6-26-2018

Catalytic Bases and Stereocontrol in Lamiaceae Class II Diterpene Cyclases

Samuel Schulte
Iowa State University

Kevin C. Potter
Iowa State University

Cody Lemke
Iowa State University, clemke@iastate.edu

Reuben J. Peters
Iowa State University, rjpeters@iastate.edu

Catalytic Bases and Stereocontrol in Lamiaceae Class II Diterpene Cyclases

Abstract

Plants from the widespread Lamiaceae family produce many labdane-related diterpenoids, a number of which serve medicinal roles, and whose biosynthesis is initiated by class II diterpene cyclases (DTCs). These enzymes utilize a general acid-base catalyzed cyclo-isomerization reaction to produce various stereoisomers of the eponymous labdaenyl carbocation intermediate, which can then undergo rearrangement and/or the addition of water prior to terminating deprotonation. Identification of the pair of residues that cooperatively serve as the catalytic base in the DTCs that produce ent-copalyl diphosphate (CPP) required for gibberellin phytohormone biosynthesis in all vascular plants has led to insight into the addition of water as well as rearrangement. Lamiaceae plants generally contain an additional DTC that produces the enantiomeric normal CPP, as well as others that yield hydroxylated products derived from the addition of water. Here the catalytic base in these DTCs was investigated. Notably, changing two adjacent residues that seem to serve as the catalytic base in the normal CPP synthase from Salvia miltiorrhiza (SmCPS) to the residues found in the closely related perigrinol diphosphate synthase from Marrubium vulgare (MvPPS), which produces a partially rearranged and hydroxylated product derived from the distinct syn stereoisomer of labdaenyl+, altered product outcome in an unexpected fashion. Specifically, the relevant SmCPS:H315N/T316V double mutant produces terpentedienyl diphosphate, which is derived from complete substituent rearrangement of syn rather than normal labdaenyl+. Accordingly, alteration of the residues that normally serve as the catalytic base surprisingly can impact stereo-control.

Disciplines

Biochemistry | Biophysics | Molecular Biology | Plant Sciences | Structural Biology

Comments

This is a manuscript of an article published as Schulte, Samuel, Kevin C. Potter, Cody Lemke, and Reuben J. Peters. "Catalytic bases and stereocontrol in Lamiaceae class II diterpene cyclases." *Biochemistry* 57, no. 25 (2018): 3473-3479. doi: 10.1021/acs.biochem.8b00193. Posted with permission.

Published in final edited form as:

Biochemistry. 2018 June 26; 57(25): 3473–3479. doi:10.1021/acs.biochem.8b00193.

Catalytic bases and stereo-control in Lamiaceae class II diterpene cyclases

Samuel Schulte[†], **Kevin C. Potter**[†], **Cody Lemke**, and **Reuben J. Peters**[*]

Roy J. Carver Department of Biochemistry, Biophysics & Molecular Biology, Iowa State University, Ames, IA 50011, USA

Abstract

Plants from the widespread Lamiaceae family produce many labdane-related diterpenoids, a number of which serve medicinal roles, and whose biosynthesis is initiated by class II diterpene cyclases (DTCs). These enzymes utilize a general acid-base catalyzed cyclo-isomerization reaction to produce various stereoisomers of the eponymous labdaenyl carbocation intermediate, which can then undergo rearrangement and/or the addition of water prior to terminating deprotonation. Identification of the pair of residues that cooperatively serve as the catalytic base in the DTCs that produce *ent*-copalyl diphosphate (CPP) required for gibberellin phytohormone biosynthesis in all vascular plants has led to insight into the addition of water as well as rearrangement. Lamiaceae plants generally contain an additional DTC that produces the *enantiomeric* normal CPP, as well as others that yield hydroxylated products derived from the addition of water. Here the catalytic base in these DTCs was investigated. Notably, changing two adjacent residues that seem to serve as the catalytic base in the normal CPP synthase from *Salvia miltiorrhiza* (SmCPS) to the residues found in the closely related perigrinol diphosphate synthase from *Marrubium vulgare* (MvPPS), which produces a partially rearranged and hydroxylated product derived from the distinct *syn* stereoisomer of labdaenyl$^+$, altered product outcome in an unexpected fashion. Specifically, the relevant SmCPS:H315N/T316V double mutant produces terpentedienyl diphosphate, which is derived from complete substituent rearrangement of *syn* rather than normal labdaenyl$^+$. Accordingly, alteration of the residues that normally serve as the catalytic base surprisingly can impact stereo-control.

TOC image

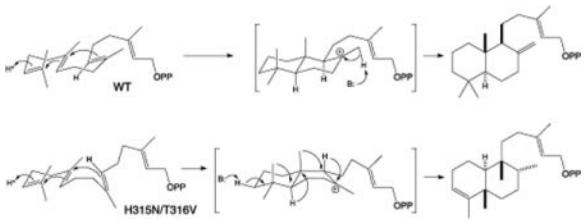
[*] rjpeters@iastate.edu.

[†]These authors contributed equally to this study

Current address: Biology Department, University of North Carolina, Chapel Hill, NC 27599, USA (K.C.P.); Medical Scientist Training Program, University of California, Los Angeles, and California Institute of Technology (S.S.)

SUPPORTING INFORMATION

Supporting figures for additional mutants. Available free of charge at the ACS Publications website.



INTRODUCTION

The labdane-related diterpenoid super-amily consists of over 7,000 known compounds, the vast majority of which are produced in plants as secondary or more specialized metabolites.[1] This natural product superfamily is defined by biosynthetic origin, namely the initiating (bi)cyclization reaction catalyzed by class II diterpene cyclases (DTCs). These react with the general diterpenoid precursor (*E*,*E*,*E*)-geranylgeranyl diphosphate (GGPP, **1**) and prototypically form a characteristic *trans*-decalin ring structure, although it has recently been shown that they can produce alternative ring structures as well.[2]

More specifically, DTCs utilize a general acid-base catalyzed reaction in which initiating protonation of the terminal carbon-carbon double-bond (C=C) of **1** leads to sequential *anti* addition from the two internal C=C to form the eponymous bicyclic intermediate labda-13*E*-en-8-yl$^+$ diphosphate (**2**$^+$). This can undergo a series of 1,2-shifts of the hydride and methyl substituents of the decalin ring structure, with the full series of migrations forming the clerodane-type skeletal structure (Scheme 1). The reaction can be terminated by deprotonation of any of the resulting carbocations, leading to an olefin – e.g., immediate deprotonation of **2**$^+$ to form labdadienyl/copalyl diphosphate (CPP, **2**). Alternatively, water can be added first, with terminating deprotonation then leading to a hydroxylated product instead – e.g., with **2**$^+$ to form copal-8-ol diphosphate (CLP, **3**). Notably, depending on the initial prochiral conformation of **1**, **2**$^+$ can be formed in any of four different configurations, which are distinguished by the stereochemistry at C9 and C10, although no DTC producing the *ent-syn* (9*S*,10*R*) isomer of **2**$^+$ has yet been identified.[1]

The evolutionary origins of the expanded labdane-related diterpenoid super-family in plants stems from the requisite production of gibberellins, as these phytohormones are needed for normal growth and development.[3] Accordingly, all vascular plants have a DTC that produces the relevant *ent*-**2** (9*R*,10*R*), termed a CPP synthase (CPS). Through gene duplication and neo-functionalization these CPSs have repeatedly given rise to DTCs that produce alternative products. For example, Lamiaceae plants seem to generally contain additional DTC(s), with a conserved clade that generally produce normal **2** (9*S*,10*S*) or the hydroxylated (normal) derivative **3** for more specialized metabolism.

It has been demonstrated that the DTC catalytic acid is the ‘middle’ aspartic acid in a highly conserved DxDD motif.[4] While there is no similarly broad conservation of the catalytic base, this is consistent with the necessary variation in positioning imposed by differing product outcome. Previous work, assisted by the determination of crystal structures for the CPS from *Arabidopsis thaliana* (AtCPS),[5, 6] identified the catalytic base in the *enantiomer* of

2 (*ent*-**2**) producing CPSs from gibberellin phytohormone biosynthesis. This is a dyad, consisting of a conserved histidine and asparagine that assist ligation of a water molecule that seems to serve as the actual base.[7] Strikingly, substitution of alanine at these positions can lead to the addition of the water and production of *ent*-**3**.[7, 8] By contrast, substitution of phenylalanine or tyrosine for the catalytic histidine in AtCPS leads to the full series of substituent migrations and production of *ent*-cleroda-3,13*E*-dienyl/kolavenyl diphosphate (*ent*-KPP).[9] Notably, these latter results were found to be predictive and enabled reverse engineering of two recently discovered *ent*-KPP synthases to produce *ent*-**2** instead.[10, 11] Moreover, similar results have been reported with other DTCs as well.[8, 12]

These previous reports demonstrate the importance of DTC catalytic base positioning for product outcome, at least the addition of water and/or substituent migration. Here the catalytic base for the conserved clade of DTCs from more specialized labdane-related diterpenoid metabolism in the Lamiaceae was investigated. This led to the surprising discovery that changes to these residues can alter stereochemical outcome as well as lead to substituent migration.

MATERIALS AND METHODS

Unless stated otherwise, reagents utilized here were purchased from Fisher Scientific. Sequence analyses were carried out using CLC Main Workbench 8.0 (Qiagen), with the **2** producing CPSs from the following Lamiaceae plant species, *Coleus forkohlii* (CfCPS, KF444506),[13] *Marrubium vulgare* (MvCPS3, AIE77092),[14] *Rosmarinus officinalis* (RoCPS, KF805857),[15] *Salvia fruticosa* (SfCPS, KP091840),[16] *Salvia miltiorrhiza* (SmCPS, ABV57835;[17] and SmCPS2; AHJ59322),[18] along with **3** producing CLSs from *Coleus forskohlii* (CfCLS, KF444507)[13] and *Salvia sclarea* (SsCLS, AET21247),[19] as well as the peregrinol diphosphate synthase from *M. vulgare* (MvPPS, AIE77090),[14] and AtCPS (NP_192187)[20] as the ancestral out-group sequence. A model for SmCPS was built using the known high-resolution AtCPS structure complexed with an early reaction intermediate analog (4LIX),[6] via the SWISS-MODEL webtool.[21]

SmCPS and AtCPS were recombinantly expressed using the previously described pseudomature constructs.[17, 22] Site-directed mutagenesis was conducted by whole plasmid PCR amplification of the previously described Gateway (Invitrogen) pENTR derived constructs,[17] with overlapping forward and reverse primers incorporating the desired mutation (Table S1). In each case the entire gene was sequence verified before transfer via directional recombination to a pACYC-Duet (Novagen) derived pGG-DEST expression vector, which contains a GGPP synthase for the production of **1**, as previously described.[23]

The resulting pGG-DEST expression constructs were transformed into the C41 OverExpress strain of *Escherichia coli* (Lucigen) along with the pIRS plasmid, which overexpresses three upstream enzymes to increase flux toward isoprenoid biosynthesis, as previously described.[24] For product analysis the resulting recombinant bacteria were grown in 45 mL liquid Terrific Broth (TB) medium (13.3 g/L casein, 26.7 g/L yeast extract, 0.36% glycerol, 34 mg/L chloramphenicol, 250 mg/L spectinomycin) at 37 °C with shaking at 200 RPM until reaching an optical density at 600 nm (OD_{600}) of 0.6 – 0.8. Cultures were then transferred to

16 °C with shaking at 200 RPM for 30 minutes followed by addition of phosphate buffer (pH 7.5, to 100 mM), sodium pyruvate (to 50 mM), $MgCl_2$ (to 1 mM), and IPTG (to 0.5 mM) to induce expression. The cultures were then further incubated at 16 °C with shaking at 200 RPM for approximately 72 hours, with addition of pyruvate (again to 50 mM) after the first 24 hours. The DTC products, observed as the dephosphorylated primary alcohols (i.e., with hydroxyl group at carbon number 15) derived from the action of endogenous phosphatases (as indicated by the prime' notation), were extracted with an equal volume of hexanes, with gentle swirling and an overnight rest at 4 °C to let any emulsion settle out. The hexane extract was separated out and dried completely under N_2 gas, with the resulting residue was resuspended in 200 μL fresh hexanes for analysis by gas chromatography with mass spectrometric detection (GC-MS).

For in vitro assays WT and the H315N/T316V double mutant of SmCPS were transferred from pENTR to pET100 via PCR amplification and directional topoisomerization. The resulting constructs were transformed into the C41 strain of *E. coli*, and starter cultures (50 mL NZY medium) grown overnight (200 RPM at 37° C), which were used to inoculate 1 L NZY cultures that were grown to OD_{600} of 0.5 – 0.6 (200 RPM at 37° C), transferred to 16° C (200 RPM) for 1 hour before inducing with IPTG (1 mM) and then shaken for an additional ~16 hours. The cells were harvested by centrifugation (5,000 *g* for 20 min.), and the cell pellet resuspended in 20 mL lysis buffer (50 mM MOPSO (pH 6.9), 300 mM NaCl, 10% v/v glycerol), with 20 μL of 1M DTT (1 mM) and 200 μL of 100 mM PMSF (1 mM) were added. Cells were lysed via homogenization, with a pinch of DNase added. Proteins were purified on a gravity flow column using 1 mL Ni-NTA resin equilibrated in the lysis buffer. This was rotated for 30 minutes with the lysate to maximize binding before washing. After letting the lysate through, the column was washed with 10 bed volumes of wash buffer (50 mM MOPSO (pH 6.9), 150 mM NaCl, 10% glycerol, 25 mM imidazole), and then eluted with 10 bed volumes of elution buffer (50 mM MOPSO (pH 6.9), 150 mM NaCl, 10% glycerol, 250 mM imidazole). Fractions (0.5 mL) that contained SmCPS were combined and dialyzed at 4 °C for 4 hours against 1 L of dialysis buffer (50 mM MOPSO (pH 6.9), 2 mM DTT), which was then replaced with a fresh 1 L of dialysis buffer, with further dialysis overnight.

Reactions were carried out in 1 mL assay buffer (50 mM HEPES (pH 7.75), 100 mM KCl, 0.1 mM $MgCl_2$, 10% glycerol), with 22.5 μL of 1 mg/mL GGPP (50 μM) added as substrate, and the reactions begun via the addition of 50 μL purified SmCPS. These assays were incubated in a 30 °C water bath overnight and then 130 μL of 10x CIAP phosphatase buffer (500 mM Tris-HCl (pH 9.3), 100 mM $MgCl_2$, 10 mM $ZnCl_2$) was added, followed by 15 μL of phosphatase. These reactions were then let shake (200 RPM) at 37 °C overnight. The assays were then extracted thrice by vortexing with 1 mL of hexanes, the resulting organic extracts pooled, dried under N_2, and then resuspended in 100 μL of hexanes for GC-MS analysis. Labeling with ^2H_2O (D_2O) was carried out in the same assays, only the enzyme assay buffer was made in D_2O rather than H_2O, and the proteins exchanged into this D_2O assay buffer via a 20 kDa MWCO centrifugation concentrator, with triply repeated 10-fold concentration and dilution.

GC-MS analysis was carried out over a HP-5MS column (Agilent, 0.25 μm, 0.25ID, 30 m) with a 3900 GC with a Saturn 2100T ion trap MS (Varian). Samples (1 μL) were introduced via splitless injection at 250 °C. A 1.2 mL/minute helium flow rate was utilized with an oven temperature program consisting of a 50 °C hold for 3 minutes, 15 °C/minute increase to 300 °C, and a 300 °C hold for 3 minutes. MS data from an *m/z* range of 90 to 300 was collected from 13 minutes onward. Products were identified either by already known product outcome – i.e., the production of (normal) **2** by the SmCPS,[17] and *ent*-**2** by AtCPS[20] – or by comparison of retention time and mass spectra to other known enzymatic products – i.e., *syn*-CPP from a rice (*Oryza sativa*) CPS (OsCPS4),[25] (normal) *endo*-CPP from the class I activity blocking mutant of a bifunctional enzyme from *Selaginella moellendorfii* (SmCPS/KSL1:D500A/D504A),[26] (normal) **3** from a CLS from *Nicotiana glutinosa* (NgCLS),[12] peregrinol diphosphate from MvPPS,[14] *syn*-halima-5,13*E*-dienyl diphosphate from the OsCPS4:H501F mutant,[27] *ent*-KPP from the AtCPS:H263Y mutant,[9] and *syn*-KPP from a bacterial DTC.[28]

RESULTS AND DISCUSSION

Lamiaceae plants produce a variety of labdane-related diterpenoids, and generally contain a number of DTCs involved in more specialized metabolism. The first to be identified was a CPP synthase from *Salvia miltiorrhiza* (SmCPS),[17] which is involved in production of the tanshinones that are active constituents of this traditional Chinese medicinal herb (Danshen).[29] Several homologous DTCs that also produce (normal) **2** for similar phenolic abietane diterpenoid biosynthesis have been identified from *S. miltiorrhiza* and other Lamiaceae plants (i.e., CPSs).[13, 15, 16, 18] Two other close homologs that produce (normal) **3**, termed CLP synthases (CLSs) also have been identified.[13, 19, 30, 31] In addition, another close homolog that produces peregrinol diphosphate (PPP, **4**) from *M. vulgare* (MvPPS) for production of the anti-diabetic marrubiin, has been identified as well.[14] Notably, the production of **4** stems from the alternative *syn* (9*R*,10*S*) configuration of $\mathbf{2}^{+}$, with initial hydride migration (C9 → C8) followed by the addition of water and deprotonation (see Scheme 1).[4] Thus, MvPPS differs from closely related DTCs in both stereochemical outcome and carrying out partial substituent migration. Nevertheless, alignment demonstrated that these DTCs share ≥50% amino acid sequence identity with each other, and are distinct from other DTCs such as AtCPS, with which they share ≤40% identity (Figure S1).

Given the established importance of the ancestral catalytic base dyad in the CPSs from gibberellin biosynthesis, as represented by AtCPS, these positions were examined in this conserved clade of Lamiaceae DTCs from more specialized metabolism (Figure 1). Intriguingly, in place of the catalytic histidine, the Lamiaceae DTCs producing **2** all have a phenylalanine at this first position, whereas those producing hydroxylated derivatives (**3** or **4**) have a tyrosine instead. This immediately suggested the possibility that the tyrosine might serve to help position a water molecule for addition to (normal) $\mathbf{2}^{+}$ to produce (normal) **3**, which was investigated by site-directed mutagenesis with SmCPS. However, when incorporated into a previously described metabolic engineering system,[23] the resulting

SmCPS:F256Y continued to produce **2** – i.e., the same as wild-type (WT) SmCPS (c.f., Figure 2A&B) – observed here as the dephosphorylated derivative (**2'**) by GC-MS.

In place of the catalytic asparagine, the Lamiaceae DTCs producing (normal) **2** and **3** all have a histidine. By contrast, MvPPS contains an asparagine at this second position. To determine if this was involved in the distinct product outcome catalyzed by MvPPS, the corresponding histidine in SmCPS was substituted with asparagine. However, the resulting SmCPS:H315N mutant also produced **2** (Figure 2C). In addition, even when the residues at both positions were changed to those found in MvPPS, the resulting SmCPS:F256Y/H315N double mutant also continued to produce **2** (Figure 2D).

The reduced yield observed with SmCPS:F256Y/H315N was somewhat surprising, as previous work with AtCPS indicated that tyrosine would be more favorable than phenylalanine. In particular, modeling suggested that a tyrosine at this first position could hydrogen-bond to the catalytic asparagine, and the corresponding tyrosine mutant provides higher yield than does substitution with phenylalanine.[9] This suggests some difference in configuration of these residues in the Lamiaceae DTCs from more specialized metabolism.

Intriguingly, upon reexamination of the alignment co-conservation of the residue immediately following the second position was noted. In particular, the Lamiaceae DTCs with histidine at the second position invariably had an immediately following threonine. Moreover, a recent report indicated that this adjacent position can impact DTC product outcome as well.[10]

Modeling of SmCPS indicated that the side chains of these adjacent residues hydrogen-bond to each other (Figure 3). This further suggests that these residues might cooperatively serve as the catalytic base in these **2** or **3** producing DTCs. Accordingly, the relevance of this position also was examined in SmCPS. Substituting valine, as found in MvPPS (as well as AtCPS), essentially inactivated SmCPS, as very little product (only trace amounts of **2**) was observed with the resulting SmCPS:T316V mutant (Figure 2E).

Perhaps more interestingly, tandem substitution of the two adjacent residues led to a striking change in product outcome. Specifically, the SmCPS:H315N/T316V double mutant was found to largely yield a different product (Figure 2F). Comparison of the mass spectrum and retention time of this major product to authentic standards revealed it to be *syn*-KPP (also known as terpentedienyl diphosphate, TPP, **5**; see Figure S2). Moreover, SmCPS:H315N/T316V also produced appreciable amounts (~15%) of *syn*-halima-5,13*E*-dienyl diphosphate (*syn*-HPP, **6**). Reexamination of the product profiles of the other mutants indicated that only SmCPS:H315N produces **5** and **6**, and only in trace amounts. Notably, these products are similar to the **4** produced by MvPPS in that both are derived from initial formation of *syn*-$\mathbf{2}^{+}$, followed by substituent migrations, with terminating deprotonation yielding an olefin either after the full series (producing **5**) or before the final methyl migration (producing **6**). This reflects a change in folding of the substrate **1**, from the *pro*-chair-chair configuration that leads to (normal) $\mathbf{2}^{+}$ to a *pro*-chair-boat configuration that leads to *syn*-$\mathbf{2}^{+}$ instead. While this effect is surprising, it is consistent with the expected positioning of the

catalytic base near carbon-8 (C8) of the '**B**' ring whose configuration is altered by this double mutation.

Production of **5** requires deprotonation of the same carbon protonated to initiate bicyclization, which was hypothesized to occur by removal of the originally added proton. This mechanistic hypothesis was investigated by labeling assays with D_2O. Both WT and the H315N/T316V double-mutant of SmCPS were assayed in both H_2O (as the control) and D_2O. While the **2** produced in D_2O was labeled with deuterium, the **5** produced by SmCPS:H315N/T316V was not, consistent with removal of the originally added proton following the full series of substituent migrations (Figure 4). This deprotonation is presumably mediated by the same aspartate that acted as the catalytic acid and initiated bicyclization. Given that energetic barriers exist for each of these 1,2-shifts from *syn*-**2**$^+$,[27] it seems likely that this aspartate (D372) represents the general-base best positioned to deprotonate any of the relevant carbocation intermediates, with the observed minor production of **6** potentially stemming from alternative position of the reactants within the active site.

Despite the resulting nominal resemblance to SmCPS:H315N/T316V or MvPPS (respectively), it has already been reported that changing the first position of the catalytic base dyad to a phenylalanine or tyrosine in AtCPS (H263F/Y) results in formation of *ent*-KPP (**7**) rather than alteration of stereochemical outcome.[9] To further investigate any effect of the second position (and immediately following residue) on stereochemical outcome in AtCPS, these were changed to the histidine-threonine pairing found in the rest of this Lamiaceae DTC clade, either alone or in combination with H263Y/F. However, this (N322H ± V323T ± H263Y/F) does not seem to change stereochemical outcome either. Instead these changes generally decrease specificity (e.g., AtCPS:N322H/V323T produces a mixture of *ent*-**2** and *ent*-**3**), with deleterious effect on yield as well (Figure S3). Accordingly, the effect that altering these residues has on stereochemical outcome seems to be specific to this clade of Lamiaceae DTCs from more specialized metabolism.

To further investigate the basis for the ability of the H315N/T316V change to alter stereochemical outcome in SmCPS, it was hypothesized that other changes to the active site cavity might help promote the *pro*-chair-boat conformation of **1** required to form *syn*-**2**$^+$. This was investigated by altering aromatic residues that line the active site of the SmCPS model and are otherwise conserved in the Lamiaceae DTCs from more specialized metabolism, but differ in MvPPS, in the context of the H315N/T316V double-mutant (Figure S4). However, none of these additional mutations increased the overall yield of **5**. Moreover, changes to these residues on their own do not appear to alter stereochemical outcome either (Figure S5).

Hypothesizing that introduction of the ancestral histidine at the first position might lead to production of *syn*-CPP, this substitution was made in the context of the H315N/T316V double mutation. However, this SmCPS:F256H/H315N/T316V triple mutant was found to be inactive (Figure S5). Indeed, this F256H mutation alone led to greatly reduced activity (only trace amounts of **2**), which is consistent with the previously reported requirement for an electron-rich group such as a strongly aromatic residue near the incipient C8-carbocation

formed in cyclization of **1** to $\mathbf{2}^{+}$.[9] Similarly, while a histidine found in place of an otherwise highly-conserved tyrosine has been suggested to be the catalytic base in the rice *syn*-CPP synthase OsCPS4,[27] additional substitution of histidine at this position in SmCPS also blocked activity (i.e., the resulting SmCPS:H315N/T316V/Y505H triple mutant was inactive). By contrast, this Y505H mutation alone led to deprotonation of (normal) $\mathbf{2}^{+}$ at an alternative site (i.e., C7), leading to some production of *endo*-CPP (**8**) by SmCPS:Y504H, albeit in quite small amounts (Figure S5).

Given the influence of the neighboring residue for the second position of the ancestral catalytic base pair, the residues surrounding the first position were examined here as well. In particular, these residues in SmCPS were changed to their equivalents in MvPPS (see Figure 1), both on their own and in combination with the other residues investigated here. While many of the resulting SmCPS mutants were inactive, consistent with the importance of the positioning of the residue at this first position (e.g., to be near the C8-carbocation of $\mathbf{2}^{+}$), SmCPS:L255M/F256Y/Y505F produces significant amounts of **3**, stemming from the addition of water to $\mathbf{2}^{+}$ prior to deprotonation (Figure S5).

CONCLUSION

The results reported here suggest that the catalytic base in the conserved clade of Lamiaceae DTCs involved in more specialized labdane-related diterpenoid metabolism generally consists of an adjacent hydrogen-bonded histidine and threonine pair (e.g., H315 and T316 in SmCPS). In this pair the threonine plays a surprisingly important role, as changing it to the valine found at this position not only in MvPPS but also all the *ent*-**2** producing CPSs from gibberellin biosynthesis, essentially blocks catalysis. This is presumably due, at least in part, to altered positioning of the histidine side-chain. By contrast, changing both residues in tandem (i.e., to the ancestral asparagine-valine pairing) enables catalysis to proceed. However, this appears to occur much more readily if the substrate **1** is folded into the distinct *pro*-chair-boat conformation, leading to predominant formation of *syn*-$\mathbf{2}^{+}$, rather than the *pro*-chair-chair conformation that leads to (normal) $\mathbf{2}^{+}$ and, hence, the usual production of **2** or **3** by the DTCs from this clade. Although unexpected and seemingly limited to these DTCs, perhaps reflecting the likely production of *syn*-$\mathbf{2}^{+}$ by the ancestral member of this Lamiaceae clade of DTCs,[32] this effect is consistent with the expected location of the catalytic base near the altered '**B**' ring of $\mathbf{2}^{+}$. The previously noted requirement for an electron-rich group near the incipient C8-carbocation of $\mathbf{2}^{+}$ is presumably fulfilled by the phenylalanine occupying the first position of the ancestral (gibberellin CPS) catalytic base pair. This enables the exergonic formation of *syn*-$\mathbf{2}^{+}$, which is thermodynamically trapped, leading to the substituent migrations required for the observed production of **5** and **6**. While surprising, the impact of these residues on both stereochemical outcome, as well as substituent migration (Scheme 2), further illustrates the plasticity of DTCs, driven in no small part by the previously noted intrinsic reactivity of terpenoids themselves.[33]

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by a grant from the NIH (GM076324 to R.J.P.).

References

1. Peters RJ. Two rings in them all: The labdane-related diterpenoids. Nat Prod Rep. 2010; 27:1521–1530. [PubMed: 20890488]
2. Xu M, Jia M, Hong YJ, Yin X, Tantillo DJ, Proteau PJ, Peters RJ. Premutilin Synthase: Ring Rearrangement by a Class II Diterpene Cyclase. Org Lett. 2018; 20:1200–1202. [PubMed: 29388775]
3. Zi J, Mafu S, Peters RJ. To Gibberellins and Beyond! Surveying the Evolution of (Di)Terpenoid Metabolism. Annu Rev Plant Biol. 2014; 65:259–286. [PubMed: 24471837]
4. Prisic S, Xu J, Coates RM, Peters RJ. Probing the role of the DXDD motif in class II diterpene cyclases. ChemBioChem. 2007; 8:869–874. [PubMed: 17457817]
5. Köksal M, Hu H, Coates RM, Peters RJ, Christianson DW. Structure and mechanism of the diterpene cyclase ent-copalyl diphosphate synthase. Nat Chem Biol. 2011; 7:431–433. [PubMed: 21602811]
6. Köksal M, Potter K, Peters RJ, Christianson DW. 1.55A-resolution structure of ent-copalyl diphosphate synthase and exploration of general acid function by site-directed mutagenesis. Biochim Biophys Acta. 2014; 1840:184–190. [PubMed: 24036329]
7. Potter K, Criswell J, Peters RJ. Novel product chemistry from mechanistic analysis of ent-copalyl diphosphate synthases from plant hormone biosynthesis. Angew Chem Int Ed. 2014; 53:7198–7202.
8. Mafu S, Potter KC, Hillwig ML, Schulte S, Criswell J, Peters RJ. Efficient heterocyclisation by (di)terpene synthases. Chem Commun (Camb). 2015; 51:13485–13487. [PubMed: 26214384]
9. Potter KC, Zi J, Hong YJ, Schulte S, Malchow B, Tantillo DJ, Peters RJ. Blocking Deprotonation with Retention of Aromaticity in a Plant ent-Copalyl Diphosphate Synthase Leads to Product Rearrangement. Angew Chem Int Ed. 2016; 55:634–638.
10. Hansen NL, Nissen JN, Hamberger B. Two residues determine the product profile of the class II diterpene synthases TPS14 and TPS21 of Tripterygium wilfordii. Phytochemistry. 2017; 138:52–56. [PubMed: 28279524]
11. Pelot KA, Mitchell R, Kwon M, Hagelthorn DM, Wardman JF, Chiang A, Bohlmann J, Ro DK, Zerbe P. Biosynthesis of the psychotropic plant diterpene salvinorin A: Discovery and characterization of the Salvia divinorum clerodienyl diphosphate synthase. Plant J. 2017; 89:885–897. [PubMed: 27865008]
12. Criswell J, Potter K, Shephard F, Beale MB, Peters RJ. A single residue change leads to a hydroxylated product from the class II diterpene cyclization catalyzed by abietadiene synthase. Org Lett. 2012; 14:5828–5831. [PubMed: 23167845]
13. Pateraki I, Andersen-Ranberg J, Hamberger B, Heskes AM, Martens HJ, Zerbe P, Bach SS, Moller BL, Bohlmann J, Hamberger B. Manoyl Oxide (13R), the Biosynthetic Precursor of Forskolin, Is Synthesized in Specialized Root Cork Cells in Coleus forskohlii. Plant Physiol. 2014; 164:1222–1236. [PubMed: 24481136]
14. Zerbe P, Chiang A, Dullat H, O'Neil-Johnson M, Starks C, Hamberger B, Bohlmann J. Diterpene synthases of the biosynthetic system of medicinally active diterpenoids in Marrubium vulgare. Plant J. 2014; 79:914–927. [PubMed: 24990389]
15. Brückner K, Bozic D, Manzano D, Papaefthimiou D, Pateraki I, Scheler U, Ferrer A, de Vos RC, Kanellis AK, Tissier A. Characterization of two genes for the biosynthesis of abietane-type diterpenes in rosemary (Rosmarinus officinalis) glandular trichomes. Phytochemistry. 2014; 101:52–64. [PubMed: 24569175]
16. Bozic D, Papaefthimiou D, Bruckner K, de Vos RC, Tsoleridis CA, Katsarou D, Papanikolaou A, Pateraki I, Chatzopoulou FM, Dimitriadou E, Kostas S, Manzano D, Scheler U, Ferrer A, Tissier A, Makris AM, Kampranis SC, Kanellis AK. Towards Elucidating Carnosic Acid Biosynthesis in Lamiaceae: Functional Characterization of the Three First Steps of the Pathway in Salvia fruticosa and Rosmarinus officinalis. PLoS One. 2015; 10:e0124106. [PubMed: 26020634]

17. Gao W, Hillwig ML, Huang L, Cui G, Wang X, Kong J, Yang B, Peters RJ. A functional genomics approach to tanshinone biosynthesis provides stereochemical insights. Org Lett. 2009; 11:5170–5173. [PubMed: 19905026]

18. Cui G, Duan L, Jin B, Qian J, Xue Z, Shen G, Snyder JH, Song J, Chen S, Huang L, Peters RJ, Qi X. Funcation divergence of diterpene synthases in the medicinal plant *Salvia miltiorrhiza* Bunge. Plant Physiol. 2015; 169:1607–1618. [PubMed: 26077765]

19. Schalk M, Pastore L, Mirata MA, Khim S, Schouwey M, Deguerry F, Pineda V, Rocci L, Daviet L. Towards a Biosynthetic Route to Sclareol and Amber Odorants. J Am Chem Soc. 2012; 134:18900–18903. [PubMed: 23113661]

20. Sun TP, Kamiya Y. The Arabidopsis GA1 locus encodes the cyclase *ent*-kaurene synthetase A of gibberellin biosynthesis. Plant Cell. 1994; 6:1509–1518. [PubMed: 7994182]

21. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Gallo Cassarino T, Bertoni M, Bordoli L, Schwede T. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res. 2014; 42:W252–258. [PubMed: 24782522]

22. Prisic S, Peters RJ. Synergistic substrate inhibition of ent-copalyl diphosphate synthase: A potential feed-forward inhibition mechanism limiting gibberellin metabolism. Plant Physiol. 2007; 144:445–454. [PubMed: 17384166]

23. Cyr A, Wilderman PR, Determan M, Peters RJ. A Modular Approach for Facile Biosynthesis of Labdane-Related Diterpenes. J Am Chem Soc. 2007; 129:6684–6685. [PubMed: 17480080]

24. Morrone D, Lowry L, Determan MK, Hershey DM, Xu M, Peters RJ. Increasing diterpene yield with a modular metabolic engineering system in E. coli: comparison of MEV and MEP isoprenoid precursor pathway engineering. Appl Microbiol Biotechnol. 2010; 85:1893–1906. [PubMed: 19777230]

25. Xu M, Hillwig ML, Prisic S, Coates RM, Peters RJ. Functional identification of rice *syn*-copalyl diphosphate synthase and its role in initiating biosynthesis of diterpenoid phytoalexin/allelopathic natural products. Plant J. 2004; 39:309–318. [PubMed: 15255861]

26. Jia M, Potter KC, Peters RJ. Extreme promiscuity of a bacterial and a plant diterpene synthase enables combinatorial biosynthesis. Metab Eng. 2016; 37:24–34. [PubMed: 27060773]

27. Potter KC, Jia M, Hong YJ, Tantillo DJ, Peters RJ. Product rearrangement from altering a single residue in the rice syn-copalyl diphosphate synthase. Org Lett. 2016; 18:1060–1063. [PubMed: 26878189]

28. Hamano Y, Kuzuyama Y, Itoh N, Furihata K, Seto H, Dairi T. Functional analysis of eubacterial diterpene cyclases responsible for biosynthesis of a diterpene antibiotic, terpentecin. J Biol Chem. 2002; 277:37098–37104. [PubMed: 12138123]

29. Cheng Q, Su P, Hu Y, He Y, Gao W, Huang L. RNA interference-mediated repression of SmCPS (copalyldiphosphate synthase) expression in hairy roots of Salvia miltiorrhiza causes a decrease of tanshinones and sheds light on the functional role of SmCPS. Biotechnology letters. 2014; 36:363–369. [PubMed: 24078134]

30. Caniard A, Zerbe P, Legrand S, Cohade A, Valot N, Magnard JL, Bohlmann J, Legendre L. Discovery and functional characterization of two diterpene synthases for sclareol biosynthesis in *Salvia sclarea* (L.) and their relevance for perfume manufacture. BMC Plant Biol. 2012; 12:119. [PubMed: 22834731]

31. Günnewich N, Higashi Y, Feng X, Choi KB, Schmidt J, Kutchan TM. A diterpene synthase from the clary sage *Salvia sclarea* catalyzes the cyclization of geranylgeranyl diphosphate to (8R)-hydroxy-copalyl diphosphate. Phytochemistry. 2013; 91:93–99. [PubMed: 22959531]

32. Heskes AM, Sundram TCM, Boughton BA, Jensen NB, Hansen NL, Crocoll C, Cozzi F, Rasmussen S, Hamberger B, Hamberger B, Staerk D, Moller BL, Pateraki I. Biosynthesis of bioactive diterpenoids in the medicinal plant Vitex agnus-castus. Plant J. 2018; 93:943–958. [PubMed: 29315936]

33. Tantillo DJ. Importance of Inherent Substrate Reactivity in Enzyme-Promoted Carbocation Cyclization/Rearrangements. Angew Chem Int Ed Engl. 2017; 56:10040–10045. [PubMed: 28349600]

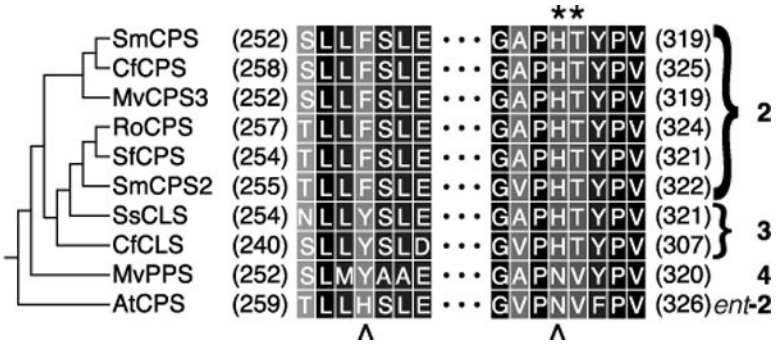


Figure 1.

Conservation of ancestral base dyad positions (indicated by '^' underneath, with AtCPS as representative of the ancestral CPSs from gibberellin phytohormone biosynthesis) in the conserved clade (cladogram shown on left) of Lamiaceae DTCs (with products, numbered as in the text, shown on the right) involved in more specialized metabolism (names as defined in the text), and their catalytic base pair as identified here (indicated by '*' above).

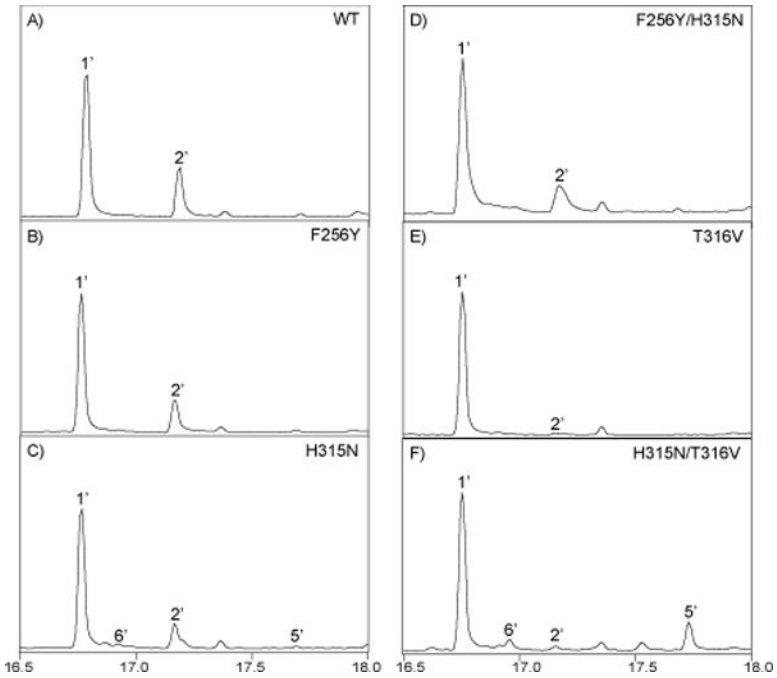


Figure 2.

Identification of catalytic base in conserved clade of Lamiaceae DTCs from more specialized metabolism. GC-MS chromatograms (total ion counts) of dephosphorylated product(s) of SmCPS, either wild-type (WT) or the indicated mutant (**1'**, geranylgeraniol; **2'**, copalol; **5'**, *syn*-kolavenol; **6'**, *syn*-halima-5,13*E*-dien-15-ol).

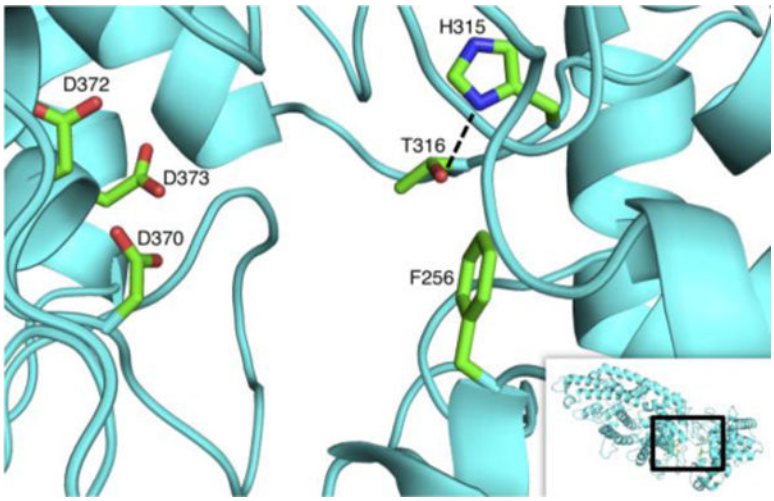


Figure 3.

SmCPS active site. Modeled structure, with the side chains of the key residues targeted here, along with the aspartates of the DxDD motif that defines the active site, shown and labeled. Also depicted is the hydrogen bond predicted between H315 and T316.

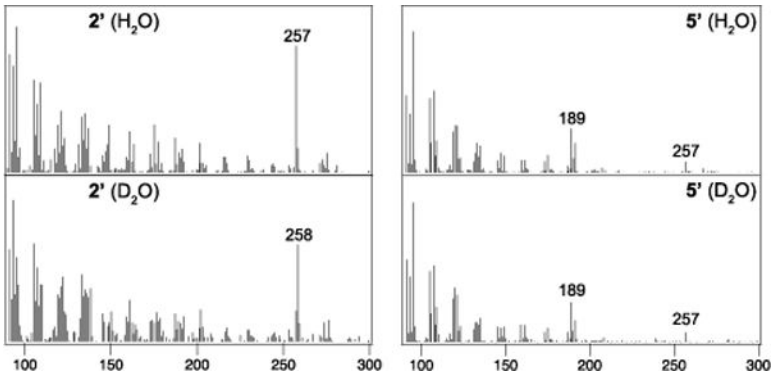
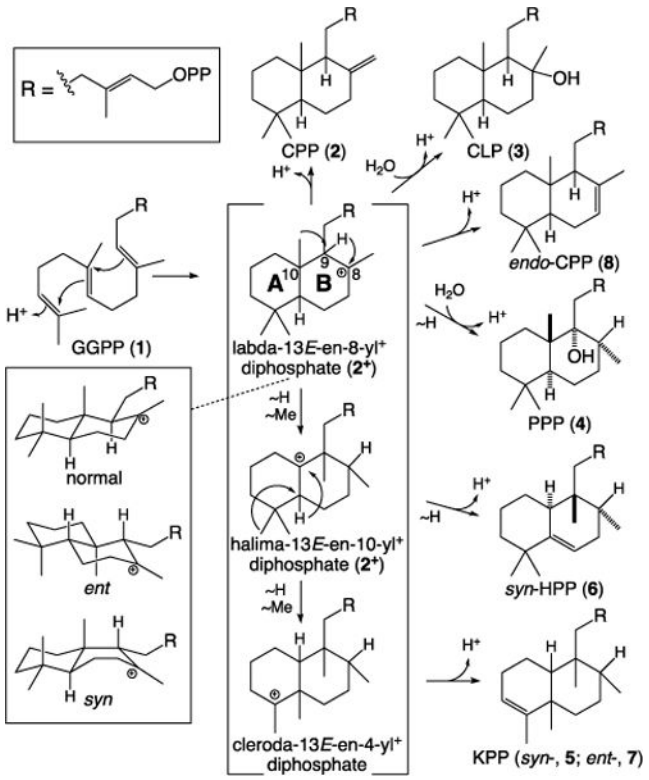


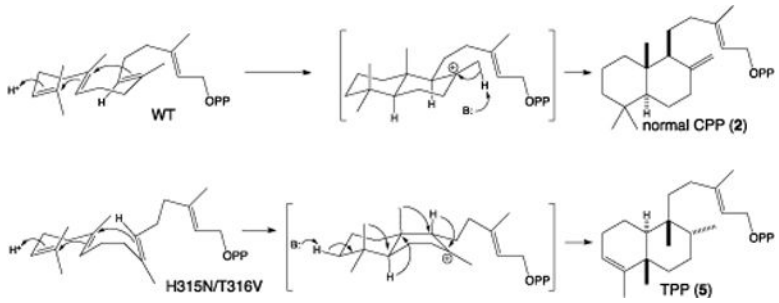
Figure 4.

Initially added proton is removed in formation of **5**. Mass spectra for (dephosphorylated) products from assays in H_2O or D_2O (as indicated).



Scheme 1.

DTC cyclization depicting formation of the general (**2**$^+$, **2**, **3** and **8**) and more specific (**4**, **5**, **6** and **7**) products mentioned in the text, along with the known stereoisomers for the initial bicyclic intermediate **2**$^+$ (with numbering of relevant carbons and lettering for the decalin ring structure).



Scheme 2.

Reactions catalyzed by wild-type (WT) and H315N/T316V double-mutant of SmCPS.