Metabolic Engineering ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Metabolic Engineering

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

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

René Höfer [a], Lemeng Dong [b], François André [c], Jean-François Ginglinger [a], Raphael Lugan [a], Carole Gavira [a], Sebastien Grec [d], Gerhard Lang [e], Johan Memelink [f], Sander Van Der Krol [b], Harro Bouwmeester [b], Danièle Werck-Reichhart [a,*]

[a] Institut de Biologie Moléculaire des Plantes, CNRS UPR 2357-Université de Strasbourg, France
[b] Laboratory of Plant Physiology, Wageningen University, 6700 AR Wageningen, The Netherlands
[c] CEA, iBiTecS, Service de Bioénergétique Biologie Structurale et Mécanismes (SB2SM), Gif-sur-Yvette, France
[d] Université Lille Nord de France, Lille 1, France
[e] Philip Morris Products S.A., Research & Development, Quai Jeanrenaud 5, CH-2000 Neuchâtel, Switzerland
[f] Institute of Biology, Leiden University, Sylvius Laboratory, P.O. Box 9505, 2300 RA Leiden, The Netherlands

ARTICLE INFO

Article history:
Received 28 March 2013
Received in revised form
10 June 2013
Accepted 1 August 2013

Keywords:
Catharanthus roseus
Cytochrome P450
Geraniol oxidation
Multifunctional enzyme
In planta enzyme assay
Pathway reconstitution

ABSTRACT

The geraniol-derived (seco)iridoid skeleton is a precursor for a large group of bioactive compounds with diverse therapeutic applications, including the widely used anticancer molecule vinblastine. Despite of this economic prospect, the pathway leading to iridoid biosynthesis from geraniol is still unclear. The first geraniol hydroxylation step has been reported to be catalyzed by cytochrome P450 enzymes such as CYP76B6 from *Catharanthus roseus* and CYP76C1 from *Arabidopsis thaliana*. In the present study, an extended functional analysis of CYP76 family members was carried-out to identify the most effective enzyme to be used for pathway reconstruction. This disproved CYP76C1 activity and led to the characterization of CYP76C4 from *A. thaliana* as a geraniol 9- or 8-hydroxylase. CYP76B6 emerged as a highly specialized multifunctional enzyme catalyzing two sequential oxidation steps leading to the formation of 8-oxogeraniol from geraniol. This dual function was confirmed *in planta* using a leaf-disc assay. The first step, geraniol hydroxylation, was very efficient and fast enough to outcompete geraniol conjugation in plant tissues. When the enzyme was expressed in leaf tissues, 8-oxogeraniol was converted into further oxidized and/or reduced compounds in the absence of the next enzyme of the iridoid pathway.

1. Introduction

Catharanthus roseus or Madagascar periwinkle (formerly known as *Vinca rosea*) contains a complex mixture of about 130 terpenoid indole alkaloids (TIAs) or Vinca alkaloids with a broad range of biological activities (van der Heijden et al., 2004; O'Connor and Maresh, 2006). Among them, vinblastine and the related compound vincristine, or the synthetic derivatives vindesine and vinorelbine are major drugs used for the treatment of various cancers and as leads for further drug development (van der Heijden et al., 2004; Roepke et al., 2010). Despite their low *in planta* concentration and the presence of a complex mixture of other alkaloids, these compounds and their monomeric precursors are still extracted from the periwinkle plant, with low yield and at high cost. For these reasons *C. roseus* became one of the most studied medicinal plants and a model for biotechnological developments in plant secondary metabolism (see *e.g.* Montiel et al., 2007; Peebles et al., 2009, 2011; O'Connor, 2012; Zhao et al., 2013).

The formation of the complex vinblastine molecule is thought to be catalyzed by around 30 enzymes with intricate subcellular localization and tissue-specific expression patterns (Guirimand et al., 2011; Verma et al., 2012). All TIAs found in plants are derived from a common precursor, strictosidine, which is formed by condensation of tryptamine and the iridoid secologanin. While tryptamine biosynthesis and many of the steps in the formation of the vinblastine precursor vindoline have been described, the pathway that leads to secologanin is still partly unclear. This is surprising since iridoids are quite widespread in dicotyledenous plants and exhibit a range of interesting bioactivities that make them candidates for therapeutic treatments as anti-inflammatory,

* Corresponding author. Fax: +33 354861921.
E-mail address: daniele.werck@ibmp-cnrs.unistra.fr (D. Werck-Reichhart).

1096-7176/$ - see front matter © 2013 Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.ymben.2013.08.001

anticancer, antimicrobial, neuroprotective, immunomodulating, hypoglycemic, hypolipidemic, antispasmodic and anti-oxidative compounds (Dinda et al., 2011; Tundis et al., 2008; Viljoen et al., 2012). Moreover, iridoids could be of great interest as agrochemicals to repel insects (Birkett et al., 2011; Khan et al., 2012) and are one of the main classes of compounds responsible for the bitter taste of plants or plant extracts, as exemplified by olive oil (Mateos et al., 2004). With the objective to improve secologanin and vinblastine production in *C. roseus* or to introduce this pathway in heterologous engineered host plants, we focused on the iridoid segment of the pathway. The first dedicated step in this pathway is the formation of geraniol from geranyl diphosphate (Simkin et al., 2013), followed by 8-hydroxylation (commonly misnamed geraniol 10-hydroxylation). The cytochrome P450 enzyme CYP76B6 isolated from *C. roseus* was previously reported to catalyze geraniol 8-hydroxylation (Collu et al., 2001). The same activity was reported for the closely related CYP76B10 from *Swertia mussotii* (Gentianaceae) that accumulates the iridoid monoterpenoid swertiamarin (Wang et al., 2010). Geraniol 8-hydroxylase activity was also claimed for CYP76C1 of *Arabidopsis thaliana* in a patent by Otah and Mizutani (1998), which suggests that geraniol hydroxylation could be a common property of CYP76s from different plant species. These reports led us to investigate and compare geraniol hydroxylase activities of a set of CYP76s from different plant species in order to identify the most suitable candidate for metabolic engineering and optimal production of 8-hydroxygeraniol *in planta*. In contrast to what was reported in the patent by Otah and Mizutani (1998), we show that CYP76C1 does not have geraniol hydroxylase activity. However, we identify CYP76C4, also from *A. thaliana*, as an active geraniol 8-, 9-hydroxylase. In addition, our data show that *C. roseus* CYP76B6 is a highly efficient and multifunctional geraniol-8-oxidase, that catalyzes the first two committed oxidation steps of the iridoid pathway, the regioselective geraniol 8-hydroxylation and the further oxidation to 8-oxogeraniol, both *in vitro* and *in planta*. In the absence of the next enzyme in the iridoid pathway, engineering of the first two steps of the pathway in *Nicotiana benthamiana* leads to the accumulation of 8-oxogeraniol, and particularly reduced, further oxidized, and conjugated metabolites of 8-oxogeraniol, which suggests that intermediate channeling will be critical to reconstruct the entire TIA pathway in heterologous plant hosts.

2. Material and methods

2.1. Generation of expression vectors

The CYP76C1, CYP76C2, CYP76C5 and CYP76C7 constructs, were generated by PCR amplification of the desired fragment from cDNA prepared from different tissues of *A. thaliana* with primers harboring appropriate enzyme restriction sites. The digested PCR fragments were then integrated into the yeast expression vector pYeDP60 (Table S3). For CYP76C3, CYP76C4, CYP76C6 and CYP76B6 the USER™ cloning technique (New England Biolabs) was used according to Nour-Eldin et al. (2006). Their coding sequences were amplified from cDNA prepared from *A. thaliana* flowers (CYP76Cs) or *C. roseus* (CYP76B6) using PCR primers designed to introduce a USER cloning site (Table S3). The resulting fragments were integrated into pYeDP60u2. The USER compatible vector pYeDP60u2 was prepared as described in Hamann and Lindberg Møller (2007) by integration of complementary oligonucleotides (5′-phos-AATTCGCTGAGGTTTAATTAAGGATCCTTAATTAAACCT-CAGCG-3′ and 5′-phos-GATCCGCTGAGGCTTAATTAAGGATCCT-TAATTAAGCCTCAGCG-3′) into a BamHI/EcoRI digested pYeDP60. The CYP76B1::pYeDP60 construction is described in Didierjean et al. (2002). Plant expression vectors for CYP76C4 and CYP76B6 were produced by integration of the PCR fragments into pCAMBIA330035Su (kindly provided by B. A. Halkier, University of Copenhagen). Correct amplification of all constructs was confirmed by sequencing.

2.2. Heterologous expression in yeast and enzyme activity

The yeast expression constructs were transformed into the *Saccharomyces cerevisiae* strain WAT11 (Pompon et al., 1996) as described by Gietz and Schiestl (2007). One transformed yeast colony was grown and protein expression induced as described by Pompon et al. (1996). Microsomal membranes were prepared as described in Heitz et al. (2012) and their cytochrome P450 content was measured by differential spectrophotometry according to Omura and Sato (1964). Standard enzyme assays were carried out in 100 μL of 20 mM phosphate citrate buffer (pH 7.4) containing varying concentrations of substrate, 600 μM NADPH and adjusted amounts of P450 enzyme. After the addition of NADPH, samples were incubated at 27 °C for 6–20 min. The reaction was stopped by addition of 10 μL of 1 M HCl and 500 μL ethyl acetate. After centrifugation, the ethylacetate phase was transferred to a new vial and extraction was repeated with another 500 μl ethyl acetate. The combined organic phase was dried over anhydrous Na_2SO_4, concentrated under argon and analyzed by GC-FID and GC–MS. Enzyme assays were performed in triplicates.

To generate amounts of products large enough for NMR analysis, the standard enzyme assay was scaled-up to a volume of 10 mL containing 400 μM substrate. After a first incubation for 15 min at 27 °C, a second aliquot of P450 enzyme was added and incubated for another 15 min. The reaction was stopped by adding 1 mL of 1 M HCl, vortexing and cooling on ice. Several scaled-up assays were pooled to ensure proper NMR detection of the products. For the extraction of the products, SPE columns (Oasis® HLB extraction cartridges, Waters) were equilibrated with $CHCl_3$, methanol and water prior to gradual extraction of up to 15 mL of combined samples. After drying, the columns were eluted with $CDCl_3$, and the combined organic phase was dried over anhydrous Na_2SO_4, and concentrated under argon before NMR analysis.

2.3. GC-FID and GC–MS analysis

Capillary GC was performed on a Varian 3900 gas chromatograph (Agilent technologies) equipped with a flame ionization detector and a DB-5 column (30 m, 0.25 mm, 0.25 μm; Agilent technologies) with splitless-injection, at 250 °C injector temperature, and with a temperature program of 0.5 min at 50 °C, 10 °C/min to 320 °C, and 5 min at 320 °C. Terpenoids were identified on the basis of their retention time and using electron-impact MS spectra (70 eV, *m*/*z* 50–600) with a PerkinElmer Clarus 680 gas chromatograph coupled to a PerkinElmer Clarus 600T mass spectrometer (PerkinElmer). Capillary GC–MS was performed as described above, prior to a second GC–MS run optimized for better separation of the products and enhanced spectra quality with splitless-injection at 250 °C injector temperature, and 0.5 min at 50 °C, 10 °C/min to 100 °C, 4 °C/min to 160 °C, 20 °C/min to 320 °C and 4 min at 320 °C in the GC-oven. Reference standards of 8-hydroxygeraniol, 8-hydroxygeranial, 8-oxogeraniol and 8-oxogeranial were synthesized by Chiralix B.V. (Nijmegen, The Netherlands), 8-carboxygeraniol and 8-carboxygeranic acid by Synthelor (Vandoeuvre-les-Nancy, France).

2.4. NMR analysis

NMR was conducted on a 500 MHz Bruker Avance spectrometer equipped with a 5 mm DCH dual cryoprobe or a 5 mm BBI

probe with *z*-gradient operating at 500.13 MHz for ^{1}H and 125.758 MHz for ^{13}C. A number of different spectra including 1D ^{1}H, ^{1}H–^{1}H COSY, edited ^{1}H–^{13}C HSQC and ^{1}H–^{13}C HMBC were recorded for each sample, adding ^{1}H–^{1}H NOESY and 1D ^{13}C when required. Pulse sequences were taken from the Bruker library. All experiments were acquired at 300 K with a minimal relaxation delay of 2 s and a mixing time of 600 or 800 ms for NOESY experiments. Coupling constants were assumed to be around 145 Hz and 8 Hz for $^{1}J(^{13}C–^{1}H)$ and $^{n}J(^{13}C–^{1}H)$ respectively. Acquisition parameters were adjusted when necessary but typically spectral windows were set to 7 KHz for ^{1}H and 27 or 31 KHz for ^{13}C. For 2D spectra, data size was at least 2048 points in the direct dimension and varied between 128 and 1024 points in the indirect dimension according to the required resolution.

2.5. Leaf disc assay

For transient expression of the genes of interest in *N. benthamiana* leaves, the plant expression constructs (pCAMBIA330035Su derived) were introduced into the *Agrobacterium tumefaciens* strain LBA4404(*vir*GN54D) (van der Fits et al., 2000) by electroporation as described by Weigel and Glazebrook (2006). Transformed *A. tumefaciens* strains were grown in 5 mL Luria broth containing the appropriate antibiotics. Following overnight growth at 28 °C, bacteria were pelleted by centrifugation for 20 min at 4000 g, washed twice with water and resuspended in tap-water to $OD_{600}=0.2$. An *A. tumefaciens* strain carrying a binary vector for expression of the viral p19 silencing suppressor protein was prepared in the same way (Voinnet et al., 2003). The bacterial suspensions were mixed in a ratio 1/5 (v/v; p19/gene of interest). Young, fully expanded leaves of 6-week-old *N. benthamiana* plants, cultivated in a growth chamber at 20 °C with a photoperiod of 16/8 h light/dark cycle, were infiltrated with the bacterial mix on the abaxial side using a needleless syringe. After 5 days, 15 mm diameter discs were excised from the infiltrated leaves with a cork borer. The leaf discs (20 per sample) were placed on 10 mL 20 mM phosphate citrate buffer pH 7.4 containing 400 μM geraniol in a Petri dish, exposed to vacuum twice, and placed in a growth chamber for2 h. After removal of the leaf discs, terpenes were extracted from the incubation medium with two times 5 mL ethyl acetate. The combined organic phase was dried over anhydrous Na_2SO_4, concentrated under argon and analyzed by GC-FID and GC–MS. Assays were carried out in triplicates.

2.6. UPLC–MS analysis

The leaf discs were directly ground with mortar and pestle in 5 mL methanol and incubated at 20 °C for 1 h. The extracts were centrifuged 1 min at 2000 rpm and the supernatant was transferred into a new vial. The cleared samples were concentrated under argon and analyzed by UPLC–MS. Analyses were performed using a Waters Quattro Premier XE (Waters) equipped with an electrospray ionization source (ESI) and coupled to an Acquity UPLC system (Waters). Chromatographic separation was achieved using an Acquity UPLC BEH C_{18} column (100 × 2.1 mm, 1.7 μm; Waters) and pre-column. The mobile phase consisted of (A) water and (B) methanol, both containing 0.1% formic acid. The run started by 2 min of 95% A. Then a linear gradient was applied to reach 100% B at 12 min, followed by isocratic run using B during 2 min. Equilibration to initial conditions was achieved in 3 min, giving a total run time of 17 min. The column was operated at 35 °C with a flow-rate of 0.35 mL/min, injecting 3 μL samples. Nitrogen was used as the drying and nebulizing gas. The nebulizer gas flow was set to approximately 50 L/h, and the desolvation gas flow to 900 L/h. The interface temperature was set at 400 °C and the source temperature at 135 °C. The capillary voltage was set to 3.4 kV and the cone voltage to 25 V, the ionization was in positive or negative mode. Low mass and high mass resolution were 15 for both mass analyzers, ion energies 1 and 2 were 0.5 V, entrance and exit potential were 50 V and detector (multiplier) gain was 650 V. Data acquisition and analysis were performed with the MassLynx software (version 4.0). Full-scan and Selected Ion Recording (dwell time 0.1 s) mode were used for qualitative analyses.

2.7. Homology modeling of CYP76C4 and CYP76B6 enzymes and docking

3D models of CYP76B6 and CYP76C4 were obtained using Modeller 9v8. For structural template identification, PSI-BLAST of CYP76B6 and CYP76C4 was performed against the PDB database, and several multi-template rebuilt models were tested. Final models were based on an alignment of the crystal structures of CYP1B1 (□NF-bound, PDB code 3pm0), CYP2C9 (flurbiprofen-bound, PDB 1r9o), CYP2C8 (free, PDB 1pq2), CYP1A2 (□NF-bound, PDB 2hi4), and CYP2A13 (indole-bound, PDB 2p85) as templates for CYP76C4, and an alignment of the crystal structures of CYP1A2 (□NF bound, PDB 2hi4), CYP2C9 (warfarin-bound, PDB 1og2), CYP2C5 (free, PDB 1dt6), CYP2A6 (coumarin bound, PDB 1z10), and CYP2E1 (4-Methylpyrazole-bound, PDB 3e4e) as templates for CYP76B6. Structures with bound substrates in the active site were included in the template selection to allow for flexible active site topology. Multiple alignments were performed using MUSCLE (Multiple Sequence Comparison by Log-Expectation) within Jalview 2.7 or using MUSTANG (Multiple Structural AligNment AlGorithm) for structure-based alignments. For each enzyme, 500 models were generated and evaluated by their DOPE score (Discrete Optimized Protein Energy generated by Modeller, see Eswar et al. (2006)). The best model according to the DOPE score was then submitted to optimization steps consisting in several cycles of minimization and molecular dynamic (MD) simulations (20 ns runs, truncated octahedron periodic boundary conditions, 17,800 water molecules). Counter ions were added to the solvent bulk of the protein water complex in order to maintain neutrality of the complex. The AMBER db99 force field parameters were used for the amino acids and TIP3P model employed for explicit water molecules. The parameters applied for the heme were obtained from Oda et al. (2005). A time step of 2 fs was used in all MD runs, which were carried out with the AMBER 9.0 program suite. Prior to starting MD simulations, three consecutive minimizations of the entire ensembles were performed, setting the maximum number of cycles to 1000, the cut off to 8 Å, and switching from steepest descent to conjugate gradient method after 100 minimization cycles. After heating to 300 K (50 ps), 50 ps equilibration of the water shell at 300 K, 2 ns of equilibration of the entire ensemble at 300 K, two steps of 10 ns of MD simulation were performed in the NTP ensemble for both models, in periodic boundary conditions.

Molecular docking experiments of geraniol in the active site were performed using AutoDock 4.2 in the semi-flexible mode. The docking box, in which grid maps were computed, included the active site with the protoporphyrin group on one edge, and part of the access channels. Ligand molecules were prepared under Maestro 9.2 (Schrödinger, LLC, New York) molecular modeling suite and saved under mol2 format as input files for Autodock. Partial charges were assigned using OPLS 2005 force field, for ligand and receptor structures. Ten different conformers of geraniol were generated using ConfGen module. The docking procedure showed that the ligand and some selected residues of the protein active site are flexible, whose torsion angles were identified to generate their PDBQT files using AutoDockTools (ADT). For each ligand conformer, 100 conformations were generated using the recent Lamarckian genetic algorithm, generating a total of 1000 conformations.

Please cite this article as: Höfer, R., et al., Geraniol hydroxylase and hydroxygeraniol oxidase activities of the CYP76 family of cytochrome P450 enzymes and potential for.... Metab. Eng. (2013), http://dx.doi.org/10.1016/j.ymben.2013.08.001

The poses were assigned a score calculated by Autodock that can be considered as an estimated free energy of ligand binding (indicative of binding affinity), then clustered as a function of the closeness of their positions and conformations (with RMSD set at 2.0 Å), and finally ranked by their binding score (for the best pose in the cluster). The grid built by AutoGrid included 126, 126, and 126 points in *x*, *y*, and *z* directions, with a grid spacing of 0.20 Å to allow a good compromise between resolution of the explored surface and the size of the binding area. The default settings were used for all other parameters. The best clusters, corresponding to the lowest (*i.e.* the most negative) energetic binding scores, were generally well-resolved in the histogram of scores.

2.8. Transient expression of multiple genes in N. benthamiana

Agro-infiltration of whole *N. benthamiana* leaves was performed as described in (van Herpen et al., 2010). A suspension of different combinations of *A. tumefaciens* cultures harboring expression constructs (35S:GES, 35S:CYP76C4, 35S:CYP76B6, or empty vector) were infiltrated into leaves of six-week-old *N. benthamiana* plants by pressing a 1 mL syringe without metal needle against the abaxial side of the leaf and slowly injecting the bacterium suspension into the leaf. After agro-infiltration the plants were grown under greenhouse conditions until further analyses (16/8 h light/dark cycle, 28 °C day, 25 °C night). Leaves were harvested 10 days post-agroinfiltration for analysis. Experiments were performed in triplicate.

2.9. Analysis by LC-PDA-QTOF-MS and LTQ-Orbitrap-MSn

Products in agro-infiltrated *N. benthamiana* leaves were analyzed by liquid chromatography coupled to quadrupole time-of-flight mass spectrometry (LC-QTOF-MS) (De Vos et al., 2007). Frozen leaf samples were ground to powder in liquid nitrogen. An aliquots of 200 mg powdered material was extracted in a 1.5 ml Eppendorf vial with 0.6 ml 99.867% MeOH/0.133% formic acid. After short vortexing and 15 min sonication, the extracts were centrifuged and filtered through 0.45 μm filters (SRP4, Sartorius) and 5 μl of the filtered extract was injected into the LC-QTOF-MS. Measurements were in negative ionization mode and leucine encephalin ($[M-H]^-=554.2620$) was used as a lock mass for online mass calibration.

For further identification, selected compounds were targeted for fragmentation by an Accela HPLC tower connected to a LTQ Orbitrap hybrid mass spectrometer (Thermo Fisher Scientific). The instrument settings of HPLC and LTQ Orbitrap were as described (van der Hooft et al., 2012). The LTQ was programmed to use a window of 10 D to isolate the mass of interest in MS1. The data-dependent fragmentation was set as follows: MS2 fragmentation of most intense ion in MS1; MS3 fragmentation of the 5 most intense fragment ions in MS2; MS4 fragmentation of the 5 most intense fragment ions in each MS3.

2.10. LC-QTOF-MS data processing and multivariate analysis

MassLynx 4.0 (Waters) was used for visualization and manual processing of LC-QTOF-MS data. For comparison between samples, mass data were automatically aligned using MetAlign version 1.0 (www.metalign.nl). Baseline and noise calculations were performed from scan number 70 to 2480. The maximum amplitude was set to 25,000 and peaks lower than three times the local noise were discarded. Multiple mass signals derived from the same compound were grouped using MSClust (biotools.wurnet.nl) with the Multivariate Mass Spectra Reconstruction (MMSR) model (Tikunov et al., 2005). The selected mass intensities were log 2 transformed and normalized using range scaling, in which each value in a row is divided by the intensity range observed for this row throughout all samples analyzed. Each row was then mean centered. Finally, the normalized and log-transformed data matrix was used for Principle Component Analysis (PCA) implemented in GeneMath XT version 2.1 (Applied Mathematics).

2.11. Annotation of metabolites

Significant differences between transgenic and wild type plant tissues in intensity of each aligned mass signal were determined using the Student *t*-test (level of significance set at 0.05). In addition, the masses which showed significant difference were checked manually in MassLynx. Metabolites were identified by calculating the elemental composition, selected for the presence of C, H and O using the MassLynx software. A list of possible molecular formulas was obtained using a Mass tolerance of 5 ppm. Best matches were searched in the Dictionary of Natural Products and SciFinder databases for possible structures.

3. Results

3.1. Candidate screening identifies a novel geraniol hydroxylase

As geraniol 8-hydroxylases reported so far belong to the CYP76 family of cytochrome P450 enzymes (Collu et al., 2001; Wang et al., 2010; Otah and Mizutani, 1998), we assumed that other and possibly more active geraniol 8-hydroxylases could be found among related proteins. Based on available data, we selected candidates from the CYP76B and CYP76C subfamilies. Besides the reference enzyme CYP76B6 from *C. roseus*, we included the closely related CYP76B1 from *Helianthus tuberosus* – so far described as a herbicide detoxifying enzyme (Robineau et al., 1998; Didierjean et al., 2002) – as well as the set of seven members of the CYP76C subfamily in *A. thaliana* in our study. The 10 genes were expressed in yeast under the control of a galactose-inducible promoter in the strain WAT11 also expressing the P450-reductase ATR1 from *A. thaliana*. Enzyme expression and activity were tested with the microsomal fractions. Evaluation of protein expression by differential spectrophotometry (Omura and Sato, 1964), confirmed that CYP76B1, CYP76B6, CYP76C1, CYP76C2, and CYP76C4 were expressed at significant levels in yeast. Geraniol hydroxylase activity in the presence of NADPH was detected as expected for CYP76B6, but not for CYP76C1 (Fig. 1), despite repeated attempts with different microsomal preparations and despite the fact

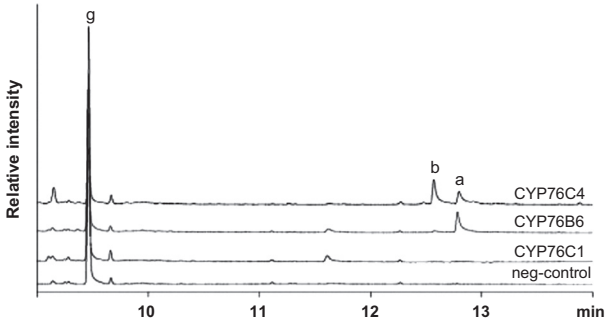

Fig. 1. Geraniol metabolism by CYP76C1, CYP76C4 and CYP76B6. GC-FID chromatograms of the reaction mixtures of geraniol incubated with yeast-expressed *CYP76C1*, *CYP76C4* or *CYP76B6*. Microsomal preparations from recombinant yeast transformed with the *CYP76C1*, *CYP76C4*, *CYP76B6* expression vector or with the empty vector (empty-control) were incubated with 200 μM geraniol for 20 min in the presence of NADPH. No NADPH was added to the negative control (neg-control). Compounds were identified based on their retention time and GC–MS spectra. With g=geraniol; a=8-hydroxygeraniol; b=9-hydroxygeraniol.

Please cite this article as: Höfer, R., et al., Geraniol hydroxylase and hydroxygeraniol oxidase activities of the CYP76 family of cytochrome P450 enzymes and potential for.... Metab. Eng. (2013), http://dx.doi.org/10.1016/j.ymben.2013.08.001

that enzyme activity was detected with other monoterpenols. No activity was detected for the other candidates, except for CYP76C4. GC-FID and GC–MS analysis of the reaction products showed that CYP76B6 catalyzed the formation of one major compound from geraniol with a retention time and fragmentation pattern identical to the reference standard 8-hydroxygeraniol and minor product (*x*), while CYP76C4 generated two products in a ratio 1.3/1, the minor likely 8-hydroxygeraniol, the major with a different retention time but identical mass spectrum (Figs. 1 and S1). For unambiguous identification, we scaled-up the microsome assay to allow NMR analysis of the products. Their structures were determined as 9- (major) and 8- (minor) hydroxygeraniol (Figs. S3, S4, S6–S9).

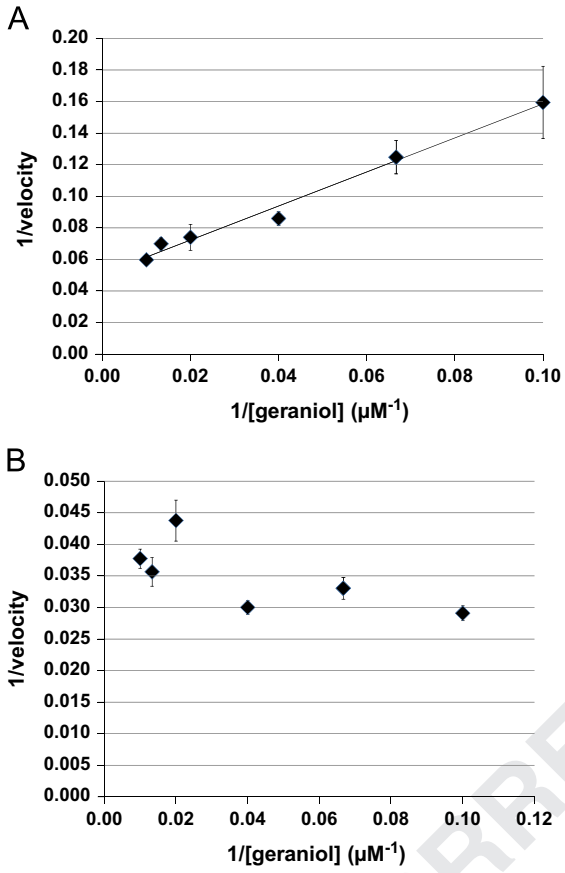

Fig. 2. Lineweaver–Burk plots for conversion of geraniol by CYP76C4 or CYP76B6. Microsomal preparations from recombinant yeast expressing CYP76C4 (A) or CYP76B6 (B) were incubated with different concentrations of geraniol in the presence of NADPH. The amounts of product formed were calculated from the peak areas in GC-FID analysis. In the case of CYP76C4, activity was calculated from the sum of 9- and 8-hydroxygeraniol. Incubations were performed in triplicate, error bars represent the standard deviation.

3.2. *CYP76B6 is a bifunctional enzyme that catalyzes two oxidation steps in the iridoid pathway*

To compare the relative efficiencies of CYP76B6 and CYP76C4 for geraniol hydroxylation, their catalytic parameters were compared. Typical Michaelis–Menten kinetics were observed for CYP76C4 (Fig. 2) allowing calculation of a $K_{m,app}$ of 21 ± 6 μM and a k_{cat} of 20 ± 1 min^{-1}. CYP76B6 however provided aberrant kinetics whatever the concentration and incubation time range used in the experiments (Fig. 2). This suggested that 8-hydroxygeraniol was further converted and therefore competed with geraniol hydroxylation. To test this hypothesis, we first reinvestigated the minor products obtained after longer incubations and higher substrate concentrations. With a refined analytical method, we detected two additional products with CYP76B6 (Fig. 3A), while no additional product was detected with CYP76C4. To test if the additional products resulted from further oxidation of 8-hydroxygeraniol, the latter was assayed as a substrate of yeast-expressed CYP76B6. We observed an efficient conversion of 8-hydroxygeraniol into a compound with a GC retention time and fragmentation pattern similar to reference 8-oxogeraniol (Figs. 3B, S1) and small amounts of the unknown product *x* that was also detected upon incubation of the enzyme in the presence of geraniol. Scaling-up the enzyme assay allowed NMR analysis of the products and confirmed 8-oxogeraniol as the main product (Figs. S3, S5–S9), but the amounts of *x* were not sufficient for structure determination. CYP76B6 did not catalyze any further conversion when 8-oxogeraniol was used as substrate. Hence, *x* most likely derives from 8-hydroxygeraniol. In the presence of 8-hydroxygeraniol, the CYP76B6 reaction showed biphasic kinetics (Fig. S2) with estimated $K_{m,app}$ ranging between 3.5 and 151 μM and and k_{cat} between 15 and 37 min^{-1}. CYP76B6 did not metabolize other potential downstream products such as

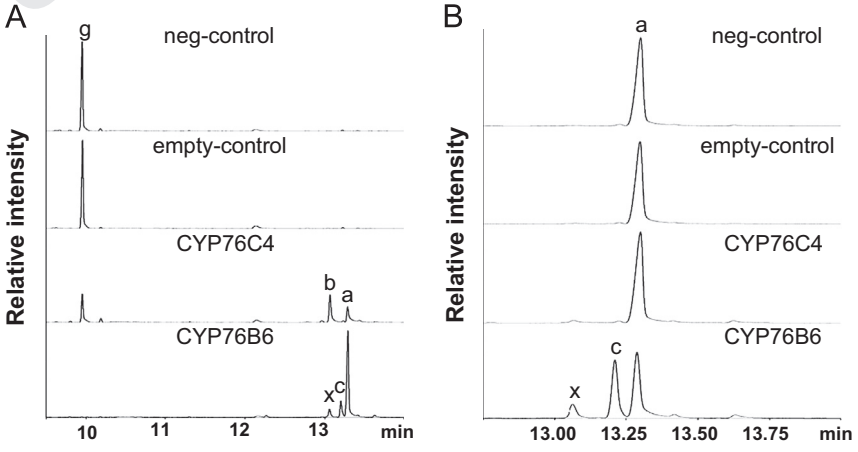

Fig. 3. Metabolism of geraniol and 8-hydroxygeraniol by CYP76B6 and CYP76C4. GC-FID chromatograms of (A) geraniol and (B) 8-hydroxygeraniol conversion by microsomal preparations from recombinant yeast expressing CYP76C4, CYP76B6 or transformed with the empty expression vector (empty-control). Incubations contained 200 μM geraniol or 8-hydroxygeraniol and NADPH. No NADPH was added to the negative control (neg-control). Based on retention time and fragmentation pattern, compound (c) was identified as 8-oxogeraniol, compound (x) could not be identified. With g=geraniol; a=8-hydroxygeraniol; b=9-hydroxygeraniol. c=8-oxogeraniol; *x*=unknown compound.

Please cite this article as: Höfer, R., et al., Geraniol hydroxylase and hydroxygeraniol oxidase activities of the CYP76 family of cytochrome P450 enzymes and potential for.... Metab. Eng. (2013), http://dx.doi.org/10.1016/j.ymben.2013.08.001

8-hydroxygeranial and 8-oxogeranial or geranic acid. No significant inhibition of geraniol hydroxylation was observed in the presence of 8-oxogeraniol (data not shown).

To enable a rough comparisons of CYP76B6 and CYP76C4 turnovers and of the conversion rates of geraniol and 8-hydroxygeraniol by CYP76B6, we performed short enzyme assays at saturating substrate concentration (200 μM), assuming linear rates for the reactions. In these assays, estimated CYP76C4 turnover rate was $16 \pm 1\ min^{-1}$ for production of 8- and 9-hydroxygeraniol, and CYP76B6 catalyzed the formation of 8-hydroxygeraniol with a turnover rate of $170 \pm 6\ min^{-1}$ and the formation of 8-oxogeraniol from 8-hydroxygeraniol with a turnover of $64 \pm 1\ min^{-1}$.

3.3. Binding of geraniol in CYP76C4 and CYP76B6 active sites

While CYP76C4 and CYP76B6 are functionally related P450 isoforms, they share a rather moderate similarity in protein sequence (46%, alignment shown in Fig. S10). It was therefore interesting to locate the structural differences in their respective active sites using homology models of the two enzymes. Homology modeling was based on the crystal structures of the five most similar P450s available in PDB for each of them. Their divergence in sequence led to the selection by PDB-BLAST of slightly different sets of structural templates (see alignments in Figs. S11 and S12). The resulting structures, after refinement by MD simulations, revealed a strikingly conserved active site environment consisting mainly of hydrophobic residues (white-colored in Fig. S13B, C, E, F), including a conserved tryptophan residue in close proximity of the heme (TRP115 in CYP76B6 and TRP123 in CYP76C4). Two aspartate residues (ASP293/300 in CYP76B6 and 307/314 in CYP76C4) are also conserved and positioned similarly above the heme plane.

Despite this similarity of the active site environment, the shape and size of the two active sites, as characterized by VOIDOO, were different with CYP76C4 exhibiting a more narrow active site than CYP76B6 (Fig. S13A, D). As a result, the docking experiments of geraniol into the active sites by Autodock revealed different positioning in the vicinity of heme (Fig. 4). In CYP76C4, geraniol is positioned above the heme plane, almost perpendicular, presenting either the C8 or the C9 carbon to the heme iron, while in CYP76B6, the geraniol molecule is positioned closer to protoporphyrin with the double bond C6–C7 parallel to the heme plane.

In CYP76C4, as shown in Fig. 4A, the most favored cluster (clusters were ranked as a function of their energy) of poses of geraniol in the active site (two poses were considered the same when the calculated RMSD was less than 1.5 Å) contained positionings with equivalent distances from C8 and C9 to the heme iron, alternating with distances C(8/9)–Fe typically found at around 3.5 Å. The hydroxyl terminal group of geraniol was not found stabilized by H-bonding, although Asp307 could act as a potential H-bonding partner. Trp123 lines the cavity, forming a hydrophobic contact along the rest of the geraniol molecule.

Conversely, in the best ranked cluster of poses in CYP76B6, geraniol preferentially exposed the double bond C6–C7 to the heme iron at an average distance of 4.5 Å, which suggests a favored attack on C6, while C8 and C9 were at about 5.2 Å (Fig. 4B). As in CYP76C4, Trp115 is in hydrophobic contact with geraniol, but, in CYP76B6, geraniol positioning is stabilized by H-bonding with Asp300 (the equivalent of Asp314 in CYP76C4). The regiospecific 8-hydroxylation observed with CYP76B6 thus suggests that either a more appropriate conformation is achieved upon oxygen binding, or that regioselectivity results from easier oxidation of the allylic C8. The docking with the C8 position not directly pointing to the heme allows further oxidation of 8-hydroxygeraniol by CYP76B6, which is not observed in the case of CYP76C4.

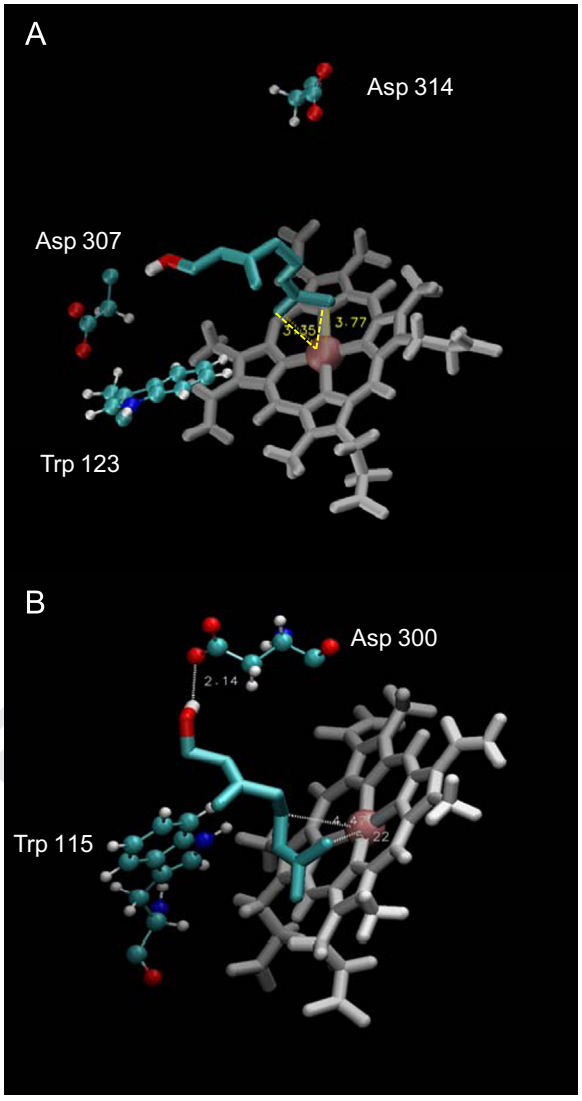

Fig. 4. Docking of geraniol in the active site of CYP76C4 and CYP76B6. The heme catalytic center is represented with iron in orange. The orientation of geraniol relative to the heme of CYP76C4 and CYP76B6 is shown. A and B show the two possible positions of geraniol in the active site of CYP76C4. C illustrates the positioning of geraniol in the active site of CYP76B6. Distances of the substrate C6 and C8 carbon atoms to the heme iron and of the C1–OH to the side chain of Asp300 are indicated in Angströms. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.4. Characterization of products formed in planta

3.4.1. Leaf disc substrate feeding assay

As our ultimate goal is reconstruction of the TIA pathway in a plant for production of anticancer drugs, the activity of CYP76B6 and CYP76C4 was also tested *in planta*. To this end, we developed a fast leaf-disc assay for evaluation of individual substrate conversion and fate in plant tissues. Each P450 was transiently expressed by agroinfiltration in leaves of *N. benthamiana*. Leaf discs were excised 5 days post infiltration, vacuum-infiltrated with and incubated in buffer containing substrate. Products were analyzed in the ethyl acetate extract of the incubation medium, and in the leaf-disc methanol extract.

GC analysis of the incubation buffer extract confirmed the formation of 9- and 8-hydroxygeraniol by CYP76C4 in the plant in a 2/1 ratio (Figs. 5, S14). In the case of CYP76B6-transformed leaves, 8-hydroxygeraniol and 8-oxogeraniol were detected in the incubation medium together with the unknown compound *x*

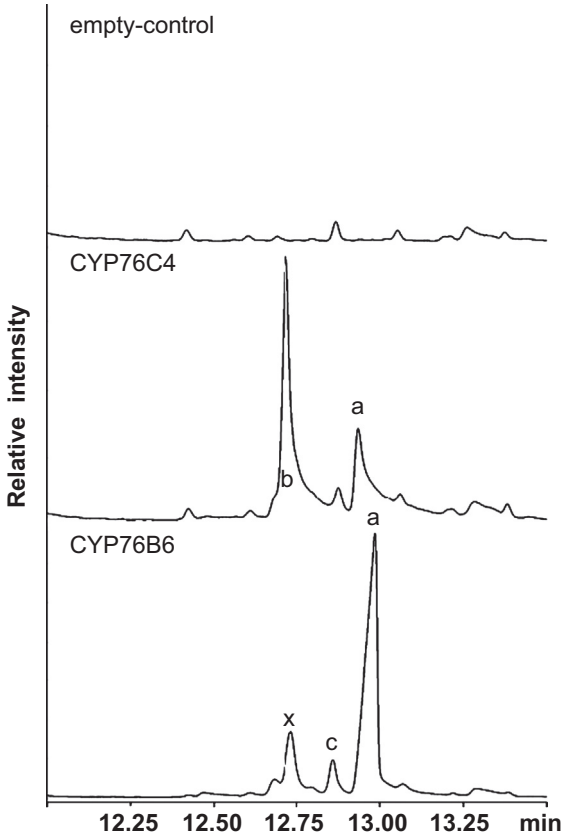

Fig. 5. Products of geraniol conversion in extracts of agro-infiltrated *N. benthamiana* leaf-disc incubation medium. *N. benthamiana* leaf discs transiently expressing *CYP76C4*, *CYP76B6* or the empty pCAMBIA330035Su plant expression vector (neg-control) were incubated in phosphate buffer containing 400 μM geraniol. The ethyl acetate extract of the incubation medium was analyzed by GC-FID and the products were identified by GC–MS: a=8-hydroxygeraniol; b=9-hydroxygeraniol; c=8-oxogeraniol; x=unknown product.

(Figs. 5, S14). The 9/8-hydroxygeraniol and the 8-oxogeraniol/*x* ratios observed *in vivo* were different from those recorded in yeast microsome incubations. This can be the result of a differential membrane permeability or transport of the products, or might reflect differential conversion/conjugation and storage in the plant tissues.

To clarify this point, leaf disc methanol extracts were analyzed by UPLC–MS/MS, providing a more complex picture. In a first approach, targeted analysis by single-ion recording (SIR) was performed. Target ions were selected after analysis of commercial or synthetic standards of geraniol, 8-hydroxygeraniol and 8-oxogeraniol. The most prominent peak in all samples was geraniol (main fragment with *m*/*z* 137), with the highest abundance in leaf discs infiltrated with the empty vector, and lower levels in CYP76C4- and CYP76B6-infiltrated discs (Fig. 6A). Also in the control, infiltrated with empty vector, geraniol was metabolized to some extent, likely by *N. benthamiana* enzymes, to form several compounds with a fragment of *m*/*z* 137 and RT between 9.5 and 10.3 min (y in Fig. 6A). In the CYP76C4-transformed discs, free 9- and 8-hydroxygeraniol (main fragment of *m*/*z* 135) could be clearly identified, while in the CYP76B6 expressing leaf discs only 8-hydroxygeraniol was detected (b and c in Fig. 6A). In addition, we detected 8-hydroxygeranyl hexose in similar amounts in CYP76C4 and CYP76B6. This may indicate that the formation of this compound was limited by the activity of the glycosyl transferase catalyzing this conversion and not by the amount hydroxygeraniol produced. In the CYP76B6-expressing discs, only small amounts of 8-oxogeraniol were detected (d in Fig. 6A), and two peaks e and f not present in the control (Fig. 6A).

To identify these peaks and other related products, we performed untargeted UPLC–MS analysis (Fig. S15). Based on available standards and the untargeted UPLC–MS analysis, we were able to show the presence of free 8-carboxygeraniol (or foliamenthoic acid) (h in Fig. 6B) and 8-carboxygeranic acid (j in Fig. 6B) in the extract of the CYP76B6-transformed leaves. Taking these compounds as starting points to look for related products, we were able to propose the tentative structures for most of the major compounds detected in disc extracts (Fig S16). They include the reduced forms 8-hydroxy-2,6-dimethyloct-2-enoic acid (e in Fig. 6B), 8-hydroxy-2,6-dimethyloctanoic acid (g in Fig. 6B), a more oxygenated compound with tentative structure proposed in f in Fig. 6B and their glycosylated derivatives. The 8-hydroxy-2,6-dimethyloct-2-enoic acid (e in Fig. 6) corresponds to one of the peaks visible in the SIR chromatogram shown in Fig. 6A. Complex metabolism of 8-oxogeraniol including oxidation of the alcohol and aldehyde at the 1 and/or 8 positions, double bond reduction and glycosylation thus occurs in *N. benthamiana* leaves (Fig. S18). Taken together, these results confirm the *in planta* geraniol 8-oxidase activities of CYP76B6 and geraniol 9- and 8-hydroxylase activity of CYP76C4. The enzyme products do not accumulate in the plant but are further converted by endogenous enzymes of *N. benthamiana*.

3.4.2. Pathway reconstitution in N. benthamiana

In addition to leaf disc assays with substrate feeding, the first two steps of the iridoid pathway were also reconstituted *in planta* for determination of product formation during prolonged *in planta* gene expression. *CYP76C4* and *CYP76B6* were transiently co-expressed with a geraniol synthase (*VoGES*) from *Valeriana officinalis* (Dong et al., submitted for publication) via agro-infiltration in 6-week-old *N. benthamiana* leaves. To be able to discriminate products generated from endogenous substrates, infiltrated plants with single genes (*VoGES*, *CYP76C4*, *CYP76B6* and empty vector (*EV*)) and non-infiltrated plants (wild type) were used as control. Previous experiments had shown that 10 days after agro-infiltration of *N. benthamiana* leaves with *VoGES*, the levels of free geraniol and geraniol derivatives were below the level of detection by GC–MS. Therefore, the semi-polar, non-volatile metabolites were extracted with methanol and analyzed by LC-QTOF-MS. Fig. S17A shows a representative example of the chromatograms obtained for leaves infiltrated with *VoGES+CYP76B6*, *VoGES +CYP76C4*, *VoGES* and the *EV* control. For comparison of the different samples, the LC–MS chromatograms were processed and the resulting mass lists used for PCA. This shows clustering of the replicate groups and separation between the extracts from *GES*, *VoGES+CYP76C4* and *VoGES+CYP76B6* agro-infiltrated plants in the first component (59.3% of the variation; Fig. S17B). The samples of *N. benthamiana* leaves agro-infiltrated with only *CYP76C4*, *CYP76B6* or *EV* also showed separation in the PCA plot, indicating that CYP76C4 and CYP76B6 catalyzed the oxidation of endogenous *N. benthamiana* compounds (Fig. S17B). Comparison of the chromatograms identified in total 19 geraniol related masses (compounds) with significant differential peak intensities in the extracts of *VoGES* infiltrated leaves when compared with the controls of *CYP76C4*, *CYP76B6* or *EV* infiltrated leaves (Table S2). For identification, these masses were targeted for fragmentation by LTQ Orbitrap MSn. The resulting MSn spectra show that these compounds most likely are conjugates of geraniol, hydroxygeraniol, foliamenthoic acid or carboxygeranic acid with hexose, pentose, malonyl or acetyl groups (Table S2).

In *N. benthamiana* agro-infiltration with *VoGES*, eight geraniol glycosides and three hydroxygeraniol glycosides were found.

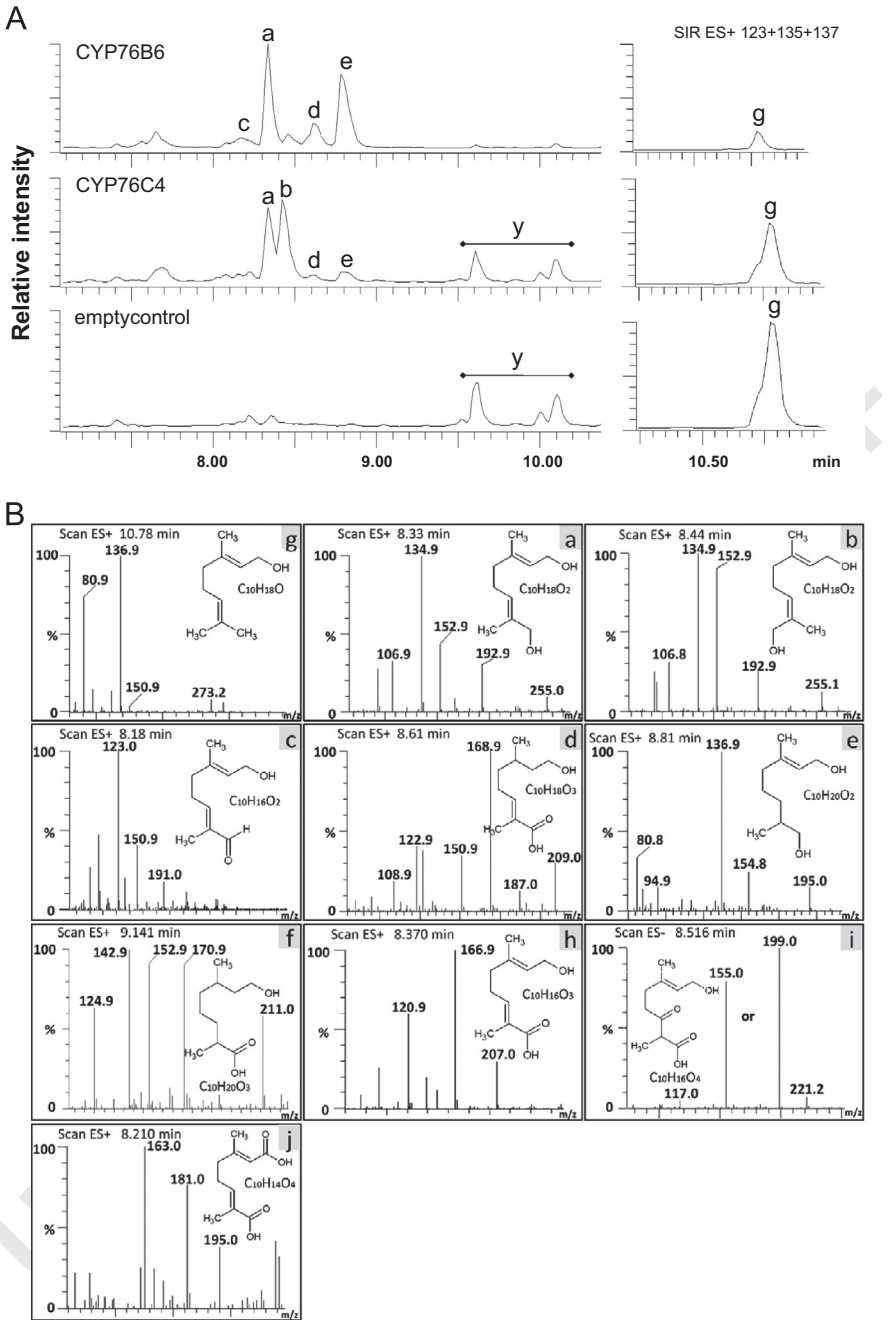

Fig. 6. Products of geraniol conversion detected in the methanol extract of agro-infiltrated *N. benthamina* leaf-discs. *N. benthamiana* leaf discs transiently expressing *CYP76C4*, *CYP76B6* or the empty pCAMBIA330035Su plant expression vector (neg-control) incubated with geraniol, were extracted with methanol and the extracts were analyzed by UPLC–MS. (A) SIR chromatograms showing variation in geraniol (a), geraniol glycosides (y), 8-hydroxygeraniol (b) and 9-hydroxygeraniol (c) contents in the leaf tissues. (B) Representative MS fragmentation spectra of the differentially accumulated compounds identified. The spectra of a=geraniol, b=8-hydroxygeraniol, c=9-hydroxygeraniol, d=8-oxogeraniol, h=foliamenthoic acid and j=8-carboxygeraniol are similar to those of authentic standards. The molecular formulas determined for e=$C_{10}H_{18}O_3$, f=$C_{10}H_{20}O_2$, g=$C_{10}H_{20}O_3$ and i=$C_{10}H_{16}O_4$, were tentatively allocated the following structures according to their UPLC–MS fragmentation patterns: e=8-hydroxy-2,6-dimethyloct-2-enoic acid, f=3,7-dimethyloct-2-ene-1,8-diol, g=8-hydroxy-2,6-dimethyloctanoic acid and i=2,6-dimethyl-3-oxooct-6-enoic acid. The main fragments for each molecule are detailed in Table S1.

When *CYP76C4* was co-infiltrated with *GES*, the level of geraniol glycosides decreased by half compared with those in leaves only infiltrated with *VoGES* (Fig. 7). In addition, the levels of the three hydroxygeraniol glycosides were increased, while also new conjugates of hydroxygeraniol glycosides with an acetyl group were formed.

The geraniol glycosides had completely disappeared in leaves infiltrated with *VoGES+CYP76B6*, suggesting that the geraniol is converted more efficiently by CYP76B6 than by CYP76C4. For both *VoGES+CYP76C4* and *VoGES+CYP76B6* transformed tissues, the glycosides of further oxidized products foliamenthoic acid and carboxygeranic acid were formed, while these were not present in leaves infiltrated with *VoGES* alone. Again, the levels of these products were higher in the plants co-infiltrated with *VoGES +CYP76B6* than in the plants expressing *VoGES+CYP76C4* (Fig. 7), suggesting that geraniol is oxidized more efficiently by CYP76B6 is more efficiently than by CYP76C4.

Please cite this article as: Höfer, R., et al., Geraniol hydroxylase and hydroxygeraniol oxidase activities of the CYP76 family of cytochrome P450 enzymes and potential for.... Metab. Eng. (2013), http://dx.doi.org/10.1016/j.ymben.2013.08.001

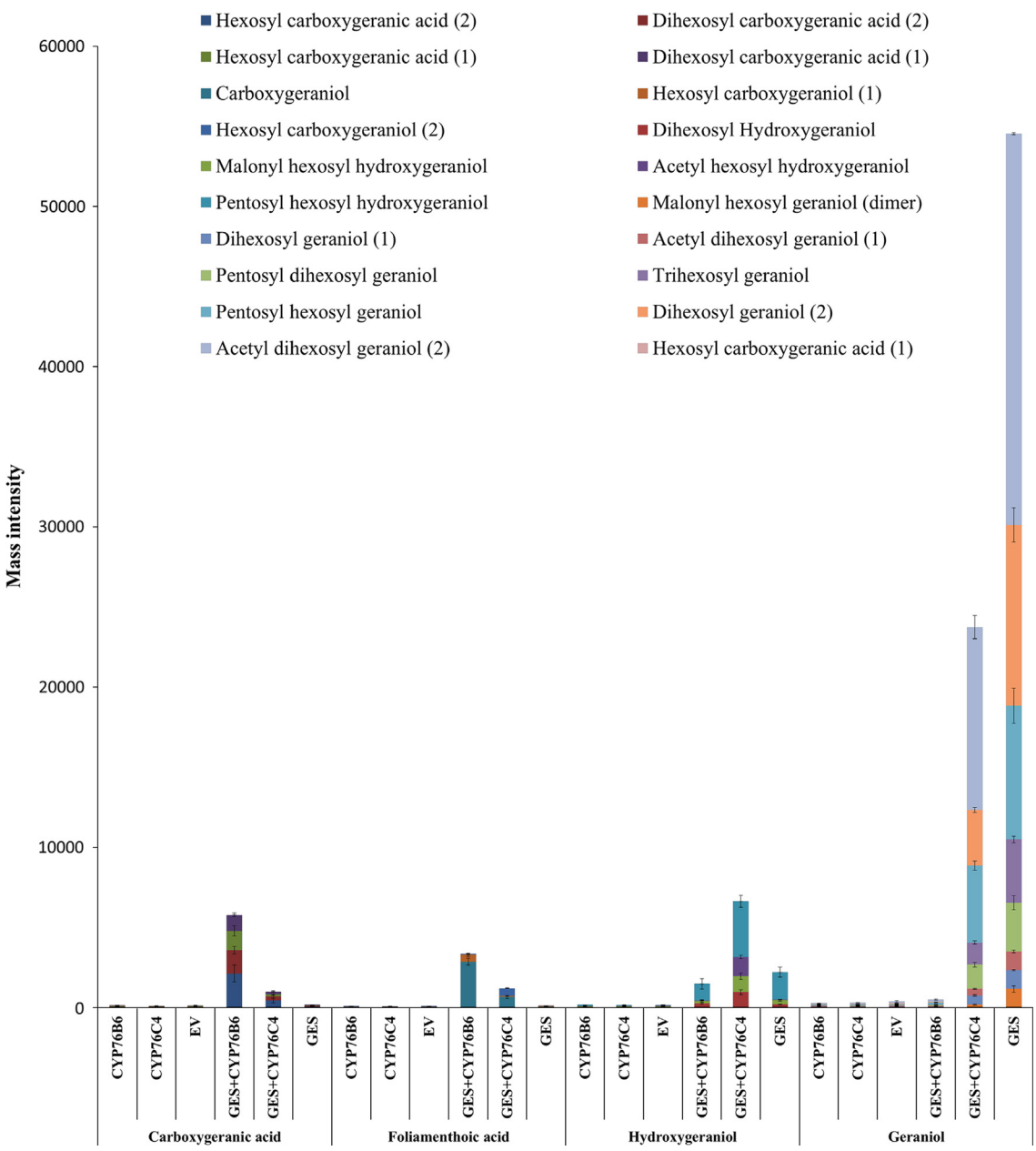

Fig. 7. Summary of reactions catalyzed by CYP76B6 and CYP76C4. In a first step CYP76B6 acts as G8H to convert geraniol into 8-hydroxygeraniol, in a second step CYP76B6 is a 8HGO to produce 8-oxogeraniol from 8-hydroxygeraniol. The G8H and G9H activities of CYP76C4 result in the production of 8- and 9-hydroxygeraniol from geraniol.

4. Discussion

4.1. *CYP76B6 is a regiospecific geraniol oxidase*

Previous reports of geraniol hydroxylase activity of cytochrome P450 enzymes belonging to different CYP76 subfamilies (CYP76B6/B10 and CYP76C1) suggested that this function might be shared by a fairly large set of CYP76 enzymes. Our data show that this is not the case. We detected geraniol hydroxylase activity for the CYP76B6 and CYP76C4 enzymes, but not for the other members of the CYP76B and CYP76C subfamilies that we investigated. This rather suggests that the geraniol hydroxylase function evolved in different phyla as the result of a convergent evolution, favored by a common functional background inherited from a common ancestor. This is supported by the different regiospecificities and catalytic properties of the two enzymes (Fig. 8), and by the different volumes and properties of their active sites (Fig. S13). CYP76C4 from *A. thaliana* is a geraniol hydroxylase oxidizing the 9 as well as the 8 position. Such relaxed regiospecificity, together with moderate enzyme turnover rates, suggest that geraniol is likely not the physiological substrate of CYP76C4 in arabidopsis. So far, no geraniol synthase has been reported in arabidopsis. In addition, geraniol was not detected in the head-space collected from arabidopsis plants and isolated organs (Rohloff and Bones, 2005), so it is possible that geraniol does not occur in this plant species. Hydroxygeraniol is not further converted by CYP76C4. Homology modeling, based on the resolved structures of mammalian P450 enzymes, indicates that geraniol can bind the CYP76C4 active site in varios similar positions that are found in the same dominant cluster of poses, placing either C8 or C9 at distances suitable for oxidative attack, but with no obvious anchoring of the alcohol function (Fig. 4A). In all cases, the C8 or C9 of geraniol

Please cite this article as: Höfer, R., et al., Geraniol hydroxylase and hydroxygeraniol oxidase activities of the CYP76 family of cytochrome P450 enzymes and potential for.... Metab. Eng. (2013), http://dx.doi.org/10.1016/j.ymben.2013.08.001

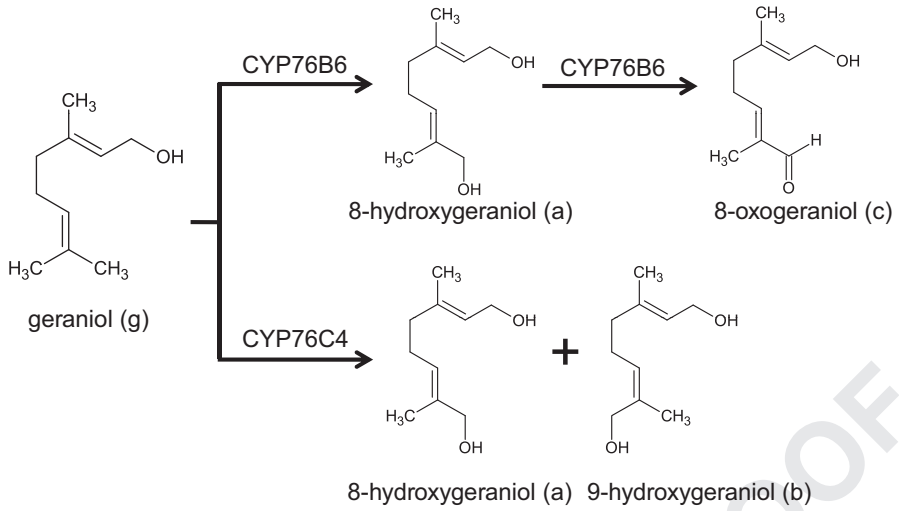

Fig. 8. LC-QTOF-MS analysis of oxidized derivatives of geraniol and their conjugates in *GES+CYP76B6*, *GES+CYP76C4* and *GES* agro-infiltrated *N. benthamiana* leaves.

points straight to the heme. In this orientation, likely due to steric hindrance constraints, hydroxygeraniol binding for further oxidation would not be favored. Alternatively, coordination of the alcohol oxygen with the heme iron would be expected, resulting in enzyme activity inhibition (Ortiz de Montellano and Correia, 1995). Modeling thus provides a reasonable explanation for the observed CYP76C4 activity and the duality of its regiospecificity. Conversely, CYP76B6 is a highly regiospecific and effective geraniol 8-hydroxylase with a catalytic turnover rate of 150 min^{-1} in the range of CYPs with selection-optimized physiological functions. In addition to this high regioselective geraniol 8-hydroxylase activity, CYP76B6 efficiently converts 8-hydroxygeraniol into 8-oxogeraniol. Homology modeling of CYP76B6 and docking of geraniol in its active site suggests that, in this active site, the substrate is stabilized in the active site by the alcohol H-bonding with Asp300 (Fig. 4B). The oxidation at C8 appears favored by both the allylic position and its proximity to the heme. As geraniol lies almost parallel to the heme, no steric hindrance is expected for 8-hydroxygeraniol binding and further oxidation. The 8-hydroxygeraniol oxidase activity, which was not previously detected, corresponds to the next step in the proposed pathway to secologanin. This activity was so far attributed to a soluble NADPH-dependent oxidoreductase, unspecifically oxidizing allylic primary alcohols at the 1 and 8 positions (Ikeda et al., 1991; Hallahan et al., 1995). This would suggest that either the two enzymes cooperate to the formation of iridodial or that the described oxidoreductase, which so far has not been characterized at the molecular level, has another function in the plant.

We were not able to accurately determine the catalytic parameters of the geraniol hydroxylation step by CYP76B6 due to the further conversion of the product and resulting competition. It was not possible either to reliably evaluate the affinity of the CYP76B6 enzyme for geraniol through spectral changes induced by substrate binding, due to the rather low expression of CYP76B6 in yeast. A K_m of 16 μM and a k_{cat} of 8 min^{-1} were previously reported for the enzyme expressed in insect cells also expressing the *A. thaliana* P450 reductase ATR1 (Sung et al., 2011). The k_{cat} measured with insect cell membranes is about 20 times lower than the approximate turnover rate that we measured when the enzyme was expressed in yeast, which points to an importance of the membrane context (yeast or insect) for achieving optimal activity. Taking as a reference the K_m for geraniol measured in insect cells, a strong competition between geraniol and 8-hydroxygeraniol would be expected. The very low geraniol hydroxylation obtained with insect cell membranes, most likely explains why no further conversion of the product was observed. The high turnover rate measured for geraniol hydroxylation with yeast microsomes would however favor the hydroxylation reaction over further oxidation of 8-hydroxygeraniol, explaining the accumulation of 8-hydroxygeraniol observed in the *in vitro* enzyme assays.

4.2. CYP76B6 metabolites are further processed in N. benthamiana

Analysis of the free forms of geraniol and oxidized geraniol products formed in the *CYP76B6*-transformed leaf-disc feeding assay and the transient expression of *VoGES* with *CYP76B6* in intact leaves of *N. benthamiana* suggest a preferential formation of products as shown in Fig. S18. A number of different products were detected in the leaf-disc assay (2 h geraniol conversion) and the intact-leaf infiltration (10 days expression of *VoGES* and *CYP76B6*). The leaf-disc results show that early 8-oxogeraniol conversion mainly consists of double bond reduction (formation of 8-hydroxy-2,6-dimethyloct-2-enoic acid, 8-hydroxy-2,6-dimethyloctanoic acid and 2,6-dimethyl-3-oxooct-6-enoic acid and the corresponding glycosylated forms; Fig. S18). These compounds are not detected after 10 days of *VoGES* and *CYP76B6* expression *in planta* (Fig. 7). Instead, the main long-term accumulation products are 8-carboxygeranic acid and several different conjugates. It is thus likely that the 8-hydroxy-2,6-dimethyloct-2-enoic acid, 8-hydroxy-2,6-dimethyloctanoic acid and 2,6-dimethyl-3-oxooct-6-enoic acid derivatives are further processed in the plant. The leaf disc assay thus provides a picture of the early modifications by endogenous enzymes of geraniol, whereas pathway reconstitution in the plant rather identifies final steady-state products.

4.3. Towards TIA production in heterologous hosts

In a first step of the TIA pathway reconstruction, the *N. benthamiana* leaves were agro-infiltrated with *VoGES* alone, resulting in production of geraniol, which was mostly accumulated as geraniol glycosides after conversion by endogenous glycosyltransferases (Dong et al., submitted for publication). The second step of the TIA pathway is the conversion of geraniol to 8-hydroxygeraniol by a geraniol 8-hydroxylase (G8H). CYP76B6 is here confirmed to have a high G8H activity and the different experiments *in planta* show that it can effectively compete with the glycosylation – by endogenous *N. benthamiana* glycosyl transferases – of the first product of the pathway, geraniol. Even when the geraniol concentration was very high, *e.g.* upon feeding of geraniol to leaf discs expressing *CYP76B6*, only hydroxygeraniol glycosides and no geraniol glycosides were found (Fig. S16). In contrast, geraniol glycosides were still present upon expression of *CYP76C4*, which

indicates, in good agreement with the *in vitro* data, that the enzyme turnover is not high enough to outcompete the *N. benthamiana* glycosyl transferases that use geraniol as substrate.

From the *in planta* expression results we can in principle not conclude whether CYP76B6 acts only on free geraniol or also on the geraniol glycosides. However, it's very high turnover rates measured with free geraniol *in vitro* strongly supports the former. In addition, the difference in glycosylation profile of geraniol glycosides and oxidated geraniol glycosides seems to rule out a sequential product relationship, but rather suggests independent pathways (Fig. 7). Moreover, in the short term leaf-disc assay, oxidized geraniol is found with only one hexose (Figs. S15, S16), suggesting that CYP76B6 acts only on free geraniol. While this is consistent with the evolution of CYP76B6 of *C. roseus* as a very efficient geraniol hydroxylase, the subcellular localization and proximity to GES in the plant cell may also contribute to preferential geraniol oxidation over conjugation. For instance, VoGES was shown to be located on plastidial stromules (Dong et al., submitted for publication), and stromules have been suggested to interact with the ER (Schattat et al., 2011).

The next step towards TIA production in *N. benthamiana* would be the conversion of 8-oxogeraniol to 8-oxogeranial. In *N. benthamiana* this enzymatic step needs to compete with step 3 and step 6 in the pathway (Fig. S18). At present we cannot exclude that some 8-oxogeraniol is already converted to 8-oxogeranial in our *in vivo* experiments as this double aldehyde may be very reactive and therefore not detectable as intermediate. This compound (lacking free OH groups) can also not be sequestered and stabilized as a glycoside. The intermediate 8-carboxygeranial may be formed via foliamenthoic acid or via 8-oxogeranial, but was not detected, presumably because of rapid conversion to 8-carboxygeranic acid. In the same way, 8-oxogeraniol may be readily converted into foliamenthoic acid in the plant in the absence of a competing enzyme. Further introduction of the iridoid synthase which converts 8-oxogeranial to iridodial (Geu-Flores et al., 2012) will be the next step and will tell if an intermediate enzyme is still required.

Q2 Uncited references

Collu et al. (1999), Lommen (2009), Murata et al. (2008).

Acknowledgments

The research described here was funded by the European Union Seventh Framework Programme FP7/2007–2013 under Grant agreement no. 222716 – SMARTCELL. J-FG PhD was supported by a BDI from CNRS and the Région Alsace. CG PhD was supported by CIFRE a grant from the Agence Nationale de la Recherche Technique and MANE fils. We are grateful to L. Miesch and M. Miesch from the Laboratoire de Chimie Organique Synthétique, UMR 7177, CNR-University of Strasbourg for very helpful discussion, and to Lionel DUCASSOU, former PhD student at University René Descartes, who helped in the design of Modeller and Autodock scripts for CYP76 modeling studies.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.ymben.2013.08.001.

References

Birkett, M.A., Hassanali, A., Hoglund, S., Pettersson, J., Pickett, J.A., 2011. Repellent activity of catmint, *Nepeta cataria*, and iridoid nepetalactone isomers against Afro-tropical mosquitoes, ixodid ticks and red poultry mites. Phytochemistry 72, 109–114.

Collu, G., Bink, H.H.J., Moreno, P.R.H., van der Heijden, R., Verpoorte, R., 1999. Determination of the activity of the cytochrome P450 enzyme geraniol 10-hydroxylase in plants by high-performance liquid chromatography. Phytochem. Anal. 10, 314–318.

Collu, G., Unver, N., Peltenburg-Looman, A.M., van der Heijden, R., Verpoorte, R., Memelink, J., 2001. Geraniol 10-hydroxylase, a cytochrome P450 enzyme involved in terpenoid indole alkaloid biosynthesis. FEBS Lett. 508, 215–220.

De Vos, R.C.H., Moco, S., Lommen, A., Keurentjes, J.J.B., Bino, R.J., Hall, R.D., 2007. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. Nat. Protoc. 2, 778–791.

Didierjean, L., Gondet, L., Perkins, R., Lau, S.M.C., Schaller, H., O'Keefe, D., Werck-Reichhart, D., 2002. Engineering herbicide metabolism in tobacco and Arabidopsis with CYP76B1, a cytochrome P450 enzyme from Jerusalem artichoke. Plant Physiol. 130, 179–189.

Dinda, B., Debnath, S., Banik, R., 2011. Naturally occurring iridoids and secoiridoids. An updated review, part 4. Chem. Pharm. Bull. (Tokyo) 59, 803–833.

Dong, L., Goedbloed, M., Verstappen, F.W.A., Van der Krol, S., Jongsma, M.A., Bouwmeester, H.J., 2013. Different subcellular location of geraniol synthases with little effect on product level. Metab. Eng.(submitted for publication). Q3

Eswar, N., Marti-Renom, M.A., Webb, B., Madhusudhan, M.S., Eramian, D., Shen, M., Pieper, U., Sali, A., 2006. Comparative Protein Structure Modeling with MODELLER. Current Protocols in Bioinformatics. John Wiley & Sons, Inc.. (Supplement 15, 5.6.1–5.6.30). Q9

Gietz, R.D., Schiestl, R.H., 2007. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat. Protoc. 2, 31–34.

Guirimand, G., Guihur, A., Poutrain, P., Héricourt, F., Mahroug, S., St-Pierre, B., Burlat, V., Courdavault, V., 2011. Spatial organization of the vindoline biosynthetic pathway in *Catharanthus roseus*. J. Plant Physiol. 168, 549–557.

Geu-Flores, F., Sherden, N.H., Courdavault, V., Burlat, V., Glenn, W.S., Wu, C., Nims, E., Cui, Y., O'Connor, S.E., 2012. An alternative route to cyclic terpenes by reductive cyclization in iridoid biosynthesis. Nature 492, 138–142.

Hallahan, D.L., West, J.M., Wallsgrove, R.M., Smiley, D.W., Dawson, G.W., Pickett, J.A., Hamilton, J.G., 1995. Purification and characterization of an acyclic monoterpene primary alcohol:NADP+ oxidoreductase from catmint (*Nepeta racemosa*). Arch. Biochem. Biophys. 318, 105–112.

Hamann, T., Lindberg Møller, B., 2007. Improved cloning and expression of cytochrome P450s and cytochrome P450 reductase in yeast. Protein. Expr. Purif. 56, 121–127.

Heitz, T., Widemann, E., Lugan, R., Miesch, L., Ullmann, P., Désaubry, L., Holder, E., Grausem, B., Kandel, S., Miesch, M., Werck-Reichhart, D., Pinot, F., 2012. Cytochrome P450 CYP94C1 and CYP94B3 catalyze two successive oxidation steps of the plant hormone jasmonoyl-isoleucine for catabolic turnover. J. Biol. Chem. 287, 6296–6306.

Ikeda, H., Esaki, N., Nakai, S., Hashimoto, K., Uesato, S., Soda, K., Fujita, T., 1991. Acyclic monoterpene primary alcohol:NADP+ oxidoreductase *of Rauwolfia serpentina* cells: the key enzyme in biosynthesis of monoterpene alcohols. J. Biochem. 109, 341–347.

Khan, M.A., Jones, I., Loza-Reyes, E., Cameron, M.M., Pickett, J.A., Birkett, M.A., 2012. Interference in foraging behaviour of European and American house dust mites *Dermatophagoides pteronyssinus* and *Dermatophagoides farinae* (Acari: Pyroglyphidae) by catmint, *Nepeta cataria* (Lamiaceae). Exp. Appl. Acarol. 57, 65–74.

Lommen, A., 2009. MetAlign: interface-driven, versatile metabolomics tool for hyphenated full-scan mass spectrometry data preprocessing. Anal. Chem. 81, 3079–3086.

Mateos, R., Cert, A., Pérez-Camino, C.M., García, J.M., 2004. Evaluation of virgin olive oil bitterness by quantification of secoiridoid derivatives. J. Am. Oil Chem. Soc. 81, 71–75. Q10

Montiel, G., Breton, C., Thiersault, M., Burlat, V., Jay-Allemand, C., Gantet, P., 2007. *Catharanthus roseus* suspension cells. Transcription factor Agamous-like 12 from Arabidopsis promotes tissue-like organization and alkaloid biosynthesis. Metab. Eng. 9, 125–132.

Murata, J., Roepke, J., Gordon, H., De Luca, V., 2008. The leaf epidermome of *Catharanthus roseus* reveals its biochemical specialization. Plant Cell 20, 524–542.

Nour-Eldin, H.H., Hansen, B.G., Nørholm, M.H., Jensen, J.K., Halkier, B.A., 2006. Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. Nucleic Acids Res. 34, e122.

O'Connor, S.E., Maresh, J.J., 2006. Chemistry and biology of monoterpene indole alkaloid biosynthesis. Nat. Prod. Rep. 23, 532–547.

O'Connor, S.E., 2012. Strategies for engineering plant natural products: the iridoid-derived monoterpene indole alkaloids of *Catharanthus roseus*. Methods Enzymol. 515, 189–206.

Oda, A., Yamaotsu, N., Hirono, S., 2005. New AMBER force field parameters of heme iron for cytochrome P450s determined by quantum chemical calculations of simplified models. J. Comput. Chem. 26, 818–826.

Omura, T., Sato, R., 1964. The carbon monoxide-binding pigment of liver microsomes. J. Biol. Chem. 239, 2370–2378.

Ortiz de Montellano, P.R., Correia, M.A., 1995. Inhibition of cytochrome P450 enzymes. In: Ortiz de Montellano, P. (Ed.), Cytochrome P450: Structure, Mechanism, and Biochemistry, 2nd ed. Plenum, New York, pp. 305–364.

Otah, D., Mizutani, M., 1998. Plant Geraniol 10-Hydroxylase and DNA Coding Therof. US patent 5,753,507.

Peebles, C.A., Sander, G.W., Hughes, E.H., Peacock, R., Shanks, J.V., San, K.Y., 2011. The expression of 1-deoxy-D-xylulose synthase and geraniol-10-hydroxylase or

anthranilate synthase increases terpenoid indole alkaloid accumulation in *Catharanthus roseus* hairy roots. Metab. Eng. 13, 234–240.

Peebles, C.A., Hughes, E.H., Shanks, J.V., San, K.Y., 2009. Transcriptional response of the terpenoid indole alkaloid pathway to the overexpression of ORCA3 along with jasmonic acid elicitation of *Catharanthus roseus* hairy roots over time. Metab. Eng. 11, 76–86.

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

Robineau, T., Batard, Y., Nedelkina, S., Cabello-Hurtado, F., Leret, M., Sorokine, O., Didierjean, L., Werck-Reichhart, D., 1998. The chemically-inducible plant cytochrome P450 CYP76B1 actively metabolizes phenylureas and other xenobiotics. Plant Physiol. 118, 1049–1056.

Roepke, J., Salim, V., Wu, M., Thamm, A.M., Murata, J., Ploss, K., Boland, W., De Luca, V., 2010. Vinca drug components accumulate exclusively in leaf exudates of Madagascar periwinkle. Proc. Natl. Acad. Sci. 107, 15287–15292.

Rohloff, J., Bones, A.M., 2005. Volatile profiling of *Arabidopsis thaliana* – putative olfactory compounds in plant communication. Phytochemistry 66, 1941–1955.

Schattat, M., Barton, K., Baudisch, B., Klösgen, R.B., Mathur, J., 2011. Plastid stromule ranching coincides with contiguous endoplasmic reticulum dynamics. Plant Physiol. 155, 1667–1677.

Simkin, A.J., Miettinen, K., Claudel, P., Burlat, V., Guirimand, G., Courdavault, V., Papon, N., Meyer, S., Godet, S., St-Pierre, B., Giglioli-Guivarc'h, N., Fischer, M.J., Memelink, J., Clastre, M., 2013. Characterization of the plastidial geraniol synthase from Madagascar periwinkle which initiates the monoterpenoid branch of the alkaloid pathway in internal phloem associated parenchyma. Phytochemistry 85, 36–43.

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, 4637–4643.

Tikunov, Y., Lommen, A., de Vos, C.H.R., Verhoeven, H.A., Bino, R.J., Hall, R.D., Bovy, A.G., 2005. A novel approach for nontargeted data analysis for metabolomics. Large-scale profiling of tomato fruit volatiles. Plant Physiol. 139, 1125–1137.

Tundis, R., Loizzo, M.R., Menichini, F., Statti, G.A., Menichini, F., 2008. Biological and pharmacological activities of iridoids: recent developments. Mini Rev. Med. Chem. 8, 399–420.

van der Fits, L., Deakin, E.A., Hoge, J.H.C., Memelink, J., 2000. The ternary transformation system: constitutive *vir*G on a compatible plasmid dramatically increases Agrobacterium-mediated plant transformation. Plant Mol. Biol. 43, 495–502.

van der Heijden, R., Jacobs, D.I., Snoeijer, W., Hallard, D., Verpoorte, R., 2004. The Catharanthus alkaloids: pharmacognosy and biotechnology. Curr. Med. Chem. 11, 607–628.

van der Hooft, J.J.J., Vervoort, J., Bino, R.J., de Vos, R.C.H., 2012. Spectral trees as a robust annotation tool in LC–MS based metabolomics. Metabolomics 8, 691–703.

van Herpen, T.W.J.M., Cankar, K., Nogueira, M., Bosch, D., Bouwmeester, H.J., Beekwilder, J., 2010. *Nicotiana benthamiana* as a production platform for artemisinin precursors. PloS One 5, e14222.

Verma, P., Mathur, A.K., Srivastava, A., Mathur, A., 2012. Emerging trends in research on spatial and temporal organization of terpenoid indole alkaloid pathway in *Catharanthus roseus*: a literature update. Protoplasma 249, 255–268.

Viljoen, A., Mncwangi, N., Vermaak, I., 2012. Anti-inflammatory iridoids of botanical origin. Curr. Med. Chem. 19, 2104–2127.

Voinnet, O., Rivas, S., Mestre, P., Baulcombe, D., 2003. An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant J. 33, 949–956.

Wang, J., Liu, Y., Cai, Y., Zhang, F., Xia, G., Xiang, F., 2010. Cloning and functional analysis of geraniol 10-hydroxylase, a cytochrome P450 from *Swertia mussotii* Franch. Biosci. Biotechnol. Biochem. 74, 1583–1590.

Weigel, D., Glazebrook, J., 2006. Transformation of Agrobacterium using electroporation. Cold Spring Harb. Protoc. 7, http://dx.doi.org/10.1101/pdb.prot4666. **Q11**

Zhao, L., Sander, G.W., Shanks, J.V., 2013. Perspectives of the metabolic engineering of terpenoid indole alkaloids in *Catharanthus roseus* hairy roots. Adv. Biochem. Eng. Biotechnol. **Q12**

Please cite this article as: Höfer, R., et al., Geraniol hydroxylase and hydroxygeraniol oxidase activities of the CYP76 family of cytochrome P450 enzymes and potential for.... Metab. Eng. (2013), http://dx.doi.org/10.1016/j.ymben.2013.08.001