Biosynthesis of cannflavins A and B from Cannabis sativa L

Kevin A Rea, José A. Casaretto, M. Sameer Al-Abdul-Wahid, Arjun Sukumaran, Jennifer Geddes-McAlister, Steven J. Rothstein, Tariq A. Akhtar

ABSTRACT

In addition to the psychoactive constituents that are typically associated with Cannabis sativa L., there exist numerous other specialized metabolites in this plant that are believed to contribute to its medicinal versatility. This study focused on two such compounds, known as cannflavin A and cannflavin B. These prenylated flavonoids specifically accumulate in C. sativa and are known to exhibit potent anti-inflammatory activity in various animal cell models. However, almost nothing is known about their biosynthesis. Using a combination of genomics and biochemical approaches, an aromatic prenyltransferase from Cannabis sativa (CsPT3) was identified that catalyzes the regiospecific addition of either geranyl diphosphate (GPP) or dimethylallyl diphosphate (DMAPP) to the methylated flavone, chrysoeriol, to produce cannflavins A and B, respectively. Further evidence is presented for an O-methyltransferase (CsOMT21) encoded within the C. sativa genome that specifically converts the widespread plant flavone known as luteolin to chrysoeriol, both of which accumulate in C. sativa. These results therefore imply the following reaction sequence for cannflavins A and B biosynthesis: luteolin → chrysoeriol → cannflavin A and cannflavin B. Taken together, the identification of these two unique enzymes represent a branch point from the general flavonoid pathway in C. sativa and offer a tractable route towards metabolic engineering strategies that are designed to produce these two medicinally relevant Cannabis compounds.

1. Introduction

Since antiquity, Cannabis sativa L. (Cannabaceae) has been cultivated for a variety of industrial, medicinal and recreational uses. In recent years, a worldwide socio-political movement aimed towards the legalization of C. sativa has spurred a resurgence of interest in this versatile plant and have afforded researchers the opportunity to explore its metabolic diversity (McPartland and Russo, 2001; van Bakel et al., 2011; Sawler et al., 2015; Booth et al., 2017; Russo and Marcu, 2017; Sexton et al., 2018). Apart from the psychoactive constituents that are typically associated with Cannabis sativa L., there exist a plethora of specialized metabolites in this plant species that are believed to contribute to its medicinal properties (Elsoby and Slade, 2005; Radwan et al., 2008a,b; Flores-Sanchez and Verpoorte, 2008; McPartland and Russo, 2014; Booth et al., 2017). One such class of compounds are the prenylated flavonoids, known as cannflavins A and B (Barrett et al., 1985).

The interest surrounding cannflavins A and B within the Cannabis community stems from three seminal studies: First, Fairbairn and Pickens (1981) demonstrated that THC and CBD-free extracts from C. sativa could reduce the cataleptic effects of THC in mice and that this effect could be reversed by prostaglandin E2 (PGE2) administration. Barrett and colleagues reportedly identified the causal agent in these extracts as cannflavins A and B and verified that these prenylated flavonoids could inhibit the production of PGE2 in human rheumatoid synovial cells and provide anti-inflammatory benefits that were approximately thirty times more effective than aspirin (Barrett et al., 1985, 1986). It was later demonstrated that the underlying basis for their potent anti-inflammatory properties was that cannflavins A and B act to inhibit the in vivo production of two pro-inflammatory mediators, prostaglandin E2 and the leukotrienes (Werz et al., 2014). Surprisingly, however, since these striking reports of two non-psychoactive constituents from C. sativa that have medicinal potential, little attention has been focused on how these unique flavonoids are actually synthesized within C. sativa (Andre et al., 2016; Pollastro et al., 2018).

Cannflavins A and B belong to the class of plant flavonoids known as flavones, which occur in several plant lineages (Winkel-Shirley, 2001; Ross and Kasum, 2002; Andersen and Markham, 2005; Jiang et al., 2016). Flavones perform a myriad of in planta functions that range from...
regulators of auxin transport to mediators in plant-pathogen interactions (Johnson et al., 2007; Zhang et al., 2009; Falcone Ferreyra et al., 2012). In addition, the dietary consumption of various plant flavonoids is well established to offer neuroprotective, antioxidant, and anticancer properties in several animal models (Pietta, 2000; Nahavi et al., 2015; Madunic et al., 2018). While the flavone biosynthetic pathway has been extensively studied in several plants, almost nothing is known about this process in cannabis (Flores-Sanchez and Verpoorte, 2008; Andre et al., 2016). It is expected that the core flavone pathway is similarly embedded within C. sativa, given that this plant also accumulates three widespread flavonoids (apigenin, luteolin and chrysoeriol) as well as their glycosylated derivatives (Turner et al., 1980; Radwan et al., 2008a,b).

Cannflavins A and B appear to be Cannabis specific (Vanhoenacker et al., 2002; Ross et al., 2005; Werz et al., 2014) and their unique bioactivity appear to be linked to two key modifications of their parent flavone backbone: First is their distinct prenylation pattern in which a prenyl side-chain, in the form of a geranyl (C10) or a dimethylallyl (C5) group, are affixed to the 6 position of the flavone A-ring, respectively (Barrett et al., 1985; Choi et al., 2004). These prenyl moieties impart lipophilicity to the parent flavone, which are believed to enhance uptake and bioaccumulation into cells and promote their interaction with membrane-bound enzymes and receptors that are involved in numerous cell-signalling pathways (Milligan et al., 1999; Botta et al., 2005; Wajten et al., 2007; Werz et al., 2014; Vychyňová et al., 2017). Second, both cannflavins A and B are modified at the 3’ position of the flavone B-ring with a methoxy group, which also increases lipophilicity and may therefore enhance their cellular retention and access to various cellular targets (Ibrahim, 2005; Walle, 2007; Berin and Gang, 2016). In which order these two unique modifications (prenylation and methoxylation) of the parent flavone occur en route to cannflavin A and cannflavin B biosynthesis, however, is not known.

We therefore reasoned that the biosynthesis of cannflavins A and B occurs via a specific branch point from the conserved C. sativa flavone biosynthetic pathway in which a central flavone must be methoxylated at the 3’ position of the flavone B-ring and prenylated at the 6 position of the A-ring. Using a combination of phylogenomic and biochemical approaches, we report the identification and characterization of two enzymes that catalyze these penultimate and final steps of cannflavin A and cannflavin B synthesis in C. sativa.

2. Results and discussion

2.1. Phylogenetic analysis of C. sativa prenyltransferases

To synthesize cannflavins A and B, a prenyl moiety must be added to position 6 of a flavone that typically accumulates in C. sativa. Therefore, we first searched for gene sequences that were putatively annotated as flavonoid or related aromatic prenyltransferases in the Transcriptome Shotgun Assembly (TSA) database for C. sativa, which is accessible through NCBI. A previously described flavone prenyltransferase from Glycyrrhiza uralensis (GuA6DT; GenBank AIT11912.1) was used as a query in these searches (Li et al., 2014). GuA6DT prenylates apigenin which is a widespread plant flavone that also accumulates in C. sativa (McPartland and Russo, 2001). This search uncovered eight full-length cDNA sequences from C. sativa that exhibited 22–53% identity at the amino acid level to GuA6DT and were putatively annotated as C. sativa prenyltransferases (CsPT1–8; Fig. S1). One of the prenyltransferases that were identified in this search (CsPT1) matched a previously reported enzyme from C. sativa that is known to be involved in the prenylation of olivetolic acid to cannabigerolic acid in the cannabinoid biosynthesis pathway (Page and Boubakir, 2014). We next performed a phylogenetic analysis that included CsPT1 and these seven other prenyltransferases from C. sativa along with all known plant prenyltransferases that have been previously shown to accommodate aromatic substrates (Fig. 1; Fig. S1). This analysis demonstrated that plant aromatic prenyltransferases fall into six distinct groups, which are conveniently defined by the specific branch of aromatic metabolism in which they participate. The eight CsPTs occupy three of these six groups: CsPT2 and CsPT6 reside in a unique clade of prenyltransferases (Group 2) which have been shown to participate in the tocopherol biosynthetic pathway (Collakova and DellaPenna, 2001; Savidge et al., 2002; Tian et al., 2007). CsPT5 appears to be orthologous to homogenisate solanesyltransferases (Group V) that function in plastoquinone biosynthesis (Venkatesh et al., 2006; Tian et al., 2007). The five remaining CsPTs (CsPT1, 3, 4, 7, and 8) formed a third and distantly related group (Group VI) that includes two prenyltransferases from Humulus lupulus (hops), which are involved in the aromatic prenylation reactions required for terpenophenolic biosynthesis (Nagel et al., 2008; Tsurumaru et al., 2012; Li et al., 2015). Surprisingly, this analysis revealed that none of the CsPTs were closely related to any of the flavonoid or coumarin prenyltransferases (Groups I and IV, respectively) that have been previously identified in various plant species (Sasaki et al., 2008, 2011; Akashi et al., 2009; Shen et al., 2012; Wang et al., 2014; Munakata et al., 2016; Yoneyama et al., 2016; Yang et al., 2018). Interestingly, in silico analysis of each CsPT predicted that they are all targeted to plastids (Table S2). We return to this point below.

2.2. Biochemical characterization of recombinant CsPTs identifies a regiospecific chrysoeriol-6-prenyltransferase

In pursuit of identifying a flavone prenyltransferase(s) that is involved in cannflavin A and B biosynthesis, we focused on the sixth group of plant PTs that was identified from our phylogenetic analysis, which included CsPT1, 3, 4, 7, and 8. These particular CsPTs drew our attention for two reasons: First, they represent the only prenyltransferases in our search of the C. sativa genome that are evolutionarily distinct from those CsPTs that appear to function in tocopherol and plastoquinone biosynthesis, which are central plant compounds (Sattier et al., 2004; Kriese et al., 2004). Second, this group of CsPTs are closely related to two enzymes from hops (a Cannabaceae family relative), which are believed to prenylate naringenin chalcone, a widespread intermediate in the general flavonoid pathway (Nagel et al., 2008; Li et al., 2015).

We therefore introduced each of these CsPTs into a well-established yeast expression system that is typically used to characterize this class of enzymes (Sasaki et al., 2011; Shen et al., 2012; Li et al., 2014; Wang et al., 2014). Principle component analysis of whole cell proteomes from the strains expressing each of the five CsPTs revealed a clear clustering that was separate from the strain harboring the vector alone negative control (Figure S2). Moreover, a Pearson correlation analysis followed by Hierarchical clustering by Euclidean distance confirmed distinct clustering between the strains expressing the CsPTs compared to the strain harboring the empty vector negative control, which together indirectly indicate that each CsPT is active and expressed in this system. As a positive control for prenyltransferase enzyme activity, we also introduced the open-reading frame of GuA6DT from G. uralensis into this host yeast strain, which was previously shown to prenylate not only apigenin, but a variety of flavones at position 6 of the A ring using dimethylallyl diposophate (DMAPP) as a substrate (Li et al., 2014). As expected, in assays with microsomes expressing GuA6DT together with apigenin and DMAPP as a prenyl donor, we observed a single reaction product whose mass-to-charge ratio (m/z 339) was consistent with 6-dimethylallyl apigenin (Fig. S2). As previously demonstrated, luteolin and chrysoeriol were also converted by GuA6DT to their corresponding 6-dimethylallyl flavones using DMAPP as a substrate (Li et al., 2014), and these monoprenylated flavones were subsequently purified by HPLC to serve as prenyflavone standards in subsequent exploratory assays with recombinant CsPTs (Fig. S3). Accordingly, the microsomal fractions from each yeast strain expressing the open-reading frames of CsPT1, 3, 4, 7, and 8 were recovered and tested for prenyltransferase activity with three flavone substrates that are present in C. sativa: apigenin, chrysoeriol, and luteolin (McPartland and Russo, 2001;
As potential prenyl donors, DMAPP, IPP and GPP were also included as co-substrates in each of these enzyme assays. The reaction products from these assays were extracted with ethyl acetate, analyzed by reverse-phase HPLC, and compared to the authentic prenylated flavones that were isolated from in vitro assays with GuA6DT (see above). No detectable prenylated flavone products were observed in assays with microsomes obtained from yeast cells harbouring the empty vector or with CsPT1, 4, or 7 under the above conditions. However, this analysis revealed that microsomes containing CsPT3 readily converted apigenin and chrysoeriol, but not luteolin, to their corresponding prenylated flavones using DMAPP as a prenyl donor, and kinetic analysis demonstrated that chrysoeriol was the preferred flavone substrate (Table 1). Assays with microsomes containing CsPT8 revealed that this enzyme also prenylated apigenin using DMAPP as a substrate, but not with any of the other flavones (Fig. S4). IPP was not accommodated as a prenyl donor in assays with microsomes containing either CsPT3 or CsPT8. The prenylated product from assays with CsPT8 did not match that of 6-dimethylallyl apigenin, based on HPLC retention time (Fig. S4). On the other hand, the enzymatic product that was observed in assays with CsPT3, chrysoeriol and DMAPP exhibited the same HPLC retention time and fragmentation pattern, as determined by tandem mass spectrometry, to that of 6-dimethylallyl chrysoeriol, also known as cannflavin B (Fig. 2A). We therefore chose to focus solely on CsPT3. Interestingly, CsPT3 also accommodated GPP as a prenyl donor in assays with apigenin and chrysoeriol, and kinetic analysis again revealed that chrysoeriol was the preferred substrate (Table 1). That CsPT3 accommodated both DMAPP and GPP fits well with its predicted subcellular location, as the enzyme would have access to these two end products of the plastidial methylerythritol phosphate pathway. The reaction product that was observed with assays including CsPT3, chrysoeriol and GPP exhibited less polarity than 6-dimethylallyl chrysoeriol/cannflavin B (as indicated by HPLC-UV analysis) and exhibited a mass-to-charge ratio (m/z 437) that was consistent with the addition of a geranyl moiety onto the chrysoeriol backbone, and was therefore assumed to be cannflavin A (Fig. 2B).

The structures of the two enzymatic products that were obtained...
from assays with CsPT3, chrysoeriol, and either DMAPP or GPP as prenyl donors, were further analyzed by 1H and 13C NMR. The 1H NMR spectra of cannflavin A exhibited three peaks of area 3H in the region between 1.5 and 2.0 ppm and 2 peaks of area 1H in the region between 5.0 and 5.5 ppm, consistent with three methyl groups and two vinylic protons, respectively, suggesting the presence of a geranyl group. The 1H NMR spectra of cannflavin B exhibited only two peaks of area 3H in the region between 1.5 and 2.0 ppm and a single peak of area 1H in the region between 5.0 and 5.5 ppm, consistent with the two methyl groups and one vinylic proton of a single prenyl group. COSY, TOCSY, and HMBC spectra also demonstrated peak patterns that were consistent with a geranyl group for cannflavin A and a prenyl group for cannflavin B (Crombie and Crombie, 1982; Barrett et al., 1986; Choi et al., 2004).

Initial NMR assignments for both cannflavin A and B were performed using the COSY, TOCSY, and HSQC spectra to assign all proton resonances as well as those of proton-bearing carbons and an HMBC experiment was then used to complete the carbon assignments (Fig. 3). In both cannflavin A and cannflavin B, HMBC correlations were observed from both the 5-OH hydroxyl proton and the 1” proton on the geranyl or prenyl group to the same carbon resonance.

Table 2

| NMR assignments for cannflavin A and cannflavin B. |
|-------------------------------|-----------------|-----------------|-----------------|-----------------|
| Position      | Cannflavin A     | Cannflavin B     |     |     |
|               | δH   | δC   | δH   | δC   |
| 2             | –     | 164.4| –     | 164.7|
| 3             | 6.68  | 104.2| 6.69  | 104.5|
| 4             | –     | 182.8| –     | 183.2|
| 5             | –     | 159.8| –     | 160.2|
| 6             | –     | 112.0| –     | 112.3|
| 7             | –     | 162.1| –     | 162.3|
| 8             | 6.62  | 93.9 | 6.62  | 94.1 |
| 9             | –     | 156.3| –     | 156.6|
| 10            | –     | 104.9| –     | 105.3|
| 1’            | –     | 123.4| –     | 123.8|
| 2’            | 7.76  | 110.0| 7.61  | 110.5|
| 3’            | –     | 148.5| –     | 148.8|
| 4’            | –     | 151.0| –     | 151.3|
| 5’            | 7.00  | 115.8| 7.00  | 116.3|
| 6’            | 7.57  | 121.0| 7.59  | 121.3|
| 1’’           | 3.36  | 21.8 | 3.36  | 22.0 |
| 2’’           | 5.29  | 122.9| 5.28  | 121.2|
| 3’’           | –     | 135.0| –     | 131.7|
| 4’’           | 1.96  | 40.1 | 1.78  | 17.9 |
| 5’’           | 2.05  | 27.6 | 1.65  | 25.9 |
| 6’’           | 5.07  | 124.8| –     | –    |
| 7’’           | –     | 131.3| –     | –    |
| 8’’           | 1.59  | 25.3 | –     | –    |
| 9’’           | 1.54  | 17.3 | –     | –    |
| 10’’          | 1.79  | 16.0 | –     | –    |
| O-CH3         | 3.98  | 56.2 | 3.99  | 56.6|
| 5-OH         | 13.30 | –    | 13.30 | –    |
2.3. Phylogenetic analysis of C. sativa O-methyltransferases involved in the methylation of luteolin to chrysoeriol

The observation that CsPT3 preferentially prenylates chrysoeriol, in vitro, and that prenylated luteolin is apparently absent in extracts from C. sativa implies, a priori, that methylation of luteolin to chrysoeriol must occur first in the cannflavin A and/or B pathway. We reasoned that the alleged enzyme named methylates luteolin at the 3'-hydroxyl position of the flavone B-ring to yield chrysoeriol would likely fall into the class of S-adenosyl-L-methionine (AdoMet)-dependent O-methyltransferases (OMTs), which are widely distributed throughout the plant kingdom (Ibrahim et al., 1998; Ibrahim, 2005; Kim et al., 2010). We focused our initial searches of the TSA database for C. sativa on type 1 OMTs, which specifically methylate hydroxylation moieties of phenylpro-panoid-based compounds (Noel et al., 2003). Using a previously char-acterized flavonoid-O-methyltransferase from Oryza sativa (OsOMT9) that methylates the 3'-hydroxyl group on a variety of flavonoids as a query (Kim et al., 2006), this search uncovered 40 unique nucleotide sequences corresponding to partial or full-length transcripts that were loosely annotated as 'cafeic acid-O-methyltransferases'. We next compared these transcript sequences via BLASTn searches against the Cannabis whole genome contig database to confirm their corresponding full-length open reading frames. This analysis revealed twenty-four unique protein sequences (Fig. S5) which were subsequently annotated as C. sativa O-methyltransferases (CsOMT-24).

A phylogenetic analysis of the CsOMT family was then performed to establish their evolutionary relatedness to various plant OMTs that have been previously identified to act on aromatic substrates. This analysis revealed that type 1 CsOMTs are distributed into four general groups (Fig. 4). It should be noted however, as Schroder et al. (2002) and Lam et al. (2007) previously pointed out, that assigning substrate preference based on sequence similarity alone for this class of plant enzymes is precarious. For example, the first group of type 1 plant OMTs depicted in our phylogenetic analysis includes enzymes that utilize a broad array of aromatic substrates - from simple phenolic compounds, such as chavicol, guaiacol and orcinol (Gang et al., 2002; Scalliet et al., 2006; Akhtar et al., 2013), to more complex heterocyclic aromatics, such as homoeriodictyol, myricetin, and resveratrol (Schroder et al., 2004; Schmidt et al., 2008; Schmidt et al., 2011). We found nine CsOMT family members present within this group. The second group of type 1 OMTs appear specific to the Cannabaceae family and include seven CsOMTs along with two OMTs from Humulus lupulus that are involved in the synthesis xanthohumol (Nagel et al., 2008). The third and fourth groups represent two closely related sister clades of type 1 plant OMTs and contain the remaining members of the CsOMT family. Strikingly, all plant OMTs that are known to methylate the 3'-hydroxyl position of various flavonoids are confined to group three and include re-presentatives from Arabidopsis, peppermint, rice, wheat, and American golden saxifrage (Gauthier et al., 1996; Muzac et al., 2000; Willits et al., 2004; Kim et al., 2006; Zhou et al., 2006). We found three CsOMTs (CsOMT6, 12 and 21) that fell into this group.

2.4. Identification and biochemical characterization of a luteolin O-methyltransferase

While recognizing that phylogenetic-driven predictions of plant OMT function has its caveats, we were nevertheless intrigued by the presence of CsOMT6, 12 and 21 amongst a group of evolutionary conserved OMTs that exhibit regioselective methylation activity for the 3'-position on a variety of flavonoids. We therefore chose to survey the enzymatic activities of these three CsOMTs to elucidate if their encoded proteins could methylate luteolin at the 3'-hydroxyl position of the flavone B-ring to yield chrysoeriol. In accordance with our hypothesis, this reaction would represent the penultimate step in cannflavin A and B biosynthesis.

We first introduced the open-reading frames of CsOMT6, 12 and 21 into E. coli cells as N-terminal fusion proteins with a His6 tag and subsequently assayed desalted protein extracts from cells expressing each protein for O-methyltransferase enzyme activity. We chose to test three flavonones (apigenin, luteolin, and chrysoeriol) and two flavonols (quercetin and kaempferol) that typically accumulate in C. sativa as potential substrates for each CsOMT, along with the universal methyl donor \textsuperscript{14}C-labeled S-adenosyl methionine as a co-substrate (Fig. 5A). After extracting and quantifying the amount of radiolabel within the enzymatic products that were obtained from these initial assays, it was determined that only CsOMT6 and CsOMT21 exhibited appreciable OMT activity with the flavonoid substrates that were provided. Recombinant CsOMT6 exhibited strict substrate specificity towards quercetin (Fig. S6), while CsOMT21 methylated luteolin primarily (Fig. 5B), yet also accommodated quercetin as a substrate, albeit with less efficiency (57% activity compared to luteolin). Notably, neither apigenin, nor chrysoeriol, nor kaempferol, which lack a free 3'-hydroxyl group on their B-ring, were used as substrates by either enzyme. This observation therefore lends further support to the view (Lam et al., 2007) that CsOMT6, CsOMT21 and the other type 1 OMTs that are present in group three (defined by our phylogenetic analysis) encompass an evolutionarily conserved group of enzymes with regiospecificity for 3'-hydroxy groups on a variety of flavonoid compounds. It also implies that CsOMT21 is a 3'-O-methyltransferase and likely catalyzes the penultimate step in cannflavin A and B biosynthesis by converting luteolin to chrysoeriol. To test this possibility, recombinant CsOMT21 was purified via Ni\textsuperscript{2+} affinity chromatography (Fig. 5B) and assayed with luteolin and unlabelled S-adenosyl methionine as co-substrates. Indeed, the identity of the reaction product was confirmed to be chrysoeriol, based on its identical HPLC retention time and mass spectral fragmentation pattern with the authentic standard (Fig. 5C). Recombinant CsOMT21 exhibited Michaelian kinetics with kinetic constants that were similar to those that have been previously reported for this class of enzymes (Fig. 5D). While our analysis cannot exclude the possibility that the other twenty-one CsOMT family member(s) may convert luteolin to chrysoeriol, these results do provide evidence for a branch point from the general flavonoid pathway that is present in C. sativa, whereby luteolin can be first methylated to chrysoeriol en route towards cannflavin A and cannflavin B synthesis.

3. Concluding remarks

Guided by previously published metabolomic data from C. sativa, a targeted phylogenomics approach combined with in-vitro biochemical assays was employed in this study to explore the biosynthetic pathway towards cannflavin A and B (Fig. 6). We provide evidence that a unique branch point from the general plant flavonoid pathway has evolved in C. sativa in which the widespread plant flavone, luteolin, is converted into cannflavin A and B via regiospecific methylation and prenylation reactions. The identification of these two enzymatic steps opens new opportunities for the metabolic engineering of cannflavin A and cannflavin B biosynthesis and underscores the value of phylogenomics driven gene discovery, an approach that is largely underutilized in the Cannabis research space.

4. Experimental

4.1. Chemicals and reagents

Authentic flavonoid standards for apigenin and luteolin were purchased from Indofine Chemical Company, kaempferol and quercetin were purchased from Sigma-Aldrich, and chrysoeriol was from Toronto Research Chemicals. The trans-prenyl diphosphates, isopentenyl diphosphate, dimethylallyl diphosphate, and geranyl diphosphate, were obtained from Echelon Biosciences. Radiolabeled S-[Methyl-\textsuperscript{14}C] adenosyl-L-methionine (58.0 mCi mmol\textsuperscript{-1}) was from PerkinElmer. Synthetic drop-out media lacking histidine for cultivating yeast was
obtained from US Biological. All primers were synthesized by Sigma-Aldrich and are listed in Table S1. All other chemicals were obtained from Sigma-Aldrich, BioBasic, or Fisher Scientific.

4.2. Phylogenetic analysis

The DNA sequences encoding for putative type 1 O-methyltransferases (CsOMTs) and aromatic prenyltransferases (CsPTs) were first identified in the Cannabis sativa genome available at NCBI (https://blast.ncbi.nlm.nih.gov/). Search was performed in the Transcriptome Shotgun Assembly (TSA) database available for the strain Purple Kush (Bioproject Accession PRJNA74271; van Bakel et al., 2011) via tBLASTn (https://blast.ncbi.nlm.nih.gov/Blast.cgi) for transcripts that exhibited homology to the Oryza sativa O-methyltransferase, OsOMT9, and the Glycyrrhiza uralensis prenyltransferase, Gu6ADT. These sequences were then matched against the Cannabis whole genome contig database (https://www.ncbi.nlm.nih.gov genomes/?term = cannabis + sativa; van Bakel et al., 2011), via BLASTn searches, to retrieve the full-length open reading frames for the C. sativa OMT and PT gene families (Supplemental Fig. S1 and S4). The in-silico assembled amino acid sequences from these gene sequences were then used to construct phylogenetic relationships using the MEGA software package (version 6.0) by the neighbor-joining method with bootstrap analysis of 1000 replicates.

4.3. Cloning and recombinant protein expression of CsOMTs in E. coli

The C. sativa L. O-methyltransferase open reading frames were synthesized by Genscript. These cDNAs for CsOMT6, 12 and 21 were amplified by PCR using the KOD Hot Start DNA polymerase (Novagen) and then ligated between the NdeI/AseI and HindIII sites of the pET28b vector system (Novagen) which introduces an N-terminal 6 x His tag to each coding sequence. These constructs were then introduced into E. coli BL21-CodonPlus (DE3)-RIPL cells. Bacterial cells expressing recombinant CsOMT6, 12 and 21 were cultured in LB media at 37 °C to an OD600 of 0.6. Isopropyl-β-D-thiogalactoside was then added to a final concentration of 1 mM and the cells were incubated at 16 °C for an additional 18 h. The bacterial cells were collected by centrifugation, resuspended in buffer A (20 mM Tris-HCl, pH 8.0, 500 mM KCl), and then disrupted by sonication. Crude protein extracts were centrifuged at 12,000×g for 10 min at 4 °C to remove unbroken cells and debris and then applied to a 1 mL HisTrap HP column (GE Healthcare) equilibrated in buffer A. Proteins bound to the Ni2+ affinity matrix were washed with five column volumes of buffer A containing 20 mM imidazole, eluted with one column volume of buffer A containing 400 mM imidazole, and then immediately desalted on PD-10 columns (GE Healthcare) equilibrated with 50 mM Tris-HCl, pH 7.5, 5 mM MgCl2, and 10% (v/v) glycerol. Protein concentration was determined by the method of Bradford (1976) using BSA as a standard.
4.4. Cloning and recombinant protein expression of CsPTs in yeast

The *C. sativa* and *G. uralensis* prenyltransferase open reading frames were synthesized by Genscript and included CsPT1, 3, 4, 7, 8 and GuA6DT (Li et al., 2014). These cDNAs were applied by PCR and ligated between the *Bam*HI and *Xho*I sites of pESC-HIS (Agilent). The sequence-verified constructs were introduced into the *Saccharomyces cerevisiae* YPH499 yeast strain (ura3–52 lys2–801 amber ade2–101 ochre trp1–Δ63 his3–Δ200 leu2–Δ1) using the method outlined by Gietz and Schiestl (2007) and transformants were selected on synthetic drop-out media lacking histidine, supplemented with 0.67% yeast nitrogen base and 2% glucose. For recombinant protein expression, the yeast transformants were cultured as above at 28 °C to an OD600 of 1.0. Yeast cells were then pelleted by centrifugation (5,000 x rpm, 10 min), washed twice with sterile water, and re-suspended in the same media as above containing 2% galactose instead of glucose. Cells were incubated for an additional 18 h at 28 °C to induce protein expression.

4.5. O-methyltransferase enzyme assays

Assays for determining O-methyltransferase enzyme activity were performed using ~2 μg of purified recombinant protein incubated in a final reaction volume of 100 μL containing 1 mM substrate and 6.9 μM S-[Methyl-14C]-adenosyl-L-methionine in 50 mM Tris-HCl, pH 7.5, 5 mM MgCl2, and 10% (v/v) glycerol for 30 min at 37 °C. The enzymatic products were extracted with four volumes of ethyl acetate and quantified using a scintillation counter (Model LS6500, Beckman). For reaction product identification, assays were scaled up to a final volume of 500 μL containing ~50 μg of recombinant protein, 2 mM substrate and 2 mM S-adenosyl-L-methionine in 50 mM Tris-HCl, pH 7.5, 5 mM MgCl2, and 10% (v/v) glycerol for 60 min at 37 °C. Enzymatic products were extracted as above, evaporated to dryness under N2 gas, and
resuspended in 100 μL of methanol. Samples were applied to a Spherisorb ODS2 reverse-phase column (250 mm × 4.6 mm, 5 μm; Supelco) and resolved by HPLC using a non-linear potassium phosphate buffer (50 mM, pH 3.0) and acetonitrile gradient at a flow rate of 1 mL min⁻¹. The amount of acetonitrile in the mobile phase under the starting conditions was 10% and then increased to 30% during the first 10 min. After being maintained at 30% for an additional 10 min, the acetonitrile concentration was then increased by 5% increments every 5 min during the next 30 min such that the final concentration was 60% at the 50 min mark. The eluted products were detected by absorption at 340 nm and quantified relative to authentic standards. Mass spectral analysis of the enzymatic products was performed as described below.

4.6. Microsome extraction and prenyltransferase enzyme assays

Yeast cells expressing the various prenyltransferases were isolated as described above. The cell pellets were re-suspended in 100 mM Tris-HCl, pH 9.0 and disrupted with one-half volume of acid-washed glass beads (425–600 μm, Sigma-Aldrich) for a total of four min (30 s vortex; 30 s on ice). Following lysis, cell debris and glass beads were removed by centrifugation (1,500 × g, 20 min, 4 °C) and microsomes were pelleted from the supernatant by ultracentrifugation (160,000 × g, 90 min, 4 °C). The resulting supernatant was removed and the pelleted microsomes were then re-suspended in 100 mM Tris-HCl, pH 9.0 and protein concentration was determined by the method of Bradford (1976) using BSA as a standard. Prenytransferase enzyme assays were conducted with ~200 μg of microsomal protein in a final reaction volume of 1 mL containing 200 μL of prenyl acceptor substrate and 400 μM DMAPP, GPP or IPP in 100 mM Tris-HCl, pH 9.0 and 10 mM MgCl₂. Assays were allowed to proceed for 60 min at 37 °C and then terminated with the addition of 10 μL of 20% formic acid. Prenylated reaction products were extracted with two volumes of ethyl acetate, evaporated to dryness under N₂ gas, and then re-suspended in 100 μL of methanol. The samples were applied to a Spherisorb ODS2 reverse-phase column (250 mm × 4.6 mm, 5 μm; Supelco) and eluted with a 20 min linear gradient from 45% to 95% methanol in water, pH 2.7 containing 0.1% formic acid (v/v). The mobile phase was maintained at 100% methanol for an additional 10 min. Products were detected by absorption at 340 nm and quantified relative to authentic standards.

4.7. Synthesis of 6-dimethylallyl flavone standards

A flavonoid prenyltransferase from G. uralensis (GuA6DT) that catalyzes the regiospecific addition of DMAPP onto position 6 of the A-ring on a variety of flavones (Li et al., 2014) was expressed in yeast microsomes, as described above. According to the in vitro prenyltransferase assay conditions outlined above, yeast microsomes expressing GuA6DT were supplied with apigenin, chrysoeriol, or luteolin as flavone substrates, along with DMAPP as a prenyl donor. The enzymatic reaction products from these assays were resolved by HPLC and compounds corresponding to 6-dimethylallyl apigenin, 6-dimethylallyl chrysoeriol, and 6-dimethylallyl luteolin were collected off-line at retention times of 19.58, 19.93, and 18.49 min, respectively.

4.8. Mass spectrometry analysis of enzymatic reaction products

The prenylated flavones that were produced by GuA6DT or CsP3T, in vitro, were purified by HPLC as described above. Samples were then subjected to liquid chromatography mass spectrometry analysis performed on an Agilent 1200 HPLC liquid chromatograph interfaced with an Agilent UHD 6530 Q-TOF mass spectrometer. A C₁₈ cartridge column (Agilent Rapid Resolution 2.1 × 30 mm, 3.5 μm) at 30 °C was used with the following solvents: 1:1 water and acetonitrile both with 0.1% formic acid. The first 2 and last 5 min of the isocratic flow were sent to waste and not to the spectrometer. The flow rate was maintained at 0.4 mL min⁻¹. The mass spectrometer electrospray capillary voltage was maintained at 4.0 kV and the drying gas temperature at 250 °C with a flow rate of 8 L/min. Nebulizer pressure was 30 psi and the fragmentor was set to 160 V. Nitrogen was used as both nebulizing, drying gas, and collision-induced dissociation gas. The mass-to-charge ratio was scanned across a range of 100–3000 m/z in 4 GHz extended dynamic range positive-ion MS mode. The instrument was externally calibrated with the ESI TuneMix (Agilent). The sample injection volume was 10 μL. Chromatograms were analyzed within Agilent Qualitative Analysis software B 08.0 finding compounds by the Molecular Feature algorithm and generating possible compound formulas including elements C, H, O, and N. Fragmentation patterns of the various parent (molecular) ions were obtained using collision energies of 5, 10 and 20 eV, with 20 eV being optimal.

4.9. NMR characterization of enzymatic reaction products

The enzymatic reaction products from assays with GuA6DT were resolved by HPLC as described above. Compounds suspected to be cannflavin A and B eluted at 23.35 min and 20.03 min, respectively, and were subsequently collected. Approximately 0.5 mg of each compound was evaporated to dryness under N₂ gas, resuspended in acetone-d₆ and analyzed using ¹H and ¹³C NMR. NMR spectra were collected on a Bruker AVANCE III 600 MHz spectrometer equipped with a 5 mm TCI cryoprobe. The sample temperature was regulated at 298 ± 1 K. Peak assignments for cannflavin A and B were determined using standard 2D pulse sequences (COSY: cosygaf, TOCSY: ddp32gphps, HSQC: hsqctgspip2p.2, HMBC: hmbcgp2ndqf). The HMBC was collected with 768 increments in the indirect dimension; all other experiments were collected with 256 indirect increments. The TOCSY mixing time was set to 80 msec, and the HMBC coupling constant was set to 10 Hz.

Cannflavin A: ¹H NMR (Acetone-d₆, 600 MHz): δH 6.68 (1H, s, H-3), 6.62 (1H, s, H-8), 7.60 (1H, s, H-2), 7.00 (1H, d, J = 8.3 Hz, H-5'), 7.57 (1H, d, J = 8.3 Hz, H-6') 3.36 (2H, d, J = 7.1 Hz, H-1'), 5.29 (1H, dt, J = 1.1 Hz, 7.2 Hz, H-2'), 1.96 (2H, t, t, J = 7.1 Hz, H-4'), 2.05 (2H, m, H-5'), 5.07 (1H, t, J = 7.1 Hz, H-6'), 1.59 (3H, s, H-8'), 1.54 (3H, s, H-9'), 1.79 (3H, s, H-10'), 3.98 (3H, s, O-CH₃). 13.3 (5H, bs, 5-OH), ¹³C NMR (Acetone-d₆, 150 MHz): δC 164.4 (C-2), 104.2 (C-3), 182.8 (C-4), 159.8 (C-5), 121.0 (C-6'), 12.61 (C-7), 93.9 (C-8), 156.3 (C-9), 104.9 (C-10), 123.4 (C-11), 110.0 (C-20), 148.5 (C-3'), 151.0 (C-4'), 115.8 (C-5'), 121.0 (C-6'), 21.8 (C-1'), 122.9 (C-2'), 135 (C-3'), 40.1 (C-4'), 27.6 (C-5'), 124.8 (C-6'), 131.3 (C-7'), 25.3 (C-8'), 17.3 (C-9'), 6.0 (C-10'), 56.2 (O-CH₃).

Cannflavin B: ¹H NMR (Acetone-d₆, 600 MHz): δH 6.69 (1H, s, H-3), 6.62 (1H, s, H-8), 7.61 (1H, d, J = 2.1 Hz, H-2'), 7.00 (1H, d, J = 8.3 Hz, H-5'), 7.59 (1H, dd, J = 8.3 Hz, 2.1 Hz, H-6'), 3.36 (2H, d, J = 7.2 Hz, H-1'), 5.28 (1H, m, H-2'), 1.78 (3H, s, H-4'), 1.65 (3H, d, J = 0.9 Hz, H-5'), 3.99 (3H, s, O-CH₃), 13.3 (5H, bs, 5-OH). ¹³C NMR (Acetone-d₆, 150 MHz): δC 164.7 (C-2), 104.5 (C-3), 183.2 (C-4), 160.2 (C-5), 112.3 (C-6), 162.3 (C-7), 94.1 (C-8), 156.6 (C-9), 105.3 (C-10), 123.8 (C-11), 110.5 (C-20), 148.8 (C-3'), 151.3 (C-4'), 116.3 (C-5'), 121.3 (C-6'), 22.0 (C-1'), 123.2 (C-2'), 131.7 (C-3'), 17.9 (C-4'), 25.9 (C-5'), 56.6 (O-CH₃).

Declarations of interest


Acknowledgments

We thank Drs. Armen Charchoglyan and Dyanne Brewer for their expertise with mass spectrometry analysis and Dr. Gale Bozzo for fruitful discussions regarding flavonoid biochemistry.
Med. 84, 234–241.


