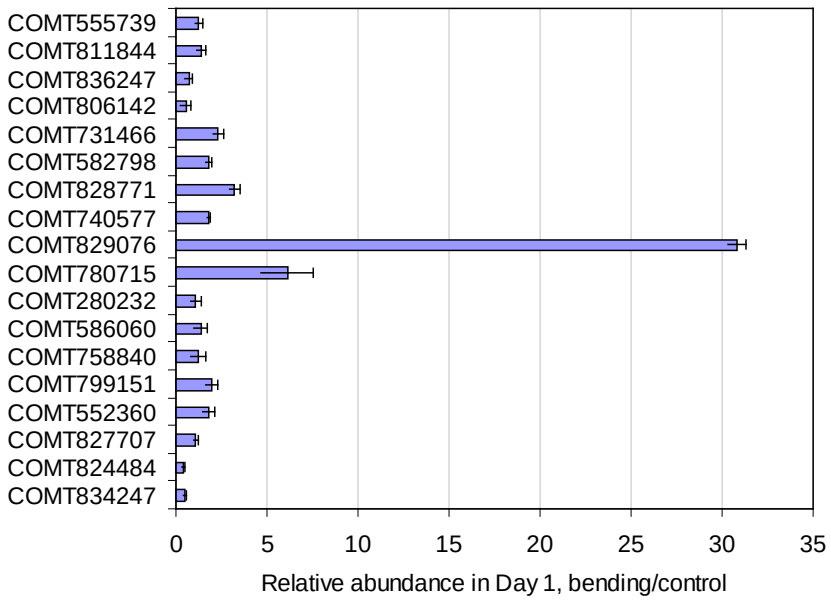

Figure 6 – Relative abundance of *PtcCOMT*s in one day of bending (mean±SD, n=3).

Detailed analysis of motifs in these genes may contribute to a better understanding of this on-off control and, consequently, of the control related to the methylation involved in monolignol biosynthesis. It is likely that the observed ectopic expression of *Ptc-TWCOMT* may provide an early boost of S monolignol biosynthesis for a high S/G ratio in TW lignin. Knowledge on the transcription control mechanisms, as well as the biochemical function of *Ptc-TW-COMT* are essential for a better understanding of S monolignol and thus of lignin biosynthesis in plant growth and adaptation.

CONCLUSIONS

The study was focused on *COMT* genes with annotated functions in phenylpropanoid biosynthesis pathway. 24 such *COMT* genes were identified in the genome of *P. trichocarpa*. *PtcCOMT1*, *PtcCOMT2* and *PtcCOMT3* are possibly associated with monolignol biosynthesis because of their xylem-specific expression patterns. However, only *PtcCOMT1* is known for its authentic functions in monolignol biosynthesis, on the basis of its biochemical functions of its ortholog in *P. tremuloides*.[23,24] Such a genome-based study, revealing the tissuespecific abundances of all possible involving *COMT*s and a possible key S monolignol biosynthetic gene, the *Ptc-TW-COMT*, provides a comprehensive platform for the development of a rigorous understanding of the methylation functions in monolignol biosynthesis.

REFERENCES

S. A. Brown and A. C. Neish, Shikimic acid as a precursor in lignin biosynthesis, *Nature*, **175**, 688(1955).

S. A. Brown and A. C. Neish, Studies of lignin biosynthesis using isotopic carbon. IV. formation from some aromatic monomers, *Canadian Journal of Biochemistry and Physiology*, **33**, 948(1955).

S. A. Brown and A. C. Neish, Studies of lignin biosynthesis using isotopic carbon. V. comparative studies on different plant species, *Canadian Journal of Biochemistry and Physiology*, **34**, 769(1956).

S. A. Brown and A. C. Neish, Studies of lignin biosynthesis using isotopic carbon. VIII. isolation of radioactive hydrogenolysis products of lignin, *Journal of the American Chemical Society*, **81**, 2419(1959).

T. Higuchi, Studies on the biosynthesis of lignin, in "Biochemistry of Wood", edited by Forth Internation Congress of Biochemistry, Pergamon Press, 1959, 161-188.

T. Higuchi and S. A. Brown, Studies of lignin biosynthesis using isotopic carbon. XII. the biosynthesis and metabolism of sinapic acid, *Canadian Journal of Biochemistry and Physiology*, **41**, 613(1963).

T. Higuchi and S. A. Brown, Studies of lignin biosynthesis using isotopic carbon. XIII. the phenylpropanoid system in lignification, *Canadian Journal of Biochemistry and Physiology*, **41**, 621(1963).

D. Wright, S. A. Brown and A. C. Neish, Studies of lignin biosynthesis using isotopic carbon. VI. formation of the side chain of the phenylpropane monomer, *Canadian Journal of Biochemistry and Physiology*, **36**, 1037(1958).

T. Higuchi, M. Shimada and H. Ohashi, Role of O-methyltransferase in lignification of bamboo, *Agricultural and Biological Chemistry*, **31**, 1459(1967).

T. Higuchi, "Biosynthesis and Biodegradation of Wood Components", Academic Press, 1985.

T. Higuchi, "Biochemistry and Molecular Biology of Wood", Springer, 1997.

T. Higuchi, M. Shimada, F. Nakatsubo and M. Tanahashi, Differences in biosynthesis of guaiacyl and syringyl lignins in woods, *Wood Science and Technology*, **11**, 153(1977).

H. Kuroda and T. Higuchi, O-Methyltransferase as a tool to evaluate the lignin evolution, *Wood Research*, **68**, 1(1982).

H. Kuroda, M. Shimada and T. Higuchi, Purification and properties of O-methyltransferase involved in the biosynthesis of gymnosperm lignin, *Phytochemistry*, **14**, 1759(1975).

H. Kuroda, M. Shimada and T. Higuchi, Characterization of a lignin-specific O-methyltransferase in aspen wood, *Phytochemistry*, **20**, 2635(1981).

Y. Nakamura, H. Fushiki and T. Higuchi, Metabolic differences between gymnosperms and angiosperms in the formation of syringyl lignin, *Phytochemistry*, **13**, 1777(1974).

M. Shimada, H. Fushiki and T. Higuchi, O-Methyltransferase activity from Japanese black pine, *Phytochemistry*, **11**, 2657(1972).

M. Shimada, H. Fushiki and T. Higuchi, O-Methyltransferase as a key enzyme in the biosynthesis of guaiacyl and syringyl lignins, *Mokuzai Gakkaishi*, **18**, 43(1972).

M. Shimada, H. Fushiki and T. Higuchi, Mechanism of biochemical formation of the methoxyl groups in softwood and hardwood lignins, *Mokuzai Gakkaishi*, **19**, 13(1973).

M. Shimada and T. Higuchi, Properties of O-methyltransferase from bamboo shoot, *Mokuzai Kenkyu*, **50**, 19(1970).

M. Shimada, H. Kuroda and T. Higuchi, Evidence for the formation of methoxyl groups of ferulic and sinapic acids in *Bambusa* by the same O-methyltransferase, *Phytochemistry*, **12**, 2873(1973).

M. Shimada, H. Ohashi and T. Higuchi, O-Methyltransferases involved in the biosynthesis of lignins, *Phytochemistry*, **9**, 2463(1970).

R. C. Bugos, V. L. C. Chiang and W. H. Campbell, cDNA cloning, sequence analysis and seasonal expression of lignin-bispecific caffeic acid / 5-hydroxyferulic acid O-methyltransferase of Aspen, *Plant Molecular Biology*, **17**, 1203(1991).

L. G. Li, J. L. Popko, T. Umezawa and V. L. Chiang, 5-Hydroxyconiferyl aldehyde modulates enzymatic methylation for syringyl monolignol formation, a new view of monolignol biosynthesis in angiosperms, *Journal of Biological Chemistry*, **275**, 6537(2000).

K. Osakabe, C. C. Tsao, L. G. Li, J. L. Popko, T. Umezawa, D. T. Carraway, R. H. Smeltzer, C. P. Joshi and V. L. Chiang, Coniferyl aldehyde 5-hydroxylation and methylation direct syringyl lignin biosynthesis in angiosperms, *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 8955(1999).

T. Higuchi, Pathways for monolignol biosynthesis via metabolic grids: coniferyl aldehyde 5-hydroxylase, a possible key enzyme in angiosperm syringyl lignin biosynthesis, *Proceedings of the Japan Academy Series B-Physical and Biological Sciences*, **79**, 227(2003).

T. Higuchi, Look back over the studies of lignin biochemistry, *Journal of Wood Science*, **52**, 2(2006).

G. Scurfield, Reaction wood: its structure and function, *Science*, **179**, 647(1973).

A. B. Wardrop and G. W. Davies, The structure of reaction wood: The structure and differentiation of compression wood, *Australian Journal of Botany* **12**, 24(1964).

B. J. Fergus and D. A. I. Goring, Location of guaiacyl and syringyl lignins in birch xylem tissue, *Holzforschung*, **24**, 113(1970).

B. J. Fergus and D. A. I. Goring, Distribution of lignin in birch wood as determined by ultraviolet microscopy, *Holzforschung*, **24**, 118(1970).

Y. Musha and D. A. I. Goring, Distribution of syringyl and guaiacyl moieties in hardwoods as indicated by ultraviolet microscopy, *Wood Science and Technology*, **9**, 45(1975).

S. Saka and D. A. I. Goring, Localization of lignin in wood cell walls, in "Biosynthesis and Biodegradation of Wood Components", edited by T. Higuchi, Academic Press, 1985, 141-160.

S. Andersson-Gunneras, E. J. Mellerowicz, J. Love, B. Segerman, Y. Ohmiya, P. M. Coutinho, P. Nilsson, B. Henrissat, T. Moritz and B. Sundberg, Biosynthesis of cellulose-enriched tension wood in *Populus*: global analysis of transcripts and metabolites identifies biochemical and developmental regulators in secondary wall biosynthesis, *Plant Journal*, **45**, 144(2006).

E. Paux, V. Carocha, C. Marques, A. M. de Sousa, N. Borralho, P. Sivadon and J. Grima-Pettenati, Transcript profiling of *Eucalyptus* xylem genes during tension wood formation, *New Phytologist*, **167**, 89(2005).

R. Whetten, Y. H. Sun, Y. Zhang and R. Sederoff, Functional genomics and cell wall biosynthesis in loblolly pine, *Plant Molecular Biology*, **47**, 275(2001).

J. Quackenbush, Microarrays - Guilt by association, *Science*, **302**, 240(2003).

S. F. Altschul, T. L. Madden, A. A. Schaffer, J. H. Zhang, Z. Zhang, W. Miller and D. J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Research*, **25**, 3389(1997).

J. D. Thompson, D. G. Higgins and T. J. Gibson, CLUSTAL-W - Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Research*, **22**, 4673(1994).

S. Suzuki, L. G. Li, Y. H. Sun and V. L. Chiang, The cellulose synthase gene superfamily and biochemical functions of xylem-specific cellulose synthase-like genes in *Populus trichocarpa*, *Plant Physiology*, **142**, 1233(2006).

J. Van Doorsselaere, B. Dumas, M. Baucher, B. Fritig, M. Legrand, M. Vanmontagu and D. Inze, One-step purification and characterization of a lignin-specific O-methyltransferase from poplar, *Gene*, **133**, 213(1993).

L. G. Li, J. L. Popko, X. H. Zhang, K. Osakabe, C. J. Tsai, C. P. Joshi and V. L. Chiang, A novel multifunctional O-methyltransferase implicated in a dual methylation pathway associated with lignin biosynthesis in loblolly pine, *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 5461(1997).

L. G. Li, X. F. Cheng, J. Leshkevich, T. Umezawa, S. A. Harding and V. L. Chiang, The last step of syringyl monolignol biosynthesis in angiosperms is regulated by a novel gene encoding sinapyl alcohol dehydrogenase, *Plant Cell*, **13**, 1567(2001).

C. P. Joshi and V. L. Chiang, Conserved sequence motifs in plant S-adenosyl-L-methionine-dependent methyltransferases, *Plant Molecular Biology*, **37**, 663(1998).