Running title: Biosynthesis and intracellular storage of bioactive diterpenes

To whom author correspondence should be sent:

Björn Hamberger
Plant Biochemistry Laboratory
Department of Plant and Environmental Sciences, University of Copenhagen
Thorvaldsensvej 40, DK-1871 Frederiksberg C, Copenhagen, Denmark
Phone: +45 3533 3328
Email: bjoernh@life.ku.dk

Research area: Biochemistry and metabolism

Key words:
Plant specialized metabolism, high-value bioactive diterpenes, pathway discovery, diterpene synthases,
lipid bodies, Coleus forskohlii, forskolin, manoyl oxide
Manoyl oxide (13R), the biosynthetic precursor of forskolin, is synthesized in specialized root cork cells in Coleus forskohlii.

Irini Pateraki2,*, Johan Andersen-Ranberg2,4,*, Britta Hamberger2, Allison Maree Heskes3, Helle Juel Martens2, Philipp Zerbe5, Søren Spanner Bach2, Birger Lindberg Møller2,4, Jørg Bohlmann5, Björn Hamberger2,4,5

2Department of Plant and Environmental Sciences, Faculty of Science, University of Copenhagen, Thorvaldsensvej 40, 1871 Copenhagen, Denmark
3School of Botany, University of Melbourne, Victoria 3010, Australia
4Center for Synthetic Biology bioSYNergy, Denmark
5Michael Smith Laboratories, University of British Columbia, Vancouver, BC, Canada V6T 1Z4

One sentence summary: The first two steps of the biosynthesis of forskolin, a complex bioactivelabdane diterpene, are active in Coleus forskohlii specialized root cork cells harboring a uniquehydrophobic intracellular compartment used for terpenoid storage.
Footnotes

1Financial support was provided by the UNIK [Universitetsforskningens Investeringskapital (Investment Capital for University Research)] Center for Synthetic Biology “bioSYNergy” at the University of Copenhagen, funded by the Research Initiative of the Danish Ministry of Science, Technology and Innovation (BH and BLM), and the Novo Nordisk Foundation Center for Biosustainability (BH and BLM). Additional support was provided by a personal postdoctoral stipend awarded by the FP7 PEOPLE MARIE CURIE ACTIONS Intra-European Fellowships (IP). This project was supported by the PhytoMetaSyn Project (funds to JB) funded by Genome Canada, Genome Alberta, Government of Alberta, Genome British Columbia, Genome Prairie, Genome Quebec, National Research Council of Canada, Canada Foundation for Innovation, Ontario Ministry of Research and Innovation, McGill Innovation Centre, MDEIE (Ministère du Développement Économique, Innovation et Exportation). Additional financial support was provided from the Villum Foundation to the research center ‘Plant Plasticity’ (BLM), from the Danish Council on Technology and Production Sciences (BLM), and by the Natural Sciences and Engineering Research Council of Canada (NSERC) (funds to JB).

*These authors have equally contributed to this work

5Corresponding author: bjoernh@life.ku.dk

To whom correspondence should be addressed: Björn Hamberger, Plant Biochemistry Laboratory, Department of Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensvej 40, DK-1871 Frederiksberg C, Copenhagen, Denmark; phone: +45 3533 3328, fax: +45 3533 3300, email: bjoernh@life.ku.dk
ABSTRACT

Forskolin, a complex labdane diterpenoid found in the root of Coleus forskohlii (Lamiaceae), has received attention for its broad range of pharmacological activities, yet the biosynthesis has not been elucidated. We detected forskolin in the root cork of C. forskohlii in a specialized cell type containing characteristic structures with histochemical properties consistent with oil bodies. Organelle purification and chemical analysis confirmed the localization of forskolin and of its simplest diterpene precursor backbone, (13R) manoyl oxide, to the oil bodies. The labdane diterpene backbone is typically synthesized by two successive reactions catalyzed by two distinct classes of diterpene synthases. We have recently described the identification of a small gene family of diterpene synthase candidates (CfTPSs) in C. forskohlii. Here, we report the functional characterization of four CfTPSs using in vitro and in planta assays. CfTPS2 (which synthesizes the intermediate copal-8-ol diphosphate) in combination with CfTPS3 resulted in the stereospecific formation of (13R) manoyl oxide, while the combination of CfTPS1 and CfTPS3 or CfTPS4 led to formation of miltiradiene, precursor of abietane diterpenoids in C. forskohlii. Expression profiling and phylogenetic analysis of the CfTPS family further support the functional diversification and distinct roles of the individual diterpene synthases and the involvement of CfTPS1-4 in specialized metabolism and of CfTPS14 and CfTPS15 in general metabolism. Our findings pave the way towards the discovery of the remaining components of the pathway to forskolin, likely localized in this specialized cell type and support a role of oil bodies as storage organelles for lipophilic bioactive metabolites.
INTRODUCTION

Coleus forskohlii (synonym: Plectranthus barbatus) is a perennial medicinal shrub of the mint family (Lamiaceae) indigenous to the subtropical and temperate climate zones of India and south-east Asia (Kavitha et al., 2010). The plant has been used since ancient times in Hindu and Ayurvedic traditional medicine for treating a broad range of human health disorders (Valdés et al., 1987; Lukhoba et al., 2006). The main active compound of C. forskohlii is forskolin, a heterocyclic labdane type diterpene found in the roots of the plant (Ammon and Müller, 1985). The diverse known and potential pharmaceutical applications of forskolin extend from alleviation of glaucoma (Wagh et al., 2012), anti-HIV or antitumor activities (Li and Wang, 2006) to treatment of hypertension and heart failure (Yoneyama et al., 2002). The efficacy of forskolin relies on activation of the adenylate cyclase enzyme (Seamon et al., 1981; Daly, 1984) leading to a marked increase of the intracellular level of cAMP (3'-5'-cyclic adenosine monophosphate) in mammalian in vitro and in vivo systems (Alasbahi and Melzig, 2010). The semi-synthetic forskolin derivative NKH477 has been approved for commercial use in Japan for treatment of cardiac surgery complications, heart failure, and cerebral vasospasm (Toya et al., 1998; Kikura et al., 2004), while a forskolin eye drop solution was recently approved as an effective treatment for glaucoma (Wagh et al., 2012). Despite much effort, the full stereospecific synthesis of forskolin has not yet been achieved with current methods yielding racemic mixtures (Ye et al., 2009).

The chemical complexity of C. forskohlii has been well studied and a total of 68 different diterpenoids have been isolated and identified from different tissues of the plant, of which 25 belong to the class of abietanes and 43 to the class of labdanes (Alasbahi and Melzig, 2010). While the tricyclic abietane diterpenes have been reported to accumulate predominantly in the aerial parts, labdane diterpenoids with a bicyclic decalin core were detected primarily in the roots. Forskolin is a representative of an unusual series of tricyclic (8,13)-epoxy-labdanes, characteristic for this plant. Given its importance as a pharmaceutical, we set out to discover genes involved in the biosynthesis of forskolin. Our general approach utilized and built on a recently established diterpene pathway gene discovery strategy in non-model systems (Zerbe et al., 2013). As guiding principles we considered possible spatial separation of different diterpenoid pathways in C. forskohlii, the large diversity of diterpenoids found in C. forskohlii and the particular heterocyclic structure of forskolin. Based on the identification of putative labdane-
intermediates en route to forskolin in hairy root cultures of *C. forskohlii*, a pathway from trans-
eranylgeranyl diphosphate (GGPP) to forskolin has recently been hypothesized (Asada et al., 2012).

GGPP is synthesized in plant plastids by the plastidial methylenetetrahydropyran 4-phosphate (MEP) pathway
(Rodriguez-Concepcion and Boronat, 2002). For the formation of labdane-type diterpenoids in
angiosperms, GGPP undergoes an initial protonation-initiated cationic cycloisomerization to a
labdadienyl/copalyl diphosphate intermediate, catalyzed by class II diterpene synthase (diTPS)
enzymes (Peters, 2010; Chen et al., 2011). The class II diTPS reaction, which forms the bicyclic
decalin core of labdane-type diterpenoids, is terminated either by deprotonation or by water capture of
the diphosphate carbocation. Subsequently, class I diTPSs catalyze cleavage of the diphosphate group
and often additional cyclization or rearrangement reactions on the resulting carbocation. As with the
class II diTPSs, deprotonation or water capture terminate the class I diTPS reaction. Water capture as a
mechanism of diTPSs that leads to oxygen functionalities in the diterpene products has been described
for the class II active site of bifunctional class I/II gymnosperm diTPSs (Keeling et al., 2011; Zerbe et
al., 2012; Zerbe et al., 2013), class II angiosperm diTPSs (Falara et al., 2010; Caniard et al., 2012;
Zerbe et al., 2013) and class I angiosperm diTPS (Caniard et al., 2012). Further oxidative
functionalization of diterpenoids is typically catalyzed by cytochrome P450-dependent monooxygenase
enzymes (P450s) (Ro et al., 2005; Swaminathan et al., 2009; Hamberger et al., 2011; Wang et al.,
2011; Guo et al., 2013; Hamberger and Bak, 2013), which then provide molecular handles for addition
of auxiliary functional groups, leading ultimately to highly complex and decorated products such as
forskolin and its derivatives.

Recently, Zerbe et al. (2013) reported the sequencing and assembly of a root transcriptome of *C.
forskohlii* using 454 and Illumina sequencing technologies. Mining of the transcriptome databases
resulted in the identification of a panel of candidate *CfTPSs*. Here, we describe the functional
characterization of five of them. Specifically, we demonstrate the function of *CfTPS2* in combination
with *CfTPS3* in the stereospecific biosynthesis of (13R) manoyl oxide, the putative precursor of
forskolin. We show that forskolin accumulates within a specific cell type of the root, the cork cells. Oil
body-like structures, characteristic and unique for this type of cell are found in the cytosol. We propose
that these oil body-like structures facilitate the accumulation of high amounts of lipophilic diterpene
metabolites.
RESULTS

Identification of unique lipophilic organelles in *C. forskohlii* root cork cells

Although high-level accumulation of forskolin-related labdanes in the root of *C. forskohlii* is well established, the localization of the biosynthetic pathway and specialized anatomical structures for the storage of diterpenes have not been reported. When transverse sections of *C. forskohlii* root (Fig. 1A) were examined using light microscopy, we found that cells of the root cork contained oil body-like structures (hereafter termed oil bodies) with a typical distribution of one oil body per cell of the root cork (Fig. 1B). These oil bodies appeared to be highly characteristic of the cork tissue, since they were not found in other cell types of the root. In sections of younger parts of the root, the oil bodies appeared yellow and occurred as single compartments within the cytoplasm of the immature cork cells (Fig. 1B). The color and pigmentation of oil bodies was observed to change with tissue maturation, possibly indicating developmental changes in their metabolite profile (Fig. 1B). In addition, cells containing more than one oil body were occasionally seen in older tissue (Fig. 1C). Further microscopy of other tissues revealed only sporadic occurrence of oil bodies in stem cork tissue (Supplemental Fig S1).

To probe the nature of the oil bodies, histochemical staining was performed with Nile Red (9-diethylamino-5H-benzo-[α]-phenoxazine-5-one). Nile Red is a selective lipid-specific dye that is strongly fluorescent only in a hydrophobic environment. Its emission spectrum shifts depending on the polarity of its microenvironment, fluorescing magenta in the presence of polar lipids and green in the presence of neutral lipids (Diaz et al., 2008). Confocal laser scanning microscopy of *C. forskohlii* root cork stained with Nile Red indicated that the observed structures were indeed oil bodies and that the composition of the lipophilic content was heterogeneous. Both neutral (Fig. 1D) and polar (Fig. 1E) lipophilic compounds were observed to be non-uniformly distributed in the oil bodies with globules of neutral lipids dispersed in a predominantly polar lipid matrix (Fig. 1D-F).

Localization of forskolin and (13R) manoyl oxide in root cork oil bodies

As a first step to test whether the occurrence of forskolin is associated with the presence of oil bodies, we confirmed the localization of forskolin in the root cork. Separate methanol extracts of the root cork
and the root cortex and stele were analyzed by high-performance liquid chromatography (HPLC) using an evaporative light scattering detector (ELSD) and compared with flowers, leaves and stems. Forskolin was primarily detected in the root cork, and was not found in the root cortex and stele, leaves or flowers (Fig. 2). Traces of forskolin were detected in the stem, consistent with the observed sporadic presence of a small number of oil bodies in stem cork tissue (Supplemental Fig. S1). To further examine if forskolin was present specifically in the oil bodies, methanol extracts of isolated oil bodies purified to apparent homogeneity (Fig. 3C) were subjected to HPLC-mass spectrometry (MS). Forskolin was detected in these structures along with ions matching forskolin related compounds (Fig. 3A; Table 1). Ion peaks consistent with abietane type diterpenes previously reported from *C. forskohlii* were also present in oil bodies (Table 1).

In order to search for diterpenoid backbones that could serve as precursors to forskolin or to different abietane diterpenoids, isolated oil bodies as well as root cork, root cortex and stele, stem, leaves, and flowers were extracted with hexane and the resulting extracts analyzed by gas chromatography-mass spectrometry (GC-MS) (Figs. 3 and 4). (13R) manoyl oxide was detected in root cork and, more specifically, in oil bodies isolated from this tissue (Fig. 3C), as well as in the root cortex and stele, stem and flowers (Figs. 3B, 3E and 4). The localization of this compound to oil bodies is noteworthy given that it is considered the simplest diterpenoid backbone structure of forskolin and contains the correct stereochemical configuration (Fig. 3D). The abietane type diterpenes miltiradiene, abietadiene and dehydroabietadiene were also detected in various tissues (Fig. 4). Dehydroabietadiene was predominantly found in both types of root tissue while abietadiene was mainly detected in the root cork tissue and miltiradiene in the stem and leaf tissue of *C. forskohlii* (Fig. 4).

*C. forskohlii* diTPSs constitute a small gene family

To investigate the molecular underpinnings of the diversity of diterpenoids found in the roots of *C. forskohlii*, we mined the root transcriptome of *C. forskohlii* for the identification of *CfTPS* candidates (Zerbe et al., 2013). A panel of six diterpene synthases was identified, *CfTPS1, CfTPS2, CfTPS3, CfTPS4, CfTPS14* and *CfTPS15* which, with exception of *CfTPS15*, represented full-length cDNAs with predicted N-terminal plastidial transit peptides. *CfTPS1, CfTPS2*, and *CfTPS15* contained the Asp-
rich conserved motif DxDD characteristic of class II diTPS, while *CfTPS3*, *CfTPS4* and *CfTPS14* carried the DDxxD signature motif of class I diTPS (Supplemental Fig. S2). To further substantiate the classification, relationships and evolution of the CfTPSs, we performed separate phylogenetic analyses of class II and class I CfTPSs including functionally characterized representatives from the Lamiaceae and other angiosperm species. Included in the phylogenies were representative gymnosperm class II and class I diTPSs [PgCPS and PgEKS; (Keeling et al., 2010)] and the bifunctional diTPS from the moss *Physcomitrella patens* [PpCPS/EKS; (Hayashi et al., 2010)], as it is considered an ancestral archetype of plant diTPSs (Fig. 5).

Among the class II diTPSs, two distinct clades were apparent for dicotyledon species. One clade includes the single copy Arabidopsis copalyl-diphosphate synthase (AtCPS) involved in general metabolism of gibberellic acid formation and orthologous from other species with characterized CPS functions. Also in this clade are paralogous gene pairs from *Cucurbita maxima* (CmCPS1 and CmCPS2) and *Isodon eryocalyx* (IeCPS1 and IeCPS2), which resulted from gene duplications potentially allowing for neofunctionalization of diTPS in the biosynthesis of specialized metabolites (Li et al., 2012a). In the second and apparently more divergent clade, enzymes of *Cistus creticus* (CcCLS), tobacco (*Nicotiana tabacum*, NtLPPS) and two closely related Lamiaceae (*Salvia sclarea*, SsLPPS and *S. miltiorrhiza*, SmCPS1) are involved in specialized metabolism and have been shown to produce copal-8-ol diphosphate and (+)-copalyl diphosphate, respectively (Gao et al., 2009; Caniard et al., 2012). In *C. forskohlii* we identified CfTPS15, a single copy class II diTPSs grouping with the bona fide CPS (general metabolism), and CfTPS1 and CfTPS2, a pair of diTPSs grouping with the class II enzymes involved in specialized metabolism.

The overall topology of the class I diTPS phylogeny resembles the class II phylogeny. One clade includes the ent-kaurene synthase (EKS) of general metabolism of gibberellic acid formation together with a set of rice (*Oryza sativa*) class I diTPS which evolved through species-specific duplications associated with neofunctionalization in rice specialized metabolism (Morrone et al., 2011). CfTPS14 falls into this clade. A separate clade contains members involved in specialized metabolism, including tobacco NtCAS, converting copal-8-ol diphosphate to the specialized metabolite *cis*-abienol (Sallaud et al., 2012). The enzymes from *S. sclarea* and *S. miltiorrhiza*, producing the labdane specialized metabolites sclareol (SsSCS) and the abietane miltiradiene (SmMRS) respectively (Gao et al., 2009;
Caniard et al., 2012), are members of a subgroup consisting of Lamiaceae enzymes. A second pair of C. forskohlii diTPS, CfTPS3 and CfTPS4, is clustered with this clade. All members of this group of Lamiaceae sequences represent bi-domain (βα) class I diTPS, lacking the γ-domain.

Transcript levels of C. forskohlii diTPSs in various tissues

To correlate the transcript levels of CfTPS genes with accumulation of forskolin related labdane-diterpenoids and abietane diterpenoids in C. forskohlii tissues, quantitative (q)RT-PCR analysis was performed using cDNA templates derived from total RNA extracted from root cork, root cortex and stele (root without cork), leaves, stems and flowers. CfTPS1, CfTPS2 and CfTPS3 shared similar transcript profiles patterns across all tissues, with high transcript accumulation in root cork cells, up to 1000-fold in comparison with all other tissues tested (Fig. 6). These data provide support for the involvement of CfTPS1, CfTPS2 and CfTPS3 in the formation of specialized metabolites in the root cork. In contrast, the transcript levels of CfTPS4, CfTPS14 and CfTPS15 were relatively low across all tissues tested. Despite the close phylogenetic relation of CfTPS3 and CfTPS4 (Fig. 5), they show surprisingly different expression patterns. In contrast to CfTPS3, CfTPS4 transcripts were mostly detected in the aerial parts of the plant, especially in the leaves, while its transcripts accumulate only to very low levels in the root (Fig. 6).

Transcript profiles of the individual CfTPSs, including both spatial patterns of distribution across tissues and relative transcript abundance, support a role for CfTPS14 and CfTPS15 in general metabolism and the involvement of CfTPS1, CfTPS2 and CfTPS3 in localized, active biosynthesis of specialized metabolites. These patterns are also in agreement with the results of the phylogenetic analysis (Fig. 5). Transcript profiles did not provide indications for a role of CfTPS4.

In vitro functional characterization of C. forskohlii diTPSs

For the functional characterization of the CfTPSs described here (except for CfTPS15, for which no full length sequence could be retrieved), cDNAs were heterologously expressed in E. coli with a C-terminal 6xHis epitope tag. Purified recombinant proteins were tested individually in single or coupled
in vitro assays, supplied with appropriate substrates and the reaction products were analyzed by GC-MS. Products of the in vitro single assays with the class II diTSPs, CfTPS1 and CfTPS2, were treated with alkaline phosphatase before GC-MS analysis.

Enzyme assays with CfTPS1 yielded a diterpene with a mass spectrum matching copal-15-ol, indicating that the primary product before dephosphorylation is copalyl diphosphate (Fig. 7A). Assays of CfTPS2 resulted in the formation of 13(16)-14-labdien-8-ol and labd-13-en-8,15-diol as major products (Fig. 7B), supporting a function as labda-13-en-8-ol (or copal-8-ol) diphosphate synthase, similar to the functions of previous reported similar enzymes (Falara et al., 2010; Caniard et al., 2012; Sallaud et al., 2012; Zerbe et al., 2013). We also detected the non-stereoselective formation of the (13R) and (13S) epimers of manoyl oxide, which were previously observed in in vitro reactions of similar class II diTPSs and were suggested to be the result of a non-enzymatic reaction (Caniard et al., 2012; Zerbe et al., 2013). These results indicate that CfTPS1 and CfTPS2 represent functionally distinct class II diTPSs, both necessary and sufficient to form the diphosphate intermediates required for the abietane and labdane classes of diterpenoids detected in the root of C. forskohlii.

The overall expression level of CfTPS14 in E. coli cells was low, with little activity in coupled assays with CfTPS1 and no detectable product formation in combination with CfTPS2 (data not shown). However, we have previously suggested a function based on co-expression with a diTPS from Euphorbia peplus (Zerbe et al., 2013), which supports a function of CfTPS14 in the general metabolism of gibberellin phytohormones. Assays of CfTPS1 coupled to either CfTPS3 or CfTPS4 resulted in formation of miltiradiene (Fig. 7A), an abietane diterpene formed from a copalyl diphosphate intermediate (Gao et al., 2009) and is consistent with the results of the single enzyme assay of CfTPS1. Coupled assays with CfTPS2 and CfTPS3 showed the formation of the pure (13R) enantiomer of manoyl oxide (Fig. 7B), supporting a role of CfTPS2 in formation of the copal-8-ol diphosphate intermediate as detected in the single enzyme assay. In coupled assays of CfTPS2 with CfTPS4, both (13R) and (13S) manoyl oxide epimers were detected, albeit at a ratio different from the dephosphorylation product of CfTPS2 alone (Fig. 7B). The stereospecific production of the (13R) manoyl oxide epimer when CfTPS2 (class II) and CfTPS3 (class I) enzymes were combined, along with their transcriptional co-expression in root cork tissue where (13R) manoyl oxide and forskolin were detected, supports their involvement in the formation of the intermediate of forskolin in planta.
Single enzyme assays using CfTPS3 and CfTPS4 with GPP as substrate did not result in any product formation (Supplemental Fig. S3).

**In planta heterologous expression and functional characterization of C. forskohlii diTPSs**

To validate the results obtained with CfTPS produced as recombinant proteins in *E. coli*, we investigated the *in planta* function of the CfTPSs by transient heterologous expression in *Nicotiana benthamiana* leaves after agroinfiltration. GC-MS analyses of extracts from *N. benthamiana* leaves transiently expressing the individual class I CfTPS3, CfTPS4 and CfTPS14 did not result in detectable accumulation of additional metabolites compared to control plants (data not shown). Extracts from *N. benthamiana* expressing the class II CfTPS1 alone showed only trace amounts of additional diterpenes compared to the controls, none of which could be accurately identified (Fig. 8B). Consistent with the *in vitro* enzyme assays, both (13R) and (13S) epimers of manoyl oxide were identified in the extracts from *N. benthamiana* expressing the class II CfTPS2 (Fig. 8A). Co-expression of CfTPS2 and CfTPS14 did not change the product profile compared to expression of CfTPS2 alone, suggesting that CfTPS14 does not accept the copal-8-ol diphosphate as substrate (Fig. 8A). In extracts of plants co-expressing CfTPS1 with CfTPS3 or CfTPS4 miltiradiene was observed as the main product together with minor traces of dehydroabietadiene and abietadiene (Fig. 8B).

In extracts from *N. benthamiana* co-expressing CfTPS2 with CfTPS3 or CfTPS4, only the (13R) epimer of manoyl oxide was identified (Fig. 8A), consistent with the stereochemical conformation of forskolin and the related series of labdane-type diterpenoids. This result suggests that the class I CfTPS3 and CfTPS4 can accept the copal-8-ol diphosphate synthesized by CfTPS2 and catalyze the stereospecific formation of (13R) manoyl oxide.
DISCUSSION

Intracellular accumulation of diterpenes in the *C. forskohlii* root cork may be facilitated by unique hydrophobic oil-body structures: A suitable means for near-surface sequestration of terpenoids in roots.

Plants have evolved both specialized mechanisms and specialized anatomical structures for the secretion, sequestration and accumulation of defense-related and potentially toxic molecules (Morant et al., 2008; Schilmiller et al., 2008; Sirikantaramas et al., 2008). These metabolites may otherwise display adverse activities for the producing plant cell. Intracellular storage of such biologically active metabolites in the vacuole is well established for water soluble compounds and for compounds which become water soluble through conjugation (Marinova et al., 2007; Ferreres et al., 2011; Zhao et al., 2011; Li et al., 2012b). Similarly, for the large class of often lipophilic terpenoids, it has been suggested that their intracellular accumulation may be limited by non-specific interference with cellular processes and structures, such as interaction with membrane integrity (Uribe et al., 1985; Gershenzon and Dudareva, 2007; Sirikantaramas et al., 2008; Zore et al., 2011). Plant anatomical features and cellular structures typically associated with the biosynthesis and storage of large amounts of terpenoids are well studied and include glandular trichomes (Gershenzon et al., 2000; Iijima et al., 2004; Siebert, 2004; Schilmiller et al., 2008; Xie et al., 2008; Chatzopoulou et al., 2010; Lane et al., 2010), laticifer cells (Mahlberg, 1993; Post et al., 2012), resin cells, resin blisters or resin ducts (Martin et al., 2002; Zulak and Bohlmann, 2010), as well as glandular cavities lined by epithelial cells (Heskes et al., 2012; Voo et al., 2012).

Here, we showed accumulation of the partly lipophilic diterpenoid forskolin and the presence of its precursor, (13R) manoyl oxide, together with a series of other labdane and abietane type diterpenoids, in root cork cells of *C. forskohlii*. In these unique cell types, large oil bodies represent dominant and highly characteristic intracellular compartments which are not found in other tissues of the plant. Interestingly, the presence of both forskolin and (13R) manoyl oxide in these oil bodies indicates that the biosynthetic route to forskolin is active in this specific cell type. Accumulation of forskolin specifically in these oil bodies may facilitate the effective storage of terpenoids in a localized intracellular and lipophilic environment that is compartmentalized from the rest of the cytosol. Our
designation of these oil bodies as terpenoid accumulation compartments in the cork cells of *C. forskohlii* roots extends the known anatomical repertoire of specialized cells and cell compartments for producing and accumulating biologically active terpenoids in plants. Given the biological activities of diterpenoids such as forskolin and their high level of accumulation in oil bodies of root cork cells, the diterpenoid enriched root cork could provide a barrier against soil born pests or pathogens. For the roots of *C. forskohlii*, the near-to-the-surface sequestration of terpenoids in a cork layer may serve a similar protective function below ground as is attributed to glandular trichomes on surfaces of the above ground parts of plants. Glandular trichomes exist on the above ground parts of *C. forskohlii* but are generally not known to exist on root surfaces. The exclusive accumulation of forskolin, and detection of manoyl oxide and a series of related labdane diterpenoids in the cells of the root cork and the fact that labdane type diterpenoids have been implied before in plant defense against microbes (Fragoso-Serrano et al., 1998; Habibi et al., 2000; Peters, 2006; da Silva et al., 2008), suggest a specific role of these compounds in the interaction of the plant root with the soil ecosystem and defense.

### A new role for oil bodies

Oil bodies are intracellular organelles mainly consisting of neutral lipids such as triacylglycerides, which have been suggested to function as carbon and energy reserves activated under starvation or rapid growth conditions (Penno et al., 2013). In some plant species oil bodies have also been associated with stress responses, hormone signaling, plant growth and development, and sterol biosynthesis (Fujimoto and Parton, 2011; Chapman et al., 2012; Silvestro et al., 2013). In liverworts, these types of organelles have been suggested to co-occur with terpenoids (He et al., 2013). In vascular plants, oil body structures have been detected in several parts of the plant including seeds, leaves, flowers, pollen and fruits. Oil bodies are considered to be derived from the endoplasmic reticulum (ER), resulting from the accumulation of triacylglycerides between the ER bilayer, followed by budding of the cytoplasm oriented membrane to form the newly generated oil bodies (Beller et al., 2010; Chapman et al., 2012; Murphy, 2012). Our findings of oil bodies with a heterogeneous composition including diterpenoids
and localization to the cork of the C. forskohlii root suggest additional functions of oil bodies in the
sequestration of specialized metabolites and possibly plant defense (as discussed above).

Evolution of the C. forskohlii diTPS family towards chemical diversity

Screening of the C. forskohlii root transcriptome resulted in the identification of a small CfTPS gene family. Single-copy class I and class II diTPS candidates likely required for general metabolism, as well as pairs of duplicated class I and class II diTPS candidates potentially involved in specialized metabolism have been identified. Based on phylogenetic relations and transcript profiles we hypothesized involvement of CfTPS1, CfTPS2, CfTPS3 and CfTPS4 in specialized metabolism and a role for CfTPS14 and CfTPS15 in general metabolism of gibberellin phytohormone biosynthesis. DiTPS genes of general metabolism can serve as a template for gene duplication which in turn could facilitate the evolution of genes with novel functions in specialized metabolism [for review, see (Chen et al., 2011)]. Here, we traced such events from the phylogenetic relationships of the six different CfTPSs and substantiated them with functional characterization.

Originating from a class I/II bifunctional diTPS progenitor, angiosperm and gymnosperm diTPSs are thought to have evolved in two distinct clades, the class II or CPS and CPS-derived and the class I or EKS and EKS-derived enzymes (Chen et al., 2011). Additional clades of bifunctional or mono-functional diTPSs exist outside of the angiosperms (Zerbe et al., 2013). Within the two clades of class I and class II diTPSs, several branches of angiosperm diTPSs involved in specialized metabolism indicate events of lineage-specific diversification through gene duplication and neo-functionalization. Events of neo-functionalization may be limited to recruitment of the enzyme to a novel pathway, yet retaining its original CPS or EKS enzymatic function, or may have led to novel enzymatic functions (with or without loss of the original function). Examples are class II diTPSs involved in the metabolism of specialized diterpenes, such as Isodon eriocalyx IeCPS2 (Li et al., 2012a) derived from a recent duplication of IeCPS1 and retaining ent-CPS activity. In contrast, Salvia miltiorrhiza SmCPS, Nicotiana tabacum NtLPPS and Salvia sclarea ScSCS represent CPS-like diTPS with novel enzymatic functions (Gao et al., 2009; Caniard et al., 2012; Sallaud et al., 2012). Results from our work in C.
*C. forskohlii* support a model of evolution through gene duplication, neo-functionalization and loss of ancestral *ent*-CPS activity leading to CfTPS1 and CfTPS2.

For the class I clade of diTPS, separation of functionalities in general and specialized metabolism may be more challenging to assign. Enzymes of this clade may acquire the capacity to accept novel substrates while also retaining activity towards copalyl diphosphate. CfTPS14 appears to be a *bona fide* EKS with a likely function in gibberellin biosynthesis. Our results with both CfTPS3 and CfTPS4, when coupled with functionally distinct class II diTPS CfTPS1 or CfTPS2, indicate substrate promiscuity and possible redundancy between these two EKS-like enzymes in specialized metabolism. However, the distinct transcript profiles of *CfTPS3* and *CfTPS4* also indicated that the two enzymes may be involved in different functional contexts or pathways in *C. forskohlii*. *CfTPS4* is mainly expressed in the aerial tissues of the plant, where forskolin does not accumulate. It cannot be excluded though, that CfTPS4 is coupled, in these tissues, with specific type II diTPSs not identified in our study.

Taken together, results from *in planta* and *in vitro* assays show that a relatively small family of diTPS can generate an array of different diterpene core structures (Fig. 9) when individual enzymes are combined in higher level functional modules of pairs of class II and class I enzymes. The diterpene structures detected in our assays match those known to exist in *C. forskohlii* and represent the starting points for oxidative decorations in the plant.

**Involvement of CfTPS2 and CfTPS3 in the formation of (13R) manoyl oxide *en route* to forskolin in root cork cells**

We showed here that the two class I diTPS CfTPS3 and CfTPS4 can each be effectively coupled with either of the two class II enzymes, CfTPS1 and CfTPS2 (Fig. 9), highlighting the remarkable modularity of diterpene metabolism. *CfTPS1* was also found highly expressed in the same tissue as *CfTPS2* and *CfTPS3*, indicating that cork cells may have the capacity to produce a suite of both labdane and abietane type diterpenoids. Following recent gene duplication, resulting in the *CfTPS3* and *CfTPS4* pair, a change in the expression pattern could have resulted in novel functional modules of diTPSs in...
different tissues. This finding is supported by the gene expression studies together with the non-
targeted metabolomic HPLC-MS and GC-MS analyses of the root cork metabolites, showing highly
complex profiles with accurate masses characteristic of diterpenoids of labdane and abietane classes
(Fig. 4 and Table 1).

Detection of both forskolin and (13R) manoyl oxide along with the accumulation of CfTPS2 and
CfTPS3 transcripts in the specialized root cork cell type, together with the in vitro and in vivo
functional characterization of CfTPS2 and CfTPS3 suggest that these two enzymes are involved in the
early steps of forskolin biosynthesis via (13R) manoyl oxide in the root of C. forskohlii (Fig. 9).
Transcript profiles of CfTPS4 would make a similar conclusion for this diTPS less obvious. Racemic
mixtures of (13R) and (13S) manoyl oxide have previously been reported as artifacts or side products
of enzyme assays (Caniard et al., 2012; Zerbe et al., 2013), a result also described here in our in vitro
assay with CfTPS2 and CfTPS4. However, the stereospecific cyclisation of GGPP into the (13R)
epimer of manoyl oxide by the functional module of CfTPS2 coupled with either CfTPS3 (or CfTPS4)
in vivo in the N. benthamiana expression system is consistent with the stereochemical conformation of
forskolin and the series of forskolin-related labdanes detected in C. forskohlii (Asada et al. 2012), as
well as with the absence of the (13S) epimer of manoyl oxide in the plant tissues analyzed (Fig. 4).

While the significance of (13R) manoyl oxide is attributed to its role as a forskolin precursor, manoyl
oxide itself exhibits a number of important properties, like antibacterial or anti-inflammatory activities
or as a potent anticancer agent (de las Heras and Hoult, 1994; Demetzos et al., 1994; Dimas et al.,
1999; Angelopoulou et al., 2001).

Functions for CfTPS1 and CfTPS3 or CfTPS4 in the formation of miltiradiene

Assays with modules of CfTPS1 coupled to either the CfTPS3 or CfTPS4 resulted both in vitro and in
vivo in the formation of miltiradiene, an intermediate in the biosynthesis of tanshinones in the closely
related S. miltiorrhiza (Gao et al., 2009). While miltiradiene readily rearranged to dehydroabietadiene
in our experimental setup, the in planta route to the aromatic dehydroabietadiene-derived diterpenoids,
which are the second dominant type in C. forskohlii (Alasbahi and Melzig, 2010), remains unclear.
Recently, enzymatic conversion of miltiradiene to ferruginol, a 10-hydroxy-dehydroabietadiene, by a member of the CYP76 family was reported from *S. miltiorrhiza* (Guo et al., 2013). The identification of cytochrome P450s involved in the oxidative decoration of (13R) manoyl oxide *en route* to forskolin and related labdane diterpenoids is now underway using existing *C. forskohlii* P450 candidate resources and through development of new transcriptome and proteome resources for the specialized cell type of the *C. forskohlii* root cork cells.

**CONCLUSIONS**

Here we reported the functional characterization of a panel of class I and II diTPSs, including a specific pair of diTPS involved in the biosynthesis of (13R) manoyl oxide in *C. forskohlii*. Manoyl oxide, itself an anticancer compound, is a structurally unusual diterpenoid carrying an oxygen containing-heterocycle and represents the molecular core for a large series of bioactive derivatives including the cAMP modulator, forskolin. We have shown that both manoyl oxide and forskolin, accumulate together in a specific root cork cell type in *C. forskohlii*, indicating a role of these cells in protective plant-environment interactions in the rhizosphere. Biosynthesis and accumulation of specialized diterpenoids in these cells may be supported by the presence of oil bodies, providing the equivalent of a compartment specialized for storage of hydrophobic bioactive metabolites. Our discovery of *C. forskohlii* diTPSs participating in the early steps of forskolin biosynthesis, as well as the involvement of oil bodies in the storage of diterpenoids, may have implications for the discovery of other biosynthetic pathways of terpenoids in non-model plants and in biotechnological applications aiming at the production of high-value terpenoid-based pharmaceuticals, flavors, fragrances and other terpenoid bioproducts.

**MATERIALS AND METHODS**

**Plant Growth and Microscopy**

*Coleus forskohlii* (Lamiaceae) plants were grown in the greenhouse at the University of Copenhagen, Denmark, under ambient photoperiod and 24°C day/17°C night temperatures. Transverse sections of
roots (diameter of approximately 1 to 5 mm) were prepared for histochemical analysis. Sections were performed by hand or by vibrating blade microtome (100 µm) (Microm HM 650 V) and observed unstained with a Leica DM 5000B or a Nikon Eclipse 80i light and fluorescence microscope. Additionally, root samples were fixed in a solution containing 2.5% glutaraldehyde, 2% paraformaldehyde, and 0.1 M sodium cacodylate buffer, pH 7.2 for 24 h, thereafter surface sections and cross sections from the root cork were incubated in 0.1 µg/mL Nile Red for identification of lipid components (Eltgroth et al., 2005). Images from intact cells were recorded on a Leica SP5X confocal laser scanning microscope (Leica Microsystems, Mannheim, Germany). A 20× water-immersion objective was used for all images. Nile red was excited with the 514 nm line from the argon laser and the emitted light was collected at 525-567 nm and at 648-698 nm, respectively. Autofluorescence was tested without dye using the same microscope settings. Contrast adjustments were carried out to improve clarity of images but did not alter overall appearance. Final image processing, cropping and mounting of the images were done with Adobe Photoshop CS2 and Illustrator CS2.

Diterpene profiling and forskolin quantification in C. forskohlii tissues

Tissue was extracted as described in (De Vos et al., 2007). Cold methanol acidified with formic acid (0.125%) was added to ground and frozen tissue samples in a ratio of 3:1 (solvent:tissue). Samples were sonicated in an ultrasonic bath at 23°C for 15 min at 40 kHz (Branson, 3510), filtered using 96 well filter plates and analyzed by HPLC (High-Performance Liquid Chromatography) equipped with an evaporative light scattering detector (ELSD). All tissue types were extracted in triplicate. The HPLC-ELSD system was comprised of a Shimadzu LC-20AT pump, SIL-20A HT autosampler and ELSD-LTI detector (Shimadzu Corporation). Samples were separated on a Synergi 2.5 µm Fusion-RP C18 column (50 × 2 mm; Phenomenex) at a flow rate of 0.2 ml min⁻¹ with column temperature held at 25°C. The mobile phase consisted of water with 0.1% formic acid (v/v; solvent A) and acetonitrile with 0.1% formic acid (v/v; solvent B). The gradient program was 20-100% B over 35 min, 100% B for 1 min followed by a return to starting conditions over 0.25 min which was then held for 15 min to allow the column to re-equilibrate. The ELSD drift tube temperature was held at 50°C and the nitrogen drying gas pressure was 3.8 bar. Forskolin was quantified by comparison to a standard series of forskolin (Sigma).
For the diterpene profiling of isolated oil bodies, extracts were lyophilized and then dissolved in methanol and analyzed by HPLC-ESI-HRMS. Separation was carried out on an Agilent 1100 Series HPLC unit (Agilent Technology) with column and gradient as described above. The LC unit was coupled to a Bruker microTOF mass spectrometer (Bruker Daltronics) for accurate mass measurements.

Isolation of *C. forskohlii* root cork oil bodies

For isolation of oil bodies from root cork tissue, approximately 15 g of tissue was gently ground in 100 mL extraction buffer (EB; 20 mM Tricine, 250 mM sucrose, 0.2 mM PMSF, pH 8.5); the homogenate was filtered through Miracloth (Calbiochem) and centrifuged at 3500 rpm for 10 min for separation of cellular debris. The supernatant was collected and transferred in centrifugation tubes. Buffer B (20% sucrose, 20 mM HEPES, 100 mM KCl, 2 mM MgCl<sub>2</sub>, pH 10.5) was overlaid (5 mL of B for 25 mL of supernatant) and samples were centrifuged for 40 min at 5000 g. The resulting floating oil bodies were collected carefully from the surface layer.

Identification and cloning of full length diTPS genes

Mining of the *C. forskohlii* databases was performed as previously described (Zerbe et al., 2013), using tBLASTn software and known angiosperm diTPSs as query (CPS and EKS) and guided full-length cloning of a number of putative class I and class II diTPS genes. Total RNA from *C. forskohlii* roots, extracted as previously described (Hamberger et al., 2011), was used for cDNA synthesis. First strand cDNA was synthesized using the Takara PrimeScript 1<sup>st</sup> strand cDNA synthesis kit and oligo-dT primer. Cloning of the putative diTPS genes was achieved after PCR amplification using gene specific primers that were designed based on the *in silico* sequences of the identified CfTPS genes (Supplemental Table S1). PCR products were cloned into the pJET1.2 vector and verified by sequencing.

Phylogenetic analysis
The phylogenetic analyses are based on manually inspected amino acid alignments (DIALIGN-TX) which were analyzed by PhyML 3.1 (four rate substitution categories, LG substitution model, BIONJ starting tree, 100 bootstrap repetitions) and visualization in treeview (rooted with the bifunctional copalyl diphosphate synthase/kaurene synthase from the moss Physcomitrella patens PpCPS/EKS [BAF61135]). The abbreviations and accession numbers of sequences used are given in Supplemental Table S2.

RNA extraction and Quantitative Real-Time PCR

Total RNA from C. forskohlii root cork was extracted according to Hamberger et al., (2011) and further purified using the Spectrum Plant Total RNA Kit (Sigma) while total RNA from leaves, flowers, stems and root cortex and stele was extracted using the Spectrum Plant Total RNA Kit (Sigma). RNA extraction was followed by on-column DNase I digestion. The integrity of RNA samples was evaluated using the RNA-nano assay on the Agilent 2100 Bioanalyser (Agilent Technologies). First-strand cDNAs were synthesized from 0.5 μg of total RNA, with oligo-dT primer, using the “SuperScript III First-Strand Synthesis System for RT-PCR” (Invitrogen). The resulting cDNA was diluted 10-fold for the qRT-PCR reactions. The absence of genomic DNA contamination was verified by primers designed in exon-intron spanning regions of the TIF4a (Translation Initiation Factor 4a) and EF1a (Elongation Factor 1a) reference genes. Quantitative real-time PCR reactions were performed with gene specific primers (Supplemental Table S1) and Maxima SYBR Green/Fluorescein qPCR Master Mix (Fermentas) on a Rotor-Gene Q cycler (Qiagen). The PCR reactions were performed using the following cycling parameters: 95°C for 7 min (enzyme activation), 35 cycles of 95°C for 15 s, 60°C for 30 s and 72°C for 30 s followed by a melting curve cycle from 60°C to 90°C. TIF4a and EF1a were used as reference genes as they showed the lowest variation across different tissues. No statistically significant differences were observed between the results obtained from the 2 different reference genes. The results were normalized with TIF4a. Relative transcript abundance was calculated as the mean of three biological replications (three different plants), while the reactions were performed in three technical replicates. Amplification efficiency was calculated with the “Real Time PCR Miner” (http://www.miner.ewindup.info/Version2). Efficiency-corrected ΔC_T values were used to quantify relative differences in target gene transcript accumulation. Primer specificity was assessed by agarose
gel analysis and sequencing of amplicons from representative reactions, as well as from melting curve analysis of every reaction.

**Functional characterization of CfTPS - *in vitro* assays**

For the expression of the CfTPS1, CfTPS2, CfTPS3, CfTPS4 and CfTPS14 in *E. coli*, pseudomature variants lacking predicted plastidial target sequences were cloned into the pET28b+ vector. The software ChloroP was used for prediction of the plastidial target sequence ([http://www.cbs.dtu.dk/services/ChloroP/](http://www.cbs.dtu.dk/services/ChloroP/)) (Emanuelsson et al., 1999). As the expression levels of the recombinant CfTPS3 was very poor, a codon optimized version was synthesized by GenScript USA Inc. and subsequently cloned into the same vector (sequence is given in Supplemental Table S3). pET28b+ constructs were transformed into *E. coli* BL-21DE3-C41 cells and inoculated in a starter culture with LB (Lysogeny Broth) media and 50 µg/mL kanamycin. A starter culture was diluted 1:100 in a 50 mL TB (Terrific Broth) medium with 50 µg/mL kanamycin and grown at 37°C and 180 rpm until the OD<sub>600</sub> reached 0.3-0.4. Cultures were cooled to 16°C and expression was induced at OD<sub>600</sub>~0.6 with 0.2 mM IPTG. Expression was done overnight and cells were harvested by centrifugation. Binding buffer (20 mM HEPES (pH 7.5), 0.5 M NaCl, 25 mM Imidazole, 5% glycerol), one protease inhibitor cocktail tablet per 100 mL (Roche) and 0.1 mg/L lysozyme were added to the cell pellet which was gently shaken for 30 min and subsequently lysed by sonication (Branson sonifier 250 duty cycle 30%, output control 2-3, 1/4” tip). The cell lysate was centrifuged for 25 min at 12000 g and the supernatant was subsequently used for purification of the recombinant proteins. Proteins encoded by *C. forskohlii* diTPS and the characterized monoterpene synthase SfCIN from *Salvia fruticosa* (Kampranis et al., 2007) were purified on 1 mL His SpinTrap™ columns (GE healthcare) using elution buffer (binding buffer with 325 mM Imidazole and 5 mM DTT) and desalted on PD MiniTrap™ G-25 columns (GE healthcare) with a desalting buffer (20 mM HEPES (pH 7.2), 350 mM NaCl, 5 mM DTT, 1 mM MgCl<sub>2</sub>, 5% glycerol). *In vitro* TPS assays were performed by adding 15 µM GGPP or geranyl diphosphate (GPP) and 100 µg purified CfTPS (or SfCIN) enzymes in 397 µL enzyme assay buffer (50 mM HEPES (pH 7.2), 7.5 mM MgCl<sub>2</sub>, 5% (v/v) glycerol, 5 mM DTT). Onto the reaction mix, 500 µL n-hexane (Fluka GC-MS grade) was gently added as an overlay. Assays were incubated for 60 min at 30°C and ~70 rpm and the hexane overlay was subsequently removed for GC-MS analysis.
Functional characterization of CfTPS - transient expression in *Nicotiana benthamiana*

Full length CfTPS cDNAs were cloned into the agrobacterium binary vector for plant transformation pCAMBIA1300_35Su with gene specific primers (see Supplemental Table I) by USER™ cloning described in (Nour-Eldin et al., 2006). Transient expression of CfTPSs with the gene silencing suppressor p19 protein (Voinnet et al., 2003) in *N. benthamiana* leaves and extraction of diterpenes were performed as recently described (Spanner et al., 2013). Hexane extracts of *N. benthamiana* expressing the gene silencing suppressor p19 protein alone were used as controls. Compounds of interest were identified by comparison of GC-MS total ion chromatogram (TIC) and extracted ion chromatograms (EIC) of \( m/z \) 275 and 272 from samples. The ion \( m/z \) 275 is characteristic of several labdane type diterpenes including manoyl oxide whereas \( m/z \) 272 is characteristic of several other non-labdane type diterpenes such as abietane like diterpenes. All extractions from *N. benthamiana* transiently expressing diTPSs were carried out in biological triplicates (different leaves/plants infiltrated with the same agrobacteria mixture).

Metabolite analysis from *in vitro* and *in planta* assays

For the gas chromatography-mass spectrometry (GC-MS) analysis of *N. benthamiana* leaves expressing the CfTPSs and specific *C. forskohlii* tissues, 500 µL GC-MS grade hexane was added to 2 leaf discs (Ø=3 cm) in a 1.5 mL glass vial. Samples were incubated at room temperature for 1 h in a Roto-Shake Genie® revolving at 40 cycles/min. After extraction, the solvent was transferred into new 1.5 mL glass vials and stored at -20°C until GC-MS analysis. 1 µL of hexane extract was injected into a Shimadzu GC-MS-QP2010 Ultra. Separation was carried out using an Agilent HP-5MS column (30 m × 0.250 mm i.d., 0.25 µm film thickness) with purge flow of 1 mL min⁻¹, using helium as carrier gas. The GC temperature program was: 50°C for 2 min, ramp at rate 4°C min⁻¹ to 110°C, ramp at rate 8°C min⁻¹ to 250°C, ramp at rate 10°C min⁻¹ to 310°C and hold for 5 min. Injection temperature was set at 250°C in splitless mode. For the GC-MS analysis of hexane extracts from *in vitro* assay the following GC-program was used: 100°C for 1 min, ramp at rate 10°C min⁻¹ to 250, ramp at rate 20°C min⁻¹ to 310°C and hold for 2 min. Compound identification was done by comparison to authentic standards (dehydroabietadiene, abietadiene), reference spectra from literature, databases and comparison of
retention time (miltiradiene, manoyl oxide, copalol, labd-13-en-8,15-diol and 13(16)-14-labdien-8-ol) (Sons, 2006; Adams, 2007; Lane et al., 2010) (Wiley Registry of Mass Spectral Data, 8th Edition, July 2006, John Wiley & Sons, ISBN: 978-0-470-04785-9). The differentiation of the C-13 epimers (13R) and (13S) manoyl oxide was performed as previously described (Demetzos et al., 2002). For the GC-MS analysis of hexane extracts from in vitro assays using GPP as substrate, the parameters used were: 2 min at 50 °C, ramp to 140 °C with 20°C/min, ramp to 320 with 10°C/min to 320 and hold for 3 min. 1µL was injected in splitless mode at 250°C and the system was set in constant velocity mode with a linear velocity of 59.8 cm/sec using H2 as carrier gas. Solvent cutoff was set to 3.5 min.

Accession numbers

Nucleotide sequences of characterized enzymes have been submitted to the GenBankTM/EBI Data Bank with accession numbers: CfTPS1, KF444506; CfTPS2, KF444507; CfTPS3, KF444508; CfTPS4, KF444509; CfTPS15, KF471011.

Acknowledgements

We thank Dr. Damian Drew for critical reading of the manuscript and Dr. Carl Erik Olsen for excellent technical assistance with the LC-MS metabolomics. We also thank the greenhouse service personnel of the University of Copenhagen, Faculty of Science, Department of Plant and Environmental Sciences. We are grateful to Jason Goodger (University of Melbourne, School of Botany, Australia) for providing a clone of SfCIN. Imaging data were collected at the Center for Advanced Bioimaging (CAB) Denmark, University of Copenhagen.

Abbreviations

LPP, labd-13-en-8,15-diol diphosphate; LC-MS, liquid chromatography-mass spectrometry; GPP, geranyl diphosphate
LITERATURE CITED


Chen F, Tholl D, Bohlmann J, Pichersky E (2011) The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. Plant J 66: 212-229


Murphy D (2012) The dynamic roles of intracellular lipid droplets: from archaea to mammals. Protoplasma 249: 541-585


Silvestro D, Andersen TG, Schaller H, Jensen PE (2013) Plant Sterol Metabolism. Δ^7-Sterol-C^5-Desaturase (STE1/DWARF7), Δ^5^2-Sterol-Δ^5^-Reductase (DWARF5) and Δ^24-Sterol-Δ^24^-Reductase (DIMINUTO/DWARF1) Show Multiple Subcellular Localizations in Arabidopsis thaliana (Heynh) L. PLoS ONE 8: e56429


Valdés LJ, Ill, Mislankar SG, Paul AG (1987) Coleus barbatus (C. forskohlii) (Lamiaceae) and the Potential New Drug Forskolin (Coleonol). Econ Botany **41**: 474-483


Xie Z, Kapteyn J, Gang DR (2008) A systems biology investigation of the MEP/terpenoid and shikimate/phenylpropanoid pathways points to multiple levels of metabolic control in sweet basil glandular trichomes. Plant J **54**: 349-361


**Figure Legends**

**Figure 1.** Localization of oil bodies within the root cork of *C. forskohlii*. (A) Cross section of entire root with thick fissured cork. Lower right inset: the location of cork cells. (B) Rows of cork cells each with one prominent oil body. (C-F) Confocal imaging of Nile Red labeled oil bodies. (C) Transmitted light image of a cork cell with two oil bodies. Fluorescence images of the same oil bodies showing discrimination between neutral lipids (green fluorescence, D) and polar lipids (magenta fluorescence, E). (F) Overlay of the two fluorescence images. Scale bar represents 200 µm (A) and 10 µm (B-F).

**Figure 2.** Forskolin content (mg g⁻¹ DW) as determined by HPLC-ELSD analysis from different tissues of *C. forskohlii*. (Ck), root cork; (CS), root cortex and stele; (Fl), flowers; (St) stems and (Lv) leaves. Data are the mean ± SE of three independent biological replicates.

**Figure 3.** Detection of forskolin (a) and (13R) manoyl oxide (b) in *C. forskohlii* root cork and isolated oil bodies. (A) LC-MS extracted ion chromatogram (EIC) of *m/z* 433.2197 [forskolin+Na]⁺ in isolated oil bodies and in root cork tissue from *C. forskohlii*. (B) GC-MS EIC of *m/z* 275 corresponding to manoyl oxide in isolated oil bodies and in root cork tissue from *C. forskohlii*. (C) Bright field microscope image of isolated *C. forskohlii* oil bodies. Scale bar represents 5 µm. (D) Molecular structure of (13R) manoyl oxide. (E) Mass spectrum obtained from manoyl oxide identified in root cork tissue (top) and reference spectrum (bottom) from Wiley mass spectrum database.

**Figure 4.** GC-MS analysis of hexane extracts from *C. forskohlii* tissues. (b), (13R) manoyl oxide; (h), dehydroabietadiene; (f), miltiradiene; (i), abietadiene. IS, internal standard (1 ppm 1-eicosene).

**Figure 5.** Phylogenetic classification of *C. forskohlii* diterpene synthases with known class II (A) and class I (B) sequences. The phylograms are rooted with the bifunctional ent-copalyl diphosphate synthase/ent-kaurene synthase from the moss *Physcomitrella patens*. Asterisks indicate nodes supported by >80% bootstrap confidence and the scale bar indicates 0.1 amino acid changes. The numbers indicated at each enzyme refer to their respective enzymatic products, the structures of which are given on the right. Species and GenBank accession are given in Supplemental Table S2.

**Figure 6.** Relative expression of *CfTPS* genes in *C. forskohlii* tissues. (Ck), root cork; (CS), root cortex and stele; (Fl), flowers; (St), stems and (Lv), leaves. Transcript abundance of *CfTPS* genes expressed in...
arbitrary units was measured by qPCR using the translation initiation factor (TIF4a) for normalization. Each value represents the average of three biological replicates, each of which was performed in at least three technical replicates.

**Figure 7.** GC-MS analysis of *in vitro* assays with *C. forskohlii* diTPS. (A) *In vitro* assays with CfTPS1 alone and coupled assays with CfTPS1 and CfTPS3 and CfTPS4. Extracts of CfTPS1 assays were treated with calf intestinal alkaline phosphatase (CIP). (B) *In vitro* assays with CfTPS2 and coupled with CfTPS3 and CfTPS4. Extracts of CfTPS2 were treated with CIP. (b), (13R) manoyl oxide; (c), (13S) manoyl oxide; (d), labd-13-en-8,15-diol and (e), labden-8-ol; (f), miltiradiene and (g), copal-15-ol. IS, internal standard (1ppm 1-eicosene). (C) Mass spectra of compounds identified from assays. Structures tentatively identified as described in Materials and Methods.

**Figure 8.** GC-MS analysis of hexane extracts from *N. benthamiana* transiently expressing *C. forskohlii* diTPSs. (A) Extracted ion chromatogram (EIC) of *m/z* 275 and (B) EIC of *m/z* 272. (b), (13R) manoyl oxide; (c), (13S) manoyl oxide; (h), dehydroabietadiene; (f), miltiradiene and trace amount of (i), abietadiene.

**Figure 9.** Scheme of the biosynthetic routes from GGPP to specialized and general diterpenoids of the abietane, labdane and ent-kaurene class. Dashed arrows indicate reactions without experimental evidence in *C. forskohlii*. Detection of (+)-ferruginol in *C. forskohlii* was reported earlier (Hurley, 1998); 2CYP76AH1 from the close relative *Salvia miltiorrhiza* was proposed to convert miltiradiene to ferruginol (Guo et al., 2013).
Table 1. Abietane and labdane type diterpenoids detected by LC-MS analysis of oil bodies isolated from *C. forskohlii* root cork. Compounds were identified based on accurate mass measurements with the exception of forskolin which was also identified based on co-elution with an authentic standard.

<table>
<thead>
<tr>
<th>Ret. Time (min)</th>
<th>m/z</th>
<th>Adduct</th>
<th>Predicted Formula</th>
<th>Mass Error (ppm)</th>
<th>Compound</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.69</td>
<td>331.1901</td>
<td>[M+H]^+</td>
<td>C_{20}H_{26}O_{4}</td>
<td>0.87</td>
<td>14-deoxycoleon U</td>
<td>Xu et al., 2005</td>
</tr>
<tr>
<td>9.64</td>
<td>375.2159</td>
<td>[M+Na]^+</td>
<td>C_{20}H_{32}O_{5}</td>
<td>-4.84</td>
<td>deoxy-deacetyl-forskolin</td>
<td>Asada et al., 2012</td>
</tr>
<tr>
<td>9.67</td>
<td>433.2194</td>
<td>[M+Na]^+</td>
<td>C_{22}H_{34}O_{7}</td>
<td>0.67</td>
<td>forskolin</td>
<td>Bhat et al., 1977</td>
</tr>
<tr>
<td>10.88</td>
<td>401.2298</td>
<td>[M+Na]^+</td>
<td>C_{22}H_{34}O_{5}</td>
<td>0.12</td>
<td>dideoxy-forskolin</td>
<td>Asada et al., 2012</td>
</tr>
<tr>
<td>14.20</td>
<td>317.2102</td>
<td>[M+H]^+</td>
<td>C_{20}H_{28}O_{3}</td>
<td>2.91</td>
<td>20-deoxocarnosol</td>
<td>Kelecom, 1984</td>
</tr>
<tr>
<td>16.11</td>
<td>417.2252</td>
<td>[M+Na]^+</td>
<td>C_{22}H_{34}O_{5}</td>
<td>-1.12</td>
<td>deoxy-forskolin</td>
<td>Asada et al., 2012</td>
</tr>
<tr>
<td>18.29</td>
<td>385.2341</td>
<td>[M+Na]^+</td>
<td>C_{22}H_{34}O_{4}</td>
<td>2.16</td>
<td>trideoxy-forskolin</td>
<td>Asada et al., 2012</td>
</tr>
<tr>
<td>22.95</td>
<td>323.1967</td>
<td>[M+Na]^+</td>
<td>C_{20}H_{28}O_{2}</td>
<td>4.83</td>
<td>barbatusol</td>
<td>Kelecom, 1983</td>
</tr>
</tbody>
</table>
Supplemental Material

The following materials are available in the online version of this article.

Supplemental Figure 1. Bright field microscopy of C. forskohlii stem cross section. Arrows indicate oil bodies. Scale bar represents 10 μm.

Supplemental Figure 2. Alignment of protein sequences of the CfTPSs together with PpCPS/EKS. The alignment was performed using ClustalW2 and visualized by BOXSHADE 3.21. Identical residues are highlighted in black while similar residues are highlighted in gray. The aspartate-rich motifs DxDD and DDxxD are underlined with a red solid line. Residues marked in red indicate the cleavage site for the removal of the plastidial transit peptide that was performed for the E. coli expression experiments.

Supplemental Figure 3. GC-MS analysis of hexane extract from in vitro assays. No monoterpene compounds were identified with GC-MS analysis when CfTPS3 or CfTPS4 enzymes were tested in vitro using GPP as substrate. In vitro assays with CfTPS2 coupled to CfTPS3 and CfTPS4 respectively were used as positive controls for CfTPS3 and CfTPS4 enzymes activity. As positive controls of the assays the characterized monoterpene synthase SfCIN from Salvia fruticosa (Kampranis et al., 2007) and geranyl diphosphate (GPP) dephosphorylated to geraniol was used. All compounds were confirmed by reference spectra in Wiley 8, with an identity >90%.

Supplemental Table S1. List of primers used in this study.

Supplemental Table S2. Annotation and GenBank accession numbers of the proteins used in the phylograms.

Supplemental Table S3. Alignment of the nucleotide sequences of the native CfTPS3 (natCfTPS3) and the synthetic variant (synCfTPS3) used for heterologous expression in E. coli. The alignment has been performed using the software ClustalW2 (http://www.ebi.ac.uk/).
Figure 1. Localization of oil bodies within the root cork of *C. forskohlii*. (A) Cross section of entire root with thick fissured cork. Lower right inset: the location of cork cells. (B) Rows of cork cells each with one prominent oil body. (C-F) Confocal imaging of Nile Red labeled oil bodies. (C) Transmitted light image of a cork cell with two oil bodies. Fluorescence images of the same oil bodies showing discrimination between neutral lipids (green fluorescence, D) and polar lipids (magenta fluorescence, E). (F) Overlay of the two fluorescence images. Scale bar represents 200 μm (A) and 10 μm (B-F).
Figure 2. Forskolin content (mg g⁻¹ DW) as determined by HPLC-ELSD analysis from different tissues of *C. forskohlii*. (Ck), root cork; (CS), root cortex and stele; (Fl), flowers; (St) stems and (Lv) leaves. Data are the mean ± SE of three independent biological replicates.
Figure 3. Detection of forskolin (a) and (13R) manoyl oxide (b) in *C. forskohlii* root cork and isolated oil bodies. (A) LC-MS extracted ion chromatogram (EIC) of m/z 433.2197 [forskolin+Na]+. in isolated oil bodies and in root cork tissue from *C. forskohlii*. (B) GC-MS EIC of m/z 275 corresponding to manoyl oxide in isolated oil bodies and in root cork tissue from *C. forskohlii*. (C) Bright field microscope image of isolated *C. forskohlii* oil bodies. Scale bar represents 5 μm. (D) Molecular structure of (13R) manoyl oxide. (E) Mass spectrum obtained from manoyl oxide identified in root cork tissue (top) and reference spectrum (bottom) from Wiley mass spectrum database.
Figure 4. GC-MS analysis of hexane extracts from *C. forskohlii* tissues. (b), (13R) manoyl oxide; (h), dehydroabietadiene; (f), miltiradiene; (i), abietadiene. IS, internal standard (1 ppm 1-eicosene).
Figure 5. Phylogenetic classification of *C. forskohlii* diterpene synthases with known class II (A) and class I (B) sequences. The phylograms are rooted with the bifunctional ent-copalyl diphosphate synthase/ent-kaurene synthase from the moss *Physcomitrella patens*. Asterisks indicate nodes supported by >80% bootstrap confidence and the scale bar indicates 0.1 amino acid changes. The numbers indicated at each enzyme refer to their enzymatic products, the structures of which are given on the right. Species and GenBank accession are given in Supplemental Table S2.
Figure 6. Relative expression of CfTPS genes in *C. forskohlii* tissues. (Ck), root cork; (CS), root cortex and stele; (Fl), flowers; (St), stems and (Lv), leaves. Transcript abundance of *CfTPS* genes expressed in arbitrary units was measured by qPCR using the translation initiation factor (TIF4a) for normalization. Each value represents the average of three biological replicates, each of which was performed in at least three technical replicates.
Figure 7. GC-MS analysis of in vitro assays with C. forskohlii diTPS. (A) In vitro assays with CfTPS1 and coupled with CfTPS3 and CfTPS4. Extracts of CfTPS1 were treated with CIP. (B) In vitro assays with CfTPS2 alone and coupled assays with CfTPS2 and CfTPS3 and CfTPS4. Extracts of CfTPS2 assays were treated with calf intestinal alkaline phosphatase (CIP). (b), (13R) manoyl oxide; (c), (13S) manoyl oxide; (d), labd-13-en-8,15-diol (e), labdien-8-ol; (f), miltiradiene and (g), copal-15-ol. IS, internal standard (1ppm 1-eicosene). (C) Mass spectra of compounds identified from assays. Structures tentatively identified as described in Materials and Methods.
Figure 8. GC-MS analysis of hexane extracts from N. benthamiana transiently expressing C. forskohlii diTPSs. (A) Extracted ion chromatogram (EIC) of m/z 275 and (B) EIC of m/z 272. (b), (13R) manoyl oxide; (c), (13S) manoyl oxide; (h), dehydroabietadiene; (f), miltiradiene and trace amount of (i), abietadiene.
Figure 9. Scheme of the biosynthetic routes from GGPP to specialized and general diterpenoids of the abietane, labdane and ent-kaurene class. Dashed arrows indicate reactions without experimental evidence in *C. forskohlii*. 1: Detection of (+)-ferruginol in *C. forskohlii* was reported earlier (Kelecom, 1983); 2: CYP76AH1 from the close relative *Salvia miltiorrhiza* was shown to convert miltiradiene to ferruginol (Guo et al., 2013).