Published in final edited form as:
Chem Biol. 2011 January 28; 18(1): 32–37. doi:10.1016/j.chembiol.2010.11.008.

Genome Mining in *Streptomyces clavuligerus*. Expression and Biochemical Characterization of Two Cryptic Sesquiterpene Cyclases, (−)-δ-Cadinene Synthase and (+)-T-Muurolol Synthase

Yunfeng Hu, **Wayne K. W. Chou**, **Russell Hopson**, and **David E. Cane***
Department of Chemistry, Brown University, Box H, Providence, Rhode Island 02912-9108, USA

SUMMARY

Two presumptive terpene synthases of unknown biochemical function encoded by the *sscg_02150* and *sscg_03688* genes of *Streptomyces clavuligerus* ATCC 27074 were individually expressed in *Escherichia coli* as N-terminal-His_6-tag proteins, using codon-optimized synthetic genes. Incubation of recombinant SSCG_02150 with farnesyl diphosphate (**1**, FPP) gave (−)-δ-cadinene (**2**) while recombinant SSCG_03688 converted FPP to (+)-T-muurolol (**3**). Individual incubations of (−)-δ-cadinene synthase with [1,1-$^2H_2$]FPP (**1a**), (1*S*)-[1-^{2}H]-FPP (**1b**), and (1*R*)-[1-^{2}H]-FPP (**1c**) and NMR analysis of the resulting samples of deuterated (−)-δ-cadinene supported a cyclization mechanism involving the intermediacy of nerolidyl diphosphate (**4**) leading to a helminthogermacradienyl cation **5**. Following a 1,3-hydride shift of the original H-1_{si} of FPP, cyclization and deprotonation will give (−)-δ-cadinene. Similar incubations with recombinant SSCG_03688 supported an analogous mechanism for the formation of (+)-T-muurolol (**3**), also involving a 1,3-hydride shift of the original H-$1_{si}$of FPP.

INTRODUCTION

Terpenoids constitute the largest and most varied group of natural products. To date, more than 50,000 terpenoid compounds with a wide range of chemical structures have been reported, most of which have been isolated from both terrestial and marine plants, as well as numerous fungi and a relatively small but increasing number of bacteria. The C_{15} sesquiterpene hydrocarbons and alcohols, which display the greatest range of structural diversity among the terpenes, include more than 300 identified carbon skeletons, each of which is derived by cyclization of the universal acyclic precursor farnesyl diphosphate (**1**, FPP). Extensive further metabolic transformations generate the enormous range of sesquiterpenoid metabolites.

Streptomyces are already noteworthy for the production of an enormous variety of natural products, many of which play important roles in human and animal medicine and in agriculture. More than 20 partial or complete *Streptomyces* genome sequences together harbor several dozen presumed terpene synthases, the vast majority of which are of

*Correspondence: David_Cane@brown.edu.

SUPPLEMENTAL INFORMATION
Full experimental details. SDS-PAGE gels, ^{1}H and ^{13}C NMR assignments, GC-MS data, and additional mechanistic schemes.

unknown biochemical function. The two recently independently determined genomic sequences of *S. clavuligerus* ATCC 27604, the producer of the widely used β-lactamase inhibitor clavulanic acid, indicate the presence of as many as 20 presumptive terpene synthase genes, most of which are located on a giant, 1.8-Mb linear plasmid, pSCL4 (Fischbach, et al., 2008; Medema, et al., 2010). Of these 20 ORFs, the biochemical function of only *sclav_p0159* can be predicted with any certainty, based on its 62-64% identity and 74-75% similarity (E value 0.0) over 724 amino acids (aa) to the known *S. coelicolor* and *S. avermitilis* geosmin synthases (Jiang, et al., 2006; Cane, et al., 2006).

Of the remaining *S. clavuligerus* terpene synthase genes, two in particular attracted our attention. Each is part of an apparent 3-4 gene operon, suggesting the encoded terpene synthases catalyze the first, committed step in the biosynthesis of oxidized and methylated sesquiterpene metabolites of as yet unknown structure. Thus *sscg_02150* (*sclav_p0328*; UniProt ID B5GS26) encodes a predicted 327-aa protein that is flanked downstream by a unidirectionally transcribed putative cytochrome P450 and upstream by a predicted methyl transferase and a divergently transcribed oxidoreductase (Figure 1A). Similarly, the *sscg_03688* gene (*sclav_p0068*, UniProt ID B5GW45) encodes a predicted 418-aa terpene synthase protein flanked immediately upstream and downstream by two unidirectionally transcribed cytochrome P450s and a downstream methyl transferase (Figure 1B). The cytochrome P450s and methyl transferases are presumed to carry out an unspecified series of oxidations and methylations of the individual sesquiterpene products initially produced by the parent terpene synthase from each operon. Both terpene synthases display the two highly conserved Mg^{2+}-binding motifs characteristic of essentially all known and presumed terpene synthases, the aspartate-rich ^{84}**DD**R**I**D motif and the downstream "NSE/DTE" triad ^{223}**N**DLM**T**VDK**E** in SSCG_02150 and the corresponding ^{83}**DD**EYC**D** and ^{225}**N**DIL**S**HHK**E** motifs in SSCG_03688. Similar to all other terpene synthases, these conserved motifs are each separated by 139 and 142 amino acids, respectively. On the other hand, neither protein shows any significant match in pairwise sequence alignments either to each other or to any terpene synthase of known biochemical function, and at best low-level (<33%) identity to predicted proteins generically annotated only as "terpene synthase" or erroneously as "pentalenene synthase". We have now determined the biochemical function of both enzymes, demonstrating that the SSCG_02150 protein catalyzes the cyclization of FPP to (−)-δ-cadinene (**2**) while the SSCG_03688 protein is shown to be a (+)-T-muurolol (**3**) synthase. We have also investigated the mechanism of both enzyme-catalyzed FPP cyclizations and determined the stereochemistry of the 1,3-hydride transfers involved in each reaction. Although both (+)-δ-cadinene and (−)-T-muurolol are widely occurring plant metabolites, neither (−)-**2** nor (+)-**3**, nor any derived sesquiterpenoid metabolites have previously been reported as constituents of *S. clavuligerus* or any other Streptomycete.

RESULTS and DISCUSSION

Expression and Biochemical Characterization of Recombinant SSCG_02150 and SSCG_03688

Synthetic genes for *S. clavuligerus sscg_02150* and *sscg_03688*, optimized for heterologous expression in *Escherichia coli*, were ligated into the pET-28a(+) expression vector and the derived plasmids, pET-02150 and pET-03688, were each transformed into *E. coli* BL21(DE3). The resultant recombinant *N*-terminal His_6-tagged proteins, after purification to homogeneity by Ni-NTA chromatography (Figure S1), were each incubated with 60 μM FPP in the presence of 15 mM Mg^{2+} for 16 h at 30 °C. GC-MS analysis of the pentane extracts of the SSCG_02150 incubation mixture indicated the formation of a single (>95 %) sesquiterpene hydrocarbon, $C_{15}H_{24}$, *m/z* 204, while recombinant SSCG_03688 yielded a sesquiterpene alcohol, $C_{15}H_{26}O$, *m/z* 222 as the predominant (>95%) product (Figure S2).

Steady-State Kinetic Parameters

Kinetic parameters were measured for each synthase by carrying out 30-min incubations over a range of concentrations of [1-^{3}H]-FPP from 0.25 to 7.5 μM giving a k_{cat} 1.14 ± 0.03 × 10^{-3} s^{-1} and K_m 1.4 ± 0.1 μM for recombinant SSCG_02150, while recombinant SSCG_03688 had a k_{cat} 1.63 ± 0.09 × 10^{-3} s^{-1} and K_m 2.7 ± 0.3 μM.

Production in *E. coli* of Sesquiterpenes Generated by *S. clavuligerus* SSCG_02150 and SSCG_03688

In order to obtain sufficient quantities of the sesquiterpene products of each of the two *S. clavuligerus* terpene synthases, we utilized a metabolically engineered *E. coli* host designed for the efficient *in vivo* synthesis of terpenoid metabolites (Harada, et al., 2009). The two plasmids, pAC-Mev/Scidi/AacI and pET-02150 were used to co-transform *E. coli* JM109(DE3). Cultures of the resultant transformant, *E. coli* JM109(DE3)/pAC-Mev/Scidi/AacI/pET-02150, were induced with isopropyl–β-D-thiogalactopyranoside (IPTG) and supplemented with lithium acetoacetate and then incubated at 18 °C for 48 h. Extraction of the harvested cells with acetone followed by SiO_2 column chromatographic purification of the derived pentane-soluble products gave 10 mg/L of a sesquiterpene hydrocarbon, *m/z* 204, identical to that generated *in vitro* by SSCG_02150-catalyzed cyclization of FPP. In like manner, co-transformation of *E. coli* JM109(DE3) with pAC-Mev/Scidi/AacI and pET-03688 and pentane extraction of IPTG-induced, lithium acetoacetate-supplemented cultures of *E. coli* JM109(DE3)/pAC-Mev/Scidi/AacI/pET-03688 gave, after chromatographic purification of the concentrated pentane extract, 20 mg/L of a sesquiterpene alcohol, *m/z* 222, identical to the product previously obtained by SSCG_03688-catalyzed cyclization of FPP.

Identification of (−)-δ-Cadinene (2) and (+)-T-Muurolol (3)

The structures of the sesquiterpene hydrocarbon produced by SSCG_02150 and the sesquiterpene alcohol generated by SSCG_03688 were deduced as (−)-δ-cadinene (**2**) and (+)-T-muurolol (**3**), respectively, by a combination of 1D and 2D NMR (Figure S3), as well as GC-MS analysis. The ^{1}H NMR spectrum of **2** displayed one olefinic proton (δ 5.43 (s)) while one C-H (^{13}C 124.7 ppm) and 3 quaternary olefinic carbon signals (134.2, 129.9, 124.5 ppm) were apparent in the ^{13}C NMR spectrum, indicating the presence of 2 double bonds and therefore 2 rings (Table S1). Of the four methyl signals, two belonged to an isopropyl group (δ 0.77 (d) and 0.94 (d); ^{13}C 15.7 and 21.8 ppm) and two showed typical resonances for allylic methyl groups (δ 1.63 (bs) and 1.65 (bs); ^{13}C 18.5 and 23.6 ppm). The remaining connectivity was readily deduced from the HSQC and HMBC spectra, and the deduced δ-cadinene structure was confirmed by comparison with the published ^{1}H and ^{13}C NMR spectra of (+)-δ-cadinene (Davis, et al., 1996). The absolute configuration of the enzymatically generated (−)-δ-cadinene (**2**) was assigned by direct GC-MS comparison with authentic (+)-δ-cadinene ((+)-**2**), a major component of commercial cade oil (Davis, et al., 1996). While the two samples showed identical retention times (Retention Index [RI] 1520) and mass spectra when analyzed by capillary GC-MS, they were clearly distinguished by chiral capillary GC-MS.

^{1}H and ^{13}C NMR analysis of the purified sesquiterpene alcohol produced by *E. coli* JM109(DE3)/pAC-Mev/Scidi/AacI/pET-03688 showed the presence of a single trisubsituted double bond as evidenced by the proton signal at δ 5.54 (d) correlated with the ^{13}C resonance at 124.8 ppm and a second olefinic carbon at 133.5 ppm attached to an allylic methyl (δ 1.63 (bs), ^{13}C 23.6 ppm), indicating the presence of a bicyclic structure. A pair of geminal methyl groups indicated the presence of an isopropyl group (δ 0.82 (d) and 0.87 (d); ^{13}C 21.6 and 15.4 ppm), while a fourth methyl group (δ 1.18 (s), ^{13}C 29.3 ppm) was attached to the quaternary carbinol carbon (^{13}C 72.4 ppm). The remaining connectivity and

relative stereochemistry was deduced from the 2D NMR analysis and the T-muurolol structure was confirmed by the precise match to the reported NMR spectra of (−)-T-muurolol in both $CDCl_3$ and C_6D_6 (Ding, et al., 2009) and (+)-T-muurolol in $CDCl_3$ (Nagashima, et al., 1994), which all differed in several significant details from those reported for α-cadinol, amorphenol, and related diastereomers of **3**. The (+)-T-muurolol had an RI value of 1639 (cf. RI 1633 for (−)-**3**, MassFinder 4.0,). The absolute configuration of enzymatically-generated **3** was assigned from the $[\alpha]_D$ +85.79°, corresponding to that reported for (+)-T-muurolol (Nagashima, et al., 1994).

Mechanism and Stereochemistry of the Cyclization of FPP to (−)-δ-Cadinene (2) and (+)-T-Muurolol (3)

To elucidate the mechanism and stereochemistry of the enzymatic formation of formation of both (−)-δ-cadinene (**2**) and (+)-T-muurolol (**3**), we carried out preparative-scale incubations of each protein with individual samples of [1,1-$^2H_2$]-FPP (**1a**), (1*S*)-[1-^{2}H]-FPP (**1b**), and (1*R*)-[1-^{2}H]-FPP (**1c**). The resultant deuterated samples of (−)-δ-cadinene (**2a**-**2c**) (Figure 3) and (+)-T-muurolol (**3a**-**3c**) (Figure 4), were analyzed by GC-MS and NMR (Figures S4 and S5). (−)-δ-Cadinene (**2a**) derived from [1,1-$^2H_2$]-FPP (**1a**) retained both deuterium atoms, as evidenced by the parent peak at *m/z* 206 [M^+] and the base peak at *m/z* 162 [M–44] due to loss of the deuterated isopropyl side chain (Figure S4). The mass spectrum of [11-^{2}H]-**2b** derived from (1*S*)-[1-^{2}H]-FPP (**1b**) exhibited a parent peak of *m/z* 205 and a base peak *m/z* 161 [M–44] indicating the presence of the deuterium in the isopropyl side chain, while [5-^{2}H]-**2c** derived from (1*R*)-[1-^{2}H]-FPP (**1c**) had the complementary pattern with a parent peak of *m/z* 205 and a base peak *m/z* 162 [M–43]. The ^{1}H NMR spectrum of [11-^{2}H]-**2b** lacked the multiplet at δ 2.03 corresponding to the H-11 isopropyl methine proton, while the two neighboring isopropyl methyl protons each now appeared as singlets (Figure 3A). In the complementary experiment, [5-^{2}H]-**2c** lacked the bridgehead H-5 allylic proton resonance at δ 2.50 (Figure 3B). Consistent with these results, the HSQC spectrum of [5,11-$^2H_2$]-**2a** lacked both the H-5/C-5 and H-11/C-11 cross-peaks that were present in the spectrum of unlabeled **2** (data not shown).

The SSCG_03688-catalyzed formation of (+)-T-muurolol (**3**) showed the same stereochemistry for the 1,3-hydride shift as that determined for (−)-δ-cadinene synthase (Figure 2). Thus [5,11-$^2H_2$]-(+)-T-muurolol (**3a**) derived from [1,1-$^2H_2$]-FPP (**1a**) retained both deuterium atoms, as evidenced by the parent peak at *m/z* 224 [M^+], the [M – H_2O] peak at *m/z* 206, and a peak at *m/z* 162 [M – 64] due to loss of the deuterated isopropyl side chain plus water (Figure S5). Similarly, the mass spectrum of [11-^{2}H]-**3b** ([M^+] *m/z* 223) derived from (1*S*)-[1-^{2}H]-FPP (**1b**) displayed a peak at *m/z* 161 resulting from loss of the deuterated isopropyl side chain plus water, while [5-^{2}H]-**3c** ([M^+] *m/z* 223) derived from (1*R*)-[1-^{2}H]-FPP (**1c**) retained the deuterium in the *m/z* 162 fragment. The ^{1}H NMR spectrum of [5,11-$^2H_2$]-**3a** lacked the signals for both H-5 (δ 2.25) and H-11 (δ 1.96), while the corresponding spectrum of [11-^{2}H]-**3b** lacked only the signal for H-11, with the isopropyl methyl signals appearing as singlets at δ 0.80 and 0.86 (Figure 4).

Occurrence of δ-Cadinene and T-Muurolol

Although both (+)-δ-cadinene and (−)-T-muurolol, the enantiomorphs of the sesquiterpene cyclization products produced by the two *S. clavuligerus* terpene synthases, are common plant metabolites, the isolation of either enantiomer of **2** and **3** from bacterial cultures is very rare. Trace amounts of δ-cadinene of undetermined absolute configuration have been detected in the volatile organic products of the North Sea *Streptomyces* strain GWS-BW-H5 in which epicubenol (**8**) of unspecified configuration is the major metabolite (Dickschat, et al., 2005). Recently (−)-T-muurolol ((−)-**3**) has been isolated from the marine *Streptomyces* sp. M491, along with several oxidized T-muurolol derivatives, including 3-oxo-T-muurolol

and 3-oxo-15-hydroxy-T-muurolol (Ding et al., 2009). The only previously documented occurrence of (+)-T-muurolol (**3**) is its isolation from the Belgian liverwort *Scapania undulata* along with the closely related (+)-epicubenol ((+)-**8**) (Nagashima and Asakawa, 2001; Nagashima, et al., 1994). The latter sesquiterpene alcohol has previously been isolated from both *Streptomyces* sp. LL-B7 (Gerber, 1971) and the liverwort *Heteroscyphus planus* (Nabeta, et al., 1995; Nabeta, et al., 1994). The discovery that *S. clavuligerus* SSCG_02150 is a (−)-δ-cadinene synthase and that the SSCG_03688 protein is a (+)-T-muurolol synthase is the first documented evidence for the biosynthesis of the specific enantiomeric forms of each of these sesquiterpenes in *Streptomyces*. Since both of these *S. clavuligerus* sesquiterpene synthases are located within apparent gene clusters that also harbor genes for 1-2 additional cytochrome P450s and a methyl transferase, it is unlikely that either the parent sesquiterpene hydrocarbon **2** or the sesquiterpene alcohol **3** accumulates in *S. clavuligerus*.

Mechanism and Stereochemistry of the Cyclization Reactions

The enzyme-catalyzed formation of both (−)-δ-cadinene and (+)-T-muurolol mediated by the *S. clavuligerus* synthases encoded by the *sscg_02150* and *sscg_03688* genes both involve a 1,3-hydride shift of the original H-1_{si} of farnesyl diphosphate, presumably involving a common (*Z,E*)-helminthogermacradienyl cation intermediate **5** (Figure 2). Since the bridgehead stereochemistry at C-10 of the derived bicyclic cation **6** is uncertain in the case of (−)-δ-cadinene biosynthesis, while (+)-T-muurolol must be formed from the corresponding *cis*-fused muurolyl cation **7**, it is unclear at which point the two enzymatic cyclization pathways diverge. We have previously established the stereochemistry of the analogous 1,3-hydride shift in the cyclization of FPP to (+)-epicubenol (**8**) catalyzed by the *Streptomyces* sp. LL-B7 (+)-epicubenol synthase (Cane and Tandon, 1995) (Figure S6). The formation of (+)-**8** has been shown to involve the intermediacy of (3*R*)-nerolidyl diphosphate (**4**, NPP) that is generated by an initial *syn* allylic rearrangement of FPP. The cisoid conformer of (3*R*)-NPP then undergoes cyclization to the helminthogermacradienyl cation **5**. A similar 1,3-hydride shift has also been implicated in the formation of the enantiomeric (+)-δ-cadinene by *Gossypium arboreum* (cotton) cadinene synthase (Benedict, et al., 2001). Interestingly, intermolecular competition experiments using the cotton cyclase demonstrated that (3*R*)-NPP (**4**) is an intermediate in the formation of (+)-δ-cadinene, the enantiomer of the product generated by the *S. clavuligerus* SSCG_02150 enzyme. Assuming that this (3*R*)-NPP would be generated from (*E,E*)-FPP by the usual *syn* allylic rearrangement, it can be inferred that the 1,3-hydride shift in the biosynthesis of (+)-δ-cadinene also involves the H-1_{si} (H_b) of FPP. In fact, as Arigoni has elegantly demonstrated in a penetrating analysis of FPP cyclization reactions involving 1,3-hydride shifts that generate products with *cis* double bonds and isopropyl side chains, there is a strict correlation between the identity of the enantiotopic H-1 proton of FPP that undergoes the 1,3-hydride shift and the relative *cis*- or *trans*-configuration of the bridgehead proton and the isopropyl side chain in the ultimately formed sesquiterpene product (Arigoni, 1975). On the other hand, the absolute configuration of the enzyme-bound NPP intermediate cannot be deduced solely from the absolute or relative configuration of the eventually formed cadinene, muurolene, or amorphene product. We have provisionally invoked the intermediacy of (3*R*)-NPP (**4**) in the SSCG_02150-catalyzed formation of (−)-δ-cadinene (**2**) and SSCG_03688-catalyzed formation of (+)-T-muurolol (**3**) (Figure 2), based on analogy to the demonstrated role of (3*R*)-NPP (**4**) in the closely related reaction catalyzed by (+)-epicubenol (**8**) synthase from *Streptomyces* LL-B7 (Cane and Tandon, 1995) (Figure S6). An alternative cyclization pathway for (−)-δ-cadinene synthase involving (3*S*)-NPP that would give the same observed labeling results is illustrated in Figure S7. These two theoretical possibilities will have to be distinguished experimentally.

In spite of the fact that the *S. clavuligerus* (−)-δ-cadinene synthase (SSCG_02150) and (+)-T-muurulol synthase (SSCG_03688) are derived from the same organism and catalyze cyclization reactions that are closely related both mechanistically and stereochemically, the two proteins are remarkably different, showing <25% identity in pairwise alignment and possessing very different Mg^{2+}-binding motifs. The *S. clavuligerus* (−)-δ-cadinene synthase also has insignificant similarity either at the amino acid level or in the identity of presumptive active site residues to the well-studied *G. arboreum* (+)-δ-cadinene synthase (UniProt ID Q39761) (Benedict, et al., 2001; David, et al., 1996). Although the cotton enzyme has a typical aspartate-rich motif ^{308}**DD**TY**D**, it lacks the canonical NSE/NTE Mg^{2+}-binding triad, displaying instead a second aspartate-rich motif, ^{451}**DD**VA**E**. The *S. clavuligerus* (−)-δ-cadinene synthase is also distinct in both amino acid sequence and Mg^{2+}-binding motifs from the recently described δ-cadinene synthase of the basidiomycete *Coprinus cinereus* (Cop4, UniProt ID A8NU13) which generates δ-cadinene of unspecified absolute configuration (Agger, et al., 2009). *S. clavuligerus* SSCG_03688, the first reported T-muurolol synthase, also has insignificant amino acid sequence similarity to the γ-muurolene synthase of *C. cinereus* (Cop3, UniProt ID A8NE23). In spite of the substantial differences in both overall amino acid sequence and specific Mg^{2+}-binding motifs, both SSCG_02150 and SSCG_03688 are predicted by protein homology modeling using the Modweb server (http://modbase.compbio.ucsf.edu/modbase-cgi/index.cgi, (Pieper, et al, 2009)) to have the common all-α-helical class I terpenoid synthase fold that has been established by the 2.4 Å crystal structure of the cotton (+)-δ-cadinene synthase (Gennadios, et al., 2009) and the prototype bacterial sesquiterpene synthase pentalenene synthase (Lesburg, et al, 1997), as well as all other terpene synthases analyzed to date (Christianson, 2006).

Since terpene synthase sequences cluster in multiple sequence alignments predominantly according to phylogenetic origin rather than biochemical reaction type (Bohlmann, et al., 1998), bioinformatic analysis of deduced amino acid sequence alone is incapable of assigning the specific biochemical function of any newly recognized terpene synthase gene. Direct biochemical characterization of terpene synthases is therefore essential for assignment of the actual terpene cyclization product and even the native biological substrate (FPP or geranyl diphosphate) (Bohlmann and Gershonzon, 2009). Of the 20 or so sequences in the *S. clavuligerus* 27604 genome that are currently annotated as “terpene synthase”, only two of these ORFs are now of experimentally defined biochemical function: (−)-δ-cadinene synthase (SSCG_02150) and (+)-T-muurolol synthase (SSCG_03688).

SIGNIFICANCE

Genome mining has provided a powerful and versatile new paradigm for the discovery of new biochemical transformations and natural product pathways (Corre and Challis, 2009; Walsh and Fischbach, 2010). Although generic terpene synthases from both bacterial and eukaryotic genomes can be readily recognized by the presence of the characteristic pair of conserved Mg^{2+}-binding motifs, separated in all cases by 140±5 aa, a major challenge in the annotation of presumptive terpene synthases is the lack of an obvious correlation between detailed protein sequence and any specific terpene cyclization pathway. Direct biochemical characterization is therefore indispensable for the identification of both the actual acyclic terpene substrate and the derived cyclization product. The *sscg_02150* and *sscg_03688* gene products of *S. clavuligerus* have now been established to be (−)-δ-cadinene synthase and (+)-T-muurolol synthase, respectively. This represents the first isolation and characterization of either sesquiterpene synthase from any biological source and the first documented evidence for the occurrence of the specific enantiomers of either sesquiterpene in a bacterial host. The role of the two sesquiterpenes in *S. clavuligerus* metabolism and the structure of the ultimately derived terpenoid metabolites produced by the corresponding biosynthetic

gene clusters remain unknown, as does the function of the impressive number of remaining uncharacterized terpene synthases that make up the *S. clavuligerus* terpenome.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Norihiko Misawa of Kirin Holdings Co. Ltd. Japan for providing plasmid pAC-Mev/Scidi/AacI. This work was supported by National Institutes of Health Grant GM30301 to D. E. C. We thank Dr. Tun-Li Shen for assistance with mass spectrometry.

REFERENCES

Agger S, Lopez-Gallego F, Schmidt-Dannert C. Diversity of sesquiterpene synthases in the basidiomycete *Coprinus cinereus*. Molecular microbiology. 2009; 72:1181–1195. [PubMed: 19400802]

Arigoni D. Stereochemical aspects of sesquiterpene biosynthesis. Pure Appl. Chem. 1975; 41:219–245.

Benedict CR, Lu JL, Pettigrew DW, Liu J, Stipanovic RD, Williams HJ. The cyclization of farnesyl diphosphate and nerolidyl diphosphate by a purified recombinant δ-cadinene synthase. Plant. Physiol. 2001; 125:1754–1765. [PubMed: 11299356]

Bohlmann J, Gershenzon J. Old substrates for new enzymes of terpenoid biosynthesis. Proc. Natl. Acad. Sci. U S A. 2009; 106:10402–10403. [PubMed: 19553206]

Bohlmann J, Meyer-Gauen G, Croteau R. Plant terpenoid synthases: Molecular biology and phylogenetic analysis. Proc. Natl. Acad. Sci. U S A. 1998; 95:4126–4133. [PubMed: 9539701]

Cane DE, Tandon M. Epicubenol synthase and the stereochemistry of the enzymic cyclization of farnesyl and nerolidyl diphosphate. J. Am. Chem. Soc. 1995; 117:5602–5603.

Cane DE, He X, Kobayashi S, Omura S, Ikeda H. Geosmin biosynthesis in *Streptomyces avermitilis*. Molecular cloning, expression, and mechanistic study of the germacradienol/geosmin synthase. J. Antibiot. (Tokyo). 2006; 59:471–479. [PubMed: 17080683]

Christianson DW. Structural biology and chemistry of the terpenoid cyclases. Chem. Rev. 2006; 106:3412–3442. [PubMed: 16895335]

Corre C, Challis GL. New natural product biosynthetic chemistry discovered by genome mining. Nat. Prod. Rep. 2009; 26:977–986. [PubMed: 19636446]

Davis EM, Tsuji J, Davis GD, Pierce ML, Essenberg M. Purification of (+)-δ-cadinene synthase, a sesquiterpene cyclase from bacteria-inoculated cotton foliar tissue. Phytochemistry. 1996; 41:1047–1055. [PubMed: 8728715]

Davis GD, Essenberg M, Berlin KD, Faure R, Gaydou EM. Complete ^{1}H and ^{13}C NMR spectral assignment of δ-Cadinene, a bicyclic sesquiterpene hydrocarbon. Magn. Reson. Chem. 1996; 34:156–161.

Dickschat JS, Martens T, Brinkhoff T, Simon M, Schulz S. Volatiles released by a *Streptomyces* species isolated from the North Sea. Chem. Biodiversity. 2005; 2:837–865.

Ding L, Pfoh R, Ruhl S, Qin S, Laatsch H. T-muurolol sesquiterpenes from the marine *Streptomyces* sp. M491 and revision of the configuration of previously reported amorphanes. J. Nat. Prod. 2009; 72:99–101. [PubMed: 19072130]

Fischbach M, Ward D, Young S, Jaffe D, Gnerre S, Berlin A, Heiman D, Hepburn T, Sykes S, Alvarado L, et al. Annotation of *Streptomyces clavuligerus* ATCC 27064; whole genome shotgun sequencing project. 2008 NCBI Nucleotide Accession Nr. ABJH00000000.

Gennadios HA, Gonzalez V, Di Costanzo L, Li A, Yu F, Miller DJ, Allemann RK, Christianson DW. Crystal structure of (+)-δ-cadinene synthase from *Gossypium arboreum* and evolutionary divergence of metal binding motifs for catalysis. Biochemistry. 2009; 48:6175–6183. [PubMed: 19489610]

Gerber NN. Sesquiterpenoids from Actinomycetes: Cadin-4-ene-1-ol. Phytochemistry. 1971; 10:185–189.

Harada H, Yu F, Okamoto S, Kuzuyama T, Utsumi R, Misawa N. Efficient synthesis of functional isoprenoids from acetoacetate through metabolic pathway-engineered *Escherichia coli*. Appl. Microbiol. Biotechnol. 2009; 81:915–925. [PubMed: 18836713]

Jiang J, He X, Cane DE. Geosmin biosynthesis. *Streptomyces coelicolor* germacradienol/germacrene D synthase converts farnesyl diphosphate to geosmin. J. Am. Chem. Soc. 2006; 128:8128–8129. [PubMed: 16787064]

Lesburg CA, Zhai G, Cane DE, Christianson DW. Crystal structure of pentalenene synthase: mechanistic insights on terpenoid cyclization reactions in biology. Science. 1997; 277:1820–1824. [PubMed: 9295272]

Medema MH, Trefzer A, Kovalchuk A, van den Berg M, Muller U, Heijne W, Wu L, Alam MT, Ronning CM, Nierman WC, et al. The Sequence of a 1.8-Mb bacterial linear plasmid reveals a rich evolutionary reservoir of secondary metabolic pathways. Genome Biol. Evol. 2010; 2:212–224. [PubMed: 20624727]

Nabeta K, Kigure K, Fujita M, Nagoya T, Ishikawa T, Okuyama H, Takasawa T. Biosynthesis of (+)-cubenene and (+)-epicubenol by cell-free extracts of cultured cells of Heteroscyphus planus and cyclization of [2H]farnesyl diphosphates. J. Chem. Soc., Perkin Trans. 1. 1995:1935–1939.

Nabeta K, Mototani Y, Tazaki H, Okuyama H. Biosynthesis of sesquiterpenes of cadinane type in cultured cells of Heteroscyphus planus. Phytochemistry. 1994; 35:915–920.

Nagashima F, Asakawa Y. Sesqui- and diterpenoids from two Japanese and three European liverworts. Phytochemistry. 2001; 56:347–352. [PubMed: 11249099]

Nagashima F, Suda K, Asakawa Y. Cadinane-type sesquiterpenoids from the liverwort Scapania undulata. Phytochemistry. 1994; 37:1323–1325.

Pieper U, Eswar N, Webb BM, Eramian D, Kelly L, Barkan DT, Carter H, Mankoo P, Karchin R, Marti-Renom MA, et al. MODBASE, a database of annotated comparative protein structure models and associated resources. Nucl. Acids Res. 2009; 37:D347–354. [PubMed: 18948282]

Walsh CT, Fischbach MA. Natural products version 2.0: connecting genes to molecules. J. Am. Chem. Soc. 2010; 132:2469–2493. [PubMed: 20121095]

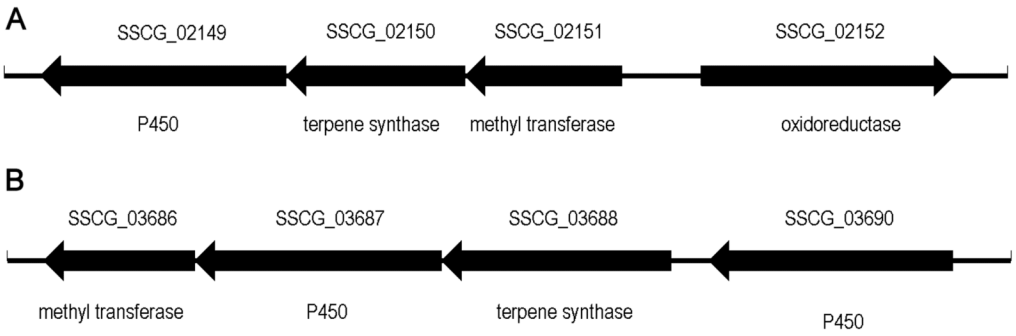

Figure 1.

S. clavuligerus terpene synthase gene clusters. A. SSCG_02150 biosynthetic gene cluster. B. SSCG_03688 biosynthetic gene cluster. (See also Figures S1 and S2).

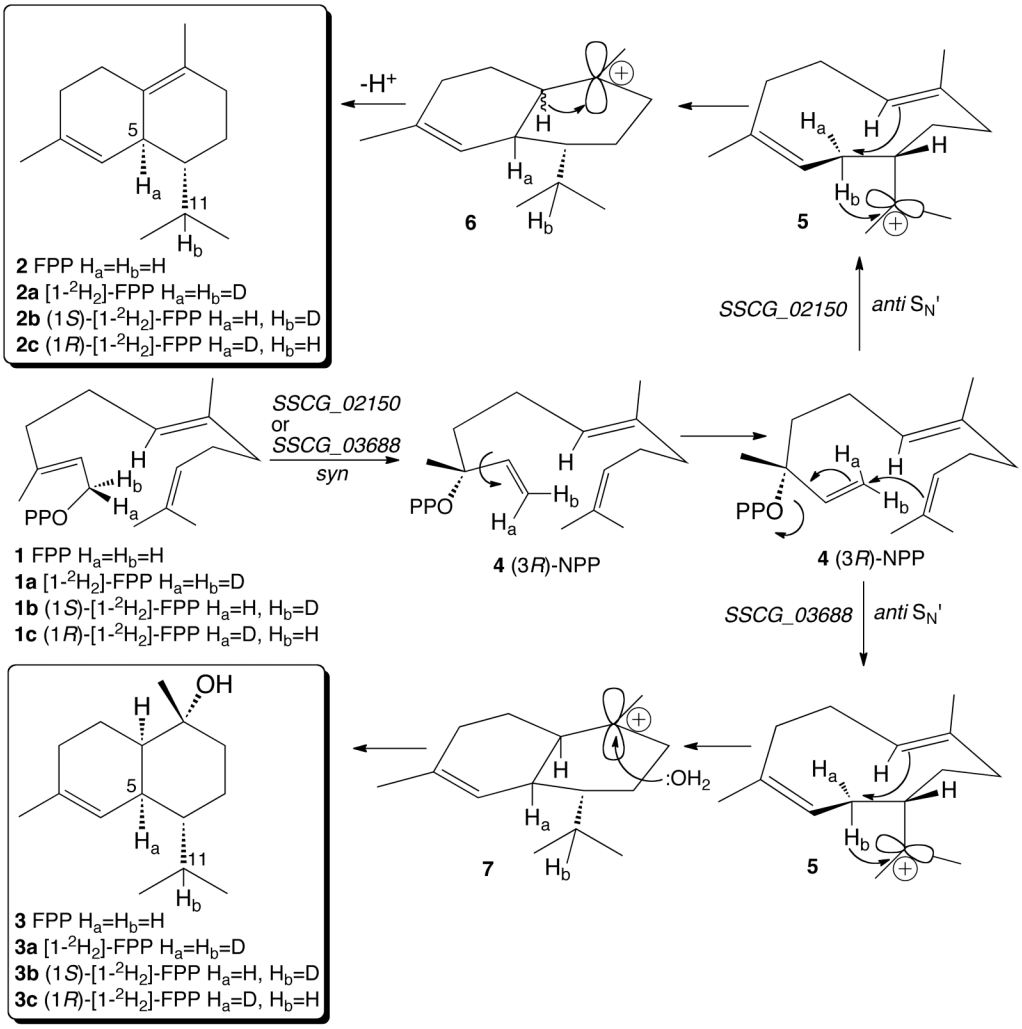

Figure 2.

Mechanism and stereochemistry of the enzymatic cyclization of chirally deuterated FPP. A. SSCG_02150-catalyzed cyclization of FPP to (−)-δ-cadinene (**2**). B. SSCG_03688-catalyzed cyclization of FPP to (+)-T-muurolol (**3**). (See also Figures S3, S6, and S7 and Table S1).

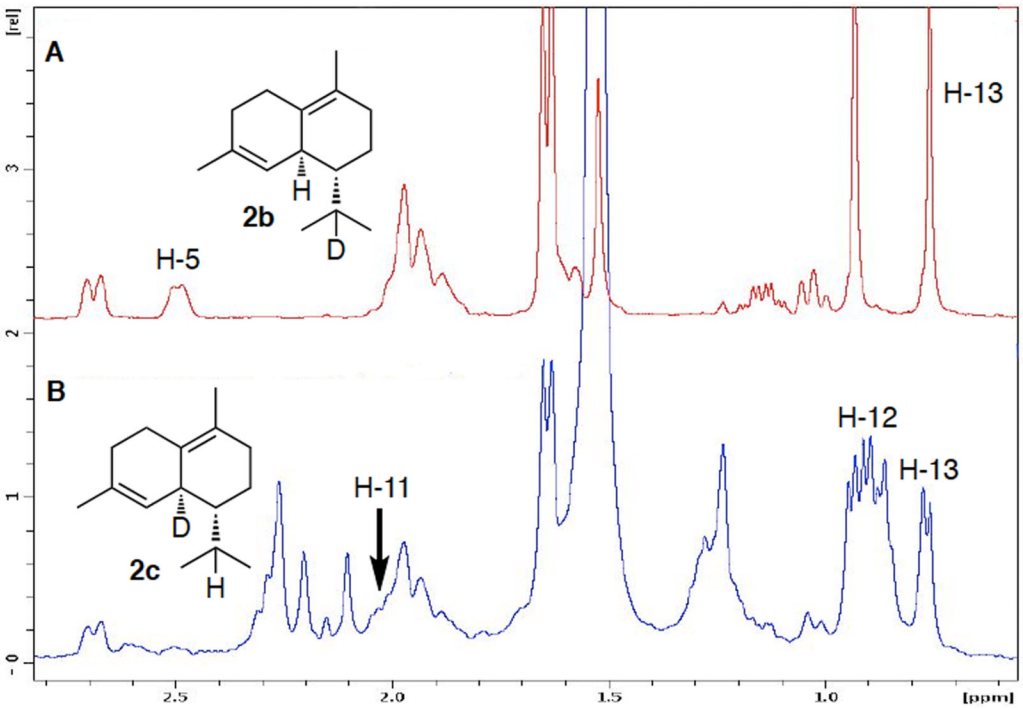

Figure 3.

^{1}H NMR analysis of (−)-δ-cadinene (**2**) generated by SSCG_02150-catalyzed cyclization of chirally deuterated FPP. A. [11-^{2}H]-(−)-δ-cadinene (**2b**) from incubation with (1*S*)-[1-^{2}H]-FPP (**1b**); B. [5-^{2}H]-(−)-δ-cadinene (**2c**) from incubation with (1*R*)-[1-^{2}H]-FPP (**1c**). (See Figure S4)

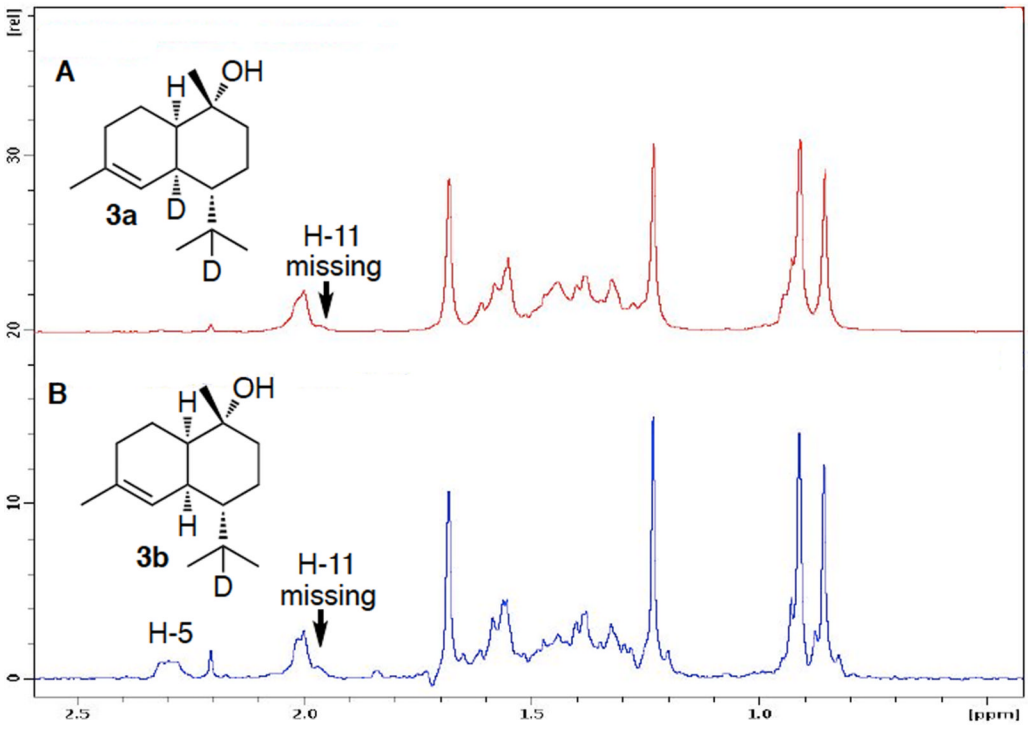

Figure 4.

^{1}H NMR analysis of (+)-T-muurolol (**3**) generated by SSCG_03688-catalyzed cyclization of chirally deuterated FPP. A. [5,11-$^{2}H_{2}$]-(+)-T-muurolol (**3a**) from incubation with [1,1-$^{2}H_{2}$]-FPP (**1a**); B. [11-^{2} H]-(+)-T-muurolol (**3b**) from incubation with (1*S*)-[1-^{2}H]-FPP (**1b**). (See Figure S5)