PHENYLPROPANOID METABOLISM IN POPLAR: CHARACTERIZATION OF THE 4-COUMARATE:CoA LIGASE GENE FAMILY AND A PUTATIVE HYDROXYCINNAMOYL-CoA THIOESTERASE

by

SANDRA MARY ALLINA

B.Sc., The University of British Columbia, 1992

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Department of Botany)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

November, 1998

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of BOTANY

The University of British Columbia
Vancouver, Canada

Date November 13, 1998

ABSTRACT

The work presented in this thesis is part of a larger study designed to examine the possibility of altering lignin content and/or composition by genetic engineering. Hybrid poplar was used as a model tree system. Lignin precursors (cinnamyl alcohols) are derived from phenylpropanoid metabolism. 4-coumarate:CoA ligase (4CL) activity is required for the activation of hydroxycinnamic acids prior to their being fed into the lignin specific branch pathway. The purpose of this research was to test the hypothesis that multiple genes encode 4CL isoforms that are differentially expressed and preferentially utilize one or more methoxylated hydroxycinnamic acid substrate.

4CL is encoded by a gene family in hybrid poplar. Two full-length cDNA clones (4CL-216 and 4CL-9) representing two different genes (*4CL1* and *4CL2*) were isolated from a young leaf library and characterized. Southern blot analysis demonstrated that other *4CL* genes exist in the poplar genome. Northern blot analysis demonstrated that *4CL1* and *4CL2* are differentially expressed. *4CL1* is preferentially expressed in old leaves, while *4CL2* is preferentially expressed in young leaves. Neither gene is expressed at high levels in developing xylem and they may not play a major role in lignification of this tissue.

The *4CL* cDNA clones were expressed in a insect cell/baculovirus system. Partially purified recombinant 4CL proteins were tested for substrate preference. Both of the recombinant 4CL proteins used hydroxycinnamic acid substrates in a similar manner. The substrate preferences were similar to those found for each of the partially purified native 4CL isoforms (Allina et al., 1998). This suggests that 4CL isoforms with large differences in their substrate utilization profiles do not, in fact, exist in poplar.

The recombinant 4CL proteins use cinnamic acid while the native proteins do not. It was hypothesized that a cellular factor was responsible for modifying native 4CL such that it could no longer use cinnamic acid as a substrate. However, the results of this study strongly suggest that 4CL activity is not being modified but rather, that a thioesterase activity is responsible for the observed phenomena. The thioesterase is able to use all hydroxycinammoyl-CoA thioesters as substrates. This was previously masked by the fact that recombinant 4CL could convert most hydroxycinnamic acids to hydroxycinnamoyl-CoA thioesters faster than the thioesterase could perform the reverse reaction. The possible role of a hydroxycinnamoyl-CoA thioesterase in phenylpropanoid metabolism is discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

4CL	4-coumarate:coenzyme A ligase
AFLP	amplified fragment length polymorphisms
C4H	cinnamate-4-hydroxylase
CAD	cinnamyl alcohol dehydrogenase
CHI	chalcone isomerase
CHS	chalcone synthase
CoA	coenzyme A
CoCOMT	caffeoyl-CoA 3-*O*-methyltransferase
COMT	bi-functional caffeic acid *O*-methyltransferase
CCR	cinnamoyl:CoA reductase
EST	expressed sequence tags
F5H	ferulate 5-hydroxylase
GUS	β-glucuronidase
HPLC	high performance liquid chromatography
MEC	multi-enzyme complex
nkat	nanokatals - nmoles of substrate converted to product per second
pkat	picokatals - pmoles of substrate converted to product per second
PAGE	polyacrylamide gel electrophoresis
PAL	phenylalanine ammonia-lyase
PCR	polymerase chain reaction
PIPES	piperazine-N,N'-bis[ethanaesulfonic acid]
RAPD	random amplified polymorphic DNA
RFLP	restriction fragment length polymorphism
RT-PCR	reverse transcriptase-polymerase chain reaction
SDS	sodium dodecyl sulphate
STS	stilbene synthase
TLC	thin layer chromatography
TRIS	tris(hydroxymethyl)aminomethane

ACKNOWLEDGMENTS

This thesis is the result of the encouragement and support shown to me by many people and I want to express my gratitude to them.

A big thank you goes to Dr. Carl Douglas who accepted me into his lab as an undergrad and inspired me to become a scientist. He taught me to do the very best job that I could do and supported my choices. He also introduced me to the mountains. Another big thank you goes to Dr. Brian Ellis, whose good humour, patience and encouragement has allowed me to become an independent thinker and to trust in my own judgment.

Dr. David Theilmann generously allowed me perform much of the baculovirus/insect cell work in his lab, and was interested in my project as a whole. I would like to thank the members of my committee, past and present, Dr. John Carlson, Dr. Edith Camm and Dr. Ljerka Kunst. In addition, I would like to thank the members of my examining committee, Dr. Tom Beatty, Dr. Tom Grigliatti and Dr. George Haughn.

Of course, I want to thank all the past and present members of the Douglas, Ellis and Theilmann labs! And also, the staff of Botany and Plant Science have been a great help.

I would like to acknowledge my parents and family who always hoped I would do something like this. Well, I DID IT.

None of this would have been possible without the support of my good friends Robin Taylor and Jenn Norton. Nor would it have been possible without the love of my son, Joe. He probably doesn't remember a time when I wasn't a student. Joe made it easy to love him, especially since he never complained when I said on the weekends, "I just have to go to the lab for a minute..."

CHAPTER ONE

GENERAL INTRODUCTION

1.1 PHENYLPROPANOID METABOLISM

Phenylpropanoid metabolism is a major secondary metabolic pathway in plants. It serves to convert phenylalanine to a vast collection of compounds with diverse functions. Some of these compounds are produced during development in a cell type-specific manner. For example, lignin is deposited in the cell walls of specialized cells such as tracheary elements and sclereids. Flavonoids may be synthesized in plant organs such as seeds and flowers where they act as pigments (Douglas et al., 1992). Flavonoids synthesized in roots of legumes may function as plant-microbe signaling molecules by inducing transcription of nodulation genes (*nod*) in nitrogen fixing bacteria (Phillips, 1992).

Phenylpropanoid compounds may be synthesized in response to stress. Phytoalexins are antimicrobial compounds synthesized in response to pathogen attack. The level of these phytoalexins increase around the site of infection at concentrations that are toxic to pathogens. Phenylpropanoid-derived phytoalexins include (furano)coumarins, stilbenes, flavonols and isoflavonoids (Dixon and Paiva, 1995). Salicylic acid, derived from cinnamic acid, is produced during resistance responses and is thought to be a signaling molecule in the induction of systemic acquired resistance (Malamy and Klessig, 1992). High levels of UV irradiation may induce the production of (iso)flavonoids, sinapate esters and furanocoumarins (Dixon and Paiva 1995). Lignin and suberin may be deposited after wounding or pathogen attack to seal off these sites (Douglas et al., 1992).

Phenylpropanoid metabolism can be divided into a general pathway that serves to convert phenylalanine to activated substrates, and specific branch pathways that produce the phenylpropanoid derivatives. Many phenylpropanoid compounds are derived from activated hydroxycinnamoyl-CoA thioesters. The conversion of phenylalanine to these thioesters requires, at minimum, the action of three enzymes, phenylalanine ammonia-lyase (PAL), cinnamate 4-hydroxylase (C4H) and 4-coumarate:CoA ligase (4CL) (see Figure 1.1). In addition, cinnamate 3-hydroxylase (C3H), ferulate 5-hydroxylase (F5H) and caffeate O-methyltransferase (COMT) are required to produce differentially hydroxy- and methoxylated cinnamic acids which may serve as substrates for 4CL. Many of these enzymes have been purified or partially purified, and most of the genes that encode these enzymes have been isolated from various plants.

Lignin and flavonoids are major end products of phenylpropanoid metabolism. The biosynthesis of lignin has been well studied and requires the action of both cinnamoyl-CoA reductase (CCR) and cinnamyl alcohol dehydrogeanse (CAD). Flavonoid biosynthesis which has also been well characterized, requires chalcone synthase (CHS) and chalcone isomerase (CHI) activities. The biosynthesis of phenylpropanoid compounds will be discussed in section 1.3.

1.2 GENERAL PHENYLPROPANOID METABOLISM

As mentioned in section 1.1, general phenylpropanoid metabolism converts phenylalanine to 4-coumaroyl-CoA and other hydroxycinnamoyl-CoA thioesters. The three required enzymes will be discussed in detail.

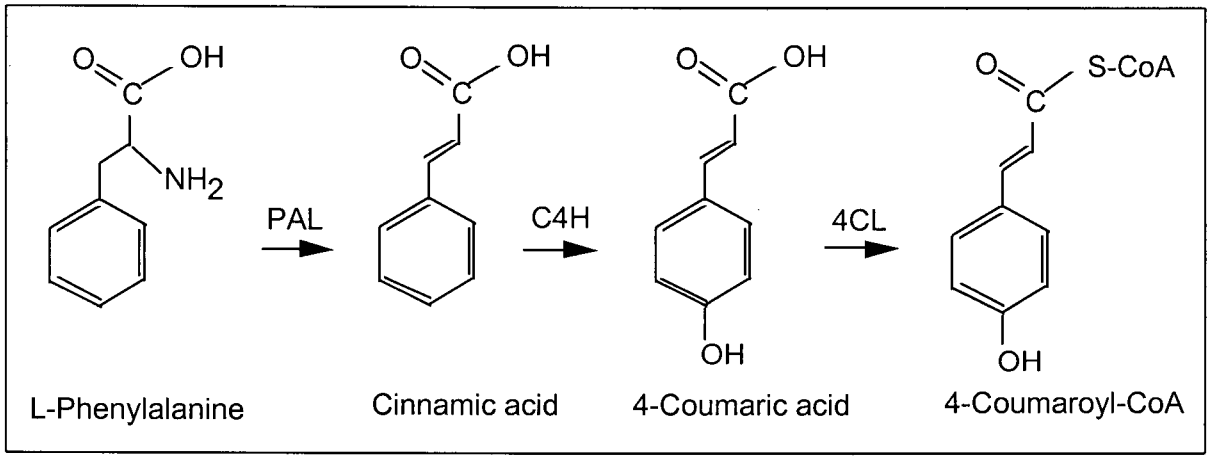


Figure 1.1 The core reactions of general phenylpropanoid metabolism. PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate:CoA ligase.

1.2.1 Phenylalanine Ammonia Lyase

Phenylalanine ammonia lyase (PAL) is the first enzyme of general phenylpropanoid metabolism. PAL activity is required for the deamination of L-phenylalanine to *trans*-cinnamic acid and NH_4^+ (Figure 1.1). The enzyme functions as a tetramer and each monomer is approximately 77 kDa in size. PAL has been widely studied; over 550 papers dealing with PAL have been published in the last ten years. PAL has been characterized from numerous plants, as well as some bacteria and fungi. In monocots and fungi PAL protein preparations may also show activity against L-tyrosine. A maize *PAL* gene expressed in *E. coli* was able to catalyze the deamination of both L-phenylalanine and tyrosine (Roesler et al., 1997). PAL activity can be inhibited by *trans*-cinnamic acid (Bolwell et al., 1986) and chemical analogs such as α-aminooxy-β-phenylpropionic acid (AOPP). PAL activity has also been suppressed *in vivo* using antisense or co-suppression technology (Elkind et al., 1990; Pallas et al., 1996).

Bean PAL activity is induced after cell suspension cultures are treated with an elicitor. However, if *trans*-cinnamic acid is then added to these cells PAL activity and the

rate of PAL synthesis rapidly declines (Bolwell et al., 1986). Down-regulation may be due to phosphorylation which marks the protein for degradation to an inactive form (Bolwell, 1992). Mung bean seedlings grown in the presence of AOPP developed normally but there was a decrease in anthocyanin (Amrhein and Holländer, 1979) and lignin synthesis (Amrhein et al., 1983). Tobacco plants with an ectopic bean *PAL2* transgene had suppressed levels of PAL activity due to co-suppression. Various phenotypes were observed, including altered flower pigmentation and lower levels of pollen germination. There was also reduced lignin content and stunted growth (Elkind et al., 1990). PAL-suppressed tobacco plants also produce four-fold lower levels of salicylic acid which leads to the plants failing to develop systemic acquired resistance (Pallas et al., 1996).

A nucleotide search of GenBank in July 1998 located 131 *PAL* entries. PAL is encoded by a small gene family in all plants examined, such as bean, parsley and poplar (Cramer et al., 1989; Lois et al., 1989; Subramaniam et al., 1993). However, as many as 50 *PAL* genes may exist in potato (Joos and Hahlbrock, 1992). PAL activity and mRNA levels have been shown to increase in bean, poplar and pine cell suspension cultures after treatment with elicitor (Kuhn et al., 1984; Bolwell et al., 1985; Moniz de Sá et al., 1992; Campbell and Ellis, 1992).

Poplar *PAL* will be discussed as an example of the *PAL* gene family. One cDNA and two independently segregating genes, *PAL1* and *PAL2* have been isolated from hybrid poplar (*P. trichocarpa* X *P. deltoides*) (Subramaniam et al., 1993). The *PAL* cDNA has been expressed in an insect cell/baculovirus system and high levels of active enzyme were produced. Mutagenesis of the putative active site serine^{202} to alanine resulted in a complete

loss of activity, confirming the importance of this specific serine residue (McKegney et al., 1996).

The highest levels of poplar *PAL1/2* mRNA were observed in developing stem and leaf tissue, which is consistent with the accumulation of phenylpropanoid compounds in these tissues. Lower mRNA levels were observed in tissues undergoing secondary growth such as the developing xylem of mature stems. *PAL* mRNA was undetectable in mature leaves. These results were confirmed by *in situ* hybridization (Subramaniam et al., 1993). *PAL* mRNA levels are drastically increased after poplar cell suspension cultures are treated with an elicitor (Moniz de Sá et al., 1992). Differential expression is observed in the *PAL* genes isolated from a hybrid aspen (*P. sieboldii* X *P. grandidentata*). *pal g1* was highly expressed in young tissues near the shoot bud and probably corresponds to *PAL 1/2* from poplar. *pal g2a and b* are divergent forms of *PAL*. *pal g2b* is highly expressed in mature stems with *pal g2a* accumulating at a lower level in the same tissue. The expression pattern of *pal g4* is not well characterized (Osakabe et al., 1995ab).

1.2.2 Cinnamate-4-Hydroxylase

Cinnamate-4-hydroxylase (C4H) is the second enzyme of general phenylpropanoid metabolism, and functions to hydroxylate cinnamate at the 4' position (Figure 1.1). It is a member of the cytochrome P450 family of mono-oxygenases, using molecular oxygen and NADPH as an electron donor. An NADPH-cytochrome P450 reductase activity is also required for C4H activity (Benveniste et al., 1977). C4H activity is increased in response to wounding (Tanaka et al., 1974; Benviste et al., 1977), fungal infection (Maule and Ride, 1983) and by UV light (Beneviste et al., 1978). Characterization of C4H has been difficult

due to the enzyme being membrane bound, located on the endoplasmic reticulum (Beneviste et al., 1977). Within recent years the C4H protein has been purified to homogeneity, C4H genes have been cloned and active C4H proteins expressed in heterologous systems.

The C4H proteins from Jerusalem artichoke, bean and soybean are about 58 kDa in size (Kochs et al., 1992; Rodgers et al., 1993; Gabriac et al., 1991). PCR was first used to isolate a cDNA clone encoding C4H from mung bean (Mizutani et al., 1993) and from *Catharanthus roseus* (Hotze et al., 1995). Other *C4H* clones have been isolated using antibody screening of an expression library (Fahrendorf et al., 1993; Teutsch et al., 1993), or previously isolated *C4H* clones (Mizutani et al., 1997). A single gene encodes C4H in Arabidopsis, from which both cDNA and genomic clones have been isolated (Batard et al. 1997; Bell-Lelong et al., 1997). A *C4H* cDNA has been isolated from aspen (Ge and Chiang, 1996) and hybrid poplar (N. Mah and C. Douglas, unpublished results). In poplar, *C4H* is expressed in tissues involved in producing phenylpropanoid compounds, and is also highly expressed in poplar cell suspension cultures treated with an elicitor (N. Mah and C. Douglas, unpublished results).

There is a growing number of DNA sequences that are described as encoding plant P450s, based on sequence similarities in regions related to common catalytic properties, that is, heme and oxygen binding, electron transfer and oxygen activation. However, few of these have been shown to encode C4H (Pierrel et al., 1994). To establish the authenticity of *C4H* cDNA clones, it is essential that they be expressed in heterologous systems. Because C4H activity requires NADPH-P450 reductase activity, this activity must be exogenously supplied to the expressed C4H if the cells of the expression system cannot supply it. The Arabidopsis

C4H expressed in insect cells was shown to bind to *trans*-cinnamic acid, as detected by an appropriate spectral shift, and not to other potential P450 substrates, suggesting that it is, indeed, C4H (Mizutani et al., 1997). Instead of adding NADPH-P450 reductase activity, the *C. roseus C4H* cDNA was fused to a P450 reductase (P_{450} *RED*) and then expressed in the *E. coli* system. Some modifications had to be made to the N-terminal sequence (which is very hydrophobic, due to a membrane integration signal peptide) but C4H activity was detected (Hotze et al., 1995). The putative C4H cDNA from Jerusalem artichoke was expressed in yeast microsomes. This P450 activity appears to be devoted to the hydroxylation of cinnamate although it does have the ability to oxygenate some other small molecules such as 7-methoxycoumarin and 7-ethoxycoumarin (Pierrel et al., 1994). The aspen *C4H* cDNA expressed in yeast also showed authentic C4H activity (Ge and Chiang et al., 1996).

Wounding results in a transient increase in the *C4H* mRNA in Arabidopsis (Mizutani et al., 1997) and also in the protein and the mRNA levels in Jerusalem artichoke (Batard et al., 1997). In Arabidopsis, the expression pattern of *C4H*, *PAL1* and *4CL* were similar in different tissues and they followed the same time course after treatment with light or wounding (Mizutani et al., 1997). Treatment with an elicitor induced a transient increase in *C4H* mRNA levels in *Catharanthus roseus* cell suspension culture (Hotze et al., 1995). Thus *C4H* appears to respond to the same type of stimuli as other phenylpropanoid genes.

1.2.3 4-Coumarate:CoA Ligase

4-coumarate:CoA ligase (4CL) activity is required to convert hydroxycinnamic acids to activated CoA-thioesters. 4CL has been studied at both the protein and the DNA level. 4CL has been purified or partially purified from a wide range of plants including both

angiosperms and gymnosperms. Most reports suggest the protein is approximately 60 kDa in size and functions as a monomer. In general, 4CL activity is specific for the activation of the *trans*-isomer of differently substituted hydroxycinnamic acids and requires ATP, Mg^{2+} and coenzyme A as co-factors (Figure 1.2).

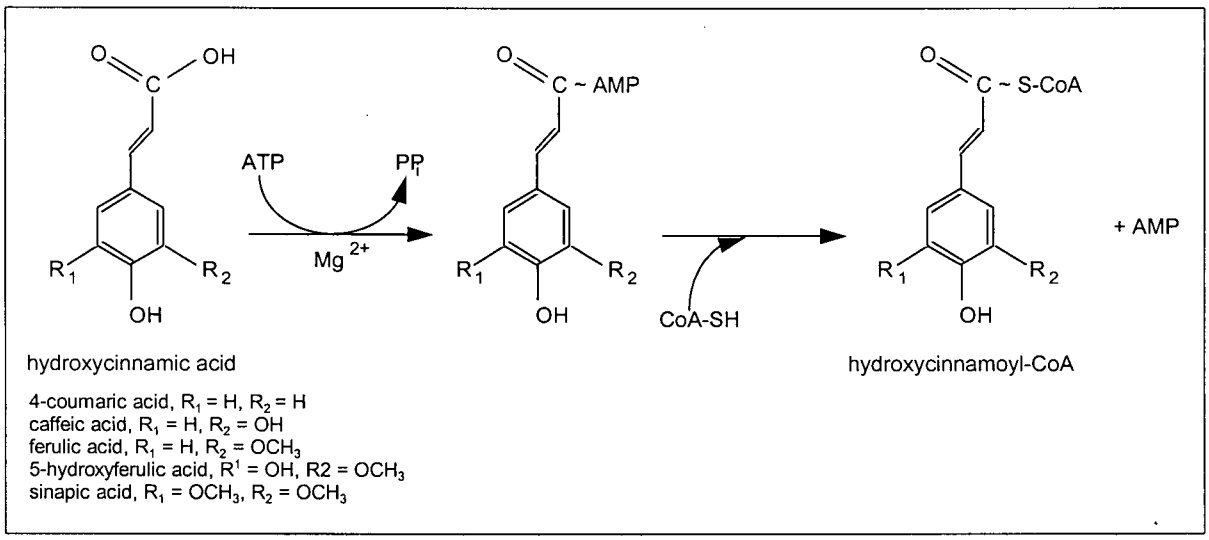


Figure 1.2 The reaction catalyzed by 4-coumarate:CoA ligase.

4CL may exist as a single form in some species, such as *Forsythia* (Gross and Zenk, 1974). However, most plants contain multiple 4CL isoforms. 4CL isoforms from soybean cell suspension cultures show differences in their substrate preferences (Knobloch and Hahlbrock, 1975), but the two very similar isoforms isolated from parsley show no differences in their substrate preferences (Lozoya et al., 1988). Conflicting data exists in the literature as to the substrate usage of *Populus* 4CL. 4CL proteins have been studied in *P. euramericana* (poplar) (Kutsuki et al., 1982; Grand et al., 1983) and *P. tremuloides* (aspen) (Meng and Campbell, 1997; Hu et al., 1998). Grand et al. (1983) reported that three partially purified 4CL isoforms isolated from poplar demonstrated difference substrate preference

patterns. They also demonstrated that one isoform could use sinapic acid as a substrate. However, Kutsuki et al. (1982) tested a variety of species for 4CL activity and found no activity towards sinapic acid in crude xylem extracts of poplar, as did Meng and Campbell (1997) in aspen.

Recombinant tobacco 4CL enzymes have been produced using an *E. coli* system. Both enzymes showed the same pattern of substrate usage, which was similar to the native 4CL activity detected in crude tobacco protein extracts. Like the native proteins, the recombinant proteins did not utilize sinapic acid, but they readily converted cinnamic acid to cinnamoyl-CoA whereas the native activity did not (Lee and Douglas, 1996).

When this project was started, the genes or cDNAs encoding 4CL had only been isolated from parsley (Douglas et al., 1987; Lozoya et al., 1988), rice (Zhao et al., 1990) and potato (Becker-André et al., 1991). Analysis of the parsley *4CL1* promoter was in progress (Douglas et al., 1991; Hauffe et al., 1991). Since then, the genes or cDNAs encoding 4CL have been isolated from soybean (Uhlmann and Ebel, 1993), Arabidopsis (Lee et al., 1995a; Ehlting et al., 1998), loblolly pine (Voo et al., 1995; Zhang and Chiang, 1997), *Lithospermum* (Yazaki et al., 1995) *Vanilla* (Brodelius and Xue, 1997), tobacco (Lee and Douglas, 1996), and recently, aspen (Hu et al., 1998).

4CL appears to be encoded by a small gene family in all the previously mentioned plants, including Arabidopsis (Ehlting et al., 1998), and loblolly pine (Zhang and Chiang, 1997), two species in which it was previously thought to be encoded by single genes (Lee et al., 1995a; Voo et al., 1995). The sequence similarity between most 4CL sequences is quite high, which probably reflects that fact that most *4CL* cDNA or genomic clones were isolated

using parsley or potato cDNA clones as the hybridization probe. A few clones have been isolated using PCR (Zhang and Chiang, 1997; Hu et al., 1998) and 4CL antibodies (Voo et al., 1995).

The known 4CL amino acid sequences have been subjected to phylogenetic analysis and appear to group into two distinct classes (Ehlting et al, 1998). Class I contains the majority of the 4CL sequences, including both parsley and potato sequences, all three tobacco sequences and the single *Vanilla* sequence. Class I also contains one each of the soybean, aspen and *Lithospermum* 4CL sequences, and two of the three Arabidopsis sequences. Class II contain both rice sequences and one each of soybean, aspen, *Lithospermum*, and Arabidopsis. Interestingly, both forms of the recombinant proteins encoded by the parsley and tobacco cDNAs use substrates in essentially a similar manner. However, the recombinant proteins encoded by aspen cDNAs (one in each class) appear to show a divergence in substrate utilization.

Two highly conserved motifs exist in all known 4CL sequences. PYSSGTTGLPKG is part of the putative AMP binding site. This sequence is conserved in other, unrelated enzymes such as luciferase, peptide synthetases, and ENT E, that also bind and hydrolyze ATP to activate aromatic carboxylic acids (luciferin, phenylalanine and 2.3-dihydroxybenzoic acid, respectively) (Becker-André et al., 1991).

The second conserved motif is GEIRCG. The cysteine (Cys^{390}) embedded within this motif is found in the previously mentioned enzyme sequences and is one of six cysteines conserved in 4CL sequences. This cysteine is thought to be important in maintaining catalytic activity. To further characterize this domain, Becker-André et al. (1991) used site-directed

mutagenesis of potato 4CL-1 to change Cys^{390} to Ser^{390}, and also, in another experiment Arg^{392} to Gly^{392} but both mutations rendered the protein unstable in *E. coli* cells.

Brodelius and Xue (1997) suggest that the high identity in the C-terminal half of the 4CL proteins may indicate that it is important for enzyme structure and/or function. Aspen Pt4CL2, rice 4CL and *Lithospermum* 4CL2 have approximately 20 extra amino acids at the N-terminus. Only recombinant Pt4CL2 proteins have been examined, but Hu et al. (1998) suggest this sequence may be involved in substrate binding specificity. This is based on the observation that Pt4CL1, which does not contain the extra amino acids at the N-terminus has a different substrate utilization pattern than Pt4CL2.

4CL, like most other phenylpropanoid metabolism genes, is expressed in a cell type-specific manner during development and in a non-cell type-specific manner after induction by environmental stresses. *4CL* is expressed by 3 days post-germination in developing Arabidopsis seedlings. The expression in the cotyledons is associated with lignin deposition, and in the roots with flavonoid synthesis. The highest levels of *4CL* expression in the mature plants are seen in the bolting stem, which is highly lignified. Transgenic Arabidopsis with a parsley *4CL1-GUS* transgene show a similar GUS expression pattern, indicating that the parsley *4CL1* promoter contains all the necessary information to direct correct developmental *4CL* expression (Lee et al., 1995a). The flower-specific expression of endogenous tobacco *4CL* and an introduced parsley *4CL1* was examined in tobacco. In situ hybridization demonstrated that both genes were expressed in a similar cell type-specific manner, consistent with the accumulation of phenylpropanoid compounds such as flavonoids and lignin (Reinold et al., 1993).

In the plants tested, *4CL* mRNA levels increase transiently, as does 4CL activity, in response to UV light, fungal attacks and wounding. After treatment with a plant-specific elicitor, cell suspension cultures of parsley (Kuhn et al., 1984) and parsley protoplasts (Dangl et al., 1987), potato cell suspension cultures and leaves (Becker-André et al., 1991), poplar (Moniz de Sá et al., 1992) and soybean cell suspension cultures (Uhlmann and Ebel, 1993) all show a transient increase in *4CL* mRNA levels by 1 to 5 hours post elicitation, followed by an increase in 4CL activity several hours later. In parsley there is a concomitant increase in furanocoumarins (Tietjen et al., 1983) and in poplar, accumulation of phenolic material esterified to the cell wall (Moniz de Sá et al., 1992). Both *4CL* genes in the parsley and the potato systems respond to elicitor treatment, while only one of the soybean genes is activated (Douglas et al., 1987; Becker-André et al., 1991; Uhlmann and Ebel, 1993). Because a cell wall was not necessary for elicitor activation of parsley *4CL*, Dangl et al. (1987) suggested that the elicitor may bind to a plasma membrane receptor which then initiates a signaling pathway that activates *4CL* and other phenylpropanoid gene expression. In support of this hypothesis, Nürnberger et al. (1994) identified a 13 amino acid oligopeptide from a fungal glycoprotein elicitor that also acts as a parsley elicitor. Furthermore, they identified a high affinity binding site for the oligopeptide that is localized to the plasma membrane.

4CL expression is also activated by wounding. In wounded parsley roots *4CL1* but not *4CL2* was activated transiently 2 hours post wounding (Lois and Hahlbrock, 1992). A similar response was observed after wounding potato leaves (Becker-André et al., 1991). Transgenic Arabidopsis and tobacco plants homozygous for a parsley *4CL1* promoter-GUS transgene, were mechanically wounded. Two hours after wounding a transient increase in

both *4CL* and *GUS* mRNA levels was observed which disappeared by 24 hours after wounding (Lee et al., 1995a; Ellard-Ivey and Douglas, 1996). Jasmonates may play a role in mediating activation of *4CL* and other phenylpropanoid genes in response to stress. It has been suggested that after perception of a wound-generated signal at the cell surface linoleic acid is released from membranes and is then converted to jasmonic acid. In parsley cell suspension cultures and transgenic tobacco (parsley *4CL1-GUS*), *4CL* and *GUS* mRNA levels increased after exogenous jasmonic acid, methyl jasmonate or their precursor, linoleic acid, were added. However, methyl jasmonate treatment only increased the level of furanocoumarins to one third of the level after elicitor treatment, and *ELI7* (a parsley, elicitor inducible gene) was not activated by jasmonates. Thus while jasmonates may mediate the activation of *4CL* by wounding, other signaling pathways, in addition to the jasmonate pathway, may be responsible for mediating some of the elicitor responses (Ellard-Ivey and Douglas, 1996).

4CL genes are activated in parsley suspension-cultured cells and protoplasts in response to treatment with UV light. The cells respond to UV light by accumulating flavonoids (which strongly absorb UV light) in their vacuoles, as a means of protection (Chappell and Hahlbrock, 1984). *4CL* genes were also transiently activated in parsley plants 6 hours after treatment with ozone. This was correlated to an increase in the production of flavone glycosides and furanocoumarins (Eckey-Kaltenbach et al., 1994).

The developmental expression of poplar *4CL* has been determined using in situ hybridization. *4CL* is highly expressed in young leaves of the apical bud, in epidermal cells

(as was *PAL*), and at lower levels in differentiating xylem. In young stems, *4CL* was mainly expressed in developing xylem (Douglas et al., 1992).

1.2.4 Coordinate Regulation of Phenylpropanoid Metabolism Enzymes

It is a well-established fact that phenylpropanoid compounds are produced in response to externally applied stresses. This requires the coordinate expression of genes encoding the enzymes required to produce phenylpropanoid compounds. A large body of work has been devoted to this field. For example, parsley cells exposed to UV irradiation, elicitor and wounding respond with coordinated increases in mRNA levels for *PAL*, *C4H* and *4CL* (Logemann et al., 1995). This has been demonstrated to be due to an increase in transcription rates (Chappel and Hahlbrock, 1984).

Other examples of coordinated regulation are found in alfalfa, bean, Arabidopsis, poplar and old-man-cactus. Alfalfa cell suspension cultures accumulate high levels of the isoflavonoid, medicarpin, after elicitation. *PAL*, *CHS and CHI* mRNA levels increased transiently after elicitation (Dalkin et al., 1990). Elicitation of bean cell suspension cultures resulted in rapid increases in *PAL*, *CHS* and *CHI* mRNA levels (Cramer et al., 1985). A similar result was observed in Arabidopsis suspension-cultured cells treated with PGA lyase (an elicitor). Both *PAL* and *4CL* mRNA reached a transient maximum at 8 - 10 hrs followed by an increase of *COMT* at 24 hours (Davis and Ausubel, 1989). Following fungal elicitation of poplar cell suspension cultures, cell wall-bound phenolic compounds accumulate. This is preceded by increases in *PAL* and *4CL* mRNA and enzyme activity levels. However, CHS, CAD, and coniferin β-glucosidase activities remained unchanged (Moniz de Sá et al., 1992). *PAL*, *4CL* mRNA levels were rapidly and transiently increased after potato leaves were

treated with a fungal pathogen. This could be mimicked by treating suspension-cultured cells with either elicitor or arachidonic acid (Fritzemeier et al., 1987). Old-man-cactus, treated with an elicitor accumulates the aurone phytoalexin, cephalocerone. Increased levels of PAL, CHI, and CHS activities accompany the synthesis of cephalocerone (Paré et al., 1992).

Other stimuli also lead to coordinate expression of phenylpropanoid genes. Parsley cell suspension cultures treated with ozone produce both furanocoumarins and flavone gylcosides. This is preceded by an increase in the mRNA levels of *PAL*, *4CL*, *CHS*, pathogenesis related proteins *PR1* and *PR2* and *Eli 6* of unknown function (Eckey-Kaltenbach et al., 1994). *PAL* has been shown to increase in ozone-treated Arabidopsis cell suspension cultures (Davis and Ausubel, 1989). Both *PAL* and *CHS* mRNAs accumulate in the leaves of Arabidopsis plants that have been exposed to low temperatures (Leyva et al., 1995). In tobacco plants, methyl jasmonate treatment led to a transient increase in *PAL* (Sharan et al., 1998) and *4CL* (Ellard-Ivey and Douglas, 1996) mRNA levels.

To elucidate the mechanism by which the genes are coordinately regulated, phenylpropanoid gene promoters have been isolated and analyzed. 290 bp of the *PAL1* promoter from Arabidopsis is sufficient to establish tissue-specific expression. An extra 250 bp was required for response to environmental stimuli (Ohl et al., 1990). 210 bp of the parsley *4CL-1* promoter is sufficient to direct tissue-specific and cell-specific expression of a GUS reporter gene in transgenic tobacco and Arabidopsis (Hauffe et al., 1991; Lee et al., 1995a). This expression is dependent on combinatorial interactions between distinct promoter domains (Hauffe et al., 1993). An additional 280 bp of promoter and exonic sequences are required for expression of the parsley *4CL-1* gene after fungal elicitor or light treatment of

transformed suspension-cultured parsley cells or transgenic tobacco. It is unknown whether the exonic sequences act as transcriptional enhancers or mRNA stability elements (Douglas et al., 1991).

Lois et al. (1989) identified three inducible sites of DNA/protein interactions (boxes P, A and L) in the parsley *PAL-1* promoter using in vivo footprinting. These putative cis-elements alone are conserved in many phenylpropanoid gene promoters, including *4CL* and *C4H* (Lois et al., 1989; Logemann et al., 1995) but are not alone sufficient for activation of GUS reporter gene expression by light or elicitor treatment in parsley cells. This suggests that other elements are involved in this aspect of regulation (Logemann et al., 1995). da Costa e Silva et al. (1993) demonstrated that the parsley *PAL-1* gene is controlled by transcriptional activation. They found a Box P binding factor (BPF-1) activity in nuclear extracts, and a BPF-1 cDNA was isolated from an expression library. Sequence analysis showed no similarity to other known DNA-binding proteins, but domains present in other transcription factors were identified. *BPF-1* mRNA accumulates in parsley cell suspension cultures that have been treated with an elicitor and in leaves that have been exposed to a pathogen. There was a strong correlation between the amount of *BPF-1* and *PAL* mRNA (da Costa e Silva et al., 1993). However, the importance of *BPF*-1 for the regulation of *PAL* expression in vivo is unclear, and no further work has been presented.

Boxes P and L are the same as AC-rich elements that are conserved in the promoters of many phenylpropanoid genes (Lois et al., 1989; Sablowski et al., 1994). It is likely that they are myb binding sites. Grotewald et al. (1994) showed that the maize *P* gene (which specifies red pigmentation in certain organs) is a myb homolog and that it binds to the

promoter of the *A1* gene (required for both anthocyanin and phlobaphene pigment biosynthesis). *A1* binds to a series of preferred binding sites which include AC elements. Thus, myb-related transcription factors (myb factors) may play an important role in coordinate regulation of phenylpropanoid metabolism genes.

While a family of genes exist in plants that encode myb factors, the functions are known for only a few of the family members. Several of these genes encode transcription factors that control flavonoid biosynthesis. Ectopic expression of *Myb305*, an *Antirrhinum* petal-specific gene, increased the level of expression of bean *PAL2* in tobacco leaf protoplasts (Sablowski et al., 1994). Transgenic tobacco plants overexpressing *Antirrhinum* AmMYB308 and AmMYB330 myb factors had several phenotypes including stunted growth and reduced lignin content. The myb factors appear to inhibit expression of genes encoding enzymes for all steps between C4H and CAD, and many of the genes have promoters which contain motifs that are recognized by MYB proteins. The authors suggest these myb factors may encode weak activators/repressors that, when overexpressed, compete with stronger myb proteins to control target gene expression (Tamagnone et al., 1998).

Finally, another candidate transcription factor has been found in tobacco. A cDNA clone (ACBF) encoding a DNA-binding factor was isolated from a tobacco expression library using complementary oligonucleotides bearing a bean *PAL1* AC-rich motif (Séguin et al., 1997). ACBF is expressed in all tobacco tissues examined. A cDNA (56BF) with very high sequence identity to ACBF also binds to a different sequence (not homologous to the AC elements) in the parsley *4CL1* promoter (S. Lee and C. Douglas, unpublished results). Both cDNA clones share sequence homology with RNA-binding proteins. Tobacco plants with

suppressed levels of 56BF do not show a novel phenotype and it is unclear what role, if any, 56BF plays in vivo (J. Jones and C. Douglas, unpublished results).

1.2.5 Phenylpropanoid Metabolism Enzymes as a Multi-Enzyme Complex

Stafford (1974) was one of the first to suggest that phenylpropanoid metabolism enzymes might exist as a multi-enzyme complex (MEC). Few experiments have been done to test this hypothesis, but those that have been done will be presented. Stafford (1974) isolated several different molecular weight forms of *Sorghum bicolor* F5H and suggested that they are associated with different MECs. Furthermore, she described a large molecular weight complex that contained F5H, PAL/TAL and C4H activities.

Hrazdina and Wagner (1985) studied the enzyme induction kinetics of PAL, C4H, 4CL and CHS in buckwheat seedlings and *Hippeastrum* (Dutch red hybrid) petals. They demonstrated that these enzyme activities appear coordinately and are correlated with flavonoid production. A proportion of the enzyme activity of C4H, PAL and C4H elute together in the void volume from a size exclusion column, suggesting that they are associated together in a MEC. This includes all of the C4H activity, approximately 35% of the PAL activity and 5-15% of the CHS activity. Both PAL and C4H appear to be associated with the ER membrane. A feeding experiment using [^{3}H] phenylalanine and [^{14}C] cinnamic acid showed that phenylalanine rather than cinnamic acid, is preferentially incorporated into 4-coumaric acid, providing biochemical evidence for a MEC.

Deshpande et al. (1993) also suggested that PAL, C4H and 4CL exist in potato as a MEC. They demonstrated that activity of each enzyme elutes in two peaks from a size exclusion column. The first peak is in the same position for each enzyme (in the void

volume) and the second peak for each is distinct (at the predicted mw size). Re-chromatography of the first peak showed the same pattern of two peaks. This suggests that the first peak includes all the enzymes in a MEC and the second peak may be the specific enzyme disassociated from the MEC or the free native form. They also showed that after supplying phenylalanine, and the required co-factors, to the putative MEC, cinnamic acid, 4-coumaric acid, and 4-coumaroyl-CoA were produced.

Another line of evidence to suggest that phenylpropanoid metabolism enzymes exist as a MEC come from work by Sewalt et al. (1997). Transgenic tobacco plants with decreases in either PAL or C4H activity both had an overall decrease in lignin content. However, depending on which activity was decreased the transgenic plants contained lignin with differences in their monomeric composition. The authors suggested this may be due to different isoforms of the enzymes being involved in different MECs, perhaps also involving enzymes such as COMT or 4CL (Sewalt et al., 1997).

Finally, Shirley et al. (1998) have provided evidence that the enzymes specific to flavonoid metabolism in Arabidopsis may exist as a MEC. Using affinity chromatography, immunoprecipitation and two-hybrid analysis they have shown that CHS and CHI interact with each other. They have also shown that CHS and CHI are both localized around vacuoles and at the cell wall. Taken together, these experiments might suggest that phenylpropanoid metabolism enzymes exist as a MEC. However, at this time there is limited little direct evidence of phenylpropanoid metabolism enzymes existing as a MEC.

1.3 BIOSYNTHESIS OF PHENYLPROPANOID PRODUCTS

1.3.1 Lignin

In the past few years, the biology of lignin has been extensively reviewed (Whetten and Sederoff, 1991; Sederoff et al., 1994; Boudet et al., 1995; Douglas, 1996; Boudet, 1998 Whetten et al., 1998). These reviews discuss the current status of research concerning the biosynthesis, analysis and feasibility of genetically engineering a change in the composition of lignin. The analysis of lignin usually requires chemical degradation to release phenolic residues with substitution patterns similar to the monolignols. Complete information of lignin composition cannot be obtained by any one method. Recently, solution and solid state ^{13}C-NMR has been employed to analyze the composition of lignin. ^{13}C-NMR is not destructive, does not modify the lignin monomer and only small amounts of material are required (Whetten et al., 1998).

Lignin represents approximately 25% of the dry weight of wood, which makes it the second most abundant biomolecule, after cellulose. Lignin imparts both strength and water impermeability to specialized cells such as tracheary elements and sclereids. It is also deposited in response to stress where it helps to seal off the sites of pathogen attack and wounding (Douglas, 1996). Lignin is a complex biopolymer traditionally believed to be composed of a heterogeneous mixture of hydroxycinnamoyl alcohols (Figure 1.3). These alcohols or monolignols are linked together in three dimensions by more than twenty different possible intersubunit linkages. The monolignols are thought to be derived from hydroxycinnamoyl-CoA thioesters that have undergone two reductive steps (Figure 1.3), first to aldehydes by the action of cinnamoyl-CoA reductase (CCR) and then to the alcohols by

the action of cinnamyl alcohol dehydrogenase (CAD). After the monolignols are made, they may be stored until required, then transported to the cell wall and polymerized.

Different plant species have lignins of different compositions. For example, angiosperm lignin has both guaiacyl lignin and syringyl lignin derived from coniferyl and sinapyl monolignols, respectively. Gymnosperm lignin is mostly guaiacyl lignin with small amounts of *p*-hydroxyphenyl lignin, derived from 4-coumaryl alcohol. The reason for the absence of syringyl lignin in gymnosperms may be due to the lack of hydroxylase and *O*-methyltrasnferase enzymatic activity towards ferulic and 5-hydroxyferulic acids, respectively (Whetten and Sederoff, 1995). To add to the level of complexity, different cell types within a single plant may have lignin with differing compositions.

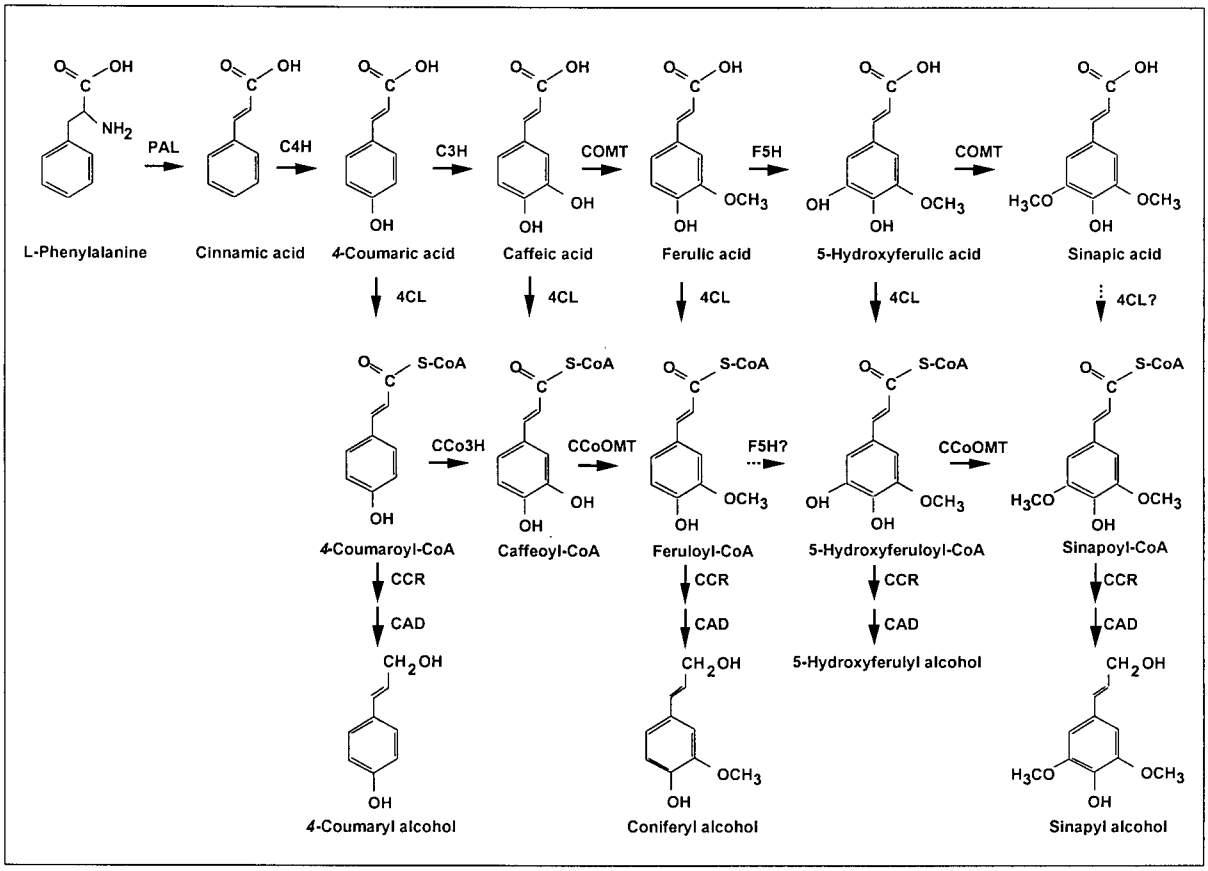


Figure 1.3 Proposed lignin biosynthetic pathway. Enzymes that are thought to catalyze each step (shown by an arrows) are indicated, ? represents enzymatic reactions under debate, at least in some plants. Abbreviations: PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; C3H, coumarate 3-hydroxylase; COMT, bi-functional caffeic acid *O*-methyltransferase; F5H, ferulate 5-hydroxylase; 4CL, 4-coumarate:CoA ligase; CCo3H, coumaroyl-CoA-3-hydroxylase; CCR, cinnamoyl:CoA reductase; CAD, cinnamyl alcohol dehydrogenase.

Because the field of lignin biology is so large, I will restrict my discussion to current research that is making use of mutants and transgenic plants with an altered lignin content or composition to dissect the lignin biosynthetic pathway. The traditional view of lignin biosynthesis is being reassessed. Earlier literature has suggested that the lignin biosynthetic pathway is a well-ordered linear pathway. However, it is becoming apparent that lignin biosynthesis is highly complex and potentially "plastic". That is, it may be able to respond to

changes in the availability or activity of biosynthetic enzymes by incorporating non-traditional phenolic monomers into the lignin polymer (Whetten et al., 1998).

Many of the genes encoding the lignin biosynthetic enzymes have been cloned. cDNAs and/or genomic clones have been isolated for PAL, C4H, COMT, F5H, 4CL, CAD and CCR from numerous plants including Arabidopsis, bean, parsley, poplar and pine. Sense and antisense suppression of lignin biosynthetic genes is producing plants with changed lignin composition and content. In some cases, conflicting results have been reported, when the same gene has been manipulated in different plants. In addition, at least three different mutants with changes in lignin composition are known.

Conflicting results were reported in two studies that involved engineering decreased levels of 4CL activity in Arabidopsis (Lee et al., 1997) and tobacco (Kajita et al., 1996). Arabidopsis lines with only residual 4CL activity showed a decrease in extractable lignin. There was a significant decrease in the amount of guaiacyl lignin, but there was no change in syringyl lignin amount (Lee et al., 1997). These results suggest that 4CL does not play a role in activating sinapic acid. Tobacco plants with decreased 4CL activity had lower levels of lignin. However, in some of the transgenic tobacco plants the monomeric ratio of syringyl to guaiacyl units was decreased and in others it was increased (Kajita et al., 1996). Clearly there is either a difference between the response of Arabidopsis and tobacco plants to decreased levels of 4CL activity or this difference may be due to different gene family members being suppressed. The results from decreasing 4CL activity are encouraging. Suppressing 4CL activity could be used to produce plants with a changed lignin composition.

The Arabidopsis *fah1* phenotype results from a mutation in the ferulate 5-hydroxylase (*F5H*) gene (Figure 1.3). The lignin of the *fah1* mutant contains only guaiacyl residues and no sinapic acid derivatives are detected (Chapple et al., 1992). Overexpression of the *F5H* gene in Arabidopsis results in an increase in syringyl lignin with a relative decrease in guaiacyl lignin (Whetten et al., 1998). This is probably due to all the ferulic acid being hydroxylated and being shunted away from the coniferyl alcohol pathway. Thus, F5H activity is required for synthesis of sinapyl monolignols. The results of Lee et al. (1997) and Chapple et. al (1992) strongly suggest that 5-hydroxyferulic acid and/or sinapic acid are the precursors of sinapyl alcohol, but it is still unclear whether 4CL activity is required in its synthesis.

The effect of decreased COMT (Figure 1.3) levels has been tested in transgenic aspen, poplar and tobacco using antisense suppression or co-suppression (Dwivedi et al., 1994; van Doorsselaere et al., 1995a; Atanassova et al., 1995). As with the 4CL-suppressed plants, there were conflicting results. In COMT-suppressed tobacco there was a marked decrease in syringyl subunits while the amount of guaiacyl lignin remained unchanged (Astanassova et al., 1995). COMT-suppressed poplar plants also had a decrease in syringyl lignin, accompanied by an increase in guaiacyl lignin (van Doorsselaere et al., 1995a). In both tobacco and poplar COMT-suppressed plants a new type of monomer, 5-hydroxy guaiacyl, was detected. The brown-midrib 3 mutant (bm3) of maize, which is deficient in COMT activity, also has lignin with a lowered syringyl content and accumulates the 5-hydroxy guaiacyl monolignol (Vignols et al., 1995). A different pattern was observed when aspen COMT was used to decrease the COMT levels in transgenic tobacco. While there was an overall decrease in lignin content some plants had a normal syringyl/guaiacyl ratio, others

had suppressed syringyl levels with either no change or enhanced levels of guaiacyl units (Dwivedi et al., 1994). This may be a problem associated with using a heterologous gene for suppressing the COMT levels.

Cinnamoyl-CoA reductase (CCR) (Figure 1.3) may be an important control point in the regulation of monolignol production because it is the first committed step of lignin biosynthesis. Piquemal et al. (1998) produced tobacco plants transformed with eucalyptus CCR antisense constructs. The overall lignin content was drastically reduced. The lignin composition was reduced in the amount of guaiacyl subunits, while syringyl subunits were unchanged. While there appears to be only one CCR gene in tobacco, it is still possible that additional divergent CCR genes are present and that this complement of genes could preferentially use feruloyl or sinapoyl-CoA as a substrate. Additionally, ferulic and sinapic acid were detected attached to the cell wall of the transgenic plants. The lignin of the most severely CCR suppressed plants was orange brown in colour which is suggestive of an increase in ferulic acid deposition in the cell wall. Because these transgenic plants have a changed lignin composition, CCR might be a good target for the genetic modification of lignin.

Cinnamyl alcohol dehydrogenase (CAD) (Figure 1.3) suppressed plants have also been generated. The lignin of tobacco and poplar plants with suppressed levels of CAD had a red colouration, suggestive of hydroxycinnameldehydes being incorporated into the lignin. A loblolly pine heterozygous for a mutation in CAD (*cad-n1*) accumulates hydroxy-cinnamaldehydes and also contains unusual subunits, including, 2-methoxybenzaldehydes and dihydroconiferyl alcohol (MacKay et al., 1997). Again, different plants showed

differences in their response to enzyme suppression. CAD-suppressed tobacco plants had a decreased lignin content as well as a decrease in syringyl content, but displayed only a slightly lowered guaiacyl content (Yahiaoui et al., 1997). This is in contrast to the finding that the lignin of CAD-suppressed poplar plants had increased aldehyde units, but showed no modification of guaiacyl/syringyl composition or lignin content (Baucher et al., 1996).

These results clearly demonstrate that while we have a good basic understanding of the biosynthesis of lignin, there is still much to be learned. It is possible that the biosynthesis of such an important polymer has been designed to withstand alterations to the basic biosynthetic steps. What is observed in these transgenic and mutant plants may merely reveal previously unforeseen complexities in the biosynthesis of lignin.

1.3.2 Suberin

Suberin is a complex polymeric substance containing both aliphatic and aromatic components. Suberin forms both protective and wound-healing layers (Bernards et al., 1995). It is found within the cell walls of vascular plant tissues such as the periderm, the epidermis and hypodermis of roots, bundle sheaths of grasses, idioblast sheaths between secretory glands, pigment strands of grains and connections between vascular tissue and seed coats (Lewis, 1993). Suberin is essential for water retention and may function in the control of water movement. Like the evolution of lignin biosynthesis, the development of suberin biosynthesis probably played an important role in the transition of plants from an aquatic to terrestrial environment. Most information about suberization has been acquired from the wound healing process. However, relatively little information about the induction, regulation and actual biosynthesis of suberin is available (Bernards et al., 1995).

Historically, the model for suberin structure suggests that the aromatic monomers and aliphatic components are covalently linked and that the aromatic monomers may be derived from monolignols (e.g. caffeoyl alcohol) and also ferulic acid. The aliphatic component consists mainly of long-chain (C_{18} to C_{28}) fatty acids or alcohols esterified to each other and to the aromatic domain. Unfortunately, no method is available to isolate suberin in its native form (Lewis, 1993). Different methods of lignin detection have been used in an attempt to determine the chemical composition of suberin but the results have been conflicting. Thioacidolysis analysis revealed only minor amounts of lignin-type material, about one-sixth of the amount detected by alkaline nitrobenzene oxidation (Bernards et al., 1995).

Recently, Bernards et al. (1995) used solution-state NMR spectroscopy methods to study suberized material. Both ^{1}H- and ^{13}C-NMR failed to detect any lignin-like material in this material. Solid state ^{13}C-NMR demonstrated only a limited amount of lignin while most of the labeled phenylalanine was detected in hydroxycinnamic acid derivatives. Moreover, the phenolic component was highly cross-linked, possibly due to stable C-C or C-O cross-links between the phenolic monomers, as well as possible amide linkages. This result strongly suggests that hydroxycinnamic acids rather than monolignols are the precursors of the phenolic constituents of suberin. If this is the case, 4CL may not play a major role in suberization.

In contrast to Bernards et al. (1995), Lofty et al. (1996) have evidence to suggest that hydroxycinnamoyl-CoA thioesters and thus, potentially, hydroxycinnamyl alcohols may be precursors of the phenolic constituents of suberin. Lofty et al. (1996) identified an acyltransferase, hydroxycinnamoyl-CoA: ω-hydroxypalmitic acid *O*-hydroxy-cinnamoyl-

transferase (HHT), in tobacco cell suspension cultures. Induction and tissue distribution of an HHT from potato tuber discs was previously shown to be coincident with the deposition of suberin. Feruloyl-CoA was found to be the best substrate followed by sinapoyl-CoA and 4-coumaroyl-CoA. No other hydroxycinnamoyl-CoA compounds were accepted as donor substrates. 16-hydroxyhexadecanoic acid and 15-hydroxypentadecanoic acid appeared to be the best acceptor substrates for HHT (Lofty et al., 1996). If HHT activity is a required biosynthetic step for the production of suberin, then activated hydroxycinnamoyl-CoA thioesters, produced by the action of 4CL, are required.

The work of Bernards et al. (1995) and Lofty et al. (1996) clearly demonstrates that the biosynthesis of suberin is still under debate. The emergence of improved detection and analytical systems will be key to clarifying this process.

1.3.3 C_6-C_1 Compounds

Salicylic acid is an endogenous regulator of plant disease resistance (Lee et al., 1995b). It is presently unclear whether 4CL plays a role in the biosynthesis of salicylic acid, which is derived from cinnamic acid by side chain shortening of cinnamic acid to benzoic acid followed by 2-hydroxylation on the ring (Yalpani et al. 1993). A soluble cytochrome P450 enzyme specifically hydroxylates the *ortho*-position of benzoic acid to produce salicylic acid (León et al., 1995).

The mechanism by which cinnamic acid is converted to benzoic acid has not yet been resolved. One suggestion is that a non-oxidative chain-shortening reaction occurs. Schnitzler et al. (1992) demonstrated that 4-hydroxybenzoate is produced from 4-coumaric acid. The reaction requires NAD or NADP as cofactors, with neither ATP nor CoA being required.

Further supporting the suggestion that 4CL activity is not required in the biosynthesis of benzoic acid are the results obtained by Funk and Brodelius (1990ab), who examined the biosynthesis of vanillic acid (3-methoxy-4-hydroxy benzoic acid) in *Vanilla planifolia.* Inhibition of 4CL activity by 3,4-(methylenedioxy)-cinnamic acid resulted in cinnamic acid being incorporated into vanillic acid (Funk and Brodelius, 1990a). Furthermore, Funk and Brodelius (1990b) demonstrated that methoxycinnamic acids (which are not substrates for 4CL) were converted to the corresponding 4-hydroxybenzoic acids through a glycosylated intermediate.

Evidence also exists that the side-chain shortening occurs via a β-oxidation reaction similar to fatty acid degradation. Löscher and Heide (1994) demonstrated that, in cell suspension cultures of *Lithospermum*, 4-coumaroyl-CoA produced via the action of 4CL is an intermediate in 4-hydroxybenzoate biosynthesis. The 4-coumaroyl-CoA is then converted to 4-hydroxybenzoate and acetate. The preceding results suggest that possibly two different, species-specific, mechanisms may be responsible for the side-chain shortening of cinnamic acid to benzoic acid.

1.3.4 Flavonoids

At least 2000 structurally different flavonoids have been described (Harborne, 1980). Flavonoids are based on a C_{15} skeleton. The biosynthesis of most flavonoids begins by the condensation of 4-coumaroyl-CoA with three acetate units derived from malonyl-CoA. B-ring deoxy flavonoids are derived from cinnamoyl-CoA; e.g. pinocembrin in *Populus deltoides* (Shain and Miller, 1982) and old-man-cactus (Liu et al., 1995). The condensation reaction is catalyzed by the action of chalcone synthase (CHS) which initially produces a

chalcone (Hahlbrock, 1981). Chalcone isomerase (CHI) activity converts chalcones to flavones by forming the six-member heterocyclic "C" ring (Hahlbrock, 1981). The flavone can be further modified by hydroxylation, methylation and glycosylation. Flavonoids may function as pigments, phytoalexins and UV protectants (Hahlbrock and Scheel, 1989). Flavonoids may also be synthesized in roots where they function as plant-microbe signaling molecules in N_2-fixing associations between rhizobial bacteria and legumes (Phillips, 1992).

Isoflavonoids are a sub-group of flavonoids that serve as phytoalexins in legumes (Dixon et al., 1992). Isoflavonoids are produced by the action of isoflavone synthase (IFS) which rearranges the flavonoid carbon skeleton (Dixon and Paiva, 1995). IFS is a P450 enzyme and is very unstable *in vitro*. IFS activity has been detected in cell suspension cultures of *Pueraria lobata* (Hakamatsuka et a., 1990) but no further information on this enzyme is available.

1.3.5 Stilbenes

Stilbenes, like flavonoids, are compounds derived from the condensation of (hydroxy)cinnamoyl-CoA thioesters and three acetate units from malonyl CoA. Unlike flavonoid biosynthesis, stilbene synthesis requires an additional decarboxylation (Schröder and Schröder, 1992). Lunularic acid, a dihydrostilbene, is the most widespread natural stilbene. It is produced in all liverworts and one angiosperm, *Hydrangea macrophylla*; it functions as a growth inhibitor. Other dihydrostilbenes also act as plant growth inhibitors (Harborne, 1980). Stilbenes with fungicidal potential exist in several unrelated plant species, such as resveratrol from both peanut and grapevine, and pinosylvin from pine (Hain et al., 1993).

Stilbene biosynthesis only specifically requires the presence of one enzyme: stilbene synthase (STS). STS is responsible for the stepwise condensation of the precursor molecules and cyclization to form the stilbenes. The amino acid sequences of scots pine STS and CHS show high similarity and antiserum against parsley CHS reacts with both proteins (Fliegmann et al., 1992). The amino acid sequence surrounding the putative active site cysteine is well conserved in both STS and CHS genes with exception of the -2 and -3 positions. Site-directed mutagenesis of these positions in CHS and STS from scots pine and STS from peanut did not convert the activity of CHS to the activity of STS, or *vice versa*. However, site-directed mutation of the resveratrol-forming peanut STS was able to either create a total loss of activity or a change in substrate preference from 4-coumaroyl-CoA to phenylpropionyl-CoA (Shröder and Schröder, 1992). This result is especially relevant to the question of the possibility of substrate specificities of 4CL isoforms.

Due to the absence of STS activity in some plants, not all plants synthesize stilbenes. However, the precursor molecules malonyl-CoA and p-coumaroyl-CoA are both present in plants. A stilbene synthase gene isolated from grapevine and transferred into tobacco rendered the transgenic tobacco plants more resistant to infection by *Botrytis cinerea* (Hain et al., 1993). This was the first report of increased disease resistance in a transgenic plant due to the addition of a gene required for phytoalexin biosynthesis.

1.3.6 Coumarins and Furanocoumarins

Coumarins are derived from *o*-hydroxycinnamic acid by cyclization and ring closure. Coumarins often occur in plant tissues as glycosides of o-hydroxycinnamic acid derivatives. If the tissue is damaged the glycosides loses the sugar moiety, the side-chain isomerases and

the molecule undergoes ring closure (Harborne, 1980). Coumarin itself is volatile and upon release has "the odour of new-mown hay". 4-coumaric acid, caffeic acid and ferulic acid are the starting acids for the three most common hydroxycoumarins, umbelliferone, aesculetin and scopoletin, respectively. Hydroxycoumarins derived from sinapic acid are very rare, although isofraxidin is found in the bark of *Fraxinus* (Harborne, 1980).

Furanocoumarins are derived from 7-hydroxycoumarin (umbelliferone) by condensation with a C_5 isoprenoid unit and subsequent modification of the side chain. Furanocoumarins display a number of interesting biological properties. For example, psoralen inhibits the seed germination of *Psoralea subacaulis* until it is leached out into the soil where it may then inhibit the germination of other seeds (Harborne, 1980). Many furanocoumarins are antibiotically active. Protein synthesis in *E. coli* cells is inhibited in the presence of psoralen, possibly mediated by the binding of the furanocoumarin to tRNA. Furanocoumarins also bind to chromosomal DNA, blocking the movement of DNA polymerase, and thus inhibiting DNA synthesis (Murray et al., 1982). Furanocoumarins, administered either orally or topically can induce melanin formation in cases of vitilago in mammals (Murray et al., 1982).

The furanocoumarin biosynthetic pathway has not been well characterized. It is unclear if 4CL is required for hydroxycinnamic acid activation prior to isomerization of the side-chain. It has been suggested that either hydroxycinnamoyl-CoA thioesters or hydroxycinnamoyl-glucose esters are furanocoumarin precursors. Three membrane associated enzymes are required for formation of the furan ring. Activity of these enzymes increases in elicitor-treated parsley cell suspension cultures. The central ring is hydroxylated

by the action of SAM:xanthotoxol O-methylatransferase (XMT) and SAM:bergaptol O-methyltransferase (BMT) (Hahlbrock and Scheel, 1989). The activity of each enzyme is increased in parsley cell suspension cultures treated with an elicitor.

1.4 POPLAR

Poplar is rapidly emerging as a useful model system for studying the molecular biology of trees. F1 hybrids derived from a cross between *Populus trichocarpa* and *P. deltoides* are of commercial interest on the west coast of North America because of their desirable characteristics, including accelerated growth (Heilman and Stettler, 1985). A three-generation pedigree is now available for some lines (Bradshaw et al., 1994). Other features that make poplar a useful model woody system are its small genome size (680 MB; Parsons et al., 1989), the ability to generate clones by vegetative propagation and the potential to transform and regenerate transgenic trees (Fillatti et al., 1987; Van Doorsselaere et al., 1995a). The *Populus* genome is being mapped using a combination of RFLP, sequence-tagged sites, RAPD and AFLP markers (Bradshaw et al., 1994; Boerjan et al., 1997). The genome maps are being used to determine the position of quantitative trait loci that control traits such as spring bud flush and stem volume (Bradshaw, 1996). Floral homeotic genes are being characterized with the goal of genetically engineering sterility in transgenic poplars expressing other desirable traits (Strauss et al., 1997). In addition, a collection of ESTs is being generated from wood-forming tissues (Regan et al., 1997).

The genus *Populus* is rich in phenolic compounds, and a few examples will be discussed. The bud exudate of poplars contains a complex mixture of phenolic compounds, the composition of which is characteristic of a species. The bud exudate is collected by bees

and incorporated into propolis which is used as a glue in their hives (Scaysbrook et al., 1991). Propolis is also collected by humans and used in homeopathatic and medical cosmetic preparations (Greenway et al., 1988).

The bud exudates of a series of poplars with resistance to *Melampsora medusae* (which causes leaf rust) and *Xanthomonas populi* (which causes bacterial canker) have been analyzed for phenolic composition. Scaysbrook et al. (1991) found no correlation between the phenolic composition and resistance to either pathogen, but they did, find lower levels of 4-coumaric acid in more resistant trees and suggested that 4-coumaric acid is converted to phytoalexins which were not detected in this study (Scaysbrook et al., 1991). Pinocembrin, a 5,7-dihydroxyflavanone, is the major constituent of leaf resin of *P. deltoides*. It has been shown to have antifungal activity; it inhibits spore germination of fungal pathogens including *M. medusae* (Shain and Miller, 1982). Phenolic glycosides are produced in freshly wounded stem sections of *P. tremuloides* after inoculation with mycelium from two pathogenic fungi *Hypoxylon mammatum* and *Alternaria sp.* (Flores and Hubbes, 1980).

Many of the cDNA clones or genes encoding lignin biosynthetic enzymes have been isolated from poplar, including *PAL* (Subramaniam et al., 1993; Osakabe et al., 1995ab), *C4H* (Ge and Chiang, 1996; N. Mah and C. Douglas, unpublished results) and *CAD* (van Doorsselaere et al., 1995b). At the time my project was started, only poplar 4CL proteins had been studied. Because of the availability of these lignin biosynthesis genes and the other advantageous features, poplar is becoming increasingly useful as a system in which to study the biosynthesis of lignin.

1.5 PROJECT OBJECTIVES

When this project was started, reports in the literature suggested that 4CL isoforms exist in many plants and some of these isoforms may differentially use hydroxycinnamic acids as substrates. These isoforms could be encoded by a gene family that is differentially expressed. If this were the case in hybrid poplar, it might offer the possibility to genetically engineer transgenic poplar trees with a changed lignin composition.

Therefore, the purpose of this project was to determine if a *4CL* gene family exists in hybrid poplar, and if so, to test the hypothesis that multiple *4CL* genes encode enzymatically distinct proteins in poplar, as suggested by Grand et al. (1983). This was the first step in evaluating the potential for manipulating phenylpropanoid metabolism, specifically lignin biosynthesis, in poplar via genetic engineering of *4CL*. A poplar hybrid clone, H11, an F_1 hybrid derived from a cross between *P. trichocarpa* and *P. deltoides* (Bradshaw et al., 1994) was chosen as the plant material.

CHAPTER TWO

MOLECULAR ANALYSIS OF *4CL* GENES

2.1 INTRODUCTION

4CL activity is required to catalyze the conversion of hydroxycinnamic acids to the corresponding thioesters which are then directed into specific phenylpropanoid biosynthetic branch pathways. When this project was started 4CL proteins had been widely studied, but only parsley (Lozoya et al., 1988; Douglas et al., 1987) and potato (Becker-André et al., 1991) clones had been isolated and studied. However, during the past few years *4CL* genes have been cloned and characterized from many plants including Arabidopsis, *Lithospermum*, loblolly pine, soybean, tobacco, vanilla, and very recently, aspen (Lee et al., 1995a; Ehlting et al., 1998; Yazaki et al., 1995; Voo et al., 1995; Zhang and Chiang, 1997; Uhlmann and Ebel, 1993; Lee and Douglas, 1996; Brodelius and Zue, 1997; Hu et al., 1998). Various *4CL* gene sequences have also been deposited in GenBank.

It is becoming apparent that *4CL* is encoded by a small gene family in all plants. Both parsley and potato appear to have two, almost identical, *4CL* genes (*4CL1* and *4CL2*). The potato genes show no evidence of differential expression (Becker-André et al., 1991). The parsley genes are expressed in a similar manner (Douglas et al., 1987), however *4CL1* is wound-inducible in the roots and *4CL2* is expressed in the flowering stem (Lois and Hahlbrock, 1992). At least, three *4CL* genes exist in Arabidopsis (*At4CL1*, *At4CL2* and *At4CL3*) (Lee et al., 1995a; Ehlting et al., 1998). They appear to be differentially expressed. *At4CL1* is highly expressed in bolting stems, roots (Lee et al., 1995a), siliques and to a lesser extent leaves (Ehlting et al., 1998). *At4CL2* is highly expressed in roots and siliques and

At4CL3 appears to be expressed at low levels in roots bolting stems, leaves, flowers and siliques (Ehlting et al., 1998). Two divergent *4CL* cDNA clones were isolated from tobacco; they are expressed in a similar manner, but at slightly different levels (Lee and Douglas, 1996). Two 4CL cDNA clones isolated from soybean cell suspension cultures appear to be differentially regulated with respect to elicitor treatment (Uhlmann and Ebel, 1993), while two 4CL genes from *Lithospermum* cell suspension cultures show a difference in the level of expression after treatment with light (Yazaki et al., 1995). Voo et al. (1995) found a single *4CL* cDNA in differentiating xylem of loblolly pine, but Zhang and Chiang (1997) demonstrated that two loblolly pine genomic clones containing *4CL* genes of identical sequence were non-allelic.

A report from Grand et al. (1983) suggested that multiple 4CL isoforms exist in *Populus eumaricana* (poplar), and that these display different substrate-utilization patterns. If, indeed, multiple 4CL isoforms exist in poplar they may play an important role in channeling carbon through the various phenylpropanoid biosynthetic branch pathways. At the time this thesis was started neither *4CL* cDNA nor genomic clones had been isolated and characterized. It was unknown whether the reported 4CL isoforms were encoded by different genes or were a result of post-translational modifications of one gene product.

This project was undertaken to determine whether *4CL* is encoded by a gene family in poplar and to test the hypothesis that multiple genes encode enzymatically distinct proteins, as suggested by Grand et al. (1983). This was the first step in evaluating the potential for manipulating phenylpropanoid metabolism, specifically lignin biosynthesis, via genetic engineering of *4CL*. A poplar hybrid clone, H11, an F_1 hybrid derived from a cross between

P. trichocarpa and *P. deltoides* (Bradshaw et al., 1994) was chosen as the plant material. Similar hybrids are of commercial interest on the west coast of North America. A previously constructed cDNA library made from poplar H11 young leaves was available (Subramaniam *et al.*, 1993), as was a previously isolated 1.6 kb putative 4CL clone, isolated from this library (Douglas et al., 1992).

The work described in this chapter was published in Allina et al. (1998).

2.2 MATERIALS AND METHODS

2.2.1 cDNA library screening

Approximately 2 X 10^5 p.f.u from a previously described cDNA library constructed in λZAPII (Stratagene) using poly(A)$^+$ from poplar H11 young leaves (Subramaniam et al., 1993) were screened using a probe generated from a 1.5 kb poplar 4CL cDNA previously isolated from this library (Douglas et al., 1992). Filters were washed at low stringency (2X SSC, 0.1% SDS. 65°C) as described by Sambrook et al. (1989).

2.2.2 Sequencing and sequence analysis

Both strands of 4CL cDNA clones 4CL-2, 4CL-9 and 4CL-B2 were sequenced using the T7 Sequencing Kit™ (Pharmacia); some sequence information was obtained from the University of British Columbia Nucleic Acid-Protein Service Unit using the PRISM Ready Reaction DyeDeoxy Terminator Cycle Sequencing Kit. DNA and predicted amino acid sequences, predicted molecular weights, and isoelectric points were analyzed using the University of Wisconsin Genetics Computer Group software.

2.2.3 Nucleic acid methods

Standard molecular biology techniques were performed as described by Sambrook et al. (1989). Poplar genomic DNA was isolated as described (Subramaniam et al., 1993). Total poplar RNA was extracted using the method of Hughes and Galau (1988). For northern blots, ten μg of RNA was electrophoresed in a 1.2% agarose gel containing formaldehyde, rinsed with water for 45 minutes, and stained with 0.5 μg/ml ethidium bromide. The RNA was blotted overnight onto Hybond-N (Amersham). The blots were washed at moderate stringency (0.5x SSC, 0.1% SDS, 65°C). To ensure that all lanes were evenly loaded, the

blots were stripped and re-hybridized with a probe for a pea rRNA gene (Jorgensen et al., 1982).

DNA fragments were radioactively labeled using the Random Primers DNA Labeling System (Gibco BRL).

2.2.4 Construction of chimeric 4CL-216

PCR was used to engineer a chimeric 4CL2-like full length cDNA clone. Briefly, two 4CL specific primers, primer A (5'-ATAAGAATGCGGCCGCTCTTTCATTCTCTGTT CCAGA-3') and primer B (5'-AACTTGTTGTGCCACAC-3') were used to amplify a predicted 670 bp PCR fragment from H11 genomic DNA. Amplified fragments were digested with SpeI and NotI to enable the cloning of the fragment into the SpeI and NotI restriction sites of the vector Bluescript KS^+. Any 4CL-9-like clones were eliminated by digestion with SacI and ApaI, as 4CL-9 has an internal SacI restriction site and both restriction enzymes cut externally of all PCR clones. Ten clones, similar to 4CL-2 on the basis of restriction digests were fully sequenced in one direction. PCR fragment A16 was chosen to replace the 4CL-2 5'end and create 4CL-216.

2.2.5 Plant material

Clonally propagated individuals of poplar hybrid (*Populus trichocarpa* Torr. & Gray X *P. deltoides* Marsh) H11 were used for tissue isolation. Secondary xylem was isolated from 4- to 6-year old trees grown in the field at the University of British Columbia. Trees were harvested in May, debarked, and developing xylem was scraped off the stems using razor blades. Harvested material was placed immediately in liquid N_2 and stored at -80°C until use.

2.3 RESULTS

2.3.1 Cloning and characterization of Poplar 4CL cDNAs

A cDNA library made from the young leaves from hybrid poplar (*Populus trichocarpa* X *P. deltoides*, clone H11) was screened using a 1500-bp fragment of a putative *4CL* cDNA (4CL-B2), previously isolated from this library using a potato *4CL* cDNA clone (Becker-André, 1991), as a hybridization probe (E. Molitor and C. Douglas, unpublished). Sixteen putative clones were identified. Restriction digest and Southern blot analysis allowed these clones, including the original 4CL-B2, to be placed into 3 groups. Group 1 clones contained a ClaI site 1000 bp from the 3' end, group 2 clones contained a ClaI site 1450 bp from the 3' end and multiple EcoRV sites, and group 3 contained clones with a unique DraI site at the 3' end. Each group contained clones of different lengths. The longest clones within each group 4CL-2, 4CL-9, and 4CL-B2, were approximately 2200 bp, 1950 bp and 1500 bp, respectively. A detailed restriction map of these clones is shown in Figure 2.1.

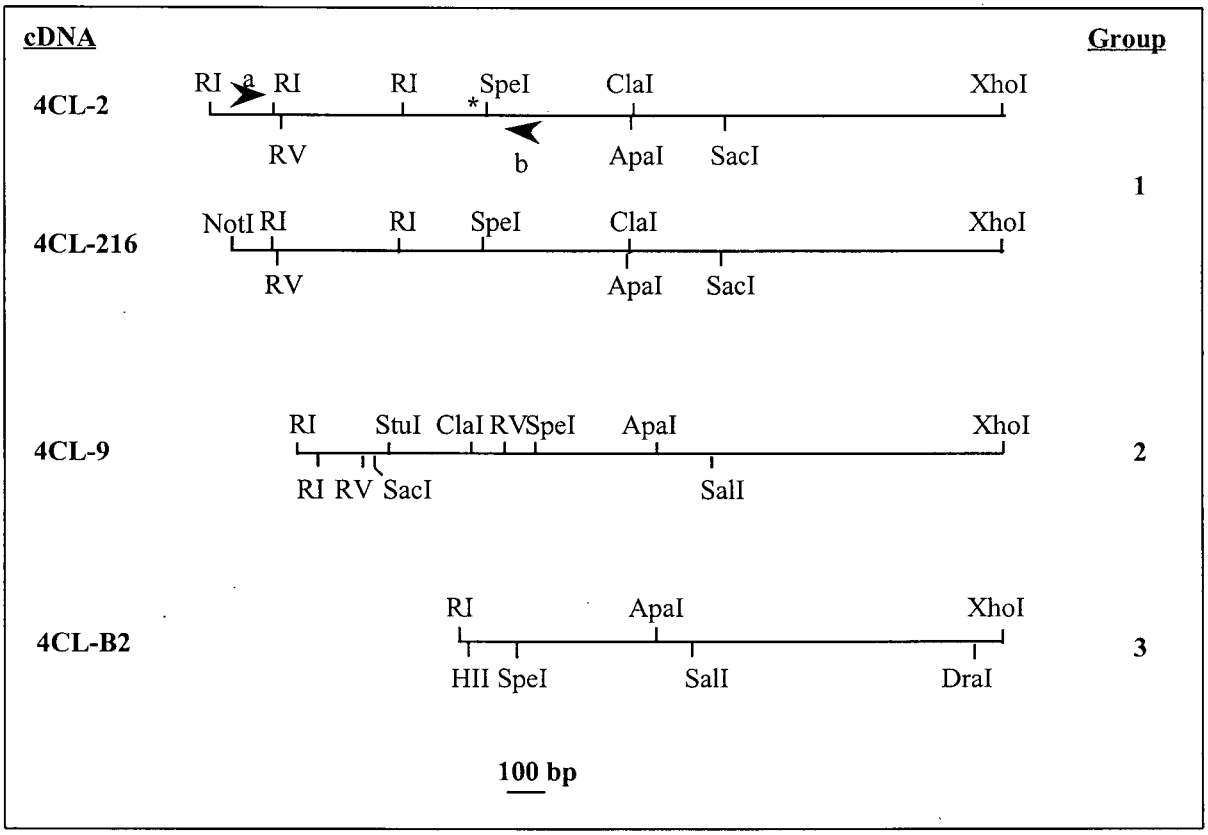


Figure 2.1 Restriction maps of poplar 4CL cDNA clones. RI, EcoRI; RV, EcoRV. Arrowheads represent placement of *4CL*-specific primers (a and b) used to amplify a 670-bp H11 genomic DNA fragment to replace the 5' end of 4CL-2, which contained an erroneous stop codon. The asterisk represents the approximate position of that stop codon.

The three longest clones were fully sequenced and the predicted nucleotide and amino acid sequences were determined. 4CL-2 and 4CL-9 clones contained in-frame methionines (the putative start codon) and in 4CL-2 an in-frame stop codon was located upstream of the start codon. 4CL-9 had the potential to encode a protein of 548 amino acids, while 4CL-B2 contained an open reading frame of 400 amino acids. The nucleotide identity, within the coding region, between 4CL-9 and 4CL-B2 was 98% and the identity between the predicted amino acid sequences was 96%. However, there was some divergence between the 3'

untranslated regions and several gaps were required to maintain sequence alignment. With such a high degree of identity, it is possible that 4CL-9 and 4CL-B2 are allelic. 4CL-2 contained two open reading frames with extensive similarity to known 4CL proteins; the open reading frames were, however, interrupted by an in-frame stop codon at amino acid position 170 (660 bp). Comparison of the 4CL-2 and 4CL-9 nucleotide sequence in this region (Figure 2.1) showed that to maximize alignment between the two clones a gap had to be introduced into the 4CL-9 sequence. This suggested the stop codon in 4CL-2 had been generated by a single-base-pair frame-shift mutation. I hypothesized that 4CL-2 either represented an mRNA transcribed from a non-functional pseudogene or that the stop codon was an artifact of the cloning process.

```
                .         .         .         .         .
492 TTACAGAAAACTGCTTGCA.TTTCTCAGAGTTGACAAGCTCTGATGAGAA 540
      |||||||||| ||||||  ||||||  ||||||||||   |||||||||
619 CACCAGAAAACTACTTGCACTTTCTCGGAGTTGACAAATTCTGATGAGGA 668
```

Figure 2.2 Sequence alignment of 4CL-9 (top) and 4CL-2 (bottom) in the region of the stop codon. A gap introduced to maximize the alignment is indicated by a dot in the 4CL-9 sequence. The stop codon in 4CL-2 is underlined.

To determine if the stop codon was an artifact of cloning, we used *4CL*-specific primers (approximate position shown in Figure 2.1) to amplify H11 genomic DNA by PCR. The expected size of the PCR product was 670 bp, and covered the predicted 5' end of the 4CL genes, within the first exon. Of the ten 4CL-2 like PCR products fragments cloned and sequenced (the full sequence of each of the PCR products can be found in the Appendix, none contained an in-frame stop codon and all showed a high degree of sequence similarity to 4CL-2 but not to 4CL-9. This suggests that the original 4CL-2 stop codon was an artifact

of cloning and did not originate from a non-functional gene containing a stop codon at this position. However, none of the 4CL-2-like clones were identical in nucleotide sequence to 4CL-2, possibly due to errors in PCR amplification. Fragment A16 was chosen to replace the 5' sequence of 4CL-2 because it was the most similar in sequence to the original 4CL-2, with the only amino acid change at position 198. This change was to a proline from an alanine and all known 4CL sequences at that time contained a conserved proline at this position. This new chimeric clone, 4CL-216, was used in all subsequent work as a representative of 4CL-2-like genes.

A comparison of the nucleotide and predicted amino acid sequences of 4CL-216 and 4CL-9 and other known 4CL amino acid sequences is shown in Table 2.1.

Table 2.1 Comparison of 4CL-9 and 4CL-216 nucleotide and predicted amino acid sequence to each other and to other full-length 4CL sequences..

	4CL-216			4CL-9			
	% NT Ide.	% AA Ide.	% AA Sim.	% NT Ide.	% AA Ide.	% AA Sim.	ACCESSION NUMBER
4CL-216	-	-	-	86.9	85.9	92.7	AF008184
4CL-9	86.9	85.9	92.7	-	-	-	AF008183
VP4CL	71.7	78.5	88.6	72.9	79.2	88.9	X75542
Nt4CL19	72.3	78.1	88.1	72.1	78.6	89.0	U50846
Nt4CL1	72.9	77.4	88.3	72.6	77.1	88.5	U50845
Tob4CL	72.1	77.5	87.9	71.7	77.1	88.1	D43773
St4CL[a]	nd	77.4	87.8	nd	76.4	88.7	M62755
Pc4CL	72	75.5	86.3	71.7	74.7	85.2	X13324
Pt4CL1	72.4	74.1	86.3	72.8	74.6	85.3	AF041049
Le4CL1	70.7	73.6	85.2	71.3	73.8	84.8	D49366
At4CL2	69.3	71.3	85.0	69.2	71.3	84.5	[na]Ehlting et al., 1998
At4CL	67.7	71.1	83.1	67.7	70.6	82.7	U18675
PT4CL2	64.1	67.2	82.0	64.7	68.4	82.0	U12013
PT4CL1	64.2	66.9	82.1	64.8	68.2	82.0	U12012
Pt4CL2	65.5	66.9	80.7	65.1	64.8	79.9	AF041050
At4CL3	65.1	66.3	80.7	65.1	66.4	79.8	[na]Ehlting et al., 1998
Os4CL1[a]	nd	60	78.6	nd	60.6	76.9	X52623
Os4CL2	60.9	60.4	79.5	60.2	59.6	78.6	L43362
Le4CL2	65.5	61.5	79.2	65.2	59.3	77.4	D49367
Tob4CL2[b]	75.7	81.1	91.7	91.7	75.3	89.4	D50033
SB4CL[b]	58.7	55.6	70.2	58.6	54.3	71.7	U23787
GM4CL14[b]	74	78.9	88.9	74.1	80.2	88.5	X69954
GM4CL16[b]	65.2	70.3	85.4	64.5	66.2	82.6	X69955

[a]Genomic clones, nucleotide identity not determined
[b]Comparison to partial length sequences
[na]Not available, reference given

Within the coding region, the nucleotide identity between 4CL-216 and 4CL-9 is 87% and the amino acid identity is 86% (93% similar). At the nucleotide and amino acid level, both 4CL-216 and 4CL-9 show the highest identity to a full-length *Vanilla* 4CL clone (Brodelius and Xue et al., 1997) and the lowest identity to a partial clone from *Sorghum bicolor* (Hipskind, 1995) and to two full-length rice *4CL* clones (Zhao *et al.*, 1990; Zhao et al., unpublished). It is expected that, poplar, as a dicot, would show highest similarity to other dicot clones (such as *Vanilla*) and less similarity to monocot clones (such as rice). Figure 2.3 illustrates the similarity between all the known 4CL sequences. Two sequences found in 4CL-216 and 4CL-9 are conserved in all known 4CL sequences. The sequence PYSSGTTGLPKG (motif I) is postulated to be part of the AMP-binding domain that is common to both eukaryotic and prokaryotic ATP-dependent enzymes (Schröder, 1989). GEICIRG (motif II) is hypothesized to be important in the catalytic activity of the enzyme (Becker-André et al., 1991). Six conserved cysteines, including one in the GEIRCG motif are also found in the poplar 4CL sequences.

Figure 2.3 Amino acid sequence comparison of 4CL-9, 4CL-216 and other known 4CL sequences. The full predicted amino acid sequence of 4CL-9 is given; amino acid in 4CL-216 are shown only where they differ from 4CL-9. Amino acids identical in all known 4CL sequences, including those of 4CL-216 and 4CL-9, are shown (IDE). A gap introduced to maximize the alignment is indicated by a dot in the 4CL-9 sequence. Motifs I (part of the AMP-binding site) and II (putative catalytic site) are boxed; conserved cysteine residues and the putative stop codons are indicated by asterisks.

```
                                                  *           *
4CL-9    MEANKDQVQE FIFRSKLPDI YIPNHLPLHT YCFEKLSQFK DNPCLINGPT      50
4CL-216     KN  A              H              N  R
IDE      ---------- ---RSKL-DI -I---L-LH- Y--------- ---C-I----


4CL-9    GDIYTYADVE LTSRKVASGL YKLGLQQGDV ILLLLQNSPE FVFAFLGASF     100
4CL-216   E H   E              N   IK                          I
IDE      ---------- ---R------ ---G------ ----L----E ----F--AS-


4CL-9    IGAISSTANP FYTSAEIAKQ ATASKAKLII THAAYAEKVQ QFAQENDHVK     150
4CL-16        T        P  V                Q V       E VK  V
IDE      -G------NP -----E---- ---------- -------K-- ----------


                                                        I
4CL-9    IMTIDSLTEN CLHFSELTSS DENEIPTVKI KPDDIMALPY SSGTTGLPKG     200
4CL-216   V V  PP    Y        N    DD  A E  N   VV
IDE      ---------- ---------- ---------- ------A-P- SSGTTGL-KG


                                                 *               *
4CL-9    VMLTHKGLVT SVAQQVDGEN PNLYFHERDV ILCVLPLFHI YSLNSVFLCG     250
4CL-216                                 K                       L
IDE      V-LTH----- -VA--VDG-- -N------DV --C-LP-FHI ---N------


4CL-9    LRAGSAILVM QKFDTVSLMD LVQKYKVTIA PLVPPICLAI AKSPVVDQYD     300
4CL-216    V     L     EI T  E             F   VV  V   C    K
IDE      -R-------M --F------- ---------- ---PP----- -K-------D


4CL-9    LSSIRTVLSG AAPLGKELED TVRAKLPNAK LGQGYGMTEA GPVIAMCLAF     350
4CL-216         M      M                                  LS
IDE      -S--R---SG AAP-G----- ------P-A- -GQGYGMTEA GPV--M---F


                   *                                 II*
4CL-9    AKEPFEIKSG ACGTVVRNAE MKIVDPETGE SQPRNKTGEI CIRGCQIMKG     400
4CL-216                                 D  R  L   QS         S
IDE      -K-P---K-G -CGTVVRN-- -K-----T-- S---N--GEI CIRG--IMK-


4CL-9    YLNDPEATER TIDKDGWLHT GDIGYID.ED ELFIVDRLKE LIKYKGFQVA     450
4CL-216             V N              GD
IDE      Y-N----T-- T----GWL-T G-----D--- E-FIVDR-KE -IK--GFQV-


4CL-9    PAELEAMLIA HPNISDAAVV PMKDEAAGEV PVAFVVRSNG SKITEDEIKQ     500
4CL-216              D   C               I     A
IDE      P-ELE----- HP---D-AVV --------E- P-AFV----- ----E---K-


4CL-9    YISKQVIFYK RIGRVFFTEA IPKAPSGKIL RKDLRARVSA GDLPCTSDS*     550
4CL-216       V      S                          LAT   FLIKFQ..
IDE      ---K-V--YK ----V----- --K------- ---------- ----------

4CL-9     ----------
4CL-216  HDTYMQKQQ*
IDE      ----------
```

2.3.2 Genomic organization of *4CL*

Southern blot analysis was used to assess the size of the potential *4CL* gene family in poplar and to examine the relationship between the genes from which cDNA clones 4CL-216 and 4CL-9 were derived. Genomic DNA was isolated from parental (*P. trichocarpa* 93-968 and *P. deltoides* ILL129) and F_1 hybrid (53-246) poplar clones that are part of a three-generation pedigree (Bradshaw et al., 1994). Thus, the inheritance of 4CL-specific restriction fragments could be followed so that allelic DNA polymorphisms could be distinguished from duplicate genes.

The same blot containing restriction-digested DNA from these clones was hybridized sequentially to the 4CL-216, 4CL-9 and 4CL-B2 probes (Figure 2.4). Only EcoRV could digest the 4CL-216 and 4CL-9 cDNAs, but all enzymes could potentially digest within the *4CL* genes. High-stringency hybridization to the 4CL-216 and 4CL-9 probes revealed that each probe hybridized strongly to a distinct set of restriction fragments in parental clones 93-968 and Ill129, and that the F_1 hybrid, 53-246, contained combinations of these fragments. This indicates that the two cDNA clones were derived from separate genes, each present in the parental genotypes. Additionally, there were some weakly hybridizing fragments present. Some of these fragments could be attributed to cross-hybridization of the respective probes to fragments that strongly hybridized to the other probe, but others were distinct from these. These fragments were most easily seen after low-stringency hybridization to the 4CL-B2 probe.

These data indicate that 4CL-216 and 4CL-9 cDNA clones are derived from two single genes and that 4CL-B2 is potentially allelic to 4CL-9. To distinguish the cDNA clones

from the genes they represent, the clones (and their encoded proteins) are referred to as 4CL-216 and 4CL-9, and the respective genes are called *4CL1* and *4CL2*.

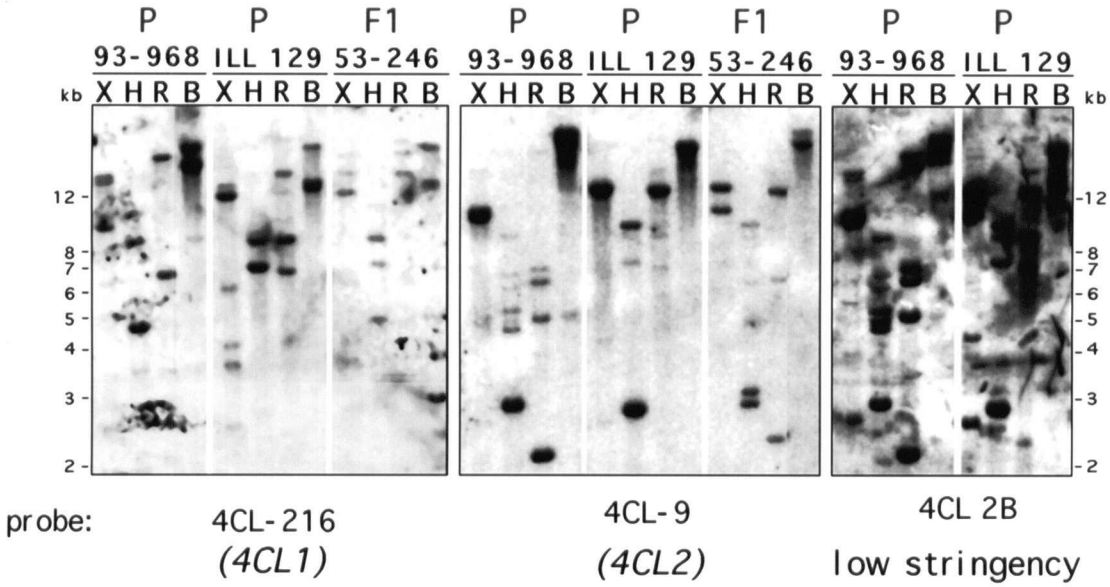


Figure 2.4 Genomic Southern blot analysis of 4CL genes in parental and hybrid poplar genotypes in a controlled cross. Genomic DNA (10 μg) from parental genotypes *P. trichocarpa* 93-969 and *P. deltoides* Ill129 and F_1 progeny 53-246 was cut with the four enzymes indicated and a Southern blot was prepared. The blot was hybridized sequentially to the cDNA probes indicated. X, XbaI; H, HindIII; R, EcoRV; and B, BamHI. Migration of size markers (in kilobases) is shown.

2.3.3 Expression of *4CL* genes

To examine the developmentally regulated and inducible expression patterns of the cloned poplar *4CL* genes, duplicate northern blots were prepared using RNA isolated from various H11 tissues and from an H11 cell suspension culture that had been treated with a PGA lyase elicitor, which activates defense responses in poplar cell cultures (Moniz de Sá et al., 1991). The blots were hybridized with 4CL-216 (*4CL1*) or 4CL-9 (*4CL2*) under conditions which allowed some cross-hybridization between the genes (not shown).

Figure 2.5 shows that the steady-state *4CL1* mRNA was highest in old leaves, with lower levels seen in young leaves, green stem and xylem. In comparison, the highest level of steady-state *4CL2* mRNA levels was in young leaves, with lower levels seen in old leaf, green stem and xylem. *4CL1* expression in green stem and xylem appeared to be stronger than *4CL2* expression in these tissues. Because there was some cross-hybridization between the probes, it is possible that some of the apparent low levels of *4CL2* expression is attributable to *4CL1* transcripts and that some of the *4CL1* hybridization is attributable to cross-hybridization to *4CL2* transcripts. Thus, it appears that both genes are expressed in the same tissues, but at different levels. Surprisingly, neither gene was induced after elicitor treatment, although similar blots loaded with the same RNA samples and hybridized to poplar *PAL* and *C4H* probes showed that the expression of these genes is strongly stimulated by the elicitor treatment (N. Mah and C. Douglas, unpublished data).

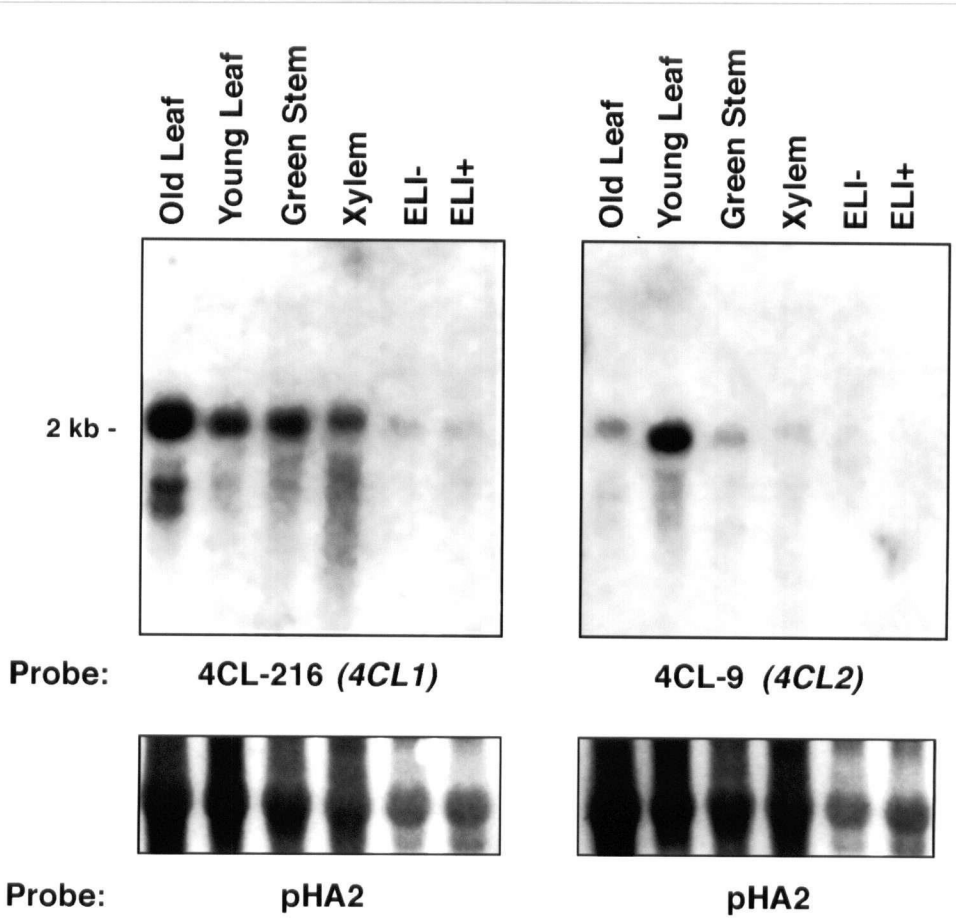


Figure 2.5 Northern blot analysis of poplar *4CL* mRNA levels. Total RNA (10 μg) from different tissues and organs was separated on duplicate formaldehyde agarose gels, transferred to nylon membranes, and hybridized to either 4CL-216 or 4CL-9 cDNA probes. ELI+, Elicitor-treated cell suspension culture; ELI-, untreated cells. The membranes were stripped and rehybridized with an rRNA probe (pHA2) to examine the eveness of loading.

2.4 DISCUSSION

4CL appears to be encoded by a relatively small gene family in all species examined to date. Although poplar 4CL proteins were characterized by Grand et al. (1983), this is the first study of the poplar *4CL* gene family. At least three classes of cDNAs were isolated from an H11 young leaf library. Two classes were represented by apparently full-length clones, 4CL-216 and 4CL-9. Based on sequence analysis they appear to be genuine *4CL* clones. 4CL-216 and 4CL-9 have the potential to encode proteins predicted to be 86% identical (93% similar). Both clones showed significant similarity to 4CL genes of other species, from 79.2 to 54.3% amino acid identity, (88.9 to 71.2% similarity).

Both 4CL-216 and 4CL-9 contain two motifs conserved in *4CL* sequences. The first is composed of a serine/threonine/glycine rich motif followed by a proline/lysine/glycine triplet (PYSSGTTGLPKG). However the proline (underlined) may not be an absolute requirement because the original poplar 4CL-2 clone isolated (with the in-frame stop codon) has an alanine in this position as does a published sequence for *Oryza sativa* (rice) (Zhao et al., 1995). Various authors suggest this is part of the AMP-binding domain because it is found in other (unrelated) genes that encode proteins that bind and hydrolyze ATP to activate molecules for further reaction (Becker-André et al., 1991). Motif II, GEIRCG, is also found in both 4CL-216 and 4CL-9. It is conserved in all known 4CL sequences; and it is thought to be associated with catalytic activity (Becker-André et al., 1991). The cysteine embedded within this motif is one of six cysteines conserved in both of the poplar clones and all other *4CL* sequences.

Southern blot analysis using the first two generations of a controlled *P. trichocarpa* X *P. deltoides* cross (Bradshaw et al., 1994) showed that 4CL-216 and 4CL-9 represented two different genes, *4CL1* and *4CL2*, both of which are present in the parents of the hybrid clone. Additional restriction fragments that weakly cross-hybridize to the cDNA clones are present in the parental and hybrid genomes, suggesting that the poplar *4CL* gene family includes divergent members in addition to *4CL1* and *4CL2*.

Recently, Hu et al. (1998) isolated two divergent 4CL cDNA clones (Pt4CL1 and Pt4CL2) from *Populus tremuloides* (aspen). The 4CL-216 and 4CL-9 sequences are distinct from both of the aspen sequences. Because *P. deltoides* X *P. trichocarpa* and *P. tremuloides* are members of the same genus it is possible that they have a similar gene complement. This suggests that the *4CL* gene family in *Populus* consists of at least 4 members. Supporting this hypothesis, two additional *4CL* clones, 4CL-6 and 4CL-10, have been isolated from *P. trichocarpa* and are quite divergent from 4CL-216 and 4CL-9. 4CL-6 and 4CL-10 show high identity to Pt4CL1 and Pt4CL2, respectively (D. Cukovic and C. Douglas, unpublished data).

The tissue, developmental and inducible expression patterns of the poplar *4CL* gene family members were examined to identify *4CL* genes directly involved in supplying hydroxycinnamoyl-CoA thioesters to specific branch pathways. Northern blot analysis demonstrated that *4CL1* and *4CL2* are differentially expressed. *4CL1* is preferentially expressed in old leaves with lower levels seen in green stem and xylem. Because of the fairly low expression in developing xylem, *4CL1* may not play a major role in the biosynthesis of lignin. *4CL1* may be important in the biosynthesis of leaf phenolics such as pinocembrin (Shain and Miller, 1982) salicin, tremuloidin, tremulacin, catechin, trichocarpin and

salireposide (Pearl and Darling, 1971). *4CL2* is preferentially expressed in young leaves and may be important in the lignification of the developing leaf vascular system and/or flavonoid/phenolic synthesis in that tissue.

Neither gene is activated in H11 cell suspension cultures after treatment with the elicitor, PGA lyase. This was surprising, because both poplar *PAL* and *C4H* (Moniz de Sá et al., 1992; Subramaniam et al., 1993; N. Mah and Douglas, unpublished data) are highly expressed after this treatment. Earlier studies by Moniz de Sá et al. (1992) demonstrated that *4CL* gene expression is induced after PGA lyase treatment of the cell suspension cultures. Results from this same study suggested that treatment with PGA lyase leads to activation of a subset of 4CL isoforms that preferentially use 4-coumaric acid and ferulic acid. In a later study, Dr. Aviva Pri-Hadash demonstrated that a single native 4CL isoform accumulates in response to this treatment (Allina et al. 1998). It is therefore likely that another poplar *4CL* gene exists that is inducible by elicitation.

The poplar *4CL* genes tested in this study do not respond in a similar manner to elicitor as GM4CL16, from soybean (Uhlmann and Ebel, 1993). The high divergence between the poplar and GM4CL16 coding sequences (between 66 and 70% amino acid identity) may be accompanied by a high divergence in their promoter sequences, such that they respond to different regulators. The poplar and GM4CL14 (a non-elicitor inducible soybean *4CL*) sequences are more closely related (approximately 80% identical) and show a more similar expression pattern.

Of considerable interest is the report that *Pt4CL1* (from aspen) is preferentially expressed in secondary xylem (Hu et al. 1998). Three cis-acting sequences (box P, box A and

box L), known to be found within all known *PAL* and *4CL* gene promoters, were found in the promoter sequence of *Pt4CL1* while none were found within the promoter of *Pt4CL2,* which is not expressed in secondary xylem. These differences in promoter sequence are likely responsible for the different expression patterns. Hauffe et al. (1991) showed that a single cis-acting element (from -174 to - 210 bp) in the parsley *4CL-1* promoter is sufficient to direct tissue and cell-specific expression of a GUS reporter gene in transgenic tobacco. It is possible that the differences mentioned above exist in poplar and are sufficient for the differential regulation of the *4CL1* and *4CL2* genes I identified. Differential expression of *4CL* family members has also been shown in parsley (Lois and Hahlbrock, 1992) and soybean (Uhlmann and Ebel, 1993).

Two members (*4CL1* and *4CL2*) of the poplar *4CL* gene family were isolated in this study, but it is clear that other poplar *4CL* family members exist and that they are differentially regulated. The roles of the *4CL1* and *4CL2* in phenylpropanoid metabolism are unclear. Based on their expression patterns, they may supply hydroxycinnamic thioesters to the flavonoid and/or phenolic glycoside branch pathways in leaves. Perhaps the compounds they help to produce are important in UV-protection or as feeding deterrents. *4CL2* may play a role in lignification of the leaf veins, but does not appear to have any appreciable role in the biosynthesis of lignin in developing xylem of the stem. To characterize differences between the proteins encoded by *4CL1* and *4CL2*, the cDNA clones were used to produce recombinant proteins. The properties of the recombinant proteins are examined in the next chapter.

CHAPTER THREE

RECOMBINANT 4CL PROTEINS

3.1 INTRODUCTION

The enzymatic properties of 4CL have been studied in a wide variety of plants including both angiosperms and gymnosperms. 4CL enzymes are normally found to use a number of hydroxycinnamic acids as substrates (Gross and Zenk, 1974; Knobloch and Hahlbrock, 1975, 1977; Ranjeva et al., 1976; Wallis and Rhodes, 1977; Lüderitz et al., 1991; Kutsuki et al., 1982; Maul and Ride, 1982; Grand et al., 1983; Lozoya et al., 1988; Hipskind et al., 1993; Liu et al., 1995; Voo et al., 1995; Lee and Douglas, 1996; Meng and Campbell, 1997). While some species, such as *Forsythia*, are reported to have only one 4CL form (Gross and Zenk, 1974), most plants contain multiple 4CL isoforms. Some 4CL isoforms show differences in their substrate preferences. For example, form I isolated from soybean cell suspension cultures, is most highly active towards ferulic acid, followed by sinapic acid, 5-hydroxyferulic acid, 4-coumaric acid, caffeic acid, and cinnamic acid. Form II prefers caffeic acid, followed by 4-coumaric acid, 5-hydroxyferulic acid, cinnamic acid, ferulic acid and has no activity towards sinapic acid (Knobloch and Hahlbrock, 1974). However, the two isoforms isolated from parsley show no differences in their substrate preferences (Lozoya et al., 1987).

Conflicting data exists in the literature as to the substrate usage of *Populus* 4CL. 4CL proteins have been studied in *P. euramericana* (poplar) (Kutsuki et al., 1982; Grand et al., 1983) and *P. tremuloides* (aspen) (Meng and Campbell, 1997; Hu et al., 1998). Grand et al. (1983) reported that three partially purified 4CL isoforms isolated from poplar demonstrated

difference substrate preference patterns. Each isoform could act on 4-coumaric acid and ferulic acid, form I preferentially used 5-hydroxyferulic acid and sinapic acid, whereas form III preferred caffeic acid. These 4CL isoforms were hypothesized to help determine lignin monomer composition in different tissues, based on their potential to supply the appropriate hydroxycinnamoyl-CoA thioesters to the lignin biosynthetic pathway. However, Kutsuki et al. (1982) tested a variety of species for 4CL activity and found no activity towards sinapic acid in crude xylem extracts of poplar, as did Meng and Campbell (1997) in aspen.

Recombinant proteins were produced in an effort to clarify the enzymatic properties of the proteins encoded by the poplar *4CL* gene family members. While characterization of the recombinant proteins was underway, Dr. Aviva Pri-Hadash, working in parallel in the lab of Dr. Brian Ellis at UBC, determined the properties of the native poplar 4CL isoforms (Allina et al., 1998). Previously, an *E. coli* expression system had been used to produce recombinant tobacco 4CL enzymes encoded by two distinct genes. Both enzymes showed the same pattern of substrate usage which was somewhat similar to the native 4CL activity detected in crude tobacco protein extracts. However, unlike the native proteins, the recombinant proteins readily converted cinnamic acid to cinnamoyl-CoA (Lee and Douglas, 1996).

Thus, to ensure faithful processing of foreign gene products, an insect cell/baculovirus expression system (eukaryotic in nature) was chosen to produce recombinant poplar 4CL proteins. The baculovirus protein polyhedrin is essential in the natural viral life cycle, but is dispensable in cultured Sf9 cells derived from ovarian cells of *Spodoptera frugiperda* (the fall army worm). The polyhedrin promoter is extremely strong and in

cultured Sf9 cells proteins produced under its direction can accumulate to 50 - 75 % of the total "stainable" proteins of the cell (as detected on SDS-PAGE). Large amounts of recombinant proteins produced can be purified and characterized. One disadvantage of the system is that baculovirus infection of insect cells results in the death of the host and hence the necessity to reinfect fresh cultures of Sf9 cells for each round of protein synthesis. Another disadvantage is that the process of glycosylation is somewhat different to that in vertebrates and it has not been examined in detail in plants. Finally, approximately 6 to 8 weeks of preparation time are required before the first recombinant proteins are ready for analysis.

The work described in this chapter was published in Allina et al. (1998). All work on the native 4CL isoforms reported in that paper was performed by Dr. Aviva Pri-Hadash.

3.2 MATERIALS AND METHODS

3.2.1 Expression of recombinant 4CL proteins in baculovirus-infected insect cells

The full length 4CL-9 and 4CL-216 cDNAs inserts were isolated from a Bluescript plasmid by digesting with PstI or NotI, respectively and XhoI. A blunt end was created at the XhoI site by filling in with Klenow polymerase. pVL1392, a baculovirus vector (Webb and Summers, 1990), was digested with PstI or NotI and SmaI. The respective cDNA inserts and pVL1392 were ligated together to create the plasmids pVL1392::4CL-9 and pVL1392::4CL-216. $2x10^6$ cells Sf9 cells (American Type Culture collection, Accession Number CRL-1711) were co-transfected with 2 μg of each plasmid along with 0.25 μg of AcNPV viral DNA (BaculoGold™, PharmMingen). Recombinant baculovirus were plaque-purified (Summers and Smith, 1987) and two of each recombinant virus were chosen for production of high-titre virus stocks. The viral titre was calculated either by plaque assays or by end point dilution (Summers and Smith, 1987). These were tested for production and activity over time. A single virus stock expressing each recombinant protein was selected for further use.

3.2.2 Insect cell extracts

$1x10^8$ Sf9 cells were infected at a multiplicity of infection of 1 and harvested approximately 48 hours post infection. Cells were centrifuged at 1000 xg, for 5 minutes at 4°C. The cell pellet was washed twice with Dulbecco's Phosphate Buffered Saline (no $MgCl_2$, no $CaCl_2$), (Sigma) and then resuspended in 4.2 ml of 200 mM Tris, pH 7.8 for general usage. The cells were lysed using a 15 ml Wheaton Homogenizer. Cellular debris was removed through centrifugation at 15,000 xg, for 15 minutes. Glycerol (30%, v/v) and 14 mM β-mercaptoethanol were added to the supernatant; samples were placed at -20°C for

storage. Infected cells used for subsequent purification of 4CL enzyme activity (using anion exchange chromatography) were prepared as described, but were resuspended in 50 mM Tris, pH 7.8. Before loading on the column, the cell extract was filtered through a 0.22 μm filter. Protein content of cell extracts was quantified by the Bradford (1976) method using BSA as a standard and the Bio-Rad Protein Assay Kit™ (Bio-Rad).

3.2.3 Purification of recombinant proteins using anion exchange chromatography

To partially purify 4CL enzymes from insect cell extracts, crude lysates were subjected to anion-exchange chromatography using an analytical HR5/5 Mono Q column (Pharmacia). Solvent A contained 50 mM Tris pH 7.8, 5% glycerol and 14 mM β-mercaptoethanol; solvent B additionally contained either 500 mM or 1 M KCl. The gradient conditions using 1M KCl consisted of 0% B for 2 ml, 0 - 7% for 5 ml, 7% for 5 ml, 7 - 10% for 15 ml, 10% for 5 ml, 10 - 30% for 30 ml, 30 - 50% for 5 ml, with a flow rate of 0.5 ml/min or 1.0 ml/min. Fractions (1 ml) were collected and glycerol was added immediately to 30% (v/v) and were stored at -20°C.

3.2.4 4CL enzyme assays

4CL enzyme assays were performed as described by Knobloch and Hahlbrock (1977), using a direct spectrophotometric assay to measure the accumulation of hydroxycinnamoyl-CoA thioesters. Briefly, final concentrations in the reaction mixture were: 5 mM ATP, 5 mM $MgCl_2$, 470 mM Tris, pH 7.8, 0.33 mM CoA. The pH of the reaction mixture was adjusted to pH 7.8 with 4 M KOH (0.5 ml per 10 ml). Assays performed after chromatography to identify fractions containing 4CL activity were performed using 0.5 mM 4-coumarate; the assays done to compare substrate specificities were performed using 0.2 mM

hydroxycinnamic acids, and assays to determine the Km's of recombinant 4CL-9 were performed using hydroxycinnamic acids in the range of 0.01 to 0.6 mM, depending on the substrate. Reaction mixtures were pre-incubated for one minute and were started with the addition of CoA, which had been warmed to room temperature for 30 seconds. Thirty seconds after addition, the accumulation of hydroxycinnamoyl-CoA thioester was determined by measuring the change in absorbance at the appropriate wavelengths (311, 333, 346, 345, 352 and 356 nm, according to the absorption maxima for cinnamoyl-CoA, 4-coumaroyl-CoA, caffeoyl-CoA, feruloyl-CoA, sinapoyl-Co and 5-hydroxyferuloyl-Co, respectively; Stöckigt and Zenk, 1975). All assays were performed at room temperature. Enzyme activity is expressed as nkat/mg of protein or pkat/ml of protein solutions.

3.2.5 SDS-PAGE and Western blot anlaysis

Protein samples were electrophoresed using 4% resolving and 10% separating SDS polyacrylamide gels according to the method of Laemmli (1970). The proteins were transferred onto nylon membranes (PVDF, Schleicher & Schuell), blocked with 5% (w/v) non-fat powdered milk, reacted with antisera raised against parsley 4CL (Ragg et al., 1981) at a 1:5000 dilution, and incubated with goat anti-rabbit IgG conjugated to alkaline phosphatase at a 1:500 dilution (Gibco BRL). Alkaline phosphatase activity was visualized using nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate as the substrates (Gibco BRL).

3.2.6 Detection of glycoproteins

Recombinant 4CL proteins were tested for glycosylation using the DIG Glycan Detection Kit (Boehringer). Briefly, recombinant proteins were transferred onto Immobilon

(Millipore) membranes. Glycoprotein detection was performed as described (Boehringer) with development for 1 hour or overnight.

3.2.7 Antibody generation

Polyclonal antibodies were raised in New Zealand White rabbits, cared for at the Animal Care Facility at the University of British Columbia. Two rabbits were pre-bled and the sera were tested for cross-reactivity (at a 1:1000 dilution) against native proteins from poplar xylem, tissue culture and young leaves, Arabidopsis and tobacco stems, against crude cell lysates of recombinant 4CL-216 and 4CL-9, and also against partially purified recombinant 4CL-216 and 4CL-9. The two rabbits were injected subcutaneously with 200 μg anion-exchange-purified recombinant 4CL-9 (mixed with Freund's complete adjuvant, Sigma) and then received two intramuscular booster injections of 200 μg (mixed with Freund's incomplete adjuvant) at approximately 4-week intervals. A final booster injection of 200 μg was given 6 weeks later. The rabbits were sacrificed and their blood was recovered to 50 ml Falcon tubes. The blood was placed at 37°C for 1 hour to facilitate clotting. The edge of the tube was rimmed with a pipet and the tubes of blood were placed at 4°C overnight. The serum was decanted and centrifuged at 1,000 xg for 20 min to remove blood cells (pre-immune blood was treated in a similar manner). Half of the recovered serum was filter sterilized through a 0.2 μm filter and the other half was treated with 0.2 % sodium azide. The sera were stored at -20°C.

3.2.8 Plant material

Plant material was collected and stored as previously described in Chapter 2 (2.2.5).

3.3 RESULTS

3.3.1 Analysis of recombinant 4CL proteins

To study the enzymatic properties of the proteins encoded by 4CL-216 and 4CL-9, recombinant proteins were generated using a baculovirus expression system. In a typical experiment, 50 ml spinner flasks of Sf9 insect cells infected with 4CL-216 and 4CL-9 baculovirus stocks yielded 7.8 and 6.7 mg of total protein, respectively. The 4CL specific activities of the respective crude extracts towards 4-coumaric acid were 4.7 and 17.1 nkat/mg protein (Table 3.1). Both proteins appeared to be quite stable at room temperature, maintaining activity over the 60 minutes tested (about the time taken to isolate the 4CL activity from the anion-exchange column) (data not shown). They both responded in a similar manner to a thermal stability test (Figure 3.1). After treatment at the appropriate temperature for 10 minutes, 4CL-216 activity started to decrease at 40°C and was lost by 45°C, while 4CL-9 activity started to decrease by 45°C and was lost at 50°C.

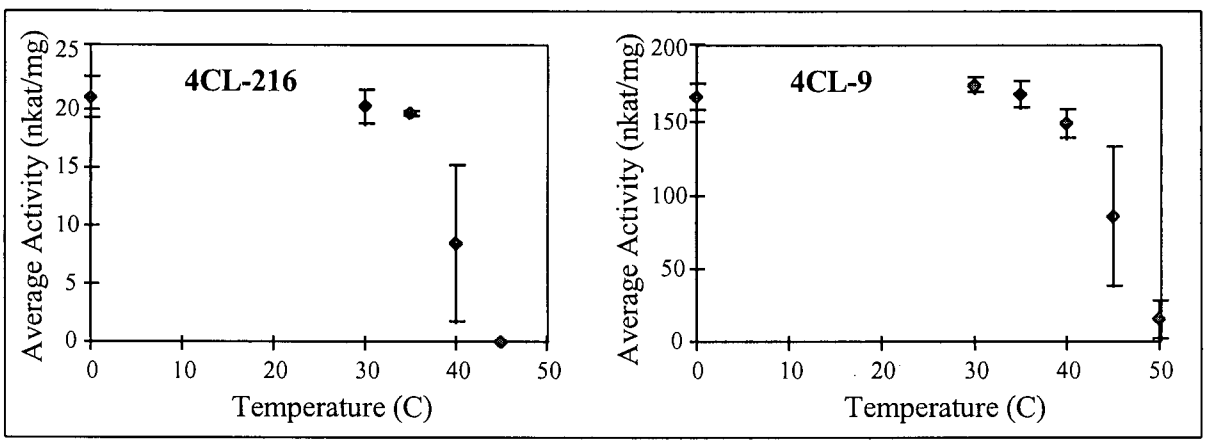


Figure 3.1 Thermal Stability test of 4CL-216 and 4CL-9. For each trial the protein was held at the appropriate temperature for 10 minutes and then cooled briefly on ice. A standard 4CL enzyme assay was then performed. Results are the average of three trials ± SD.

Anion-exchange liquid chromatography was used to separate the recombinant proteins from insect cell and viral proteins. Each protein extract yielded a single peak of 4CL enzyme activity (Figure 3.2), and recombinant 4CL-9 consistently eluted earlier in the KCl gradient than recombinant 4CL-216. Approximately 18 and 45% of the original 4CL activity in crude extracts was recovered in the most active fractions of 4CL-216 and 4CL-9. These pooled, peak fractions contained 117 and 408 μg protein, respectively, with specific activities of 55.6 and 126 nkat/mg of protein (Table 3.1).

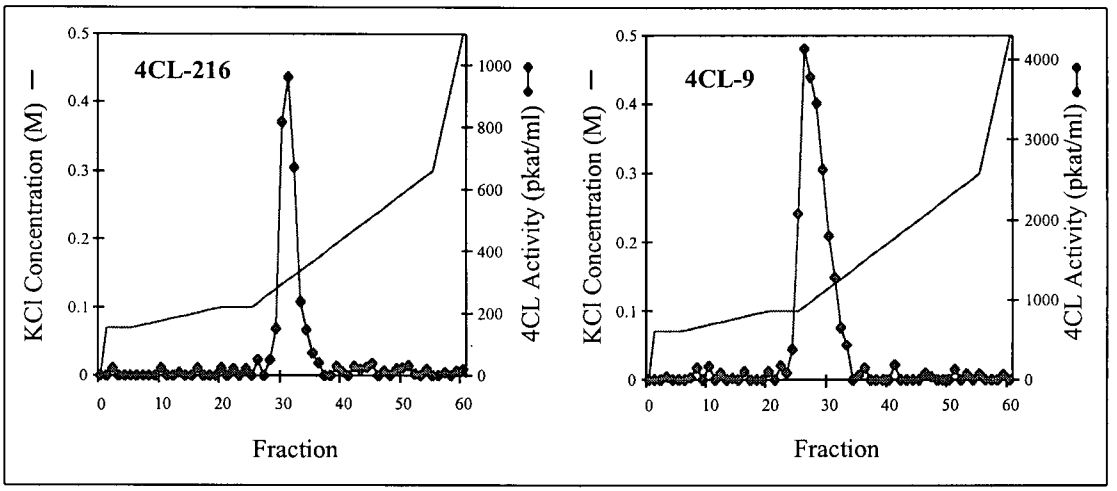


Figure 3.2 Elution profile of 4CL activity after anion-exchange chromatography. 4CL-216 (left hand graph) peak of activity elutes at fraction 31, 4CL-9 (right-hand graph) peak of activity elutes at fraction 27. 4CL enzyme activity was measured using 0.5 mM 4-coumaric acid.

Table 3.1 Partial purification of recombinant 4CL-216 and 4CL-9

		Protein mg	Total Activity nkat	Specific Activity nkat/mg
4CL-216	Crude extract	7.8	36.66	4.7
	Anion exchange	0.117	6.51	55.6
4CL-9	Crude extract	6.7	113.90	17.0
	Anion exchange	0.408	51.41	126.0

The anion exchange-purified recombinant proteins were tested for their abilities to use different hydroxycinnamic acids as substrates. Figure 3.3 shows the substrate utilization profiles obtained using these substrates at a concentration of 0.2 mM. This concentration was used because previously it was shown that 4CL activity was highest at this concentration when working with the native 4CL proteins (Pri-Hadash, unpublished data). The two recombinant proteins showed a similar pattern of substrate usage, with a strong preference for using 4-coumaric acid and decreasing activities towards ferulic acid, caffeic acid, cinnamic acid and 5-hydroxyferulic acid. Neither recombinant protein had any detectable activity towards sinapic acid. Thus, at this concentration, the enzymes are indistinguishable from each other in their abilities to differentially utilize these substrates. However, the specific activity of 4CL-9 using 4-coumaric acid as a substrate was approximately 4 time higher than the specific activity of 4CL-216 towards 4-coumaric acid. This pattern was consistent for all substrates tested.

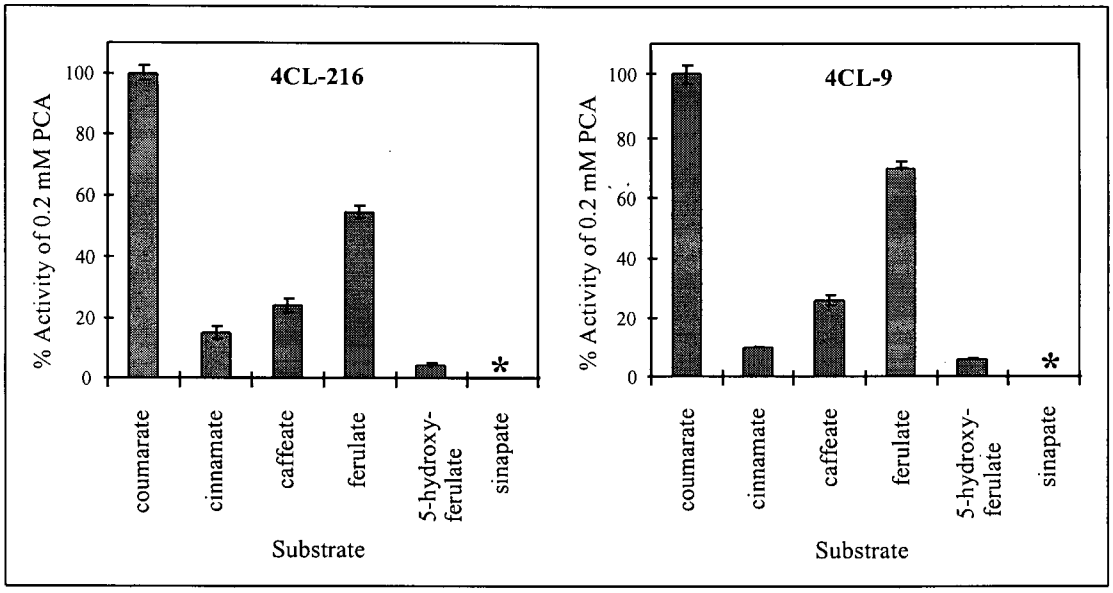


Figure 3.3 Substrate-utilization profiles of recombinant 4CL proteins. Recombinant 4CL-216 (left hand graph), and recombinant 4CL-9 (right-hand graph) enzyme activity was measured using anion-exchange purified proteins and 0.2 mM concentrations of hydroxycinnamic acids. Results are the average of three trials ± SD. Asterisks indicate the absence of detectable activity. Results are reported as a percentage of the activity against 4-coumaric acid (shown as 100%), which was 32.0 nkat/mg protein for 4CL-216 and 119.2 nkat/mg protein for 4CL-9 in the experiment shown.

Anion exchange-purified recombinant 4CL-216 and 4CL-9 protein was used to determine the kinetic properties of the recombinant enzymes. Table 3.2 lists the apparent K_m, V_{max} and V_{max}/K_m determined using 4-coumaric acid, ferulic acid and cinnamic acid as substrates. All of these substrates provided kinetic responses which fit the Michaelis-Menton equation. Caffeic acid did not follow classical Michaelis-Menton kinetics so apparent K_m and V_{max} values were not determined. Insufficient 5-hydroxyferulic acid was available for kinetic analysis and there was no detectable 4CL activity with sinapic acid as a substrate.

Table 3.2 Kinetic properties of recombinant 4CL-216 and 4CL-9

	Apparent K_m (μM)	Relative K_m	Apparent V_{max} (nkat/mg)	Apparent V_{max}/K_m (nkat/mg/μM)	Relative V_{max}/K_m [a]
4CL-216					
4-coumaric acid	11 ± 1 [b]	1	30 ± 3	2.7 ± 0.2	1
ferulic acid	13 ± 3	1.2	20 ± 2	1.5 ± 0.2	0.55
cinnamic acid	1141 ± 465	103.7	33 ± 9	0.03 ± 0.004	0.01
4CL-9					
4-coumaric acid	80 ± 9	1	353 ± 36	4.4 ± 0.1	1
ferulic acid	102 ± 6	1.3	190 ± 11	1.9 ± 0.1	0.43
cinnamic acid	1048 ± 43	13.1	126 ± 3	0.1 ± 0.01	0.02

[a]Values are relative to the 4-coumaric acid value (taken as one).
[b]Results are the average of three trials ± SD

The apparent K_m values showed that each recombinant protein had high affinity for 4-coumaric acid and ferulic acid. Interestingly, 4CL-216 had higher affinities for 4-coumaric acid and ferulic acid than 4CL-9. However, the apparent V_{max}/K_m values indicated that 4CL-9 may be more catalytically efficient with 4-coumaric acid and ferulic acid than 4CL-216. This is consistent with the study of substrate specificity at a fixed concentration (Figure 3.3).

To verify the authenticity of the recombinant 4CL proteins, and to assess the purity of anion-exchange purified 4CL-216 and 4CL-9, the pooled peak fractions were separated by SDS-PAGE in parallel with crude Sf9 and poplar extracts. A duplicate gel was blotted and reacted with an antiserum specific to parsley 4CL (Ragg et al., 1981). Figure 3.4A shows the migration of crude protein extracts, and anion exchange-purified recombinant 4CL proteins in an SDS-PAGE gel. Lanes 1, 2 and 3 contain crude extracts of poplar elicited cell culture, xylem and young leaves, lanes 4 and 5, crude insect cell extracts containing the recombinant

proteins, and lanes 6 and 7 anion exchange-purified recombinant protein. No 4CL protein was detected in the Sf9 insect cells containing the wild type baculovirus (Figure 3.4A, lane 9). The two recombinant 4CL proteins in crude insect extracts (lanes 4 and 5) and after anion-exchange purification (lanes 5 and 6) migrated with mobilities similar to those of the 4CL proteins in poplar tissue extracts, with apparent molecular weights of 60,000. Interestingly, 4CL-216 reproducibly migrated somewhat more slowly than 4CL-9 in SDS-PAGE gels, although the predicted size of the 4CL-216 protein was only 9 amino acids (1235 kd) larger than 4CL-9. Figure 3.4B shows an immuno-blot showing the cross-reaction of these same extracts as in Figure 3.4A to a parsley 4CL-specific antibody (Ragg et al., 1981). Both the recombinant 4CL proteins and the 4CL proteins in poplar crude extracts were recognized by the 4CL antibodies, but there was no reaction of the crude insect cell and wild type baculovirus proteins to this antibody.

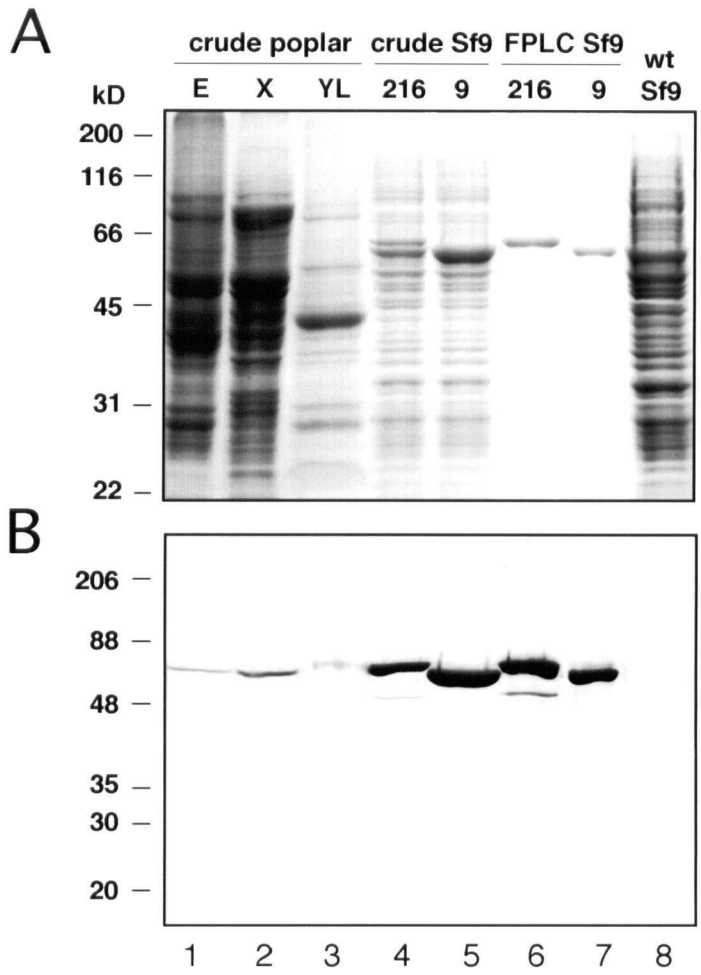


Figure 3.4 SDS-PAGE and immunoblot analysis of recombinant 4CL proteins. SDS-PAGE gel (A) and an immuno-blot of a gel run in parallel (B) reacted with antiserum specific to parsley 4CL (Ragg et al., 1981). Crude extracts (20 μg protein) of Sf9 cells infected with 4CL-216 and 4CL-9 baculovirus constructs, and 4CL-216 and 4CL-9 (1μg each) anion-exchange purified recombinant proteins from infected Sf9 cells, were loaded as shown. Wt Sf9: Wild-type (uninfected) Sf9 cells. Crude poplar extracts were derived from elicitor-treated suspension-cultured cells (E), differentiating xylem (X), and young leaves (YL). Molecular mass standards (in kilodatons) are shown to the left.

The predicted size difference of 4CL-9 and 4CL-216 is 1.2 kD. However, using a calibrated SDS-PAGE gel, the proteins appear to be at least 6 kD different in size. Several experiments were performed in an effort to verify the exact size of the two 4CL proteins and to determine the cause of the difference in apparent size. The expected size of the 4CL proteins can be calculated from the nucleotide sequence, but it is possible that the translation start site of one or both of the recombinant 4CL proteins in insect cells is not as expected. For this reason, N-terminal peptide sequencing was attempted. Unfortunately, both proteins were N-terminally blocked and efforts to de-block were unsuccessful. Electro-spray ionization mass spectrometry was also attempted to determine if there was an actual mass difference between 4CL-216 and 4CL-9. This is a powerful technique because it directly measures the mass of the of a protein, rather than inferring the mass from a size-based gel assay with standards. It is very sensitive (less material required) and accurate (for small peptides to +/- 0.1 Da). However, the protein samples were not sufficiently pure to determine a correct mass. Thus, using these two methods, no data was gained as to the actual size of the 4CL proteins.

It was hypothesized the two proteins could differ in apparent size due to a difference in glycosylation or phosphorylation. The 4CL proteins tested negative for glycosylation (data not shown). They were also tested for possible differences in phosphorylation. Both proteins were treated with 20 units of calf intestinal phosphatase (CIP) and subjected to SDS-PAGE analysis. Both the CIP treated and control proteins migrated in a similar manner (data not shown). Thus, the cause for the difference in apparent size (as determined by a difference in migration on SDS-PAGE) of the 4CL proteins is not known.

3.3.2 Poplar 4CL antibodies

Polyclonal antibodies to recombinant 4CL-9 were produced in two rabbits (A and B). The pre-immune sera were tested for cross-reactivity with a variety of poplar tissues, elicited tissue culture cells, xylem and young leaf, Arabidopsis and tobacco stem and recombinant 4CL-216 and 4CL-9 (crude insect protein extracts and anion exchange-purified). The pre-immune sera detected only a small number of antigenic proteins, and produced very weak interactions even at a 1:1000 dilution (data not shown).

The final sera were tested titrated against partially purified recombinant 4CL-216 and 4CL-9 at both a 1:1000 and a 1:10,000 dilution. Sera from both rabbits could detect as little as 25 ng of protein at a 1:10,000 dilution (not shown). These sera was also tested against the same samples as the pre-immune sera (Figure 3.5). Both sera detected 4CL bands in all samples at a 1:5000 dilution.

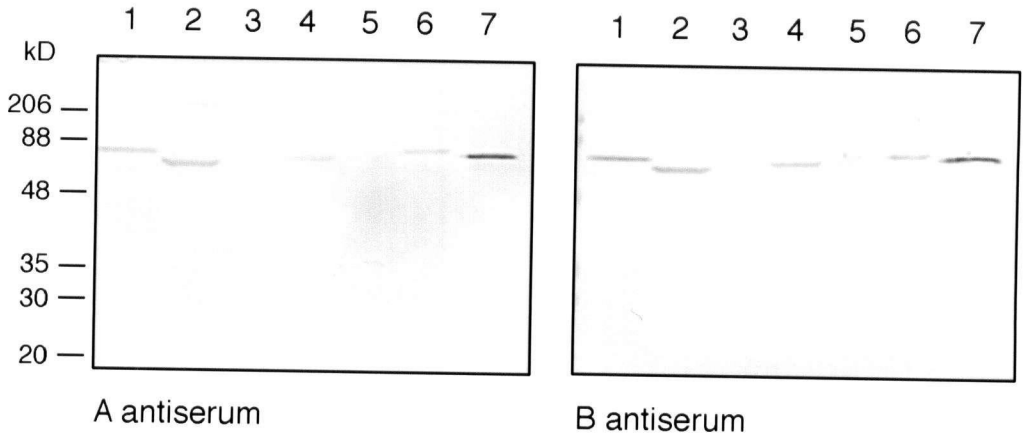


Figure 3.5 Immunoblot of proteins from selected tissues and recombinant 4CL proteins. Crude extracts (20 μg) of (1) Arabidopsis stem, (2) tobacco stem, (3) poplar elicited cell culture, (4) poplar stem, (5) poplar young leaves, anion exchange purified (50 ng) (6) 4CL-216 and (7) 4CL-9. Molecular mass standards (in kilodaltons) are shown to the left.

3.4 DISCUSSION

To examine the properties of the proteins encoded by the poplar 4CL-216 and 4CL-9 cDNA clones, corresponding to *4CL1* and *4CL2* genes, recombinant proteins were produced in an insect cell/baculovirus expression system. The authenticity of the recombinant 4CL proteins was demonstrated by their ability to use hydroxycinnamic acids as substrates (Figure 3.3), and their recognition by antisera raised against parsley 4CL (Figure 3.4B). Antisera raised against recombinant 4CL-9 reacted with an approximately 60 kD band in Arabidopsis, tobacco and poplar (Figure 3.5) that corresponded in size to the 4CL protein. The antisera did not cross-react with proteins from wild-type (non-infected) Sf9 cells.

The baculovirus-expressed proteins encoded by 4CL-216 and 4CL-9 were active in both crude insect cell extracts and in anion-exchange purified fractions. Thus, like other recombinant 4CL proteins from tobacco and aspen (Lee and Douglas, 1996; Hu et al., 1998) it is evident that plant-specific modifications of poplar 4CL are not required for enzymatic activity. While both proteins could be eluted off the anion exchange column with approximately 100 mM KCl, 4CL-9 reproducibly eluted slightly before 4CL-216 (Figure 3.2), indicating that the *4CL1* and *4CL2* genes encode physically separable 4CL isoforms.

The 4CL-216 and 4CL-9 recombinant proteins each displayed similar preferences for the various hydroxycinnamic acids substrates (at a concentration of 0.2 mM). They both preferred 4-coumaric, acid followed by ferulic, caffeic, cinnamic and 5-hydroxyferulic acids. Neither protein used sinapic acid as a substrate (Figure 3.3). This substrate utilization pattern is similar to that observed for the hybrid poplar H11 native 4CL isoforms (Allina et al., 1998). Three physically distinct native 4CL isoforms were detected in a variety of tissues. At

a substrate concentration of 0.2 mM, each isoform preferred 4-coumaric acid, followed by ferulic acid and caffeic acid. Insufficient 5-hydroxyferulic acid was available for extensive assays, and it was therefore not tested as a substrate in that study. Traces of activity detected towards cinnamic and sinapic acids were within the noise limits of the spectrophotometric assay and could not be confirmed by HPLC analysis (Allina et al., 1998).

The substrate usage of the poplar recombinant proteins resembles that of *P. euramericana* and aspen determined by Kutsuki et al. (1982) and Meng and Campbell (1997), respectively. In both those studies, crude protein extracts displayed 4CL activity towards 4-coumaric acid, ferulic acid, caffeic acid and 5-hydroxyferulic acid. There was no activity towards sinapic acid and no mention of cinnamic acid usage.

Thus, no evidence was obtained from my experiments on the enzymatic activity of the recombinant 4CL proteins, nor from the activity of partially purified native proteins in the parallel study (Allina et al., 1998), to suggest the presence of poplar 4CL isoforms with distinct substrate-utilization profiles. This is in contrast to a previous study by Grand et al. (1983) who reported the partial purification of three 4CL isoforms. Each isoform could use 4-coumaric acid and ferulic acid. Form I could, additionally, use 5-hydroxyferulic acid and sinapic acid as substrates, while Form III specifically activated caffeic acid. Apparently none of the isoforms could use cinnamic acid as a substrate. In contrast to Grand et al. (1983) no evidence was found for the conversion of sinapic acid to sinapoyl-CoA via the action of the recombinant 4CL proteins or via the action of any of the native isoforms (Allina et al., 1988). Thus, the recombinant and native poplar 4CL activity seen in our laboratory is more like that reported by Kutsuki et al. (1982) and Meng and Campbell (1997) than the activity reported

by Grand et al. (1983). The accumulating evidence suggests the results of Grand et al. (1983) are anomalous.

Recombinant 4CL proteins have now been produced from Arabidopsis, aspen, and tobacco (Ehlting et al., 1998; Hu et al., 1998; Lee and Douglas, 1996). All were expressed in an *E. coli* system. The recombinant poplar 4CL proteins use hydroxycinnamic acids in a manner similar to a recombinant Arabidopsis 4CL protein (At4CL3), which prefers 4-coumaric acid, followed by ferulic acid, caffeic acid, and cinnamic acid. At4CL1 is most active against caffeic acid. At4CL2 is also most active with caffeic acid, but does not utilize ferulic acid. The recombinant Arabidopsis proteins use cinnamic acid with a very low efficiency, with At4CL1 and At4CL2 catalyzing the reaction at a slower rate than At4CL3. None of the recombinant Arabidopsis 4CL proteins use sinapic acid as a substrate (Ehlting et al., 1998).

In contrast to 4CL-216 and 4CL-9, recombinant proteins expressed from two divergent *4CL* cDNA clones from aspen show some difference in substrate usage. Most notably, recombinant Pt4CL1 showed activity with 5-hydroxyferulic acid while Pt4CL2 did not. In view of the fact that Pt4CL1 is highly expressed in developing xylem, the authors suggested that Pt4CL1 might be directly involved in the biosynthesis of lignin, whereas Pt4CL2 could be responsible for the production of other phenylpropanoid derivatives (Hu et al., 1998). This analysis was based on the specific activity of histidine-tag purified 4CL proteins but, unfortunately, kinetic data was not presented to further support their suggestion of a divergence in substrate preferences.

The poplar recombinant proteins show the greatest similarity in substrate usage to the two recombinant tobacco 4CL proteins. Like recombinant poplar 4CL proteins, both tobacco 4CL proteins prefer 4-coumaric acid, followed by ferulic acid, caffeic acid and cinnamic acid, with no detectable activity towards sinapic acid (Lee and Douglas, 1996).

This comparison of substrate usage between recombinant 4CL proteins is somewhat surprising. Both recombinant 4CL poplar proteins are most similar in primary amino acid sequence, to tobacco sequences (see Table 2.1), and show the most similarity in substrate utilization patterns. However, when compared to the Arabidopsis 4CL sequences, the poplar 4CL recombinant proteins are most similar to At4CL2 but least similar in substrate usage to this form. This suggests that overall primary protein structure may not be a good predictor of substrate utilization patterns. Hu et al. (1998) suggest that high divergence in the first 180 aa of the N-terminal sequence and a shorter amino acid sequence (by 24 aa) is sufficient to allow Pt4CL1 to use 5-hydroxyferulic acid while Pt4CL2 cannot. They suggest these amino acids might be involved in substrate binding specificity, but this remains to be tested.

Kinetic analysis of the recombinant poplar proteins indicates the relative abilities of the enzymes to use differently substituted hydroxycinnamic were similar to that of native forms from other species. Reported K_m values of purified or partially purified 4CL proteins (Knobloch and Hahlbrock, 1975, 1977; Grand et al., 1983; Voo et al., 1995) range from 6.8 to 32 μM for 4-coumaric acid and from 9.1 to 130 μM for ferulic acid. Thus, the affinities of 4CL-216 for these substrates (about 11 and 13 μM, respectively) were well within this range, while the affinities of 4CL-9 for these substrates (4-coumaric acid, 80 μM, ferulic acid, 130 μM) were, on average, several times lower than those reported for these native enzymes.

The data presented here and elsewhere (Allina et al., 1998; Wallis and Rhodes, 1977; Lozoya et al., 1988; Voo et al., 1995; Lee and Douglas, 1996; Hu et al., 1998) demonstrate the absence of 4CL activity with sinapic acid as a substrate in many plants. Based on the kinetic analysis and the substrate utilization profiles of 4CL-216, 4CL-9 and native poplar 4CL isoforms, H11 poplar 4CL enzymes appear to be typical of those described for many plants (Wallis and Rhodes, 1977; Lozoya et al., 1988; Voo et al., 1995; Lee and Douglas, 1996) with highest activity against 4-coumaric acid, followed by ferulic acid and caffeic acid, and no activity against sinapic acid. Angiosperm lignin, including that of poplar, is composed of both guaiacyl and syringyl subunits based on coniferyl and sinapyl alcohol subunits, respectively. Thus it is unlikely that differential expression of *4CL* gene family members encoding enzymes with different substrate utilization profiles is a mechanism used in poplar (with the possible exception of aspen Pt4CL1 and Pt4CL2), or most other plants, to partition carbon into guaiacyl and syringyl lignin. The mechanisms by which 4CL participates in the regulation of phenylpropanoid carbon flow into differently methylated lignin precursors may need to be re-examined.

Because 4CL is apparently inactive against sinapic acid, the biosynthetic route to sinapyl alcohol, conventionally thought to be derived from sinapoyl-CoA via the actions of cinnamoyl-CoA reductase and cinnamyl alcohol dehydrogenase, is unclear. In the literature, evidence has been presented that suggests that at least two different biosynthetic routes might exist.

Evidence exists in several plants to suggest that an alternative pathway may exist that uses the enzyme CCoAOMT. CCoAOMT but not COMT (see Figure 1.3) is induced in

Zinnia cell cultures that are undergoing tracheary element development (Ye at al., 1994). Aspen Pt4CL1 utilizes 5-hydroxyferulic acid but not sinapic acid, and is located in lignin forming tissues (Hu et al., 1998). Taken together, these results suggest that sinapyl alcohol monomers could be derived from an alternative pathway involving hydroxylation and methoxylation (probably by CCoAOMT) of hydroxycinammoyl CoA thioesters, most likely feruloyl-CoA or 5-hydroxyferuloyl-CoA.

Arabidopsis mutants lacking ferulate-5-hydroxylase (F5H) (see Figure 1.3) activity produced no sinapic acid, sinapyl-malate or syringyl lignin (Chapple et al., 1992). Overexpression of this same gene results in a decrease in guaiacyl lignin with a concomitant increase in syringyl lignin (reviewed in Whetten et al., 1998). This suggests that sinapyl alcohol monomers are derived from either 5-hydroxyferulic acid or sinapic acid.

However, Lee and Douglas (1996) demonstrated that Arabidopsis plants with reduced 4CL activity show an overall decrease in lignin content, with a much greater decrease in ferulic acid than sinapic acid-derived monomers. This suggests that sinapic acid is not a substrate for 4CL. However, if CCoAOMT were responsible for converting 5-hydroxyferuloyl-CoA to sinapoyl-CoA (as suggested above) then there would also be a decrease in the amount of syringyl lignin in 4CL antisense plants. This suggests that sinapoyl-CoA moieties may well be the precursors for syringyl lignin but that they may be formed by the action of a novel enzymatic reaction. Finally, in transgenic poplars with severely reduced COMT levels the syringyl/guaiacyl lignin ratio was reduced six-fold (decreased syringyl and increased guaiacyl). Also, a new lignin component, 5-hydroxy-guaiacyl residue was detected (Van Doorsselaere et al., 1995a). This suggests that

CCoAOMT does not function in *O*-methylation of 5-hydroxyferuloyl-CoA, in poplar, in vivo. It also suggests that while sinapoyl-CoA may be the precursor of sinapyl alcohol monomers (possibly formed by the action of a novel enzyme), sinapic acid is not formed via the action of what was thought to be a bispecific enzyme (COMT) but by a distinct 5-hydroxyferulic acid OMT.

Obviously, the biosynthetic route to sinapyl alcohol monomers is under debate. The transgenic plant work suggests that enzymatic functions elucidated in vitro may not function in a similar manner in vivo. Also, different plants may use different routes. The best way to solve this puzzle would be to isolate mutants for each of the steps and then cross them to make double and even triple mutants. Because most, if not all, of the genes are members of gene families, this will probably prove to be very difficult to do.

The substrate preferences of 4CL from compression wood of loblolly pine are different from those of the enzyme from other tissues. It was suggested that overproduction of cinnamic acid during compressional stress induces a cellular factor which facilitates this change in activity (Zhang and Chiang, 1997). If this hypothesis is correct, perhaps, 4CL activity towards sinapic acid could by analogy be modulated by cellular factors. Thus 4CL may only use sinapic acid as a substrate in certain environmental conditions or during specific developmental processes. Purification of 4CL proteins would inevitably lead to the loss of the putative cellular factor, such that it no longer has activity towards sinapic acid. In addition, 4CL activity towards sinapic acid in crude extracts may only exist if the starting material is collected under appropriate conditions.

It was surprising to observe that the 4CL recombinant proteins were able to use cinnamic acid as a substrate while the H11 native isoforms did not (Allina et al., 1998). 4CL-216 and 4CL-9 readily converted cinnamic acid to cinnamoyl-CoA at efficiencies of about 20% of that of 4-coumaric acid (at 0.2 mM), while partially purified native 4CL isoforms from poplar are unable to carry out this reaction (Allina et al., 1998). A similar phenomenon described by Lee and Douglas (1996) showed that recombinant tobacco proteins expressed in *E. coli* convert cinnamic acid to cinnamoyl-CoA, while such activity was lacking in crude protein extracts of tobacco stem. While recombinant Arabidopsis 4CL proteins use cinnamic acid as a substrate (Ehlting et al., 1998) crude protein extracts from Arabidopsis bolting stems show no activity towards cinnamic acid (Lee et al., 1997). Kutsuki et al. (1982) and Meng and Campbell (1997) did not report on the possible use of cinnamic acid by poplar and aspen 4CL proteins, and Grand et al. (1983) reported an absence of activity toward cinnamic acid by all poplar 4CL isoforms. Unfortunately, there was no mention of cinnamic acid usage in the study by Hu et al. (1998).

Cinnamic acid is not normally thought of as a substrate for 4CL even though some of the earliest forms of 4CL identified did have activity, albeit with relatively low efficiency, towards cinnamic acid. Native 4CL forms showing activity towards convert cinnamic acid include *Forsythia* (Gross and Zenk, 1974), soybean (Knobloch and Hahlbrock, 1974), forms I and II from petunia (Ranjeva et al., 1976), parsley (Knobloch and Hahlbrock, 1977), pea, form I (Wallis and Rhodes, 1977), spruce (Lüderitz et al., 1982), and loblolly pine (Voo et al., 1995). The apparent affinities for cinnamic acid reported native enzymes range from 80 to 2850 μM and the affinities of recombinant 4CL-216 and 4CL-9 fall within this range (1141

and 1048 μM, respectively). The affinity of 4CL-9 for cinnamic acid was about 13-fold lower than the value for 4-coumaric acid (relative K_m; Table 3.2). In contrast, the affinity of 4CL-216 for cinnamic acid was about 104-fold lower than the value for 4-coumaric acid (relative K_m; Table 3.2).

If some of the 4CL enzymes from some plants do use cinnamic acid as a substrate, what might the end products be? Unsubstituted B-ring-deoxy flavonoids, derived from cinnamoyl-CoA appear to function as phytoalexins. A single 4CL isoform, partially-purified from elicited cell cultures of "old man" cactus, has roughly equal activity towards cinnamic acid and 4-coumaric acid. These cell cultures accumulate high levels of unsubstituted B-ring deoxy flavonoids. (Liu et al., 1995). Pinocembrin, a 5,7-dihyroxyflavone derived from cinnamoyl-CoA, has antifungal properties and is a major constituent of leaf and bud resins of *P. deltoides* (Shain and Miller, 1982). Salicylic acid, an endogenous regulator of disease resistance (Lee et al., 1995b) may be derived from cinnamic acid by one of two pathways, one of which would require cinnamic acid to be first activated to its CoA-thioester.

Native 4CL activity toward cinnamic acid has not been detected in maize (Hipskind et al., 1993), in form II from pea (Wallis and Rhodes, 1977) or form III from petunia (Ranjeva et al., 1976), in poplar (Grand et al., 1983; Allina et al., 1998), or in wheat (Maule and Ride, 1983). It is obvious that 4CL proteins from some plants use cinnamic acid as a substrate while others do not. Thus, cinnamic acid usage by some 4CL proteins and not others may be due to the requirement for specific end products by each plant. Lee et al. (1996) hypothesized that post-translational modifications may change the ability of native tobacco 4CL to use cinnamic acid. It might then be expected that recombinant 4CL proteins, expressed in a

heterologous cellular environment where these modifications cannot take place would be able to use cinnamic acid as a substrate. The apparent discrepancy in cinnamic acid usage by native and recombinant 4CL poplar proteins is the subject of the next chapter.

CHAPTER FOUR

THIOESTERASE

4.1 INTRODUCTION

In the previous chapter (3), a discrepancy between the ability of recombinant 4CL-216 and 4CL-9 to utilize cinnamic acid as a substrate, and the inability of partially purified poplar native 4CL isoforms to do so (Allina et al., 1998) was noted. This was interesting because it suggested that, given the proper conditions, native 4CL could use cinnamic acid to produce cinnamoyl-CoA. Some important phenylpropanoid metabolites are thought to be derived from cinnamoyl-CoA, such as B-ring-deoxy flavonoids and possibly, salicylic acid. It was decided that this phenomenon warranted further study.

Lee and Douglas (1996) noted a similar discrepancy in the usage of cinnamic acid by native and recombinant 4CL proteins from tobacco. They demonstrated that addition of tobacco protein extracts to the tobacco recombinant protein 4CL assays modified the usage of cinnamic acid as a substrate of 4CL. Their results suggested that the modifying agent was probably a protein (it was heat-labile and of high molecular weight) and hypothesized that 4CL activity was post-translationally modified by this protein. This could occur by a covalent modification, such as phosphorylation, or by non-covalent interaction with 4CL, for example as part of a multienzyme complex.

The purpose of this work was to test the hypothesis that 4CL activity is modified by a protein in the xylem extract. The work in this chapter will be presented chronologically. It was decided to partially purify the modifying protein as a first step in determining the nature

of its action on the 4CL enzyme. In the course of this work, it became apparent that an alternative explanation for the effect of xylem extract on apparent 4CL activity was more likely. According to this hypothesis, the modifying activity did not alter 4CL activity itself, but instead modified the product of the reaction, namely hydroxycinnamoyl-CoA thioester moieties, and thus affected the values generated by the spectrophotometric 4CL assay. The modifying activity was more fully characterized and shown to be a thioesterase which converts hydroxycinnamoyl-CoA thioesters back to free hydroxycinnamic acids.

4.2 METHODS

4.2.3 Poplar protein extracts

Poplar xylem tissue, collected as described (Chapter 2, 2.2.5), was homogenized in liquid N_2. The frozen powder was mixed with 100 mM Pipes, pH 6.0, 14 mM β-mercaptoethanol and Dowex AG 1-X2 Resin (0.1 g of resin per gram of tissue in 2 ml buffer; Bio-Rad). This mix was rotated at 4°C for 20 minutes, filtered through one thickness of Miracloth (Calbiochem), and centrifuged to remove resin and debris. The thioesterase activity was precipitated from the supernatant with solid $(NH_4)_2SO_4$ (60-80% saturation) and dissolved in a minimal volume of buffer. The sample was desalted by chromatography through a Sephadex G-25 (Pharmacia) column.

4.2.2 4CL/xylem protein extract assay

Known amounts of crude or partially purified xylem protein extracts were added to recombinant 4CL and a standard 4CL enzyme assay was performed (Chapter 3, 3.1.5). Cinnamic acid (1 mM) was used as a substrate except for the experiment shown in Figure 4.2 where all hydroxycinnamic acids were used at a 0.2 mM concentration.

4.2.3 Variation of the 4CL/thioesterase assay

A standard 4CL enzyme assay was performed and the accumulation of the appropriate hydroxycinnamoyl-CoA thioester was monitored at the appropriate wavelength for 5 minutes. In some experiments the 4CL reaction was stopped by boiling the reaction mixture for 15 minutes. The reaction mixture was cooled on ice for 10 minutes and centrifuged at 15,000 xg for 15 minutes. Poplar protein extract was added and the disappearance of the CoA thioester was monitored spectrophotometrically for a further 5 minutes. Alternatively, poplar

protein extract was added to chemically synthesized cinnamoyl-CoA or 4-coumaroyl-CoA and the disappearance of the CoA ester was monitored at the appropriate wavelength.

4.2.4 Size exclusion chromatography

200 µl of the 60 - 80 % $(NH_4)_2SO_4$ protein fraction were loaded onto a Superose 12 HR 10/30 column (Pharmacia) pre-equilibrated with 100 mM Pipes, pH 6.0, 50 mM KCl, and eluted isocratically with the same buffer at 0.5 ml/min. Fractions were collected every two minutes. Molecular weight protein standards were used to calibrate the Superose 12 column, including cytochrome C (12.5 kDa), chymotrypsinogen (25 kDa), albumin (45 and 68 kDa), aldolase (158 kDa) (Boehringer). Blue Dextran 2000 (approx. 2 x 10^6 kDa) was used to estimate the void volume (V_o).

4.2.5 Preparation and analysis of [^{14}C]-cinnamic acid and [^{14}C]-cinnamoyl-CoA samples

[^{14}C]-cinnamic acid was prepared by mixing 0.5 µCi of L[U-^{14}C]phenylalanine (Amersham) and 5 µmoles of unlabeled L-phenylalanine with 500 µl of crude recombinant poplar PAL (McKegney et al., 1996) and allowing the reaction to proceed for 90 minutes at room temperature. A 100 µl aliquot of this mixture was added to 500 mM Tris and reserved as control 1 (cinnamic acid). To make [^{14}C]-cinnamoyl-CoA, 50 µl of anion-exchange purified recombinant 4CL-9, 500 mM Tris, pH 7.8, and co-factors (see 3.2.4) were added to another 200 µl aliquot of the PAL incubation mix. This mixture was incubated at room temperature for 60 minutes to allow conversion of the [^{14}C]-cinnamic acid to [^{14}C]-cinnamoyl-CoA. The reaction mix and control 1 were then each boiled for 15 minutes to inactivate the recombinant enzymes. The reaction mix was separated into a further two aliquots (control 2 and experimental). Xylem protein extract (60 - 80 % ammonium sulphate

fraction) (500 μg) was added to the experimental reaction mix and incubated for 20 minutes, while the second aliquot (^{14}C-cinnamoyl-CoA control) was left untreated. The three aliquots were then subjected to acid hydrolysis (2 M HCl) for 60 minutes and then chilled on ice. Each of the aliquots (^{14}C-cinnamic acid control, ^{14}C-cinnamoyl-CoA control and xylem treated ^{14}C cinnamoyl-CoA) was extracted three times with diethyl ether and the organic phase was dried over anhydrous $MgSO_4$. The supernatant was decanted, the ether was allowed to evaporate and the resulting residue was dissolved in 25 μl 95% ethanol.

The ethanol solutions were loaded onto Polygram^{R} SIL G TLC plates (Macherey-Nagel, GmbH & Co.) and run in three different solvent systems: toluene:acetic acid (4:1), chloroform:methanol (98:2) and ethylacetate:toluene (8:12). Radioactive bands were visualized by overnight exposure with Kodak Diagnostic Film (X-OMAT^{TM}AR).

4.2.6 Preparation and analysis of enzymatically prepared hydroxycinnamoyl-CoA samples

Cinnamoyl-CoA and other hydroxycinnamoyl-CoA samples for HPLC analysis were prepared in a similar manner to those used for TLC analysis, except that each reaction started from the appropriate unlabeled acid. Briefly, 0.2 mM solutions of various hydroxycinnamic acids (1 mM for cinnamic acid) were converted to the CoA thioester by incubating for 1 hour with recombinant 4CL-9, ATP and CoA. The reaction mixtures were boiled for 15 minutes to denature the recombinant 4CL-9. The reaction mixtures were separated into 2 aliquots (control and experimental) and the experimental aliquot was treated with 500 μg of poplar protein extract (60 - 80 % ammonium sulphate fraction). Each aliquot was passed through a 0.22 μm filter before HPLC analysis.

20μl of each aliquot was applied to a Bondapak reverse-phase C18 column and subjected to gradient elution. Buffer A consisted of 1% phosphoric acid in water and buffer B was 1% phosphoric acid in acetonitrile. Flow rate was 1 ml/min. The reaction products were eluted using a non-linear gradient of 100% A for 5 minutes, 50% A for 2 minutes, 20 % A for 3 minutes, 100 % for 10 minutes. The eluate was monitored at 270 nm and each peak detected at that wavelength was scanned at wavelengths from 200 - 600 nm. The resulting data were then analyzed using Millenium software (Waters). Data for cinnamoyl-CoA are reported at 311 nm and for 4-coumaroyl-CoA at 333 nm.

4.3 RESULTS

4.3.1 Cinnamic acid usage is inhibited with the addition of poplar xylem protein extracts

To determine whether a xylem protein was able to alter the ability of recombinant 4CL proteins to use cinnamic acid as a substrate, an experiment analogous to that of Lee and Douglas (1996) was performed. All work in this chapter was carried out using recombinant 4CL-9. Figure 4.1 shows the 4CL activity of recombinant 4CL-9, of crude xylem protein extracts and of the two mixed together, using different hydroxycinnamic acid substrates and following the reactions spectrophotometrically. The control 4CL assays use recombinant 4CL-9 or crude xylem protein extracts in standard 4CL assays. To determine whether crude xylem protein extract could modify 4CL activity towards a specific hydroxycinnamic acid the xylem extract was incubated with 4CL-9 at the beginning of the assay, a standard 4CL assay was performed, and the 4CL activity was compared to that in the control assay. It was observed that crude xylem protein extract alone had no activity towards cinnamic acid and that adding it to recombinant 4CL-9 greatly reduced the activity of the recombinant 4CL towards cinnamic acid. This suggested that less cinnamoyl-CoA was produced when the xylem protein extract was added to recombinant 4CL-9. The same experiment was performed using 4-coumaric acid, caffeic acid, ferulic acid and sinapic acid as substrates. Both the recombinant 4CL-9 protein and the crude xylem protein extract displayed 4CL activity towards 4-coumaric acid, caffeic acid and ferulic acid, but not to sinapic acid. Furthermore, when the recombinant 4CL-9 protein and crude xylem protein extracts were added together, the resulting activity was approximately additive. The results of this experiment suggested

that modification of 4CL activity, via addition of crude xylem protein extract, was specific to cinnamic acid.

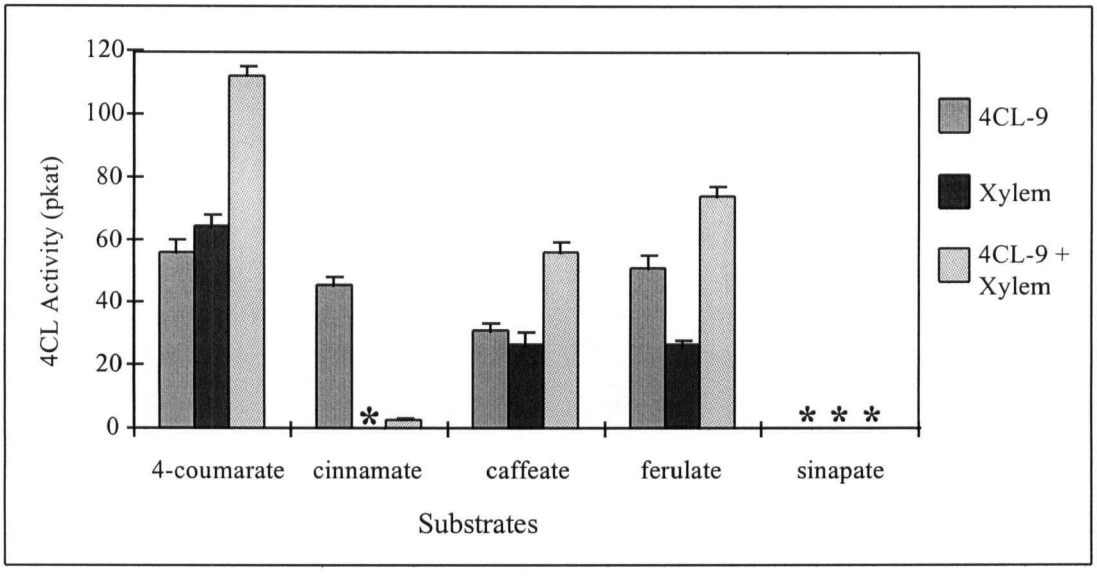


Figure 4.1 The effect of adding xylem protein extract to recombinant 4CL-9 and assaying with different hydroxycinnamic acids. Results are the average of three trial ± SD. Asterisks represents no activity detected.

The modifying activity in the crude xylem protein extract was shown to likely be a protein because the modifying activity was found in the high molecular mass eluate after de-salting the extract through a G-25 Sephadex column (approximately 5 kDa exclusion limit), and the activity was lost after boiling the extract for 10 minutes (not shown). Moreover, the modifying activity appeared to be dose-dependent (Figure 4.2). Increasing amounts of the crude xylem protein extract further decreased the activity of recombinant 4CL-9 with cinnamic acid as a substrate.

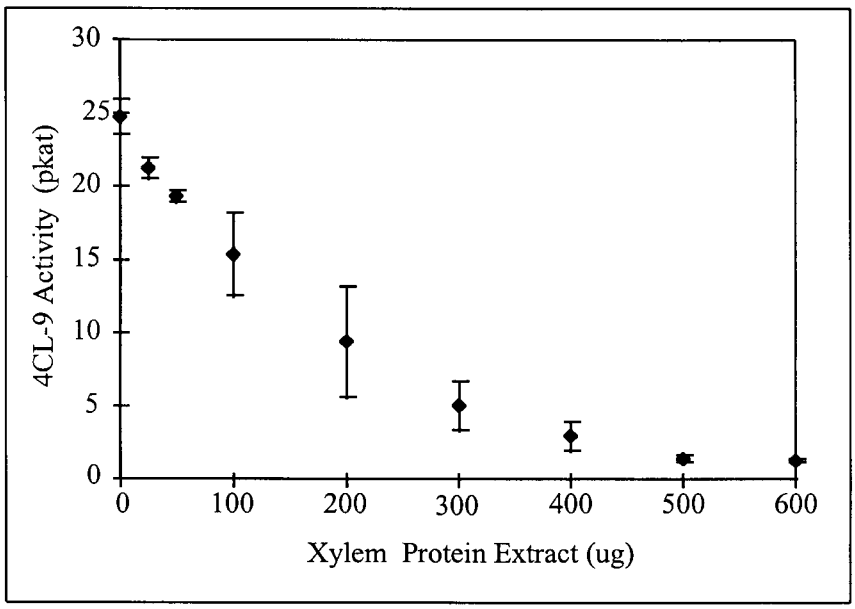


Figure 4.2 The effect of adding increasing amounts of crude poplar xylem protein extract to recombinant 4CL-9. Increasing amounts of crude xylem protein extract was added to recombinant 4CL-9 and 4CL activity was then measured. Results are the average of three trials ± SD.

4.3.2 Partial purification of modifying activity

A variety of techniques was used in an effort to separate the modifying activity from other proteins in the crude xylem extract. The assay for modifying activity was performed as in 4.3.1. Briefly, 4CL-9, a xylem protein fraction, cinnamic acid and the other co-factors were added together and the change in absorbance was measured. A control 4CL-9 reaction in the absence of xylem protein was performed in parallel to determine the amount of 4CL activity present. The difference in absorbance (representing accumulation of cinnamoyl-CoA) between the experimental and the control reactions was used as a measure of modifying activity, as shown in Figure 4.3.

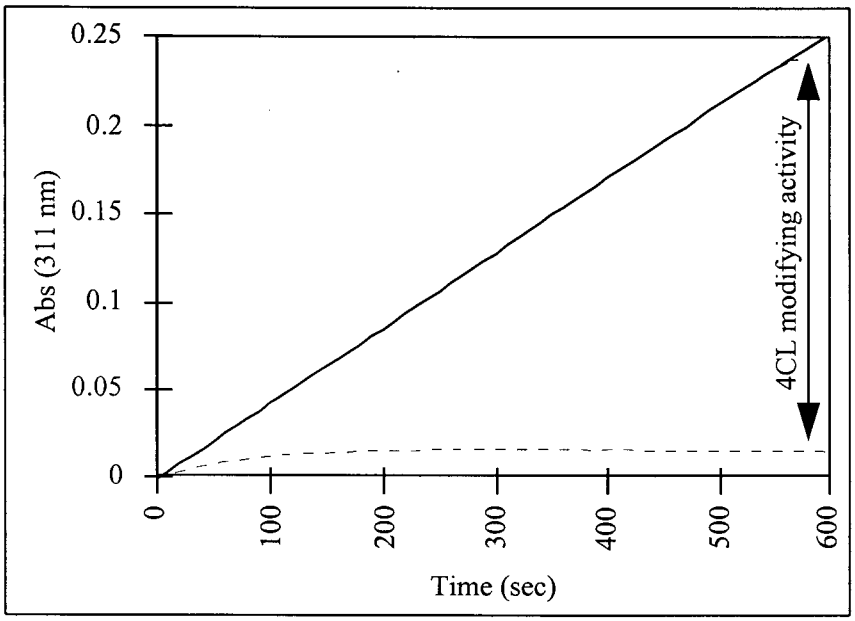


Figure 4.3 Visual representation of the change in absorbance of a control 4CL and a modified 4CL reaction using cinnamic acid as a substrate. Straight line represents control 4CL reaction. Dotted line represents modified 4CL reaction.

In crude extracts, the modifying activity was shown to be insensitive to either changes of pH between 6 and 9 or salt concentration between 50 and 200 mM but was decreased with the addition of glycerol (not shown). Therefore, 100 mM PIPES, pH 6.0, 15 mM β-mercaptoethanol was used for all extractions. As a first purification step, crude xylem protein extracts were precipitated using solid $(NH_4)_2SO_4$. The modifying activity was recovered in the 60 - 80% fraction and was thus separated from native 4CL activity which was recovered in the 40 - 60% fraction (not shown). After de-salting, the 60 - 80% fraction was subjected to size exclusion chromatography. As shown in Figure 4.4 the modifying activity eluted in a 1 ml fraction (Fraction 14) from a size exclusion column (Superose 12 HR 10/30, optimal separation of 1,000 - 3 X 10^5, Pharmacia). The recovered activity showed a 7.5-fold increase in specific activity over a crude extract.

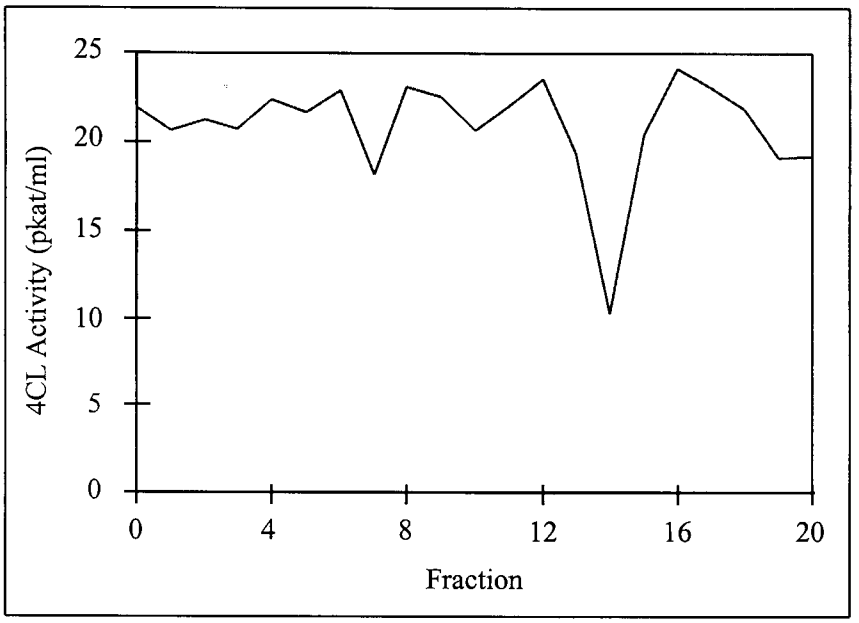


Figure 4.4 Elution profile of modifying activity after size exclusion chromatography. Modifying activity (apparent decrease in 4CL activity towards cinnamic acid) peak elutes at fraction 14 from a Superose-12 column. 1 mM cinnamic acid used as a substrate.

Further attempts at purifying the activity by use of a variety of separation methods were unsuccessful. The activity bound tightly to a cation matrix (Sepharose S, Pharmacia), at pH values lower than 5.5 and could not be recovered whereas the activity did not bind at pH values higher than this. At all pH values above 6 the activity bound tightly to an anion exchange matrix (Sepharose Q, Pharmacia) and could not be eluted with high salt. The poplar protein extract was also fractionated on a chromatofocusing column (Mono P, Pharmacia). The loaded protein was eluted within the range of pH 7 to 4 using Pharmacia Polybuffer (a buffer that generates a pH gradient), but no activity was recovered from the column. Fractions were pooled and concentrated to test the possibility that the modifying activity required more than one protein that had been separated during the purification step, but no activity was detected in the pooled fractions. It is probable that the activity was bound tightly

to the column, suggesting the pI of the protein is higher than 7. Attempts at purifying the activity were abandoned and attention was focused on understanding the nature of the activity found in the poplar protein extract.

4.3.3 Modified 4CL assays

A series of modified 4CL assays was performed to further investigate the effect of the putative modifying activity on the reaction catalyzed by recombinant 4CL. As previously mentioned in 4.3.1, the standard (control) 4CL assay was carried out by initially mixing recombinant 4CL-9, cinnamic acid and the other required co-factors (see 3.1.5), and assaying the accumulation of cinnamoyl-CoA on the basis of an increase in absorbance of the reaction mix at 311 nm over time (Figure 4.5A). The method for assaying 4CL-modifying activity involved supplementing the initial reaction mixture with xylem protein extract at the beginning of the assay and similarly measuring the increase in absorbance at 311 nm over time (Figure 4.5B). Comparison of these absorbance values relative to those measured in the control assay provided a measure of the putative modifying activity. It was observed that the addition of xylem protein extract greatly reduced the change in absorbance, which suggested that less cinnamoyl-CoA was produced when the xylem protein extract was added to the reaction mix. However, close inspection of the kinetics showed there was an initial linear accumulation of cinnamoyl-CoA which reached a plateau after 120 seconds (Figure 4.5B). It should be noted that the accumulation of cinnamoyl-CoA in the reaction containing xylem protein extract was less than one quarter of the level attained in the control reaction in the first minute. I decided to further investigate this phenomenon using a modified 4CL/xylem reaction.

In the modified assay, a standard 4CL assay was first performed during which cinnamic acid was converted to cinnamoyl-CoA over a period of 5 minutes. The accumulation of cinnamoyl-CoA was recorded over this time by measuring the absorbance of the reaction mix at 311 nm. A small amount (12.5 μg protein) of xylem protein extract was then added to the reaction mix and the absorbance at 311 nm was followed for another 5 minutes. This caused an immediate lowering of the rate of increase in absorbance at 311 nm (Figure 4.5C), which continued at this new rate for at least 5 minutes. This modified assay was repeated with increasing amounts of poplar protein extract (25 - 100 μg protein), each of which led to more marked decreases in absorbance at 311 nm over time. These effects on the kinetics of the absorption changes were directly related to the amount of poplar protein extract added to the reaction mixes (Figure 4.5D-F). In control assays, the accumulation of cinnamoyl-CoA was linear over a period of 10 minutes in the absence of xylem extract (Figure 4.5A).

These results suggested that, rather than altering 4CL enzyme activity, the "modifying activity" observed in the xylem protein extract was actually using cinnamoyl-CoA as a substrate, converting it to a compound with lower absorbance at 311 nm. The simplest explanation would be that the previously hypothesized "modifying activity" is a thioesterase capable of converting cinnamoyl-CoA back to free cinnamic acid. Alternatively, the activity could be converting the cinnamoyl-CoA to another compound via an unknown mechanism.

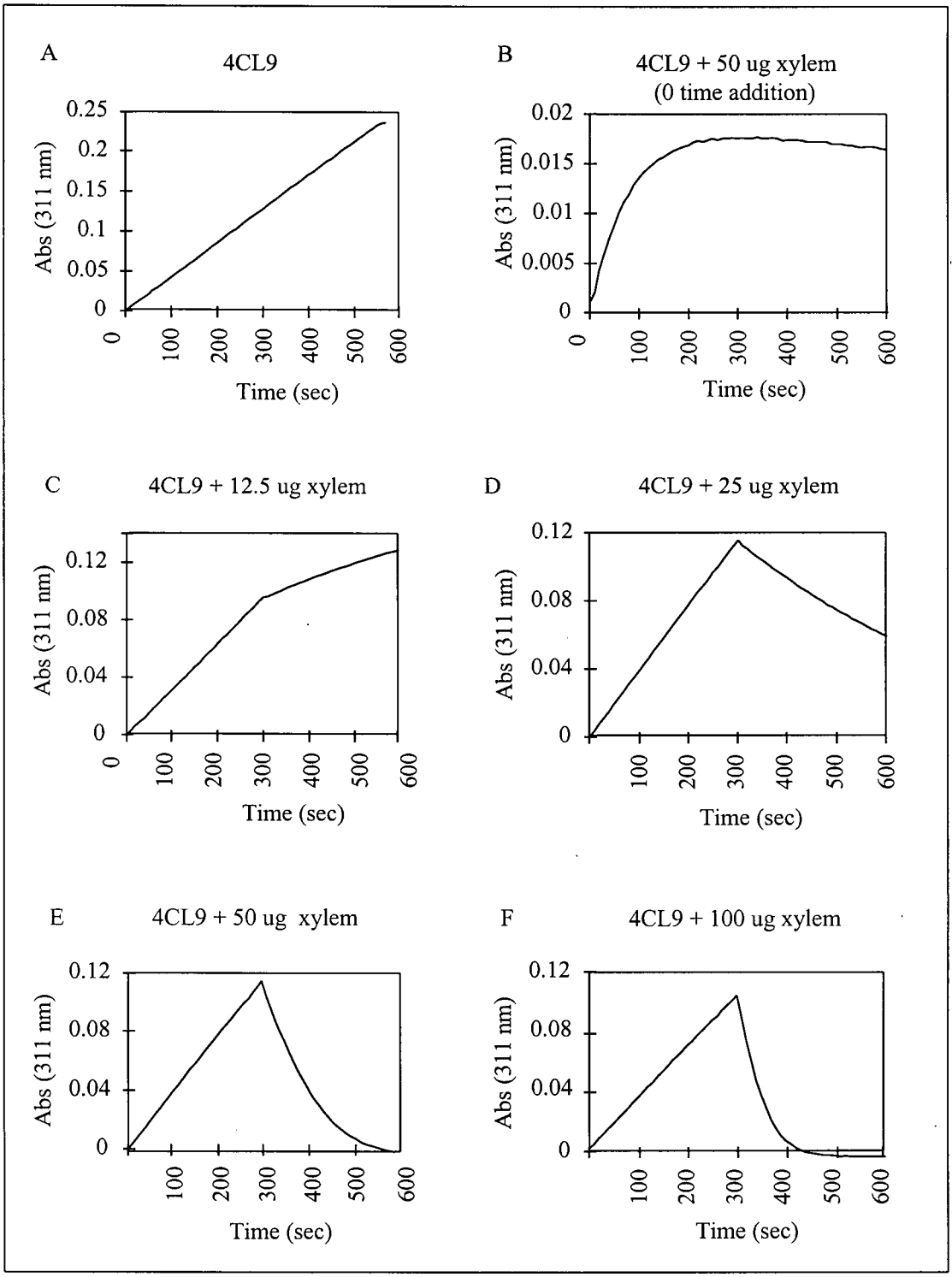


Figure 4.5 The effect of adding modifying activity after the accumulation of cinnamoyl-CoA. A, 4CL-9 control assay; B, 4CL-9 and xylem protein extract added together to the reaction mix; C - F, increasing amounts of xylem protein extract added after accumulation of cinnamoyl-CoA (at 300 seconds) in the reaction vessel.

4.3.4 TLC analysis of [^{14}C]-cinnamic acid, [^{14}C]-cinnamoyl-CoA and xylem treated [^{14}C]-cinnamoyl-CoA

To distinguish between the two alternative explanations for the cinnamoyl-CoA consuming activity (a thioesterase activity or the conversion of cinnamoyl-CoA to another product), TLC was used to analyze the products accumulating in the modified 4CL enzyme reaction, upon addition of the xylem protein extract. In order to synthesize [^{14}C]cinnamoyl-CoA to use as a tracer substrate for these experiments, L[U-^{14}C]phenylalanine was converted enzymatically to [^{14}C]cinnamoyl-CoA via [^{14}C]cinnamic acid, using recombinant poplar PAL (McKegney et al., 1996) and recombinant 4CL-9. [^{14}C]cinnamoyl-CoA prepared in this manner was treated with the poplar protein extract or left untreated as a control. The control [^{14}C]cinnamoyl-CoA and the xylem treated [^{14}C]cinnamoyl-CoA preparations were then acid hydrolysed to facilitate partitioning into ether. The ether soluble products were fractionated using TLC in three different solvents and the developed plates were subjected to autoradiography (Figure 4.6; not shown). Only one band, corresponding to cinnamic acid, was observed in each lane. Thus, no other ether-extractable compounds were present after treatment of cinnamoyl-CoA with the xylem protein extract. This result suggests that the modifying activity may, in fact, be a thioesterase capable of hydrolyzing cinnamoyl-CoA to generate free cinnamic acid.

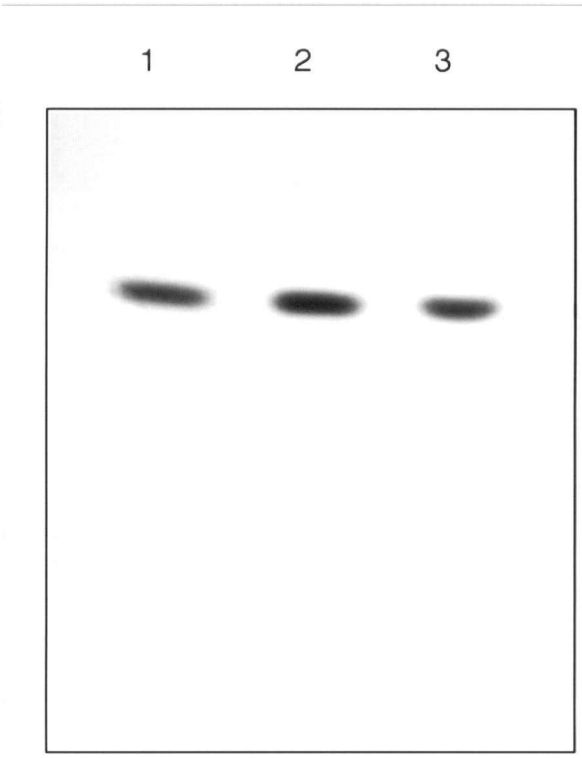


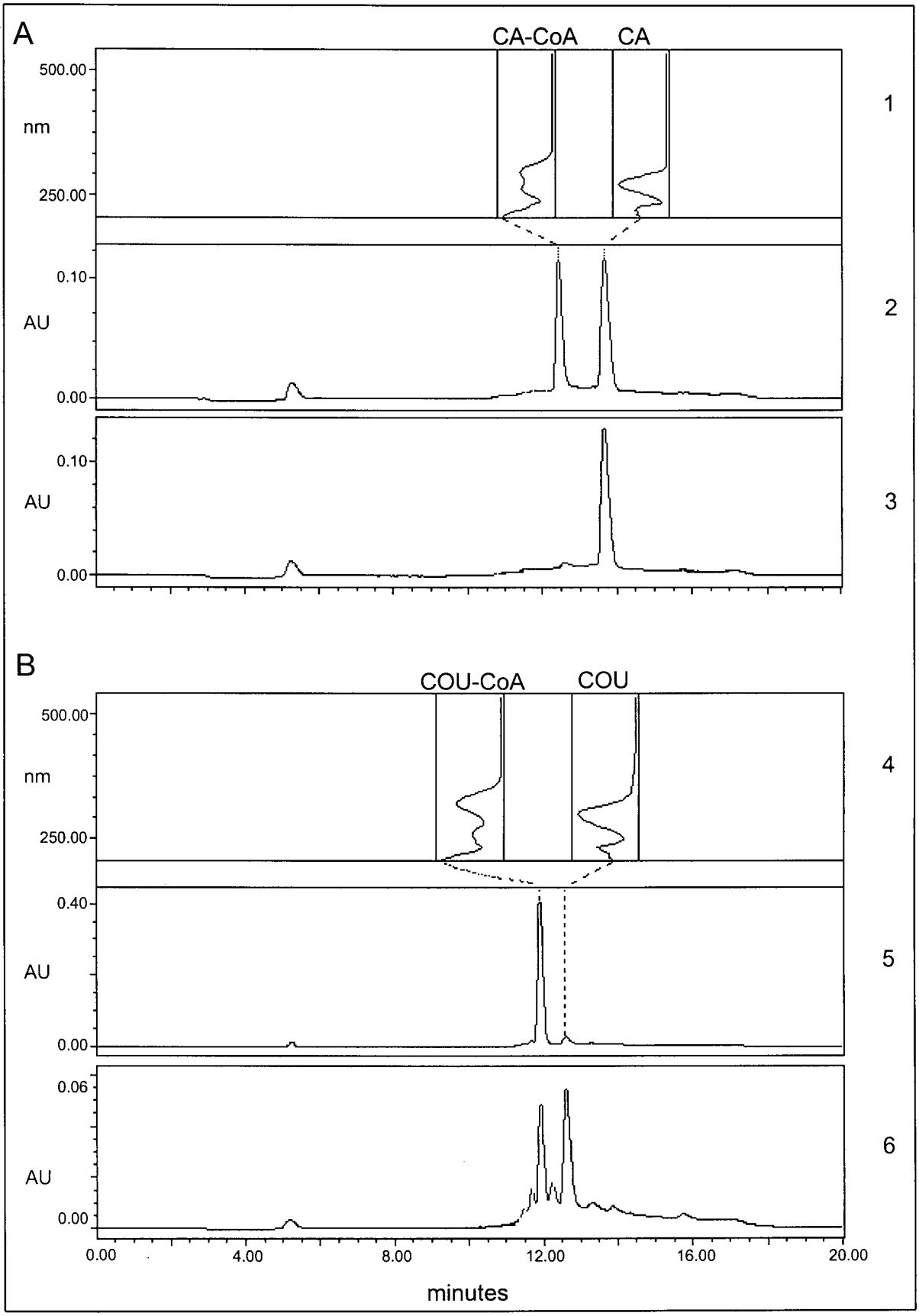
Figure 4.6 Thin-layer chromatography of [^{14}C] products. Lane 1, control cinnamic acid; lane 2, acid hydrolysed cinnamoyl-CoA; lane 3, cinnamoyl-CoA treated with xylem protein extract and further acid hydrolysed.

4.3.5 HPLC analysis of hydroxycinnamoyl-CoA thioesters and xylem treated hydroxycinnamoyl-CoA thioesters

To test the hypothesis that cinnamoyl-CoA and possibly other hydroxycinnamoyl-CoA thioesters were being hydrolyzed by a putative thioesterase present in xylem protein extracts, HPLC was employed to separate the thioesters from their corresponding free acids. Cinnamoyl-CoA and other hydroxycinnamoyl-CoA thioesters were synthesized enzymatically using recombinant 4CL-9 and then treated with the xylem protein extract. Figure 4.7A2 shows the chromatogram for the enzymatically synthesized cinnamoyl-CoA, reported at 311 nm. The retention time for cinnamoyl-CoA is 12.43 minutes and that of the remaining, unconverted, cinnamic acid is 13.65 minutes. Figure 4.7A3 shows the chromatogram of cinnamoyl-CoA (retention time, 12.43 minutes) treated with the xylem protein extract. Figure 4.7A1 presents the UV spectra for the peaks of cinnamoyl-CoA and cinnamic acid. These results show that, after treatment with the xylem protein extract, the cinnamoyl-CoA peak is virtually eliminated. Figure 4.7B shows the results for a similar analysis of enzymatically synthesized 4-coumaroyl-CoA, reported at 333 nm. The retention time for 4-coumaroyl-CoA is 11.91 minutes and for the unconverted 4-coumaric acid 12.6 min (Figure 4.7B5). Figure 4.7B6 shows the chromatogram obtained from analysis of xylem protein extract-treated 4-coumaroyl-CoA. The peak of 4-coumaroyl-CoA is markedly diminished, although the hydrolysis appears to be less complete that was observed with cinnamoyl-CoA. Two additional minor peaks become evident after treatment with the xylem protein extract. The retention times for these peaks are 11.69 min and 12.24 min. Figure 4.7B4 presents the UV spectra for the peaks of 4-coumaroyl-CoA and 4-coumaric acid. Similar results were obtained for caffeic acid/caffeoyl-CoA and ferulic acid/feruloyl-CoA.

Since sinapic acid was not converted to sinapoyl-CoA by recombinant 4CL, it was not possible to examine the ability of the thioesterase to hydrolyze the trisubstituted CoA ester.

Figure 4.7 HPLC analysis of cinnamoyl-CoA and 4-coumaroyl-CoA. (A) CA-CoA, cinammoyl-CoA; CA, cinnamic acid. 1, UV spectra of major peaks; 2, enzymatically synthesized cinnamoyl-CoA control; 3, cinnamoyl-CoA treated with xylem protein extract. The results of the run are reported at 311 nm. (B) COU-CoA, 4-coumaroyl-CoA; COU, 4-coumaric acid. 4, UV spectra of major peaks; 5, 4-coumaroyl-CoA control; 6, 4-coumaroyl-CoA treated with xylem protein extract. The results of the run are reported at 333 nm. Retention times are shown in minutes.


A
CA-CoA
CA
500.00
nm
250.00
1
0.10
AU
0.00
2
0.10
AU
0.00
3
B
COU-CoA
COU
500.00
nm
250.00
4
0.40
AU
0.00
5
0.06
AU
0.00
6
0.00
4.00
8.00
12.00
16.00
20.00
minutes

4.4 DISCUSSION

The results in this chapter strongly suggest that what was thought to be a cellular factor that modified 4CL activity towards cinnamic acid is, in fact, a CoA-thioesterase that hydrolyzes all hydroxycinnamoyl-CoA thioesters tested, converting them back to free hydroxycinnamic acids. This was established in the following ways. First, using a spectrophotometric assay, it was demonstrated that the activity in xylem protein extracts converted cinnamoyl-CoA (Figure 4.5) to a compound with low absorption at 311 nm. Second, using TLC, it was shown that after treatment with the xylem protein extract followed by acid hydrolysis no compound other than cinnamic acid was produced. Third, using HPLC it was shown that the peak of cinnamoyl-CoA was eliminated after treatment with the xylem protein extract and no new UV-absorbing product was obtained.

Originally, it was thought that the modifier/thioesterase activity was specific to cinnamic acid/cinnamoyl-CoA because, as shown in Figure 4.1, there was no decrease in 4CL activity towards the other hydroxycinnamic substrates when xylem protein extract was co-incubated with 4CL-9. However, subsequent experiments using HPLC analysis showed that all hydroxycinnamoyl-CoA thioesters tested were converted to the corresponding free hydroxycinnamic acids by the putative thioesterase, although probably at different rates.

The HPLC results were confirmed using the spectrophotometric assay. After enzymatic conversion of 4-coumaric acid, caffeic acid and ferulic acid to their corresponding hydroxycinnamoyl-CoA thioesters the reactions were boiled to denature the 4CL-9 protein and then xylem protein extract was added. This resulted in a decrease in absorption at the specific wavelength (data not shown). This suggests that in the previous spectrophotometric

assays 4CL-9 was converted the free acids (except for cinnamic acid) to their hydroxycinnamoyl-CoA thioesters faster than the putative thioesterase could convert them back to free acids. However, when the 4CL-9 protein was denatured, the true nature of the xylem factor (the putative thioesterase) was revealed. That is, it is able to use all tested hydroxycinnamoyl-CoA thioesters as substrates.

In the experiment shown in Figure 4.7, a mixture of free acids and CoA esters was used as substrate for the putative thioesterase. To confirm that the cellular factor is a thioesterase one could treat pure hydroxycinnamoyl-CoA thioesters with the xylem protein extract and run the samples on HPLC. If, indeed, it is a thioesterase, a decrease in the amount of hydroxycinnamoyl-CoA and a corresponding accumulation of the free hydroxycinnamic acid would be expected. It is unknown if the preferred substrates of the putative thioesterase are hydroxycinnamoyl-CoA thioesters or other thioesters such as fatty acyl-CoA derivatives. Comparative kinetic analysis of the hydrolysis of hydroxycinnamoyl-CoA thioesters and acetyl-CoA thioesters would make it possible to determine if hydroxycinnamoyl-CoA thioesters are the preferred substrates. It would be expected that the putative thioesterase would have the highest affinity for its preferred substrate. Unfortunately, kinetic data could not be collected due to the unavailability of pure hydroxycinnamoyl-CoA thioesters.

A rich literature exists dealing with thioesterases. Thioesterases play a role in chain termination during fatty acid biosynthesis. In plants, fatty acids have been reported to be found in seeds (Leonard et al., 1997, Voelker et al., 1997) embryos (Fuhrmann and Heise, 1993), developing cotyledons (Davies et al., 1991), maturing olive fruits (Sanchez et al., 1992), and surface waxes (Liu and Beittenmiller, 1995). Fatty acid synthesis is localized in

plastids, and the nascent thioesterases include a transit peptide to direct them to that location (Knutzon et al., 1992; Loader et al., 1993). Fatty acids produced by this system are not normally longer than 18 carbons in length (Lassner et al., 1996).

Some plants accumulate shorter fatty acids such as laurate (12:0) and caprate (10:0). A 12:0-ACP thioesterase was purified from *Umbellularia californica* (Davies et al., 1991). The gene encoding it was isolated and expressed in Arabidopsis and *Brassica napus* where it intercepted the fatty acid chains destined to become 16- or 18- carbon fatty acids and facilitated the subsequent accumulation of 12:0 fatty acids (Voelker et al. 1992). Similar results were obtained using a thioesterase gene from *Cuphea hookeriana* transformed into canola (Dehesh et al., 1996). Thus, it appears that a variety of thioesterases exists, each with different substrate preferences.

Thioesterases play a similar role in fatty acid biosynthesis in animals. Outside of that pattern, two examples of thioesterases are worthy of comment. Long-chain fatty acids are covalently attached to the side chain of cysteine residues of some proteins, including those of the nervous system (Bizzozero, 1997). Palmitoyl-protein thioesterases located in lysozymes act to remove the fatty acid groups from these proteins (Verkruyse and Hoffman, 1996). A defect in this palmitoyl-protein thioesterase function causes a hereditary neurological disorder (Vesa et al., 1995).

A human thioesterase (hTE), homologous to thioesterase II from *E. coli* and distinct from other known human thioesterases, interacts with Nef, a protein conserved in primate lentiviruses (HIV-1, HIV-2, SIMV). Nef facilitates virus replication and enhances the infectivity of the virus. In helper T cells, CD4, a palmitoylated protein, is the surface receptor

which enables HIV to bind to and infect these cells, often resulting in the complete destruction of their population (Roitt, 1988). There is a correlation between the phenomenon of ability of CD4 down regulation and the ability of Nef to interact with CD4. Consistent with this correlation, a Nef mutant which does not interact with hTE is unable to down-regulate CD4 (Liu et al., 1997).

Thioesterases have also been isolated from bacteria, where they are involved in chain termination during fatty acid synthesis, as in plants and animals (Jiang and Cronan Jr., 1994). Surprisingly, Cho and Cronan (1994) demonstrated that *E. coli* "protease I" which had previously been shown to hydrolyze proteins such as polynucleotide phosphorylase, oxidized bovine insulin and a model substrate N-acetyl-DL-phenylalanine-2-naphthyl ester (NAPNE) also functions as a fatty acid thioesterase. Lee et al. (1996) subsequently showed that it also had arylesterase activity, using aromatic esters as substrates. The highest specific activity was seen with p-nitrophenyl esters, with much less activity shown towards α-naphthyl acetate and butyrate, phenyl acetate, tributyrin and triacetin as substrates. Using kinetic analysis, they demonstrated the preferred substrates of the multi-functional protease/thioesterase/arylesterase are the model substrate NAPNE, and palmitoyl-CoA, and that the enzyme's affinity for the p-nitrophenyl esters was at least 100-fold lower. The authors further demonstrated that deletion of the gene had no effect on cell growth suggesting that the protein is non-essential and may function as a scavenger enzyme *in vivo*.

Two bacterial thioesterases are known that use substrates similar to the hydroxycinnamoyl-CoA thioesters. *Pseudomonas* sp. strain CBS3 converts 4-chlorobenzoate, an environmental pollutant, to 4-hydroxybenzoate via a multi-step process. 4-chlorobenzoate

is first converted to a CoA thioester and then the chlorine is removed from the ring with the addition of water. This requires the action of a dual-function 4-chlorobenzoate:CoA ligase-dehalogenase. The CoA is then removed via the action of a thioesterase, an α_4 tetramer of 16 kDa polypeptides (Scholten et al., 1991). Küver et al. (1995) isolated a CoA ligase/thioesterase system from *Rhodopseudomonas palustris* which appears to break down carboxy cyclohexane and other carboxy derivatives. Interestingly, the CoA ligase had no activity towards cinnamic acid or 4-hydroxybenzoate. The partially purified thioesterase was found to be very hydrophobic with a tendency to bind to surfaces, and required glycerol for extended storage. The molecular weight was estimated to be 30 kDa and it acts as a monomer (Küver et al., 1995). The authors suggest that one function for the CoA thioesterase may be to maintain a certain level of free CoASH in the cell, and to control intracellular retention of some CoA intermediates.

The cellular function of the putative thioesterase that I identified, which converts hydroxycinnamoyl-CoA thioesters to free hydroxycinnamic acids, is unclear. Although the thioesterase reaction itself is energetically favorable, the process of hydrolyzing the thioesters has an overall energy cost to the cell since the 4CL reaction that generates the thioesters requires ATP hydrolysis (Figure 1.2). The thioesterase may have a much higher affinity for cinnamoyl-CoA than the other hydroxycinnamoyl-CoA thioesters which could inhibit the biosynthesis of phenylpropanoid compounds using cinnamic acid. Alternatively, a signal transduction pathway may exist which activates the thioesterase if the levels of hydroxycinnamoyl-CoA thioesters become elevated. In the literature, the only examples of thioesterases characterized in plants are those associated with the production of fatty acids. It

could be that the poplar thioesterase is normally involved in fatty acid biosynthesis and is compartmentalized in the plastids so that *in vivo* it would not normally affect the accumulation of hydroxycinnamoyl-CoA thioesters. Thus, the hydroxycinnamoyl-CoA thioesterase activity may be an artifact of my 4CL assay procedure or it may be a multi-functional enzyme, much like the protease/thioesterase/arylesterase found in *E. coli* (Cho and Cronan, 1994b; Lee et al., 1996).

The biological role of the thioesterase could be determined in a number of ways. The protein could be purified and kinetic analysis could be performed to determine possible cellular substrates. Immunolocalization using thioesterase-specific antibodies could show the cellular location of the protein which might suggest a function. N-terminal and internal amino acid sequences could be obtained. The sequences could be used to determine the nucleotide sequence and PCR primers could be used to obtain, at least, a partial clone of the thioesterase which could then be used to screen a cDNA library for a full length clone. This clone could be expressed using an *E. coli* or an insect/baculovirus system and the recombinant protein further characterized.

My attempts at purifying the modifying activity, now thought to be a thioesterase, were only partly successful. Part of the difficulty arose from use of an indirect assay for "modifying activity" rather than a direct assay for thioesterase activity. Modifying activity was determined by incubating a fraction with putative activity with recombinant 4CL-9 and performing a 4CL spectrophotometric assay. Modifying activity was judged to be present if the kinetics of the 4CL reaction were not linear as compared to a control assay. The activity was precipitated with 60 - 80% $(NH_4)_2SO_4$ which effectively separated it from native 4CL

activity. The activity consistently eluted as a single peak from a Superose 12 column, and on a calibrated column the protein was shown to have a size between 25 and 60 kDa.

Further attempts at purification were unsuccessful. The activity bound tightly to an anion exchange matrix and could not be eluted, even with high salt. At a pH of 5.5 or lower the activity was bound tightly to a cation exchange matrix and at a higher pH the activity would not bind. Crude extracts were subjected to chromatofocusing at pHs between 7 and 4, but activity was not recovered from the column and was probably bound tightly. This possibility is supported by the fact that the predicted pIs of two medium-chain thioesterase from *Cuphea palustris* are 9.3 and 8.9 (Dehesh et al., 1996). However, the pI of a medium-chain thioesterase from *Umbellularia californica* was 6.5 (Davies et al., 1991). Thus the pI of the poplar xylem thioesterase may be higher than 7. If so, the activity might be recovered by chromatofocusing at a higher pH.

Lee and Douglas (1996) previously studied a similar 4CL "modifier" in tobacco. The modifier appeared to be specific for 4CL activity with cinnamic acid as a substrate. They demonstrated that the modifier was probably a protein, and they hypothesized that the protein post-translationally modified 4CL either by phosphorylation or another modification. Alternatively, the protein could interact with 4CL to form a multienzyme complex. No further work has been done with this tobacco 4CL modifier, but it is most likely that it, too, is a thioesterase. This could be tested for by performing a modified 4CL reaction (4.2.3).

The fact that crude protein extracts from poplar, and possibly tobacco, contain a thioesterase activity using hydroxycinnamoyl-CoA thioesters as substrates suggests that this activity may also exist in other plants. If this is true, then the results of 4CL assays from

crude extracts should be re-examined. This would include work by Knogge et al. (1981), Kutsuki et al. (1982), Löscher and Heide (1994) and Meng and Campbell (1997). Interestingly, Löscher and Heide (1994) synthesized *p*-hydroxybenzoate in cell-free extracts of *Lithospermum erythrorhizon*. They converted 4-coumaric acid to 4-coumaroyl-CoA using crude cell extracts and mentioned that an equilibrium between synthesis and hydrolysis of 4-coumaroyl-CoA occurred. They attributed this equilibrium to a 4-coumaroyl-CoA hydrolyzing activity (possibly a thioesterase). When attempting to purify 4CL proteins, most protocols precipitate 4CL activity with 40 - 80% $(NH_4)_2SO_4$. In my hands, most native 4CL activity is readily precipitated with 40 - 60% $(NH_4)_2SO_4$ whereas the thioesterase activity is precipitated with 60 - 80% $(NH_4)_2SO_4$. It might be wise for this procedure to be adopted by others working with 4CL. There is no evidence to suggest that the thioesterase is strongly associated with 4CL and thus, partial purification of 4CL using anion exchange chromatography should be sufficient to eliminate any thioesterase activity remaining after $(NH_4)_2SO_4$ precipitation.

CHAPTER FIVE

SUMMARY AND FUTURE DIRECTIONS

The work presented in this thesis was part of a larger study designed to examine the possibility of altering lignin content and/or composition by genetic engineering. It has been speculated that 4CL plays an important role in the channeling of carbon into specific lignin pathways to produce specific monolignols. A report in the older literature (Grand et al., 1983) suggested that multiple 4CL isoforms from poplar may preferentially use specific hydroxycinnamic acids as substrates. If this is true then the possibility exists that genetically engineered trees, either over- or under-expressing a particular 4CL isoform, could produce lignin with altered properties.

My work has confirmed that *4CL* is encoded by a gene family in hybrid poplar. Two full-length cDNA clones (4CL-216 and 4CL-9) representing two different genes (*4CL1* and *4CL2*) were isolated and characterized. One of the cDNA clones is a chimera, since the 3' end was isolated from a cDNA library while the 5' end was isolated using PCR. Southern blot analysis demonstrated that still other *4CL* genes exist in the poplar genome. Two of these have since been isolated in another study (D. Cukovic and C. Douglas, unpublished data) and are similar to the two characterized aspen cDNA clones recently reported by He et al. (1998). At this time, further work is being done to fully characterize the other 4CL cDNA clones. They will also be expressed in the insect cell/baculovirus system and the properties of the recombinant proteins will be determined (D. Cukovic and C. Douglas, unpublished data).

Northern blot analysis demonstrated that *4CL1* and *4CL2* are differentially expressed. *4CL1* is preferentially expressed in old leaves, with lower levels seen in green stem and xylem, while *4CL2* is preferentially expressed in young leaves. Neither gene is expressed at high levels in developing xylem and they therefore may not play a major role in lignification of this tissue.

The cDNA clones representing 4CL-216 and 4CL-9 were expressed in a insect cell/baculovirus system. The recombinant proteins were partially purified and tested for substrate preference. Both of the recombinant proteins used hydroxycinnamic acid substrates in a similar manner. They each preferred 4-coumaric acid followed by ferulic acid, caffeic acid, cinnamic acid, and 5-hydroxyferulic acid. Neither of the recombinant proteins used sinapic acid. The substrate preferences were also similar to those found for each of the partially purified native poplar 4CL isoforms (Allina et al., 1998). This could indicate that 4CL isoforms with large differences in their substrate utilization profiles do not in fact exist in poplar, but it will be interesting to compare the enzymatic properties of the proteins encoded by the newly discovered divergent poplar *4CL* cDNAs, since it has been suggested that one of these is distinct in its inability to use 5-hydroxyferulic acid as a substrate (Hu et al., 1998).

Most 4CL proteins have a conserved proline found in domain I (PYSSGTTGLPKG), but the original 4CL-2 sequence was found to have an alanine in this position, as does a published sequence for *Oryza sativa* (rice) (Zhao et al., 1995). It would be interesting to examine the possibility that 4CL with alanine in this position displays different substrate preferences. A single amino acid change could potentially alter the substrate preference of

4CL in a manner analogous to the changed substrate specificity of the resveratrol-forming peanut STS (Shröder and Schröder, 1992). This experiment could be done by replacing the 5' end of 4CL-2 with PCR fragment C31 (Appendix 1) and expressing this new clone in the insect/baculovirus system. It might also be possible to determine the authenticity of the putative active site. This could be accomplished by site-directed mutagenesis and expression of the mutated protein in the insect/baculovirus system.

The recombinant proteins examined in this study used cinnamic acid as a substrate while the native proteins did not. It was hypothesized that a cellular factor was responsible for modifying native 4CL in such a fashion that it could no longer use cinnamic acid as a substrate. The results of my work to characterize this potential 4CL modifier strongly suggest that a 4CL modifier is not responsible for the apparent inability of native 4CL isoforms to use cinnamic acid. In fact, it appears that a thioesterase activity is responsible for the observed phenomenon. The thioesterase is able to use a variety of hydroxycinammoyl-CoA thioesters as substrates. This was previously masked by the fact that recombinant 4CL could convert most hydroxycinnamic acids to hydroxycinnamoyl-CoA thioesters faster than the thioesterase could perform the reverse reaction. However, if a reaction containing recombinant 4CL was boiled before the addition of the thioesterase, it became apparent that each hydroxycinnamoyl-CoA thioester was a potential substrate for the thioesterase.

A plant thioesterase with activity towards hydroxycinnamoyl-CoA thioesters has not been previously reported. There is much work that needs to be completed before the thioesterase is fully characterized and its biological role elucidated. The most important thing is to determine the substrate preference(s) of the thioesterase, since it is not clear at this point

whether the thioesterase preferentially uses hydroxycinnamoyl-CoA thioesters as substrates. To test this, pure hydroxycinnamoyl-CoA thioesters need to be chemically or enzymatically synthesized. Kinetic analyses could then be performed using these and other putative substrates (e.g. thioester of fatty acids) and the properties compared. This should help to reveal the actual role of the thioesterase in cellular metabolism. The thioesterase could also be further purified, using the knowledge gained about the nature of the preferred substrates (unknown when purification was attempted; section 4.3.2).

If the thioesterase is specific to hydroxycinnamoyl-CoA thioesters then it will be important to elucidate its biological function. It can be speculated that the thioesterase plays an important role in controlling the availability of different hydroxycinnamoyl-CoA thioesters for different biosynthetic pathways. If the thioesterase can be sufficiently purified, amino acid sequencing could be performed or antibodies could be prepared. Either of these could be used to isolate cDNA clones of the thioesterase.

It will be important to correlate the expression of the thioesterase with that of *4CL*. Initially, a survey of different tissues and treatments could be carried out to establish where and when the thioesterase is expressed. This could be followed by either immunolocalization (using thioesterase antibodies) or northern blot analysis (using a cDNA as the hybridization probe).

Finally, it should be determined what role the thioesterase activity may play in controlling the phenotype of poplar. For this purpose the thioesterase activity could be suppressed by producing transgenic poplar plants with either antisense or sense thioesterase constructs, and the phenotypes assessed.

BIBLIOGRAPHY

Allina SM, Pri-Hadash A, Theilmann D, Ellis BE and Douglas CJ (1998) 4-coumarate:CoA ligase in hybrid poplar. Properties of native enzymes, cDNA cloning, and analysis of recombinant enzymes. Plant Physiology **116**: 743-754

Amrhein N and Holländer H (1979) Inhibition of anthocyanin formation in seedlings and flowers by the enantiomers of α-aminooxy-β-phenylpropionic acid and their N-benzyloxycarbonyl derivative. Planta **144**: 385-389

Amrhein N, Frank G, Lemm G and Luhmann H-B (1983) Inhibiiton of lignin formation by L-α-aminooxy-β-phenylpropionic acid, an inhibitor of phenylalanine ammonia-lyase. European Journal of Cell Biology **29**: 139-144

Atanassova R, Favet N, Martz F, Chabbert B, Tollier M-T, Monties B, Fritig B and Legrand M (1995) Altered lignin composition in transgenic tobacco expressing *O*-methyltransferase sequences in sense and antisense orientation. The Plant Journal **8**: 465-477

Batard Y, Schalk M, Pierral M-A, Zimmerlin A, Durst F and Werk-Reichhart D (1977) Regulation of cinnamate 4-hydroxylase (CYP73A1) in Jerusalem artichoke tubers in response to wounding and chemical treatments. Plant Physiology **113**: 951-959

Baucher M, Chabbert B, Pilate G, Vandoorsselaere J, Tollier MT, Petitconil M, Cornu D, Monties B, Van Montagu M, Inze D, Jouanin L and Boerjan W (1996) Red xylem and higher lignin extractability by down-regulating a cinnamyl alcohol dehydrogenase in poplar. Plant Physiology **112**: 1479-1490

Becker-André M, Schultze-Lefert P and Hahlbrock K (1991) Structural comparison, modes of expression, and putative *cis*-acting elements of the two 4-coumarate:CoA ligase genes in potato. The Journal of Biological Chemistry **266**: 8551-8559

Bell-Lelong DA, Cusuman JC, Meyer K and Chapple, C (1997) Cinnamate-4-hydroxylase expression in Arabidopsis. Plant Physiology **113**: 729-738

Bernards MA, Lopez ML, Zajicek J and Lewis NG (1995) Hydroxycinnamic acid-derived polymers constitute the polyaromatic domain of suberin. The Journal of Biological Chemistry **270**: 7382-7386

Benveniste I, Salaun J-P and Durst F (1977) Wounding induced cinnamic acid hydroxylase in Jerusalem artichoke tuber. Phytochemistry **16**: 669-73

Bizzozero OA (1997) The mechanism and functional roles of protein palmitoylation in the nervous system. Neuropediatrics **28**: 23-26

Boerjan W, Cervera M-T, Steenackers M, Storme V, Ivens B, Steenackers J, Broeck AV, Michiels B, Stevens M, Slijcken JV and Van Montagu M (1997) AFLP™ genome mapping of poplar and analysis of quantitative traits. Joint meeting of the IUFRO working parties 2.04-07 and 2.04-06. Somatic cell genetics and molecular genetics of trees Abstract 138

Bolwell GP, Bell JN, Cramer CL, Schuch W, Lamb CJ and Dixon RA (1985) L-phenylalanine ammonia-lyase from *Phaseolus vulgaris*: Characterisation and differential induction of multiple forms from elicitor-treated cell suspension cultures. European Journal of Biochemsistry **149**: 411-419

Bolwell GP, Cramer CL, Lamb CJ, Schuch W and Dixon RA (1986) L-phenylalanine ammonia-lyase from *Phaseolus vulgaris*: Modulation of the levels of active enzyme by *trans*-cinnamic acid. Planta **169**: 97-107

Bolwell GP (1992) A role for the phosphorylation in the down-regulation of phenylalanine ammonia-lyase in suspension-cultured cells of french bean. Phytochemistry **31**: 4081-4086
Boudet AM, Lapierre C and Grima-Pettenati J (1995) Tansley Revie No. 80. Biochemistry and molecular biology of lignification. New Phytologist **129**: 203-236

Boudet AM (1998) A new view of lignification. Trends in Plant Science **3**: 67-71

Bradshaw Jr HD and Stettler RF (1993) Molecular genetics of growth and development in *Populus* I. Triploidy in hybrid poplars. Theoretical and Applied Genetics **86**: 301-307

Bradshaw Jr H, Villar M, Watson B, Otto K, Stewart S and Stettler R (1994) Molecular genetics of growth and development in *Populus* III. A genetic linkage map of a hybrid poplar composed of RFLP, STS, and RAPD markers. Theoretical and Applied Genetics **89**: 167-178

Brodelius, PE and Xue, Z-T (1997) Isolation and characterization of a cDNA from cell suspension cultures of *Vanilla planifolia* encoding 4-coumarate:coenzyme A ligase. Plant Physiology and Biochemistry **35**: 497-506

Campbell MM and Ellis BE (1992) Fungal elicitor-mediated responses in pine cell cultrures. I. Induction of phenylpropanoid metabolism. Planta **186**: 409-417

Chappell J and Hahlbrock K (1984) Transcription of plant defence genes in response to UV light or fungal elicitor. Nature **311**: 76-78

Chapple CCS, Vogt T, Ellis BE and Somerville CR (1992) An arabidopsis mutant defective in the general phenylpropanoid pathway. The Plant Cell **4**: 1413-1424

Cho H and Cronan, Jr. JE (1994) "Protease I" of *Escherichia coli* functions as a thioesterase in vivo. Journal of Bacteriology **176**: 1793-1795

Cramer CL, Bell JN, Ryder TB, Bailey JA, Schuch W, Bolwell GP, Robbin MP, Dixon RA and Lamb CJ (1985) Co-ordinated synthesis of phytoalexin biosynthetic enzymes in biologically-stressed cells of bean (*Phaseolus vulgaris* L.) The EMBO Journal **4**: 285-289

Cramer CL, Edwards K, Dron M, Liang X, Dildine SL, Bolwell GP, Dixon RA, Lamb CJ and Schuch W (1989) Phenylalanine ammonia-lyase gene organization and structure. Plant Molecular Biology **12**: 367-383

da Costa e Silva O, Klein L, Schmelzer E, Trezzini GF and Hahlbrock K (1993) BPF-1, a pathogen-induced DNA-binding protein involved in the plant defense response. The Plant Journal **4**: 125-135

Dalkin K, Edwards R, Edington B and Dixon RA (1990) Stress responses in alfalfa (*Medicago sativa* L.) I. Induction of phenylpropanoid biosynthesis and hydrolytic enzymes in elicitor-treated cell suspension cultures. Plant Physiology **92**: 440-446

Dangl JL, Hauffe KD, Lipphardt S, Hahlbrock K and Scheel D (1987) Parsley protoplasts retain differential responsiveness to u.v. light and fungal elicitor. The EMBO Journal **6**: 2551-2556

Davies HM, Anderson, L, Fan C and Hawkins DJ (1991) Developmental induction, purification, and further characterization of 12:0-ACP thioesterase from immature cotyledons of *Umbellularia californica.* Archives of Biochemistry and Biophysics **290**: 37-45

Davis KR and Ausubel FM (1989) Characterization of elicitor-induced defenxe responses in suspension-cultured cells of Arabidopsis. Molecular Plant-Microbe Interactions **2**: 363-368

Dehesh K, Jones A, Knutzon DS and Voelker TA (1996) Production of high levels of 8:0 and 10:0 fatty acids in transgenic canola by overexpression of *Ch FatB2*, a thioesterase cDNA from *Cuphea hookeriana*. The Plant Journal **9**: 167-172

Deshpande AS, Surendranathan KK and Nair PM (1993) The phenylpropanoid pathway enzymes in *Solanum tuberosum* exist as a multienzyme complex. Indian Journal of Biochemistry and Biophysics **30**: 36-41

Dixon RA, Choundhary AD, Dalkin K, Edwards R, Fahrendorf T, Gowri G, Harrison MJ, Lamb CJ, Loake GJ, Maxwell CA, Orr J and Paiva NL (1992) Molecular biology of stress-induced phenylpropanoid and isoflavonoid biosynthesis in alfalfa. In: Phenolic Metabolism in Plants. (HA Stafford and RK Ibrahim, eds) Plenum Press, New York, pp 91-137.

Dixon RA and Paiva NL (1995) Stress-induced phenylpropanoid metabolism. Plant Cell **7**: 1085-1097

Douglas C, Hoffmann H, Schulz W and Hahlbrock, K (1987) Structure and elicitor or u.v.-light stimulated expression of two 4-coumarate:CoA ligase genes in parsley. The EMBO Journal **6**: 1189-1195

Douglas CJ, Hauffe KD, Ites-Morales M-E, Ellard M, Paszkowski U, Hahlbrock K and Dangl JL (1991) Exonic sequences are required for elicitor and light activation of a plant defense gene, but promoter sequences are sufficient for tissue specific expression. The EMBO Journal **10**: 1767-1775

Douglas CJ, Ellard M, Hauffe KD, Molitor E, Moniz de Sá M, Reinold S, Subramaniam R, and Williams F (1992) General phenylpropanoid metabolism: regulation by environmental and developmental signals. In: Phenolic Metabolism in Plants. (HA Stafford and RK Ibrahim, eds) Plenum Press, New York, pp 63-89.

Douglas CJ (1996) Phenylpropanoid metabolism and lignin biosynthesis: from weeds to trees. Trends in Plant Science **1**: 171-178

Dwivedi UN, Campbell WH, Yu J, Datla RSS, Bugos RC, Chiang VL and Podila GK (1994) Manipulation of lignin biosynthesis in transgenic tobacco through expression of an antisense *O*-methyltransferase gene from quaking aspen. Plant Molecular Biology **26**: 61-71

Eckey-Kaltenbach H, Groβkopf E, Sandermann Jr. H and Ernst D (1994) Induction of pathogen defence genes in parsley (*Petroselinum crispum* L.) plants by ozone. Proceeding of the Royal Society of Endinburgh **102B**: 63-74

Ehlting J, Büttner D, Li Q, Douglas CJ, Kombrink E and Somssich IE (1998) The 4-coumarate:CoA ligase (4CL) gene family in Arabidopsis thaliana. Gene structures, enzymatic properties and modes of expression. 9th International conference on Arabidopsis research Abstract 510

Elkind Y, Edwards R, Mavandad M, Hedrick SA, Ribak O, Dixon RA and Lamb CJ (1990) Abnormal plant development and down-regulation of phenylpropanoid biosynthesis in transgenic tobacco containing a heterologous phenylalanine ammonia-lyase gene. Proceedings of the National Academy of Science, USA **87**: 9057-9061

Ellard-Ivey M and Douglas CJ (1996) Role of jasmonates in the elicitor- and wound-inducible expression of defense genes in parsley and transgenic tobacco. Plant Physiology **112**: 183-192

Fahrendorf T and Dixon RA (1993) Stress responses in alfalfa (*Medicaga sativa* L.) XVIII: Molecular cloning and expression of the elicitor-inducible cinnamic acid 4-hydroxylase cytochrome P450. Archives of Biochemistry and Biophysics **305**: 509-515

Fillatti JJ, Sellmet J, McCown B, Haissig B and Comai L (1987) *Agrobacterium* mediated transformation and regeneration of *Populus*. Molecular and General Genetics **206**: 192-199

Fliegmann J, Schröder G, Schanz S, Britsch L and Schröder J 1992) Molecular analysis of chalcone and dihydropinosylvin synthase from Scots pine (*Pinus sylvestris*), and differential regulation of these and related enzyme activities in stressed plants. Plant Molecular Biology **18**: 489-503

Flores G and Hubbes M (1980) The nature and role of phytoalexin produced by aspen (*Populus tremuloides* Michx.) European Journal of Forest Pathology **10**: 95-103

Fritzemeier K-H, Cretin C, Kombrink E, Rohwer F, Taylor J, Schee D and Hahlbrock K (1987) Transient induction of phenylalanine ammonia-lyase and 4-coumarate:CoA ligase mRNAs in potato leaves infected with virulent or avirulent races of *Phytophthora infestans*. Plant Physiology **85**: 34-41

Funk C and Brodelius, PE (1990a) Phenylpropanoid metabolism in suspension cultures of *Vanilla planifolia* Andr. II. Effects of precursor feeding and metabolic inhibitors. Plant Physiology **94**: 95-101

Funk C and Brodelius, PE (1990b) Phenylpropanoid metabolism in suspension cultures of Vanilla planifolia Andr.III. Conversion of 4-methoxycinnamic acids into 4-hydroxybenzoic acids. Plant Physiology **94**: 102-108

Fuhrmann J and Heise KP (1993) Factors controlling medium-chain fatty acid synthesis in plastids from maturing cuphea embryos. Zeitschrift fuer Naturforschung Section C Biosciences **48c**: 616-622

Gabriac B, Werk-Reichhart D, Teutsch H and Durst F (1991) Purification and immunocharacterization of a plant cytochrome P450: The cinnamic acid 4-hydroxylase. Archives of Biochemistry and Biophysics **288**: 302-309

Ge L and Chiang V (1996) A full length cDNA encoding *trans*-cinnamate 4-hydroxylase from developing xylem of *Populus tremuloides* (Accession No U47293) Plant Physiology **112**: 861

Grand C, Boudet A and Boudet AM (1983) Isoenzymes of hydroxycinnamate:CoA ligase from poplar stems: properties and tissue distribution. Planta **158**: 225-229

Greenaway W, Wollenweber E, Scaysbrook T and Whatley FR (1988) Esters of caffeic acid with aliphatic alcohols in bud exudate of *Populus nigra*. Zeitschrift fuer Naturforschung Section C Biosciences **43c**: 795-798

Gross GG and Zenk MH (1974) Isolation and properties of hydroxycinnamate:CoA ligase from lignifying tissue of *Forsythia*. European Journal of Biochemistry **42**: 453-459

Grotewold E, Drummond BJ, Bowen B and Peterson T (1994) The *myb*-homologous *P* gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset. Cell **76**: 543-553

Hahlbrock K (1981) Flavonoids. In: The Biochemistry of Plants, Volume 7. (EE Conn, ed) Academic Press, New York, pp 425-455

Hahlbrock K and Scheel D (1989) Physiology and molecular biology of phenylpropanoid metabolism. Annual Review of Plant Physiology and Plant Molecular Biology **40**: 347-469

Hain R, Reif H-H, Krause E, Langebartels R, Kindl J, Vornam B, Wiese W, Schmelzer E, Schreier PH, Stöcker RH and Stenzel K (1993) Disease resistance results from foreign phytoalexin expression in a novel plant. Nature **361**: 153-156

Hakamatsuka T, Noguchi H, Ebizuka Y and Sankawa U (1990) Isoflavone synthase from cell suspension cultures of *Pueraria lobata*. Chemical & Pharmaceutical Bulletin (Tokyo) **38**: 1942-1945

Harborne JB (1980) Plant Phenolics. In: Encyclopedia of Plant Physiology, New Series, Volume 8, Secondary Plant Products. (EA Bell and BV Charlwood, eds) Springer-Verlag, NY, pp. 329-402

Hauffe KD, Paszkowski U, Schulze-Lefert P, Hahlbrock K, Dangl JL and Douglas CJ (1991) A parsley 4CL-1 promoter fragment specifies complex expression patterns in transgenic tobacco. The Plant Cell **3**: 435-443

Hauffe KD, Lee SP, Subramaniam R and Dogulas CJ (1993) Combinatorial interactions between positive and negative *cis*-acting elements control spatial patterns of *4CL-1* expression in transgenic tobacco. The Plant Journal **4**: 235-253

Heilman PE and Stettler RF (1985) Genetic variation and productivity of *Populus trichocarpa* and its hybrids. II. Biomass production a 4-year plantation. Canadian Journal of Forest Research **15**: 384-388

Hipskind JD, Vincent JR, Butler LG and Nicholson (1993) Partial purification and characterization of 4-hydroxycinnamic:CoA ligase from maize leaves infected with *Bipolaris maydis*. Physiological and Molecular Plant Pathology **43**: 365-377

Holton TA and Cornish EC (1995) Genetics and biochemistry of anthocyanin biosynthesis. The Plant Cell **7**: 1071-1083

Hotze M, Schröder G and Schröder J (1995) Cinnamate 4-hydroxylase from *Catharanthus roseus*, and a strategy for the functional expression of plant cytochrome P_{450} proteins as translational fusions with P_{450} reductase in *Escherichia coli*. FEBS Letters **374**: 345-350

Hrazdina G and Wagner GJ (1985) Metabolic pathways as enzyme complexes: evidence for the synthesis of phenylpropanoids and flavonoids on membrane associated enzyme complexes. Archives of Biochemistry and Biophysics **237**: 88-100

Hu W-J, Kawaoka A, Tsai C-J, Lung J, Osakabe K, Ebinuma H and Chiang V. (1998) Compartmentalized expression of two structurally and functionally distinct 4-coumarate:CoA ligase genes in aspen (*Populus tremuloides*). Proceedings of the National Academy of Science, USA **95**: 5407-5412

Hughes WD and Galau G (1988) Preparations of RNA from cotton leaves and pollen. Plant Molecular Biology Reporter **6**: 253-257

Jiang P and Cronan, Jr. JE (1994) Inhibition of fatty acid synthesis in *Escherichia coli* in the absence of phospholipid synthesis and release of inhibition by thioesterase action. Journal of Bacteriology **176**: 2814-2821

Joos H-J and Hahlbrock K (1992) Phenylalanine ammonia-lyase in potato (*Solanum tuberosum* L.). Genomic complexity, structural comparison of two selected genes and modes of expression. European Journal of Biochemistry **204**: 621-629

Jorgensen A, Cuellar E and Thompson F (1982) Modes and tempos in the evolution of nuclear-encoded ribosomal RNA genes in legumes. Carnegie Inst Wash Year Book **81**: 98-101

Kajita S, Katayama Y and Omori S (1996) Alterations in the biosynthesis of lignin in transgenic plants with chimeric genes for 4-coumarate:coenzymeA ligase. Plant and Cell Physiology **37**: 957-965

Knobloch K-H and Hahlbrock K (1975) Isoenzymes of *p*-coumarate:CoA ligase from cell suspensions of *Glycine max*. European Journal of Biochemistry **52**: 311-320

Knobloch K-H and Hahlbrock K (1977) 4-coumarate:CoA ligase from cell suspension cultures of *Petroselinum hortense* Hoffm. Archives of Biochemistry and Biophysics **184**: 237-248

Knogge W, Beulen C and Weissenböck G (1981) Distribution of phenylalanine ammonia-lyase and 4-coumarate:CoA ligase in oat primary leaf tissues. Zeitschrift fuer Naturforschung Section C Biosciences **36c**: 389-395

Knutzon DS, Bleibaum JL, Nelsen J, Kridl JC and Thompson GA (1992) Isolation and characterization of two safflower oleoyl-acyl carrier protein thioesterase cDNA clones. Plant Physiology **100**: 1751-1758

Kochs G, Werk-Reichhart D and Griesbach H (1992) Further characterization of cytochrome P450 involved in phytoalexin synthesis in soybean: cytochrome P450 cinnamate 4-hydroxylase and 3,9-dihyroxypterocarpan 6a-hydroxylase. Archives of Biochemistry and Biophysics **293**: 187-194

Kuhn DN, Chappell J, Boudet A and Hahlbrock K (1984) Induction of phenylalanine ammonia-lyase and 4-coumarate:CoA ligase mRNAs in cultured plant cells by UV light or fungal elicitor. Proceedings of the National Academy of Science, USA **81**: 1102-1106

Kutsuki H, Shimada M and Higuchi T (1982) Distribution and roles of *p*-hydroxycinnamate:CoA ligase in lignin biosynthesis. Phytochemistry **21**: 267-271

Küver J, Xu Y and Gibson J (1995) Metabolism of cyclohexane carboxylic acid by the photosynthetic bacterium *Rhodopseudomonas palustris*. Archives of Microbiology **164**: 337-345

Laemmli U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**: 680-685

Lassner MW, Lardizabal K and Metz JG (1996) A jojoba β-ketoacyl-CoA synthase cDNA complements the canola fatty acid elongation mutation in transgenic plants. The Plant Cell **8**: 281-292

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

Lee D and Douglas CJ (1996) Two divergent members of a tobacco 4-coumarate:CoenzymeA ligase (*4CL*) gene family. Plant Physiology **112**: 193-205

Lee D, Meyer K, Chapple C and Douglas CJ (1997) Antisense suppression of 4-coumarate:CoenzymeA ligase activity in *Arabidopsis* leads to altered lignin subunit composition. The Plant Cell **9**: 1985-1998

Lee H-I, León J and Raskin, I (1995b) Biosynthesis and metabolism of salicylic acid. Proceedings of the National Academy of Science, USA **92**: 4076-4079

León J, Shulaev V, Yalpani N, Lawton M A and Raskin I (1995) Benzoic acid 2-hydroxylase, a soluble oxygenase from tobacco, catalyzes salicylic acid biosynthesis. Proceedings of the National Academy of Science, USA **92**: 10413-10417

Leonard JM, Slabaugh MB and Knapp SJ (1997) *Cuphea wrightii* thioesterases have unexpected broad specificities on saturated fatty acids. Plant Molecular Biology **34**: 669-679

Lewis, NG (1993) Plant phenolics. In: Antioxidants in Higher Plants. (RG Alscher and JL Hess, eds) CRC Press, Ann Arbor, Michigan, USA, pp. 136-169

Leyva A, Jarillo JA, Salinas J and Martinez-Aapater JM (1995) Low temperature induces the accumulation of *phenylalanine ammonia-lyase* and *chalcone synthase* mRNAs of *Arabidopsis thaliana* in a light-dependent manner. Plant Physiology **108**: 39-46

Liu D and Post-Beittenmiller D (1995) Discovery of an epidermal stearoyl-acyl carrier protein thioesterase: Its potential role in wax biosynthesis. The Journal of Biological Chemistry **270**: 16962-16969

Liu LX, Margottin F, Le Gall S, Schwartz O, Selig L, Benarous R and Benichou S (1997) Binding of HIV-1 Nef to a novel thioesterase enzyme correlates with Nef-mediated CD4 down-regulation. The Journal of Biological Chemistry **272**: 13779-13785

Liu Q, Boness MS, Liu M, Seradge E, Dixon, RA, Mabry TJ. (1995) Enzymes of B-ring-deoxy flavonoid biosynthesis in elicited cell cultures of "old man" cactus (*Cephalocereus senilis*). Archives of Biochemistry and Biophysics **321**: 397-404

Loader NM, Woolner EM, Hellyer A, Slabas AR and Safford R (1993) Isolation and characterization of two *Brassica napus* embryo acyl-ACP thioesterase cDNA clones. Plant Molecular Biology **23**: 769-778

Lofty S, Javelle F and Negrel J (1996) Purification and characterization of hydroxycinnamoyl-CoenzymeA:ω-hydroxypalmitic acid *O*-hydroxycinnamoyltransferase from tobacco (*Nicotiana tabacum* L.) cell-suspension cultures. Planta **199**: 475-480

Logemann E, Parniske M and Hahlbrock K (1995) Modes of expression and common structural features of the complete phenylalanine ammonia-lyase gene family in parsley. Proceedings of the National Academy of Science, USA **92**: 5905-5909

Lois R, Dietrich A, Hahlbrock K and Schulz W (1989) A phenylalanine ammonia-lyase gene from parsley: structure, regulation and identification of elicitor and light responsiveness elements. The EMBO Journal **8**: 1641-1648

Lois R and Hahlbrock K (1992) Differential wound activation of members of the phenylalanine ammonia-lyase and 4-coumarate:CoA ligase gene families in various organs of parsley plants. Zeitschrift fuer Naturforschung Section C Biosciences **47c**: 90-94

Löscher R and Heide L (1994) Biosynthesis of *p*-hydroxybenzoate from *p*-coumarate and *p*-coumaroyl-coenyzme A in cell-free extracts of *Lithospermum erythrorhizon* cell cultures. Plant Physiology **106**: 271-279

Lozoya E, Hoffmann H, Douglas CJ, Schulz W, Scheel D and Hahlbrock K (1988) Primary structure and catalytic properties of isoenzymes encoded by the two 4-coumarate:CoA ligase genes in parsley. European Journal of Biochemistry **176**: 661-667

Lüderitz T, Schatz G and Griesbach H (1982) Enzymatic synthesis of lignin precursors. Purification and properties of 4-coumarate:CoA ligase from cambial sap of Spruce (*Picea abies* L.). European Journal of Biochemistry **123**: 583-586

MacKay JJ, O'Malley DM, Presnell T, Booker FT, Campbell MM, Whetten RW and Sederoff RR (1997) Inheritance, gene expression, and lignin characterization in a mutant pine deficient in cinnamyl alcohol dehydrogenase. Proceedings of the National Academy of Science, USA **94**: 8255-8260

Malamy J and Klessig DF (1992) Salicylic acid and plant disease resistance. The Plant Journal **2**: 643-654

Maule AJ and Ride JP (1983) Cinnamate 4-hydroxylase and hydroxycinnamate:CoA ligase in wheat leaves infected with *Botrytis cinerea*. Phytochemistry **22**: 1113-1116

McKegney GR, Butland, SL, Theilmann D and Ellis BE (1996) Expression of poplar phenylalanine ammonia-lyase in insect cell cultures. Phytochemistry **41**: 1259-1263

Meng H and Campbell, WH (1997) Facile enzymatic synthesis of caffeoyl CoA. Phytochemistry **44**: 605-608

Mizutani M, Ward E, DiMaio J, Ohta D, Ryals J and Sato R (1938) Molecular cloning and sequencing of a cDNA encoding Mung Bean cytochrome P450 ($P450_{C4H}$) possessing cinnamate 4-hydroxylase activity. Biochemical and Biophysical Research Comunications **190**: 875-880

Mizutani M, Ohta D and Sato R (1997) Isolation of a cDNA and a genomic clone encoding cinnamate 4-hydroxylase from Arabidopsis and its expression manner in planta. Plant Physiology **113**: 755-763

Moniz de Sá M, Subramaniam R, Williams FE and Douglas CJ (1992) Rapid activation of phenylpropanoid metabolism in elicitor-treated hybrid poplar (*Populus trichocarpa* Torr & Gray X *Populus deltoides* Marsh) suspension-cultured cells. Plant Physiology. **98**: 728-737

Murray RDH, Méndez J and Brown SA (1982) The Natural Coumarins. Occurrence, Chemistry and Biochemistry. John Wiley & Sons, Ltd., Toronto, ON, pp. 304-307

Nürnberger T, Nennstiel D, Jabs T, Sacks WR, Hahlbrock K and Scheel D (1994) High affinity binding of a fungal oligopeptide elicitor to parsley plasma membranes triggers multiple defense responses. Cell **78**: 449-460

Ohl S, Hedrick SA, Chory J and Lamb CJ (1990) Functional properties of phenylalanine ammonia-lyase promoter from Arabidopsis. The Plant Cell **2**: 837-848

Osakabe Y, Ohtsubo Y, Kawai S, Katayama Y and Morohoshi N (1995a) Structure and tissue-specific expression of genes for phenylalanine ammonia-lyase from a hybrid aspen, *Populus kitakamiensis*. Plant Science **105**: 217-226

Osakabe Y, Osakabe K, Kawai S, Katayama Y and Morohoshi N (1995b) Characterization of the structure and determination of mRNA levels of the phenylalanine ammonia-lyase gene family from *Populus kitakamiensis*. Plant Molecular Biology **28**: 1133-41

Pallas JA, Paiva NL, Lamb C and Dixon RA (1996) Tobacco plants epigenetically suppressed in phenylalanine ammonia-lyase expression do not develop systemic acquired resistance in response to infection by tobacco mosaic virus. The Plant Journal **10**: 281-293

Paré PW, Mischke CF, Edwards R, Dixon RA, Norman HA and Mabry TJ (1992) Induction of phenylpropanoid pathway enzymes in elicitor-treated cultures of *Cephalocereus senilis*. Phytochemistry **31**: 149-153

Parsons TJ, Bradshaw Jr. HD and Gordon MP (1989) Systemis accumulation of specific mRNAs in response to wounding in poplar trees. Proceedings of the National Academy of Science, USA **86**: 7895-7899

Pearl IA and Darling, SF (1971) Phenolic extractives of the leaves of *Populus balsamifera* and of *P. trichocarpa*. Phytochemistry **10**: 2844-2847

Phillips DA (1992) Flavonoids: plant signals to soil microbes. In: Phenolic metabolism in plants. (HA Stafford and RK Ibahim, eds.) Plenus Press, New York, pp. 201-231

Piquemal J, Lapierre C, Myton K, O'Connell A, Schuch W, Grima-Pettenati J and Boudet A-M (1998) Down-regulation of cinnamoyl-CoA reductase induces significant changes of lignin profiles in transgenic tobacco plants. The Plant Journal **13**: 71-83

Pierrel MA, Batard Y, Kazamaier M, Mignotte-Vieux C, Durst F and Werk-Reichhart D (1994) Catalytic properties of the plant cytochrome *P*450 CYP73 expressed in yeast. Substrate specificity of a cinnamate hydroxylase. European Journal of Biochemistry **224**: 835-844

Ragg H, Kuhn D and Hahlbrock K (1981) Coordinated regulation of 4-coumarate:CoA ligase and phenylalanine ammonia-lyase mRNAs in cultured plant cells. The Journal of Biological Chemistry **256**: 10061-10065

Ranjeva R, Boudet AM and Faggion R (1976) Phenolic metabolism in *Petunia* tissues. IV. Properties of *p*-coumarate:coenzyme A ligase isozymes. Biochemie **58**: 1255-1262

Reinold S, Hauffe KD and Douglas CJ (1993) Tobacco and parsley 4-coumarate:CoA ligase genes are temporally and spatially regulated in a cell type-specific manner during tobacco flower development. Plant Physiology **101**: 373-383

Regan S, Sterky F, Herzber M, Amini B, Karlsson J, Berg S, Lundeberg J, Sandberg G, Gustafsson P, Sundberg B and Uhlén M (1997) Knocking on wood: large-scale production of ESTs from the wood-forming tissues of poplar. Joint meeting of the IUFRO working parties 2.04-07 and 2.04-06. Somatic cell genetics and molecular genetics of trees Abstract 93

Rodgers MW, Zimmerlin A, Werk-Reichhart D and Bolwell GP (1993) Microsomally associated heme proteins from french bean: Characterization of the cytochrome P450 cinnamate-4-hydroxylase and two peroxidases. Archives of Biochemistry and Biophysics **304**: 74-80

Roesler J, Krekel F, Amrhein N and Schmid J (1997) Maize phenylalanine ammonia-lyase has tyrosine ammonia-lyase activity. Plant Physiology **113**: 175-179

Roitt, IM (1988) Essential Immunology. Blackwell Scientific Publications, Boston, MA, p. 188

Sablowski RWM, Moyano E, Culianez-Macia A, Schuch W, Martin C and Bevan M (1994) A flower-specific Myb protein activates transcription of phenylpropanoid biosynthetic genes. The EMBO Journal **13**: 128-137

Sambrook J, Fritsch EF and Maniatis T (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sanchez J, Del Cuvillo MT and Harwood JL (1992) Glycerolipid biosynthesis by microsomal fractions from olive fruits. Phytochemistry **31**: 129-134

Scaysbrook T, Greenaway W and Whatley FR (1991) Relation of "antimicrobial" compounds present in poplar bud exudates to disease resistance by poplars. Zeitschrift fuer Naturforschung Section C Biosciences **47c**: 197-201

Schnitzler J-P, Madlung J, Rose A and Sietz HU (1992) Biosynthesis of *p*-hydroxybenzoic acid in elicitor-treated carrot cell cultures. Planta **188**: 594-600

Scholten JD, Chang K-H, Babitt C, Charest H, Sylvestre M and Dunaway-Mariano D (1991) Novel enzymic hydrolytic dehalogenation of a chlorinated aromatic. Science **25**: 182-185

Schröder J (1989) Protein sequence homology between plant 4-coumarate:CoA ligase and firefly luciferase. Nucleic Acids Research **17**: 460

Schröder G and Schröder J (1992) A single change of histidine to glutamine alters the substrate preference of a stilbene synthase. The Journal of Biological Chemistry **267**: 20558 - 20560

Sederoff R, Campbell M, O'Malley, D and Whetten R (1994) Genetic regulation of lignin biosynthesis and the potential modification of wood by genetic engineering in loblolly pine. In: Genetic Engineering of Plant Secondary Metabolism. (B.E. Ellis et al., eds.) Plenum Press, NY pp. 313-355

Séguin A, Laible G, Leyva A, Dixon RA and Lamb CJ (1997) Characterization of a gene encoding a DNA-binding protein that interacts *in vitro* with vascular specific *cis* elements of the phenylalanine ammonia-lyase promoter. Plant Molecular Biology **35**: 281-291

Sewalt VJH, Ni W, Blount JK, Jung HG, Masoud SA, Howles PA, Lamb C and Dixon RA (1997) Reduced lignin content and altered lignin composition in transgenic tobacco down-regulated in expression of L-phenylalanine ammonia-lyase or cinnamate 4-hydroxylase. Plant Physiology **115**: 41-50

Shain L and Miller JB (1982) Pinocembrin: An antifungal compound secreted by leaf glands of eastern cottonwood. Phytophathology **72**: 877-880

Shirley BW, Bandara ABPA, Burbulis IE, Pelletier MK, Santos MO and Saslowsky DE (1998) Exploring the flavonoid enzyme metabolon. 9th International conference on Arabidopsis research Abstract 5

Stafford HA (1974) Possible multienzyme complexes regulating the formation of C_6-C_3 compounds and lignins in higher plants. Recent Advances in Phytochemistry **8**: 53-79

Stöckigt J and Zenk M (1975) Chemical synthesis and properties of hydroxycinnamoyl-CoenzymeA derivatives. Zeitschrift fuer Naturforschung Section C Biosciences **30c**: 352-358

Strauss SH, Sheppard LA, Rottman WH, Brunner AM and Meilan, R (1977) Floral homeotic gene expression and genetic engineering of sterility in *Populus*. Joint meeting of the IUFRO working parties 2.04-07 and 2.04-06. Somatic cell genetics and molecular genetics of trees Abstract 23

Subramaniam R, Reinold S, Molitor EK and Douglas CJ (1993) Structure, inheritance, and expression of hybrid poplar (*Populus trichocarpa* X *Populus deltoides*) phenylalanine ammonia-lyase genes. Plant Physiology **102**: 71-83

Summers MD and Smith GE (1987) A manual of methods for baculovirus vectors and insect cell culture procedures. Texas Agricultural Experimental Bulletin No. 1555

Tamagnone L, Merida A, Parr A, Mackay S, Culianez-Macia F, Roberts K and Martin, C (1998) The AmMYB308 and AmMYB330 transcription factors from Antirrhinum regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. The Plant Cell **10**: 135-154

Tanaka Y, Kojima M and Uritani I (1974) Properties, development and cellular localization of cinnamic acid 4-hydroxylase in cut-injured sweet potato. Plant and Cell Physiology **15**: 843-854

Teutsch H G, Hasenfratz MP, Lest A, Stoltz C, Garnier J-M, Jeltsh J--M, Durst F and Werk-Reichhart DW (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. Proceedings of the National Academy of Science, USA **90**: 4102-4106

Tietjen KG, Hunkler D and Matern U (1983) Differential responses of cultured parsley cells to elicitors from two non-pathogenic strains of fungi. I. Identification of induced products as coumarin derivative. European Journal of Biochemistry **131**: 404-407

Uhlmann A and Ebel J (1993) Molecular cloning and expression of 4-coumarate:coenzyme A ligase, an enzyme involved in the resistance of soybean (*Glycine max* L) against pathogen infection. Plant Physiology **102**: 1147-1156

Van Doorsselaere J, Baucher M, Chognot E, Chabbert B, Tollier M-T, Petit-Conil M, Leplé J-C, Pilate G, Cornu D, Monties B, Van Montagu M, Inzé D and Jouanin L (1995a) A novel lignin in poplar trees with reduced caffeic acid/5-hydroxyferulic acid *O*-methyltransferase activity. The Plant Journal **8**: 855-864

Van Doorsselaere J, Baucher M, Feuillet C, Boudet AM, Van Montagu M and Inzé D (1995b) Isolation of cinnamyl alcohol dehydrogenase cDNAs from two important economic species: alfalfa and poplar. Demonstration of a high homology of the gene within angiosperms. Plant Physiology and Biochemistry **33**: 105-109

Verkruyse LA and Hofmann SL (1996) Lysosomal targeting of palmitoyl-protein thioesterase. The Journal of Biological Chemistry **271**: 15831-15836

Vesa J, Hellsten E, Verkruyse LA, Camp LA, Rapola J, Santavuori P, Hoffmann SL and Peltonen L (1995) Mutations in the palmitoyl protein thioesterase gene causing neuronal ceroid lipofuscinosis. Nature **376**: 584-587

Vignols F, Rigau J, Torres MA, Capellades M and Puigdomènech (1995) The *brown midrib3* (*bm3*) mutation in maize occurs in the gene encoding caffeic acid *O*-methyltransferase. The Plant Cell **7**: 407-416

Voelker TA, Worrell AC, Anderson L, Bleibaum J, Fan C, Hawkins DJ, Radke S and Davies HM (1992) Fatty acid biosynthesis redirected to medium chains in transgenic oilseed plants. Science **257**: 72-73

Voelker TA, Jones A, Crammer AM, Davies HM and Knutzon DS (1997) Broad-range and binary-range acyl-acyl-carrier-protein thioesterase suggest an alternative mechanism for medium-chain production in seeds. Plant Physiology **114**: 669-677

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

Wallis PJ and Rhodes MJC (1977) Multiple forms of hydroxycinnamate:CoA ligase in etiolated pea seedlings. Phytochemistry **16**: 1891-1894

Webb NR and Summers MD (1990) Expression of proteins using recombinant baculoviruses. Technique **2**: 173-188

Whetten R and Sederoff R (1995) Lignin biosynthesis. Plant Cell **7**: 1001-1013

Whetten RW, MacKay JJ and Sederoff RR (1998) Recent advances in understanding lignin biosynthesis. Annual Review of Plant Physiology and Plant Molecular Biology **49**: 585-609

Yalpani N, León J, Lawton MA and Raskin, I (1993) Pathway of salicylic acid biosynthesis in healthy and virus-inoculated tobacco. Plant Physiology **103**: 315-321

Yahiaoui N, Marque C, Myton KE, Negrel and Boudet AM (1998) Impact of different levels of cinnamyl alcohol dehydrogenase down-regulation on lignins of transgenic tobacco plants. Planta **204**: 8-15

Yazaki K, Ogawa A and Tabata M (1995) Isolation and characterization of two cDNAs encoding 4-coumarate:CoA ligase in *Lithospermum* cell cultures. Plant and Cell Physiology **36**: 1319-1329

Ye ZH, Kneusel RE, Matern U and Varner JE (1994) An alternative methylation pathway in lignin biosynthesis in *Zinnia*. The Plant Cell **6**: 1427-1439

Zhang X-H and Chiang V (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 Physiology **113**: 65-74

Zhao Y, Kung SD, Dube SK (1990) Nucleotide sequence of rice 4-coumarate:CoA ligase gene, 4-CL1. Nucleic Acids Research **18**: 6144

APPENDIX

Alignment of 4CL2 (top line), and translations of PCR products and 4CL9 (bottom line). Underlined amino acids of the translated PCR products are those that differ from 4CL2.

```
4CL2    1  MEAKNDQAQE FIFRSKLPDI HIPNHLPLHT YCFENLSRFK DNPCLINGPT
A16     1  MEAKNDQAQE FIFRSKLPDI HIPNHLPLHT YCFENLSRFK DNPCLINGPT
B2      1  MEAENDQAQE FIFRSKLPDI HIPNHLPLHT YCFENLSRFK DNPCLINGPT
B7      1  MEAKNDQAQE FIFRSKLPDI HIPNHLPLHT YCFENLSRFK DNPCLINGPT
B10     1  MEAENDQAQE FIFRSKLPDI HIPNHLPLHT YCFENLSRFK DNPCLINGPT
C11     1  MEVENDQAEE FIFRSKLPDI HIPNHLPLHT YCFENLSRFK DSPCLINGPT
C31     1  MEAKNDQAQE FIFRSKLPDI HIPNHLPLHT YCFENLSRFK DNPCLINGPT
C34     1  MEAENDQAQE FIFRSKLPDI HIPNHLPLHT YCFENLSRFK DNPCLINGPT
C3      1  MEVENDQAQE FIFRSKLPDI HIPNHLPLHT YCFENLSRFK DNPCLINGPT
C42     1  MEVENDQAQE FIFRSKLPDI HIPNHLPLHT YCFENLSRFK DSPCLINGPT
4CL9    1  MEANKDQVQE FIFRSKLPDI YIPNHLPLHT YCFEKLSQFK DNPCLINGPT

4CL2   51  GEIHTYAEVE LTSRKVASGL NKLGIKQGDV ILLLLQNSPE FVFAFLGASI
A16    51  GEIHTYAEVE LTSRKVASGL NKLGIKQGDV ILLLLQNSPE FVFAFLGASI
B2     51  GEIHTYAEVE LTSRKVASGL NKLGIKQGAV ILLLLQNSPE FVFAFLGASI
B7     51  GEIHTYAEVE LTSRKVASGL NKLGIKQGDV ILLLLQNSPE FVFAFLGASI
B10    51  GEIHTYAEVE LTSRKVASGL NKLGIKQGDV ILLLLQNSPE FVFAFLGASI
C11    51  GEIHTYAEVE LTSRKVASGL NKLGIKQGDV ILLLLQNSPE FVFAFLGASI
C31    51  GEIHTYAEVE LTSRKVASGL NKLGIKQGDV ILLLLQNSPE FVFAFLGASI
C34    51  GEIHTYAEVE LTSRKVASGL NKLGIKQGDV ILLLLQNSPE FVFAFLGASI
C3     51  GEIHTYAEVE LTSRKVASGL NKLGIKQGDV ILLLLQNSPE FVFAFLGASI
C42    51  GEIHTYAEVE VTSRKVASGL NKLGIKQGDV ILLLLQNSPE FVFASLGASI
4CL9   51  GDIYTYADVE LTSRKVASGL YKLGLQQGDV ILLLLQNSPE FVFAFLGASF

4CL2  101  IGAISTTANP FYTPAEVAKQ ATASKAKLII TQAVYAEKVQ EFVKENVHVK
A16   101  IGAISTTANP FYTPAEVAKQ ATASKAKLII TQAVYAEKVQ EFVKENVHVK
B2    101  _GAISTTANP FYTPAEVAKQ ATASKAKLII TQAVYAEKVQ EFVKENVHVK
B7    101  IGAISTTANP FYTPAEVAKQ ATATKAKLII TQAVYAEKVQ EFVKENVHVK
B10   101  IGAISTTANP FYTPAEVAKQ ATASKAKLII TQAVYAEKVQ EFVKENDHVK
C11   101  IGAISTTANP FYTPAEVAKQ ATASKAKLII TQAVYAEKVQ EFVKENVHVK
C31   101  IGAISTTANP FYTPAEVAKQ ATASKAKLII TQAVYAEKVQ EFVKENVHVK
C34   101  IGAISTTANP FYTPAEVAKQ ATASKAKLII TQAVYAEKVQ EFVKENVHVK
C3    101  IGAISTTANP FYTPAEVAKQ ATASKAKLYI TQAVYAEKVQ EFVKENVHVK
C42   101  IGAISTTANP FYTPAEVAKQ ATASKAKLYI TQAVYAEKVQ EFVKENVHVK
4CL9  101  IGAISSTANP FYTSAEIAKQ ATASKAKLII THAAYAEKVQ QFAQENDHVK
```

```
4CL2 151  IVTVDSPPEN YLHFSELTNS DEDDIPAVEI NPDDVVALPY SSGTTGLAKG
A16  151  IVTVDSPPEN YLHFSELTNS DEDDIPAVEI NPDDVVALPY SSGTTGLPKG
B2   150  IVTVDSPPEN YLHFSELTNS DEDDIPAVEI NPDDVVALPY SSGTTGLPKG
B7   151  IVTVDSPPEN YLHFSELTNS DEDDIPAVEI NPDDVVALPY SSGTTGLPKG
B10  151  IVTVDSPPEN YLHFSELTNS DEDDIPAVEI NPDDVVALPY SSGTTGLPKG
C11  151  IVTVDSPPEN YLHFSELTNS DEDDIPAVEI NPDDVVALPY SSGTTGLPKG
C31  151  MVTVDSPPEN YLHFSELTNP DEDDIPAVEI NPDDVVALPY SSGTTGLAKG
C34  151  IATVDSPPEN YLHFSELTNS DEDDIPAVEI NPDDVVALPY SSGTTGLPKG
C3   151  IVTVDSPPEN YLHFSELTNS DEDDIPAVEI NPDDVVALPY SSGTTGLPKG
C42  151  IVTVDSPPEN YLHFSELTNS DEDDIPAVEI NPDDVVALPY SSGTTGLPKG
4CL9 151  IMTIDSLTEN CLHFSELTSS DENEIPTVKI KPDDIMALPY SSGTTGLPKG
```

```
4CL2 201  VMLTHKGLVT S
A16  201  VMLTHKGLVT S
B2   201  VMLTHKGCYE
B7   201  VMLTHKGLVT S
B10  201  VMLTHKGLVT S
C11  201  VMLTHKGLVT S
C31  201  VMLTHKGLVT S
C34  201  VMLTHKGLVT S
C3   201  VMLTHKGLVT S
C42  201  VMLTHKGLVT S
4CL9 201  VMLTHKGLVT S
```