Phytochemistry 118 (2015) 204–215

Contents lists available at ScienceDirect

Phytochemistry

journal homepage: www.elsevier.com/locate/phytochem

Functional expression of a putative geraniol 8-hydroxylase by reconstitution of bacterially expressed plant CYP76F45 and NADPH-cytochrome P450 reductase CPR I from *Croton stellatopilosus* Ohba

Siriluk Sintupachee [a,b], Worrawat Promden [b,c], Nattaya Ngamrojanavanich [d], Worapan Sitthithaworn [e], Wanchai De-Eknamkul [b,f,*]

[a] Program in Biotechnology, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand
[b] Research Unit for Natural Product Biotechnology, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Bangkok 10330, Thailand
[c] Division of General Science, Faculty of Education, Buriram Rajabhat University, Buriram 31000, Thailand
[d] Department of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand
[e] Department of Pharmacognosy, Faculty of Pharmacy, Srinakharinwirot University, Nakonnayok 26120, Thailand
[f] Department of Pharmacognosy and Pharmaceutical Botany, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Bangkok 10330, Thailand

ARTICLE INFO

Article history:
Received 28 June 2014
Received in revised form 3 August 2015
Accepted 10 August 2015
Available online 20 August 2015

Keywords:
Croton stellatopilosus Ohba
Euphorbiaceae
Geraniol-8-hydroxylase
CYP76F45
P450 reductase type I
Functional expression
19-*E*-vallesamine
Monoterpenoid indole alkaloid

ABSTRACT

While attempting to isolate the enzyme geranylgeraniol 18-hydroxylase, which is involved in plaunotol biosynthesis in *Croton stellatopilosus* (Cs), the cDNAs for a cytochrome P450 monooxygenase (designated as CYP76F45) and an NADPH-cytochrome P450 reductase (designated as CPR I based on its classification) were isolated from the leaf. The *CYP76F45* and *CsCPR I* genes have open reading frames (ORFs) encoding 507- and 711-amino acid proteins with predicted relative molecular weights of 56.7 and 79.0 kDa, respectively. Amino acid sequence comparison showed that both CYP76F45 (63–73%) and CsCPR I (79–83%) share relatively high sequence identities with homologous proteins in other plant species. Phylogenetic tree analysis confirmed that CYP76F45 belongs to the CYP76 family and that CsCPR I belongs to Class I of dicotyledonous CPRs, with both being closely related to *Ricinus communis* genes. Functional characterization of both enzymes, each expressed separately in *Escherichia coli* as recombinant proteins, showed that only simultaneous incubation of the membrane-bound proteins with the substrate geraniol (GOH) and the coenzyme NADPH could form 8-hydroxygeraniol. The enzyme mixture could also utilize acyclic sesquiterpene farnesol (FOH) with a comparable substrate preference ratio (GOH:FOH) of 54:46. The levels of the CYP76F45 and CsCPR I transcripts in the shoots, leaves and twigs of *C. stellatopilosus* were correlated with the levels of a major monoterpenoid indole alkaloid, identified tentatively as 19-*E*-vallesamine, that accumulated in these plant parts. These results suggested that CYP76F45 and CPR I function as the enzyme geraniol-8-hydroxylase (G8H), which is likely to be involved in the biosynthesis of the indole alkaloid in *C. stellatopilosus*.

1. Introduction

Geraniol 8-hydroxylase (G8H, EC 1.14.13.152), previously known as geraniol 10-hydroxylase, is a membrane-bound cytochrome-P450 monooxygenase (CYP) belonging to the CYP76 family. It catalyzes the hydroxylation of geraniol (GOH) (**1**) at the C-8 position to form 8-hydroxygeraniol (**2**). G8H has been suggested to be a potential site for regulation of the production of secologanin, which is involved in the biosynthesis of monoterpenoid indole alkaloids and iridoid monoterpenoids (Hofer et al., 2013). The enzyme from *Catharanthus roseus* has also been shown to catalyze the 3′ hydroxylation of naringenin to produce eriodictyol in the flavonoid biosynthetic pathway (Sung et al., 2011).

The 8-hydroxylation reaction of GOH (**1**) catalyzed by G8H has been observed in only a few plant species. These include CYP76B6 isolated from *C. roseus* (Collu et al., 2001; Ginis et al., 2012; Meijer et al., 1993), CYP76B10 from *Swertia mussotii* (Wang et al., 2010), CYP76A4 from *Petunia x hybrida* (Tamaki et al., 2005) and CYP76C1 (Mizutani and Ohta, 1998) and CYP76C4 (Hofer et al., 2013) from *Arabidopsis thaliana*. However, although it has been suggested that GOH hydroxylation could be a common characteristic of CYP76s

* Corresponding author at: Department of Pharmacognosy and Pharmaceutical Botany, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Bangkok 10330, Thailand.
E-mail address: dwanchai@chula.ac.th (W. De-Eknamkul).

http://dx.doi.org/10.1016/j.phytochem.2015.08.005
0031-9422/

(Mizutani and Ohta, 1998), more evidence from different plant species is needed to confirm this conclusion.

CYPs are located in the endoplasmic reticulum, and their catalytic activities strictly rely on electron supply from NADPH-cytochrome P450 reductases (CPRs) (Rana et al., 2013). Due to the difficulties usually encountered in fractioning individual proteins from endogenous CYP pools, characterization of CYPs has primarily depended on gene cloning and heterologous expression in eukaryotic systems, such as *Pichia pastoris* yeast (Wang et al., 2010), insect cells (Sung et al., 2011), and *Saccharomyces cerevisiae* (Tamaki et al., 2005). The use of eukaryotic systems allows CYP enzyme activities to be defined in cell lysates or purified microsomes with endogenous or co-expressed CPRs (Sung et al., 2011). For prokaryotic systems, however, there have been fewer cases where *Escherichia coli* was successfully used to express functional CYPs and CPRs to characterize their activities (for review, (Duan and Schuler, 2006)).

The enzyme geranylgeraniol 18-hydroxylase, which catalyzes geranylgeraniol (GGOH) 18-hydroxylation to form plaunotol in *Croton stellatopilosus* Ohba (Tansakul and De-Eknamkul, 1998), has been of great interest. While searching for genes that encode geranylgeraniol 18-hydroxylase, cDNAs of putative *CsG8H's* (designated as *CYP76F45* by D. Nelson), and its putative catalytic partner *CsCPR* (designated as CPR I based on its classification), were identified during the gene cloning process. Both genes were expressed separately in the simple *E. coli* prokaryotic system and were catalytically active when the recombinant CYP76F45 and CsCPR I preparations were incubated together with NADPH. Here, bacterial expression of functionally active CYP76F45 and CsCPR I isreported, these believed to be involved in the hydroxylation of GOH (**1**) in *C. stellatopilosus* Ohba.

2. Results

2.1. CYP76F45 and CsCPR I gene cloning and phylogenetic analysis

Degenerated primers designed from the conserved regions covering various CYP-family clans known to be involved in hydroxylation of different substrates (Nelson and Werck-Reichhart, 2011) were used in this study. The clans included CYP51 (sterols), CYP71/76 (shikimate/monoterpenes), CYP72 (isoprenoid hormones), CYP74 (allene), CYP85 (sterol and acyclic terpenes), CYP86 (fatty acids) and CYP97 (carotenoids). Using the degenerate primers, partial fragments of the CYP97, CYP76 and CYP86 genes could be amplified from the cDNAs isolated from the *C. stellatopilosus* leaves. The partial fragments of CYP97 and CYP76, which are likely to be involved in terpenoid pathways, were used to obtain full-length genes. Both genes, designated as CYP97C27 and CYP76F45 following the rules for classification of P450s (Nelson, 2009), were subsequently expressed in *E. coli* and showed clear differences in their catalytic functions. CYP97C27 acted as a geranylgeraniol 18-hydroxylase (Sintupachee et al., 2014), and CYP76F45 acted as a putative geraniol 8-hydroxylase (this study).

The *CYP76F45* (GenBank ID: KF738255) and *CsCPR I* (ID: KF738254) cDNAs had 1521 and 2,300 bp open reading frames (ORFs) encoding 507- and 711-amino acid proteins, respectively. The sequences of both genes were then investigated for their monomeric molecular weights and pIs using the ExPASy-Compute pI/Mw tool (http://web.expasy.org/compute_pi/). CYP76F45 was predicted to have a molecular weight of 56.7 kDa and a pI of 8.63, whereas CsCPR I had a molecular weight of 78.7 kDa and a pI of 5.27. The obtained CYP76F45 showed an amino acid identity of 73% with *Ricinus communis* CYP76F36, 68% with *Vitis vinifera* CYP76C1, 63% with *Populus trichocarpa* CYP76F3, and 56% with *C. roseus* CYP76B6. CsCPR I showed an amino acid identity of 83% with *R. communis* CYP76F36, 79% with *P. trichocarpa* CYP76C1, 77% with *V. vinifera* CYP76F3, and 73% with *C. roseus* CYP76B6. All of the compared species belong to dicotyledonous CPR Class I. Typical conserved cofactor- and substrate-binding domain characteristics for both CYP76F45 (oxygen- and heme-binding sites, Fig. 1) and CsCPR I (two FMN-, FAD-, and NADPH-binding sites and one P450-binding site, Fig. 2) were identified.

Based on the phylogenetic analysis, it was confirmed that CYP76F45 was indeed closely related to *R. communis* CYP76F36 but far away from the *C. roseus* CYP76B6 on the phylogenetic tree (Fig. 3). Similarly, CsCPR I also appeared to be closely related to *R. communis* CPR I (Fig. 4).

2.2. Heterologous expression of CYP76F45 and CsCPR I in E. coli

The *CYP76F45* and *CsCPR I* genes were constructed separately in the pET-32a expression vector, and protein expression was placed under the control of a T7-promoter in the *E. coli* BL21 (DE3) strain. The CYP76F45 and CsCPR I recombinant proteins could each be detected via SDS–PAGE as a protein band with increasing intensity corresponding to the induction time-course (Fig. 5) and could be confirmed by Western blot (Fig. 5C). The molecular weights appeared to be 74 kDa for CYP76F45 and 96 kDa for CsCPR I, which include a histidine tag at the N-terminus of each recombinant protein. CYP76F45 reached a maximum expression level after induction with 1 mM IPTG at 30 °C for 18 h (Fig. 5B). CsCPR I reached a maximum expression level after induction with 1 mM IPTG at 30 °C for 40 h (Fig. 5A). The bacteria carrying the empty vector expressed no equivalent proteins.

2.3. Functional identification of CYP76F45 and CsCPR I

To examine whether recombinant CYP76F45 and CsCPR I were both catalytically active as G8Hs, an *in vitro* assay for determining enzyme activities and product detection was performed. The reaction mixtures containing either CYP76F45 or CsCPR I alone or both proteins together were incubated with NADPH and GOH (**1**) for 0 and 18 h. Each reaction mixture was then extracted with ethyl acetate and analyzed for product formation by TLC (visualized under UV light at 254 nm). In the presence of GOH (**1**) and NADPH, active formation of the reaction product could only be observed with the CYP76F45 and CsCPR I proteins together, although a minute amount of product could also be detected by adding CsCPR I alone. As shown in Fig. 6A, the substrate GOH (**1**) was almost completely converted to the product during the 18-h incubation. Absolutely no product formation was observed in either the *E. coli* transformed with the empty vector control or the boiled enzyme preparation. The formation of the enzymatic product was also confirmed by a time-course study using a more accurate technique of GC–MS for product detection and identification. The results of GC chromatograms clearly showed that the reaction product was detected with a continuous increase of its peak area from 2 to 18 h of the incubation. Its peak eluted at the retention time of 15.2 min appeared to have a pattern of mass fragmentation (*m*/*z* 67, 81, 109, 121 and 152, Fig. 6B) that was very similar to the reported mass fragmentation of 8-OH geraniol (**2**) (Collu et al., 2001; Sung et al., 2011). However, attempts to purify both CYP76F45 and CsCPR I from the 20,000 g microsomal fractions for kinetic studies were not successful due to the loss of enzyme activity in the process. Therefore, a preliminary kinetic study of CYP76F45 activity toward GOH (**1**) was performed using the unpurified CYP76F45 and CsCPR I preparations. The results showed that CYP76F45 had an apparent *Km* value of 0.066 μM and *Vmax* of 4.60 nmol min^{-1} (Fig. 6C).

In terms of substrate specificity, recombinant CYP76F45 was examined using three different but related acyclic terpenoid substrates, GOH (**1**), FOH (**3**) and GGOH (**5**), as well as naringenin

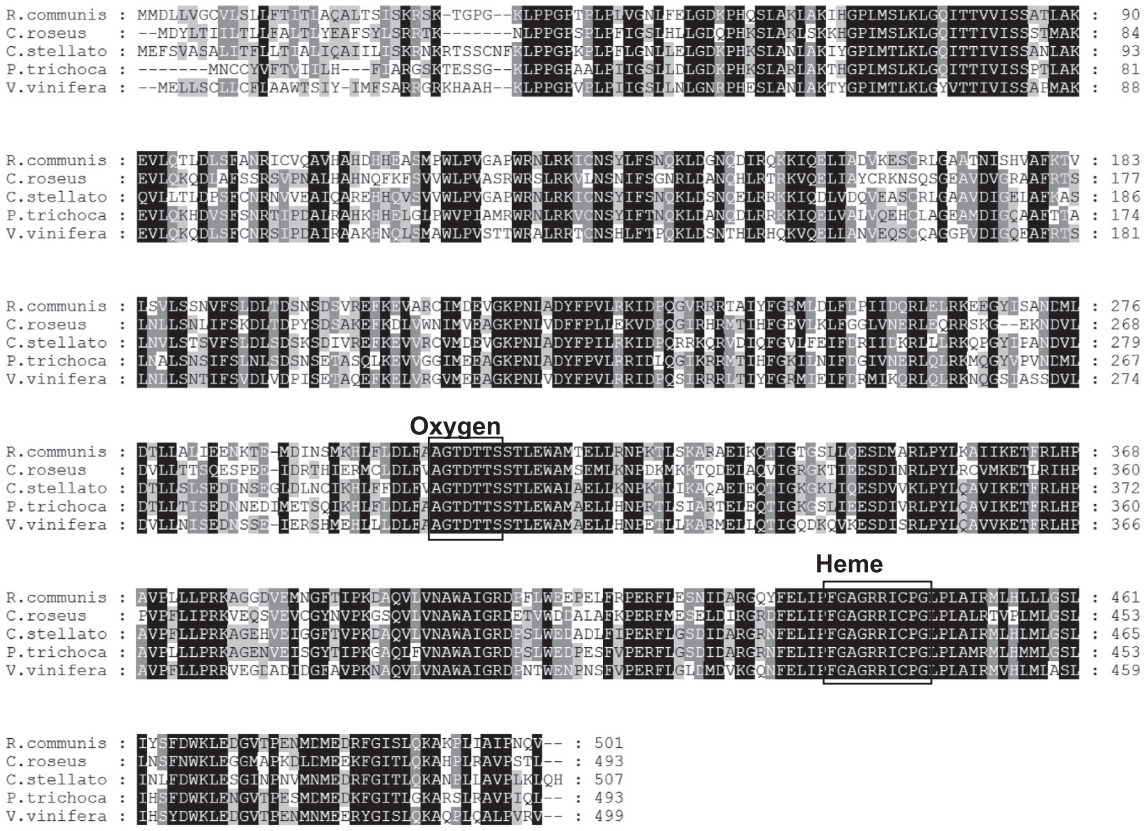

Fig. 1. Alignment of the CYP76 family amino acid sequences. The alignment profile was performed using the homologous amino acid sequences retrieved from other plant hosts using BLASTX (nucleotide translated blast) from the NCBI database with the GeneDoc program. The deduced amino acid sequence of CsG8H represent the CYP76s isolated from *C. stellatopilosus* (C. stellato; GenBank ID: KF738255), *R. communis* (R. communis; GenBank ID: XP002509653), *V. vinifera* (V. vinifera; GenBank: XP002281735), *P. trichocarpa* (P. trichocar; GenBank: XP002299579) and *C. roseus* (C. roseus; GenBank ID Q8VWZ7). The boxes indicate the cytochrome P450 characteristic motifs, which include heme-binding (FXXGXRXCXG) and oxygen-binding motifs (A/G)GX(D/E)T(T/S). Identical amino acid residues are shaded in black, and similar residues are shaded in gray. Gaps are inserted to maximize the homology.

(**7**), the flavanone previously reported to be a possible *C. roseus* G8H substrate (Sung et al., 2011). The same method of product detection by TLC was used to determine the enzyme activity toward the selected substrates. As shown in Fig. 7A, naringenin (**7**) could not be utilized by the CYP76F45 and CsCPR I enzyme mixture, while GOH (**1**), FOH (**3**) and GGOH (**5**) could all be converted with different activities to their corresponding products, 8-OH GOH (**2**), presumably 12-OH FOH (**4**) and 16-OH GGOH (**6**), respectively. The relative reaction rates, as determined from the intensity ratios of the enzymatic products after an 18-h incubation, appeared to be 50:42:8 for GOH (**1**), FOH (**3**) and GGOH (**5**), respectively (Fig. 7B). With a shorter incubation time of 4 h, the product formation rate was 54:46:0 (data not shown).

2.4. Transcription profile analysis of CYP76F45 and CPR I and detection of involved products

To understand the possible function of CYP76F45 and CPR I in *C. stellatopilosus*, the transcription profiles of the genes and the presence of potentially involved secondary products were examined. Three different plant parts including the shoot, leaf and twig were chosen for this study. Because G8H is known to be involved in the biosynthesis of monoterpenoid indole alkaloids, detection of this compound group in these three plant parts was also carried out. The results showed that both genes were expressed in these plant tissues, with the highest level of CYP76F45 expression in shoot and leaf followed by twig and the highest level of CPR I expression in the leaf followed by shoot and twig (Fig. 8A). Similarly, the amount of an alkaloid commonly present in these tissues, as detected by TLC using the specific Dragendorff's alkaloidal reagent, also showed the same pattern of product distribution (Fig. 8B). Identification by GC–MS established that the detected alkaloid, eluted at 54.5 min, had a pattern of mass fragmentation similar to that of the monoterpenoid indole alkaloid, 19-*E*-vallasamine (Fig. 8C), based on the comparison with the NIST libraries (NIST 2011) and the mass data reported previously (340 M+, 310, 267, 208, 180, 169, 154, 143, 130) (Nielsen et al., 1994; Zeches et al., 1987).

3. Discussion

In this study, the cDNAs for a P450, designated as CYP76F45, and a CPR, designated as CsCPR I, were successfully isolated from the young leaf portion of the *C. stellatopilosus* plant. The amino acid sequences deduced for both genes confirmed that the CYP76F45 had the characteristics of the plant cytochrome P450 monooxygenase superfamily (heme and oxygen binding) (Werck-Reichhart et al., 2002) and that CsCPR I had the characteristics of NADPH-cytochrome reductases (CYP and FMN, FAD and NADPH binding) (Jensen and Moller, 2010). The results of the phylogenetic analysis established that *C. stellatopilosus* CYP76F45 was closely related to *R. communis* CYP76F36, *P. trichocarpa* CYP76F3, and a series of recently reported *Santalum album* CYP76F[37-43] proteins (Diaz-

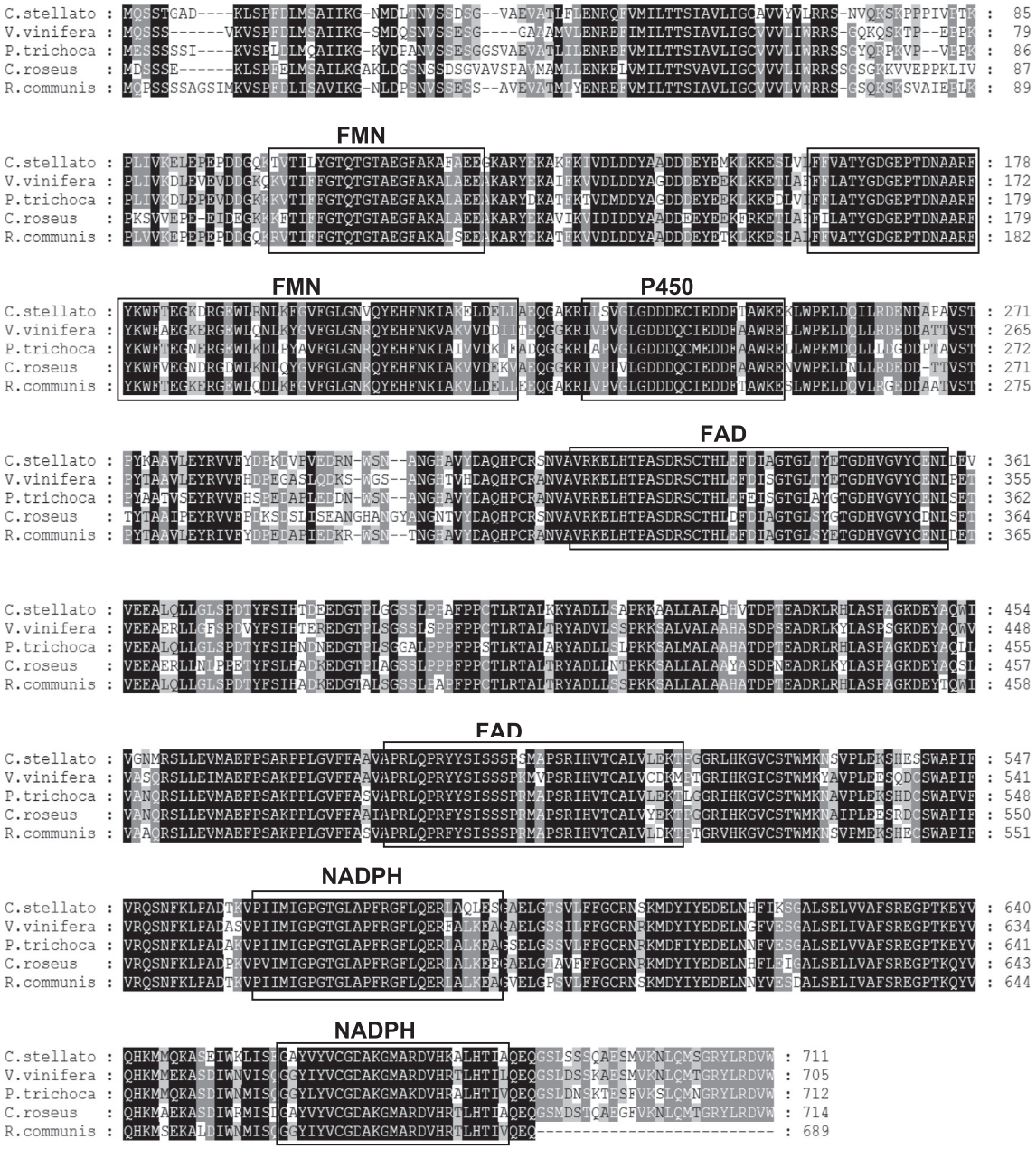

Fig. 2. Alignment of the plant CPR amino acid sequences. The alignment was performed using the homologous amino acid sequences retrieved from other plant hosts using BLASTX (nucleotide translated blast) from the NCBI database with the GeneDoc program. The deduced amino acids sequences of CsCPR represent the CPRs isolated from *C. stellatopilosus* (C. stellate; GenBank ID: KF738254), *V. vinifera* (V. vinifera; GenBank ID: CAN77401); *P. trichocarpa* (P. trichoca; GenBank ID: AAK15261|AF302498); *C. roseus* (C. roseus; GenBank ID: CAA49446) and *R. communis* (R. communis, GenBank ID: XP002534464). The boxes indicate the CPR characteristic motifs, which include two FMN-, FAD-, and NADPH-binding motifs and one P450-binding motif. Identical amino acid residues are shaded in black, and similar residues are shaded in gray. Gaps are inserted to maximize the homology.

Chavez et al., 2013) but was relatively distant from the well-known *C. roseus* CYP76B6. Similarly, CsCPR I was also closely related to the *R. communis* and *P. trichocarpa* CPRs, but more distant from the *C. roseus* CPR. All of the proteins used in this comparison belong to species in the dicotyledonous CPR Class I.

To date, many plant CYPs and CPRs have been expressed in eukaryotic yeasts and insect cells compared with prokaryotic bacteria (for review, see (Duan and Schuler, 2006)). *S. cerevisiae* yeast and *E. coli* bacterial systems are known to differ from one another with respect to the presence of endogenous CYPs and CPRs. *E. coli* has no endogenous CYP or CPR genes, whereas yeast contains three CYPs and one associated CPR gene (Barnes et al., 1991). Therefore, it has been hypothesized that the expression of unmodified plant CYPs in *E. coli* may be more complicated than in yeast because of the nature of the eukaryotic membrane-bound system, which is not present in bacteria (Chen et al., 2003; Zelasko et al., 2013). Our results, however, have shown that these possible drawbacks are not the case for functional expression of membrane-bound CYP76F45 and CsCPR I in *E. coli*. Both proteins were presumably expressed as plasma membrane-bound enzymes because *E. coli* does not possess cellular compartmentalization and their activities were only detected in the insoluble 20,000 g pellet fraction. The

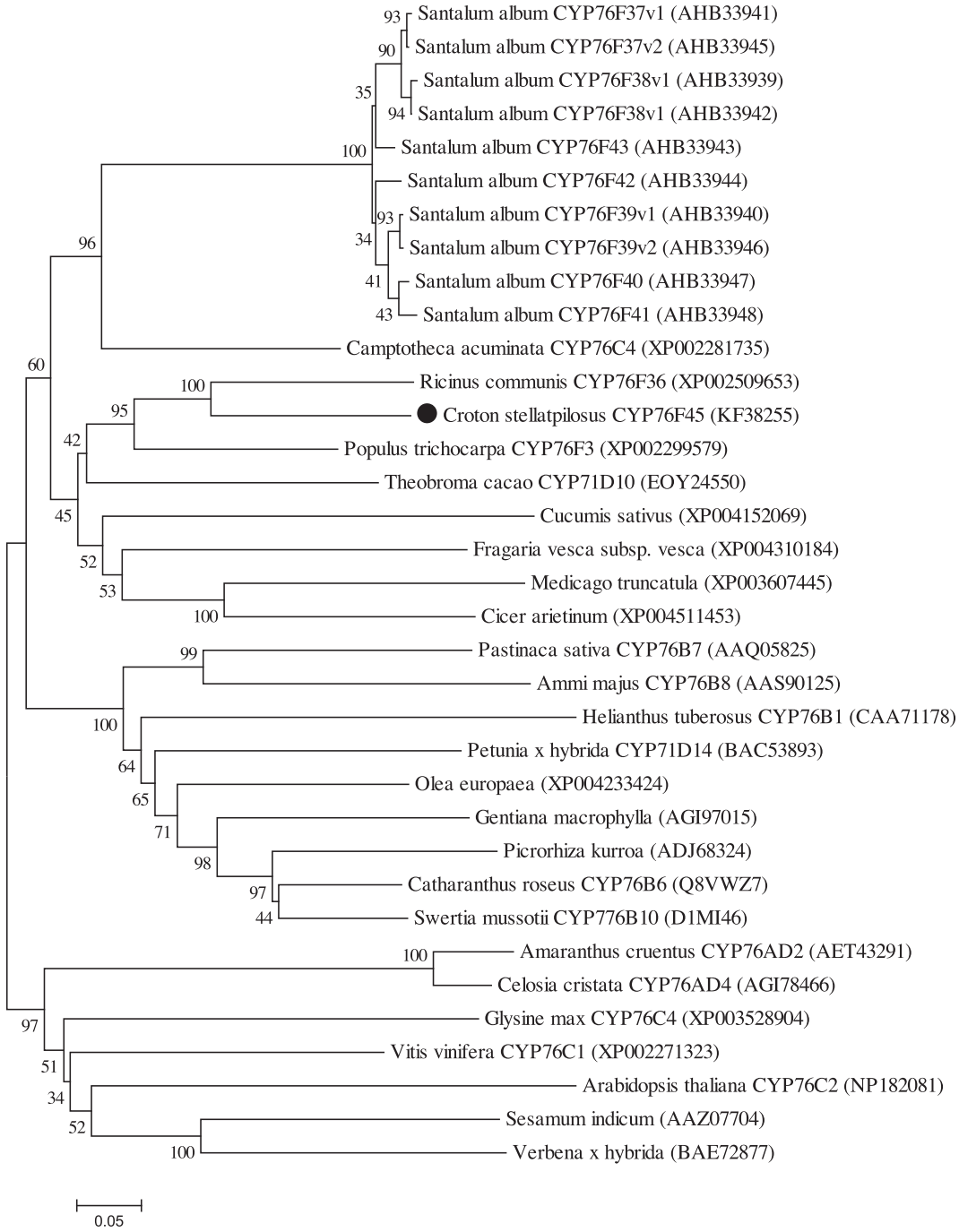

Fig. 3. Phylogenetic analysis of the CYP76 family based on 35 amino acid sequences. The mid-point root neighbor-joining (NJ) tree was generated using the MEGA6 program (http://www.megasoftware.net) (Saitou and Nei, 1987; Tamura et al., 2013). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein, 1985). The evolutionary distances were computed using the JTT matrix-based method (Jones et al., 1992a). The accession number of the CYP76 amino acid sequence from each species is represented at the end of the scientific name. The CYP76F45 isolated from *C. stellatopilosus* is indicated by a dark circle. The tree was drawn to scale with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree (scale bar).

successful functional expression might be because the *E. coli* expression system (pET32a/BL21 (DE3)) used in this study has been developed to avoid the toxic effects of overexpression associated with membrane proteins by using the T7 bacteriophage promoter in the bacterial BL21 strain (Zelasko et al., 2013). In addition to CYP76F54, other plant CYPs, including *A. thaliana* CYP79F1 and CYP79F2 (Chen et al., 2003; Hansen et al., 2001) and *C. stellatopilosus* CYP97C27 (Sintupachee et al., 2014), have been reported to have enzyme activities after their recombinant expression in *E. coli*.

Attempts to purify the obtained membrane-bound, His-tagged CYP76F45 and CsCPR I enzymes from the 20,000 g insoluble fraction (using 0.1% Triton X-100 and Ni–NTA resins) in active soluble forms were not successful. No enzyme activities were detected after the solubilization and purification steps. Therefore, the unpurified preparations of the 20,000 g insoluble fractions of both enzymes were used for the enzyme assays. The enzyme assays clearly demonstrated that only the preparation containing both CYP76F45 and CsCPR I could form the reaction product in the reaction mixture containing GOH (**1**) and NADPH. The results from GC–

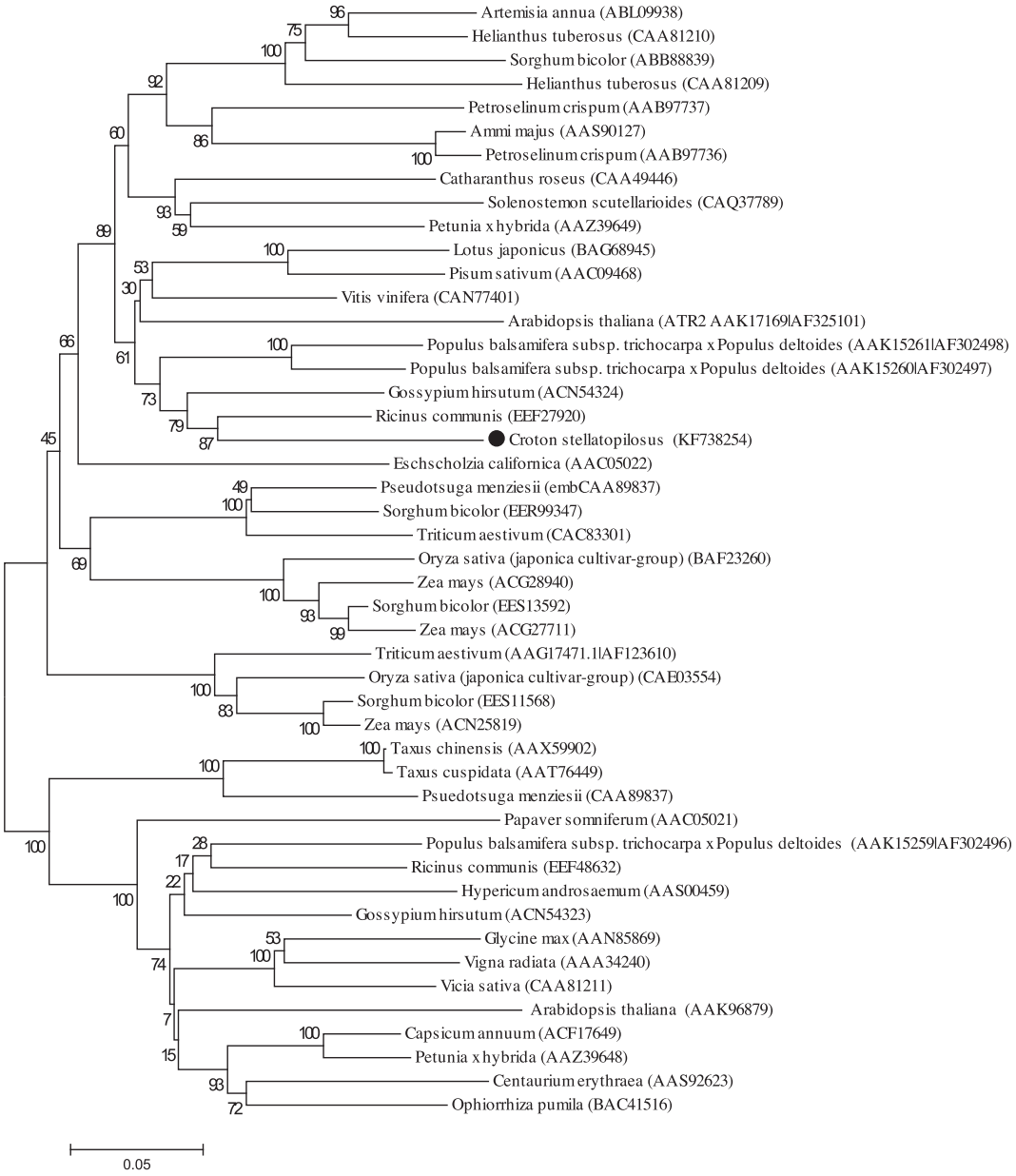

Fig. 4. Phylogenetic analysis of CPR family members based on 47 amino acid sequences (28 species). The mid-point root neighbor-joining (NJ) tree was constructed using the MEGA6 program (Saitou and Nei, 1987; Tamura et al., 2013). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein, 1985). The evolutionary distances were computed using the JTT matrix-based method (Jones et al., 1992b). The accession number of the CPR amino acid sequence from each species is represented at the end of the scientific name. The CPR isolated from *C. stellatopilosus* is indicated by a dark circle. The tree was drawn to scale with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree (the scale bar).

MS analysis confirmed that the geraniol 8-hydroxylation reaction took place in the incubation mixture based on the data of product identification reported previously (Collu et al., 2001; Sung et al., 2011). Therefore, it was concluded that CYP76F45/CsCPR I enzymes function as a G8H in *C. stellatopilosus*. As mentioned earlier, there have only been a few subfamilies of CYP76 reported able to catalyze geraniol 8-hydroxylation, including *C. roseus* and *S. mussotii* CYP76Bs, *P. hybrida* CYP76A and *A. thaliana* CYP76Cs. The *C. stellatopilosus* CYP76F described here is another CYP76 subfamily that is capable of catalyzing this reaction.

However, in addition to GOH (**1**), our results from the substrate specificity studies showed that the combined CYP76F45 and CsCPR I preparations could also utilize the related acyclic sesquiterpene FOH (**3**) as a substrate; GGOH (**5**) was also a usable substrate to a much lesser extent, while naringenin (**7**) was not a usable substrate. These enzyme properties were clearly demonstrated, but it should be noted that both recombinant CYP76F45 and CsCPR I carry additional N-terminal amino acids adding close to 20 kDa of various tags and proteolytic cleavage sites to the enzymes. Although both enzymes retained the capacity to support a function of the P450, this region constituted over 30% of the P450, and it cannot be excluded that the enzyme activity, specificity, or kinetics of CYP76F45/CsCPR I are possibly changed. Nevertheless, the findings that CYP76F45 has a relatively equal substrate preference between GOH (**1**) and FOH (**3**) have led us to question whether CYP76F45 functions physiologically in the metabolism of monoterpenoids, sesquiterpenoids or both. Recently, a closely related series of CYP76F subfamily members has been identified in *S. album* (Diaz-Chavez et al., 2013). In that study, 10 CYP76Fs, CYP76F37 to CYP76F43, were cloned, and nine were found to be functionally

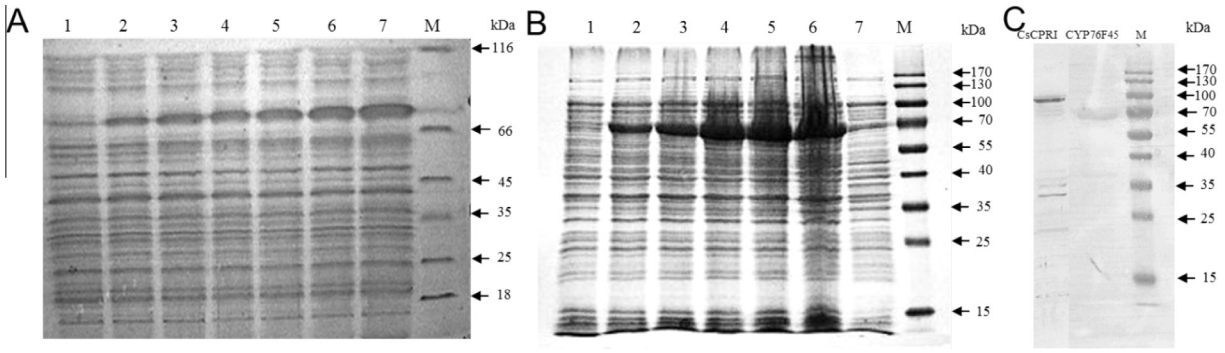

Fig. 5. Heterologous expression of the recombinant CsCPR I (A) and CYP76F45 (B) proteins in *E. coli* analyzed by 12% SDS–PAGE. (A) Lane 1, empty pET32a in *E. coli*; lanes 2–7 show protein patterns of cell lysates obtained after inducing with IPTG for 1, 2, 4, 6, 18 and 40 h, respectively. (B) Lane 1, empty pET32a in *E. coli*; lanes 2–6 show protein patterns of cell lysates obtained after inducing with IPTG for 1, 2, 6, 18 and 40 h, respectively; lane 7 shows the pattern of un-induced cell lysate for 40 h. Lane M shows the molecular markers. (C) Immunoblot with alkaline phosphatase (AP)-conjugated antibodies for the 6His-tag at the N-terminal of CsCPR I and the C-terminal of CYP76F45.

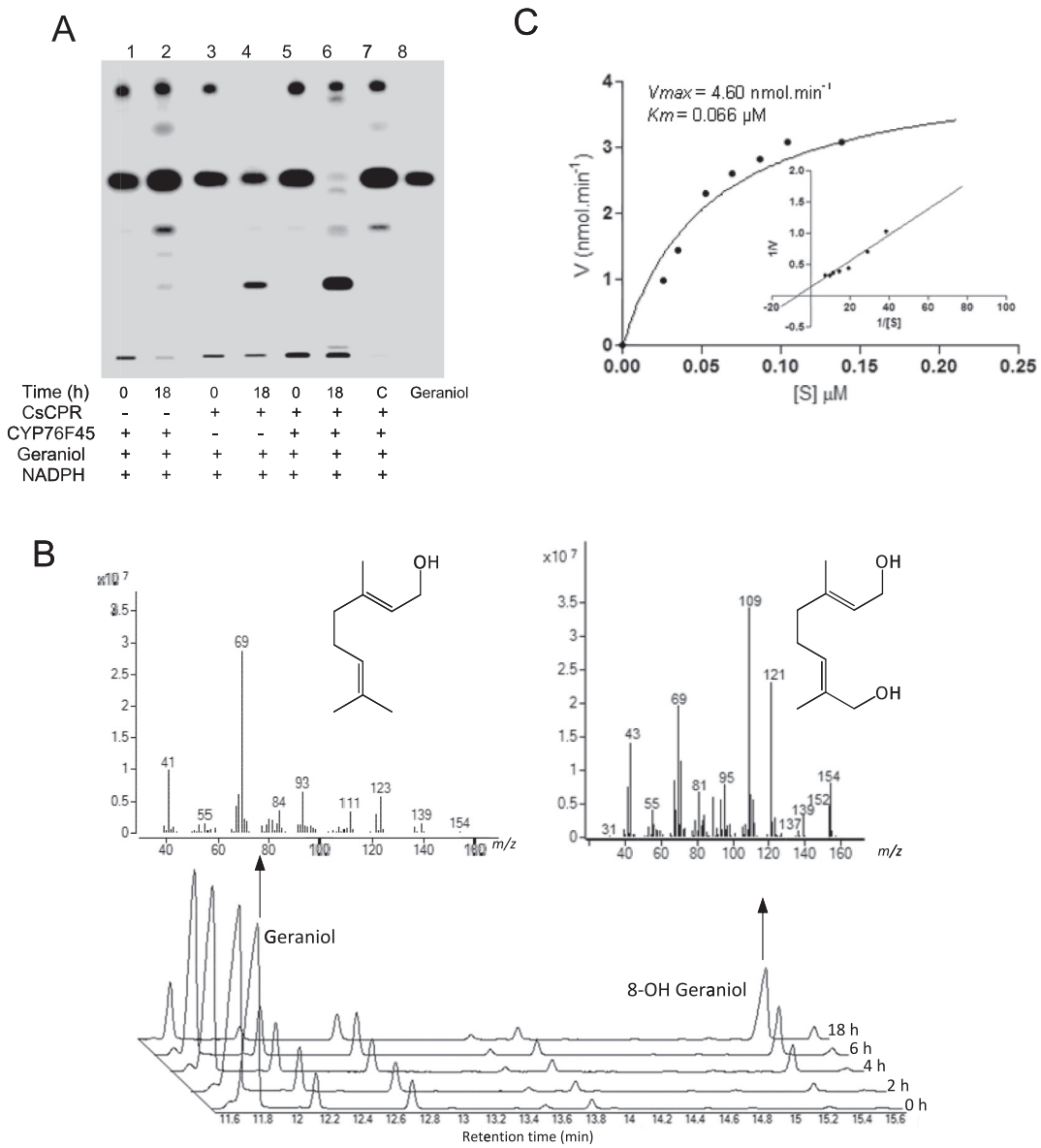

Fig. 6. (A) TLC detection of the CYP76F45 enzymatic product in the reaction mixtures in the presence and absence of CsCPR I under UV light at 254 nm. The microsomal fraction (250 µg) of CYP76F45, CsCPR I or both was used in each enzyme assay containing GOH (**1**) and NADPH. Each reaction mixture and the empty-vector control were incubated at 0 and 18 h at 30 °C before being analyzed by TLC. (B) GC–MS analysis showing the time-course formation of the CYP76F45/CsCPR I-catalyzed reaction product which was identified as 8-OH GOH (**2**). (C) Enzyme kinetic analysis of the CYP76F45/CsCPR I-catalyzed reaction using the substrate GOH (**1**). The *Km* and *Vmax* values for GOH (**1**) of the CYP76F45/CsCPR were determined based on the Michaelis–Menten plot (V_0 versus [S]) and Lineweaver–Burk plot (insert).

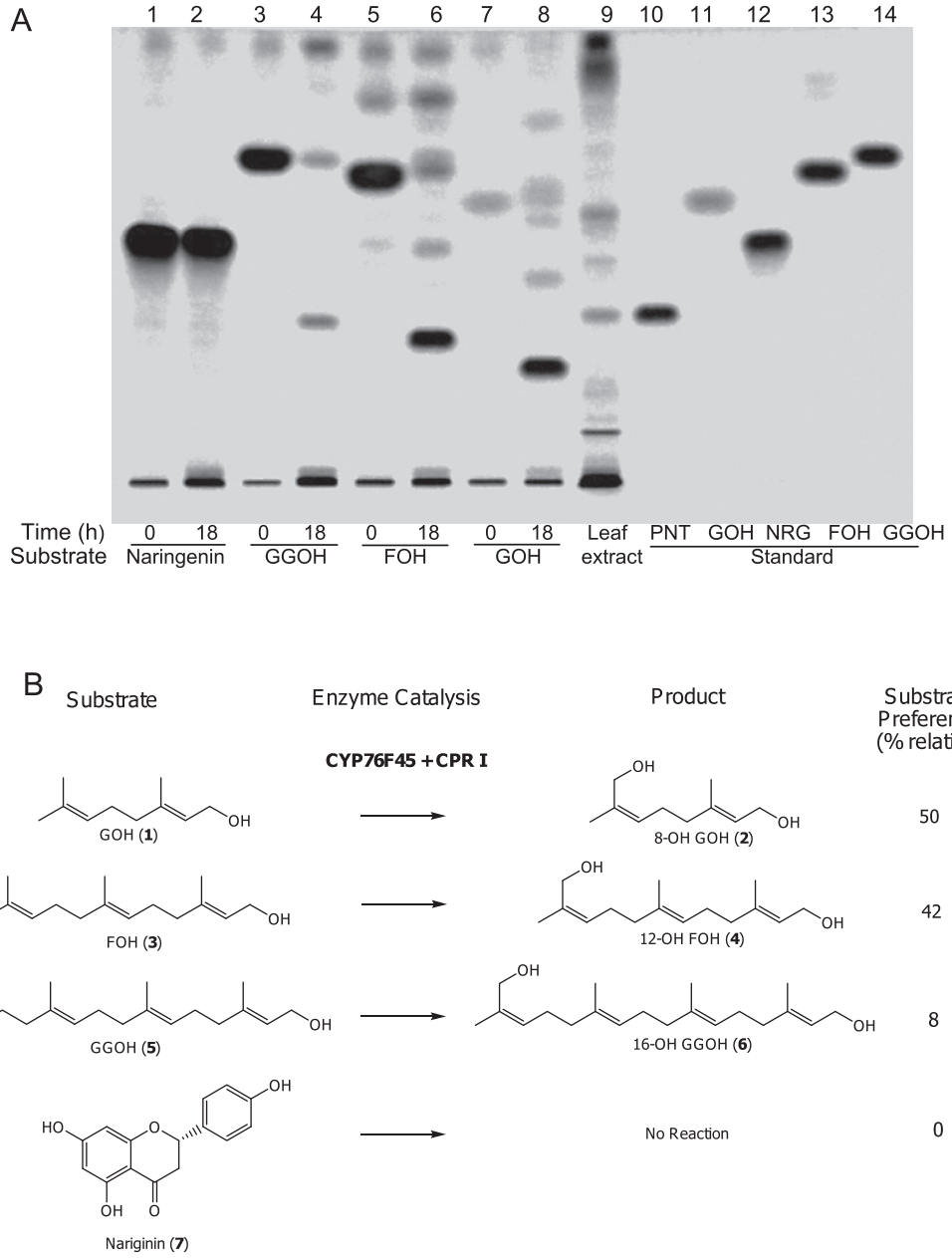

Fig. 7. (A) TLC detection at 254 nm of the CYP76F45/CsCPR I-catalyzed reactions using the substrates naringenin (NRG) (**7**), GOH (**1**), FOH (**3**), and GGOH (**5**). The reaction contained each substrate, CYP76F45, CsCPR I, and NADPH and was incubated at 30 °C for 18 h. (B) Utilization of some potential substrates by the enzymes CYP76F45/CsCPR and formation of potential products based on their catalytic reaction of GOH (**1**) 8-hydroxylation. The value of the relative substrate preference of each substrate was determined from the area of each product peak obtained from densitometric scanning (λ 210 nm) and was quantitated using the standard curve of its substrate.

active in hydroxylating sesquiterpenes santalene and bergamotene. This group of oxygenated-C15 compounds has been recently reported to be present as essential oil constituents in some *Croton* species (Turiel et al., 2013). However, our thorough analysis by GC–MS of the essential oil constituents in *C. stellatopilosus* leaves did not lead to the identification of any similar oxygenated-C15 compounds and thus does not support this possibility of FOH hydroxylation.

Therefore, in order to identify the secondary products that might potentially be biosynthesized through the function of CYP76F45/CsCPR I, further examination of the group of monoterpenoid indole alkaloids was performed to determine if the compounds are present in *C. stellatopilosus*. In the literature, there have been no reports on the isolation of this group of compounds in any *Croton* species, presumably due to their presence as minor constituents in the plant tissues. Therefore, in the study, a method of acid-base extraction for specific isolation of alkaloids from crude extracts was used. Interestingly, the resulting alkaloid preparation appeared to contain essentially a single type of alkaloid that could be visualized as a major band on TLC after spraying with the alkaloid-specific Dragendorff's reagent. This alkaloid band was subsequently analyzed by GC–MS and was found to have a mass spectrum that was similar to the known compound, 19-*E*-vallasamine (Nielsen et al., 1994; Walser and Djerassi, 1964; Zeches et al., 1987). Although complete elucidation is still needed to confirm the structure, its inclusion in the group of monoterpenoid indole alkaloids is highly possible. This is supported by the recent findings that the expressed sequence tag (EST)-based simple sequence repeat (SSR) markers in *C. roseus* have high cross-species transferability (31%) in a related *Croton* species, namely *C. macro-*

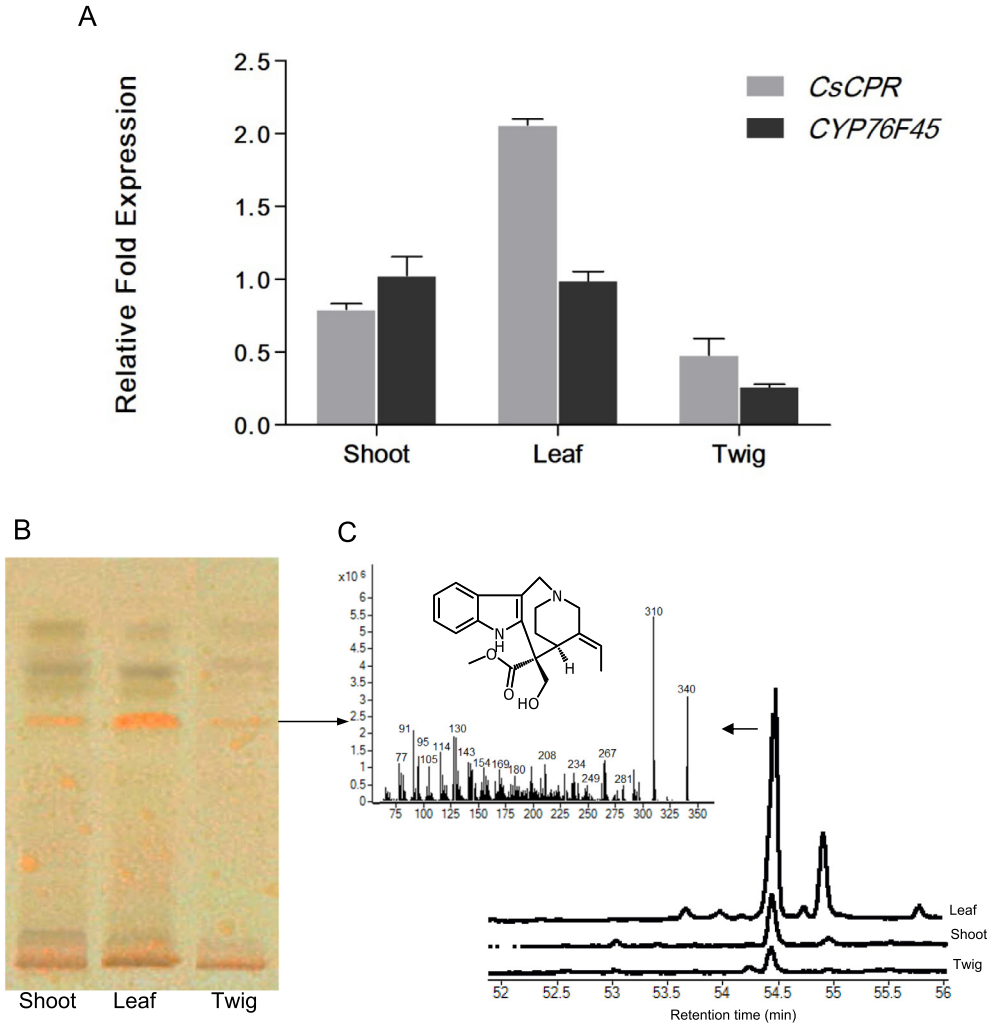

Fig. 8. (A) mRNA-expression levels of CYP76F45 and CsCPR I genes (relative to 18S rRNA) in shoot, leaf and twig of *C. stellatopilosus*. (B) TLC analysis showing the presence in all plant parts of an alkaloid visualized as orange bands after spraying the plate with Dragendorff's reagent. (C) GC–MS analysis showing the GC chromatograms and MS profile of the eluted alkaloid, which was identified tentatively as 19-*E*-vallesamine.

stachyus (Mishra et al., 2011). Among this transferability, some SSR loci appear to have homology with the enzymes strictosidine-D-glucosidase in monoterpenoid indole alkaloid metabolism, flavonoid-O-methyltransferase in flavonoid metabolism, and serine hydroxymethyltransferase/glycine hydroxymethyltransferase in primary metabolism. This information suggests that the actual function of the enzymes CYP76F45/CsCPR I as a G8H is in fact involved in the biosynthetic pathway of monoterpenoid indole alkaloids in *C. stellatopilosus*. In addition, the relationship between the putative 19-*E*-vallasamine and CYP76F45/CsCPR I was further supported by the results of real-time expression analysis of the two genes in the shoots, leaves, and twigs, which clearly showed correlation with the accumulation of the monoterpenoid indole alkaloids.

4. Concluding remarks

In conclusion, our observed enzymatic formation of 8-hydroxygeraniol from GOH (**1**) by the combined CYP76F45 and CsCPR I enzymes isolated from the same plant clearly confirms the importance of the cooperation between the CYP and CPR enzymes in catalyzing reactions in *C. stellatopilosus*. Whether both are localized closely to one other *in vivo* to catalyze such a hydroxylation reaction in the biosynthesis of the putative 19-*E*-vallesamine remains to be clarified.

5. Experimental

5.1. Plant collection

C. stellatopilosus Ohba (Cs) leaves were collected from an open field in the Faculty of Pharmaceutical Sciences, Chulalongkorn University, Bangkok, Thailand. The location of the plant is 13° 44′33″ North latitude, 100°31′49″ East longitude. The *C. stellatopilosus* (No. 184779) voucher specimen is deposited at the Office of the Forest Herbarium, Department of National Parks, Wildlife and Plant Conservation, Bangkok, Thailand.

5.2. Total-RNA isolation and cDNA synthesis

Total RNA was isolated from mature *C. stellatopilosus* green leaves by using the Trizol® method (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions and was purified using the RNeasy® Mini Kit (Qiagen, Valencia, CA). RNA was evaluated for quantity and quality at A_{260} and A_{280} and on agarose formaldehyde gels. The isolated RNA sample was used for cDNA

synthesis employing the RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific, Waltham, MA) with oligo-$dT_{(18)}$ primers.

5.3. PCR cloning of partial sequences of the CYP76F45 and CsCPR I genes from C. stellatopilosus

Partial sequences of both CYP76F45 and CsCPR I were amplified by using the degenerated primers CYP76F45-de-F/CYP76F45-de-R and CsCPR I-de-F/CsCPR I-de-R (Table 1), respectively. The degenerated primers were designed from the conserved regions of amino acid sequences from the CYP76 family (for CYP76F45) or cytochrome P450 reductases (for CsCPR I) from *V. vinifera*, *P. trichocarpa*, *C. roseus*, and *R. communis*.

PCR was performed according to the manufacturer's protocol using Platinum® Taq DNA Polymerase High Fidelity (Invitrogen, Carlsbad, CA) with the cDNA as the template. The PCR reaction was completed using a PCR MyCycler™ Thermal Cycler (Bio-Rad, Hercules, CA) for 30 cycles with the following reaction conditions: 30 s at 94 °C, 30 s at 55 °C and 1 min at 68 °C. PCR products were cloned into the pGEM-T Easy vector (Promega, Madison, WI), transformed into *E. coli* TOP10 cells (Invitrogen, Carlsbad, CA) and then confirmed by sequencing.

5.4. RACE-PCR

To obtain the full-length cDNA sequences, the Rapid Amplification of cDNA Ends polymerase chain reaction (RACE-PCR) technique was used to amplify the 3′- and 5′-termini of the CYP76F45 and CsCPR I genes. RACE-PCR was performed with the GeneRacer™ kit (Invitrogen, Grand Island, NY) according to the manufacturer's instructions. Gene-specific primers for 3′- and 5′-RACE were designed from the partial sequences of the CYP76F45 and CsCPR I genes (listed in Table 1).

Table 1
Primer sequences. Standard IUPAC ambiguity codes are underlined and the restriction enzyme sites are bolded.

Primer name	Sequence 5′→3′
Degenerated PCR	
CsCPR I-de-F	TTTARARTTGTTGATWTGGATGATTATGC
CsCPR I-de-R	GCCATRCYTTGGCATCACCACASACATA
CYP76F45-de-F	CTHGCCMARAWWYAYGGTCCAHTNATG
CYP76F45-de-R	MAYBCCWYYYTCWAGYTTCCAATYRAA
RACE-PCR	
CsCPR I-3RACE	AACACCAGGAGGACGGCTACACAA
CsCPR I-3RACEnested	GGTTGCAGGAACAGTAAGATGG
CsCPR I-5RACE	CCTGCTCAGCAAGCAACTCATCCAACTC
CsCPRI-5RACEnested	CTGTTGGCTCGCCATCTCCATATGT
CYP76F45-3RACE	AAGATGCTCAAGTGCTGGTCAATGC
CYP76F45-3RACEnested	ATGGGCGATTGGTAGAGACCCAAGT
CYP76F45-5RACE	TTCCTTAGCAGCAACCGCTTGTCA
CYP76F45-5RACEnested	CTCCGCAATTCCTGATTAGAATCAAGC
GeneRacer™3′primer	GCTGTCAACGATACGCTACGTAACG
GeneRacer™3′nested primer	CGCTACGTAACGGCATGACAGTG
GeneRacer™5′primer	CGACTGGAGCACGAGGACACTGA
GeneRacer™5′nested primer	GGACACTGACATGGACTGAAGGAGTA
Full-length PCR	
CsCPR I-ORF-F-*BamH*I	**GGATC**CATGCAATCTTCGACTGGTGCAG
CsCPR I-ORF-R-*Sal*I	**GTCGAC**GTCGACTTAGAAATTGCCTGCTTC
CYP76F45-ORF-F-*Not*I	GCTAACTG**GCGGCCGC**ATGGAGTTCTCTGTGGCTTCTG
CYP76F45-ORF-R-*Xho*I	AGTAGC**CTCGAG**TAATGTTGCAATTTCAAAGGAA
Real-time PCR	
CsCPR I-RT-F	ATGCAACCATCATCATCGGCAGGTTGG
CsCPR I-RT-R	GTGCCGTCCTCTTCATCAGTA
CYP76F45-RT-F	GAAGTAGTGAGATGCGTGTGG
CYP76F45-RT-R	CTTAGCAGCAACCGCTTGTC
18s rRNA-F	CAAAGCAAGCCTACGCTCTG
18s rRNA-R	CGCTCCACCAACTAAGAACG

The primary 3′-RACE-PCR was performed using the GeneRacer™ 3′ primer and either the CYP76F45-3RACE for CYP76F45 or the CsCPR I-3RACE for CsCPR I. The primary 5′-RACE-PCR was performed using the GeneRacer™ 5′ primer and either the CYP76F45-5RACE for CYP76F45 or the CsCPR I-5RACE for CsCPR I. Nested PCR was performed using the primary PCR product diluted 1:100 as the template. The nested PCR for 3′ RACE was performed using the GeneRacer™ 3′ nested primer and either the CYP76F45-3RACEnested or the CsCPR I-3RACEnested. The nested PCR for 5′ RACE was performed using the GeneRacer™5′nested primer and either the CYP76F45-5RACEnested or the CsCPR I-5RACEnested. The RACE-PCR underwent 30 cycles with the following reaction conditions: (1) 30 s at 94 °C, (2) 30 s at 55 °C and (3) 1 min at 68 °C. The purified PCR products were ligated into the pGEM-T easy vector, transformed into *E. coli* TOP10 cells and then confirmed by sequencing. The nucleotide sequences from the 3′- and 5′-RACE along with the CYP76F45 or CsCPR I partial sequences were assembled into full-length cDNA sequences using Clone Manager 9.2 software (Scientific and Educational Software).

5.5. DNA sequencing and sequence analysis

Plasmids for sequencing were prepared using the QIAprep Spin Miniprep Kit (Qiagen, GmbH, Hilden, Germany). All of the DNA fragments were sequenced using a 96-capillary ABI 3730XL DNA Analyzer filled with liquid polymer POP-7™ (Applied Biosystems, Foster City, CA). Sequences and chromatographs underwent quality assurance using Applied Biosystems Sequence Scanner Software v1.0. Homology search analysis and gene alignment were performed using BLAST (http://www.ncbi.nlm.nih.gov/BLAST/). Phylogenetic analyses were constructed using Molecular Evolutionary Genetics Analysis, version 6.0 (MEGA6) (Tamura et al., 2013). A bootstrap resampling analysis with 1000 replicates was performed to evaluate the inferred tree topology (Felsenstein, 1985, 1987). The amino acid sequence motifs were aligned and analyzed using GeneDoc software (Nicholas et al., 1997). The sequences described in this work have been deposited into GenBank under the following accession numbers: (1) KF738255 for CYP76F45 and (2) KF738254 for CsCPR I.

5.6. Construction of the recombinant C. stellatopilosus CYP76F45 and CsCPR I genes

On the basis of the full-length cDNA sequences of CYP76F45 and CsCPR I, gene-specific primers were designed to clone the open reading frames (ORF) of either CYP76F45 or CsCPR I with introduced artificial restriction sites. The CYP76F45-ORF-F-NotI/CYP76F45-ORF-R-XhoI primers were used for cloning CYP76F45 and the CsCPR I-ORF-F-BamHI/CsCPR I-ORF-R-SalI primers were used for cloning CsCPR I (listed in Table 1). PCR reactions were run for 30 cycles using the following conditions: (1) 30 s at 94 °C, (2) 30 s at 55 °C and (3) 2 min at 68 °C. Subsequently, the PCR products were ligated into the NotI/XhoI sites for CYP76F45 or the BamHI/SalI sites for CsCPR I in the pET32a(+) expression vector (Novagen, Madison, WI). The constructed plasmids were transformed into *E. coli* BL21 (DE3) cells (Novagen, Madison, WI). The cells were plated on LB agar containing ampicillin (100 $\mu g\ ml^{-1}$) and the isolated plasmid DNA was purified and examined by DNA sequencing.

5.7. Expression of recombinant proteins and enzyme preparation

For gene expression, *E. coli* BL21 (DE3) cells containing the plasmid encoding the target gene were cultivated at 37 °C in Luria–Bertani (LB) medium containing ampicillin with shaking (250 rpm). The OD_{600} was monitored and the cells were induced with 1 mM

IPTG and 0.005 μM δ-ALA upon reaching an OD_{600} of 0.6, followed by continued shaking at 37 °C (for CYP76F45) or 30 °C (for CsCPR I). Growth was monitored hourly, and equivalent amounts of cells were taken as samples for SDS–PAGE analysis up to 40 h after induction. The cultures were harvested by centrifugation at 8000*g* at 4 °C and washed once with 50 mM Tris–HCl, pH 7.5. The cell pellets were resuspended in 3.0 ml of homogenization buffer consisting of 50 mM Tris–HCl, pH 7.5 and 0.1 M NaCl. The samples were centrifuged at 20,000*g* for 20 min at 4 °C. The pellets (microsomal fraction) were resuspended in o homogenization buffer (1.0 ml). The total protein concentration was determined using Bradford's reagent (Bio-Rad, Hercules, CA) with bovine serum albumin (BSA) as the protein standard.

5.8. Protein qualitative analysis by SDS–PAGE and Western blot

The optimization of the recombinant protein expression was focused on induction time. Upon reaching an OD_{600} equal to 0.6, the protein-expressing cells were induced with 1 mM IPTG and maintained at 37 °C with continued shaking at 250 rpm. Culture samples were collected at 1, 2, 4, 6, 18 and 40 h after induction. The total protein from each time point was qualified by 12% SDS–PAGE (Mini-PROTEAN™ Tetra Cell, Bio-Rad). Western blot analysis began by transferring the proteins separated by SDS–PAGE to a polyvinylidene fluoride (PVDF) membrane (Mini Trans-Blot® Electrophoretic Transfer Cell, Bio-Rad). CYP76F45 and CsCPR I were visualized using alkaline phosphatase (AP)-conjugated antibodies diluted 1:2000 (Invitrogen, Carlsbad, CA) and were developed in an NBT/BICP solution.

5.9. Western blot analysis

To detect the 6×His-tagged recombinant CYP76F45 and CsCPR I proteins, the extracted proteins were resolved by 12% SDS–PAGE using the Mini-PROTEAN® Tetra Cell (Bio-Rad, Hercules, CA) and blotted onto PVDF membranes using the Mini Trans-Blot® Electrophoretic Transfer Cell (Bio-Rad, Hercules, CA). The membrane was blocked with 5% (w/v) skim milk in 0.1% (v/v) TBS-Tween20. The membrane was probed with Anti-His (C-term)/AP Ab (Invitrogen) mouse anti-His monoclonal antibody (diluted to 1:2000 in a dilution buffer containing 1% (w/v) skim milk in 0.1% (v/v) TBS-Tween 20). After being washed, the membrane was incubated with alkaline phosphatase ((AP)-conjugated to the anti-His (C-term) monoclonal antibody). Lastly, immunoreactive bands were visualized using the BCIP/NBT reaction (BCIP/NBT kit, Invitrogen).

5.10. Enzyme assay

The enzyme assay was carried out to examine whether the recombinant CYP76F45 and CsCPR I could be reconstituted to be catalytically active in converting GOH (**1**) to 8-hydroxygeraniol (**2**). Briefly, 100 μg of both CYP76F45 and CsCPR I (insoluble 20,000 g fractions) were added to enzymatic reaction buffer (80 mM Tricine pH 7.8, 0.8 mM NADPH, and 60 mM GOH (**1**)). To study the enzyme substrate specificity, other substrates of the same concentration (60 mM), including farnesol (**3**) (FOH), geranylgeraniol (**5**) (GGOH), and naringenin (**7**), were used for comparison. The reaction was incubated with vigorous shaking at 30 °C for 18 h. Products from the reaction were extracted twice with EtOAc (300 μl) and dried under vacuum for 2 h at 37 °C. The product was then dissolved in EtOAc (5 μl) and spotted onto a TLC silica gel 60 F254 aluminum plate CAMAG Linomat 5 (CAMAG, Muttenz, Switzerland), which was developed in a toluene:MeCN:EtOAc:glacial AcOH (35:5:15:0.15) solvent system. The spots were detected and scanned at λ 210 nm by the CAMAG TLC Scanner 3 under UV light at a wavelength of 254 nm. For GC–MS analysis, the EtOAc fractions containing the reaction products was analyzed directly as described below (Section 5.12) without further preparation of the samples.

To determine of the kinetic parameter of *Km* for GOH (**1**), the same amount of the CYP76F45 and CsCPR I proteins (insoluble 20,000 g fractions) used in the incubation mixture described above were incubated with varying GOH (**1**) concentrations (from 0.01–0.15 μM, each in triplicate) and each reaction was allowed to proceed at 30 °C for 4 h. The product obtained from each reaction was extracted, dried, redissolved, and subjected to the same TLC system for separation and analysis as described above. The enzymatic product (8-hydroxygeraniol, **2**) was quantified by TLC using a standard curve calculated with various known concentrations of authentic geraniol (**1**) (in triplicate). The kinetic *Km* and *Vmax* were then estimated by nonlinear regression with GraphPad Prism5 (GraphPad software).

5.11. Expression patterns of CYP76F45 and CPR I in C. stellatopilosus

The expression of the genes encoding CYP76F45 and CPR I in *C. stellatopilosus* was determined by semi-quantitative real-time PCR and was compared to the accumulation of the putative 19-*E*-vallesamine in *C. stellatopilosus*. The plant parts of the shoot, leaf and twig were ground and extracted for total RNA according to the method described in Section 5.2. Real-time PCR reactions were performed with a LightCycler®480 SYBR Green I Master (Roche) with gene-specific RT primers (Table 1). The reaction consisted of 1X Light Cycler480 SYBR GreenI Master mix (Roche), 20 ng cDNA and 1 mM specific primer for 5′- and 3′-ends of the template (as listed in Table 1). The reaction progress was monitored in a Light Cycler 480 multi-well plate (Roche) with triplicate determination. The relative quantitative real-time PCR condition defined as pre-incubation at 95 °C (1 min) for the first cycle, and then amplification during 44 repeated cycles at 95 °C (10 s), 55 °C (30 s), and 72 °C (2 min), and the reaction was completed with the determination of a melting curve at 65 °C (5 s) and 95 °C (5 s). The relative fold-expression was analyzed by the CFX manager program (Bio-Rad).

5.12. Preparation crude alkaloid extracts

Crude alkaloid extracts were prepared from three different plant parts of *C. stellatopilosus* shoot, leaf and twig. Each freshly collected plant part (60 g) was extracted under conditions of reflux with MeOH (300 ml) for 2 h. After filtration, the solvent was removed under reduced pressure at 60 °C. The crude alkaloid mixture was then obtained by initial extraction with aqueous AcOH followed by CH_2Cl_2 extraction and then basification of the aqueous solution and further CH_2Cl_2 extraction. The alkaloid mixture was checked by TLC with a silica gel 60 F254 plate (Merck) using the solvent system of hexane and EtOAc (1:2). The alkaloids on the plate were detected by spraying with Dragendorff's reagent.

5.13. Identification of enzymatic products and plant alkaloids by GC–MS

GC–MS was performed in electron ionization mode on an Agilent 7890B GC–MS system (Triple Quadrupole GC/MS (GC-QQQ), Agilent Technologies). For GOH (**1**) reaction product identification, the EtOAc extract of each reaction mixture (1 μl) was analyzed on a HP-5 ms (30 m × 0.25 mm, 0.25 μM) column using He as a carrier gas, with a constant flow rate of 1 ml min^{-1}. The injector was operated in split mode (20:1 ratio) with an inlet temperature of 280 °C. The oven program for the HP5 column was 60 °C; ramp of 5 °C min^{-1} to 220 °C, 10 °C min^{-1} to 300 °C (5 min hold) for 45 min. For alkaloid detection, the constant flow rate of the He carrier was at 1.2 ml min^{-1} with a 10:1 inlet split ratio at 280 °C. The oven

condition was 120 °C (4 min hold); ramp of 3 °C min^{-1} to 280 °C (20 min hold) for 77.33 min, 300 °C (2 ml min^{-1}) for 3 min. The reaction product and alkaloid peaks were identified by comparison of retention time (70 eV, mass scan range: m/z 30–500 and 33–650 for GOH reaction and alkaloid, respectively) and spectra with those of standards and the Wiley and National Institute of Standards and Technology Libraries (NIST2011).

Acknowledgments

We would like to thank Professor Dr. David Nelson for help in naming the new CYP76 family gene and Dr. Supakarn Chamni for helpful discussions on the GC–MS results. This work was partially supported by Thailand's National Center for Genetic Engineering and Biotechnology (BIOTEC) and the Office of the Higher Education Commission (OHEC).

References

Barnes, H.J., Arlotto, M.P., Waterman, M.R., 1991. Expression and enzymatic activity of recombinant cytochrome P450 17 alpha-hydroxylase in *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. 88 (13), 5597–5601.

Chen, S. et al., 2003. CYP79F1 and CYP79F2 have distinct functions in the biosynthesis of aliphatic glucosinolates in Arabidopsis. Plant J. 33 (5), 923–937.

Collu, G. et al., 2001. Geraniol 10-hydroxylase, a cytochrome P450 enzyme involved in terpenoid indole alkaloid biosynthesis. FEBS Lett. 508 (2), 215–220.

Diaz-Chavez, M.L. et al., 2013. Biosynthesis of Sandalwood Oil: Santalum album CYP76F cytochromes P450 produce santalols and bergamotol. PLoS ONE 8 (9), e75053.

Duan, H., Schuler, M.A., 2006. Heterologous expression and strategies for encapsulation of membrane-localized plant P450s. Phytochem. Rev. 5, 507–523.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Felsenstein, J., 1987. Estimation of hominoid phylogeny from a DNA hybridization data set. J. Mol. Evol. 26 (1–2), 123–131.

Ginis, O. et al., 2012. Molecular cloning and functional characterization of *Catharanthus roseus* hydroxymethylbutenyl 4-diphosphate synthase gene promoter from the methyl erythritol phosphate pathway. Mol. Biol. Rep. 39 (5), 5433–5447.

Hansen, C.H. et al., 2001. Cytochrome p450 CYP79F1 from arabidopsis catalyzes the conversion of dihomomethionine and trihomomethionine to the corresponding aldoximes in the biosynthesis of aliphatic glucosinolates. J. Biol. Chem. 276 (14), 11078–11085.

Hofer, R. et al., 2013. Geraniol hydroxylase and hydroxygeraniol oxidase activities of the CYP76 family of cytochrome P450 enzymes and potential for engineering the early steps of the (seco)iridoid pathway. Metab. Eng. 20, 221–232.

Jensen, K., Moller, B.L., 2010. Plant NADPH-cytochrome P450 oxidoreductases. Phytochemistry 71 (2–3), 132–141.

Jones, D.T., Taylor, W.R., Thornton, J.M., 1992a. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8 (3), 275–282.

Jones, D.T., Taylor, W.R., Thornton, J.M., 1992b. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8 (275–282), 275.

Meijer, A.H. et al., 1993. Isolation and characterization of a cDNA clone from *Catharanthus roseus* encoding NADPH:cytochrome P-450 reductase, an enzyme essential for reactions catalysed by cytochrome P-450 mono-oxygenases in plants. Plant J. 4 (1), 47–60.

Mishra, R.K., Gangadhar, B.H., Yu, J.W., Kim, D.H., Park, S.W., 2011. Development and characterization of EST based SSR markers in Madagascar periwinkle (*Catharanthus roseus*) and their transferability in other medicinal plants. Plant Omic J. 4 (3), 154–162.

Mizutani, M., Ohta, D., 1998. Two isoforms of NADPH:cytochrome P450 reductase in *Arabidopsis thaliana*. Gene structure, heterologous expression in insect cells, and differential regulation. Plant Physiol. 116 (1), 357–367.

Nelson, D., Werck-Reichhart, D., 2011. A P450-centric view of plant evolution. Plant J. 66 (1), 194–211.

Nelson, D.R., 2009. The cytochrome p450 homepage. Hum. Genomics 4 (1), 59–65.

Nicholas, K., Nicholas, H., Deerfield, D., 1997. GeneDoc: Analysis and Visualization of Genetic Variation 4. EMBNEW.

Nielsen, H.B., Hazell, A., Hazell, R., Ghia, F., Torssell, K.B.G., 1994. Indole alkaloids and terpenoids from *Tabernaemontana markgrafiana*. Phytochemistry 37 (6), 1729–1735.

Rana, S. et al., 2013. NADPH-cytochrome P450 reductase: molecular cloning and functional characterization of two paralogs from *Withania somnifer*a (L.) dunal. PLoS ONE 8 (2), e57068.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4 (4), 406–425.

Sintupachee, S., Ngamrojanavanich, N., Sitthithaworn, W., De-Eknamkul, W., 2014. Molecular cloning, bacterial expression and functional characterisation of cytochrome P450 monooxygenase and NADPH-cytochrome P450 reductase from *Croton stellatopilosus* Ohba. Plant Sci. 229, 131–141.

Sung, P.H., Huang, F.C., Do, Y.Y., Huang, P.L., 2011. Functional expression of geraniol 10-hydroxylase reveals its dual function in the biosynthesis of terpenoid and phenylpropanoid. J. Agric. Food Chem. 59 (9), 4637–4643.

Tamaki, K., Imaishi, H., Ohkawa, H., Oono, K., Sugimoto, M., 2005. Cloning, expression in yeast, and functional characterization of CYP76A4, a novel cytochrome P450 of petunia that catalyzes (omega-1)-hydroxylation of lauric acid. Biosci. Biotechnol. Biochem. 69 (2), 406–409.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. 30 (12), 2725–2729.

Tansakul, P., De-Eknamkul, W., 1998. Geranlygeraniol-18-hydroxylase: the last enzyme on the plaunotol biosynthetic pathway in *Croton sublyratus*. Phytochemistry 47 (7), 1241–1246.

Turiel, N.A. et al., 2013. Essential oils composition of croton species from the Amazon. Nat. Prod. Commun. 8 (10), 1471–1472.

Walser, W., Djerassi, C., 1964. Alkaloid-studien XLIX die strukturen von vallesamin und O-acetyl-vallesamin. Helv. Chim. Acta 47 (7), 2072–2086.

Wang, J. et al., 2010. Cloning and functional analysis of geraniol 10-hydroxylase, a cytochrome P450 from *Swertia mussotii* Franch. Biosci. Biotechnol. Biochem. 74 (8), 1583–1590.

Werck-Reichhart, D., Bak, S., Paquette, S., 2002. Cytochromes p450. Arabidopsis Book 1, e0028.

Zeches, M. et al., 1987. Some new vallesamine-type alkaloids. J. Nat. Prod. 50 (4), 714–720.

Zelasko, S., Palaria, A., Das, A., 2013. Optimizations to achieve high-level expression of cytochrome P450 proteins using *Escherichia coli* expression systems. Protein Expr. Purif. 92 (1), 77–87.