Geraniol 10-hydroxylase[1], a cytochrome P450 enzyme involved in terpenoid indole alkaloid biosynthesis

Graziella Collu[a,b], Nehir Unver[a,b], Anja M.G. Peltenburg-Looman[a,b], Robert van der Heijden[a], Robert Verpoorte[a], Johan Memelink[b,*]

[a] *Division of Pharmacognosy, Leiden/Amsterdam Center for Drug Research, Leiden University, Gorlaeus Laboratories, P.O. Box 9502, 2300 RA Leiden, The Netherlands*
[b] *Institute of Molecular Plant Sciences, Clusius Laboratory, Wassenaarseweg 64, 2333 AL Leiden, The Netherlands*

Received 7 September 2001; accepted 15 October 2001

First published online 26 October 2001

Edited by Marc Van Montagu

Abstract Geraniol 10-hydroxylase (G10H) is a cytochrome P450 monooxygenase involved in the biosynthesis of iridoid monoterpenoids and several classes of monoterpenoid alkaloids found in a diverse range of plant species. *Catharanthus roseus* (Madagascar periwinkle) contains monoterpenoid indole alkaloids, several of which are pharmaceutically important. Vinblastine and vincristine, for example, find widespread use as anticancer drugs. G10H is thought to play a key regulatory role in terpenoid indole alkaloid biosynthesis. We purified G10H from *C. roseus* cells. Using degenerate PCR primers based on amino acid sequence information we cloned the corresponding cDNA. The encoded CYP76B6 protein has G10H activity when expressed in *C. roseus* and yeast cells. The stress hormone methyljasmonate strongly induced *G10h* gene expression coordinately with other terpenoid indole alkaloid biosynthesis genes in a *C. roseus* cell culture.

Key words: Alkaloid; Cytochrome P450; Geraniol 10-hydroxylase; Indole alkaloid; Methyljasmonate; Secondary metabolism; Terpenoid indole alkaloid; *Catharanthus roseus*

1. Introduction

The enzyme geraniol 10-hydroxylase (G10H) is a cytochrome P450 monooxygenase (P450), which hydroxylates the monoterpenoid geraniol at the C-10 position (Fig. 1) [1]. This reaction forms the first committed step in the formation of iridoid monoterpenoids in a diverse range of plant species. In some plant species, iridoid monoterpenoids accumulate as major end products, whereas in others, the iridoids are further metabolized to form different classes of monoterpenoid alkaloids of mixed biosynthetic origin, such as the terpenoid indole, the pyrroloquinoline, the quinoline, and the terpenoid tetrahydroisoquinoline alkaloids.

The terpenoid indole alkaloids (TIAs) are found in a number of plant species belonging to the plant families Apocynaceae, Loganiaceae, Rubiaceae and Nyssaceae. TIAs form a large group of structurally diverse molecules that include compounds with pharmaceutically interesting activities. Several TIAs are used in modern medicine, such as the anti-neoplastic agents vincristine and vinblastine, and the anti-hypertensive drugs reserpine and ajmalicine and the anti-arrhythmic drug ajmaline. Central to the biosynthesis of all TIAs is the intermediate strictosidine formed by coupling of tryptamine and secologanin, which are derived directly from the amino acid tryptophan and via several steps from (10-hydroxy-)geraniol, respectively (Fig. 1). Strictosidine is the branch point from which biosynthesis routes diverge towards the various alkaloids found in different TIA-producing plant species [2,3]. Several studies using *Catharanthus roseus* indicate that the formation of secologanin can be a limiting factor in TIA production [4,5]. G10H was suggested to have a regulatory effect on alkaloid production [6,7]. Insight into the regulatory role of G10H requires a thorough understanding of its expression characteristics and enzymatic properties, and would benefit greatly from the availability of a *G10h* cDNA clone.

G10H cloning has been previously attempted using different approaches. Differential screening of a *C. roseus* cDNA library resulted in the isolation of two nearly identical P450 cDNAs (CYP72A1) [8]. Hybridization with *CYP72A1* cDNA detected an mRNA species with an induction pattern similar to that of G10H activity in alkaloid production medium. However, when CYP72A1 was expressed in yeast, no activity towards geraniol or other P450 substrates tested was found [8]. Recently, CYP72A1 was shown to be involved in the terminal step of secologanin biosynthesis from its immediate precursor loganin [9]. In another approach to clone the *G10h* cDNA, a set of degenerate oligonucleotide primers based on the conserved heme-binding domain of P450 proteins was used for polymerase chain reaction (PCR) amplification on reverse-transcribed mRNA. Partial cDNA sequences of 16 different P450s from *C. roseus* were isolated [10], but functions could not be attributed to any of the corresponding proteins. Application of a similar PCR strategy resulted in the isolation of multiple cDNAs encoding P450s from *Arabidopsis*

*Corresponding author. Fax: (31)-71-5275088.
E-mail address: memelink@rulbim.leidenuniv.nl (J. Memelink).

[1] The nucleotide sequence newly reported in this paper has been deposited in the GenBank/EMBL data banks with the accession number AJ251269

Abbreviations: CaMV, cauliflower mosaic virus; CPR, cytochrome P450 reductase; GC/MS, gas chromatography/mass spectrometry; G10H, geraniol 10-hydroxylase; MeJA, methyljasmonic acid; P450, cytochrome P450 monooxygenase; ORF, open reading frame; PCR, polymerase chain reaction; Rps9, ribosomal protein S9; *Sgd*, strictosidine β-D-glucosidase; *Sls*, secologanin synthase; *Str*, strictosidine synthase; *Tdc*, tryptophan decarboxylase; TIA, terpenoid indole alkaloid

0014-5793/01/$20.00
PII: S0014-5793(01)03045-9

thaliana [11]. One of the encoded P450 proteins, assigned as CYP76C1, was expressed using the baculovirus expression system, and systematic exploration of its function revealed that it can hydroxylate geraniol [12]. It is unclear why *A. thaliana* would possess a G10H enzyme, since it is not known to produce secondary metabolites derived from iridoid monoterpenoids. One possibility is that geraniol hydroxylation by CYP76C1, although detectable in vitro, is not its physiological function in vivo.

Here we report the molecular cloning and the functional expression of the cDNA encoding G10H from a *C. roseus* cell culture. This clone was isolated by a conventional approach consisting of purification and sequencing of the enzyme to design oligonucleotide probes for cDNA library screening via PCR. This approach resulted in the isolation of a full-length cDNA clone, encoding a protein that is a new member of the CYP76B plant P450 subfamily. Expression in transgenic *C. roseus* and yeast cells proved that the CYP76B6 protein catalyzes the hydroxylation of geraniol and nerol at their C-10 position.

2. Materials and methods

2.1. Chemicals

Geraniol (>99.5%) was obtained from Fluka, nerol (95%) from Roth, 10-hydroxygeraniol (97%) from Aldrich and 10-hydroxynerol was synthesized as described previously [13]. All other chemicals were of the highest purity available. Organic solvents were of analytical grade.

2.2. Cell cultures

C. roseus (L.) G. Don cell suspension lines $9Cr58A_{12}A_2$ [14] and MP183L [15] were handled as described.

2.3. G10H purification and enzyme assay

G10H was purified according to [16] with the following minor modifications (Table 1). The membrane fraction containing G10H activity was isolated from the crude cellular homogenate via differential centrifugation between 200 and $20\,000 \times g$. Other minor changes were the omission of the hydroxyapatite chromatography step, and the use of the same basic buffer (50 mM Tris–HCl pH 7.8, 1 mM EDTA, 1 mM DTT, 15% (v/v) glycerol) for all chromatographic purification steps. G10H activities were determined as described [14] after reconstitution of the P450 fraction with NADPH–cytochrome P450 reductase (CPR) and a crude lipid fraction, which were both prepared according to [16]. Cytochrome P450 concentrations were determined according to [16].

2.4. Protein gel electrophoresis and internal amino acid sequence analysis

Gel electrophoresis was performed using the PhastSystem Electrophoresis Apparatus (Amersham Pharmacia Biotech). SDS–PAGE was carried out using pre-cast PhastGel Gradient 10–15 and isoelectric focussing in pre-cast PhastGel IEF 3–9. After electrophoresis the proteins were visualized by silver staining. Molecular mass and isoelectric point were determined using the Pharmacia LMW calibration kit and the p*I* calibration kit, respectively. Internal sequence analysis of the purified G10H enzyme was performed by Eurosequence B.V. (Groningen, The Netherlands).

2.5. cDNA cloning of G10h

A cDNA library was constructed in the λ-ZAPII vector (Stratagene) with 5 μg poly-adenylated RNA from a 5-day-old $9Cr58A_{12}A_2$ cell culture grown in induction medium [14]. PCR was performed with 10^9 phage from the cDNA library, 200 pmol of a degenerate primer (Table 2) and 25 pmol of the vector-specific T7 primer. Reactions were heated for 5 min at 94°C, followed by 35 cycles of 1.5 min at 94°C, 1.5 min at 30°C, and 2 min at 72°C. PCR with a new primer (G10H5), based on the sequence of a 620 bp fragment obtained with primer G10H3, in combination with the vector-specific T3 primer using the same conditions, except that the annealing temperature was raised to 55°C, amplified a 1160 bp fragment. The full-length open reading frame (ORF) was amplified with primers G10H7 (CCCTCTAGAGATCTCCATGGATTACCTTACCATAATAT-TAAC) and G10H8 (CCCAGATCTCATAACGCGAAATAGGA-CAGTTTGG) and cloned in pGEM-T (Promega). Its sequence was submitted to the P450 nomenclature committee for the assignment of a P450 family name (D.R. Nelson, WebPage address http://drnelson.utmem.edu/nelsonhomepage.html).

2.6. Expression in transgenic yeast and C. roseus cells

The *CYP76B6* ORF was excised as a *Xba*I/*Sac*I or *Bgl*II fragment from pGEMT-G10H and cloned in the *Xba*I/*Sac*I sites of the yeast expression vector YCpIF1 [17] or the *Bam*HI site of plant expression vector pMOG463, respectively. The *CYP76B6* derivatives and the corresponding empty vectors were introduced into *Saccharomyces cerevisiae* strain YPH500 [18] via the standard lithium acetate procedure, or into *C. roseus* cell line MP183L via particle bombardment [19]. Transgenic yeast strains were grown in 250 ml of selective synthetic dropout (SD) medium with 2% glucose as a carbon source for 5 days at 30°C, and GAL1 promoter-controlled gene expression was then induced by shaking the pelleted and washed cells for 1 day at 30°C in 250 ml of selective SD medium with 2% galactose as a sole carbon source. Transgenic *C. roseus* MP183L cells were cultured as described [15], and harvested after 7 days. Cells were extracted for G10H enzyme assays as described [14].

2.7. Mass spectrometry

The G10H reaction mixture with protein extract from *C. roseus* cell line G17, prepared according to [14], was subjected to capillary gas chromatography/mass spectrometry (GC/MS), using a Varian Saturn 2000 ion trap mass spectrometer run in electron ionization mode. The reaction mixture dissolved in methanol was separated on a Varian 3800 gas chromatograph equipped with a DB-5 capillary column (30 m × 0.25 mm, film thickness of 0.25 μm) (J&W Scientific) using helium as carrier gas at a flow rate of 1.2 ml/min. The separation conditions were: injection split ratio 1:50, injector temperature 250°C, initial oven temperature 50°C for 0.5 min, then linear gradient to 300°C at a rate of 10°C/min. The mass scan range was 50–300 u with a scan range time of 1 s.

2.8. Northern blot analysis

Three-day-old *C. roseus* cells (MP183L) were treated for different lengths of time with 50 μM methyljasmonic acid (MeJA) (Bedoukian Research) diluted in Me_2SO (0.1% final concentration of Me_2SO). RNA extractions and Northern blot analysis were performed as described [20]. DNA probes for tryptophan decarboxylase (*Tdc*), strictosidine synthase (*Str*), strictosidine β-D-glucosidase (*Sgd*), 40S ribosomal protein S9 (*Rps9*) [21] and *Cpr* [22] were described.

Table 1
Purification of the cytochrome P450 enzyme G10H from *C. roseus*

Purification step	Protein (mg)	P450 (nmol)	G10H (nkat)	Purification (fold)	Yield (%)
Membrane fraction	834.5	127.7	301.0	1.00	100.0
Solubilizate	212.5	22.8	40.5	0.53	13.4
DEAE–Sephacel	35.7	10.6	10.4	0.81	3.5
ω-Aminooctyl agarose	11.1	6.4	4.5	1.11	1.5
TSK Phenyl-5PW	0.8	0.9	2.4	8.53	0.8

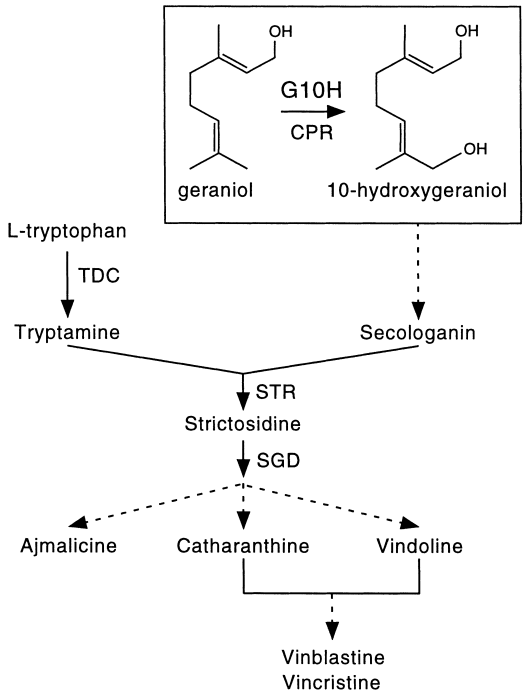

Fig. 1. Biosynthetic route to TIAs in *C. roseus*. Dashed arrows indicate multiple enzymatic steps.

3. Results

3.1. Purification of G10H

Screening of cell lines available in our laboratory in different plant media resulted in the selection of the *C. roseus* cell line $9Cr58A_{12}A_2$, which, when grown in induction medium [16], exhibited a high specific G10H activity of 240 pkat/mg protein. Using this cell line as an enzyme source, G10H was purified according to [16] with a few minor modifications, leading to a highly purified preparation (Table 1; Fig. 2). Differential centrifugation of the crude homogenate resulted in a recovery of 87% of the total G10H activity in the 200–$20\,000 \times g$ pellet. The protein content of this membrane fraction was low (835 mg membrane proteins from 500 g of *C. roseus* cells with a specific content of P450 of 0.15 nmol/mg protein) with a high specific G10H activity of 350 nkat/mg protein.

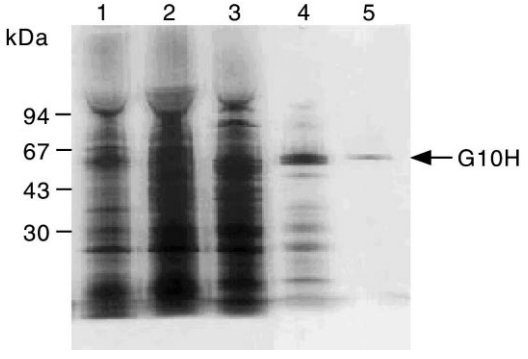

Fig. 2. SDS–PAGE of protein samples from different G10H purification steps visualized by silver staining. Lane 1, microsomal fraction; lane 2, cholate solubilizate; lane 3, DEAE–Sephacel; lane 4, ω-aminooctyl agarose; lane 5, TSK Phenyl-5PW. Positions of size markers are indicated.

The cholate-solubilized membrane fraction was further purified by a number of chromatographic steps according to [16] with omission of the hydroxyapatite ultragel chromatographic step (Table 1). SDS–PAGE of the fraction with the highest G10H activity from the last chromatographic step (TSK Phenyl-5PW) showed a single protein band on SDS–PAGE with a molecular mass of 56.3 kDa (Fig. 2, lane 5). This molecular mass is close to the values reported for other plant P450s (between 47 and 57 kDa). The p*I* of this protein was 8.7 as determined by isoelectric focusing. The resulting G10H preparation had a specific P450 content of 1.2 nmol/mg protein (Table 1). In a reconstituted system with NADPH–CPR and 100 μg crude *C. roseus* lipids the purified G10H enzyme was able to hydroxylate geraniol (3.1 nkat/mg protein) and its *cis*-isomer nerol (1.9 nkat/mg protein).

Amino acid sequencing of tryptic fragments of the purified G10H protein provided the amino acid sequence IEES-DINRLPYLR. Comparison of this sequence with amino acid sequences from the EMBL database revealed that it had homology to plant P450s. The obtained amino acid sequence appeared to be located in the C-terminal half of a typical P450 protein.

```
ACAAATTGTACACTTCCATTCCATGGATTACCTTACCATAATATTAACTTTACTATTTGCCTTGACTCTCTATGA    75
                  M  D  Y  L  T  I  I  L  T  L  L  F  A  L  T  L  Y  E    18

AGCCTTCAGCTACCTATCCAGAAGAACCAAAAACCTTCCTCCAGGACCATCGCCATTGCCGTTCATCGGAAGCCT   150
 A  F  S  Y  L  S  R  R  T  K  N  L  P  P  G  P  S  P  L  P  F  I  G  S  L    43
                                     PROLINE-RICH REGION

CCATTTATTAGGCGACCAACCACACAAATCCTTAGCAAAACTTTCCAAAAAACACGGTCCAATTATGAGTCTCAA   225
 H  L  L  G  D  Q  P  H  K  S  L  A  K  L  S  K  K  H  G  P  I  M  S  L  K    68

ATTAGGCCAGATCACTACAATCGTCATATCTTCATCAACAATGGCGAAAGAAGTTCTTCAAAAACAGGATTTAGC   300
 L  G  Q  I  T  T  I  V  I  S  S  S  T  M  A  K  E  V  L  Q  K  Q  D  L  A    93

ATTTTCAAGCAGATCAGTTCCAAACGCACTCCACGCTCACAATCAATTCAAATTCTCCGTTGTATGGCTTCCGGT   375
 F  S  S  R  S  V  P  N  A  L  H  A  H  N  Q  F  K  F  S  V  V  W  L  P  V   118

AGCCTCACGATGGAGAAGTCTTCGAAAAGTTTTGAATTCTAATATATTTTCCGGCAATCGGCTCGACGCTAATCA   450
 A  S  R  W  R  S  L  R  K  V  L  N  S  N  I  F  S  G  N  R  L  D  A  N  Q   143

ACATTTGAGAACTAGAAAAGTACAGGAACTAATTGCGTATTGCCGGAAAAATAGCCAGAGCGGAGAAGCGGTTGA   525
 H  L  R  T  R  K  V  Q  E  L  I  A  Y  C  R  K  N  S  Q  S  G  E  A  V  D   168
                                  ← G10H6 ←
CGTCGGCCGAGCTGCTTTTAGAACTTCGTTGAATTTGTTGTCGAATTTGATTTTTTCAAAGGATTTGACGGATCC   600
 V  G  R  A  A  F  R  T  S  L  N  L  L  S  N  L  I  F  S  K  D  L  T  D  P   193

TTATTCGGATTCTGCCAAGGAATTCAAGGATTTGGTTTGGAATATAATGGTTGAGGCGGGGAAACCTAATTTGGT   675
 Y  S  D  S  A  K  E  F  K  D  L  V  W  N  I  M  V  E  A  G  K  P  N  L  V   218

CGATTTTTTCCCCTGCTTGAAAAAGTTGATCCTCAAGGTATACGACATCGTATGACGATTCACTTTGGGGAAGT   750
 D  F  F  P  L  L  E  K  V  D  P  Q  G  I  R  H  R  M  T  I  H  F  G  E  V   243

TCTTAAGCTTTTTGGTGGACTTGTTAATGAAAGATTGGAGCAAAGAAGATCAAAAGGGGAAAAAAATGATGTGTT   825
 L  K  L  F  G  G  L  V  N  E  R  L  E  Q  R  R  S  K  G  E  K  N  D  V  L   268

GGATGTACTTCTAACTACCAGCCAAGAAAGCCCTGAGGAAATCGATAGAACTCACATTGAGCGAATGTGCTTGGA   900
 D  V  L  L  T  T  S  Q  E  S  P  E  E  I  D  R  T  H  I  E  R  M  C  L  D   293

CCTGTTTGTAGCAGGGACGGACACAACATCAAGCACATTAGAATGGGCAATGTCAGAAATGCTTAAAAACCCAGA   975
 L  F  V  A  G  T  D  T  T  S  S  T  L  E  W  A  M  S  E  M  L  K  N  P  D   318
           CENTRAL HELIX I

CAAAATGAAGAAAACCCAAGATGAACTTGCACAAGTAATCGGCAGAGGAAAAACAATAGAAGAATCCGATATTAA  1050
 K  M  K  K  T  Q  D  E  L  A  Q  V  I  G  R  G  K  T  I  E  E  S  D  I  N   343
→G10H3→
CCGCTTACCTTACTTAAGATGCGTTATGAAAGAAACCTTAAGGATACATCCACCAGTTCCCTTCTTAATTCCTCG  1125
 R  L  P  Y  L  R  C  V  M  K  E  T  L  R  I  H  P  P  V  P  F  L  I  P  R   368
 ← G10H5 ←
CAAAGTGGAACAAAGTGTTGAGGGTTTGTGGATACAATGTCCCTAAAGGATCACAAGTTCTTGTGAATGCTTGGGC  1200
 K  V  E  Q  S  V  E  V  C  G  Y  N  V  P  K  G  S  Q  V  L  V  N  A  W  A   393

AATTGGACGTGATGAAACTGTTTGGGATGATGCTTTGGCATTCAAACCCGAGAGATTTATGGAATCTGAATTGGA  1275
 I  G  R  D  E  T  V  W  D  D  A  L  A  F  K  P  E  R  F  M  E  S  E  L  D   418

TATCCGTGGAAGAGATTTCGAGCTGATTCCGTTCGGTGCTGGCCGAAGAATTTGCCCAGGGTTGCCATTGGCACT  1350
 I  R  G  R  D  F  E  L  I  P  F  G  A  G  R  R  I  C  P  G  L  P  L  A  L   443
                            HEME BINDING DOMAIN

AAGGACTGTGCCTTTGATGCTTGGTTCCTTTGTTGAACTCTTTTAATTGGAAGCTTGAAGGTGGGATGGCTCCAAA  1425
 R  T  V  P  L  M  L  G  S  L  L  N  S  F  N  W  K  L  E  G  G  M  A  P  K   468

AGATTTGGATATGGAGGAGAAGTTTGGTATTACACTGCAGAAGGCTCATCCTTTGCGTGCTGTACCAAGCACCCT  1500
 D  L  D  M  E  E  K  F  G  I  T  L  Q  K  A  H  P  L  R  A  V  P  S  T  L   493

TTAAAATTCTTGCCCTTTTTTTTTTTAAGCTTGCATAAAATAAGGGCAGTGCCAAACTGTCCTATTTCGCGTTAT  1575

GTACTCTCTCCCTCATGGTTGGAATACGTTTAATATATGTATTCATCTTTTTTTTTTTAAAAAAAAAAAAAAAAA  1650

AAAAAAAAAAAAA                                                              1664
```

Fig. 3. Nucleotide sequence of the *CYP76B6* cDNA encoding G10H from *C. roseus*. The deduced amino acid sequence is written underneath the nucleotide sequence. The annealing sites of the primers used for PCR are indicated above the nucleotide sequence and underlined, with the direction of priming shown by the arrows. Conserved domains found in plant P450s are boxed. The underlined N-terminal amino acid sequence was previously determined from a G10H preparation by [23].

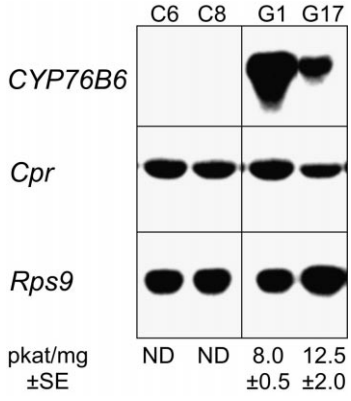

Fig. 4. Analysis of transgenic *C. roseus* cell lines expressing CYP76B6. Two transgenic MP183L control (C) cell lines and two lines transformed with CaMV 35S-*CYP76B6* (G) were analyzed for mRNA levels corresponding to *CYP76B6*, *Cpr*, and *Rps9*. G10H activities measured in these cell lines are shown at the bottom in pkat/mg protein ± S.E.M..

3.2. cDNA cloning of G10H

Based on the obtained amino acid sequence several degenerate oligonucleotide primers were designed. G10H1 and G10H2 corresponded to the first part of the internal sequence, and G10H3 and G10H4 to the last part (Table 2). Since PCR on available cDNA libraries from *C. roseus* roots or cell line MP183L did not result in specific amplification products, we constructed a λ-ZAPII cDNA library of the cell line $9Cr58A_{12}A_2$ used for G10H purification. Only the G10H3 primer in combination with the vector-specific T7 primer resulted in the amplification of a PCR fragment of 620 bp with this library. The deduced amino acid sequence contained the sequence PFGAGRRICPG that corresponds to the highly conserved heme-binding domain PFGxGRRxCxG found in most plant P450s. A new primer (G10H5) was developed based on the nucleotide sequence of this PCR fragment and used for PCR on the cDNA library in combination with the vector-specific T3 primer. The deduced amino acid sequence of the resulting 1160 bp PCR fragment contained the entire internal sequence (IEESDINRLPYLR) as determined from the purified G10H enzyme. Furthermore, it contained the sequence AGTDTT, which is highly conserved in plant P450s and which is thought to interact with both the substrate and the iron-bound oxygen. These findings indicated that the PCR fragments corresponded to a P450 protein containing the internal amino acid sequence. Combining the sequences of the two PCR fragments indicated that the complete coding sequence for a P450 protein was probably present. To ensure that the sequence was complete at the N-terminus, another primer (G10H6) was designed based on the 1160 bp PCR fragment and used in a PCR reaction with the vector-specific T3 primer. This resulted in the amplification of a single 555 bp fragment with identical sequence to the 1160 bp fragment except that it contained a few additional bp at the 5′ untranslated end. Noteworthy, the N-terminal amino acid sequence deduced from the P450 cDNA exactly matched the N-terminal sequence as determined from a G10H preparation purified by [23], providing additional evidence that the cDNA clone contains a complete ORF. As shown in Fig. 3, the isolated P450 cDNA sequence contained a 1482 bp ORF encoding a protein of 493 amino acids with a molecular mass of 55.7 kDa, which is consistent with the apparent molecular mass of 56.3 kDa estimated for purified G10H (Fig. 2). In addition, a relatively basic p*I* of 9.1 was predicted, which is also in close agreement with the p*I* of 8.7 for purified G10H as determined by isoelectric focussing.

Comparison of the deduced amino acid sequence with plant P450 sequences from the GenBank/EMBL database showed amino acid identities of 39–67% with P450s of the CYP76 family. Highest identities ranging from 53 to 67% were observed with the CYP76B subfamily. According to the P450 nomenclature system, the cloned P450 protein belongs to the CYP76B subfamily and was designated CYP76B6.

3.3. Expression of CYP76B6 in transgenic C. roseus and yeast cells

To test the enzymatic function of the *CYP76B6* protein, its full-length coding region was amplified by PCR with specific primers. The ORF was cloned behind the cauliflower mosaic virus (CaMV) 35S promoter in the plant expression vector pMOG463. The *CYP76B6* construct, as well as the empty expression vector, were introduced together with a selectable hygromycin resistance marker in the *C. roseus* cell line MP183L by particle bombardment. Cell line MP183L had no detectable G10H activity (data not shown), did not have detectable levels of *CYP76B6* gene expression under normal growth conditions [5], and a cDNA library prepared from elicited MP183L cells did not allow the amplification of a PCR product corresponding to the *CYP76B6* cDNA (see above). Therefore, cell line MP183L appeared to be a suitable background for testing CYP76B6 enzymatic activity. Cell suspension cultures established from hygromycin-resistant calli were screened for high expression levels of the co-transformed CaMV 35S-*CYP76B6* transgene by Northern blot hybridization. Cell lines G1 and G17 contained a high level of *CYP76B6* transcript, whereas control cell lines C6 and C8 did not contain detectable *CYP76B6* mRNA amounts (Fig. 4). All cell lines showed equally high levels of *Cpr* expression (Fig. 4), indicating that NADPH–CPR activity is unlikely to be a limiting factor for the activity of any P450 enzyme in this cell line. Protein extracts prepared from cell lines G1 and G17 showed G10H activities of 8.0 ± 0.5 and 12.5 ± 2.0 pkat/mg

Table 2
DNA sequences of the degenerate primers aligned with the peptide sequence from which they were designed

Primer	Amino acid sequence and codon selection						Degeneracy
	I	E	E	S	D	I	
G10H1	5′-ATH	GAR	GAR	WSN	GAY	AT -3′	256
G10H2	5′-ATH	GAR	GAR	WSI	GAY	AT -3′	64
	N	R	L	P	Y	L	
G10H3	5′-AAY	MGN	YTN	CCN	TAY	YT -3′	2048
G10H4	5′-AAY	MGI	YTI	CCI	TAY	YT -3′	32

Coding redundancies are H = T, C or A; R = A or G; W = T or A; S = C or G; Y = C or T; M = A or C; N = C,T,A or G; I = inosine.

protein, respectively, whereas G10H activities in protein extracts from the control cell lines were undetectable. GC/MS profiles of a G10H activity assay with G17 protein extract showed the formation of one product with the same retention time as authentic 10-hydroxygeraniol, which was not observed in a control assay with C6 extract (Fig. 5A–C). No other peaks were detected when compared to the control incubation. In both incubations an intermediate peak with a retention time of 11.8 min was formed, which was also formed in a blank incubation without protein extract (data not shown), and therefore did not represent an enzymatically derived product from geraniol. Mass spectral analysis positively identified the product running at a retention time of 12.8 min as 10-hydroxygeraniol by comparison with a commercial standard (Fig. 5D,E). These results unequivocally identify CYP76B6 as the elusive G10H. Expression of CYP76B6 in the yeast *S. cerevisiae* gave essentially similar results (data not shown), albeit that G10H activity levels were lower

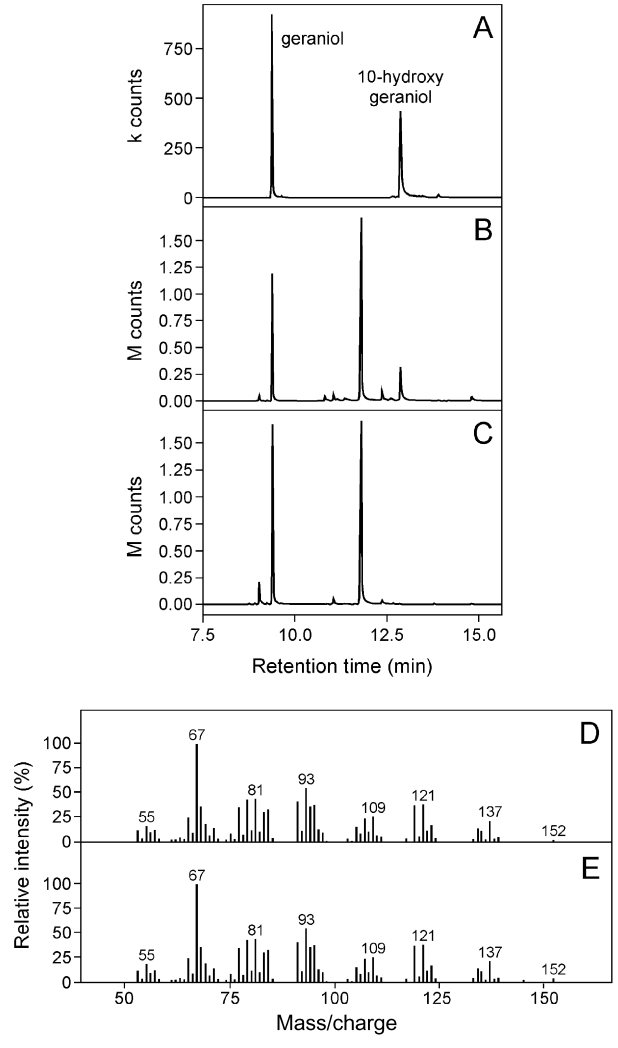

Fig. 5. Analysis of the CYP76B6-catalyzed reaction product with the substrate geraniol. GC-MS analysis of authentic geraniol and 10-hydroxygeraniol (A), and reaction products with the substrate geraniol and protein extracts from cell line G17 expressing CYP76B6 (B) and from control cell line C6 (C). Mass spectra of authentic 10-hydroxygeraniol (D) and the peak from B with retention time 12.8 min (E).

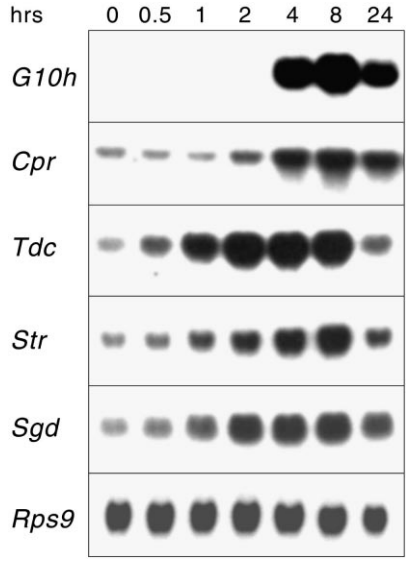

Fig. 6. Methyljasmonate induces *G10h* gene expression coordinately with other TIA biosynthetic genes. *C. roseus* cell line MP183L was treated for various times (indicated in h) with 50 μM MeJA. Northern blots were hybridized with *G10h*, *Cpr*, *Tdc*, *Str*, *Sgd* and *Rps9* cDNAs.

(1.5 ± 0.2 pkat/mg protein). Also nerol was hydroxylated to 10-hydroxynerol by extracts from CYP76B6-expressing yeast, consistent with the substrate specificity of purified *C. roseus* G10H. The results further indicate that yeast CPR is able to support the G10H reaction to at least some degree. Control yeast cells containing the empty expression vector YCpIF1 had undetectable G10H activity. The lower G10H activity in CYP76B6-expressing yeast compared to *C. roseus* cells could be due to low expression or instability of the CYP76B6 protein, or to a less efficient interaction between the yeast CPR and the CYP76B6 protein.

3.4. Jasmonate-responsive expression of G10h

The plant stress hormone (methyl)jasmonic acid ((Me)JA) induces the formation of TIAs in cell suspension cultures [24] and in seedlings [25]. It does so by coordinately switching on the expression of multiple genes in the TIA biosynthetic pathway [5,20,21] and precursor pathways [5]. Therefore it was of interest to study the effect of MeJA on *G10h* expression. Cell line MP183L was treated for various times with 50 μM MeJA, and the expression of TIA biosynthetic genes including *G10h* was measured by Northern blot analysis. As shown in Fig. 6, *G10h* gene expression was undetectable in untreated cells. MeJA strongly induced *G10h* mRNA accumulation with a maximum level at 8 h. *Cpr* gene expression was detectable in unstimulated cells, and was also induced by MeJA with very similar kinetics as *G10h* gene expression, peaking at 8 h. The expression of three other TIA biosynthetic genes, *Tdc*, *Str* and *Sgd*, was induced with similar kinetics compared to each other, but slightly different from the kinetics of *G10h* and *Cpr* gene expression in that the induction was measurable at earlier time points from 1 h onwards. Control treatment with 0.1% Me_2SO had no effect on expression of TIA biosynthesis genes (data not shown). The level of *Rps9* mRNA, encoding the 40S Rps9, was not affected by MeJA, indicating that gene expression in general was not affected.

4. Discussion

Application of an established purification protocol for the P450 enzyme G10H from *C. roseus* with minor modifications yielded a highly purified protein preparation that enabled partial internal amino acid sequence determination, which served as a starting point for the cloning of the corresponding cDNA via PCR. G10H belongs to the CYP76B subfamily, and was designated CYP76B6. Amino acid identities between G10H and other P450s of the CYP76 family in the database ranged from 39 to 67%. Highest identities were observed with members of the CYP76B subfamily, which have unknown enzymatic activities. CYP76B6 had amino acid identities of 67% with CYP76B2 and 66% with CYP76B5 from *Petunia hybrida*, 61% with CYP76B1 from *Helianthus tuberosus*, and 53% with CYP76B3 and CYP76B4 from *Medicago sativa*.

Remarkably, G10H belongs neither to the CYP71 nor to the CYP72 family, as could be expected since both P450 families possess members which were thought to have G10H activity. CYP71A5 from *Nepeta racemosa* was speculated to be G10H [26], but this function was never directly confirmed. CYP72A1 from *C. roseus* was also thought to be G10H [8], but was recently shown to have secologanin synthase (*Sls*) activity [9]. Interestingly, G10H belongs to the CYP76 family of which also CYP76C1 is a member. CYP76C1 is a P450 enzyme from *A. thaliana* that is able to hydroxylate geraniol in vitro [12]. G10H shows 47% amino acid identity to CYP76C1, and especially in the C-terminal part of the protein the homology is quite high.

Analysis of the *G10h* gene expression showed that it was strongly induced by MeJA, together with the expression of a number of other TIA biosynthetic genes. Kinetics of *G10h* mRNA induction closely corresponded to those of *Cpr* mRNA. The undetectable level of expression in control cell cultures indicate that *G10h* gene expression is likely to be a limiting factor in TIA biosynthesis in cell suspension cultures.

Genes for a large number of enzymes of the TIA pathway have now been cloned. These include enzymes acting in the vindoline-specific branch of the pathway, such as the P450 enzyme tabersonine 16-hydroxylase [27], desacetoxyvindoline-4-hydroxylase [28] and acetyl-CoA:4-*O*-deacetylvindoline 4-O-acetyltransferase [29], as well as genes acting earlier in the pathway, such as *Tdc* [30], *Str* [15], *Cpr* [22], *Sls* [9] and *Sgd* [21]. The availability of the G10H cDNA now allows overexpression of combinations of *G10h* with other genes in plants and cell cultures. This may provide more insight in alkaloid biosynthesis and its regulation. Such studies may also uncover rate-limiting steps in TIA biosynthesis. In addition, the coordinated induction of many genes including *G10h* by MeJA indicates that strategies to improve TIA production using jasmonate-responsive transcription factors [5,31,32] are worthwhile to explore.

Acknowledgements: We thank A.H. Meijer for valuable advice on G10H purification, M. Soto Hernández for assistance with *C. roseus* tissue culture, A. Oudin for performing RNA blot hybridizations, and B. Pauw for assistance with GC-MS analysis. N. Unver was financially supported by TUBITAK-NATO (Turkish Scientific and Technical Research Council).

References

[1] Meehan, T.D. and Coscia, C.J. (1973) Biochem. Biophys. Res. Commun. 53, 1043–1048.
[2] Verpoorte, R., van der Heijden, R. and Moreno, P.R.H. (1997) Biosynthesis of Terpenoid Indole Alkaloids in *Catharanthus roseus* Cells, in: The Alkaloids, Vol. 49 (Cordell, G.A., Ed.), pp. 221–229, Academic Press, San Diego, CA.
[3] Kutchan, T.M. (1998) Molecular Genetics of Plant Alkaloid Biosynthesis, in: The Alkaloids, Vol. 50 (Cordell, G.A., Ed.), pp. 257–316, Academic Press, San Diego, CA.
[4] Whitmer, S., Canel, C., Hallard, D., Gonçalves, C. and Verpoorte, R. (1998) Plant Physiol. 116, 853–857.
[5] van der Fits, L. and Memelink, J. (2000) Science 289, 295–297.
[6] McFarlane, J., Madyastha, K.M. and Coscia, C.J. (1975) Biochem. Biophys. Res. Commun. 66, 1263–1269.
[7] Schiel, O., Witte, L. and Berlin, J. (1987) Z. Naturforsch. 42c, 1075–1081.
[8] Vetter, H.P., Mangold, U., Schröder, G., Marner, F.J., Werck-Reichhart, D. and Schröder, J. (1992) Plant Physiol. 100, 998–1007.
[9] Irmler, S., Schröder, G., St-Pierre, B., Crouch, N.P., Hotze, M., Schmidt, J., Strack, D., Matern, U. and Schröder, J. (2000) Plant J. 24, 797–804.
[10] Meijer, A.H., Souer, E., Verpoorte, R. and Hoge, J.H.C. (1993) Plant Mol. Biol. 22, 379–383.
[11] Mizutani, M., Ward, E. and Ohta, D. (1998) Plant Mol. Biol. 37, 39–52.
[12] Mizutani, M., Ward, E. and Ohta, D. (1997) Plant geraniol/nerol 10-hydroxylase and DNA coding therefor, PCT International Patent Application, WO 97/11184.
[13] Madyastha, K.M., Meehan, T.D. and Coscia, C.J. (1976) Biochemistry 15, 1097–1102.
[14] Collu, G., Bink, H.H.J., Moreno, P.R.H., van der Heijden, R. and Verpoorte, R. (1999) Phytochem. Anal. 10, 314–318.
[15] Pasquali, G., Goddijn, O.J.M., de Waal, A., Verpoorte, R., Schilperoort, R.A., Hoge, J.H.C. and Memelink, J. (1992) Plant Mol. Biol. 18, 1121–1131.
[16] Meijer, A.H., de Waal, A. and Verpoorte, R. (1993) J. Chromatogr. 635, 237–249.
[17] Foreman, P.K. and Davis, R.W. (1994) Gene 144, 63–68.
[18] Sikorski, R.S. and Hieter, P. (1989) Genetics 122, 19–27.
[19] van der Fits, L. and Memelink, J. (1997) Plant Mol. Biol. 33, 943–946.
[20] Menke, F.L.H., Parchmann, S., Mueller, M.J., Kijne, J.W. and Memelink, J. (1999) Plant Physiol. 119, 1289–1296.
[21] Geerlings, A., Martinez-Lozano Ibanez, M., Memelink, J., van der Heijden, R. and Verpoorte, R. (2000) J. Biol. Chem. 275, 3051–3056.
[22] Meijer, A.H., Lopes Cardoso, M.I., Voskuilen, J.T., de Waal, A., Verpoorte, R. and Hoge, J.H.C. (1993) Plant J. 4, 47–60.
[23] Meijer, A.H. (1993) Cytochrome P-450 and secondary metabolism in *Catharanthus roseus*, PhD thesis, Leiden University, Leiden.
[24] Gantet, P., Imbault, N., Thiersault, M. and Doireau, P. (1998) Plant Cell Physiol. 39, 220–225.
[25] Vazquez-Flota, F.A. and De Luca, V. (1998) Phytochemistry 49, 395–402.
[26] Clark, I.M., Forde, B.G. and Hallahan, D.L. (1997) Plant Mol. Biol. 33, 875–885.
[27] Schröder, G., Unterbusch, E., Kaltenbach, M., Schmidt, J., Strack, D., De Luca, V. and Schröder, J. (1999) FEBS Lett. 458, 97–102.
[28] Vazquez-Flota, F., De Carolis, E., Alarco, A.-M. and De Luca, V. (1997) Plant Mol. Biol. 34, 935–948.
[29] St-Pierre, B., Laflamme, P., Alarco, A.-M. and De Luca, V. (1998) Plant J. 14, 703–713.
[30] De Luca, V., Marineau, C. and Brisson, N. (1989) Proc. Natl. Acad. Sci. USA 86, 2582–2586.
[31] Menke, F.L.H., Champion, A., Kijne, J.W. and Memelink, J. (1999) EMBO J. 18, 4455–4463.
[32] Memelink, J., Verpoorte, R. and Kijne, J.W. (2001) Trends Plant Sci. 6, 212–219.