Plant Physiol. (1996) 112: 193–205

Two Divergent Members of a Tobacco 4-Coumarate:Coenzyme A Ligase (*4CL*) Gene Family[1]

cDNA Structure, Gene Inheritance and Expression, and Properties of Recombinant Proteins

Diana Lee and Carl J. Douglas*

Department of Botany, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4

Several cDNA clones encoding 4-coumarate:coenzyme A ligase (4CL) were isolated from a tobacco (*Nicotiana tabacum*) cDNA library and grouped into two classes. Sequencing of one cDNA from each class showed that the clones were similar to other *4CL* genes and about 80% identical with each other. Genomic Southern blots using DNA from *Nicotiana sylvestris*, *Nicotiana tomentosiformis*, and *N. tabacum* demonstrated the presence of both classes of *4CL* sequences (*4CL1* and *4CL2*) in the progenitor species and in tobacco. Northern blots indicated that *4CL* mRNA transcripts are highest in old stems and higher in the unpigmented corolla tubes than in the pigmented limbs of tobacco flowers. The *4CL* genes are developmentally regulated and are wound and methyl jasmonate inducible. The relative abilities of recombinant 4CL1 and 4CL2 proteins to utilize 4-coumarate, ferulate, and caffeate as substrates were similar and comparable with that of 4CL in tobacco stem extracts. Surprisingly, both recombinant 4CL proteins utilized cinnamate as a substrate, an activity not observed in stem extracts. This activity was inhibited by a heat-labile, high-molecular-weight factor found in tobacco stem extracts, suggesting that the substrate specificity of 4CL is, in part, determined by the activity of proteinaceous cellular components.

The flow of carbon from primary metabolism into the biosynthesis of an array of phenylpropanoid secondary products involves a minimum of three enzymatic steps, catalyzed by the actions of PAL, cinnamate 4-hydroxylase, and 4CL, which collectively form the general phenylpropanoid pathway. The phenylpropanoid products formed by the action of specific downstream branch pathways include coumarins, flavonoids, lignin, suberin, tannins, and other phenolic compounds, which serve diverse functions as phytoalexins, UV protectants, floral and fruit pigments, structural components of cell walls, and signaling molecules (Hahlbrock and Scheel, 1989).

As the last enzyme in the general phenylpropanoid pathway, 4CL converts 4-coumaric acid and other hydroxy- or methoxy- derivatives of cinnamic acid, such as caffeic acid, ferulic acid, and sinapic acid, into the corresponding CoA esters. These esters serve as substrates for entry into various branch pathways (Hahlbrock and Scheel, 1989). Because of its terminal position in the general phenylpropanoid pathway and its ability to utilize a number of related substrates, 4CL may play a key role in regulating carbon flow into specific branch pathways of phenylpropanoid metabolism. It has been postulated that 4CL isoforms with distinct abilities to use differently substituted hydroxycinnamic acids as substrates could participate in directing carbon into different branch pathways (Knobloch and Hahlbrock, 1975; Grand et al., 1983). For example, 4-coumaroyl:CoA could be directed into flavonoid biosynthesis and different amounts of feruloyl and sinapoyl:CoA directed into the biosynthesis of lignin via distinct 4CL isoforms. In support of this, physically distinct 4CL isoforms have been reported from soybean, petunia, poplar, maize, and parsley (Knobloch and Hahlbrock, 1975; Ranjeva et al., 1976; Grand et al., 1983; Vincent and Nicholson, 1987; Lozoya et al., 1988). Furthermore, in soybean, petunia, and poplar, partially purified 4CL isoforms exhibit different substrate specificities toward substituted cinnamic acids (Knobloch and Hahlbrock, 1975; Ranjeva et al., 1976; Grand et al., 1983). In contrast, a single 4CL form was purified from loblolly pine (Voo et al., 1995), and two slightly different 4CL isoforms with similar substrate preferences are encoded by two parsley *4CL* genes (Lozoya et al., 1988), suggesting that 4CL does not participate in directing metabolites into different branch pathways in these species.

The analysis of *4CL* cDNAs and/or genes from a limited number of plants supports the view that 4CL is encoded by multiple, divergent genes in some plants (rice, soybean, and poplar; Zhao et al., 1990; Uhlmann and Ebel, 1993; Allina and Douglas, 1994), very similar duplicated genes in others such as parsley and potato (Lozoya et al., 1988; Becker-André et al., 1991), and apparently single genes in others, such as Arabidopsis and pine (Lee et al., 1995; Voo et al., 1995). Based on earlier biochemical studies cited above, divergent *4CL* genes in plants such as soybean and poplar are predicted to encode 4CL enzymes with distinct enzymatic properties, but this prediction has not yet been

[1] Supported by a Natural Sciences and Engineering Research Council (NSERC) of Canada research grant to C.J.D. and by NSERC graduate and University of British Columbia University Graduate fellowships to D.L.

* Corresponding author; e-mail cdouglas@unixg.ubc.ca; fax 1–604–822–6089.

Abbreviations: 4CL, 4-coumarate:CoA ligase; CHS, chalcone synthase.

addressed in the literature. To our knowledge, with the exception of a single study of the two nearly identical parsley 4CL isoforms (Lozoya et al., 1988), no information about the properties of recombinant proteins encoded by *4CL* gene family members has been published.

4CL gene expression is regulated developmentally and is activated by external stimuli such as pathogen infection, elicitor treatment, wounding, and UV light irradiation (Douglas et al., 1987; Schmelzer et al., 1989; Wu and Hahlbrock, 1992). In tobacco flowers in situ hybridizations showed that endogenous tobacco *4CL* transcripts and those of an introduced parsley *4CL1* gene accumulate in a cell-type-specific manner and that the patterns of accumulation are generally consistent with the sites of phenylpropanoid natural product accumulation (Reinold et al., 1993). Also, *4CL* expression in tobacco is activated by wounding, light, and methyl jasmonate treatment (Douglas et al., 1991; Ellard-Ivey and Douglas, 1996).

We have begun an investigation of *4CL* genes in tobacco (*Nicotiana tabacum* L.) and report here that 4CL is encoded by multiple genes in this plant and its diploid progenitors. Full-length or near-full-length cDNA clones specific to two of the genes (*4CL1* and *4CL2*) were cloned and sequenced. Northern blots showed that, although the two genes are relatively divergent in sequence, they are similarly expressed in various tobacco organs. Another interesting point is that *4CL* RNA accumulates to much higher levels in the unpigmented corolla tube of tobacco flowers than in the limb, where anthocyanin pigments accumulate. Recombinant 4CL1 and 4CL2 proteins produced in *Escherichia coli* were active and had similar relative abilities to use different hydroxycinnamic acids as substrates. However, in contrast to 4CL activity measured in tobacco extracts, both recombinant proteins efficiently utilized cinnamate as a substrate. This activity of the recombinant proteins was suppressed by incubation with tobacco extracts. Suppression was dependent on the presence of a high-molecular-weight, heat-labile component in the extracts, providing evidence that the 4CL enzyme in tobacco is posttranslationally modified by, or interacts with, other protein(s) in vivo.

MATERIALS AND METHODS

Tobacco (*Nicotiana tabacum* L. cv Xanthi SR1) plants were grown in growth chambers at 23°C with an 8-h dark and 16-h light regime. Mature, fully expanded leaves were wounded by slicing the leaves into 2-mm-wide strips and placing them on a piece of Whatman filter paper moistened with Murashige-Skoog medium for 24 h before RNA isolation. Methyl jasmonate-treated plants were sprayed with 1 mM methyl jasmonate (in 1% Triton X-100 [Sigma]) until the solution ran off the leaves. The treated plants were covered with a bell jar and placed under constant light for 24 h, after which mature, fully expanded leaves were excised and used for RNA isolation. Young shoot tips were no more than 8 cm long. Young stems, approximately 2 cm long, were harvested just below the apical meristems. Old stems, greater than 1 cm in diameter and approximately 6 cm in length, were harvested from the base of mature plants. Floral tissues were harvested from developmental stages 1 to 6 as described by Reinold et al. (1993). The floral tissues were divided into sectors as described by Drews et al. (1992): the limb (corresponding to the pigmented part of the corolla), the tube (corresponding to the white part of the corolla), the base (corresponding to the corolla that is surrounded by the sepals), and the sepal. Stem sections approximately 1 cm in diameter, between the third and sixth nodes above the base of mature tobacco plants, were used for enzyme assays.

Nicotiana sylvestris and *Nicotiana tomentosiformis* seeds were kindly donated by Société Nationale d'Exploitation Industrielle des Tobacs et Allumettes (Institut Expérimental du Tabac, Bergerac, France), and plants were grown as described for *N. tabacum*.

cDNA Library Construction and Screening

A cDNA library was constructed in λZAPII (Stratagene) using poly(A)$^+$ RNA isolated from young tobacco shoot tips (S. Lee and C. Douglas, unpublished data). The primary library, consisting of approximately 10^7 independent recombinants, was amplified and approximately 10^5 plaque-forming units were screened using a probe generated from a 2-kb potato *4CL* cDNA (Becker-André et al., 1991). Hybridized filters were washed at low stringency (2× SSC, 65°C) as described by Sambrook et al. (1989).

DNA Sequencing and Sequence Analysis

All putative *4CL* cDNA clones were partially sequenced at the 5′ and 3′ ends using the T7 Sequencing Kit (Pharmacia). Two clones, Nt4CL-1 and Nt4CL-19, were subcloned by conventional methods, and both strands were completely sequenced by the University of British Columbia Nucleic Acid-Protein Service Unit using the PRISM Ready Reaction DyeDeoxy Terminator Cycle Sequencing Kit (Applied Biosytems). DNA and predicted amino acid sequences were analyzed using University of Wisconsin Genetics Computer Group (Madison) software. Predicted phosphorylation and glycosylation sites, pIs, and molecular weights were determined using computer programs PRO-SITE and pI/MW from the University of Geneva (Switzerland).

Nucleic Acid Methods

Standard molecular biology techniques were performed as described by Sambrook et al. (1989). Genomic DNA from *Nicotiana* sp. was isolated using hexadecyltrimethylammonium bromide as described by Doyle and Doyle (1990). Genomic DNA (10 μg) was digested with restriction enzymes having sites not found in the cDNA clones, electrophoresed in a 0.8% agarose gel, and blotted onto a Hybond nylon membrane (Amersham). Southern blots were hybridized to random primer-labeled probes generated from either Nt4CL-1 or Nt4CL-19 and washed at high stringency (0.2× SSC, 65°C).

RNA was extracted using guanidinium hydrochloride as described by Logemann et al. (1987). RNA (10 μg) was electrophoresed in a 1.2% agarose gel containing formaldehyde, rinsed in water for 1 h, and then blotted onto

a nylon membrane. Northern blots probed with Nt4CL were washed at moderate stringency (0.5× SSC, 65°C), whereas those probed with StPAL were washed at low stringency. To demonstrate evenness of loading, blots were stripped with 0.1% SDS and re-probed with a gene encoding for rRNA from pea (Jorgensen et al., 1982) with high-stringency washes.

DNA fragments were radioactively labeled using the Random Primer DNA Labeling System (GIBCO-BRL) according to the manufacturer's specifications.

Construction of Recombinant 4CL Proteins

PCR was used to engineer *Sph*I restriction sites into the 5′ ends of the Nt4CL-1 and Nt4CL-2 cDNAs using a T7 primer and cDNA-specific oligonucleotides (5′-GGGGCATGCCA-ATGGAGACTACTAC-3′ for Nt4CL-1 and 5′-GGGGCAT-GCATGGAGAAAGATACAAAACAG-3′ for Nt4CL-19). Nt4CL-1 and Nt4CL-19 PCR products were cloned into the QIAexpressionist (QiaGen, Chatsworth, CA) expression plasmids pQE-32 and pQE-30, respectively. The 5′ and 3′ portions of the DNA constructs were sequenced to ensure fidelity of the PCR. The center core of each PCR product was excised and replaced with the corresponding portion from the original cDNA to eliminate the possibility of single base-pair changes from the PCR. Expression plasmids were propagated and expressed in *E. coli* strain M15 as recommended by the manufacturer. The expression plasmid (pQE-30) without a cDNA insert was generated in strain M15 as a negative control.

Bacterial and Plant Extracts

A bacterial culture (A_{600} approximately 0.7; 10 mL) was induced with 2 mM isopropylthio-β-galactoside for 4 h. After the sample was centrifuged the bacterial pellet was resuspended in 2 mL of 200 mM Tris, pH 7.8, and cells were disrupted in a French press. Cellular debris was removed by centrifugation, and glycerol was added to the supernatant to a final concentration of 30%. Samples were frozen in liquid nitrogen and stored at −80°C.

Tobacco stem extracts were prepared essentially as described by Knobloch and Hahbrock (1977) for 4CL enzyme assays. Briefly, stem sections were ground into a fine powder, resuspended in 200 mM Tris, pH 7.8, and 15 mM β-mercaptoethanol, rotated for 15 min in the presence of 10% (w/w) Dowex (Bio-Rad), centrifuged to remove debris, concentrated in 10K cutoff microconcentrators (Microsep; Filtron, Northborough, MA), made to 30% glycerol, frozen in liquid nitrogen, and stored at −80°C. Protein content in plant and bacterial extracts was quantified by the Bradford (1976) method using BSA as a standard and the Bio-Rad Protein Assay Kit.

Enzyme Assays and Western Blot Analysis

4CL activity was measured at room temperature spectrophotometrically as described by Knobloch and Hahlbrock (1977) using 5 mM ATP, 5 mM $MgCl_2$, 0.33 mM CoA, and 0.2 mM cinnamic acid derivatives as substrates. The change in absorbance of the reaction mixtures was monitored at wavelengths of 311, 333, 346, 345, and 352 nm, according to the absorption maxima for cinnamoyl:CoA, 4-coumaroyl:CoA, caffeoyl:CoA, feruloyl:CoA, and sinapoyl:CoA, respectively (Stöckigt and Zenk, 1975).

Protein extracts were electrophoresed in 10% SDS-polyacrylamide separating gels as described by Laemmli (1970). The proteins were blotted onto nylon membranes, blocked with 5% (w/v) nonfat powdered milk, reacted with antisera raised against parsley 4CL (Ragg et al., 1981), reacted with goat anti-rabbit IgG conjugated to alkaline phosphatase, (GIBCO-BRL), and then visualized using Fast-Red R and Naphtol AS-MX phosphate as substrates (Sigma).

RESULTS

Cloning and Characterization of 4CL cDNAs

A young shoot tip cDNA library made from *N. tabacum* was screened using a potato *4CL* cDNA clone (Becker-André et al., 1991) as a hybridization probe. Eleven clones were identified, which, by preliminary sequence analysis, encoded 4CL. Sequencing and restriction digests showed that the 11 cDNAs could be placed into four groups. Each group contained cDNAs that were apparently identical except in length and thus represented copies of four different *4CL* genes or alleles. Detailed restriction maps of clones Nt4CL-1, Nt4CL-5, Nt4CL-17, and Nt4CL-19, the longest representatives of each group, are illustrated in Figure 1A. These restriction maps showed that the four cDNAs fell into two classes. One class (Nt4CL-1 and Nt4CL-17) had distinctive *Pst*I and *Kpn*I sites, and the other class (Nt4CL-19 and Nt4CL-5) had *Ava*II, *Bam*HI, and *Sac*I sites. Sequencing of the 5′ region of these clones showed that restriction site polymorphisms between Nt4CL-17 and Nt4CL-1 and between Nt4CL-19 and Nt4CL-5 were due to single base-pair changes. In Figure 1B, 3′ sequences of the four cDNAs are compared. These comparisons demonstrate that the four cDNAs again fall into two classes, with Nt4CL-1 and Nt4CL-17 in one class (94% sequence identity) and Nt4CL-19 and Nt4CL-5 in a second class (97% sequence identity). In contrast, the 3′ sequences of Nt4CL-1 and Nt4CL-19 are only 67% identical; comparison of any other pair of clones between the two classes gives similar results (<70% sequence identity, not shown). It should be noted that, although Nt4CL-19 and Nt4CL-5 are 97% identical at the 3′ end, there is a 36-bp region within Nt4CL-5 that is absent in Nt4CL-19.

Cytogenetic (Smith, 1968), biochemical (Gray et al., 1974; Obokata et al., 1990; Takahashi et al., 1991), and, more recently, molecular (van Buuren et al., 1992; Hua et al., 1993; Kronenberger et al., 1993; Pellegrini et al., 1993) studies strongly suggest that the allotetraploid *N. tabacum* (tobacco) arose from the hybridization of *N. sylvestris* and *N. tomentosiformis*, followed by chromosome duplication. Thus, the heterogeneity in tobacco *4CL* cDNA clones can be explained in two ways. The two classes of cDNAs (Nt4CL-1/Nt4CL-17 and Nt4CL-5/Nt4CL-19) could represent distinct *N. sylvestris* and *N. tomentosiformis* *4CL* genes, brought together in the hybrid tobacco genome, with the heteroge-

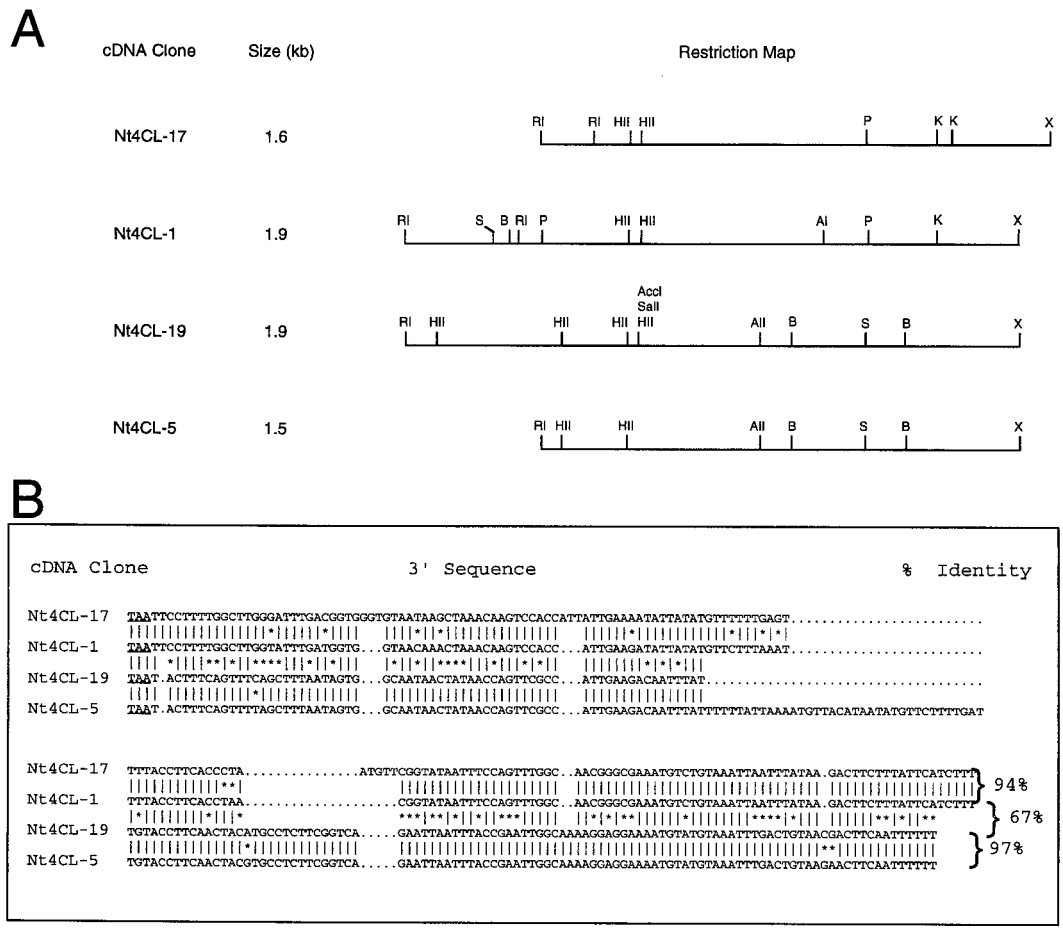

Figure 1. Structure of tobacco *4CL* cDNA clones. A, Restriction maps of representative clones. AI, *Ava*I; AII, *Ava*II; B, *Bam*HI; HII, *Hin*cII; K, *Kpn*I; P, *Pst*I; RI, *Eco*RI; S, *Sac*I; X, *Xho*I. B, 3′ sequences of the cDNAs in A, starting at the translation stop codons (underlined). Vertical lines indicate identical nucleotides, asterisks indicate different nucleotides, and dots indicate gaps. All sequences terminate with a poly(A) tail (not shown) with the exception of Nt4CL-17, which extends for an additional 136 bp, and Nt4CL-19, which extends for an additional six nucleotides prior to the polyadenylation. The percentage of sequence identity between pairs of clones is shown at the bottom right. GenBank accession nos. for Nt4CL-1 and Nt4CL-19 are U50845 and U50846, respectively.

neity within each class (e.g. Nt4CL-1 versus Nt4CL-17) representing allelic differences. Alternatively, the two classes could represent two divergent *4CL* gene family members (*4CL1* and *4CL2*), each of which was present in the two progenitors of tobacco. In this case, the heterogeneity within each cDNA class could be explained by the presence of two *4CL1* and two *4CL2* genes (one each from each of the progenitors) in the hybrid genome of tobacco. For example, Nt4CL-1 and Nt4CL-17 could represent two *4CL1* genes, one from *N. sylvestris* and one from *N. tomentosiformis*.

Clones Nt4CL-1 and Nt4CL-19 were used in subsequent experiments since they were the longest cDNAs of the two classes cloned, and initial sequence analysis indicated the presence of MET initiation codons. The top and bottom strands of both clones were completely sequenced, and the nucleotide and predicted amino acid sequences of the single open reading frames from each clone were compared with each other and with other *4CL* sequences available in the databases (Table I). The Nt4CL-1 and Nt4CL-19 nucleotide sequences were 82% identical with each other and were most similar to potato *4CL*, followed by decreasing identity to the parsley, soybean, Arabidopsis, pine, and rice *4CL* sequences. The sequence of Nt4CL-1 was more similar to the potato *4CL* sequence (88%) than to the Nt4CL-19 sequence (82%). Nt4CL-1 and Nt4CL-19 cDNA sequences had open reading frames predicted to encode proteins of 547 and 542 amino acids, with molecular masses of 59.8 and 59.4 kD and pIs of 5.40 and 5.69, respectively (results not shown). The predicted amino acid identity between the two cDNA clones was 81% (Table I). Both predicted peptides (results not shown) contained the conserved GEICIRG amino acid motif believed to be essential for 4CL activity (Becker-André et al., 1991), an AMP-binding signature (Schröder, 1989), and five other conserved Cys residues also found in other cloned *4CL*s (Uhlmann and Ebel, 1993). These results indicate that the cDNAs encode two divergent forms of the 4CL enzyme and are likely to have originated from two different *4CL* genes. We ini-

Table I. *Comparison of Nt4CL-1 and Nt4CL-19 nucleotide and predicted amino acid sequences to each other and to other 4CL sequences*

cDNA	Nt4CL-1			Nt4CL-19			Ref.
	DNA identity	Amino acid identity	Amino acid similarity	DNA identity	Amino acid identity	Amino acid similarity	
	%	%	%	%	%	%	
Nt4CL-1[a]	–[b]	–	–	82	81	89	This paper
Nt4CL-19[c]	82	81	89	–	–	–	This paper
Potato 4CL1	88	86	92	80	80	88	Becker-André et al., 1991
Parsley 4CL1	75	73	83	76	72	81	Lozoya et al., 1988
Soybean 4CL14	73[d]	73	83	73[d]	73	83	Uhlmann and Ebel, 1993
Arabidopsis 4CL	67	66	80	70	69	82	Lee et al., 1995
Soybean 4CL16	66[d]	62	80	69[d]	65	79	Uhlmann and Ebel, 1993
Pine 4CL	65	63	79	66	65	80	Voo et al., 1995
Rice 4CL	66[d]	59	80	61[d]	61	77	Zhao et al., 1990

[a] GenBank accession no. U50845. [b] –, Comparison not applicable. [c] GenBank accession no. U50846. [d] Comparisons are based on partial cDNA sequences.

tially designated the tobacco genes from which Nt4CL-1 and Nt4CL-19 originated as *4CL1* and *4CL2*, respectively.

Inheritance of *4CL* Genes

We used Southern blots to determine whether *4CL* genes corresponding to the two divergent cDNAs (*4CL1* and *4CL2*) were both present in the progenitors of tobacco and to assess the size of the potential *4CL* gene family in tobacco. In a preliminary experiment, restriction enzyme-digested Nt4CL-1 and Nt4CL-19 cDNAs were blotted and hybridized separately to probes generated from the insert of each plasmid. Figure 2A shows that at moderate stringency (0.5× SSC), little cross-hybridization was observed,

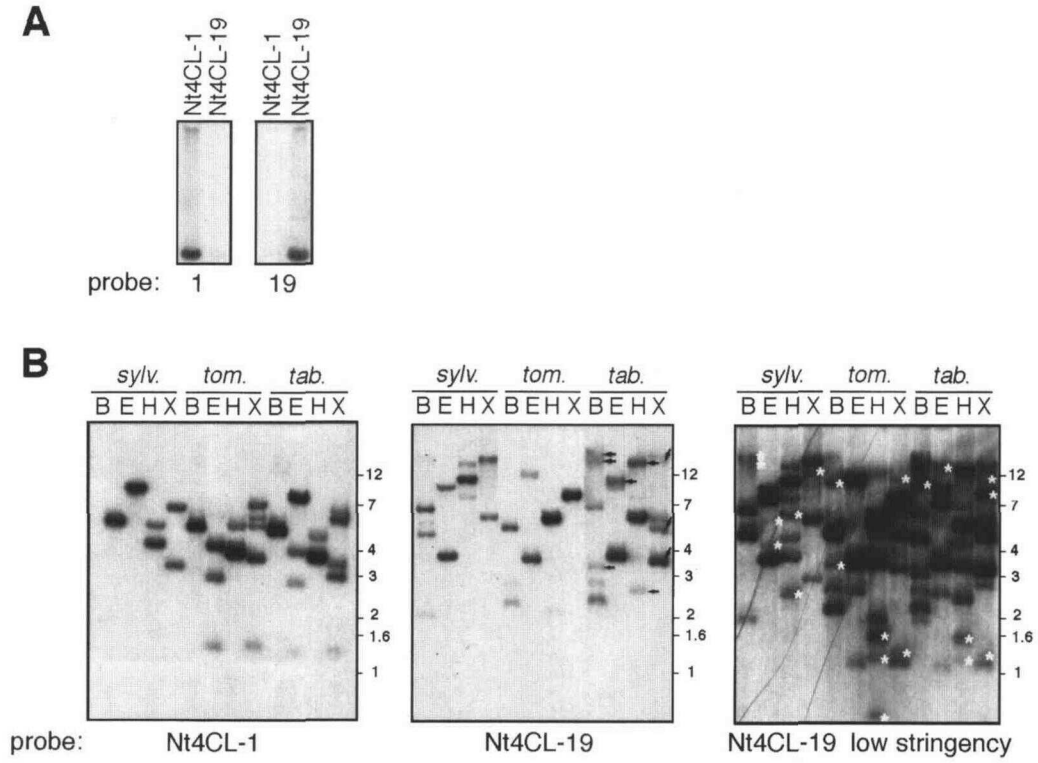

Figure 2. Inheritance of *4CL* genes in *Nicotiana* species. A, Southern blot analysis of restriction enzyme digests of Nt4CL-1 and Nt4CL-19 plasmids hybridized to either Nt4CL-1 or Nt4CL-19 probes (1 or 19), washed at moderate stringency (0.5× SSC). B, Southern blots of genomic DNA (10 μg) from *N. sylvestris* (*sylv.*), *N. tomentosiformis* (*tom.*), and *N. tabacum* (*tab.*) were digested with different restriction enzymes. A single blot was hybridized sequentially with Nt4CL-1 and Nt4CL-19 probes and washed at high stringency (0.2× SSC). Arrows indicate hybridization restriction fragments that are unique to *N. tabacum* and not found in the progenitors. The same blot was re-hybridized to an Nt4CL-19 probe and washed at low stringency (Nt4CL-19 low stringency). Asterisks indicate hybridizing restriction fragments not detected after high-stringency washes. Migration of size standards (in kb) are shown on the right. B, *Bgl*II; E, *Eco*RV; H, *Hin*dIII; and X, *Xba*I.

demonstrating that the cDNAs could be used as gene-specific probes to detect the presence of *4CL1* and *4CL2* sequences in tobacco and its progenitors.

Genomic Southern blots were prepared with DNA from *N. sylvestris*, *N. tomentosiformis*, and *N. tabacum* and then hybridized sequentially to Nt4CL-1 and Nt4CL-19 probes. Figure 2B shows that both probes hybridized to restriction fragments from the progenitors and from tobacco and that each probe hybridized to a unique set of restriction fragments in the three species. This demonstrates that at least two divergent *4CL* genes, represented by the Nt4CL-1 and Nt4CL-19 cDNA clones (*4CL1* and *4CL2*), are both present in the tobacco progenitors as well as in tobacco itself.

The small number of fragments that hybridized to each probe in *N. sylvestris* and *N. tomentosiformis* suggests that there are one to two tobacco *4CL1*- and *4CL2*-like genes in these progenitor species. The DNA hybridization patterns in *N. tabacum* were more complex, as would be predicted from the acquisition of a set of *4CL1*- and *4CL2*-like genes from each progenitor. Indeed, Nt4CL-1-hybridizing (*4CL1*) restriction fragments in *N. tabacum* can be traced directly from the restriction fragments of the parental species (Fig. 2B). For example, a single 9-kb *Eco*RV *4CL1* fragment was observed in the *N. sylvestris* genome, three *Eco*RV fragments (approximately 1.3, 3, and 4 kb in size) were observed in *N. tomentosiformis*, and all four *Eco*RV fragments were observed in the *N. tabacum* genome. Nt4CL-19-hybridizing (*4CL2*) fragments in *N. tabacum* also appeared to be derived from a combination of the *N. sylvestris* and *N. tomentosiformis 4CL2* gene complements, but about one-half of the *N. tabacum 4CL2* restriction fragments were polymorphic with respect to fragments in the progenitor species (Fig. 2B, arrows). In summary, the results of the Southern blots suggest that the allotetraploid tobacco genome contains at least two *4CL1* genes, contributed by *N. sylvestris* and *N. tomentosiformis* and represented by Nt4CL-1 and Nt4CL-17, and at least two *4CL2* genes, likewise contributed by *N. sylvestris* and *N. tomentosiformis* and represented by Nt4CL-5 and Nt4CL-19.

Restriction fragments specific to both *4CL1* and *4CL2* were detected when the genomic Southern blot was hybridized to Nt4CL-19 and washed at low stringency (2× SSC; Fig. 2B). In addition, however, additional hybridizing restriction fragments were revealed (Fig. 2B, asterisks), which were not apparent after high-stringency washes with either probe. These bands may represent additional *4CL* genes divergent from *4CL1* and *4CL2*.

Expression of *4CL1* and *4CL2*

To investigate the inducible and developmentally regulated expression of *4CL1* and *4CL2*, duplicate northern blots were prepared using RNA isolated from wounded or methyl jasmonate-treated tobacco leaves and from various tobacco organs. Blots were hybridized to Nt4CL-1 or Nt4CL-19 probes under conditions that allowed little cross-hybridization between probes (Fig. 2A). Subsequent stripping and re-hybridization of the blots to a pea rRNA gene probe confirmed that RNA loading between samples and between gels was similar. Figure 3 shows that the expression of both classes of *4CL* genes was inducible by wounding and methyl jasmonate treatment. In this and similar experiments, however, *4CL2* was less responsive to methyl jasmonate treatment than *4CL1* but was approximately equivalent in its response to wounding. The steady-state levels of both *4CL1* and *4CL2* mRNA transcripts were highest in old stems, followed by young stems and ovaries. RNA levels were lowest in shoot tips, the pigmented limb of the petals, and untreated mature leaves (Fig. 3). No major differences in the expression of *4CL1* and *4CL2* were observed, although *4CL2* RNA was reproducibly less abundant in young stems and ovaries relative to *4CL1*.

The low expression of *4CL* in the pigmented limb of tobacco petals was unexpected since transgenic studies using the parsley *4CL* promoter, bean *PAL* and *CHS* promoters, and a tobacco *CHS* promoter fused to the *GUS* reporter gene showed high GUS activity in petal limbs (Bevan et al., 1989; Schmid et al., 1990; Hauffe et al., 1991; Drews et al., 1992). This led us to examine *4CL* expression in petals more closely. Tobacco flowers from developmental stages 1 through 6 (the corolla is just emerging from the sepals at stage 1, pigmentation of limb tissue is detectable at stage 4, and anthers are dehiscing at stage 6; Reinold et al., 1993) were harvested, and RNA was isolated from the sepals and the base, tube, and limb sectors of the corolla (see "Materials and Methods"). Figure 4 shows a northern blot of this RNA hybridized to an Nt4CL-19 probe. *4CL2* RNA accumulation was highest in the unpigmented portions of the petals (tube and base), where its accumulation was temporally regulated. High transcript levels were first evident in stage 4 flowers, and *4CL2* RNA had decreased significantly in stage 6 flowers. An identical pattern of

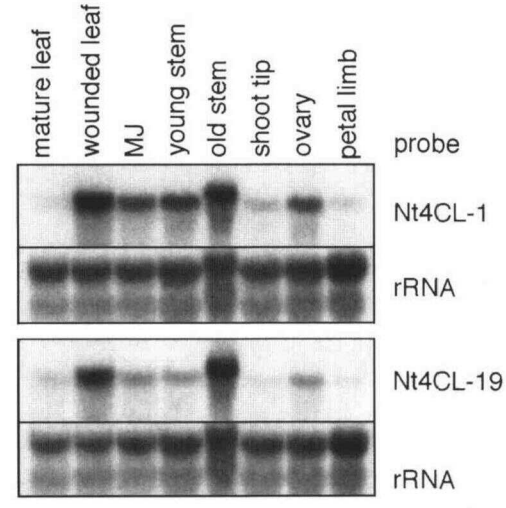

Figure 3. Northern blot analysis of *4CL* RNA levels in tobacco. Duplicate samples of total RNA (10 μg) from various organs were separated on formaldehyde gels, blotted onto nylon membranes, hybridized to Nt4CL-1 or Nt4CL-19 probes, and washed at moderate stringency (0.5× SSC). The blots were subsequently stripped and hybridized to rRNA probes to demonstrate evenness of loading. RNA was isolated from fully expanded leaves (mature leaf), fully expanded leaves that had been wounded (wounded leaf), leaves treated with methyl jasmonate (MJ), and other tissues as indicated.

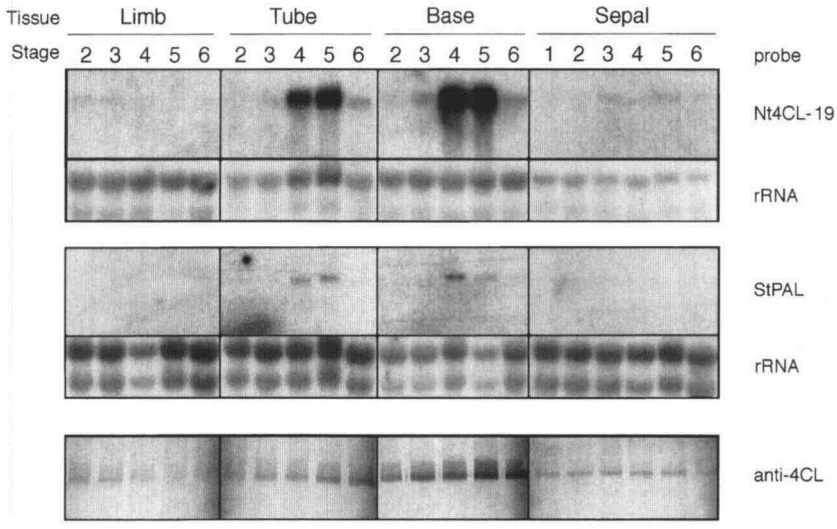

Figure 4. Steady-state *4CL* and *PAL* RNA and 4CL protein levels in tissues from developing petals and sepals. Tobacco flowers in developmental stages 1 through 6 (Reinold et al., 1993) were harvested and dissected into four sectors: the limb, the tube, the base, and the sepals (see "Materials and Methods"). Duplicate northern blots prepared using 10 μg of total RNA from these samples were hybridized to Nt4CL-19 and potato PAL (StPAL) probes. The blots were subsequently stripped and re-hybridized to rRNA probes to demonstrate evenness of loading. The bottom panel shows an immunoblot analysis of protein extracts from the petal and sepal samples described above. The blot was reacted with a polyclonal antibody raised against parsley 4CL.

expression was observed when blots were hybridized to an Nt4CL-1 probe (results not shown). To determine whether *PAL* gene expression was coordinately regulated with *4CL* in petals, a northern blot was hybridized to a potato *PAL* probe (Fig. 4, StPAL). Although the heterologous probe did not hybridize strongly, an identical pattern of *PAL* expression was observed, with high expression in tube and base segments of stage 4 and stage 5 flowers.

We used an anti-parsley 4CL antiserum to detect tobacco 4CL proteins on western blots to determine whether tobacco 4CL proteins accumulate specifically in the unpigmented petal regions of high *4CL* transcript accumulation. This 4CL antiserum is highly specific to 4CL in extracts derived from parsley (Ragg et al., 1981; Lozoya et al., 1988) and reacts with single bands corresponding to the molecular weight of 4CL on SDS-PAGE gels of crude protein extracts from tobacco, Arabidopsis, and poplar (Fig. 5, lane 4; D. Lee and C. Douglas, unpublished data). Western blot analysis of tobacco petal and sepal protein extracts showed

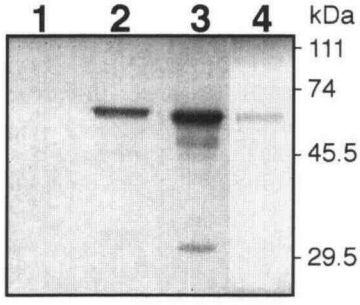

Figure 5. Immunoblot analysis of recombinant tobacco 4CL expressed in *E. coli*. Crude bacterial protein extracts (5 μg) were separated by SDS-PAGE and blotted onto a nylon membrane, and the blot was reacted with an antibody raised against parsley 4CL. Lane 1, Extract from bacteria harboring the expression plasmid pQE-30 without an insert; lane 2, extract from bacteria harboring pQE-1, expressing Nt4CL-1-encoded protein (4CL1); lane 3 extract from bacteria harboring pQE-19, expressing Nt4CL-19-encoded protein (4CL2); and lane 4, extract of total protein (25 μg) from tobacco old stem. Molecular mass standards (in kD) are shown on the right.

that, as with *4CL* RNA levels, 4CL protein was most abundant in the tube and base segments of tobacco petals by stages 5 and 6 (Fig. 4, anti-4CL), consistent with the increased accumulation of 4CL protein in these areas subsequent to RNA accumulation. However, the difference in accumulation of 4CL protein in the petal tube and base relative to that in limbs and sepals was not as pronounced as the differences in *4CL* RNA accumulation, suggesting that 4CL protein accumulation could be controlled by factors other than transcript level. Also, 4CL protein was detected in limbs and sepals at all stages and in tubes and limbs at stages 2 and 3, although RNA was only weakly detectable in these organs. This suggests that 4CL proteins may be fairly stable and thus accumulate to detectable steady-state levels in the absence of high transcript levels in these developing floral organs. Alternatively, the 4CL antiserum could detect the products of divergent *4CL* genes expressed in these organs, whose transcripts were not detected in RNA blots. Together, these results show that during flower development phenylpropanoid gene expression is very active in the unpigmented portions of tobacco petals and appears to be much higher than in pigmented petal limbs.

Enzyme Activities of Recombinant 4CL Proteins

The 4CL1 and 4CL2 proteins encoded by Nt4CL-1 and Nt4CL-19 were expressed in *E. coli* to examine the biochemical characteristics of the two proteins. Two plasmids, designated pQE-1 and pQE-19, were constructed in which the coding portions of the Nt4CL-1 and Nt4CL-19, respectively, were fused to N-terminal His tags in expression plasmids pQE-32 and pQE-30 (QiaGen). Figure 5 shows that the recombinant proteins produced in bacterial extracts were approximately 60 kD, reacted with the antibody raised against parsley 4CL, and migrated to the same location in an SDS-polyacrylamide gel as the 4CL found in tobacco stem extracts. pQE-1-generated protein (4CL1) was localized in the cytosol, whereas pQE-19-generated protein (4CL2) was present in both cytosolic and membrane-bound

fractions (results not shown). No 4CL protein (Fig. 5, lane 1) or 4CL activity (not shown) was detected in a bacterial strain containing the pQE-30 plasmid alone.

The two recombinant proteins were tested for their relative abilities to utilize differently substituted hydroxycinnamic acids as substrates. Figure 6 shows that bacterial extracts from the pQE-1 strain expressing 4CL1 exhibited relative 4CL activities of 100, 21, 17, and 73% toward 4-coumarate, cinnamate, caffeate, and ferulate, respectively, and extracts of the pQE-19 strain expressing 4CL2 had relative activities of 100, 29, 25, and 62%, respectively. Neither recombinant protein had detectable activity toward sinapate as a substrate. In parallel, crude extracts prepared from tobacco stems were assayed for 4CL activity and were found to have relative activities of 100, 17, and 60 toward *p*-coumarate, caffeate, and ferulate, respectively; no detectable activity toward cinnamate; and a small amount of activity (4%) toward sinapate (Fig. 6). Taken together, these results show that the two recombinant 4CL proteins have nearly identical substrate specificities compared with that of 4CL in tobacco extracts. However, the lack of activity toward sinapate and relatively high activity toward cinnamate, not generally regarded as a cellular target for the 4CL enzyme, distinguished the recombinant enzymes from the native 4CL protein(s) in stem extracts.

Since the activity of the recombinant proteins toward cinnamate seemed particularly enigmatic, we considered possible explanations for this finding. Conceivably, the bacterial extracts that were used to assay recombinant 4CL activity may have contained an activity capable of modifying cinnamate (e.g. by hydroxylation), leading to an apparent activity toward this substrate. This was tested by assaying 4CL activity toward cinnamate in tobacco extracts to which crude extracts of the bacterial strain harboring the empty expression plasmid (pQE-30) had been added. No 4CL activity toward cinnamate was detectable, indicating that the bacterial extract itself was not the cause of the apparent 4CL activity toward cinnamate (results not shown).

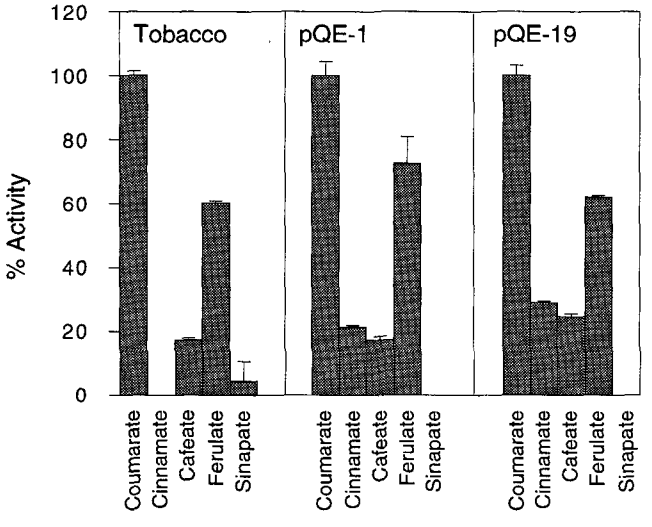

Figure 6. Substrate specificities of tobacco and recombinant 4CL proteins. 4CL enzyme activity was measured from crude tobacco stem extracts (Tobacco), crude bacterial extracts expressing Nt4CL-1 (pQE-1), and crude bacterial extracts expressing Nt4CL-19 (pQE-19). 4CL activity is expressed as a percentage of the activity of the preparation using 4-coumarate as a substrate. Results are averaged from three determinations using the hydroxycinnamate substrates (0.2 mM). Error bars represent SDs. The average specific enzyme activities using coumarate as a substrate, which were taken as 100%, are as follows: 0.326 mkat kg^{-1} for the tobacco stem extract, 22.9 mkat kg^{-1} for the pQE-1 extract, and 25.1 mkat kg^{-1} for the pQE-19 extract.

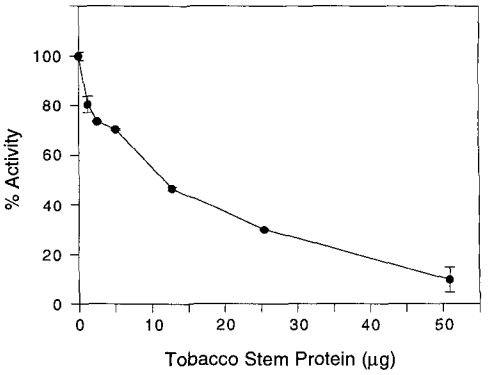

Figure 7. Effect of tobacco stem extracts on recombinant 4CL activity toward cinnamate. Recombinant bacterial extracts expressing Nt4CL-19 (pQE-19) were assayed for 4CL enzyme activity against 0.2 mM cinnamate in the presence of different amounts of a tobacco stem extract. Activity is expressed as a percentage of the 4CL activity toward cinnamate in the absence of added tobacco extract. Results are averaged from three determinations; error bars represent SDs. The specific enzyme activity using cinnamate as a substrate in the absence of added tobacco extract, which was taken as 100%, was 5.24 mkat kg^{-1}.

We next considered the possibility that recombinant 4CL protein has bona fide activity toward cinnamate but that the native tobacco protein is posttranslationally modified or interacts with other proteins, repressing its activity toward cinnamate in plant extracts. To test this, crude tobacco extracts were incubated together with crude bacterial extracts containing recombinant 4CL1 or 4CL2, and the mixture was assayed for activity toward cinnamate. Figure 7 shows that the activity of the pQE-19 (*4CL2*) extract toward cinnamate decreased exponentially in the presence of increasing amounts of tobacco extract, consistent with the hypothesis that such extracts indeed contain a component capable of modifying the substrate specificity of recombinant 4CL.

Furthermore, as shown in Figure 8, this apparent modification was specific to the activity toward cinnamate, and no decrease in recombinant 4CL2 activity toward 4-coumarate, caffeate, or ferulate was observed following incubation with the tobacco extract. Boiled tobacco extract was no longer capable of modifying the activity of the recombinant 4CL (Fig. 8). In contrast, desalting the tobacco extract by passage through a G-50 Sephadex column (Pharmacia) did not abolish its ability to modify recombinant 4CL activity toward cinnamate (Fig. 8). Neither plant extraction buffer alone (Fig. 8) nor BSA (not shown) decreased the recombinant 4CL activity toward cinnamate.

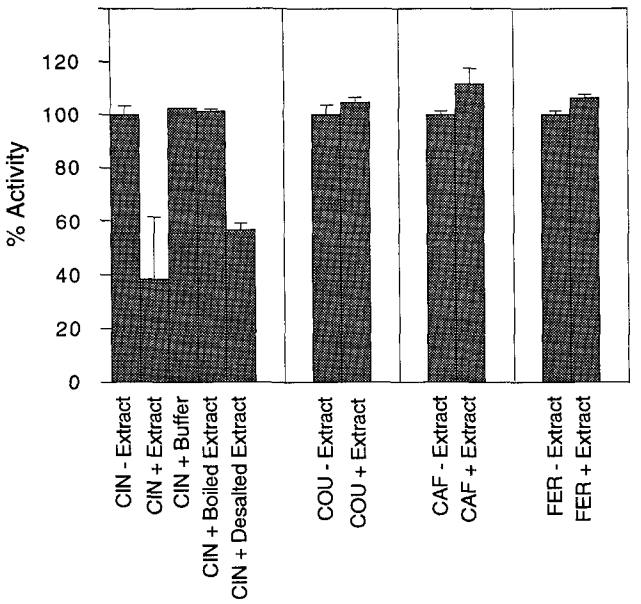

Figure 8. Characterization of the 4CL-modifying activity in tobacco stem extracts. Recombinant 4CL (pQE-19) was assayed for 4CL activity toward cinnamate (CIN), 4-coumarate (COU), caffeate (CAF), or ferulate (FER) in the absence (− Extract) or presence (+ Extract) of 25 μg of stem extract. The stem extracts used in assays with cinnamate as a substrate were untreated (CIN + Extract), heated at 100°C for 15 min (CIN + Boiled Extract), or passed through a Sephadex G-50 column (CIN + Desalted Extract) before being added to the enzyme assays. Activity is expressed as a percentage of the 4CL activity toward each hydroxycinnamic acid in the absence of added tobacco extract. Results are averages of three determinations; error bars represent the SDs. The average specific enzyme activities for the four substrates in the absence of added extracts or buffer, which were taken as 100%, are as follows: 6.2 mkat kg^{-1} for cinnamate, 20.8 mkat kg^{-1} for coumarate, 4.8 mkat kg^{-1} for caffeate, and 12.5 mkat kg^{-1} for ferulate.

Although these experiments were performed using pQE-19 extracts containing recombinant 4CL2 protein, a similar effect of tobacco extracts on 4CL1 in extracts of the pQE-1 *E. coli* strain was also observed (results not shown). Taken together, these results suggest that tobacco stem extracts contain a large heat-labile component, possibly a protein, which is capable of modifying the substrate specificity of recombinant 4CL.

DISCUSSION

The data presented in this paper show that 4CL is encoded by a gene family in tobacco, an allotetraploid hybrid. Genomic Southern blot analysis (Fig. 2B) demonstrated the presence of tobacco *4CL* gene family members in *N. sylvestris* and *N. tomentosiformis*, the *Nicotiana* species most closely related to the presumed progenitors of tobacco, strongly suggesting that *4CL* gene duplication and divergence had occurred in *Nicotiana* prior to the emergence of these species. The sequences of full-length cDNA clones specific to two of these genes, which we call *4CL1* and *4CL2*, show that they are relatively divergent, sharing about 80% nucleotide and amino acid sequence identity. Consistent with the allotetraploid nature of the *N. tabacum* genome, genomic Southern blots (Fig. 2) show that the tobacco *4CL* gene family contains *4CL1* and *4CL2* copies from both parental species, for a total of at least four genes. This is further supported by the existence of two classes of polymorphic cDNA clones, with each class containing two cDNA types (Fig. 1). It seems likely that Nt4CL-1/Nt4CL-17 and Nt4CL-19/Nt4CL-5 represent *4CL1* and *4CL2* genes, respectively, that originated from the two parental species. Similar evolutionary schemes have been proposed for the origins of tobacco gene families encoding nitrite reductase (Kronenberger et al., 1993), *psbP* (Hua et al., 1993), and endochitinases (van Buuren et al., 1992). We conclude that tobacco, like soybean (Uhlmann and Ebel, 1993) and poplar (Allina and Douglas, 1994; S. Allina, A. Pri-Hadash, B. Ellis, and C. Douglas, unpublished results), contains a family of divergent *4CL* genes. In contrast, parsley (Lozoya et al., 1988), potato (Becker-André et al., 1991), pine (Voo et al., 1995), and Arabidopsis (Lee et al., 1995) apparently contain single *4CL* genes or very similar duplicated genes.

The tobacco *4CL* nucleotide and amino acid sequences deduced from cDNA clones Nt4CL-1 and Nt4CL-19 were most similar to potato *4CL*, which is in the same order (Polemoniales) and family (Solanaceae) as tobacco. Tobacco sequences are increasingly divergent from the *4CL* sequences in plants of the orders Umbellales (parsley), Fabales (soybean), Papaverales (Arabidopsis), Pinales (pine), and Poales (rice) (Table I). Although there are only a few *4CL* sequences available, this pattern of sequence divergence is consistent with proposed taxonomic relationships between these groups (Heywood, 1978). It is interesting that the Nt4CL-1 (*4CL1*) cDNA and amino acid sequences are more similar to potato *4CL* than to Nt4CL-19 (*4CL2*), which suggests that *4CL1* may be homologous to an ancestral *4CL* gene from which all solanaceous *4CL* genes (including potato *4CL*) evolved, whereas *4CL2* represents a *4CL* gene that was duplicated in the *Nicotiana* evolutionary line and has undergone greater divergence from the ancestral *4CL*. *4CL1* restriction fragments are conserved in *N. tabacum* with respect to its proposed progenitors *N. sylvestris* and *N. tomentosiformis* (Fig. 2B), whereas many of the *4CL2*-specific restriction fragments are polymorphic (Fig. 2B, arrows). This is consistent with a more rapid divergence of the *4CL2* locus than the *4CL1* locus subsequent to the hybridization event giving rise to tobacco, an event thought to have occurred within the last 6 million years (Okamuro and Goldberg, 1985).

Under low-stringency washes, the Nt4CL-19 probe hybridized to numerous restriction fragments in tobacco, and its progenitors, which were distinct from Nt4CL-1 and Nt4CL-19 fragments, were observed at high stringency (Fig. 2B, asterisks). These DNA sequences may represent a third class of *4CL* genes, distinct from *4CL1* and *4CL2*. Two partial-length tobacco *4CL* cDNA sequences have recently been deposited in the EMBL data bank. The sequence of one clone, TOBTCL2 (accession no. D50033), is 99% identical with Nt4CL-1 at the nucleotide level. In contrast, clone TOBTCL6 (accession no.

D50034) is only 75 and 74% identical with Nt4CL-1 and Nt4CL-19, respectively. Thus, TOBTCL6 may represent a member of the third most highly divergent class of tobacco *4CL* that was detected under low-stringency hybridization.

Previous studies in which a heterologous probe was utilized have shown that tobacco *4CL* genes are activated by wounding and methyl jasmonate treatment (Douglas et al., 1991; Ellard-Ivey and Douglas, 1996) and are expressed in specific cell types in developing tobacco flowtl;1ers (Reinold et al., 1993). Expression of heterologous *PAL* and *4CL* gene promoter-*GUS* fusions in transgenic tobacco suggests that these genes are actively expressed in pigmented limbs of the flowers, developing primary and secondary xylem, and root tips (Bevan et al., 1989; Liang et al., 1989; Hauffe et al., 1991). These sites of expression are correlated with the sites of accumulation of phenylpropanoid products such as lignin and anthocyanins. The results of northern blots (Fig. 3) on which Nt4CL-1 (*4CL1*) and Nt4CL-19 (*4CL2*) probes were used are generally consistent with the patterns of *4CL* and *PAL* expression in tobacco deduced from the experiments cited above. Highest expression was observed in stems, ovaries, and wounded or methyl jasmonate-treated leaves. Although these experiments provide little evidence for differential expression of the two genes, Nt4CL-19 (*4CL2*) mRNA accumulation appeared to be lower than that of Nt4CL-1 (*4CL1*) in ovaries, young stems, and methyl jasmonate-treated leaves, relative to the accumulation of RNA specific to the two genes in old stem (Fig. 3). Although these differences were reproducible, the significance of the 2- to 3-fold lower (*4CL2*) levels is unclear. Since the ovary is a complex organ, and *4CL* is expressed in ovules, in the carpel wall, and in nectaries (Hauffe et al., 1991; Reinold et al., 1993), we cannot exclude the possibility that the two genes are differentially regulated at the tissue level in this organ. Further experiments using gene-specific probes and in situ hybridization will be necessary to resolve this question.

The high levels of *4CL1* and *4CL2* expression in the unpigmented corolla tube and base portions of tobacco flowers and the very weak expression in the pigmented limb (Fig. 4) are in striking contrast to the high expression of heterologous *PAL-GUS*, *CHS-GUS*, and *4CL-GUS* fusions specifically in the limb portions of transgenic tobacco flowers reported in several studies (Bevan et al., 1989; Liang et al., 1989; Schmid et al., 1990, Hauffe et al., 1991) and to the accumulation of tobacco *CHS* transcripts and expression of a tobacco *CHS-GUS* fusion specifically in epidermal cells of the limb where anthocyanin pigments accumulate (Drews et al., 1992). However, the temporally regulated expression of *4CL* in the petal tube and base is consistent with previous results (Reinold et al., 1993) in which in situ hybridization showed that tobacco *4CL* mRNA, as well as mRNA specific to a parsley *4CL* transgene, accumulates to high levels in epidermal and mesophyll cell layers of the petal base in developing tobacco flowers. Furthermore, the results in Figure 4 show that tobacco *PAL* mRNA detected using a heterologous potato *PAL* probe accumulates in a similar manner and that 4CL protein accumulates to higher levels in the tube and base than in the limb. These results clearly indicate that phenylpropanoid metabolism is strongly activated in the unpigmented petal during flower development.

Possible functions for the high activity of 4CL and PAL in the unpigmented tube and base regions of the petal are in the biosynthesis of the colorless flavonols kaempferol and quercetin, which accumulate in tobacco flower petals (Holton et al., 1993) and in the white petals of plants such as Arabidopsis (Shirley et al., 1995). Recent observations indicate that colorless flavonoids are abundant in the base and tube regions of tobacco flowers (Reinold, 1995) and we note that in petunia, *CHS* expression in unpigmented petal cells may be correlated with the accumulation of uncolored flavonoids (Koes et al., 1990). However, biosynthesis of flavonoids in unpigmented tobacco petal regions would require the expression of *CHS* and other flavonoid biosynthetic genes. How can the pattern of *CHS* expression in tobacco petals (high in petal limb, low in tube and base) be reconciled with the pattern of *4CL* expression (high in tube and base, low in limb) reported here and with the apparent accumulation of flavonoids in both the unpigmented and pigmented portions of tobacco petals? One possible explanation is that uncharacterized divergent *4CL* or *CHS* gene family members are expressed specifically in unpigmented petal cells and pigmented limb cells. As discussed above, there is evidence for divergent tobacco *4CL* gene family members in addition to *4CL1* and *4CL2* that could be expressed in the limb, and to our knowledge, the extent of the tobacco *CHS* gene family is unknown. Since petal epidermal cells, which accumulate anthocyanin pigments, are highly specialized morphologically (Drews et al., 1992; Noda et al., 1994), distinct sets of phenylpropanoid gene family members may be expressed in these cells. A second possibility is that, in the unpigmented corolla tube, high *PAL* and *4CL* expression is related to the biosynthesis of phenylpropanoid products not requiring CHS activity, in addition to flavonoids. The nature of such possible secondary metabolites is unknown and deserves investigation.

Given the high *4CL* and *PAL* expression in unpigmented portions of tobacco petals, why is the activity of phenylpropanoid gene promoter-*GUS* fusions transferred into tobacco apparently much higher in the pigmented limb than in the white tube of the tobacco flowers corolla? The heterologous *PAL*, *4CL*, and *CHS* promoters studied in some investigations may lack the necessary *cis* elements to confer expression in the unpigmented regions of tobacco petals and thus not truly report the activity of the respective tobacco genes. Alternatively, there may be peculiarities of the *GUS* reporter gene that affect its expression in unpigmented cells. Since the *GUS* gene has been shown to contain cryptic elements that activate ectopic *GUS* expression (Uknes et al., 1993), a cryptic negative element may be present that precludes expression in other cells.

The expression of recombinant 4CL1 and 4CL2 proteins in *E. coli* allowed us to ask whether these divergent 4CL isoforms have distinct abilities to utilize differently substituted hydroxycinnamic acids as substrates, as has been

hypothesized for other plants (Knobloch and Hahlbrock, 1975; Grand et al., 1983). The recombinant 4CL proteins were most active toward 4-coumarate, followed by ferulate and caffeate, and the 4CL1 and 4CL2 patterns of substrate utilization were very similar to each other. Thus, it is unlikely that the two forms play a role in controlling the biosynthesis of different phenylpropanoids in tobacco. The relative activities of the recombinant tobacco 4CL proteins toward hydroxycinnamate substrates are quite similar to those of the two partially purified 4CL forms from parsley (Lozoya et al., 1988). In parsley, these two forms have similar substrate specificities and are encoded by two genes with predicted amino acid sequences that are 99.5% identical. In soybean, two divergent classes of *4CL* cDNAs (70% amino acid sequence identity) have been cloned. However, because the cDNAs were not full length, enzyme activities of the encoded proteins were not tested (Uhlmann and Ebel, 1993). In poplar, two significantly different *4CL* cDNAs have been cloned, and preliminary evidence suggests that they encode enzymes with indistinguishable activities (Allina and Douglas, 1994; S. Allina and C. Douglas, unpublished data). Thus, despite a number of studies in which 4CL forms exhibiting distinct substrate specificities were partially purified (Knobloch and Hahlbrock, 1975; Ranjeva et al., 1976; Wallis and Rhodes, 1977; Grand et al., 1983), to our knowledge, there is yet to be a documented example of recombinant 4CL isoforms having different substrate specificities.

The activities of recombinant 4CL1 and 4CL2 were distinct from the 4CL activity in crude tobacco stem extracts: both 4CL1 and 4CL2 had relatively high levels of activity toward cinnamate, whereas crude tobacco stem extracts lacked detectable activity toward this compound and had low but measurable activity toward sinapate. To explain the lack of sinapate activity of the recombinant enzymes, the existence of a third 4CL form present in stems with activity toward sinapate could be postulated. As discussed above, one or more additional classes of *4CL* genes may exist in tobacco that could account for this activity.

We found it more difficult to explain the lack of cinnamate activity in stem extracts, because the high levels of *4CL1* and *4CL2* expression in this organ (Fig. 3) would be predicted to specify activity toward cinnamate in stems. One formal possibility is that this unique property of the recombinant proteins is due to the presence of a His tag at the N terminus of the recombinant proteins, a consequence of the bacterial system we used to express *4CL1* and *4CL2*. However, we considered this unlikely, since initial analysis showed that several kinetic properties of the recombinant enzymes were similar to those of endogenous 4CL in tobacco extracts (results not shown). Also, unmodified recombinant poplar 4CL protein produced in a eukaryotic (baculovirus) expression system exhibited activity toward cinnamate, an activity not found in poplar extracts (A. Pri-Hadash, S. Allina, B. Ellis, and C. Douglas, unpublished data). Alternatively, results shown in Figure 7 suggest that the discrepancy between the recombinant and endogenous 4CL activities toward cinnamate may be attributed to the presence of a 4CL-modifying factor(s) present in stem extracts. The factor(s) specifically affects recombinant 4CL activity toward cinnamate, reducing it by 5- to 10-fold (Fig. 7), but has no effect on activity toward other substrates (Fig. 8). Thus, we hypothesize that the in vivo enzymatic properties of 4CL1 and 4CL2 in tobacco are partially determined posttranslationally by interaction with a factor or factors absent in the *E. coli* host used to produce the recombinant proteins.

What is the nature of the 4CL-modifying activity in tobacco stems? Since the activity depends on high-, but not low-, molecular-weight molecules and is heat labile (Fig. 8), our working hypothesis is that one or more proteins constitute the modifying activity. At least two mechanisms can be postulated by which activity of the 4CL protein could be modified by a protein factor or factors. First, the 4CL protein could be posttranslationally altered by phosphorylation or other modifications. Both Nt4CL-1 and Nt4CL-19 sequences have predicted phosphorylation sites (results not shown), and several of these are conserved in all known 4CL protein sequences. Although 4CL has never been implicated as a substrate for posttranslational modification, there have been some reports suggesting that PAL is modified posttranslationally (Bolwell and Rodgers, 1991; Bolwell, 1992). Second, 4CL activity could be altered by interactions with other proteins that form a multienzyme complex, including 4CL. There is a limited amount of evidence suggesting that enzymes of phenylpropanoid biosynthesis do form such complexes (Hrazdina and Wagner, 1985; Deshpande et al., 1993), and proteins that interact specifically with 4CL in a potential phenylpropanoid multienzyme complex would be predicted to be present in the tobacco stem extracts. Further experiments are in progress to elucidate the nature of the modifying factor or factors in tobacco stem extracts, to determine whether such activity is present in other plants, and to test the ability of the tobacco factor(s) to modify the activity of 4CL proteins from other plants.

ACKNOWLEDGMENTS

We thank Brian Ellis for helpful suggestions and comments regarding this work, Dave Neustaedter for insightful discussions, Elizabeth K. Molitor for critical reading of this manuscript, Stephen Lee for preparation of the tobacco cDNA library, and the Institut Expérimental du Tabac for providing *N. sylvestris* and *N. tomentosiformis* seeds.

Received March 8, 1996; accepted May 29, 1996.
Copyright Clearance Center: 0032-0889/96/112/0193/13.

LITERATURE CITED

Allina SM, Douglas CJ (1994) Isolation and characterization of the 4-coumarate:CoA ligase gene family in a poplar hybrid (abstract no. 852). Plant Physiol **105:** S-154

Becker-André M, Schulze-Lefert P, Hahlbrock K (1991) Structural comparison, modes of expression, and putative *cis*-acting elements of the two 4-coumarate:CoA ligase genes in potato. J Biol Chem **266:** 8551–8559

Bevan M, Shufflebottom D, Edwards K, Jefferson R, Schuch W (1989) Tissue- and cell-specific activity of a phenylalanine

ammonia-lyase promoter in transgenic plants. EMBO J **8:** 1899–1906

Bolwell GP (1992) A role for phosphorylation in the down-regulation of phenylalanine ammonia-lyase in suspension-cultured cells of French bean. Characterisation and differential expression of antigenic multiple M_r forms. Phytochemistry **31:** 4081–4086

Bolwell GP, Rodgers MW (1991) L-Phenylalanine ammonia-lyase from French bean (*Phaseolus vulgaris* L.). Biochem J **279:** 231–236

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principles of protein-dye binding. Anal Biochem **72:** 248–254

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

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

Douglas CJ, Hauffe KD, Ites-Morales M-E, Ellard M, Paszkowski U, Hahlbrock K, 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. EMBO J **10:** 1767–1775

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus **12:** 13–15

Drews GN, Beals TP, Bui AQ, Goldberg RB (1992) Regional and cell-specific gene expression patterns during petal development. Plant Cell **4:** 1383–1404

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

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

Gray JC, Kung SD, Wildman SG, Sheen SJ (1974) Origin of *Nicotiana tabacum* L. detected by polypeptide composition of fraction I protein. Nature **252:** 226–227

Hahlbrock K, Scheel D (1989) Physiology and molecular biology of phenylpropanoid metabolism. Annu Rev Plant Physiol Plant Mol Biol **40:** 347–369

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

Heywood VH (1978) Flowering Plants of the World. Mayflower Books, New York, pp 13–15

Holton TA, Brugliera F, Tanaka Y (1993) Cloning and expression of flavonol synthase from *Petunia hybrida*. Plant J **4:** 1003–1010

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

Hua S-B, Dube SK, Kung S-D (1993) Molecular evolutionary analysis of the *psbP* gene family of the photosystem II oxygen-evolving complex in *Nicotiana*. Genome **36:** 483–488

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

Knobloch K-H, Hahlbrock K (1975) Isoenzymes of *p*-coumarate: CoA ligase from cell suspension cultures of *Glycine max*. Eur J Biochem **52:** 311–320

Knobloch K-H, Hahlbrock K (1977) 4-Coumarate:CoA ligase from cell suspension cultures of *Petroselinum hortense* Hoffm. Partial purification, substrate specificity and further properties. Arch Biochem Biophys **184:** 237–248

Koes RE, van Blockland R, Quattrochio F, van Tunen AJ, Mol JNM (1990) Chalcone synthase promoters in petunia are active in pigmented and unpigmented cell types. Plant Cell **2:** 379–392

Kronenberger J, Lepingle A, Caboche M, Vaucheret H (1993) Cloning and expression of distinct nitrite reductases in tobacco leaves and roots. Mol Gen Genet **236:** 203–208

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

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

Liang X, Dron M, Schmid J, Dixon RA, Lamb CJ (1989) Developmental and environmental regulation of a phenylalanine ammonia-lyase-β-glucuronidase gene fusion in transgenic tobacco plants. Proc Natl Acad Sci USA **86:** 9284–9288

Logemann J, Schell J, Willmitzer L (1987) Improved method for the isolation of RNA from plant tissues. Anal Biochem **163:** 16–20

Lozoya E, Hoffmann H, Douglas C, Schulz W, Scheel D, Hahlbrock K (1988) Primary structures and catalytic properties of isoenzymes encoded by the two 4-coumarate:CoA ligase genes in parsley. Eur J Biochem **176:** 661–667

Noda K, Glover BJ, Linstead P, Martin C (1994) Flower color intensity depends on specialized cell shape controlled by a Myb-related transcription factor. Nature **369:** 661–664

Obokata J, Mikami K, Hayashida N, Sugiura M (1990) Polymorphism of a photosystem I subunit caused by alloploidy in *Nicotiana*. Plant Physiol **92:** 273–275

Okamuro JK, Goldberg RB (1985) Tobacco single-copy DNA is highly homologous to sequences present in the genomes of its diploid progenitors. Mol Gen Genet **198:** 290–298

Pellegrini L, Geoffroy P, Fritig B, Legrand M (1993) Molecular cloning and expression of a new class of ortho-diphenol-*O*-methyltransferases induced in tobacco (*Nicotiana tabacum* L.) leaves by infection or elicitor treatment. Plant Physiol **103:** 509–517

Ragg H, Kuhn DN, Hahlbrock K (1981) Coordinated regulation of 4-coumarate:CoA ligase and phenylalanine ammonia-lyase mRNAs in cultured plant cells. J Biol Chem **256:** 10061–10065

Ranjeva R, Boudet AM, Faggion R (1976) Phenolic metabolism in petunia tissue. Biochimie **58:** 1255–1262

Reinold S (1995) Vergleichende Untersuchungen zur Lokalisation des Phenylpropanoidstoffwechsels in ausgewählen Pflanzen. PhD thesis. Universität zu Köln, Cologne, Germany

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

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Schmelzer E, Kruger-Lebus S, Hahlbrock K (1989) Temporal and spatial patterns of gene expression around sites of attempted fungal infection in parsley leaves. Plant Cell **1:** 993-1001

Schmid J, Doerner PW, Clouse SD, Dixon RA, Lamb CJ (1990) Developmental and environmental regulation of a bean chalcone synthase promoter in transgenic tobacco. Plant Cell **2:** 619–631

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

Shirley BW, Kubasek WL, Storz G, Bruggemann E, Koornneef M, Ausubel FM, Goodman HM (1995) Analysis of *Arabidopsis* mutants deficient in flavonoid biosynthesis. Plant J **8:** 659–671

Smith HH (1968) Recent cytogenetic studies in the genus *Nicotiana*. Adv Genet **14:** 1–54

Stöckigt J, Zenk MH (1975) Chemical syntheses and properties of hydroxycinnamoyl-coenzyme A derivatives. Z Naturforsch **30c:** 352–358

Takahashi H, Ehara Y, Hirano H (1991) A protein in the oxygen-evolving complex in the chloroplast is associated with symptom expression on tobacco leaves infected with cucumber mosaic virus strain Y. Plant Mol Biol **16:** 689–698

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

Uknes S, Dincher S, Friedrich L, Negrotto D, Williams S, Thompson-Taylor H, Potter S, Ward E, Ryals J (1993) Regula-

tion of pathogenesis-related protein-1A gene expression in tobacco. Plant Cell **5:** 159–169

van Buuren M, Neuhaus J-M, Shinshi H, Ryals J, Meins F Jr (1992) The structure and regulation of homeologous tobacco endochitinase genes of *Nicotiana sylvestris* and *N. tomentosiformis* origin. Mol Gen Genet **232:** 460–469

Vincent JR, Nicholson RL (1987) Evidence for isoenzymes of 4-hydroxycinnamic acid:CoA ligase in maize mesocotyls and their response to infection by *Helminthosporium maydis* race O. Physiol Mol Plant Pathol **30:** 121–129

Voo KS, Whetten RW, O'Malley DM, Sederoff RR (1995) 4-Coumarate coenzyme A ligase from loblolly pine xylem. Plant Physiol **108:** 85–97

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

Wu SC, Hahlbrock K (1992) *In situ* localization of phenylpropanoid-related gene expression in different tissues of light-grown and dark-grown parsley seedlings. Z Naturforsch **47c:** 591–600

Zhao Y, Kung SD, Dube SK (1990) Nucleotide sequence of rice 4-coumarate:CoA ligase gene, 4-CL.1. Nucleic Acids Res **18:** 6144