The Plant Journal (2001) **28**(1), 95–104

Molecular cloning and functional identification of (+)-δ-cadinene-8-hydroxylase, a cytochrome P450 mono-oxygenase (CYP706B1) of cotton sesquiterpene biosynthesis

Ping Luo[1], Yan-Hong Wang[2], Guo-Dong Wang[1], Margaret Essenberg[2] and Xiao-Ya Chen[1,*]
[1]*National Laboratory of Plant Molecular Genetics, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 300 Fenglin Road, Shanghai 200032, People's Republic of China, and* [2]*Department of Biochemistry and Molecular Biology, Oklahoma State University, Stillwater, OK 74078-3035, USA*

Received 22 March 2001; revised 25 June 2001; accepted 30 June 2001
*For correspondence (fax: +86 21 64042385; email: xychen@iris.sipp.ac.cn)

Summary

In cotton, gossypol and related sesquiterpene aldehydes are present in the glands of aerial tissues and in epidermal cells of roots. A cytochrome P450 was found to be expressed in aerial tissues of glanded cotton cultivars, but not or at an extremely low level in the aerial tissues of a glandless cultivar. Its cDNA was then isolated from *Gossypium arboreum* L. After expression in *Saccharomyces cerevisiae*, the P450 was found to catalyse the hydroxylation of (+)-δ-cadinene, forming 8-hydroxy-(+)-δ-cadinene. This P450 mono-oxygenase has been classified as CYP706B1, and is the first member of the CYP706 family for which a function has been determined. Sesquiterpene aldehydes and *CYP706B1* transcripts were detected in roots of both the glanded and glandless cultivars and in aerial tissues of the glanded cultivar. In suspension cultured cells of *G. arboreum*, elicitors prepared from the phytopathogenic fungus *Verticillium dahliae* caused a dramatic induction of *CYP706B1* expression. The expression pattern of CYP706B1 and the position at which it hydroxylates (+)-δ-cadinene suggest that it catalyses an early step in gossypol biosynthesis. Southern blotting revealed a single copy of *CYP706B1* in the genome of *G. arboreum*. *CYP706B1* holds good potential for manipulation of gossypol levels in cottonseed via genetic engineering.

Keywords: cytochrome P450, *Gossypium*, gossypol, sesquiterpene, cadinene, phytoalexin.

Introduction

Cotton plants (*Gossypium* spp.) accumulate sesquiterpenes in sub-epidermal glands of aerial tissues and in root epidermal cells (Bell, 1986). These defence compounds may also function as phytoalexins, with their formation induced by fungal and bacterial infection and by other stress factors (Bell, 1986; Essenberg *et al.*, 1990). The majority of cotton sesquiterpenoids, including gossypol, are derived from a common parent compound (+)-δ-cadinene (Davis and Essenberg, 1995). The cotton (+)-δ-cadinene synthase (CAD1, also referred to as CDN1), a sesquiterpene cyclase, has been investigated at both enzymatic and molecular genetic levels (Alchanati *et al.*, 1998; Chen *et al.*, 1995; Chen *et al.*, 1996; E.M. Davis *et al.*, 1996; Meng *et al.*, 1999; Tan *et al.*, 2000). Little is known about enzymes catalysing subsequent biosynthetic steps. Desoxyhemigossypol-*O*-methyl transferase, which catalyses one of the late steps, has been purified (J. Liu *et al.*, 1999). However, until now, no enzymes that modify (+)-δ-cadinene itself have been reported.

In plants, cytochrome P450 mono-oxygenases participate in many biochemical pathways, including secondary metabolism, hormone biosynthesis and detoxification of xenobiotics (Bak *et al.*, 2001; Chapple, 1998; Kraus and Kutchan, 1995; Nielsen and Møller, 1999). A number of P450s of the phenylpropanoid pathway have been identified from plants (Akashi *et al.*, 1999; Humphreys *et al.*, 1999; Jung *et al.*, 2000; Martens and Forkmann, 1999; Steele *et al.*, 1999; reviewed by Chapple, 1998). In terpenoid pathways,

P450 mono-oxygenases are involved in the biosynthesis of various classes of compounds (Hedden and Kamiya, 1997; Hefner *et al.*, 1996). Microsomes prepared from *Mentha* spp. were demonstrated to catalyse hydroxylation of the monoterpene (–)-4*S*-limonene (Karp *et al.*, 1990; Lupien *et al.*, 1995), and recently cDNAs encoding two regiospecific P450 limonene hydroxylases, (–)-4*S*-limonene-3-hydroxylase and (–)-4*S*-limonene-6-hydroxylase, were reported (Lupien *et al.*, 1999). For biosynthesis of taxol, a diterpenoid found in *Taxus* spp., the first oxygenation step was catalysed by a cytochrome P450 taxane 10β-hydroxylase, and the cDNA was recently cloned (Schoendorf *et al.*, 2001). P450s were also shown to be involved in resin biosynthesis of conifer trees (Funk and Croteau, 1994). Sesquiterpenes constitute the largest group of natural terpenoids (Bohlmann *et al.*, 1998), and P450s have been proposed to play a major role in their biosynthetic pathways (Mihaliak *et al.*, 1993; de Kraker *et al.*, 2001). Although great progress has been made in the investigation of plant sesquiterpene cyclases, which catalyse the first committed steps in sesquiterpene biosynthesis (Bohlmann *et al.*, 1998; Chappell, 1995), up to now P450 enzymes catalysing subsequent oxidative reactions of sesquiterpene phytoalexins have not been characterized at the molecular level.

The cotton (+)-δ-cadinene synthases are encoded by a gene family. On the basis of sequence similarities, the family has been divided into two subfamilies, *CAD1-A* and *CAD1-C*. The diploid genome of *G. arboreum* contains about six members of *CAD1-C*, and a single copy of *CAD1-A* (Tan *et al.*, 2000). Our previous investigation indicated that both *CAD1-C* and *CAD1-A* members are actively transcribed in developing seeds of glanded cotton cultivars, but neither are transcribed in seeds of a glandless cultivar, of which the seeds are gossypol-free (Meng *et al.*, 1999). Therefore, there is reason to assume that genes encoding other enzymes of the gossypol pathway are also silent in developing seeds of glandless cultivars. By using a strategy comprising PCR and differential hybridization, we isolated a P450 cDNA from *G. arboreum*. Microsomal proteins prepared from yeast cells expressing this P450 catalysed 8-hydroxylation of (+)-δ-cadinene, which is a precursor to desoxyhemigossypol, and through it to hemigossypol and related sesquiterpene aldehydes (Benedict *et al.*, 1995; Davis and Essenberg, 1995). This cotton P450 has been placed in a new subfamily as CYP706B1, and it is the first member of the CYP706 family for which the function has been determined.

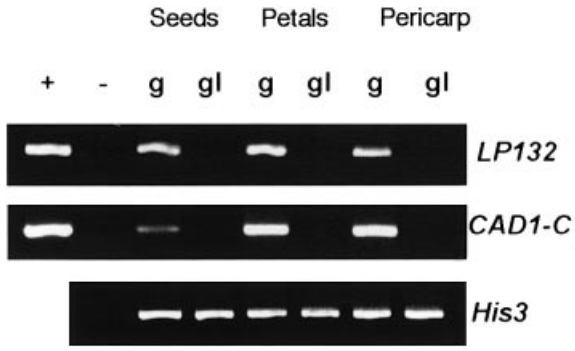

Figure 1. RT–PCR analysis (30 cycles of amplification) of expression of *LP132* and *CAD1-C* in developing seeds (27 days post-anthesis), petals (3 days pre-anthesis) and pericarp (4 days post-anthesis) of glanded and glandless cotton cultivars.
His3, amplification of *histone 3* as an internal control; *+*, positive control with corresponding plasmid DNA as the template; *–*, negative control, no template DNA; g, glanded cultivar of *G. hirsutum* cv. Zhong-12; gl, glandless cultivar of *G. hirsutum* cv. Hai-1.

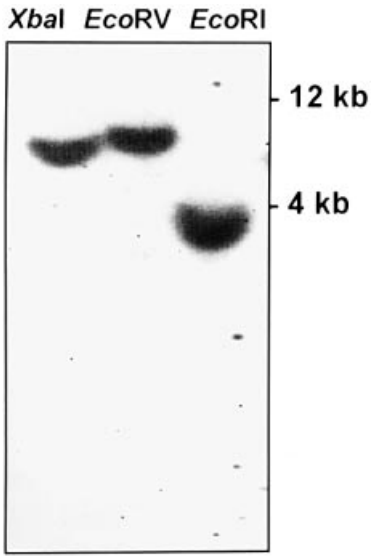

Figure 2. Southern blot of *CYP706B1* in the genome of *Gossypium arboreum* L. DNA (20 μg) was digested with *Xba*I, *Eco*RV or *Eco*RI, separated on an 0.8% agarose gel, and hybridized with a ^{32}P-labelled *CYP706B1* probe.

Results

cDNA cloning and analysis

The 3′-terminal cDNA fragments of P450 mono-oxygenases were amplified by PCR from a *G. arboreum* library that was constructed from mRNAs of fungal elicitor-treated suspension cells (Chen *et al.*, 1995). This resulted in a mixture of DNA fragments with different lengths ranging from 300–600 bp, possibly due to a large number of P450 genes expressed in cotton cells after elicitation. About 100 individual clones of the PCR products were then used in differential dot-blot hybridization with cDNA probes prepared from developing seeds of a glanded cultivar *G. hirsutum* cv. Zhong-12 and a glandless cultivar *G. hirsutum* cv. Hai-1, respectively. A clone with an

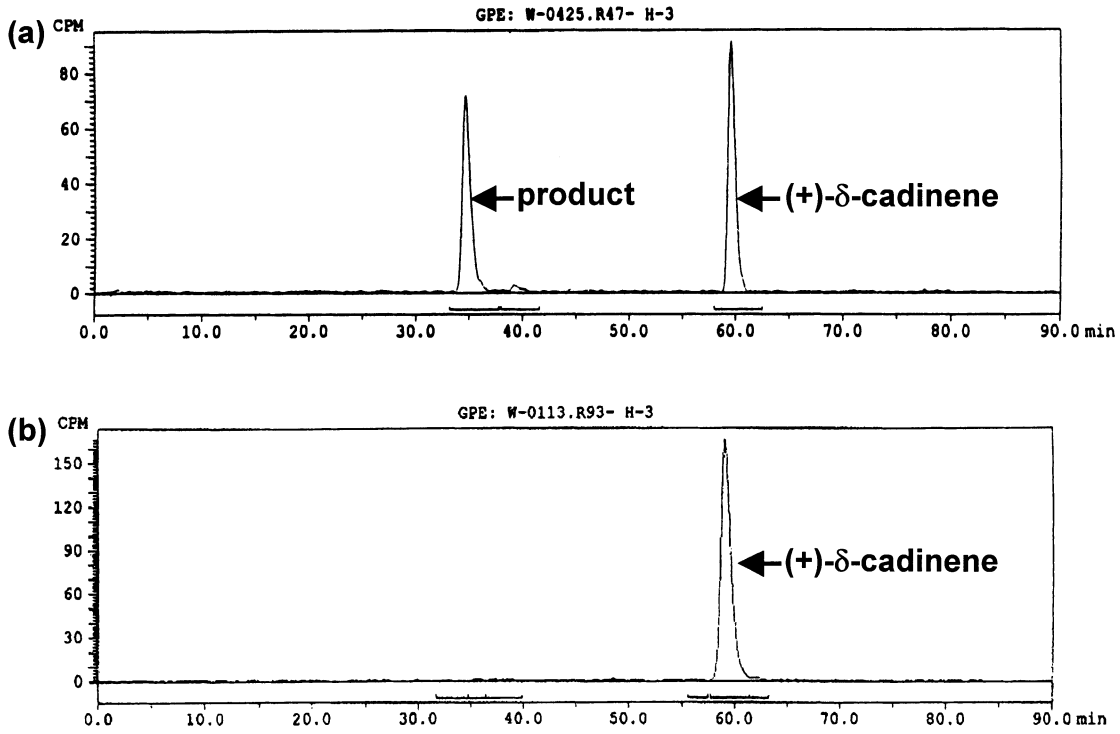

Figure 3. Reversed-phase HPLC of reaction mixtures.
[^{3}H]-(+)-δ-cadinene was incubated with microsomes of yeast cells harbouring CYP706B1 (a) or the empty expression vector pYeDP60 (b). [^{3}H]-(+)-δ-cadinene eluted at 59–60 min; the product eluted at 33.5 min.

approximately 500 bp insert, *LP132*, showed a clear hybridization signal with probes from glanded cotton only (data not shown), and its nucleotide sequence was then determined. A search of the database with the BLASTX program revealed that *LP132* had high sequence similarities with plant P450 mono-oxygenases (30–45% identity at the amino acid sequence level). Subsequent RT–PCR with primers specific to *LP132* indicated that, while a significant level of its transcript was present in developing seeds, petals and pericarp of the glanded cultivar, this transcript was indeed undetectable in these tissues of the glandless cultivar. As shown in Figure 1 and also reported previously (Meng *et al.*, 1999; Tan *et al.*, 2000), this expression pattern was identical to that of (+)-δ-cadinene synthase, the sesquiterpene cyclase that catalyses the first committed step in biosynthesis of cotton sesquiterpene phytoalexins.

A cDNA clone corresponding to *LP132* was isolated from the *G. arboreum* library. It contained a reading frame coding for a protein of 536 amino acid residues, with a calculated molecular mass of 60.1 kDa. A search of the database with its deduced protein sequence revealed greatest sequence identity (up to 47%) with six putative P450 proteins from *Arabidopsis thaliana* (CYP706A1–CYP706A6; AL080318 and AL161533, genomic sequences in chromosome 4). Among proteins of known functions, it has the highest amino acid sequence identity (34.2%) with a flavonoid 3′-hydroxylase from *Petunia hybrida*, a member of the CYP75 family (Brugliera *et al.*, 1999). This cotton P450 has been placed in a new subfamily as CYP706B1 (accessed via dnelson@utmenl.utmen.edu).

Sequence analysis revealed several structural motifs characteristic of eukaryotic P450s. The highly conserved haem-binding motif FxxGxRxCxG (Chapple, 1998) was found in CYP706B1 as FGSGRRMCAG, 71 amino acid residues from the C-terminus. In most plant P450s, there is a proline residue immediately after the invariant haem-binding cysteine (Schalk *et al.*, 1999); however, in CYP706B1, this proline is replaced by alanine. The proline-rich region immediately after the N-terminal signal anchor sequence (Nelson and Strobel, 1988), with a consensus of (P/I)PGPx(G/P)xP (Schalk *et al.*, 1999), was completely conserved in this cotton P450 as PPGPRGLP, amino acids 66–73. In addition, the threonine-containing pocket for binding an oxygen molecule, with a consensus sequence of (A/G)Gx(D/E)T(T/S) (Durst and Nelson, 1995), was found as GGTDTT, amino acids 332–337.

Hybridization of the genomic DNA of *G. arboreum* with a *CYP706B1* probe revealed single bands in *Eco*RI-, *Eco*RV- and *Xba*I-digested DNA samples (Figure 2). This hybridiza-

tion pattern indicated a single-copy gene encoding CYP706B1 in the genome of *G. arboreum*, a diploid cotton species.

Sesquiterpene hydroxylase activity

Microsomal proteins prepared from yeast cells expressing CYP706B1 showed clear hydroxylase activity *in vitro* with tritium-labelled (+)-δ-cadinene as a substrate. HPLC revealed a single product peak, more polar than (+)-δ-cadinene (Figure 3a). No product was detected by radio-HPLC when the microsomes from yeast cells harbouring an empty pYeDP60 vector were used as the catalyst (Figure 3b). In a 15 min assay, the yeast microsomes containing CYP706B1 showed a specific activity of 42.6 ± 2.8 nmol product mg^{-1} protein h^{-1}. The highest activity was achieved when the reaction was supported by 1 mM NADPH; when NADPH was replaced by NADH,

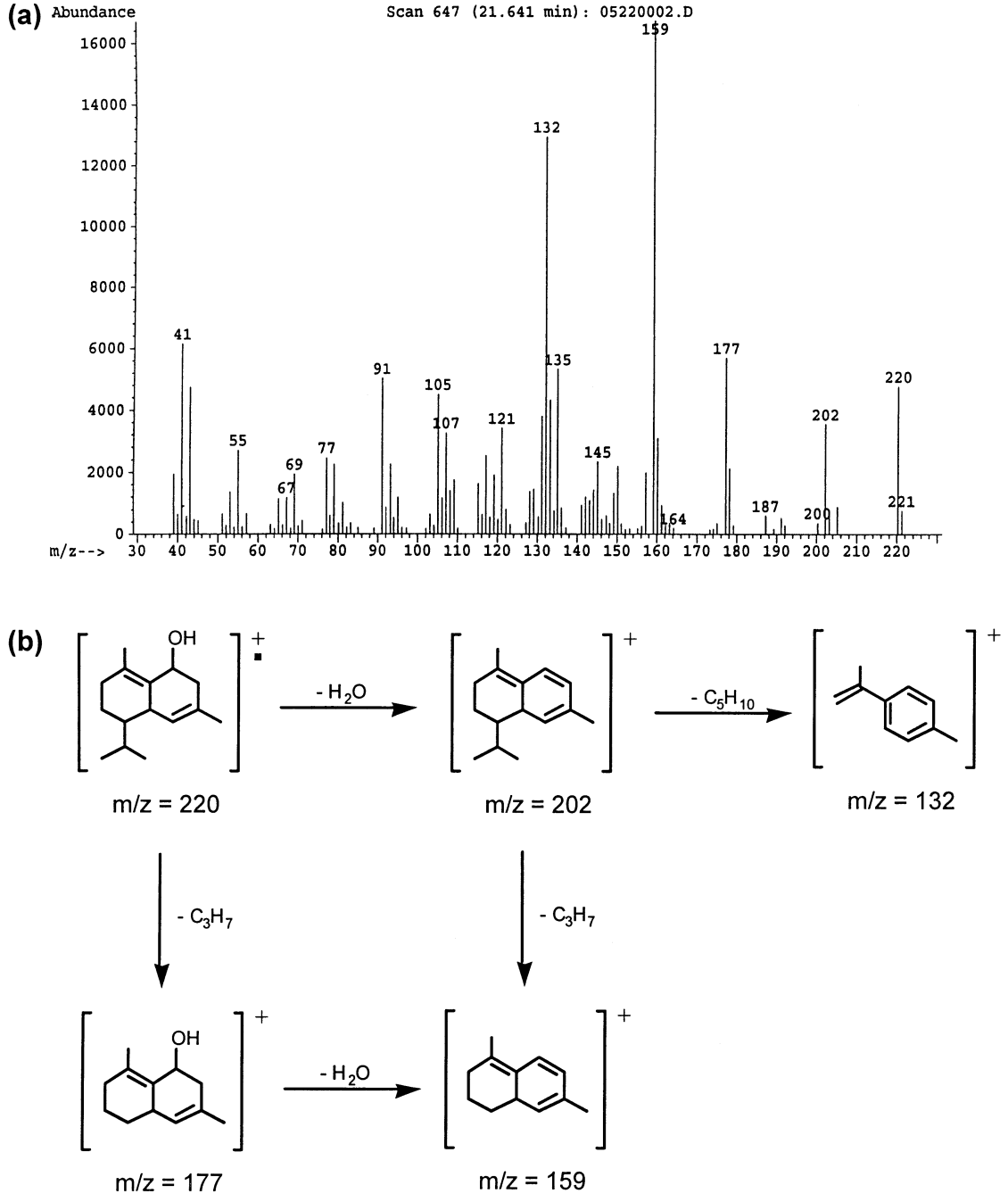

Figure 4. (a).

the activity decreased by about 65%. Treating the reaction mixture with a slow stream of CO for about 2 min inhibited the hydroxylase activity by 70%.

Product identification by GC/MS and NMR

GC analysis also revealed a single product peak (data not shown). The mass spectra of the product showed a molecular mass of 220 (Figure 4a), consistent with that of a hydroxy-δ-cadinene (the molecular mass of (+)-δ-cadinene is 204.). The MS fragmentation pattern indicated the loss of water, loss of the isopropyl group (C_3H_7) and retro Diels-Alder cleavage (-C_5H_{10}) that are typical of sesquiterpenes with this carbon skeleton (Figure 4b) (Davis and Essenberg, 1995). The ^{1}H-NMR spectrum was very similar to that of δ-cadinene (Table 1). The differences in chemical shifts and multiplicities indicated that the product is hydroxylated at C8. A two-dimensional ^{1}H–^{1}H COSY (correlation spectroscopy) analysis revealed all the expected connectivities between hydrogen atoms on the same or adjacent carbon atoms except that for H13 to H4. Detailed elucidation of the structure of this compound will be published elsewhere. The reaction catalysed by CYP706B1 of *Gossypium arboreum* is given in Figure 4(c). Some preparations of the product appeared to be conjugated through the 8-hydroxyl group to a moiety that has not yet been identified. However, when the microsomes were prepared by ultracentrifugation rather than with polyethylene glycol and the non-ionic detergent *n*-octyl-β-D-glucoside was included in the enzyme reaction mixture, (+)-δ-cadinene was reproducibly converted to the free alcohol, 8-hydroxy-(+)-δ-cadinene.

Expression pattern

In cotton roots, gossypol and related sesquiterpene aldehydes are stored in epidermal tissues, rather than in sub-epidermal glands, as found in aerial tissues. RT–PCR indicated that *CYP706B1* was expressed in roots of both the glanded *G. arboreum* and the glandless cultivar of *G. hirsutum* cv. Hai-1 (Figure 5a), although, according to Northern analysis, roots of the glanded cultivar had a higher steady-state mRNA level than roots of the glandless cultivar (Figure 5b). Similarly, roots of *G. arboreum* had a higher level of sesquiterpene aldehydes than roots of *G. hirsutum* cv. Hai-1 (Figure 5c). Transcripts of *CYP706B1* were also detected in cotyledons and hypocotyls of *G. arboreum* seedlings by RT–PCR, but not or at an extremely

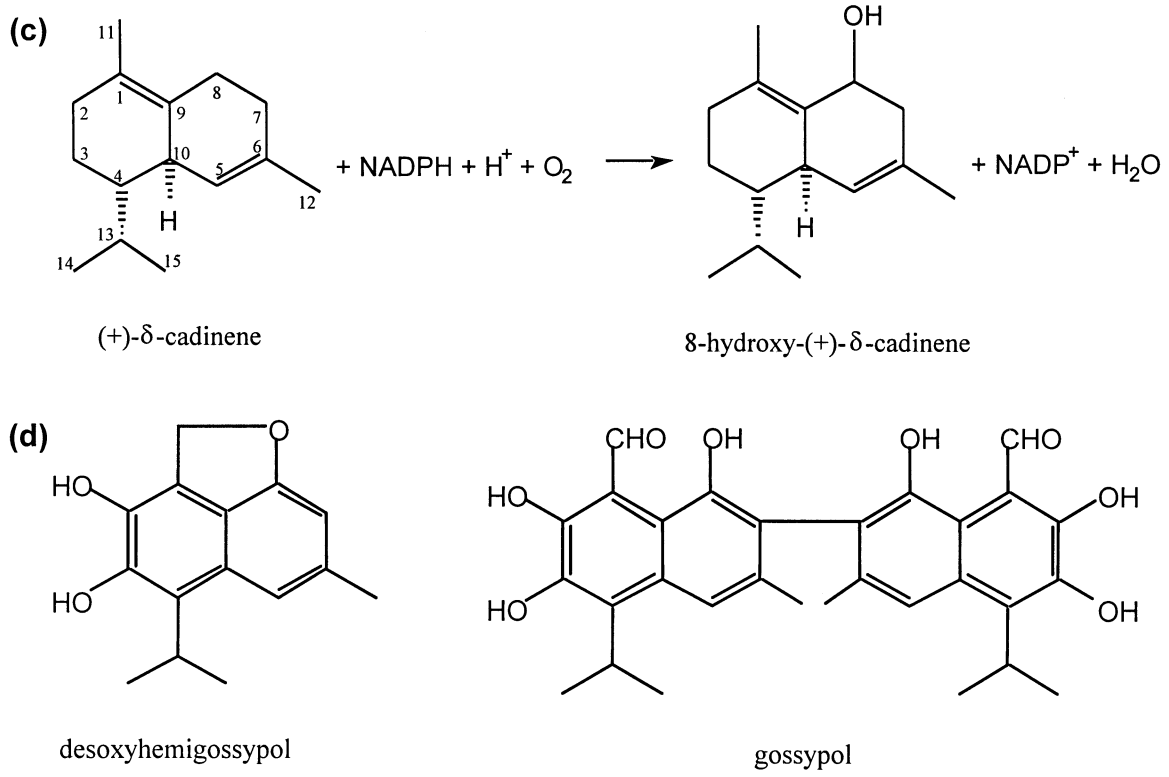

Figure 4. Mass spectral analysis of 8-hydroxy-(+)-cadinene and the proposed reaction catalysed by CYP706B1.
(a) Mass spectrum of the product obtained by incubation of (+)-δ-cadinene with yeast microsomes containing CYP706B1. (b) Proposed fragmentation scheme of the molecular ion. (c) The reaction catalysed by the cotton CYP706B1. (d) Structures of desoxyhemigossypol and gossypol.

Table 1. ^{1}H-NMR chemical shifts of δ-cadinene and 8-hydroxy-δ-cadinene

H	δ-cadinene[a]	8-hydroxy-δ-cadinene[b]
H_2-2	1.95 (m)	
H-2a		2.02
H-2b		1.96
H-3a	1.61 (m)	1.63
H-3b	1.16 (m)	1.16
H-4	1.05 (m)	1.12 (m)
H-5	5.45 (br s)	5.60 (br s)
H_2-7	2.00 (m)	
H-7a		2.40 (br d, 12.5 Hz)
H-7b		2.21 (br d, 12.5 Hz)
H-8a	2.72 (m)	3.45
H-8b	1.90 (m)	
H-10	2.52 (br d, 9.0 Hz)	2.25 (d, 12.2 Hz)
H-13	2.06 (m)	1.91
11-Me	1.66 (s)	1.67 (s)
12-Me	1.68 (s)	1.69 (s)
14-Me	0.96 (d, 6.9 Hz)	1.10 (d, 6.8 Hz)
15-Me	0.78 (d, 6.9 Hz)	1.00 (d, 6.8 Hz)

[a]400 MHz, $CDCl_3$ (G.D. Davis *et al.*, 1996).
[b]600 MHz, d_6-benzene.
Signal multiplicities are indicated as follows: m, multiplet; br s, broad singlet; br d, broad doublet; s, singlet; d, doublet.

low level in cotyledons and hypocotyls of the glandless *G. hirsutum* cv. Hai-1 (Figure 5a). Accordingly, sesquiterpene aldehydes were detected in glanded cotyledons and hypocotyls (Figure 5c), but not in glandless cotyledons and hypocotyls (data not shown). In developing seeds of *G. arboreum*, *CYP706B1* transcripts were detected at 20 days post-anthesis and afterwards (Figure 6a), followed by sesquiterpene aldehyde accumulation (Figure 6b), a pattern similar to *CAD1*-C expression and sesquiterpene accumulation in developing seeds of another glanded cultivar, *G. hirsutum* cv. Sumian-6 (Meng *et al.*, 1999).

When suspension cultured cells of *G. arboreum* were treated with elicitors of *Verticillium dahliae*, a phytopathogenic fungus responsible for a vascular wilt disease of cotton, transcription of *CYP706B1* was significantly induced (Figure 7a), followed by increased production of sesquiterpene phytoalexins (Figure 7b). After induction of *CYP706B1* transcription within 4 h of elicitation, the mRNA steady-state level peaked again at around 20 h post-elicitation (Figure 7a). The elicitation experiment was repeated, and two peaks of transcript level were again detected by Northern hybridization (data not shown).

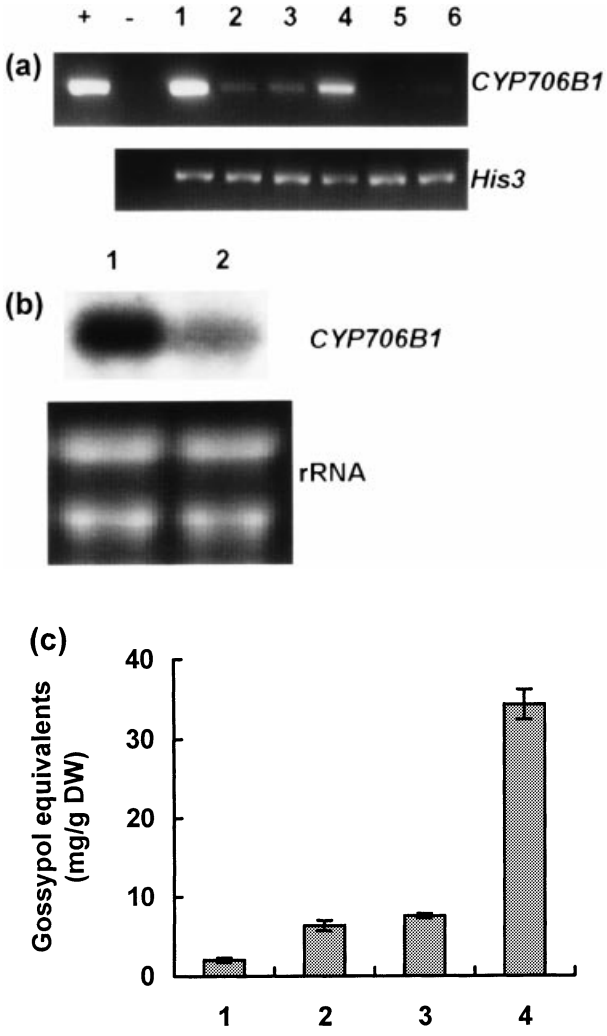

Figure 5. Expression of CYP706B1 and sesquiterpene aldehyde levels in seedlings of glanded *G. arboreum* and glandless *G. hirsutum* cv. Hai-1.
(a) RT–PCR analysis (30 cycles of amplification) of *CYP706B1* transcripts. RNAs were isolated from roots, hypocotyls and cotyledons, respectively, of glanded seedlings (lanes 1–3), and from those tissues of glandless seedlings (lanes 4–6). Controls were as in Figure 1. (b) Northern blot of *CYP706B1* transcripts in roots of glanded (lane 1) and glandless (lane 2) seedlings. (c) Sesquiterpene aldehydes in roots of glandless seedlings (lane 1), and in roots, hypocotyls and cotyledons, respectively, of glanded seedlings (lanes 2–4).

Discussion

There are two groups of cadinane-type sesquiterpenoids in cotton, differentiated by their positions of hydroxylation. The 7-hydroxylated cadinanes, such as 2,7-dihydroxycadalene and lacinilene C, are induced to accumulate in foliar tissues after bacterial infection (Davila-Huerta *et al.*, 1995; Essenberg *et al.*, 1990; Pierce *et al.*, 1996). The 8-hydroxylated cadinanes, such as gossypol, desoxyhemigossypol and related sesquiterpene aldehydes, are the largest group of cotton sesquiterpenoids and are distributed in healthy roots, seeds and glanded green tissues; their formation may also be elicited by fungal or bacterial infection (Bell, 1986; Abraham *et al.*, 1999; Tan *et al.*, 2000). Hydroxylation of (+)-δ-cadinene at C8 is a critical step for the formation of desoxyhemigossypol, which is a key

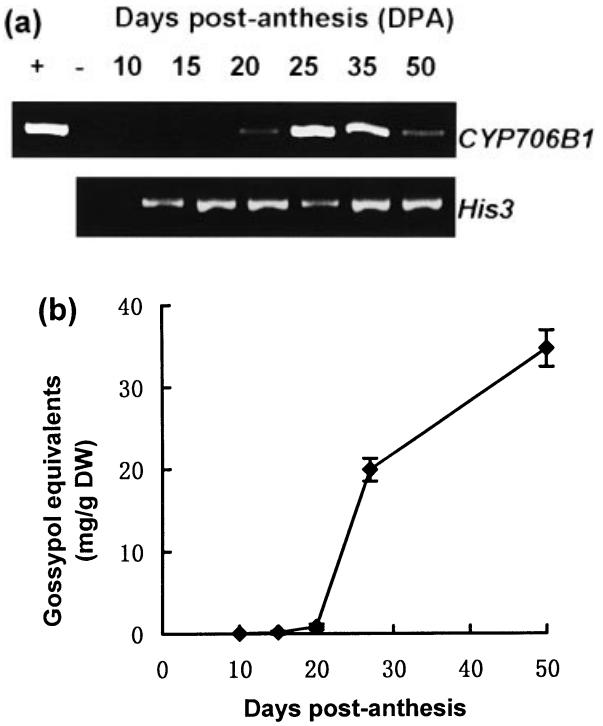

Figure 6. Expression of *CYP706B1* and accumulation of sesquiterpene aldehydes in developing seeds of *G. arboreum*.
(a) RT–PCR analysis (25 cycles of amplification) of *CYP706B1* transcripts.
(b) Sesquiterpene aldehydes.

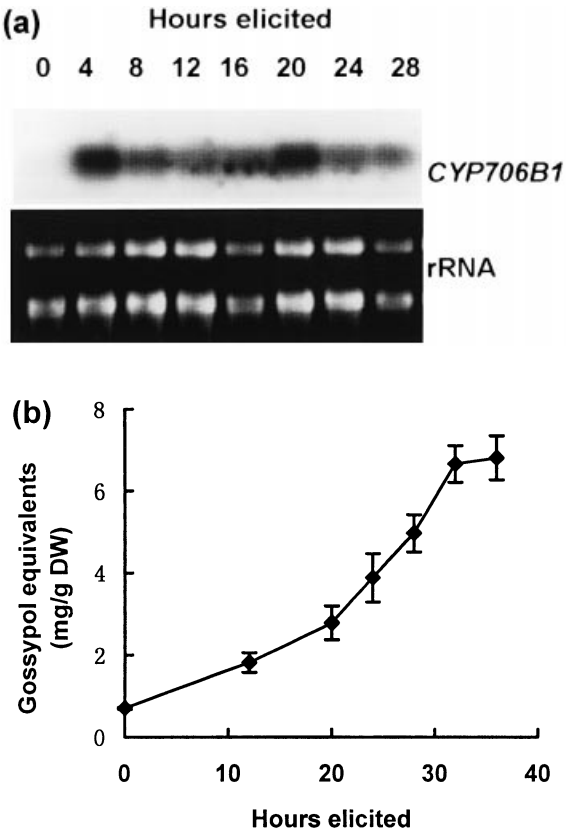

Figure 7. Induced expression of *CYP706B1 and* accumulation of sesquiterpene aldehydes in *G. arboreum* suspension-cultured cells treated with *V. dahliae* elicitors.
(a) Northern blot of *CYP706B1* transcripts.
(b) Sesquiterpene aldehydes.

intermediate in biosynthesis of the sesquiterpene aldehydes of cotton (Figure 4d; Bell, 1986; J. Liu *et al.*, 1999). Therefore, cloning of an enzyme catalysing this hydroxylation of (+)-δ-cadinene is important progress in elucidation of gossypol biosynthesis. Still needed to confirm the biosynthetic role of CYP706B1 is evidence that 8-hydroxy-(+)-δ-cadinene can be incorporated into gossypol by cotton tissues or cells.

In various tissues of seedlings and mature plants, the expression pattern of CYP706B1 reported herein is similar to that of the sesquiterpene cyclase CAD1-C (Meng *et al.*, 1999; Tan *et al.*, 2000). In suspension cultured cells, expression of CYP706B1 is inducible by fungal elicitors (Figure 7), as are farnesyl diphosphate synthase and CAD1 (Chen *et al.*, 1995; Chen *et al.*, 1996; C.J. Liu *et al.*, 1999), the two enzymes upstream of this P450 mono-oxygenase. These enzymes may be coordinately regulated, directing isoprenoid intermediates into sesquiterpene aldehydes, including gossypol. Coordinate regulation of enzymes involved in phenylpropanoid phytoalexin synthesis has been reported previously (Logemann *et al.*, 2000). As genes encoding the two enzymes catalysing early steps in the gossypol pathway were found not to be expressed in developing seeds and other aerial tissues of healthy plants of *G. hirsutum* cv. Hai-1, this glandless cultivar will be a valuable screening tool for cloning other enzymes involved in gossypol biosynthesis. It seems that, in cotton plants, a common factor(s) controls development of glands and biosynthesis of sesquiterpene aldehydes. However, in roots, which accumulate sesquiterpene aldehydes in epidermal cells rather than in sub-epidermal glands, expression of CYP706B1 and biosynthesis of sesquiterpene aldehydes was detected in both the glanded and glandless cultivars. This suggests that mechanisms regulating sesquiterpene aldehyde biosynthesis in roots and in aerial tissues are at least partly different.

Cotton seed contains on average 30% oil and over 30% protein, and is potentially useful as a food additive. However, the nutritional value of the seed is limited because of a high content of sesquiterpene aldehydes, mainly gossypol, which are toxic to monogastric animals. CYP706B1 appears to catalyse an early step in gossypol biosynthesis, directing (+)-δ-cadinene into toxic sesquiterpene aldehydes. It is encoded by a single-copy gene in the diploid *G. arboreum*. In comparison with (+)-δ-cadinene

synthase, which is encoded by a complex gene family (Tan *et al.*, 2000), this enzyme may provide a better target for suppression of gossypol formation in cotton seed through genetic engineering.

Experimental procedures

Materials

Gossypium arboreum L. cv. Qingyangxiaozhi, *G. hirsutum* L. cv. Zhong-12, and a glandless cultivar *G. hirsutum* cv. Hai-1 were grown in a greenhouse. Flowers, peels (pericarp of the cotton boll), and seeds were collected at various developmental intervals as previously described (Meng *et al.*, 1999; Tan *et al.*, 2000). Cell suspension cultures of *G. arboreum* cv. Qingyangxiaozhi were maintained in liquid MS medium, elicitors of the fungus *Verticillium dahliae* were prepared and applied to suspension cultured cells at a final concentration of 1 μg sucrose equivalent per ml culture, as previously described (C.J. Liu *et al.*, 1999).

[^{3}H](+)-δ-cadinene was prepared from commercial [1–^{3}H]FPP (NEN, Boston, MA, USA), which had been diluted with non-radioactive FPP to a specific activity of 32 μCi μmol^{-1} (E.M. Davis *et al.*, 1996), with the CAD1-C1 fusion protein expressed in *E. coli* (Chen *et al.*, 1995). Reaction products were extracted with hexane, passed through a 2 g, 100–200 mesh silica column to remove any [1-^{3}H]farnesol, and eluted with hexane.

Cloning of cDNA

A degenerate primer 5′-GCGGATCCGA(AG)TT(CT)(AC)G(AG-CT)CC(AGCT)GA(AG)(AC)G-3′ was synthesized corresponding to a conserved peptide sequence of EEF(L/R)PERF (Frank *et al.*, 1996), about 20 amino acids upstream of the haem-binding domain of plant P450 mono-oxygenases. It was used together with a vector-specific T7 primer (Stratagene, La Jolla, CA, USA) in PCR amplification of P450 fragments from a λ-UniZap cDNA library constructed from elicitor-treated *G. arboreum* cells (Chen *et al.*, 1995). The PCR programme was 94°C for 30 sec, 55°C for 30 sec and 72°C for 30 sec, for 30 cycles. The PCR products were inserted into pGEM-T vectors (Promega, Madison, WI, USA). After amplification in *E. coli*, plasmid DNA from individual clones was spotted onto nitrocellulose membranes, which were then fixed at 80°C for 2 h for subsequent hybridization screening.

Probes were generated from total RNAs isolated from developing seeds of the glanded and glandless *G. hirsutum* cultivars. The first-strand cDNAs were synthesized, ^{32}P-labelled using a random DNA labelling kit (TaKaRa, Dalian, China), and used for dot hybridizations. Clones showing preferential hybridization with probes from glanded seeds were selected and sequenced. Specific primers LP132F (5′- TGACTGATCATGAGAAGCT-3′) and LP132R (5′-GTGCTGGAGATTTGATGGT-3′), based on the sequence of *LP132*, were used for screening the *G. arboreum* cDNA library by using a PCR 96-well plate method (C.J. Liu *et al.*, 1999). A cDNA clone, *CYP706B1*, was then isolated and sequenced.

DNA and RNA analysis

Genomic DNA was isolated from *G. arboreum* foliar tissues as described previously (Tan *et al.*, 2000). After complete digestion (4 h to overnight) with restriction enzymes *Eco*RI, *Xba*I and *Eco*RV, about 20 μg of DNA per lane were separated by electrophoresis and transferred onto a nitrocellulose membrane. For probe preparation, *CYP706B1* was digested with *Xba*I and *Eco*RV, and the 947 bp fragment released was ^{32}P-labelled. Hybridization and washing were performed following a standard protocol (Sambrook *et al.*, 1989), and the membrane was exposed to X-ray film for 1–2 days.

Total RNAs were isolated from tissues or from suspension-cultured cells by a cold phenol method (C.J. Liu *et al.*, 1999), and the transcripts were analysed by RT–PCR with primers LP132F and LP132R for *CYP706B1* (positions 1433–1689), primers 97400 (5′-CACATCC(AC)TTCGATTCCGAC-3′) and 97T580 (5′-AGGCTTAAA-TGGTGGGTGGT-3′) for *CAD1-C* (positions 398–610), and primers H3F (5′-GAAGCCTCATCGATACCGTC-3′) and H3R (5′-CTACCA-CTACCATCATGTC-3′) for *histone 3* (*his3*, positions 95–526). For Northern analysis, 20 μg of RNA per lane were separated by electrophoresis and blotted onto a nitrocellulose membrane, and the blots were hybridized with ^{32}P-labelled DNA probes of either *CYP706B1* (see above) or *CAD1-C1* (C.J. Liu *et al.*, 1999). After hybridization and washing, the blots were exposed to X-ray film for 2 days.

Expression in yeast cells and enzyme assay

The yeast *Saccharomyces cerevisiae* strain W(R), which over-expresses the yeast cytochrome P450 reductase when grown on galactose, and the expression vector pYeDP60 (Pompon *et al.*, 1996) were used for heterologous expression of CYP706B1. The cDNA of *CYP706B1* was modified by PCR with a 5′ terminal primer 5′-GGGTACCATGTTGCAAATAGCTTTCAG-3′ in which a *Kpn*I site was introduced, and a 3′ terminal primer 5′-GGGAGCTCTTAC-TTCATATAGTGCTGGA-3′ in which a *Sac*I site was introduced. PCR was conducted on plasmid DNA by using Pyrobest™ DNA polymerase (TaKaRa). After digestion with the restriction enzymes, the DNA fragment was inserted into pYeDP60. Plasmid DNA was introduced into yeast cells by a LiAc method, and transformants were cultured by a high-density procedure and broken mechanically (Pompon *et al.*, 1996). Microsomes were isolated by centrifugation with polyethylene glycol at 10 000 ***g*** (Pompon *et al.*, 1996) or by ultracentrifugation at 100 000 ***g*** for 80–90 min.

Determination of the (+)-δ-cadinene hydroxylase activity assay was largely based on a published protocol for monoterpene P450 hydroxylases (Mihaliak *et al.*, 1993). The radioactive assay was conducted in 100 μl potassium phosphate buffer (50 mM, pH 7.4) containing 1 mM EDTA (pH 8.0), 0.4 M sucrose, 2 mM DTT, 0.1% BSA, 1 mM NADPH, 5 μM flavin adenine dinucleotide (FAD), 5 μM flavin mono-nucleotide (FMN), 4 mM glucose-6-phosphate, 1 unit glucose-6-phosphate dehydrogenase, 40 μM ^{3}H-(+)-δ-cadinene (32 μCi μmol^{-1}) and yeast microsomes (100–200 μg protein). The reaction was started by adding the microsomes, incubated at 30°C for 1 h, and stopped by chilling on ice. The reaction mixture was extracted three times with 500 μl hexane:ethyl acetate (1:1), the extract was filtered through a 0.2 g, 100–200 mesh silica gel column in a Pasteur pipette, eluted with 1.5 ml hexane:ethyl acetate (1:1), and concentrated to approximately 100 μl with an argon stream. The entire sample was subjected to reversed-phase HPLC on C_{18} silica (250 × 4.6 mm, 5 μm particle diameter) with a gradient of 40:60 to 100:0 acetonitrile/water over 60 min followed by 10 min of 100% acetonitrile, flow rate 1 ml min^{-1}. Tritium was detected by an online liquid scintillation counter (β-RAM, IN/US Systems Inc., Tampa, FL, USA). The tritium-labelled product eluted at 33–34 min, followed by the substrate (+)-δ-cadinene, at 59–60 min.

Protein concentration was determined with the BioRad Protein Assay (Bio-Rad, Hercules, CA, USA). Sesquiterpene aldehydes were extracted and quantified by the phloroglucinol method (Meng *et al*., 1999).

Isolation of the reaction product

Samples for GC/MS and NMR analyses were prepared from non-radioactive (+)-δ-cadinene, which had been prepared by acid-catalysed rearrangement of commercial (–)-α-cubebene (Davis and Essenberg, 1995). To avoid problems with scaling up the reaction, a hundred 1.0 ml reaction mixtures (the same as stated above) were prepared with 200 μM non-radioactive (+)-δ-cadinene as substrate and incubated at 30°C for 30 min. *N*-octyl-β-D-glucoside (0.1%) was included in some reaction mixtures. The reactions were quenched by adding 0.2 ml diethyl ether to each tube. The mixtures were combined, the ether was removed, and products were extracted from the aqueous phase twice more with 20 ml diethyl ether. The combined ether extract was dried by passage through a silica gel column (1 g, 10–40 μm particle size, 9 mm × 30 mm) column, and evaporated to yield an 11.4 mg residue. The residue was dissolved in 100 μl diethyl ether and chromatographed on a silica gel column (3 g, 10–40 μm particle size, 9 mm × 90 mm) with hexane-ethyl ether (4:1) as eluant, collecting 2 ml fractions. Fractions 7–15 were combined and washed with 3 M sodium carbonate buffer, pH 12.

Identification of the reaction product

Samples were analysed by GC/MS on an HP 5890 series II GC equipped with a HP-ZB5 column (30 m × 0.25 mm). The helium inlet pressure was controlled by electronic pressure control to achieve a constant column flow of 1.0 ml min^{-1} during the following oven programme: initial temperature 45°C for 5 min, ramp of 5°C min^{-1} to 295°C, and 10 min at 295°C. The ionization potential was 70 eV. ^{1}H-NMR and ^{1}H–^{1}H COSY spectra were recorded on a Varian Unity Inova 600 spectrometer using tetramethylsilane as an internal standard.

Acknowledgements

Special thanks are given to Emeritus Professor P. Heinstein (Purdue University) for his various contributions. We are also grateful to Professor R. Croteau (Washington State University), Professor J. Chappell (University of Kentucky) and Emeritus Professor E. Eisenbraun (Oklahoma State University) for their helpful discussions. We thank Dr D. Pompon (Centre National de la Recherche Scientifique, Gif-sur-Yvette, France) for providing the yeast expression system, Mr K. Payne, Dr D. Jacobs and Dr F. Qiu (Oklahoma State University) for GC/flame ionization detector (FID), GC/MS and NMR analyses, respectively, Dr K. Jackson (University of Oklahoma Health Sciences Center) and Dr L. Sumner (S.R. Noble Foundation, Ardmore, Oklahoma) for electrospray ionization (ESI)/MS analyses. Funds for the 600 MHz NMR spectrometer of the Oklahoma Statewide Shared NMR Facility were provided by the National Science Foundation (BIR-9512269), the Oklahoma State Regents for Higher Education, the W.M. Keck Foundation and Conoco Inc. This work was supported by the National Natural Science Foundation of China (3992505 and 30030020), the National High Tech Program of China (Z170204), Cotton Incorporated (Agreement 98-598) and the Oklahoma Agricultural Experiment Station.

References

Abraham, K.J., Pierce, M.L. and Margaret, M. (1999) The phytoalexins deoxyhemigossypol and hemigossypol are elicited by *Xanthomonas* in *Gossypium cotyleodns*. *Phytochemistry*, **52**, 829–836.

Akashi, T., Aoki, T. and Ayabe, S. (1999) Cloning and functional expression of a cytochrome P450 cDNA encoding 2-hydroxyisoflavanone synthase involved in biosynthesis of the isoflavonoid skeleton in licorice. *Plant Physiol.* **121**, 821–828.

Alchanati, I., Patel, J.A.A., Liu, J., Benedict, C.R., Stipanovic, R.D., Bell, A.A., Cui, Y. and Magill, C.W. (1998) The enzymatic cyclization of nerolidyl diphosphate by δ-cadinene synthase from cotton stele tissue infected with *Verticillium dahliae*. *Phytochemistry*, **47**, 961–967.

Bak, S., Tax, F.E., Feldmann, K.A., Galbraith, D.W. and Feyereisen, R. (2001) CYP83B1, a cytochrome P450 at the metabolic branch point in auxin and indole glucosinolate biosynthesis in Arabidopsis. *Plant Cell*, **13**, 101–111.

Bell, A.A. (1986) Physiology of secondary products. In: *Cotton Physiology* (Mauney, J. R. and Stewart, J.McD., eds). Memphis, Tennessee: The Cotton Foundation, pp. 597–621.

Benedict, C.R., Alchanati, I., Harvey, P.J., Liu, J., Stipanovic, R.D. and Bell, A.A. (1995) The enzymatic formation of δ-cadinene from farnesyl diphosphate in extracts of cotton. *Phytochemistry*, **39**, 327–331.

Bohlmann, J., Meyer-Gauen, G. and Croteau, R. (1998) Plant terpenoid synthases: molecular biology and phylogenetic analysis. *Proc. Natl Acad. Sci. USA*, **95**, 4126–4133.

Brugliera, F., Barri-Rewell, G., Holton, T.A. and Mason, J.G. (1999) Isolation and characterization of a flavonoid 3′-hydroxylase cDNA clone corresponding to the *Ht1* locus of *Petunia hybrida*. *Plant J.* **19**, 441–451.

Chappell, J. (1995) Biochemistry and molecular biology of the isoprenoid biosynthetic pathway in plants. *Annu. Rev. Plant. Physiol. Plant Mol. Biol.* **46**, 521–547.

Chapple, C. (1998) Molecular-genetic analysis of plant cytochrome P450-dependent monooxygenases. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **49**, 311–343.

Chen, X.Y., Chen, Y., Heinstein, P. and Davisson, V.J. (1995) Cloning, expression, and characterization of (+)-δ-cadinene synthase: a catalyst for cotton phytoalexin biosynthesis. *Arch. Biochem. Biophys.* **324**, 255–266.

Chen, X.Y., Wang, M., Chen, Y., Davisson, V.J. and Heinstein, P. (1996) Cloning and heterologous expression of a second (+)-δ-cadinene synthase from *Gossypium arboreum*. *J. Nat. Prod.* **59**, 944–951.

Davila-Huerta, G., Hamada, H., Davis, G.D., Stipanovic, R.D., Adams, C.M. and Essenberg, M. (1995) Cadinane-type sesquiterpenes induced in *Gossypium* cotyledons by bacterial inoculation. *Phytochemistry*, **39**, 531–536.

Davis, E.M., Tsuji, J., Davis, G.D., Pierce, M.L. and Essenberg, M. (1996) Purification of (+)-δ-cadinene synthase, a sesquiterpene cyclase from bacteria-inoculated cotton foliar tissue. *Phytochemistry*, **41**, 1047–1055.

Davis, G.D. and Essenberg, M. (1995) (+)-δ-cadinene is a product of sesquiterpene cyclase activity in cotton. *Phytochemistry*, **39**, 553–567.

Davis, G.D., Essenberg, M., Berlin, K.D., Faure, R. and Gaydou, E.M. (1996) Complete ^{1}H and ^{13}C NMR spectral assignment of δ-cadinene, a bicyclic sesquiterpene hydrocarbon. *Magn. Reson. Chem.* **34**, 156–161.

Durst, F. and Nelson, D.R. (1995) Diversity and evolution of plant

P450 and P450-reductases. *Drug Metab. Drug Interact.* **12**, 189–206.

Essenberg, M., Grover, P.B., Jr and Cover, E.C. (1990) Accumulation of antibacterial sesquiterpenoids in bacterially inoculated *Gossypium* leaves and cotyledons. *Phytochemistry*, **29**, 3107–3113.

Frank, M.R., Deyneka, J.M. and Schuler, M.A. (1996) Cloning of wound-induced cytochrome P450 monooxygenases expressed in pea. *Plant Physiol.* **110**, 1035–1046.

Funk, C. and Croteau, R. (1994) Diterpenoid resin acid biosynthesis in conifers: characterization of two cytochrome P450-dependent monooxygenases and an aldehyde dehydrogenase involved in abietic acid biosynthesis. *Arch. Biochem. Biophys.* **308**, 258–266.

Hedden, P. and Kamiya, Y. (1997) Gibberellin biosynthesis. enzymes, genes, and their regulation. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **48**, 431–460.

Hefner, J., Rubenstein, S.M., Ketchum, R.E., Gibson, D.M., Williams, R.M. and Croteau, R. (1996) Cytochrome P450-catalyzed hydroxylation of taxa-4(5),11(12)-diene to taxa-4(20),11(12)-dien-5alpha-ol. the first oxygenation step in taxol biosynthesis. *Chem. Biol.* **3**, 479–489.

Humphreys, J.M., Hemm, M.R. and Chapple, C. (1999) New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. *Proc. Natl Acad. Sci. USA*, **96**, 10045–10050.

Jung, W., Yu, O., Lau, S.-M.C., O'Keefe, D.P., Odell, J., Fader, G. and McGonigle, B. (2000) Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes. *Nature Biotechnol.* **18**, 208–212.

Karp, F., Mihaliak, C.A., Harris, J.L. and Croteau, R. (1990) Monoterpene biosynthesis. specificity of the hydroxylations of (–)-limonene by enzyme preparations from peppermint (*Mentha piperita*), spearmint (*Mentha spicata*), and perilla (*Perilla frutescens*) leaves. *Arch. Biochem. Biophys.* **276**, 219–226.

de Kraker, J.-W., Franssen, M.C.R., Dalm, M.C.F., de Groot, A. and Bouwmeester, H.J. (2001) Biosynthesis of germacrene A carboxylic acid in chicory roots. Demonstration of a cytochrome P450 (+)-germacrene A hydroxylase and $NADP^+$-dependent sesquiterpenoid dehydrogenase(s) involved in sesquiterpene lactone biosynthesis. *Plant Physiol.* **125**, 1930–1940.

Kraus, P.F. and Kutchan, T.M. (1995) Molecular cloning and heterologous expression of a cDNA encoding berbamunine synthase, a C-O phenol-coupling cytochrome P450 from the higher plant *Berberis stolonifera*. *Proc. Natl Acad. Sci. USA*, **92**, 2071–2075.

Liu, C.J., Heinstein, P. and Chen, X.Y. (1999) Expression pattern of genes encoding farnesyl diphosphate synthase and sesquiterpene cyclase in cotton suspension-cultured cells treated with fungal elicitors. *Mol. Plant–Microbe Interact.* **12**, 1095–1104.

Liu, J., Benedict, C.R., Stipanovic, R.D. and Bell, A.A. (1999) Purification and characterization of *S*-adenosyl-L-methionine:desoxyhemigossypol-6-*O*-methyltransferase from cotton plants. An enzyme capable of methylating the defense terpenoids of cotton. *Plant Physiol.* **121**, 1017–1024.

Logemann, E., Tavernaro, A., Schulz, W., Somssich, I.E. and Hahlbrock, K. (2000) UV light selectively coinduces supply pathways from primary metabolism and flavonoid secondary product formation in parsley. *Proc. Natl Acad. Sci. USA*, **97**, 1903–1907.

Lupien, S., Karp, F., Ponnamperuma, K., Wildung, M. and Croteau, R. (1995) Cytochrome P450 limonene hydroxylases of *Mentha* species. *Drug Metab. Drug Interact.* **12**, 245–260.

Lupien, S., Karp, F., Wildung, M. and Croteau, R. (1999) Regiospecific cytochrome P450 limonene hydroxylases from mint (*Mentha*) species. cDNA isolation, characterization, and functional expression of (–)-4*S*-limonene-3-hydroxylase and (–)-4*S*-limonene-6-hydroxylase. *Arch. Biochem. Biophys.* **368**, 181–192.

Martens, S. and Forkmann, G. (1999) Cloning and expression of flavone synthase II from *Gerbera* hybrids. *Plant J.* **20**, 611–618.

Meng, Y.L., Jia, J.W., Liu, C.J., Liang, W.Q., Heinstein, P. and Chen, X.Y. (1999) Coordinated accumulation of (+)-δ-cadinene synthase mRNAs and gossypol in developing seeds of *Gossypium hirsutum* and a new member of the cad1 family from *G. arboreum*. *J. Nat. Prod.* **62**, 248–252.

Mihaliak, C.A., Karp, F. and Croteau, R. (1993) Cytochrome P-450 terpene hydroxylases. *Methods Plant Biochem.* **9**, 261–279.

Nelson, D.R. and Strobel, H.W. (1988) On the membrane topology of vertebrate cytochrome P-450 proteins. *J. Biol. Chem.* **263**, 6038–6050.

Nielsen, J.S. and Møller, B.L. (1999) Biosynthesis of cyanogenic glucosides in *Triglochin maritima* and the involvement of cytochrome P450 enzymes. *Arch. Biochem. Biophys.* **368**, 121–130.

Pierce, M.L., Cover, E.C., Richardson, P.E., Scholes, V.E. and Essenberg, M. (1996) Adequacy of cellular phytoalexin concentrations in hypersensitively responding cotton leaves. *Physiol. Mol. Plant Pathol.* **48**, 305–324.

Pompon, D., Louerat, B., Bronine, A. and Urban, P. (1996) Yeast expression of animal and plant P450s in optimized redox environments. *Methods Enzymol.* **272**, 51–64.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning. A. Laboratory Manual*, 2nd edn. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Schalk, M., Nedelkina, S., Schoch, G., Batard, Y. and Werck-Reichhart, D. (1999) Role of unusual amino acid residues in the proximal and distal heme regions of a plant P450, CYP73A1. *Biochemistry*, **38**, 6093–6103.

Schoendorf, A., Rithner, C.D., Williams, R.M. and Croteau, R.B. (2001) Molecular cloning of a cytochrome P450 taxane, 10b-hydroxylase cDNA from *Taxus* and functional expression in yeast. *Proc. Natl Acad. Sci. USA*, **98**, 1501–1506.

Steele, C.L., Gijzen, M., Qutob, D. and Dixon, R.A. (1999) Molecular characterization of the enzyme catalyzing the aryl migration reaction of isoflavonoid biosynthesis in soybean. *Arch. Biochem. Biophys.* **367**, 146–150.

Tan, X.P., Liang, W.Q., Liu, C.J., Luo, P., Heinstein, P. and Chen, X.Y. (2000) Expression pattern of (+)-δ-cadinene synthase genes and biosynthesis of sesquiterpene aldehydes in plants of *Gossypium arboreum* L. *Planta*, **210**, 644–651.

EMBL and GenBank accession number AF332974(*CYB706B1*)